

**ESCUELA POLITÉCNICA NACIONAL**

**FACULTAD DE INGENIERÍA ELÉCTRICA**

**TESIS DE GRADO**

**“MEJORAMIENTO DE IMÁGENES MONOCROMÁTICAS  
USANDO LA TRANSFORMADA DISCRETA COSENO”**

**TESIS PREVIA A LA OBTENCIÓN DEL TÍTULO DE  
INGENIERO EN LA ESPECIALIZACIÓN DE ELECTRÓNICA  
Y TELECOMUNICACIONES**

**WASHINGTON LENIN PAREDES CRUZ**

**QUITO, SEPTIEMBRE DE 1997**

Certifico que el presente trabajo ha sido  
realizado en su totalidad por el señor:

Washington Lenin Paredes Cruz

A handwritten signature in black ink, appearing to read 'M. Soledad Jiménez', written over a horizontal line.

Ing. María Soledad Jiménez  
**DIRECTOR DE TESIS**

### **DEDICATORIA**

Este trabajo está dedicado a todas aquellas personas interesadas en el procesamiento digital de imágenes, en especial, para las generaciones venideras de estudiantes del Departamento de Electrónica y Telecomunicaciones de la Facultad de Ingeniería Eléctrica de la Escuela Politécnica Nacional.

### **AGRADECIMIENTO**

Mi eterno agradecimiento a mis padres, sin cuyo apoyo moral y económico no hubiera sido posible culminar mi carrera. Así mismo, mil gracias a la Ing. María Soledad Jiménez por su invaluable guía en la consecución de este trabajo.

# ÍNDICE

INTRODUCCIÓN .....	i
--------------------	---

## CAPÍTULO 1

### TRATAMIENTO Y REPRESENTACIÓN DE IMÁGENES

1.0	INTRODUCCIÓN .....	1
1.1	PROCESAMIENTO DIGITAL DE IMÁGENES .....	2
1.2	SISTEMAS BIDIMENSIONALES .....	4
1.3	REPRESENTACIÓN DE IMÁGENES .....	6
	1.3.1 Representación mediante funciones ortogonales .....	6
	1.3.2 Representación de imágenes en el computador .....	8
1.4	TEORÍA DE MATRICES .....	10
	1.4.1 Definiciones de matrices .....	10
1.5	TEORÍA DE PROBABILIDADES .....	16
	1.5.1 Conceptos probabilísticos .....	17
	1.5.2 Procesos aleatorios .....	21
	1.5.3 Función tasa de distorsión .....	27

## CAPÍTULO 2

### TRANSFORMADAS DISCRETAS BIDIMENSIONALES

2.0	INTRODUCCIÓN .....	29
2.1	TRANSFORMADAS UNITARIAS Y ORTOGONALES .....	29
	2.1.1 Imágenes bases .....	32
	2.1.2 Propiedades de las transformadas unitarias .....	34
2.2	DESCRIPCIÓN DE LAS TRANSFORMADAS DISCRETAS .....	36
	2.2.1 La Transformada discreta de Fourier .....	36
	2.2.2 La transformada discreta coseno .....	38
	2.2.3 La transformada discreta seno .....	40
	2.2.4 La transformada discreta Hadamard .....	42
	2.2.5 La transformada discreta Haar .....	44
	2.2.6 La transformada discreta Slant .....	45
	2.2.7 La transformada Karhunen-Loeve .....	46

## CAPÍTULO 3

### LA TRANSFORMADA DISCRETA COSENO

3.0	INTRODUCCIÓN .....	54
3.1	LA TRANSFORMADA DISCRETA COSENO .....	55
3.1.1	Medidas de comportamiento de las transformadas unitarias .	56
3.1.2	Definición de la transformada discreta coseno .....	59
3.1.3	Propiedades de la transformada discreta coseno .....	63
3.2	CÁLCULO DE LA TRANSFORMADA DISCRETA COSENO, ALGORITMO .....	75
3.2.1	Partición y reordenamiento de matrices .....	77
3.2.2	Transformada discreta coseno directa .....	80
3.2.3	Transformada discreta coseno inversa .....	85
3.3	TRANSFORMADA DISCRETA COSENO DE IMÁGENES MONOCROMÁTICAS .....	89

## CAPÍTULO 4

### FILTRADO DE IMÁGENES EN EL DOMINIO DE LA TRANSFORMADA DISCRETA COSENO PARA SU MEJORAMIENTO

4.0	INTRODUCCIÓN .....	96
4.1	MEJORAMIENTO DE IMÁGENES .....	97
4.2	FILTROS .....	100
4.2.1	Operaciones en el dominio de las transformadas discretas .	102
4.2.2	Filtro lineal generalizado .....	105
4.2.3	Filtro raíz .....	109
4.2.4	Filtro homomórfico .....	112

## CAPÍTULO 5

### DESARROLLO DEL SISTEMA DE SOFTWARE

5.0	INTRODUCCIÓN .....	120
5.1	REQUERIMIENTOS COMPUTACIONALES .....	120
5.2	IMPLEMENTACIÓN DEL SOFTWARE .....	122
5.2.1	El lenguaje C .....	122
5.2.2	Programación Windows .....	125
5.2.3	Software desarrollado .....	131

## CAPÍTULO 6

### ANÁLISIS COMPARATIVO DE RESULTADOS OBTENIDOS

6.0	INTRODUCCIÓN .....	141
6.1	FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO GAUSSIANO BLANCO .....	142
6.2	FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO SAL Y PIMIENTA .....	149
6.3	FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO IMPULSIVO POSITIVO .....	154
6.4	FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO IMPULSIVO NEGATIVO .....	160
6.5	FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO DEPENDIENTE DE LA SEÑAL .....	165
6.6	FILTRADO HOMOMÓRFICO DE IMÁGENES NO CONTAMINADAS .....	170

## CAPÍTULO 7

### COMENTARIOS, CONCLUSIONES Y RECOMENDACIONES

7.1	COMENTARIOS .....	177
7.2	CONCLUSIONES .....	182
7.3	RECOMENDACIONES .....	186
	BIBLIOGRAFÍA .....	189
	APÉNDICE A: CONCEPTOS BÁSICOS DE SISTEMAS DISCRETOS ..	193
	APÉNDICE B: DEMOSTRACIONES MATEMÁTICAS .....	197
	APÉNDICE C: DEFINICIONES DE LA TDC .....	202
	APÉNDICE D: RUIDO ADITIVO .....	204
	APÉNDICE E: MANUAL DE USUARIO .....	208

# INTRODUCCIÓN

No existe en la naturaleza un sistema de imágenes que entregue una imagen de calidad perfecta. El objetivo del mejoramiento de imágenes es manipular dicha imagen de manera de perfeccionar su calidad. Por ejemplo en reconocimiento aéreo de blancos, y en fotografías de planetas a través de la atmósfera, las imágenes son degradadas por turbulencia atmosférica, imperfecciones del sistema óptico, etc. En el área médica las radiografías son usualmente de baja resolución y bajo contraste. Como estas, existen un sin número de imperfecciones que deterioran la calidad de la imagen; así como la naturaleza de las degradaciones es variada, así también se han desarrollado diferentes técnicas para su corrección. Pero una técnica que trabaje adecuadamente para un tipo de degradación puede muy bien no servir para otro tipo.

Entre las técnicas destinadas al mejoramiento de imágenes están aquellas desarrolladas para el proceso en el dominio de la frecuencia, para ello hacen uso de las llamadas transformadas imagen y entre ellas encontramos a la transformada discreta coseno, de Fourier. Hadamard, entre otras. Este trabajo de tesis se encarga de estudiar el filtraje de imágenes con propósitos de mejoramiento en el dominio de la transformada discreta coseno, en este caso las degradaciones se simularán contaminando una imagen de prueba con diferentes clases de ruido para luego filtrarlas probando la efectividad de los filtros sobre cada uno de los ruidos.

Para una ordenada consecución de la investigación se la ha dividido en 7 capítulos. El primero contiene conceptos introductorios sobre procesamiento digital, incluye además un resumen de herramientas matemáticas de mucha utilidad como teoría de matrices, y conceptos probabilísticos.



El segundo capítulo hace una breve descripción sobre las transformadas discretas bidimensionales comúnmente utilizadas en procesamiento digital de imágenes, sus definiciones y sus propiedades más importantes.

El tercer capítulo profundiza en el estudio de la transformada discreta coseno, demostración de sus propiedades y la descripción detallada de su algoritmo de cálculo para el par de transformadas; esto es, tanto para la transformada discreta coseno directa como para la transformada inversa.

El capítulo cuarto se encarga en cambio de desarrollar la matemática que describe el comportamiento de los filtros en dominio de la frecuencia. El capítulo cinco es parte complementaria al tercero y cuarto, en este se describe el software implementado, cuyas principales funciones son: calcular la transformada discreta coseno directa, realizar el proceso de filtrado en dominio de la transformada, y recuperar la imagen filtrada aplicando la transformada coseno inversa; todo en base a las deducciones matemáticas realizadas en estos dos capítulos previos.

En el capítulo sexto se analizan los resultados de las pruebas realizadas aplicando cada uno de los filtros a las imágenes contaminadas con los diferentes ruidos.

Finalmente el séptimo capítulo presenta los comentarios, las conclusiones y recomendaciones derivadas del estudio realizado en este trabajo de tesis.

Demostraciones matemáticas, conceptos y propiedades que no son estrictamente necesarios pero que vale la pena conocerlos se han desarrollado como anexos al final y constituyen temas opcionales para el lector.

# CAPÍTULO

---

# 1

## TRATAMIENTO Y REPRESENTACIÓN DE IMÁGENES

### 1.0 INTRODUCCIÓN

Este capítulo está destinado a la descripción de las imágenes, para lo cual se inicia definiendo el procesamiento digital de imágenes, qué es y qué procesos lo integran.

Como todo fenómeno necesita un sustento matemático, se incluye la parte matemática estrictamente necesaria, como son: la teoría de matrices, donde se presentan las definiciones elementales de vectores y matrices, acompañadas de su clasificación y propiedades; y, finalmente lo referente a la teoría de probabilidades, en donde se define algunas funciones básicas útiles para la exposición de temas como valores promedios, correlación, autocorrelación, etc.

## 1.1 PROCESAMIENTO DIGITAL DE IMÁGENES

El término *procesamiento digital de imágenes* se usa para referirse al procesamiento de un cuadro bidimensional generalmente realizado por un microprocesador. Una *imagen digital* es un arreglo de números reales o complejos cada uno representado por un determinado número de bits.

Una imagen puede estar dada en forma de transparencia, fotografía, slide, etc. Esta imagen, ingresada mediante un adecuado periférico de entrada, será primeramente digitalizada y almacenada en memoria de un computador para su posterior procesamiento y finalmente presentada al usuario mediante algunos de los periféricos de salida (monitor, plotter o impresora).

El procesamiento digital de imágenes tiene un amplio campo de aplicaciones como: detección remota vía satélite, procesamiento médico, procesamiento acústico, radar, sonar y entre otras inspección de partes industriales.

La adquisición de imágenes por satélite tiene mucha aplicación en rastreo de recursos naturales, predicción de cosechas agrícolas, desarrollo urbano, estado atmosférico, etc. Las aplicaciones médicas están especialmente enfocadas a la detección de tumores mediante el uso de rayos X, tomografía, resonancia nuclear magnética y rastreo ultrasónico. Los sistemas de radar y sonar tienen su aplicación para detección y reconocimiento de varios tipos de blancos. En suma, el procesamiento digital de imágenes, puede extenderse desde visión robotizada para automatización industrial hasta síntesis de imágenes para la realización de caricaturas o diseño en moda. Todas estas aplicaciones del procesamiento digital de imágenes requieren la solución de algunos de los siguientes problemas:

### **Modelación de imágenes**

En términos generales una imagen será considerada una función bidimensional que contenga información. Los modelos de una imagen dan una descripción cuantitativa de las propiedades de esta función.

Modelos estadísticos describen a una imagen como un conjunto caracterizado por las funciones promedio y de covarianza. Los modelos estocásticos caracterizan cada unidad de imagen (pixel ó pel) por una relación con sus pixeles vecinos. Uno u otro modelo son de mucha utilidad para el desarrollo de algoritmos para su tratamiento.

### **Mejoramiento de Imágenes**

Se entiende por mejoramiento de imágenes el acentuar ciertas características de la imagen, sin alterar la información contenida en sus datos, incluye mejoramiento de bordes y contraste, filtrado de ruido, conformación y magnificación, de mucha utilidad para análisis y presentación visual de la imagen. Este tópico será considerado con detenimiento en el capítulo 4.

### **Restauración de Imágenes**

La restauración de imágenes se refiere a la eliminación o minimización de degradaciones en una imagen, dichas degradaciones serán producidas por limitaciones o alinealidades de los sensores, se incluye también filtrado de ruido.

A pesar de estar estrechamente relacionadas hay una importante diferencia entre restauración y mejoramiento de imágenes, para mostrar la diferencia note que una imagen original (imagen no degradada) no necesita ser restaurada, pero puede ser mejorada.

### **Análisis de Imágenes**

Es una medida cuantitativa con el fin de lograr una descripción de la imagen, de mucha utilidad por ejemplo en lectura de etiquetas para identificar productos en un comisariato o medir el tamaño y orientación de las moléculas en la sangre para el caso de una imagen médica.

### **Reconstrucción de imágenes a partir de proyecciones**

Es un caso especial de la restauración de imágenes, en donde objetos bidimensionales son reconstruidos partiendo de varias proyecciones unidimensionales, dichas proyecciones requieren que sean tomadas desde diferentes ángulos usando radiación penetrante.

### **Compresión de datos en la imagen**

La cantidad de datos asociada con la información de tipo visual es muy grande (por ejemplo 1 hora de TV-color requiere una capacidad<sup>1</sup> de  $28 \cdot 10^9$  bytes), por ende hace uso de grandes capacidades para su almacenamiento con velocidades de acceso normalmente inversamente proporcional a su capacidad.

La compresión de datos se refiere a la reducción del número de bits requeridos para almacenar o transmitir una imagen sin pérdida apreciable de su información.

## **1.2 SISTEMAS BIDIMENSIONALES**

Un *sistema* se define como un conjunto de objetos combinados adecuadamente, capacitados para trabajar de una manera armónica con el propósito de alcanzar un

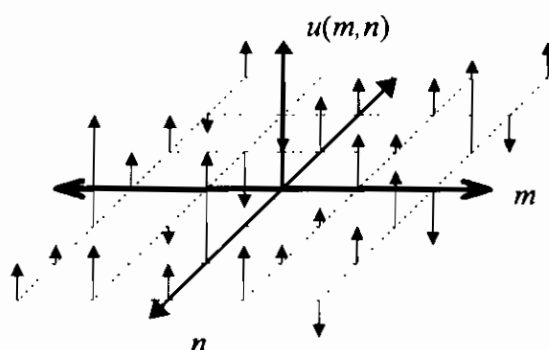
---

<sup>1</sup> La referencia [1], en su primer capítulo presenta otras fuentes de imágenes y su respectivo volumen de información.

objetivo, una *señal* en cambio será un suceso que está en capacidad de dar inicio a la acción del sistema.

Todas las propiedades de los sistemas bidimensionales son de fácil extensión partiendo del caso unidimensional.

Una secuencia bidimensional (2-D) es una función de dos variables enteras, de una manera gráfica se muestra un ejemplo en la figura 1.2.0.1 .



**Figura 1.2.0.1** Representación gráfica de una secuencia 2-D.

Implícitamente al hablar de sistemas discretos se entiende la necesidad de trabajar con secuencias. En el Apéndice A se presentan algunas funciones de uso frecuente, necesarias para describir el comportamiento de sistemas bidimensionales, así como también algunas de sus propiedades como: causalidad, linealidad, invarianza y separabilidad.

Una imagen digital es una función bidimensional de intensidad de luz  $u(m,n)$ , en la cual  $m$  y  $n$  representan las coordenadas espaciales y  $u$  indica el nivel de brillo de la imagen en el punto  $(m,n)$ .

## 1.3 REPRESENTACIÓN DE IMÁGENES

En primer lugar se ve conveniente presentar un tipo de señal de mucho uso como es la de energía y posteriormente una manera de representar señales.

Una señal de *energía* se define como aquella que dura solamente un intervalo finito de tiempo y si está definida en un intervalo de tiempo infinito su mayor parte de energía se concentra en un intervalo finito. Para el caso específico de sistemas eléctricos la señal  $x(t)$  será corriente ó voltaje. La potencia disipada por esta señal en una resistencia de valor unitario es [8],

$$p = |x(t)|^2 \quad (1.3.0.1)$$

Si se sabe que la potencia es la rapidez con que cambia la energía en el sistema,

$$p = \frac{dE}{dt} \Rightarrow E = \int_{t_1}^{t_2} |x(t)|^2 dt \quad (1.3.0.2)$$

La señal de energía se define cuando ésta es finita, aunque el intervalo considerado sea infinito, es decir,

$$\int_{-\infty}^{\infty} |x(t)|^2 dt < \infty \quad (1.3.0.3)$$

### 1.3.1 REPRESENTACIÓN MEDIANTE FUNCIONES ORTOGONALES

Una señal continua unidimensional puede representarse como una función de una sola variable como  $f(t)$ ,  $u(x)$ , etc, adicionalmente  $f(t)$  con la premisa de tener energía finita

puede ser expresada por medio de un conjunto contable de números,  $f_n$ , que no dependen explícitamente de la variable tiempo, esto es [8]:

$$f(t) = \sum_n f_n \phi_n(t) \quad (1.3.1.1)$$

donde  $\phi_n(t)$  forma un conjunto ortogonal de funciones y  $f_n$  son números no dependientes del tiempo.

Se espera que  $f(t)$  sea expresado unívocamente en función de  $\phi_n$  y  $f_n$ , para ello es necesario que cada término de la suma sea linealmente independiente. Esto es análogo a expresar un vector como una combinación lineal de sus componentes ortogonales, lo cual induce a escoger un conjunto ortogonal completo de funciones<sup>1</sup>, las  $\phi_n(t)$  que cumplirán la condición [8]:

$$\int_{t_1}^{t_2} \phi_n(t) \phi_m^*(t) dt = \begin{cases} 0 & \text{si } n \neq m \\ k_n & \text{si } n = m \end{cases} \quad (1.3.1.2)$$

Si este conjunto de funciones cumple además que  $k_n = 1$ , para todo  $n$ , se dice *ortonormal*.

La operación descrita mediante la ecuación (1.3.1.2) se conoce como el *producto interno*<sup>2</sup> de las funciones  $\phi_n(t)$  y  $\phi_m(t)$ , note que el producto interno de una función consigo misma es  $k_n$ . Al valor  $\sqrt{k_n}$  se conoce como la *norma*<sup>3</sup> de la función.

<sup>1</sup> Un conjunto se dice ORTOGONAL si todas las funciones que lo conforman son ortogonales entre si [8].

<sup>2</sup> Al producto interno dado por la ecuación (1.3.1.2) se denota con  $\langle \phi_n, \phi_m \rangle$ .

<sup>3</sup> Recuerdese que para el caso de un vector la norma se refiere a su longitud o módulo. A la norma de la función  $\phi_n(t)$  se denota con  $\|\phi_n\|$ .



Para el caso discreto se considera una secuencia de  $N$ -muestras espaciadas uniformemente sobre el intervalo  $[0, NT]$ , dada por;

$$u(nT) = u(0), u(T), u(2T), \dots, u([N - 1]T) \quad (1.3.1.3)$$

entonces de manera análoga a (1.3.1.1) se tiene,

$$U(k\Omega) = \sum_{n=0}^{N-1} u(nT)f(nTk\Omega) \quad , k = 0, \dots, N - 1 \quad (1.3.1.4)$$

Así como una señal se puede representar por una serie ortogonal (como la serie de Fourier), así mismo una imagen puede representarse mediante funciones ortogonales bidimensionales llamadas *imágenes bases*, estas imágenes bases se obtienen mediante matrices unitarias llamadas *transformadas imagen* para el caso de imágenes discretas.

### 1.3.2 REPRESENTACIÓN DE IMÁGENES EN EL COMPUTADOR

Una imagen representa los puntos luminosos en una escena, cuyo nivel de brillo cambia de un punto a otro. En las imágenes monocromáticas se refiere a los niveles de gris, mientras que en las escenas a color (policromáticas) se hace referencia a las tonalidades de color.

Una imagen para su tratamiento en el computador tiene que ser previamente digitalizada, pues este trabaja únicamente con valores discretos, tanto en el proceso de almacenamiento como en el de procesamiento. El digitalizador divide la imagen en pequeños elementos llamados *pixeles*<sup>1</sup>, asignando a cada uno una intensidad específica.

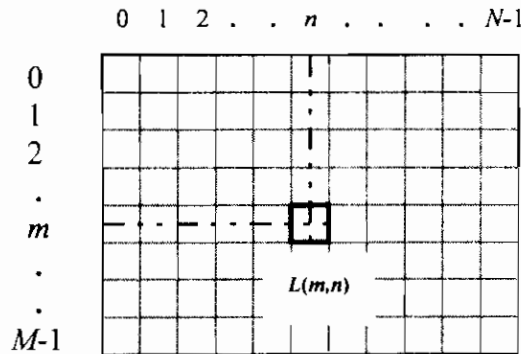
El método más común divide a la imagen en una matriz de *pixeles*, cada uno con una posición específica (fila, columna), dicha matriz constituye la entrada para el

<sup>1</sup> Originalmente el término *pixel* se utilizaba para referirse al menor elemento de fósforo que el haz de electrones podía excitar en el monitor de video. Actualmente el término se ha generalizado para referirse al menor punto direccionable en el monitor gráfico [19].

procesamiento de la imagen. La figura 1.3.2.1 muestra una matriz que representa una imagen monocromática en un computador, al desplegarla en un monitor o al presentarla mediante un dispositivo de salida. En esta matriz el valor de cada una de las celdas representa el nivel de gris de un pixel.

En esta matriz de pixeles se puede verificar que cada uno de ellos están especificados por la coordenada vertical " $m$ ", la coordenada horizontal " $n$ " y su nivel de gris " $L(m,n)$ " (para el caso policromático,  $C(m,n)$  representará la tonalidad de color en el punto  $(m,n)$ ).

El rango de los niveles de gris puede ser representado por un byte, logrando con ello 256 posibles niveles, donde el nivel 0 representa color negro y nivel 255 el color blanco.



**Figura 1.3.2.1** Imagen monocromática. El valor de cada una de las celdas representa el nivel de gris de un pixel [14].

En una imagen monocromática todos los pixeles se refieren de la siguiente forma [14]:

$$[I] = [L(m,n) ; m = 0, \dots, M - 1; n = 0, \dots, N - 1] \tag{1.3.2.1}$$

Si la imagen es policromática se tiene:

$$[I] = [C(m,n) ; m = 0, \dots, M - 1; n = 0, \dots, N - 1] \tag{1.3.2.2}$$

La tonalidad de color de cada pixel está determinada por la combinación de los tres colores primarios que son el rojo (Red), verde (Green) y azul (Blue), y, por tanto<sup>1</sup>:

$$C(m, n) = \{R(m, n), G(m, n), B(m, n)\} \quad (1.3.2.3)$$

## 1.4 TEORÍA DE MATRICES

Al desarrollar modelos matemáticos para la descripción y tratamiento de sistemas de imágenes, es muy factible encontrarse frente a sistemas de ecuaciones involucradas que pueden complicarse debido a la multiplicidad de puntos de entrada y salida, en realidad el número de ecuaciones puede llegar a los centenares. La simplicidad lograda usando representación mediante notación vectorial puede resultar muy conveniente para manipular y resolver problemas complejos mediante el concurso de un computador.

A continuación se dan algunas definiciones de mucha utilidad sobre teoría de matrices cuyas bases servirán para el subsiguiente análisis.

### 1.4.1 DEFINICIONES DE MATRICES

Con frecuencia secuencias unidimensionales y bidimensionales son representadas mediante vectores y matrices respectivamente.

#### Vector

Un vector columna  $\mathbf{u}$  de  $N$ -elementos complejos, es un arreglo de la forma<sup>2</sup>,

---

<sup>1</sup> El modelo de color especificado es el RGB, el mismo que es utilizado en monitores a color.

<sup>2</sup> Tanto los vectores como las matrices se notarán con letras en negrilla.

$$\mathbf{u} = \{u(n)\} = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N-1) \end{bmatrix} \quad (1.4.1.1)$$

el  $n$ -ésimo elemento de  $\mathbf{u}$  es denotado por  $u(n)$ ,  $u_n$ ,  $\mathbf{u}_n$  o  $[\mathbf{u}]_n$ . A menos que se especifique otra cosa los vectores siempre serán vectores columna, es decir un vector  $N \times 1$ .

### Matriz

Se define una matriz como un arreglo rectangular donde sus elementos pueden ser números reales, números complejos, matrices mismas (éste es el caso de las matrices-bloque), funciones u operadores.

Una matriz  $\mathbf{A}$  de tamaño  $M \times N$ , contiene  $M$  filas y  $N$  columnas y está definida de la siguiente forma:

$$\mathbf{A} = \{a(m,n)\} = \begin{bmatrix} a(0,0) & a(0,1) & \cdots & a(0,N-1) \\ a(1,0) & a(1,1) & \cdots & a(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ a(M-1,0) & a(M-1,1) & \cdots & a(M-1,N-1) \end{bmatrix} \quad (1.4.1.2)$$

donde,  $a(m,n) = [\mathbf{A}]_{m,n} = a_{m,n} = \mathbf{A}_{m,n} = a_n(m)$  es el  $m$ -ésimo elemento del  $n$ -ésimo vector columna. El primer elemento de una matriz tiene índice (0,0), mientras que para el caso de un vector el primer elemento tiene índice (0).

Así también, cuando no se especifique el tamaño de la matriz, se supondrá una matriz cuadrada de tamaño  $N \times N$ .

### Ordenamiento por filas y columnas

En ocasiones es conveniente escribir una matriz en forma de un vector, por ejemplo para almacenar una imagen en cinta o disco. Esto se consigue por apilamiento de cada fila o columna respectivamente a la derecha o debajo de la anterior.

Para una matriz  $M \times N$ , el vector ordenado por filas es [1],[2],

$$\mathbf{x}_r = [x(0,0) \cdots x(0,N-1) \ x(1,0) \cdots x(1,N-1) \cdots x(M-1,0) \cdots x(M-1,N-1)]^T \quad (1.4.1.3)$$

Si se ordena dicha matriz por columnas se tendrá:

$$\mathbf{x}_c = [x(0,0) \cdots x(M-1,0) \ x(0,1) \cdots x(M-1,1) \cdots x(0,N-1) \cdots x(M-1,N-1)]^T \quad (1.4.1.4)$$

### Matriz diagonal

Es aquella en la cual todos los elementos ubicados fuera de la diagonal principal son cero<sup>1</sup>,

$$\mathbf{A} = \{a(m,n)\delta(m-n,n-m)\} \quad (1.4.1.5)$$

con frecuencia a la matriz diagonal se la escribe como [15],

$$\text{diag}(a_{0,0}, a_{1,1}, \dots, a_{n,n}, \dots, a_{N-1,N-1}) \quad \text{si } 0 \leq n \leq N-1 \quad (1.4.1.6)$$

### Matriz unidad o identidad

Es la matriz diagonal con la particularidad siguiente,

$$\mathbf{I} = \text{diag}(1,1,\dots,1) \quad (1.4.1.7)$$

### Matriz no-singular

Es aquella matriz cuyo determinante es diferente de cero, haciendo factible su inversión.

---

<sup>1</sup> En la definición de la matriz diagonal se utiliza la función  $\delta(m,n)$  la cual se define en el Apéndice A.

El determinante de la matriz  $\mathbf{A}$  desarrollado en función de los elementos de la fila  $k$  es [16],

$$\det \mathbf{A} = |\mathbf{A}| = \sum_{j=1}^N (-1)^{k+j} a_{kj} \det \mathbf{A}_{k,j} \quad (1.4.1.8)$$

en donde  $\mathbf{A}_{k,j}$  es la matriz  $\mathbf{A}$  sin considerar la fila  $k$  y columna  $j$ .

### Matriz inversa

La matriz inversa de  $\mathbf{A}$  es la matriz  $\mathbf{A}^{-1}$  que cumple,

$$\mathbf{A}^{-1} \mathbf{A} = \mathbf{A} \mathbf{A}^{-1} = \mathbf{I} \quad (1.4.1.9)$$

### Traza

Se conoce como *traza* de una matriz a la sumatoria de los elementos de la diagonal principal de dicha matriz,

$$Tr[\mathbf{A}] = \sum_n a(n,n) \quad (1.4.1.10a)$$

Una propiedad de mucha utilidad es,

$$Tr[\mathbf{AB}] = Tr[\mathbf{BA}] \quad (1.4.1.10b)$$

### Matriz transpuesta

La matriz  $\mathbf{B}$  de orden  $N \times M$  es la matriz transpuesta de la matriz  $\mathbf{A}$  de orden  $M \times N$ , si cumple que  $a_{m,n} = b_{n,m}$ . Se nota  $\mathbf{B} = \mathbf{A}^t$ . La transpuesta del producto de dos matrices cumple que  $(\mathbf{AB})^t = \mathbf{B}^t \mathbf{A}^t$ .

### Matriz simétrica

La matriz  $\mathbf{A}$  es simétrica ssi  $\mathbf{A} = \mathbf{A}^t$ .

### Matriz conjugada

La matriz conjugada de  $\mathbf{A}$  es  $\mathbf{A}^*$  si cada uno de sus elementos se reemplaza por su complejo conjugado. Es decir si  $a_{m,n}^*$  es el complejo conjugado de  $a_{m,n}$ . Esta operación se distribuye normalmente sobre la multiplicación de matrices.

### Matriz hermítica

Se dice que  $\mathbf{A}$  es hermítica si y solo si  $\mathbf{A} = \mathbf{A}^{*T}$ .

Una matriz hermítica se dice *definida positiva* si ( $Q > 0$ ) o *semidefinida positiva (no negativa)* si ( $Q \geq 0$ ), donde  $Q$  es la forma cuadrática definida como [1],

$$Q = \mathbf{x}^{*T} \mathbf{A} \mathbf{x}, \quad \forall \mathbf{x} \neq \mathbf{0} \quad (1.4.1.11)$$

Así si  $\mathbf{A}$  es definida positiva, entonces todos sus valores propios  $\lambda_k$  son positivos, y el determinante de  $\mathbf{A}$  cumple la desigualdad,

$$|\mathbf{A}| = \prod_{k=0}^{N-1} \lambda_k \leq \prod_{k=0}^{N-1} a(k,k) \quad (1.4.1.12)$$

### Valores y vectores propios, diagonalización

Si  $\lambda$  es un número real y  $\mathbf{x}$  un vector  $N \times 1$  se llaman valor y vector propio respectivamente de una matriz cuadrada de tamaño  $N \times N$ , denotada  $\mathbf{R}$ , si y solo si,

$$\mathbf{R}_n \mathbf{x}_k = \lambda_k \mathbf{x}_k \quad k = 0, \dots, N-1 \quad (1.4.1.13)$$

$\lambda$  es la raíz de la ecuación característica de la matriz  $\mathbf{R}$ , encontrada como,

$$\det[\lambda \mathbf{I} - \mathbf{A}] = \det[\mathbf{A} - \lambda \mathbf{I}] = 0 \quad (1.4.1.14)$$

Si la matriz  $\mathbf{R}$  es hermítica, entonces existe una matriz ortogonal tal que [1],

$$\Lambda = \Phi^{*T} \mathbf{R} \Phi \quad (1.4.1.15)$$

donde la matriz denotada  $\Lambda$  es una matriz diagonal que contiene los valores propios de  $\mathbf{R}$ . Para una matriz hermítica vectores propios correspondientes a valores propios distintos son ortogonales. La matriz  $\Phi$  se llama matriz propia y sus columnas serán los vectores propios ortonormalizados de  $\mathbf{R}$ .

### Matriz ortogonal

La matriz  $\mathbf{A}$  es ortogonal si su matriz inversa es igual a su transpuesta, es decir  $\mathbf{A}$  es ortogonal si,

$$\mathbf{A} \mathbf{A}^T = \mathbf{A}^T \mathbf{A} = \mathbf{I} \quad (1.4.1.16)$$

### Matriz unitaria

La matriz  $\mathbf{A}$  se dice unitaria si su inversa es igual a su transpuesta conjugada, es decir,

$$\mathbf{A} \mathbf{A}^{*T} = \mathbf{A}^{*T} \mathbf{A} = \mathbf{I} \quad (1.4.1.17)$$

### Producto Kronecker

Sean las matrices  $\mathbf{A}$  de orden  $M_1 \times N_1$ ,  $\mathbf{B}$  de orden  $M_2 \times N_2$  su producto kronecker se define como [1],

$$\mathbf{A} \otimes \mathbf{B} = \{a(m,n)\mathbf{B}\} = \begin{bmatrix} a(0,0)\mathbf{B} & \cdots & a(0,N_1-1)\mathbf{B} \\ a(1,0)\mathbf{B} & \cdots & a(1,N_1-1)\mathbf{B} \\ \vdots & \ddots & \vdots \\ a(M_1-1,0)\mathbf{B} & \cdots & a(M_1-1,N_1-1)\mathbf{B} \end{bmatrix}_{M_1 M_2 \times N_1 N_2} \quad (1.4.1.18)$$

Nótese que,

$$\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A} \quad (1.4.1.19)$$



El producto kronecker es de mucha utilidad en la generación de matrices de orden superior partiendo de matrices de orden inferior. Algunas de sus propiedades se presentan en la tabla 2.7 de la referencia [1].

### Matriz Toeplitz

Es aquella que tiene elementos constantes a lo largo de la diagonal principal y de las subdiagonales. Esto es, los elementos  $(m,n)$  dependen solamente de la diferencia  $m - n$ , es decir  $t(m,n) = t_{m-n}$ , dando la siguiente matriz  $\mathbf{T}$ ,

$$\mathbf{T} = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-N+2} & t_{-N+1} \\ t_1 & t_0 & \ddots & \ddots & t_{-N+2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ t_{N-2} & \ddots & \ddots & \ddots & t_{-1} \\ t_{N-1} & t_{N-2} & \cdots & t_1 & t_0 \end{bmatrix} \quad (1.4.1.20)$$

la cual se define completamente con solo  $(2N - 1)$  elementos. Son de utilidad para describir secuencias estacionarias, secuencias definidas en el apartado 1.5.

## 1.5 TEORÍA DE PROBABILIDADES

Las señales determinísticas son aquellas que pueden especificarse de manera exacta, es decir en sus valores no hay incertidumbre y casi siempre existe una relación matemática para expresar tales señales de una manera explícita, de otro modo siempre es expresable en términos de sumatorias de funciones explícitas como por ejemplo las series de Fourier.

La teoría de probabilidades permite describir señales sobre las que se tiene cierto grado de incertidumbre antes de que ocurran, dichas señales son conocidas como "no determinísticas" o *aleatorias*.

## 1.5.1 CONCEPTOS PROBABILÍSTICOS

A continuación se expondrán selectivamente algunos conceptos de utilidad para capítulos posteriores sobre la teoría de probabilidades.

### Probabilidad

La *probabilidad* numéricamente mide los eventos favorables de un determinado experimento, una manera simple de definir la probabilidad de ocurrencia de un resultado es la siguiente, bajo la suposición (no siempre cierta) de resultados equiprobables [8],

$$P(A) = \frac{\text{número de resultados posibles favorables para el evento } A}{\text{número total de resultados equiprobables}} = \frac{N_A}{N} \quad (1.5.1.1)$$

dicha probabilidad de ocurrencia del resultado favorable cumplirá la condición,

$$0 \leq P(A) \leq 1 \quad (1.5.1.2)$$

Dos eventos son mutuamente excluyentes, si es imposible que ocurran simultáneamente, es decir  $P(AB) = 0$ .

### Variable aleatoria

Es un concepto que permite asignar de manera arbitraria (según convenga) un número real a cada uno de los posibles resultados de un experimento sin que dicho número represente la medida del resultado. Es muy frecuente usar letras mayúsculas para designar las variables aleatorias y su correspondiente minúscula para los valores dentro del intervalo.

Si el número de posibles resultados es finito la variable aleatoria será discreta, caso contrario será continua.

**Función de distribución acumulativa**

Una gráfica de las probabilidades de los resultados de un determinado experimento en función de la variable aleatoria, será de mucha utilidad para definir esta función. Notada como  $F_X(x)$  y asociada a la variable aleatoria  $X$ , se define así,

$$F_X(x) = P[X \leq x] \quad (1.5.1.3)$$

y representa la probabilidad de ocurrencia en un intervalo específico.

Esta función en base al concepto de probabilidad cumplirá las propiedades siguientes,

$$0 \leq F_X(x) \leq 1 \quad (1.5.1.4)$$

$$F_X(x_1) \leq F_X(x_2) \quad , \quad \text{si} \quad x_1 \leq x_2 \quad (1.5.1.5)$$

$$\begin{aligned} F_X(-\infty) &= 0 \\ F_X(\infty) &= 1 \end{aligned} \quad (1.5.1.6)$$

**Función de densidad de probabilidad (fdp)**

Es más conveniente definir una nueva función cuya área sea la función de distribución acumulativa; esta función denominada *función de densidad de probabilidad*, notada como  $p_X(x)$  se la define así,

$$F_X(x) = \int_{-\infty}^x p_X(x) dx \quad (1.5.1.7)$$

derivando ambos miembros de esta ecuación se tiene:

$$p_X(x) = \frac{d}{dx} F_X(x) \quad (1.5.1.8)$$

para el caso discreto, se convierte en:

$$F_X(x) = \sum_i P_i u_{-1}(x - x_i) \quad (1.5.1.9)$$

relacionando las dos últimas expresiones se llega a<sup>1</sup>,

$$p(x) = \sum_i P_i \delta(x - x_i) \quad (1.5.1.10)$$

donde,  $P_i = P(X = x_i)$ .

### Valor esperado

Se denomina al promedio estadístico de  $X$  y es el promedio numérico de los valores que puede tomar la variable aleatoria  $X$  (ó una función de ésta), se lo llama también valor medio o valor promedio de  $X$  [8],

$$E[X] = \mu_X = m_X = \int_{-\infty}^{\infty} x p(x) dx \quad (1.5.1.11)$$

para el caso de una variable aleatoria discreta el valor medio es,

$$m_X = \sum_{i=1}^N x_i P_i \quad (1.5.1.12)$$

En general el valor esperado  $E[X^n]$  se llama "*n-ésimo momento*" de la variable aleatoria  $X$ , y es,

$$E[X^n] = \int_{-\infty}^{\infty} x^n p(x) dx \quad (1.5.1.13)$$

### Varianza

Restar el primer momento  $m_X$  de la variable aleatoria, equivale en mecánica a referir los momentos al centro de gravedad más no al origen del sistema de coordenadas, los nuevos momentos así establecidos se llaman *centrales*.

<sup>1</sup> Las funciones paso-unitario  $u_{-1}(x)$  e impulso-unitario  $\delta(x)$ , están definidas en el Apéndice A.

El segundo momento central de la variable aleatoria  $X$ , se denomina *varianza*, definida como [8],

$$\sigma_X^2 = E[(X - m_X)^2] = \int_{-\infty}^{\infty} (x - m_X)^2 p(x) dx \quad (1.5.1.14)$$

La probabilidad de que  $X$  se encuentre cerca de  $m_X$  es mayor cuando  $\sigma_X$  es menor. es decir la *desviación estándar*,  $\sigma_X$ , es un indicador del ancho efectivo de la *fdp* de la variable  $X$ .

Para el caso discreto la varianza de  $X$ , con  $N$  posibles resultados es,

$$\sigma_X^2 = \sum_{i=1}^N (x_i - m_X)^2 P_i \quad (1.5.1.15)$$

### **Funciones de densidad conjunta y condicional**

En muchas ocasiones para describir un proceso físico se requiere el uso de dos variables aleatorias. En este caso se habla del sistema bidimensional que no es más que la extensión del caso unidimensional.

La probabilidad de encontrar simultáneamente  $X$  en el rango:  $x_1 < X \leq x_2$  e  $Y$  en:  $y_1 < Y \leq y_2$  es [8],

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} p_{XY}(x, y) dx dy \quad (1.5.1.16)$$

donde  $p_{XY}(x, y)$  es la *fdp* conjunta de las variables aleatorias  $X$  e  $Y$ .

Las variables aleatorias  $X, Y$  son ortogonales si,

$$E[XY^*] = 0 \quad (1.5.1.17)$$

### Correlación entre variables aleatorias

La *covarianza* de dos variables aleatorias, denotada como “ $r$ ”, se define en forma análoga a la varianza de una sola variable así [8],

$$r = E[(X - m_x)(Y - m_y)] \quad (1.5.1.18)$$

Si  $X$  e  $Y$  son estadísticamente independientes<sup>1</sup>, se tiene  $r = 0$ .

Si  $X$  e  $Y$  son completamente dependientes, es decir  $X = Y$  ó  $X = -Y$ , entonces,

$$|r| = \sigma_x \sigma_y \quad (1.5.1.19)$$

El *coeficiente de correlación* se define como la covarianza normalizada,

$$\rho = \frac{r}{\sigma_x \sigma_y}, \quad -1 \leq \rho \leq 1 \quad (1.5.1.20)$$

y determina hasta que punto las dos variables aleatorias son dependientes, de esto se puede concluir que si dos variables aleatorias son estadísticamente independientes entonces no están correlacionadas. La afirmación recíproca no es verdadera.

Desarrollando la ecuación (1.5.1.18) se tiene,

$$r = E[XY] - m_x m_y$$

de este desarrollo el primer término se conoce como la *correlación* de  $X$  e  $Y$ , entonces,

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy p_{xy}(x, y) dx dy \quad (1.5.1.21)$$

## 1.5.2 PROCESOS ALEATORIOS

Para determinar las probabilidades de ocurrencia de los posibles resultados de un experimento, este habrá que repetirlo varias veces (uno a continuación de otro), pudiendo

<sup>1</sup> Si la ocurrencia de  $X$  no influye de ninguna manera en la ocurrencia de  $Y$  se dice que hay independencia estadística entre las variables  $X$  e  $Y$ .

considerarse como un generador de números aleatorios. Supóngase un conjunto  $X(t)$  de estos generadores accionados por un reloj común, esta colección  $\{X_i(t), i = 1, 2, \dots\}$  es conocida como *conjunto*. A cada una de estas señales generadas se conoce como *funciones de muestra* y su agrupación conforma el *proceso aleatorio*,  $X(t)$ . En general para  $k$  instantes de tiempo  $\{t_1, t_2, \dots, t_k\}$  se define el evento conjunto  $\{X(t_1) \leq x_1, \dots, X(t_k) \leq x_k\}$  donde  $X(t_i)$  es una variable aleatoria tomada del proceso  $X(t)$  al tiempo  $t = t_i$ , la función de distribución conjunta es:  $F_{X(t)}(x) = P(X(t_1) \leq x_1, \dots, X(t_k) \leq x_k)$ . El promedio de conjunto es el promedio estadístico obtenido a partir de la evaluación de todas las funciones de muestra en un instante determinado. La evaluación del proceso  $X(t)$  en  $k$  instantes de tiempo forma  $k$  variables aleatorias, en realidad  $k$  puede ser infinito, una variable aleatoria para cada número real. Los promedios de tiempo son los observados en una sola función de muestra (a una sola función muestra se la nombra en ocasiones como *una realización* del proceso) por período de tiempo (esto equivale a lanzar una moneda 100 veces sucesivamente, a diferencia de lo anterior de lanzar 100 monedas simultáneamente).

Un proceso aleatorio se dice *estacionario en el sentido amplio*, cuando las características estadísticas de las funciones de muestra son constantes en el tiempo ( $E[X(t_j)] = E[X(t_k)] = m_X$ , y los estadísticos de segundo orden son funciones de la diferencia de tiempos). Un proceso aleatorio es *ergódico* si los promedios de tiempo y de conjunto son iguales, entonces todos sus estadísticos pueden determinarse a partir de una sola función de muestra, pues en la práctica es lo único disponible. Un proceso ergódico es necesariamente estacionario pero no a la inversa. Con este supuesto igualando los promedios de conjunto con los de tiempo se tendrá,

$$E[X] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt \quad (1.5.2.1)$$

$$E[X^2] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \quad (1.5.2.2)$$

con estos conceptos sobre procesos aleatorios se puede ya definir la autocorrelación.

### Autocorrelación

La función  $a_{XX}(t_j, t_k)$  definida como se muestra a continuación se denomina función de *autocorrelación* del proceso aleatorio  $X(t)$  evaluada en  $t = (t_j, t_k)$  [8],

$$a_{XX}(t_j, t_k) = E[X(t_j)X(t_k)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_j x_k p_{X_j, X_k}(x_j, x_k) dx_j dx_k \quad (1.5.2.3)$$

para señales complejas, de  $X_k$  se tomará su conjugado.

Para el caso estacionario, el primer momento es constante en el tiempo y el segundo momento  $E[X_j X_k]$  depende solamente de la diferencia entre los tiempos y no del valor absoluto de  $t$ , es decir,

$$a_{XX}(t_j, t_k) = a_{XX}(t_k - t_j) \quad (1.5.2.4)$$

haciendo  $\tau = t_k - t_j$ , se tendrá lo siguiente,

$$a_{XX} = E[x(t)x(t + \tau)] \quad (1.5.2.5)$$

Un proceso aleatorio será *estacionario en el sentido estricto* si la *fdp* conjunta solo depende de  $\tau$  y no de los valores absolutos de  $t_j$  y  $t_k$ .

Este segundo momento no describe completamente al fenómeno, pero si da mucha información, particularmente proporciona una medida de la dependencia del valor de una función de muestra respecto a otro valor desplazado  $\tau$  unidades de tiempo. Si  $\tau$  es



suficientemente grande como para que  $X(t_j)$  y  $X(t_k)$  sean estadísticamente independientes la función de autocorrelación se simplifica a,

$$a_{xx}(\tau) = E[x(t)x(t+\tau)] = m_x^2 \quad (1.5.2.6)$$

La función de autocorrelación da una idea de como los valores presentes dependen de los valores pasados de  $X(t)$ , consecuentemente que tan rápido puede variar la función de muestra dada.

### Procesos aleatorios discretos en el tiempo

Una simple realización muestreada del proceso  $X(t)$  se representa por:

$$\{x(n), n = \dots, -1, 0, 1, \dots\} \quad (1.5.2.7)$$

suponiendo el intervalo de observación del proceso  $X(t)$  estacionario y representado por el conjunto de muestras:

$$\{x(n), n = 0, 1, \dots, N-1\} \quad (1.5.2.8)$$

la matriz de autocorrelación correspondiente se define como [6]:

$$a_x = E[\mathbf{X}\mathbf{X}^T], \quad \mathbf{X} = \begin{bmatrix} x(0) \\ \vdots \\ x(N-1) \end{bmatrix} \quad (1.5.2.9)$$

$$a_x = \begin{bmatrix} E[x(0)x(0)] & \dots & E[x(0)x(N-1)] \\ \vdots & \ddots & \vdots \\ E[x(N-1)x(0)] & \dots & E[x(N-1)x(N-1)] \end{bmatrix} \quad (1.5.2.10)$$

### Campos aleatorios discretos

Cuando el proceso aleatorio  $X(t)$  es *discreto bidimensional*, a este se conoce como “*Campo aleatorio discreto*”. Denótese a un infinito número de variables aleatorias complejas por  $\{u(m,n)\}$ , en donde  $u(m,n)$  es una variable aleatoria evaluada en el punto

$(m,n)$ . El campo aleatorio  $\{u(m,n)\}$  se caracteriza completamente mediante una *fdp* conjunta de todas las variables aleatorias, para tamaños prácticos de imágenes esto es una tarea muy difícil, pues el número de variables aleatorias es grande, por ejemplo para imágenes  $512 \times 512$  pixeles dicho número es 262144. Una alternativa inmediata es especificar este conjunto usando momentos de primer y segundo orden (esto es, valor promedio y función covarianza).

Si se obtiene una muestra (función muestra ó realización) del campo aleatorio  $\{u(m,n)\}$  constituido por un conjunto de funciones muestras bidimensionales, el resultado será una secuencia 2-D, señal aleatoria denotada también por  $u(m,n)$ . Por tanto una imagen  $u(m,n)$  se considera como una función muestra de un conjunto de imágenes, en este caso al conjunto se lo llama imagen aleatoria.

La función de autocorrelación del campo  $\{u(m,n)\}$ ,  $a_u(m,n;m',n')$  se define como se muestra a continuación:

$$a_u(m,n;m',n') = E[u(m,n) u^*(m',n')] \quad (1.5.2.11)$$

La función covarianza de  $u(m,n)$ , denotada  $r_u(m,n;m',n')$  se define así;

$$\begin{aligned} r_u(m,n;m',n') &= E[(u(m,n) - E[u(m,n)])(u(m',n') - E[u(m',n')])^*] \\ &= a_u(m,n;m',n') - E[u(m,n)] E[u^*(m',n')] \end{aligned} \quad (1.5.2.12)$$

El proceso aleatorio se llama de *valor promedio nulo* si;

$$E[u(m,n)] = 0, \quad \forall(m,n)$$

Un campo aleatorio  $\{u(m,n)\}$  se dice estacionario en el *sentido estricto*, si su función densidad de probabilidad es invariante a los desplazamientos del origen del índice  $(m,n)$ , es decir, para  $k$  y  $l$  enteros fijos [3]:

$$P_{u(m,n),u(m',n'),\dots}(u_1, u_2, \dots) = P_{u(m+k,n+l),u(m'+k,n'+l),\dots}(u_1, u_2, \dots) \quad (1.5.2.13)$$

Para un campo aleatorio estacionario discreto se tiene:

$$\begin{aligned} E[u(m,n)] &= m_u = \mu_u = \text{constante} \quad \forall(m,n) \\ a_u(m,n;m',n') &= a_u(m-m',n-n';0,0) \\ r_u(m,n;m',n') &= r_u(m-m',n-n';0,0) \end{aligned} \quad (1.5.2.14)$$

luego, las dos últimas expresiones de (1.5.2.14) pueden escribirse como:

$$\begin{aligned} a_u(m,n) &= a_u(m',n';m'-m,n'-n) \\ r_u(m,n) &= r_u(m',n';m'-m,n'-n) \end{aligned} \quad (1.5.2.15)$$

Entonces un campo aleatorio que cumple (1.5.2.14) y no necesariamente (1.5.2.13) se dice estacionario en el *sentido amplio* (referido simplemente como *estacionario*).

De (1.5.2.11) y (1.5.2.12) puede obtenerse que:

$$r(m,n;m',n') = r^*(m',n';m,n) \quad (1.5.2.16)$$

para el caso estacionario será:

$$r(m,n) = r^*(-m,-n) \quad (1.5.2.17)$$

Supóngase que se desea estimar  $m_u = E[u(m,n)]$ , puesto que  $m_u$  representa el promedio de conjunto, se requerirá todo un conjunto de  $u(m,n)$  para cada punto particular  $(m,n)$ , pero en la práctica de lo único que se dispone es de una simple realización del campo  $\{u(m,n)\}$ . Si el proceso es *Ergódico*, entonces sus estadísticos pueden calcularse desde la única realización como a continuación [3]:

$$m_u = E[u(m,n)] = \lim_{N \rightarrow \infty} \frac{1}{(2N+1)^2} \sum_{m'=-N}^N \sum_{n'=-N}^N (u(m,n)) \quad (1.5.2.18)$$

$$r_u(m,n) = \lim_{N \rightarrow \infty} \frac{1}{(2N+1)^2} \sum_{m'=-N}^N \sum_{n'=-N}^N (u(m',n') - m_u)(u(m'-m,n'-n) - m_u)^* \quad (1.5.2.19)$$

Nota: "Ergodicidad implica estacionaridad, pero estacionaridad no implica ergodicidad".

El espectro de Potencia de un campo aleatorio estacionario  $u(m,n)$ , denotado  $P_u(\omega_1, \omega_2)$  está definido por:

$$P_u(\omega_1, \omega_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} r_u(m,n) \exp(-j(\omega_1 m + \omega_2 n)) \quad (1.5.2.20)$$

usando (1.5.2.17) y (1.5.2.20) se obtiene:

$$P_u(\omega_1, \omega_2) = P_u^*(\omega_1, \omega_2) \quad (1.5.2.21)$$

La potencia promedio del campo aleatorio  $u(m,n)$ , es  $\sigma_u^2$ , es decir [3],

$$r_u(0,0) = \sigma_u^2 = E[(u(m,n) - m_u)(u(m,n) - m_u)^*] \quad (1.5.2.22)$$

En la práctica normalmente se dispone apenas de una realización de corta duración de un campo  $\{u(m,n)\}$ , supóngase que se dispone de la secuencia de variables aleatorias  $\{u(m,n), m = 0, \dots, M-1; n = 0, \dots, N-1\}$  del proceso, para caracterizar el proceso es de suponer que el proceso es ergódico, es decir, que sus parámetros estadísticos  $m_u$  y  $r_u$  pueden determinarse partiendo de una sola realización (con medias espaciales), estas cantidades serán estimadas desde la imagen dada, reemplazando los promedios de conjunto por los promedios espaciales. Para una imagen de tamaño  $M \times N$ , según la referencia [1]:

$$m_u = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n) \quad (1.5.2.23)$$

$$r_u(m,n) = \frac{1}{(M-m)(N-n)} \sum_{m'=0}^{M-1-m} \sum_{n'=0}^{N-1-n} (u(m',n') - m_u)(u(m+m',n+n') - m_u) \quad (1.5.2.24)$$

### 1.5.3 FUNCIÓN TASA DE DISTORSIÓN

La función tasa de distorsión de una variable aleatoria Gaussiana "x" de varianza  $\sigma^2$ , da la razón promedio mínima  $R_D$  (en bits/muestra) requerida para representarla, dada una

distorsión  $D$  en su valor reproducido “ $y$ ”, dicha distorsión se mide como el valor cuadrático medio de su diferencia, esto es,

$$D = E[(x - y)^2] \quad (1.5.3.1)$$

y la función tasa de distorsión según [1] se define como:

$$R_D = \max\left\{0, \frac{1}{2} \log_2\left(\frac{\sigma^2}{D}\right)\right\} \quad (1.5.3.2)$$

para el caso de un vector Gaussiano  $\mathbf{x} = [x(0) \ x(1) \ \dots \ x(N-1)]^T$  y su vector reproducido  $\mathbf{y} = [y(0) \ y(1) \ \dots \ y(N-1)]^T$ , la distorsión cuadrática media promedio es:

$$D = \frac{1}{N} \sum_{k=0}^{N-1} E[(x(k) - y(k))^2] \quad (1.5.3.3)$$

Para un valor fijo  $D$ , la función tasa de distorsión  $R_D$ , se encuentra como:

$$R_D = \frac{1}{N} \sum_{k=0}^{N-1} \max\left\{0, \frac{1}{2} \log_2\left(\frac{\sigma_k^2}{\theta}\right)\right\} \quad (1.5.3.4)$$

donde  $\theta$  se encuentra resolviendo la ecuación siguiente:

$$D = \frac{1}{N} \sum_{k=0}^{N-1} \min\{\theta, \sigma_k^2\} \quad (1.5.3.5)$$

si  $\theta \leq \sigma_k^2 \ \forall k$  entonces  $D = \theta$  y [35],

$$R_D = \frac{1}{2N} \sum_{k=0}^{N-1} \log_2(\theta_k^2) - \frac{1}{2} \log_2(D) \quad (1.5.3.6)$$

# CAPÍTULO

---

# 2

## **TRANSFORMADAS DISCRETAS BIDIMENSIONALES**

### **2.0 INTRODUCCIÓN**

En este capítulo se incluye la descripción matemática de las transformadas discretas bidimensionales, en un primer punto se presenta las definiciones generales para transformadas unitarias y ortogonales así como también las propiedades más representativas que regulan su comportamiento, en un segundo punto se procede a describir cada una de las transformadas que son usadas en el tratamiento de imágenes, de las cuales se expone brevemente su definición y sus propiedades más relevantes. Parte complementaria a este capítulo constituye el Apéndice B, el cual contiene los desarrollos matemáticos de las ecuaciones cuya obtención no es del todo evidente.

### **2.1 TRANSFORMADAS UNITARIAS Y ORTOGONALES**

Una señal unidimensional puede ser representada por una serie ortogonal de funciones bases, tal como se señaló en la sección 1.3.1, de igual manera una imagen puede ser

expresada en términos de arreglos discretos denominadas *imágenes bases*, éstas últimas serán generadas por matrices *unitarias* denominadas *transformadas imagen*.

Para una secuencia unidimensional representada por un vector  $\mathbf{u}$  de tamaño  $N$   $\{u(n), 0 \leq n \leq N-1\}$ . El vector transformado  $\mathbf{v}$ , usando una matriz de *transformación unitaria*<sup>1</sup>  $\mathbf{A}$ , será escrito según la referencia [1] como:

$$\mathbf{v} = \mathbf{A}\mathbf{u} \Rightarrow v(k) = \sum_{n=0}^{N-1} a(k,n)u(n) \quad , 0 \leq k \leq N-1 \quad (2.1.0.1)$$

haciendo uso de la correspondiente transformada inversa se puede recuperar  $\mathbf{u}$ , esto es,

$$\mathbf{u} = \mathbf{A}^{*T}\mathbf{v} \Rightarrow u(n) = \sum_{k=0}^{N-1} a^{*}(k,n)v(k) \quad , 0 \leq n \leq N-1 \quad (2.1.0.2)$$

los vectores columnas de  $\mathbf{A}^{*T}$ , esto es  $\mathbf{a}_k^{*} = \{a^{*}(k,n) \quad , 0 \leq n \leq N-1\}^T$  son llamados *vectores bases de A* (nótese que los vectores bases corresponden a las filas de la matriz  $\mathbf{A}^{*}$ ). La ecuación (2.1.0.2) proporciona una representación en serie de la secuencia  $u(n)$ .

Las ecuaciones (2.1.0.1) y (2.1.0.2) forman el llamado par de transformadas (al hablar de par de transformadas se incluye la transformada directa e inversa), para el caso discreto unidimensional. Con este sustento se procede inmediatamente a su extensión para el caso bidimensional.

Para una imagen  $u(m,n)$  de tamaño  $M \times N$  el par de transformadas de acuerdo a las referencias [1], [2] y [3], toma la siguiente forma:

$$v(k,l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n)a_{k,l}(m,n) \quad , 0 \leq k \leq M-1, 0 \leq l \leq N-1 \quad (2.1.0.3)$$

$$u(m,n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} v(k,l)a_{k,l}^{*}(m,n) \quad , 0 \leq m \leq M-1, 0 \leq n \leq N-1 \quad (2.1.0.4)$$

<sup>1</sup> Una matriz se dice unitaria si cumple con la ecuación (1.4.1.17).

en estas dos últimas ecuaciones, a los elementos  $v(k,l)$  se les conoce como los *coeficientes de la transformada* y la matriz que estos conforman  $\mathbf{V} = \{v(k,l)\}$  es la *imagen transformada*.

Debe indicarse además que tanto la transformada directa como la inversa son funciones, en consecuencia tienen un dominio (conjunto de partida) y un recorrido (conjunto de llegada). Para la transformada directa a su dominio se conoce como “dominio espacial” cubierto por las variables  $m$  y  $n$ , y a su recorrido como “dominio de la frecuencia espacial” cubierto por las variables  $k$  y  $l$ . La transformada inversa al ser la función inversa de la transformada directa tendrá como dominio el dominio de la frecuencia espacial y como recorrido el dominio espacial.

A la matriz  $\{a_{k,l}(m,n)\}$  se la conoce como *transformada imagen* y está conformada por un conjunto ortonormal discreto de funciones bases, las mismas que satisfacen las dos propiedades siguientes, ortonormalidad y suficiencia, respectivamente:

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{k,l}(m,n) a_{k',l'}^*(m,n) = \delta(k - k', l - l') \quad (2.1.0.5)$$

$$\sum_{k=0}^{M-1} \sum_{l=0}^{N-1} a_{k,l}(m,n) a_{k',l'}^*(m',n') = \delta(m - m', n - n') \quad (2.1.0.6)$$

Note que la propiedad de ortonormalidad se define en el dominio de la frecuencia espacial (variables  $k,l$ ) y la de suficiencia en el dominio espacial (variables  $m,n$ ).

En el cálculo de los coeficientes  $v(k,l)$  haciendo uso de la ecuación (2.1.0.3) el número de sumas y multiplicaciones requeridas para una imagen es del orden de  $(MN)^2$ . Si la



transformada es separable<sup>1</sup> la dimensionalidad del problema se ve reducida a un orden  $MN(M + N)$ , es decir si,

$$a_{k,l}(m,n) = a_k(m)b_l(n) = a(k,m)b(l,n) \quad (2.1.0.7)$$

Las matrices  $\mathbf{A} = \{a(k,m)\}$  y  $\mathbf{B} = \{b(l,n)\}$  son unitarias por tanto cumplen con la ecuación (1.4.1.17) y su obtención se fundamenta en la propiedad de separabilidad.

Nótese que la matriz  $\mathbf{A}$  relaciona la variable “ $m$ ” del plano  $m,n$  del dominio espacial con la correspondiente variable “ $k$ ” del plano  $k,l$  del dominio de la frecuencia espacial, de manera análoga la matriz  $\mathbf{B}$  relaciona las variables “ $n$ ” y “ $l$ ” que respectivamente son ortogonales a las variables “ $m$ ” y “ $k$ ”.

Por facilidad y considerando que las dos direcciones en cada plano en si son similares se toma  $\mathbf{A} = \mathbf{B}$  con esto (2.1.0.3) y (2.1.0.4) de acuerdo con [1] y [3] se reduce a:

$$v(k,l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a(k,m)u(m,n)a(l,n) \Leftrightarrow \mathbf{V} = \mathbf{AUA}^T \quad (2.1.0.8)$$

$$u(m,n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} a^*(k,m)v(k,l)a^*(l,n) \Leftrightarrow \mathbf{U} = \mathbf{A}^{*T} \mathbf{VA}^* \quad (2.1.0.9)$$

La sección B-1 del apéndice B, presenta el desarrollo para la obtención de (2.1.0.9) partiendo de (2.1.0.8).

### 2.1.1 IMÁGENES BASES

Denótese con  $\mathbf{a}_k^*$  a la  $k$ -ésima columna de la matriz  $\mathbf{A}^{*T}$ , se define  $\mathbf{A}_{k,j}^*$  como:

$$\mathbf{A}_{k,j}^* = \mathbf{a}_k^* \mathbf{a}_j^{*T} \quad (2.1.1.1)$$

<sup>1</sup> La propiedad de separabilidad se define en el apéndice A.

La siguiente ecuación expresa la imagen  $\mathbf{U}$  como una combinación lineal de  $MN$  matrices  $\mathbf{A}_{k,l}^*$ ,  $0 \leq k \leq M-1$ ,  $0 \leq l \leq N-1$ , a las mismas que se les llama *imágenes bases*,

$$\mathbf{U} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} v(k,l) \mathbf{A}_{k,l}^* \quad (2.1.1.2)$$

El proceso matemático que permite comprobar la veracidad de (2.1.1.2) se presenta en la sección B-2 del Apéndice B.

Mediante el uso de las imágenes bases se presenta a continuación una ecuación que permite determinar los coeficientes  $v(k,l)$ , así,

$$v(k,l) = \langle \mathbf{U}, \mathbf{A}_{k,l}^* \rangle \quad (2.1.1.3)$$

en conclusión los coeficientes  $v(k,l)$  según la ecuación (2.1.1.3) son simplemente el producto interno entre la  $k,l$ -ésima imagen base  $\mathbf{A}_{k,l}^*$  y la imagen  $\mathbf{U}$  (la demostración de esta igualdad se presenta en la sección B-3 del Apéndice B). En esta relación se interpreta a  $v(k,l)$  como la proyección de la matriz  $\mathbf{U}$  sobre la  $k,l$ -ésima imagen base.

Si se expresan las matrices  $\mathbf{U}$  y  $\mathbf{V}$  en vectores ordenados por filas ( $\mathbf{u}_r$  y  $\mathbf{v}_r$ , respectivamente), las ecuaciones (2.1.0.8) y (2.1.0.9) puede presentarse usando el producto Kronecker según [1] así:

$$\mathbf{v}_r = (\mathbf{A} \otimes \mathbf{A}) \mathbf{u}_r = \mathbf{A}_A \mathbf{u}_r \quad (2.1.1.4)$$

$$\mathbf{u}_r = \mathbf{A}_A^{*T} \mathbf{v}_r \quad (2.1.1.5)$$

en dichas expresiones,  $\mathbf{A}_A$  es matriz unitaria la misma que se calcula como el producto Kronecker de la matriz unitaria  $\mathbf{A}$  consigo misma, es decir:

$$\mathbf{A}_A = \mathbf{A} \otimes \mathbf{A} \quad (2.1.1.6)$$

### 2.1.2 PROPIEDADES DE LAS TRANSFORMADAS UNITARIAS

En este punto se presentan las propiedades comunes más importantes de las transformadas unitarias; en la descripción de cada transformada se incluirán sus propiedades particulares.

#### a).- Conservación de la Energía y Rotación

Una Transformada unitaria gráficamente corresponde a una rotación de las coordenadas bases y las componentes de  $\mathbf{v}$  constituyen las proyecciones de  $\mathbf{u}$  en las nuevas bases (las nuevas bases están dadas por las filas de la matriz  $\mathbf{A}^*$ ).

Para la siguiente transformada unitaria, se cumplirá la implicación señalada,

$$\mathbf{v} = \mathbf{A} \mathbf{u} \Rightarrow \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 \quad (2.1.2.1)$$

Para demostrar esta propiedad se debe recurrir a la definición de norma, adaptando de la expresión (1.3.1.2), para el caso discreto se llega a:

$$\|\mathbf{v}\|^2 = \sum_{n=0}^{N-1} |v(k)|^2 = \mathbf{v}^* \mathbf{v} = (\mathbf{A} \mathbf{u})^* \mathbf{A} \mathbf{u} = \mathbf{u}^* \mathbf{A}^* \mathbf{A} \mathbf{u} = \mathbf{u}^* \mathbf{I} \mathbf{u} = \mathbf{u}^* \mathbf{u} = \|\mathbf{u}\|^2$$

esta propiedad físicamente indica, que al aplicar la transformada unitaria a una señal se conserva la energía de la señal. De manera análoga para las transformadas unitarias 2-D, se cumple que:

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |u(m, n)|^2 = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |v(k, l)|^2 \quad (2.1.2.2)$$

#### b).- Compresión de Energía, varianza de los coeficientes de la Transformación

Algunas de las transformadas unitarias tienden a concentrar gran parte de su energía en pocos componentes de los coeficientes de la transformada [1], en este caso la gran parte de los coeficientes restantes contendrán la pequeña cantidad de energía excedente.

Si  $\mathbf{m}_u$  y  $\mathbf{R}_u$  representan al valor promedio o valor esperado y covarianza del vector  $\mathbf{u}$  respectivamente, sus correspondientes magnitudes del vector  $\mathbf{v}$ , serán:

$$\mathbf{m}_v = E[\mathbf{v}] = E[\mathbf{A} \mathbf{u}] = \mathbf{A} E[\mathbf{u}] = \mathbf{A} \mathbf{m}_u \quad (2.1.2.3)$$

$$\mathbf{R}_v = \mathbf{A} \mathbf{R}_u \mathbf{A}^{*T} \quad (2.1.2.4)$$

La sección B-4 del Apéndice B muestra el procedimiento para la obtención de la ecuación (2.1.2.4).

La varianza de los coeficientes de la transformada, está dada por los elementos de la diagonal principal de la matriz covarianza  $\mathbf{R}_v$ , esto es,

$$\begin{aligned} \sigma_v^2(k) &= E\left[|v(k) - m_v(k)|^2\right] \\ &= E[(v(k) - m_v(k))(v(k) - m_v(k))^*] \\ &= r_v(k, k) \\ &= [\mathbf{R}_v]_{k,k} \end{aligned} \quad (2.1.2.5)$$

Adicionalmente tomando la matriz  $\mathbf{A}$  como unitaria, se tiene:

$$\|\mathbf{m}_v\|^2 = \|\mathbf{m}_u\|^2 \quad (2.1.2.6)$$

para la demostración de la ecuación (2.1.2.6) se seguirá un procedimiento similar al usado para el caso de la ecuación (2.1.2.1), además usando la ecuación (1.4.1.10b) en (2.1.2.4) se puede afirmar lo siguiente:

$$\sum_{k=0}^{N-1} \sigma_v^2(k) = \text{Tr}[\mathbf{R}_v] = \text{Tr}[\mathbf{A} \mathbf{R}_u \mathbf{A}^{*T}] = \text{Tr}[\mathbf{A}^{*T} \mathbf{A} \mathbf{R}_u] = \text{Tr}[\mathbf{R}_u] = \sum_{n=0}^{N-1} \sigma_u^2(n) \quad (2.1.2.7)$$

Para un campo aleatorio 2-D  $u(m,n)$ , donde su promedio y covarianza son respectivamente  $m_u(m,n)$  y  $r_u(m,n;m',n')$ , los correspondientes coeficientes  $v(k,l)$  cumplen lo siguiente:

$$m_v(k,l) = \sum_m \sum_n a(k,m) m_u(m,n) a(l,n) \quad (2.1.2.8)$$

$$\sigma_v^2(k,l) = \sum_m \sum_n \sum_{m'} \sum_{n'} a(k,m) a(l,n) r_u(m,n;m',n') a^*(k,m') a^*(l,n') \quad (2.1.2.9)$$

La ecuación (2.1.2.9) se obtiene partiendo de la definición de varianza como se presenta en el apéndice B.

## 2.2 DESCRIPCIÓN DE LAS TRANSFORMADAS DISCRETAS

### 2.2.1 LA TRANSFORMADA DISCRETA DE FOURIER (TDF)

La Matriz Transformada Discreta de Fourier, denotada con  $\mathbf{F}$ , utilizada en aplicaciones de procesamiento digital de señales e imágenes está dada por [1],

$$[\mathbf{F}]_{k,m} = \frac{1}{\sqrt{N}} W_N^{km} \quad , \quad k, m = 0, \dots, N-1 \quad (2.2.1.1)$$

dicha matriz es una matriz unitaria de tamaño  $N \times N$ , en la cual,

$$W_N = \exp\left(-\frac{j2\pi}{N}\right) \quad (2.2.1.2)$$

El par de transformadas de Fourier 2-D se conseguirá haciendo  $\mathbf{A} = \mathbf{F}$  de (2.2.1.1) en (2.1.0.8) y (2.1.0.9) obteniéndose por tanto:

$$\mathbf{V} = \mathbf{F}\mathbf{U} \quad , \quad \mathbf{U} = \mathbf{F}^* \mathbf{V} \quad (2.2.1.3)$$

de manera similar, operando sobre las ecuaciones unidimensionales dadas por (2.1.0.1) y (2.1.0.2) se obtiene,

$$\mathbf{v} = \mathbf{F}\mathbf{u} \quad , \quad \mathbf{u} = \mathbf{F}^* \mathbf{v}$$

#### Propiedades de la Transformada Discreta de Fourier

##### 1.- La matriz unitaria $\mathbf{F}$ es simétrica

Por definición la matriz  $\mathbf{F}$  es simétrica, esto es,

$$\mathbf{F}^{-1} = \mathbf{F}^{*'} = \mathbf{F}^* \quad (2.2.1.4)$$

## 2.- Las extensiones son periódicas

Las extensiones de la transformada discreta de Fourier (TDF) de una secuencia  $u(m,n)$  de tamaño  $N \times N$  son periódicas<sup>1</sup> con período  $N$ , esto es [1],

$$\begin{aligned} v(k+N, l+N) &= v(k, l) \quad \forall (k, l) \\ u(m+N, n+N) &= u(m, n) \quad \forall (m, n) \end{aligned} \quad (2.2.1.5)$$

A pesar que la TDF es aplicada a secuencias no periódicas, en el cálculo de la misma se asume implícitamente que la secuencia se torna periódica.

## 3.- Simetría conjugada

Esta propiedad es característica de la Transformada discreta de Fourier para imágenes reales y establece que [1]:

$$v(N/2 \pm k, N/2 \pm l) = v^*(N/2 \mp k, N/2 \mp l) \quad , 0 \leq k, l \leq N/2 - 1 \quad (2.2.1.6)$$

ó también,

$$v^*(N - k, N - l) = v(k, l) \quad , 0 \leq k, l \leq N - 1 \quad (2.2.1.7)$$

4.- Teorema de Convolución.- establece que, la transformada discreta de Fourier de la convolución de dos secuencias es igual al producto de las transformadas de Fourier de cada una de las secuencias, esto es<sup>2</sup>,

$$y(m, n) = \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} h(m-m', n-n') x(m', n') \quad , 0 \leq m, n \leq N-1 \quad (2.2.1.8)$$

$$\Rightarrow \text{TDF}\{y(m, n)\}_N = \text{TDF}\{h(m, n)\}_N \text{TDF}\{x(m, n)\}_N$$

<sup>1</sup> En general, una secuencia  $u(m,n)$  se dice periódica con período  $M \times N$  si  $u(m,n)$  cumple con la siguiente condición según [3]:

$$u(m, n) = u(m+M, n) = u(m, n+N) = u(m+M, n+N) \quad \forall (m, n)$$

<sup>2</sup> La notación  $\text{TDF}\{y(m, n)\}_N$  se interpreta como la Transformada Discreta de Fourier del arreglo  $y(m, n)$  de tamaño  $N \times N$ . Además la función  $h(m, n)$  es la respuesta del sistema al impulso  $\delta(m, n)$ .

### 2.2.2 LA TRANSFORMADA DISCRETA COSENO (TDC)

La matriz Transformada Coseno llamada también Transformada Discreta Coseno (TDC) y denotada como  $C = \{c(k,m)\}$  se define según [1] de la siguiente manera<sup>1</sup>,

$$c(k,m) = \sqrt{\frac{2}{N}} e(k) \cos\left(\frac{(m+1/2)k\pi}{N}\right), \quad e(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } k = 0 \\ 1 & \text{si } k \neq 0 \end{cases} \quad (2.2.2.1)$$

El tamaño de la matriz Transformada discreta Coseno es  $N \times N$ , consecuentemente los índices  $k$  y  $m$  tomarán valores desde 0 hasta  $N-1$ , de manera análoga a como ocurre con la TDF.

El par de transformadas discretas coseno bidimensionales, se obtiene haciendo  $A = A^* = C$  en las ecuaciones (2.1.0.8) y (2.1.0.9)<sup>2</sup>, esto es,

$$V = CUC^T, \quad U = C^TVC \quad (2.2.2.2)$$

#### Propiedades de la Transformada Coseno.

1.- La Transformada Discreta Coseno es real y ortogonal.

Si la matriz Coseno es real y ortogonal, cumple también con ser unitaria, es decir,

$$C = C^* \quad , \quad C^{-1} = C^T \Rightarrow C^{*T} = C^{-1} \quad (2.2.2.3)$$

2.- La Transformada Discreta Coseno no es la parte real de la Transformada Discreta de Fourier. Esto se puede verificar relacionando las igualdades (2.2.1.1) y (2.2.2.1). Sin embargo existe una expresión matemática que las relaciona, la misma que se desarrolla en el capítulo 3.

<sup>1</sup> Esta matriz relaciona la coordenada en el dominio de la frecuencia espacial  $k$  con su correspondiente en el dominio espacial  $m$ .

<sup>2</sup> Para el caso 1-D, se hará el mismo reemplazo, pero en este caso en las ecuaciones (2.1.0.1) y (2.1.0.2).

3.- La TDC de un vector de tamaño  $N$  puede calcularse mediante aproximadamente  $N \log_2 N$  operaciones<sup>1</sup> mediante el concurso de la Transformada rápida de Fourier (TRF), esta técnica constituye una de los métodos alternativos de cálculo de la transformada coseno mencionados en el capítulo 3. Para esto se procede a separar a la secuencia  $u(n)$  en sus partes: par  $u(2n)$  e impar  $u(2n+1)$ , de la siguiente manera:

$$\left. \begin{aligned} \tilde{u}(n) &= u(2n) \\ \tilde{u}(N-1-n) &= u(2n+1) \end{aligned} \right\}, \quad 0 \leq n \leq N/2-1 \quad (2.2.2.4)$$

trabajando matemáticamente la ecuación (2.1.0.1) después de hacer  $A = C$ , se llegará como se presenta en la sección 6 del apéndice B a la siguiente igualdad:

$$v(k) = \text{Real} \left\{ e^{jk} W_{2N}^{k/2} \text{TDF} \{ \tilde{u}(n) \}_N \right\} \quad (2.2.2.5)$$

4.- Los vectores propios de la matriz tridiagonal simétrica  $Q_C$ , definida según [1] como sigue, representan los vectores bases de la transformada coseno y además son las filas de la matriz  $C$ ,

$$Q_C = \begin{bmatrix} 1-\alpha & -\alpha & & & 0 \\ -\alpha & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 & -\alpha \\ 0 & & & & -\alpha & 1-\alpha \end{bmatrix} \quad (2.2.2.6)$$

donde:  $\alpha = \frac{\rho}{1+\rho^2}$ ,  $\rho =$  coeficiente de correlación

5.- La transformada Coseno se aproxima a la Transformada Karhunen-Loeve para una secuencia estacionaria de Markov<sup>2</sup> de primer orden de longitud  $N$ , donde la matriz

<sup>1</sup> Una operación se considera una multiplicación más una suma.

<sup>2</sup> Una secuencia se llama de Markov de orden  $k$ , si conociendo  $\{u(j), n-k \leq j \leq n-1\}$ , entonces, el pasado  $\{u(j), j \leq n-k-1\}$  y el futuro  $\{u(j), j \geq n\}$  son independientes.



covarianza correspondiente está dada por la matriz  $\mathbf{R}$  con coeficiente de correlación cercano a 1.  $\mathbf{R}^{-1}$  es tridiagonal simétrica, en la cual,  $\beta = \frac{1-\rho^2}{1+\rho^2}$  y  $\alpha = \frac{\rho}{1+\rho^2}$  y cumple la siguiente relación<sup>1</sup>:

$$\beta^2 \mathbf{R}^{-1} = \begin{bmatrix} 1-\rho\alpha & -\alpha & & & 0 \\ -\alpha & 1 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 1 & -\alpha \\ 0 & & & -\alpha & 1-\rho\alpha \end{bmatrix} \quad (2.2.2.7)$$

finalmente las propiedades 4 y 5 se relacionan de la siguiente manera:

$$\text{Si } \rho \rightarrow 1 \Rightarrow \mathbf{Q}_c = \beta^2 \mathbf{R}^{-1} \quad (2.2.2.8)$$

### 2.2.3 LA TRANSFORMADA DISCRETA SENOS (TDS)

La matriz transformada discreta Seno llamada también transformada discreta Seno se define según [1] como:

$$[\Psi]_{k,m} = \psi(k,m) = \sqrt{\frac{2}{N+1}} \operatorname{sen}\left(\frac{(m+1)(k+1)\pi}{N+1}\right), \quad 0 \leq k, m \leq N-1 \quad (2.2.3.1)$$

análogamente a la transformada coseno,  $k$  es una variable en el dominio de la frecuencia espacial y  $m$  su correspondiente en el dominio espacial.

El par de transformadas seno bidimensionales se obtiene reemplazando la ecuación (2.2.3.1) en las ecuaciones (2.1.0.8) y (2.1.0.9), obteniéndose de esta manera:

<sup>1</sup> La matriz covarianza está dada por [1]:

$$\mathbf{R} = \{r(m,n)\} = \{r(m-n)\} = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ \rho & \ddots & \ddots & \ddots & \vdots \\ \rho^2 & \ddots & \ddots & \ddots & \rho^2 \\ \vdots & \ddots & \ddots & \ddots & \rho \\ \rho & \dots & \rho^2 & \rho & 1 \end{bmatrix}$$

una secuencia puede ser llamada estacionaria *ssi* si sus matrices covarianza y autocorrelación son Toeplitz.

$$\mathbf{V} = \Psi \mathbf{U} \Psi^T, \mathbf{U} = \Psi^T \mathbf{V} \Psi \quad (2.2.3.2)$$

Para el caso unidimensional, se hará la misma sustitución pero esta vez en las ecuaciones (2.1.0.1) y (2.1.0.2).

### Propiedades de la Transformada Discreta Seno

1.- La Transformada Seno es real, simétrica y ortogonal, es decir [1],

$$\Psi^* = \Psi = \Psi^{-1} = \Psi^T \quad (2.2.3.3)$$

Esto se comprobará analizando la expresión (2.2.3.1).

2.- La Transformada Discreta Seno no constituye la parte imaginaria de la transformada Discreta de Fourier

Esta afirmación se puede comprobar comparando las definiciones de la TDS con la TDF.

3.- Los vectores bases de la TDS son los vectores propios de la matriz  $\mathbf{Q}$ , la misma que cumple con ser Toeplitz tridiagonal y simétrica, en la cual  $\alpha = \frac{\rho}{1 + \rho^2}$ , siendo  $\rho$  el coeficiente de correlación, entonces [1]:

$$\mathbf{Q} = \begin{bmatrix} 1 & -\alpha & & 0 \\ -\alpha & 1 & & \\ & & \ddots & \\ & & & 1 & -\alpha \\ 0 & & & -\alpha & 1 \end{bmatrix} \quad (2.2.3.4)$$

4.- La transformada Seno se aproxima a la Transformada KL para secuencias de Markov de primer orden, con matriz covarianza  $\mathbf{R}$  (referirse a la propiedad cinco de la TDC), y coeficiente de correlación ubicado en el intervalo (-0.5, 0.5) [1].

## 2.2.4 LA TRANSFORMADA DISCRETA HADAMARD

En este caso los elementos que conforman la matriz que describe esta transformada toman valores binarios, solamente  $\pm 1$ , adaptándose perfectamente al procesamiento digital.

La matriz Hadamard notada  $\mathbf{H}_n$ , es una matriz de tamaño  $N \times N$  ( $N = 2^n$ ) generada en forma recurrente mediante el uso del producto Kronecker, como se muestra a continuación y según lo manifiestan las referencias [1] y [12],

$$\mathbf{H}_n = \mathbf{H}_{n-1} \otimes \mathbf{H}_1 = \mathbf{H}_1 \otimes \mathbf{H}_{n-1}, \quad \mathbf{H}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (2.2.4.1)$$

por tanto,

$$\mathbf{H}_n = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{H}_{n-1} & \mathbf{H}_{n-1} \\ \mathbf{H}_{n-1} & -\mathbf{H}_{n-1} \end{bmatrix} \quad (2.2.4.2)$$

Haciendo  $\mathbf{A} = \mathbf{A}^* = \mathbf{A}^T = \mathbf{H}$  en las ecuaciones (2.1.0.8) y (2.1.0.9) se obtiene el par de transformadas Hadamard para una imagen de tamaño  $N \times N$ .

El número de transiciones (cambios de signos entre elementos de una misma fila) en la transformada Hadamard es llamada su secuencia. En la matriz generada por (2.2.4.2), los vectores fila no están ordenados secuencialmente y cuando si lo están se llama *Hadamard ordenada*.

### Propiedades de la Transformada Hadamard

1.- La Transformada Hadamard  $\mathbf{H}$ , es real, ortogonal y simétrica, es decir,

$$\mathbf{H} = \mathbf{H}^* = \mathbf{H}^T = \mathbf{H}^{-1} \quad (2.2.4.3)$$

2.- La matriz transformada Hadamard  $\mathbf{H}$  de orden  $N$ , determinada por (2.2.4.1) y (2.2.4.2), está constituida únicamente por valores  $\pm 1$ , puede escribirse según referencia [1] como el producto de  $n$  matrices menos densas, esto es,

$$\mathbf{H} = \mathbf{H}_n = \tilde{\mathbf{H}}^n, \quad n = \log_2 N \quad (2.2.4.4)$$

donde:

$$\tilde{\mathbf{H}}_{N \times N} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & & & & \dots & 0 & 0 \\ 0 & 0 & & & 0 & 0 & \dots & 1 & 1 \\ 1 & -1 & 0 & 0 & & & \dots & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & & & & \dots & 0 & 0 \\ 0 & 0 & & & 0 & 0 & \dots & 1 & -1 \end{bmatrix}$$

Para el caso unidimensional  $\mathbf{v} = \mathbf{H}\mathbf{u}$ , el número de sumas o restas requeridas para su cálculo se reduce desde  $N^2$  hasta  $N \log_2 N$ , entonces<sup>1</sup>,

$$\mathbf{v} = \underbrace{\tilde{\mathbf{H}}\tilde{\mathbf{H}}\dots\tilde{\mathbf{H}}}_{n\text{-términos}} \mathbf{u} \quad (2.2.4.5)$$

Para el caso bidimensional:

$$\mathbf{V} = \underbrace{\tilde{\mathbf{H}}\tilde{\mathbf{H}}\dots\tilde{\mathbf{H}}}_{n\text{-términos}} \mathbf{U} \underbrace{\tilde{\mathbf{H}}\tilde{\mathbf{H}}\dots\tilde{\mathbf{H}}}_{n\text{-términos}} \quad (2.2.4.6)$$

el número de operaciones se reduce desde  $2N^3$  hasta  $2N^2 \log_2 N$ .

3.- La Transformada Hadamard tiene una buena compresión energética para imágenes altamente correlacionadas<sup>2</sup>.

<sup>1</sup> Pues solamente  $N$  sumas o restas se requieren por cada una de las  $n$ -veces que se opera sobre el vector  $\mathbf{u}$ , dando un total de  $Nn = N \log_2 N$  operaciones.

<sup>2</sup> El término altamente correlacionado se usa para indicar un coeficiente de correlación cercano a uno.

Según se afirma en la referencia [1], la fracción de energía empaquetada en los  $N/2$  primeros coeficientes ordenados secuencialmente es del 97.5% de la energía total, bajo las condiciones: secuencia de entrada  $\{u(n), 0 \leq n \leq N-1\}$  y  $\rho = 0,95$ .

### 2.2.5 LA TRANSFORMADA DISCRETA HAAR

Las funciones Haar denotadas como  $h_k(x)$  están definidas sobre un intervalo cerrado continuo  $x \in [0, 1]$ , para  $k = 0, \dots, N-1$ , donde  $N = 2^n$ . El número  $k$  puede ser descompuesto de manera única según la referencia [1] en:

$$k = 2^p + q - 1 \quad (2.2.5.1)$$

en donde:

$$\begin{aligned} q &= 0, 1 \quad \text{si } p = 0 \\ 1 \leq q \leq 2^p & \quad \text{si } 1 \leq p \leq n-1 \end{aligned} \quad (2.2.5.2)$$

Se pueden redefinir las funciones Haar representando  $k$  por el par  $(p, q)$  como:

$$h_0(x) = h_{0,0}(x) = \frac{1}{\sqrt{N}} \quad , \quad 0 \leq x \leq 1 \quad (2.2.5.3a)$$

$$h_k(x) = h_{p,q}(x) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & \text{si } x \in \left[ \frac{q-1}{2^p}, \frac{q-1/2}{2^p} \right[ \\ -2^{p/2} & \text{si } x \in \left[ \frac{q-1/2}{2^p}, \frac{q}{2^p} \right[ \\ 0 & \text{en otro caso en } [0, 1] \end{cases} \quad (2.2.5.3b)$$

#### Propiedades de la Transformada Discreta Haar

1.- La Transformada Haar es real y ortogonal [1],

$$\begin{aligned} \mathbf{H}_r &= \mathbf{H}_r^* \\ \mathbf{H}_r^{-1} &= \mathbf{H}_r^T \end{aligned} \quad (2.2.5.4)$$

2.- La Transformada Haar es una transformada muy rápida [1]

Para un vector  $N \times 1$  el número de operaciones requeridas para implementarla es del orden de  $N$ .

3.- Para imágenes tiene una pobre compresión energética [1].

4.- Los vectores bases de la matriz Haar están ordenados secuencialmente [1].

### 2.2.6 LA TRANSFORMADA SLANT

La Transformada Slant  $N \times N$  notada  $S_n$  ( $2^n = N$ ), en forma recursiva se define en [1] de la siguiente manera:

$$S_n = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & & 1 & 0 & & 0 \\ a_n & b_n & & -a_n & b_n & & 0 \\ & & \mathbf{I}_{\frac{N}{2}-2} & & & & \\ 0 & 1 & & 0 & -1 & & 0 \\ -b_n & a_n & & b_n & a_n & & \\ & & & & & & \\ 0 & & \mathbf{I}_{\frac{N}{2}-2} & & & & -\mathbf{I}_{\frac{N}{2}-2} \end{bmatrix} \begin{bmatrix} S_{n-1} & 0 \\ 0 & S_{n-1} \end{bmatrix} \quad (2.2.6.1)$$

en dicha matriz:

$$S_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{I}_M \text{ es la matriz identidad } M \times M$$

$$a_n = \frac{2a_{n-1}}{\sqrt{1+4a_{n-1}^2}}, \quad b_n = \frac{a_n}{2a_{n-1}} \quad \text{y} \quad a_1 = 1$$

#### Propiedades de la Transformada Slant

1.- La Matriz Transformada Slant es real y ortogonal, es decir [1],

$$S = S^* \quad , \quad S^{-1} = S^T \quad (2.2.6.2)$$

- 2.- La Transformada Slant es una Transformada rápida, pues, puede ser implementada en  $N \log_2 N$  operaciones aproximadamente cuando opera sobre un vector  $N \times 1$  [1].
- 3.- Tiene una excelente compresión de energía para imágenes [1].
- 4.- Los vectores bases de la matriz transformada Slant no están ordenados secuencialmente para  $n \geq 3$ . Si se quiere que lo estén habrá que redefinirla [1].

### 2.2.7 LA TRANSFORMADA KARHUNEN-LOEVE (KL)

Fue introducida originalmente por Karhunen y Loeve como una expansión en series de procesos aleatorios continuos. Se la conoce también como transformada Hotelling o Método de las Componentes Principales, pues así se denominó a los primeros estudios sobre expansión en series de secuencias aleatorias.

Los vectores bases de la Transformada KL para un vector real  $\mathbf{u}$  de tamaño  $N \times 1$  están dados por los vectores propios ortonormalizados de la matriz de autocorrelación  $\mathbf{R}$  (Si  $\mathbf{u}$  tiene un valor esperado nulo, solo en este caso la covarianza coincide con la autocorrelación, hecho asumido por simplicidad en los desarrollos realizados en la referencia [1]), entonces:

$$\mathbf{R}\phi_k = \lambda_k \phi_k \quad , \quad 0 \leq k \leq N - 1 \quad (2.2.7.1)$$

en donde  $\phi_k$  es la  $k$ -ésima columna<sup>1</sup> de  $\Phi$ . Las matrices  $\Phi$  y  $\mathbf{R}$  son tales que cumplen con la ecuación (1.4.1.15), esto es:

$$\Phi^{*T} \mathbf{R} \Phi = \Lambda = \text{Diag}(\lambda_k) \quad (2.2.7.2)$$

<sup>1</sup>  $\phi_k$  es el  $k$ -ésimo vector propio correspondiente al  $k$ -ésimo valor propio  $\lambda_k$  de la matriz  $\mathbf{R}$ .

El par de Transformadas KL haciendo  $\mathbf{A} = \Phi^{*T}$  en (2.1.0.1) y (2.1.0.2) será:

$$\begin{aligned} \mathbf{v} &= \Phi^{*T} \mathbf{u} \\ \mathbf{u} &= \Phi \mathbf{v} = \sum_{k=0}^{N-1} v(k) \phi_k \end{aligned} \quad (2.2.7.3)$$

La transformada Coseno se aproxima a la KL si se cumple que [1]:

$$\mathbf{R} \mathbf{Q}_C = \mathbf{Q}_C \mathbf{R}$$

cosa similar ocurre con la transformada Seno y la matriz  $\mathbf{Q}$  dada por la ecuación (2.2.3.4).

Para una imagen cuadrada  $u(m,n)$  de tamaño  $N \times N$ , el par de transformadas KL lo conforman las siguientes igualdades:

$$\begin{aligned} \mathbf{V} &= \Phi_1^{*T} \mathbf{U} \Phi_2^* \\ \mathbf{U} &= \Phi_1 \mathbf{V} \Phi_2^T \end{aligned} \quad (2.2.7.4)$$

las matrices  $\Phi_1$  y  $\Phi_2$  existen siempre que cumplan con la expresión:

$$\Psi = \Phi_1 \otimes \Phi_2 \quad (2.2.7.5)$$

$$\Psi = \{\psi_i\} \quad , \quad \mathbf{R}_r \psi_i = \lambda_i \psi_i \quad , \quad i = 0, \dots, N^2 - 1 \quad (2.2.7.6)$$

donde,

$$\mathbf{R}_r = E[\mathbf{u}_r \mathbf{u}_r^T] = \mathbf{R}_1 \otimes \mathbf{R}_2 \quad (2.2.7.7)$$

en esta expresión  $\mathbf{u}_r$  es un vector de tamaño  $N^2 \times 1$  resultado de un ordenamiento por filas de la matriz  $\mathbf{U}$ .

La transformada KL de  $\mathbf{u}_r$  es [1]:

$$\mathbf{v}_r = \Psi^{*T} \mathbf{u}_r \quad (2.2.7.8)$$

Adicionalmente debe cumplirse:

$$\Lambda_i = \Phi_i \mathbf{R}_i \Phi_i^{*T} \quad i = 1, 2 \quad (2.2.7.9)$$



## Propiedades de la transformada KL

### 1.- Decorrelación

Los coeficientes transformados  $\{v(k), k= 0, \dots, N-1\}$  no están correlacionados y tienen un valor esperado nulo, es decir,

$$\begin{aligned} E[v(k)] &= 0 \\ E[v(k)v^*(l)] &= \lambda_k \delta(k-l) \end{aligned} \quad (2.2.7.10)$$

esta última igualdad se puede demostrar fácilmente como sigue:

$$E[\mathbf{v}\mathbf{v}^{*T}] = E[\Phi^{*T}\mathbf{u}\mathbf{u}^{*T}\Phi] = \Phi^{*T} E[\mathbf{u}\mathbf{u}^{*T}]\Phi = \Phi^{*T}\mathbf{R}\Phi = \Lambda$$

Cuando los elementos del vector de entrada están altamente correlacionados, los elementos de la transformada no están correlacionados. Esto significa que los elementos de la diagonal principal de la matriz covarianza  $\mathbf{R}_v$  son grandes comparados con los elementos restantes de la matriz. Esto hace de la transformada KL óptima, empaquetando el máximo de la energía en un número dado de coeficientes no correlacionados entre sí.

### 2.- Control de Error Cuadrático Medio

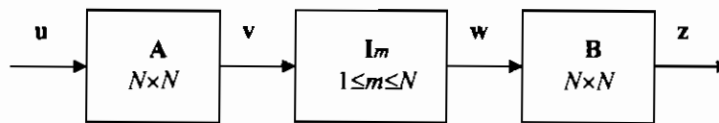


Figura 2.2.7.1 Control de error en TKL.

Refiriéndose a las operaciones indicadas mediante la figura 2.2.7.1,  $\mathbf{A}$  y  $\mathbf{B}$  son matrices de tamaño  $N \times N$ ,  $\mathbf{I}_m$  es una matriz que tiene 1s a lo largo de los primeros  $m$ -elementos de la diagonal principal y ceros en las posiciones restantes,  $\mathbf{u}$  se transforma en  $\mathbf{v}$  con la aplicación de  $\mathbf{A}$  y  $\mathbf{w}$  en  $\mathbf{z}$  con la aplicación de  $\mathbf{B}$ , mientras  $\mathbf{w}$  cumplirá con tener los siguientes elementos:

$$w(k) = \begin{cases} v(k), & 0 \leq k \leq m-1 \\ 0, & k \geq m \end{cases} \quad (2.2.7.11)$$

Se define la cantidad *control de error* notado como  $J_m$  al promedio del error cuadrático medio entre las secuencias  $u(n)$  y  $z(n)$  de [1] como:

$$J_m = \frac{1}{N} \text{Tr} [E\{(\mathbf{u} - \mathbf{z})(\mathbf{u} - \mathbf{z})^{*T}\}] \quad (2.2.7.12)$$

El teorema 5.1 desarrollado en [1], establece que el error  $J_m$  se minimiza cuando se hace cumplir las condiciones siguientes,

$$\begin{aligned} \mathbf{A} &= \Phi^{*T} \\ \mathbf{B} &= \Phi \\ \mathbf{AB} &= \mathbf{I} \end{aligned} \quad (2.2.7.13)$$

las ideas precedentes a la conclusión de este teorema son de utilidad, por tanto se expondrán brevemente, estos criterios fueron tomados de la demostración descrita en la referencia [1].

$$J_m = \frac{1}{N} \text{Tr} [\mathbf{R} - \mathbf{I}_m \mathbf{A} \mathbf{R} \mathbf{A}^{*T}] \quad (2.2.7.14)$$

$J_m$  es minimizado si la cantidad  $\tilde{J}_m = \frac{1}{N} \text{Tr} [\mathbf{I}_m \mathbf{A} \mathbf{R} \mathbf{A}^{*T}]$  es maximizada, para lo cual se

requiere  $\tilde{J}_m = \sum_{k=0}^{m-1} \lambda_k$ , en donde el conjunto  $\{\lambda_k, 0 \leq k \leq m-1\}$  corresponden a los  $m$

valores propios más grandes de la matriz covarianza  $\mathbf{R}$  de  $\mathbf{u}$ . Puesto que  $\tilde{J}_m$  debe ser maximizada para todos los  $m$  valores propios, es necesario la siguiente relación:

$$\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N-1}$$

entonces las filas de  $\mathbf{A}$  son los vectores transpuestos conjugados, de los vectores propios correspondientes a los valores propios de la matriz  $\mathbf{R}$  y, consecuentemente  $\mathbf{A}$  es la transformada KL.

3.- Distribución de Varianza

De entre todas, la Transformada KL ( $\Phi^{*T}$ ) empaqueta el máximo de energía promedio en  $m \leq N$  coeficientes unidimensionales en el dominio de la frecuencia [1].

Por definición:  $\sigma_k^2$ , la varianza de la transformada unitaria  $A$ , está dada por:

$$\sigma_k^2 = E[|v(k)|^2], \quad \sigma_0^2 \geq \sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_{N-1}^2, \quad S_m(A) = \sum_{k=0}^{m-1} \sigma_k^2 \quad (2.2.7.15)$$

Se trata de demostrar:

$$S_m(\Phi^{*T}) \geq S_m(A) \quad (2.2.2.16)$$

*Demostración:*

$$\begin{aligned} S_m(A) &= \sum_{k=0}^{m-1} [A R A^{*T}]_{k,k}, \text{ por definición y (2.1.2.4)} \\ &= Tr[\mathbf{I}_m A R A^{*T}] \text{ , por propiedad anterior} \\ &= \tilde{J}_m \end{aligned}$$

de la propiedad anterior  $\tilde{J}_m$  es maximizado cuando  $A$  es la TKL, es decir  $A = \Phi^{*T}$  en consecuencia  $\sigma_k^2 = \lambda_k$ , finalmente para cualquier transformada unitaria  $A$  se tiene:

$$\sum_{k=0}^{m-1} \lambda_k \geq \sum_{k=0}^{m-1} \sigma_k^2 \quad (2.2.7.17)$$

4.- Función Tasa de Distorsión

Considérese para este caso la figura 2.2.7.2,

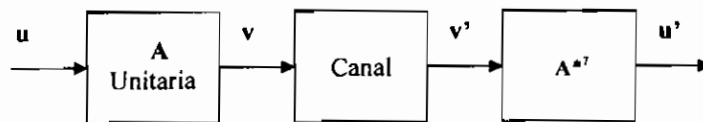


Figura 2.2.7.2 Transmisión de coeficientes de la transformada unitaria  $A$ .

en dicha figura, la transformada unitaria del vector  $\mathbf{u}$ , es decir  $\mathbf{v}$ , es aplicado a un canal de comunicaciones, luego del cual se obtiene  $\mathbf{v}'$  para finalmente aplicando la transformada inversa obtener  $\mathbf{u}'$ . Se asume que  $\mathbf{u}$ ,  $\mathbf{u}'$ ,  $\mathbf{v}$  y  $\mathbf{v}'$  son Gaussianos.

La distorsión promedio de  $\mathbf{u}$  es según [1] y [11],

$$D = \frac{1}{N} E[(\mathbf{u} - \mathbf{u}')^*{}^T (\mathbf{u} - \mathbf{u}')] \quad (2.2.7.18)$$

aplicando a esta expresión la ecuación (2.1.0.1) se obtiene finalmente,

$$D = \frac{1}{N} E[\delta\mathbf{v}^*{}^T \delta\mathbf{v}] \quad (2.2.7.19)$$

en esta expresión  $\delta\mathbf{v} = \mathbf{v} - \mathbf{v}'$  representa el error ocurrido en la reproducción de  $\mathbf{v}$ .

De (2.2.7.19) se ve que  $D$  es invariante a los cambios de la transformada unitaria, la función tasa de distorsión está dada por:

$$R = \frac{1}{N} \sum_{k=0}^{N-1} \max \left[ 0, \frac{1}{2} \log_2 \left( \frac{\sigma_k^2}{\theta} \right) \right] \quad (2.2.7.20)$$

$$D = \frac{1}{N} \sum_{k=0}^{N-1} \min[\theta, \sigma_k^2] \quad (2.2.7.21)$$

en las cuales:

$$\sigma_k^2 = E[|v(k)|^2] = [\mathbf{A} \mathbf{R} \mathbf{A}^*{}^T]_{k,k} \quad (2.2.7.22)$$

como la matriz  $\mathbf{R}$  depende de los datos, mientras  $\mathbf{A}$  depende de la transformada, entonces

$R$  es una función de  $\mathbf{A}$ , esto es:

$$R = R(\mathbf{A}) \quad (2.2.7.23)$$

Finalmente para cada valor de  $D$ ,

$$R(\Phi^*{}^T) \leq R(\mathbf{A}) \quad (2.2.7.24)$$

La igualdad (2.2.7.24) se interpreta según la referencia [1] como: “*para un nivel fijo de distorsión, el número de bits requerido para transferir coeficientes KL será menor que el requerido para transmitir coeficientes de cualquier otra transformada*”.

**RESUMEN.-** Una propiedad deseable de una transformación es la reducción de la correlación entre los coeficientes transformados, dicha propiedad es conocida como: “propiedad de reducción de correlación” [3]. Dicho coeficiente de correlación podrá ser reducido mediante un escogitamiento adecuado de las funciones bases (transformación adecuada).

La *compresión energética* es una propiedad deseable y está relacionada con la anterior. Al concentrar gran parte de la energía dentro de pocos coeficientes es posible descartar muchos de éstos sin alterar seriamente a la imagen.

El cumplimiento de la primera propiedad ayuda a la segunda pero en si no es suficiente, por ejemplo dos muestras de una señal aleatoria de ruido blanco no están correlacionadas pero la señal no tiene compresión energética [3].

De todas la transformadas lineales la transformada KL es la mejor en lo referente a compresión energética, con sus coeficientes KL no correlacionados. Teóricamente es ideal, pero presenta una seria dificultad práctica, diferentes imágenes tendrán diferentes función covarianza consecuentemente diferentes funciones bases<sup>1</sup>.

Adicionalmente la TKL se basa en la suposición que la secuencia  $u(m,n)$  es un proceso aleatorio (lo cual para imágenes puede no ser válido) y su función covarianza es conocida

---

<sup>1</sup> Calcular la transformada KL de una secuencia  $M \times N$  sin algoritmos eficientes requiere aproximadamente  $M^2 \times N^2$  operaciones[3].

(en la práctica debe ser calculada), si el cálculo no es correcto sus propiedades teóricamente óptimas no se mantienen.

Una transformada que mantiene el conjunto de funciones bases, algoritmos de computación eficientes y buena compresión energética para imágenes es la TDF, pero cuando la propiedad de compresión energética y los requerimientos computacionales son considerados simultáneamente la TDC será por lo general considerada como la mejor. Las transformadas como Haar y Hadamard requieren para su implementación menor número de cálculos que la coseno pero su compresión energética no se compara siquiera.

# CAPÍTULO

---

# 3

## LA TRANSFORMADA DISCRETA COSENO

### 3.0 INTRODUCCIÓN

Este capítulo está destinado a la descripción de la Transformada Discreta Coseno, se inicia indicando algunas de sus aplicaciones más importantes, se presenta la definición de la Transformada Discreta Coseno así como también algunas de sus propiedades más importantes.

Seguidamente se desarrolla un algoritmo que permite su cálculo de una manera muy rápida comparado con algunos métodos existentes. Este algoritmo se presenta básicamente en dos partes, una destinada al cálculo de la transformada en sentido directo y otra para el cálculo en sentido inverso, las dos partes se fundamentan en las mismas bases. Finalmente en el último literal de este capítulo se presentan varias imágenes monocromáticas acompañadas de su respectivo espectro de la transformada coseno.

### 3.1 LA TRANSFORMADA DISCRETA COSENO

Una *transformada imagen* puede ser vista como la descomposición de la imagen dato en un espectro bidimensional, en dicho dominio cada una de las componentes espectrales corresponde a una cantidad de energía de la función espectral (ecuación 1.5.2.20). El primer coeficiente (siempre positivo) corresponde a la brillantez<sup>1</sup> de la imagen (componente DC), los restantes coeficientes (componentes AC) dan una indicación de los detalles de la imagen [36].

La frecuencia espacial indica la tendencia de iluminación, una diferencia pronunciada entre pixeles representa una alta frecuencia espacial e indicará un límite o contorno definido en la imagen; de otro modo una baja frecuencia espacial representará un contorno borroso o poco definido. Una imagen, entonces puede ser clasificada en dos categorías [39]:

- imágenes conformadas por áreas con intensidad de pixeles aproximadamente constante, en este caso los coeficientes Coseno serán mejor modelados por la distribución Gamma<sup>2</sup>.

<sup>1</sup> Brillantez.- Atributo de la percepción visual por lo que una cierta área de una imagen da la apariencia que emite mayor o menor intensidad luminosa [23].

<sup>2</sup> La función densidad de probabilidad Gamma, está dada por:

$$f_X(x) = \frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\int_0^{\infty} x^{\alpha-1} e^{-x} dx} \quad x, \alpha \text{ y } \lambda \text{ positivos}$$

además:  $E[X] = \alpha/\lambda$  y  $VAR[X] = \alpha/\lambda^2$



- imágenes con muchos bordes y textura áspera, en este caso en cambio los coeficientes Coseno serán mejor modelados por la distribución Laplaciana<sup>1</sup>.

De entre todas, la Transformada Discreta de Fourier (TDF) es la más usada en procesamiento digital, por la existencia de algoritmos eficientes de cálculo, buena compresión energética para imágenes típicas y en los inicios del procesamiento digital fue la primera en ser considerada. Sin embargo la Transformada Discreta Coseno (TDC) presenta de entre todas el mejor comportamiento<sup>2</sup> [32], además su cálculo (real) es más simple y rápido que el cálculo (complejo) de la TDF [29].

### 3.1.1 MEDIDAS DE COMPORTAMIENTO DE LAS TRANSFORMADAS UNITARIAS

El comportamiento de las transformadas discretas es comparado con la Transformada Karhunen-Loeve, la misma que es conocida por ser óptima con respecto a las siguientes medidas de comportamiento [24]:

#### a) Distribución de varianza

Este parámetro de comparación se define en la propiedad número tres de la transformada KL en la sección (2.2.7). Observando la gráfica 3.1.1.1 se puede comprobar que para

<sup>1</sup> La función densidad de probabilidad Laplaciana, está dada por:

$$f_X(x) = \frac{\alpha}{2} e^{-\alpha|x|}, \quad x \text{ es real y } \alpha \text{ positivo}$$

además:  $E[X] = 0$  y  $VAR[X] = \frac{2}{\alpha^2}$ . Fórmulas tomadas de, Probability and Random Processes for Electrical Engineering de Alberto Leon-Garcia, publicado en mayo de 1989 por Addison-Wesley Publishing Company.

<sup>2</sup> Se ha observado que la transformada coseno de imágenes típicas tienen gran parte de su energía concentrada en una pequeña región cerca del origen en el dominio de la frecuencia. Una razón para que la energía se concentre cerca del origen es que las imágenes típicamente tienen grandes regiones donde la intensidad cambia lentamente [3]

coeficientes de correlación cercanos a 1, por ejemplo para  $\rho=0,95$  la TDC se comporta como la TKL.

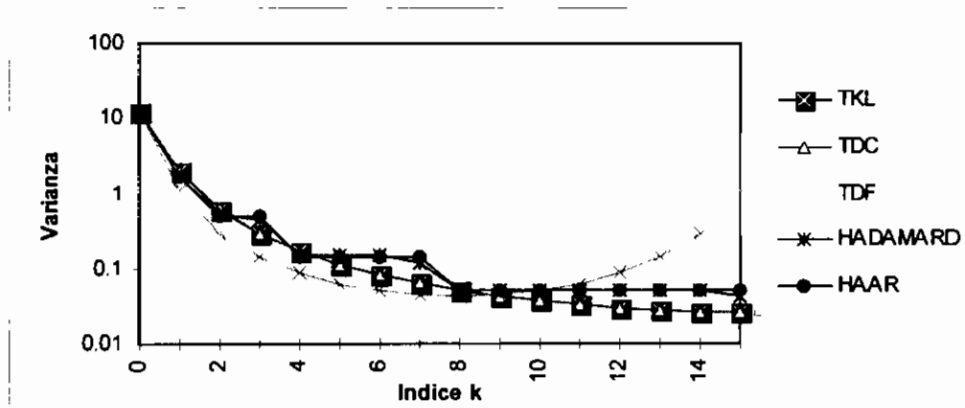


Figura 3.1.1.1 Varianza en el dominio de la transformada  $16 \times 16$ ,  $\rho=0,95$  [24].

**b) Criterio del error cuadrático medio**

En la figura 3.1.1.2  $Z$  es un vector  $M \times 1$  y resulta de la suma del vector  $X$  y del vector  $N$ .  $X$  se considera un proceso aleatorio cuya matriz covarianza se define en la propiedad cinco de la sección 2.2.2 .

El ruido aditivo  $N$  se considera blanco (valor promedio nulo y no correlacionado con los datos).

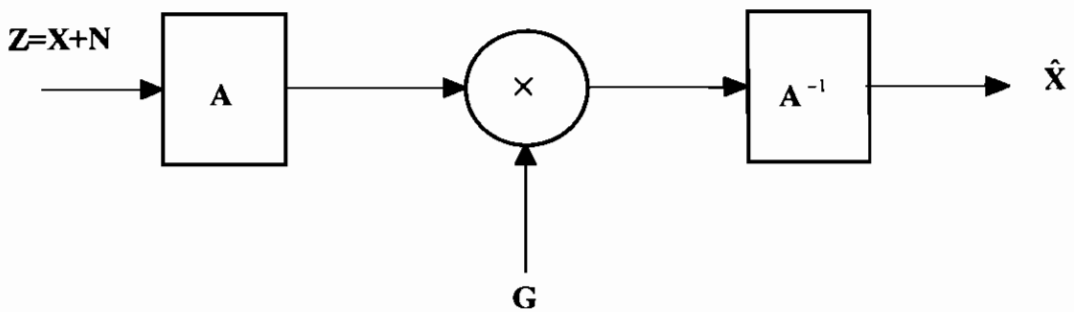


Figura 3.1.1.2 Filtraje usando filtro Wiener [24].

El filtro *WIENER* es una matriz de tamaño  $M \times M$ , denotada como  $\mathbf{G}$ , si se restringe la matriz a ser diagonal<sup>1</sup> esta se denota como  $\mathbf{Q}$ .  $\mathbf{A}$  y  $\mathbf{A}^{-1}$ , representan a una transformación ortogonal y a su inversa respectivamente,  $\hat{\mathbf{X}}$  denota el valor estimado de  $\mathbf{X}$  usando el criterio del *error cuadrático medio*.

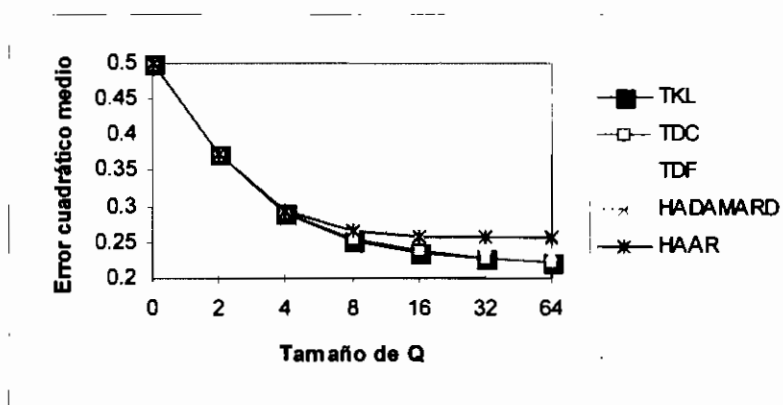


Figura 3.1.1.3 Error cuadrático medio de varias transformadas para filtro escalar Wiener con  $\rho=0,9$  [24].

El error cuadrático medio denotado  $e_Q$ , debido a un filtro escalar está dado por [24], mediante la siguiente expresión:

$$e_Q = 1 - \frac{1}{M} \sum_{j=0}^{M-1} \frac{\mathbf{R}_x(j, j)}{\mathbf{R}_x(j, j) + \mathbf{R}_n(j, j)} \quad (3.1.1.1)$$

en donde  $\mathbf{R}_x$  y  $\mathbf{R}_n$ , denotan respectivamente las matrices covarianza de datos y ruido en el dominio de la transformada. La figura 3.1.1.3 grafica los valores de  $e_Q$  para  $\rho = 0,9$  y diferentes tamaños de transformadas unitarias.

### c) Función tasa de distorsión

Esta característica se muestra en la figura 3.1.1.4. Esta propiedad se definió en la cuarta propiedad de la sección 2.2.7.

<sup>1</sup> Si  $\mathbf{G}$  es diagonal el filtrado se llama *escalar*, el cual toma individualmente cada componente espectral del vector de entrada. Si  $\mathbf{G}$  no es diagonal, el filtrado se llama *vectorial* y combina todas las componentes espectrales del vector de entrada. Los requerimientos computacionales determinarán el uso de tal o cual matriz [11].

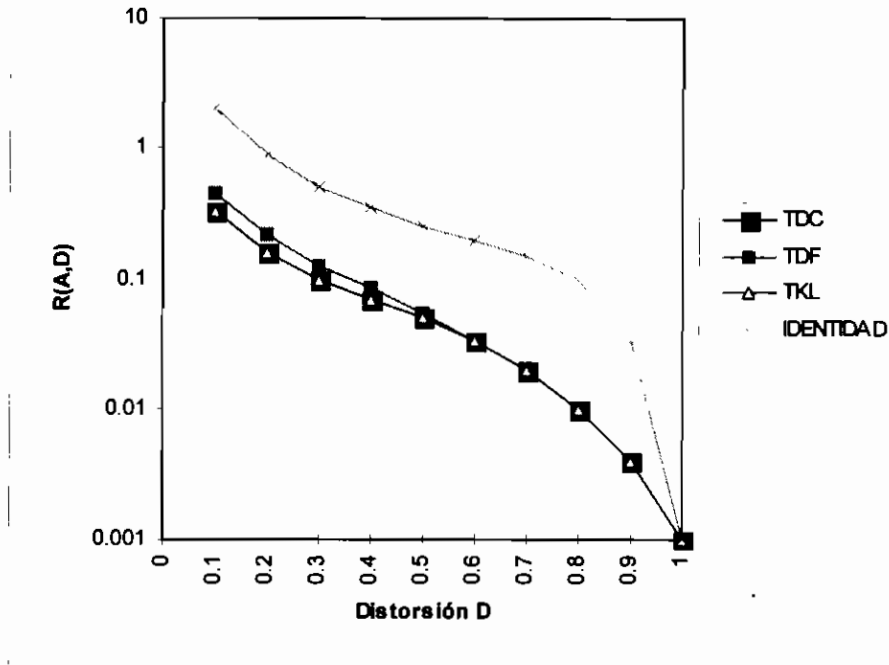


Figura 3.1.1.4 Tasa versus distorsión para transformadas  $16 \times 16$  y  $\rho = 0,9$  [24].

Por su cercanía a la transformada KL la transformada coseno ha encontrado aplicación en compresión y codificación de imágenes [25], [27], [29] y [34], para imágenes altamente correlacionadas<sup>1</sup> con un modelo de Markov de primer orden. La transformada coseno tiene amplia aplicación en procesamiento digital de imágenes y señales, en extracción de rasgos, en la realización de bancos de filtros en sistemas TDM-FDM ,etc [27] y [29].

### 3.1.2 DEFINICIÓN DE LA TRANSFORMADA DISCRETA COSENO

La ecuación (2.2.2.1) presenta la definición de una versión de la matriz transformada discreta coseno (otras versiones de la transformada discreta coseno se presentan en el Apéndice C), dicha definición se describe en el apartado 2.2.2. La definición de esta

<sup>1</sup> Alta correlación representará, coeficiente de correlación mayor de 0,9.

matriz se reproduce a continuación, la misma que como se demostrará más adelante cumple con ser unitaria:

$$\mathbf{C} = \left\{ \sqrt{\frac{2}{N}} e^{(k)} \cos\left(\frac{(m+1/2)k\pi}{N}\right), e^{(k)} = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } k = 0 \\ 1 & \text{si } k \neq 0 \end{cases} \quad 0 \leq k, m \leq N-1 \right\} \quad (3.1.2.1)$$

los vectores propios de la matriz covarianza  $\mathbf{R}$  (referirse a la quinta propiedad de la sección 2.2.2) para  $N = 8$ , con coeficiente de correlación  $\rho = 0.9$  están dados por el siguiente conjunto [1], [24]:

$$\left\{ \frac{1}{\sqrt{8}}, \frac{1}{2} \cos\left(\frac{(2m+1)k\pi}{16}\right); k = 1, 2, \dots, 7 \text{ y } m = 0, 1, \dots, 7 \right\} \quad (3.1.2.2)$$

dichos vectores constituyen las funciones bases de la TDC dada por (3.1.2.1). y gráficamente se presenta en la figura 3.1.2.1 .

La figura 3.1.2.2 muestra las imágenes bases para la transformada discreta coseno  $8 \times 8$ , para el cálculo de dichas imágenes bases se utiliza la ecuación (2.1.1.1).

La TDC 2-D puede calcularse recurriendo a la TDC 1-D y a la propiedad de separabilidad, pero estos métodos requieren incrementar el número de cálculos. En algoritmos bidimensionales, el número de operaciones para el caso de arreglos de muestras de tamaño  $M \times N$  es del orden de  $MN \log_2 MN$  con  $M$  y  $N$  potencias de 2 [25], número que es menor al requerido mediante el uso de Transformadas Coseno rápidas TDCR 1-D o transformada rápida de Fourier unidimensionales TDFR 1-D.

Sea  $u(m,n)$  el  $m$ -ésimo elemento del  $n$ -ésimo vector columna de un arreglo espacial 2-D de tamaño  $M \times N$ , de manera similar  $v(k,l)$  es el correspondiente elemento en el dominio de la frecuencia, tomando a  $u(0,0)$  como el primer elemento del arreglo  $\mathbf{U}$  [39], el par de

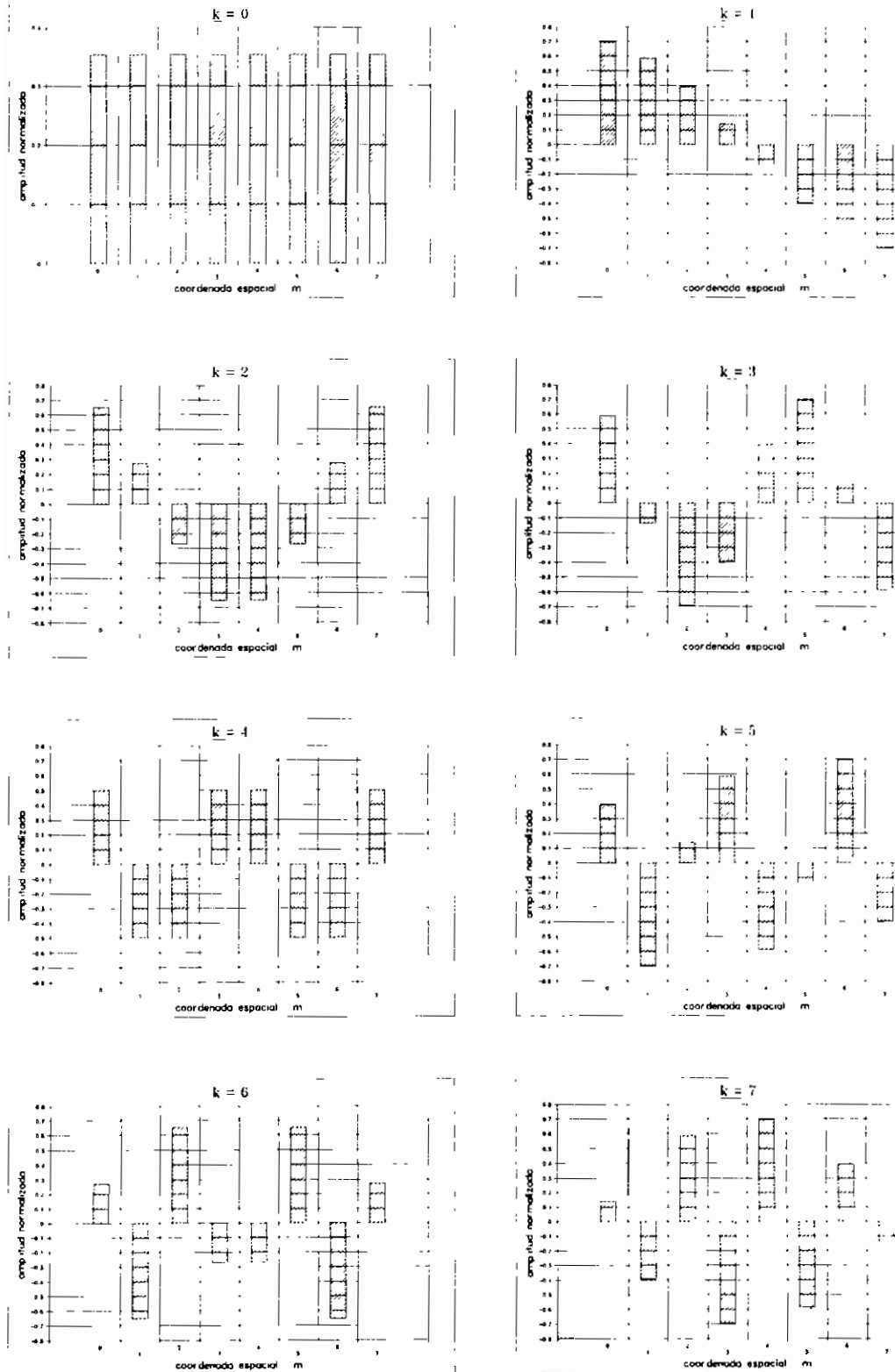


Figura 3.1.2.1 Vectores bases de la TDC cuando  $N=8$ .

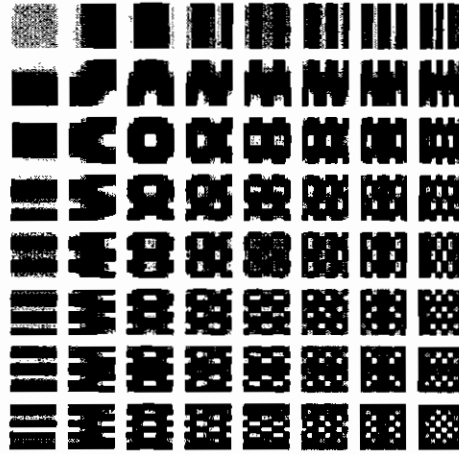


Figura 3.1.2.2 Imágenes bases de la Transformada Discreta Coseno 8x8.

transformadas coseno bidimensionales con coordenadas espaciales  $m,n$  y coordenadas en el dominio de la frecuencia  $k,l$  con  $k,m = 0,1,\dots, M-1$  y  $l,n = 0,1,\dots, N-1$  está dado por:

$$v(k,l) = \frac{2e(k)e(l)}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n) \cos\left(\frac{(2m+1)k\pi}{2M}\right) \cos\left(\frac{(2n+1)l\pi}{2N}\right) \quad (3.1.2.3)$$

$$u(m,n) = \frac{2}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} e(k)e(l)v(k,l) \cos\left(\frac{(2m+1)k\pi}{2M}\right) \cos\left(\frac{(2n+1)l\pi}{2N}\right) \quad (3.1.2.4)$$

estas dos relaciones pueden reescribirse de la siguiente manera:

$$v(k,l) = \frac{2e(k)e(l)}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n) C_{2M}^{(2m+1)k} C_{2N}^{(2n+1)l} \quad (3.1.2.5)$$

$$u(m,n) = \frac{2}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} e(k)e(l)v(k,l) C_{2M}^{(2m+1)k} C_{2N}^{(2n+1)l} \quad (3.1.2.6)$$

en estas expresiones  $C'_j = \cos\frac{\pi i}{j}$ , luego haciendo los siguientes remplazos:

$$\tilde{v}(k,l) = v(k,l) \frac{\sqrt{MN}}{2e(k)e(l)} \quad (3.1.2.7a)$$

$$A_{k,m}^M = C_{2M}^{(2m+1)k} \quad (3.1.2.7b)$$

$$\hat{v}(k, l) = v(k, l) \frac{2 e(k) e(l)}{\sqrt{MN}} \quad (3.1.2.7c)$$

el par de transformadas coseno bidimensionales para una imagen rectangular de tamaño  $M \times N$  expresado en términos matriciales toma la forma [1], [2] y [25]:

$$\tilde{\mathbf{V}} = \mathbf{A}^M \mathbf{U} (\mathbf{A}^N)^T \quad (3.1.2.8a)$$

$$\mathbf{U} = (\mathbf{A}^M)^T \tilde{\mathbf{V}} \mathbf{A}^N \quad (3.1.2.8b)$$

### 3.1.3 PROPIEDADES DE LA TRANSFORMADA DISCRETA COSENO

De entre todas las transformadas unitarias, la transformada Karhunen-Loeve (TKL) empaqueta la máxima cantidad de energía.

Para coeficientes de correlación menores a 0.5, la transformada discreta seno y para coeficientes de correlación ubicados en el intervalo [0.5, 1] la transformada discreta coseno son las mejores substitutas para la KL [2]. Para imágenes monocromáticas, el coeficiente de correlación se acerca a la unidad, entonces la TDC es la preferida.

En el apartado (2.2.2) se presentan algunas de las propiedades de la transformada discreta coseno, a continuación se reproducen estas propiedades y otras no incluidas en el capítulo 2, cada una con su respectiva demostración:

1.- La matriz transformada discreta coseno es real y ortogonal

Para la demostración de la ortogonalidad de la matriz coseno dada por  $\mathbf{C} = \{c(k, m)\}$  en la cual:



$$c(k, m) = \sqrt{\frac{2}{N}} e(k) \cos\left(\frac{(m+1/2)k\pi}{N}\right), \quad e(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } k = 0 \\ 1 & \text{si } k \neq 0 \end{cases} \quad (3.1.3.1)$$

habrá que verificar las siguientes dos igualdades:

$$\mathbf{C}\mathbf{C}^T = \mathbf{C}^T\mathbf{C} = \mathbf{I} \quad (3.1.3.2)$$

en las cuales  $\mathbf{I}$  es la matriz identidad.

a) Demostración de  $\mathbf{C}\mathbf{C}^T = \mathbf{I}$

Denótese el  $m$ -ésimo elemento de la  $k$ -ésima fila de la matriz  $\mathbf{C}$  por:

$$c(k, m) = \sqrt{\frac{2}{N}} e(k) \cos\left(\frac{(m+1/2)k\pi}{N}\right) \quad 0 \leq k, m \leq N-1 \quad (3.1.3.3)$$

Sea el  $m$ -ésimo elemento de la  $l$ -ésima columna de la matriz  $\mathbf{C}^T$  el denotado por:

$$c'(m, l) = \sqrt{\frac{2}{N}} e(l) \cos\left(\frac{(m+1/2)l\pi}{N}\right) \quad 0 \leq m, l \leq N-1 \quad (3.1.3.4)$$

entonces el  $k, l$ -ésimo elemento del producto  $\mathbf{C}\mathbf{C}^T$  vendrá dado por:

$$\begin{aligned} a(k, l) &= \sum_{m=0}^{N-1} c(k, m) c'(m, l) \\ &= \sqrt{\frac{2}{N}} e(k) \sqrt{\frac{2}{N}} e(l) \sum_{m=0}^{N-1} \cos\left(\frac{(m+1/2)k\pi}{N}\right) \cos\left(\frac{(m+1/2)l\pi}{N}\right) \\ &= \frac{2}{N} e(k) e(l) \sum_{m=0}^{N-1} \cos\left(\frac{(m+1/2)k\pi}{N}\right) \cos\left(\frac{(m+1/2)l\pi}{N}\right) \end{aligned} \quad (3.1.3.5)$$

usando la identidad trigonométrica  $2 \cos a \cos b = \cos(a+b) + \cos(a-b)$ , (3.1.3.5) se

convierte en:

$$\begin{aligned} a(k, l) &= \frac{1}{N} e(k) e(l) \sum_{m=0}^{N-1} \left\{ \cos\left(\frac{(m+1/2)(k+l)\pi}{N}\right) + \cos\left(\frac{(m+1/2)(k-l)\pi}{N}\right) \right\} \\ &= \frac{1}{N} e(k) e(l) (A1 + A2) \end{aligned} \quad (3.1.3.6)$$

Sea la función  $A(b) = \sum_{m=0}^{N-1} \cos\left(\frac{(m+1/2)b}{N}\right)$ , esta expresión puede reescribirse como:

$$A(b) = \sum_{m=0}^{N-1} \operatorname{Re}\left\{\exp\left(\frac{j(m+1/2)b}{N}\right)\right\} = \operatorname{Re}\left\{\exp\left(\frac{jb}{2N}\right) \sum_{m=0}^{N-1} \left[\exp\left(\frac{jb}{N}\right)\right]^m\right\} \quad (3.1.3.7)$$

utilizando la equivalencia de la suma geométrica, la función  $A(b)$  se simplifica para obtener:

$$A(b) = \operatorname{Re}\left\{\exp\left(\frac{jb}{2N}\right) \frac{(1 - \exp(jb))}{(1 - \exp(jb/N))}\right\} \quad (3.1.3.8)$$

luego se procede a racionalizar, entonces:

$$A(b) = \operatorname{Re}\left\{\exp\left(\frac{jb}{2N}\right) \frac{(1 - \exp(jb)) (1 - \exp(-jb/N))}{(1 - \exp(jb/N)) (1 - \exp(-jb/N))}\right\} \quad (3.1.3.9)$$

desarrollando íntegramente en sumas tanto el numerador como el denominador y tomando en consideración las dos siguientes relaciones,

$$2 \cos a = \exp(ja) + \exp(-ja) \quad \text{y} \quad j2 \sin a = \exp(ja) - \exp(-ja)$$

se llega a:

$$A(b) = \operatorname{Re}\left\{\frac{j \sin(b/2N)(1 - \exp(jb))}{1 - \cos(b/N)}\right\} = \operatorname{Re}\left\{\frac{j \sin(b/2N)(1 - \cos b - j \sin b)}{1 - \cos(b/N)}\right\} \quad (3.1.3.10)$$

desarrollando el numerador y extrayendo la parte real de dicha expresión se obtiene el valor de  $A(b)$ , y este es:

$$A(b) = \frac{\sin(b/2N) \sin b}{1 - \cos(b/N)} \quad (3.1.3.11)$$

Nótese que cuando  $b = 0$ , se tiene una indeterminación del tipo  $\frac{0}{0}$ , aplicando la regla de

L'Hospital, se concluye que cuando  $b \rightarrow 0$   $A(0) = N$ .

De lo anterior se llega a lo siguiente:

$$A1 = A((k+l)\pi) = \begin{cases} N & \text{si } k=l=0 \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.3.12)$$

y

$$A2 = A((k-l)\pi) = \begin{cases} N & \text{si } k=l \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.3.13)$$

finalmente de (3.1.3.12) y (3.1.3.13) se encuentra el valor de (3.1.3.6):

$$a(k,l) = \frac{e(k)e(l)}{N} (A1 + A2) = \begin{cases} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} (N+N) = 1 & \text{si } k=l=0 \\ \frac{1}{N} (0+N) = 1 & \text{si } k=l \neq 0 \\ 0 & \text{en otro caso} \end{cases} \quad (3.1.3.14)$$

en consecuencia:  $\{a(k,l), 0 \leq k, l \leq N-1\} = \mathbf{I}$ .

**b) Demostración de  $\mathbf{C}^T \mathbf{C} = \mathbf{I}$**

Sea el  $k$ -ésimo elemento de la  $m$ -ésima fila de la matriz  $\mathbf{C}^T$  y denotado por:

$$c(m,k) = \sqrt{\frac{2}{N}} e(k) \cos\left(\frac{(m+1/2)k\pi}{N}\right) \quad 0 \leq m, k \leq N-1 \quad (3.1.3.15)$$

Sea el  $k$ -ésimo elemento de la  $n$ -ésima columna de la matriz  $\mathbf{C}$  denotado por:

$$c'(k,n) = \sqrt{\frac{2}{N}} e(k) \cos\left(\frac{(n+1/2)k\pi}{N}\right) \quad 0 \leq k, n \leq N-1 \quad (3.1.3.16)$$

entonces el producto  $\mathbf{C}^T \mathbf{C}$  se encontrará como:

$$\begin{aligned} a(m,n) &= \sum_{k=0}^{N-1} c(m,k) c'(k,n) \\ &= \sum_{k=0}^{N-1} \left[ \sqrt{\frac{2}{N}} e(k) \right]^2 \cos\left(\frac{(m+1/2)k\pi}{N}\right) \cos\left(\frac{(n+1/2)k\pi}{N}\right) \\ &= \frac{1}{N} \sum_{k=0}^{N-1} [e(k)]^2 \left\{ \cos\left(\frac{(m+n+1)k\pi}{N}\right) + \cos\left(\frac{(m-n)k\pi}{N}\right) \right\} \end{aligned} \quad (3.1.3.17)$$

luego:

$$\begin{aligned}
 a(m, n) &= \frac{1}{N} \left[ \left( \frac{1}{\sqrt{2}} \right)^2 (1+1) + \sum_{k=1}^{N-1} \left\{ \cos \left( \frac{(m+n+1)k\pi}{N} \right) + \cos \left( \frac{(m-n)k\pi}{N} \right) \right\} \right] \\
 &= \frac{1}{N} \left[ \left\{ \frac{1}{2} + \sum_{k=1}^{N-1} \cos \left( \frac{(m+n+1)k\pi}{N} \right) \right\} + \left\{ \frac{1}{2} + \sum_{k=1}^{N-1} \cos \left( \frac{(m-n)k\pi}{N} \right) \right\} \right] \\
 &= \frac{1}{N} (A3 + A4)
 \end{aligned} \tag{3.1.3.18}$$

se conoce que<sup>1</sup>:

$$\sigma_n(z) = \frac{1}{2} + \sum_{k=1}^n \cos kz = \frac{\operatorname{sen} \frac{(2n+1)z}{2}}{2 \operatorname{sen} \frac{z}{2}} \tag{3.1.3.19}$$

en consecuencia, haciendo  $z = (m+n+1)\pi/N$  en (3.1.3.19) y simplificando se tiene:

$$A3 = \sigma_{N-1}((m+n+1)\pi/N) = \frac{\operatorname{sen} \left[ (m+n+1)\pi - \frac{(m+n+1)\pi}{2N} \right]}{2 \operatorname{sen} \left[ \frac{(m+n+1)\pi}{2N} \right]} \tag{3.1.3.20}$$

y para el caso  $z = (m-n)\pi/N$ , se obtiene el valor de A4:

$$A4 = \sigma_{N-1}((m-n)\pi/N) = \frac{\operatorname{sen} \left[ (m-n)\pi - \frac{(m-n)\pi}{2N} \right]}{2 \operatorname{sen} \left[ \frac{(m-n)\pi}{2N} \right]} \tag{3.1.3.21}$$

evaluando (3.1.3.20) y (3.1.3.21) se obtiene:

$$A3 = \begin{cases} -\frac{1}{2} & \text{si } m+n+1 = \text{par o } m+n = \text{impar} \\ \frac{1}{2} & \text{si } m+n+1 = \text{impar o } m+n = \text{par} \end{cases} \tag{3.1.3.22}$$

<sup>1</sup> PISKUNOV, Cálculo Diferencial e Integral, Editorial Mir Moscú, Sexta Edición, 1977, tomo 2, pag. 348.

$$A4 = \begin{cases} -\frac{1}{2} & \text{si } m - n = \text{par} \neq 0, \\ \frac{1}{2} & \text{si } m - n = \text{impar} \\ \frac{\pi - \pi/2N}{2\pi/2N} = N - 1/2 & \text{si } m = n \end{cases} \quad (3.1.3.23)$$

finalmente, considerando que tanto  $(m+n)$  como  $(m-n)$  tienen la misma paridad para las condiciones de paridad de  $m$  y  $n$ , y reemplazando (3.1.3.22) y (3.1.3.23) en (3.1.3.18) se concluye:

$$a(m,n) = \frac{1}{N}(A3 + A4) = \begin{cases} \frac{1}{N}\left(\frac{1}{2} - \frac{1}{2}\right) = 0 & \text{si } m + n = \text{par } m \neq n, \\ \frac{1}{N}\left(-\frac{1}{2} + \frac{1}{2}\right) = 0 & \text{si } m + n = \text{impar} \\ \frac{1}{N}\left(\frac{1}{2} + N - \frac{1}{2}\right) = 1 & \text{si } m = n \end{cases} \quad (3.1.3.24)$$

en consecuencia  $\{a(m,n), 0 \leq m, n \leq N-1\} = \mathbf{I}$ , por tanto de (3.1.3.14) y (3.1.3.24) se concluye que la matriz TDC es ortogonal.

## 2.- La transformada discreta coseno puede ser particionada en cuatro cuadrantes

- Para filas pares los cuadrantes izquierdo y derecho son iguales.
- Para filas impares los cuadrantes izquierdo y derecho son opuestos en signo.

*Demostración:*

Sea  $\tilde{t}_{k,m} = \cos\left(\theta_k + \frac{2km\pi}{N}\right)$  y  $\theta_k = \frac{k\pi}{2N}$ , desarrollando el coseno de la suma se tiene:

$$\tilde{t}_{k,m} = \cos\theta_k \cos\left(\frac{2km\pi}{N}\right) - \text{sen}\theta_k \text{sen}\left(\frac{2km\pi}{N}\right) \quad 0 \leq k, m \leq N-1 \quad (3.1.3.25)$$

separando el dominio de  $m$  en dos intervalos, esto es:

$$0 \leq m_1 \leq \frac{N}{2} - 1 \quad ; \quad \frac{N}{2} \leq m_2 \leq N-1 \quad \Rightarrow \quad m_2 = m_1 + \frac{N}{2}$$

para el primer intervalo:

$$\tilde{t}_{k,m_1} = \cos\theta_k \cos\left(\frac{2km_1\pi}{N}\right) - \text{sen}\theta_k \text{sen}\left(\frac{2km_1\pi}{N}\right) \quad (3.1.3.26)$$

en el segundo intervalo, en cambio:

$$\tilde{t}_{k,m_2} = \cos\theta_k \cos\left(\frac{2km_2\pi}{N}\right) - \text{sen}\theta_k \text{sen}\left(\frac{2km_2\pi}{N}\right) \quad (3.1.3.27)$$

en esta última expresión reemplazando el valor de  $m_2$  y simplificando:

$$\tilde{t}_{k,m_2} = \cos\theta_k \cos\left(\frac{2km_1\pi}{N} + k\pi\right) - \text{sen}\theta_k \text{sen}\left(\frac{2km_1\pi}{N} + k\pi\right) \quad (3.1.3.28)$$

como conclusión analizando (3.1.3.26) y (3.1.3.28) se concluye lo siguiente:

$$\tilde{t}_{k,m_2} = \begin{cases} \tilde{t}_{k,m_1} & \text{si } k \text{ par} \\ -\tilde{t}_{k,m_1} & \text{si } k \text{ impar} \end{cases} \quad (3.1.3.29)$$

Gracias a esta propiedad se puede desplazar todas las filas numeradas pares ( $k = 0, 2, 4, \dots$ ) hacia la mitad superior de la matriz coseno y presentarla de la siguiente manera:

$$\mathbf{C} = \begin{pmatrix} \mathbf{P}\left(\frac{N}{2}\right) & \mathbf{P}\left(\frac{N}{2}\right) \\ \mathbf{I}\left(\frac{N}{2}\right) & -\mathbf{I}\left(\frac{N}{2}\right) \end{pmatrix}_{N \times N} \quad (3.1.3.30)$$

en donde cada una de las 4 submatrices son cuadradas de tamaño  $\frac{N}{2} \times \frac{N}{2}$ .

### 3.- Propiedad de desplazamiento

Considérese una secuencia de entrada dado por el vector  $\mathbf{u}$  y la versión desplazada (desplazada una muestra en retraso) del mismo vector denotado por  $\mathbf{u}_+$ , definidos como a continuación:

$$\mathbf{u} = \begin{pmatrix} u(0) \\ u(1) \\ \vdots \\ u(N-1) \end{pmatrix} \quad \mathbf{u}_+ = \begin{pmatrix} u(1) \\ u(2) \\ \vdots \\ u(N) \end{pmatrix} \quad (3.1.3.31)$$

esta propiedad proporciona una relación entre sus respectivas transformadas coseno  $v$  y  $v_+$ , ayudando de esta manera a reducir la complejidad computacional haciendo factible su implementación en tiempo real [33].

*Demostración:*

Sean las igualdades:

$$v(k) = \sqrt{\frac{2}{N}} e(k) \sum_{m=0}^{N-1} u(m) \cos\left(\frac{(2m+1)k\pi}{2N}\right) \quad (3.1.3.32)$$

$$v_+(k) = \sqrt{\frac{2}{N}} e(k) \sum_{m=0}^{N-1} u(m+1) \cos\left(\frac{(2m+1)k\pi}{2N}\right) \quad (3.1.3.33)$$

$$v_s(k) = \sqrt{\frac{2}{N}} e(k) \sum_{m=0}^{N-1} u(m) \sin\left(\frac{(2m+1)k\pi}{2N}\right) \quad (3.1.3.34)$$

haciendo  $m = m+1-1$  en  $v_+(k)$ , esto es,

$$\begin{aligned} v_+(k) &= \sqrt{\frac{2}{N}} e(k) \sum_{m=0}^{N-1} u(m+1) \cos\left(\frac{(2(m+1-1)+1)k\pi}{2N}\right) \\ &= \sqrt{\frac{2}{N}} e(k) \sum_{m=0}^{N-1} u(m+1) \cos\left(\frac{(2(m+1)+1)k\pi}{2N} - \frac{2k\pi}{2N}\right) \end{aligned} \quad (3.1.3.35)$$

aplicando en (3.1.3.35) la identidad trigonométrica  $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$ .

extrayendo factor común y agrupando términos semejantes, se obtiene:

$$\begin{aligned}
 v_+(k) &= \sqrt{\frac{2}{N}} e(k) \left\{ \cos\left(\frac{k\pi}{N}\right) \sum_{m=0}^{N-1} u(m+1) \cos\left(\frac{(2(m+1)+1)k\pi}{2N}\right) + \right. \\
 &\quad \left. + \operatorname{sen}\left(\frac{k\pi}{N}\right) \sum_{m=0}^{N-1} u(m+1) \operatorname{sen}\left(\frac{(2(m+1)+1)k\pi}{2N}\right) \right\} \\
 &= \sqrt{\frac{2}{N}} e(k) \left\{ \cos\left(\frac{k\pi}{N}\right) T_1 + \operatorname{sen}\left(\frac{k\pi}{N}\right) T_2 \right\} \tag{3.1.3.36}
 \end{aligned}$$

en donde:

$$T_1 = \sum_{m=0}^{N-1} u(m+1) \cos\left(\frac{(2(m+1)+1)k\pi}{2N}\right) \text{ y } T_2 = \sum_{m=0}^{N-1} u(m+1) \operatorname{sen}\left(\frac{(2(m+1)+1)k\pi}{2N}\right)$$

reescribiendo adecuadamente  $T_1$  y  $T_2$  en función de  $v(k)$ :

$$\begin{aligned}
 T_1 &= \sum_{m=0}^{N-1} u(m+1) \cos\left(\frac{(2(m+1)+1)k\pi}{2N}\right) = \sum_{m=1}^N u(m) \cos\left(\frac{(2m+1)k\pi}{2N}\right) \\
 &= u(0) \cos\left(\frac{k\pi}{2N}\right) + \sum_{m=1}^{N-1} u(m) \cos\left(\frac{(2m+1)k\pi}{2N}\right) - u(0) \cos\left(\frac{k\pi}{2N}\right) + (-1)^k u(N) \cos\left(\frac{k\pi}{2N}\right) \\
 \Rightarrow T_1 &= \frac{v(k)}{e(k)} \sqrt{\frac{N}{2}} + (-1)^k u(N) \cos\left(\frac{k\pi}{2N}\right) - u(0) \cos\left(\frac{k\pi}{2N}\right) \tag{3.1.3.37}
 \end{aligned}$$

procediendo de igual manera para  $T_2$ :

$$T_2 = \frac{v_s(k)}{e(k)} \sqrt{\frac{N}{2}} + (-1)^k u(N) \operatorname{sen}\left(\frac{k\pi}{2N}\right) - u(0) \operatorname{sen}\left(\frac{k\pi}{2N}\right) \tag{3.1.3.38}$$

reemplazando los valores de  $T_1$  y  $T_2$  dados por (3.1.3.37) y (3.1.3.38) respectivamente se obtiene, la siguiente expresión:

$$v_+(k) = \sqrt{\frac{2}{N}} e(k) \left\{ \cos\left(\frac{k\pi}{N}\right) \left[ \frac{v(k)}{e(k)} \sqrt{\frac{N}{2}} + \cos\left(\frac{k\pi}{2N}\right) \{(-1)^k u(N) - u(0)\} \right] + \right. \\
 \left. + \operatorname{sen}\left(\frac{k\pi}{N}\right) \left[ \frac{v_s(k)}{e(k)} \sqrt{\frac{N}{2}} + \operatorname{sen}\left(\frac{k\pi}{2N}\right) \{(-1)^k u(N) - u(0)\} \right] \right\} \tag{3.1.3.39}$$



luego,

$$v_+(k) = \cos\left(\frac{k\pi}{N}\right)v(k) + \operatorname{sen}\left(\frac{k\pi}{N}\right)v_s(k) + \\ + \sqrt{\frac{2}{N}}e(k)\{(-1)^k u(N) - u(0)\} \left\{ \cos\left(\frac{k\pi}{N}\right)\cos\left(\frac{k\pi}{2N}\right) + \operatorname{sen}\left(\frac{k\pi}{N}\right)\operatorname{sen}\left(\frac{k\pi}{2N}\right) \right\}$$

operando trigonómicamente sobre esta expresión:

$$v_+(k) = \cos\left(\frac{k\pi}{N}\right)v(k) + \operatorname{sen}\left(\frac{k\pi}{N}\right)v_s(k) + \sqrt{\frac{2}{N}}e(k)\{(-1)^k u(N) - u(0)\} \cos\left(\frac{k\pi}{N} - \frac{k\pi}{2N}\right)$$

finalmente:

$$v_+(k) = \cos\left(\frac{k\pi}{N}\right)v(k) + \operatorname{sen}\left(\frac{k\pi}{N}\right)v_s(k) + \sqrt{\frac{2}{N}}e(k)\{(-1)^k u(N) - u(0)\} \cos\left(\frac{k\pi}{2N}\right) \quad (3.1.3.40)$$

De esto se puede afirmar que la TDC de una secuencia desplazada de datos puede obtenerse desde la transformada respectiva de la secuencia no desplazada. Esta propiedad será de utilidad en aplicaciones donde los datos ingresan en forma continua; evitando la necesidad de calcular la transformada del grupo completo de  $N$ -puntos en cada instante de tiempo, pues esta propiedad permitirá actualizar instantáneamente la transformada.

#### 4.- La transformada Discreta Coseno no es la parte Real de la Transformada Discreta de Fourier [1]

El elemento  $c(k,m)$  de la matriz TDC está definida de la siguiente manera:

$$c(k,m) = \sqrt{\frac{2}{N}}e(k)\cos\left(\frac{(2m+1)k\pi}{2N}\right) \quad (3.1.3.41)$$

Escribiendo en forma exponencial y separando términos se llega a la siguiente expresión:

$$\begin{aligned}
c(k, m) &= \sqrt{\frac{2}{N}} e(k) \operatorname{Re} \left\{ \exp \left( \frac{-j(2m+1)k\pi}{2N} \right) \right\} \\
&= \sqrt{2} e(k) \operatorname{Re} \left\{ \exp \left( \frac{j(2m-1)k\pi}{2N} \right) \frac{1}{\sqrt{N}} \exp \left( \frac{-j2\pi km}{N} \right) \right\}
\end{aligned} \quad (3.1.3.42)$$

tomando en consideración (2.2.1.1) y (2.2.1.2) se obtiene finalmente:

$$c(k, m) = \sqrt{2} e(k) \operatorname{Re} \left\{ \exp \left( \frac{j(2m-1)k\pi}{2N} \right) [\mathbf{F}]_{k,m} \right\} \neq \operatorname{Re} \{ [\mathbf{F}]_{k,m} \} \quad (3.1.3.43)$$

5.- La Transformada Discreta Coseno tiene una excelente compresión energética para datos altamente correlacionados

Esta afirmación se desprende del cumplimiento de las propiedades cuarta y quinta de la sección 2.2.2.

La demostración consistirá en probar que efectivamente la TDC es la mejor substituta de la TKL para secuencias de Markov de primer orden altamente correlacionadas, que es lo que manifiestan las propiedades en mención.

Para la demostración se parte tomando de la referencia [1], los valores y vectores propios de la matriz covarianza de la matriz TKL, los cuales son:

$$\lambda_k = \frac{1 - \rho^2}{1 - 2\rho \cos \omega_k + \rho^2} \quad (3.1.3.44)$$

$$\phi_k(m) = \phi(m, k) = \sqrt{\frac{2}{N + \lambda_k}} \operatorname{sen} \left( \omega_k \left( m + 1 - \frac{N+1}{2} \right) + \frac{(k+1)\pi}{2} \right) \quad 0 \leq m, k \leq N-1 \quad (3.1.3.45)$$

en donde  $\{\omega_k\}$  son las raíces positivas de la siguiente ecuación:

$$\operatorname{tg}(N\omega) = \frac{(\rho^2 - 1) \operatorname{sen} \omega}{\cos \omega - 2\rho + \rho^2 \cos \omega} \quad (3.1.3.46)$$

la matriz TKL está definida entonces como:  $\Phi' \triangleq \{\phi(k, m)\}$ . El siguiente paso es demostrar que cuando  $\rho$  tiende a 1, se cumpla la siguiente igualdad:

$$\alpha(k, m) = \sqrt{\frac{2}{N + \lambda_k}} \operatorname{sen} \left( \omega_k \left( m + 1 - \frac{N+1}{2} \right) + \frac{(k+1)\pi}{2} \right), \quad 0 \leq k, m \leq N-1 \quad (3.1.3.47)$$

*Demostración:*

En la ecuación (3.1.3.46) haciendo que  $\rho$  tienda a 1 y bajo la condición  $\cos \omega_k \neq 1$ , se obtienen las raíces, y estas son:

$$\operatorname{tg}(N\omega_k) = 0 \Rightarrow N\omega_k = k\pi \Rightarrow \omega_k = \frac{k\pi}{N} \neq 0 \quad \wedge \quad \lambda_{k \neq 0} = 0$$

luego reemplazando esta solución en el transpuesto del vector dado por (3.1.3.45), se encuentra una relación con los vectores bases de la matriz TDC, procediendo de la siguiente manera:

$$\begin{aligned} \phi(k, m) &= \sqrt{\frac{2}{N + \lambda_k}} \operatorname{sen} \left( \frac{k\pi}{N} \left( \frac{2m+1-N}{2} \right) + \frac{k\pi}{2} + \frac{\pi}{2} \right) \\ &= \sqrt{\frac{2}{N + \lambda_k}} \operatorname{sen} \left( \frac{k\pi(2m+1)}{2N} + \frac{\pi}{2} \right) = \sqrt{\frac{2}{N + \lambda_k}} \cos \left( \frac{k\pi(2m+1)}{2N} \right) \end{aligned} \quad (3.1.3.48)$$

el siguiente paso es buscar el valor de  $\lambda$  cuando  $k=0$ , que se dirá cuando  $\omega_k \rightarrow 0$ , bajo esta condición,  $\operatorname{tg}(N\omega) \approx N\omega$ ,  $\operatorname{sen} \omega \approx \omega$  entonces (3.1.3.46) se convierte en (3.1.3.49),

esto es:

$$N\omega = \frac{(\rho^2 - 1)\omega}{\cos \omega - 2\rho + \rho^2 \cos \omega} \quad (3.1.3.49)$$

despejando a continuación  $\cos \omega$ , se obtiene la siguiente expresión:

$$\cos \omega = \frac{\rho^2 - 1 + 2N\rho}{N(1 + \rho^2)} \quad (3.1.3.50)$$

reemplazando (3.1.3.50) en (3.1.3.44) se obtiene  $\lambda_0$ , del modo siguiente:

$$\lambda_0 = \frac{1 - \rho^2}{1 - 2\rho \frac{\rho^2 - 1 + 2N\rho}{N(1 + \rho^2)} + \rho^2} = \frac{N(1 - \rho^4)}{N(1 + \rho^2)^2 - 2\rho(\rho^2 - 1 + 2N\rho)} \quad (3.1.3.51)$$

evaluando  $\lambda_0$  para  $\rho = 1$ , se obtiene una indeterminación del tipo  $\frac{0}{0}$ , salvando la indeterminación por derivación con respecto al coeficiente de correlación, se obtiene,

$$\lambda_0 = \frac{-4N\rho^3}{2 - 4N\rho - 6\rho^2 + 4N\rho^3} \quad (3.1.3.52)$$

evaluando luego (3.1.3.52) en  $\rho = 1$ ,

$$\lambda_0 = \frac{-4N}{2 - 4N - 6 + 4N} = N \quad (3.1.3.53)$$

resumiendo:

$$\phi(k, m) = \sqrt{\frac{2}{N + \lambda_k}} \cos\left(\frac{(2m+1)k\pi}{2N}\right), \lambda_k = \begin{cases} 0 & \text{si } k \neq 0 \\ N & \text{si } k = 0 \end{cases} \quad 0 \leq k, m \leq N-1 \quad (3.1.3.54)$$

en consecuencia el elemento  $\phi(k, m)$  de (3.1.3.45) constituye el  $k, m$ -ésimo elemento de la matriz TDC.

## 3.2 CÁLCULO DE LA TRANSFORMADA DISCRETA COSENO, ALGORITMO

Desde la introducción de la Transformada Discreta Coseno y por la diversidad de aplicaciones dadas, se han desarrollado muchos algoritmos<sup>1</sup> para su implementación, estos algoritmos pueden ser ubicados en una de las siguientes categorías, dependiendo del método de aproximación [27]:

<sup>1</sup> Algoritmo.- Conjunto de reglas, procesos o ecuaciones encaminadas a la resolución de un problema en un número finito de pasos [23].

1. Cálculo indirecto.- En este caso se usan tanto la TDF como la transformada Walsh-Hadamard para la implementación de la transformada discreta coseno, pero en su cálculo participan muchas operaciones innecesarias.
2. Factorización directa.- Estos algoritmos requieren factorización matricial; teóricamente una matriz unitaria puede factorarse siempre en matrices menos densas, obteniéndose ganancia en velocidad en el proceso de cálculo de la TDC, pues requiere un número menor de operaciones.
3. Cálculo recursivo.- Este método persigue generar una matriz TDC de orden superior partiendo de una matriz TDC de orden inferior.

El algoritmo que se describe a continuación se ubica en la tercera categoría. Es una versión extendida para el caso 2-D de algoritmos de cálculo de la Transformada Discreta Coseno 1-D. Este algoritmo bidimensional de la Transformada rápida coseno requiere sumas y multiplicaciones reales solamente y trabaja directamente sobre la imagen dato bidimensional.

Las ecuaciones (3.1.2.3)-(3.1.2.8) representan la definición de la Transformada Rápida Coseno TRC y el algoritmo descrito requiere para su implementación  $3/8 \times MN \log_2 MN$  operaciones reales sin normalizaciones iniciales [25] para una imagen dato  $\{U_{m,n}\}_{M \times N}$  de tamaño  $2^m \times 2^n$ .

El par de transformadas coseno, se define de la siguiente manera:

$$v(k,l) = \frac{2e(k)e(l)}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n) C_{2M}^{(2m+1)k} C_{2N}^{(2n+1)l} \quad (3.2.0.1)$$

$$u(m,n) = \frac{2}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} e(k)e(l) v(k,l) C_{2M}^{(2m+1)k} C_{2N}^{(2n+1)l} \quad (3.2.0.2)$$

en estas expresiones  $C_j^i = \cos \frac{\pi i}{J}$  y,  $e(k)$  como se definió en la ecuación (3.1.2.1),

asumiendo además las siguientes notaciones:

$$\tilde{v}(k, l) = v(k, l) \frac{\sqrt{MN}}{2e(k)e(l)} \quad (3.2.0.3)$$

$$\hat{v}(k, l) = v(k, l) \frac{2e(k)e(l)}{\sqrt{MN}} \quad (3.2.0.4)$$

$$\mathbf{A}_{k,m}^M = C_{2M}^{(2m+1)k} \quad (3.2.0.5)$$

$$\mathbf{B}_{l,n}^N = C_{2N}^{(2n+1)l} \quad (3.2.0.6)$$

con los dominios de  $m, n, k$  y  $l$  establecidos en (3.2.0.1) y (3.2.0.2) el par de transformadas coseno en forma matricial es:

$$\tilde{\mathbf{V}} = \mathbf{A}^M \mathbf{U} (\mathbf{A}^N)^T \quad (3.2.0.7)$$

$$\mathbf{U} = (\mathbf{A}^M)^T \hat{\mathbf{V}} \mathbf{A}^N \quad (3.2.0.8)$$

El primer paso en el desarrollo del algoritmo de cálculo de la TDC consiste en particionar y reordenar convenientemente las matrices  $\mathbf{A}$ ,  $\mathbf{U}$  y  $\mathbf{V}$ , de las expresiones (3.2.0.7) y (3.2.0.8), como se procede a continuación;

### 3.2.1 PARTICIÓN Y REORDENAMIENTO DE MATRICES

1.- Matriz de datos espaciales.- La matriz de datos espaciales  $\mathbf{U}$  de tamaño  $M \times N$  se particiona y reordena en cuatro submatrices de tamaño  $M/2 \times N/2$  de la siguiente manera:

$$\mathbf{U} = \{u(m, n)\} = \begin{bmatrix} \hat{\mathbf{p}} & \hat{\mathbf{q}} \\ \hat{\mathbf{r}} & \hat{\mathbf{s}} \end{bmatrix} \quad (3.2.1.1)$$

para ello debe cumplirse con las siguientes condiciones [25] y [26]:

$$\begin{aligned}
 \hat{\mathbf{p}}_{m,n} &= u(m,n) \\
 \hat{\mathbf{q}}_{m,n} &= u(m, N-1-n) \\
 \hat{\mathbf{r}}_{m,n} &= u(M-1-m, n) \\
 \hat{\mathbf{s}}_{m,n} &= u(M-1-m, N-1-n)
 \end{aligned}
 \tag{3.2.1.2}$$

para:

$$\begin{aligned}
 m &= 0, 1, \dots, M/2 - 1 \\
 n &= 0, 1, \dots, N/2 - 1
 \end{aligned}$$

2.- Matriz en dominio de la frecuencia espacial.- La matriz en el dominio de la frecuencia espacial denotada por  $\mathbf{V}$  de tamaño  $M \times N$  se particiona y reordena en cuatro submatrices de tamaño  $M/2 \times N/2$  de la siguiente manera:

$$\mathbf{V} = \{v(k,l)\} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix}
 \tag{3.2.1.3}$$

las mismas que cumplen:

$$\begin{aligned}
 \mathbf{P}_{k,l} &= v(2k, 2l) \\
 \mathbf{Q}_{k,l} &= v(2k, 2l+1) \\
 \mathbf{R}_{k,l} &= v(2k+1, 2l) \\
 \mathbf{S}_{k,l} &= v(2k+1, 2l+1)
 \end{aligned}
 \tag{3.2.1.4}$$

para:

$$\begin{aligned}
 k &= 0, 1, \dots, M/2 - 1 \\
 l &= 0, 1, \dots, N/2 - 1
 \end{aligned}$$

3.- Matrices Coseno.- La matriz de coeficientes  $\mathbf{A}^M$  de tamaño  $M \times M$  es particionada y reagrupada usando las cuatro submatrices  $\mathbf{A1}$ ,  $\mathbf{A2}$ ,  $\mathbf{A3}$  y  $\mathbf{A4}$  de tamaño  $M/2 \times M/2$ ,

$$\mathbf{A}^M = \begin{bmatrix} \mathbf{A1} & \mathbf{A3} \\ \mathbf{A2} & \mathbf{A4} \end{bmatrix}
 \tag{3.2.1.5}$$

las mismas que cumplen:

$$\begin{aligned}
 \mathbf{A1}_{k,m} &= \mathbf{A}_{2k,m}^M \\
 \mathbf{A2}_{k,m} &= \mathbf{A}_{2k+1,m}^M \\
 \mathbf{A3}_{k,m} &= \mathbf{A}_{2k,M-1-m}^M \\
 \mathbf{A4}_{k,m} &= \mathbf{A}_{2k+1,M-1-m}^M
 \end{aligned} \tag{3.2.1.6}$$

dentro de los intervalos  $k, m = 0, 1, \dots, M/2 - 1$ , desarrollando **A3** y **A4**, usando (3.2.0.5) y (3.2.1.6), aplicando las propiedades de paridad y periodicidad de la función coseno, se tiene:

$$\begin{aligned}
 \mathbf{A3}_{k,m} &= \mathbf{A}_{2k,M-1-m}^M = \cos\left(\frac{(2(M-1-m)+1)2k\pi}{2M}\right) \\
 &= \cos\left(\frac{2M \times 2k\pi}{2M} - \frac{(2m+1)2k\pi}{2M}\right) = \cos\left(\frac{(2m+1)2k\pi}{2M}\right) \\
 &= \mathbf{A}_{2k,m}^M = \mathbf{A1}_{k,m}
 \end{aligned} \tag{3.2.1.7}$$

de manera similar **A4** será:

$$\begin{aligned}
 \mathbf{A4}_{k,m} &= \mathbf{A}_{2k+1,M-1-m}^M = \cos\left(\frac{(2(M-1-m)+1)(2k+1)\pi}{2M}\right) \\
 &= \cos\left(\frac{2M(2k+1)\pi}{2M} - \frac{(2m+1)(2k+1)\pi}{2M}\right) = -\cos\left(\frac{(2m+1)(2k+1)\pi}{2M}\right) \\
 &= -\mathbf{A}_{2k+1,m}^M = -\mathbf{A2}_{k,m}
 \end{aligned} \tag{3.2.1.8}$$

De acuerdo a (3.2.1.6), (3.2.1.7) y (3.2.1.8),

$$\mathbf{A}^M = \begin{bmatrix} \mathbf{A1} & \mathbf{A1} \\ \mathbf{A2} & -\mathbf{A2} \end{bmatrix} \tag{3.2.1.9}$$

de manera similar la matriz de coeficientes **B** tomará la siguiente forma:

$$\mathbf{B}^N = \begin{bmatrix} \mathbf{B1} & \mathbf{B1} \\ \mathbf{B2} & -\mathbf{B2} \end{bmatrix} \tag{3.2.1.10}$$

en donde:

$$\begin{aligned}
 \mathbf{B1}_{l,n} &= \mathbf{B}_{2l,n}^N \\
 \mathbf{B2}_{l,n} &= \mathbf{B}_{2l+1,n}^N
 \end{aligned} \tag{3.2.1.11}$$



El siguiente paso en el desarrollo del algoritmo utiliza el proceso matricial realizado en el numeral anterior y lo aplica en la expresión de la TDC directa, de manera, que esta quede expresada en función de 4 transformadas directas de orden inmediatamente inferior, es decir de tamaño  $M/2 \times N/2$ .

### 3.2.2 TRANSFORMADA DISCRETA COSENO DIRECTA

La ecuación (3.2.0.1) usando (3.2.0.3), (3.2.0.5) y (3.2.0.6) puede reproducirse de la siguiente forma:

$$\begin{aligned} \tilde{v}(k, l) = & \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} u(m, n) \mathbf{A}_{k,m}^M \mathbf{B}_{l,n}^N + \sum_{m=0}^{M/2-1} \sum_{n=N/2}^{N-1} u(m, n) \mathbf{A}_{k,m}^M \mathbf{B}_{l,n}^N + \\ & + \sum_{m=M/2}^{M-1} \sum_{n=0}^{N/2-1} u(m, n) \mathbf{A}_{k,m}^M \mathbf{B}_{l,n}^N + \sum_{m=M/2}^{M-1} \sum_{n=N/2}^{N-1} u(m, n) \mathbf{A}_{k,m}^M \mathbf{B}_{l,n}^N \end{aligned} \quad (3.2.2.1)$$

si se cambian de variables adecuadamente de manera que  $m$  y  $n$  cubran el dominio de (3.2.1.2),

$$\tilde{v}(k, l) = \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} \left[ \begin{aligned} & u(m, n) \mathbf{A}_{k,m}^M \mathbf{B}_{l,n}^N + u(m, N-1-n) \mathbf{A}_{k,m}^M \mathbf{B}_{l, N-1-n}^N + \\ & + u(M-1-m, n) \mathbf{A}_{k, M-1-m}^M \mathbf{B}_{l,n}^N + \\ & + u(M-1-m, N-1-n) \mathbf{A}_{k, M-1-m}^M \mathbf{B}_{l, N-1-n}^N \end{aligned} \right]$$

mediante el uso de (3.2.1.2), se tiene:

$$\tilde{v}(k, l) = \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} \left[ \begin{aligned} & \hat{\mathbf{p}}_{m,n} \mathbf{A}_{k,m}^M \mathbf{B}_{l,n}^N + \hat{\mathbf{q}}_{m,n} \mathbf{A}_{k,m}^M \mathbf{B}_{l, N-1-n}^N + \\ & + \hat{\mathbf{r}}_{m,n} \mathbf{A}_{k, M-1-m}^M \mathbf{B}_{l,n}^N + \hat{\mathbf{s}}_{m,n} \mathbf{A}_{k, M-1-m}^M \mathbf{B}_{l, N-1-n}^N \end{aligned} \right]$$

de acuerdo con (3.2.1.4), se llega a:

$$\tilde{\mathbf{P}}_{k,l} = \tilde{v}(2k, 2l) = \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} \left[ \begin{aligned} & \hat{\mathbf{p}}_{m,n} \mathbf{A}_{2k,m}^M \mathbf{B}_{2l,n}^N + \hat{\mathbf{q}}_{m,n} \mathbf{A}_{2k,m}^M \mathbf{B}_{2l, N-1-n}^N + \\ & + \hat{\mathbf{r}}_{m,n} \mathbf{A}_{2k, M-1-m}^M \mathbf{B}_{2l,n}^N + \hat{\mathbf{s}}_{m,n} \mathbf{A}_{2k, M-1-m}^M \mathbf{B}_{2l, N-1-n}^N \end{aligned} \right]$$

usando luego (3.2.1.6), (3.2.1.7) y (3.2.1.11), se obtiene:

$$\begin{aligned} \tilde{\mathbf{P}}_{k,l} &= \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} [\mathbf{A1}_{k,m}(\hat{\mathbf{p}}_{m,n} + \hat{\mathbf{q}}_{m,n} + \hat{\mathbf{r}}_{m,n} + \hat{\mathbf{s}}_{m,n})\mathbf{B1}_{l,n}] \\ \text{ssi } \tilde{\mathbf{P}} &= \mathbf{A1}(\hat{\mathbf{p}} + \hat{\mathbf{q}} + \hat{\mathbf{r}} + \hat{\mathbf{s}})(\mathbf{B1})^T \end{aligned} \quad (3.2.2.2)$$

$$\tilde{\mathbf{Q}}_{k,l} = \tilde{v}(2k,2l+1) = \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} \left[ \begin{array}{l} \hat{\mathbf{p}}_{m,n} \mathbf{A}_{2k,m}^M \mathbf{B}_{2l+1,n}^N + \hat{\mathbf{q}}_{m,n} \mathbf{A}_{2k,m}^M \mathbf{B}_{2l+1,N-1-n}^N + \\ + \hat{\mathbf{r}}_{m,n} \mathbf{A}_{2k,M-1-m}^M \mathbf{B}_{2l+1,n}^N + \\ + \hat{\mathbf{s}}_{m,n} \mathbf{A}_{2k,M-1-m}^M \mathbf{B}_{2l+1,N-1-n}^N \end{array} \right]$$

usando en este caso (3.2.1.6), (3.2.1.7), (3.2.1.8) y (3.2.1.4), se obtiene la siguiente expresión:

$$\begin{aligned} \tilde{\mathbf{Q}}_{k,l} &= \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} [\mathbf{A1}_{k,m}(\hat{\mathbf{p}}_{m,n} - \hat{\mathbf{q}}_{m,n} + \hat{\mathbf{r}}_{m,n} - \hat{\mathbf{s}}_{m,n})\mathbf{B2}_{l,n}] \\ \text{ssi } \tilde{\mathbf{Q}} &= \mathbf{A1}(\hat{\mathbf{p}} - \hat{\mathbf{q}} + \hat{\mathbf{r}} - \hat{\mathbf{s}})(\mathbf{B2})^T \end{aligned} \quad (3.2.2.3)$$

análogamente a como se obtuvo (3.2.2.2) y (3.2.2.3), se procede para encontrar (3.2.2.4) y (3.2.2.5), llegándose a:

$$\tilde{\mathbf{R}} = \mathbf{A2}(\hat{\mathbf{p}} + \hat{\mathbf{q}} - \hat{\mathbf{r}} - \hat{\mathbf{s}})(\mathbf{B1})^T \quad (3.2.2.4)$$

$$\tilde{\mathbf{S}} = \mathbf{A2}(\hat{\mathbf{p}} - \hat{\mathbf{q}} - \hat{\mathbf{r}} + \hat{\mathbf{s}})(\mathbf{B2})^T \quad (3.2.2.5)$$

luego se procede a escribir las ecuaciones (3.2.2.2)-(3.2.2.5) en forma matricial de la siguiente manera:

$$\begin{bmatrix} \hat{\mathbf{g}}1 \\ \hat{\mathbf{g}}2 \\ \hat{\mathbf{g}}3 \\ \hat{\mathbf{g}}4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{p}} \\ \hat{\mathbf{q}} \\ \hat{\mathbf{r}} \\ \hat{\mathbf{s}} \end{bmatrix} \quad (3.2.2.6)$$

en donde:

$$\begin{aligned} \tilde{\mathbf{P}} &= \mathbf{A1} \hat{\mathbf{g}}1(\mathbf{B1})^T \\ \tilde{\mathbf{Q}} &= \mathbf{A1} \hat{\mathbf{g}}2(\mathbf{B2})^T \\ \tilde{\mathbf{R}} &= \mathbf{A2} \hat{\mathbf{g}}3(\mathbf{B1})^T \\ \tilde{\mathbf{S}} &= \mathbf{A2} \hat{\mathbf{g}}4(\mathbf{B2})^T \end{aligned} \quad (3.2.2.7)$$

Desde este punto en adelante el objetivo es expresar el cálculo de  $\tilde{\mathbf{P}}, \tilde{\mathbf{Q}}, \tilde{\mathbf{R}}$  y  $\tilde{\mathbf{S}}$ , mediante el uso de la transformada coseno de orden inmediatamente inferior, esto es, de orden  $M/2 \times N/2$ .

Para ello tómesese en consideración desde (3.2.0.5), (3.2.0.6) y (3.2.1.6) que:

$$\mathbf{A1}_{k,m} = C_{2M}^{(2m+1)2k} = \cos\left(\frac{(2m+1)2k\pi}{2M}\right) = \cos\left(\frac{(2m+1)k\pi}{2(M/2)}\right) = C_{2(M/2)}^{(2m+1)k}$$

en consecuencia para las matrices  $\mathbf{A}$  y  $\mathbf{B}$  se cumple que:

$$\begin{aligned} \mathbf{A1}_{k,m} &= \mathbf{A}_{k,m}^{M/2} \\ \mathbf{B1}_{l,n} &= \mathbf{B}_{l,n}^{N/2} \end{aligned} \tag{3.2.2.8}$$

y, considerando la identidad trigonométrica:

$$2 \cos(a) \cos(b) = \cos(a - b) + \cos(a + b)$$

se deriva la siguiente relación:

$$\begin{aligned} 2C_{2M}^{(2m+1)(2k+1)} C_{2M}^{(2m+1)} &= C_{2M}^{(2m+1)2k} + C_{2M}^{(2m+1)2(k+1)} \\ &= C_{2(M/2)}^{(2m+1)k} + C_{2(M/2)}^{(2m+1)(k+1)} \end{aligned} \tag{3.2.2.9}$$

entonces, tomando en cuenta (3.2.2.8) se tendrá:

$$\mathbf{A2}(\mathbf{W1})^{-1} = \mathbf{A}^{M/2} + (\mathbf{A}^{M/2})^1 \tag{3.2.2.10}$$

$$\mathbf{B2}(\mathbf{W2})^{-1} = \mathbf{B}^{N/2} + (\mathbf{B}^{N/2})^1 \tag{3.2.2.11}$$

para ello  $\mathbf{W1}$  y  $\mathbf{W2}$  se definen como sigue, asumiendo siempre que:

$$k, m = 0, 1, \dots, M/2 - 1$$

$$l, n = 0, 1, \dots, N/2 - 1$$

$$(\mathbf{W1})^{-1} = \begin{bmatrix} 2C_{2M}^1 & & & \\ & 2C_{2M}^3 & & \\ & & \ddots & \\ & & & 2C_{2M}^{M-1} \end{bmatrix} \tag{3.2.2.12}$$

$$(\mathbf{W2})^{-1} = \begin{bmatrix} 2C_{2N}^1 & & & \\ & 2C_{2N}^3 & & \\ & & \ddots & \\ & & & 2C_{2N}^{N-1} \end{bmatrix} \quad (3.2.2.13)$$

adicionalmente,

$$\begin{aligned} (\mathbf{A}^{M/2})_{k,m}^1 &= \mathbf{A}_{k+1,m}^{M/2} \\ (\mathbf{B}^{N/2})_{l,n}^1 &= \mathbf{B}_{l+1,n}^{N/2} \end{aligned} \quad (3.2.2.14)$$

con esto se prueba que  $\mathbf{A2}$  y  $\mathbf{B2}$  puede expresarse respectivamente en función de  $\mathbf{A1}$  y  $\mathbf{B1}$ , es decir (3.2.0.7) puede calcularse partiendo de 4 transformadas coseno de orden inferior según (3.2.2.6), (3.2.2.7), (3.2.2.10) y (3.2.2.11) y cada una de estas a su vez en función de las 4 transformadas directas de orden  $M/4 \times N/4$  y así sucesivamente lográndose por tanto un método recursivo de cálculo de la transformada coseno directa de orden  $M \times N$ .

En (3.2.2.7) la primera ecuación, es decir la que calcula  $\mathbf{P}$ , ya está reducida, es decir usa (3.2.2.8),

$$\tilde{\mathbf{P}} = \mathbf{A1} \hat{\mathbf{g}}1(\mathbf{B1})^T = \mathbf{A1} \mathbf{g}1(\mathbf{B1})^T \quad (3.2.2.15)$$

seguidamente:

$$\tilde{\mathbf{Q}} = \mathbf{A1} \hat{\mathbf{g}}2(\mathbf{B2})^T$$

pero desde (3.2.2.8) y (3.2.2.11),

$$\mathbf{B2} = (\mathbf{B1} + (\mathbf{B1})^1) \mathbf{W2} \Rightarrow \mathbf{B2}^T = \mathbf{W2}(\mathbf{B1} + (\mathbf{B1})^1)^T$$

entonces,

$$\tilde{\mathbf{Q}} = \mathbf{A1} \hat{\mathbf{g}}2 \mathbf{W2}((\mathbf{B1})^T + ((\mathbf{B1})^1)^T) = \tilde{\mathbf{Q}}1 + \tilde{\mathbf{Q}}2 \quad (3.2.2.16)$$

en donde,

$$\begin{aligned} \tilde{\mathbf{Q}}1 &= \mathbf{A1} \mathbf{g}2(\mathbf{B1})^T \\ \mathbf{g}2 &= \hat{\mathbf{g}}2 \mathbf{W2} \end{aligned} \quad (3.2.2.17)$$

$$\tilde{\mathbf{Q}}_2_{k,l} = \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} \mathbf{A}1_{k,m} \mathbf{g}2_{m,n} \mathbf{B}1_{l+1,n} = \tilde{\mathbf{Q}}1_{k,l+1} \quad (3.2.2.18)$$

análogamente desde (3.2.2.7),

$$\tilde{\mathbf{R}} = \mathbf{A}2 \hat{\mathbf{g}}3(\mathbf{B}1)^T$$

pero desde (3.2.2.8) y (3.2.2.10),

$$\mathbf{A}2 = (\mathbf{A}1 + (\mathbf{A}1)^t) \mathbf{W}1$$

entonces,

$$\tilde{\mathbf{R}} = (\mathbf{A}1 + (\mathbf{A}1)^t) \mathbf{W}1 \hat{\mathbf{g}}3(\mathbf{B}1)^T = \tilde{\mathbf{R}}1 + \tilde{\mathbf{R}}2 \quad (3.2.2.19)$$

en donde,

$$\begin{aligned} \tilde{\mathbf{R}}1 &= \mathbf{A}1 \mathbf{g}3(\mathbf{B}1)^T \\ \mathbf{g}3 &= \mathbf{W}1 \hat{\mathbf{g}}3 \end{aligned} \quad (3.2.2.20)$$

$$\tilde{\mathbf{R}}2_{k,l} = \sum_{m=0}^{M/2-1} \sum_{n=0}^{N/2-1} \mathbf{A}1_{k+1,m} \mathbf{g}3_{m,n} \mathbf{B}1_{l,n} = \tilde{\mathbf{R}}1_{k+1,l} \quad (3.2.2.21)$$

finalmente desde (3.2.2.7), tomando como base (3.2.2.15) - (3.2.2.21), se concluye,

$$\tilde{\mathbf{S}} = \tilde{\mathbf{S}}1 + \tilde{\mathbf{S}}2 + \tilde{\mathbf{S}}3 + \tilde{\mathbf{S}}4 \quad (3.2.2.22)$$

en donde:

$$\mathbf{g}4 = \mathbf{W}1 \hat{\mathbf{g}}4 \mathbf{W}2 \quad (3.2.2.23)$$

$$\tilde{\mathbf{S}}1 = \mathbf{A}1 \mathbf{g}4(\mathbf{B}1)^T \quad (3.2.2.24)$$

$$\tilde{\mathbf{S}}2 = \tilde{\mathbf{S}}1_{k+1,l} \quad (3.2.2.25)$$

$$\tilde{\mathbf{S}}3_{k,l} = \tilde{\mathbf{S}}1_{k,l+1} \quad (3.2.2.26)$$

$$\tilde{\mathbf{S}}4_{k,l} = \tilde{\mathbf{S}}1_{k+1,l+1} \quad (3.2.2.27)$$

**NOTA.-** Se asumen nulos todos los datos fuera del bloque  $M/2 \times N/2$

En el siguiente numeral, con las mismas bases matemáticas de partición y reordenamiento matricial utilizadas en el desarrollo de la TDC directa se opera sobre la TDC inversa de manera que esta sea calculable aplicando 4 veces la transformada inversa de orden inmediatamente inferior.

### 3.2.3 TRANSFORMADA DISCRETA COSENO INVERSA

La ecuación (3.2.0.2) usando (3.2.0.4), (3.2.0.5) y (3.2.0.6) según se procede en [25] y [26] puede escribirse de la siguiente manera:

$$u(m,n) = \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \left[ \hat{V}_{2k,2l} \mathbf{A}_{2k,m}^M \mathbf{B}_{2l,n}^N + \hat{V}_{2k,2l+1} \mathbf{A}_{2k,m}^M \mathbf{B}_{2l+1,n}^N + \right. \\ \left. + \hat{V}_{2k+1,2l} \mathbf{A}_{2k+1,m}^M \mathbf{B}_{2l,n}^N + \hat{V}_{2k+1,2l+1} \mathbf{A}_{2k+1,m}^M \mathbf{B}_{2l+1,n}^N \right] \quad (3.2.3.1)$$

luego usando (3.2.1.4) y (3.2.1.6),

$$u(m,n) = \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \left[ \hat{P}_{k,l} \mathbf{A}_{1,k,m} \mathbf{B}_{1,l,n} + \hat{Q}_{k,l} \mathbf{A}_{1,k,m} \mathbf{B}_{2,l,n} + \right. \\ \left. + \hat{R}_{k,l} \mathbf{A}_{2,k,m} \mathbf{B}_{1,l,n} + \hat{S}_{k,l} \mathbf{A}_{2,k,m} \mathbf{B}_{2,l,n} \right]$$

a continuación, usando (3.2.1.2) y dentro de los rangos establecidos para estas igualdades, se tiene:

$$\hat{\mathbf{p}}_{m,n} = u(m,n) \Leftrightarrow \hat{\mathbf{p}} = \mathbf{A1}^T (\hat{\mathbf{P}}\mathbf{B1} + \hat{\mathbf{Q}}\mathbf{B2}) + \mathbf{A2}^T (\hat{\mathbf{R}}\mathbf{B1} + \hat{\mathbf{S}}\mathbf{B2}) \quad (3.2.3.2)$$

usando luego (3.2.1.7), (3.2.1.8) y (3.2.1.11):

$$\hat{\mathbf{q}}_{m,n} = u(m, N-1-n) = \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \left[ \hat{P}_{k,l} \mathbf{A}_{1,k,m} \mathbf{B}_{1,l,N-1-n} + \hat{Q}_{k,l} \mathbf{A}_{1,k,m} \mathbf{B}_{2,l,N-1-n} + \right. \\ \left. + \hat{R}_{k,l} \mathbf{A}_{2,k,m} \mathbf{B}_{1,l,N-1-n} + \hat{S}_{k,l} \mathbf{A}_{2,k,m} \mathbf{B}_{2,l,N-1-n} \right]$$

matricialmente esta ecuación tomará la siguiente forma:

$$\hat{\mathbf{q}} = \mathbf{A1}^T (\hat{\mathbf{P}}\mathbf{B1} - \hat{\mathbf{Q}}\mathbf{B2}) + \mathbf{A2}^T (\hat{\mathbf{R}}\mathbf{B1} - \hat{\mathbf{S}}\mathbf{B2}) \quad (3.2.3.3)$$

de manera similar:

$$\hat{\mathbf{r}}_{m,n} = u(M-1-m,n) = \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \left[ \hat{P}_{k,l} \mathbf{A}_{1,k,M-1-m} \mathbf{B}_{1,l,n} + \hat{Q}_{k,l} \mathbf{A}_{1,k,M-1-m} \mathbf{B}_{2,l,n} + \right. \\ \left. + \hat{R}_{k,l} \mathbf{A}_{2,k,M-1-m} \mathbf{B}_{1,l,n} + \hat{S}_{k,l} \mathbf{A}_{2,k,M-1-m} \mathbf{B}_{2,l,n} \right]$$

entonces,

$$\hat{r}_{m,n} = \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \left[ \begin{array}{l} \hat{P}_{k,l} A1_{k,m} B1_{l,n} + \hat{Q}_{k,l} A1_{k,m} B2_{l,n} + \\ + \hat{R}_{k,l} (-A2_{k,m}) B1_{l,n} + \hat{S}_{k,l} (-A2_{k,m}) B2_{l,n} \end{array} \right]$$

su forma matricial será:

$$\hat{r} = A1^T (\hat{P}B1 + \hat{Q}B2) - A2^T (\hat{R}B1 + \hat{S}B2) \tag{3.2.3.4}$$

finalmente de (3.2.3.1), (3.2.3.2), (3.2.3.3) y (3.2.3.4):

$$\hat{s} = A1^T (\hat{P}B1 - \hat{Q}B2) - A2^T (\hat{R}B1 - \hat{S}B2) \tag{3.2.3.5}$$

las ecuaciones (3.2.3.2), (3.2.3.3), (3.2.3.4) y (3.2.3.5) pueden expresarse de la siguiente manera:

$$\begin{bmatrix} \hat{p} \\ \hat{q} \\ \hat{r} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} g1 \\ g2 \\ g3 \\ g4 \end{bmatrix}$$

en donde:

$$\begin{aligned} g1 &= A1^T \hat{P}B1 \\ g2 &= A1^T \hat{Q}B2 \\ g3 &= A2^T \hat{R}B1 \\ g4 &= A2^T \hat{S}B2 \end{aligned} \tag{3.2.3.6}$$

El cálculo de **g1** puede realizarse desde la expresión dada en (3.2.3.6), pues está ya expresada en función de la transformada inversa de orden inmediatamente inferior; entonces solamente se operará sobre las otras tres expresiones, para ello teniendo siempre presente las relaciones dadas por (3.2.2.8) y (3.2.2.14), prosiguiendo con la ecuación para **g2**:

$$g2 = A1^T \hat{Q}B2 \tag{3.2.3.7}$$

para aplicar (3.2.2.11), se procede desde (3.2.3.7):

$$\hat{\mathbf{g}}_2 = \mathbf{g}_2(\mathbf{W}_2)^{-1} = \mathbf{A}_1^T \hat{\mathbf{Q}} \mathbf{B}_2(\mathbf{W}_2)^{-1} \quad (3.2.3.8)$$

luego,

$$\hat{\mathbf{g}}_2 = \mathbf{A}_1^T \hat{\mathbf{Q}}(\mathbf{B}_1 + (\mathbf{B}_1)^1) = \mathbf{A}_1^T \hat{\mathbf{G}}_2 \mathbf{B}_1 \quad (3.2.3.9)$$

expresando (3.2.3.9) en sumatoria y asumiendo siempre que todos los datos fuera del bloque  $M/2 \times N/2$  son nulos,

$$\begin{aligned} \hat{\mathbf{g}}_{2,m,n} &= \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \mathbf{A}_{1,k,m} \hat{\mathbf{V}}_{2k,2l+1} (\mathbf{B}_{1,l,n} + \mathbf{B}_{1,l+1,n}) \\ &= \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} (\mathbf{A}_{1,k,m} \hat{\mathbf{V}}_{2k,2l+1} \mathbf{B}_{1,l,n} + \mathbf{A}_{1,k,m} \hat{\mathbf{V}}_{2k,2l+1} \mathbf{B}_{1,l+1,n}) \end{aligned}$$

en el segundo sumando se hace el cambio de variable  $l+1=l'$ , para en un paso posterior hacer  $l'=l$ ,<sup>1</sup>

$$\begin{aligned} \hat{\mathbf{g}}_{2,m,n} &= \sum_k \sum_l \mathbf{A}_{1,k,m} \hat{\mathbf{V}}_{2k,2l+1} \mathbf{B}_{1,l,n} + \sum_k \sum_{l'=1}^{N/2} \mathbf{A}_{1,k,m} \hat{\mathbf{V}}_{2k,2l'-1} \mathbf{B}_{1,l',n} \\ &= \sum_k \sum_l \mathbf{A}_{1,k,m} \hat{\mathbf{V}}_{2k,2l+1} \mathbf{B}_{1,l,n} + \sum_k \sum_l \mathbf{A}_{1,k,m} \hat{\mathbf{V}}_{2k,2l-1} \mathbf{B}_{1,l,n} \\ &= \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \mathbf{A}_{1,k,m} (\hat{\mathbf{V}}_{2k,2l+1} + \hat{\mathbf{V}}_{2k,2l-1}) \mathbf{B}_{1,l,n} \end{aligned}$$

entonces en (3.2.3.9),

$$\hat{\mathbf{G}}_{2,k,l} = \hat{\mathbf{V}}_{2k,2l+1} + \hat{\mathbf{V}}_{2k,2l-1} = \hat{v}(2k,2l+1) + \hat{v}(2k,2l-1) \quad (3.2.3.10)$$

A continuación se procede con la ecuación para  $\mathbf{g}_3$  de (3.2.3.6) de la misma forma que se lo hizo en la segunda, multiplicando por  $\mathbf{W}_1$ ,

$$\hat{\mathbf{g}}_3 = (\mathbf{W}_1)^{-1} \mathbf{g}_3 = (\mathbf{W}_1)^{-1} \mathbf{A}_2^T \hat{\mathbf{R}} \mathbf{B}_1 \quad (3.2.3.11)$$

usando (3.2.2.8) y (3.2.2.10) en (3.2.3.11) se tiene,

$$\hat{\mathbf{g}}_3 = (\mathbf{A}_2(\mathbf{W}_1)^{-1})^T \hat{\mathbf{R}} \mathbf{B}_1$$

<sup>1</sup> En algunas expresiones se utiliza la siguiente notación:  $\sum_k = \sum_{k=0}^{M/2-1}$  y  $\sum_l = \sum_{l=0}^{N/2-1}$ .



$$\hat{\mathbf{g}}\mathbf{3} = (\mathbf{A}\mathbf{1} + (\mathbf{A}\mathbf{1})^l)^T \hat{\mathbf{R}}\mathbf{B}\mathbf{1} = \mathbf{A}\mathbf{1}^T \hat{\mathbf{G}}\mathbf{3}\mathbf{B}\mathbf{1} \quad (3.2.3.12)$$

expresando (3.2.3.12) en sumatoria y asumiendo nuevamente que todos los datos fuera del bloque  $M/2 \times N/2$  son nulos,

$$\hat{\mathbf{g}}\mathbf{3}_{m,n} = \sum_k \sum_l \mathbf{A}\mathbf{1}_{k,m} \hat{\mathbf{V}}_{2k+1,2l} \mathbf{B}\mathbf{1}_{l,n} + \sum_k \sum_l \mathbf{A}\mathbf{1}_{k+1,m} \hat{\mathbf{V}}_{2k+1,2l} \mathbf{B}\mathbf{1}_{l,n}$$

haciendo el cambio de variable  $k+1=k'$  en el segundo sumatorio, para en un paso posterior hacer  $k'=k$ , se tiene,

$$\begin{aligned} \hat{\mathbf{g}}\mathbf{3}_{m,n} &= \sum_k \sum_l \mathbf{A}\mathbf{1}_{k,m} \hat{\mathbf{V}}_{2k+1,2l} \mathbf{B}\mathbf{1}_{l,n} + \sum_l \sum_{k'=1}^{N/2} \mathbf{A}\mathbf{1}_{k',m} \hat{\mathbf{V}}_{2k'-1,2l} \mathbf{B}\mathbf{1}_{l,n} \\ &= \sum_k \sum_l \mathbf{A}\mathbf{1}_{k,m} \hat{\mathbf{V}}_{2k+1,2l} \mathbf{B}\mathbf{1}_{l,n} + \sum_k \sum_l \mathbf{A}\mathbf{1}_{k,m} \hat{\mathbf{V}}_{2k-1,2l} \mathbf{B}\mathbf{1}_{l,n} \\ &= \sum_{k=0}^{M/2-1} \sum_{l=0}^{N/2-1} \mathbf{A}\mathbf{1}_{k,m} (\hat{\mathbf{V}}_{2k+1,2l} + \hat{\mathbf{V}}_{2k-1,2l}) \mathbf{B}\mathbf{1}_{l,n} \end{aligned}$$

entonces en (3.2.3.12),

$$\hat{\mathbf{G}}\mathbf{3}_{k,l} = \hat{\mathbf{V}}_{2k+1,2l} + \hat{\mathbf{V}}_{2k-1,2l} = \hat{v}(2k+1,2l) + \hat{v}(2k-1,2l) \quad (3.2.3.13)$$

Como paso final, operando sobre la ecuación para  $\mathbf{g}\mathbf{4}$  de (3.2.3.6) y multiplicando por  $\mathbf{W}\mathbf{1}$  y  $\mathbf{W}\mathbf{2}$ , se tendrá:

$$\hat{\mathbf{g}}\mathbf{4} = (\mathbf{W}\mathbf{1})^{-1} \mathbf{g}\mathbf{4} (\mathbf{W}\mathbf{2})^{-1} = (\mathbf{W}\mathbf{1})^{-1} \mathbf{A}\mathbf{2}^T \hat{\mathbf{S}} \mathbf{B}\mathbf{2} (\mathbf{W}\mathbf{2})^{-1} \quad (3.2.3.14)$$

reemplazando las equivalencias dadas por (3.2.2.10) y (3.2.2.11):

$$\hat{\mathbf{g}}\mathbf{4} = (\mathbf{A}\mathbf{1} + (\mathbf{A}\mathbf{1})^l)^T \hat{\mathbf{S}} (\mathbf{B}\mathbf{1} + (\mathbf{B}\mathbf{1})^l) = \mathbf{A}\mathbf{1}^T \hat{\mathbf{G}}\mathbf{4}\mathbf{B}\mathbf{1} \quad (3.2.3.15)$$

desarrollando en forma similar a lo realizado para la obtención de (3.2.3.10) y (3.2.3.13) se obtiene para (3.2.3.15),

$$\hat{\mathbf{G}}\mathbf{4}_{k,l} = \hat{\mathbf{V}}_{2k+1,2l+1} + \hat{\mathbf{V}}_{2k+1,2l-1} + \hat{\mathbf{V}}_{2k-1,2l+1} + \hat{\mathbf{V}}_{2k-1,2l-1} \quad (3.2.3.16)$$

De esta manera una TDC de tamaño  $M \times N$  puede ser escrita como la combinación lineal de cuatro submatrices de tamaño  $M/2 \times N/2$  todas ellas escaladas en un factor  $W1$  y  $W2$ .

Haciendo recursivo este método de descomposición se concluye en el algoritmo para el cálculo de la Transformada Rápida Coseno.

El algoritmo de cálculo del par de transformadas coseno así descrito se ha implementado en el sistema de software. Los cálculos de la TDC en las dos direcciones (directa e inversa) que se realicen en las siguientes secciones de este trabajo de tesis, será en base a la aplicación del algoritmo aquí desarrollado.

### 3.3 TRANSFORMADA DISCRETA COSENO DE IMÁGENES MONOCROMÁTICAS

Como se ha mencionado varias veces la transformada discreta coseno, concentra gran parte de su energía en una pequeña región en el dominio de la frecuencia, cerca del origen y a lo largo de los ejes  $k$  y  $l$ . Además formas discontinuas como bordes contribuyen tanto a componentes de alta como de baja frecuencia. La concentración energética a lo largo de los ejes  $k$  y  $l$  se debe en parte a la ventana rectangular usada para obtener una imagen finita, creando discontinuidades en los cuatro lados [3]. La presencia de detalles verticales se manifiesta con componentes de AC a lo largo del borde vertical [36].

En las figuras siguientes se presentan ejemplos de imágenes monocromáticas de diferentes características (a), acompañadas de la imagen respectiva que representa su espectro en el dominio de la TDC (b).

Debido a la alta compresión energética que presenta la TDC no es posible visualizar el espectro<sup>1</sup> generado por el conjunto de elementos  $|v(k,l)|$ , por tal razón en su lugar se ha graficado la función:

$$v'(k,l) = K_1 \log(1 + K_2 |v(k,l)|) \quad K_1, K_2 \geq 0 \quad (3.3.0.1)$$

esta operación tiene la propiedad de comprimir las amplitudes grandes (componentes de baja frecuencia en imágenes típicas) mientras expande las amplitudes pequeñas haciendo posible mostrar claramente regiones de alta frecuencia. En dicha operación el valor de  $K_2$  es experimental y lo determinaran las necesidades visuales, produciendo mayor o menor relación entre máxima y mínima amplitudes. La constante  $K_1$  en cambio permite controlar la iluminación del píxel, adoptando para este caso un nivel máximo de 255 para la componente DC (primer coeficiente en el dominio de la frecuencia, ubicado en la esquina superior izquierda), es decir  $K_1$  se ha tomado de tal manera que  $v'(0,0) = 255$ , bajo esta condición<sup>2</sup>:

$$K_1 = \frac{255}{\log(1 + K_2 v(0,0))} \quad (3.3.0.2)$$

Note además que:

$$\begin{aligned} v'(k,l) = K_1 \log(1 + K_2 |v(k,l)|) &\geq 0 \quad \forall (k,l) \\ \text{si } |v(k,l)| = 0 &\Rightarrow v'(k,l) = 0 \end{aligned} \quad (3.3.0.3)$$

en resumen tanto  $v'(k,l)$  como  $|v(k,l)|$  monótonamente se comportan de la misma forma, lo único que se ha conseguido es atenuar las bajas frecuencias y resaltar las

<sup>1</sup> A la magnitud  $|v(k,l)|$  se conoce como el *espectro coseno* de  $u(m,n)$ , en cambio a la magnitud  $|v(k,l)|^2$  como el *espectro de potencia* o también como la *densidad espectral* de  $u(m,n)$ .

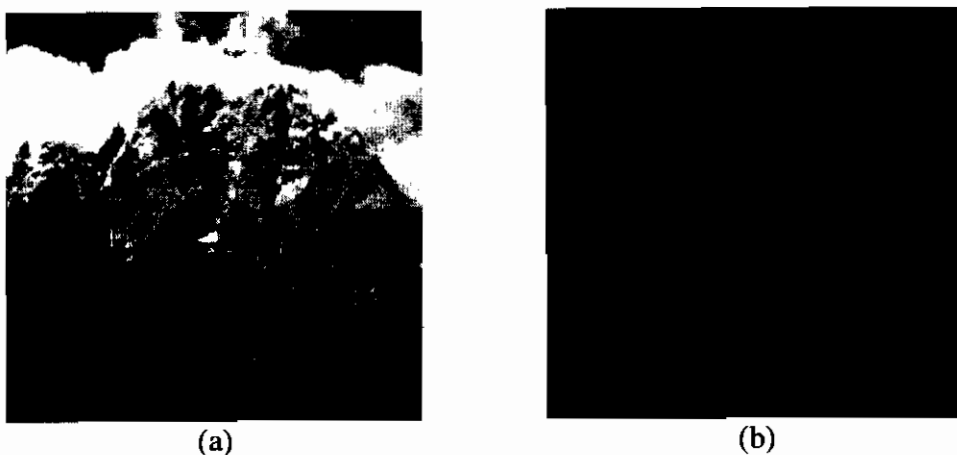
<sup>2</sup> Para imágenes típicas el valor más brillante del espectro (valor máximo de los coeficientes transformados) está dado por su componente DC, el factor de escala  $K_1$  hace corresponder el valor más brillante al nivel de gris de valor 255 (color blanco).

componentes de alta frecuencia pero solamente para efectos de visualización, en sí el espectro de la imagen no se verá afectado en absoluto .

Las tres primeras de las imágenes siguientes (figuras 3.3.0.1 - 3.3.0.3) son catalogadas como *típicas*, las mismas que tienen su energía concentrada cerca del origen en el dominio de la frecuencia; en aquellas imágenes presentadas a continuación que representan el espectro coseno, los puntos blancos indican a aquellos con mayor contenido energético, mientras que los puntos negros aquellos con mínima energía,



**Figura 3.3.0.1** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a), en donde un tono oscuro representa un coeficiente de valor mínimo y un tono claro representa un valor máximo para el coeficiente transformado, aplicando la operación (3.3.0.1).

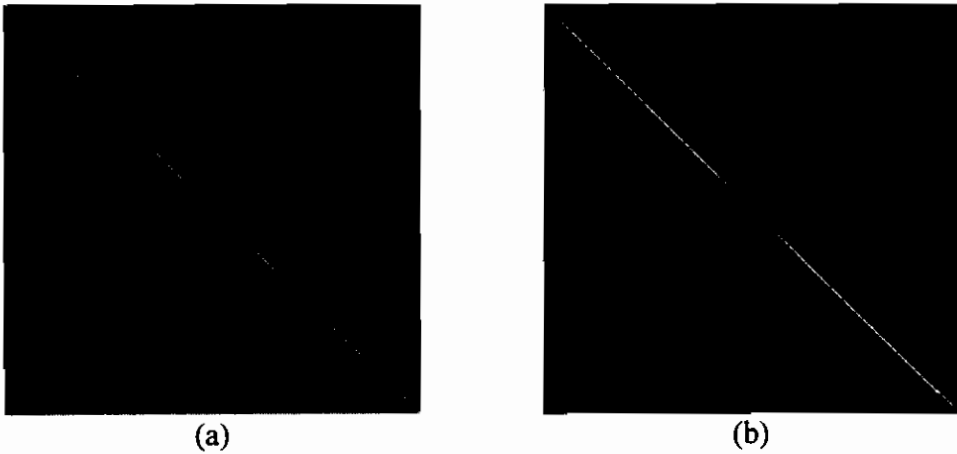


**Figura 3.3.0.2** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).



**Figura 3.3.0.3** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).

Las siguientes imágenes binarias (figuras 3.3.0.4 - 3.3.0.7) son no típicas, pues su energía no se concentra alrededor del origen del plano de la frecuencia espacial:

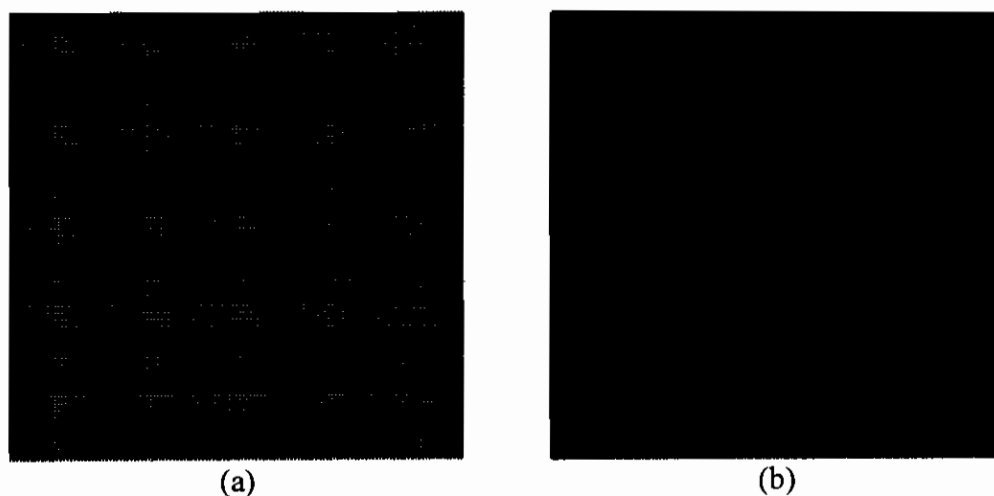


**Figura 3.3.0.4** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).

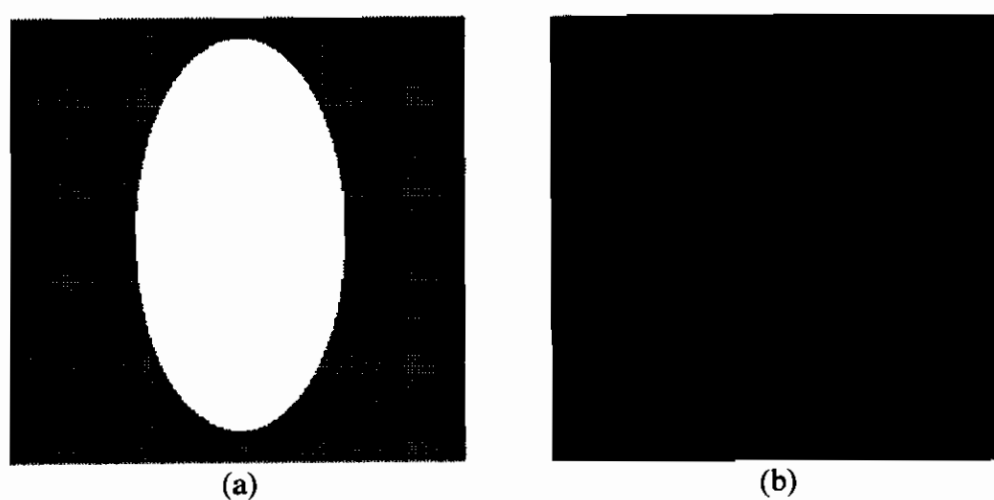


**Figura 3.3.0.5** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).

Las imágenes de las figuras No. 3.3.0.6a y 3.3.0.7a contienen componentes significativas de alta frecuencia en las dos direcciones; la primera está conformada por secuencias de puntos alternantes entre blanco y negro, en consecuencia contiene componentes insignificantes de baja frecuencia (si tiene componente DC). La segunda imagen en cambio si tiene una región suave lo que contribuye con contenido espectral de baja frecuencia.

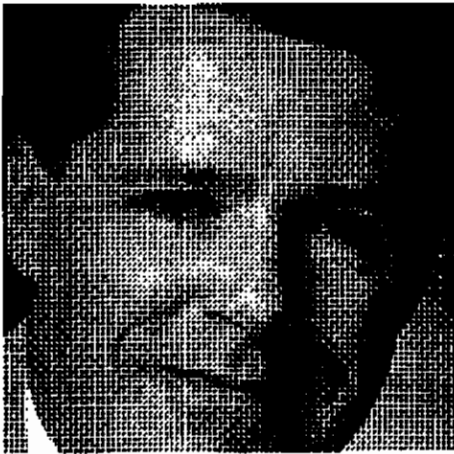


**Figura 3.3.0.6** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).

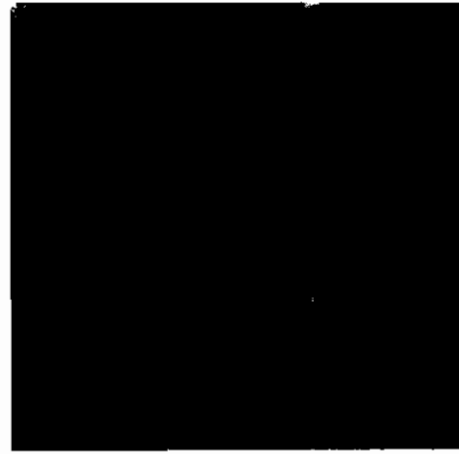


**Figura 3.3.0.7** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).

Las imágenes de las figuras No. 3.3.0.8a y 3.3.0.9a son imágenes binarias con características combinadas de los dos grupos anteriores (es decir imágenes típicas y aquellas que no lo son), en este caso los detalles suaves son conseguidos con variaciones lentas de la densidad de puntos negros ó blancos. La presencia de puntos ya sean blancos o negros dan a la imagen la característica de no típica.



(a)

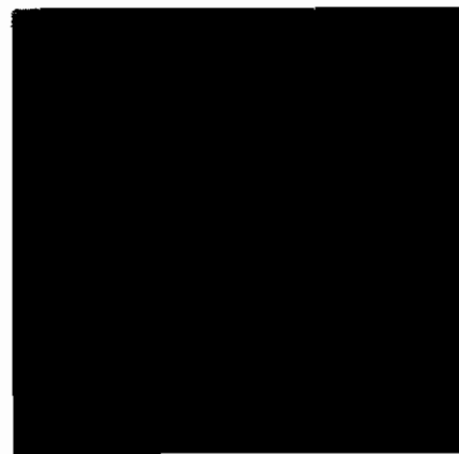


(b)

**Figura 3.3.0.8** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).



(a)



(b)

**Figura 3.3.0.9** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).

Finalmente las imágenes dadas por las figuras 3.3.0.10a y 3.3.0.11a son binarias típicas; binarias porque están descritas solamente por dos niveles de gris y típicas porque gran parte de su energía se concentra alrededor del origen en el dominio de la frecuencia:



**Figura 3.3.0.10** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).



**Figura 3.3.0.11** (a) Imagen en dominio espacial. (b) Imagen que representa el espectro de la imagen mostrada en (a).



# CAPÍTULO

---

# 4

## FILTRADO DE IMÁGENES EN EL DOMINIO DE LA TRANSFORMADA DISCRETA COSENO PARA SU MEJORAMIENTO

### 4.0 INTRODUCCIÓN

Este capítulo esta destinado a cubrir la teoría de mejoramiento de imágenes. Se inicia describiendo los conceptos que la sustentan, así como también algunas aplicaciones, posteriormente se estudian algunos filtros recomendados para el efecto en el dominio de la frecuencia y aplicados en este caso a la transformada discreta coseno. La teoría de cada una de las técnicas está acompañada de ejemplos que muestran el efecto de la aplicación de dichos filtros. El análisis comparativo de estos resultados se lo realiza en un capítulo posterior.

## 4.1 MEJORAMIENTO DE IMÁGENES

En mejoramiento de imágenes el objetivo es perfeccionar su calidad<sup>1</sup>, calidad degradada por la no existencia de un sistema que entregue una imagen perfecta.

Existen dos posibles caminos para mejoramiento de imágenes [28]. El primer método es el denominado *a priori*, el cual busca diseñar sistemas de imagen que minimicen las degradaciones, por ejemplo la inclusión de servomecanismos realimentados en los sensores constituye una solución para corregir el movimiento entre la cámara y el objeto y, disminuir las degradaciones en imágenes así obtenidas.

El segundo método para mejoramiento de imágenes se llamará *a posteriori*, para este caso no importa el como se ha generado la imagen. La variedad de estos métodos está limitada solamente por la imaginación de los inventores. Este capítulo se centrará en la descripción de algunos métodos a posteriori.

La figura No. 4.1.0.1 muestra un diagrama de bloques de un sistema de mejoramiento de imágenes,

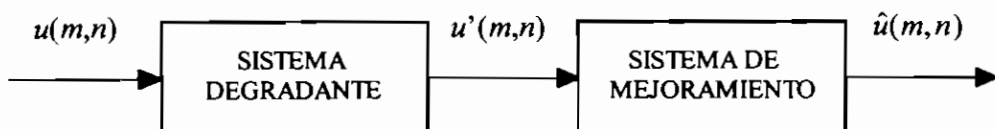


Figura 4.1.0.1 Sistema de mejoramiento de imágenes [28].

Se asume  $u(m,n)$  como la imagen ideal y es la que se obtendría si el sistema de imágenes fuera perfecto, pero como no lo es, se obtiene una imagen degradada  $u'(m,n)$ . El propósito del mejoramiento de imágenes es trabajar sobre la imagen degradada  $u'(m,n)$

<sup>1</sup> No existe una teoría unificada de mejoramiento de imágenes por la dificultad de definir un criterio que caracteriza la calidad de la imagen [2]. En muchas aplicaciones la determinación de la calidad de una imagen esta únicamente bajo responsabilidad del observador [28].

hasta obtener una imagen perfeccionada  $\hat{u}(m,n)$  la cual se acerca a la imagen ideal tanto como sea posible según el criterio de error cuadrático medio<sup>1</sup> [28].

El *mejoramiento* y la *restauración* de imágenes están estrechamente relacionadas. Cuando una imagen es degradada, la restauración de la original a menudo resulta en mejoramiento. Sin embargo hay una importante diferencia entre *restauración* y *mejoramiento*. En *restauración* si una imagen ideal ha sido degradada el objetivo es hacer que la imagen procesada se asemeje a la imagen original tanto como sea posible. En *mejoramiento* el objetivo es hacer a la imagen procesada en cierto sentido mejor que la imagen no procesada.

Para ilustrar la diferencia, note que una imagen original (imagen no degradada) no necesita ser restaurada, pero puede ser mejorada incrementando agudeza por medio de un filtro pasa-altos [3].

El mejoramiento de imágenes se refiere a la acentuación, agudizamiento de las facciones de la imagen tales como: filos, límites y contraste, buscando mejorar la apariencia visual o una forma más tratable para el análisis humano o computacional.

El proceso de mejoramiento no incrementa el contenido informático de los datos, pero si incrementa el rango dinámico<sup>2</sup> de las características escogidas para así detectarlas fácilmente.

---

<sup>1</sup> El problema del mejoramiento de imágenes a posteriori en determinadas situaciones se ve desde el punto de vista de restauración de imágenes, en este caso es aplicable el criterio del error cuadrático medio.

<sup>2</sup> Se define **Rango Dinámico** a la relación: máxima a mínima amplitud permisible de la señal en consideración, normalmente expresada en **db**.

El mejoramiento de imágenes incluye [1]:

- manipulación del nivel de gris y contraste,
- reducción del ruido
- agudizamiento de bordes
- filtrado
- interpolación y magnificación
- pseudocoloración, etc.

Como la naturaleza de las imperfecciones es variada existen por tanto diferentes técnicas de mejoramiento. Así una técnica que trabaja adecuadamente para una degradación determinada puede muy bien no ser de utilidad para otros tipos de degradaciones. Entonces un gran número de técnicas de mejoramiento son empíricas y requieren procesos iterativos para obtener resultados satisfactorios.

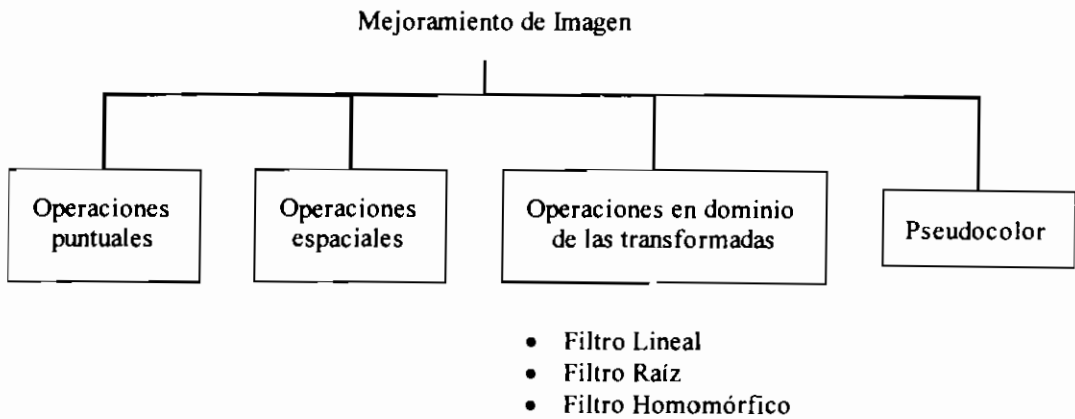
En una importante clase de problemas una imagen es mejorada modificando su contraste y/o rango dinámico. Por ejemplo una imagen aun cuando no esté degradada, a menudo aparecerá mejor cuando sus bordes sean afilados; además los bordes en un objeto contienen información muy importante de utilidad para aplicaciones de entendimiento de imágenes (identificación de objetos).

En otra clase de problemas de mejoramiento, una imagen degradada puede ser mejorada reduciendo la degradación (borrosidades, fondo ruidoso, manchas ruidosas y ruido de cuantización<sup>1</sup>) en este caso el mejoramiento se solapa con la restauración [3].

---

<sup>1</sup> En el proceso de **Cuantización**, muestras con amplitudes diferentes pero dentro del mismo intervalo de cuantización se codifican con la misma palabra código, esto generará en decodificación un ruido denominado de cuantización.

La figura 4.1.0.2 muestra algunas de las técnicas más comunes de mejoramiento de imágenes.



**Figura 4.1.0.2** Mejoramiento de imágenes [1].

Esta investigación se centra en el estudio de las técnicas de mejoramiento de imágenes en el dominio de la frecuencia espacial; según se muestra en la figura No 4.1.0.2, los filtros usados para el efecto son: el filtro lineal, filtro raíz y el filtro homomórfico. El dominio de la frecuencia a que se hace referencia en este caso es el dominio de la transformada discreta coseno.

## 4.2 FILTROS

Se define como *filtro digital* a un sistema capaz de filtrar señales muestreadas, cuyas muestras están representadas en forma binaria, realizando operaciones sobre dichas señales [4] (el objetivo final del diseño de filtros es alcanzar alguna respuesta de frecuencia específica).

Los filtros digitales multidimensionales han generado considerable investigación durante los últimos años por su potencial aplicación en: procesamiento de imágenes,

tratamiento de señales geosísmicas, análisis de señales de radar y de sonar, *mejoramiento de imágenes*, restauración de degradación lineal de imágenes, y procesamiento de imágenes biomédicas, etc [41], [43].

Los filtros pueden ser usados para el mejoramiento de imágenes degradadas por el ruido (ruido dependiente de la señal y ruido aditivo<sup>1</sup>). Las técnicas lineales poseen simplicidad matemática para remover ruido. Sin embargo, ruido dependiente de la señal y ruido impulsivo no es suprimido lo suficiente por filtros lineales, en tales casos se recurre a filtros no lineales.

Existen tres métodos comunes de implementación de filtros digitales [41]:

1. Convolucion directa<sup>2</sup>.
2. Técnicas en el dominio de la frecuencia, mediante el uso de transformadas discretas.
3. Implementación mediante métodos recursivos.

Las técnicas recursivas son las más eficientes en términos de requerimientos de memoria y rapidez de implementación, se destinan a satisfacer condiciones de magnitud y fase. Este trabajo de tesis analiza el comportamiento de filtros digitales implementados en el dominio de la frecuencia.

---

<sup>1</sup> El Apéndice D presenta un resumen sobre el tema de ruidos aditivos, entre ellos el ruido Gaussiano blanco, y los ruidos impulsivos; la teoría del ruido dependiente de la señal no se incluye, pues esta se cubre en el apartado que analiza el filtro homomórfico.

<sup>2</sup> La referencia [10] indica que las técnicas de procesamiento de imágenes en el dominio de la frecuencia son la aplicación del teorema de convolución.

## 4.2.1 OPERACIONES EN EL DOMINIO DE LAS TRANSFORMADAS DISCRETAS

Los filtros son usualmente clasificados de acuerdo a sus características en el dominio de la frecuencia como: filtro pasa-bajos, filtro pasa-altos y filtro pasa-banda.

Como se verá mas adelante en este capítulo cualquiera de los filtros (dependiendo de sus parámetros) aplicados en el dominio de la frecuencia se ubicarán en alguna de las clases mencionadas.

### **Filtro pasa-bajos**

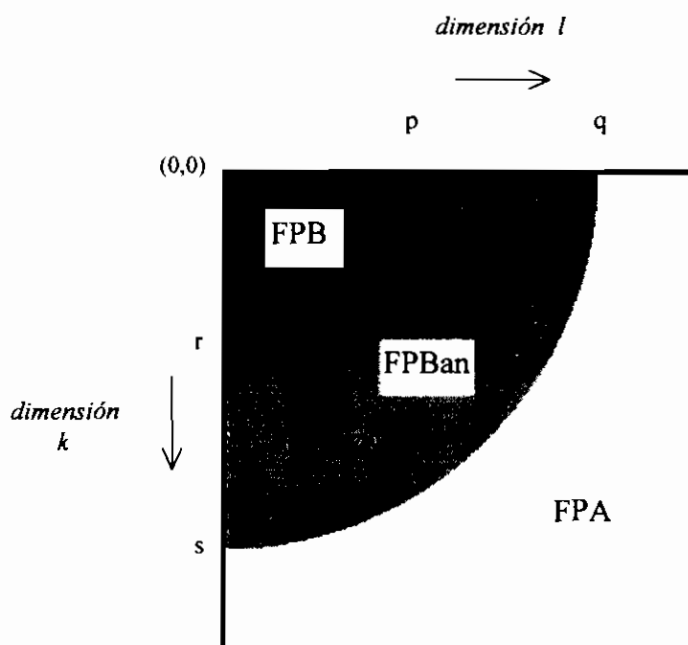
La energía de una imagen típica esta principalmente concentrada en sus componentes de baja frecuencia, esto es debido a la alta correlación espacial entre pixeles vecinos. El filtro pasa-bajos reduce componentes de alta frecuencia, mientras preserva componentes de baja frecuencia, este filtro reduce gran cantidad de ruido a expensas de reducir una pequeña cantidad de señal.

### **Filtro pasa-altos**

El filtro pasa-altos enfatiza las componentes de alta frecuencia de una señal, mientras reduce componentes de baja frecuencia. Los bordes y detalles finos de una imagen son los contribuidores primarios de las componentes de alta frecuencia de una imagen, entonces este tipo de filtro normalmente agudiza la imagen. Como el ruido de fondo tiene típicamente componentes significativos de alta frecuencia, el filtro pasa-altos incrementa el contraste local y también la potencia de ruido de fondo por lo tanto la imagen procesada aparece mas ruidosa que la no procesada [3].

El filtro pasa-banda agrupará características tanto del filtro pasa-bajos como del pasa-altos, es decir, incluirá en la imagen detalles suaves (baja frecuencia), así como también detalles finos (alta frecuencia).

La característica de respuesta de frecuencia de estos tipos de filtros se presenta en la figura 4.2.1.1. Como se muestra estos filtros ideales tienen una ganancia constante (tomada usualmente como ganancia unitaria).



**Figura 4.2.1.1** Máscara zonal  $g(k,l)$  para transformadas ortogonales, para filtro pasa-bajos FPB, filtro pasa-banda FPBan y filtro pasa-altos FPA. La función  $g(k,l)$  es cero fuera de la región de acción de cada filtro particular [1].

En las técnicas de mejoramiento mediante operaciones en el dominio de las transformadas discretas, se realizan operaciones de memoria nula<sup>1</sup> sobre la imagen transformada, seguida por la operación de la transformada inversa como se muestra en la figura 4.2.1.2.

<sup>1</sup> Operaciones de memoria nula significa que cada coeficiente transformado es procesado independiente sin considerar a los demás.



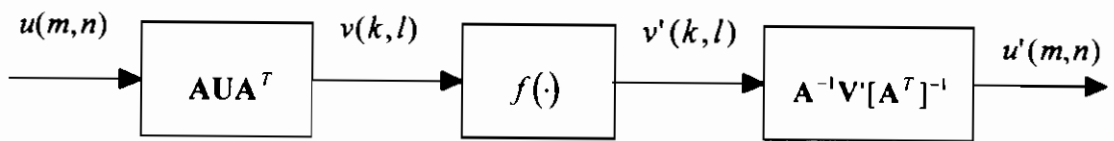


Figura 4.2.1.2 Mejoramamiento de imágenes por filtrado en el dominio de las transformadas [1].

La imagen transformada según lo descrito en el segundo capítulo es  $\mathbf{V} = \{v(k,l)\}$  y se calcula mediante la siguiente expresión:

$$\mathbf{V} = \mathbf{A}\mathbf{U}\mathbf{A}^T \quad (4.2.1.1)$$

donde  $\mathbf{U} = \{u(m,n)\}$  es la imagen de entrada. Entonces la transformada inversa de:

$$v'(k,l) = f(v(k,l)) = f(\mathbf{V}, k,l) \quad (4.2.1.2)$$

se calcula como:

$$\mathbf{U}' = \mathbf{A}^{-1}\mathbf{V}'[\mathbf{A}^T]^{-1} \quad (4.2.1.3)$$

y es la que proporciona la imagen mejorada, denotada  $\mathbf{U}' = \{u'(m,n)\}$ .

Los filtros digitales según la característica de linealidad de  $f$  respecto a  $\mathbf{V}$  pueden clasificarse en dos grandes grupos:

1.- Filtros lineales.- en este caso la función  $f$  cumple con la propiedad de linealidad<sup>1</sup> y para el caso de una imagen es una multiplicación pixel a pixel<sup>2</sup>, esto es;

$$v'(k,l) = g(k,l)v(k,l) \quad (4.2.1.4)$$

<sup>1</sup> Para probar que la ecuación (4.2.1.4) cumple con la propiedad de linealidad, debe probarse lo siguiente:

$$f(av_1(k,l) + bv_2(k,l)) = af(v_1(k,l)) + bf(v_2(k,l))$$

*Demostración:*

$$\begin{aligned} f(av_1(k,l) + bv_2(k,l)) &= g(k,l)[av_1(k,l) + bv_2(k,l)] \\ &= ag(k,l)v_1(k,l) + bg(k,l)v_2(k,l) \\ &= af(v_1(k,l)) + bf(v_2(k,l)) \end{aligned}$$

<sup>2</sup> En este caso específico la operación en cuestión se realiza individualmente sobre cada uno de los coeficientes transformados.

en donde  $g(k,l)$  es la llamada máscara zonal<sup>1</sup> mostrada en la figura 4.2.1.1. Constituyen filtros lineales los filtros: pasa-bajos, pasa-banda, pasa-altos y Gaussiano inverso.

2.- Filtros no lineales.- en los filtros no lineales en cambio la función  $f$  es no lineal, constituyen filtros no lineales el filtro raíz y el filtro homomórfico, que serán estudiados en las secciones 4.2.3 y 4.2.4 respectivamente.

## 4.2.2 FILTRO LINEAL GENERALIZADO

Es aquel filtro que resulta de la aplicación de la operación dada por la ecuación (4.2.1.4), la misma que según se probó es una función lineal de  $v(k,l)$ .

De acuerdo a la definición de la función  $g(k,l)$ , este filtro lineal puede clasificarse en: filtro pasa-bajos, filtro pasa-altos, filtro Gaussiano inverso.

### (a) Filtro pasa-bajos ideal

Este tipo de filtro reduce ruido aditivo, pero al mismo tiempo empaña o mancha la imagen, este efecto constituye entonces una limitación del filtro pasa bajos [3].

Se deberá seleccionar para este tipo de filtro la función  $g$  de la ecuación (4.2.1.4) de manera que  $v'(k,l)$  resulte de la atenuación de las componentes de alta frecuencia de  $v(k,l)$ , para satisfacer esta condición se recurre al filtro pasa bajos ideal bidimensional dado por la siguiente función de transferencia [10]:

$$g(k,l) = \begin{cases} 1 & \text{si } d(k,l) \leq D \\ 0 & \text{en otro caso} \end{cases} \quad (4.2.2.1)$$

<sup>1</sup> A la función  $g(k,l)$  se la conoce también como *función de transferencia* del filtro [10].

en donde  $d(k,l)$  es la distancia desde el origen del plano en dominio de la frecuencia espacial al punto  $(k,l)$ . El punto de transición entre  $g(k,l) = 1$  y  $g(k,l) = 0$  se conoce como *frecuencia de corte* del filtro.

El conjunto formado por los puntos dados por la expresión (4.2.2.1) y según la gráfica 4.2.1.1 traza un cuarto de círculo cuyo perímetro está limitado por un conjunto de frecuencias de corte. Todas las frecuencias dentro del círculo son pasadas sin atenuación, mientras todas las frecuencias de su exterior son atenuadas completamente de lo cual se desprende el término *ideal*, filtro que puede ser simulado mediante un computador pero que es imposible realizarlo con componentes electrónicos. La figura 4.2.2.1 muestra una imagen monocromática (a) y su respectivo espectro en dominio de la transformada coseno (b); en cambio la figura 4.2.2.2a presenta una imagen acompañada de su espectro, resultante de la aplicación de un filtro pasa-bajos a la imagen de la figura anterior, en esta se puede notar el deterioro de los detalles de la imagen (eliminación de alta frecuencia) y un empañamiento de la misma, característico de la aplicación de dicho tipo de filtro.



**Figura 4.2.2.1** (a) imagen monocromática, (b) Espectro coseno de la imagen original dada en (a).



**Figura 4.2.2.2** (a) Imagen filtrada, resultante de aplicar el filtro máscara pasa-bajos a la original dada en la figura 4.2.2.1a, (b) Imagen que representa el espectro de la imagen filtrada dada en (a).

### (b) Filtro pasa-altos ideal

El filtraje pasa-altos consiste en atenuar el contenido energético de baja frecuencia sin atenuar la información de alta frecuencia, un filtro pasa altos ideal se define por medio de la siguiente función de transferencia:

$$g(k,l) = \begin{cases} 0 & \text{si } d(k,l) \leq D \\ 1 & \text{en otro caso} \end{cases} \quad (4.2.2.2)$$

tanto en la expresión (4.2.2.1) como en la (4.2.2.2)  $d(k,l)$  se define como la distancia de corte medida desde el origen del plano en dominio de la frecuencia al punto  $(k,l)$  y se calcula de la siguiente manera [10]:

$$d(k,l) = \sqrt{k^2 + l^2} \quad (4.2.2.3)$$

Este filtro es contrario al filtro pasa bajos ideal, puesto que atenúa completamente todas las frecuencias dentro del un cuarto de círculo de radio  $D$ , permitiendo el paso sin atenuar de las frecuencias localizadas fuera del círculo. La figura 4.2.2.3a muestra la imagen dada por la figura 4.2.2.1a luego de la aplicación de un filtro pasa-altos ideal

como se comprueba observando su respectivo espectro, en este caso en cambio se nota un realce de los perfiles, característica que muestra la presencia mayoritaria de componentes de alta frecuencia.



**Figura 4.2.2.3** (a) Imagen original dada en la figura 4.2.2.1a filtrada con filtro máscara pasa-altos en el dominio de la frecuencia, (b) Imagen que representa el espectro resultante de la imagen filtrada.

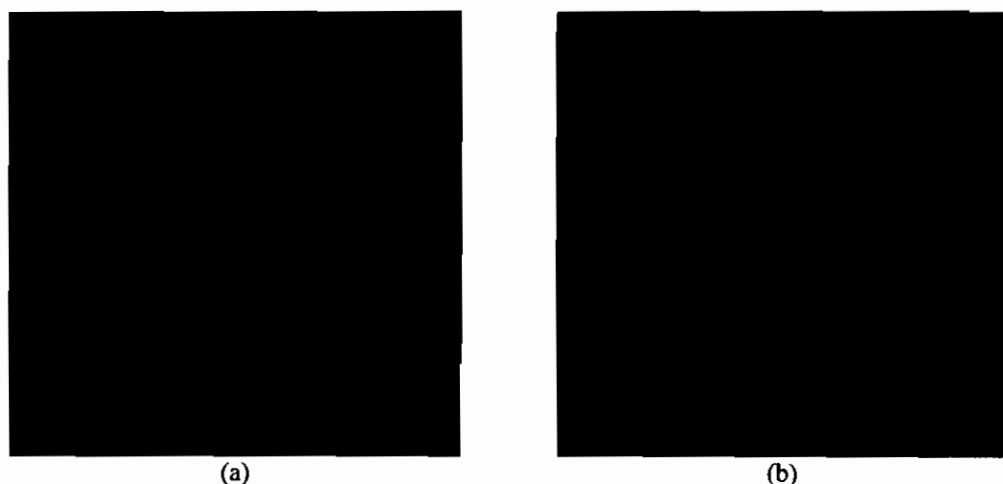
### (c) Filtro Gaussiano inverso

El filtro lineal generalizado de especial interés es el llamado *filtro Gaussiano inverso*, cuya función de transferencia o máscara zonal  $g(k,l)$  de la expresión (4.2.1.4) está dada para una imagen de tamaño  $N \times N$  mediante la siguiente expresión [1]:

$$g(k,l) = \exp\left\{\frac{k^2 + l^2}{2\sigma^2}\right\}, \quad 0 \leq k, l \leq N-1 \quad (4.2.2.4)$$

en donde  $\sigma^2 = r_v(0,0)$  es la varianza del conjunto formado por los coeficientes transformados  $v(k,l)$  cuyo cálculo se realiza usando la expresión (1.5.2.24).

Note que la función  $g$  es exponencialmente creciente con  $k$  y  $l$ , lo que equivale a decir que enfatiza las componentes de alta frecuencia permitiendo por ejemplo la restauración de imágenes empañadas por turbulencia atmosférica u otros fenómenos que pueden ser modelados por funciones espectrales Gaussianas [1].



**Figura 4.2.2.4** (a) Imagen filtrada aplicando filtro Gaussiano inverso a la imagen original, figura 4.2.2.1a, (b) Espectro de la imagen filtrada dada en (a).

### 4.2.3 FILTRO RAÍZ

Los coeficientes transformados (coeficientes en dominio de la frecuencia) pueden ser escritos usando notación polar, es decir en función de su magnitud y fase, de la siguiente manera:

$$v(k, l) = |v(k, l)| \exp\{j\theta(k, l)\} \quad (4.2.3.1)$$

El filtro no-lineal denominado *filtro raíz* se define como:

$$v'(k, l) = |v(k, l)|^\alpha \exp\{j\theta(k, l)\} \quad \alpha \geq 0 \quad (4.2.3.2)$$

relacionando las ecuaciones (4.2.1.2) con (4.2.3.2) se concluye que:

$$f\{v(k, l)\} = |v(k, l)|^\alpha \exp\{j\theta(k, l)\} \quad (4.2.3.3)$$

de donde se afirma que  $f$  es una función no lineal<sup>1</sup> de  $v$ .

<sup>1</sup> El filtro raíz es un filtro no lineal porque:  $|v_1(k, l) + v_2(k, l)|^\alpha \neq |v_1(k, l)|^\alpha + |v_2(k, l)|^\alpha$

La referencia [2] establece que,  $0 \leq \alpha \leq 2$ . La referencia [1] manifiesta que para valores de  $\alpha$  menores que la unidad, el filtro actúa como un filtro pasa-altos, puesto que la magnitud de  $v(k,l)$  es relativamente pequeña para componentes de alta frecuencia<sup>1</sup>, el efecto es por tanto mejorar frecuencias espaciales altas respecto a frecuencias espaciales bajas.

La aplicación de un filtro raíz ( $0 < \alpha < 1$ ) resulta entonces en la compresión de amplitudes grandes (baja frecuencia) mientras expande amplitudes pequeñas (alta frecuencia), por tanto  $v'(k,l)$  muestra más claramente regiones de alta frecuencia [1] y [3]. Para valores de  $\alpha$  mayores que la unidad amplitudes grandes (componentes de baja frecuencia) son amplificadas respecto a amplitudes bajas (alta frecuencia) dando un efecto de filtrado pasa bajos para imágenes típicas [1].

Con la finalidad de aclarar las afirmaciones realizadas, se procede de la siguiente forma: sea  $v_b$  una componente cualquiera de baja frecuencia y sea en cambio  $v_a$  una componente cualquiera de alta frecuencia, entonces por la propiedad de compresión energética aplicada a imágenes típicas, se tiene:

$$v_b \geq v_a \quad (4.2.3.4)$$

por la monotonía de la función exponencial es posible afirmar lo siguiente:

$$\frac{v_b}{v_a} \geq 1 \quad \wedge \quad \left( \frac{v_b}{v_a} \right)^\alpha \geq 1 \quad (4.2.3.5)$$

y

$$\text{si } \alpha \leq 1 \Rightarrow \frac{v_b}{v_a} \geq \left( \frac{v_b}{v_a} \right)^\alpha \quad (4.2.3.6)$$

---

<sup>1</sup> Téngase presente que la afirmación de que la magnitud de las componentes de alta frecuencia es relativamente pequeña con respecto a la magnitud de las componentes de baja frecuencia es válida únicamente para el caso de imágenes típicas.

reubicando términos se concluye en la siguiente expresión:

$$\Rightarrow \frac{v_a^\alpha}{v_a} \geq \frac{v_b^\alpha}{v_b} \quad (4.2.3.7)$$

la misma que establece que para coeficientes  $\alpha$  menores que la unidad, el grado de amplificación de un coeficiente respecto al coeficiente original (es decir coeficiente no filtrado) es mayor para componentes de alta frecuencia con respecto a la relación coeficiente filtrado a coeficiente no filtrado de las componentes de baja frecuencia.

En cambio:

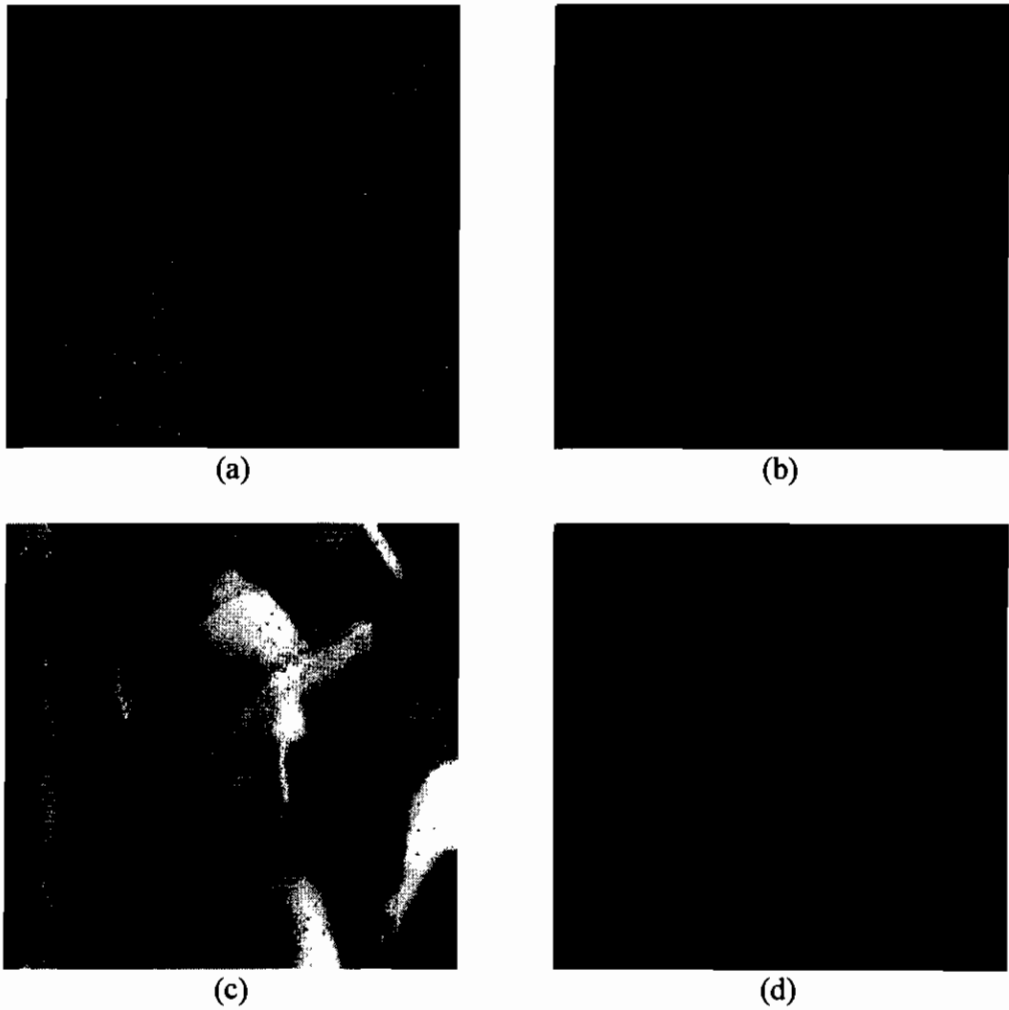
$$\text{si } \alpha \geq 1 \Rightarrow \left( \frac{v_b}{v_a} \right)^\alpha \geq \frac{v_b}{v_a} \quad (4.2.3.8)$$

luego para coeficientes  $\alpha$  mayores que la unidad las componentes de baja frecuencia se ven amplificadas con respecto a las de alta frecuencia, según lo muestra la siguiente expresión:

$$\Rightarrow \frac{v_b^\alpha}{v_b} \geq \frac{v_a^\alpha}{v_a} \quad (4.2.3.9)$$

En conclusión, para imágenes típicas solamente: el filtro raíz con coeficiente  $\alpha$  menor que la unidad actúa como filtro pasa-altos, en cambio el filtro raíz con coeficiente  $\alpha$  mayor que la unidad actúa como filtro pasa-bajos, ésta afirmación se puede comprobar analizando los ejemplos siguientes.





**Figura 4.2.3.1** Imagen resultante de la aplicación del filtro raíz con: (a)  $\alpha=0.5$ , (c)  $\alpha=1.3$  sobre la imagen de la figura 4.2.2.1a. (b) y (d) Espectros coseno de las imágenes dadas en (a) y (c) respectivamente.

#### 4.2.4 FILTRO HOMOMÓRFICO

Los filtros lineales no son capaces de remover ruido multiplicativo y ruido dependiente de la señal, pues éstas señales se combinan de una manera relativamente complicada con la imagen, estos casos se presentan en aplicaciones de procesamiento de señales sísmicas, procesamiento de voz, restauración digital de imágenes, con combinaciones de señales de una manera no lineal. Las clásicas técnicas lineales de procesamiento digital

no son de mucha utilidad porque las expresiones que describen dichas señales no cumplen la propiedad de linealidad, frente a este problema surge la necesidad de convertir no linealidades en linealidades y esto se consigue mediante la aplicación de operaciones no lineales (principalmente logarítmicas), éstos son los llamados filtros homomórficos.

### **Filtro homomórfico en mejoramiento de imágenes**

Una imagen  $u(m,n)$  de tamaño  $M \times N$  puede ser expresada en función de su iluminación  $e(m,n)$  y su reflectancia  $r(m,n)$  en términos multiplicativos<sup>1</sup> de la siguiente manera:

$$u(m,n) = e(m,n) r(m,n) \quad (4.2.4.1)$$

si la iluminación es constante  $e(m,n) = e$  sobre una escena entera, la imagen obtenida  $u(m,n)$  se relacionan linealmente con la reflectancia. Sin embargo la iluminación no es controlable especialmente en escenas externas, pudiendo describirla como un proceso aleatorio de variación lenta, su espectro usa entonces rangos de baja frecuencia, al contrario, la componente de reflectancia ocupa componentes de alta frecuencia [5].

Las técnicas lineales de filtraje no son aplicables para remover efectos de iluminación puesto que está inmerso en un proceso multiplicativo. Pero la función logarítmica puede ser usada para convertir ruido multiplicativo en ruido aditivo, esto es:

$$\ln[u(m,n)] = \ln[e(m,n)] + \ln[r(m,n)] \quad (4.2.4.2)$$

El mejoramiento de imágenes usando esta técnica se resume en la figura No. 4.2.4.1,

---

<sup>1</sup> En este modelo de imagen, una parte es la cantidad de luz disponible para la iluminación de objetos y la otra es la habilidad de estos objetos de reflejar la luz (reflectancia) [48].

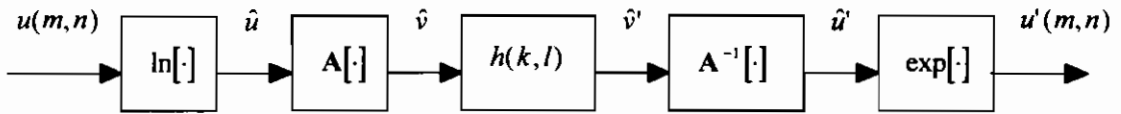


Figura 4.2.4.1.- Sistema homomórfico para mejoramiento de imágenes [5].

La respuesta de frecuencia  $h(k, l)$  del filtro está dada por la figura No. 4.2.4.2,

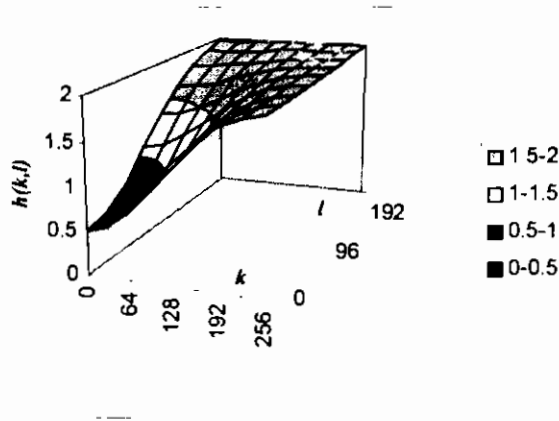


Figura 4.2.4.2.- Respuesta de frecuencia de la parte lineal del filtro homomórfico [3], [5] y [50].

Como se dijo anteriormente el espectro de iluminación  $E(k, l) = A[\ln[e(m, n)]]$  ocupa la región de bajas frecuencias, mientras que el espectro de reflectancia  $R(k, l) = A[\ln[r(m, n)]]$  ocupa las regiones de alta frecuencia, ambos espectros se asumen espacialmente isotrópicos.

La respuesta de frecuencia del filtro deberá ser más grande que la unidad y más pequeña que la unidad en la región de altas y bajas frecuencias respectivamente, pudiendo tomar la forma descrita por la siguiente ecuación<sup>1</sup>:

$$h(k, l) = 0.5 + 1.5 \left[ 1 - \exp \left[ \frac{-3.5(k^2 + l^2)}{MN} \right] \right], \quad 0 \leq k \leq M-1, \quad 0 \leq l \leq N-1 \quad (4.2.4.3)$$

<sup>1</sup> La ecuación (4.2.4.3) se obtuvo experimentalmente, tratando de alcanzar la respuesta de frecuencia mostrada en la figura No 4.2.4.2 y verificando su comportamiento pasa-altos mediante pruebas con imágenes monocromáticas no ruidosas.

de manera que para baja frecuencia  $h(k,l) = 0.5$  y para altas frecuencias  $h(k,l) = 2$ , la aplicación de esta operación concluye en la imagen mejorada aproximadamente de la siguiente manera,  $u'(m,n) = [e(m,n)]^{0.5} [r(m,n)]^2$ , atenuando la iluminación y resaltando la reflectancia [5].

La aplicación del filtro homomórfico persigue encontrar una imagen resultante con atractivo visual, lo cual se consigue mejorando el contraste mediante la aplicación de una característica pasa-altos según se solicita en la figura No. 4.2.4.2 y se la proporciona mediante la función de transferencia dada por la ecuación 4.2.4.3. La figura No. 4.2.4.3 presenta un ejemplo en el cual se aplica este tipo de filtro, en el se puede apreciar que efectivamente se resalta las componentes de alta frecuencia, componentes de frecuencia dadas por los contornos.



Figura 4.2.4.3.- (a) Imagen original, (b) Imagen resultante de aplicar filtro homomórfico.

#### **Filtrado homomórfico de ruido dependiente de la señal**

El ruido dependiente de la señal normalmente se encuentra en formación de imágenes, los tipos especiales encontrados son: el ruido de gránulo de película y el fotoelectrónico,

el primero es un proceso Gaussiano debido a los gránulos de plata que se precipitan durante la exposición de la película, el segundo también es un proceso Gaussiano pero producido por la fluctuación electrónica del número de fotones en la superficie sensible del detector.

Una representación útil del ruido dependiente de la señal está determinada por la siguiente relación [5]:

$$u_r = t(u) + nr(u) \quad (4.2.4.4)$$

en la cual:

- $u$  denota la imagen original,
- $t(\cdot), r(\cdot)$  son funciones 2-D no lineales.
- $n$  es un proceso Gaussiano, no correlacionado con la señal  $u$ , con valor medio *cero* y *varianza unitaria*, y
- $u_r$  es la imagen observada.

El ruido multiplicativo se consigue haciendo  $t(u) = cr(u)$  en (4.2.4.4) obteniéndose [5]:

$$u_r = cr(u) + r(u)n \quad c = \text{constante} \quad (4.2.4.5)$$

si en cambio,  $t(u) = c_1 u^l$ ,  $r(u) = c_2 u^m$  en (4.2.4.4) se obtiene el ruido fotoelectrónico. dado por la siguiente expresión [5]:

$$u_r = c_1 u^l + c_2 u^m n \quad (4.2.4.6)$$

El filtro homomórfico ha sido usado para remover ruido multiplicativo y de gránulo de película, tales técnicas tratan de desacoplar el ruido desde la señal para obtener ruido aditivo. Este desacoplamiento puede obtenerse pasando la imagen observada  $u_r$  de (4.2.4.4) a través de una no linealidad  $g(x)$  del siguiente modo [49]:

$$g(u_r) = g[t(u) + nr(u)] = u(u) + N(n, u) \quad (4.2.4.7)$$

en dicha expresión  $u(u)$  es una función no lineal de la imagen  $u$  y  $N(n, u)$  es un término de ruido dependiente de la señal  $u$ , la expansión en serie de Taylor de la función  $g$  alrededor del punto  $t(u)$  es [5]:

$$g(u_r) = \sum_{k=0}^{\infty} \frac{g^{(k)}[t(u)]}{k!} r^k(u) n^k \quad (4.2.4.8)$$

en donde  $g^{(k)}(x)$  es la  $k$ -ésima derivada de  $g(x)$ . Igualando (4.2.4.7) a (4.2.4.8) se encuentra que:

$$\begin{aligned} u(u) &= g[t(u)] \\ N(n, u) &= \sum_{k=1}^{\infty} \frac{g^{(k)}[t(u)]}{k!} r^k(u) n^k \end{aligned} \quad (4.2.4.9)$$

La parte ruidosa de  $g(u_r)$  se hace independiente de  $u$ , si se satisface el siguiente grupo de ecuaciones [5], [49]:

$$\frac{g^{(k)}[t(u)]}{k!} r^k(u) = a_k, \quad \forall k \quad (4.2.4.10)$$

en donde  $a_k$  son constantes. Este conjunto de ecuaciones se satisface solamente si:

$$r(u) = b t(u) \quad b = \text{constante} \quad (4.2.4.11)$$

de (4.2.4.11) se encuentra que (4.2.4.4) se convierte en ruido multiplicativo dado por la expresión (4.2.4.5), o por:

$$u_r = t(u) + b t(u) n \quad (4.2.4.12)$$

De lo anterior, el *ruido multiplicativo* es el único que puede ser completamente transformado a ruido aditivo, en este caso la función no lineal que satisface (4.2.4.10) es

$g(x) = \ln(x)$ . En todos los demás casos, esto es,  $g(x) \neq \ln(x)$  el ruido dependiente de la señal no es completamente transformado a ruido aditivo.

Como se puede apreciar en (4.2.4.9)  $N(n, u)$  consiste de varios términos cuyas contribuciones decrecen al aumentar el orden. El término más importante es  $g^{(1)}[t(u)]r(u)n$ , el siguiente objetivo es buscar una función  $g$  que haga de éste un término independiente de la señal, por tanto:

$$g^{(1)}[t(u)]r(u) = 1 \tag{4.2.4.13}$$

Así mismo la función  $g$  deberá satisfacer (4.2.4.10), luego (4.2.4.9) se convierte en:

$$N(n, u) = n + \sum_{k=2}^{\infty} \frac{g^{(k)}[t(u)]}{k!} r^k(u) n^k = N(n) \tag{4.2.4.14}$$

despreciando los términos de orden superior de (4.2.4.14), la ecuación (4.2.4.7) usando (4.2.4.9) toma la forma:

$$g(u_r) = g[t(u)] + n \tag{4.2.4.15}$$

el término ruido aditivo  $n$  puede ser removido por una técnica lineal como un filtro lineal pasa-bajos, obteniéndose:

$$g(u_r) = g[t(u)] \tag{4.2.4.16}$$

para en un paso posterior aplicar otra función no lineal  $f$  tal que [5]:

$$f(g[t(u)]) = u \tag{4.2.4.17}$$

finalmente la estructura del filtro homomórfico para remover ruido dependiente de la señal toma la forma de la figura siguiente:

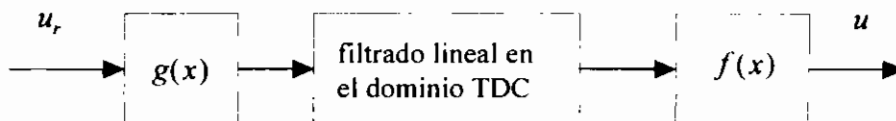


Figura 4.2.4.4.- Estructura del filtro homomórfico para remover ruido dependiente de la señal [5].

Para el caso del ruido dependiente de la señal dado por (4.2.4.6) las funciones no lineales están dadas por las siguientes expresiones [5]:

$$g(x) = \frac{l}{c_2(l-m)c_1^{-m/l}} x^{(l-m)/l} \quad (4.2.4.18)$$

$$f(x) = \left[ \frac{(l-m)c_2}{l c_1} x \right]^{l/(l-m)} \quad (4.2.4.19)$$

Las características de una cámara de TV comercial pueden muy bien ser descritas por la siguiente ecuación de una imagen contaminada con ruido dependiente de la señal:

$$u_r = u^{0.7} + u^{0.35} n \quad (4.2.4.20)$$

esta ecuación se encuentra haciendo  $l = 0.7$ ,  $m = 0.35$ ,  $c_1 = c_2 = 1$  en la ecuación (4.2.4.6). La figura No. 4.2.4.5 muestra que una imagen contaminada con este tipo de ruido, tiene su rango dinámico comprimido y el ruido es más notorio en regiones de alto brillo. Las no linealidades  $g$  y  $f$  se encuentran reemplazando respectivamente los valores de  $l, m, c_1$  y  $c_2$  señalados en las ecuaciones (4.2.4.18) y (4.2.4.19) [5].

A continuación se presenta un ejemplo de una imagen contaminada con ruido dependiente de la señal, y de ésta luego de haberla filtrado usando filtro homomórfico.



**Figura 4.2.4.5.-** (a) Imagen contaminada con ruido dependiente de la señal, (b) Imagen filtrada con filtro homomórfico.



# CAPÍTULO

---

# 5

## DESARROLLO DEL SISTEMA DE SOFTWARE

### 5.0 INTRODUCCIÓN

Este capítulo está constituido por dos partes, la primera encargada de señalar los requerimientos computacionales para la implementación de la aplicación, mientras la segunda comienza señalando brevemente las características más importantes del lenguaje de programación C que es el utilizado para la actividad en mención. Seguidamente se dan conceptos importantes para el desarrollo de aplicaciones ejecutables en entornos Windows, para concluir con la descripción de cada módulo en que está dividido el programa resultante.

### 5.1 REQUERIMIENTOS COMPUTACIONALES

En lo que a este punto se refiere debe hacerse una diferenciación entre el software y hardware requeridos, ya sea para la instalación del software de los utilitarios necesarios para la implementación del programa y aquellos para la instalación del software creado.

- 1.- Requisitos para la elaboración del Software:
- a) Microsoft Windows versión 3.0 o superior.
  - b) Compilador Turbo C++ versión 3.0 o superior.
  - c) Compilador de recursos versión 1.0 o superior.
  - d) Librería de clases ObjectWindows de Borland.
  - e) Controles propios de Borland.
  - f) Compilador de ayudas versión 3.1.

Los requerimientos marcados como b,c,d y e son proporcionados por Borland, los mismos que corren sobre la familia de computadores 286 o superior con al menos 2 Mbytes de memoria RAM y con monitor compatible con Windows (EGA o superior).

2.- Requisitos para la operación del software implementado

Para que el software desarrollado trabaje adecuadamente se sugiere el siguiente perfil:

- a) Microsoft Windows 95 o superior.
- b)  $16MN$  bytes de memoria RAM<sup>1</sup>.
- c)  $24MN$  bytes de disco fijo.
- d) Monitor compatible con Windows (SVGA o superior).
- e) Computadora 80486.

Los literales b y c muestran las capacidades mínimas de memoria y disco que la aplicación usará mientras ésta se encuentre en ejecución.

---

<sup>1</sup> En esta y la siguiente expresión  $M$  representa el alto y  $N$  el ancho de la imagen, dimensionadas en pixeles.

## 5.2 IMPLEMENTACIÓN DEL SOFTWARE

### 5.2.1 EL LENGUAJE C

Con frecuencia se denomina al lenguaje C como un lenguaje de computador de nivel medio sin que esto signifique que sea de menor potencia, de difícil uso o no tan evolucionado como lenguajes de alto nivel tales como Basic o Pascal. C se presenta como lenguaje de nivel medio porque combina características de lenguajes de alto nivel con la funcionalidad del lenguaje ensamblador, permitiendo la manipulación de bits, bytes y direcciones, elementos básicos con los que funciona la computadora, además de que el código C es muy portable<sup>1</sup>.

El término lenguaje estructurado no es aplicable estrictamente al lenguaje C mas bien se lo califica como lenguaje sencillamente estructurado. Técnicamente un lenguaje estructurado permite declarar funciones dentro de funciones, como C no permite ésta particularidad no puede ser calificado como tal.

Los términos que a continuación se definen se usa frecuentemente dentro de la teoría de programación:

- **Código Fuente.**- El código de un programa que un usuario puede leer, constituye el código de entrada al compilador de C.
- **Código Objeto.**- es el código de máquina resultante de traducir el código fuente del programa y es aquel que la computadora ejecuta directamente, es la entrada al enlazador.

---

<sup>1</sup> La portabilidad se refiere a la posibilidad de adaptar software escrito para un tipo de computador en otro.

- **Enlazador.-** Es un programa que enlaza funciones que han sido compiladas separadamente para obtener un sólo programa, combinando las bibliotecas de C con el código escrito. La salida del enlazador es el programa ejecutable.
- **Biblioteca.-** Denominado así al archivo que contiene determinadas funciones que son utilizadas en los programas.
- **Compilador.-** Convierte el programa fuente en código objeto que puede ser ejecutado directamente por la computadora.

C es un lenguaje flexible y potente y ha sido usado para crear algunos de los productos más importantes de software de las dos últimas décadas, pero cuando un proyecto excede un cierto tamaño, C alcanza sus límites, dando paso a C++ con la programación orientada a objetos. Un objeto es simplemente una entidad lógica que contiene datos y un código que manipula estos datos [19].

### **Ventajas del C**

- **Tamaño pequeño.-** C tiene un menor número de reglas de sintaxis, con la posibilidad de escritura de compiladores de alta calidad de C con requerimientos pequeños de memoria. En realidad C tiene mas operadores y combinaciones de operadores que palabras claves.
- **Velocidad.-** Un programa producido por C tiende a ser muy eficiente, siendo un lenguaje cercano al hardware la velocidad de ejecución de un programa escrito en C se aproxima a la velocidad de sus equivalentes escritos en lenguaje ensamblador.
- **Lenguaje Estructurado.-** C incluye todas las estructuras de control que se esperaría en un lenguaje moderno, pues incluye lazos *for* y *while*, sentencias *if-else*, sentencias

*case*. También permite control de ámbito de datos con la presencia de variables locales y globales.

- Programación Modular.- C permite la compilación separada para enlaces posteriores, reduciendo así el tiempo de compilación cuando solamente ciertas partes del proyecto han cambiado durante el desarrollo.
- Lenguaje C y Ensamblador.- La compilación y enlace separados facilita la inclusión de rutinas de código de ensamblador en códigos escritos en C.
- Manejo de Bits.- Como el origen del C está ligado al Sistema Operativo UNIX, el lenguaje incluye operadores de manipulación de bits.
- Punteros.- Algunos lenguajes manejan punteros pero, C se destaca por su aritmética de punteros, esta capacidad de direccionar áreas específicas de memoria aumenta la velocidad de ejecución de un programa.

### **Inconvenientes del C**

- Ausencia de Tipos Estrictos.- Como C es un lenguaje poco tipificado, permite una gran flexibilidad para manipular datos. C permite asignar una variable “**int**” a una tipo “**char**” y viceversa, de ahí la necesidad de usar las variables adecuadamente para evitar efectos colaterales. Por ejemplo el operador “= ” puede aparecer varias veces en una expresión, expresión que podría no tener un valor definido. Restringiendo el uso de este y otros operadores se eliminarían los efectos secundarios, pero C perdería su potencia y su atractivo de lenguaje ensamblador de alto nivel [18].
- Ausencia de Verificación.- El costo de optimizar un compilador buscando velocidad y eficiencia es el apareamiento de efectos secundarios por la no verificación en tiempo de ejecución, por ejemplo el sistema no verifica si se ha excedido los límites de un arreglo.

## 5.2.2 PROGRAMACIÓN WINDOWS

Como es lógico la programación Windows debe estar orientada a objetos, por ejemplo una ventana (las ventanas son objetos rectangulares sobre la pantalla) recibe entradas desde el teclado o desde el ratón y muestra resultados en la pantalla, se tiene entonces un estado (los datos) y un comportamiento (las funciones miembro) ante las acciones del usuario, características básicas de un objeto [21].

La mayor parte de las operaciones realizadas por una aplicación Windows se implementan en base a llamadas a funciones predefinidas de Windows. Dichas funciones se encuentran en el Interfaz de Programación de Aplicaciones Windows **API** (Application Programming Interface).

Las aplicaciones que llaman a estas funciones ejecutan directamente Windows, las funciones ejecutables del **API** de Windows están ubicadas dentro de las denominadas bibliotecas de enlace dinámico **DLL** (Dynamic Link Library). Existen tres **DLLs** distintas contenidas en el **API** que son: **KERNEL.EXE**, **USER.EXE** Y **GDI.EXE**. Son las únicas funciones con extensión **.EXE** todas las demás tendrán la extensión **DLL** [20]:

- **KERNEL.EXE** Responsable de la administración de memoria, recursos e interacción con el Sistema Operativo.
- **USER.EXE** Tiene bajo su responsabilidad la administración de ventanas, entradas de usuario y comunicaciones.
- **GDI.EXE** Responsable de la presentación de gráficos en la pantalla e impresora.

Estas tres bibliotecas se denominan de enlace dinámico porque no son enlazadas a la aplicación sino hasta el instante de la ejecución, en contraste con el enlace estático en el

cual se inserta el código objeto de las funciones de biblioteca en el programa ejecutable resultante.

Una *Biblioteca de Importación* contiene información relativa a la ubicación de las funciones ejecutables incluidas en el API. La Biblioteca de Importación en Borland C++ se llama IMPORT.LIB.

### **Componentes de una Aplicación Windows**

Normalmente un proyecto de aplicación Windows está constituido por cuatro tipos de archivos fuente. Archivo de definición de módulo .DEF, un archivo de cabecera .H, un archivo de especificación de recursos .RC y un archivo fuente .CPP.

#### **Archivo de Definición de Módulo**

Contiene información técnica sobre la estructura del archivo ejecutable [22]. información utilizada por el enlazador.

#### **Archivo de cabecera (# include)**

Este archivo incluye: los prototipos de las funciones usadas por el programa principal e identificadores<sup>1</sup>. La siguiente tabla muestra el significado de prefijos de algunos identificadores [21].

#### **Archivo de Especificación de Recursos**

Describe aceleradores, elementos de menú, íconos, mapas de bits (bitmaps), cajas de diálogos, cadena de caracteres.

---

<sup>1</sup> Los identificadores son constantes definidas, su nombre se escribe siempre con letras mayúsculas, y debe ser significativo. Consta de dos partes: el prefijo formado por dos o tres letras que indica la categoría a la que pertenece el identificador, y la descripción, las dos partes están separadas por el símbolo subrayado ( \_ ).

PREFIJO	CATEGORÍA
ID	Identificador en general
CM	Identificador de un elemento del menú
CS	Estilo de la clase
IDI	Identificador de un Icono
IDC	Identificador de un cursor
WS	Estilo de la ventana
WM	Gestión de Windows
DT	Mostrar texto

**Tabla 5.1** Prefijos de identificadores Windows [21].

### **El Archivo Fuente CPP**

Contiene las funciones que constituyen el núcleo de la aplicación, está encargado de la gestión de los mensajes generados por la aplicación o por el usuario.

### **El Interfaz de dispositivos Gráficos GDI**

Cualquier aplicación que se ejecute bajo Windows puede llamar a las funciones GDI con propósitos de enviar gráficos a la pantalla, a la memoria, a la impresora, a un fichero o a otra aplicación.

### **Dispositivos**

El teclado, el ratón, el reloj se dicen dispositivos de entrada. En cambio la pantalla, la impresora, un modem y una unidad de disco son llamados dispositivos de salida. Por tanto los dispositivos son los lugares donde se realiza la entrada y la salida de la información

### **Contextos de Dispositivo**

Los contextos de dispositivo representan a un dispositivo de salida y su respectivo driver<sup>1</sup>.

<sup>1</sup> Un driver es un conjunto de rutinas ejecutables responsables directos de la gestión del hardware [22].



### **El Contexto de Visualización**

El contexto de visualización representa una ventana en la pantalla y es un tipo de contexto de dispositivo. Cuando se quiera dibujar sobre un dispositivo de salida debe obtenerse primero un handle<sup>1</sup>, handle que será incluido como primer parámetro en la llamada a cualquier función de la biblioteca GDI.

### **El Color en Windows**

Windows utiliza un tipo especial de datos de 32 bits denominado COLORREF para la representación de un color, los tres bytes menos significativos representan los colores básicos rojo, verde y azul, cada uno tomará valores desde 0 a 255, esta combinación proporciona cerca de 16 millones de colores. Cuando los tres colores valen cero se genera el color negro, cuando valen 255 se genera el color blanco y cuando toman el mismo valor se genera un nivel de gris, en consecuencia existen 256 niveles posibles de gris.

### **La Librería ObjectWindows**

Esta librería trata la mayor parte de los elementos Windows como objetos con conductas predefinidas, suministrando una estructura base para aplicaciones como: comportamiento de ventanas, gestión de mensajes, etc. Algo muy importante de esta librería es la respuesta automática de los mensajes de la aplicación recibidos desde Windows. En un programa tradicional C para responder a los mensajes debe usarse las sentencias **switch ... case** tomando difícil la lectura del código resultante. En cambio en ObjectWindows los mensajes se gestionan llamando a funciones miembro apropiadas

---

<sup>1</sup> Un handle es un valor numérico entero utilizado por Windows en la identificación de un objeto en la aplicación como por ejemplo una ventana, una región de memoria, un ícono, etc.

de una clase, sin embargo esto no evita la necesidad de utilizar muchas de las casi 600 funciones integrantes del API de Windows.

### **Ejecución de los Programas Bajo Windows**

Como al principio puede parecer paradójico, Windows ejecuta los programas a la vez que los programas ejecutan Windows. Windows se comporta como un árbitro en el que participan muchos programas llevando a cabo tres tareas que permitan su ejecución, [18] y [22].

- Carga y ejecuta el programa.
- Gestiona la memoria de la computadora, garantizando que cada aplicación disponga de la memoria necesaria.
- Administra las entradas de bajo nivel realizadas por la aplicación, proporcionando servicios de entrada y salida para aislar la aplicación físicamente de la máquina.

La gráfica 5.2.2.1 esquemáticamente presenta la forma de interacción entre Windows y DOS.

### **Proyectos**

Un archivo proyecto .PRJ contiene la información necesaria sobre los diversos archivos que conforman el proyecto, así como también las opciones de compilación establecidas.

Los archivos que conforman una aplicación Windows son: los archivos fuente con extensión .CPP y el archivo de definición de módulo (con extensión .DEF); dado que ObjectWindows proporciona un archivo de definición de módulo por defecto denominado OWL.DEF no habrá necesidad de definir uno propio.

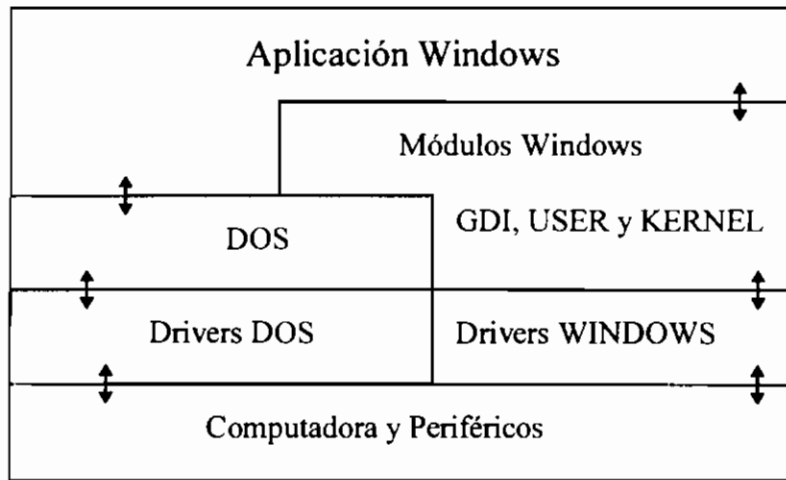


Figura 5.2.2.1 Interacción Windows-DOS [20].

La figura 5.2.2.2 describe gráficamente el proceso de compilación y enlazamiento de los distintos componentes que conforman el proyecto de la aplicación Windows: archivos .CPP, de definición de módulo, librerías y archivo de recursos.

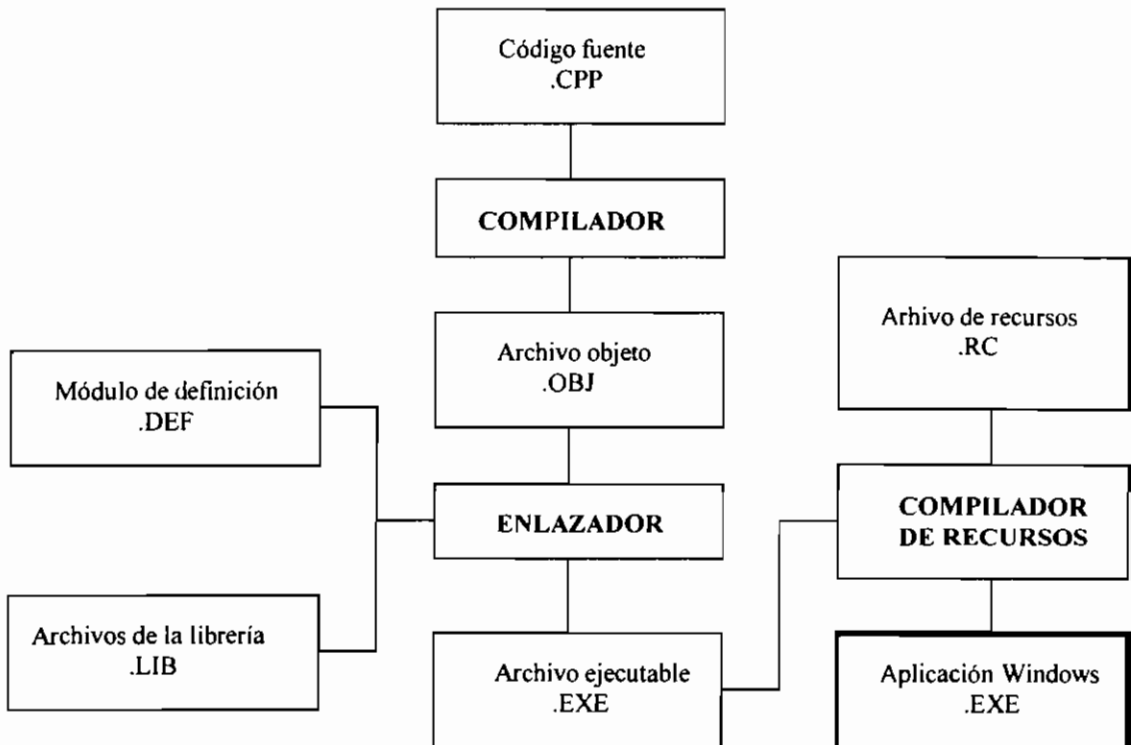


Figura 5.2.2.2 Proceso de compilación y enlazado en un proyecto Windows [21].

### **Archivo Proyecto de Aplicaciones MDI**

Una aplicación **Multiple Document Interface (MDI)** es aquella aplicación que maneja simultáneamente más de una ventana hija. La librería `ObjectWindows` permite desarrollar aplicaciones MDI usando unas pocas líneas de código. El archivo `.PRJ` incluye los archivos que contienen el código fuente de los programas a ejecutarse en cada uno de las ventanas hijas de la aplicación, así como también el programa fuente del programa principal que es el encargado de administrar la aplicación. El archivo de recursos contiene por lo menos un menú, incluirá también los recursos requeridos por cada subprograma si este es el caso.

### **5.2.3 SOFTWARE DESARROLLADO**

El programa desarrollado está constituido por los siguientes módulos: `MIMTDC.EXE`, `TDCOSEN.DLL` y `AYUDA.HLP`.

El programa ejecutable `MIMTDC.EXE` requerirá en su funcionamiento continuamente de dos módulos DLL, el módulo `BWCC.DLL` que es proporcionado por el paquete computacional Turbo C++ para Windows y da un aspecto mas atractivo a los componentes de una caja de diálogo. El módulo `TDCOSEN.DLL` calcula la transformada coseno de una imagen tanto en sentido directo como en sentido inverso. El módulo `AYUDA.HLP` será invocado cuando sea solicitada la presencia de las ayudas por parte del programa principal (El desarrollo del sistema de ayudas no se describirá, pues no presenta inconvenientes, las referencias [20] y [21] detallan el procedimiento a seguir para su implementación).

### 5.2.3.1 MÓDULO MIMTDC.EXE

El proyecto MIMTDC.PRJ está constituido por los siguientes archivos:

- ***mimtdc.def***: archivo de definición de módulo proporcionado por defecto por la librería ObjectWindows.
- ***especorg.cpp***: archivo que incluye el código necesario para la graficación de la imagen en el dominio de la frecuencia, imagen que representa el espectro de la imagen original.
- ***imgmejor.cpp***: contiene la información necesaria para la graficación de la imagen mejorada.
- ***espectro.cpp***: archivo que incluye la información necesaria para graficar el espectro de la transformada coseno.
- ***pegar.cpp***: contiene la información requerida para pegar una imagen desde el portapapeles.
- ***mimtdc.cpp***: contiene el código del programa principal que es el encargado de administrar la aplicación, y de las operaciones matemáticas que involucra el proceso de filtrado. Este es el programa encargado de solicitar la ejecución de la aplicación TDCOSEN.DLL, el código necesario para ello es el siguiente:

```
#include <owl.h>
#include <mdi.h>
#include <filedial.h>
#include <inputdia.h>
#include <...h>
```

```
    //prototipos de funciones residentes en DLL
void CosDir(int,int,float huge *,float huge *);
void CosInv(int,int,float huge *,float huge *);
```

-declaración de variable globales -

-definición de funciones de la aplicación-

```
//calcula transformada discreta coseno directa, llamando a DLL
.
.
CosDir(M,N,u,v);
.
.
//calcula transformada discreta coseno inversa, llamando a DLL
.
.
CosInv(M,N,u,v);
.
.
```

- **abrir.cpp**: archivo que contiene el código necesario para abrir un archivo (solamente será posible leer archivos en formato BMP) y luego desplegarlo en pantalla.
- **imgruido.cpp**: contiene el código necesario para añadir ruido a una imagen dato, ya sea obtenida por proceso de pegado desde el portapapeles o mediante el menú Abrir.
- **especrui.cpp**: archivo que incluye el código necesario para la graficación de la imagen que representa el espectro de la imagen ruidosa (contaminada por el usuario) si ésta existe.
- **mimtdc.rc**: archivo que contiene la descripción de recursos como: íconos, mapas de bits, cajas de diálogo, menús y aceleradores de menú.
- **tdcoseno.lib**: biblioteca de importación que contiene información de las funciones de propiedad del modulo TDCOSEN0.DLL accesibles para el módulo MIMTDC.EXE.
- **bwcc.lib**: librería de importación correspondiente del modulo BWCC.DLL.

A continuación se indican las opciones de compilación recomendadas cuando se construyen aplicaciones windows.

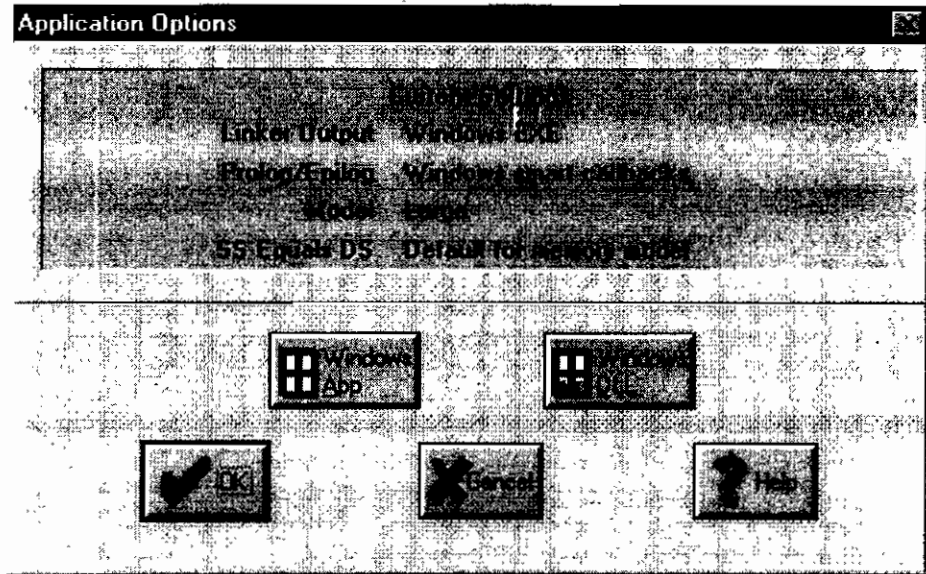


Figura 5.2.3.1.1 Caja de diálogo de Options/Application...

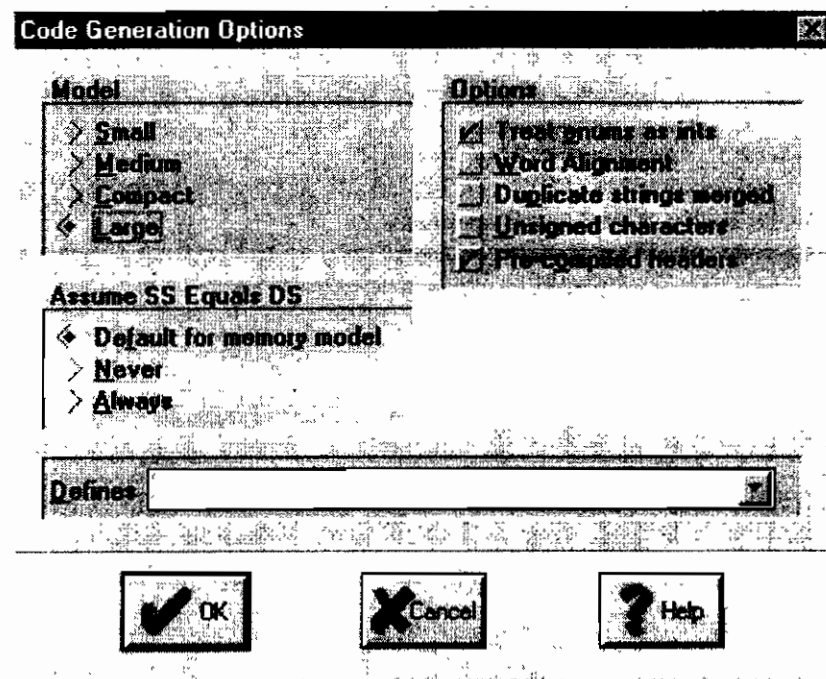


Figura 5.2.3.1.2 Caja de diálogo de Options/Compiler/Code Generation...

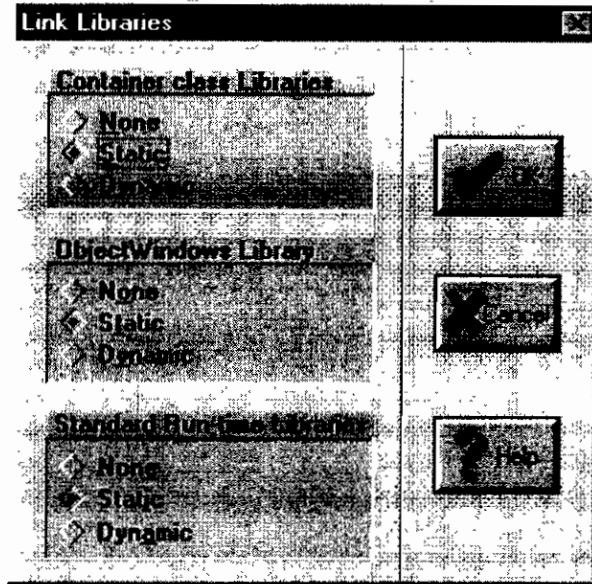


Figura 5.2.3.1.3 Caja de diálogo de Options/Linker/Libraries...

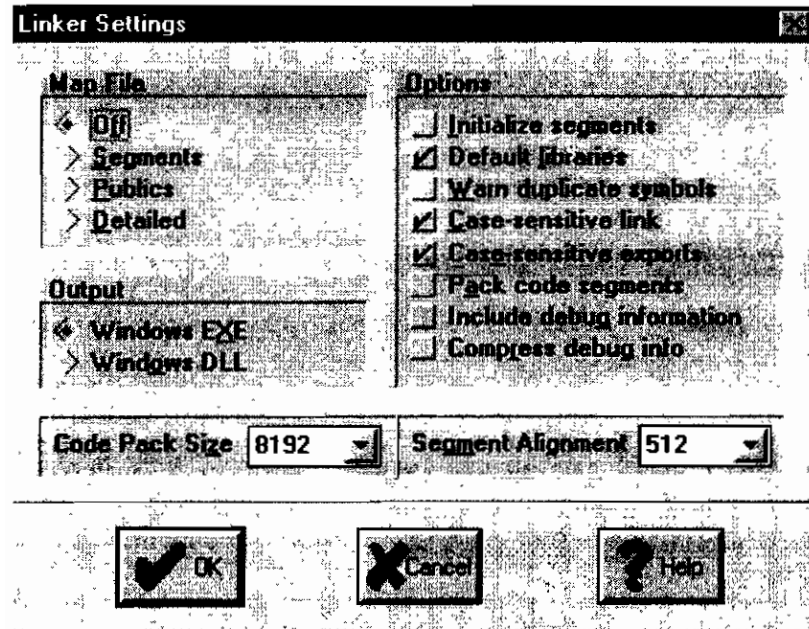


Figura 5.2.3.1.4 Caja de diálogo de Options/Linker/Settings...



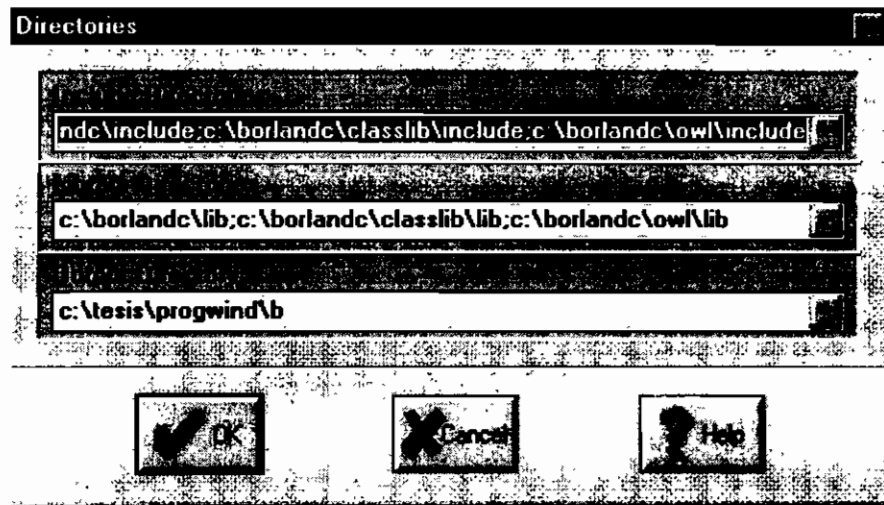


Figura 5.2.3.1.5 Caja de diálogo de Options/Directories...

### 5.2.3.2 MÓDULO TDCOSENSO.DLL

La información accesible sobre la creación de bibliotecas de enlace dinámico es sumamente limitada, es por esto que las ideas que se presentan a continuación resultan como conclusión de un análisis tedioso de escasos ejemplos disponibles en el paquete Borland C++.

El archivo proyecto correspondiente TDCOSENSO.PRJ está constituido de los siguientes archivos:

**entradll.cpp:** contiene la definición de LibMain y WEP que constituyen respectivamente las funciones de entrada y salida de la biblioteca de enlace dinámico. El código usado en este caso es:

```
#include "windows.h"
int FAR PASCAL LibMain(HANDLE hInstance,WORD, WORD,LPSTR)
{
    return 1;
}
int FAR PASCAL WEP(int)
{
    return 1;
}
```

*tdcoseno.cpp*: este archivo hace referencia al archivo *tdcoseno.h* que contiene la declaración de las funciones y clases definidas en *tdcoseno.cpp* y que pueden ser llamadas desde el programa *mimtdc.exe*. Las únicas funciones accesibles por el programa *mimtdc.exe* serán aquellas calificadas con las palabras clave **FAR \_export** tanto en su declaración como en su definición, a continuación se reproduce lo más importante del archivo de cabecera de la biblioteca *tdcoseno.dll* el archivo *tdcoseno.h*:

```
#ifndef __TDCOSENSO_H
#define __TDCOSENSO_H

#include <windows.h>
#include <...h>

-aquí se declara las variables globales-

void FAR _export CosDir(int,int,float huge *,float huge *);
void FAR _export CosInv(int,int,float huge *,float huge *);

class TCoseno
{
-miembros privados de la clase-
public:
    TCoseno(int,int,float huge *,float huge * ); //constructor
    ~TCoseno(){}; //destructor
-funciones miembro adicionales-
};

#endif
```

*tdcoseno.cpp* define cada una de las funciones declaradas en su archivo (*.h*). Es importante recalcar que este archivo solamente contiene la definición de las funciones declaradas en el archivo cabecera, en ningún caso se hace referencia a funciones como: LibMain,WEP, main o WinMain.

*tdcoseno.def*: el modelo de este archivo ha sido tomado de ejemplos existentes de DLLs, cambiando solamente el nombre de la librería de importación en la sección LIBRARY, el código es el siguiente:

```
LIBRARY TDCOSENSO
DESCRIPTION 'DLL Tcoseno'
EXETYPE     WINDOWS
CODE PRELOAD MOVEABLE DISCARDABLE
DATA PRELOAD MOVEABLE SINGLE
HEAPSIZE 1024
```

El archivo *tdcoseno.def* podría incluir una sección denominada EXPORT la cual listaría las funciones exportables, esto es, las funciones accesibles por otra aplicación DLL o EXE, otra solución es generar una biblioteca de importación TDCOSENSO.LIB e incluirla como parte del proyecto que usará la aplicación DLL generada, en este caso en MIMTDC.PRJ. La alternativa escogida es la segunda, para ello se ejecuta el utilitario IMPLIBW.EXE de Borland que solicita como archivo de entrada el archivo TDCOSENSO.DLL y devuelve el archivo TDCOSENSO.LIB.

Una vez conformado el proyecto de la aplicación DLL el siguiente paso es compilarlo, para lo cual se recomienda las siguientes opciones de compilación:

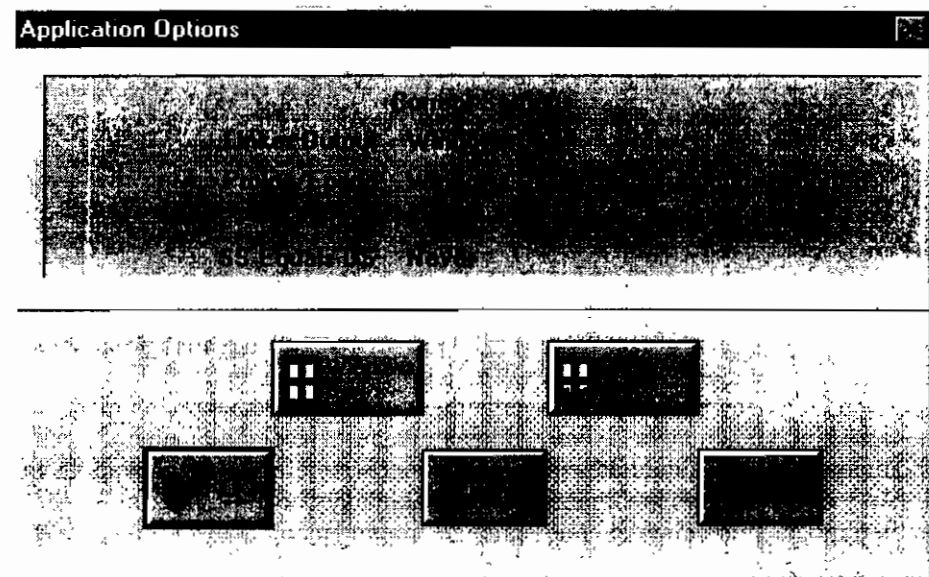


Figura 5.2.3.2.1 Caja de diálogo de Options/Application...

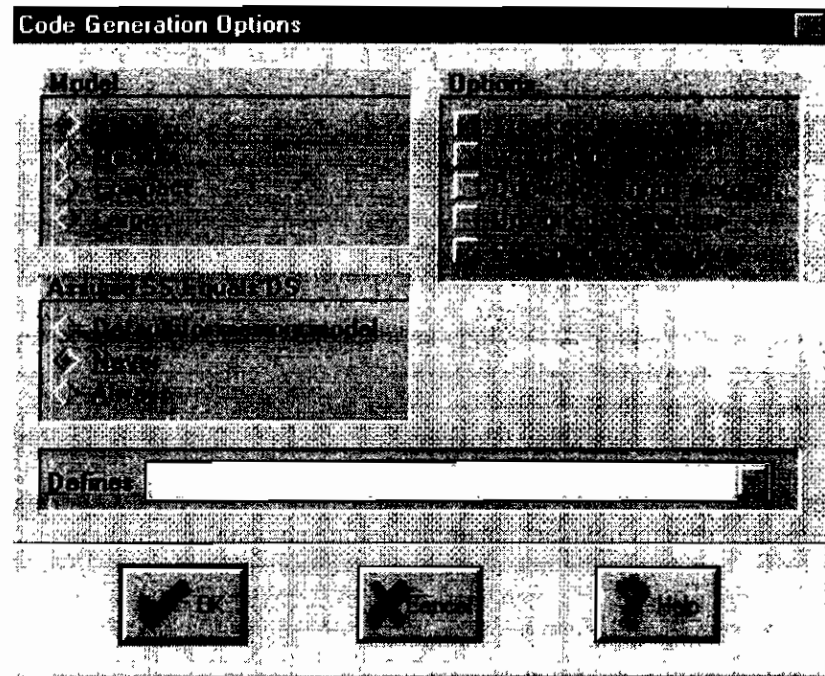


Figura 5.2.3.2.2 Caja de diálogo de Options/Compiler/Code Generation...

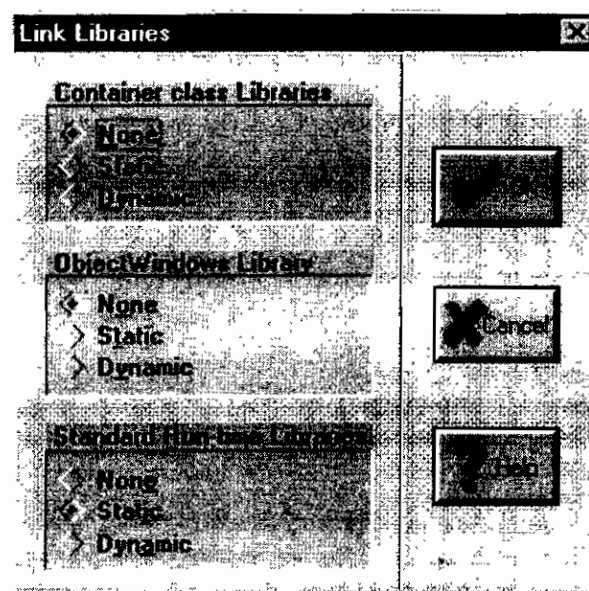


Figura 5.2.3.2.3 Caja de diálogo de Options/Linker/Libraries...

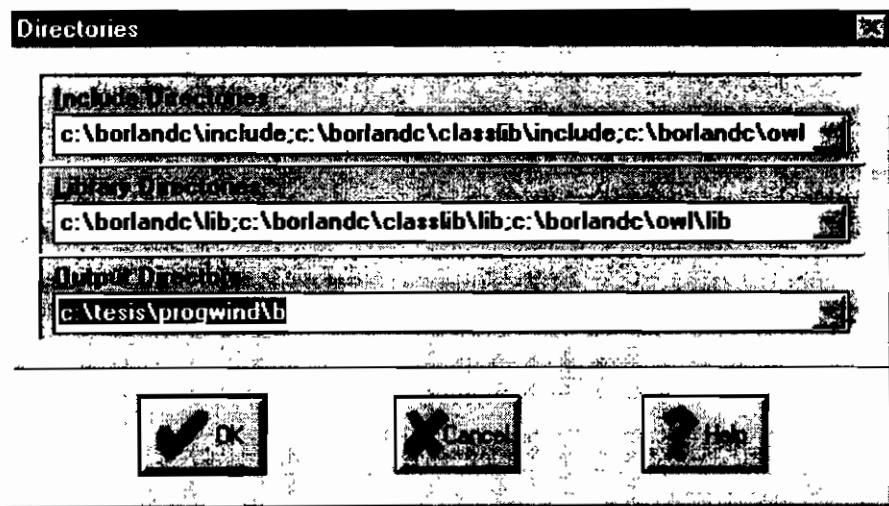


Figura 5.2.3.2.4 Caja de diálogo de Options/Directories

Las referencias [18], [20], [21] y [22] exponen la teoría básica de la programación de aplicaciones Windows en cambio [17] y [19] exponen las bases del lenguaje de programación C.

# CAPÍTULO

---

# 6

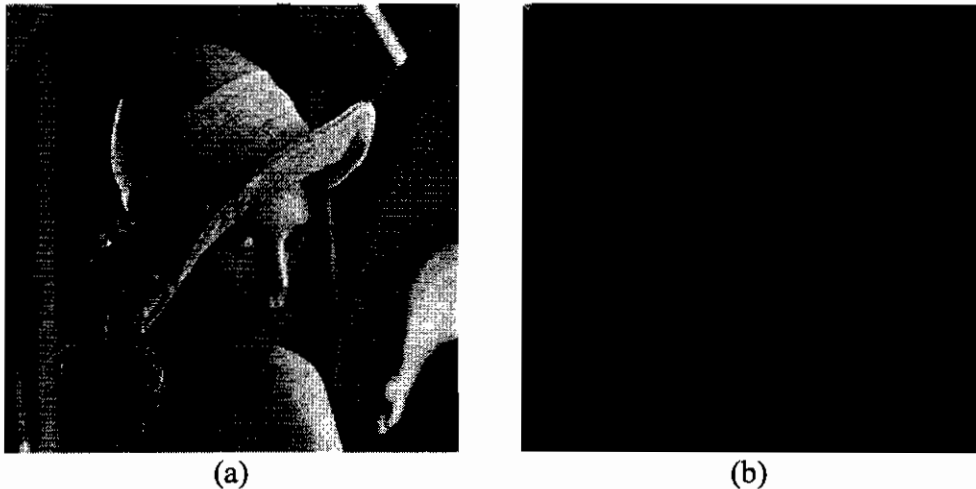
## ANÁLISIS COMPARATIVO DE RESULTADOS OBTENIDOS

### 6.0 INTRODUCCIÓN

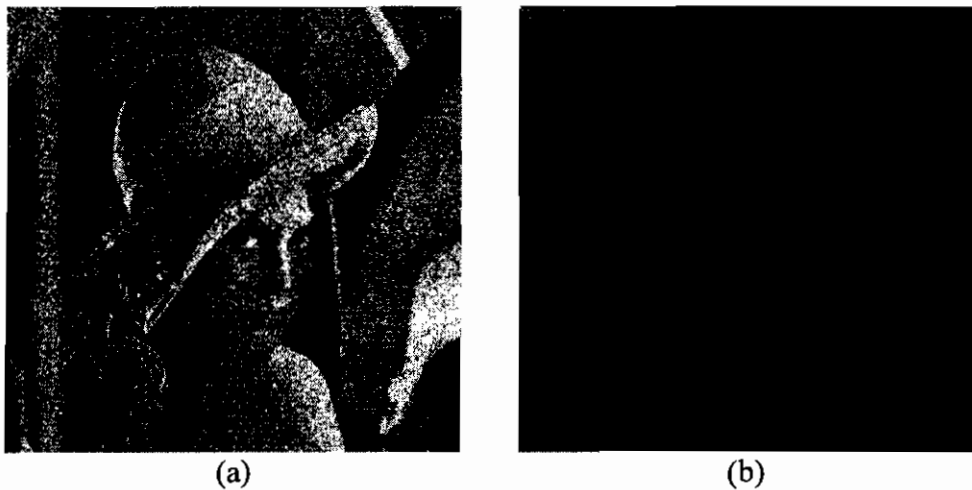
La finalidad de este capítulo es tabular los resultados obtenidos, para lo cual se ha tomando una imagen típica como modelo de análisis, se ha contaminado dicha imagen con los ruidos disponibles en el sistema de software implementado; y, luego se ha aplicado cada uno de los filtros y analizado la efectividad de los mismos sobre cada uno de los ruidos. Las imágenes se contaminarán con dos tipos de ruidos y estos son: ruido aditivo y ruido dependiente de la señal; dentro del ruido aditivo se encuentran el Gaussiano blanco y el impulsivo (dentro de este tipo a su vez se encuentran el sal y pimienta, impulsivo positivo e impulsivo negativo), como ruido dependiente de la señal se ubican al multiplicativo y al propiamente llamado dependiente de la señal.

## 6.1 FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO GAUSSIANO BLANCO

La figura 6.1.0.1 presenta la imagen original con su respectivo espectro de la transformada coseno, mientras que la figura 6.1.0.2 presenta la imagen contaminada con ruido aditivo gaussiano blanco con desviación estándar de 30, acompañada de su respectivo espectro coseno. En las gráficas que representan el espectro, un punto blanco representará máximo nivel energético, en cambio un punto negro indicará la ausencia de energía.



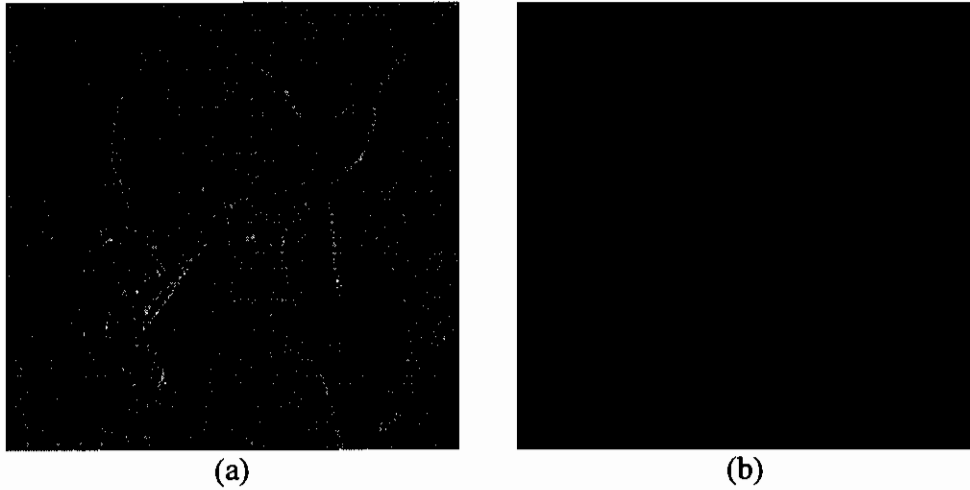
**Figura 6.1.0.1** (a) Imagen original, (b) Espectro de la transformada coseno de la imagen dada en (a).



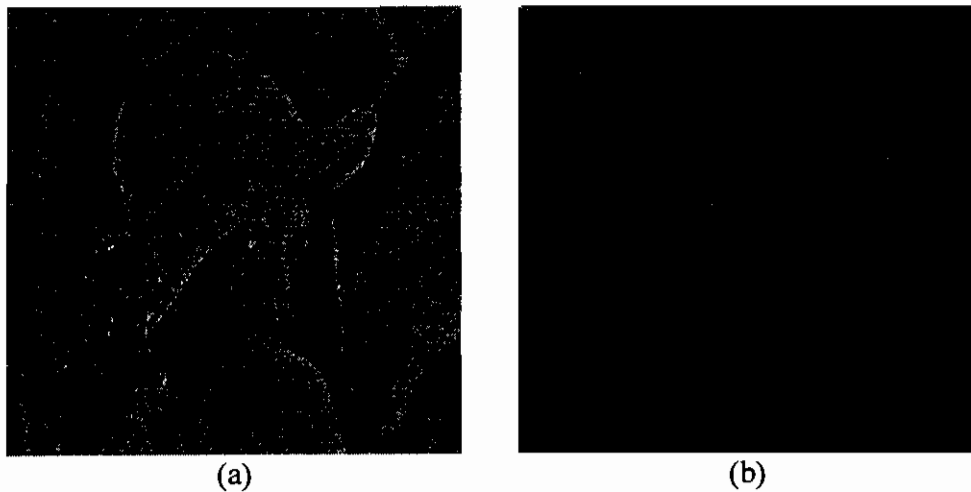
**Figura 6.1.0.2** (a) Imagen original ruidosa, ruido aditivo Gaussiano blanco, desviación estandar = 30, SNR=19.6 dB, (b) Espectro de la transformada coseno de la imagen dada en (a).

### 6.1.1 FILTRO RAÍZ

Se ha procedido a filtrar la imagen ruidosa dada por la figura 6.1.0.2a con cuatro valores de coeficientes alpha, obteniéndose los siguientes resultados:

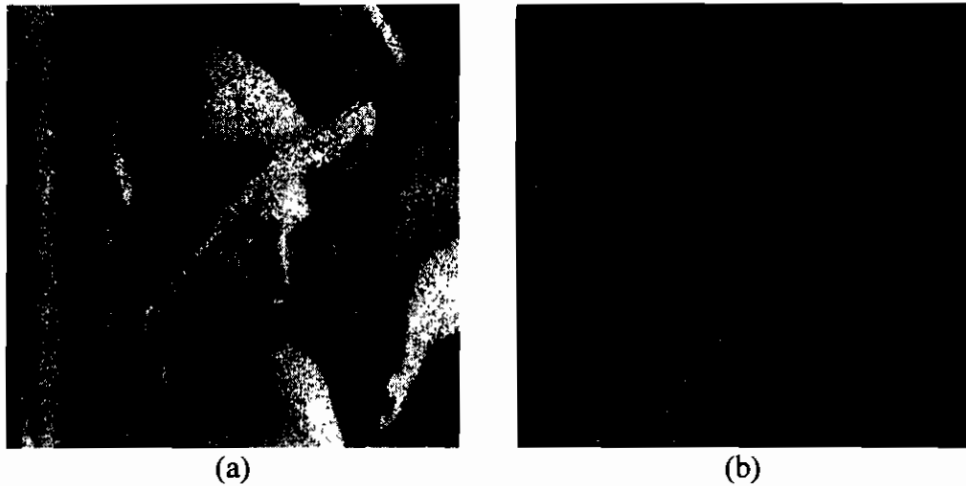


**Figura 6.1.1.1** (a) Imagen filtrada usando filtro raíz con  $\alpha=0.5$ , SNR=15.3 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

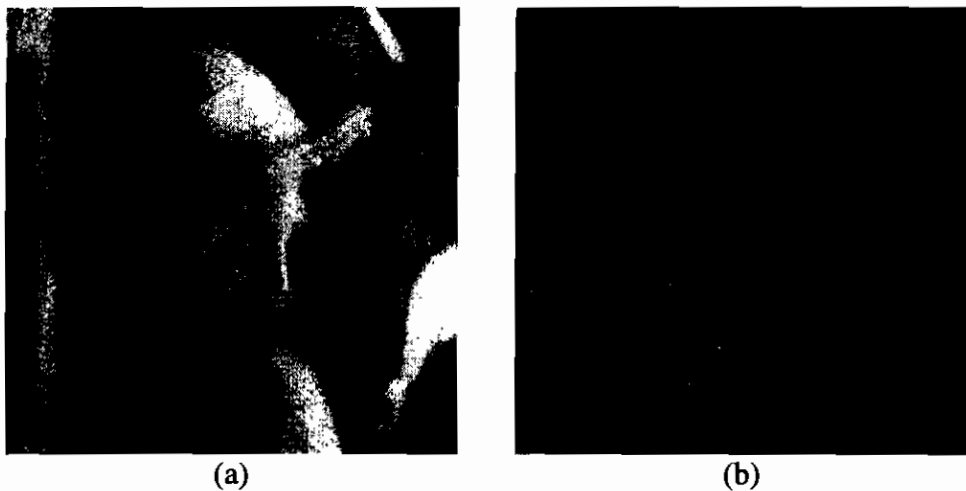


**Figura 6.1.1.2** (a) Imagen filtrada usando filtro raíz con  $\alpha=0.7$ , SNR=16.3 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).





**Figura 6.1.1.3** (a) Imagen filtrada usando filtro raíz con  $\alpha=1.1$ , SNR=21.3 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

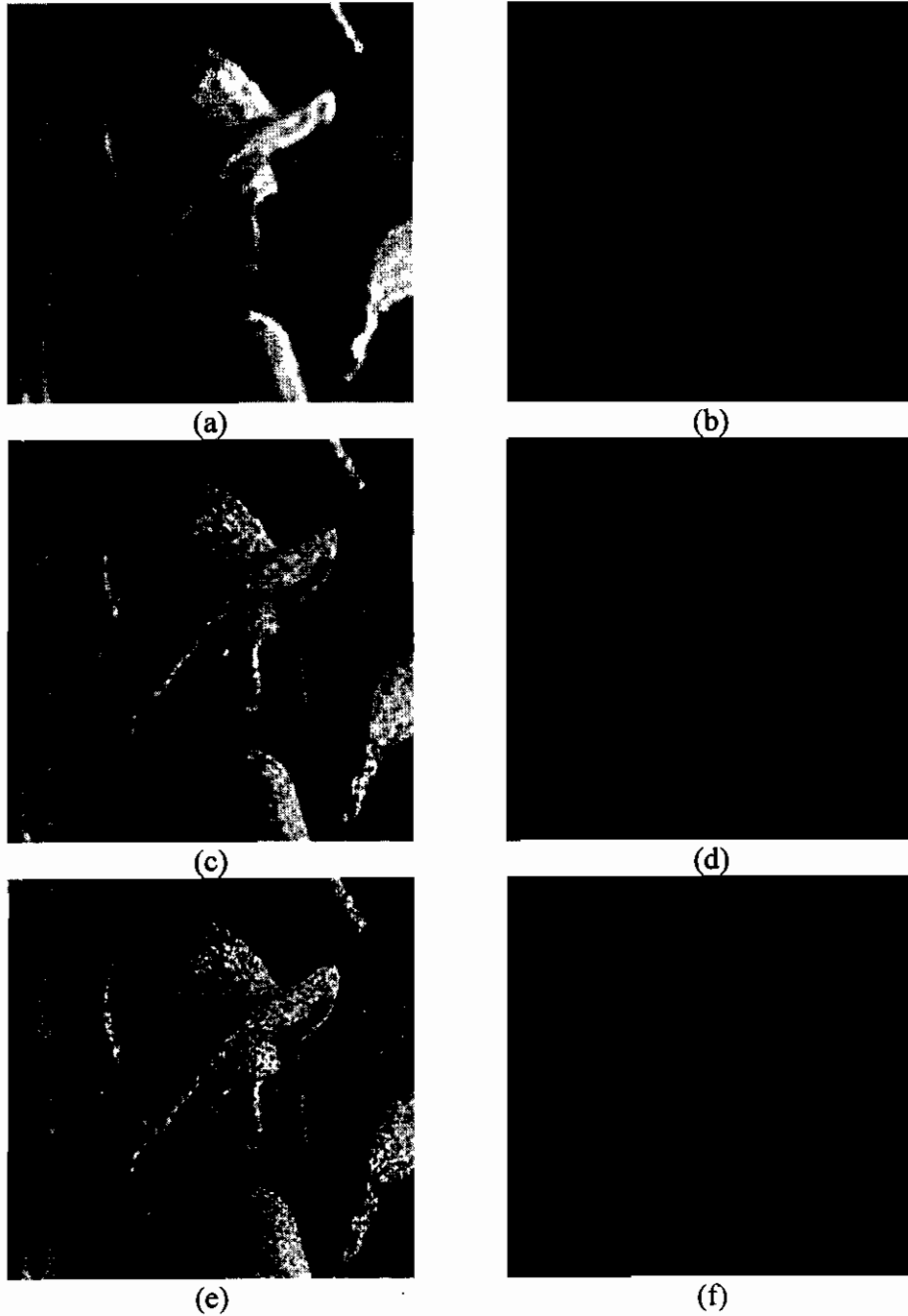


**Figura 6.1.1.4** (a) Imagen filtrada usando filtro raíz con  $\alpha=1.3$ , SNR=22.3 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

Se verifica que para coeficientes alpha menores que la unidad se resaltan componentes de alta frecuencia determinadas principalmente por pixeles ruidosos. Para coeficientes alpha mayores que 1 en cambio la imagen filtrada tiende a suavizarse opacando de esta manera la presencia de pixeles ruidosos. En el primer caso al resaltar el ruido la relación señal a ruido disminuye, situación contraria a lo que ocurre en el segundo caso.

### 6.1.2 FILTRO PASA-BAJOS IDEAL

En este caso se filtra la imagen ruidosa dada por la figura 6.1.0.2a con tres valores de radio de filtro pasa-bajos ideal, obteniéndose los siguientes resultados:

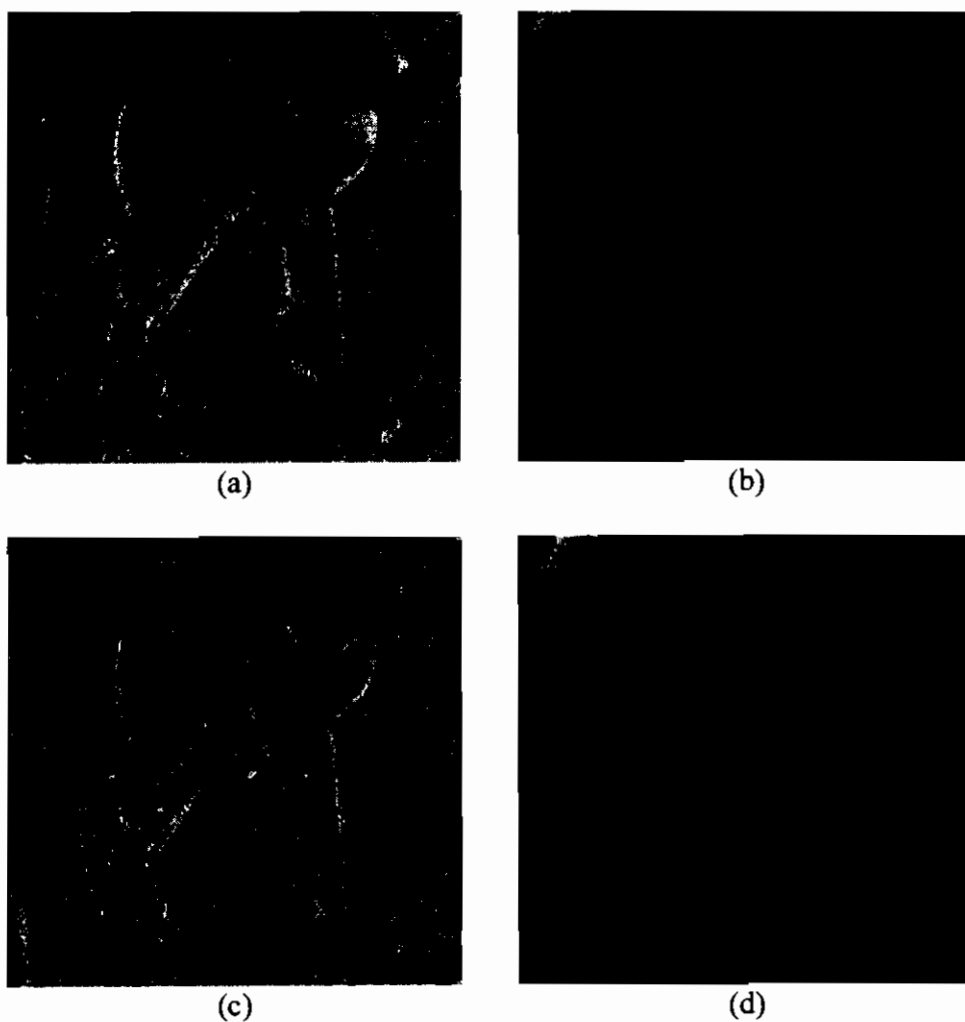


**Figura 6.1.2.1** (a) Imagen filtrada usando filtro pasa-bajos con radio = 50, SNR=23.4 dB (b) Espectro de la transformada coseno de la imagen dada en (a). (c) Imagen filtrada usando filtro pasa-bajos con radio = 100, SNR=24 dB (d) Espectro de la imagen dada en (c). (e) Imagen filtrada usando filtro pasa-bajos con radio =150, SNR=22.2 dB (f) Espectro de la imagen dada en (e).

Al disminuir el radio pasa-bajos se atenúa en cierta cantidad la presencia de ruido, a costo de deteriorar la imagen misma.

### 6.1.3 FILTRO PASA-ALTOS IDEAL

Las imágenes resultantes de filtrar la imagen ruidosa dada por la figura 6.1.0.2a con dos valores de radio de filtro pasa-altos ideal, son las siguientes:

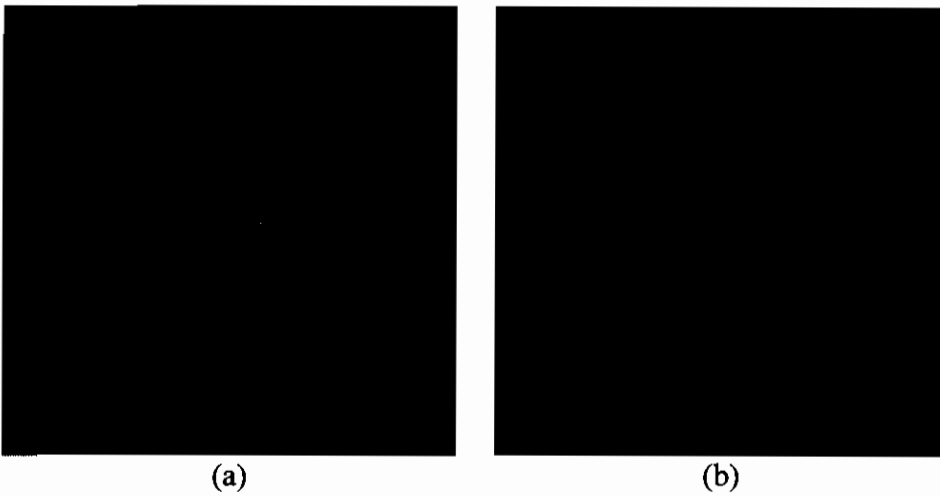


**Figura 6.1.3.1** (a) Imagen filtrada usando filtro pasa-altos con radio = 10, SNR=15.3 dB (b) Espectro de la transformada coseno de la imagen dada en (a). (c) Imagen filtrada usando filtro pasa-altos con radio = 20, SNR=14.6 dB (d) Espectro de la imagen dada en (c).

Conforme aumenta el radio pasa-altos se resaltan componentes de alta frecuencia determinadas en este caso por contornos y pixeles ruidosos, perjudicando a la relación señal a ruido.

#### 6.1.4 FILTRO GAUSSIANO INVERSO

Para este caso se ha procedido a filtrar la imagen ruidosa dada por la figura 6.1.0.2a con filtro Gaussiano inverso, obteniéndose la siguiente imagen resultante:

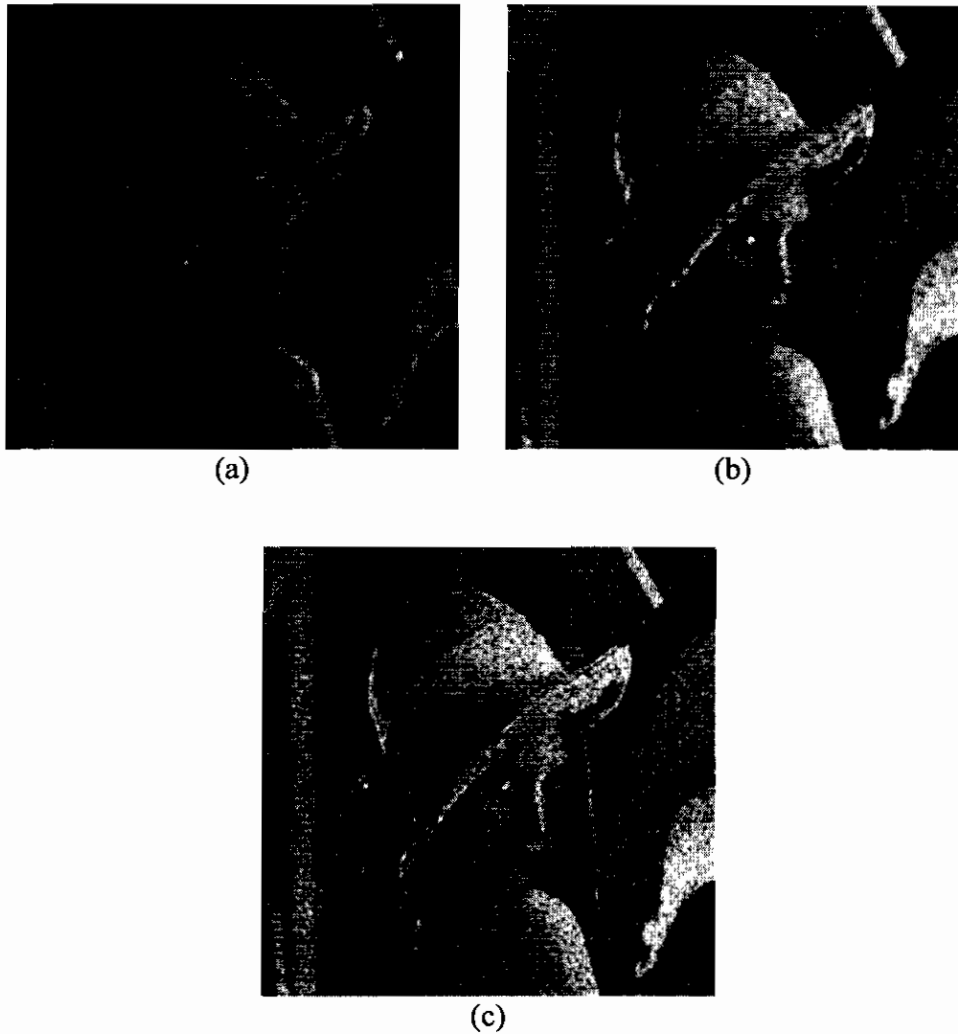


**Figura 6.1.4.1** (a) Imagen filtrada usando filtro Gaussiano inverso, SNR=14.6 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

La imagen filtrada se deteriora significativamente con respecto a la no filtrada, porque este filtro al comportarse como pasa-altos, enfatiza la presencia de ruido; esto se produce debido a que el ruido tiene componentes importantes de alta frecuencia.

#### 6.1.5 FILTRO HOMOMÓRFICO

Las imágenes resultantes obtenidas luego de aplicar el filtro homomórfico a la imagen ruidosa dada por la figura 6.1.0.2a son las siguientes:

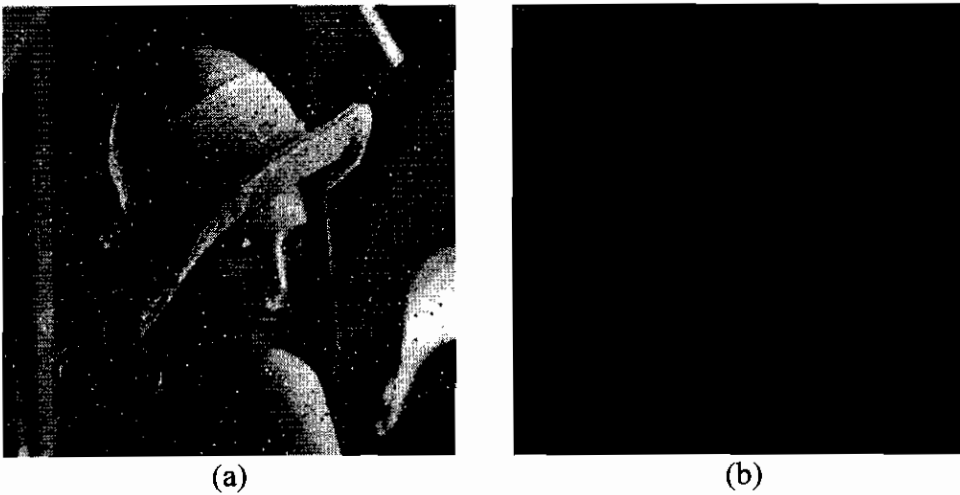


**Figura 6.1.5.1** (a) Imagen filtrada usando filtro homomórfico y pasa-bajos con radio = 50, SNR=17.0 dB  
(b) Imagen filtrada usando filtro homomórfico y pasa-bajos con radio = 100, SNR=17.6 dB. (c) Imagen filtrada usando filtro homomórfico y pasa-bajos con radio = 150, SNR=15.8 dB.

Como en este proceso de filtrado, se usa también el filtro pasa-bajos ideal, su efecto es significativo, es decir, disminuyendo el radio disminuye la presencia de ruido, pero empaña la imagen eliminando sus detalles. En cambio, al aumentar el radio del filtro lineal se recuperan los detalles de la imagen, pero esta se vuelve a contaminar.

## 6.2 FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO SAL Y PIMIENTA

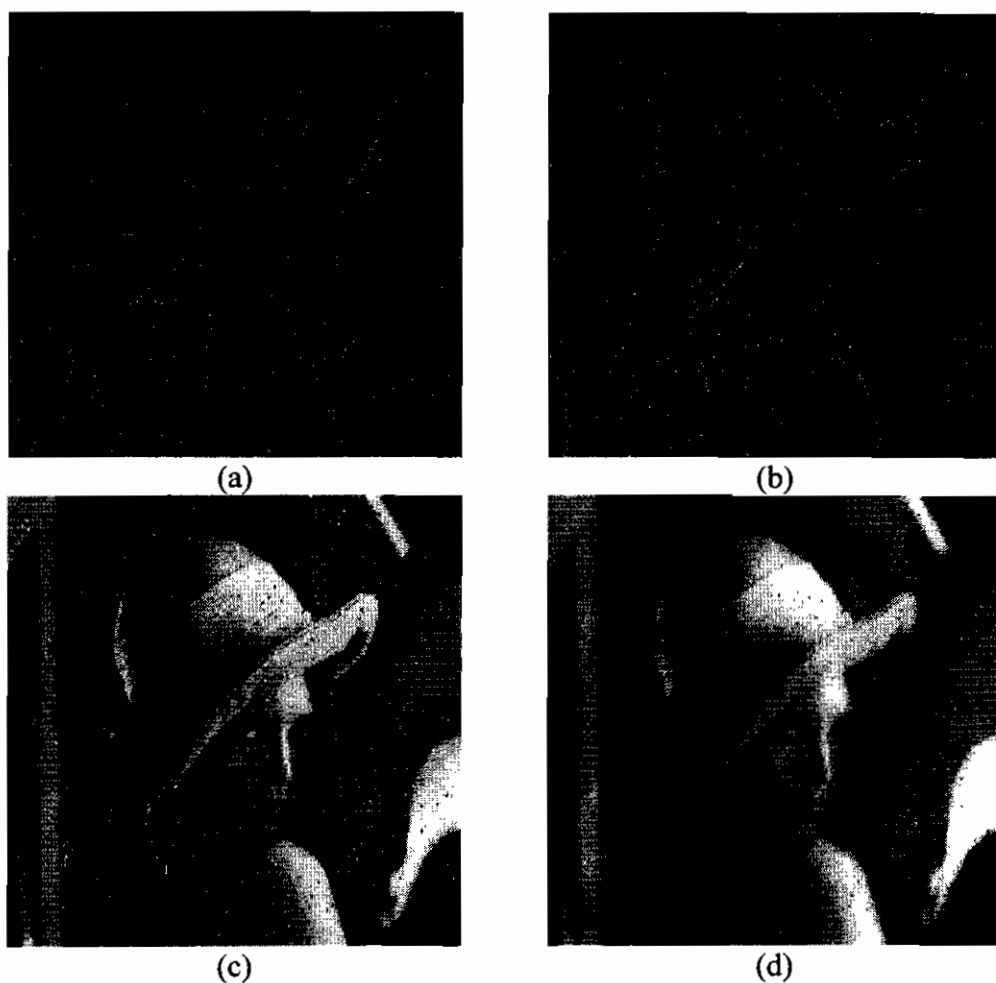
La figura 6.2.0.1a presenta la imagen original (figura 6.1.0.1a) contaminada con ruido sal y pimienta con probabilidad de ocurrencia del ruido de 0.01, acompañada de su respectivo espectro coseno. En las subsiguientes secciones no se incluyen las imágenes que representan el espectro de la imagen respectiva, pues en si la forma del espectro es similar al de los ejemplos de la sección anterior, respecto al espectro de la imagen en análisis.



**Figura 6.2.0.1** (a) Imagen original ruidosa, con ruido sal y pimienta con probabilidad de ocurrencia de 0.01, SNR=24.9 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

### 6.2.1 FILTRO RAÍZ

Filtrando la imagen ruidosa dada por la figura 6.2.0.1a con cuatro valores de coeficientes alpha, se obtienen los siguientes resultados:

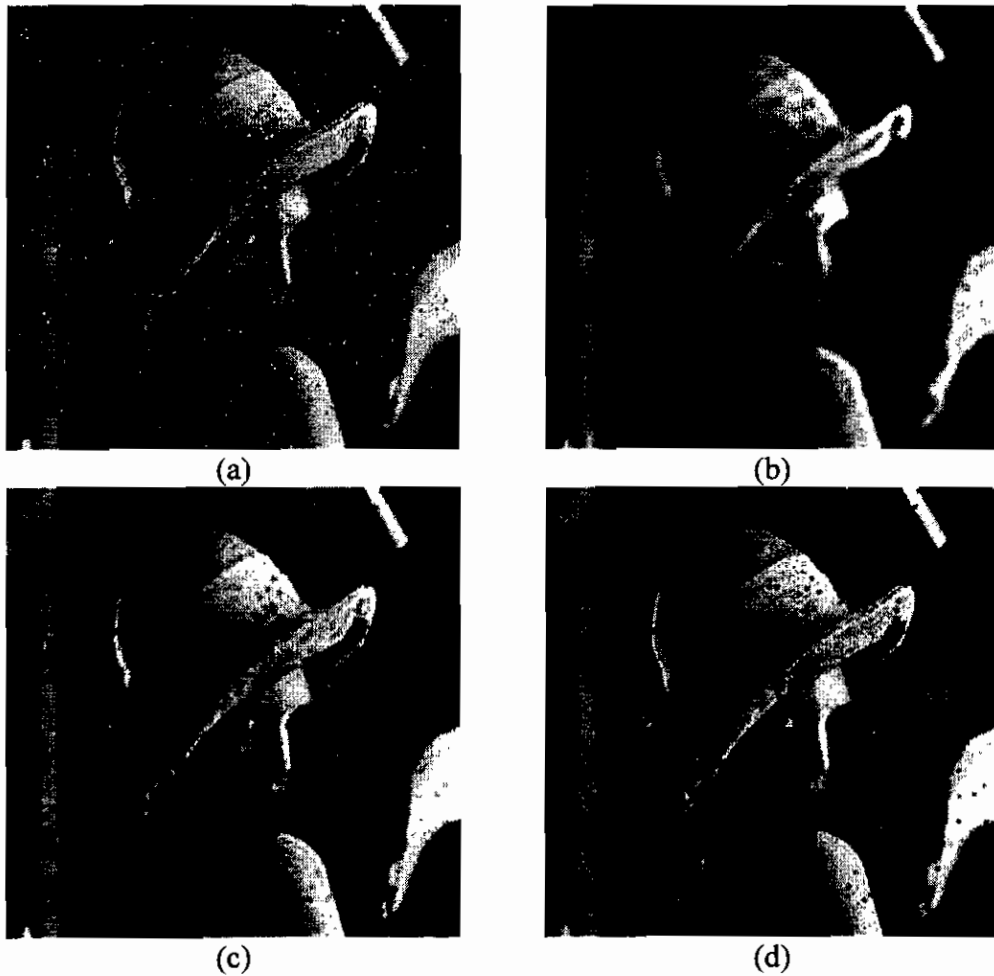


**Figura 6.2.1.1** Imagen filtrada usando filtro raíz con: (a)  $\alpha=0.5$ , SNR=15.4 dB. (b)  $\alpha=0.7$ , SNR=16.4 dB. (c)  $\alpha=1.1$ , SNR=26.0 dB. (d)  $\alpha=1.3$ , SNR=23.8 dB.

Los impulsos positivos (puntos blancos) en zonas oscuras y los impulsos negativos (puntos negros) en zonas claras son los que determinan los pixeles ruidosos más difíciles de remover. El filtro raíz en modalidad pasa-altos resalta estos pixeles, en cambio en modalidad pasa-bajos trata de atenuar el efecto de contraste provocado por estos pixeles dentro de la imagen.

## 6.2.2 FILTRO PASA-BAJOS IDEAL

Luego de aplicar el filtro pasa-bajos ideal a la imagen ruidosa dada por la figura 6.2.2.1a con tres valores de radio de filtro, se obtienen los siguientes resultados:



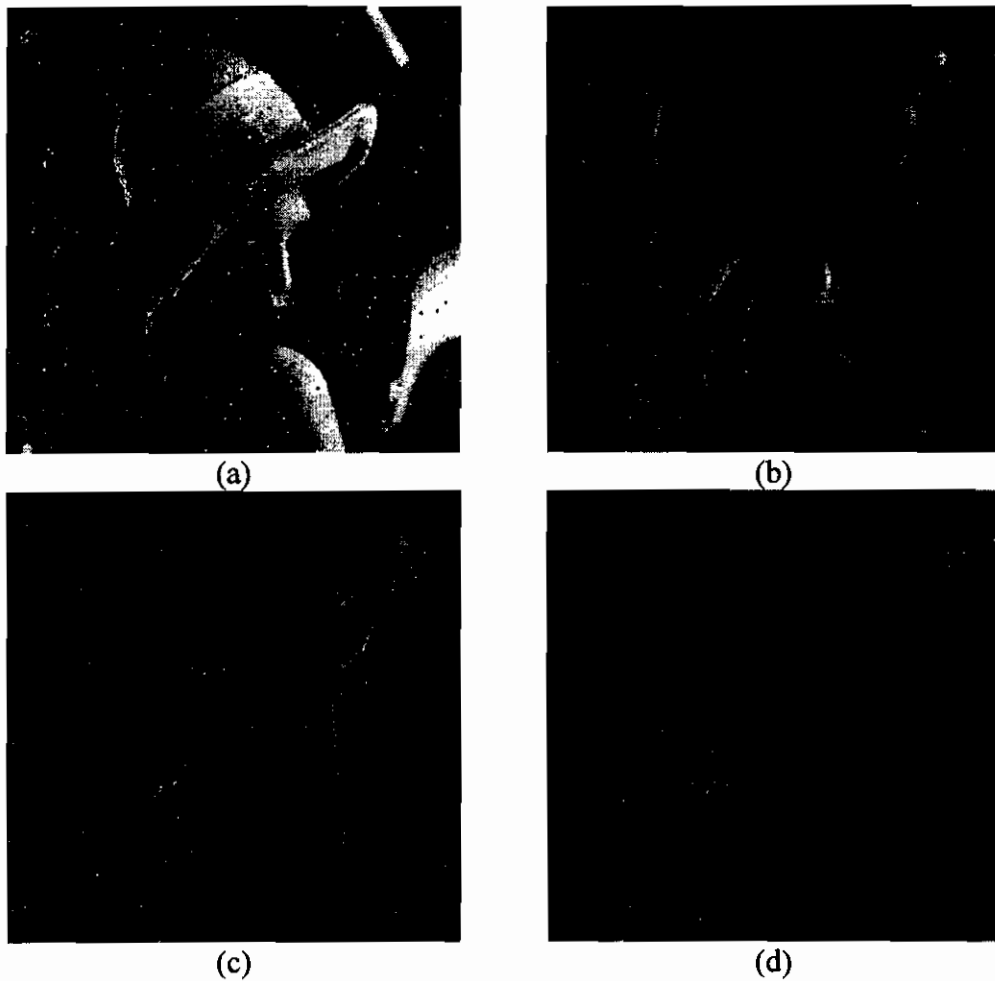
**Figura 6.2.2.1** (a) Imagen original ruidosa, con ruido sal y pimienta con probabilidad de ocurrencia de 0.01, SNR=24.9 dB. Imagen filtrada usando filtro pasa-bajos con: (b) radio = 50, SNR=23.5 dB. (c) radio = 100, SNR=25.6 dB. (d) radio =150, SNR=25.8 dB.

Dependiendo de la probabilidad de ocurrencia del ruido (concentración del ruido) los altos contrastes provocados determinan grandes cantidades energéticas concentradas en zonas de alta frecuencia lo que hace menos difícil la tarea de filtrar usando este tipo de filtro y para el caso de este ruido impulsivo. En estas imágenes se constata una reducción apreciable del ruido sal y pimienta.

### 6.2.3 FILTRO PASA-ALTOS IDEAL

Filtrando la imagen ruidosa dada por la figura 6.2.3.1a con tres valores de radio de filtro pasa-altos ideal, se obtienen los siguientes resultados:



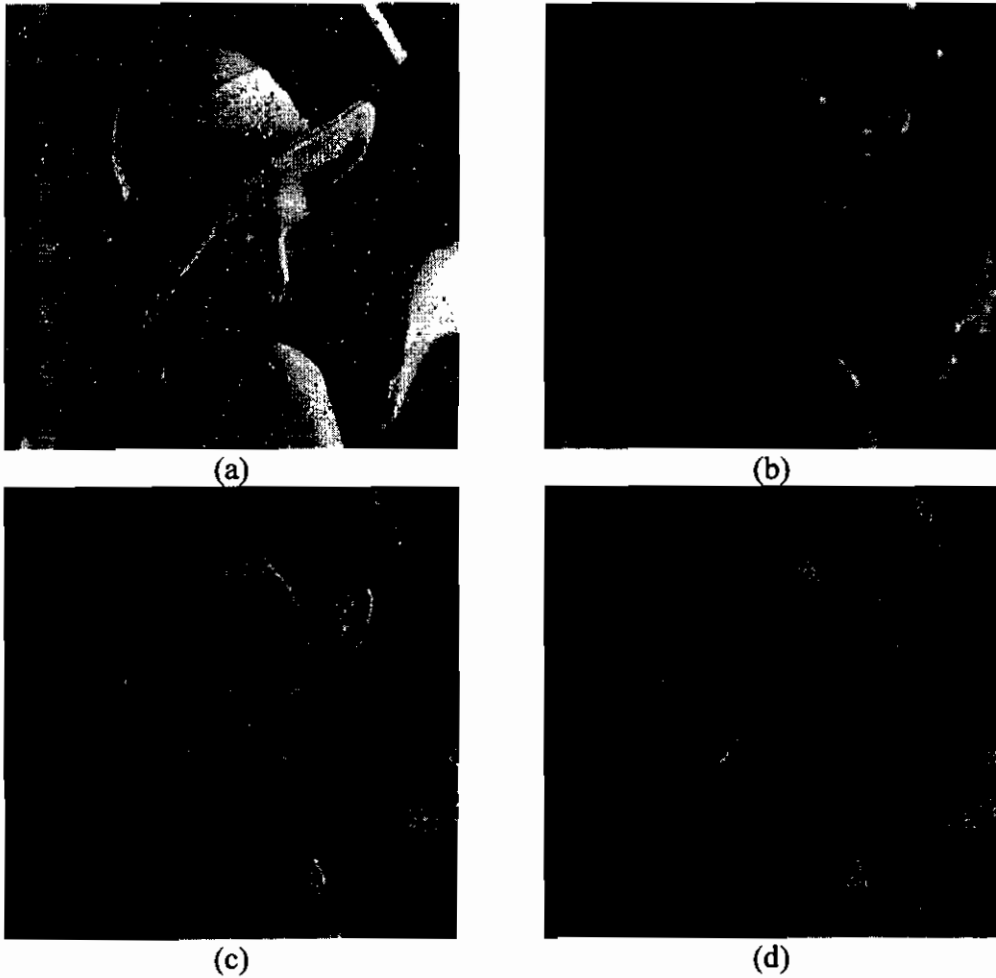


**Figura 6.2.3.1** (a) Imagen original ruidosa, con ruido sal y pimienta con probabilidad de ocurrencia de 0.01, SNR=24.9 dB. Imagen filtrada usando filtro pasa-altos con: (b) radio = 10, SNR=15.7 dB. (c) radio = 20, SNR=15.2 dB. (d) radio = 50, SNR=14.6 dB.

Lo contrario a lo ocurrido con el filtro pasa-bajos, el filtro pasa-altos resalta de manera clara el ruido en la imagen, imagen que resulta afectada en sus detalles suaves por la aplicación de este tipo de filtro.

#### 6.2.4 FILTRO HOMOMÓRFICO

Al aplicar el filtro homomórfico sobre la imagen 6.2.4.1a, se obtienen las imágenes siguientes:

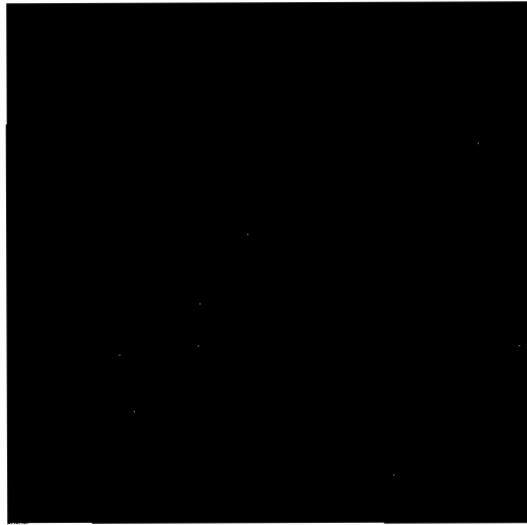


**Figura 6.2.4.1** (a) Imagen original ruidosa, con ruido sal y pimienta con probabilidad de ocurrencia de 0.01, SNR=24.9 dB. Imagen filtrada usando filtro homomórfico y filtro pasa-bajos con: (b) radio = 50, SNR=19.1 dB. (c) radio = 100, SNR=19.0 dB. (d) radio =150, SNR=18.5 dB.

Este filtro atenúa apreciablemente los impulsos positivos, no así los impulsos negativos que más bien distorsionan la imagen, manchando de negro el área cercana a la posición original de los puntos negros.

### 6.2.5 FILTRO GAUSSIANO INVERSO

En este caso se ha filtrado la imagen ruidosa dada por la figura 6.2.0.1a con filtro Gaussiano inverso, obteniéndose el siguiente resultado:

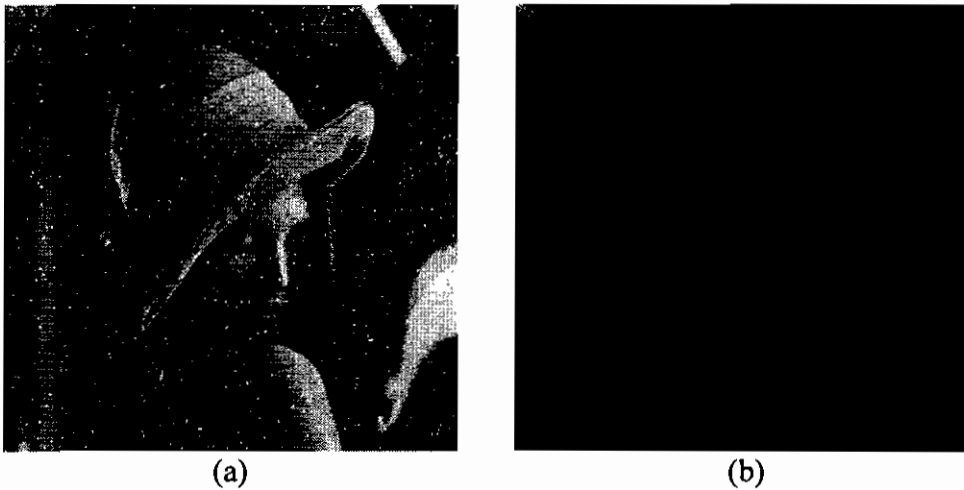


**Figura 6.2.5.1** Imagen filtrada usando filtro Gaussiano inverso, SNR=15.1 dB.

La imagen resultante de aplicar este tipo de filtro es una imagen ruidosa y sobre todo distorsionada respecto a la no filtrada. Esto se debe a que el filtro gaussiano inverso al tener característica pasa-altos enfatiza las componentes que mayoritariamente definen al ruido.

### **6.3 FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO IMPULSIVO POSITIVO**

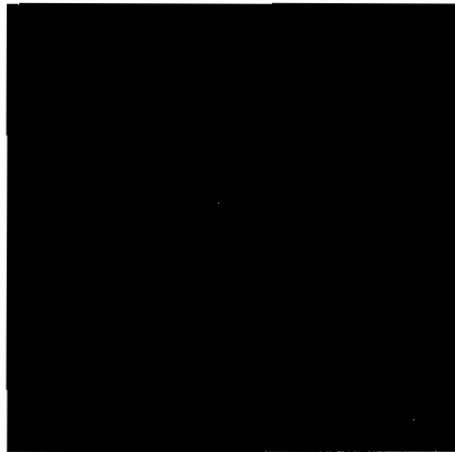
La figura 6.3.0.1a presenta la imagen original (la imagen original se muestra en la figura 6.1.0.1a) contaminada con ruido impulsivo positivo con probabilidad de ocurrencia del ruido de 0.01, acompañada de su respectivo espectro coseno.



**Figura 6.3.0.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo positivo con probabilidad de ocurrencia de 0.01, SNR=25.2 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

### 6.3.1 FILTRO GAUSSIANO INVERSO

Para este caso se ha procedido a filtrar la imagen ruidosa dada por la figura 6.3.0.1a con filtro Gaussiano inverso, obteniéndose el siguiente resultado:



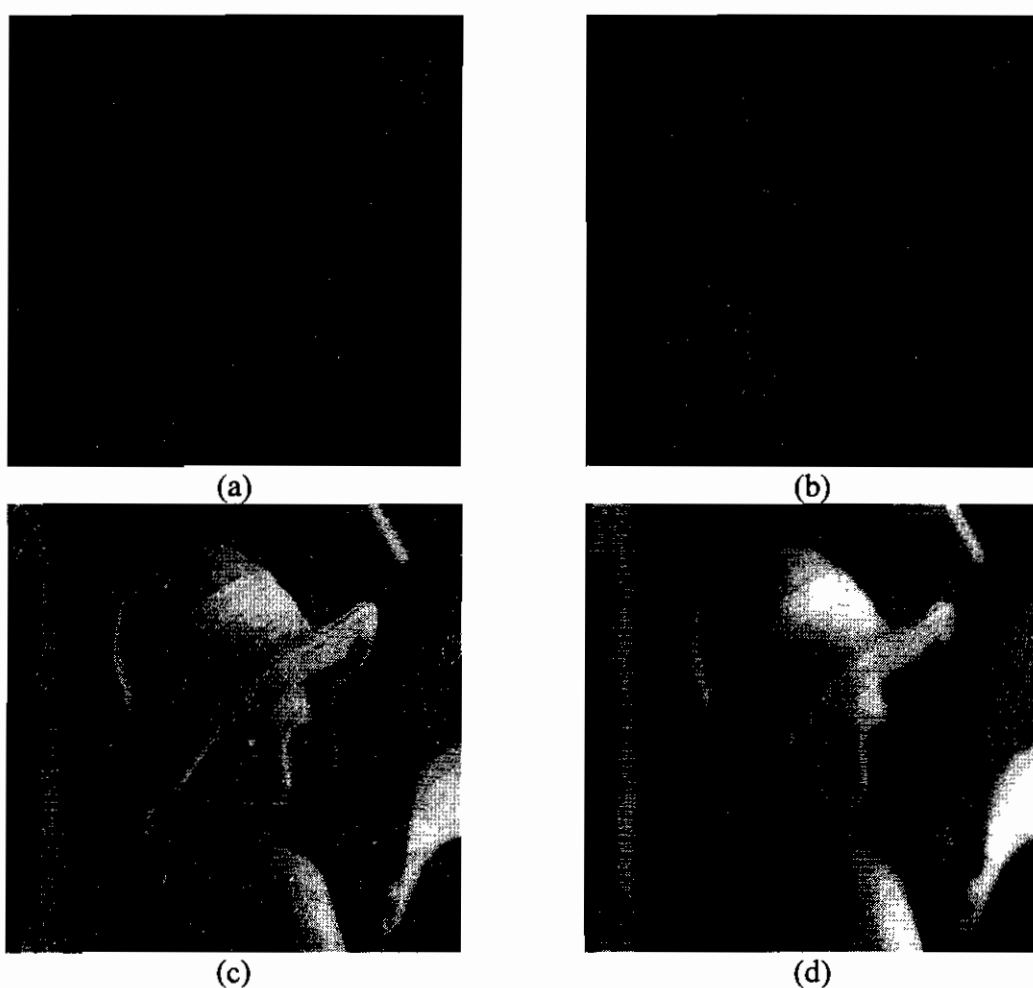
**Figura 6.3.1.1** Imagen filtrada usando filtro Gaussiano inverso, SNR=15.3 dB.

Como en los casos anteriores (con otro tipo de ruido contaminante) este filtro distorsiona la imagen resaltando el ruido; esto ocurre porque este filtro no solo deja pasar las componentes de alta frecuencia, sino que también las amplifica, esta

afirmación se desprende del análisis de la expresión que regula su comportamiento (ecuación- 4.2.2.4).

### 6.3.2 FILTRO RAÍZ

Filtrando la imagen ruidosa dada por la figura 6.3.0.1a con cuatro valores de coeficientes alpha del filtro raíz, se obtienen los siguientes resultados:

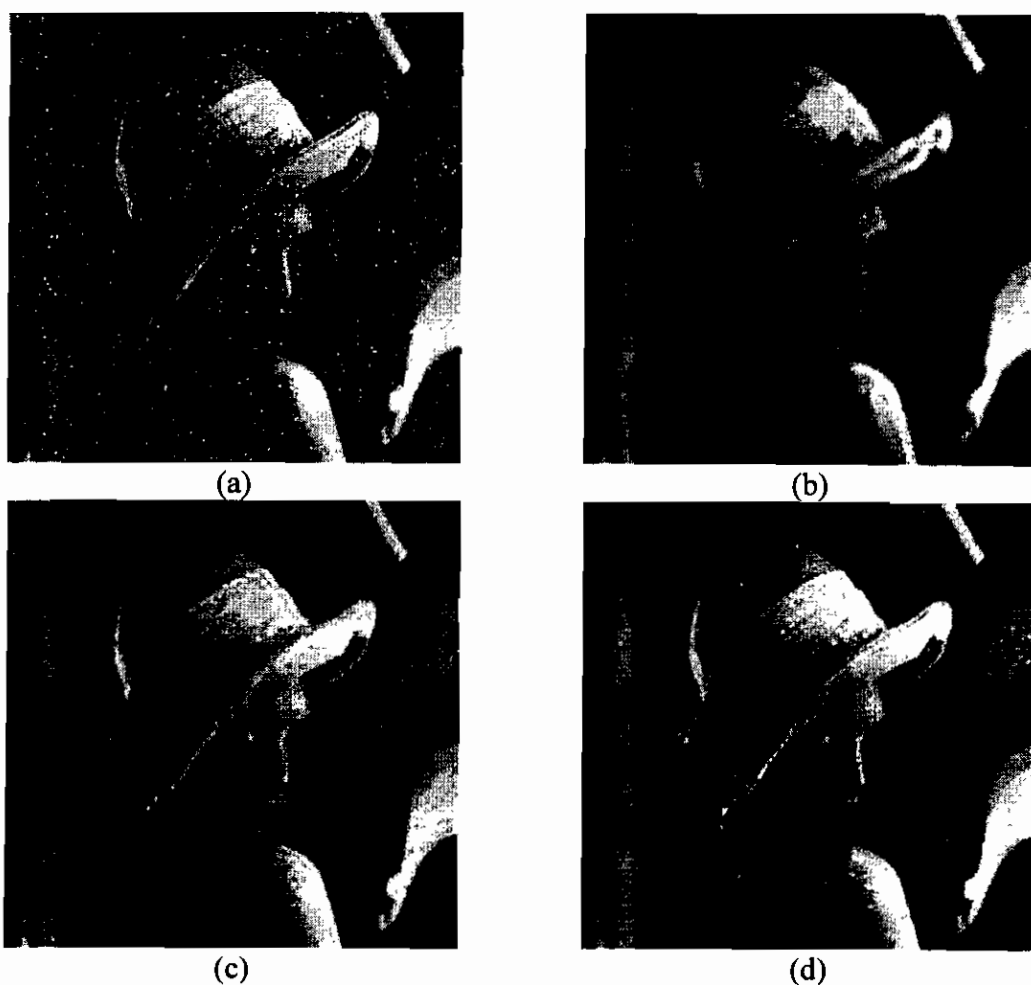


**Figura 6.3.2.1** Imagen filtrada usando filtro raíz con: (a)  $\alpha=0.5$ , SNR=13.8 dB. (b)  $\alpha=0.7$ , SNR=15.2 dB. (c)  $\alpha=1.1$ , SNR=26.3 dB. (d)  $\alpha=1.3$ , SNR=24.3 dB.

El filtro raíz actuando como pasa-altos resalta significativamente los pixeles que denotan alta frecuencia, como pasa-bajos trata de suavizar la imagen atenuando el contraste provocado por los impulsos positivos.

### 6.3.3 FILTRO PASA-BAJOS IDEAL

Luego de aplicar el filtro pasa-bajos ideal a la imagen ruidosa dada por la figura No 6.3.3.1a con tres valores de radio del filtro, se obtienen los siguientes resultados:

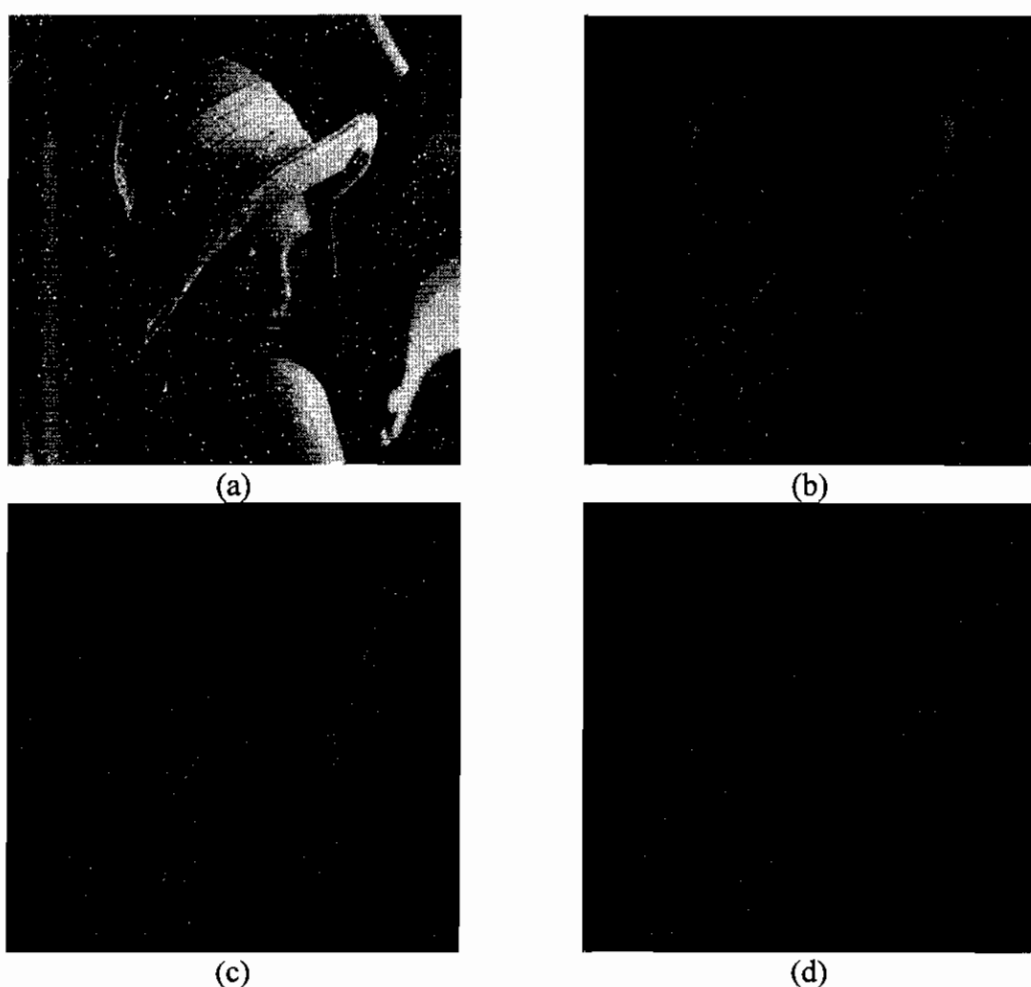


**Figura 6.3.3.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo positivo con probabilidad de ocurrencia de 0.01, SNR=25.2 dB. Imagen filtrada usando filtro pasa-bajos con: (b) radio = 50, SNR=23.2 dB. (c) radio = 100, SNR=24.8 dB. (d) radio =150, SNR=25.3 dB.

El filtro pasa-bajos suaviza el área formada por los impulsos y su contorno más no el área sin ruido, que aproximadamente se mantiene sin ningún efecto.

### 6.3.4 FILTRO PASA-ALTOS IDEAL

Filtrando la imagen ruidosa dada por la figura 6.3.4.1a con tres valores de radio del filtro pasa-altos ideal, se obtienen los siguientes resultados:

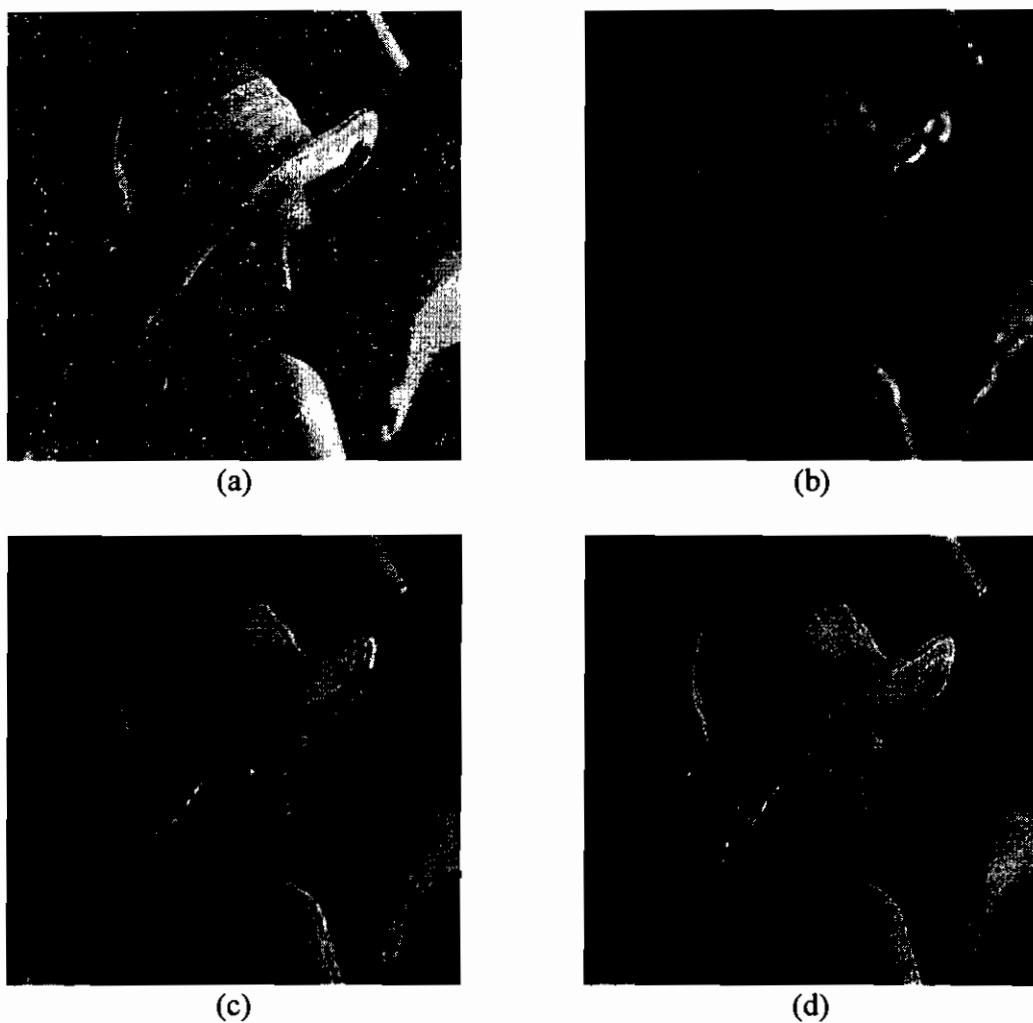


**Figura 6.3.4.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo positivo con probabilidad de ocurrencia de 0.01, SNR=25.2 dB. Imagen filtrada usando filtro pasa-altos con: (b) radio = 10, SNR=15.9 dB. (c) radio = 20, SNR=14.6 dB. (d) radio =50, SNR=13.2 dB.

El área resaltada por la aplicación de este filtro es aquella que en la imagen original tiene grandes contrastes

### 6.3.5 FILTRO HOMOMÓRFICO

Al aplicar el filtro homomórfico sobre la imagen de la figura 6.3.5.1a, y si el filtro lineal utilizado es el filtro pasa-bajos ideal se obtienen las siguientes imágenes resultantes,



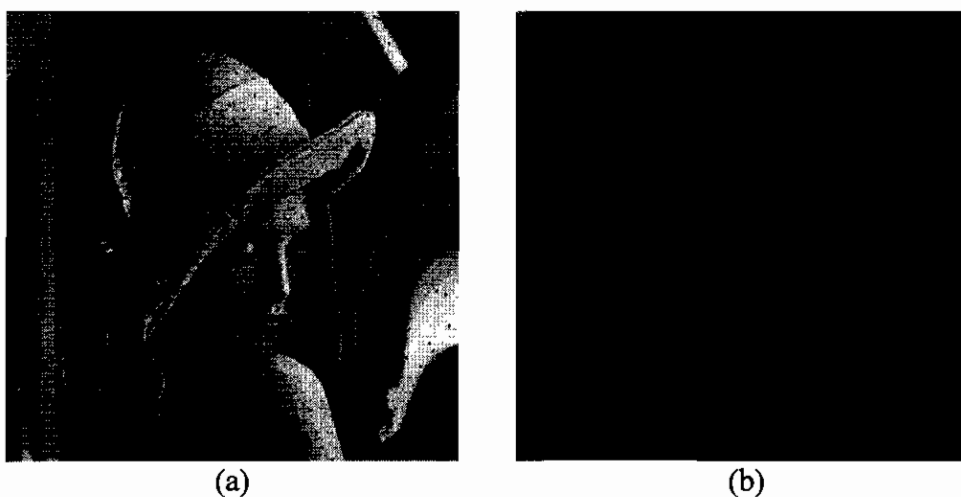
**Figura 6.3.5.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo positivo con probabilidad de ocurrencia de 0.01, SNR=25.2 dB. Imagen filtrada usando filtro homomórfico y filtro pasa-bajos con: (b) radio = 50, SNR=20.0 dB. (c) radio = 100, SNR=22.7 dB. (d) radio =150, SNR=24.4 dB.

El filtro homomórfico atenúa apreciablemente el efecto brillante causado por los impulsos positivos.



## 6.4 FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO ADITIVO IMPULSIVO NEGATIVO

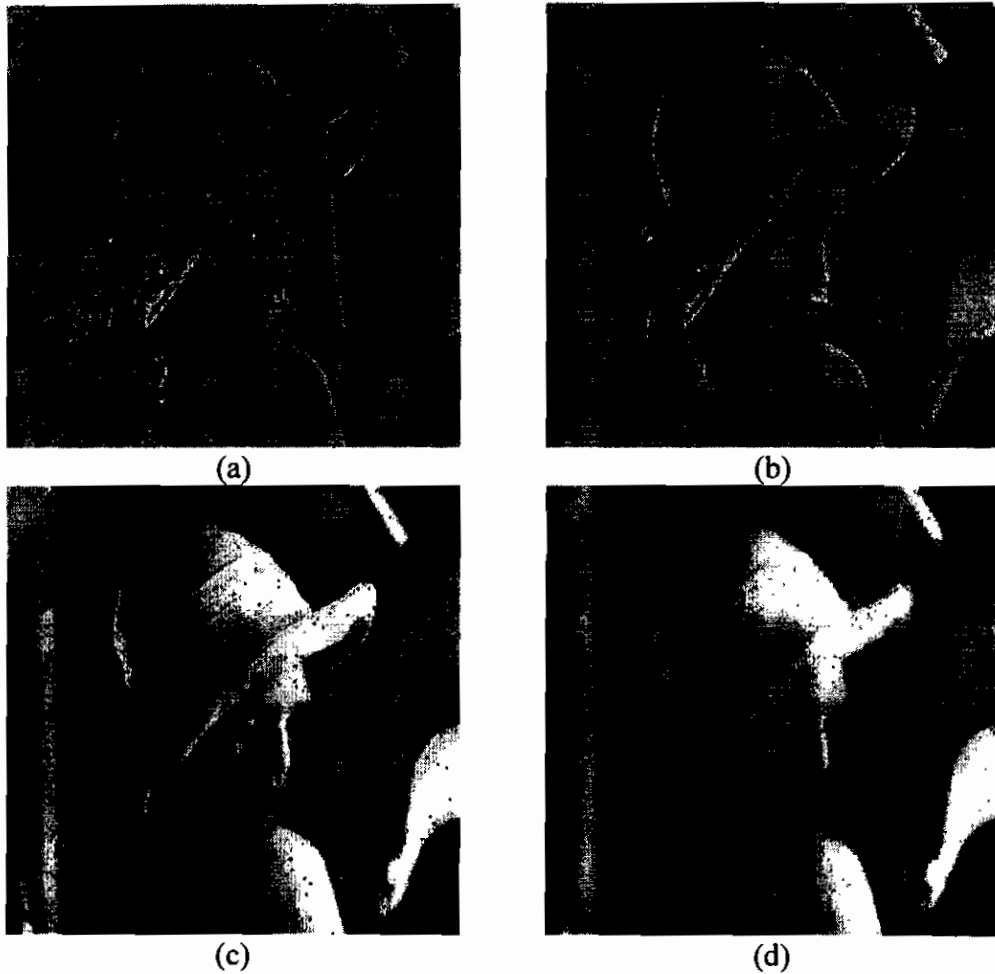
La figura 6.4.0.1a presenta la imagen original (la imagen original se muestra en la figura 6.1.0.1a) contaminada con ruido impulsivo negativo con probabilidad de ocurrencia del ruido de 0.01, acompañada de su respectivo espectro coseno.



**Figura 6.4.0.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo negativo con probabilidad de ocurrencia de 0.01, SNR=23.7 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

### 6.4.1 FILTRO RAÍZ

Filtrando la imagen ruidosa dada por la figura 6.4.0.1a con cuatro valores diferentes de coeficientes  $\alpha$  del filtro raíz, se obtienen los siguientes resultados:

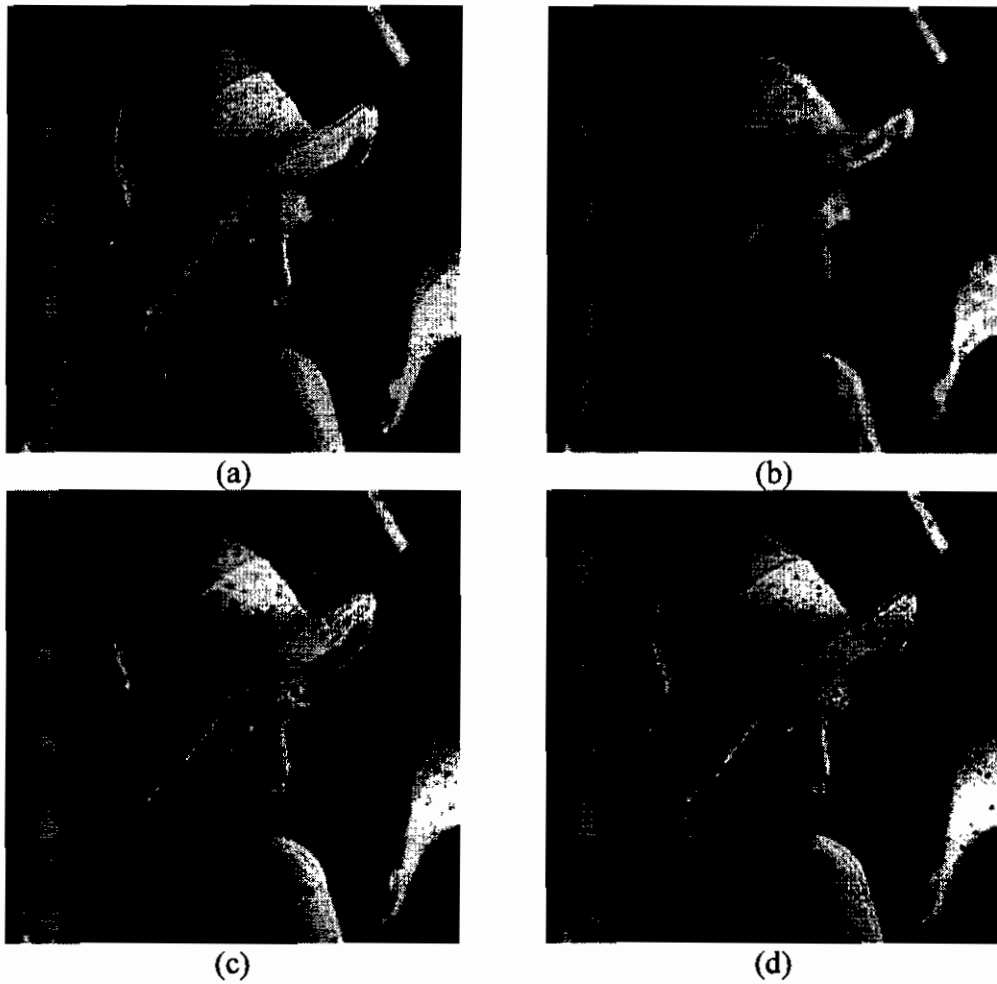


**Figura 6.4.1.1** Imagen filtrada usando filtro raíz con: (a)  $\alpha=0.5$ , SNR=12.9 dB. (b)  $\alpha=0.7$ , SNR=14.1 dB. (c)  $\alpha=1.1$ , SNR=22.8 dB. (d)  $\alpha=1.3$ , SNR=22.6 dB.

El filtro raíz así como en los casos anteriores resalta componentes de alta frecuencia o suaviza la imagen, dependiendo si se aplica como pasa-altos o pasa-bajos respectivamente.

## 6.4.2 FILTRO PASA-BAJOS IDEAL

Luego de aplicar el filtro pasa-bajos ideal a la imagen ruidosa dada por la figura 6.4.2.1a con tres valores de radio del filtro, se obtienen las siguientes imágenes:

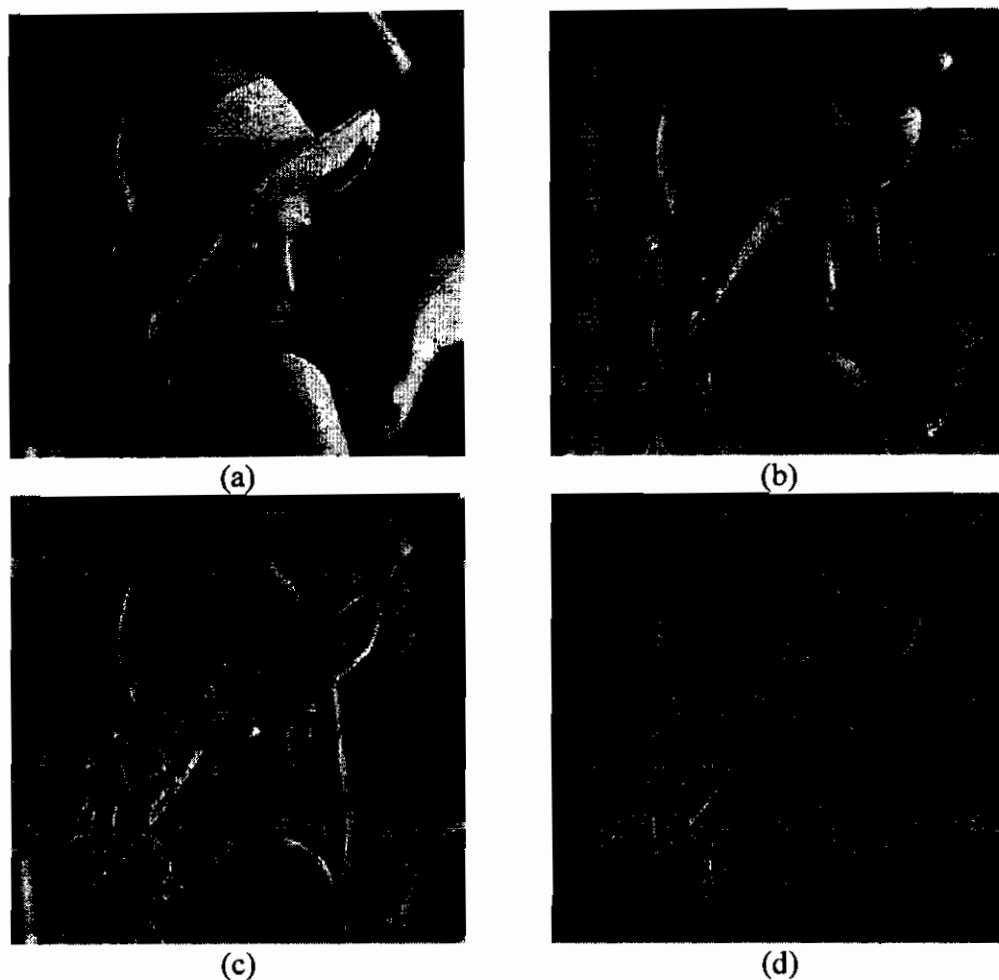


**Figura 6.4.2.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo negativo con probabilidad de ocurrencia de 0.01, SNR=23.7 dB. Imagen filtrada usando filtro pasa-bajos con: (b) radio = 50, SNR=23.5 dB. (c) radio = 100, SNR=25.8 dB. (d) radio = 150, SNR=26.8 dB.

Este filtro suaviza el área contaminada con los impulsos negativos, opacando la presencia de estos sobre áreas brillantes.

### 6.4.3 FILTRO PASA-ALTOS IDEAL

Filtrando la imagen ruidosa dada por la figura 6.4.3.1a con tres valores de radio del filtro pasa-altos ideal, se obtienen los siguientes resultados:

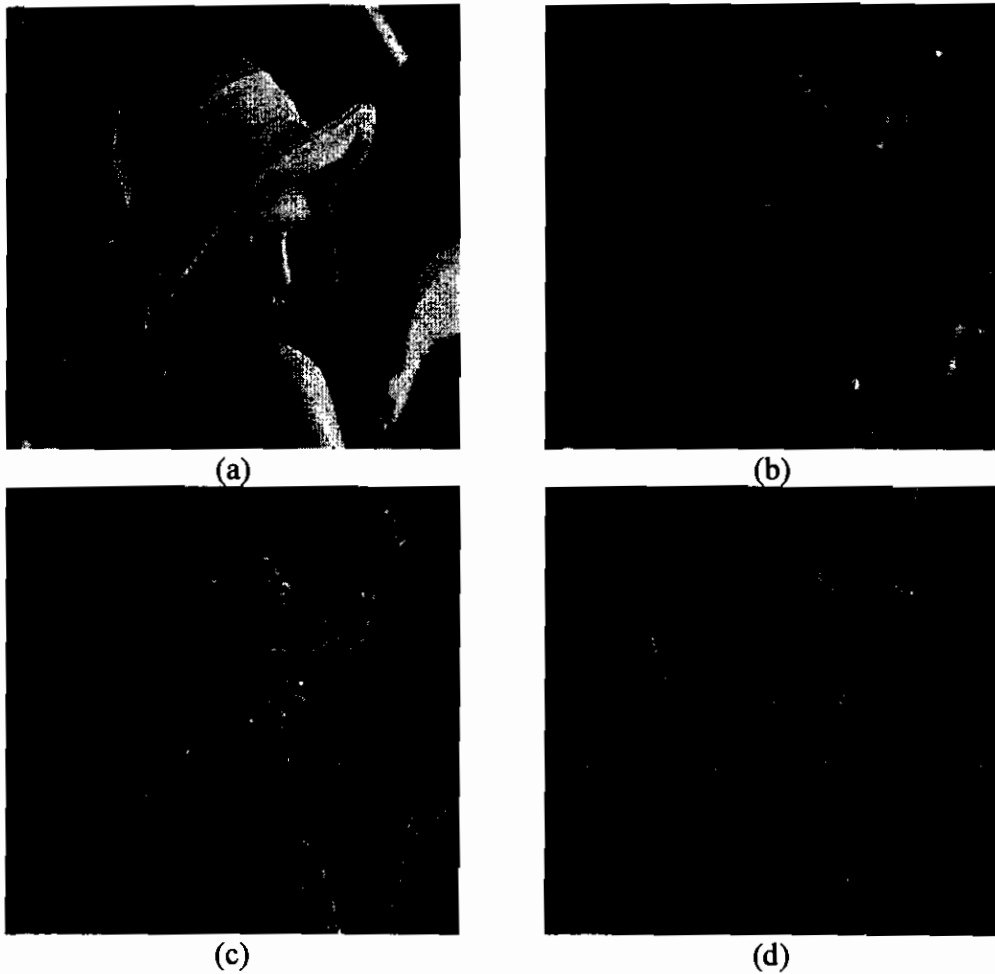


**Figura 6.4.3.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo negativo con probabilidad de ocurrencia de 0.01, SNR=23.7 dB. Imagen filtrada usando filtro pasa-altos con: (b) radio = 10, SNR=14.5 dB. (c) radio = 20, SNR=13.3 dB. (d) radio =50, SNR=13.1 dB.

Este filtro además de resaltar los contornos de la imagen, resalta el ruido en áreas donde su presencia es visible, esto es, en áreas claras.

#### 6.4.4 FILTRO HOMOMÓRFICO

Al aplicar el filtro homomórfico sobre la imagen 6.4.4.1a, y si el filtro lineal utilizado es el filtro pasa-bajos ideal se obtienen las siguientes imágenes resultantes,

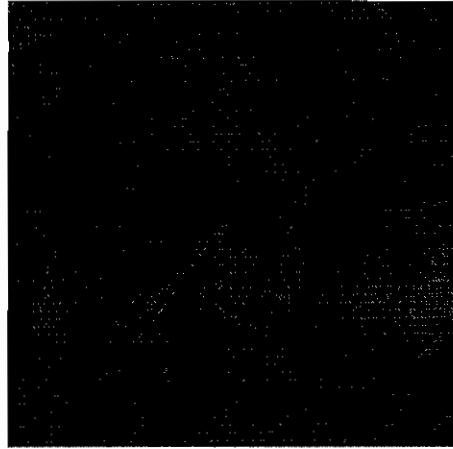


**Figura 6.4.4.1** (a) Imagen original ruidosa, con ruido aditivo impulsivo negativo con probabilidad de ocurrencia de 0.01, SNR=23.7. Imagen filtrada usando filtro homomórfico y filtro pasa-bajos con: (b) radio = 50, SNR=18.2 dB. (c) radio = 100, SNR=17.9 dB. (d) radio = 150, SNR=16.4 dB.

Filtro no apto para filtrar ruido impulsivo negativo, porque al tratar de atenuar los impulsos más bien empaña áreas aledañas a los impulsos.

### 6.4.5 FILTRO GAUSSIANO INVERSO

Para este caso se ha filtrado la imagen ruidosa dada por la figura 6.4.0.1a con filtro Gaussiano inverso, obteniéndose el siguiente resultado:

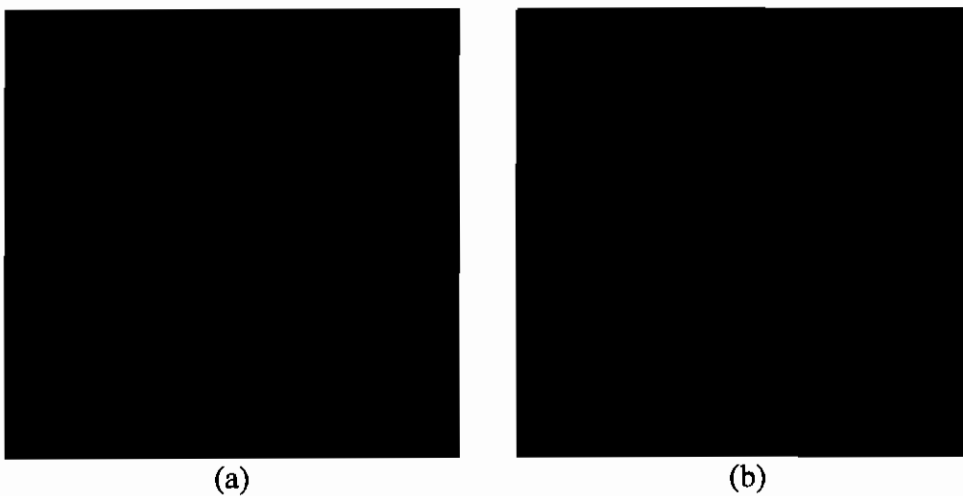


**Figura 6.4.5.1** Imagen filtrada usando filtro Gaussiano inverso, SNR=13.8 dB.

Este filtro no resulta de utilidad para filtrar imágenes contaminadas con ruido impulsivo

## 6.5 FILTRADO DE IMÁGENES CONTAMINADAS CON RUIDO DEPENDIENTE DE LA SEÑAL

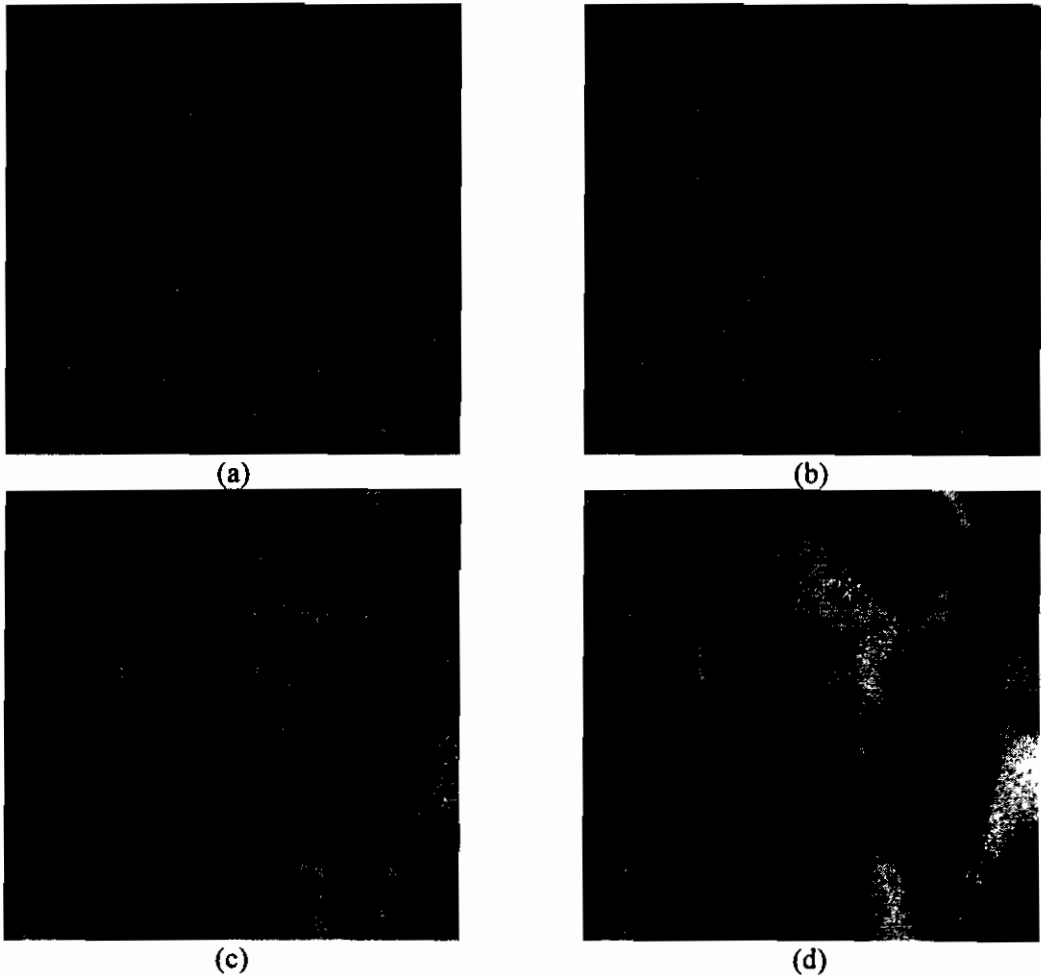
La figura 6.5.0.1a presenta la imagen original contaminada con ruido dependiente de la señal, generada según la expresión  $u_r = u^{0.7} + nu^{0.35}$ , la misma que se comenta en la sección (4.2.4).



**Figura 6.5.0.1** (a) Imagen original ruidosa, contaminada con ruido dependiente de la señal, SNR=19.2 dB. (b) Espectro de la transformada coseno de la imagen dada en (a).

### 6.5.1 FILTRO RAÍZ

Filtrando la imagen ruidosa dada por la figura 6.5.0.1a con cuatro diferentes valores de coeficientes alpha del filtro raíz, se obtienen las siguientes imágenes resultantes:



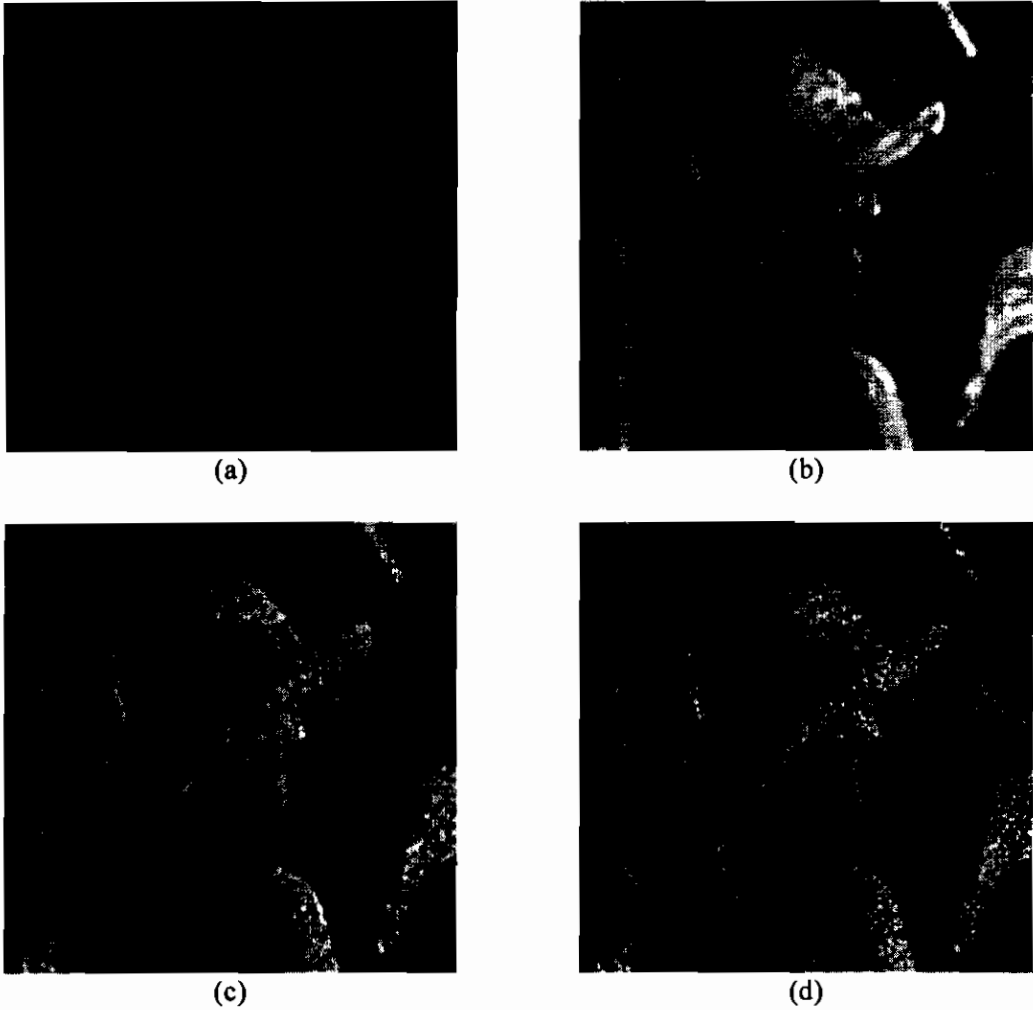
**Figura 6.5.1.1** Imagen filtrada usando filtro raíz con: (a)  $\alpha=0.5$ , SNR=15.2 dB. (b)  $\alpha=0.7$ , SNR=16.2 dB. (c)  $\alpha=1.1$ , SNR=20.7 dB. (d)  $\alpha=1.3$ , SNR=22.1 dB.

El ruido dependiente de la señal es el que mayor daño causa a la imagen, la imagen así contaminada tiene un contenido energético elevado de alta frecuencia y un rango dinámico reducido. El filtro raíz actuando como pasa-altos resalta el ruido, pero cuando actúa como pasa-bajos suaviza la imagen eliminando gran cantidad de ruido; en los dos casos se tiene un aumento del rango dinámico.

## 6.5.2 FILTRO PASA-BAJOS IDEAL

Luego de aplicar el filtro pasa-bajos ideal a la imagen ruidosa dada por la figura

6.5.0.1a con tres valores de radio del filtro, se obtienen los siguientes resultados:



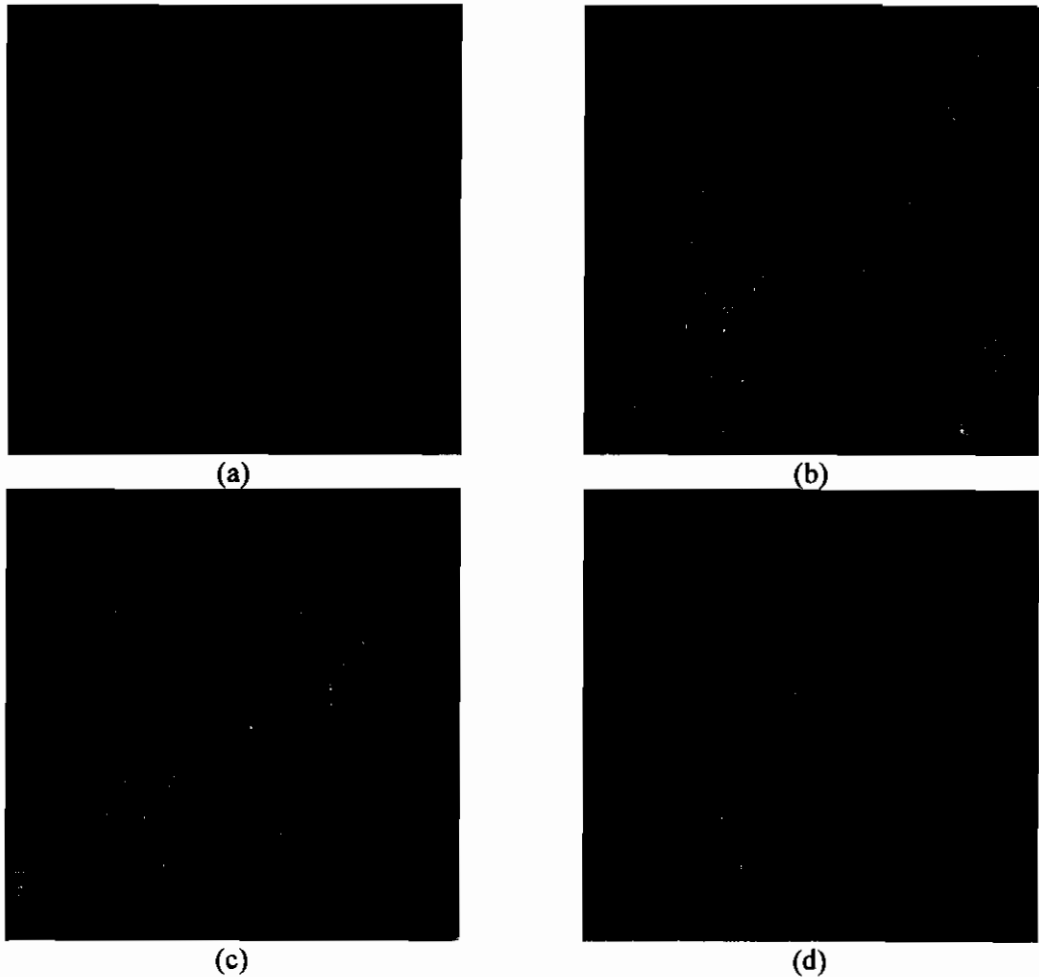
**Figura 6.5.2.1** (a) Imagen original ruidosa, contaminada con ruido dependiente de la señal, SNR=19.2 dB. Imagen filtrada usando filtro pasa-bajos con: (b) radio = 50, SNR=21.5 dB. (c) radio = 100, SNR=23.1 dB. (d) radio =150, SNR=22.0 dB.

Además de aumentar el rango dinámico, elimina gran cantidad de ruido pero no logra restaurar la imagen debido a la severa degradación causada por el ruido dependiente de la señal.



### 6.5.3 FILTRO PASA-ALTOS IDEAL

Filtrando la imagen ruidosa dada por la figura 6.5.0.1a con tres valores de radio del filtro pasa-altos ideal, se obtienen los siguientes resultados:

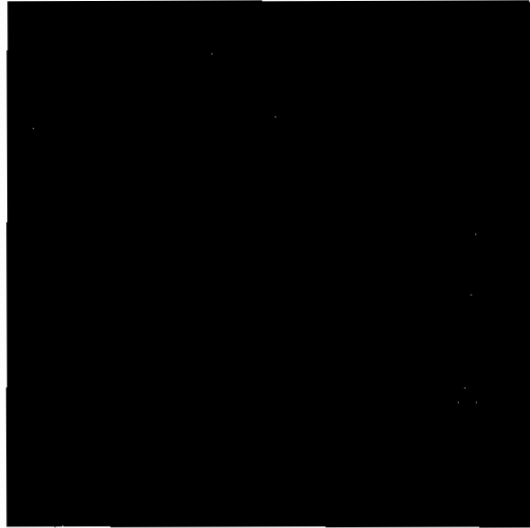


**Figura 6.5.3.1** (a) Imagen original ruidosa, contaminada con ruido dependiente de la señal, SNR=19.2 dB. Imagen filtrada usando filtro pasa-altos con: (b) radio = 10, SNR=15.6 dB. (c) radio = 20, SNR=14.8 dB. (d) radio =50, SNR=14.0 dB.

La imagen filtrada con filtro pasa-altos tiene tal cantidad de ruido que tiende a confundirse con los contornos propios de la imagen, pues este filtro por su naturaleza deja pasar gran cantidad de ruido a la vez que atenúa componentes energéticas de la imagen misma (baja frecuencia).

### 6.5.4 FILTRO GAUSSIANO INVERSO

En éste caso se ha encontrado que filtrando la imagen ruidosa dada por la figura 6.5.0.1a con filtro Gaussiano inverso, se obtiene una imagen seriamente alterada, según se puede constatar en la figura resultante mostrada a continuación:

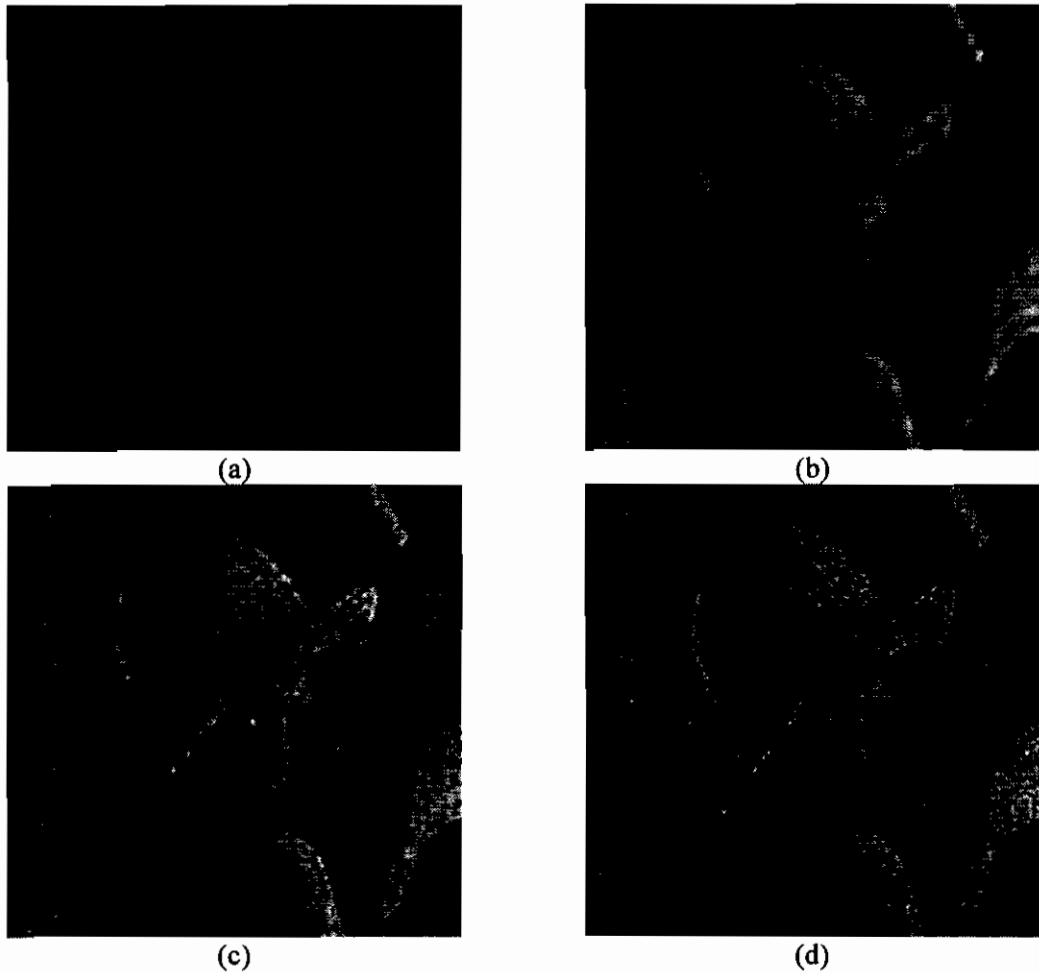


**Figura 6.5.4.1** Imagen filtrada usando filtro Gaussiano inverso, SNR=13.2 dB.

La severidad con que el ruido dependiente de la señal daña a la imagen original, se ve complementada con el efecto de la característica pasa-altos de este filtro, de tal manera que, con este proceso prácticamente se ha recuperado solamente ruido. Si se analiza la función de transferencia del filtro gaussiano inverso se pueden explicar satisfactoriamente los resultados obtenidos al aplicarlo.

### 6.5.5 FILTRO HOMOMÓRFICO

Al aplicar el filtro homomórfico sobre la imagen 6.5.0.1a, y si el filtro lineal utilizado es el filtro pasa-bajos ideal se obtienen las siguientes imágenes resultantes,



**Figura 6.5.5.1** (a) Imagen original ruidosa, contaminada con ruido dependiente de la señal, SNR=19.2 dB. Imagen filtrada usando filtro homomórfico y filtro pasa-bajos con: (b) radio = 50, SNR=20.0 dB. (c) radio = 100, SNR=21.0 dB. (d) radio =150, SNR=19.4 dB.

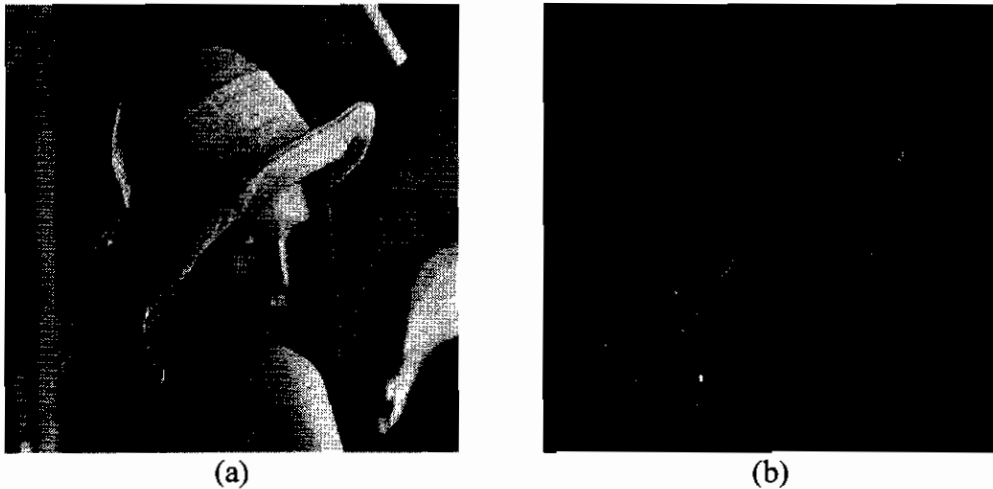
El filtro homomórfico reduce gran cantidad de ruido, los detalles finos se ven opacados y el rango dinámico aumentado, las características resultantes son en cierta medida ventajosas con respecto a las proporcionadas por los demás filtros.

## 6.6 FILTRADO HOMOMÓRFICO DE IMÁGENES NO CONTAMINADAS

En el capítulo cuarto se presenta la teoría del filtro homomórfico, en dicha exposición se divide el tema en dos partes: la primera encargada de mejorar la imagen, resaltando las componentes de alta frecuencia; y, la segunda encargada de filtrar el ruido contaminante

cuando este ha sido añadido, como cualquier ruido contribuye significativamente con componentes de alta frecuencia, cuando hay presencia de ruido no se ve conveniente resaltarlas, en este caso se aplicará la teoría que hace referencia específica a la eliminación del ruido. En las secciones anteriores de este capítulo se aplica la teoría de eliminación de ruido, pues las pruebas se las hace con imágenes contaminadas, pero queda pendiente exponer los resultados de aplicar el filtro homomórfico para resalte de altas frecuencias; esto es lo que se hace a continuación.

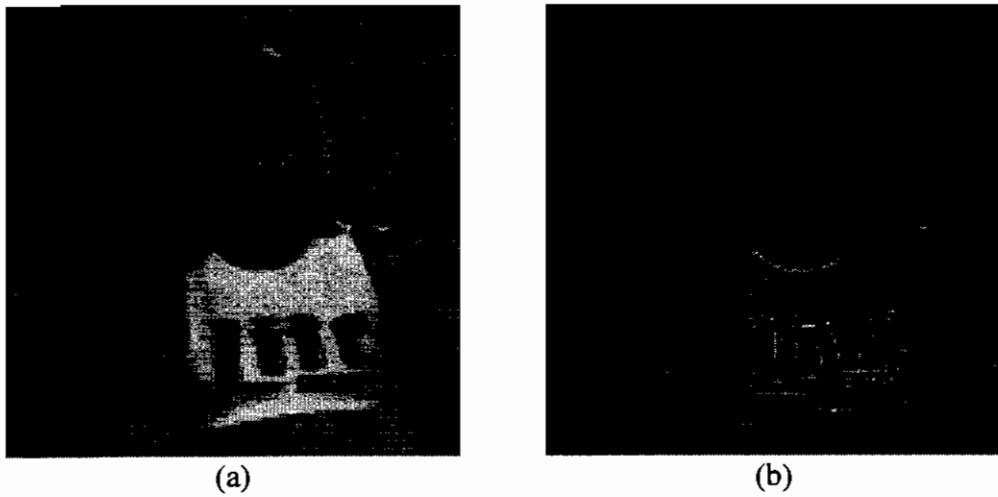
En las siguientes gráficas se presentan tanto las imágenes originales como aquellas resultantes de la aplicación del filtro homomórfico, en dicho filtro se utiliza la función logarítmica como la encargada de convertir la no linealidad en linealidad y la función de transferencia dada por la figura No. 4.2.4.2.



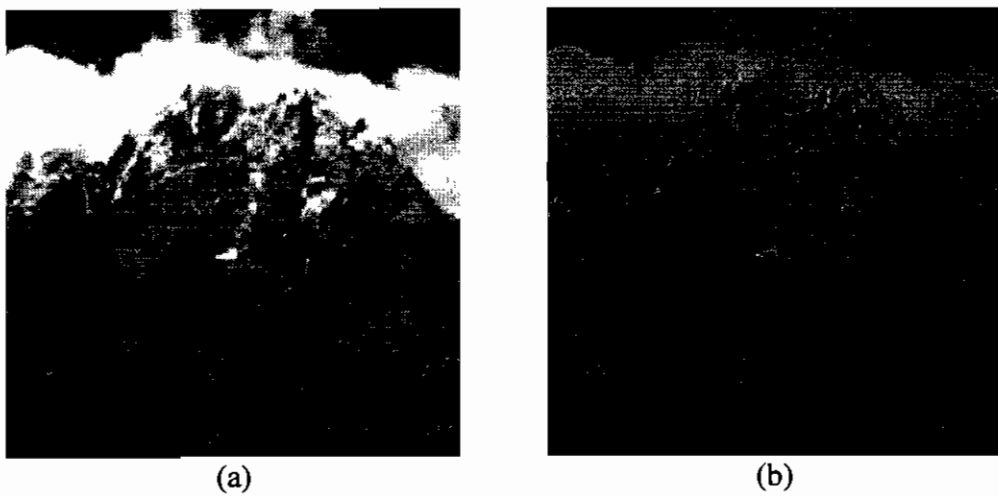
**Figura 6.6.0.1.** (a) Imagen original. (b) Imagen resultante de aplicar filtro homomórfico a la imagen original.



**Figura 6.6.0.2.** (a) Imagen original. (b) Imagen original filtrada usando filtro homomórfico



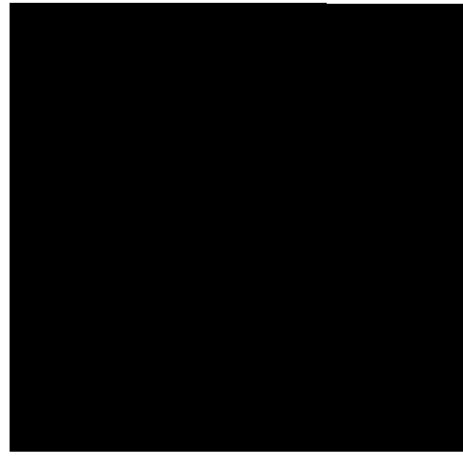
**Figura 6.6.0.3.** (a) Imagen original. (b) Imagen original filtrada usando filtro homomórfico



**Figura 6.6.0.4.** (a) Imagen original. (b) Imagen original filtrada usando filtro homomórfico

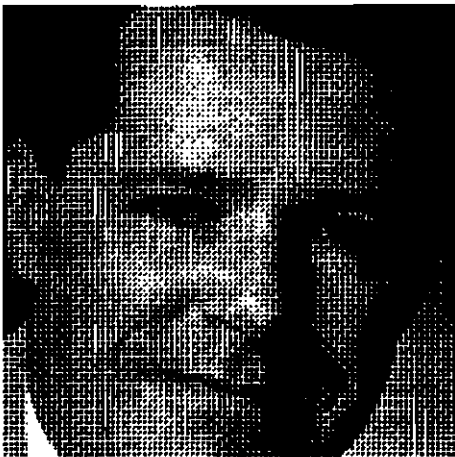


(a)



(b)

**Figura 6.6.0.5.** (a) Imagen original. (b) Imagen original filtrada usando filtro homomórfico



(a)



(b)

**Figura 6.6.0.6.** (a) Imagen original. (b) Imagen original filtrada usando filtro homomórfico



(a)



(b)

**Figura 6.6.0.7.** (a) Imagen original. (b) Imagen original filtrada usando filtro homomórfico

La función de transferencia dada por la figura No. 4.2.4.2 atenúa componentes de baja frecuencia mientras realza componentes de alta frecuencia, en los ejemplos anteriores se puede comprobar que cuando la imagen es suave, ésta es resaltada de una manera homogénea; en cambio cuando la imagen es áspera solamente se ven resaltados los puntos de más alta frecuencia (contornos más pronunciados).

## ANÁLISIS DE RESULTADOS

Con propósito de análisis y de acuerdo a la respuesta de frecuencia, a los filtros se los clasifica en pasa-bajos y pasa-altos. Como pasa-bajos encontramos: al filtro raíz, cuando su coeficiente  $\alpha$  es mayor que la unidad, al filtro pasa-bajos ideal y al filtro homomórfico; como filtros pasa-altos se tiene: al filtro raíz cuando el coeficiente  $\alpha$  es menor que la unidad, al pasa-altos ideal y al filtro gaussiano inverso.

Filtro pasa-bajos ideal.- este filtro atenúa completamente todas aquellas frecuencias espaciales mayores que la frecuencia de corte (radio pasa-bajos), mientras que las frecuencias menores o iguales a esta tienen una ganancia unitaria. Como se manifiesta en la teoría y se comprueba con los ejemplos su inconveniente radica en el hecho que empaña la imagen debido a la transición brusca en la frecuencia de corte. A parte de este problema suprime en buena medida ruido gaussiano blanco y ruido impulsivo.

Existe un radio pasa-bajos óptimo para el cual la relación señal a ruido es máxima (este radio aumenta cuando la concentración de ruido es menor), al alejarse a cualquier lado de dicho radio la relación señal a ruido disminuye. Cuando aumenta lo hace porque la cantidad de ruido que pasa es mayor, en cambio cuando disminuye se recorta señal válida (datos de imagen pura).

Filtro pasa-altos ideal.- este filtro tiene un comportamiento totalmente opuesto al anterior, esto es, atenúa completamente las frecuencias menores a la frecuencia de corte dejando pasar sin ganancia las frecuencias superiores. Al aumentar el radio pasa-altos (frecuencia de corte) resalta el ruido y los contornos de la imagen, porque estos están principalmente descritos por componentes de alta frecuencia y esto se verifica con la reducción de la relación señal a ruido.

Filtro raíz.- Dependiendo del valor de su coeficiente alpha este puede actuar como filtro pasa-bajos o como pasa-altos.

- Cuando el coeficiente es menor que la unidad (siempre positivo) las componentes de alta frecuencia se ven resaltadas (debe quedar claro que tanto las componentes de alta como de baja frecuencia se atenúan, lo que ocurre es que las de alta frecuencia se ven menos atenuadas con respecto a la amplitud de ellas antes de aplicar el filtro). Por este motivo en los ejemplos se pueden notar la presencia también de detalles suaves, solamente que disminuidos en contraste, es decir blancos siguen con tendencia a este tono, igual los tonos negros tienden a seguir negros (si hay puntos negros contaminantes entonces la imagen se torna blanca y al contrario con puntos contaminantes blancos la imagen tiende a volverse negra).
- Si el coeficiente alpha es mayor que la unidad se tiene un comportamiento pasa-bajos, en estas condiciones tanto las bajas como las altas frecuencias se ven amplificadas, pero en una proporción mayor las bajas frecuencias, la relación señal a ruido aumenta con alpha hasta cierto punto donde dicha relación comienza a disminuir y al mismo tiempo la imagen tiende a ponerse blanca; en consecuencia reduce de mejor manera el ruido gaussiano blanco y el impulsivo positivo.



Filtro gaussiano inverso.- El comportamiento de este filtro depende de factores como: el tamaño y la varianza de la imagen. Al aumentar el tamaño de la imagen la inclinación hacia la respuesta pasa-altos es más pronunciada; cuando la imagen tiene tonalidad oscura su varianza es baja, en este caso también se comporta como pasa-altos. El efecto de este filtro es amplificar todas las frecuencias, pero el rango de alta frecuencia con mayor pronunciación, por esto es inadecuado utilizarlo para cualquier tipo de ruido.

Filtro homomórfico.- El filtro homomórfico se separa en dos partes: una destinada para resaltar componentes de alta frecuencia (en consecuencia debe aplicarse sobre imágenes sin ruido) este filtro resalta las componentes de más alta frecuencia en concordancia con la respuesta de frecuencia presentada en su análisis en el capítulo cuarto. El filtro homomórfico para eliminación del ruido, según estudio realizado en el mismo capítulo utiliza alinealidades del tipo  $x^n$  para ruido dependiente de la señal y  $\ln(x)$  para el ruido multiplicativo y aditivo, en ambos casos se complementa con el uso del filtro lineal pasa-bajos ideal. Este filtro para ruido sal y pimienta e impulsivo negativo definitivamente no es adecuado, para ruido gaussiano blanco quita detalles, para impulsivo positivo es aceptable así como también para dependiente de la señal.

Resumiendo, para ruido gaussiano blanco, el filtro pasa-bajos ideal se ve más efectivo que el homomórfico, pero el raíz (actuando como pasa-bajos) conserva los detalles mejor que todos; para ruido sal y pimienta el filtro pasa-bajos ideal y el raíz lo atenúan en cierta medida; para ruido impulsivo positivo el homomórfico es bastante aceptable; para ruido impulsivo negativo el filtro raíz y pasa-bajos ideal lo suprimen en cierto grado; para ruido dependiente de la señal resulta mejor contraste y menos empañamiento al aplicar el filtro homomórfico.

# CAPÍTULO

---



## COMENTARIOS, CONCLUSIONES Y RECOMENDACIONES

### 7.1 COMENTARIOS

- Debe recordarse que de entre todas las transformadas imagen, la transformada discreta coseno tiene la más alta compresión energética (después de la transformada Karhunen-Loeve, pero la TKL es un tanto ideal), de ahí que en el dominio de la frecuencia la energía está concentrada cerca del origen. Por tanto el atenuar componentes de frecuencia cerca del origen afecta gravemente la característica suave de las imágenes típicas, resultando en su lugar cambios bruscos de intensidad, indicadores principales de perfiles o contornos.
- Idealmente la transformada KL es la mejor en lo que se refiere a la propiedad de compresión energética, pero el cálculo de esta presenta un serio inconveniente práctico, pues las funciones bases dependen de los datos, porque la matriz TKL está

calculada en base a la matriz covarianza de la matriz de entrada que representa a la imagen dato. Afortunadamente para imágenes altamente correlacionadas la TKL puede ser sustituida por la transformada discreta coseno.

- Sobre la transformada discreta de Fourier se dice también que tiene una buena compresión energética, afirmación confirmada en el análisis realizado sobre el tema en la referencia [10]. Sin embargo, realizando una comparación tomando como base un mismo número de coeficientes transformados iniciales se encuentra que la TDC tiene una mejor compresión energética.
- En las propiedades descritas en la sección 3.1.1 se verifica que para coeficientes de correlación  $\rho \geq 0.9$  la transformada coseno se aproxima a la KL, encontrando aplicación en compresión y codificación de imágenes.
- La propiedad de decorrelación de la TKL establece que cuando los elementos del vector de entrada están altamente correlacionados, esta empaqueta el máximo de energía en un número dado de coeficientes transformados no correlacionados entre sí, es decir que la información contenida en un coeficiente no depende de aquella contenida en otro coeficiente transformado.
- La propiedad Función Tasa de Distorsión podría también interpretarse de la siguiente manera: para un número de bits constante, la transferencia de coeficientes transformados provocará la menor distorsión cuando dichos coeficientes pertenezcan a la TKL, redundando nuevamente en la propiedad de compresión energética.

- El coeficiente de normalización de la matriz TDC tanto en sentido directo como en sentido inverso es  $\frac{2}{\sqrt{MN}}$ , escribiéndolo en un sólo sentido este se convierte en  $\frac{4}{MN}$ , la primera modalidad involucra mayor número de operaciones ( $2MN$ ) pues se realiza la misma operación en las dos direcciones, la segunda alternativa solamente requiere ( $MN$ ) operaciones de normalización. Sin embargo se ha optado por utilizar la primera porque esta resulta en la aplicación de la matriz transformada coseno unitaria.
- Se demuestra como la primera propiedad de la TDC que su matriz es ortogonal, tomando en cuenta que la matriz TDC es real se concluye que es también unitaria. Así mismo cualquier transformada real representada por una matriz ortogonal cumplirá también con ser unitaria.
- Adicionalmente a lo que se menciona en la segunda propiedad que establece que para filas pares de la matriz TDC los cuadrantes izquierdo y derecho son iguales y para filas impares son iguales pero opuestas en signo, debe indicarse que dichos cuadrantes no son cualquiera sino son aquellos resultantes de tal manera que la matriz dato queda ordenada en bloques par-impar.
- La TDC no es la parte real de la TDF, sin embargo existen expresiones matemáticas que las relacionan, expresiones que han sido utilizadas para desarrollar algoritmos de cálculo de la TDC partiendo de la TDF, estos algoritmos no son tan eficientes como los implementados en este trabajo que opera directamente sobre el arreglo 2-D y no

sobre filas y columnas separadamente como lo hace el implementado en base a la TDF.

- Puesto que gran parte de la energía está concentrada en la región de baja frecuencia en una imagen, esta puede ser reconstruida sin pérdida significativa de su calidad e inteligibilidad con sólo una fracción de los coeficientes transformados. El software desarrollado permite la opción de filtrado pasa-bajos ideal, esta opción puede utilizarse para averiguar el contenido energético de una imagen reconstruida con determinado número de coeficientes.
- El rango dinámico del espectro coseno es normalmente mucho mayor de aquel que los dispositivos de visualización pueden desplegar fielmente, en cuyo caso sólo aquellas partes más brillantes son visibles. Una solución a este problema es graficar en lugar del espectro una función de aquel dada por la expresión (3.3.0.1), manipulando sus constantes es posible ubicar el rango dinámico dentro de los límites requeridos, límites establecidos de 0 y 255 para el caso de una imagen monocromática de 256 tonos de gris. Durante las pruebas para encontrar el valor adecuado de las constantes se encontró que la constante  $K_1$  cumple con la ecuación (3.3.0.2); el problema surgió durante la selección de  $K_2$ , se ha encontrado que este valor depende tanto de la imagen como de la capacidad visual del observador.
- Así como existen imágenes típicas que tienen grandes regiones donde la intensidad se mantiene aproximadamente constante, generando un espectro con gran parte de su energía concentrada en regiones cercanas al origen en el dominio de la frecuencia, existen también imágenes con espectro de la transformada coseno de difícil

predicción como son las imágenes binarias que representan fotografías, en las cuales similares de variación de tonalidad se generan cambiando la densidad de puntos blancos.

- El espectro de imágenes que no es posible visualizar debido a su estructura fina es el de aquellas formadas por líneas horizontales ó verticales, la primera genera un espectro formado por una línea vertical (esto es,  $\{v(k,0) \quad 0 \leq k \leq M - 1\}$ ) indicando discontinuidades o contornos sólo en sentido vertical. El espectro formado sólo por líneas verticales en cambio está dado por el conjunto  $\{v(0,l) \quad 0 \leq l \leq N - 1\}$  indicando discontinuidades solamente en dirección horizontal.
- Una imagen puede ser de menor calidad por varias razones por ejemplo, proceso fotográfico imperfecto, ruido del canal, ruido de cuantización, etc; como la naturaleza de las imperfecciones son variadas existen también diferentes técnicas de mejoramiento, sin embargo técnicas que trabajen adecuadamente para un tipo de imperfecciones pueden muy bien no ser de utilidad para otras. Los ruidos utilizados para contaminar las imágenes de prueba tratan de simular algunos tipos de imperfecciones.
- El filtro raíz en sus dos modalidades y el filtro Gaussiano inverso producen cambios significativos en el rango dinámico de la imagen filtrada. Para el cálculo de la relación señal a ruido todos los rangos desviados son desplazados al rango inicial dado por el rango de la imagen de referencia, en este caso la imagen no contaminada.

## 7.2 CONCLUSIONES

- El tiempo de procesamiento para el cálculo tanto de la transformada directa como de la transformada inversa depende de algunos factores como por ejemplo: el método de aproximación, algoritmo, normalizaciones, lenguaje de programación (acceso a disco o memoria) y microprocesador utilizado. Entonces para comparar la rapidez de cálculo al menos se requeriría en la implementación de las transformadas lenguaje de programación e igual microprocesador. Adicionalmente debe indicarse que en el cálculo de la TDC se requiere solamente de la realización de operaciones reales en contraste con lo que ocurre por ejemplo en el caso de la TDF que necesita operaciones con números complejos.
- El ruido aditivo gaussiano blanco, el ruido impulsivo y el ruido dependiente de la señal tienen espectro coseno con componentes de frecuencia dispersos por todo el dominio de la frecuencia espacial de una manera prácticamente uniforme. Al aumentar la varianza o la probabilidad de ocurrencia del ruido según el caso, el espectro está así mismo uniformemente distribuido dentro del dominio de la frecuencia, pero el número de componentes de frecuencia no nulos aumenta. En todos los casos siempre habrá cierta cantidad de energía adicional concentrada en regiones de baja frecuencia.
- Lo expresado hasta el momento se puede resumir diciendo que, el espectro coseno de una imagen típica contaminada con cualquiera de los ruidos puede dividirse en dos zonas: *regiones de baja frecuencia* con espectro tanto de la imagen como del ruido, con componentes importantes debidas a la imagen pura que predominan porque los píxeles de esta si están mutuamente correlacionados, a diferencia de los píxeles del

ruido que no lo están. Las *regiones de alta frecuencia* en cambio tienen contribución energética pobre debido a la imagen pura, mientras que las contribuciones debidas a los pixeles del ruido son significativas.

- Lo anterior explica por qué al aplicar un filtro pasa-altos resalta los perfiles cuando la imagen no es ruidosa, y el ruido más los perfiles cuando si es ruidosa; reduciendo la relación señal a ruido tanto más cuanto más alto es el radio pasa-altos.
- El filtro raíz con coeficiente alpha menor que uno se comporta como filtro pasa-altos resaltando por tanto componentes de alta frecuencia en consecuencia conforme se acentúa la respuesta pasa-altos (disminuyendo alpha) disminuye la relación señal a ruido porque el ruido se resalta al mismo tiempo que la característica suave de la imagen se deteriora. En cambio cuando alpha es mayor que 1 el filtro raíz se comporta como pasa-bajos suavizando la imagen ruidosa conforme dicho coeficiente se aleja de la unidad.
- El comportamiento pasa-bajos o pasa-altos del filtro raíz para los valores del coeficiente alpha indicado es válido solamente para imágenes típicas, para las demás imágenes no se garantiza ningún tipo de comportamiento.
- Dependiendo de la concentración del ruido contaminante y de la imagen misma existe un punto óptimo donde la relación señal a ruido es máxima (mayor que la relación en la imagen ruidosa) para el filtro pasa-bajos ideal, si el radio del filtro se aleja del óptimo disminuye la relación señal a ruido. Cuando el radio es menor que el óptimo se suprimen también datos de la imagen, en cambio cuando el radio es mayor pasa el ruido existente desde el inicio.



- El filtro gaussiano inverso es un filtro eminentemente pasa-altos consecuentemente al aplicarlo a una imagen ruidosa resulta más bien en un deterioro de la imagen.
- El filtro homomórfico se ha dividido en dos tipos: uno destinado al mejoramiento aplicable exclusivamente a imágenes no contaminadas, puesto que su característica pasa-altos resalta contornos, si la imagen es contaminada más bien se recupera sólo el ruido; el segundo está destinado a filtrar ruido, para el efecto debe ser acompañado de un filtro pasa-bajos lineal, para este caso el pasa-bajos ideal (por ser el único filtro pasa-bajos lineal aquí analizado), su comportamiento es análogo al pasa-bajos ideal simple con una relación señal a ruido un tanto inferior.
- Se ha dicho que no existe una teoría unificada de mejoramiento de imágenes por la dificultad de definir un criterio que califique la calidad de la imagen, y muchas veces esta está definida en términos subjetivos, dejando tal responsabilidad a cargo del observador. La relación señal a ruido es un parámetro utilizado para evaluar el rendimiento de un filtro digital sobre una imagen ruidosa, criterio que necesita un parámetro de referencia y este es la imagen no contaminada. Por lo tanto en mejoramiento de imágenes el encontrar una relación señal a ruido alta no es suficiente para calificar a la imagen como mejorada, necesariamente el observador se reserva el derecho a calificar a la imagen mejorada como tal. Así mismo una imagen mejorada calificada por un observador puede no serla para otro, debido al carácter subjetivo del criterio.
- El coeficiente alpha correspondiente al filtro raíz actuando como pasa-bajos tiene un valor empírico, pues para valores cercanos a 2 a pesar de remover gran cantidad de ruido, la imagen tiende a volverse totalmente blanca debilitando los detalles finos.

Existe además un valor de  $\alpha$  óptimo donde la relación señal a ruido es máxima y a partir del cual dicha relación disminuye, este valor de  $\alpha$  es mayor cuando la concentración de ruido es mayor.

- Todos aquellos filtros pasa-altos no resultan adecuados para remover ruido, puesto que todos los ruidos tienen componentes significativas de alta frecuencia, pero en imágenes no ruidosas resultan adecuados para detección de contornos.
- El comportamiento del filtro gaussiano inverso resulta ser dependiente del tamaño de la imagen, para imágenes pequeñas ( $128 \times 128$  píxeles) es un filtro pasa-altos suave, en cambio para imágenes más grandes produce una severa alteración de la información contenida en la imagen.
- El filtro homomórfico resulta adecuado para filtrar ruido impulsivo positivo, reduciendo en la imagen el contraste producido por los puntos blancos, no así ante los impulsos negativos para los cuales resulta totalmente inadecuado, pues produce manchas negras apreciables en la ubicación original de los puntos negros.
- El ruido dependiente de la señal produce un severo deterioro de las imágenes a más de reducir su rango dinámico, el filtro pasa-bajos resulta relativamente bueno para remover este tipo de ruido, pero el filtro homomórfico se ve un tanto más beneficioso en la eliminación del ruido y en mejorar el contraste.
- De conformidad con lo expuesto en el cuarto capítulo sobre filtrado del ruido dependiente de la señal usando el filtro homomórfico, se espera una eliminación

considerable de este ruido, pero ello no ocurre, la causa más probable es que el filtro lineal pasa-bajos utilizado en el dominio homomórfico no sea el adecuado.

### 7.3 RECOMENDACIONES

Como recomendaciones se presentan algunas sugerencias sobre el uso del software implementado, así como algunos tópicos aquí mencionados pero no profundizados que quedarían a consideración para futuras investigaciones.

- El software implementado está capacitado para abrir solamente archivos gráficos con formato BMP mediante la opción de menú Archivo/Abrir; en caso de disponer de imágenes con otro formato gráfico deberá convertirse a formato BMP, o utilizando algún utilitario ubicar la imagen en el portapapeles y capturarla mediante la opción Edición/Pegar.
- Así mismo para obtener una copia de alguna de las imágenes presentes en pantalla ya sea en archivo o papel, debe copiarse al portapapeles mediante la opción Edición/Exportar y disponer de la imagen según convenga.
- Para imágenes con niveles de gris bajos es muy posible que el filtro gaussiano inverso provoque un sobreflujo matemático, truncando el funcionamiento del programa lo que producirá pérdida de la información, por tanto no deberá usarse indiscriminadamente este filtro.
- Si cuando el programa solicita algún parámetro se proporcionan valores incoherentes el programa responderá no ejecutando la orden encomendada.

- Como el programa está diseñado para trabajar bajo windows, se ha optado por no incluir un sistema de impresión o de salida hacia archivo por la facilidad de interactuar mediante el portapapeles con editores de imágenes. Una posible mejora consistiría en añadir los módulos en cuestión.
- La propiedad de la transformada coseno titulada *propiedad de desplazamiento* demostrada en el capítulo 3, presenta una manera de obtener la transformada coseno de una secuencia desplazada tomando como base la transformada coseno de la secuencia no desplazada; para ello es necesario calcular la TDC y la TDS de la secuencia original, desde este punto en adelante lo que se debe hacer es ir actualizando la transformada para cada uno de los nuevos desplazamientos. El inconveniente del cálculo inicial es encontrar la transformada seno dada por la ecuación (3.1.3.34). Como tiene el mismo tamaño y los mismos argumentos que la TDC se sugiere encontrar alguna relación con esta, de manera que su cómputo no resulte problemático. Para obtener el algoritmo de cálculo de la TDC unidimensional, se debe proceder de igual manera a como se lo hizo para el caso bidimensional en el capítulo tres, pero limitando el análisis a una sola dimensión.
- La teoría del filtro homomórfico se presenta en las referencias [1], [3], [5] y [50] aquí se ha limitado a describir solamente la parte aplicable en el dominio de la frecuencia y según la estructura dada por la figura 4.2.4.1, tanto para el filtro pasa-altos como para la eliminación del ruido dependiente de la señal, un análisis minucioso del filtro homomórfico requeriría un trabajo de investigación adicional.
- Las transformadas imagen pueden aplicarse en campos como: mejoramiento, restauración y compresión de imágenes. Un tema recomendable para futuras

investigaciones sería el realizar un estudio comparativo de la aplicación de cada una de las transformadas en estos campos; en especial en compresión de imágenes, aprovechando la propiedad de compresión energética de la cual gozan algunas de las transformadas

- El usuario puede disponer de la librería TDCOSEN0.DLL para el cálculo del par de transformadas coseno bidimensionales de tamaño  $M \times N$ , donde  $M$  y  $N$  son potencias de 2. En el manual de usuario se muestra como proceder para su utilización.

# BIBLIOGRAFÍA

## Textos

- [1] JAIN A, Fundamentals of Digital Image Processing, Prentice-Hall International Inc, USA, 1989.
- [2] CHELLAPPA y SAWCHUCK, Digital Image Processing and Analysis, IEEE Computer, IEEE Catalog No. EHO231-1, 1985.
- [3] LIM J, Two-Dimensional Signal and Image Processing, P T R Prentice-Hall, New Jersey, 1990.
- [4] BELOVE CH, Enciclopedia de la Electrónica, Ingeniería y Técnica, 8 Vol, Grupo Editorial Océano, Barcelona, 1984.
- [5] PITAS y VENETSANOPOULOS, Nonlinear digital filters: principles and applications, Kluwer Academic Publishers, Second Edition, USA, 1991.
- [6] LEON R, Introducción al Procesamiento Digital de Señales Estocásticas, Facultad de Ingeniería Electrónica, Escuela Politécnica del Ejército, 1993.
- [7] OPPENHEIM y SCHAFER, Digital Signal Processing, Prentice Hall, NY, 1975.
- [8] STREMLER F, Introducción a los Sistemas de Comunicación, Editorial Addison-Wesley Iberoamericana, Massachusetts, 1990.
- [9] AGUILAR G, Procesamiento Digital de Imágenes utilizando Filtros Morfológicos, Tesis previa a la obtención del Título de Ingeniero en Electrónica y Telecomunicaciones, EPN, Quito, 1995.
- [10] ROSERO G, Mejoramiento de Imágenes Digitales Monocromáticas en el Dominio Espacial y usando la Transformada de Fourier, Tesis previo a la obtención del Título de Ingeniero Electrónico en la Especialización de Telecomunicaciones, EPN, Quito, 1996.
- [11] ARTURO J, Transformada Coseno, Tesis previa a la obtención del Título de Ingeniero en Electrónica y Telecomunicaciones, EPN, 1983.
- [12] HIDALGO G, La Transformada Hadamard, Tesis previa a la obtención del Título de Ingeniero en Electrónica y Telecomunicaciones, EPN, 1974.
- [13] PALACIOS H, Introducción al Procesamiento de Imágenes. Un enfoque práctico. Tesis previa a la obtención del Título de Ingeniero de Sistemas en Informática y Computación, EPN, Quito, 1991.

- [14] OÑATE H, Transformaciones Geométricas con puntos de control en el procesamiento digital de imágenes multiespectrales, Tesis previa a la obtención del Título de Ingeniero de Sistemas en Informática y Computación, EPN, 1993.
- [15] OGATA y KATSUHITO, Ingeniería de Control Moderna, Prentice-Hall, México, 1980.
- [16] APOSTOL T, Calculus, 2 Vol, Blaisdell Publishing Company, Segunda Edición, Massachusetts, 1967.
- [17] KERNIGHAN y RITCHIE, El Lenguaje de Programación C, Prentice-Hall Hispanoamericana, S.A, México, 1988.
- [18] PAPPAS y MURRAY, Manual de Borland C++ 4.0, Osborne/McGraw-Hill, Madrid, 1994..
- [19] SCHILDT H, Turbo C/C++. Manual de Referencia, Osborne/McGraw-Hill, Madrid, 1992.
- [20] BORLAND INTERNACIONAL, Turbo C++ for Windows, Programmer's Guide, Version 3.0, USA, 1991.
- [21] FRANCO A, Programación de Aplicaciones Windows con Borland C++ y ObjectWindows, McGraw-Hill, España, 1994.
- [22] ADAMS L, Programación avanzada de gráficos en C para Windows, Windcrest/McGraw-Hill, Madrid, 1993.
- [23] ELECTRÓNICA FÁCIL, Diccionario Electrónico, Ed. DITEL, Medellín, 1984.

## Revistas

- [24] IEEE Transactions on Computers, Discrete Cosine Transform, AHMED, NATARAJAN y RAO, Correspondence, January 1974, pp. 90-93.
- [25] IEEE Transactions on Acustics, Speech, and Signal Processing, A Two-Dimensional Fast Cosine Transform, HAQUE M, Vol. ASSP-33 Num. 6 December 1985, pp. 1532-1539.
- [26] IEEE Transactions on Acustics, Speech, and Signal Processing, A New Algorithm to Compute the Discrete Cosine Transform, LEE B, Vol. ASSP-32 Num. 6, December 1984, pp. 1243-1245.
- [27] IEEE Transactions on Acustics, Speech, and Signal Processing, A Fast Recursive Algorithm For Computing the Discrete Cosine Transform, HOU H, Vol. ASSP-35 Num. 10, October 1987, pp. 1455-1461.

- [28] Proceedings of the IEEE, Image Processing, HUANG, SCHREIBER y TRETIAK, Vol. 59 Num. 11, November 1971, pp. 1586-1609.
- [29] IEEE Transactions on Acoustics, Speech, and Signal Processing, Input and Output Index Mappings for a Prime-Factor Decomposed Computation of Discrete Cosine Transform, LEE B, Vol. 37 Num. 2, February 1989, pp. 237-244.
- [30] IEEE Transactions on Acoustics, Speech, and Signal Processing, A Fast Cosine Transform in One and Two Dimensions, MAKHOUL J, Vol. ASP-28 Num. 1, February 1980, pp. 27-34.
- [31] IEEE Transactions on Computers, Fast Algorithms for the 2-D Discrete Cosine Transform, KAMANGAR y RAO, Vol. C-31, Num. 9, September 1982, pp. 899-906
- [32] IEEE Transactions on Acoustics, Speech, and Signal Processing, Fast Computation of the Discrete Cosine Transform and the Discrete Hartley Transform, MALVAR H, Vol. ASSP-35 Num. 10 October 1987, pp. 1484-1485.
- [33] IEEE Transactions on Acoustics, Speech, and Signal Processing, On the Shift Property of DCT's and DST's, YIP y RAO, Vol. ASSP-35 Num. 3 March 1987, pp. 404-406.
- [34] IEEE Transactions on Acoustics, Speech, and Signal Processing, On Relating Discrete Fourier Sine and Symmetric Cosine Transforms, ERSOY O, Vol. ASSP-33 Num. 1 February 1985, pp. 219-222.
- [35] IEEE Transactions on Acoustics, Speech, and Signal Processing, Comparative Performance of Two Different Versions of the Discrete Cosine Transform, WANG y HUNT, Vol. ASSP-32 Num. 2 April 1984, pp. 450-452.
- [36] IEEE Transactions on Acoustics, Speech, and Signal Processing, Image Display Techniques Using the Cosine Transform, NGAN K, Vol. ASSP-32 Num. 1 February 1984, pp. 173-177.
- [37] IEEE Transactions on Acoustics, Speech, and Signal Processing, C-Matrix Transform, KWAK, SRINIVASAN y RAO, Vol. ASSP-31 Num. 5 October 1983, pp. 1304-1307.
- [38] IEEE Transactions on Circuits and Systems, VLSI Implementation of a 16x16 Discrete Cosine Transform, SUN, CHEN y GOTTLIEB, Vol. 36 Num. 4 April 1989, pp. 610-617.
- [39] IEEE Transactions on Communications, A Comparison of the Hartley, Cas-Cas, Fourier, and Discrete Cosine Transforms for Image Coding, PERKINS M, Vol. 36 Num. 6 June 1988, pp. 758-760.
- [40] IEEE Transactions on Acoustics, Speech, and Signal Processing, Karhunen-Loeve Multispectral Image Restoration, Part I: Theory, HUNT y KÜBLER, Vol. ASSP-32 Num. 3 June 1984, pp. 592-599.



- [41] IEEE Transactions on Acustics, Speech, and Signal Processing, Real-Time Image Processing by Distributed Arithmetic Implementation of Two-Dimensional Digital Filters, JAGGERNAUTH, LOUI y VENETSANOPOULOS, Vol. ASSP-33, Num. 6, December 1985, pp. 1546-1555.
- [42] IEEE Transactions on Circuits and Systems, Design of FIR Two-Dimensional Digital Filters by Successive Projections, ABO-TALEB y FAHMY, Vol. CAS-31, Num. 9, September 1984, pp. 801-805.
- [43] IEEE Transactions on Circuits and Systems, Frecuency-Domain Design of N-D Digital Filters, MUTLUAY y FAHMY, Vol. CAS-32, Num. 12, December 1985, pp. 1226-1233.
- [44] IEEE Transactions on Circuits and Systems, Generalized Homomorphic and adaptive Order Statistic Filters for the Removal of Impulsive and Signal-Dependent Noise, DING y VENETSANOPOULOS, Vol. CAS-34, Num. 8, August 1987, pp. 948-955.
- [45] IEEE Transactions on Acustics, Speech, and Signal Processing, A New Systolic Array for Discrete Fourier Transform, CHANG y CHEN, Vol. 36 Num. 10, October 1988, pp. 1665-1666.
- [46] IEEE Transactions on Circuits and Systems, Recursive Two-Dimensional Median Filtering Algorithms for Fast Image Root Extraction, BOLES, KANEFSKY y SIMAAN, Vol. 35 Num. 10, October 1988, pp. 1323-1326.
- [47] IEEE Transactions on Acustics, Speech, and Signal Processing, Estimating the Basis Functiones of the Karhunen-Loeve Transform, BURL J, Vol. 37, Num. 1, January 1989, pp. 99-105.
- [48] Proceedings of the IEEE, Image Processing in the Context of a Visual Model, STOCKHAM T, Vol. 60 Num. 7, July 1972, pp. 828-842.
- [49] IEEE Transactions on Acustics, Speech, and Signal Processing, Nonlinear Mean Filters in Image Processing, PITAS y VENETSANOPOULOS, Vol. ASSP-34, Num. 3, June 1986, pp. 573-584.
- [50] Proceedings of the IEEE, Nonlinear Filtering of Multiplied and Convolved Signals, OPPENHEIM, SCHAFER y STOCKHAM, Vol. 56, Num. 8, August 1968, pp. 1264-1291.

# APÉNDICE

---

## A

### CONCEPTOS BÁSICOS DE SISTEMAS DISCRETOS

Una secuencia bidimensional de uso frecuente es la función paso-unitaria, definida como a continuación:

$$u_{-1} = \begin{cases} 1 & \text{si } m \geq 0, n \geq 0 \\ 0 & \text{en otro caso} \end{cases} \quad (\text{A.1})$$

Una función íntimamente relacionada con la paso unitaria es la secuencia impulso-unitaria, definida así,

$$\delta(m,n) = \begin{cases} 1 & \text{si } m = n = 0 \\ 0 & \text{en otro caso} \end{cases} \quad (\text{A.2})$$

**Separabilidad.-** Se define la *separabilidad* como aquella propiedad de la cual gozan algunas funciones bidimensionales (2-D) y establece que pueden ser expresadas como el producto de funciones unidimensionales, por ejemplo  $u(m,n)$  puede ser expresada como:

$$u(m,n) = u_1(m) u_2(n) \quad (\text{A.3})$$

basado en esto, se define la función *delta Kronecker* de la siguiente forma:

$$\delta(m,n) = \delta(m)\delta(n) \quad (\text{A.4})$$

la misma que satisface las siguientes propiedades:

$$u(m,n) = \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} u(m',n') \delta(m-m',n-n') \tag{A.5}$$

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(m,n) = 1$$

Si  $x(m,n)$  y  $y(m,n)$  representan las secuencias de entrada y salida respectivamente del sistema bidimensional ilustrado en la figura A.1, estas se relacionan como:

$$y(m,n) = H[x(m,n)] \tag{A.6}$$

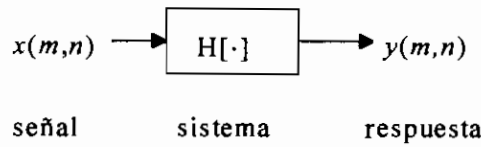


Figura A.1

**Linealidad.-** Un sistema se dice *lineal ssi*<sup>1</sup> cumple con el teorema de superposición, es decir, si para las constantes  $a_1$  y  $a_2$ , se tiene;

$$H[a_1 x_1(m,n) + a_2 x_2(m,n)] = a_1 H[x_1(m,n)] + a_2 H[x_2(m,n)]$$

$$= a_1 y_1(m,n) + a_2 y_2(m,n) \tag{A.7}$$

La respuesta impulsiva en el punto  $(m,n)$  se define cuando la entrada es la función delta Kronecker en el punto  $(m',n')$  así,

$$h(m,n;m',n') = H[\delta(m-m',n-n')] \tag{A.8}$$

el punto y coma “ ; ” se usa para separar las coordenadas entrada/salida.

<sup>1</sup> *ssi*.- Es una contracción que se lee “*si y solamente si*” y matemáticamente representa que las dos proposiciones que la rodean son simultáneamente verdaderas o falsas.

La región de soporte de la respuesta impulsiva es aquella región del plano  $m, n$  fuera de la cual la respuesta es cero. Si ésta es finita el sistema es de respuesta finita (FIR), caso contrario será de respuesta infinita (IIR). Para mayor claridad considérese una ecuación lineal en diferencias<sup>1</sup> con coeficientes constantes de la siguiente manera:

$$y(m, n) + \sum_{i=1}^p \sum_{j=1}^q a_{i,j} y(m-i, n-j) = \sum_{i=0}^r \sum_{j=0}^s b_{i,j} x(m-i, n-j) \quad (\text{A.9})$$

conocida la respuesta impulso, es posible determinar su respuesta a cualquier otra entrada, dependiendo de la duración de la respuesta impulso, se habla de:

### 1.- Sistemas de respuesta impulso finita (FIR).

Si  $a_{i,j} = 0$  en (A.9), ésta se convierte en:

$$y(m, n) = \sum_i \sum_j b_{i,j} x(m-i, n-j)$$

Entonces si un impulso ocurrió hace más de  $r, s$ -muestras, todo el miembro derecho de la ecuación es cero, en tal caso la función de transferencia de un filtro FIR tiene solo ceros y un polo en  $z = 0$ .

### 2.- Sistemas de respuesta impulso infinita (IIR).

Si alguno de los coeficientes  $a_{i,j}$  de (A.9) es diferente de cero, en este caso la respuesta impulso puede durar por siempre, pues aunque el impulso haya ocurrido hace  $r, s$ -muestras aún se recuerda en las muestras  $y(m-i, n-j)$ , entonces el sistema será de respuesta infinita y su función de transferencia tendrá polos de frecuencia finita.

<sup>1</sup> Ecuación en diferencias 2-D obtenida por extensión del caso 1-D.

La respuesta de un sistema en función de la respuesta impulsiva y de la señal de entrada es:

$$\begin{aligned}
 y(m, n) &= H[x(m, n)] = H\left[\sum_{m'} \sum_{n'} x(m', n') \delta(m - m', n - n')\right] \\
 &= \sum_{m'} \sum_{n'} x(m', n') H[\delta(m - m', n - n')] = \sum_{m'} \sum_{n'} x(m', n') h(m, n; m', n') \quad (\text{A.10})
 \end{aligned}$$

**Invarianza.-** Un sistema se llama *invariante en el espacio*, si para una entrada desplazada se obtiene una salida correspondientemente desplazada en el espacio, es decir,

$$\begin{aligned}
 h(m, n; m', n') &= H[\delta(m - m', n - n')] \\
 &= h(m - m', n - n'; 0, 0) \\
 \Rightarrow h(m, n; m', n') &= h(m - m', n - n')
 \end{aligned}$$

entonces la respuesta de un sistema lineal invariante (SLI) es la siguiente:

$$y(m, n) = \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} h(m - m', n - n') x(m', n')$$

La ecuación siguiente es conocida como la *convolución* de la entrada con la respuesta impulsiva para el caso discreto:

$$y(m, n) = h(m, n) * x(m, n) = \sum_{m'} \sum_{n'} h(m - m', n - n') x(m', n')$$

**Causalidad.-** Un sistema se dice *causal*, no anticipatorio o físicamente realizable si no responde antes de ser excitado, es decir para un sistema lineal invariante, este es causal *ssi* su respuesta impulsiva es cero para  $m$  y  $n$  negativos.

# APÉNDICE

---

## B

### DEMOSTRACIONES MATEMÁTICAS

Este apéndice contiene demostraciones matemáticas de algunas ecuaciones del capítulo dos que son importantes, pero que no se han incluido en él y que posiblemente ayuden en la comprensión al exponer los tópicos que se cubren en este apéndice

#### B-1 Demostración de la ecuación (2.1.0.9)

De la versión matricial de la ecuación (2.1.0.8), esto es:

$$\mathbf{V} = \mathbf{AUA}^T$$

realizando multiplicaciones convenientes en ambos lados de la igualdad como sigue, se obtiene la ecuación (2.1.0.9).

$$\begin{aligned}\mathbf{V} &= \mathbf{AUA}^T \\ \mathbf{A}^{*T} \cdot \mathbf{V} &= \mathbf{A}^{*T} \cdot \mathbf{AUA}^T \\ \mathbf{A}^{*T} \mathbf{V} \cdot \mathbf{A}^* &= \mathbf{A}^{*T} \mathbf{AUA}^T \cdot \mathbf{A}^* \\ \mathbf{A}^{*T} \mathbf{V} \mathbf{A}^* &= (\mathbf{A}^{*T} \mathbf{A}) \mathbf{U} (\mathbf{A}^{*T} \mathbf{A})^T \\ \mathbf{A}^{*T} \mathbf{V} \mathbf{A}^* &= \mathbf{U}\end{aligned}$$

para esta demostración se ha usado continuamente la definición de matriz unitaria, esto es,

$$\mathbf{A}^{-1} = \mathbf{A}^{*T}.$$

## B-2 Demostración de la ecuación (2.1.1.2)

Desarrollando la ecuación (2.1.0.9), se tiene,

$$\begin{aligned} u(m, n) &= \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} a^*(k, m) v(k, l) a^*(l, n) \\ &= a^*(0, m) v(0, 0) a^*(0, n) + \cdots + a^*(k, m) v(k, l) a^*(l, n) + \\ &\quad + \cdots + a^*(M-1, m) v(M-1, N-1) a^*(N-1, n) \end{aligned}$$

escribiendo luego en su forma matricial, toma la forma:

$$\begin{aligned} \{u(m, n)\}_{M \times N} &= v(0, 0) \{a^*(0, m) a^*(0, n)\} + \cdots + v(k, l) \{a^*(k, m) a^*(l, n)\} + \\ &\quad + \cdots + v(M-1, N-1) \{a^*(M-1, m) a^*(N-1, n)\} \end{aligned}$$

desarrollando íntegramente el  $k, l$ -ésimo sumando, se obtiene entonces:

$$\begin{aligned} \mathbf{U} &= \begin{bmatrix} a^*(k, 0) a^*(l, 0) & a^*(k, 0) a^*(l, 1) & \cdots & a^*(k, 0) a^*(l, N-1) \\ a^*(k, 1) a^*(l, 0) & a^*(k, 1) a^*(l, 1) & \cdots & a^*(k, 1) a^*(l, N-1) \\ \vdots & \vdots & \ddots & \vdots \\ a^*(k, M-1) a^*(l, 0) & a^*(k, M-1) a^*(l, 1) & \cdots & a^*(k, M-1) a^*(l, N-1) \end{bmatrix} \times \\ &\quad \times v(k, l) + v(0, 0) \mathbf{A}_{0,0}^* + \cdots + v(M-1, N-1) \mathbf{A}_{M-1, N-1}^* \end{aligned}$$

la  $k, l$ -ésima matriz  $M \times N$  puede ser escrita como el producto de una matriz de tamaño

$M \times 1$  por otra de tamaño  $1 \times N$ , de la siguiente manera:

$$\begin{aligned} \mathbf{U} &= \begin{bmatrix} a^*(k, 0) \\ a^*(k, 1) \\ \vdots \\ a^*(k, M-1) \end{bmatrix} \begin{bmatrix} a^*(l, 0) & \cdots & a^*(l, N-1) \end{bmatrix} \cdot v(k, l) + \\ &\quad + v(0, 0) \mathbf{A}_{0,0}^* + \cdots + v(M-1, N-1) \mathbf{A}_{M-1, N-1}^* \end{aligned}$$

este último desarrollo utiliza la definición dada por (2.1.1.1) para concluir en la igualdad

(2.1.1.2).

### B-3 Demostración de la ecuación (2.1.1.3)

Tomando la igualdad (2.1.0.8),

$$v(k,l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a(k,m) u(m,n) a(l,n)$$

en forma análoga a lo realizado en el sección B-2,

$$\mathbf{a}_k \mathbf{a}_l^T = \mathbf{A}_{k,l}$$

entonces (2.1.0.8) puede escribirse como:

$$v(k,l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n) \mathbf{A}_{k,l}(m,n)$$

en el capítulo 1 el definió el producto interno entre funciones (ecuación 1.3.1.2), dicho producto para el caso matricial se define del siguiente modo:

$$\begin{aligned} \langle \mathbf{U}, \mathbf{A}_{k,l} \rangle &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n) \mathbf{A}_{k,l}^*(m,n) \\ \Rightarrow v(k,l) &= \langle \mathbf{U}, \mathbf{A}_{k,l}^* \rangle \end{aligned}$$

### B-4 Demostración de la ecuación (2.1.2.4)

Tomando la definición de covarianza y la linealidad del valor esperado, se procede, entonces:

$$\begin{aligned} \mathbf{R}_v &= E[(\mathbf{v} - \mathbf{m}_v)(\mathbf{v} - \mathbf{m}_v)^*{}^T] \\ &= E[(\mathbf{A}\mathbf{u} - \mathbf{A}\mathbf{m}_u)(\mathbf{A}\mathbf{u} - \mathbf{A}\mathbf{m}_u)^*{}^T] \\ &= \mathbf{A}E[(\mathbf{u} - \mathbf{m}_u)(\mathbf{u} - \mathbf{m}_u)^*{}^T]\mathbf{A}^*{}^T \\ &= \mathbf{A}\mathbf{R}_u\mathbf{A}^*{}^T \end{aligned}$$

esta expresión relaciona la covarianza en el dominio espacial con su homóloga en el dominio de la frecuencia espacial.



### B-5 Demostración de la ecuación (2.1.2.9)

$$\begin{aligned}\sigma_v^2(k, l) &= E[|v(k, l) - m_v(k, l)|^2] \\ &= E[(v(k, l) - m_v(k, l))(v(k, l) - m_v(k, l))^*] \\ &= E\left[\left(\sum_m \sum_n a(k, l) a(l, n) (u(m, n) - m_u(m, n))\right) \cdot \right. \\ &\quad \left. \left[\sum_{m'} \sum_{n'} (a^*(k, m') a^*(l, n') (u^*(m', n') - m_u^*(m', n'))\right)\right]\end{aligned}$$

realizando las operaciones indicadas y factorando con otro término común, se obtendrá:

$$\begin{aligned}\sigma_v^2(k, l) &= \sum_m \sum_n \sum_{m'} \sum_{n'} (a(k, m) a(l, n) a^*(k, m') a^*(l, n') \cdot \\ &\quad \cdot E[u(m, n) u^*(m', n') - u(m, n) m_u^*(m', n') + \\ &\quad - m_u(m, n) u^*(m', n') + m_u(m, n) m_u^*(m', n') ] \\ &= \sum_m \sum_n \sum_{m'} \sum_{n'} a(k, m) a(l, n) a^*(k, m') a^*(l, n') \cdot \\ &\quad \cdot E[u(m, n) u^*(m', n') - u(m, n) m_u^*(m', n')]\end{aligned}$$

considerando la siguiente igualdad,

$$\begin{aligned}r(m, n; m', n') &= \text{Cov}(u(m, n), u^*(m', n')) \\ &= E[u(m, n)(u^*(m', n') - m_u^*(m', n'))]\end{aligned}$$

se concluye con la igualdad buscada, la cual es,

$$\sigma_v^2(k, l) = \sum_m \sum_n \sum_{m'} \sum_{n'} a(k, m) a(l, n) r(m, n; m', n') a^*(k, m') a^*(l, n')$$

### B-6 Demostración de la ecuación (2.2.2.5)

Tomando la definición de la TDC unidimensional en la cual la matriz TDC está dada por

(2.2.2.1) y haciendo  $\alpha(k) = \sqrt{\frac{2}{N}} e(k)$ , se procede de la siguiente manera:

$$\begin{aligned}v(k) &= \alpha(k) \sum_{n=0}^{N-1} u(n) \cos\left(\frac{(2n+1)k\pi}{2N}\right) \\ &= \alpha(k) \sum_{n=0}^{N/2-1} \left\{ u(2n) \cos\left(\frac{(4n+1)k\pi}{2N}\right) + u(2n+1) \cos\left(\frac{(4n+3)k\pi}{2N}\right) \right\} \\ &= \alpha(k) \sum_{n=0}^{N/2-1} \left\{ \tilde{u}(n) \cos\left(\frac{(4n+1)k\pi}{2N}\right) + \tilde{u}(N-n-1) \cos\left(\frac{(4n+3)k\pi}{2N}\right) \right\}\end{aligned}$$

haciendo el siguiente cambio de variable,

$$\begin{aligned} n' &= N - n - 1 \\ \text{si } n &= 0 & \Rightarrow n' &= N - 1 \\ \text{si } n &= N/2 - 1 & \Rightarrow n' &= N/2 \end{aligned}$$

entonces:

$$v(k) = \alpha(k) \left\{ \sum_{n=0}^{N/2-1} \tilde{u}(n) \cos\left(\frac{(4n+1)k\pi}{2N}\right) + \sum_{n=N/2}^{N-1} \tilde{u}(n) \cos\left(2k\pi - k\pi \frac{4n+1}{2N}\right) \right\}$$

tomando en consideración la paridad y periodicidad de la función coseno, se obtiene,

$$\begin{aligned} v(k) &= \alpha(k) \sum_{n=0}^{N-1} \tilde{u}(n) \cos\left(\frac{(4n+1)k\pi}{2N}\right) \\ v(k) &= \text{Re} \left\{ \alpha(k) \exp\left(-j \frac{2\pi}{2N} \times \frac{k}{2}\right) \sum_{n=0}^{N-1} \tilde{u}(n) \exp\left(-j \frac{2\pi}{N} kn\right) \right\} \\ &= \text{Re} \left\{ \alpha(k) W_{2N}^{k/2} \text{TDF}[\tilde{u}(n)]_N \right\} \end{aligned}$$

# APÉNDICE

---



## DEFINICIONES DE LA TRANSFORMADA DISCRETA

### COSENO

La referencia [33], presenta la familia de definiciones de la matriz TDC, miembro de dicha familia constituye la definición descrita en el capítulo 2 (sección 2.2.2) y en el capítulo tres (sección 3.1.2), dichas definiciones se reproducen a continuación:

$$\text{TDC-I: } [C_{N+1}^I] = \sqrt{\frac{2}{N}} k_k k_m \cos\left(\frac{mk\pi}{N}\right) \quad k, m = 0, 1, \dots, N$$

$$\text{TDC-II: } [C_N^{II}] = \sqrt{\frac{2}{N}} k_k \cos\left(\frac{(2m+1)k\pi}{2N}\right) \quad k, m = 0, 1, \dots, N-1$$

$$\text{TDC-III: } [C_N^{III}] = \sqrt{\frac{2}{N}} k_m \cos\left(\frac{(2k+1)m\pi}{2N}\right) \quad k, m = 0, 1, \dots, N-1$$

$$\text{TDC-IV: } [C_N^{IV}] = \sqrt{\frac{2}{N}} \cos\left(\frac{(2m+1)(2k+1)\pi}{4N}\right) \quad k, m = 0, 1, \dots, N-1$$

en las cuales:

$$k_i = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } i = 0 \text{ o } i = N \\ 1 & \text{en otro caso} \end{cases}$$

Para señales con coeficientes de correlación baja (cuando  $\rho \leq 0.8$  se considera baja correlación) [35] tanto la TDC-I como la TDC-II se comportan aproximadamente de la misma manera. Para datos altamente correlacionados  $\rho \geq 0.9$  la TDC-II es superior<sup>1</sup> a TDC-I [24].

Sin embargo cuando la longitud de la secuencia de datos se incrementa, el comportamiento de la TDC-I tiende a ser tan buena como la TDC-II. Puesto que la TDC-I requiere un menor número de cálculos que la TDC-II, en aplicaciones de procesamiento digital de señales e imágenes de secuencias largas de datos o cuando la correlación entre coeficientes es relativamente baja, la primera es la preferida [35]. No se dispone de ninguna información sobre el comportamiento de la TDC definidas según III y IV.

En lo referente a la TDC de una secuencia de datos tanto 1-D como 2-D es común escribir el coeficiente de normalización solamente en la dirección directa, a continuación se escribe la TDC para el caso bidimensional:

$$v(k,l) = \frac{4e(k)e(l)}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} u(m,n) \cos\left(\frac{(2m+1)k\pi}{2M}\right) \cos\left(\frac{(2n+1)l\pi}{2N}\right)$$

$$u(m,n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} e(k)e(l)v(k,l) \cos\left(\frac{(2m+1)k\pi}{2M}\right) \cos\left(\frac{(2n+1)l\pi}{2N}\right)$$

---

<sup>1</sup> En la sección 3.1.1 del capítulo 3 se exponen las medidas de comportamiento de las transformadas unitarias.

# APÉNDICE

---

## D

### RUIDO ADITIVO

Por ruido (en imágenes) se entiende a toda perturbación visual, compuesta por un conjunto de puntos de intensidad y concentración variables cuya combinación provoca una sensación visual desagradable. Al ruido en imágenes se lo clasifica en dos grandes grupos: ruido aditivo y ruido dependiente de la señal.

#### RUIDO ADITIVO

Una imagen monocromática de 256 tonos de gris contaminada con ruido aditivo tiene la siguiente formulación:

$$u_r(m, n) = u(m, n) + n(m, n) = \begin{cases} 255 & \text{si } u(m, n) + n(m, n) \geq 256 \\ 0 & \text{si } u(m, n) + n(m, n) \leq 0 \\ u(m, n) + n(m, n) & \text{en otro caso} \end{cases}$$

en la cual:

- $u(m, n)$  denota la imagen original,
- $n(m, n)$  es la función de ruido

- $u_r(m, n)$  es la imagen observada.

La función de ruido  $n(m, n)$  puede ser: de ruido gaussiano blanco ó de ruido impulsivo.

### Ruido aditivo gaussiano blanco

La función de ruido gaussiano blanco se calcula mediante la siguiente expresión:

$$n(m, n) = \sigma_n \left\{ \sum_{i=1}^{12} RND_i - 6 \right\} + m_n$$

en donde:

$\sigma_n$ : es la desviación estándar del ruido (dato proporcionado por el usuario), valor recomendado entre 10 y 30,

$RND_i$ : número aleatorio entre 0 y 1, generado por el computador, y

$m_n$ : valor medio del ruido (en este caso  $m_n = 0$ ).

El término blanco se desprende del hecho que su densidad espectral de potencia es uniforme a lo largo de todo el rango de frecuencia en consideración y gaussiano porque su función de densidad de probabilidad es gaussiana.

### Ruido aditivo impulsivo

El ruido impulsivo produce pulsos en determinados pixeles de la imagen de acuerdo a la probabilidad  $\rho$  de ocurrencia del ruido, existen tres tipos de ruido impulsivo, y estos son:

- Ruido sal y pimienta
- Ruido impulsivo positivo
- Ruido impulsivo negativo

El dato proporcionado por el usuario para la generación de este tipo de ruido es la probabilidad de ocurrencia del ruido, denotado con  $\rho$ .

### Ruido Sal y Pimienta

Una imagen contaminada con ruido sal y pimienta está representada por la siguiente función:

$$u_r(m, n) = \begin{cases} 255 & \text{si } RND_{m,n} \in ]1 - \rho, 1 - \rho/2] \\ 0 & \text{si } RND_{m,n} \in ]1 - \rho/2, 1] \\ u(m, n) & \text{si } RND_{m,n} \in [0, 1 - \rho] \end{cases}$$

### Ruido Impulsivo Positivo

De manera similar, la función que describe a la imagen contaminada con ruido impulsivo positivo toma la siguiente forma:

$$u_r(m, n) = \begin{cases} 255 & \text{si } RND_{m,n} \in ]1 - \rho, 1] \\ u(m, n) & \text{si } RND_{m,n} \in [0, 1 - \rho] \end{cases}$$

### Ruido Impulsivo Negativo

Para el caso de una imagen contaminada con ruido impulsivo negativo, la función que la representa es:

$$u_r(m, n) = \begin{cases} 0 & \text{si } RND_{m,n} \in ]1 - \rho, 1] \\ u(m, n) & \text{si } RND_{m,n} \in [0, 1 - \rho] \end{cases}$$

### Relación señal a ruido (*SNR*)

Es un parámetro utilizado para evaluar el rendimiento de un filtro digital sobre una imagen, conocida la imagen original (sin ruido), se la contamina con cualquiera de los ruidos y luego se filtra esta imagen ruidosa, para una imagen monocromática de 256 tonos de gris se calcula de la siguiente manera:

$$SNR(dB) = 20 \log \left( \frac{255}{\sqrt{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [u(m,n) - u_f(m,n)]^2}} \right)$$

La *SNR* depende tanto de la imagen original (imagen de referencia)  $u(m,n)$ , como de la imagen filtrada  $u_f(m,n)$ . El tamaño de las imágenes es  $M \times N$ .

Toda la información aquí presentada es un resumen de aquella expuesta sobre el tema de ruidos en la referencia [9].



# APÉNDICE

---

# E

## MANUAL DE USUARIO

La información contenida en este apéndice tiene por objeto familiarizar al usuario con el sistema de software implementado en este trabajo de tesis, se inicia indicando los requisitos mínimos necesarios para la ejecución del programa, para en un paso posterior indicar las funciones de menú que debe conocerse para el manejo adecuado del sistema.

### INSTALACIÓN

Para la instalación y funcionamiento adecuados del programa se recomienda el siguiente perfil:

- Microsoft Windows 95 o superior.
- 4 MBytes de memoria RAM.
- 6 MBytes de disco fijo.
- Monitor compatible con Windows (SVGA o superior, capaz de desplegar 256 tonalidades de gris).
- Computadora 80486.

Las capacidades en disco y en memoria solicitadas son las que el programa utiliza simultáneamente cuando este se encuentra en ejecución .

El archivo denominado INSTALAR.BAT, tiene bajo su responsabilidad el cargar los archivos necesarios del sistema desarrollado en el disco fijo. El archivo correspondiente para la ejecución del programa es MIMTDC.EXE, el mismo que para su correcto funcionamiento requiere de la asistencia de los módulos TDCOSENSO.DLL y BWCC.DLL cargados también durante el proceso de instalación.

## FUNCIONAMIENTO

Una vez puesto en marcha el programa se visualizará en pantalla una caja de dialogo con información relativa al programa como se muestra en la figura E-1 .



Figura E-1 Caja de diálogo mostrada cuando se inicia la ejecución del programa.

A continuación se describen cada una de las opciones de menú, indicando entre paréntesis la correspondiente combinación de teclas aceleradoras.

### Menú Archivo

- Nuevo (Ctrl + N): permite poner el menú y el programa en condiciones iniciales. Es habilitado solamente por *Archivo/Abrir* y *Edición/Pegar*.
- Abrir (Ctrl + A): permite leer un archivo gráfico con formato BMP solamente. En caso contrario puede usarse la opción *Edición/Pegar* con ayuda de otro utilitario gráfico.
- Salir (Alt + F4): abandona la aplicación

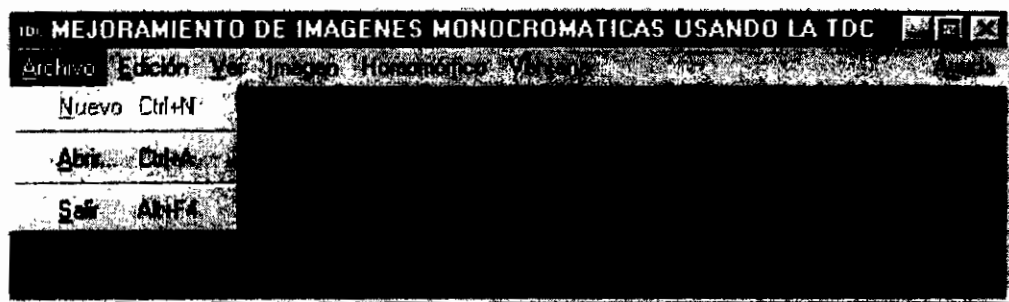


Figura E-2. Opción de menú Archivo.

### Menú Edición

- Deshacer (Alt + BkSp): deshace la acción del último filtro aplicado, opción habilitada solamente cuando se ha aplicado al menos un filtro, y desactivado luego de deshacer.
- Exportar (Ctrl + Ins): permite pegar la imagen desde la ventana principal hasta el portapapeles, opción habilitada solamente cuando hay imagen en pantalla.

- Pegar (Shift + Ins): permite pegar la imagen desde el portapapeles hasta una ventana, habilitado solamente en condiciones iniciales.



Figura E-3. Opción de menú Edición.

### Menú Ver

- Imagen Ruidosa (Ctrl + R): permite solicitar la visualización de la ventana que contiene la imagen ruidosa.
- Espectro (Ctrl + E): permite visualizar el espectro correspondiente de la imagen vigente, habilitado solamente después de haber ejecutado al menos una vez la opción de menú *Imagen/TDC*.
- Imagen Mejorada (Ctrl + M): permite visualizar la imagen mejorada, habilitado después de haber aplicado cualquier filtro.



Figura E-4. Opción de menú Ver.

## Menú Imagen

- Contaminar: permite seleccionar el tipo de ruido a usarse para la contaminación. Habilitado solamente después de usar *Archivo/Abrir* o *Edición/Pegar*.
- TDC(Ctrl + T): calcula la transformada discreta coseno de la imagen vigente. Habilitado solamente por *Archivo/Abrir* o *Edición/Pegar*.
- Atributos (Ctrl + I): muestra las dimensiones y resolución de la imagen presente.
- Filtrar: selecciona alguno de los tipos de filtros indicados. Habilitado por *Imagen/TDC*.

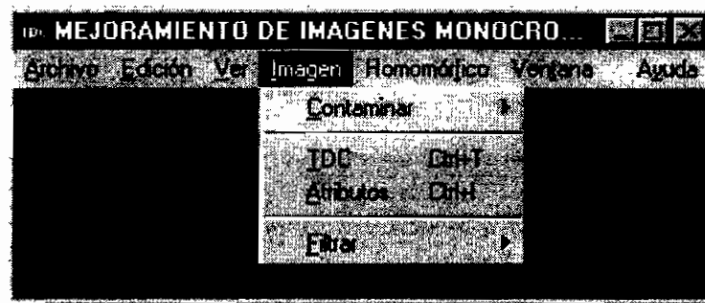


Figura E-5. Opción de menú Imagen.

## Menú Homomórfico

- Filtrar (Ctrl + H): opción de menú utilizada para iniciar el procesamiento homomórfico. Habilitado solamente después de usar *Archivo/Abrir* o *Edición/Pegar*.

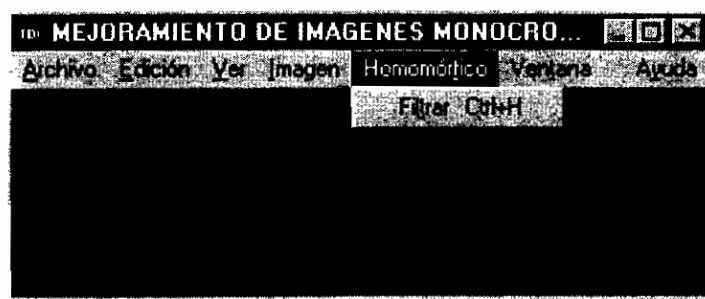


Figura E-6. Opción de menú Homomórfico.

## Menú Ventana

Un usuario familiarizado con aplicaciones Windows entenderá el significado de los elementos de esta opción.



Figura E-7. Opción de menú Ventana.

## Menú Ayuda

- Ayuda (F1): permite el acceso al sistema de ayudas.
- Acerca de...(Shift + F1): despliega la información de la aplicación.

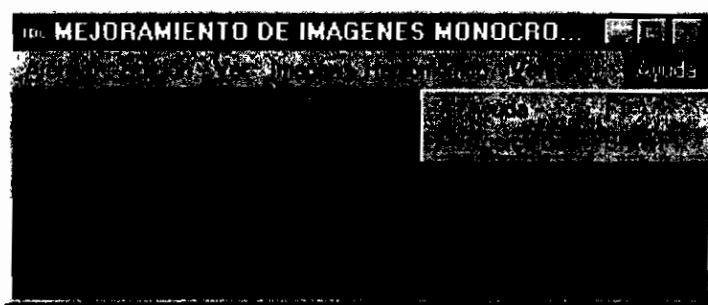


Figura E-8. Opción de menú Ayuda.

## USO DEL MOUSE

La elección de cualquier elemento de menú se lo hace normalmente con el botón izquierdo del mouse, a la opción de doble clic sobre una ventana con el botón izquierdo se le ha asignado la función de copiar el contenido de la ventana hacia el portapapeles. Un clic con

el botón derecho sobre una ventana hará desplegar una caja de diálogo con la información respectiva de la imagen presentada en aquella ventana.

## SISTEMA DE AYUDA

El sistema de ayuda contiene información de mucha utilidad especialmente cuando el usuario no está familiarizado con conceptos de procesamiento digital de imágenes en el área específica de mejoramiento de imágenes en dominio de la frecuencia. Así mismo contiene información respecto a las funciones de menú, útil para aquellas personas no iniciadas en el manejo de este sistema de software.

## UTILIZACIÓN DE LA LIBRERÍA TDCOSEN.DLL

La librería de enlace dinámico TDCOSEN.DLL contiene el código necesario para calcular el par de transformadas coseno bidimensionales de imágenes monocromáticas de tamaño  $M \times N$ , donde  $M$  y  $N$  son potencias de 2. Los prototipos de las funciones a invocarse son:

```
void CosDir(int M, int N, float huge *u, float huge *v); // calcula la TDC directa  
void CosInv(int M, int N, float huge *u, float huge *v); //calcula la TDC inversa
```

en donde  $M$  es el alto,  $N$  es el ancho de la imagen en pixeles,  $u$  es un puntero al primer elemento tipo float de un bloque de memoria que contiene la imagen en el dominio espacial,  $v$  es un puntero al primer elemento tipo float de un bloque de memoria que contiene la imagen en el dominio de la frecuencia espacial;  $\text{CosDir}()$  toma el dato de  $u$  y devuelve en  $v$ , en cambio  $\text{CosInv}()$  toma el dato de  $v$  y devuelve en  $u$ . En ambos casos los bloques de memoria deberán estar previamente reservados.