

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE INGENIERÍA ENSENADA

POSGRADO EN CIENCIAS



**RECONOCIMIENTO DE PATRONES NO LINEAL INVARIANTE A
POSICIÓN, ROTACIÓN, ESCALA Y RUIDO DE IMÁGENES
DIGITALES**

TESIS

**que para cubrir parcialmente los requisitos para obtener el grado de
DOCTOR EN CIENCIAS presenta:**

ÁNGEL CORONEL BELTRÁN

DIRECTOR DE TESIS:

DR. JOSUÉ ÁLVAREZ BORREGO

Ensenada, Baja California, México

Diciembre 2010

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE INGENIERÍA ENSENADA

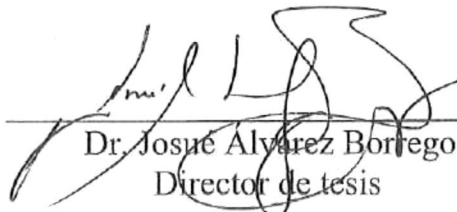
**RECONOCIMIENTO DE PATRONES NO LINEAL INVARIANTE A
POSICIÓN, ROTACIÓN, ESCALA Y RUIDO DE IMÁGENES
DIGITALES**

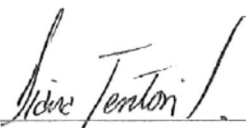
TESIS

que para cubrir parcialmente los requisitos para obtener el grado de
DOCTOR EN CIENCIAS presenta:

ÁNGEL CORONEL BELTRÁN

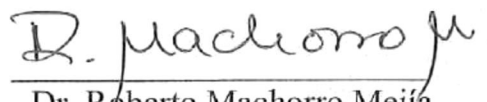
APROBADA POR:


Dr. Josué Álvarez Borrego
Director de tesis


Dra. Diana Tentori Santacruz
Miembro del comité


Dr. José Ángel González Fraga
Miembro del Comité


Dr. Alberto Leopoldo Morán y Solares
Miembro del comité


Dr. Roberto Machorro Mejía
Miembro del comité

Ensenada, Baja California, México

Diciembre 2010

RESUMEN de la Tesis de Ángel Coronel Beltrán, presentada como requisito parcial para la obtención del grado de DOCTOR EN CIENCIAS, Ensenada, Baja California, México. Diciembre de 2010.

RECONOCIMIENTO DE PATRONES NO LINEAL INVARIANTE A POSICIÓN, ROTACIÓN, ESCALA Y RUIDO DE IMÁGENES DIGITALES.

Resumen aprobado por:

Dr. Josué Álvarez Borrego
Director de tesis

RESUMEN

Desde su origen hace casi cinco décadas, el reconocimiento de patrones basado en la operación de correlación tanto óptica como digital, ha tenido gran interés por sus múltiples aplicaciones en las áreas civil y militar. En esta tesis se presenta una metodología de correlación digital no lineal invariante a posición, rotación, escala y color para la identificación de objetos cuando estos están o no inmersos en ruido mediante las operaciones de la transformada de Fourier, el mapeo de coordenadas cartesianas a polares y la transformada de escala, respectivamente. Primeramente, se realiza un experimento numérico para determinar el factor de no linealidad óptimo a utilizarse en el sistema invariante. Se elaboran los algoritmos computacionales para la realización de los filtros de correlación utilizando técnicas no lineales. Mediante el uso de dos métricas se estudia el comportamiento del sistema no lineal y se compara con un sistema invariante lineal para objetos rotados y escalados. Se implementa la correlación de Spearman dentro del sistema no lineal y se aplica a la identificación de objetos con ruido, mostrando un mejoramiento substancial. La caracterización y evaluación en el reconocimiento de patrones mediante la correlación invariante, se realiza utilizando dos criterios para determinar el desempeño de los filtros utilizados en imágenes tanto en la ausencia como en presencia de ruidos aditivo gaussiano y sal y pimienta. Se analiza el desempeño de los filtros no lineales y se comparan con los filtros lineales. Se estudia el comportamiento de la métrica pico de correlación de energía con el tiempo de reacción de lectura de individuos en el sistema no lineal invariante con diferentes tipos de letras de color y diferentes fondos de color. Se presentan las aplicaciones del sistema digital invariante no lineal con cinco investigaciones realizadas utilizando diferentes objetos con filtros lineales y no lineales. Los resultados de estos trabajos probaron la eficiencia y robustez de los algoritmos computacionales desarrollados en el sistema digital de reconocimiento invariante de patrones.

Palabras clave: Reconocimiento de patrones, correlación, filtros no lineales.

ABSTRACT of the thesis, presented by Ángel Coronel Beltrán, in order to obtain the DOCTOR OF SCIENCES degree. Ensenada, Baja California, México. December, 2010.

NONLINEAR INVARIANT PATTERN RECOGNITION TO POSITION, ROTATION, SCALE AND NOISE OF DIGITAL IMAGES.

Approved by:

Dr. Josué Álvarez Borrego
Thesis Advisor

ABSTRACT

Since its inception almost five decades ago, the pattern recognition based on the operation of both optical and digital correlation has been of great interest in many applications in civil and military areas. This thesis presents a methodology for digital nonlinear correlation invariant to position, rotation, scale and color to identify objects when they are immersed or not in noise through the operations of the Fourier transform, the mapping of Cartesian coordinates to polar, and the scale transform, respectively. First, a numerical experiment is performed to determine the optimal nonlinearity factor for use in the invariant system. Computational algorithms are developed for the implementation of correlation filters using nonlinear techniques. By using two metrics we study the behavior of the nonlinear system and is compared with a linear invariant system for rotated and scaled objects. The Spearman correlation within the nonlinear system is implemented and applied to the identification of objects in noise, showing a substantial improvement. The characterization and evaluation in pattern recognition using invariant correlation was performed using two criteria for determining the performance of the filters used in images under the absence or presence of different types of noise. We analyze the performance of nonlinear filters and they are compared with linear filters. We study the behavior of the peak to correlation energy metric with the reaction time reading of individuals in the invariant nonlinear system with different colored font letters and different colored backgrounds. We present the applications of the nonlinear invariant digital system with five realized research using different objects with linear and nonlinear filters. The results of these studies shown the efficiency and robustness of the computational algorithms developed in the invariant pattern recognition digital system.

Keywords: Pattern recognition, correlation, nonlinear filters.

DEDICATORIA

*A mi esposa María de Jesús por su amor, apoyo, comprensión
y palabras de aliento que siempre me animaron a
seguir adelante en todo momento*

*A mis queridos hijos Elisa Idalia, Ángel y Mario quienes
siempre estuvieron atentos del desarrollo de mi trabajo*

AGRADECIMIENTOS

Al Dr. Josué Álvarez Borrego, director de tesis, por su apoyo y disposición expedita en brindarme sus valiosos conocimientos para el desarrollo de mi trabajo, y más aún por su amistad, ayuda incondicional y amabilidad al invitarme a vivir durante un tiempo en su hogar en compañía de su querida familia de quienes estoy enormemente agradecido.

A los miembros del comité de tesis: Dra. Diana Tentori Santacruz, Dr. José Ángel González Fraga, Dr. Alberto Leopoldo Morán y Solares y Dr. Roberto Machorro Mejía por sus valiosos comentarios y sugerencias.

A mi esposa María de Jesús y mis hijos Elisa Idalia, Ángel y Mario, quienes a pesar de la distancia y el tiempo que estuvimos lejos, siempre me brindaron su amor y apoyo para llegar a concluir mis estudios.

A la Universidad de Sonora, por su interés de elevar su nivel académico mediante el apoyo que recibí y de permitirme realizar mis estudios doctorales.

A la Facultad de Ingeniería de la UABC de Ensenada, por facilitarme los medios para la realización de mis estudios.

Al CONACYT por su apoyo económico al otorgarme la beca para realizar mis estudios de doctorado, así como también a su financiamiento del proyecto de ciencia básica: “Correlación invariante a posición, rotación, escala, ruido e iluminación para identificación de organismos y estructuras microscópicas y macroscópicas”, número 102007 a cargo del Dr. Josué Álvarez Borrego.

ÍNDICE	Página
1. Introducción	1
1.1 Antecedentes	1
1.2 Reconocimiento de patrones óptico y digital	8
1.3 Objetivo	12
1.4 Organización de la tesis	13
2. Fundamentación matemática	15
2.1 La transformada de Fourier	15
2.1.1 Propiedades de la transformada de Fourier	17
2.2 Correlación	19
2.3 La transformada de Mellin	21
2.4 La transformada de escala	23
2.4.1 Pasos a seguir para la deducción de la transformada de escala	24
2.4.2 Relación entre la transformada de escala y la transformada de Fourier	25
2.4.3 Relación entre la transformada de escala y la transformada de Mellin	25
2.5 La transformada discreta de Fourier y su inversa en 2-D	26
2.6 Correlación discreta	28
3. Herramientas de caracterización y evaluación: ruido, modelos de imagen, filtros y métricas	29
3.1 Naturaleza del ruido	29
3.2 Ruido blanco	30
3.2.1 Definición matemática de ruido blanco	31
3.2.1.1 Vector aleatorio blanco	31
3.2.1.2 Proceso aleatorio blanco	32
3.2.2 Ruido gaussiano blanco aditivo	32
3.3 Funciones de densidad de probabilidad de algunos ruidos	33
3.3.1 Ruido gaussiano	33
3.3.2 Ruido sal y pimienta	34
3.4 Modelo aditivo	34
3.5 Modelo multiplicativo	35
3.6 Modelo disjunto	36
3.7 Filtros	36
3.7.1 Filtro clásico acoplado (CMF)	36
3.7.2 Filtro sólo de fase (POF)	37
3.7.3 Filtro no lineal (NLF)	38
3.8 Métricas de desempeño	39
3.8.1 Energía pico de correlación (PCE)	39
3.8.2 Capacidad de discriminación (DC)	39
3.8.3 Relación señal a ruido (SNR)	40
3.8.4 Eficiencia luminosa (LE)	41

4. Metodología de la correlación digital no lineal invariante	42
4.1 Correlación no lineal	44
4.2 Obtención del filtro no lineal	44
4.3 Obtención del filtro compuesto no lineal	46
4.4 Correlación no lineal invariante	47
5. Aplicaciones del sistema digital invariante no lineal	50
5.1 Filtro no lineal para el reconocimiento de patrones usando la transformada de escala	50
5.1.1 Simulaciones computacionales	51
5.1.1.1 Determinación del valor óptimo del factor de intensidad de no linealidad k usando como filtro la letra E de fuente Arial	51
5.1.1.2 Determinación del valor óptimo del factor de intensidad de no linealidad k usando imágenes reales	53
5.1.1.3 Reconocimiento del filtro E con las letras del alfabeto en fuente Arial para invariancia a rotación y escala	59
5.1.2 Comparación del desempeño de los filtros no lineales y filtros lineales para invariancia a rotación y escala	63
5.2 Reconocimiento de patrones de objetos libres de ruido y con ruido usando correlación no lineal invariante	70
5.2.1 Comparación de los filtros no lineal y sólo de fase usando un sistema digital de correlación invariante	70
5.2.2 Objetivo con ruido	75
5.2.2.1 DISNF y la correlación de Spearman	78
5.3 Sistema digital de correlación invariante aplicado a la comparación de dos fuentes de letras	82
5.4 Análisis comparativo entre diferentes tipos de fuentes y estilos de letras usando correlación digital invariante	85
5.4.1 Tipografía de las fuentes de letras TNR, Ar y CN	85
5.4.2 Relación del reconocimiento de patrones vía retino cortical y el óptico digital	86
5.4.3 Resultados experimentales	88
5.4.4 Discusión de los resultados	94
5.5 Reconocimiento invariante a rotación, escala y color de imágenes digitales usando un filtro no lineal	98
5.5.1 Modelo de color RGB	98
5.5.2 Trabajos previos	99
5.5.3 Correlación digital no lineal invariante multicanal	101
5.5.4 Resultados experimentales	105
6. Conclusiones	112
Referencias	116
Apéndices	125

Índice de figuras

<u>Figura</u>	<u>Página</u>
3.1 Gráfica de la función de densidad de probabilidad gaussiana $p(z)$.	33
3.2 Ejemplo gráfico del modelo de ruido aditivo.	35
3.3 Ejemplo gráfico del modelo de ruido multiplicativo.	35
4.1 Diagrama de bloques para obtener el filtro no lineal.	45
4.2 Diagrama de bloques para obtener el filtro compuesto no lineal.	48
4.3 Diagrama de bloques representando al sistema de correlación digital invariante.	49
4.4 Procedimiento visual para obtener un filtro no lineal con $k=0.3$ para el caso de un objeto rotado a 0° , 30° y -30° .	51
5.1 Gráfica de cajas para el <i>pico de correlación energía</i> PCE vs los valores de k no lineal para la letra E en rotación.	52
5.2 Gráfica de cajas para el <i>pico de correlación energía</i> PCE vs los valores de k no lineal para la letra E en escala.	53
5.3 (a) la imagen objetivo A, (b) las imágenes problema.	55
5.4 Gráfica de cajas para el <i>pico de correlación de energía</i> PCE versus los valores de no linealidad k para la imagen A en rotación.	56
5.5 Gráfica de cajas para el <i>pico de correlación de energía</i> PCE versus los valores de no linealidad k para la imagen A en escala.	57
5.6 Gráfica de cajas para la comparación del <i>pico de correlación de energía</i> PCE versus los valores de no linealidad k para la letra E y para la imagen A en rotación.	58
5.7 Gráfica de cajas para la comparación del <i>pico de correlación energía</i> PCE versus los valores de no linealidad k para la letra E y para la imagen A en escala.	58
5.8 Desempeño del filtro E para rotación.	61
5.9 Desempeño del filtro E para escala.	61
5.10 Correlación invariante a escala usando (a) un filtro no lineal con $k=0.3$ y PCE=0.0759, (b) un filtro sólo de fase con PCE=0.0156.	64
5.11 Correlación invariante a rotación y escala usando (a) un filtro no lineal con $k=0.3$ y PCE=0.0351, (b) un filtro sólo de fase con PCE=0.0141.	64
5.12 Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y PCE=0.4704, (b) un filtro compuesto sólo de fase con PCE=0.1932.	65
5.13 Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y PCE=0.3879, (b) un filtro compuesto solo de fase con PCE=0.1629.	65
5.14 Correlación invariante a rotación usando (a) un filtro no lineal con $k=0.3$ y PCE=0.2052, (b) un filtro sólo de fase con PCE=0.0342.	67
5.15 Correlación invariante a rotación y escala usando (a) un filtro no lineal con $k=0.3$ y PCE=0.0348, (b) un filtro sólo de fase con PCE=0.0153.	68
5.16 Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y PCE=0.4554, (b) un filtro compuesto sólo de fase con PCE=0.1316.	68
5.17 Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y PCE=0.2136, (b) un filtro compuesto sólo de fase con PCE=0.09.	69

5.18	Correlación invariante a rotación usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.0215$, (b) un filtro sólo de fase con $PCE=0.0057$, y (c) un corte a lo largo del eje de rotación de (a).	72
5.19	Correlación invariante a rotación y escala usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.025$, (b) un filtro solo de fase con $PCE=0.0061$, y (c) un transecto de escala de (a).	75
5.20	Objetivo libre de ruido (a), con ruido gaussiano aditivo de media cero y Variancia de 0.1 (b), 0.2 (c) y 0.3 (d), y con ruido S&PP de densidad de 0.1 (e), 0.2 (f) y 0.3 (g).	76
5.21	Desempeño de un filtro no lineal con $k=0.3$ en la presencia de ruido gaussiano aditivo.	77
5.22	Desempeño de un filtro no lineal con $k=0.3$ en la presencia de ruido S&P.	77
5.23	Diagrama de bloques simplificado representando al sistema de correlación invariante con un filtro no lineal usando el SDISNF.	79
5.24	Comparación del desempeño con el SDISNF cuando el objetivo está inmerso en ruido aditivo gaussiano aditivo.	80
5.25	Comparación del desempeño con el SDISNF cuando el objetivo está inmerso en ruido S&P.	80
5.26	Gráfica de cajas para el <i>pico de correlación de energía</i> PCE versus las letras del alfabeto de fuente Arial y Times New Roman (TNR) con filtros no lineal (NLF) y sólo de fase (POF).	83
5.27	Gráfica de cajas para el <i>pico de correlación de energía</i> PCE versus la letra E de Fuente Arial y Times New Roman (TNR) rotada $\pm 90^\circ$, con incrementos de 5° , con filtros no lineal (NLF) y sólo de fase (POF).	83
5.28	Tipografía de tres fuentes de letras.	86
5.29	Valores del PCE para cada una de las letras del alfabeto en fuentes Arial (Ar) y en Times New Roman (TNR) usando un filtro no lineal (NLF) y un filtro solo de fase (POF).	89
5.30	Desempeño de las letras en fuentes Arial (Ar), Courier New (CN) y Times New Roman (TNR) en estilos plano y cursivo, con cinco combinaciones de colores de frente/fondo.	90
5.31	Valores del PCE de las fuentes Arial (Ar) y Times New Roman (TNR) en estilo plano.	91
5.32	Valores del PCE de las fuentes Arial (Ar) y Times New Roman (TNR) en estilo cursivo.	92
5.33	Valores del PCE de todas las fuentes Arial (Ar) y Times New Roman (TNR) en estilos plano y cursivo.	92
5.34	Gráfica de PCE versus RT (ms) de las letras del alfabeto en fuente Ar y TNR con cinco combinaciones de colores de letra/fondo en estilo plano.	94
5.35	Diagrama esquemático de diferentes campos de estudio para el reconocimiento de patrones.	96
5.36	Cubo RGB representando las coordenadas de color.	99
5.37	Diagrama simplificado de bloques para obtener el filtro no lineal.	103
5.38	Diagrama de bloques representando al sistema de correlación digital Invariante a rotación, escala y color.	104
5.39	Galería de especies de mariposas utilizadas.	106
5.40	Gráficas de cajas del valor promedio de la correlación no normalizada de	108

- cada una de las mariposas, desde la (a) hasta la (r), con respecto al resto de las demás, en rotación desde 0° hasta 180° , con incrementos de 1° .
- 5.41 Gráficas de cajas del valor promedio de la correlación no normalizada de cada una de las mariposas, desde la (a) hasta la (r), con respecto al resto de las demás, en escala desde $\pm 25\%$, con incrementos de 1% . 110
- 5.42 Comparación entre las mariposas en color y en grises. Gráfica de cajas de la correlación no normalizada en rotación desde 0° hasta 180° , con incrementos de 1° , donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación. 111

Índice de tablas

<u>Tabla</u>	<u>Página</u>
5.1 Niveles de confianza en porcentaje de los filtros correlacionados con las 26 letras del alfabeto en rotación.	62
5.2 Niveles de confianza en porcentaje de los filtros correlacionados con las 26 letras del alfabeto en escala.	62
5.3 Comparación entre el PCE y el RT de las letras del alfabeto en fuente Ar y TNR con cinco combinaciones de colores de letra/fondo en estilos plano y cursivo. El RT está dado en milisegundos.	93

Capítulo 1

Introducción

El conocimiento es el alimento del alma

Platón

§ 1.1 Antecedentes

El reconocimiento de patrones es un campo muy amplio del conocimiento científico y tecnológico, este abarca todo el sistema sensorial de captación de señales, entre los cuales los más estudiados en la actualidad tienen que ver con los sistemas auditivos y visuales. Estos son procesos naturales del ser humano que empiezan desde muy temprana edad. En el caso de la percepción visual, el detectar formas geométricas simples y colores, que posteriormente se van desarrollando y perfeccionando para identificar, diferenciar, contar y clasificar objetos o patrones. A partir de los avances alcanzados en varias disciplinas científicas, entre las que podemos mencionar, la invención del láser en 1960, la aparición del filtro espacial complejo u holográfico del tipo de Vander Lugt en 1964, también llamado filtro acoplado clásico, empezó un interés por desarrollar nuevos filtros cada vez más eficientes. Así, con nuevos avances sustanciales en el campo del reconocimiento de patrones ópticos, la elaboración de un modelo analítico para describir y comprender la formación y transformación de imágenes en el sistema visual retino cortical en algunos animales mamíferos y los humanos, y con el desarrollo de la electrónica y la computación digital, fue posible emular en cierto grado estos procesos naturales en sistemas de identificación automatizada hasta hacerlos partes importantes en nuestra vida cotidiana.

La automatización del reconocimiento de patrones se ha convertido en una técnica muy útil para la toma de decisiones aplicada en una gran cantidad de campos, tales como la

industria civil y militar, en la vasta y variada cadena de producción industrial de productos de consumo, en sistemas de seguridad, en el apoyo de la captación, interpretación y análisis de datos experimentales en la ciencia, como la biología, la taxonomía, la física, por mencionar sólo algunas. Con esto se busca un constante mejoramiento en la capacidad de toma de decisiones de los sistemas automatizados cada vez con menos márgenes de error. En los humanos es más difícil reducir errores y en muchos casos se vuelve una tarea imposible de realizar, como sería la de encontrar un producto defectuoso de entre millones de ellos. De aquí la importancia práctica del reconocimiento de patrones debido a las necesidades y demanda por parte de los consumidores por contar con productos cada vez mejores.

La manera en que se lleva a cabo el reconocimiento de patrones es a través de una función conocida como correlación, que nos permite determinar el grado de similitud que guarda el objeto de referencia con el objeto de estudio y así poder discernir y cuantificar el resultado final. Mediante la correlación podemos no solamente identificar al objeto en cuestión, sino además ser capaces de clasificar a dicho objeto. Y con un sistema de correlación, ya sea óptico, digital o híbrido, es posible en un gran número de casos, identificar al objeto sin importar si este presenta desplazamientos espaciales, rotaciones, diferencias en tamaño, color o se encuentra inmerso en diferentes tipos de ruido, etc. En los sistemas digitales, esto se logra con el diseño de algoritmos computacionales de correlación, que son una herramienta poderosa en el reconocimiento de patrones.

Un algoritmo muy utilizado, por su alta velocidad de procesamiento de datos en los sistemas digitales, es el llamado algoritmo de la transformada rápida de Fourier (FFT)¹, que desde su introducción en la comunidad científica (Cooley y Tukey, 1965), (cabe señalar que sus creadores desconocían que tal algoritmo ya había sido elaborado a principios del siglo XIX (Gauss, 1866)), ha hecho factible el avance de nuevos métodos más eficientes de correlación de imágenes digitales. Entre las múltiples aplicaciones que ha encontrado el reconocimiento de patrones desde sus inicios, podemos mencionar las áreas de la biología, la óptica y la fisiología y neurología. Cabe mencionar, que uno de los principales atributos del reconocimiento de patrones es su propia robustez que le permite realizar una evaluación

¹ FFT: Fast Fourier Transform.

efectiva de la señal u objeto en su totalidad sin necesidad de llevar a cabo un procesamiento previo de las imágenes de estudio.

La percepción del mundo tridimensional mediante el sistema visual sensorial, desde el punto de vista fisiológico, es un acto básico de los humanos y de algunas especies del reino animal, y sólo único en la especie humana desde la perspectiva cognoscitiva, pero para los sistemas computacionales sigue siendo todavía un problema sin resolver. Esto último es debido principalmente a la gran complejidad de los objetos que se encuentran presentes en la naturaleza. Las imágenes se obtienen en diferentes situaciones de iluminación y posición de la cámara o sensor óptico que las captura, así como las variaciones de intensidad, rotación, escala y desplazamiento del objeto que se desea reconocer.

Además, en muchos casos existen diferentes tipos de ruido inmerso en la imagen, lo cual complica el reconocimiento, limitando de esta forma la efectividad de los sistemas de visión computarizada para el reconocimiento invariante de objetos. En la mayoría de los casos, el estudio está restringido a escenas visuales bidimensionales, es decir, a imágenes con un determinado patrón o forma y bajo ciertas condiciones ambientales.

En la biología, el reconocimiento de patrones encontró utilidad primeramente en la taxonomía. La identificación y conteo de microorganismos a través de un microscopio sigue siendo un método tedioso para el taxonomista, debido a que le consume bastante tiempo de trabajo, con el consecuente cansancio físico. Lo deseable sería tener un sistema automatizado que realizara esta tarea con mayor rapidez para el reconocimiento de formas. Durante años se realizaron investigaciones en este sentido, como la técnica de procesado óptico coherente en el reconocimiento de plancton, específicamente diatomeas, (Almeida e Indebetouw, 1982).

Cairns y colaboradores (1972), utilizaron la técnica de filtraje espacial óptico coherente en el reconocimiento de patrones de algunas diatomeas con filtros espaciales complejos de tipo Vander Lugt contruidos con un interferómetro de Rayleigh modificado. En ese mismo año, Almeida y colaboradores, utilizaron la técnica de microscopía holográfica para examinar ciertas especies de algas, presentando algunas fotografías de los hologramas

reconstruidos de estas, y discuten el posible estudio de los detalles de estructura de estos organismos.

Continuando con el mismo trabajo, Almeida y Kim-Tzong (1976), presentaron resultados en la identificación de especímenes biológicos aplicando la técnica de correlación óptica convencional con hologramas transformada de Fourier. De esta manera, fue posible monitorear la contaminación del agua, la cual está directamente relacionada con el número de diatomeas contenidas en una muestra dada de agua y contadas en función de las especies en un cierto período de tiempo.

Almeida y colaboradores (1978), hicieron un análisis de muestras de algas en transparencias de 35 mm utilizando un procesador híbrido, en el cual el procesamiento de la información se realiza óptica y digitalmente por medio de una computadora. Las señales de correlación son grabadas en un vidicón, digitalizadas y almacenadas en cinta magnética. Se seleccionaron 25 especies de diatomeas diferentes, y el propósito principal fue compararlas consigo mismas y formar una matriz de 25x25 correlaciones y auto-correlaciones entre ellas para probar la habilidad de discriminación del procesador.

Fujii y Almeida (1979a,b), elaboraron filtros espaciales acoplados de patrones simulados y obtuvieron correlaciones parciales entre el patrón simulado y los microorganismos, mostrando que este filtro es menos sensible a las variaciones de tamaño del objeto, lo cual reduce el número de filtros requeridos para el análisis. En este caso sólo consideraron una porción de la orilla de las diatomeas. El problema consiste en que un taxonomista no solamente reconoce el contorno del objeto, sino también su estructura interna, como son la densidad de estrías y puntos, la dirección y forma de estos. Generalmente estos organismos nunca son iguales, por ejemplo, diatomeas que pertenecen a una misma especie son muy similares en forma, pero pueden ser diferentes en tamaño y en el patrón de estrías. Por lo que es más importante muchas veces encontrar parecidos o similitudes en vez de diferencias para la identificación o clasificación de dichos organismos.

Posteriormente Fujii y colaboradores (1980), desarrollaron un método para el reconocimiento de formas microbiológicas, rotando el filtro espacial acoplado mediante un

prisma de cuña. Este filtro constaba de varias diatomeas de diferentes tamaños de tal manera que pudiera cubrir un amplio rango en cambios de escala para la identificación de una especie dada.

Los esfuerzos por lograr la realización de un dispositivo óptico para la identificación y conteo de microorganismos por los métodos anteriormente descritos, no fueron satisfactorios debido principalmente a que todos ellos efectúan la correlación óptica convencional. En otras palabras, en un sistema óptico coherente se coloca una transparencia del objeto (especimen biológico) en el plano de entrada, y en el plano de Fourier de dicho sistema se coloca un filtro espacial acoplado u holograma transformada de Fourier del objeto que se desea reconocer. La distribución de luz resultante en el plano de salida contiene la información de la correlación (Goodman, 2005). Esta técnica tiene el inconveniente de ser principalmente muy sensible a pequeños cambios de escala y orientación de los objetos (Casasent y Psaltis, 1976a).

Casasent y Psaltis (1976b), mostraron que si una de las imágenes se rota en un grado o se aumenta de tamaño en uno por ciento, entonces la razón señal a ruido del pico de correlación entre la función de entrada y el filtro espacial holográfico decae por un factor de diez. Esto significa que la habilidad de comparar las imágenes con esta técnica convencional de correlación, es muy pobre y estas aparecen muy borrosas en el plano de salida.

Para superar estos problemas, Casasent y Psaltis (1976c) propusieron el uso de la transformada de Mellin, que por sus propias características es invariante a la escala, pero no al corrimiento de la función de entrada (la transformada de Fourier sí es invariante al corrimiento pero no a la escala). Sin embargo, es posible combinar las dos transformaciones para realizar la operación de correlación óptica invariante a escala y corrimiento. También demostraron la invariancia a rotación. Sus resultados confirmaron que era posible correlacionar ópticamente dos objetos o funciones que difieran considerablemente en escala y rotación.

En lo que se refiere al trabajo realizado en el CICESE, Coronel-Beltrán (1988), basado en esta técnica, realizó el primer correlacionador óptico invariante aplicado a la

identificación de microorganismos fitoplanctónicos, para lo cual elaboró filtros holográficos del tipo de Vander Lugt. Su objetivo principal fue identificar diatomeas grabadas en una película fotográfica, utilizando un sistema óptico coherente, independientemente de su posición y orientación. Además, desde finales de la década de los 1980s hasta hoy en día, Álvarez-Borrego, J., y sus diferentes grupos de colaboración, han desarrollado y dado un fuerte impulso en las investigaciones del reconocimiento de patrones tanto óptico como digital e híbrido aplicado a la biología y otras áreas. Podemos citar los trabajos de Zavala-Hamz y Álvarez-Borrego (1997), donde aplicaron filtros armónicos circulares (CHF)², invariantes a rotación, para el reconocimiento de microorganismos planctónicos del género *Acartia* y *Calanus*. Este fue el primer trabajo donde se aplicaron los filtros CHF con fines no militares. Pech-Pacheco y Álvarez-Borrego (1998), hicieron estudios en la identificación de cinco especies de fitoplancton del género *Ceratium*, utilizando un sistema óptico digital y aplicando la técnica de correlación invariante descrita por Casasent y Psaltis (1976a,b,c). Castro-Longoria *et al* (2001) identificaron especies de *Calanoid copepods* usando el algoritmo del módulo al cuadrado de la transformada de Fourier y los patrones de difracción de estas especies se correlacionaron con filtros sólo de fase (POF, por sus siglas en inglés), con la técnica invariante, logrando discriminar las diferentes especies y el sexo. Pech-Pacheco *et al* (2001) propusieron un sistema automático para la localización e identificación de diatomeas en microscopios de transmisión. Álvarez-Borrego y Chávez (2001) utilizaron una correlación digital a color para identificar cuerpos de inclusión de virus en tejido de camarón. Álvarez-Borrego *et al* (2002) usaron la técnica de correlación invariante para identificar la bacteria del cólera *Vibrio cholerae* 01. Las imágenes de estas fueron grabadas con una cámara CCD³ y descompuestas en tres canales de colores, rojo, verde y azul, llamados RGB. Pech-Pacheco *et al* (2003) realizaron la identificación de objetos basados en correlación de fase en el dominio de la transformada de escala para el reconocimiento automático de caracteres invariante a traslación, rotación y escala. Extendieron la metodología usada por Casasent y Psaltis usando la transformada de escala que es más eficiente que la de Fourier-Mellin. Las correlaciones se realizaron usando filtros sólo de fase en un sistema digital. Recientemente,

² CHF: Circular Harmonic Filter.

³ CCD: Charge Coupled Device.

se han hecho trabajos usando firmas unidimensionales (Solorza y Álvarez-Borrego, 2010) para realizar correlación invariante a posición y rotación con un sistema digital para la identificación de las letras del alfabeto de fuente Arial y en la clasificación de fósiles de diatomeas. Lerma y Álvarez-Borrego (2009) utilizaron firmas vectoriales, basadas en las propiedades de las transformadas de escala y de Fourier, para el reconocimiento de patrones invariante a posición, rotación y escala de letras del alfabeto en fuente Arial e imágenes de varias especies de copépodos, calculando la distancia euclideana entre el objetivo y la escena de entrada. Bueno-Ibarra *et al.* (2010a,b) desarrollan una nueva técnica basada en la metodología de la ley-K no lineal de Fourier para clasificar los cuerpos de inclusión del virus del síndrome de la mancha blanca encontrados en imágenes de muestras de tejidos de camarón infectado.

Con respecto a este trabajo, es bien conocido que la escena visual que aparece sobre la retina del ojo, sufre serias distorsiones a medida que pasa a través del nervio óptico hasta llegar al cerebro. En la década de 1970s, se iniciaron los primeros estudios en la formulación de una descripción geométrica concisa de estas transformaciones. Schwartz (1977a,b) muestra en sus estudios cómo el cerebro, en ciertos primates, tiene la habilidad de reconocer las formas de los objetos, sin importar el tamaño y la orientación de los mismos, donde la imagen en la corteza visual aparecía distorsionada. La luz que incide sobre la retina, la cual es una superficie curvada compuesta de células fotosensibles, pasa a través del conjunto de nervios que componen al nervio óptico en forma de pulsos eléctricos. Este mazo de nervios une a cada célula retinal con una neurona en la corteza visual cerebral. De esta forma, la imagen en la retina es “mapeada” o “proyectada” sobre una superficie del cerebro. A cada “punto” (célula) sobre la retina le corresponde un único “punto” (neurona) sobre la corteza visual. Este mapeo de la escena visual a través del logaritmo complejo es importante para el reconocimiento de patrones y estos trabajos hacen referencia a la solución propuesta en los trabajos de Casasent y Psaltis (1976b) de utilizar la transformación de coordenadas de la función de entrada u objeto al plano polar con escalamiento logarítmico para poder realizar la operación de correlación. Los estudios realizados por Schwartz (1994) muestran que existe una fuerte evidencia fisiológica del mapeo polar logarítmico entre la retina y la corteza visual en muchos sistemas visuales incluyendo el humano. Anteriormente Fisher (1973), había hecho un tratamiento analítico

que describía la transformación de la escena visual, desde la retina hasta la corteza visual en los gatos, pero fueron los trabajos de Schwartz los que mejores resultados arrojaron para describir la transformación de la imagen de una escena visual en el proceso retino cortical.

Las contribuciones de Casasent y Psaltis en el campo del reconocimiento óptico de patrones y de Schwartz por el lado de la fisiología de la visión en sus estudios con primates, se presentaron durante el mismo período de mediados y fines de la época de los 1970s. La característica fundamental de sus estudios radica en el hecho de que ambos requieren de la realización de una transformación de la imagen en coordenadas cartesianas a coordenadas polares con escalamiento logarítmico para lograr el reconocimiento de patrones invariante a escala y rotación, que es el principio que gobierna la geometría de la visión.

§ 1.2 Reconocimiento de patrones óptico y digital

En el lenguaje común, el término *patrón*, introducido por Parsons y al que Spearman llama *forma* (Köhler, 1947), es usado generalmente como sinónimo de regularidad. Dicho de esta manera, un *patrón* es un término general que expresa la regularidad reconocible de un objeto. Desde un punto de vista más formal, en el análisis de imágenes, podemos decir que un *patrón* es una descripción estructural o cuantitativa de un objeto o de alguna otra entidad de interés en una imagen (Gonzalez y Woods, 2008d), en otras palabras, es una disposición de descriptores o características. El sentido más simple y probablemente original de la palabra forma, que nosotros llamamos patrón, hace referencia a la figura espacial de las cosas. Pero la peculiaridad del concepto consiste en la abstracción que hacemos al prescindir de la materia de las cosas y considerar la forma como algo independiente, lo que espacialmente entendemos como figura. Por otro lado, clasificamos a los objetos según sus formas espaciales; cuadrados, círculos, esferas, etc., agrupándolos por lo que tienen en común sin tener en cuenta la materia o contenido que los diferencia. Existen tres maneras diferentes para el tratamiento del concepto de forma, estas son: como figura geométrica o espacial, como principio filosófico general constitutivo de lo real, y como formalidad o punto de vista del conocimiento. De estas, es claro que la que nos interesa en nuestro estudio es la forma tratada como figura espacial, la cual es la representación gráfica de un objeto. Cuando una forma se descompone en sus partes, pierde su configuración y se

percibe como no configurada. En este sentido, decimos que “la forma es un todo”, es algo más que la suma de sus partes, de acuerdo con la teoría de la Gestalt. Si se alteran los elementos que la conforman, pierde significación. En la filosofía platónica la forma adquiere un significado más amplio, entendida como *idea* que subyace como *esencia* real e independiente de las cosas materiales y sus posibilidades.

En el reconocimiento de patrones existen dos vertientes a seguir, cada una de ellas con sus potencialidades y debilidades. Una está basada en el procesamiento óptico y la otra en el procesamiento digital de imágenes. En ambos casos el estudio está limitado a objetos estáticos, independientes del tiempo, y sólo son consideradas imágenes bidimensionales.

Procesamiento óptico de imágenes

El procesamiento óptico de imágenes consiste en utilizar los fundamentos teóricos de la óptica de Fourier, donde se utiliza la luz y sus propiedades para describir su propagación. La idea básica es obtener la transformada de Fourier de imágenes, para después trabajar con ellas haciendo operaciones matemáticas, tales como la multiplicación, convolución, correlación y la elaboración de filtros espaciales para obtener la identificación o extracción de algún elemento de la imagen. Las operaciones matemáticas se llevan a cabo en un sistema óptico coherente donde se utilizan elementos ópticos. Sin embargo, su aplicación práctica inicia apenas a principios de la década de los 1960s, cuando comenzaron a usarse los láseres y los filtros holográficos o filtros de acoplamiento clásico (Vander Lugt, 1964). La principal ventaja de este método de reconocimiento es la velocidad a la que se realizan las operaciones (el tiempo que tarda la luz en atravesar las lentes y los moduladores espaciales) sin importar la cantidad de datos de entrada, lo cual es una gran ventaja si se requiere que el proceso entero funcione en tiempo real. Sus principales desventajas son la complejidad del arreglo, costo elevado y su inestabilidad, es decir es muy sensible a ruido ambiental.

Procesamiento digital de imágenes

Las imágenes digitales son obtenidas en diferentes condiciones ambientales por un sensor óptico-digital, lo cual causa algunos problemas para la identificación y caracterización del objeto usando sistemas de visión computarizada. El estudio está limitado a objetos estáticos y sólo son consideradas imágenes bidimensionales. La visión por computadora es algo análogo al procesado óptico, ya que las operaciones realizadas sobre las imágenes, simplemente son “simulaciones digitales” de su equivalente óptico. Esto no es algo tan trivial como podría parecer, existen algunos algoritmos básicos y algunos muy poderosos y bastante complejos capaces de asistir en esta tarea, sin embargo es un área de investigación activa con una gran cantidad de artículos publicados al año con respecto a este tema. Sus ventajas y desventajas son inversamente proporcionales a las del procesado óptico, ya que el implementar un sistema digital resulta más barato y más simple. Sin embargo, a pesar de los grandes avances en cuanto a velocidad de cómputo, el coste computacional siempre es algo que se debe tener en mente especialmente si se manejara una gran cantidad de datos, ya que el tiempo de procesado aumenta significativamente cuando aumenta la cantidad de datos de entrada, es decir el método digital siempre será más lento que el óptico.

En nuestro estudio tratamos con el reconocimiento digital de patrones invariantes. El significado del término *invariante* desde la perspectiva de este trabajo lo podemos entender como la representación de una señal que es constante bajo cierta transformación. Por ejemplo, la medición S de una señal x es invariante bajo la transformación g cuando (Kyrki, 2002)

$$S\{x\}=S\{g[x]\}. \quad 1.1$$

Usando teoría de grupos, esta definición puede ser extendida de la siguiente manera: Considere que hay un grupo G que actúa sobre el conjunto X de posibles señales. Denotemos por S la acción de la medida. Si

$$S\{g[x]\}=S\{x\}h(g) \quad \forall g \in G, x \in X, \quad 1.2$$

entonces S es invariante bajo la acción del grupo. El grupo G puede ser, por ejemplo, el grupo de traslaciones, $h(g)$ es función de g solamente. Si $h(g)=1$, entonces S es un

invariante escalar, como se muestra en la ecuación (1.1). Las variaciones geométricas se categorizan basadas en las transformadas. La variación más simple es la traslación del objeto en la imagen. Cuando tenemos traslación más rotación, esto nos da como resultado el *grupo de transformadas euclidianas*, y si además introducimos la escala nos da el *grupo de transformadas de similitud*. Si esta última contiene un escalamiento no isotrópico, entonces forma el *grupo de transformadas afines*, que en los trabajos que aquí se hacen referencia sobre el reconocimiento de patrones, no se toman en cuenta. Por ejemplo, la transformada de Fourier es invariante bajo el grupo de traslaciones, mientras que su amplitud es invariante escalar. Los invariantes se categorizan como fuertes o débiles, de acuerdo a la acción de g que puede ser medida de la representación invariante. La representación se llama fuerte si contiene una componente que explícitamente codifica el grado de transformación. Así, de acuerdo a esta definición, los invariantes escalares son invariantes débiles. Se clasifican a los invariantes como globales y locales. Los primeros tienen que ver con la forma como un todo, mientras que los segundos se refieren a las propiedades locales, tales como la curvatura del objeto. Nosotros estudiamos el reconocimiento de patrones invariantes del tipo global.

En este trabajo se presentan resultados aplicando nuevos algoritmos computacionales para la identificación de objetos, independientemente del tamaño, orientación angular y ruido de dichos objetos. La metodología a seguir fue la de elegir una transformada integral digital eficiente como una alternativa a la muy conocida transformada de Fourier-Mellin, como lo es la transformada de escala (Cohen, 1995). Además, introducimos algunos filtros no lineales para realizar las operaciones de correlación digital que nos dan información de la similitud de las imágenes entre los diferentes objetos. Esta clase de filtros tienen ventajas comparados con los filtros clásicos del tipo de Vander Lugt (1964) o los filtros CHF (Zavala y Alvarez-Borrego, 1997), entre otros, debido a que tienen mayor capacidad para discriminar objetos de diferentes tamaños, el máximo valor del pico de correlación se encuentra bien localizado y con menor ruido.

La clase de sistema que utilizamos es del tipo digital. Por diversas razones, hemos elegido este sistema porque entre ellas están, el rápido y creciente desarrollo de nuevo software y hardware; y más importante, el bajo costo del equipo comparado con los

sistemas ópticos o con los híbridos que consisten de una parte óptica y otra digital, que por sus propias características tienen mayor grado de dificultad al hacer los montajes y la manipulación de los aparatos, equipos y elementos ópticos y el consecuente alto costo de adquisición, mantenimiento y servicio preventivo, entre otras cosas.

§ 1.3 Objetivos

1.3.1 Objetivo general

Elaborar una metodología de correlación no lineal invariante a posición, rotación, escala y color capaz de identificar objetos en presencia de ruido y que tenga un buen desempeño en cuanto a su habilidad de distinguir los objetos.

1.3.2 Objetivos particulares

1. Realización de un experimento numérico para determinar el factor k de no linealidad óptimo a utilizarse en el sistema invariante.
2. Elaborar los algoritmos computacionales para la realización de los filtros de correlación utilizando técnicas no lineales.
3. Incorporar invariancia a desplazamiento, rotación, escalamiento y ruido a los filtros de correlación diseñados.
4. Mediante el uso de dos métricas analizar el comportamiento del sistema no lineal y compararlo con un sistema invariante lineal para objetos rotados y escalados
5. Implementación de la correlación de Spearman dentro del sistema no lineal y aplicarla a la identificación de objetos con ruido.

6. Análisis del comportamiento de la métrica PCE en el sistema no lineal invariante con diferentes tipos de letras de color y diferentes fondos de color.
7. Partición del sistema invariante en canales de color aplicado a la identificación de imágenes reales.

§ 1.4 Organización de la tesis

Esta tesis está organizada de la siguiente manera: El Capítulo 2 presenta los fundamentos matemáticos utilizados en el procesamiento de las imágenes aquí tratadas. Entre los que podemos mencionar las transformadas integrales lineales de funciones continuas bidimensionales de las más representativas, a saber las de Fourier, Mellin y de escala con sus propiedades características. Además, se introducen las funciones de convolución y correlación.

En el Capítulo 3 se presentan las herramientas de caracterización y evaluación en el reconocimiento de patrones mediante la correlación invariante donde se han utilizado dos criterios para determinar el desempeño de los filtros utilizados en imágenes tanto en la ausencia como en presencia de diferentes tipos de ruidos.

En el Capítulo 4 se presenta la metodología utilizada en la realización de la correlación digital no lineal invariante a desplazamiento, orientación y tamaño para el reconocimiento de patrones de diferentes objetos tanto artificiales como de la vida real mediante las operaciones de la transformada de Fourier, el mapeo de coordenadas cartesianas a polares y la transformada de escala, respectivamente. Se analiza el desempeño de los filtros no lineales y se comparan con los filtros lineales.

En el Capítulo 5 se presentan las aplicaciones del sistema digital invariante no lineal con cinco investigaciones realizadas utilizando diferentes objetos con filtros lineales y no lineales. Los resultados de estos trabajos probaron la eficiencia y robustez de los algoritmos

computacionales desarrollados en el sistema digital de reconocimiento de patrones invariante. Y en el Capítulo 6 se presentan las conclusiones de este trabajo.

En los Apéndices A, B y C, se muestran las derivaciones de los teoremas y de algunas ecuaciones utilizadas.

Capítulo 2

Fundamentación matemática

El estudio profundo de la naturaleza es la fuente más fértil de descubrimientos matemáticos

Jean-Baptiste-Joseph Fourier

En este capítulo se presentan las herramientas matemáticas que sirven de fundamento y soporte teórico al trabajo aquí presentado. Estas son: las transformadas de Fourier, Mellin y de escala con sus respectivas propiedades fundamentales. Así como también se introducen los conceptos de convolución y correlación. Las demostraciones de los teoremas de Fourier y de Mellin y la deducción de la transformada de escala son simples sin la necesidad de invocar el rigor matemático característico en estudios más profundos sobre el tema y se dan en la sección de apéndices. Las definiciones y teoremas aquí tratados se consideran los más representativos y de utilidad práctica en el desarrollo de este trabajo.

§ 2.1 La transformada de Fourier

El tratamiento que se hace aquí es para funciones de dos variables independientes, que son el tipo de funciones que se utilizan para representar las imágenes digitales. Muchas de las derivaciones se pueden encontrar en los libros escritos por Bracewell (1978) y Goodman (2005).

La *transformada de Fourier* de una función continua $f(x, y)$ de dos variables continuas x y y , denotada por $\mathcal{F}[f(x, y)]$ está definida por la ecuación

$$\mathfrak{F}[f(x, y)] = \iint_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy, \quad 2.1$$

donde $j = \sqrt{-1}$, las variables u y v son también continuas y se refieren a ellas como *frecuencias*. Como $\mathfrak{F}[f(x, y)]$ es función solo de u y v escribimos $\mathfrak{F}[f(x, y)] = F(u, v)$; es decir, la transformada de Fourier de $f(x, y)$ la podemos escribir como

$$F(u, v) = \iint_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy. \quad 2.2$$

Similarmente, dado $F(u, v)$, podemos obtener $f(x, y)$ mediante la *transformada de Fourier inversa*, $f(x, y) = \mathfrak{F}^{-1}[F(u, v)]$, definida como

$$f(x, y) = \iint_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv. \quad 2.3$$

A estas dos ecuaciones se les conoce como *par de transformadas de Fourier*.

Para ciertas funciones, desde el punto de vista matemático, estas integrales no existen, por lo tanto se deben tomar en consideración algunas condiciones de existencia. Del conjunto de condiciones suficientes para la existencia de la transformada de Fourier, las más comunes son:

1. $f(x, y)$ debe ser absolutamente integrable sobre el plano infinito x - y ; esto es, $\iint_{-\infty}^{\infty} |f(x, y)| dx dy < \infty$.
2. $f(x, y)$ debe tener sólo un número finito de discontinuidades y un número finito de de máximos y mínimos en cualquier rectángulo finito.
3. $f(x, y)$ no debe tener discontinuidades infinitas.

Sin embargo, existen funciones desde el punto de vista físico que pueden violar cualquiera de estas condiciones de existencia⁴. Un ejemplo de ello es la conocida función

⁴ Como lo menciona Bracewell (1978), "...la posibilidad física es una condición suficientemente válida para la existencia de una transformada...".

delta de Dirac, la cual siendo infinita en el origen y cero en cualquier otra parte, tiene una discontinuidad infinita y por lo tanto no satisface la tercera condición de existencia. Otros ejemplos son las funciones $f(x, y) = 1$ y $f(x, y) = \cos(2\pi ux)$, ambas no satisfacen la primera condición de existencia. Afortunadamente, es posible mostrar la validez de una transformada de ciertas funciones que no satisfacen estrictamente las condiciones de existencia siempre y cuando estas funciones puedan ser definidas como el límite de una secuencia de funciones que son transformables. Al transformar cada una de las funciones de la secuencia definida, se va generando una secuencia correspondiente de transformadas, de tal manera que en el límite de esta nueva secuencia tenemos lo que se llama la *transformada de Fourier generalizada* de la función original. Para visualizar esto último tomemos como ejemplo la función delta de Dirac, que como habíamos dicho, viola la tercera condición de existencia y la cual se puede representar como $\delta(x, y) = \lim_{N \rightarrow \infty} N^2 e^{-\pi N^2(x^2+y^2)}$. Observamos que cada una de las funciones de esta secuencia satisface las condiciones de existencia y cada una de ellas tiene una transformada de Fourier dada por $\mathfrak{F}[N^2 e^{-\pi N^2(x^2+y^2)}] = e^{-\pi(u^2+v^2)/N^2}$. De acuerdo a la transformada generalizada de $\delta(x, y)$ se tiene que $\mathfrak{F}[\delta(x, y)] = \lim_{N \rightarrow \infty} e^{-\pi(u^2+v^2)/N^2} = 1$. Esto significa que el espectro de la función delta se extiende uniformemente en todo el dominio de frecuencias.

2.1.1 Propiedades de la transformada de Fourier

Algunas propiedades básicas de la transformada de Fourier están representadas con los siguientes teoremas y las demostraciones están en el APÉNDICE A.

Teorema de la suma o de la linealidad. $\mathfrak{F}[ag(x, y) + bh(x, y)] = a\mathfrak{F}[g(x, y)] + b\mathfrak{F}[h(x, y)]$, donde a y $b \in \mathbb{R}$; esto es, la transformada de la suma de dos funciones es simplemente la suma de sus transformadas individuales.

Teorema de similitud o de escalamiento. Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces $\mathfrak{F}[g(ax, by)] = \frac{1}{|ab|} G\left(\frac{u}{a}, \frac{v}{b}\right)$, donde a y $b \in \mathbb{R}$; esto es, una expansión en el dominio

espacial (x, y) resulta en una contracción en el dominio frecuencial (u, v) y viceversa, además un cambio de la amplitud en todo el espectro.

Teorema de desplazamiento. Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces $\mathfrak{F}[g(x - a, y - b)] = G(u, v)e^{-j2\pi(ua+vb)}$, donde a y $b \in \mathbb{R}$; esto es, una traslación de la función en el dominio espacial introduce un corrimiento en fase en el dominio frecuencial.

Teorema de Parseval. Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces

$$\iint_{-\infty}^{\infty} |g(x, y)|^2 dx dy = \iint_{-\infty}^{\infty} |G(u, v)|^2 du dv.$$

Esto es, la integral del cuadrado del valor absoluto de una función es igual a la integral del cuadrado del valor absoluto de su transformada de Fourier ⁵. Este teorema se interpreta como un enunciado de la conservación de la energía, donde el área bajo la curva de esta integral es la variancia total de la función.

Teorema de convolución. Si $\mathfrak{F}[g(x, y)] = G(u, v)$ y $\mathfrak{F}[h(x, y)] = H(u, v)$, entonces

$$\mathfrak{F}\left[\iint_{-\infty}^{\infty} g(\xi, \eta)h(x - \xi, y - \eta)d\xi d\eta\right] = G(u, v)H(u, v).$$

La convolución de dos funciones en el dominio espacial equivale a multiplicar sus transformadas individuales. La multiplicación es usualmente una operación mucho más simple y se visualiza de una manera más sencilla.

Teorema de autocorrelación. Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces

⁵ Esta relación procede de un teorema de 1799 sobre series, cuyo creador fue Marc-Antoine Parseval des Chênes (1755-1836), que más tarde se aplicó a las Series de Fourier. Su resultado fue publicado hasta que sus únicos cinco artículos que realizó en toda su vida, fueron publicados por la Académie des Sciences en 1806. <http://www-history.mcs.st-andrews.ac.uk/Printonly/Parseval.html>

$$\mathfrak{F} \left[\iint_{-\infty}^{\infty} g(\xi, \eta) g^*(\xi - x, \eta - y) d\xi d\eta \right] = |G(u, v)|^2.$$

Similarmente,

$$\mathfrak{F}[|g(\xi, \eta)|^2] = \iint_{-\infty}^{\infty} G(\xi, \eta) G^*(\xi + u, \eta + v) d\xi d\eta.$$

Este teorema puede considerarse como un caso especial del teorema de convolución.

§ 2.2 Correlación

Para comparar dos funciones u objetos diferentes, $f(x, y)$ y $g(x, y)$, se define la *función correlación* como

$$c_{fg}(u, v) = \iint_{-\infty}^{\infty} f(x, y) g(x + u, y + v) dx dy. \quad 2.4$$

Y la *función autocorrelación* se define como

$$c_{ff}(u, v) = \iint_{-\infty}^{\infty} f(x, y) f(x + u, y + v) dx dy. \quad 2.5$$

Ambas funciones de correlación pueden expresarse en una, dos y tres dimensiones espaciales y una temporal. Para hacer una exploración de ellas y por simplicidad tomaremos una dimensión de la forma:

$$c_{fg}(u) = \int_{-\infty}^{\infty} f(x) g(x + u) dx. \quad 2.6$$

Y similarmente para la función de autocorrelación

$$c_{ff}(u) = \int_{-\infty}^{\infty} f(x) f(x + u) dx. \quad 2.7$$

Para que la variable u pueda describir un corrimiento en la dirección del eje x positivo, la ecuación 2.6 se reescribe como

$$c_{fg}(u) = \int_{-\infty}^{\infty} f(x)g(x-u)dx. \quad 2.8$$

La definición de la función correlación se expresa también de la siguiente forma:

$$c(x) = f(x) \otimes g(x) = \int_{-\infty}^{\infty} f(x')g(x'-x)dx', \quad 2.9$$

donde $\otimes$ significa correlación y similarmente para la función de autocorrelación. En esta formulación x es el corrimiento y x' es la variable muda en la integración.

La correlación es por lo tanto parecida a la convolución, excepto que esta difiere en que g es desplazada pero no invertida o doblada.

Para la *convolución*⁶ tenemos

$$h(x) = f(x) \odot g(x) = \int_{-\infty}^{\infty} f(x')g(x-x')dx', \quad 2.10$$

donde $\odot$ significa convolución. Cuando las funciones son de variable compleja, se define la *función de correlación compleja* como

$$\gamma_{fg}(x) = f(x) \otimes g^*(x) = \int_{x'=-\infty}^{x'=+\infty} f(x')g^*(x'-x)dx' \quad 2.11$$

y la *función de autocorrelación compleja* como

$$\gamma_{ff}(x) = f(x) \otimes f^*(x) = \int_{x'=-\infty}^{x'=+\infty} f(x')f^*(x'-x)dx'. \quad 2.12$$

⁶ A la convolución también se le conoce como *producto plegado o doblado* (del Alemán *Faltung*), *producto composición* y por *integral de superposición*.

Para las funciones complejas $f(x)$ y $g(x)$, la relación entre la convolución y la correlación puede resumirse como

$$f(x) \otimes g^*(x) = f(x) \odot g^*(-x). \quad 2.13$$

Se sigue entonces que

$$f(x) \otimes g^*(x) \neq g^*(x) \otimes f(x), \quad 2.14$$

lo cual significa que la correlación no es conmutativa.

§ 2.3 La transformada de Mellin

La transformada de Mellin, llamada así en honor al matemático Hjalmar Mellin (1854-1933), es una transformada integral que está íntimamente relacionada con la teoría de las series de Dirichlet. Es usada en la teoría de números y está fuertemente relacionada con la transformada de Laplace, la transformada de Fourier y la función gamma, y forma parte de las funciones especiales. Ha sido aplicada en la restauración de imágenes espacialmente variantes (Sawchuk, 1974) y en el análisis de redes que varían con el tiempo (Gerardi, 1959), entre otras.

La transformada de Mellin de una función f está definida como (Courant y Hilbert, 1953):

$$\mathfrak{M}_f(p) = g(p) = \int_0^{\infty} f(t)t^{p-1}dt \quad 2.15$$

y su transformada inversa por:

$$f(t) = \frac{1}{2\pi j} \int_{\beta - \infty j}^{\beta + \infty j} g(p) t^{-p} dp. \quad 2.16$$

La notación implica que es una integral de línea tomada sobre una línea vertical en el plano complejo. Las condiciones en las cuales es posible esta inversión están dadas en el teorema de inversión de Mellin (Mellin, 1902). Las fórmulas de inversión de Mellin, presentadas por un par de teoremas, vienen dados en Courant y Hilbert (1953). La demostración de estas fórmulas se presenta en el APÉNDICE B.

La propiedad principal de la transformada de Mellin es la invariancia a escala, la cual podemos probar de la siguiente manera:

Si definimos $\mathfrak{M}\{f(t)\} = \mathfrak{M}_f(s) = \int_0^\infty t^{s-1} f(t) dt$ como la transformada de Mellin de la función $f(t)$, donde $s \in \mathbb{C}$, sobre el eje imaginario $s = j\omega$, y $\mathfrak{M}\{f(\alpha t)\} = \mathfrak{M}_f^\alpha(s)$ como la transformada de Mellin de la función escalada $f(\alpha t)$, donde α es el factor de escala. Entonces

$$\mathfrak{M}\{f(\alpha t)\} = \mathfrak{M}_f^\alpha(\omega) = \int_0^\infty t^{j\omega-1} f(\alpha t) dt. \quad 2.17$$

Y haciendo la sustitución $\eta = \alpha t$, $t = \eta/\alpha$, nos queda

$$\mathfrak{M}_f^\alpha(\omega) = \int_0^\infty (\eta/\alpha)^{j\omega-1} f(\eta) \frac{d\eta}{\alpha} = \alpha^{-j\omega} \int_0^\infty \eta^{j\omega-1} f(\eta) d\eta = \alpha^{-j\omega} M_f(\omega). \quad 2.18$$

Es decir,

$$\mathfrak{M}_f^\alpha(\omega) = e^{-j\omega \ln \alpha} \mathfrak{M}_f(\omega). \quad 2.19$$

Esto significa que cambios de escala se manifiestan en términos sólo de fase en el dominio de Mellin. Por lo tanto, la magnitud de la transformada de Mellin es invariante a cambios de escala,

$$|\mathfrak{M}_f(\omega)| = |\mathfrak{M}_f^\alpha(\omega)| \quad 2.20$$

Es decir, los módulos de las transformadas de Mellin de dos funciones que difieren en tamaño, son iguales. Esta invariancia en la escala no se presenta en la transformada de Fourier, como se puede ver del teorema de similitud o de escalamiento al tomar los módulos en ambos lados de la igualdad. Sin embargo, la transformada de Fourier es invariante a corrimiento como se aprecia en el teorema de desplazamiento al tomar los módulos en ambos lados de la igualdad. Esta propiedad de desplazamiento no se presenta en la transformada de Mellin.

§ 2.4 La transformada de escala

La transformada de escala es un caso especial de la transformada de Mellin, donde ésta última está definida por la ecuación 2.15:

$$\mathfrak{M}_f(p) = \int_0^{\infty} f(t)t^{p-1}dt$$

donde la variable compleja $p = -jc + \beta$, con el parámetro fijo $\beta \in \mathbb{R}$ y la variable independiente $c \in \mathbb{R}$. Esta familia de transformadas es llamada la transformada β -Mellin (De Sena y Rocchesso, 2007). La parte real de la variable compleja p está parametrizada, para el caso $\beta = \frac{1}{2}$ tenemos la transformada de escala. Es posible usar otros valores para β ; para $\beta = 0$ tenemos la transformada invariante compresión/expansión, y para $\beta = -1$ tenemos la transformada invariante a la forma. Por lo tanto, la transformada de escala es una restricción de la transformada de Mellin en la línea vertical $p = -jc + \frac{1}{2}$.

La transformada de escala es útil porque es más sensible a los cambios de escala. Si llamamos c la variable de escala, entonces la transformada de escala y su inversa están dadas por (De Sena y Rocchesso, 2004)

$$D_f(c) = (2\pi)^{-1/2} \int_0^{\infty} f(t)e^{(-jc-1/2)\ln t} dt \quad 2.21$$

y

$$f(t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} D_f(c) e^{(jc-1/2)\ln t} dc. \quad 2.22$$

En este trabajo se usó la transformada de escala en 2-D en coordenadas polares (r, θ) con el log de la coordenada radial $\rho = \ln r$, la cual está expresada como (Cristóbal y Cohen, 1998)

$$D(c_\rho, c_\theta) = (2\pi)^{-1/2} \iint_{00}^{\infty 2\pi} e^{\rho/2} f(\rho, \theta) e^{-j(\rho c_\rho + \theta c_\theta)} d\rho d\theta, \quad 2.23$$

donde se ha utilizado la implementación de la transformada de escala no-separable para mantener la invariancia a la rotación (De Sena y Rocchesso, 2007).

2.4.1 Pasos a seguir para la deducción de la transformada de escala

Para la deducción de la transformada de escala se siguen los siguientes pasos:

1. Se parte de los operadores lineales frecuencia W y tiempo T .
2. Se requiere que los operadores sean hermitianos, porque esto garantiza que las funciones propias sean completas y ortogonales y que los valores propios sean reales.
3. Del álgebra de operadores se llega al operador escala $c = WT + (1/2)j$.
4. Se resuelve el problema de valores propios para el operador escala.
5. Se llega después de hacer las sustituciones a la transformada escala.

La deducción de esta transformada está dada en el APÉNDICE C.

2.4.2 Relación entre la transformada de escala y la transformada de Fourier

Si consideramos una función $f_l(t)$, reemplazando a t por $\ln t$, de la forma

$$f_l(t) = \frac{1}{\sqrt{t}} f(\ln t), \quad 2.24$$

y considerando ahora la transformada de escala de la función $f_l(t)$, al sustituirla en la ecuación 2.21 tenemos que

$$\begin{aligned} D_l(c) &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f_l(t) \frac{e^{-jc \ln t}}{\sqrt{t}} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \left[\frac{1}{\sqrt{t}} f(\ln t) \right] \frac{e^{-jc \ln t}}{\sqrt{t}} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(\ln t) \frac{e^{-jc \ln t}}{t} dt. \end{aligned}$$

Haciendo $\tau = \ln t$; $d\tau = \frac{1}{t} dt$; $dt = t d\tau$, tenemos que

$$\begin{aligned} D_l(c) &= \frac{1}{\sqrt{2\pi}} \int f(\tau) \frac{e^{-jc\tau}}{t} t d\tau \\ &= \frac{1}{\sqrt{2\pi}} \int f(\tau) e^{-jc\tau} d\tau \\ &= F[f(\tau)]. \end{aligned} \quad 2.25$$

Por lo tanto, la transformada de escala de la función f_l es igual a la transformada de Fourier de la función f .

2.4.3 Relación entre la transformada de escala y la transformada de Mellin

De la definición de la transformada de escala de una función $f(t)$ se tiene que

$$\begin{aligned}
D_f(c) &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) e^{(-jc-1/2)\ln t} dt \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) t^{-jc-1/2} dt \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) t^{p-1} dt; \quad p = -jc + 1/2 \\
&= \frac{1}{\sqrt{2\pi}} \mathfrak{M}_f(p).
\end{aligned} \tag{2.26}$$

Aquí D_f es una restricción particular de $\mathfrak{M}_f$ en la línea vertical de la variable compleja $p = -jc + 1/2$, con $c \in \mathbb{C}$. La transformada de escala es un caso especial de la transformada de Mellin.

La relación entre las transformadas de escala y de Mellin es (Mouriño-Pérez, 1999):

$$\text{Transf. de escala} \left[g(x, y) = \frac{f(x, y)}{\sqrt{xy}} \right] \equiv \text{Trans. de Mellin}[f(x, y)]$$

$$\text{Transf. de Mellin} [\sqrt{xy}f(x, y)] \equiv \text{Trans. de escala} [f(x, y)]$$

Vía la transformada de Mellin podemos calcular la transformada de escala y viceversa.

§ 2.5 Transformada discreta de Fourier y su inversa en 2-D

La contraparte de la transformada de Fourier de funciones continuas, en funciones discretas de entrada se define como la *transformada discreta de Fourier* (DFT)⁷ que para el caso bidimensional está dada por

⁷ DFT: Discrete Fourier Transform.

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)} \quad 2.27$$

para valores de $u=0,1,2,\dots,M-1$ y $v=0,1,2,\dots,N-1$.

Y la *transformada inversa discreta de Fourier* se define como

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)} \quad 2.28$$

para valores de $x=0,1,2,\dots,M-1$ y $y=0,1,2,\dots,N-1$, donde f es una imagen digital de entrada en el dominio espacial de tamaño $M \times N$, F es la transformada de Fourier de la imagen, (x, y) son las variables espaciales discretas y (u, v) son las variables de frecuencia discretas, respectivamente. Ambas variables discretas, espaciales y frecuenciales, toman valores enteros. A estas dos ecuaciones se les llama *par de transformadas discretas de Fourier* en 2-D. Cuando la transformada discreta de Fourier se implementa en una computadora, los límites de las sumatorias van desde $u=1$ hasta $u=M$ y de $v=1$ hasta $v=N$. El centro de la transformada estará entonces en $u=(M/2)+1$ y $v=(N/2)+1$. El valor de la transformada en $(u,v)=(0,0)$ será entonces, de acuerdo con la ecuación 2.29

$$F(0,0) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \quad 2.29$$

que es el promedio de $f(x,y)$. Esto significa que si $f(x,y)$ es una imagen, el valor de la transformada de Fourier en el origen es igual al nivel de grises promedio de la imagen. Generalmente a $F(0,0)$ se le llama la componente *dc* del espectro.

El número de operaciones aritméticas requeridas para realizar el cálculo con la DFT es del orden de $(MN)^2$. Pero debido a las propiedades de la transformada, es posible reducir el número de operaciones al orden de $(MN \log_2 MN)$ mediante la aplicación del algoritmo de la Transformada Rápida de Fourier (FFT) desarrollada por Cooley y Tukey (1965).

§ 2.6 Correlación discreta

Así también, la contraparte de la correlación de funciones continuas, en funciones discretas, de entrada se define para el caso bidimensional por

$$f(x, y) \otimes g(x, y) = \frac{1}{MN} \sum_{m=0}^M \sum_{n=0}^N f(m, n) g(x + m, y + n). \quad 2.30$$

Esta expresión se puede escribir como:

$$f(x, y) \otimes g(x, y) \Leftrightarrow F(u, v) G^*(u, v) \quad 2.31$$

donde la doble flecha se usa aquí para indicar que la expresión en la derecha es obtenida tomando la transformada de Fourier de la expresión en la izquierda, mientras que la expresión de la izquierda es obtenida tomando la transformada inversa de Fourier de la expresión en la derecha. Es decir,

$$\mathfrak{F}\{f(x, y) \otimes g(x, y)\} = F(u, v) G^*(u, v) \quad 2.32$$

$$\mathfrak{F}\{f(x, y) g^*(x, y)\} = F(u, v) \otimes G(u, v) \quad 2.33$$

Por lo tanto, la correlación se puede calcular por:

$$f(x, y) \otimes g(x, y) = \mathfrak{F}^{-1}\{F(u, v) G^*(u, v)\} \quad 2.34$$

donde * indica el complejo conjugado.

Capítulo 3

Herramientas de caracterización y evaluación: ruido, modelos de imagen, filtros y métricas

Las proposiciones matemáticas, en cuanto tienen que ver con la realidad, no son ciertas; y en cuanto que son ciertas, no tienen nada que ver con la realidad

Albert Einstein

En el reconocimiento de patrones mediante correlación, los modelos de imagen más importantes son: el aditivo, el multiplicativo y el disjunto (Javidi y Horner, 1994). En este trabajo se utilizaron los dos primeros, los cuales fueron generados y aplicados a diferentes tipos de imágenes digitales. Con ellos se analizó y cuantificó, mediante métricas, el desempeño de algunos filtros utilizados, así como su poder discriminativo de las imágenes inmersas en ruido. Se presenta una breve descripción de los diferentes ruidos aquí tratados.

§ 3.1 Naturaleza del ruido

El ruido es un fenómeno perturbativo no deseado de una señal deseada, que es capaz de alterar y producir cambios, distorsiones e interferir en el verdadero significado de una señal de comunicación humana o electrónica, o bloquearla totalmente. En el procesamiento de imágenes digitales, el ruido puede considerarse como una degradación o información indeseable que aparece durante el proceso de registro de una imagen. Es un proceso estadístico con datos o información sin ningún significado cuando se transmiten, de alguna forma, a la imagen deseada. Se generan en el momento de la captura de la imagen de un objeto o escena visual, ya sea por las variaciones del medio ambiente y/o por las

variaciones en la sensibilidad del sensor detector que convierte una imagen óptica en una señal eléctrica continua. El ruido en imágenes digitales aparece comúnmente en los dispositivos de carga acoplada (CCD), por la radiación térmica, la radiación electromagnética, etc. Ya que en la práctica las imágenes tienen ruido, existen una gran cantidad de paquetes de software computacional, como el MATLAB R2008a que se usó en este trabajo, que se utilizan para añadir artificialmente ruido a una imagen. El corromper deliberadamente a una imagen con ruido, nos permite probar la resistencia al ruido usando programas de procesamiento de imágenes digitales, y evaluar el desempeño de los filtros.

Ejemplos típicos de ruido son: el ruido blanco producido por los sensores, ruido de cuantización, ruido de compresión de imágenes, ruido de manchas o moteado comúnmente llamado “speckle”. Este último es causado por las fluctuaciones turbulentas de la atmósfera terrestre generadas por las variaciones estadísticas de la temperatura, las cuales provocan inhomogeneidad aleatoria en el índice de refracción. Esto es característico en la captura de imágenes de objetos estelares mediante telescopios, donde los frentes de onda al propagarse a través de la atmósfera, son distorsionados debido a las fluctuaciones aleatorias del índice de refracción y la resolución angular del objeto estelar (~ 1 segundo de arco) impuesto por la turbulencia atmosférica y no por el límite teórico de difracción del telescopio.

§ 3.2 Ruido blanco

El ruido blanco es una señal aleatoria con una densidad espectral de potencia plana. En otras palabras, la densidad espectral de potencia de la señal tiene igual valor en cualquier banda, en cualquier frecuencia centrada, teniendo un ancho de banda dado. El ruido blanco es considerado análogo a la luz blanca que contiene todas las frecuencias.

Para un ancho de banda infinito, la señal con ruido blanco es una construcción puramente teórica. Teniendo variancia igual en todas las frecuencias, la variancia total de tal señal es infinita. En la práctica, una señal puede ser “blanca” con un espectro plano sobre una banda de frecuencia definida.

El término ruido blanco es también comúnmente aplicado a una señal de ruido en el dominio espacial que tiene una autocorrelación que puede ser representada por una función delta. La señal tiene entonces ruido “blanco” en el dominio de frecuencia espacial (esto es igualmente cierto para señales en el dominio de frecuencias angulares).

Aún una señal binaria que solo puede tomar valores de 1 ó 0, será blanca si la secuencia de ceros y unos está estadísticamente no correlacionada. El ruido que tiene una distribución continua, tal como una distribución normal, puede por supuesto ser blanco.

Generalmente es incorrecto asumir que el ruido gaussiano (i.e., ruido con una distribución de amplitud gaussiana) es necesariamente ruido blanco. La gaussianidad se refiere a la probabilidad de que la señal tenga un cierto valor en un cierto instante, mientras que el término ‘blanco’ se refiere a la manera en que la potencia de la señal (tomada sobre el tiempo) está distribuida entre las frecuencias.

Podemos por lo tanto, encontrar un ruido blanco gaussiano, pero también ruidos blancos de Poisson, de Cauchy, etc. Así, las dos palabras “gaussiano” y “blanco” están generalmente, ambas, especificadas en modelos matemáticos de sistemas. El ruido gaussiano es una buena aproximación de muchas situaciones del mundo real y genera modelos matemáticamente manejables.

3.2.1 Definición matemática de ruido blanco

3.2.1.1 Vector aleatorio blanco

Un vector aleatorio $\mathbf{w}$ es un vector aleatorio blanco si y sólo si su vector medio y la matriz de autocorrelación son las siguientes⁸ (Yates, 2009):

$$\mu_{\mathbf{w}} = \mathbb{E}\{\mathbf{w}\} = 0 \quad 3.1$$

$$R_{\mathbf{w}\mathbf{w}} = \mathbb{E}\{\mathbf{w}\mathbf{w}^T\} = \sigma^2 \mathbf{I} \quad 3.2$$

⁸ http://en.wikipedia.org/wiki/white_noise

Esto es, un vector aleatorio con media cero y su matriz de autocorrelación es un múltiplo de la matriz identidad.

3.2.1.2 Proceso aleatorio blanco

Un proceso aleatorio temporal continuo $w(\eta)$, con $\eta \in \mathbb{R}$, es un proceso de ruido blanco si y sólo si su función media y su función de autocorrelación satisfacen lo siguiente⁹ (Yates, 2009):

$$\mu_w(\eta) = \mathbb{E}\{w(\eta)\} = 0 \quad 3.3$$

$$R_{ww}(\eta_1, \eta_2) = \mathbb{E}\{w(\eta_1)w(\eta_2)\} = \left(N_0/2\right) \delta(\eta_1 - \eta_2) \quad 3.4$$

i.e., es un proceso de media cero para todo η y tiene una potencia infinita en un η igual a cero por lo que su función de autocorrelación es la función delta de Dirac. Esta función de autocorrelación implica que la siguiente densidad de espectro de potencia es:

$$S_{xx}(w) = N_0/2 \quad 3.5$$

ya que la transformada de Fourier de la función delta es igual a 1. Esta densidad espectral de potencia es la misma para todas las frecuencias, la llamamos *blanca* en analogía al espectro de frecuencias de la luz blanca.

3.2.2 Ruido gaussiano blanco aditivo

En comunicaciones, el modelo de ruido gaussiano blanco aditivo (AWGN)¹⁰ es uno en el cual se tiene solo adición lineal de ancho de banda o ruido blanco con una densidad espectral constante y una distribución gaussiana de amplitud. El modelo no toma en cuenta los fenómenos de coloración, selectividad de frecuencias, interferencia, no linealidad y dispersión. Sin embargo, produce modelos matemáticos simples.

⁹ http://en.wikipedia.org/wiki/white_noise.

¹⁰ AWGN: Additive White Gaussian Noise.

El ruido gaussiano proviene de muchas fuentes naturales, tales como las vibraciones térmicas de los átomos en antenas (referidas como ruido térmico o ruido de Johnson-Nyquist), ruido de disparo, radiación de cuerpo negro.

§ 3.3 Funciones de densidad de probabilidad de algunos ruidos

3.3.1 Ruido gaussiano

El ruido gaussiano, también llamado *normal*, tiene la siguiente función de densidad de probabilidad

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/2\sigma^2}, \quad z \in \mathbb{R} \quad 3.6$$

donde z es la variable aleatoria y representa el nivel de gris, μ es la media del valor promedio de z , σ es la desviación estándar, y σ^2 es la variancia de z . La gráfica de esta función se muestra en la Figura 3.1, donde para el intervalo $(\mu - \sigma, \mu + \sigma)$ se tiene el 68% de su valor y 95% para el intervalo $(\mu - 2\sigma, \mu + 2\sigma)$.

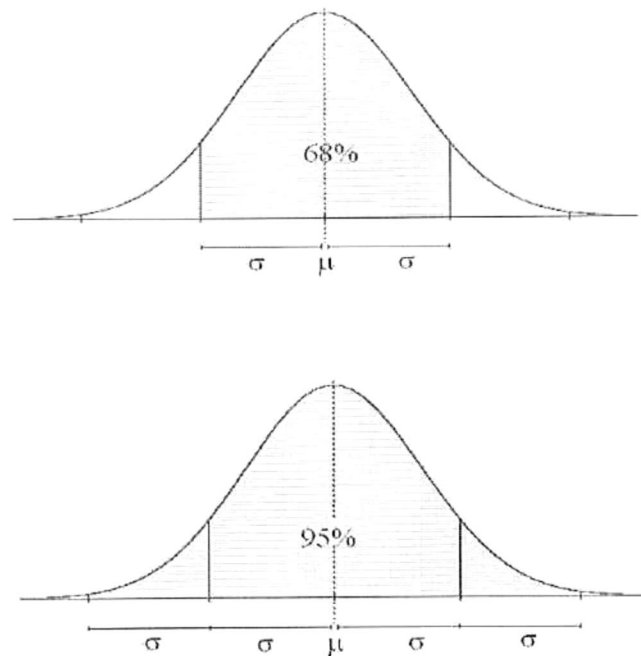


Figura 3.1. Gráfica de la función de densidad de probabilidad gaussiana $p(z)$.

3.3.2 Ruido sal y pimienta

El ruido sal y pimienta, también llamado *impulsivo* o *bipolar*, tiene la siguiente función de densidad de probabilidad

$$p(z) = \begin{cases} P_a & \text{para } z = a \\ P_b & \text{para } z = b \\ 0 & \text{otra parte} \end{cases} \quad 3.7$$

Si $a < b$, aparecerá un nivel de gris b como un punto luminoso en la imagen. En caso contrario, el nivel a aparecerá como un punto oscuro. Si P_a o P_b es cero, el ruido impulsivo se llama *unipolar*. Si ninguno de ellos tiene probabilidad cero, y especialmente si son aproximadamente iguales, los valores del ruido impulsivo se parecerán a los granos de sal y pimienta aleatoriamente distribuidos sobre la imagen. El ruido impulsivo puede estar compuesto de valores positivos o negativos, es decir, si los impulsos son negativos, en la imagen aparecen puntos negros (pimienta) y si son positivos los puntos que aparecen son blancos (sal). Así, por ejemplo, para una imagen de 8-bits se tiene $a=0$ (negro) y $b=255$ (blanco). Por esta razón, al ruido de impulso bipolar se le llama ruido *sal y pimienta*. También se usan los términos ruido de *disparo* o ruido *pico* para referirse a este tipo de ruido. En este trabajo se escoge el término ruido *sal y pimienta*.

§ 3.4 Modelo aditivo

Este modelo es útil para describir un conjunto de objetos que son capturados con una cámara. La intensidad luminosa reflejada por el objeto incide sobre el sensor de la cámara, generando así un flujo aleatorio de electrones cuyo valor esperado es proporcional a la intensidad de luz incidente. Esto significa que la imagen generada va acompañada de ruido aditivo causado por el sensor. La representación de este modelo queda expresado como

$$f_a(x, y) = s(x, y) + n_a(x, y) \quad 3.8$$

donde $s(x, y)$ es la imagen de un objeto, $n_a(x, y)$ es el ruido del sensor y $f_a(x, y)$ es la imagen ruidosa resultante. Un ejemplo de este modelo está representado en la Figura 3.2.



Figura 3.2. Ejemplo gráfico del modelo de ruido aditivo.

§ 3.5 Modelo multiplicativo

Este modelo se utiliza principalmente para representar matemáticamente algunas distorsiones de la imagen que afectan directamente los valores de los píxeles, como por ejemplo, la iluminación. Este modelo se expresa como

$$f_m(x, y) = s(x, y)n_m(x, y) \quad 3.9$$

donde $f_m(x, y)$ es la imagen de un objeto, $n_m(x, y)$ es el ruido multiplicativo y $f_m(x, y)$ es la imagen ruidosa resultante. Un ejemplo de este modelo está representado en la Figura 3.3.



Figura 3.3. Ejemplo gráfico del modelo de ruido multiplicativo.

§ 3.6 Modelo disjunto

Este modelo describe más adecuadamente las imágenes capturadas del mundo real. Consideremos el caso donde una cámara registra la imagen de un objeto en cuestión, en condiciones arbitrarias. La imagen resultante estará formada por la imagen del objeto más la porción del fondo en el que el objeto está inmerso. Este modelo se expresa como

$$f_d(x, y) = s(x, y) + w(x, y)b(x, y), \quad 3.10$$

donde $s(x, y)$ es la imagen de un objeto, $b(x, y)$ es la imagen de fondo, $w(x, y)$ es una función binaria que es igual a cero si está dentro del área de $s(x, y)$ e igual a uno en otra parte, y $f_d(x, y)$ es la imagen resultante.

§ 3.7 Filtros

3.7.1 Filtro clásico acoplado (CMF)¹¹

Este filtro, también conocido como filtro complejo u holográfico, fue propuesto y probado por Vander Lugt (1964)¹², al cual lo describió como una función valuada compleja de la forma:

$$H(u, v)_{CMF} = |H(u, v)|e^{j\phi(u, v)} \quad 3.11$$

donde (u, v) son las variables de frecuencia, $|H(u, v)|$ es la respuesta de amplitud o módulo de la transformada de Fourier del objeto y $e^{j\phi(u, v)}$ es la respuesta de fase, $\phi(u, v)$ es la fase de la transformada de Fourier del objeto. Como los filtros espaciales son elementos pasivos, es decir, no amplifican la distribución de luz; por lo tanto, $|H(u, v)| \leq 1$. En

¹¹ CMF: Classical Matched Filter.

¹² Históricamente, este filtro fue precedido por otro filtro, una técnica menos versátil, llamado *hard-clipped filter* el cual fue generado por computadora y es el primer ejemplo de lo que ahora podría llamarse *filtro sólo de fase*. Fue usado en el procesamiento de señales de radar hasta cerca de 1961, y debido a que fue un trabajo clasificado no se dio a conocer en la literatura científica hasta 1965 (Goodman, 2005).

consecuencia, los valores posibles del filtro complejo deben estar en o dentro de una circunferencia de círculo unitario centrado en el origen de un plano complejo (Vander Lugt, 1968). Esto significa que: (a) para un filtro puramente absorbente, la región está limitada en el eje real positivo entre 0 y 1, (b) para un filtro absorbente y un control de fase binario, la región se extiende desde -1 hasta 1 en el eje real, (c) para un filtro puro de fase, se tienen valores solo en el círculo unitario, y (d) para un filtro que llega a tener distribuciones arbitrarias de absorción y control de fase, la región cubre entonces el círculo unitario completo, que es la región que uno generalmente desea alcanzar. Algunas de las limitaciones de este filtro clásico son, un pico de correlación amplio y sensibilidad a distorsiones, y debido a las consideraciones tomadas en su diseño, presenta un funcionamiento inadecuado para las imágenes que contienen al objetivo espacialmente separado del fondo (Javidi y Horner, 1994). Este filtro se deriva de la maximización de la razón señal-ruido convencional (SNR, definida más adelante en 3.8.3) (Vijaya Kumar y Hassebrook, 1990).

3.7.2 Filtro sólo de fase (POF)¹³

Este filtro se expresa como (Horner y Gianino, 1984)

$$H_{POF}(u, v) = e^{j\phi(u,v)} \quad 3.12$$

La respuesta en amplitud, para este filtro sólo de fase, es unitaria. Es decir, solo toma valores en el círculo unitario del plano complejo, tal como se mostró anteriormente en el análisis que se hizo del CMF. Con el POF se obtienen picos de correlación más estrechos y es más selectivo que el CMF.

¹³ POF: Phase Only Filter.

3.7.3 Filtro no lineal (NLF)¹⁴ (ley k)

Un filtro no lineal es un filtro cuya salida es una función no lineal de la entrada. Por definición, cualquier filtro que no es un filtro lineal, es un filtro no lineal (Marion, 1991). La expresión para el filtro no lineal (ley k) está dada por (Vijaya Kumar y Hassebrook, 1990)

$$H_{NLF}(u, v) = |S(u, v)|^k e^{j\phi(u, v)} \quad 3.13$$

donde, $|S(u, v)|$ es el módulo de la transformada de Fourier del objeto a reconocer, k es el factor de no linealidad que toma valores $0 < k < 1$. Al variar los valores de k en este intervalo, podemos manipular las capacidades discriminantes del procesador no lineal y determinar el mejor valor k de los filtros no lineales. Para valores de k distintos a este intervalo, tenemos los filtros lineales; $k=1$ (filtro acoplado clásico), $k=0$ (filtro sólo de fase) y $k=-1$ (filtro inverso).

Una de las principales razones por las cuales se prefiere utilizar los filtros no lineales en vez de filtros lineales, es que estos últimos pueden ser muy sensibles a pequeños cambios de la señal de entrada. Con los filtros lineales, el plano de correlación resultante presenta lóbulos laterales (valores altos fuera del área del objeto), con lo cual, la probabilidad de las falsas alarmas se incrementa volviendo a este filtro poco confiable para aplicarse al reconocimiento de imágenes. El filtro no lineal tiene ventajas comparado con los filtros clásicos acoplados, los filtros sólo de fase (Horner y Gianino, 1984) y otros filtros lineales, debido a su gran capacidad para discriminar objetos, el máximo pico de correlación está bien localizado y el plano de salida es menos ruidoso.

¹⁴ NLF: Non Linear Filter.

§ 3.8 Métricas de desempeño

Los criterios de desempeño nos permiten evaluar la eficiencia de los filtros utilizados en el reconocimiento de patrones. Actualmente existen en la literatura científica una gran cantidad de métricas para medir el desempeño de los filtros de correlación (Vijaya Kumar y Hassebrook, 1990). De las métricas más utilizadas podemos mencionar las siguientes: la relación energía del pico de correlación (PCE), la relación de la capacidad de discriminación (DC), la relación señal a ruido (SNR), y la relación de eficiencia de luz (LE), entre otras. En este trabajo se utilizaron las métricas PCE y DC.

3.8.1 Energía pico de correlación (PCE)¹⁵

Esta métrica es una relación de la energía en el origen, la cual es medida en el plano de correlación, y la energía total que existe en el plano de correlación y está definida como (Javidi y Horner, 1994)

$$PCE = \frac{|\mathbb{E}\{c(0,0)\}|^2}{\mathbb{E}\{|c(x,y)|^2\}} , \quad 3.14$$

donde el numerador es el valor esperado de la intensidad del pico de correlación y el denominador es el valor esperado de la energía promedio en el plano de correlación.

3.8.2 Capacidad de discriminación (DC)¹⁶

Esta métrica es una relación que mide la capacidad de un filtro para rechazar cualquier objeto diferente al objeto de referencia, la habilidad de un filtro para distinguir un objeto de

¹⁵ PCE: Peak-to-Correlation Energy.

¹⁶ DC: Discrimination Capability.

entre otros objetos diferentes, cuando un objeto se encuentra incrustado en un fondo así como otros objetos falsos. La DC se define como (Vijaya Kumar y Hassebrook, 1990)

$$DC = 1 - \frac{|C_B(0,0)|^2}{|C_T(0,0)|^2}, \quad 3.15$$

donde C_B es el valor máximo en el plano de correlación sobre el área del fondo o de los objetos no deseados que se desean rechazar, y C_T es el valor máximo en el plano de correlación sobre el área de la posición del objeto de referencia. El valor máximo de esta medida es la unidad y cuando se tienen valores negativos significa que al probar el filtro el reconocimiento no tuvo éxito, es decir, el objeto no puede ser reconocido por el filtro.

3.8.3 Relación señal a ruido (SNR)¹⁷

Esta relación es una medida de la sensibilidad del filtro al ruido que nos permite caracterizar el ruido presente en el pico de correlación, y se define como (Vijaya Kumar y Hassebrook, 1990)

$$SNR = \frac{|E\{C(x,y)\}|^2}{var\{C(x,y)\}}, \quad 3.16$$

donde el numerador representa la intensidad del valor esperado del pico de correlación generado por el objeto deseado, y el denominador es la variancia del pico de correlación ocasionado por el ruido.

¹⁷ SNR: Signal-Noise Ratio.

3.8.4 Eficiencia luminosa (LE)¹⁸

Esta métrica es importante cuando el reconocimiento de patrones se realiza con procesadores ópticos. La eficiencia luminosa (Horner, 1982), es una medida que relaciona la cantidad de luz a la salida de un sistema respecto a la luz de entrada, es decir, la LE nos dice qué tanto se aprovecha la luz en un sistema óptico, y está definida como

$$\eta = \frac{I_{salida}}{I_{entrada}} = \iint_{-\infty}^{\infty} \frac{|S(u, v)|^2 |H(u, v)|^2}{|S(u, v)|^2} du dv, \quad 3.17$$

donde I es la intensidad de luz, $|S(u, v)|$ y $|H(u, v)|$ son los módulos de las transformadas de Fourier de la imagen de entrada y la función de transferencia del filtro, respectivamente.

¹⁸ LE: Light Efficiency.

Capítulo 4

Metodología de la correlación digital no lineal invariante

*La experiencia es la única fuente de la verdad;
sólo ella puede enseñarnos algo nuevo, sólo
ella puede darnos la certeza*

Henri Poincaré

En este capítulo se presenta la metodología desarrollada para realizar la correlación digital no lineal invariante a posición, rotación y escala que se realiza mediante las operaciones de la transformada de Fourier, el mapeo de coordenadas cartesianas a polares y la transformada de escala, respectivamente, para el reconocimiento de patrones de algunos objetos de imágenes creadas artificialmente con algoritmos computacionales e imágenes de objetos de la vida real. Se analiza el desempeño de los filtros no lineales y se comparan con los filtros lineales.

Desde la invención del láser en 1960 y la primera aparición de los filtros espaciales complejos u holográficos del tipo de Vander Lugt en 1964, se lograron avances sustanciales en el campo del reconocimiento de patrones ópticos, en particular con los métodos de correlación. Los esfuerzos por lograr la realización de un dispositivo óptico para la identificación y conteo de los objetos, no fueron satisfactorios debido principalmente a que se utilizaba la correlación óptica convencional, la cual tiene el inconveniente de ser muy sensible a pequeños cambios de escala y orientación de los objetos. Para vencer estos obstáculos, Casasent y Psaltis (1976a) propusieron el uso de la transformada de Mellin, que

por sus propias características es invariante a escala. Los resultados de sus investigaciones confirmaron que es posible correlacionar ópticamente dos funciones u objetos que difieran en escala y rotación. Schwartz (1977) mostró en sus estudios cómo el cerebro tiene la habilidad de reconocer las formas de los objetos. La luz del mundo exterior que incide sobre la retina, la cual es una superficie curvada compuesta por células fotosensibles, pasa a través del nervio óptico en forma de impulsos eléctricos, de tal manera que la imagen en la retina es mapeada sobre una superficie del cerebro y descrita por una transformación geométrica del espacio de coordenadas cartesianas a coordenadas polares con escalamiento logarítmico. Muestra como el mapeo de la escena visual a través del logaritmo complejo es importante para el reconocimiento de patrones y hace referencia a la solución propuesta en los trabajos de Casasent y Psaltis (1976b) de utilizar la transformación de coordenadas de la función de entrada u objeto al plano polar con escalamiento logarítmico para poder realizar la operación de correlación. Podríamos decir que ellos fueron los pioneros y su método fue clave, la *piedra angular*, en el posterior desarrollo de sistemas de reconocimiento de patrones, tanto ópticos, digitales e híbridos.

Este trabajo tiene como punto de partida el trabajo publicado por Pech-Pacheco *et al* (2003) en el cual se desarrolla un sistema de correlación invariante a posición, rotación y escala utilizando un filtro sólo de fase. Continuando en esta dirección, en esta tesis se obtienen algunos resultados con nuevos algoritmos computacionales para el reconocimiento de objetos. Se eligió la transformada de escala (Cohen, 1995) en vez de la bien conocida transformada Mellin-Fourier (Casasent y Psaltis, 1976a,b,c), porque es más sensible a los cambios de escala. El filtro no lineal (ley k) se introduce aquí para ser utilizado en la correlación digital invariante, la cual nos da información de la similitud entre diferentes objetos. Este tipo de filtro tiene ventajas comparado con los filtros clásicos acoplados (Vander Lugt, 1964), los filtros sólo de fase (Horner y Gianino, 1984) y otros filtros lineales, debido a su gran capacidad para discriminar objetos, el máximo pico de correlación está bien localizado y el plano de salida es menos ruidoso. En otros trabajos se han utilizado sistemas ópticos de correlación (Javidi, 1989a,b, 1990b), donde se usaron las características no lineales para generar un filtro no lineal (Javidi, 1989c) con un Modulador de Luz Espacial (SLM) localizado en el plano de Fourier del correlador óptico. Los

resultados de esos trabajos mostraron que los filtros no lineales tuvieron un mejor desempeño en la correlación comparados con los filtros lineales (Javidi, 1990b). En el trabajo de esta tesis, se utilizó el filtro no lineal, llamado también filtro no lineal k -ésima ley, porque se trató con sistemas digitales invariantes y no con sistemas ópticos. Ya que en un sistema óptico sería muy difícil implementar la transformada de escala para obtener las invariancias a rotación y escala.

Para el desarrollo de los algoritmos computacionales y la obtención de las gráficas, se utilizó el MATLAB Version 7.6.0 R2008a de The MathWorks, Inc., y para el análisis de datos el STATISTICA Version 6.1 de StatSoft, Inc.

§ 4.1 Correlación no lineal

La correlación no lineal en el plano de salida con filtros no lineales puede expresarse como (Javidi *et al.*, 1997)

$$c(x, y) = \iint |R(u, v)|^k |S(u, v)|^k e^{j[\phi_S(u, v) - \phi_R(u, v)]} e^{j2\pi(ux + vy)} du dv, \quad 4.1$$

donde $R(u, v)$ es la transformada de Fourier del objeto, $S(u, v)$ es la transformada de Fourier de la escena de entrada a ser reconocida, $\phi_R(u, v)$ y $\phi_S(u, v)$ son las fases de las transformada de Fourier $R(u, v)$ y $S(u, v)$, respectivamente.

§ 4.2 Obtención del filtro no lineal

Los pasos para obtener el filtro no lineal se muestran en la Figura 4.1. Primero, tenemos la imagen original u objetivo denotada por una función bidimensional de la forma $f(x, y)$, diferente de cero y de tamaño finito. El siguiente paso es calcular la transformada rápida de Fourier (FFT) y obtener el módulo de la transformada de Fourier denotada como $|F(u, v)|$, y de esta manera el desplazamiento de la imagen no es afectada en el plano de Fourier. Después de esto, realizamos un filtraje pasa altas usando un filtro parabólico al módulo de la transformada de Fourier. Esta clase de filtros atenúa las bajas frecuencias mientras deja

pasar las altas frecuencias que realzan los detalles de borde (Pech-Pacheco *et al.*, 2003). El siguiente paso es introducir un factor de escala dado por $\sqrt{r}$, donde r es la frecuencia espacial radial, que es indispensable para diferenciar la transformada de escala de la transformada de Mellin. Después, las coordenadas cartesianas son mapeadas a coordenadas polares $[F(r, \theta)]$ para obtener la invariancia a rotación. Y luego se realiza una interpolación bilineal a los primeros datos de la conversión de coordenadas para reducir la fuga de información debido al muestreo log-polar. Se hace un escalamiento logarítmico en la parte radial en coordenadas polares para realizar la invariancia a escala $[F(\exp(\rho), \theta)]$. Tomando la FFT en este paso, obtenemos un filtro invariante a posición, rotación y escala, donde $|T|$ es el módulo de la transformada de Fourier del objeto a ser reconocido, con las invariancias implícitas y φ es la fase de la transformada de Fourier del objeto.

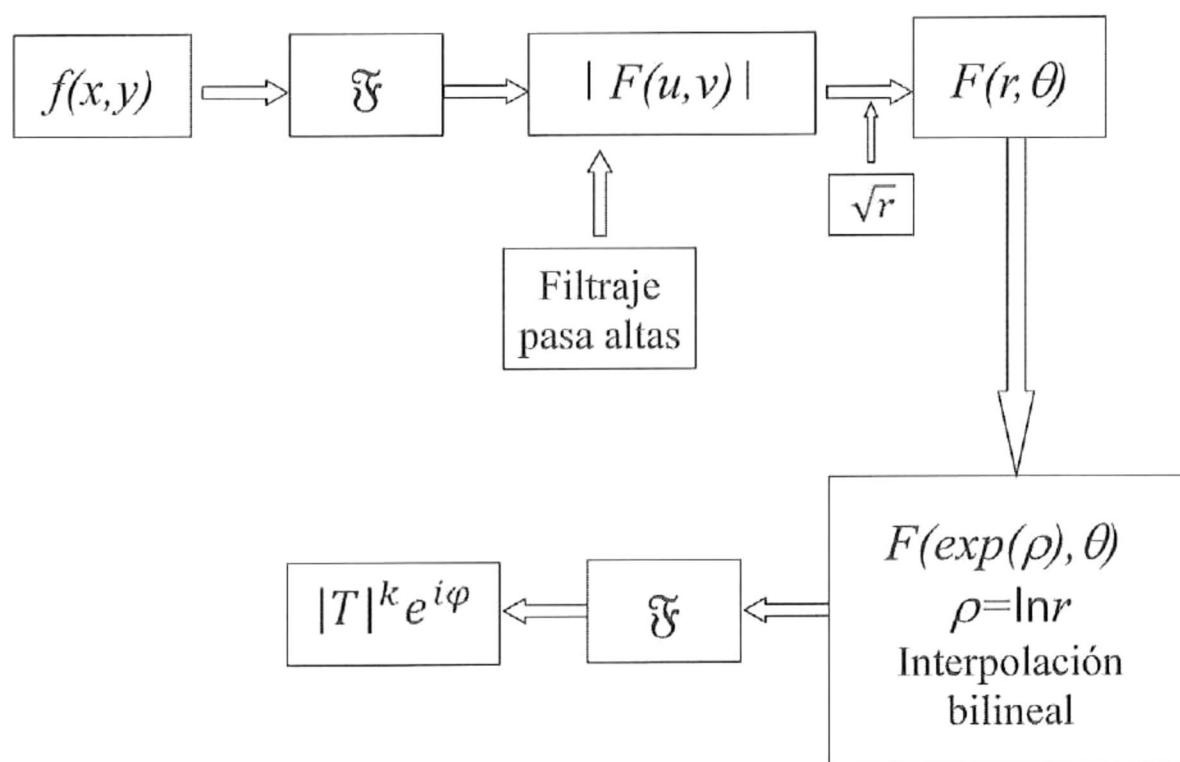


Figura 4.1. Diagrama de bloques para obtener el filtro no lineal.

§ 4.3 Obtención del filtro compuesto no lineal

La Figura 4.2 muestra los pasos para obtener un filtro compuesto no lineal. En este caso, hay un conjunto de n imágenes en la entrada del sistema, representadas como una sucesión de funciones de la forma:

$$\{f_i(x, y)\}_{i=1}^n = f_1(x, y), f_2(x, y), \dots, f_n(x, y). \quad 4.2$$

El siguiente paso es tomar la FFT de cada una de estas imágenes, y como la transformada de Fourier es una integral lineal, tomamos la suma total de las FFT expresada como:

$$|F(u, v)| = \sum_{i=1}^n |F_i(u, v)|. \quad 4.3$$

La suma es entonces procesada de la misma manera como se mencionó en la Figura 4.1.

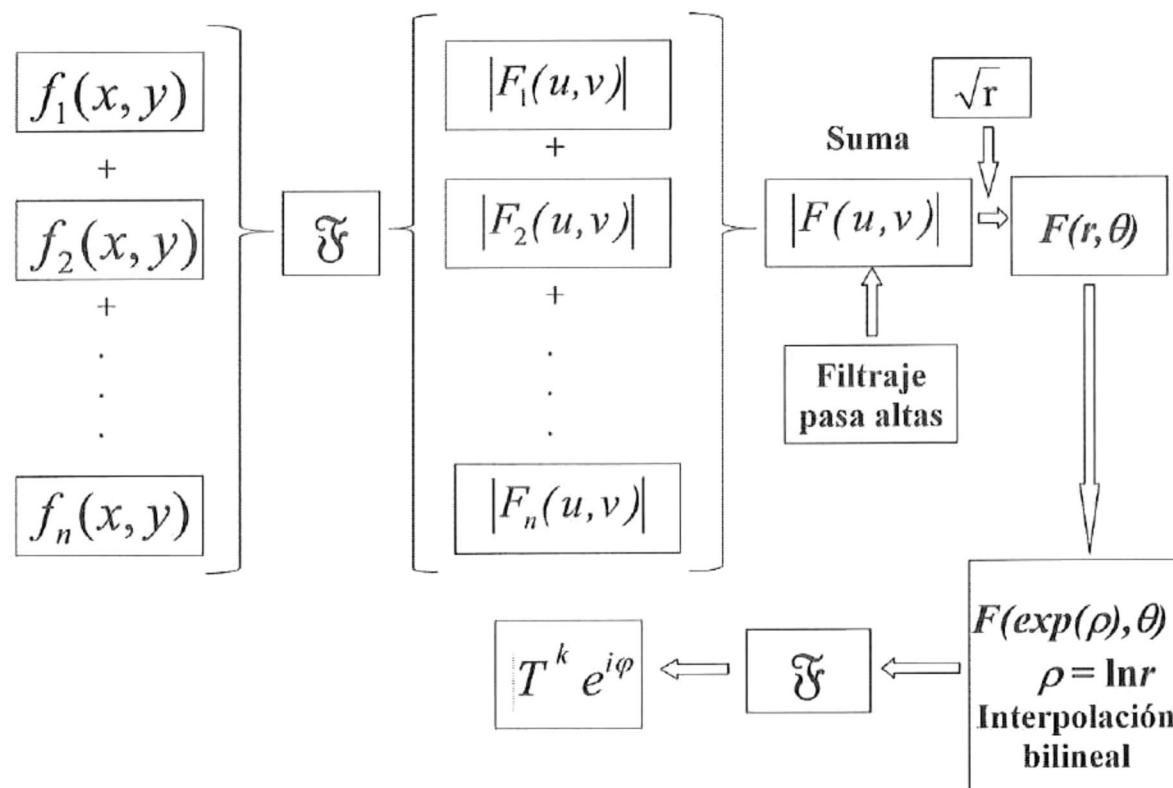


Figura 4.2. Diagrama de bloques para obtener el filtro compuesto no lineal.

En otros trabajos (Guerrero-Moreno y Álvarez-Borrego, 2009) se determinó que el valor máximo de imágenes que puede soportar un filtro compuesto no lineal para un problema específico es de aproximadamente 200, ya que el número de imágenes no puede ser infinito, porque el filtro contendría demasiada información superpuesta y perdería su capacidad de identificación de los objetos.

§ 4.4 Correlación no lineal invariante

Los pasos para realizar la correlación invariante con un filtro no lineal se muestran en la Figura 4.3. En esta figura, desde el paso (1) hasta el paso (9), el procedimiento es el mismo como en las Figuras 4.1 y 4.2. La imagen problema $g(x,y)$, la cual contiene o no al objetivo es la imagen de entrada, paso (1). En el paso (9) se muestra la información no lineal de la imagen problema. Los pasos (9) y (10) muestran el proceso de correlación para obtener la correlación digital invariante, paso (11), a posición, rotación y escala usando un filtro no lineal, donde $|T|$ y $|G|$ son los módulos de las transformadas de Fourier del objeto a ser reconocido y de la imagen problema, respectivamente, las cuales contienen la información no lineal con las invariancias implícitas y donde φ y ϕ son las fases de cada una de las transformadas de Fourier. Así, el factor de no linealidad k afecta a la magnitud de la transformada de Fourier de ambas imágenes, la de entrada y la de referencia.

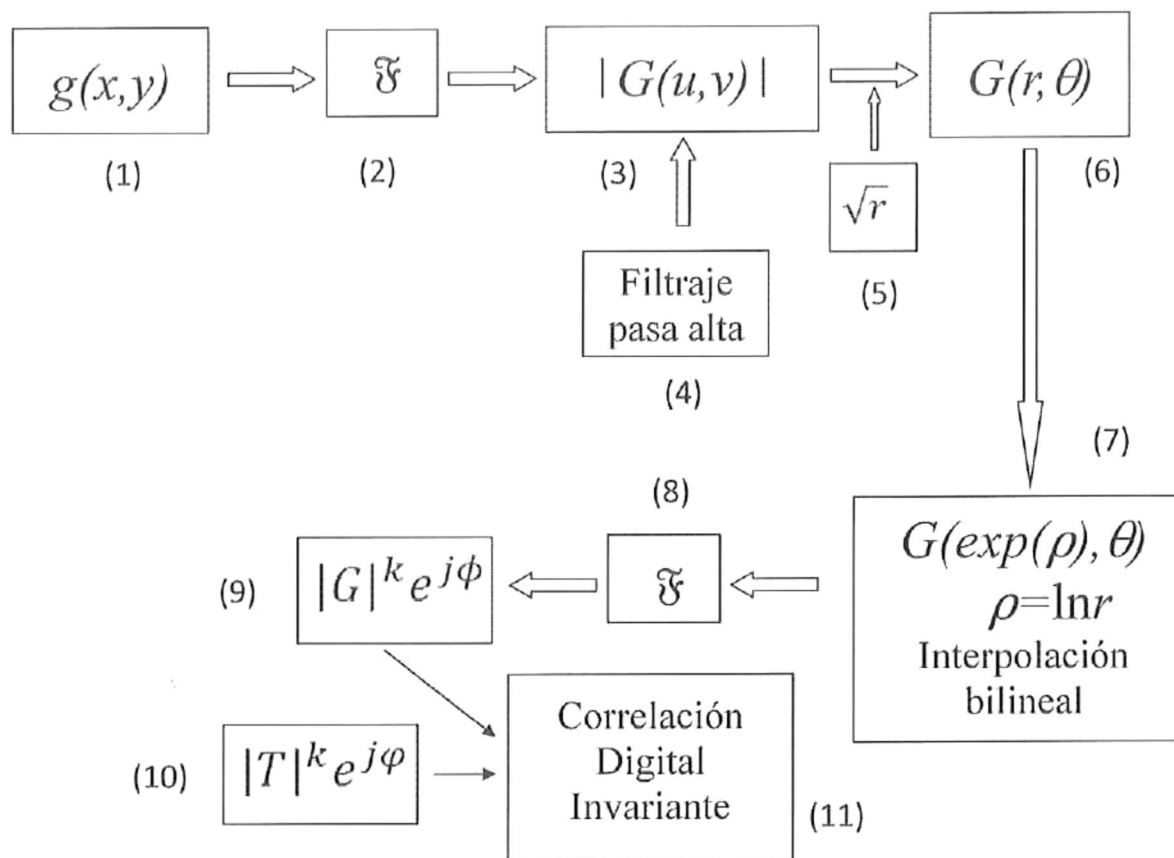


Figura 4.3. Diagrama de bloques representando al sistema de correlación digital invariante.

En la Figura 4.4 se muestran visualmente el módulo de la transformada de Fourier, el mapeo log-polar, el módulo de la transformada de escala y el filtro no lineal con $k=0.3$ para un objeto con rotaciones de 0° , 30° y -30° . De esta figura se puede observar que la transformada de Fourier gira el mismo ángulo de rotación del objeto y en el mapeo log-polar ocurre un desplazamiento en el eje horizontal de rotación para cada ángulo de rotación del objeto. La transformada de escala se mantiene igual para cada rotación del objeto y el filtro no lineal es diferente para cada caso.

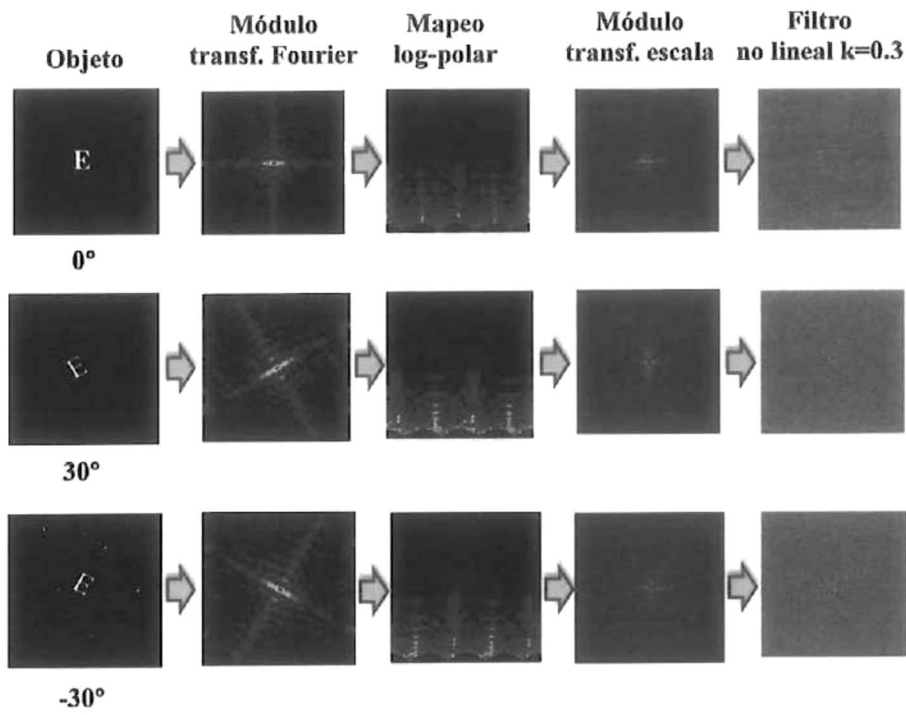


Figura 4.4. Procedimiento visual para obtener un filtro no lineal con $k=0.3$ para el caso de un objeto rotado a 0° , 30° y -30° .

Capítulo 5

Aplicaciones del sistema digital invariante no lineal

El todo es diferente de la suma de sus partes

Wolfgang Köhler

En este capítulo se presentan las contribuciones originales producidas por el trabajo de esta tesis. En total fueron cinco las aportaciones hechas, presentadas cada una por secciones, de la 5.1 a la 5.5, las cuales abarcan investigaciones realizadas aplicando el sistema digital invariante con los filtros lineales y no lineales a diferentes objetos tanto los artificiales como de la vida real. Con los resultados de estos trabajos se probó la eficacia y robustez de los algoritmos computacionales desarrollados en el sistema digital de reconocimiento invariante.

§ 5.1 Filtro no lineal para el reconocimiento de patrones usando la transformada de escala

En esta sección se presenta la aplicación del sistema digital de correlación invariante usando un filtro no lineal (ley k), de la metodología mostrada en la Figura 4.3. Se analizó el desempeño de este filtro con diferentes valores del factor de no linealidad k de acuerdo a la métrica PCE. Primeramente, en este trabajo se realizó un experimento numérico para encontrar el mejor valor de k para rotación y escala para el cual se obtiene el valor mayor del PCE, usando como filtro u objetivo la letra E la cual fue correlacionada con cada una de las letras del alfabeto en fuente Arial. Después se determinó el mejor valor de este factor utilizando imágenes de la vida real y se obtuvo el mismo valor de k del primer experimento.

Una vez obtenido el mejor valor de k para rotación y escala, se utilizó el filtro no lineal con ese valor de k y se aplicó a las 26 letras del alfabeto, donde cada letra es una imagen problema de 256x256 pixeles en tamaño. Los resultados mostraron un mejor desempeño para los filtros no lineales comparados con los filtros lineales. Los incrementos en rotación fueron de un grado y de uno por ciento para la escala, hasta completar una rotación completa de 360° y una variación de escala de $\pm 25\%$. Posteriormente, este método se aplicó a la comparación del desempeño del filtro no lineal con filtros lineales, usando letras en fuente Arial y Times New Roman.

5.1.1 Simulaciones computacionales

5.1.1.1 Determinación del valor óptimo del factor de intensidad de no linealidad k usando como filtro la letra E de fuente Arial

La primer tarea fue encontrar por un experimento estadístico numérico, el mejor valor del factor de no linealidad k para el filtro no lineal. Los resultados fueron mostrados en una gráfica de cajas, *pico de correlación energía* (PCE) *versus* k , una para la rotación y otra para la escala, con el valor de la media con un error estándar ($\pm EE$) y con dos errores estándar ($\pm 2*EE$) para el PCE usando la letra E de fuente Arial como filtro (Figuras 5.1 y 5.2), la cual fue correlacionada con las 26 letras del alfabeto. Para el caso de rotación (Figura 5.1), los incrementos fueron de un grado, desde 0 hasta 359 grados. El número de imágenes procesadas en rotación para cada valor de k fue de 9360. Pero hay 11 diferentes valores de k , esto es, para $k=0, 0.1, 0.2, \dots, 1.0$, se tiene un total de 102960 imágenes procesadas estadísticamente. De la Figura 5.1 se observa que el desempeño del filtro no lineal es mejor para $k=0.3$.

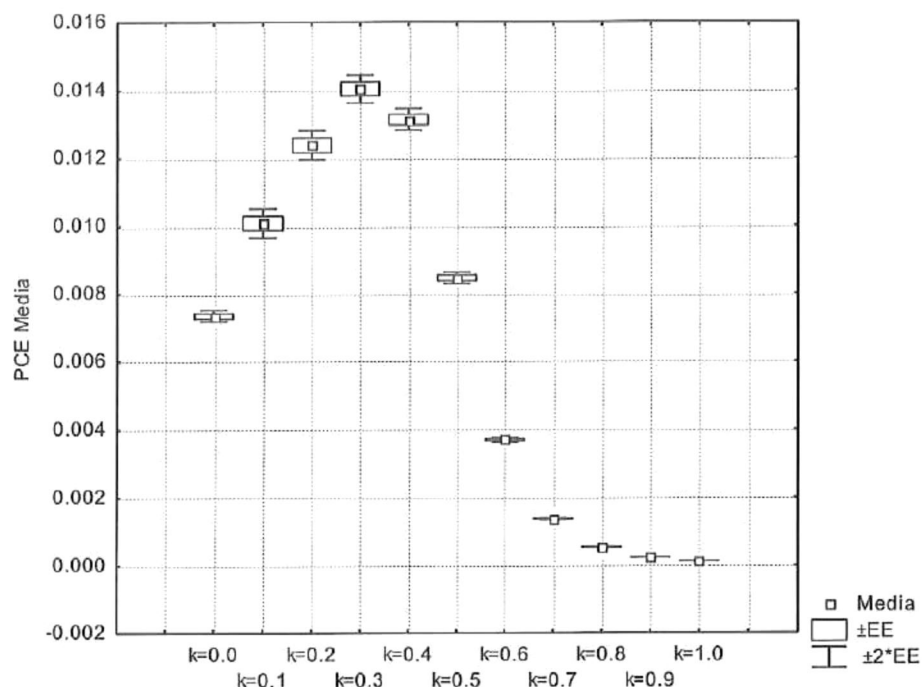


Figura 5.1. Gráfica de cajas para el *pico de correlación energía* PCE versus los valores de k no lineal para la letra E en rotación.

De manera similar, en la Figura 5.2 se muestra una gráfica del *pico de correlación energía* PCE vs los valores de k para el caso de escala, desde $k=0$ hasta $k=1$ para la letra E fuente Arial usada como filtro y correlacionada con las 26 letras escaladas del alfabeto, con incrementos de uno por ciento, desde -25% hasta +25%. Esta gráfica muestra el valor de la media con un error estándar ($\pm EE$) y dos errores estándar ($\pm 2*EE$) para el PCE. El número de imágenes procesadas para cada valor de k fue de 1326. Pero hay 11 diferentes valores k , esto es, para $k=0, 0.1, 0.2, \dots, 1.0$, se tiene un total de 14586 imágenes procesadas estadísticamente. De la Figura 5.2 se observa que el desempeño del filtro no lineal es mejor para $k=0.3$. Para ambos casos, rotación y escala, el resultado para el mejor valor de k fue el mismo, $k=0.3$.

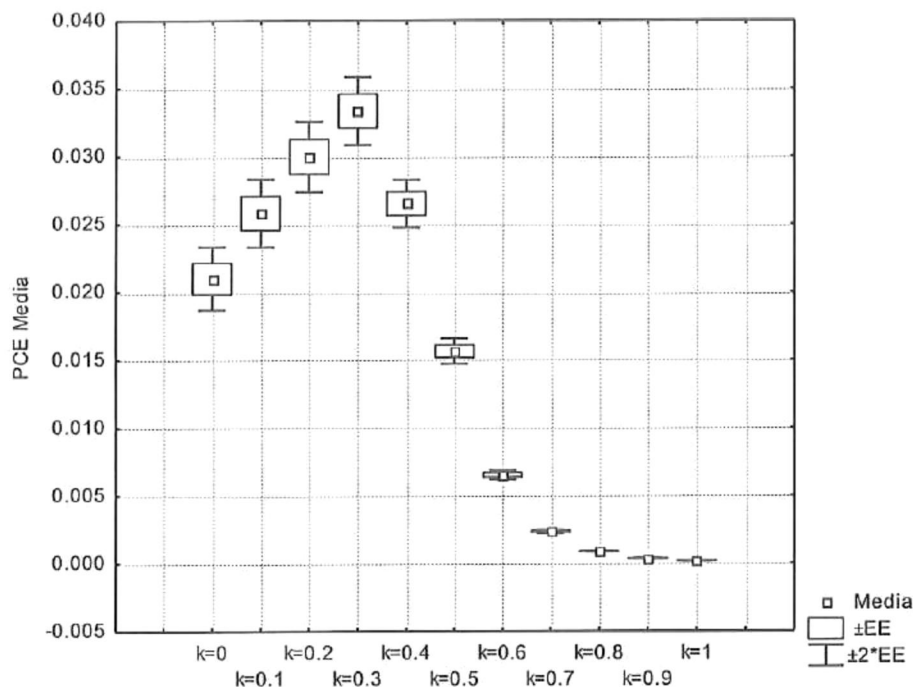


Figura 5.2. Gráfica de cajas para el *pico de correlación energía* PCE versus los valores de k no lineal para la letra E en escala.

5.1.1.2 Determinación del valor óptimo del factor de intensidad de no linealidad k usando imágenes reales

Para este experimento se eligieron varias imágenes que fueran representativas de la vida real. La imagen A, Figura 5.3(a), se tomó como filtro porque contiene mezclas de detalles, de zonas planas y una textura que hace trabajar bien a los diferentes algoritmos de compresión de imagen. Por esta razón la comunidad científica internacional la ha utilizado desde hace aproximadamente cuatro décadas como imagen de prueba para los algoritmos computacionales en el tratamiento de imágenes y se ha convertido así en un estándar en el procesamiento digital de imágenes. Esta imagen de prueba fue correlacionada con cada una de las 15 imágenes problema, Figura 5.3(b), para rotación y escala. Las imágenes son de tipo JPEG, de tamaño 256x256 píxeles y rango dinámico (0,255) niveles de grises. Los resultados se graficaron como se muestran en las Figuras 5.4 y 5.5, PCE versus k , una para la rotación y la otra para la escala, con un valor medio y dos errores estándar ($\pm 2*EE$)

usando la imagen A. En ambos casos, los resultados para el mejor valor de k fue el mismo (Coronel-Beltrán y Álvarez-Borrego, 2008), $k=0.3$.

En la Figura 5.4 se muestra una gráfica de cajas para el PCE *versus* los valores de no linealidad k para la imagen A usada como objetivo o imagen de referencia correlacionada con las 15 imágenes problema rotadas, denotadas por A,B,C,...,O, como se muestra en la Figura 5.3(b), con incrementos de un grado de rotación, desde 0 hasta 179 grados. Esta gráfica muestra el valor medio con un error estándar ($\pm EE$) y dos errores estándar ($\pm 2*EE$) para el PCE. El número de imágenes procesadas para cada valor de k fue de 2700. Pero hay 11 valores diferentes de k , esto es, para $k=0,0.1,0.2,\dots,1$, lo que da un total de 29700 imágenes procesadas estadísticamente. Esto se hizo para determinar el mejor valor de k , el cual fue $k=0.3$.



(a)



A



B



C



D



E



F



G



H



I



J



K



L



M



N



O

(b)

Figura 5.3. (a) la imagen objetivo A, (b) las imágenes problema.

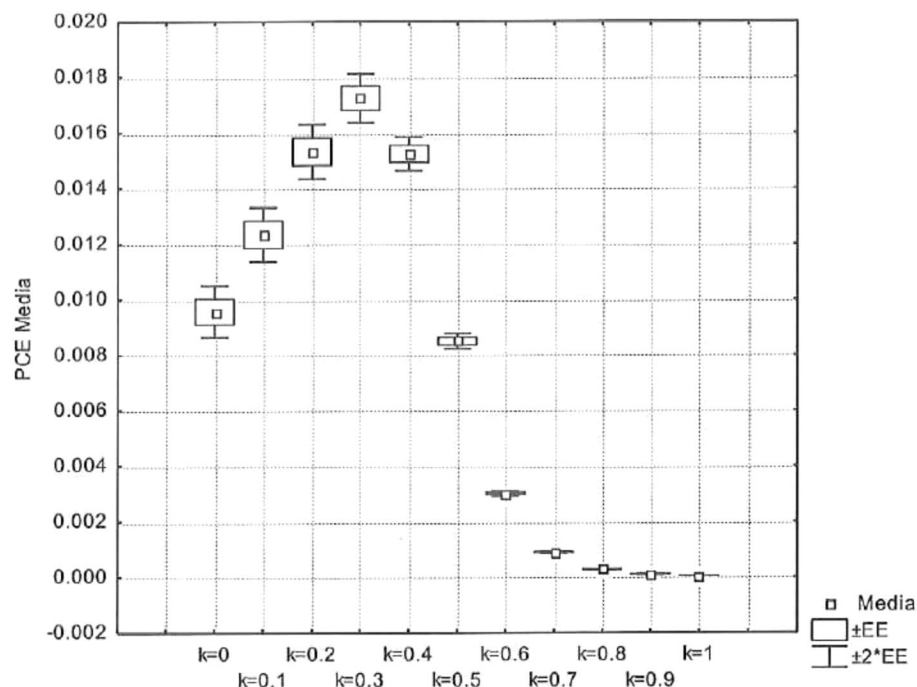


Figura 5.4. Gráfica de cajas para el *pico de correlación de energía* PCE versus los valores de no linealidad k para la imagen A en rotación.

Similarmente, en la Figura 5.5 se muestra una gráfica de cajas para el pico de correlación *versus* los valores de no linealidad k para la imagen A usada como objetivo y correlacionada con las 15 imágenes problema escaladas, con incrementos de 1%, desde 75% hasta 125%. Esta gráfica muestra el valor medio con un error estándar ($\pm EE$) y dos errores estándar ($\pm 2*EE$) para el PCE. El número de imágenes procesadas para cada valor de k fue de 765. Pero hay 11 valores diferentes para k , esto es, para $k=0, 0.1, 0.2, \dots, 1$, lo que da un total de 8415 imágenes procesadas estadísticamente. Esto se hizo para determinar el mejor valor de k , el cual fue $k=0.3$.

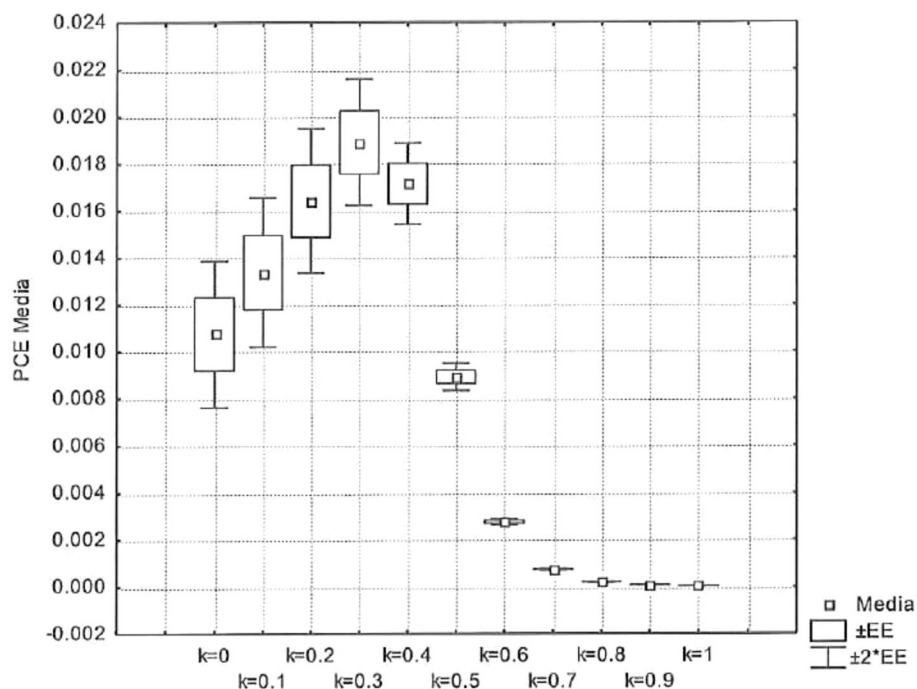


Figura 5.5. Gráfica de cajas para el *pico de correlación energía* PCE versus los valores de no linealidad k para la imagen A en escala.

Las gráficas de las Figuras 5.6 y 5.7 muestran la comparación de los PCE para el filtro E y el filtro imagen A utilizados para determinar el valor óptimo del factor de no linealidad k , para rotación y escala, respectivamente. En todos los casos, el comportamiento fue similar, tanto para el filtro E como para el filtro imagen A, el máximo valor alcanzado del PCE correspondió a $k=0.3$. En estas gráficas, Letras (en color rojo) se refiere a las 26 letras del alfabeto e Imágenes (en color azul) a las 15 imágenes.

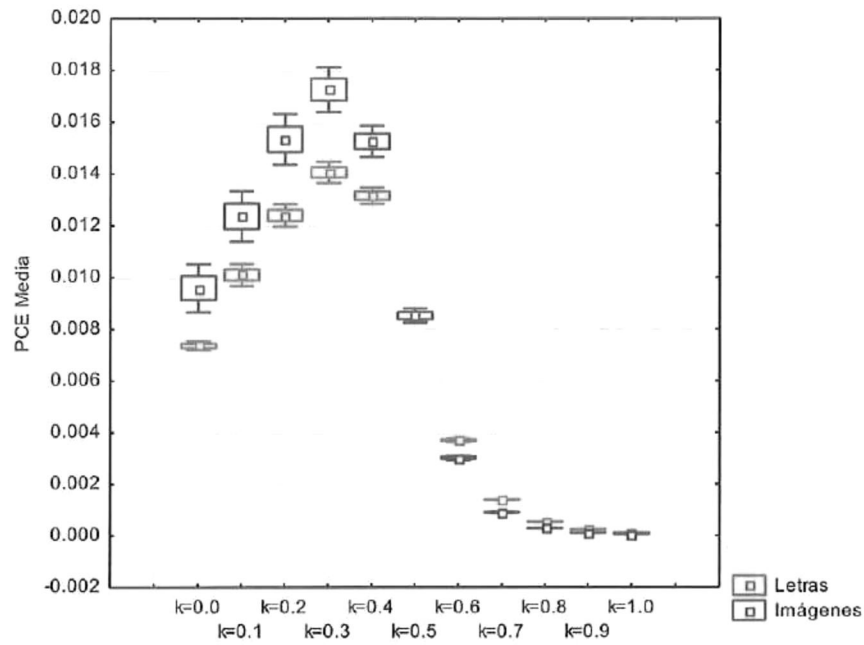


Figura 5.6. Gráfica de cajas para la comparación del *pico de correlación energía* PCE *versus* los valores de no linealidad k para la letra E y para la imagen A en rotación.

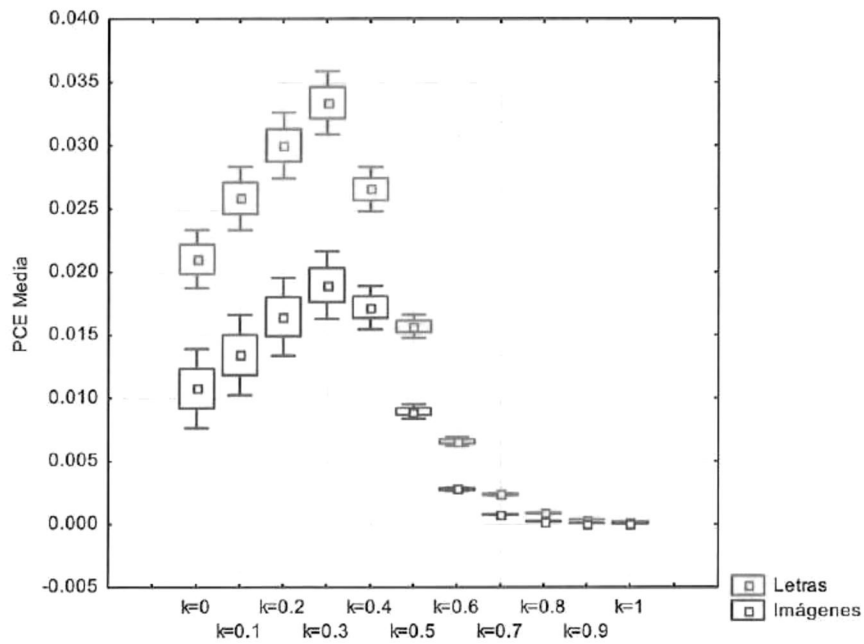


Figura 5.7. Gráfica de cajas para la comparación del *pico de correlación energía* PCE *versus* los valores de no linealidad k para la letra E y para la imagen A en escala.

5.1.1.3 Reconocimiento del filtro E con las letras del alfabeto en fuente Arial para invariancia a rotación y escala

En la Figura 5.8 se muestra el desempeño del filtro E de fuente Arial para rotación. El filtro con un factor no lineal $k=0.3$ está correlacionado con cada una de las 26 letras del alfabeto, con incrementos de un grado de rotación, desde 0° hasta 360° .

De la Figura 5.8 se observa que para la rotación no hay traslape de la letra E con el resto de las demás letras, mostrando así una excelente discriminación entre todas las letras del alfabeto.

En la Figura 5.9 se muestra el desempeño del filtro E de fuente Arial para escala. El filtro con un factor no lineal $k=0.3$ está correlacionado con cada una de las 26 letras del alfabeto, escaladas desde -25% hasta +25 con incrementos de uno por ciento.

De la Figura 5.9 se observa un traslape de la letra E con las letras F y L. Esto es debido a que la capacidad de sensibilidad a la rotación y la escala, tiene relación con la extensión y estructura de cada caracter. Por ejemplo, en este caso E es más sensible al tamaño que una O, pero no es sensible para las letras F y L. De todas maneras, a un error estándar el filtro E reconoce este carácter de la letra F. El filtro E mostró una buena discriminación entre todas las letras del alfabeto.

Se calcularon los valores de los niveles de confianza en porcentajes para los filtros no lineales en los casos de rotación y escala. Para el primer caso, la Tabla 5.1 muestra los resultados obtenidos para cada una de las letras del alfabeto como filtros, como se muestra en el primer renglón de la tabla, y correlacionados con cada una las 26 letras del alfabeto, como se muestra en la primera columna. Estos filtros fueron rotados con incrementos de 1° , desde 0° hasta 360° de rotación. El número de imágenes procesadas para los 26 filtros y correlacionados con cada una de las 26 letras del alfabeto, rotadas en un grado hasta completar 360° , da un total de 243360 imágenes procesadas estadísticamente.

Esto se hizo para determinar el nivel de confianza de cada filtro. Se encontraron solo dos traslapes entre el filtro y el caracter a ser reconocido y se denotó como No Considerado (NC), debido a la gran similitud entre ellos, el filtro y el caracter, que no fue

capaz de discriminarlo. Eso ocurre cuando hay caracteres que son altamente sensibles a rotación, como se mencionó antes debido a sus propias características de extensión y estructura.

Para el caso escalado, se calcularon los valores de los niveles de confianza en porcentajes para los filtros no lineales aplicando el mismo método para el caso de la rotación. La Tabla 5.2 muestra los resultados obtenidos para cada una de las letras como filtros, como se muestra en el primer renglón de la tabla, y correlacionados con cada una de las 26 letras del alfabeto, como se muestra en la primera columna. Estos filtros fueron escalados con incrementos de 1%, desde -25% hasta +25% en escala. El número de imágenes procesadas para los 26 filtros correlacionados con cada una de las 26 letras del alfabeto, escaladas en uno por ciento hasta completar $\pm 25\%$, da un total de 34476 imágenes procesadas estadísticamente.

De este modo, se determinó el nivel de confianza de cada filtro. Se encontraron aquí 29 traslapes entre el filtro y el carácter a ser reconocido y fue denotado como No Considerado (NC), debido a la gran similitud entre ellos. Pero los resultados mostraron un mejor nivel de confianza para la escala. Esto ocurre cuando hay caracteres que son altamente sensibles a escala, como se mencionó antes debido a sus propias características de extensión y estructura.

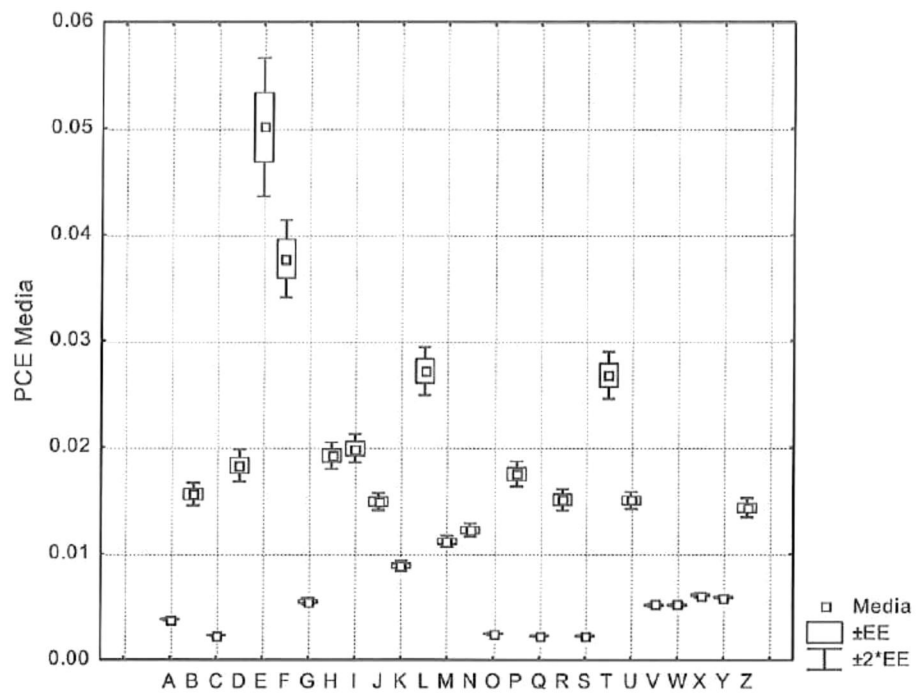


Figura 5.8. Desempeño del filtro E para rotación.

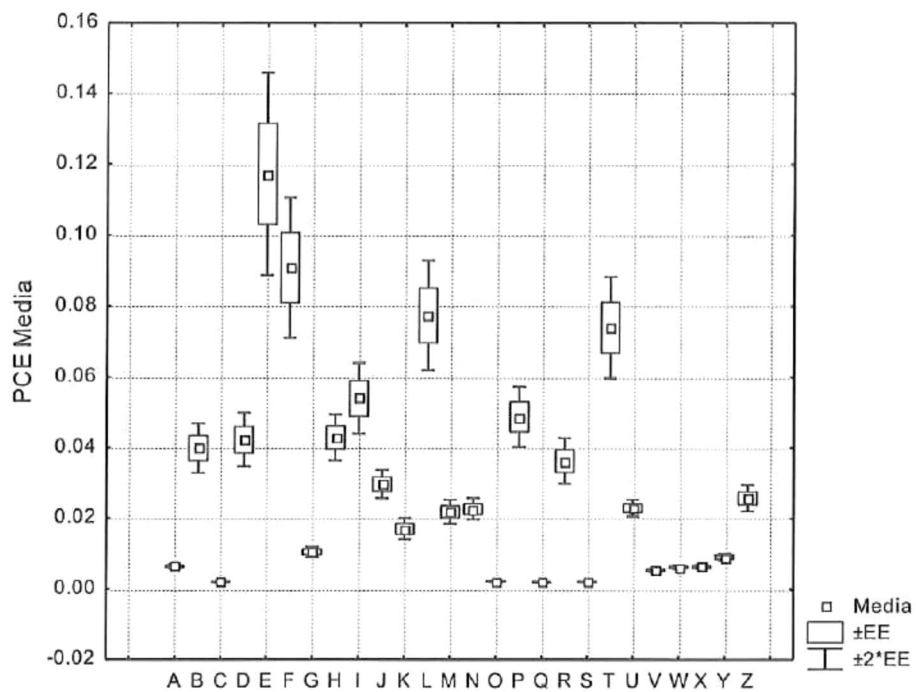


Figura 5.9. Desempeño del filtro E para escala.

Tabla 5.1. Niveles de confianza en porcentaje de los filtros mostrados en el primer renglón, correlacionados con las 26 letras del alfabeto las cuales se muestran en la primera columna, rotados desde 0 hasta 360 grados con un factor no lineal $k=0.3$. La notación NC se refiere a No Considerado, debido a que hay traslape entre el filtro y la letra en cuestión.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A		94	91	86	100	100	89	100	100	98	92	99	99	99	91	96	90	98	88	99	99	78	94	68	94	97
B	85		90	63	68	66	91	80	80	84	87	74	89	87	89	76	90	78	87	75	84	92	98	75	90	75
C	95	92		78	100	100	81	100	100	98	88	100	99	100	88	90	85	94	76	100	98	91	99	61	82	99
D	89	67	89		66	61	91	78	78	81	83	68	87	84	87	75	86	79	87	68	79	88	96	69	78	58
E	87	71	99	63		46	91	76	72	81	84	63	81	81	99	75	99	90	99	64	76	90	92	66	77	47
F	91	72	98	62	52		92	78	75	83	83	60	85	86	98	74	97	76	98	60	79	88	95	62	84	56
G	95	73	77	76	79	98		99	99	98	90	98	100	100	86	90	83	95	75	97	98	91	99	74	81	96
H	80	71	99	62	68	61	95		70	82	81	64	85	66	99	76	99	81	98	69	81	87	80	53	78	45
I	76	71	98	60	60	60	91	69		64	71	NC	77	71	97	71	98	78	99	58	69	81	90	33	64	29
J	85	72	88	64	70	66	88	73	66		78	69	82	81	88	76	87	82	88	73	75	84	91	55	75	44
K	91	73	90	68	70	83	89	79	76	79		79	85	84	89	84	89	87	89	78	82	89	97	68	80	70
L	81	66	98	59	53	55	89	69	67	76	77		80	79	98	68	92	70	98	NC	72	82	91	53	78	47
M	91	73	98	69	70	86	95	83	78	85	86	81		87	98	84	91	98	94	81	86	90	95	71	83	60
N	84	73	98	71	70	87	98	64	78	84	86	78	86		97	82	98	84	96	81	69	88	90	83	79	52
O	95	73	78	76	100	100	81	100	100	97	88	100	100	99		90	86	93	75	100	97	92	100	68	87	99
P	89	69	91	58	67	84	91	78	74	82	83	66	86	85	89		90	77	90	66	79	88	95	66	81	54
Q	95	73	79	79	100	100	79	100	100	99	99	100	100	100	87	91		93	77	100	98	91	100	67	94	99
R	90	71	89	60	68	76	91	77	80	82	84	71	87	84	90	76	89		89	72	84	90	97	73	82	69
S	95	73	80	88	100	100	79	100	100	100	89	100	100	100	88	94	83	95		100	98	93	100	64	94	100
T	86	68	98	54	57	62	89	71	71	78	81	54	84	84	96	71	96	73	95		80	85	94	63	76	54
U	82	73	93	67	69	85	91	54	69	82	84	76	83	63	91	82	92	85	90	76		89	91	71	78	51
V	69	73	89	76	99	99	90	97	93	95	94	92	96	94	91	93	89	96	89	95	94		96	89	92	89
W	78	73	93	76	97	100	90	99	99	99	94	99	99	98	93	97	92	99	91	93	97	93		66	87	97
X	95	73	89	76	98	97	91	93	89	93	90	93	97	93	92	89	92	95	89	96	88	91	99		92	69
Y	90	73	87	74	97	95	86	94	88	86	86	88	95	87	86	86	86	88	86	89	86	91	97	73		94
Z	84	73	97	74	69	86	94	84	80	83	90	80	87	85	97	83	96	87	93	84	84	90	97	85	84	

Tabla 5.2. Niveles de confianza en porcentaje de los filtros mostrados en el primer renglón, correlacionados con las 26 letras del alfabeto las cuales se muestran en la primera columna, escalados 25% con un factor no lineal $k=0.3$. La notación NC se refiere a No Considerado, debido a que hay traslape entre el filtro y la letra en cuestión.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A		100	100	100	100	100	100	100	100	100	100	100	100	100	98	100	100	100	100	100	100	96	100	100	100	100
B	100		100	78	80	80	88	80	86	84	86	80	86	86	100	73	100	84	100	82	85	100	100	100	100	100
C	100	100		100	100	100	98	100	100	100	100	100	100	100	NC	100	98	100	100	100	100	100	100	100	100	100
D	100	82	98		84	86	88	73	90	84	90	80	86	82	94	71	100	90	100	84	88	100	100	100	100	100
E	100	77	100	80		NC	96	75	86	86	88	76	84	86	100	76	100	84	100	80	82	100	100	100	100	98
F	100	69	100	NC	NC		94	67	69	71	78	NC	76	78	100	51	100	76	100	NC	76	100	100	100	100	100
G	100	98	86	90	100	100		100	100	98	100	96	100	100	88	96	100	100	100	98	100	100	100	100	100	100
H	100	62	100	75	78	73	98		78	NC	84	73	73	76	100	65	100	82	100	78	NC	100	100	100	100	100
I	100	75	100	NC	76	59	96	NC		75	84	NC	78	69	100	NC	100	75	100	NC	NC	100	100	100	100	100
J	100	92	100	88	92	94	94	71	78		94	84	88	82	96	78	100	88	100	98	82	100	100	100	100	100
K	100	90	100	90	90	90	98	76	90	88		90	90	90	88	90	100	90	100	90	90	100	100	100	100	100
L	100	69	100	NC	NC	NC	90	NC	NC	NC	82		78	75	100	NC	100	71	100	NC	73	100	100	100	100	100
M	100	84	100	86	92	86	100	80	88	84	94	84		88	100	84	100	88	100	90	78	100	100	100	100	100
N	100	84	100	80	90	82	100	73	84	78	92	80	80		100	82	100	84	100	82	71	100	100	100	100	100
O	100	100	NC	100	100	100	98	100	100	100	100	100	100	100		100	96	100	100	100	100	100	100	100	100	100
P	100	75	100	NC	80	73	86	73	75	75	86	75	84	86	96		100	80	100	82	80	100	100	100	100	100
Q	100	100	90	100	100	100	100	100	100	100	100	100	100	100	NC	100		100	100	100	100	100	100	100	100	100
R	100	88	100	82	88	82	90	78	90	88	90	84	88	82	100	75	100		100	80	90	100	100	100	100	100
S	100	100	86	100	100	100	100	100	100	100	100	100	100	100	84	100	100	100		100	100	100	100	100	100	100
T	100	77	100	NC	59	NC	90	71	73	NC	90	NC	82	80	100	61	100	88	100		80	100	100	100	100	100
U	100	80	100	75	94	80	98	75	88	80	92	82	86	86	98	78	100	90	100		100	100	100	100	100	100
V	98	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100		100	100	100	100	100
W	100	100	100	100	100	100	100	98	100	100	100	100	100	100	100	100	100	100	100	100	100		100	100	100	100
X	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100		100	100	100
Y	100	100	100	86	100	100	94	85	86	84	100	86	86	92	96	86	100	86	100	92	82	100	100	100		100
Z	100	100	100	100	76	100	100	100	100	100	100	100	98	100	100	100	100	100	100	98	100	100	100	100	100	

5.1.2 Comparación del desempeño de los filtros no lineales y filtros lineales para invariancia a rotación y escala

Las figuras de la 5.10 a la 5.13 son ejemplos de correlaciones usando filtros no lineales, con $k=0.3$, las cuales son comparadas con los filtros sólo de fase. Las últimas cuatro figuras son para los filtros compuestos no lineales y los compuestos sólo de fase. De estas figuras se observa que para los filtros no lineales el plano de correlación es menos ruidoso y los picos de correlación están más definidos y angostos que para los filtros sólo de fase. En las Figuras 5.10 y 5.11, la imagen objeto y la imagen problema son las letras E en fuente Times New Roman, y en las Figuras 5.12 y 5.13, la imagen objeto y la imagen problema son las letras EHF en fuente Arial. En la Figura 5.10, los tres picos de correlación están a lo largo del eje de escala que corresponden a cada una de las letras escaladas en la imagen problema. La Figura 5.11 muestra la escala y rotación en la imagen problema y los tres picos de correlación están sobre los ejes de escala y rotación que corresponden a cada una de las letras de la imagen problema. En la Figura 5.12 tenemos un filtro compuesto no lineal y un pico centrado en el plano de correlación que indica el reconocimiento de EH en la imagen problema. En la Figura 5.13 la letra H está rotada 90 grados en la imagen problema y hay dos picos simétricos a lo largo del eje de rotación en el plano de correlación correspondiendo a 90 y 270 grados, debido a que H está rotada 90 grados, esta tiene la misma forma que para una rotación de 270 grados. En todos estos casos, el valor del PCE para el filtro no lineal fue mejor comparado con el filtro sólo de fase.

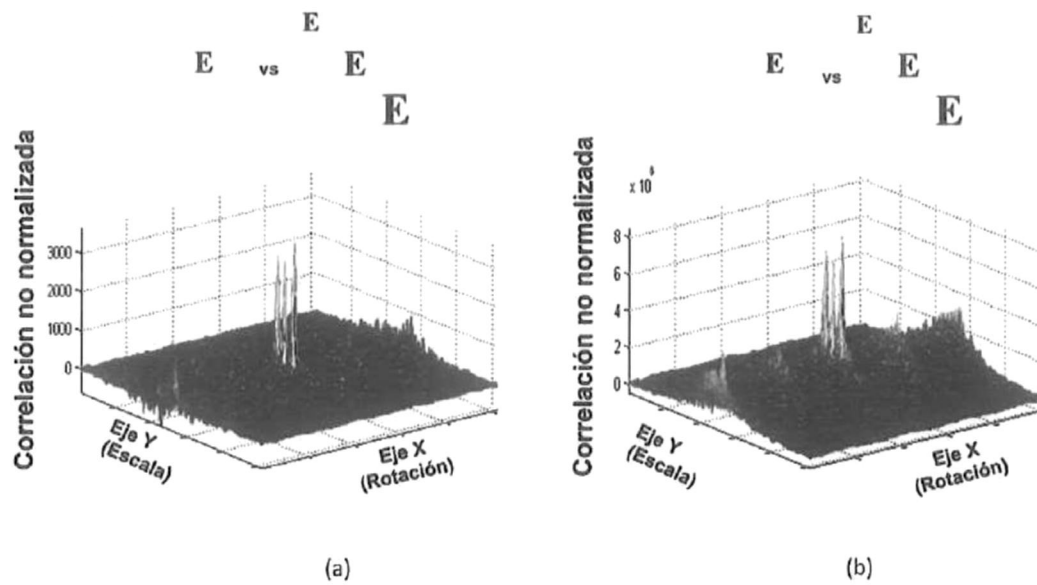


Figura 5.10. Correlación invariante a escala usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.0759$, (b) un filtro sólo de fase con $PCE=0.0156$.

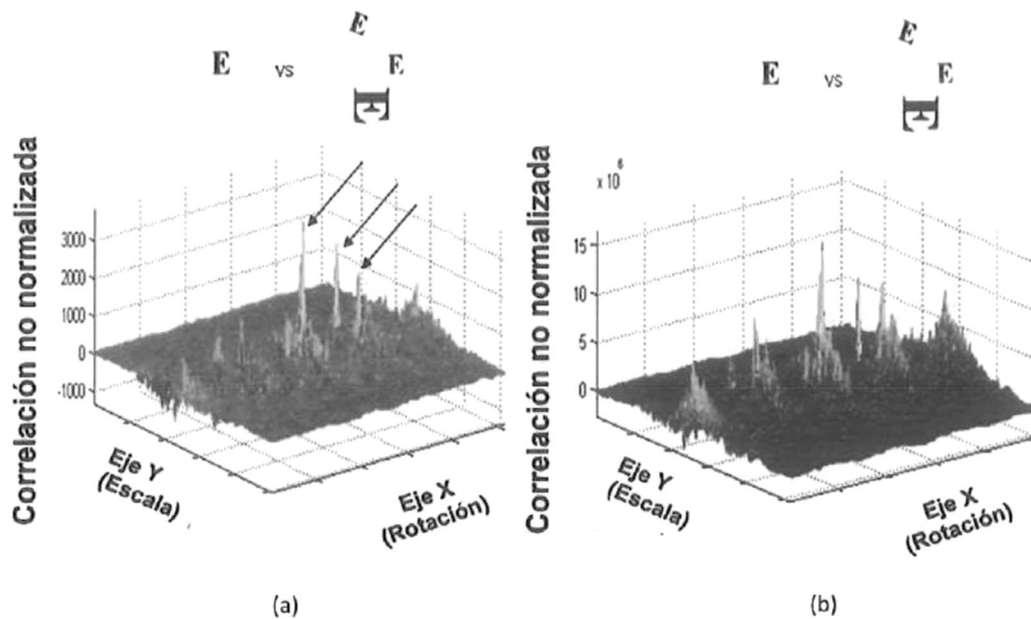


Figura 5.11. Correlación invariante a rotación y escala usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.0351$, (b) un filtro sólo de fase con $PCE=0.0141$.

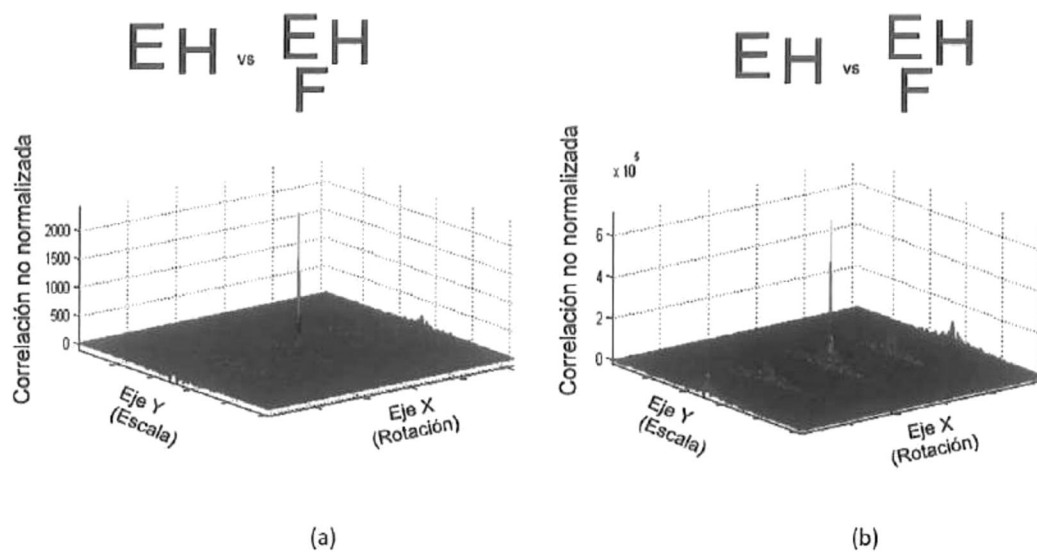


Figura 5.12. Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y $PCE=0.4704$, (b) un filtro compuesto sólo de fase con $PCE=0.1932$.

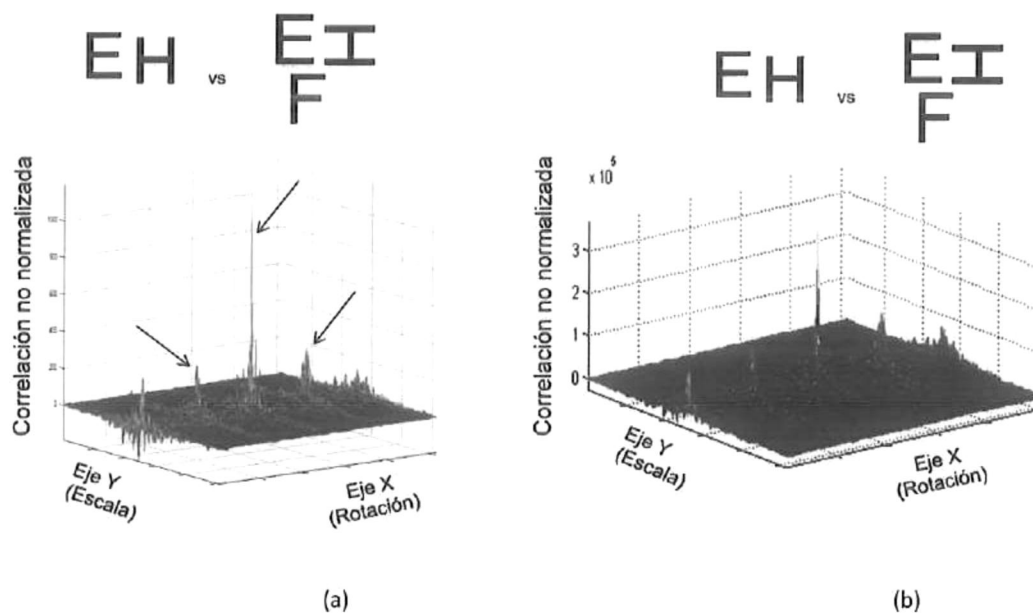


Figura 5.13. Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y $PCE=0.3879$, (b) un filtro compuesto solo de fase con $PCE=0.1629$.

De las Figuras 5.14 a la 5.17, se presentan algunos ejemplos de correlaciones usando filtros no lineales, con $k=0.3$, y son comparados con los filtros sólo de fase. Las últimas cuatro figuras son para los filtros compuesto no lineal y compuesto sólo de fase. De estas figuras se observa que para el filtro no lineal el plano de correlación es menos ruidoso y los picos de correlación están más definidos y angostos que para el filtro sólo de fase. En la Figura 5.14, el filtro y las imágenes problema son las letras E en fuente Times New Roman, y de las Figuras 5.15, 5.16 y 5.17, el filtro y las imágenes problema son letras de fuente Arial. En la Figura 5.14, los tres picos están a lo largo del eje de rotación que corresponde a cada letra E rotada en la imagen problema. La Figura 5.15 muestra la escala y rotación en la imagen problema y los tres picos de la letra A están sobre el eje de escala y el eje de rotación que corresponden a cada una de las letras de la imagen problema. En la Figura 5.16 tenemos un filtro compuesto no lineal y un pico centrado en el plano de correlación que indica el reconocimiento de EG en la imagen problema. En la Figura 5.17, la letra G está rotada 45° en la imagen problema y hay dos picos simétricos a lo largo del eje de rotación en el plano de correlación, cada uno correspondiendo a 45° y 225° respectivamente, porque la letra G rotada 45° tiene la misma forma que para 225° . Para tener una valoración cuantitativa de la comparación entre estos filtros, denotamos por PCE_{NL} , PCE_{POF} , PCE_{CNL} y PCE_{CPOF} los valores del PCE para los filtros no lineal, sólo de fase, compuesto no lineal y compuesto sólo de fase, respectivamente, de tal manera que las razones de desempeño de estos filtros fueron, de la Figura 5.14: $PCE_{NL}/PCE_{POF}=6$; de la Figura 5.15: $PCE_{NL}/PCE_{POF}=2.3$; de la Figura 5.16: $PCE_{CNL}/PCE_{CPOF}=3.5$ y de la Figura 5.17: $PCE_{CNL}/PCE_{CPOF}=2.4$. En todos estos casos, el valor PCE para el filtro no lineal fue mejor comparado con el filtro sólo de fase.

El sistema digital para la correlación invariante usando un filtro no lineal se probó con el mejor valor del factor de no linealidad $k=0.3$, el cual fue determinado experimentalmente para la escala y rotación. Se utilizaron todas las letras del alfabeto en fuente Arial, donde cada una de estas letras se tomó como filtro y fueron correlacionadas con cada una de las demás letras, con un nivel de confianza de 95%, probando así que nuestro sistema funciona eficientemente para la discriminación de objetos. Este método de correlación invariante no lineal se aplicó a varios objetos escalados y rotados usando filtros

no lineales, compuestos no lineales, sólo de fase y compuestos sólo de fase, con los cuales se encontró que el mejor desempeño del PCE correspondió a los filtros no lineales.

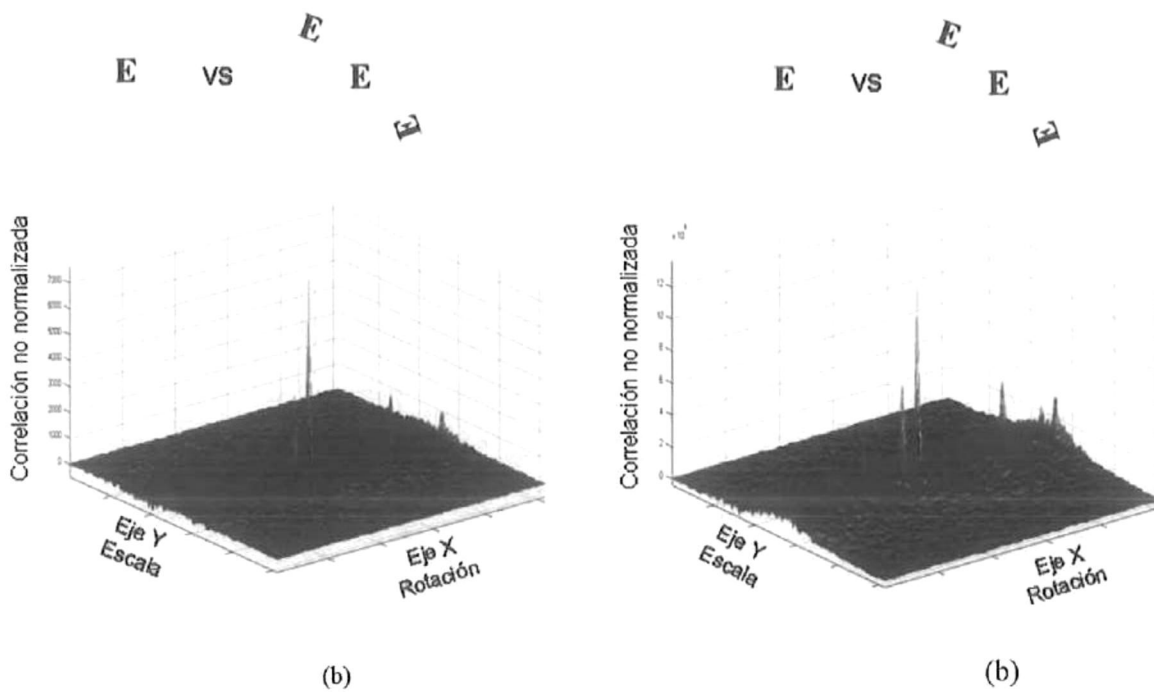


Figura 5.14. Correlación invariante a rotación usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.2052$, (b) un filtro sólo de fase con $PCE=0.0342$.

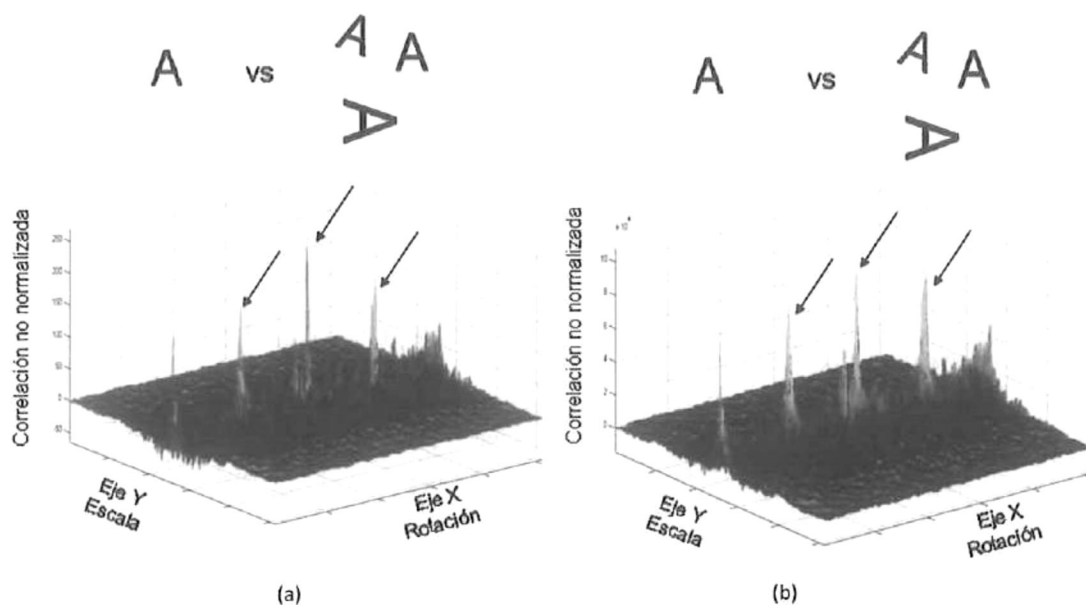


Figura 5.15. Correlación invariante a rotación y escala usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.0348$, (b) un filtro sólo de fase con $PCE=0.0153$.

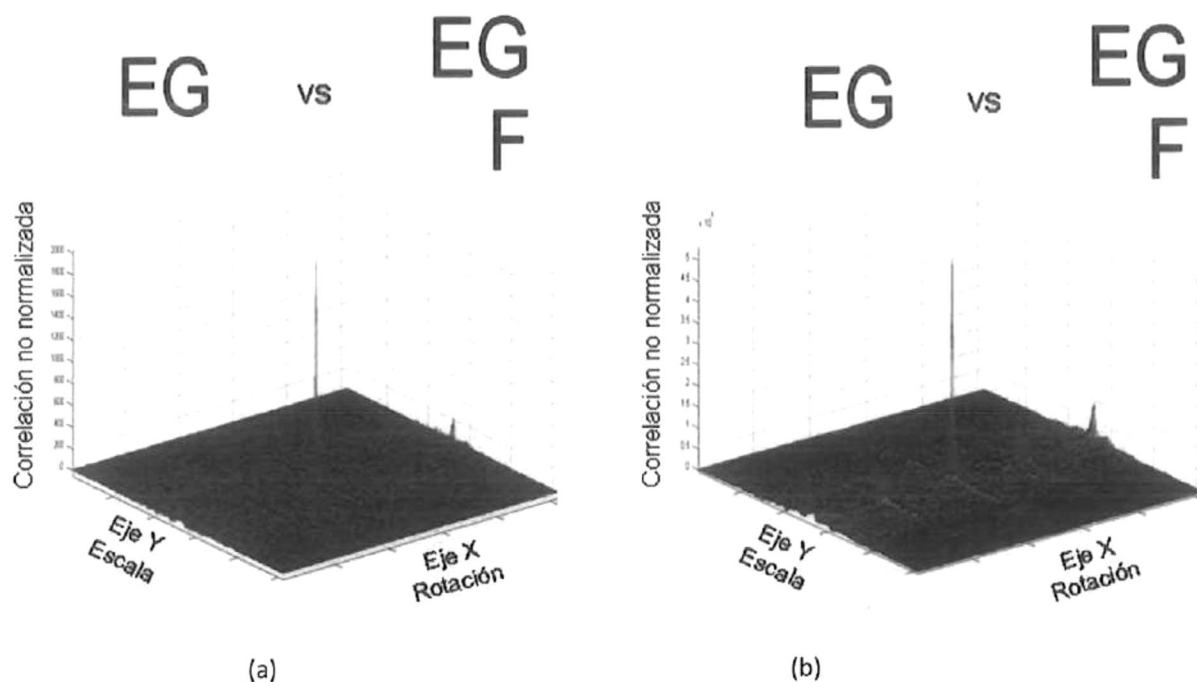


Figura 5.16. Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y $PCE=0.4554$, (b) un filtro compuesto sólo de fase con $PCE=0.1316$.

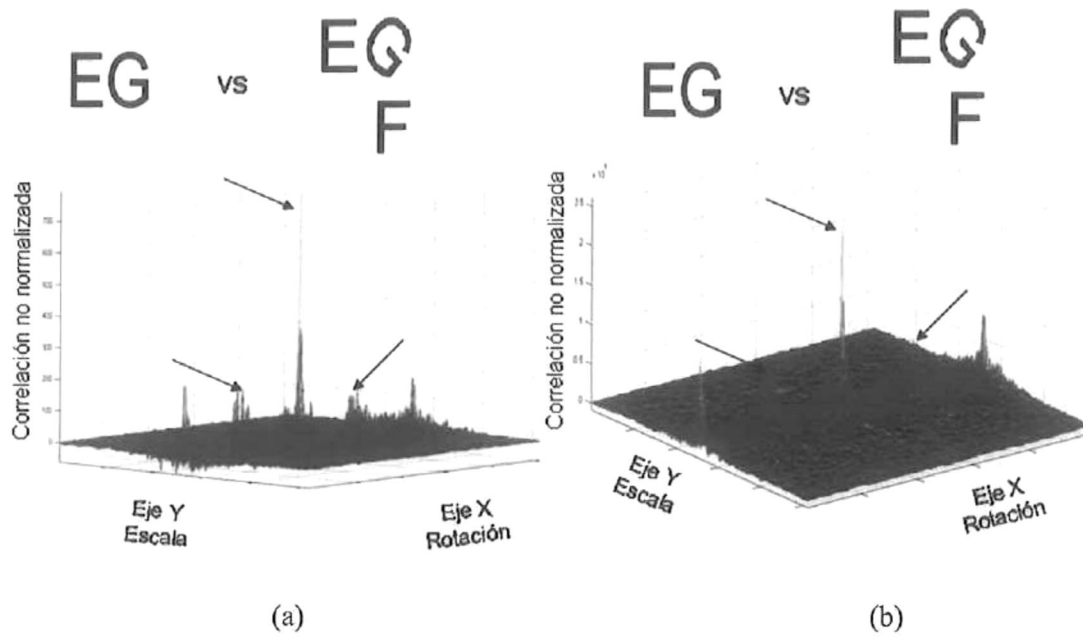


Figura 5.17. Correlación invariante usando (a) un filtro compuesto no lineal con $k=0.3$ y $PCE=0.2136$, (b) un filtro compuesto sólo de fase con $PCE=0.09$.

§ 5.2 Reconocimiento de patrones de objetos libres de ruido y con ruido usando correlación no lineal invariante

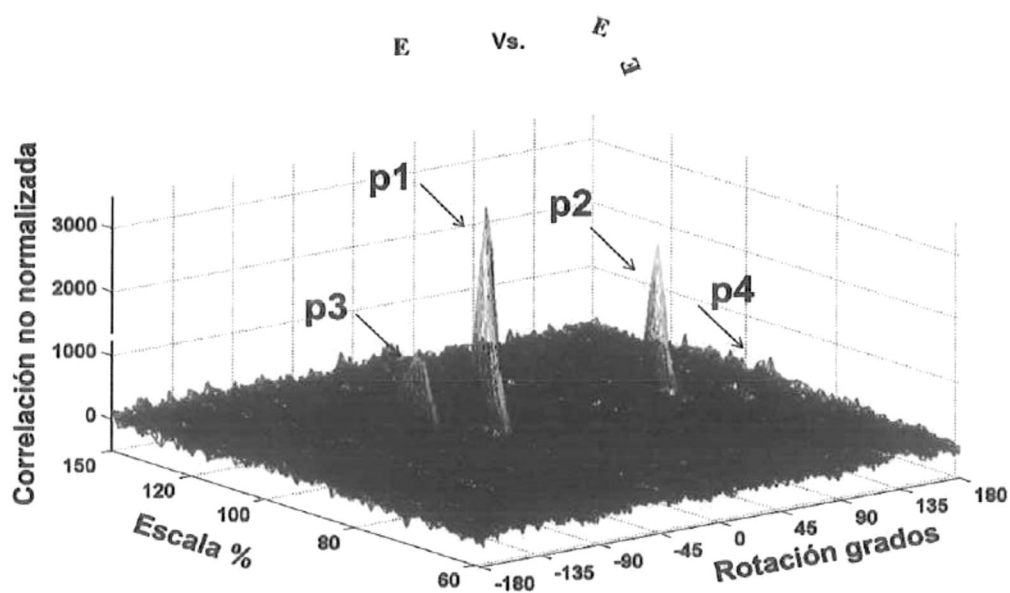
En este trabajo se obtuvieron algunos resultados aplicando la metodología descrita anteriormente en la Figura 4.3 del sistema invariante digital con un filtro no lineal, al que se denota ahora por DISNF, con un nuevo algoritmo computacional para el reconocimiento de varios objetos, independientemente de su tamaño, orientación angular, desplazamiento y ruido. Se escogió la transformada de escala, la cual ha sido aplicada en diferentes áreas, tales como en la eliminación de ruido en imágenes (Cristóbal y Cohen, 1998), en la identificación y registro de algunas letras del alfabeto y diatomeas al calcular el cepstrum de potencia del mapeo escalar log-polar (Pech-Pacheco *et al.*, 2000) y en la identificación automática de algas de fitoplancton (Pech-Pacheco *et al.*, 2001), y se introdujo un filtro que se le conoce por filtro que obedece a la k -ésima ley no lineal (Vijaya Kumar y Hassebrook, 1990) con un factor de intensidad de no linealidad dado por $k=0.3$ (Coronel-Beltrán y Alvarez-Borrego, 2008a) para realizar la correlación digital invariante la cual da información de la similitud entre objetos distintos. Se analizó el desempeño del filtro no lineal usando la métrica de capacidad de discriminación (DC) cuando la escena de entrada estaba en la presencia de ruido gaussiano aditivo y ruido sal y pimienta (S&P). Se realizaron experimentos utilizando algoritmos para encontrar la tolerancia al ruido del filtro, y se utilizó la correlación de rango de Spearman (en conjunto con el método no lineal, al que se denotó por SDISNF) para incrementar esa tolerancia. Los resultados mejoraron significativamente cuando se introdujo en el sistema digital de correlación la correlación no lineal de Spearman.

5.2.1 Comparación de los filtros no lineal y sólo de fase usando un sistema digital de correlación invariante

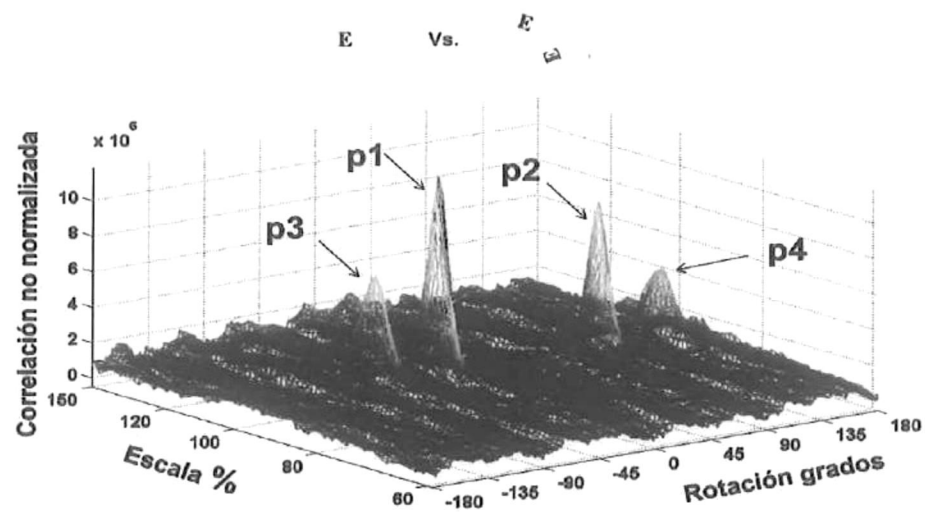
Se presentan algunos ejemplos usando un filtro no lineal con $k=0.3$, el cual es comparado con un filtro sólo de fase (Pech-Pacheco *et al.*, 2003). Las Figuras 5.18 y 5.19 muestran las gráficas de correlación a la salida del desempeño de estos filtros, donde el objetivo y la imagen problema son la letra E. En la Figura 5.18(a) se observa un plano de correlación

menos ruidoso con picos bien definidos p1, p2 y p3, cuando se usa un filtro no lineal. El valor de PCE fue 0.0215. En la Figura 5.18(b) se observa un plano de correlación de salida con un fondo más ruidoso cuando se usa un filtro lineal. Para este caso, el valor de PCE fue 0.0057.

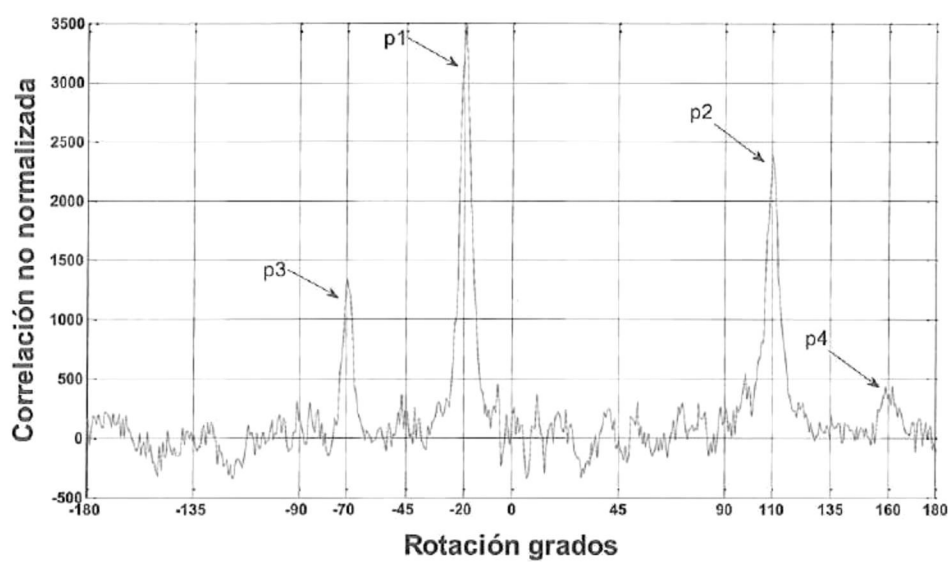
En este ejemplo, sólo se analizó la rotación angular ya que el objetivo y la imagen problema tienen la misma escala. La Figura 5.18(c) muestra un transecto en el eje de rotación de la Figura 5.18(a). Se observan tres picos, p1, p2, p3, y un pequeño pico p4.



(a)



(b)



(c)

Figura 5.18. Correlación invariante a rotación usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.0215$, (b) un filtro sólo de fase con $PCE=0.0057$, y (c) un corte a lo largo del eje de rotación de (a).

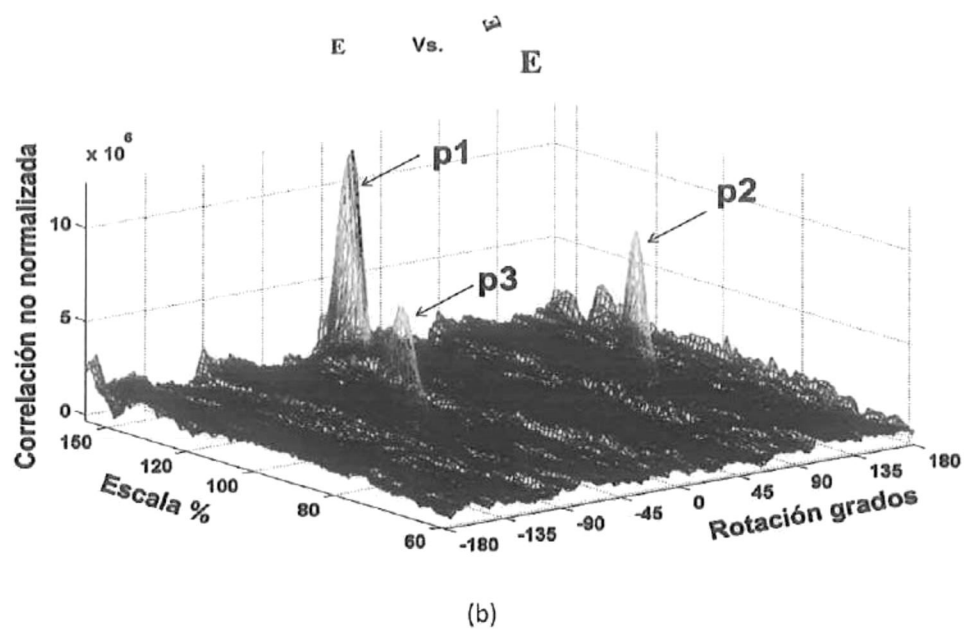
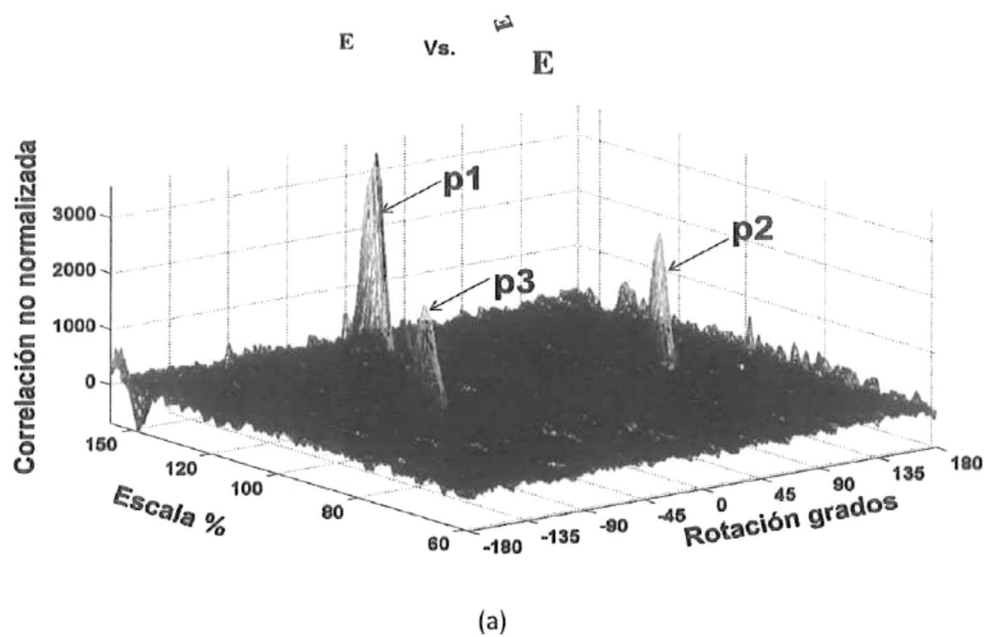
En la Figura 5.18(c) se puede observar que los picos de correlación están a lo largo del eje de rotación que corresponden a cada letra E rotada en la imagen problema. De la Figura 5.18(a) se observa que en la imagen problema se tienen dos letras E, una está rotada a -20° y la otra a 110° . El pico p1 corresponde a -20° y el pico p2 a 110° , el pico p3 a -70° y el pico p4 a 160° . Los picos p1 y p4, y p2 y p3, están a ángulos complementarios, porque cuando se rota un objeto a cierto ángulo, existe otro ángulo para el mismo objeto [Figura 5.18(c)]. Así que, en valores de correlación $p1+p4=p2+p3$. En la Figura 5.18(b) el pico p4 que está a -20° , corresponde a 160° que es complementario a -20° del pico p1. Así, de nuevo se tiene que $p1+p4=p2+p3$.

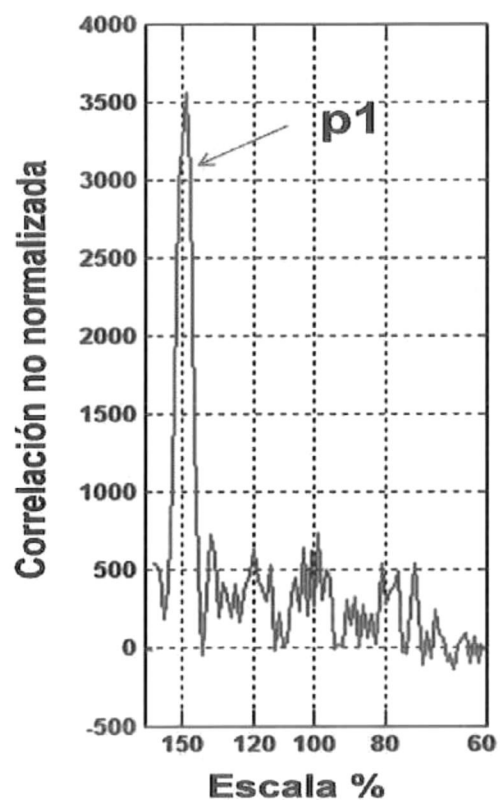
En la imagen problema de la Figura 5.19 se tienen dos letras E, una está rotada 110° y la otra a 0° y escalada a 150% con respecto al objetivo. El pico p1 está a 0° , el pico p2 a 110° y su pico correspondiente p3 está a -70° .

De la Figura 5.18(a) se puede observar al pequeño pico p4 que en la Figura 5.18(b) aparece más prominente. Esto es debido a que en el primer caso estamos tratando con un filtro no lineal y en el segundo caso con un filtro sólo de fase. Esto muestra que con el filtro no lineal tenemos menos ruido en el plano de correlación y mientras menos ruido tengamos los picos secundarios tienden a desaparecer. Las Figuras 5.19(a) y 5.19(b) muestran el plano de correlación de salida cuando en la imagen problema se tiene un caracter con diferente rotación y escala y se usa un filtro no lineal y un filtro lineal para el reconocimiento de la letra. En la Figura 5.19(c) se muestra un transecto a lo largo del eje de escala. El pico p1 aparece exactamente en el valor de 150% sobre el eje de escala.

Como se estudió el comportamiento del filtro no lineal versus el filtro lineal, se hizo una comparación cuantitativa entre estos filtros usando la métrica PCE. Un atributo deseable de un correlador es que sea capaz de producir picos de correlación angostos. Esta característica es tomada en cuenta por el PCE. Un filtro no lineal, comparado con el filtro sólo de fase, concentra la mayoría de la energía que pasa a través del filtro en el pico de correlación. Se denotan por PCE_{NLF} , PCE_{POF} los valores de PCE para el filtro no lineal y el filtro sólo de fase, respectivamente. De esta manera, las razones de estos filtros fueron, de

la Figura 5.18: $PCE_{NLF}/PCE_{POF}=3.77$; y de la Figura 5.19: $PCE_{NL}/PCE_{POF}=4.1$. En todos estos casos, el valor de PCE para el filtro no lineal fue mejor comparado con el filtro sólo de fase.





(c)

Figura 5.19. Correlación invariante a rotación y escala usando (a) un filtro no lineal con $k=0.3$ y $PCE=0.025$, (b) un filtro solo de fase con $PCE=0.0061$, y (c) un transecto de escala de (a).

5.2.2 Objetivo con ruido

El objetivo fue tratado con dos tipos de ruidos. Uno fue el ruido gaussiano aditivo y el otro fue el ruido sal y pimienta, este último denotado por S&P (Figura 5.20).

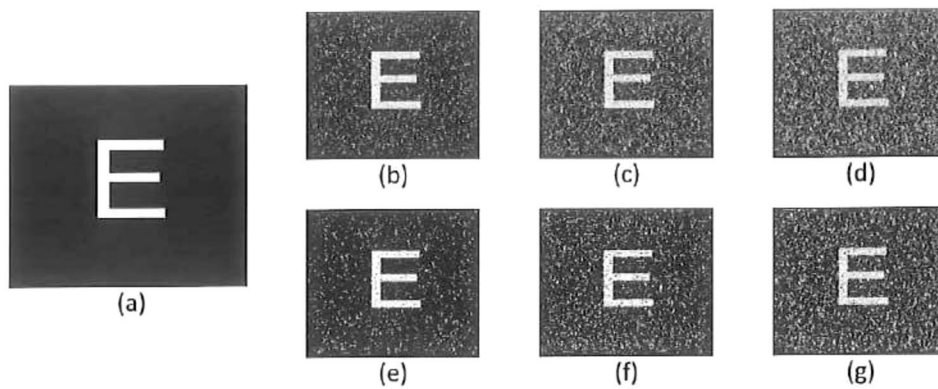


Figura 5.20. Objetivo libre de ruido (a), con ruido gaussiano aditivo de media cero y variancia de 0.1 (b), 0.2 (c) y 0.3 (d), y con ruido S&P de densidad de 0.1 (e), 0.2 (f) y 0.3 (g).

Las Figuras 5.21 y 5.22 muestran las gráficas del coeficiente de discriminación usando 30 experimentos numéricos a un 95% de nivel de confianza para un filtro no lineal con $k=0.3$ cuando se considera el ruido gaussiano aditivo con media cero, y el ruido S&P. En ambas gráficas se tiene el DC medio *versus* la variancia y la densidad de ruido. De estas gráficas, se observa que el filtro puede reconocer al objeto con una variancia de ruido de aproximadamente 0.27 y una densidad de ruido de aproximadamente 0.3 para ruidos gaussiano aditivo y S&P, respectivamente.

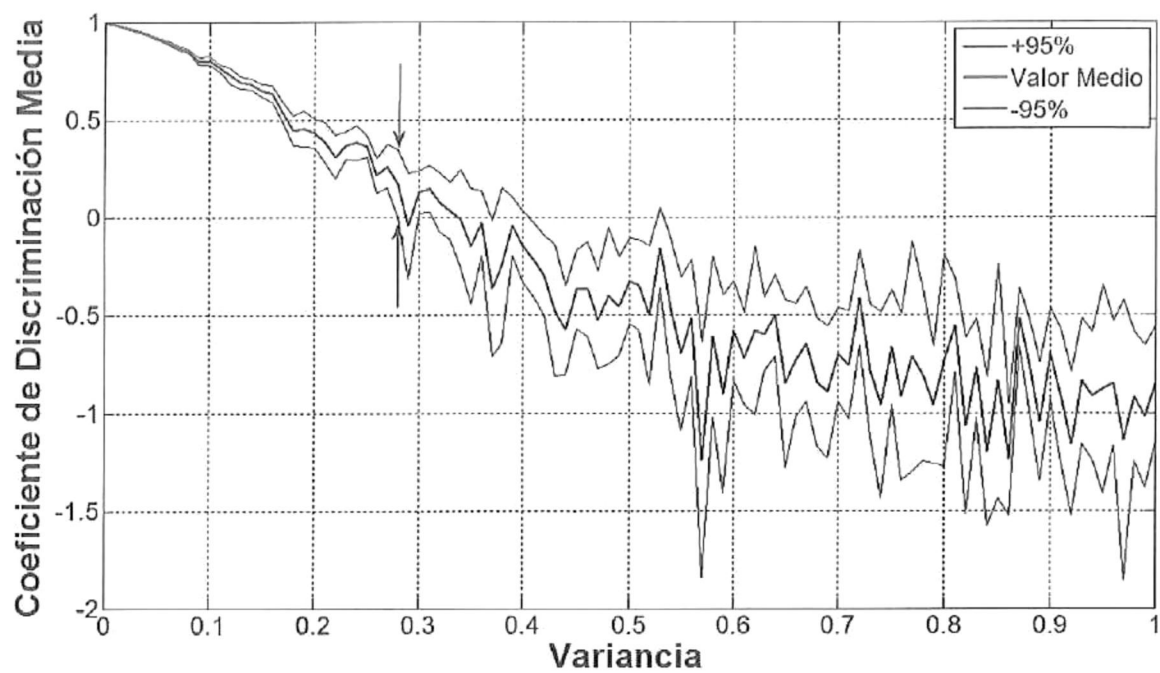


Figura 5.21. Desempeño de un filtro no lineal con $k=0.3$ en la presencia de ruido gaussiano aditivo.

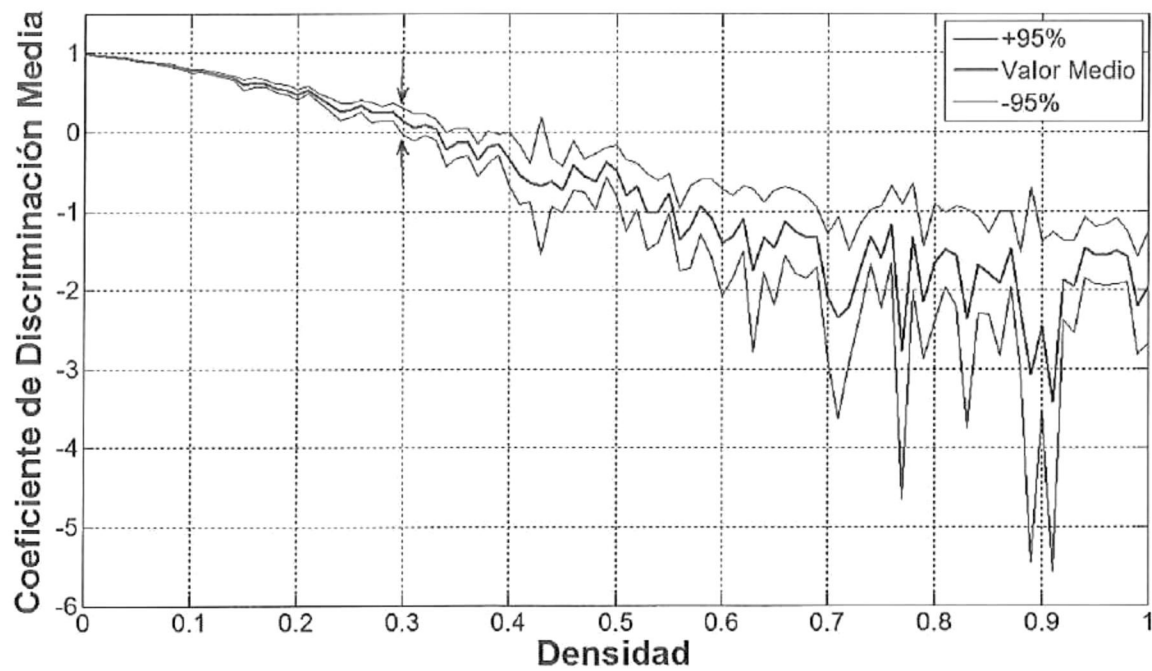


Figura 5.22. Desempeño de un filtro no lineal con $k=0.3$ en la presencia de ruido S&P.

Es importante mencionar que el máximo valor que puede obtenerse con el DC es la unidad, mientras que los valores menores que cero indican que el filtro no reconoce al objeto.

5.2.2.1 DISNF y la correlación de Spearman

Para estudiar el mejoramiento del método cuando las imágenes están inmersas en un ruido gaussiano aditivo y en un ruido S&P, se utilizó un método paramétrico llamado estadística de rango para calcular la correlación entre dos imágenes. La expresión de correlación no lineal se obtiene al sustituir el valor de cada pixel por su correspondiente rango, en la expresión de correlación normalizada [correlación de rango de Spearman (SRC)]. Tomando esto en consideración, la SRC en 2-D puede expresarse como (Guerrero-Moreno y Álvarez-Borrego, 2009; Martínez-Díaz, 2008):

$$R(k, l) = 1 - \frac{6 \sum_{m, n \in W} (r_t(m, n) - r_s(m + k, n + l))^2}{|W| \cdot (|W|^2 - 1)},$$

donde $\{r_t(m, n), m = 1, 2, \dots, N; n = 1, 2, \dots, M\}$ es el rango del objetivo, $\{r_s(m + k, n + l), k = 1, 2, \dots, N; l = 1, 2, \dots, M\}$ es el rango de la imagen problema, W es el tamaño de cada imagen, R es la correlación de rango de Spearman y m, n, k, l son las coordenadas espaciales.

La Figura 5.23 muestra un diagrama de bloques que representa a la SRC junto con el método no lineal (SDISNF) para incrementar la tolerancia al ruido. En esta figura, desde el paso (1) hasta el paso (10), el procedimiento es el mismo que el mostrado en la Figura 4.3, excepto que ahora en el paso (11) se realiza la FFT inversa para los pasos (9) y (10). El paso (12) es la SRC en el dominio espacial, y en el paso (13) se obtiene la correlación en el plano de salida.

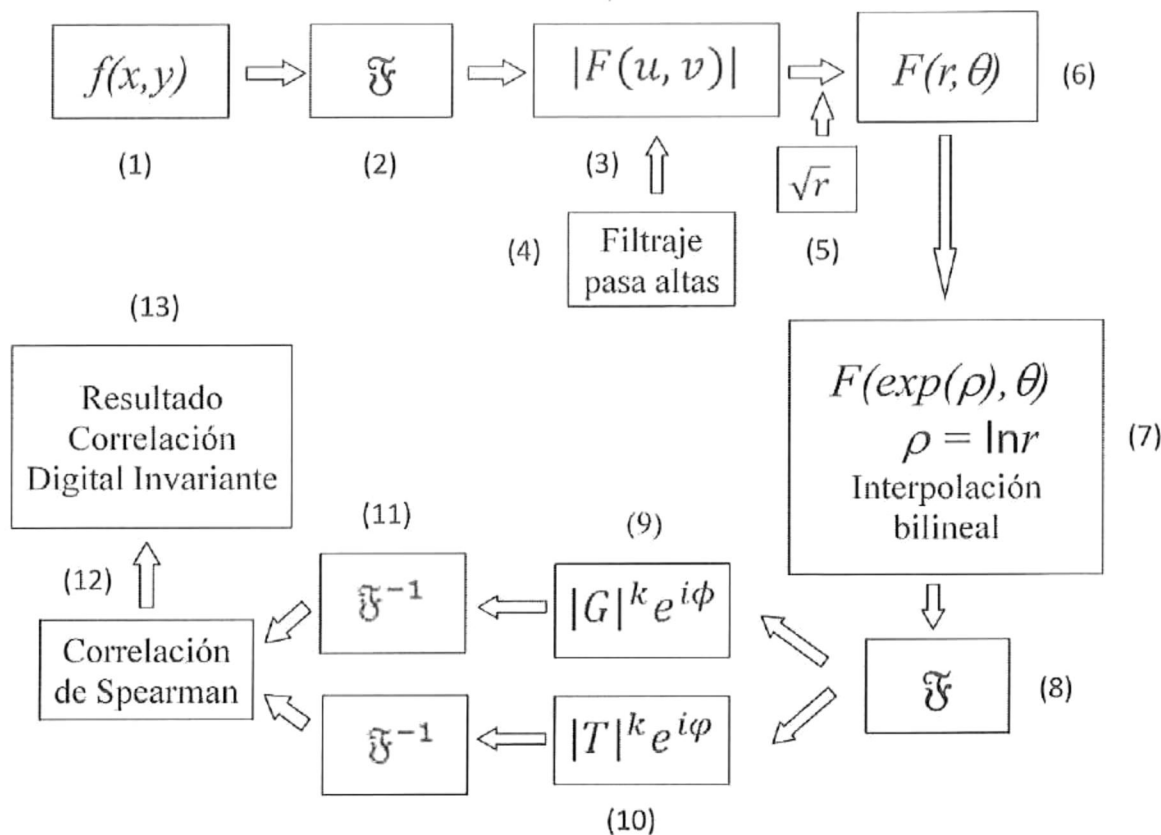


Figura 5.23. Diagrama de bloques simplificado representando al sistema de correlación invariante con un filtro no lineal usando el SDISNF.

De las Figuras 5.24 y 5.25 se observa el mejoramiento del desempeño del filtro cuando se añade ruido gaussiano aditivo con media cero y ruido S&P, respectivamente. De estas gráficas, se observa que el filtro puede soportar una variancia de ruido de aproximadamente 0.92 y una densidad de ruido de aproximadamente 0.54 para ruidos gaussiano aditivo y S&P, respectivamente. Para valores menores que estos tenemos valores negativos del coeficiente de discriminación, lo cual significa que nuestro sistema no es capaz de discriminar los objetos.

Estas figuras muestran la comparación del desempeño cuando se utiliza el SDISNF para el objetivo inmerso en ruido gaussiano aditivo y ruido S&P. De estas figuras, se puede ver que el coeficiente de discriminación es mayor por un factor de tres que cuando se tiene solo el ruido gaussiano aditivo y para el ruido S&P el factor es aproximadamente dos.

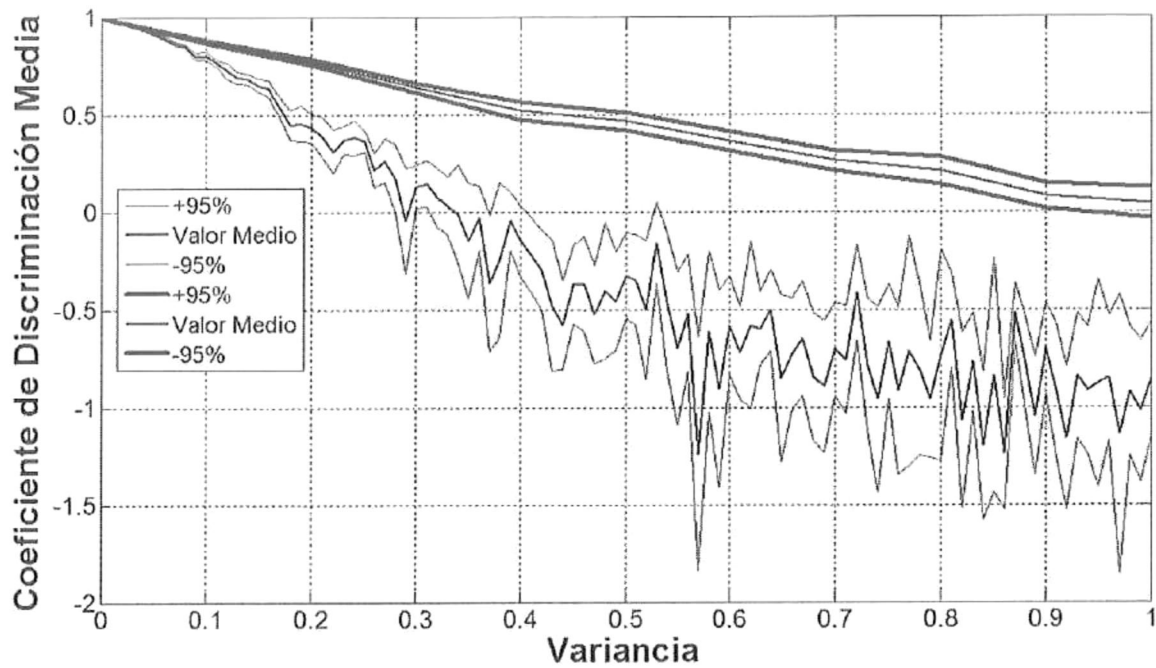


Figura 5.24. Comparación del desempeño con el SDISNF cuando el objetivo está inmerso en ruido aditivo gaussiano aditivo.

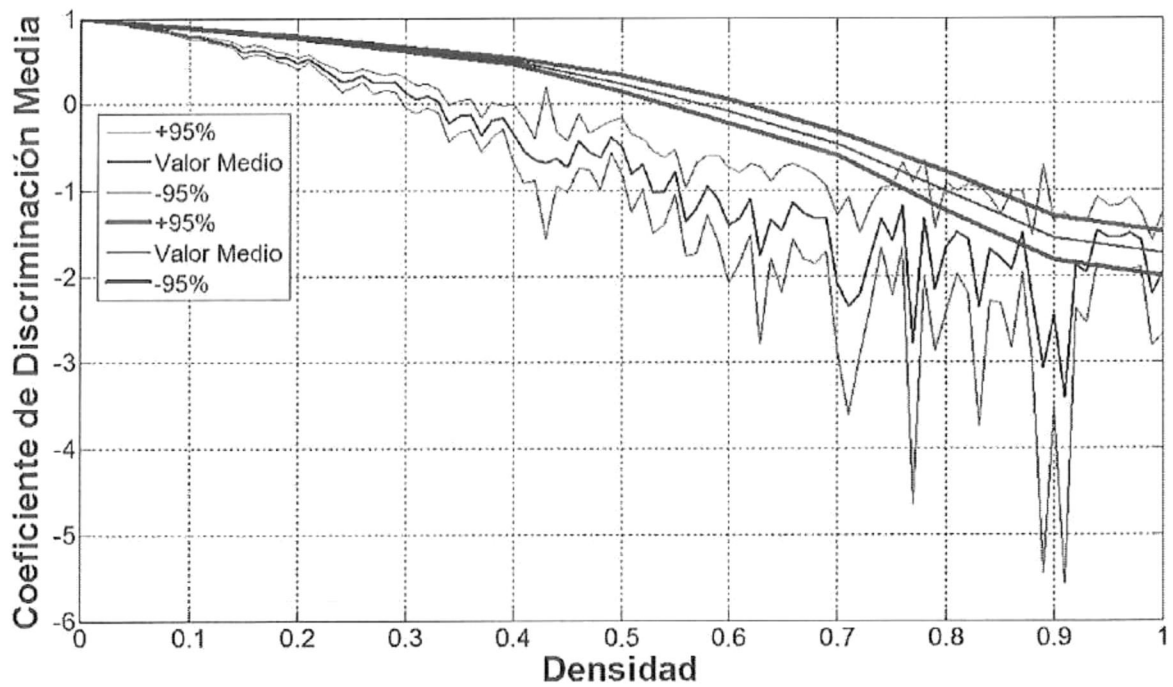


Figura 5.25. Comparación del desempeño con el SDISNF cuando el objetivo está inmerso en ruido S&P.

Se obtuvieron mejores resultados aplicando la correlación de Spearman debido a que al valor de cada pixel de la imagen se le asignó su orden correspondiente, dando como resultado un suavizado en las curvas como se muestran en las Figuras 5.24 y 5.25, ya que la diferencia del valor entre los puntos es menor que cuando no se usa la correlación de Spearman.

Las conclusiones son que se probó el sistema digital de correlación invariante usando un filtro no lineal con el máximo valor del factor de no linealidad $k=0.3$, y la correlación invariante no lineal funcionó eficientemente para la discriminación de diferentes objetos. Este método se aplicó usando filtros no lineales y sólo de fase para diferentes objetos rotados y escalados, la letra E en tipo de fuente Times New Roman, y se encontró un mejor desempeño PCE para los filtros no lineales. El desempeño DC para el objetivo inmerso en ruido gaussiano aditivo y en ruido S&P fue mejorado significativamente aplicando el SDISNF.

§ 5.3 Sistema digital de correlación invariante aplicado a la comparación de dos fuentes de letras

En esta sección se presenta un análisis de las letras de tipo fuente Times New Roman y Arial en estilo plano usando la metodología del sistema digital de correlación invariante de la Figura 4.3 con un filtro no lineal con un factor de no linealidad $k=0.3$ y un filtro sólo de fase (Coronel-Beltrán y Álvarez-Borrego, 2009; Pech-Pacheco *et al.*, 2003). Los valores de los picos de correlación con estos tipos de filtros, están dados por la métrica de desempeño PCE. El tamaño de las imágenes es de 512x512 píxeles y con formato bitmap de 256 colores. Los resultados muestran un mejor desempeño para el sistema que utiliza el filtro no lineal que para el que utiliza el filtro sólo de fase. Las letras de fuente Times New Roman mostraron valores más altos en sus picos de correlación comparadas con las de fuente Arial.

Los resultados se obtuvieron con un 95% de intervalo de confianza. En la Figura 5.26 se muestra la gráfica de cajas para el *pico de correlación energía* PCE *versus* las letras del alfabeto de fuente Arial y Times New Roman con filtros no lineal y sólo de fase con valor de la media con un error estándar ($\pm EE$) y dos errores estándar ($\pm 2*EE$) para el PCE. Se puede observar que el desempeño del filtro no lineal fue superior por aproximadamente un factor de diez comparado con el filtro sólo de fase. Y para el filtro no lineal se muestra que las letras de fuente Times New Roman tuvieron un valor promedio a un error estándar de PCE superior a las letras de fuente Arial. Para el caso del filtro sólo de fase, se observa que no hubo diferencias significativas del valor del PCE para ambas fuentes de letras, prácticamente tuvieron los mismos valores.

En la Figura 5.27 se muestra la gráfica de cajas para el *pico de correlación energía* PCE *versus* la letra E de fuente Arial y Times New Roman rotada $\pm 90^\circ$, con incrementos de 5° , con filtros no lineal y sólo de fase con valor de la media con un error estándar ($\pm EE$) y dos errores estándar ($\pm 2*EE$) para el PCE. Se puede observar que con el filtro no lineal, para las fuentes de ambas letras, los valores del PCE estadísticamente no son estadísticamente diferentes. Para el filtro sólo de fase, se observa un valor mayor del PCE para la letra de fuente Arial que para la Times New Roman. Sin embargo, el filtro no lineal tuvo un mayor desempeño que el filtro sólo de fase.

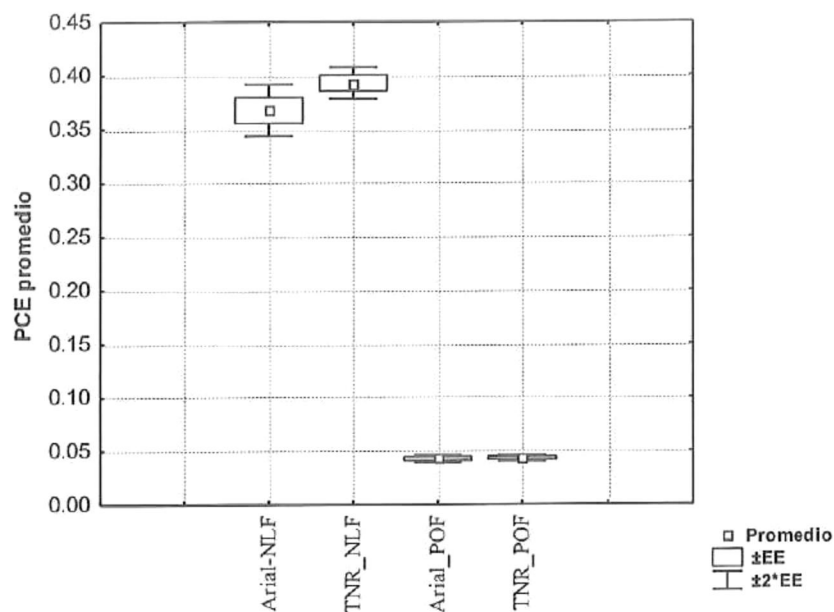


Figura 5.26. Gráfica de cajas para el *pico de correlación energía* PCE *versus* las letras del alfabeto de fuente Arial y Times New Roman (TNR) con filtros no lineal (NLF) y sólo de fase (POF).

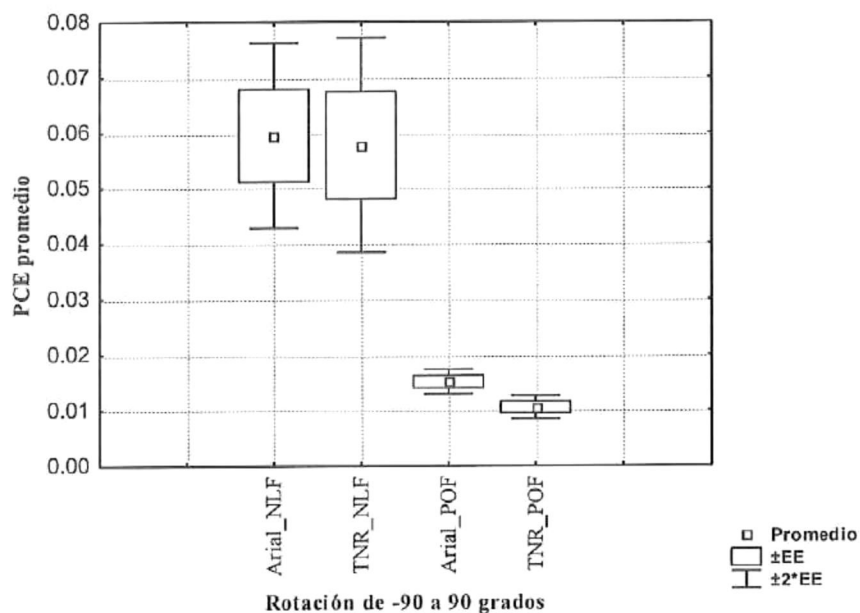


Figura 5.27. Gráfica de cajas para el *pico de correlación energía* PCE *versus* la letra E de fuente Arial y Times New Roman (TNR) rotada $\pm 90^\circ$, con incrementos de 5° , con filtros no lineal (NLF) y sólo de fase (POF).

En este trabajo se presentó un sistema digital de correlación invariante aplicado al estudio de dos fuentes de letras, Arial y Times New Roman, en estilo plano. Se utilizaron dos tipos de filtros, no lineal y sólo de fase, para analizar los efectos de estas fuentes de letras, así como comparar su desempeño. Los resultados mostraron que el filtro no lineal tuvo un desempeño considerablemente mejor comparado con el filtro sólo de fase y se obtuvo un valor promedio mayor a un error estándar para el PCE para las letras de fuente Times New Roman comparadas con las de fuente Arial con este tipo de filtro. Para rotación se utilizó la letra E en Arial y en Times New Roman y no se encontraron diferencias significativas en los valores del PCE al utilizar el filtro no lineal, más no así cuando se utilizó el filtro solo de fase.

El trabajo futuro puede ser estudiar la tipografía de otras fuentes y estilos de letras siguiendo esta misma metodología la cual sería importante para el diseño de páginas Web y textos escritos, entre otros.

§ 5.4 Análisis comparativo entre diferentes tipos de fuentes y estilos de letras usando correlación digital invariante

En esta sección se presenta un análisis comparativo de las letras en tipos de fuente Times New Roman (TNR), Courier New (CN) y Arial (Ar), en estilos plano y cursiva y los efectos de cinco combinaciones de color de frente/fondo de las letras usando un sistema de correlación digital invariante con un filtro no lineal con $k=0.3$, de acuerdo a la metodología mostrada en la Figura 4.3. La evaluación del plano de salida con este filtro está dado por la métrica PCE. Los resultados obtenidos muestran que las letras en fuente TNR tienen un mejor valor del PCE medio cuando estas se comparan con las de fuentes CN y Ar. Estos resultados están en concordancia con algunos estudios que se habían realizado sobre la legibilidad y habilidad de lectura, donde a varios individuos participantes en esos experimentos se les midió el tiempo de reacción (RT) al leer un texto. Con base en ese parámetro de medida, en este trabajo se llegó a concluir que la métrica PCE es inversamente proporcional al RT.

5.4.1 Tipografía de las fuentes de letras TNR, Ar y CN

Las diferencias principales en las características de las fuentes de letras, son los dos tipos llamados *serif* y *sans serif*. La fuente Arial es la forma más simple, tiene un trazo de ancho uniforme y la altura de las letras es más grande, al contrario de la fuente Times New Roman que tiene un trazo de ancho variable con serifs y la altura de las letras es más pequeña. La fuente Courier New es una tipografía serif espaciada en bloques (Figura 5.28). El comportamiento de estas letras ha sido estudiado desde hace algunos años hasta hoy en día (Hill, 1997) con varios individuos participantes a quienes se les mide el tiempo de reacción (RT) en la legibilidad y habilidad en la lectura de un texto escrito en diferentes fuentes y estilos de letras. En este trabajo de tesis, se hace una comparación de las fuentes Arial (*sans serif*), Courier New (*slab serif*) y Times New Roman (*serif*) en tipos plano y cursiva con cinco combinaciones de colores diferentes y se investiga la relación que existe entre el RT y los valores del PCE medio aplicando el método de correlación digital invariante a estas letras del alfabeto. Tipográficamente, la cursiva es un estilo de escritura cuyas

características más comunes son la inclinación de sus letras y la concatenación de las mismas en una palabra; por extensión, se llama cursiva (también *itálica*) a todo tipo de letra inclinada. La palabra “cursiva” proviene de *curro* (-is, -ere, *cucurri*, *cursum*), ‘correr’ en latín, por ser el tipo de letra obtenida al escribir con cierta agilidad. Contrario a la letra “plana” que no tiene inclinación.

Arial

Courier

<u>Categoría</u>	Transitional <u>sans-serif</u>
Fecha de creación	<u>1982</u>
Diseñador(es)	<u>Robin Nicholas</u> <u>Patricia Saunders</u>
<u>Empresa</u>	<u>Monotype</u>

<u>Categoría</u>	<u>Monospaced slab serif</u>
Fecha de creación	<u>1955</u>
Diseñador(es)	<u>Howard Kettler</u>
<u>Empresa</u>	<u>IBM</u>

Times New Roman

<u>Categoría</u>	Transitional <u>serif</u>
Fecha de creación	<u>1931</u>
Diseñador(es)	<u>Stanley Morison</u> <u>Starling Burgess</u> <u>Victor Lardent</u>
<u>Empresa</u>	<u>Monotype</u>

Figura 5.28. Tipografía de tres fuentes de letras.

5.4.2 Relación del reconocimiento de patrones vía retino cortical y el óptico digital

Desde el punto de vista de la fisiología, el cerebro tiene la capacidad de detectar y reconocer la forma de los objetos, independientemente de su tamaño y orientación. La percepción del mundo tridimensional a través del sistema visual, en humanos y en algunos

animales, es un proceso natural, pero los procesos mentales *cognitivos* involucrados son únicos en humanos. La luz proveniente de una escena visual es enfocada por la córnea y el cristalino del ojo sobre la retina, produciendo una reacción química en los fotorreceptores conos y bastoncitos, y después de procesar la luz en la retina, los fotones son convertidos en señales eléctricas y transmitidas a través del nervio óptico hasta llegar a la corteza visual primaria (también llamada de diferentes formas, corteza estriada, área 17 de Brodmann, o área visual 1, abreviada como V1) (Fischl, 1997).

Hace varias décadas, se hicieron algunos estudios con animales (Fisher, 1973) para encontrar un tratamiento analítico que describiera la transformación de la escena visual desde la retina hasta la corteza visual del cerebro, por ejemplo, en gatos. Las primeras funciones matemáticas para describir este mapeo fueron dadas a conocer por Schwartz (1977a) estudiando la corteza estriada en primates. Estos estudios mostraron que entre la retina y la corteza visual existe un mapeo polar logarítmico complejo, donde cada célula retinal está unida con una neurona en la corteza visual cerebral, y que este proceso está gobernado por las mismas funciones analíticas usadas en el reconocimiento de patrones óptico y digital (Schwartz, 1977b, 1980). Este descubrimiento fue muy importante para el desarrollo de nuevos métodos de reconocimiento de patrones óptico y digital. En esa época, algunos estudios en el reconocimiento de patrones óptico (Casasent y Psaltis, 1976a) mostraron que las propiedades de invariancia a rotación y escala del mapeo cortical eran las bases esenciales de la correlación para la estereopsis¹⁹ (Schwartz, 1977a). Varios años después se realizaron estudios de un mapeo en directo de la corteza visual en humanos utilizando tomografía de positrones (Schwartz, *et al.*, 1984). Se puede decir que estos estudios fueron cruciales para nuevos desarrollos de sistemas de reconocimiento de patrones ópticos, digitales e híbridos.

¹⁹ La *estereopsis* (de *stereo* que significa sólido, y *opsis* visión o vista) es el proceso dentro de la percepción visual que lleva a la sensación de profundidad a partir de dos proyecciones ligeramente diferentes del mundo físico en las retinas de los ojos. A esta diferencia en las dos imágenes retinianas se le llama disparidad horizontal, disparidad retiniana o disparidad binocular, y se origina por la diferente posición de ambos ojos en la cabeza. La estereopsis es una de las vías binoculares para la percepción de la profundidad junto con otras de carácter monocular. Obtenido de <http://es.wikipedia.org/wiki/Estereopsis>

Para la mayoría de los sistemas digitales el proceso es muy diferente, debido a la gran complejidad de la composición natural del sistema visual retino cortical. Las imágenes digitales son obtenidas por sensores ópticos-digitales en condiciones ambientales diferentes y variantes, lo cual causa severos problemas para la identificación y caracterización del objeto utilizando sistemas de visión computarizada, y el estudio está limitado a objetos estáticos e imágenes bidimensionales.

5.4.3 Resultados experimentales

Se obtuvieron algunos resultados usando nuevos algoritmos computacionales invariantes para el reconocimiento de las letras del alfabeto en diferentes fuentes y estilos y los efectos de cinco combinaciones de colores. En los experimentos digitales que se presentan en este trabajo se usó un filtro no lineal y letras mayúsculas con cero grados de rotación y sin variación en la escala. Los resultados se obtuvieron con 95% de nivel de confianza estadístico.

Los resultados, usando el sistema invariante, para las letras de fuente Arial y TNR se muestran en la Figura 5.29. Las imágenes son auto-correlacionadas, tomando a cada una de las letras del alfabeto. A diferencia de la Figura 5.27, ahora estamos comparando todas las letras del alfabeto y no solo una. De la Figura 5.29 se observan mejores resultados usando un filtro no lineal (NLF) que un POF para las letras de fuente Ar y TNR. En este caso, todas las letras del alfabeto fueron de color negro con fondo blanco (negro/blanco). La diferencia en los valores del PCE entre estos filtros fue notable: el NLF fue mejor que el POF por aproximadamente un factor de diez. Esa fue la razón por la cual se escogió el NLF en vez del POF en el sistema invariante. En general, usando el NLF, los valores del PCE fueron mejores para las letras de fuente TNR que para las letras de fuente Ar. Se puede observar que para las letras C, O, Q, S y W, los valores del PCE para las de fuente Ar fueron ligeramente mayores que para las de fuente TNR. Y para las letras A, G y V, los valores del PCE fueron los mismos para ambas fuentes de letras. Esto es debido a sus propias características de extensión y estructura de las letras.

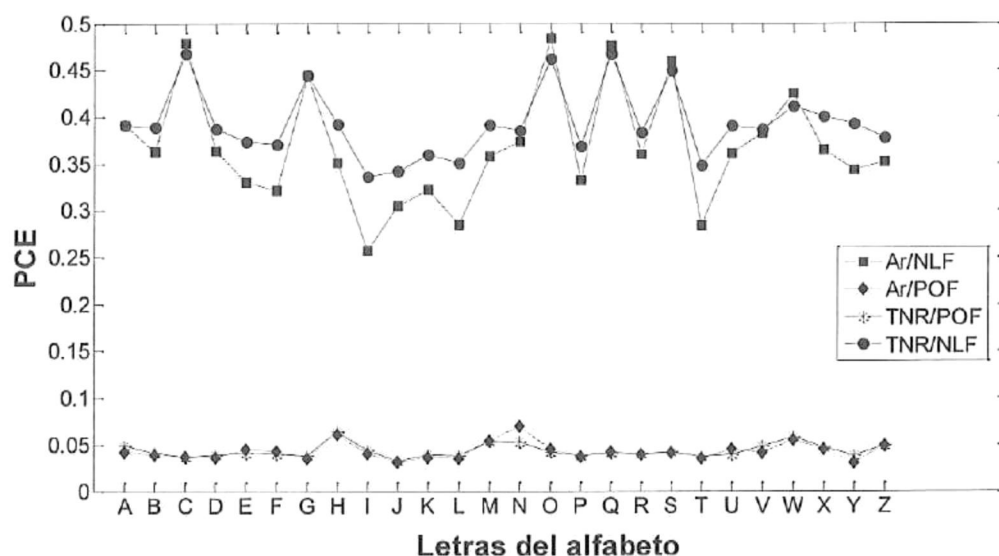


Figura 5.29. Valores del PCE para cada una de las letras del alfabeto en fuentes Arial (Ar) y en Times New Roman (TNR) usando un filtro no lineal (NLF) y un filtro sólo de fase (POF).

La Figura 5.30 muestra los valores obtenidos del PCE promedio versus las letras en tipos de fuente Arial (Ar), Courier New (CN) y Times New Roman (TNR) y en estilos plano (plain) y cursiva (italicized), con cinco combinaciones de colores de frente/fondo; negro/blanco (BK/W), verde/amarillo (GN/Y), rojo/verde (R/GN), blanco/azul (W/BL) y amarillo/azul (Y/BL), usando un sistema de correlación invariante con un filtro no lineal. Los colores del frente/fondo fueron seleccionados con las coordenadas de color RGB(r,g,b), normalizados a [0-255](bytes), como sigue; rojo: (255,0,0), verde: (0,128,0), azul: (0,0,255), amarillo: (255,255,0), negro: (0,0,0) y blanco: (255,255,255). Las imágenes son de 512x512 pixeles de tamaño y de formato bitmap con 256 colores. Los resultados muestran que las letras en fuente TNR tienen un mejor PCE promedio cuando es comparado con las de fuentes Ar y CN.

Para el estilo cursivo se puede observar de esta gráfica que la fuente TNR tuvo un valor más grande del PCE promedio comparado con las fuentes Ar y CN. Todas ellas en este estilo tuvieron el mismo desempeño y las combinaciones de colores fueron irrelevantes.

Para el estilo plano se puede observar que la fuente TNR tuvo también un valor más grande del PCE promedio comparado con las fuentes Ar y CN. Sin embargo, contrario al estilo cursivo, en el estilo plano se encontró que las combinaciones de colores afectan el valor del PCE promedio de las letras de estilos de fuente. Las combinaciones de colores negro/blanco y blanco/azul tuvieron los valores más altos del PCE promedio. Para las combinaciones de colores verde/amarillo, rojo/verde y amarillo/azul, no hubo cambios significativos en los valores del PCE promedio, y como en el caso de las cursivas, su comportamiento fue el mismo, y las combinaciones de colores fueron irrelevantes y su PCE promedio no fue estadísticamente significativo.

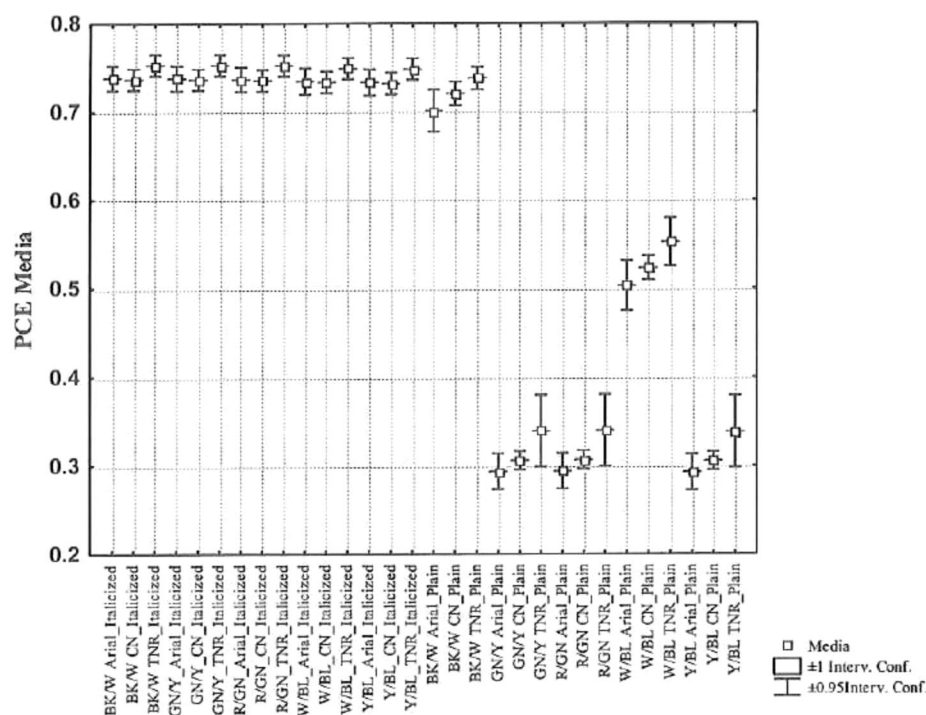


Figura 5.30. Desempeño de las letras en fuentes Arial (Ar), Courier New (CN) y Times New Roman (TNR) en estilos plano y cursivo, con cinco combinaciones de colores de frente/fondo.

Para comparar las fuentes Ar y TNR, se escogieron a todas ellas agrupándolas en estilo plano. La gráfica de cajas en la Figura 5.31 muestra que la fuente TNR tuvo valores más grandes del PCE promedio comparado con la fuente Ar.

La siguiente tarea fue comparar todas las fuentes Ar y TNR agrupadas en estilo cursivo. La gráfica de cajas de la Figura 5.32 muestra un mejor valor del PCE promedio para todas las fuentes TNR en estilo cursivo comparada con la fuente Ar. Y finalmente, Figura 5.33, se comparan todas las fuentes Ar y TNR en estilos plano y cursiva, y en este caso, del conjunto de todas las letras, Ar y TNR, el estilo cursiva tuvo un mejor valor del PCE promedio.

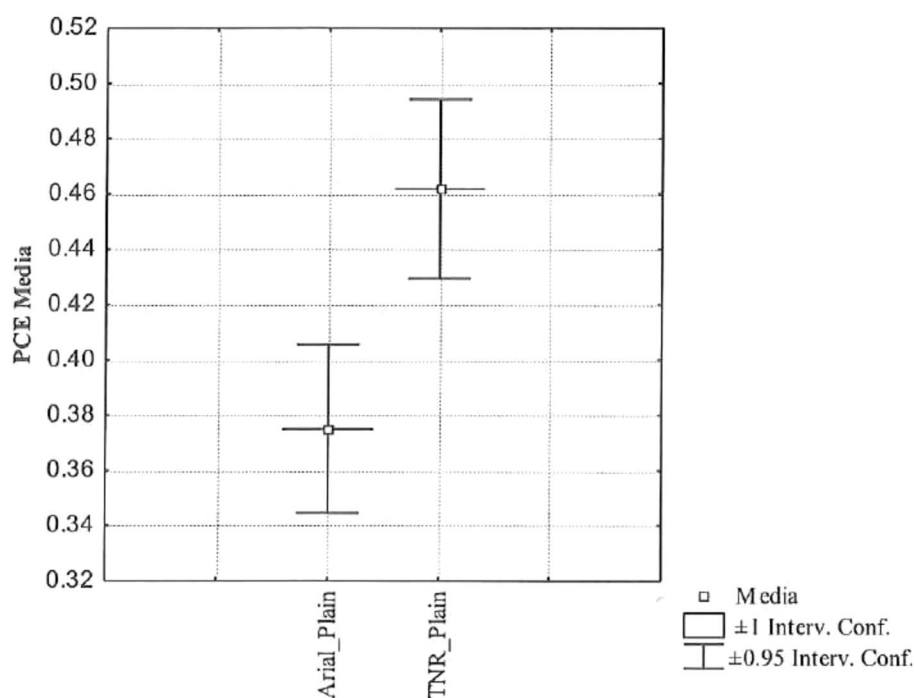


Figura 5.31. Valores del PCE de las fuentes Arial (Ar) y Times New Roman (TNR) en estilo plano.

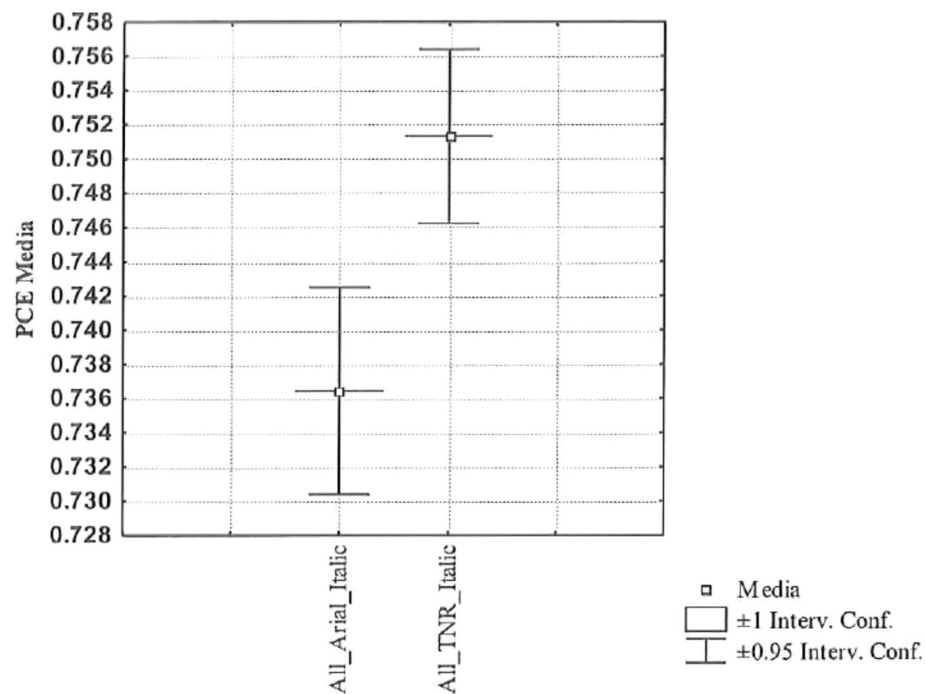


Figura 5.32. Valores del PCE de las fuentes Arial (Ar) y Times New Roman (TNR) en estilo cursivo.

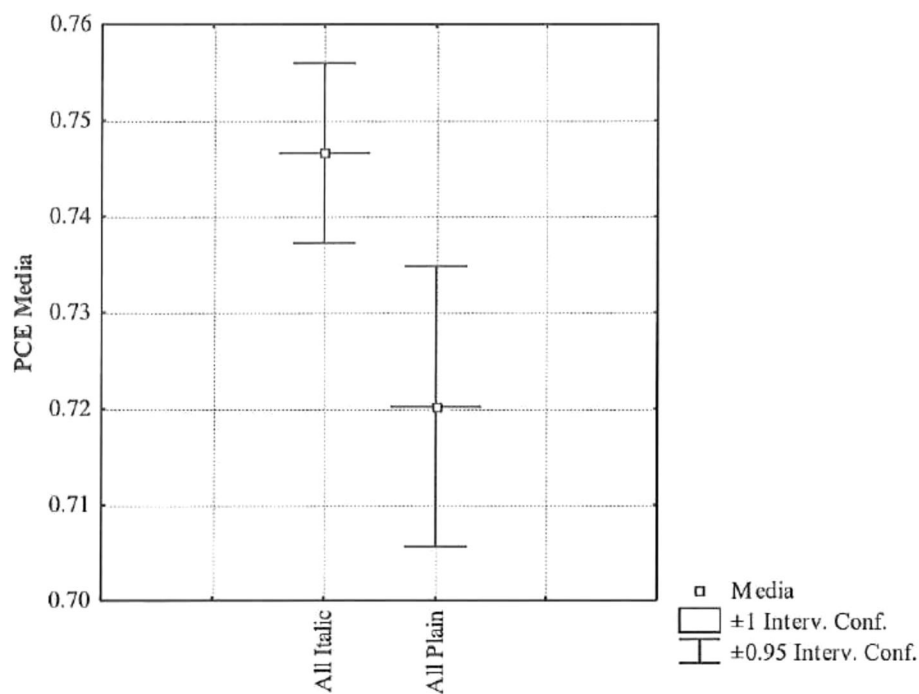


Figura 5.33. Valores del PCE de todas las fuentes Arial (Ar) y Times New Roman (TNR)

en estilos plano y cursivo.

Al comparar los resultados del PCE Medio con los resultados del RT Medio, para las letras del alfabeto de fuente Ar y TNR con cinco combinaciones de colores de letra/fondo en estilos plano y cursivo, observamos que el PCE es inversamente proporcional al RT como se muestra en la Tabla 5.3 y en la Figura 5.34 para el estilo plano.

Tabla 5.3. Comparación entre el PCE y el RT de las letras del alfabeto en fuente Ar y TNR con cinco combinaciones de colores de letra/fondo en estilos plano y cursivo. El RT está dado en milisegundos.

Estilo: Plano

Fuente	¹ PCE Medio	² RT Medio (ms)
Ar	0.375	16884.28
TNR	0.462	16856.00

Estilo: Cursivo

Fuente	¹ PCE Medio	² RT Medio (ms)
Ar	0.736	18517.42
TNR	0.751	16637.74

Estilo: Plano + Cursivo

Fuente	¹ PCE Medio	² RT Medio (ms)
Ar	0.556	17700.85
TNR	0.607	16746.87

¹Coronel-Beltrán, A. and Álvarez-Borrego, J. 2010.

²Hill, A. L., 1997.

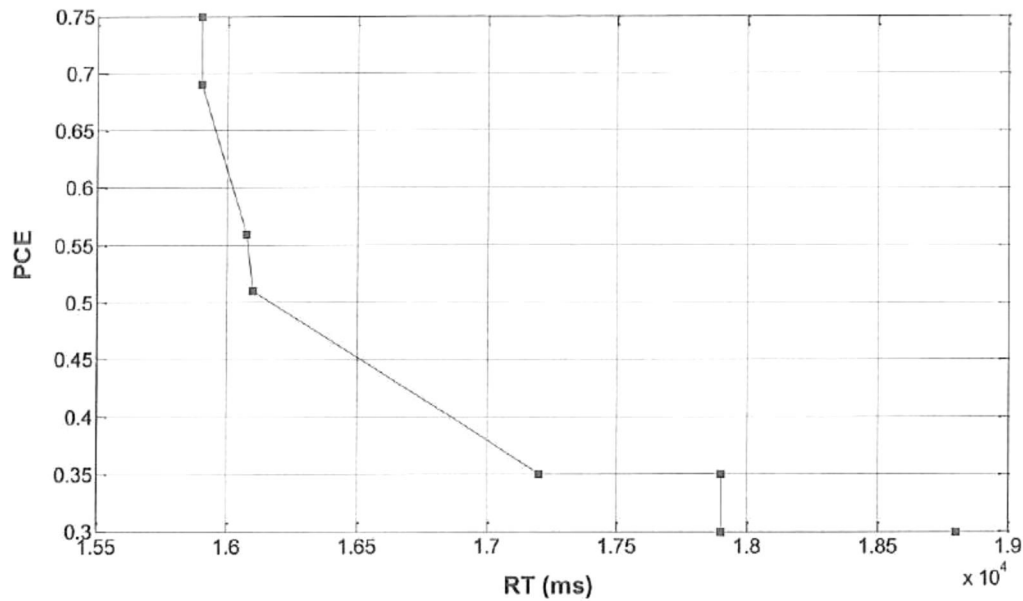


Figura 5.34. Gráfica de PCE *versus* RT (ms) de las letras del alfabeto en fuente Ar y TNR con cinco combinaciones de colores de letra/fondo en estilo plano.

5.4.4 Discusión de los resultados

El método de correlación digital invariante utilizado en este trabajo es más general, y es usado para la correlación invariante a los cambios en rotación y escala del objeto. Sin embargo, en el estudio que se hace aquí, las letras no tienen variaciones de rotación y escala, debido a que en la gran mayoría de los textos escritos, sitios de la red, correo electrónico, etc., no hay cambios de rotación o de escala de las letras. Por lo que, en este caso, el problema es más fácil de resolver con el método aquí propuesto, donde solamente la auto-correlación, con cero grados de rotación y sin cambios de escala, es tomada en cuenta.

Los efectos de las combinaciones de colores de las letras y el fondo, fuentes, estilo, tamaño del pixel, y otras variables, son muy importantes en la legibilidad y la habilidad de la lectura en textos. Se realizaron estudios (Hill, 1997) midiendo el RT de varios individuos participantes en la lectura de un texto, los efectos de las combinaciones de seis colores, y tres tipos de fuentes (Arial, Courier New y Times New Roman) en estilos plano y cursivo y

los resultados mostraron que para verde/amarillo, en cursiva, la fuente TNR tuvo mejores valores de RT que para la fuente Arial. Para todas las seis combinaciones de colores en los estilos plano y cursivo, la fuente TNR tuvo valores de RT más pequeños que para la fuente Arial, excepto para negro/blanco y rojo/verde en estilo cursivo, y para negro/blanco y blanco/azul en estilo plano. En general, se encontró que las palabras escritas en estilo plano tuvieron un mejor RT que las palabras escritas en estilo cursivo. Adicionalmente, se ha estudiado (Pastoor, 1990) el rendimiento de la legibilidad y la preferencia subjetiva para la combinación de colores en textos y no se encontraron evidencias para una influencia de polaridad de luminancia o cromaticidad en ese rendimiento. También se ha estudiado (Greco *et al.*, 2008) de qué manera las combinaciones de colores pueden mejorar la legibilidad y la comodidad de lectura de textos en una pantalla de computadora o proyectada sobre una superficie, y estos resultados mostraron que es mejor tener un texto oscuro sobre un fondo luminoso. Otros estudios (Shieh y Lai, 2008) relacionados con la combinación de colores y otras variables, tales como iluminación ambiental, contraste de luz y tipo de estímulo, a una preferencia subjetiva de una terminal de despliegue de video (VDT)²⁰, tal como el aspecto estético, la legibilidad y el confort visual, han mostrado que el negro/blanco y el azul/blanco fueron la combinación de colores más preferida con alto contraste luminoso.

En los resultados obtenidos en esta tesis, es significativo que en los estilos plano y cursivo, la TNR tuvo un mejor desempeño con valores mayores del PCE que para la Ar (Figuras 5.30-5.32). Para el análisis se eligieron solamente las letras de fuentes Ar y TNR (Figuras 5.31-5.33), debido a que estas son, en general, las más comúnmente utilizadas en la Web y en los textos de lectura. El negro/blanco y el blanco/azul fueron las combinaciones de colores “preferidas” en el sistema digital aquí utilizado. Fue muy sorprendente que los resultados obtenidos en este trabajo, basados en un método matemático analítico, están en concordancia con algunos otros estudios de carácter subjetivo. Se encontró que la salida, dada en valores de PCE, es inversamente proporcional al tiempo de reacción del individuo lector, i.e. $PCE \propto 1/RT$, como se muestra en la Tabla 5.3 y en la Figura 5.34. Esta consideración está basada en las investigaciones realizadas en el mapeo retino-cortical en humanos (Schwartz *et al.*, 1984) y específicamente en conexión

²⁰ VDT: Video Display Terminal.

con la psicología de la *Gestalt*, porque los estudios realizados en seres humanos tienen que ver con el mapeo log polar en la corteza visual que es un requisito previo para la obtención del reconocimiento de patrones óptico y digital. Esto significa tratar con percepciones, que son fenómenos *Gestalten* (Köhler, 1947) en vez de sensaciones. Algunas características notables de la percepción son las siguientes: tridimensionalidad, curvatura, rotación, forma, y contorno, entre otras. La Figura 5.35 muestra un diagrama esquemático condensado de los diferentes campos del conocimiento que conformaron los fundamentos para respaldar el resultado que se ha encontrado aquí de la relación entre el PCE y el RT. La legibilidad y la habilidad de la lectura están directamente relacionadas, la legibilidad es inversamente proporcional al RT, y la velocidad de lectura es un criterio usado para la legibilidad. Los resultados de este trabajo pueden dar una explicación más objetiva y fiable a los actuales resultados en el proceso de la lectura.

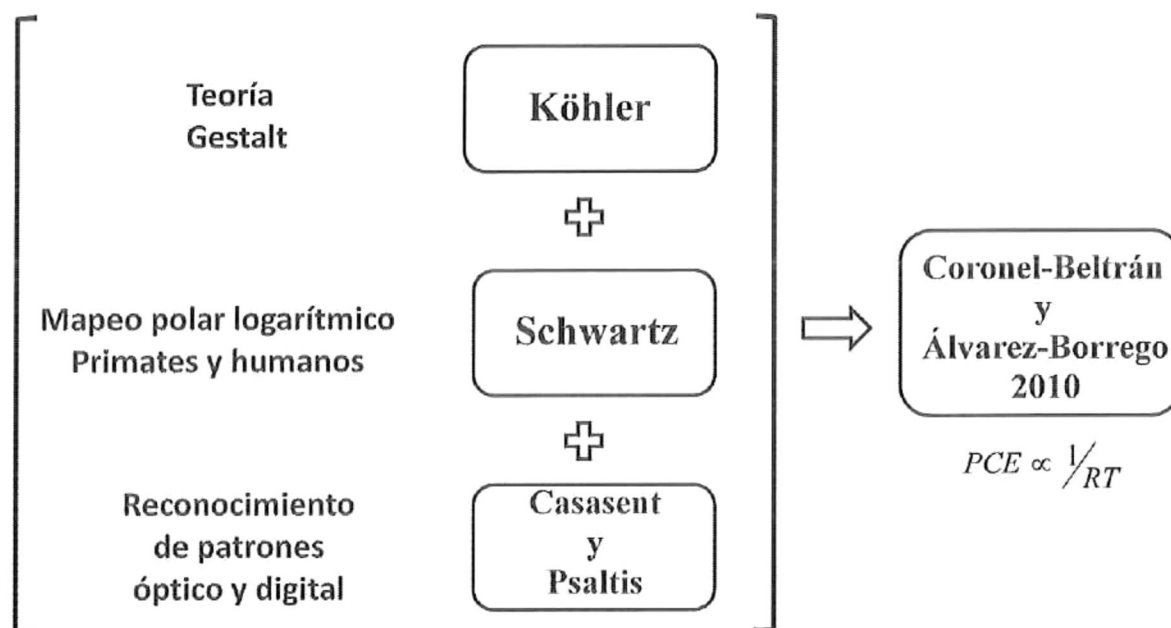


Figura 5.35. Diagrama esquemático de diferentes campos de estudio para el reconocimiento de patrones.

En conclusión, se puede decir que en este trabajo se ha presentado un nuevo método analítico –usando un sistema digital de correlación no lineal invariante – para el análisis comparativo de diferentes fuentes y estilos de letras con cinco combinaciones de colores. Los resultados obtenidos muestran una mejor salida para las letras de fuente TNR en comparación con las de fuente Ar en estilos plano y cursivo. Generalmente, la letra cursiva tuvo valores mayores del PCE que para la letra en estilo plano. Estos resultados, que estuvieron de acuerdo con otros estudios realizados con métodos subjetivos, pueden ser útiles para estudiar los efectos de la combinación de colores en el fondo de las letras, la legibilidad de un texto desplegado sobre una pantalla de computadora, de la Web, del email u otros textos escritos en libros, periódicos, revistas, etc., y su mejoramiento en la legibilidad.

§ 5.5 Reconocimiento invariante a posición, rotación y escala de imágenes digitales a color usando un filtro no lineal

En esta sección se presenta un análisis del reconocimiento de patrones de 18 imágenes a color de mariposas de distintas especies usando un sistema digital de correlación invariante con un filtro no lineal, con un factor de no linealidad $k=0.3$. Las imágenes son de 256x256 píxeles de dimensión y con formato JPEG. Cada imagen fue rotada desde 0° hasta 180° con incrementos de 1° y escalada $\pm 25\%$ con incrementos de 1%. Se utilizaron tres canales de color con el modelo de color RGB. Los resultados se obtuvieron con un 95% de intervalo de confianza para la discriminación a rotación, escala y color de cada una de las imágenes de las mariposas y con valor de la media de un error estándar ($\pm EE$) y dos errores estándares ($\pm 2*EE$). Los resultados obtenidos muestran que este sistema es capaz de reconocer imágenes de mariposas a nivel especie.

5.5.1 Modelo de color RGB

En el reconocimiento de patrones generalmente es útil introducir color para mejorar la capacidad de discriminación entre las imágenes. Se ha analizado (Attaleb *et al.*, 1999) la posibilidad de mejorar el reconocimiento óptico de patrones a color utilizando la métrica capacidad de discriminación (DC) y filtros sólo de fase (POF) basados en las componentes de RGB. El modelo de color RGB, es un arreglo de $M \times N \times 3$ píxeles de color, donde cada pixel de color es un triplete que corresponde a las componentes del rojo, verde y azul de una imagen RGB en un lugar espacial específico. Generalmente el espacio de color RGB se presenta de forma gráfica como un cubo de color RGB. Donde los colores primarios (rojo, verde, azul) y secundarios (cyan, magenta, amarillo) están representados por los vértices de este cubo (Gonzalez y Woods, 2008b) como se muestra en la Figura 5.36, donde las coordenadas de color RGB(r,g,b) están normalizadas a [0-255](bytes).

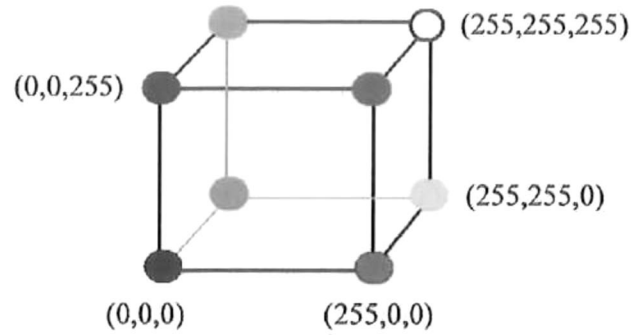


Figura 5.36. Cubo RGB representando las coordenadas de color.

En este modelo, una imagen policromática está compuesta por una combinación lineal de los tres colores primarios, los píxeles de color están representados por vectores, donde cada punto de color es un vector que se extiende desde el origen del sistema de coordenadas cartesianas (x,y) hasta un punto en el espacio RGB. Si se denota w como un vector arbitrario en el espacio de color RGB, este puede ser representado en forma de ecuación matricial como (Gonzalez y Woods, 2008c)

$$w(x,y) = \begin{bmatrix} w_R(x,y) \\ w_G(x,y) \\ w_B(x,y) \end{bmatrix} = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix}$$

Para una imagen de tamaño $M \times N$ píxeles, hay MN vectores $w(x,y)$, para $x=0,1,2,\dots,M-1; y=0,1,2,\dots,N-1$.

5.5.2 Trabajos previos

En trabajos anteriores (Millán *et al.*, 1989, 1992, 1995; Kober *et al.*, 1997), los canales de color son procesados separadamente para realizar la correlación, donde el reconocimiento de un objeto se realiza sumando los valores máximos de los picos de correlación en cada

canal. Se ha estudiado la invariancia a rotación con color multicanal usando filtros armónicos circulares (Gualdrón *et al.*, 2003). Se han obtenido mejores resultados usando correlaciones no lineales que con los filtros lineales acoplados (García-Martínez *et al.*, 2004). La utilidad de añadir color a los objetos que deseamos identificar se presenta cuando la forma o los contornos y la distribución de la intensidad luminosa no es información suficiente para poder discriminar los objetos de estudio. Un ejemplo de ello es cuando queremos reconocer una bacteria, porque la forma de las bacterias, a pesar de ser organismos muy simples, no nos provee de información suficiente ya que muchas especies comparten la misma forma, como es el caso de la bacteria *Vibrio cholerae* O1 que es igual a todo el género *Vibrio* y la familia Vibrionaceae (Mouriño-Pérez, 1998). En algunos microorganismos, como las bacterias, el color y la forma de ellos está en función de la longitud de onda con la que son iluminados, lo que significa que el color es una variable que agrega información adicional para el mejoramiento en la identificación de los objetos. Así, la descomposición de un objeto policromático en tres canales monocromáticos (rojo, verde y azul, RGB) permite realizar un reconocimiento independiente en cada uno de ellos, la suma de estos resultados produce un alto nivel de certidumbre en la identificación del objeto. Los cambios de posición, orientación y tamaño generalmente están presentes en los microorganismos así como en cualquier otro objeto en cuestión. Para vencer estos obstáculos en la identificación de objetos, es necesario utilizar un sistema de correlación invariante a posición, rotación y escala de dichos objetos, el cual puede ser óptico, digital o híbrido. Al inicio de las investigaciones del reconocimiento de imágenes policromáticas, se desarrollaron técnicas ópticas coherentes para la identificación de los objetos que difieren sólo en el color, utilizando un sistema multicanal basado en la descomposición tricromática con el modelo RGB, y del uso de filtros acoplados también llamados clásicos u holográficos de Vander Lugt (1964) para el reconocimiento de patrones mediante la correlación óptica de la imagen de entrada del objeto y de la imagen grabada en el filtro, independientemente de la posición del objeto en el plano de entrada del sistema óptico. A partir de estos trabajos, se inicia una nueva era de avances importantes en el campo del reconocimiento de patrones ópticos, se desarrollan nuevos filtros cada vez más eficientes para la identificación y conteo de objetos. El problema que se tenía hasta entonces, era que únicamente estos sistemas ópticos eran capaces de reconocer objetos a cambios de posición

en el plano de entrada pero no a variaciones de escala y orientación de los objetos debido a que se utilizaba la correlación óptica convencional. Esta desventaja se solucionó, como ya se ha mencionado en capítulos anteriores, con la introducción de la transformada de Mellin (Casasent y Psaltis, 1976a), la cual tiene la propiedad de ser invariante a los cambios de escala. Investigaciones posteriores mostraron la realización de correlaciones ópticas entre objetos que diferían en escala y rotación (Casasent y Psaltis, 1976c). Estos estudios fueron importantes para el posterior desarrollo de nuevos sistemas de reconocimiento: ópticos, digitales e híbridos. El estudio aquí presentado está basado en el reconocimiento de patrones de imágenes digitales usando un filtro no lineal invariante a rotación, escala y color que nos da información de la similitud entre objetos diferentes. En este trabajo se han obtenido algunos resultados con un nuevo algoritmo computacional para el reconocimiento de varias especies de mariposas. El sistema digital de reconocimiento de patrones se realiza mediante un proceso de correlación invariante multicanal, donde la correlación se aplica a cada canal representado por el rojo, verde y azul de la imagen a color. El resultado final del reconocimiento del objeto se obtiene por una combinación de la información de las tres correlaciones monocromáticas independientes.

5.5.3 Correlación digital no lineal invariante multicanal

Primeramente se obtiene el filtro no lineal como se muestra en la Figura 5.37, donde se tiene la imagen en color $f(x,y)$ del objeto que se quiere reconocer y la imagen problema $g(x,y)$. Estas dos funciones bidimensionales son finitas y diferentes de cero, de tamaño de 256×256 píxeles. Luego se hace una descomposición de estas imágenes en cada canal monocromático RGB, y, al igual que en el método de correlación no lineal invariante de la Figura 4.3, se calcula la transformada rápida de Fourier (FFT) del filtro y de la imagen problema. Se obtienen los módulos de las transformadas de Fourier denotadas como $|F(u,v)|$ para el objeto a reconocer y por $|G(u,v)|$ para la imagen problema, en cada canal, así se asegura que el desplazamiento de la imagen no se vea afectado en el plano de Fourier. Después de esto, se realiza un filtraje pasa altas usando un filtro parabólico al módulo de las transformadas de Fourier (Pech-Pacheco *et al.*, 2003). Esta clase de filtros

atenúa las frecuencias bajas mientras deja pasar las frecuencias altas que realza los detalles de borde. En esta parte se introduce el factor de escala dado por $\sqrt{r}$, donde r es la frecuencia espacial radial, que es indispensable para diferenciar la transformada de escala (Cohen, 1995) de la transformada de Mellin. Aquí se elige la primera transformada en vez de la segunda, debido a que es más sensible a los cambios de escala. Después, las coordenadas cartesianas son mapeadas a coordenadas polares para obtener la invariancia a rotación y luego se hace una interpolación bilineal a los primeros datos de la conversión de coordenadas para reducir la fuga de información debido al muestreo log-polar. Se hace un escalamiento logarítmico en la parte radial en coordenadas polares para realizar la invariancia a escala. Y por último, tomando la FFT se obtiene un filtro invariante a posición, rotación y escala, el cual está representado en la Figura 5.37, donde $|T|$ es el módulo de la transformada de Fourier del objeto a ser reconocido, el cual contiene la información no lineal del filtro, y $|G|$ es el módulo de la transformada de Fourier de la imagen problema después de obtener las invariancias, la cual contiene la información no lineal de la imagen problema. En la segunda parte de este método, se muestra el proceso de correlación digital invariante con un filtro no lineal como está representada en la Figura 5.38, donde se obtienen las tres correlaciones invariantes a rotación, escala y color para cada canal RGB, donde $\otimes$ indica correlación. Por último, se obtiene el promedio de estas correlaciones invariantes.

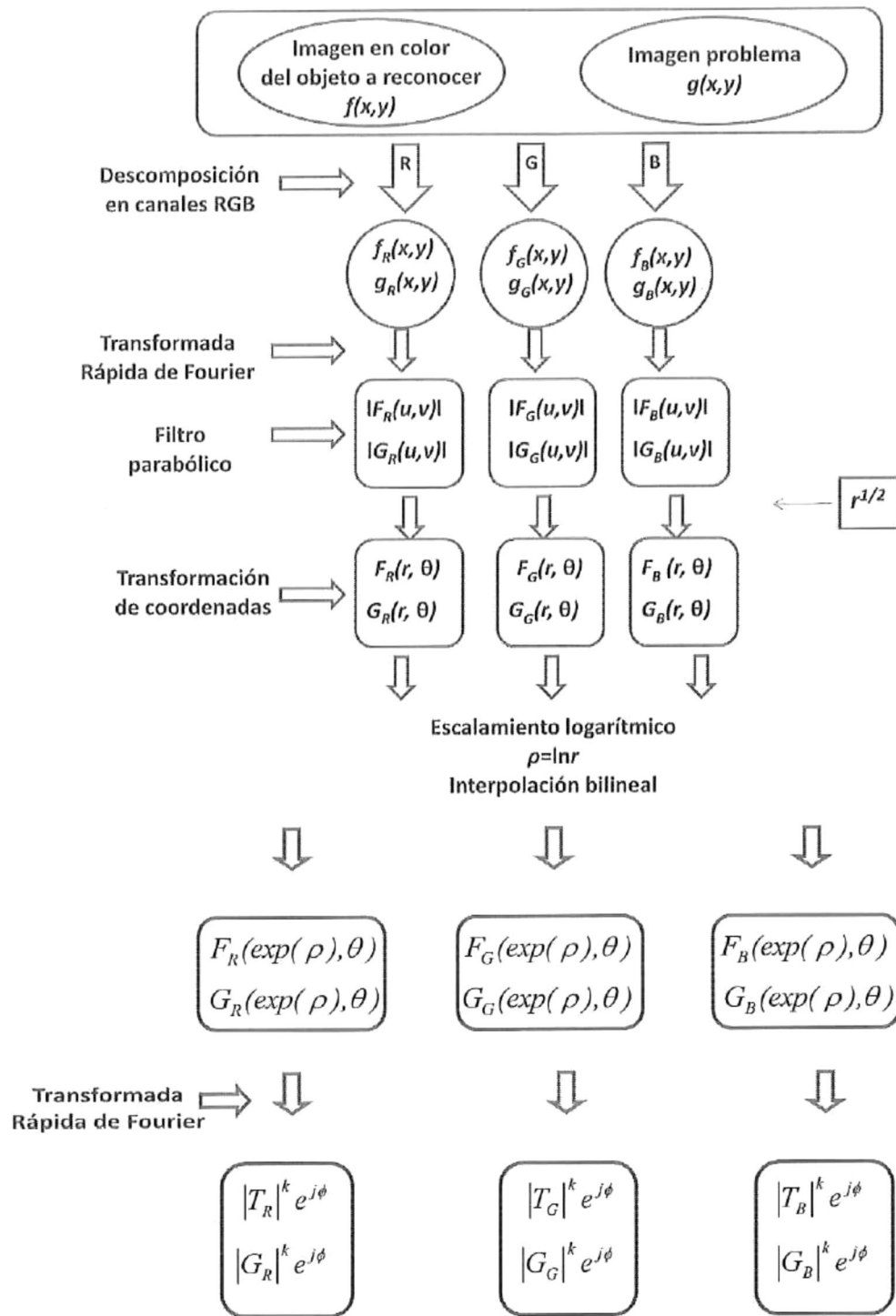


Figura 5.37. Diagrama de bloques simplificado para obtener el filtro no lineal.

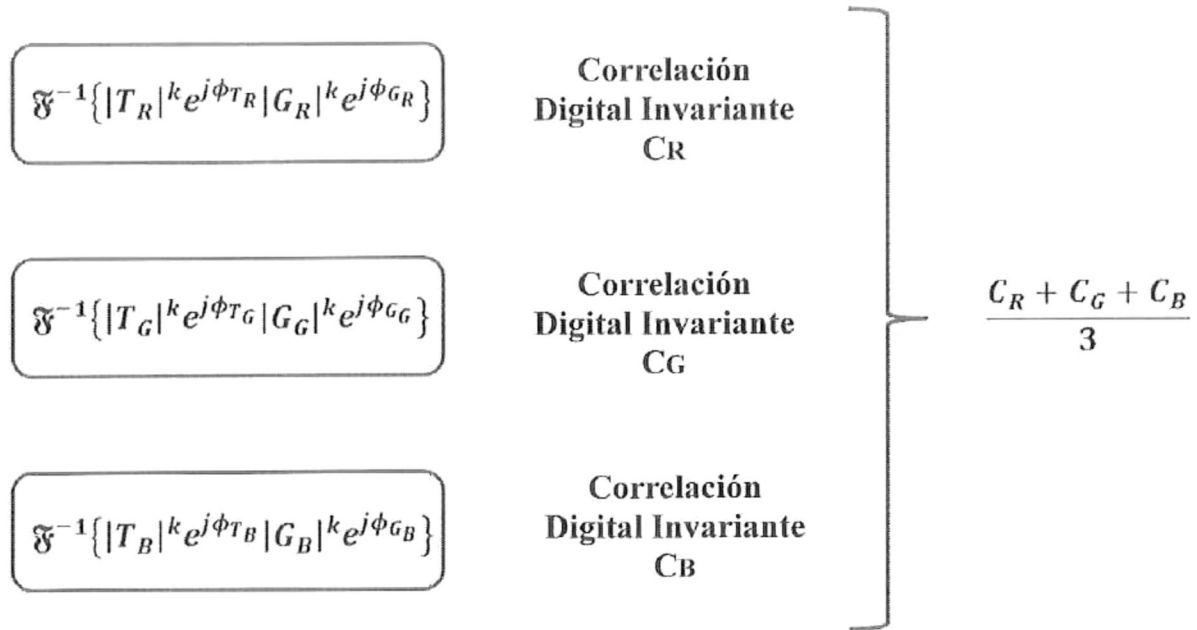


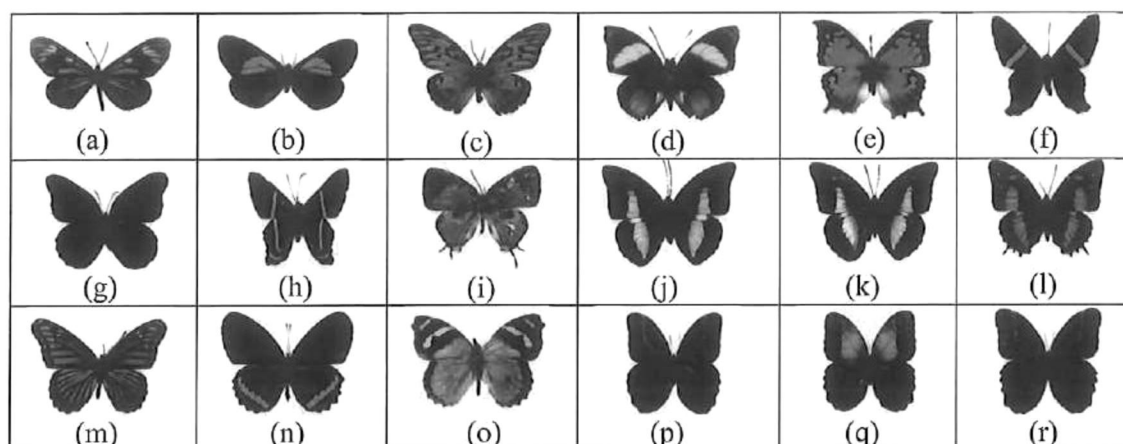
Figura 5.38. Diagrama de bloques representando al sistema de correlación digital invariante a rotación, escala y color.

5.5.4 Resultados experimentales

Los objetos estudiados fueron un conjunto de 18 especies distintas de mariposas a color de tamaño de la imagen de 256x256 píxeles y con formato JPEG, que fueron procesados en cada uno de los tres canales con el modelo de color RGB. Se obtuvo la identificación invariante a la rotación y escala de cada una de las mariposas con respecto a las demás mediante la suma de los valores máximos de los picos de correlación de cada canal monocromático. Los lepidópteros (Lepidoptera) o mariposas, son un orden de insectos que representan el segundo orden con más especies entre los insectos (superado solamente por el orden Coleoptera), existen más de 165000 especies clasificadas en 127 familias y 46 superfamilias en el mundo (de Viedma *et al.*, 1985).

En la Figura 5.39 se muestra la galería de estas mariposas estudiadas con sus nombres científicos. Los resultados se obtuvieron con un 95% de intervalo de confianza para la discriminación a rotación, escala y color de cada una de las imágenes de las mariposas. Los resultados estadísticos se muestran en las Figuras 5.40 y 5.41 con las gráficas de cajas para la correlación no normalizada para cada una de las mariposas *versus* las demás mariposas restantes con un 95% de intervalo de confianza para la discriminación a rotación, escala y color de cada una de las imágenes de las mariposas, con valor de la media con un error estándar ($\pm EE$) y dos errores estándar ($\pm 2*EE$) para la correlación, en rotación y escala, respectivamente. Cada imagen de la mariposa a reconocer fue rotada desde 0° hasta 180°, con variaciones de 1°, y escalada desde 75% hasta 125%, con variaciones de 1%. Se puede observar el excelente desempeño del filtro no lineal por su capacidad de discriminar al 100% cada una de las mariposas de todas las demás tanto en rotación y escala como se puede observar de esas figuras. De la Figura 5.39 tenemos que las mariposas (a) y (b) son del mismo género pero de diferente especie, se puede observar que estas tienen forma similar. Así también las mariposas (j) y (k) son del mismo género pero de especie diferente, presentan color y forma similares, y las mariposas (p), (q) y (r) son del mismo género pero de diferente especie, y se observa que tienen color y forma similares. En la Figura 5.40 se muestran las gráficas donde se puede observar el comportamiento estadístico para estas mariposas del mismo género en rotación las cuales nuestro sistema fue capaz de discriminarlas totalmente. Análogamente, la Figura 5.41 muestra las gráficas donde se puede apreciar también el comportamiento estadístico para

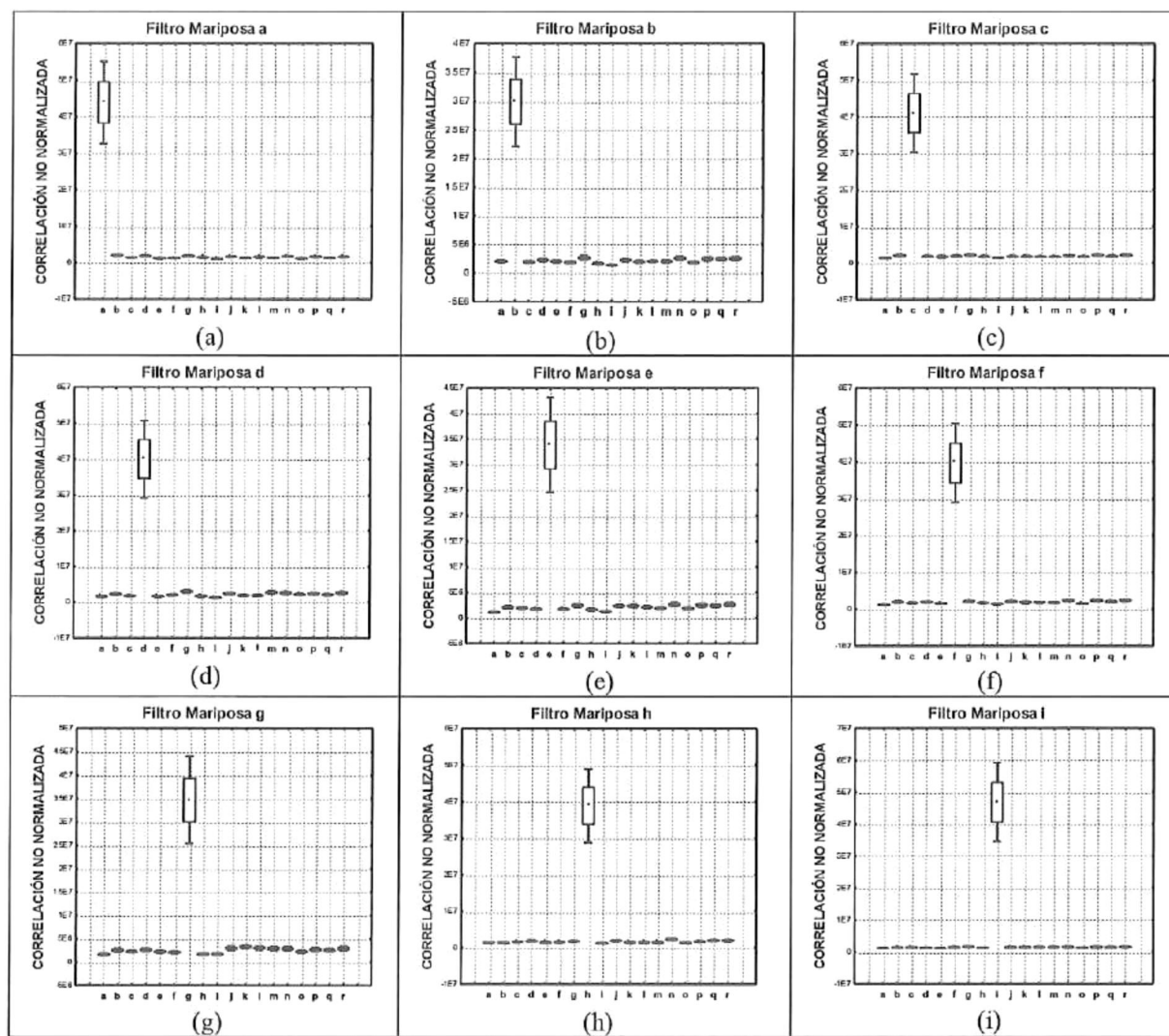
esas mismas mariposas escaladas y sin ninguna dificultad para reconocerlas. El mismo comportamiento se presentó para todas las demás mariposas que no presentaban gran parecido en color y forma entre ellas.



Fotografías: Erika Frayre

Figura 5.39. Galería de especies de mariposas utilizadas: (a) *Actinote guatemalena guerrenensis*, (b) *Actinote stratonice oaxaca*, (c) *Agraulis vanillae incarnata*, (d) *Agrias amydon oaxacata*, (e) *Anae aidea*, (f) *Ancyluris inca mora*, (g) *Anetia thirza*, (h) *Ansyluris jurgennseni*, (i) *Arawacus sito*, (j) *Archaeoprepona amphimachus amphimachus*, (k) *Archaeoprepona demophon centralis*, (l) *Baeotus baeotus*, (m) *Basilarchia archippus*, (n) *Biblis hyperia aganissa*, (o) *Bolboneura Sylphis sylphis*, (p) *Caligo eurylochus sulanos*, (q) *Caligo memnon*, (r) *Caligo oileus scamander*.

De los resultados de este trabajo, se puede concluir que el color de los objetos fue un elemento importante para poder realizar una mejor discriminación entre ellos, debido a que algunos objetos presentaban forma similar y algunos con colores parecidos y otros de forma y color semejantes. Con la separación de los colores de los objetos en los tres canales monocromáticos (rojo, verde y azul, RGB) se realizó un reconocimiento independiente en cada canal, de tal manera que la suma de ellos dio como resultado un alto nivel de certidumbre en la identificación invariante a rotación, escala y color, probando así la efectividad del algoritmo.



Continuación...

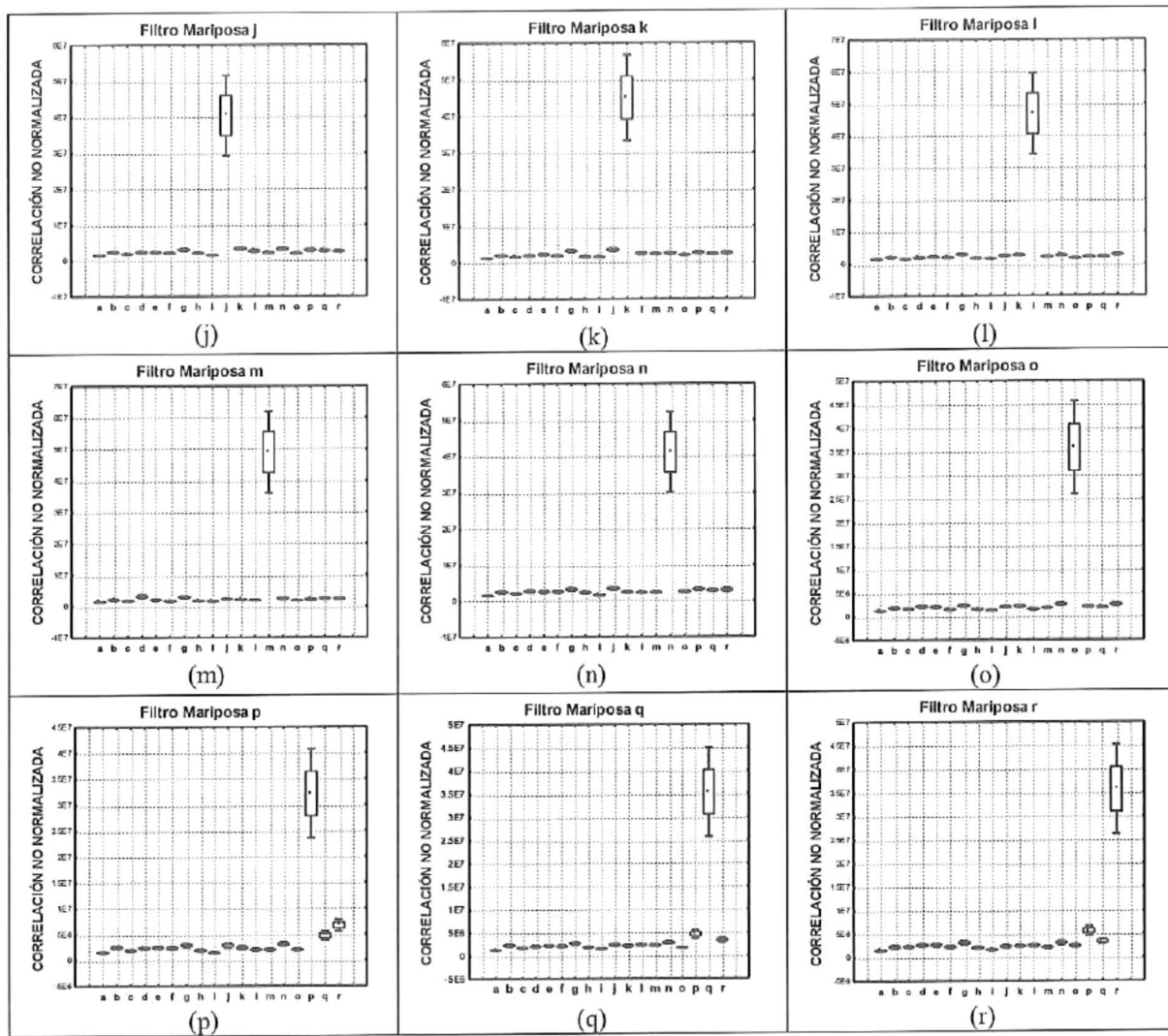
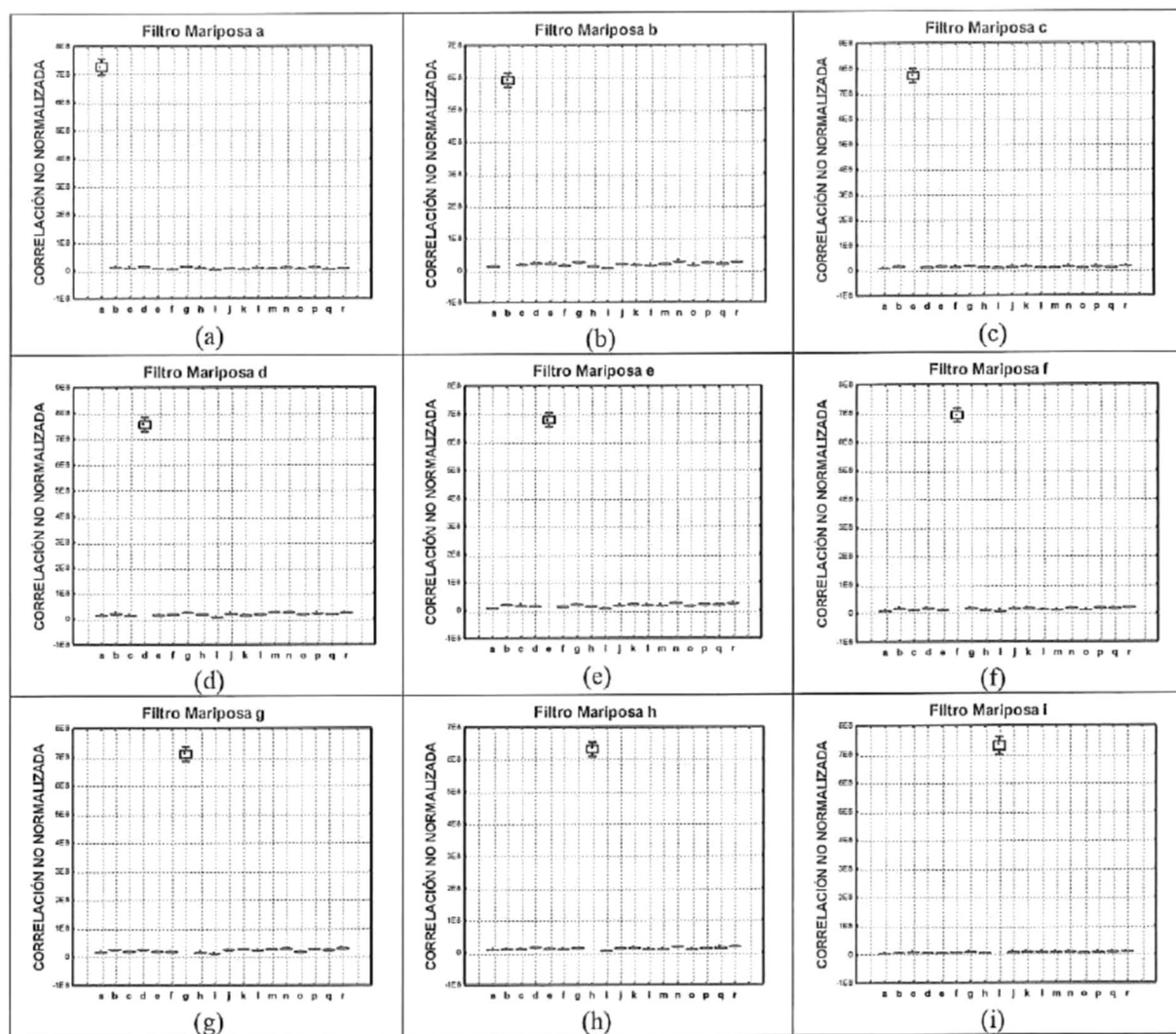


Figura 5.40. Gráficas de cajas del valor promedio de la correlación no normalizada de cada una de las mariposas, desde la (a) hasta la (r), con respecto al resto, en rotación desde 0° hasta 180° , con incrementos de 1° , donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación.



Continuación...

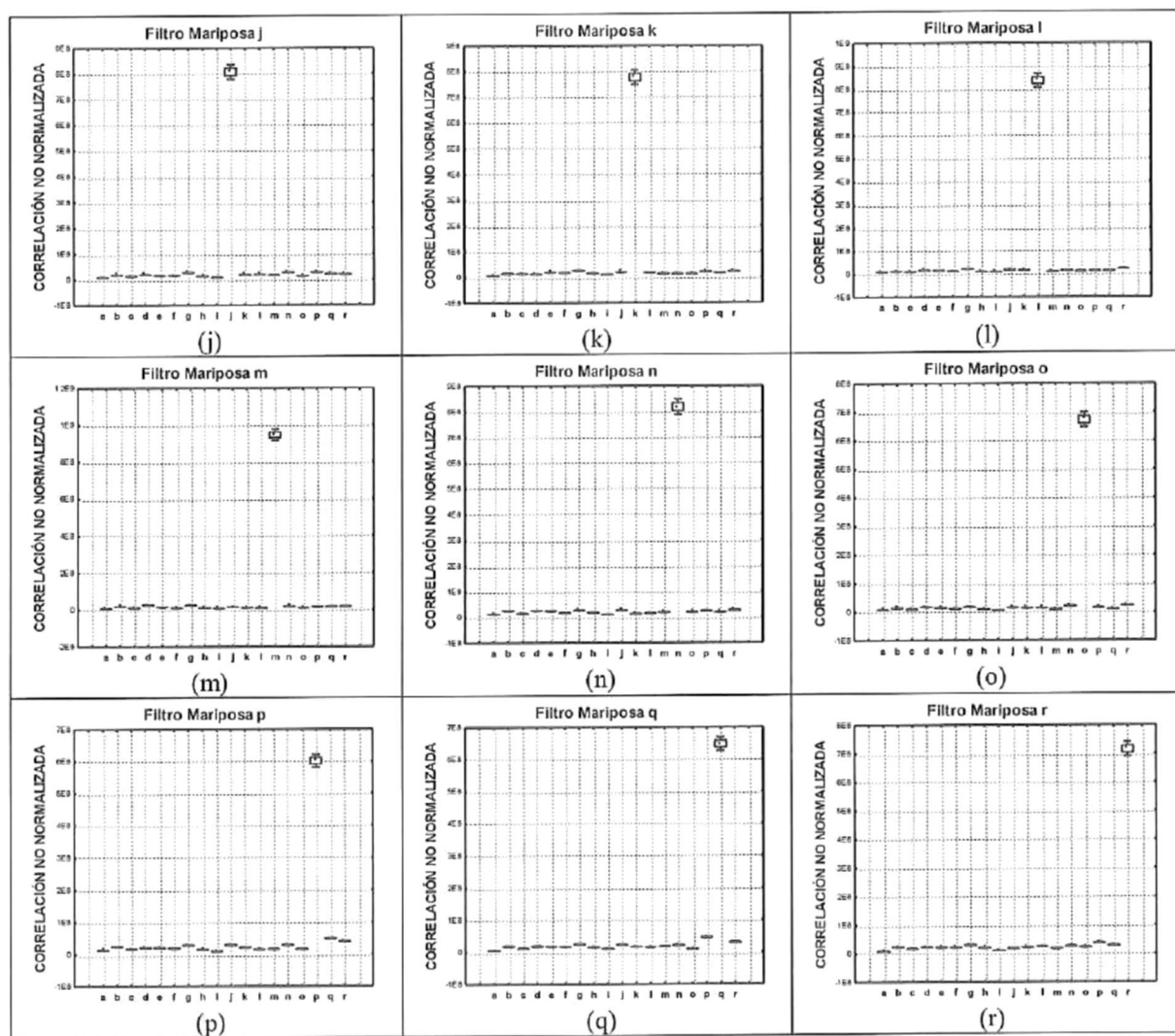


Figura 5.41. Gráficas de cajas del valor promedio de la correlación no normalizada de cada una de las mariposas, desde la (a) hasta la (r), con respecto al resto, en escala desde $\pm 25\%$, con incrementos de 1%, donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación.

En la Figura 5.42 se muestran las gráficas donde se puede observar el comportamiento estadístico para las mariposas (p), (q) y (r) del mismo género y de diferente especie en rotación, donde en la primera columna aparecen las mariposas en color y en la segunda columna las mariposas en niveles de grises. Como se puede observar, nuestro sistema fue capaz de discriminarlas totalmente en color, pero cuando las mariposas están en niveles de grises no es posible diferenciar algunas de ellas.

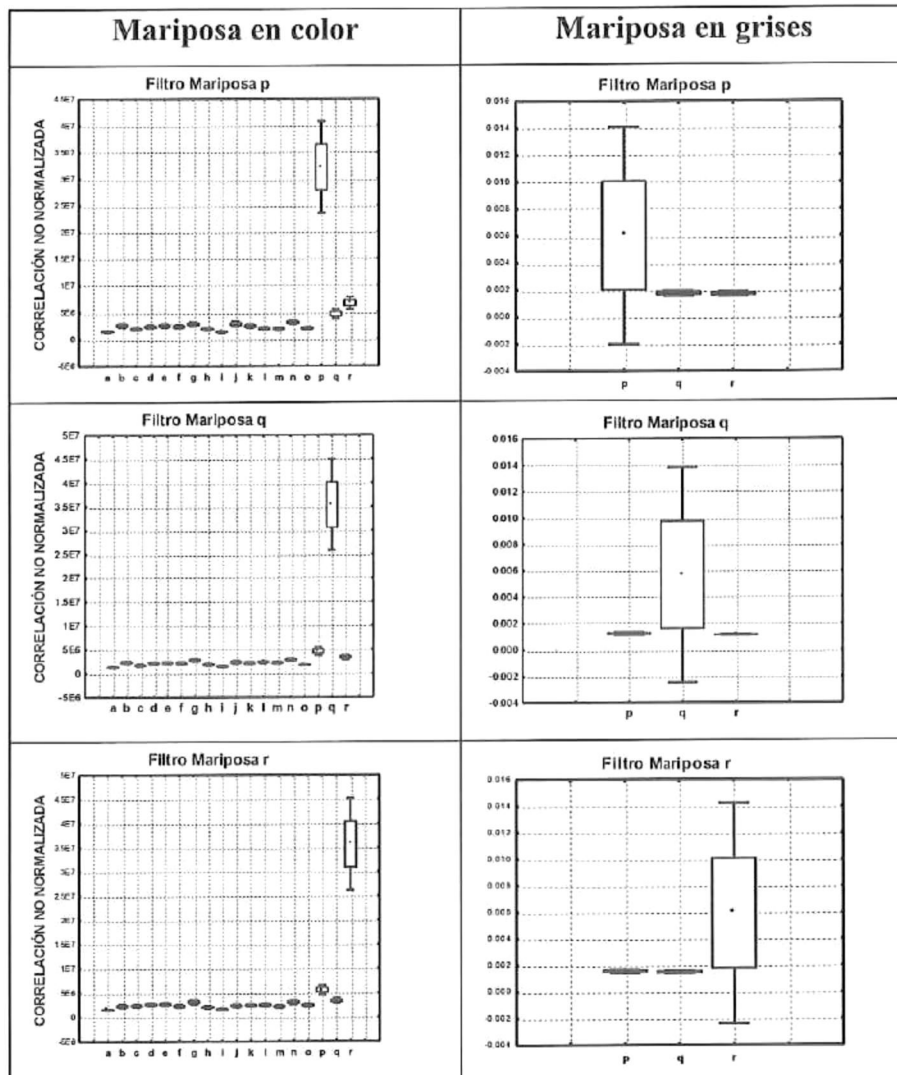


Figura 5.42. Comparación entre las mariposas en color y en grises. Gráfica de cajas de la correlación no normalizada en rotación desde 0° hasta 180° , con incrementos de 1° , donde las cajas representan un error estándar ($\pm EE$) y las barras de las cajas representan dos errores estándar ($\pm 2*EE$) para el pico de correlación.

Capítulo 6

Conclusiones

El sistema digital de reconocimiento de patrones invariante no lineal propuesto en esta tesis cumplió con los objetivos planteados. El sistema fue probado exitosamente en la identificación de diversos objetos. Su capacidad de discriminación de objetos en presencia de diferentes tipos de ruido utilizando filtros no lineales mostró un buen desempeño. En todos los trabajos realizados se obtuvo un nivel de confianza estadístico de 95%, probando así que el sistema funciona eficientemente para la discriminación de objetos.

Se determinó numéricamente el mejor valor del factor de no linealidad k , para el sistema no lineal propuesto. El sistema probó su eficacia en la discriminación de los objetos para variaciones en rotación y escala utilizando las letras del alfabeto en fuente Arial, donde cada una de estas letras se tomó como filtro y fueron correlacionadas con cada una de las demás letras.

Adicionalmente al sistema de correlación digital invariante no lineal, se introdujo la correlación de rango de Spearman a algunos objetivos inmersos en ruido gaussiano aditivo y ruido sal y pimienta. En este caso se aplicó la métrica de desempeño DC con resultados significativamente superiores. El uso de esta técnica produjo un mejoramiento considerable del sistema de reconocimiento invariante no lineal, en presencia de estos tipos de ruido.

El sistema se aplicó al análisis de dos fuentes de letras, Arial y Times New Roman, en estilo plano. Se utilizaron dos tipos de filtros, no lineal (ley k) y sólo de fase, para analizar los efectos de estas fuentes de letras, así como comparar el desempeño de estos filtros. Los resultados mostraron que el filtro no lineal tuvo un desempeño considerablemente mayor comparado con el filtro sólo de fase. Usando el sistema digital de correlación no lineal invariante, se hizo un análisis comparativo de diferentes fuentes y estilos de letras con cinco combinaciones de colores. Los resultados obtenidos mostraron

una mejor respuesta para las letras de fuente TNR en comparación con las de fuente Ar en estilos plano y cursivo; así como también, se obtuvo que la mejor combinación de colores fue para estas letras en primer lugar para negro/blanco y en segundo lugar para blanco/azul, correspondiendo a los valores más altos en el PCE promedio. Fue verdaderamente sorprendente, que estos resultados obtenidos de experimentos utilizando algoritmos computacionales con una metodología analítica basada en funciones matemáticas, estuvieran de acuerdo con otros estudios de carácter subjetivo. Fue notorio que la letra cursiva tuvo valores mayores del PCE que para la letra en estilo plano. Estos resultados, donde se encontró la relación entre el PCE y el tiempo de reacción RT, de la forma $PCE \propto 1/RT$, se justifican por los estudios realizados teórica y experimentalmente comprobados del proceso que ocurre en el mapeo de la escena visual desde la retina hasta la corteza visual V1 del cerebro en humanos y apoyada por los estudios en el campo del reconocimiento de patrones óptico y digital donde se ha mostrado que este proceso de transformación geométrica es necesario llevarlo a cabo para realizar la correlación invariante en la identificación de objetos. Esto significa que las funciones matemáticas utilizadas en los sistemas de reconocimiento óptico-digitales invariantes se manifiestan también en el mapeo retino-cortical. Desde esta perspectiva, se está tratando con percepciones, que son fenómenos “gestalten” que significa “configuraciones”, en vez de sensaciones. Por lo tanto, los resultados obtenidos en esta sección de la tesis, fueron un respaldo a otros estudios realizados con métodos subjetivos aplicados en la legibilidad y habilidad de lectura de textos escritos en individuos.

En la última aportación de los trabajos realizados, se hizo un análisis del reconocimiento de patrones de un conjunto de imágenes reales a color de mariposas de especies diferentes utilizando el sistema digital de correlación invariante con un filtro no lineal con el factor de no linealidad determinado experimentalmente, como se mencionó antes. Cada una de las imágenes a color de las mariposas fue rotada y escalada y correlacionada con las demás. Se utilizaron tres canales monocromáticos con el modelo de color RGB y los resultados obtenidos mostraron la efectividad del sistema digital de correlación invariante para el reconocimiento de patrones, logrando reconocer las imágenes de las mariposas hasta nivel especie. Las mariposas fueron del mismo género pero de especie diferente, algunas presentaban forma parecida, otras color similar y también otras

con ambos color y forma semejante. Así también para mariposas que no presentaban parecido en su color y forma. El color fue un factor importante para el mejoramiento en la discriminación de estos objetos. Con la aplicación del modelo de color RGB para separar los colores de los objetos de estudio se logró obtener una excelente identificación invariante de los objetos a rotación, escala y color, mostrando de esta manera la gran efectividad y robustez del algoritmo del sistema digital de correlación no lineal invariante.

Una de las desventajas de nuestro sistema digital de correlación no lineal invariante es el tiempo de cómputo que tarda en llevar a cabo los cálculos del mapeo polar que es del orden de cinco segundos. Otra desventaja es que con imágenes muy ruidosas nuestro sistema no funciona eficientemente como uno esperaría.

Artículos derivados del trabajo de tesis

A. Coronel-Beltrán and J. Álvarez-Borrego, "Comparative analysis between different font types and styles letters using a nonlinear invariant digital correlation," *Journal of Modern Optics*. Vol. 57, No. 1, 58–64. (2010).

A. Coronel-Beltrán and J. Álvarez-Borrego, "Invariant correlation digital system applied to the comparison of two font letters," e-gnosis, *Digital Scientific and Technological Journal ISSN 1665-5745*. (2009).

A. Coronel-Beltrán and J. Álvarez-Borrego, "Reconocimiento invariante de patrones con filtros no lineales," Memorias del LI Congreso Nacional SMF, XXI Reunión Anual AMO, Zacatecas, Zac., México (2008).

A. Coronel-Beltrán and J. Álvarez-Borrego, "Nonlinear filter for pattern recognition using the scale transform," in Applications of Digital Image Processing XXXI, Proc. of SPIE Vol. 7073 (2008).

Referencias

- Almeida, S. P., and Indebetouw, G. 1982. Applications of optical Fourier transforms. Acad. Press. Inc., 41-87.
- Almeida, S. P., and Kim-Tzong, J. 1976. Water pollution monitoring using matched spatial filters. *Appl. Opt.* 15, 510-515.
- Almeida, S. P., Case, S. K., Fournier, J. M., Fujii, H., Cairns, J. Jr., Dickson, K. L., and Pryfogle, P. 1978. Analysis of algae samples using coherent optical processing. *Proceedings of ICO-11 Conference, Madrid, Spain*, pp. 351-354.
- Almeida, S. P., Del Balzo, D. R., Cairns, J. Jr., Dickson, K. L., and Lanza, G. R. 1972. Holographic microscopy of diatoms, *Trans. Kans. Acad. Scie.* 74, 257-260.
- Alvarez-Borrego, J. and Castro Longoria, E. 2003. Discrimination between *Arcatia* (Copepoda: Calanoida) species using their diffraction pattern in a position, rotation invariant digital correlation. *Journal of Plankton Research*, Vol. 25, No. 2, pp 229-233.
- Alvarez-Borrego, J. and Chavez Sanchez, M. C. 2001. Detection of IHHN virus in shrimp tissue by digital color correlation. *Aquaculture*, Vol. 194, Issues 1-2, pp 1-9.
- Alvarez-Borrego, J., Mouriño-Pérez, R. R., Cristóbal-Pérez, G., and Pech-Pacheco, J. L. 2002. Invariant recognition of polychromatic images of *Vibrio cholerae* 01. *Opt. Eng.* 41(4), 827-833.
- Attaleb, A., Ahouzi, E., Campos, J., and Yzuel, M. J. 1999. *Opt. Rev.* 6, 42.
- Bracewell, R. N. 1978. *The Fourier transform and its applications*. McGraw-Hill, 2nd. Ed., New York, Chap.12.

Cairns, J. Jr., Dickson, K. L., and Lanza, G. R. 1972. Coherent optical spatial filtering of diatoms in water pollution monitoring. *Arch. Mikrobiol.* 83, 141-146.

Bueno-Ibarra, M. A., Chávez-Sánchez, M. C., and Álvarez-Borrego, J. (2010a). Development of a nonlinear K-law spectral signature index to classify basophilic inclusion bodies of the white spot syndrome virus. *BioSciences World*, pp. 30-33. The First International Conference on Biosciences. IEEE Computer Society.

Bueno-Ibarra, M. A., Chávez-Sánchez, M. C., and Álvarez-Borrego, J. (2010b). Nonlinear law spectral technique to analyze spot syndrome virus infection. *International Journal on Advances in Life Science*. (In press).

Casasent, D. and Psaltis, D. 1976a. Scale invariant optical correlation using Mellin transforms. *Opt. Commun.* 17, 59-63.

Casasent, D. and Psaltis, D. 1976b. Scale invariant optical transforms. *Opt. Eng.* 15, 258-261.

Casasent, D. and Psaltis, D. 1976c. Position, rotation, and scale invariant optical correlation. *Appl. Opt.* 15, 1795-1799.

Castro-Longoria, E., Alvarez-Borrego, J., and Pech-Pacheco, J. L. 2001. Identification of species of calanoid copepods using a new invariant correlation algorithm. *Crustaceana*, 74(10), 1029-1039.

Cohen, L. 1993. The scale representation. *IEEE Transactions on Signal Processing*, Vol. 41, No. 12, pp. 3275-3292.

Cohen, L. 1995. *Time-Frequency Analysis*, Prentice Hall Signal Processing Series, Chaps. 18 and 19, Upper Saddle River, NJ.

Coronel-Beltrán, A. and Álvarez-Borrego, J. 2010. Comparative analysis between different font types and letter styles using a nonlinear invariant digital correlation. *J Mod Optics*, Vol. 57, No. 1, pp. 58-64.

Coronel-Beltrán, A. and Álvarez-Borrego, J. 2009. Invariant correlation digital system applied to the comparison of two Font letters. *Digital Scientific and Technological Journal, e-Gnosis*, ISSN 1665-5742.

Coronel-Beltrán, A. and Álvarez-Borrego, J. "Nonlinear filter for pattern recognition using the scale transform", Applications of Digital Image Processing XXXI, Proc. SPIE Vol. 7073, 70732H, San Diego, CA., (2008a).

Coronel-Beltrán, Angel y Álvarez-Borrego, Josué, "Reconocimiento invariante de patrones con filtros no lineales", XXI Reunión Anual AMO, Memorias del LI Congreso Nacional de Física, Zacatecas, Zac., México (2008b).

Coronel-Beltrán, A. 1988. Correlacionador óptico invariante aplicado a la identificación de microorganismos fitoplanctónicos. Tesis de Maestría, CICESE.

Cooley, J. W., and Tukey, J. W., "An algorithm for machine calculation of complex Fourier series," *Math. Comp.*, Vol. 19, pp. 297-301 (1965).

Courant, R., and Hilbert, D. 1953. *Methods of Mathematical Physics*, Vol.1, John Wiley & Sons, Interscience Publishers, Inc. First English Edition, pp.103-105.

Cristóbal, G., and Cohen, L., "Image filtering and denoising through the scale transform", *International Symposium on Time-Frequency and Time-Scale Analysis, Proceedings of the IEEE-SP*, 617-620 (1998).

De Sena, A. and Rocchesso, D., "A Fast Mellin and Scale Transform," EURASIP Journal on Advances in Signal Processing, Hindawi Publishing Corporation, Volume 2007, Article ID 89170, 1-9 (2007).

De Sena, A. and Rocchesso, D., "A Fast Mellin Transform With Applications in DAFX," *Proc. Of the 7th Int. Conference on Digital Audio Effects (DAFx'04)*, Naples, Italy, 65-69 October 2004.

De Viedma, M. G., Baragaño, J. R., and Notario, A., "*Introducción a la entomología*", (Ed. Alhambra, ISBN 84-205-1031-9, 1985).

Fischer, B. 1973. Overlap of receptive field centers and representation of the visual field in the cat's optic tract. *Vision Research*, **13**, 2113-2120.

Fischl, B. Learning, anisotropic diffusion, nonlinear filtering and space-variant vision. Boston University Graduate School of Arts and Sciences, Ph.D. Dissertation, p. 190 (1997).

Fujii, H. and Almeida, S. P. 1979a. Partially matched spatial filtering with simulated input. *SPIE Vol. 177*, Optical Information Storage.

Fujii, H. and Almeida, S. P. 1979b. Coherent spatial filtering with simulated input. *Appl. Opt.* **18**, 1659-1662.

Fujii, H., Almeida, S. P., and Dowling, J. E., 1980. Rotational matched spatial filter for biological pattern recognition. *Appl. Opt.* **19**, 1190-1193.

Gabor, D. 1946. Theory of communication. *J. IEE (London)*, Vol. 93, Part III, No. 26, pp. 429-457.

García-Martínez, P., Otón, J., Vallés, J. J., and Arsenault, H. H. *Appl. Opt.* **43** (2004) 425.

Gauss, C. F., “Nachlass: Theoria interpolationis metodo nova tractata,” Werke, Band 3, Göttingen: Koniglichen Gesellschaft der Wissenschaften, pp. 265-327 (1866).

Gerardi, F. R. 1959. IRE Trans. Circ. Theory. 197-208.

Gonzalez, R. C. and Woods, R. E., *Digital Image Processing*, pp. 256-258, PEARSON Prentice Hall, (2008a).

Gonzalez, R. C. and Woods, R. E., *Digital Image Processing*, p. 402, PEARSON Prentice Hall, (2008b).

Gonzalez, R. C. and Woods, R. E., *Digital Image Processing*, p. 425, PEARSON Prentice Hall, (2008c).

Gonzalez, R. C. and Woods, R. E., *Digital Image Processing*, p. 861, PEARSON Prentice Hall, (2008d).

Goodman, J. W. *Introduction to Fourier Optics*. Roberts & Company Publishers, 3rd ed., Greenwood Village, CO. (2005).

Greco, M., Stucchi, N., Zavagno, D. and Marino, B. On the portability of computer-generated presentations: the effect of text-background color combinations on text legibility, *Hum Factors*, **50**(5), 821-833 (2008).

Gualdrón, O., Nicolás, J., Campos, J., and Yzuel, M. J., *Appl. Opt.* **42** (2003) 1434.

Guerrero-Moreno, R. E. and Álvarez-Borrego, J. Nonlinear composite filter performance, *Opt. Eng.* **48**(6), 067201 (June 2009).

Hill, A. L. Readability of websites with various foreground/background color combinations, font types and word styles, 1997. Stephen F. Austin State University. <http://www.laurenscharff.com/research/AHNCUR.html>

Horner, J. L. and Gianino, P. D., "Phase-only matched filtering," *Appl. Opt.*, **23**(6), 812-816 (1984).

Javidi, B. Nonlinear joint power spectrum based optical correlation, *Appl. Opt.* **28**(12), 2358-2367 (1989a).

Javidi, B. Synthetic discriminant function-based binary nonlinear optical correlator, *Appl. Opt.* **28**(13), 2490-2493 (1989b).

Javidi, B. Nonlinear matched filter based optical correlation, *Appl. Opt.* **28**(21), 4518-4520 (1989c).

Javidi, B. Generalization of the linear matched filter concept to nonlinear matched filters, *Appl. Opt.* **29**(8), 1215-1224 (1990a).

Javidi, B. Comparison of nonlinear joint transform correlator and nonlinear matched filter based correlator, *Opt. Commun.* **75**, 8-13 (1990b).

Javidi, B. and Horner, J. L., *Real-Time Optical Information Processing*, Academic Press, San Diego (1994).

Javidi, B., Wang, W. and Zhang, G. "Composite Fourier-plane nonlinear filter for distortion-invariant pattern recognition," *Opt. Eng.*, **36**(10), 2690-2696 (1997).

Kober, V., Lashin, V., Moreno, I., Campos, J., Yaroslavsky, L. P., and Yzuel, M. J., *J. Opt. Soc. Am.* **14** (1997) 2656.

Köhler, W. 1947. *Gestalt Psychology*. Liverright Press.

Kyrki, V., Local and global feature extraction for invariant object recognition. Thesis for the degree of Doctor of Science (Technology), Lappeenranta University of Technology, Finland, 2002.

Marion, A. 1991. *An introduction to image processing*. Chapman and Hall, Chap. 5.

Martinez-Diaz, S. "Métodos no lineales invariantes a distorsiones geométricas e iluminación para el reconocimiento de patrones o imágenes." p. 116, Tesis Doctoral, CICESE, 2008.

Mellin, H. 1902. Über den Zusammenhang zwischen den linearen Differential und Differenzengleichungen, *Acta Math.*, Vol. 25, pp. 139-164.

Millán, M. S., Yzuel, M. J., Campos, J., and Ferrerira, C., *Appl. Opt.* **31** (1992) 2560.

Millán, M. S., Corbalán, M., Romero, J., and Yzuel, M. J., *Opt. Lett.* **20** (1995) 1722.

Millán, M. S., Campos, J., Ferreira, C., and Yzuel, M. J., *Opt. Commun.* **73** (1989) 277.

Moreno, I., Ahouzi, E., Campos, J., and Ysuel, M. 1997. Real-time binary-amplitud phase-filter. *Appl. Opt.*, 36(29), 7428-7432.

Mouriño-Pérez, R. R. 1999. Correlación óptica a color para la identificación de diferentes estadios de *Vibrio cholerae* 01. Tesis de Doctorado, CICESE, p. 37.

Mouriño-Pérez, R. R., *Epidemiology*, **9** (1998) 355.

Parseval, M.-A. 1806. Mémoire sur les séries et sur l'intégration complète d'une équation aux différences partielle linéaires du second ordre, à coefficients constans. *Académie des Sciences* (original submission was 5 April 1799).

Pastoor, S. Legibility and subjective preference for color combinations in text, *Hum Factors*, 32(2), 157-171 (1990).

Pech-Pacheco, J. L., and Alvarez-Borrego, J. 1998. Optical-digital system applied to the identification of five phytoplankton species. *Marine Biology*, 132, 357-365.

Pech-Pacheco, J. L., Alvarez-Borrego, J., Cristóbal, G., and Keil, M. 2003. Automatic object identification irrespective to geometric changes, *Optical Engineering*, Vol. 42, No. 2, pp 551-559.

Pech-Pacheco, J., Cristobal, G., Alvarez-Borrego, J., and Cohen, L. Automatic system for phytoplanktonic algae identification. 2001. *Limnetica* 20(1), 143-158.

Pech-Pacheco, J. L., Cristobal, G., Álvarez-Borrego, J., and Cohen, L. "Power cepstral image analysis through the scale transform", in *Proc. SPIE*, vol. 4113, 2000.

Sawchuk, J. A. 1974. Space-variant image restoration by coordinate transformations. *J. Opt. Soc. Amer.* 138-144.

Schwartz, E. 1994. *Topographic mapping in primate visual cortex: history, anatomy and computation. Visual Science and Engineering. Models and Applications.* Ed. Marcel Dekker, New York.

Schwartz, E. L., Christman, D. R. and Wolf, A. P. Human Primary Visual Cortex Topography Imaged via Positron Tomography, *Brain Research*, 294, 225-230 (1984).

Schwartz, E. L., "Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding," *Vision Research*, **20**, 645-669 (1980).

Schwartz, E. L. The Development of Specific Visual Connections in the Monkey and the Goldfish: Outline of a Geometric Theory of Receptotopic Structure, *J. theor. Biol.* **69**, 655-683 (1977b).

Schwartz, E. L. Afferent geometry in the primate visual cortex and the generation of neuronal trigger features. *Biol. Cybernetics*, **28**, 1-14 (1977a).

Shieh, K. K. and Lai, Y. K. Effects of ambient illumination, luminance contrast, and stimulus type on subjective preference of VDT target and background color combinations, *Percept Mot Skills*, **107**(2), 336-352 (2008).

Solorza, S., and Álvarez-Borrego, J. 2010. Digital system of invariant correlation to position and rotation, *Opt. Commun.* **19**, 3613-3630.

Vander Lugt, A. 1964. Signal detection by complex spatial filtering. *IEEE Trans. Info. Theory*, **10**, 139-145.

Vander Lugt, A., 1968, A review of optical data-processing techniques, *Optica Acta*, **15**(1), 1-33.

Vijaya Kumar, B.V. K. and Hassebrook, L. Performance measures for correlation filters, *Appl. Opt.* **29**(20), 2997- 3006 (1990).

Yates, R. Mean of a white-noise process, *Digital Signal Labs, signal processing systems*, August 15, 2009.

Zavala-Hamz, V. A., and Alvarez-Borrego, J. 1997. Circular harmonic filters for the recognition of marine microorganisms. *Appl. Opt.*, **36**(2), 484-489.

APÉNDICE A

Demostración de los teoremas de Fourier

Teorema de la suma o de la linealidad. $\mathfrak{F}[ag(x, y) + bh(x, y)] = a\mathfrak{F}[g(x, y)] + b\mathfrak{F}[h(x, y)]$, donde a y $b \in \mathbb{R}$.

Demostración: Este teorema se deriva directamente de la propiedad de linealidad de las integrales que definen a la transformada de Fourier,

$$\begin{aligned}\mathfrak{F}[ag(x, y) + bh(x, y)] &= \iint_{-\infty}^{\infty} [ag(x, y) + bh(x, y)] e^{-j2\pi(ux+vy)} dx dy \\ &= a \iint_{-\infty}^{\infty} g(x, y) e^{-j2\pi(ux+vy)} dx dy + b \iint_{-\infty}^{\infty} bh(x, y) e^{-j2\pi(ux+vy)} dx dy \\ &= a\mathfrak{F}[g(x, y)] + b\mathfrak{F}[h(x, y)]. \blacksquare\end{aligned}$$

Teorema de similitud o de escalamiento. Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces $\mathfrak{F}[g(ax, by)] = \frac{1}{|ab|} G\left(\frac{u}{a}, \frac{v}{b}\right)$, donde a y $b \in \mathbb{R}$.

Demostración: Tomando la transformada de Fourier de $g(ax, by)$, tenemos

$$\mathfrak{F}[g(ax, by)] = \iint_{-\infty}^{\infty} g(ax, by) e^{-j2\pi(ux+vy)} dx dy$$

$$= \iint_{-\infty}^{\infty} g(ax, by) e^{-j2\pi(\frac{u}{a}ax + \frac{v}{b}by)} \frac{dax}{|a|} \frac{dby}{|b|}$$

$$= \frac{1}{|ab|} G\left(\frac{u}{a}, \frac{v}{b}\right). \blacksquare$$

Teorema de desplazamiento. Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces $\mathfrak{F}[g(x - a, y - b)] = G(u, v) e^{-j2\pi(ua + vb)}$, donde a y $b \in \mathbb{R}$.

Demostración: Tomando la transformada de Fourier de $g(x - a, y - b)$, tenemos

$$\begin{aligned} \mathfrak{F}[g(x - a, y - b)] &= \iint_{-\infty}^{\infty} g(x - a, y - b) e^{-j2\pi(ux + vy)} dx dy \\ &= \iint_{-\infty}^{\infty} g(x', y') e^{-j2\pi[u(x' + a) + v(y' + b)]} dx' dy' \\ &= G(u, v) e^{-j2\pi(ua + vb)}. \blacksquare \end{aligned}$$

Teorema de Parseval (1806). Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces

$$\iint_{-\infty}^{\infty} |g(x, y)|^2 dx dy = \iint_{-\infty}^{\infty} |G(u, v)|^2 du dv.$$

Demostración:

$$\begin{aligned}
 \iint_{-\infty}^{\infty} |g(x, y)|^2 dx dy &= \iint_{-\infty}^{\infty} g(x, y) g^*(x, y) dx dy \\
 &= \iint_{-\infty}^{\infty} dx dy \left[\iint_{-\infty}^{\infty} d\xi d\eta G(\xi, \eta) e^{j2\pi(x\xi + y\eta)} \right] \left[\iint_{-\infty}^{\infty} d\alpha d\beta G^*(\alpha, \beta) e^{-j2\pi(x\alpha + y\beta)} \right] \\
 &= \iint_{-\infty}^{\infty} d\xi d\eta G(\xi, \eta) \iint_{-\infty}^{\infty} d\alpha d\beta G^*(\alpha, \beta) \left[\iint_{-\infty}^{\infty} e^{j2\pi[(x(\xi - \alpha) + y(\eta - \beta))]} dx dy \right] \\
 &= \iint_{-\infty}^{\infty} d\xi d\eta G(\xi, \eta) \iint_{-\infty}^{\infty} d\alpha d\beta G^*(\alpha, \beta) \delta(\xi - \alpha, \eta - \beta) \\
 &= \iint_{-\infty}^{\infty} |G(\xi, \eta)|^2 d\xi d\eta. \blacksquare
 \end{aligned}$$

Teorema de convolución. Si $\mathfrak{F}[g(x, y)] = G(u, v)$ y $\mathfrak{F}[h(x, y)] = H(u, v)$, entonces

$$\mathfrak{F} \left[\iint_{-\infty}^{\infty} g(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta \right] = G(u, v) H(u, v).$$

Demostración:

$$\begin{aligned}
 \mathfrak{F} \left[\iint_{-\infty}^{\infty} g(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta \right] &= \iint_{-\infty}^{\infty} g(\xi, \eta) \mathfrak{F}[h(x - \xi, y - \eta)] d\xi d\eta \\
 &= \iint_{-\infty}^{\infty} g(\xi, \eta) e^{-j2\pi(u\xi + v\eta)} d\xi d\eta H(u, v) \\
 &= G(u, v) H(u, v). \blacksquare
 \end{aligned}$$

Teorema de auto-correlación. Si $\mathfrak{F}[g(x, y)] = G(u, v)$, entonces

$$\mathfrak{F} \left[\iint_{-\infty}^{\infty} g(\xi, \eta) g^*(\xi - x, \eta - y) d\xi d\eta \right] = |G(u, v)|^2.$$

Similarmente,

$$\mathfrak{F}[|g(\xi, \eta)|^2] = \iint_{-\infty}^{\infty} G(\xi, \eta) G^*(\xi + u, \eta + v) d\xi d\eta.$$

Demostración:

$$\mathfrak{F} \left[\iint_{-\infty}^{\infty} g(\xi, \eta) g^*(\xi - x, \eta - y) d\xi d\eta \right] = \mathfrak{F} \left[\iint_{-\infty}^{\infty} g(\xi' + x, \eta' + y) g^*(\xi', \eta') d\xi' d\eta' \right]$$

$$= \iint_{-\infty}^{\infty} d\xi' d\eta' g^*(\xi', \eta') \mathfrak{F}[g(\xi' + x, \eta' + y)]$$

$$= \iint_{-\infty}^{\infty} d\xi' d\eta' g^*(\xi', \eta') e^{j2\pi(u\xi' + v\eta')} G(u, v)$$

$$= G^*(u, v)G(u, v) = |G(u, v)|^2. \blacksquare$$

APÉNDICE B

Fórmulas de inversión de Mellin

TEOREMA 1: Sea $s = \sigma + ti$ una variable compleja. Sea la función $f(s)$ regular en la zona

$\alpha < \sigma < \beta$ y sea $\int_{-\infty}^{\infty} |f(\sigma + ti)| dt$ convergente en esa zona. Además, sea la función $f(s)$ tendiente uniformemente a cero con incrementos $|t|$ de la forma $\alpha + \delta \leq \sigma \leq \beta - \delta$ ($\delta > 0$, fijada arbitrariamente). Si para x real positiva y σ fijo definimos

$$g(x) = \frac{1}{2\pi i} \int_{\sigma - \infty i}^{\sigma + \infty i} f(s) x^{-s} ds \quad \text{B. 1}$$

entonces

$$f(s) = \int_0^{\infty} g(x) x^{s-1} dx \quad \text{B. 2}$$

en el intervalo $\alpha < \sigma < \beta$.

Demostración: Debido a la consideración de que $f(s)$ converge uniformemente a cero para $\alpha + \delta \leq \sigma \leq \beta - \delta$ cuando $|t| \rightarrow \infty$, la trayectoria de integración en ecuación B.1 puede ser desplazada paralela a si misma mientras permanece en el intervalo $\alpha < \sigma < \beta$; así que $g(x)$ no depende de σ . Si escogemos dos abscisas σ_1 y σ_2 con $\alpha < \sigma_1 < \sigma < \sigma_2 < \beta$, tenemos

$$\int_0^{\infty} x^{s-1} g(x) dx = \int_0^1 x^{s-1} dx \frac{1}{2\pi i} \int_{\sigma_1 - \infty i}^{\sigma_1 + \infty i} x^{-s_1} f(s_1) ds_1 + \int_1^{\infty} x^{s-1} dx \frac{1}{2\pi i} \int_{\sigma_2 - \infty i}^{\sigma_2 + \infty i} x^{-s_2} f(s_2) ds_2 = J_1 + J_2.$$

El orden de las integraciones en estas integrales puede ser intercambiado porque tenemos el estimado

$$|J_1| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(\sigma_1 + ti)| dt \int_0^1 x^{-1+(\sigma-\sigma_1)} dx < \infty,$$

$$|J_2| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(\sigma_2 + ti)| dt \int_1^{\infty} x^{-1+(\sigma-\sigma_2)} dx < \infty$$

para las integrales intercambiadas.

Así obtenemos

$$\int_0^{\infty} x^{s-1} g(x) dx = \frac{1}{2\pi i} \int_{\sigma_2 - \infty i}^{\sigma_2 + \infty i} \frac{f(s_2)}{s_2 - s} ds_2 - \frac{1}{2\pi i} \int_{\sigma_1 - \infty i}^{\sigma_1 + \infty i} \frac{f(s_1)}{s_1 - s} ds_1.$$

De acuerdo a la fórmula integral de Cauchy, la diferencia en el lado derecho es igual a $f(s)$, ya que las integrales sobre los segmentos horizontales conectando las dos líneas verticales $s = \sigma_1$ y $s = \sigma_2$ tienden a cero para $|t| \rightarrow \infty$ (porque $f(s) \rightarrow 0$).

Por lo tanto,

$$\int_0^{\infty} x^{s-1} g(x) dx = f(s). \blacksquare$$

TEOREMA 2: Sea $g(x)$ una función suave en trozos para $x > 0$, y sea $\int_0^{\infty} x^{\sigma-1} g(x) dx$

absolutamente convergente para $\alpha < \sigma < \beta$. Entonces la fórmula de inversión B.1 sigue de B.2.

Demostración: Hagamos $x = e^u$. Tenemos entonces

$$\frac{1}{2\pi i} \int_{\sigma - \infty i}^{\sigma + \infty i} x^{-s} f(s) ds = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-u(\sigma + ti)} dt \int_{-\infty}^{\infty} e^{v(\sigma + ti)} g(e^v) dv$$

$$= \frac{e^{-u\sigma}}{2\pi} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} e^{it(v-u)} e^{vg} g(e^v) dv \quad \text{B.3}$$

Por el teorema integral de Fourier, el cual se expresa como

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} f(t) e^{-iu(t-x)} dt$$

la expresión B.3 es igual a $e^{-u\sigma} e^{u\sigma} g(e^v) = g(x)$; por lo tanto el teorema queda demostrado.

$$\frac{1}{2\pi i} \int_{\sigma-\infty i}^{\sigma+\infty i} f(s) x^{-s} f(s) ds = g(x). \blacksquare$$

APÉNDICE C

Deducción de la transformada de escala

La *transformada de escala* (Cohen, 1993) es

$$D_f(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) e^{(-jc - 1/2) \ln t} dt$$

Para la deducción de esta transformada se utilizan dos operadores lineales llamados el operador frecuencia W y el operador tiempo T , los cuales están definidos como

$$W = \frac{1}{j} \frac{d}{dt} \text{ y } T = t.$$

Operadores en las transformadas:

- El método “operador” lo introdujo Gabor (1946) en el procesamiento de señales de la mecánica cuántica.
- Los operadores tiempo y frecuencia se definen como: $T = t$ y $W = -j(d/dt)$ en el dominio temporal.
- El operador asociado con la variable p de Mellin puede ser reescrito en términos de T y W .
- Cohen (1993) introdujo un operador de Mellin que involucra el producto del tiempo y la frecuencia, el cual se refirió a este como el operador “escala”, c ,

$$c = 1/2(TW + WT) = TW - 1/2j, \text{ por lo que } TW - WT = j.$$

Además, se requiere que estos operadores sean Hermitianos porque esto nos garantiza que las funciones propias (llamadas también *eigenfunctions*, del alemán *eigen=propio*) sean completas y ortogonales, y de que los valores propios (llamados también *eigenvalues*) sean reales.

Se dice que un operador lineal A es Hermitiano o autoadjunto si para cualquier par de funciones, $f(t)$ y $g(t)$, se cumple que

$$\int g^*(t) A f(t) dt = \int f(t) \{A g(t)\}^* dt.$$

Haciendo $A = W$, tenemos que probar que

$$\int g^*(t) W f(t) dt = \int f(t) \{W g(t)\}^* dt.$$

Sustituyendo $W = \frac{1}{j} \frac{d}{dt}$ en esta expresión e integrando por partes tenemos

$$\begin{aligned} \int g^*(t) W f(t) dt &= \int g^*(t) \frac{1}{j} \frac{d}{dt} f(t) dt = \left[\frac{1}{j} f g^* \right]_{-\infty}^{\infty} - \frac{1}{j} \int f(t) \frac{d}{dt} g^*(t) dt \\ &= \int f(t) \left(\frac{1}{j} \frac{d g(t)}{dt} \right)^* dt = \int f(t) \{W g(t)\}^* dt \end{aligned}$$

Por lo tanto, el operador frecuencia W es Hermitiano.

El problema de valores propios es $W u(\omega, t) = \omega u(\omega, t)$, y las soluciones son $u(\omega, t) = c e^{j\omega t}$, donde $\omega \in \mathbb{R}$ y en el intervalo $[-\infty, \infty]$.

Para obtener la normalización consideremos

$$\begin{aligned} \int u^*(\omega, t) u(\omega', t) dt &= c^2 \int e^{-j\omega t} e^{j\omega' t} dt \\ &= c^2 \int e^{-jt(\omega - \omega')} dt \\ &= c^2 2\pi \delta(\omega - \omega'). \end{aligned}$$

Ya que queremos normalizar este resultado a una función delta, debemos entonces tomar $c^2 = 1/2\pi$.

Así que

$$\int u^*(\omega, t) u(\omega', t) dt = \delta(\omega - \omega').$$

Y de esta manera tendremos que las funciones propias de frecuencia normalizadas son

$$u(\omega, t) = \frac{1}{\sqrt{2\pi}} e^{j\omega t}.$$

El operador tiempo en la representación temporal es t , y el problema de valores propios es

$$tu(t, t') = t' u(t, t'),$$

donde t' son los valores propios.

Es interesante hacer notar que esta última ecuación fue la que llevó a Paul A. Dirac a inventar su famosa función delta de Dirac.

Las soluciones a la ecuación $tu(t, t') = t' u(t, t')$ son $u(t, t') = \delta(t - t')$, donde t' puede ser cualquier número. Por lo tanto, los valores propios son continuos y las funciones propias son completas y ortogonales. Del álgebra de operadores tenemos que $TW - WT = j$. Esto significa que al operar con el lado izquierdo de la ecuación sobre una función arbitraria, es lo mismo que multiplicar por j . Para probar lo anterior consideremos una función $f(t)$,

$$\begin{aligned} (TW - WT)f(t) &= \left(t \frac{1}{j} \frac{d}{dt} - \frac{1}{j} \frac{dt}{dt} \right) f(t) \\ &= \frac{1}{j} \left(t \frac{df}{dt} - \frac{d}{dt} t f \right) \\ &= \frac{1}{j} \left(t \frac{df}{dt} - t \frac{df}{dt} - f \right) \\ &= -\frac{1}{j} f(t) \\ &= j f(t), \end{aligned}$$

lo cual prueba que $TW - WT = j$. Esta expresión podemos escribirla como

$$TW - WT = j$$

$$TW = WT + j$$

$$TW - \frac{1}{2}j = WT + j - \frac{1}{2}j$$

$$TW - \frac{1}{2}j = WT + \frac{1}{2}j$$

A esto último lo llamamos el *operador escala* C ,

$$C = TW - \frac{1}{2}j = WT + \frac{1}{2}j.$$

Para obtener la transformada a escala vamos a resolver el problema de valores propios para el operador escala. Usaremos c y $\gamma(c, t)$ para indicar los valores propios y las funciones propias, respectivamente.

El problema de valor propio es

$$C\gamma(c, t) = c\gamma(c, t).$$

Como el operador escala $C = TW - \frac{1}{2}j$, tenemos

$$\left(TW - \frac{1}{2}j\right)\gamma(c, t) = c\gamma(c, t)$$

$$\left(t\frac{1}{j}\frac{d}{dt} - \frac{1}{2}j\right)\gamma(c, t) = c\gamma(c, t)$$

$$t\frac{1}{j}\frac{d\gamma(c, t)}{dt} - \frac{1}{2}j\gamma(c, t) = c\gamma(c, t)$$

$$-jt\frac{d\gamma(c, t)}{dt} - \frac{1}{2}j\gamma(c, t) = c\gamma(c, t).$$

Las soluciones a esta ecuación diferencial, normalizadas a la función delta, son

$$\gamma(c, t) = \frac{1}{\sqrt{2\pi}} \frac{e^{jc\ln t}}{\sqrt{t}}; \quad t \geq 0.$$

Cualquier función puede ser expandida en términos de las funciones propias de escala, de la forma

$$f(t) = \int D(c) \gamma(c, t) dc.$$

Sustituyendo $\gamma(c, t)$ en esta última expresión, tenemos

$$f(t) = \frac{1}{\sqrt{2\pi}} \int D(c) \frac{e^{jclnt}}{\sqrt{t}} dc; \quad t \geq 0$$

y la transformación inversa es

$$D(c) = \int_0^{\infty} f(t) \gamma^*(c, t) dt = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) \frac{e^{-jclnt}}{\sqrt{t}} dt,$$

que puede escribirse como

$$D_f(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) t^{-1/2} e^{-jclnt} dt = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) e^{-1/2 lnt} e^{-jclnt} dt.$$

Por lo tanto,

$$D_f(c) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(t) e^{(-jc-1/2)lnt} dt$$

es la transformada de escala.