

7 01170  
29

Universidad Nacional Autónoma  
de México

División de Estudios de Posgrado  
de la Facultad de Ingeniería

Análisis y codificación de imágenes con  
transformada polinomial

Tesis

Que para obtener el grado de:

Maestro en Ingeniería Eléctrica  
(Comunicaciones)

Presenta

**Oscar Viveros Cancino**

Director de tesis:

Dr. Boris Escalante Ramírez.

CIUDAD UNIVERSITARIA  
-1998-

261266

TESIS CON  
FALLA DE ORIGEN



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A

Margarita Cancino y Arnulfo Viveros

mis padres

# Agradecimientos

Quiero expresar mi más profundo agradecimiento a las siguientes personas:

- Primeramente, a todos los profesores que han participado en mi formación académica, especialmente al Dr. Francisco García Ugalde, al Dr. Jorge Lira Chávez y al Dr. Rogelio Alcantara Silva.
- A los profesores Dr. Miguel Moctezuma Flores y Dr. Victor García Garduño, por los valiosos comentarios hechos a este trabajo y por haber aceptado formar parte de mi jurado.
- Un agradecimiento muy especial al Dr. Boris Escalante Ramírez, quien dirigió esta tesis y me brindó su paciencia y apoyo constante.
- A la comunidad de la DEPFI, especialmente a Juan Roman López Miranda y a Martha Angeles Meza, por los programas y la valiosa ayuda prestada.
- A todos mis amigos, pero de manera muy particular a Leticia Alvarez Castillo, Socorro Noguéz Monroy, Daniel Gatica Pérez y Cesar Jiménez Calvillo; cuya amistad agradezco infinitamente.
- A mi familia, que de manera desinteresada entienden mis ideales y me han apoyado constantemente. A mi hermana Lizbeth, por la colaboración prestada en la recta final.
- Finalmente, a todos aquellos que en este lapso han colaborado de manera directa o indirecta.

Gracias a todos ustedes.

## Resumen

El análisis de imágenes es importante si se intenta hacer procesamiento, tal como codificación y compresión. En este texto se proponen algoritmos para codificación de imágenes cuya técnica de análisis es la *transformada polinomial*.

La *transformada polinomial* tiene entre otras características que sus filtros de análisis no son ortogonales, por lo cual es redundante. Sin embargo, es un buen modelo del sistema de visión humano, por lo tanto es posible extraer de la expansión generada características perceptualmente importantes de la imagen.

Se sabe que una de las estructuras primarias importantes para el sistema de visión humano es el borde, por lo tanto, se pone especial énfasis en la búsqueda de dichas estructuras y se modela por medio de parámetros como son: la magnitud, la posición en la ventana, la orientación en la que se encuentran y el valor medio.

Los parámetros característicos de un borde que se obtienen, se pueden cuantizar. Debido a ello, se puede realizar codificación de imágenes de una manera eficiente. Se presentan resultados y se deduce que en base a las perspectivas planteadas, se pueden lograr altas tasas de compresión.

# Indice

<b>1</b>	<b>Introducción</b>	<b>10</b>
1.1	Análisis y procesamiento de imágenes . . . . .	11
1.2	Codificación de imágenes . . . . .	12
1.3	Objetivos . . . . .	13
1.4	¿Cómo alcanzar los objetivos? . . . . .	14
1.5	Contenido . . . . .	15
<b>2</b>	<b>Conceptos básicos</b>	<b>16</b>
2.1	Sistema de visión humano . . . . .	16
2.1.1	¿Cómo funciona el sistema de visión humano? . . . . .	17
2.1.2	¿Cómo se ha tratado de modelar la visión humana? . . . . .	18
2.2	Análisis de imágenes . . . . .	19
2.2.1	¿Cómo se forma una imagen digital? . . . . .	19
2.2.2	Captura de imágenes . . . . .	22
2.2.3	Técnicas de análisis de imágenes . . . . .	22
2.3	Procesamiento digital de imágenes . . . . .	24
2.3.1	¿Qué áreas componen al procesamiento digital de imágenes? . . . . .	24
2.4	Codificación y compresión de imágenes. . . . .	25
2.4.1	¿Cuándo se efectua compresión de datos? . . . . .	26
2.4.2	¿Porqué existe la necesidad de codificar y comprimir? . . . . .	27
2.4.3	Repaso a la teoría de la información . . . . .	27
2.5	Conclusiones . . . . .	31
<b>3</b>	<b>Transformada Polinomial</b>	<b>32</b>
3.1	Propiedades . . . . .	32
3.2	Transformada polinomial continua unidimensional. . . . .	33
3.2.1	Filtros de análisis y síntesis . . . . .	34
3.2.2	Función de peso o ponderación . . . . .	35
3.2.3	¿Cómo se obtienen los polinomios? . . . . .	36
3.2.4	¿Qué función ventana utilizar? . . . . .	37
3.3	Transformada polinomial para señales bidimensionales . . . . .	40
3.3.1	Transformada polinomial de Hermite bidimensional . . . . .	41
3.4	Transformada polinomial discreta . . . . .	41

3.4.1	Transformada polinomial de Hermite discreta. . . . .	42
3.5	Conclusiones . . . . .	42
<b>4</b>	<b>Codificación de imágenes a partir de la representación polinomial</b>	<b>43</b>
4.1	Representación de imágenes con coeficientes polinomiales unidimensionales . . . . .	43
4.1.1	Propiedades de la función ángulo . . . . .	45
4.1.2	Expansión unidimensional en coeficientes óptimos . . . . .	45
4.1.3	Reconstrucción de la expansión bidimensional . . . . .	46
4.2	Representación de imágenes con parámetros de bordes . . . . .	46
4.2.1	Modelo de un borde. . . . .	47
4.2.2	Extracción de parámetros característicos de un borde. . . . .	50
4.3	Algoritmos de codificación propuestos . . . . .	52
4.3.1	Codificación en una escala . . . . .	52
4.3.2	Codificación en múltiples escalas . . . . .	55
4.4	Conclusiones . . . . .	56
<b>5</b>	<b>Resultados</b>	<b>57</b>
5.1	Características generales . . . . .	57
5.1.1	De las imágenes utilizadas . . . . .	57
5.1.2	De los algoritmos de codificación propuestos . . . . .	59
5.2	Medidas de fiabilidad . . . . .	59
5.2.1	Entropía . . . . .	60
5.2.2	Relación señal a ruido . . . . .	60
5.2.3	Tasa de compresión . . . . .	60
5.3	Algoritmos de codificación en una escala . . . . .	61
5.3.1	Representación bidimensional . . . . .	61
5.3.2	Representación unidimensional. . . . .	70
5.3.3	Representación en parámetros característicos . . . . .	74
5.3.4	Gráficas de desempeño, para el algoritmo en una escala . . . . .	78
5.4	Algoritmos de codificación en múltiples escalas . . . . .	82
5.4.1	Representación bidimensional . . . . .	82
5.4.2	Representación unidimensional. . . . .	89
5.4.3	Representación en parámetros característicos . . . . .	94
5.4.4	Gráficas de desempeño, para el algoritmo en dos y tres escalas . . . . .	97
5.5	Comparación con JPEG . . . . .	104
5.6	Conclusiones . . . . .	108
<b>6</b>	<b>Perspectivas y conclusiones generales</b>	<b>110</b>
6.1	Perspectivas . . . . .	110
6.2	Conclusiones generales . . . . .	112
	<b>Bibliografía</b>	<b>113</b>

# Indice de figuras

1.1	Concepto general de formación de una imagen digital. . . . .	10
1.2	Sistema de comunicaciones digitales básico. . . . .	12
2.1	Diagrama del ojo humano . . . . .	17
2.2	Formación de una imagen digital . . . . .	20
3.1	Transformada polinomial directa e inversa, $f(x)$ es la señal original localizada por medio de los filtros de análisis $D_n$ que submuestreada por un factor T da como resultado los coeficientes $f_n$ . La señal reconstruida $\hat{f}(x)$ se obtiene de la sumatoria de los coeficientes interpolados por el mismo factor T y ponderados por los filtros $P_n$ . . . . .	35
4.1	Arriba, se observa una función borde ideal, comunmente llamada "función escalón". Abajo, se tiene a la función borde gaussiano. En los dos casos los valores son $A_b = 15$ , $\Delta f = 30$ y $x_b = 50$ ; el valor de la varianza $\sigma$ se ha multiplicado por un factor de 10 para su mejor visualización. . . . .	47
4.2	Diagrama de un borde con sus parámetros en el plano $(x, y)$ . . . . .	48
4.3	Respuesta de bordes Gaussiano e ideal a los operadores diferenciales de primero y segundo orden. Arriba a la izquierda se presenta la respuesta del borde ideal al operador diferencial de primer orden, a la derecha se tiene la respuesta al operador de segundo orden. Abajo se tienen las respuestas del borde gaussiano al operador diferencial de primero y segundo orden (izquierda y derecha respectivamente). . . . .	49
4.4	Curva típica de respuesta a los operadores diferenciales de un borde, en el plano $(f_1, f_2)$ . Siguiendo los valores que toman los operadores diferenciales en la figura 4.3 es posible imaginar una gráfica de esta forma. . . . .	50
4.5	Partición del plano $(f_1, f_2)$ . Umbrales para posición $x_b$ . . . . .	51
4.6	Esquema de codificación en una escala. . . . .	53
4.7	Esquema de codificación a múltiples escalas. . . . .	55



5.1	Arriba a la izquierda, se observa (a) la imagen Lena, a la derecha se encuentra (b) el histograma de valores contenidos en la imagen. Abajo a la izquierda, se observa (c) la imagen Casa, a la derecha (d) su histograma de valores. . . . .	58
5.2	Representación bidimensional en una escala (a) Lena y (b) Casa. El tamaño de la ventana es $M = 2$ y el submuestreo es $T = 2$ . . . . .	62
5.3	Reconstrucción de imágenes a partir de su representación bidimensional ( $M = 2$ y $T = 2$ ). Arriba a la izquierda (a), se tiene la imagen Lena reconstruida, abajo de ella se encuentra la imagen de error con respecto a la original. Arriba a la derecha (b), se encuentra la imagen de la Casa reconstruida, abajo de ella se encuentra la imagen de error con respecto a la original. . . . .	63
5.4	La representación bidimensional de la imagen Lena se presenta arriba (a), la correspondiente a la imagen Casa abajo (b), con tamaño de ventana $M = 4$ y submuestreo $T = 4$ . Los coeficientes se encuentran dispuestos (de la misma manera en las dos representaciones) de izquierda a derecha y empezando por la fila superior: $f_{0,0}, f_{0,1}, f_{0,2}, f_{0,3}, f_{0,4}$ ; en la siguiente fila se encuentran los coeficientes $f_{1,0}, f_{1,1}, f_{1,2}, f_{1,3}$ . La siguiente fila se compone de los coeficientes $f_{2,0}, f_{2,1}, f_{2,2}$ . Los coeficientes $f_{3,0}, f_{3,1}$ forman la penúltima fila y $f_{4,0}$ es la última fila. . . . .	64
5.5	Las imágenes reconstruidas a partir de sus respectivas representaciones bidimensionales ( $M = 4$ y $T = 4$ ), se observan en esta figura. Arriba a la izquierda (a), se tiene la imagen Lena reconstruida. Arriba a la derecha (b), se encuentra la imagen de la Casa reconstruida. Abajo de cada una de ellas se encuentran las imágenes de error con respecto a la original. . . . .	65
5.6	Representación bidimensional de las imágenes (a) Lena y (b) Casa, con tamaño de ventana $M = 8$ y muestreo $T = 8$ . Los coeficientes $f_{R,C}$ donde: Los renglones $R$ tienen un rango que es $(0, \dots, M)$ y las columnas $C$ tienen un rango $(0, N)$ siempre y cuando $R + C \leq M$ . Si por ejemplo se quiere saber cual es el coeficiente $f_{4,3}$ , simplemente se busca en el renglon 4 y la columna 3. El resultado es la matriz triangular superior para cada representación. . . . .	66
5.7	Las imagenes reconstruidas a partir de la representación bidimensional con parámetros $M = 8$ y $T = 8$ , (a) Lena y (b) Casa. Sus respectivos errores se observan abajo de cada una de ellas. . . . .	67
5.8	Representación polinomial unidimensional de las imágenes (a) Lena y (b) Casa, con $M = 2$ y $T = 2$ . En las dos figuras se encuentran dispuestos los coeficientes $K_n$ de la misma forma: de izquierda a derecha en la parte superior se observan los coeficientes de orden $K_1$ y $K_2$ , abajo se encuentra la imagen de orientaciones óptimas $\theta_{optima}$ y la señal de baja frecuencia o coeficiente de orden $K_0$ . . . . .	68

5.9	Reconstrucción a partir de la representación unidimensional de la figura 5.8 . Arriba se observan las reconstrucciones de la imágenes (a) Lena y (b) Casa. Abajo de ellas, su respectivo error de reconstrucción.	69
5.10	En las representaciones unidimensionales de las imágenes (a) Lena y (b) Casa, con $M = 4$ y $T = 4$ , se encuentran los coeficientes dispuestos de la siguiente manera: arriba y de izquierda a derecha, se tienen los coeficientes $K_1, K_2$ y $K_3$ . Abajo y de izquierda a derecha se observa el coeficiente $K_4$ , la imagen de orientaciones óptimas $\theta_{optima}$ y el coeficiente de orden 0, $K_0$ .	70
5.11	A partir de las representaciones de la figura 5.10, se recuperan las imágenes (a) Lena y (b) Casa. Abajo de ellas la imagen de error correspondiente.	71
5.12	Representación unidimensional con $M = 8$ y $T = 8$ de (a) Lena y (b) Casa. En las dos figuras se encuentran los coeficientes dispuestos de la misma manera: de arriba a abajo y de izquierda a derecha se encuentran los coeficientes $K_0, K_1, K_2, K_3, K_4$ y $K_5, K_6, K_7, K_8$ y orientaciones $\theta_{optima}$ .	71
5.13	Reconstrucción a partir de la representación $M = 8$ y $T = 8$ , de las imágenes (a) Lena y (b) Casa. Abajo de cada reconstrucción se encuentran las imágenes de error correspondientes.	72
5.14	Acercamiento a la reconstrucción de la imagen Lena: (a) con $M = 2, T = 2$ , (b) con $M = 4, T = 4$ , (c) con $M = 8, T = 8$ . Se observa una parte de la cara, aumentada 2 veces.	73
5.15	Representación en parámetros característicos de bordes, con $M = 2$ y $T = 2$ de la imágenes (a) Lena y (b) Casa. En la parte superior de las dos figuras se encuentra la imagen con valores de posición $x_b$ y a su derecha la imagen de magnitud de los bordes contenidos en la imagen $\Delta f$ . Abajo se tienen las imágenes de orientaciones $\theta_{optima}$ y a la derecha el valor medio $A_b$ .	73
5.16	Reconstrucción a partir de la representación en parámetros característicos que se observa en la figura 5.15. Se tiene en la parte superior la reconstrucción y abajo de ella su correspondiente error respecto de la imagen original. (a) Imagen Lena y (b) imagen Casa	75
5.17	Representación en parámetros característicos de bordes, con $M = 4$ y $T = 4$ de las imágenes (a) Lena y (b) Casa. En la parte superior de las dos figuras se encuentra la imagen con valores de posición $x_b$ y a su derecha la imagen de magnitud de los bordes contenidos en la imagen $\Delta f$ . Abajo se tienen las imágenes de orientaciones $\theta_{optima}$ y a la derecha el valor medio $A_b$ .	76
5.18	Reconstrucción de las imágenes (a) Lena y (b) Casa, a partir de sus respectivas representaciones que se muestran en la figura 5.17. Arriba se tiene la reconstrucción y abajo su correspondiente imagen de error.	77

5.19	Representación en parámetros característicos de bordes, con $M = 8$ y $T = 8$ de las imágenes (a) Lena y (b) Casa. En la parte superior de las dos figuras se encuentra la imagen con valores de posición $x_b$ y a su derecha la imagen de magnitud de los bordes contenidos en la imagen $\Delta f$ . Abajo se tienen las imágenes de orientaciones $\theta_{optima}$ y a la derecha el valor medio $A_b$ . . . . .	78
5.20	Reconstrucción de las imágenes (a) Lena y (b) Casa, con su respectiva imagen de error. Se recuperaron a partir de las representaciones mostradas en la figura 5.19. . . . .	79
5.21	Acercamiento a una parte de la imagen Lena reconstruida de las representaciones en parámetros característicos con (a) $M = 2, T = 2$ , (b) $M = 4, T = 4$ y (c) $M = 8, T = 8$ . Aumentada 2 veces. . . . .	80
5.22	Gráficas que corresponden a los valores de relación señal a ruido (PSNR) en cada representación de las imágenes (a) Lena y (b) Casa, para diferentes tamaños de ventana $M$ . Análisis realizado en una escala en las representaciones bidimensional, unidimensional y en parámetros característicos (2D, 1D y PC respectivamente). . . . .	80
5.23	Valores de Entropía en cada representación de las imágenes (a) Lena y (b) Casa con diferentes tamaños de ventana $M$ . El análisis se realizó en una escala para las representaciones 2D, 1D y PC. . . . .	81
5.24	Tasa de compresión en cada representación de las imágenes (a) Lena y (b) Casa, con diferentes tamaños de ventana $M$ . El análisis se realizó en una escala para las representaciones 2D, 1D y PC. La tasa de compresión se obtiene después de realizar codificación de Huffman. . . . .	81
5.25	Representación polinomial bidimensional en dos escalas de la imagen Lena con tamaño de ventana $M = 2$ y $T = 2$ . Los coeficientes se encuentran dispuestos de la siguiente manera: en la fila superior $f_{0,0,0}, f_{0,0,1}, f_{0,0,2}$ , en la siguiente fila $f_{0,1,0}, f_{0,1,1}$ y el conjunto de imágenes que pertenece a los coeficientes del siguiente nivel, en el mismo orden $f_{1,0,0}, f_{1,0,1}, f_{1,0,2}$ , abajo de ellos $f_{1,1,0}, f_{1,1,1}$ ; en la fila inferior se encuentran los coeficientes $f_{0,2,0}, f_{1,2,0}$ . . . . .	83
5.26	Reconstrucción de la imagen Lena a partir de la representación bidimensional de la figura 5.25. Arriba se tiene la imagen reconstruida y abajo de ella el error de reconstrucción. . . . .	84
5.27	Se observan dos conjuntos de imágenes que corresponden a la representación bidimensional de la imagen Casa en dos escalas con tamaño de ventana $M = 4$ y $T = 4$ . Los coeficientes se encuentran dispuestos de la misma manera para los dos conjuntos, formando una matriz triangular superior. . . . .	85
5.28	Reconstrucción de la imagen Casa a partir de la representación bidimensional de la figura 5.27. Arriba se tiene la imagen reconstruida y abajo de ella el error de reconstrucción. . . . .	86

5.29	Representación bidimensional de la imagen Lena para tres escalas con $M = 2$ y $T = 2$ . Los coeficientes se encuentran dispuestos de la misma manera en los tres conjuntos: $f_{0,0}, f_{0,1}, f_{0,2}$ en la primera fila, $f_{1,0}, f_{1,1}$ en la segunda fila y $f_{2,0}$ al final. . . . .	87
5.30	Representación bidimensional de la imagen Casa para tres escalas de análisis con $M = 4$ y $T = 4$ . Los coeficientes se encuentran ordenados de la misma forma para los tres conjuntos, en la primera fila $f_{0,0}, f_{0,1}, f_{0,2}, f_{0,3}, f_{0,4}$ , la segunda fila $f_{1,0}, f_{1,1}, f_{1,2}, f_{1,3}$ , la siguiente fila $f_{2,0}, f_{2,1}, f_{2,2}$ , la penúltima fila $f_{3,0}, f_{3,1}$ y la última fila $f_{4,0}$ . . . . .	88
5.31	Reconstrucción de las imágenes (a) Lena y (b) Casa a partir de las representaciones bidimensionales en tres escalas. Abajo de las imágenes se encuentra su imagen de error correspondiente. . . . .	89
5.32	Representación polinomial unidimensional de la imagen Lena con $M = 2$ y $T = 2$ . En la parte superior se tienen los coeficientes $K_{0,1}, K_{0,2}$ , abajo a la izquierda se encuentra la imagen de orientaciones $\theta_0$ , a la derecha se tiene el conjunto de imágenes que corresponde al segundo nivel, los coeficientes $K_{1,1}, K_{1,2}$ se encuentran en la parte superior y $\theta_1$ y el valor medio $A_t$ en la parte inferior. . . . .	90
5.33	Representación polinomial unidimensional en dos escalas de análisis de la imagen Casa con tamaño de ventana $M = 4$ y $T = 4$ . Los coeficientes se encuentran dispuestos de la misma forma en los dos conjuntos. En la fila superior se tienen los coeficientes $K_{0,1}, K_{0,2}, K_{0,3}$ , en la fila inferior se muestran $K_{0,4}, \theta_0$ y el conjunto de coeficientes del siguiente nivel $K_{1,1}, K_{1,2}, K_{1,3}, K_{1,4}, \theta_1, K_{1,0}$ . . . . .	91
5.34	Representación polinomial unidimensional de la imagen Lena con $M = 8$ y $T = 8$ . El conjunto de imágenes del segundo nivel se observa en la parte superior izquierda. . . . .	91
5.35	Imágenes reconstruidas (a) Lena y (b) Casa, a partir de las representaciones unidimensionales mostradas en las figuras 5.32 y 5.33 respectivamente. Abajo se encuentra la imagen de error correspondiente. . . . .	92
5.36	Representación polinomial unidimensional en tres escalas de la imagen Lena con $M = 2$ y $T = 2$ . Se observan tres conjuntos de imágenes, en la parte superior los coeficientes $K_{0,1}, K_{0,2}$ . En la parte inferior, $\theta_0$ y los coeficientes que corresponden a los siguientes niveles. En la parte inferior derecha se muestra el coeficiente de orden 0 en el último nivel, $K_{2,0}$ . . . . .	93
5.37	Representación polinomial unidimensional en tres escalas de la imagen Casa con $M = 4$ y $T = 4$ . Se observan los coeficientes que corresponden al primer nivel $K_{0,1}, K_{0,2}, K_{0,3}$ , abajo $K_{0,4}$ y $\theta_0$ . En la siguiente imagen se observan los conjuntos de coeficientes que corresponden a los siguientes niveles de análisis. . . . .	94

5.38	Reconstrucción de las imágenes (a) Lena y (b) Casa a partir de las representaciones mostradas en las figuras 5.36 y 5.37 respectivamente. Abajo de ellas se encuentra su correspondiente imagen de error. . . .	95
5.39	Representación en parámetros característicos de la imagen Lena en dos escalas de análisis con $M = 2$ y $T = 2$ . En la parte superior se observan las imágenes de distancia $x_b$ y magnitud del borde $\Delta f$ . Abajo se encuentra la imagen de orientaciones y el conjunto de imágenes que corresponden al siguiente nivel de análisis con la imagen de baja frecuencia $A_b$ . . . . .	96
5.40	Representación en parámetros característicos de la imagen Casa en dos escalas de análisis con $M = 4$ y $T = 4$ . En la parte superior se observan las imágenes de distancia $x_b$ y magnitud del borde $\Delta f$ . Abajo se encuentra la imagen de orientaciones y el conjunto de imágenes que corresponden al siguiente nivel de análisis con la imagen de baja frecuencia $A_b$ . . . . .	96
5.41	Representación en parámetros característicos de la imagen Lena en dos escalas de análisis con $M = 8$ y $T = 8$ . En la parte superior se observan las imágenes de distancia $x_b$ y magnitud del borde $\Delta f$ . Abajo se encuentra la imagen de orientaciones y el conjunto de imágenes que corresponden al siguiente nivel de análisis con la imagen de baja frecuencia $A_b$ . . . . .	97
5.42	Reconstrucción de las imágenes (a) Lena y (b) Casa a partir de las representaciones mostradas en las figuras 5.39 y 5.40 respectivamente. Abajo se muestra la imagen de error. . . . .	98
5.43	Representación en parámetros característicos de la imagen Lena en tres escalas con $M = 2$ y $T = 2$ . Se observan tres conjuntos de imágenes, en la parte superior las imágenes de distancia $x_b$ y magnitud del borde $\Delta L$ . En la parte inferior, $\theta_0$ y las magnitudes y distancias de bordes en los siguientes niveles se encuentran a la derecha. En la parte inferior derecha se muestra la señal de baja frecuencia $A_b$ en el último nivel. . . . .	99
5.44	Representación en parámetros característicos de la imagen Casa en tres niveles para $M = 4$ y $T = 4$ . Se observan tres conjuntos de imágenes, en la parte superior las imágenes de distancia $x_b$ y magnitud del borde $\Delta L$ . En la parte inferior, $\theta_0$ y las magnitudes y distancias de bordes en los siguientes niveles se encuentran a la derecha. En la parte inferior derecha se muestra la señal de baja frecuencia $A_b$ en el último nivel. . . . .	99
5.45	Reconstrucciones de las imágenes (a) Lena y (b) Casa con sus respectivas imágenes de error debajo de ellas. Se obtienen a partir de las representaciones mostradas en las figuras 5.43 y 5.44, con valores de $M = 2, 4$ y $T = M$ . . . . .	100

5.46	Gráficas que corresponden a los valores de relación señal a ruido (PSNR) en cada representación de las imágenes (a) Lena y (b) Casa, para diferentes tamaños de ventana $M$ . Análisis realizado en dos escalas en las representaciones bidimensional, unidimensional y en parámetros característicos (2D, 1D y PC respectivamente). . . . .	101
5.47	Valores de Entropía en cada representación de las imágenes (a) Lena y (b) Casa con diferentes tamaños de ventana $M$ . El análisis se realizó en dos escalas para las representaciones 2D, 1D y PC. . . . .	102
5.48	Tasa de compresión en cada representación de las imágenes (a) Lena y (b) Casa, con diferentes tamaños de ventana $M$ . El análisis se realizó en dos escalas para las representaciones 2D, 1D y PC. La tasa de compresión se obtiene después de realizar codificación de Huffman. . . . .	102
5.49	Gráficas que corresponden a los valores de relación señal a ruido (PSNR) en cada representación de las imágenes (a) Lena y (b) Casa, para diferentes tamaños de ventana $M$ . Análisis realizado en tres escalas en las representaciones bidimensional, unidimensional y en parámetros característicos (2D, 1D y PC respectivamente). . . . .	103
5.50	Valores de Entropía en cada representación de las imágenes (a) Lena y (b) Casa con diferentes tamaños de ventana $M$ . El análisis se realizó en tres escalas para las representaciones 2D, 1D y PC. . . . .	104
5.51	Tasa de compresión en cada representación de las imágenes (a) Lena y (b) Casa, con diferentes tamaños de ventana $M$ . El análisis se realizó en tres escalas para las representaciones 2D, 1D y PC. La tasa de compresión se obtiene después de realizar codificación de Huffman. . . . .	105
5.52	Imagen Lena recuperada después de la compresión con JPEG con diferente calidad de reconstrucción: (a) 100 %, (b) 75 %, (c) 50 % y (d) 25 %. . . . .	106
5.53	Imagen Lena recuperada después de la compresión con JPEG con diferente calidad de reconstrucción: (a) 10 %, (b) 5 %, (c) 3 % y (d) 1 %. . . . .	107
5.54	Gráficas de desempeño del JPEG aplicado sobre la imagen Lena. (a) Corresponde a los valores de PSNR, (b) a los valores de tasa de compresión. . . . .	108

# Capítulo 1

## Introducción

Cuando se habla de una imagen digital, se trata de una variación de intensidades dispuestas en una rejilla bidimensional que representa a una escena real. En la figura 1.1, se puede observar de manera general como se efectúa la formación de dichas imágenes.

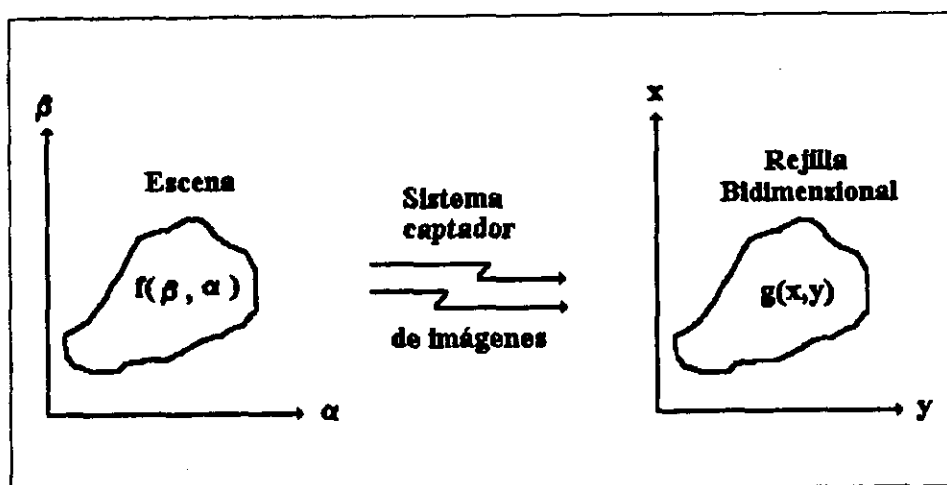


Figura 1.1: Concepto general de formación de una imagen digital.

La imagen digital se denota por  $g(x,y)$  que se formó a partir de la energía radiante captada por un sensor, esto quiere decir que intervinieron muchos factores como iluminación de la escena, posición de la cámara, etc.. Se puede lograr entonces el modelado de estas características a partir del análisis y procesamiento de una imagen. Los métodos de modelado y procesamiento de imágenes dan lugar a sistemas de comunicaciones visuales, que se encargan de presentar alguna escena dada por medios electrónicos.

Los sistemas de comunicación visual han presentado un notable auge en los últimos años, a partir del surgimiento de la interconexión de redes computacionales. Actualmente la computadora es la herramienta de comunicaciones más poderosa, con ella se realiza transmisión y recepción de todo tipo de información, voz, texto e imágenes.

Ejemplos de sistemas de comunicación visual son la televisión, el videoteléfono, videoconferencias y transmisión de imágenes por satélite. Las imágenes generadas en un sistema visual de comunicaciones se procesan para el buen aprovechamiento de los recursos del sistema, de ahí surge un área llamada *procesamiento de imágenes*.

## 1.1 Análisis y procesamiento de imágenes

El problema más común que se plantea en el área de *procesamiento de imágenes* es la búsqueda de algoritmos capaces de obtener cambios de representación eficientes con la finalidad de realizar almacenamiento, codificación y encriptado. Las técnicas de codificación, compresión y reducción de ruido, están revolucionando a los sistemas de comunicación visual, ejemplo de ello es la transformada ondícula <sup>1</sup> [Mal92a], la transformada coseno [Rab91], y la transformada polinomial [Mar90b], que es la técnica que aquí se utiliza. Una manera de lograr que todas estas técnicas realicen su función con un mejor desempeño es a través de una analogía con el comportamiento del sistema de visión humano partiendo de la idea de que la "transmisión de información no relevante para visualización resulta en pérdida de recursos del canal" [Net89], pues no logra una mejora en la calidad visual de la imagen.

Paralelamente surge la utilidad del *análisis de imágenes* en diversas áreas de la ciencia y esta muy ligado al procesamiento. De hecho, para que un sistema de comunicaciones visuales funcione eficientemente es necesario que la información que maneja, en este caso imágenes, sea analizada con la finalidad de conocer los elementos que la componen y sus propiedades, enseguida tomar decisiones sobre el tratamiento que se le dará con el objetivo de optimizar el uso de recursos. El análisis de imágenes ha tenido gran evolución, desde aquellas técnicas donde solo se investigaba acerca de la naturaleza frecuencial (transformada de Fourier), hasta las técnicas de análisis escala-espacio (transformada ondícula y polinomial), que extraen mayor información de la imagen gracias a la conjunción del dominio del tiempo y el de la frecuencia o escala de la imagen. Se sabe que en diferentes escalas de la imagen se tiene información diferente. En el plano temporal se produce la misma situación.

Una técnica que modela eficientemente el sistema de visión humano es la *transformada polinomial*, la cual realiza el análisis por medio de la localización de la imagen con bancos de filtros polinomiales. Dichos filtros se generan a partir de los polinomios y no son ortogonales entre sí, gracias a ello la información que generan es redundante. Por lo cual, la *transformada polinomial* es rica en detalles que se aprovechan para definir exactamente características perceptuales de la imagen.

El funcionamiento de la transformación de manera general es el siguiente: la imagen original se localiza por medio de filtros en las posiciones indicadas por una rejilla de muestreo (Lattice). Estos filtros se obtienen de la multiplicación de una función

---

<sup>1</sup>En inglés: "Wavelet Transform"



ventana con los polinomios de orden  $n$ , tales filtros son del tamaño de la ventana. El resultado del filtrado de la imagen son coeficientes polinomiales cuyo orden se determina por el orden del filtro, tomando en cuenta que en el caso discreto, el tamaño de la ventana determina el máximo orden del filtro. Esta información generada es mayor a la inicial debido a una característica importante del filtro, su no ortogonalidad, por lo tanto los filtros de análisis y síntesis no son iguales.

Los filtros de la *transformada polinomial*, generados a partir de los polinomios, son bidimensionales. esto resulta en una representación cuyos coeficientes son bidimensionales también. Dicha descomposición puede reducirse a partir de la proyección de los coeficientes bidimensionales a coeficientes unidimensionales. La proyección se logra buscando el ángulo donde el contraste es mayor y se almacena la información de magnitud y orientación. De esa manera se obtiene la representación unidimensional, la cual se utiliza para conocer características de las estructuras perceptualmente importantes que componen a la imagen.

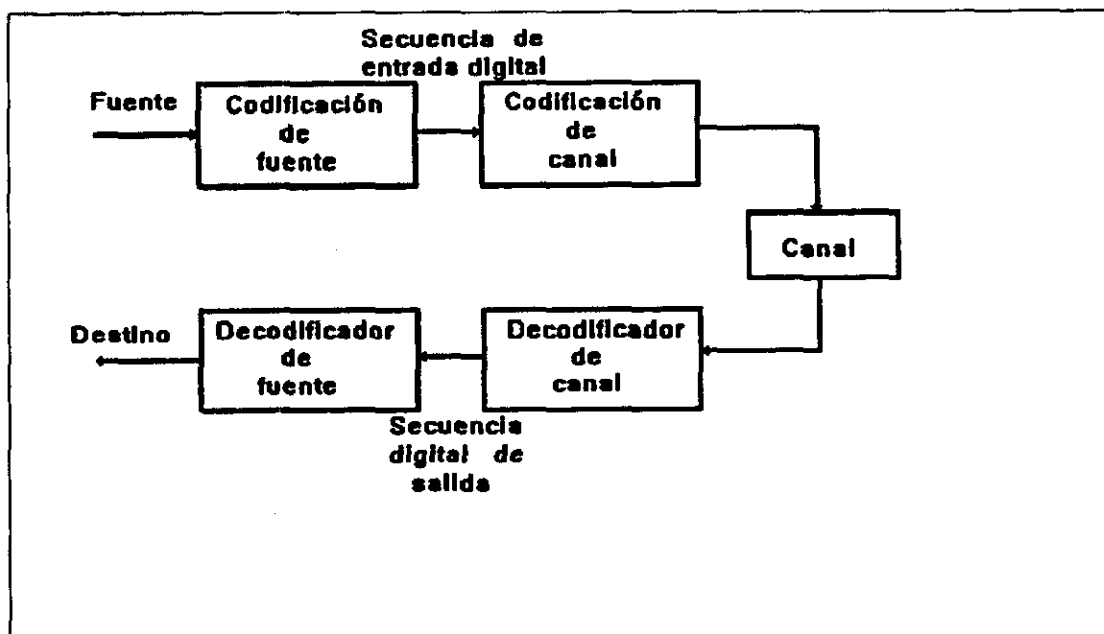


Figura 1.2: Sistema de comunicaciones digitales básico.

## 1.2 Codificación de imágenes

Existen dos problemas dentro de un sistema digital de comunicaciones común, uno es tratar de minimizar el número de bits a transmitir en el canal, el otro problema es asegurarse que los bits transmitidos en un canal se reciban a pesar de las degradaciones sufridas por ruido, interferencias o debilitamiento; de tal manera que se pueda recuperar la información con cierta fidelidad. Los problemas se resuelven con codificación de fuente y de canal respectivamente.

La figura 1.2 ilustra un sistema de comunicaciones básico, las partes que lo componen se describen a continuación:

- **La fuente**, es el lugar donde se genera la información a procesar. En un sistema de comunicaciones visuales se trata del sistema captador de imágenes, que entrega una imagen digital a la siguiente etapa.
- **Codificación de fuente**, es necesaria para lograr un cambio en la representación de la imagen por medio de alguna técnica de procesamiento de imágenes. Se logra aquí mejorar la relación señal a ruido a una tasa de bits dada o viceversa.
- **Codificación de canal**, se refiere a una serie de transformaciones que mejoran el desempeño de las comunicaciones haciendo menos sensible la información a problemas tales como deterioro de canal, ruido, interferencias. Se reduce la probabilidad de error en bits o símbolos.
- **Canal**, es el medio por el cual se transporta la información. Dependiendo de la calidad del canal, es la cantidad de transformaciones que se requieren para conservar la calidad de la información.
- **Los decodificadores**, realizan las transformaciones inversas para recuperar la imagen original.

### 1.3 Objetivos

Se ha resaltado la importancia de la codificación de fuente en un sistema de comunicaciones, por lo cual, el objetivo principal de este trabajo es lograr un cambio de representación eficiente de la imagen usando la *transformada polinomial* y sus características de imitación de las propiedades del sistema de visión humano (SVH). El término eficiente, se refiere a la capacidad del algoritmo para eliminar la redundancia, generada durante el análisis de la imagen. Los puntos a alcanzar durante el desarrollo del presente trabajo se enlistan a continuación:

- Conocimiento de la *transformada polinomial* como técnica de análisis de imágenes.
- Lograr cambios de representación eficientes con la variación de algunos parámetros, como son:
  1. Tamaño de la ventana ( $M$ ).
  2. Número de niveles.
  3. Cantidad de muestreo ( $T$ ).

- Clasificados en tres tipos de representaciones
  1. **Bidimensional**, coeficientes que entrega la *transformada polinomial*.
  2. **Unidimensional**, proyección de los coeficientes bidimensionales.
  3. **Parámetros Característicos**, extracción de estructuras perceptualmente importantes de la imagen.
- Comparación de los resultados obtenidos.
- Propuesta de nuevos algoritmos de procesamiento de imágenes usando la *transformada polinomial*.

## 1.4 ¿Cómo alcanzar los objetivos?

La clasificación de patrones o estructuras primarias en la imagen que son importantes para el sistema de visión humano da lugar a la discriminación de información que no es relevante para el mismo. Para ello se utilizan técnicas de análisis que imitan el comportamiento del SVH. Como ejemplo están las pirámides gaussianas, la transformada de Gabor, la transformada ondícula y la transformada polinomial que cuentan con filtros que tienen un comportamiento muy similar al SVH y que se basan principalmente en la teoría escala-espacio.

Con la transformada polinomial se puede lograr una representación de la imagen en la cual es posible identificar estructuras primarias importantes como esquinas, bordes, líneas, puntos, etcétera. Una vez que se han identificado se extraen valores que las caracterizan. Ya que una de las estructuras primarias más importantes es el borde, entonces se toman en cuenta los siguientes valores:

- Posición del borde.
- Magnitud del borde.
- Angulo respecto del centro de la ventana de análisis.
- Señal de baja frecuencia o valor medio.

Después de analizar la imagen con la transformada polinomial se obtiene una representación bidimensional de la imagen, esta representación se reduce a una representación unidimensional y a partir de ahí se logra la caracterización de estructuras

unidimensionales perceptualmente más importantes -en este caso bordes -, que componen a la imagen. El esquema de codificación piramidal a múltiples resoluciones se propone para disminuir la entropía de la información.

## 1.5 Contenido

El texto trata de una manera introductoria conceptos sobre el sistema de visión humano, procesamiento de imágenes y teoría de la información. La razón de esta introducción es que el algoritmo comprende estas tres áreas, pues se trata de un algoritmo de análisis y codificación de imágenes, implicando procesamiento.

En base a los objetivos planteados, se incluye en el texto una descripción del funcionamiento del algoritmo de extracción de parámetros de estructuras unidimensionales, que es la parte esencial del trabajo desarrollado. Dentro del mismo capítulo se hace mención a las pruebas realizadas con los esquemas de codificación utilizados. La extracción de parámetros se realiza a partir de una técnica de análisis llamada *transformada polinomial*, razón por la cual, se le dedica especial atención a la descripción de su funcionamiento. Se inicia con el comportamiento de la técnica con señales unidimensionales hasta el comportamiento en imágenes en un esquema a múltiples escalas de análisis.

La parte de resultados es necesaria para validar los algoritmos propuestos. Se aplican a las imágenes típicas (Casa y Lena) y se presentan gráficas de desempeño a manera de un breve estudio comparativo entre los métodos propuestos. Se indican aquí los criterios de evaluación del desempeño de los algoritmos propuestos.

Para finalizar se tienen conclusiones que se generaron en el desarrollo del trabajo y perspectivas de la presente investigación.

# Capítulo 2

## Conceptos básicos

Este capítulo presenta un panorama general del área de comunicaciones visuales; inicia con una breve descripción del sistema de visión humano y su funcionamiento, útil para el modelado posterior y su implantación computacional. Más adelante se realiza un recuento de los conceptos más importantes en el análisis de imágenes y se hace énfasis en su utilidad. Enseguida la descripción del procesamiento de imágenes, donde se mencionan la definición, utilidad y subareas que lo componen. Al final se tiene codificación y compresión de imágenes dentro del contexto de una descripción conceptual con la teoría de la información.

### 2.1 Sistema de visión humano

El sentido de la vista es importante pues significa el descubrimiento de lo que existe en el mundo y donde se ubica, a partir de él podemos percibir una gran cantidad de información por medio de imágenes. Esta información se encuentra en una pequeña parte del espectro de radiación electromagnético, con esto se quiere enfatizar que el subsistema de formación de imágenes del ser humano es bastante limitado (longitud de onda de 400 a 700 nm, del violeta al rojo); sin embargo, el sistema de visión tiene la capacidad de analizar e interpretar las imágenes percibidas a una velocidad sorprendente [Hen93].

La importancia del estudio del sistema de visión humano (SVH) radica en la posibilidad de conocer aquellas características importantes al ojo humano que se pueden tomar en cuenta para la transmisión, codificación y representación de la información de una manera eficiente. Es obvio que la transmisión y despliegue de información no importante a la vista dá como resultado una pérdida en recursos de canal y medios de despliegue, así que se debe transmitir la información con la mínima precisión requerida por el ojo humano y lograr de esta forma el funcionamiento óptimo de un sistema de comunicaciones visuales. De hecho la transmisión perfecta es imposible, como ejemplo en sistemas de televisión se incluyen distorsiones sistemáticas debido al formato propio del despliegue.

### 2.1.1 ¿Cómo funciona el sistema de visión humano?

Desde el principio de los tiempos se ha tratado de encontrar una explicación al proceso de visión; de las teorías más antiguas se conoce la que corresponde a Aristoteles, quien decía que “la visión consiste de alteraciones que el objeto realizaba al medio que se encontraba entre el objeto y el ojo propagando información, de tal manera, que al ser captada se lograba visualizarlo” [Asi84], estas ideas continuaron por mucho tiempo; fue hasta la época renacentista que se realizaron investigaciones importantes empezando con Vesalius, que realizó disecciones del ojo humano; más tarde Kepler en Alemania y Descartes en Francia, quienes realizaron profundas investigaciones acerca de la participación de la retina en la captación de la luz; y así hasta el siglo XIX con Thomas Young quien descubrió la importancia de la cornea en el proceso de visión. Actualmente se han desarrollado importantes descubrimientos que han permitido conocer de una manera más clara el proceso de visión [Bal82].

En la figura 2.1 se puede ver un diagrama de las partes que componen el ojo humano, la cornea, lente ocular, pupila, iris, retina y el nervio óptico. Las imágenes se perciben en el ojo humano con una rapidez de actualización sorprendente, basta mirar a nuestro alrededor para darnos cuenta de tal capacidad.

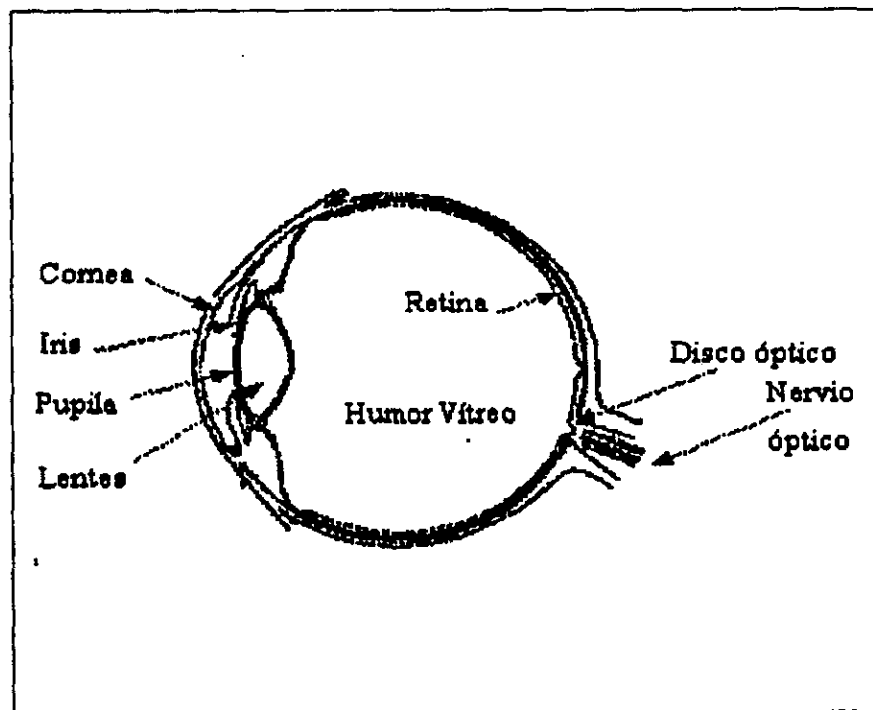


Figura 2.1: Diagrama del ojo humano

La visión es un proceso que involucra a las partes anatómicas del ojo humano y al cerebro. La retina es la parte medular del proceso de visión, las neuronas fotoreceptoras en la retina reciben la luz y la convierten en impulsos que son enviados al cerebro, que interpreta la información de lo que estamos viendo. La retina, esta

formada por tres capas de neuronas, en ella se encuentra la fovea compuesta principalmente de conos que son los responsables de la detección de color, y de bastones, responsables de la visión nocturna. La luz es reflejada por los objetos y se enfoca en la retina por medio de la cornea y los lentes oculares, en la retina se realiza la conversión de energía luminosa a impulsos eléctricos que se transportan al cerebro a través del nervio óptico, donde se interpreta la información.

### **¿Cómo se forman las imágenes en el cerebro?**

Una escena real incide sobre las células receptoras que se encuentran ubicadas en el globo ocular, quienes a su vez envían impulsos eléctricos a través del nervio óptico al cerebro, es decir que la imagen se adquiere como tal y en el cerebro solo se recibe la información comprimida para ser interpretada.

Es de gran interés conocer la naturaleza del procesamiento de la información que se lleva a cabo en el sistema visual humano, por ejemplo, la diferencia entre los conos y los bastones como células receptoras. Asimismo causa sorpresa saber que existen más de 125 millones de bastones y conos en la retina pero solo alrededor de 1 millón de neuronas (células ganglioretinales) que alimentan al nervio óptico; por lo tanto, ahí se está efectuando una compresión de imágenes por algunas células intermedias llamadas células bipolares, la compresión es mayor que 125:1 en la parte lejana de la fovea y un poco menor en el área interior. Finalmente es sorprendente que la retina esté organizada con sus tres regiones en secuencia opuesta, con las células ganglioretinales antes de las células bipolares y éstas enfrente de los conos y bastones[You91][Hen93].

### **2.1.2 ¿Cómo se ha tratado de modelar la visión humana?**

La construcción de modelos matemáticos del procesamiento de actividades fisiológicas se realiza primero conociéndolas. La mayor parte de la investigación se ha generado a través de experimentos con animales que poseen un sistema de visión similar al humano como primates y gatos. Aún así, solo se llevan a cabo mediciones en regiones limitadas del espectro [Lim90].

Uno de los modelos generados a partir de las investigaciones realizadas en el funcionamiento de los campos receptivos, indica que en cada punto del campo visual se tienen cuatro filtros o máscaras que analizan la imagen y simulan la respuesta de los campos receptivos. La forma de estos filtros se aproxima a una diferencia de dos gaussianas, los más pequeños tienen propiedades temporales importantes y los dos más grandes tienen ciertos transitorios. Después de mediciones exactas, se encontró que los filtros  $\nabla^2 G$  son los que mejor se aproximan a esos valores y que se trata de filtros pasabanda cuyos operadores son similares a la diferencia de dos gaussianas [Hen93][Bal82]. El operador también se encuentra en otro experimento realizado con los campos receptivos de las células ganglioretinales cuya organización espacial es circularmente simétrica, pues se observó que la respuesta de dichas celu-

las a la luz hace que las células centrales se exciten y las otras se inhiban, esto se debe seguramente a la superposición de pequeñas regiones excitatorias centrales, cuya respuesta se aproxima por la diferencia de dos gaussianas [Gon87].

Uno de los descubrimientos más importantes fue el rol que jugaban los cruces por cero en la detección de imágenes. Si aplicamos a una imagen el operador que obtuvimos de los experimentos anteriores vamos a notar que se obtiene un mapa de las fronteras de intensidad de una imagen, a esto se le conoce como los cruces por cero de la segunda derivada de una función, ellos nos indican el lugar donde se encuentra un borde, por ello se deduce que dada la función de análisis que se tiene en el sistema de visión humano, podemos decir que una de las estructuras visuales más importantes son los bordes o cambios de intensidad [Mar82].

## 2.2 Análisis de imágenes

Con el rápido desarrollo de instrumentos científicos computarizados surge una amplia variedad de problemas interesantes en el análisis de datos y procesamiento de imágenes, que se aplican en campos tan diversos que van desde astronomía extragaláctica a espectroscopía molecular, imágenes médicas a visión computarizada [Gud95]. Específicamente las imágenes se crean en una creciente tasa por fuentes tales como defensa y satélites civiles, vuelos de reconocimiento, experimentos científicos, imágenes biomédicas y entretenimiento doméstico.

El análisis de cualquier fenómeno es importante porque se pueden conocer así sus propiedades y se infieren los elementos que lo componen y su naturaleza, posteriormente se toman decisiones basadas en dicho análisis. De manera similar con imágenes, es necesario analizarlas para obtener información acerca de sus componentes y la naturaleza de estos [Lir87].

Desde la percepción remota hasta la microscopía y en áreas como medicina o en manufactura y hasta en sistemas de vigilancia, la generación de imágenes es sorprendente, y para realizar algún procesamiento, es necesario primero conocer la imagen y las partes que la componen [Viv96]. Ya se ha visto que la manera en que el cerebro realiza el análisis de imágenes es percibiendo el objeto y mediante un proceso abstracto extrae las propiedades básicas, creando una liga entre las propiedades y la función o naturaleza del objeto, resultando en un mayor conocimiento de sus características con lo cual se identifica su función y se facilita su reconocimiento.

### 2.2.1 ¿Cómo se forma una imagen digital?

Ya se ha tratado el concepto de imagen digital en el capítulo 1. Se añade también que el concepto de imagen digital se refiere a una representación lógica de una escena, numérica y visual. Para describir una imagen es necesario conocer como se efectúa el proceso de formación de la misma, para ello nos referiremos al diagrama de bloques de la figura 2.2, que se compone de la fuente de radiación, la escena a ser captada,



un sistema óptico, filtros o espectrorradiómetro, un convertidor analógico-digital y finalmente la imagen digital.

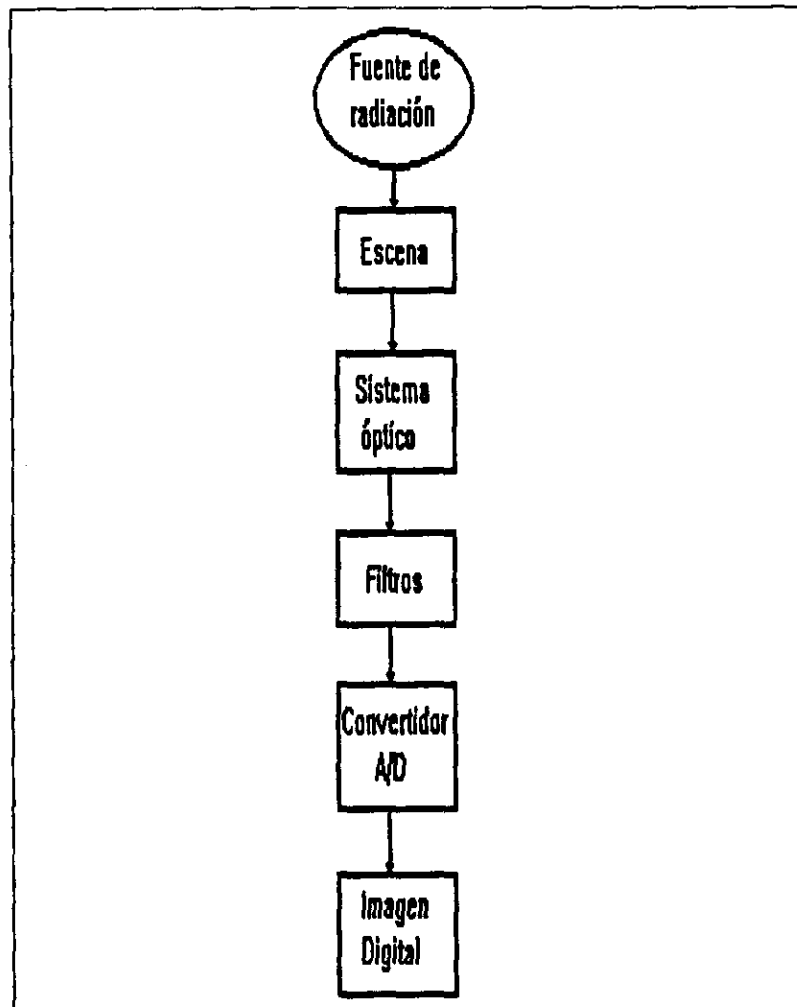


Figura 2.2: Formación de una imagen digital

El diagrama de bloques de la figura 2.2, describe el proceso de formación de una imagen digital; los elementos que la componen son

- **Fuente de radiación**, cuya naturaleza puede ser de tipo ondulatoria (electromagnética o acústica) o corpuscular (neutrones, electrones o positrones), envía la radiación a la escena. En ciertos casos como materiales radioactivos la fuente y la escena están contenidas una en la otra, o con rayos X o haces de neutrones que atraviesan la escena.
- **Sistema óptico**, que se encarga de enviar la energía radiada por la escena a la siguiente etapa; aquí se determina la manera en la que se proyecta la imagen.

- **Escena(3D)**, está compuesta por elementos básicos llamados campo instantáneo de vista, cada uno de ellos interactúa con la fuente de radiación resultando en una emisión de energía al detector.
- **Filtros**, la entrada a esta etapa es la energía radiada de la interacción de la fuente de radiación con un campo instantáneo de vista. Aquí se separan las bandas de energía para el caso de imágenes multispectrales. Para el caso de niveles de gris se determina el valor que corresponde a un pixel.
- **Convertidor A/D**, la naturaleza de las señales en el mundo real es analógica. Debido a la limitación que existe para el tratamiento de este tipo de señales, se realiza una conversión a señales discretas. El término analógico, se refiere a los sistemas y procesos en los cuales las señales con variaciones continuas y uniformes se manipulan. El término digital se refiere a los sistemas y procesos que involucran la manipulación de números discretos. Aún cuando en alguna parte de la formación de la imagen digital esta es de naturaleza continua, se convierte a señal digital por medio de una conversión analógica a digital.
- **Imagen digital**, que es la simplificación de una escena captada en sus representaciones ya mencionadas. Actualmente, la avanzada tecnología en el área de visión permite obtener imágenes generadas con radiación invisible al ojo humano, a pesar de los aspectos del funcionamiento del sistema de visión humano que no se han reproducido en sistemas de computo.

De manera general el funcionamiento común es el siguiente: a partir de la emisión de radiación a la escena por medio de una fuente, una cierta cantidad se refleja debido a la interacción con un elemento físico de la escena que es el campo instantáneo de vista, el resultado de dicha interacción es una cantidad de energía liberada a partir del campo instantáneo de vista que es tomada por el sistema de captación de imágenes a través del sistema óptico y enviada al filtro o conjunto de filtros, en esta etapa se tiene una imagen analógica, sin embargo, no es útil para su procesamiento en computadora, por lo cual, es necesario realizar una conversión a digital por medio de cuantización, finalmente el objetivo es visualizar la imagen en la computadora. Las imágenes se forman como consecuencia de 4 factores entremezclados:

- Geometría.
- Reflectancia de superficies visibles.
- Iluminación de la escena.
- Punto de vista

Se dice que una imagen ha sido analizada, cuando se reconocen e identifican las partes que componen la escena. Algunas imágenes deben sus valores a un factor o a la combinación de algunos de ellos. De tal manera que el proceso primario

de visión <sup>1</sup> determina cuales variaciones de intensidad se deben a que factor y así se crean representaciones en donde los cuatro factores estan separados [Mar82]. Estas representaciones son parte de la teoría de la visión que consiste en hacer una inferencia de las estructuras que se encuentran en una imagen. Se sabe que en terminos generales el mundo visual esta constituido en su parte primaria por bordes, contornos, líneas, puntos, etcétera, y estos son localizados espacialmente.

Desde el punto de vista de la geometría, las imágenes se pueden formar de dos maneras diferentes que son: proyección perspectiva y proyección ortográfica. La diferencia fundamental entre las dos proyecciones, es la siguiente: todo punto de la escena en la *proyección perspectiva* tendrá su punto equivalente en la imagen en la dirección que se indique cuando pase por el origen de los ejes óptico y cartesiano; en la *proyección ortográfica*, el punto de la escena tiene un equivalente en la dirección perpendicular [Hor86].

Las funciones de reflectancia tienen particular importancia en la formación de representaciones de la imagen y se ven influidas por las siguientes consideraciones físicas: (1) la *existencia de superficies*, indican que el mundo visual se encuentra compuesto por superficies suaves a diferentes escalas cuya estructura espacial puede ser elaborada; (2) *continuidad espacial* que nos indica que los bordes se dan siempre de manera suave; por último (3) la *continuidad de flujo* que nos indica las fronteras del objeto gracias al movimiento de sus discontinuidades.

## 2.2.2 Captura de imágenes

Los sistemas de captura de imágenes cubren actualmente todo el espectro de frecuencias conocido, es decir se logran captar imágenes que son invisibles al ojo humano, así por ejemplo tenemos que diferentes sensores detectan radiación en distintas regiones espectrales, para detectar rayos gamma ( $10^{-16}$  a  $10^{-5}\mu m$ ) se utilizan contadores de centelleo y fotomultiplicadores; los contadores geiger, detectan rayos X ( $10^{-5}$  a  $0.004\mu m$ ); la luz ultravioleta ( $0.004$  a  $0.38\mu m$ ) se puede captar con película fotográfica con lentes de cuarzo y fotomultiplicadores; para captar el rango de luz visible ( $0.38$  a  $0.78\mu m$ ) existen un gran número de sistemas de captura, como televisión, camaras convencionales, detectores de estado sólido; la radiación infrarroja ( $0.38$  a  $0.78\mu m$ ) se detecta con película sensible al infrarrojo y detectores de estado sólido; finalmente las microondas ( $10^3$  a  $10^6\mu m$ ) se detectan con antenas barreadoras y rastreadoras. Algunos de estos sensores realizan la entrega de información de la imagen en forma numérica [Lim90] [Gud95].

## 2.2.3 Técnicas de análisis de imágenes

El análisis y representación de señales e imágenes tiene particular importancia cuando se quiere efectuar algun tipo de procesamiento, tal como codificación, restauración, reducción de ruido, etcétera. Ya que así se puede lograr una mejor eficiencia

---

<sup>1</sup>En inglés: "Raw Primal Sketch".

en el desempeño de los algoritmos. En otras palabras, conociendo la naturaleza del fenómeno se pueden estimar o eliminar algunas características no importantes, de esta manera se logra una representación que es óptima para la aplicación final que se busca.

Las técnicas de análisis de imágenes se pueden clasificar como se indica a continuación:

- **Extracción de parámetros característicos**, es un conjunto de técnicas que trata de modelar de manera cuantitativa las estructuras de una imagen.
- **Segmentación**, separa los objetos que componen la escena para realizar un estudio de cada uno de ellos.
- **Clasificación**, los objetos que se han identificado son ahora acomodados en diferentes categorías.

Dentro del conjunto de extracción de parámetros característicos se encuentra la extracción espacial de parámetros, donde el interés consiste en conocer la magnitud de propiedades físicas de una imagen tales como reflectividad o respuesta multiespectral, un ejemplo es con las imágenes infrarrojas donde el valor de magnitud indica la temperatura y eso facilita la segmentación de nubes y parte terrestre; las características estadísticas de una imagen también son útiles para esos fines; una parte importante de estas técnicas son las transformaciones. La transformada de Fourier, por ejemplo, realiza una descripción de los componentes frecuenciales de una imagen que se pueden discriminar [Pro88], lo mismo con otras técnicas como la transformada Polinomial de donde podemos obtener valores de magnitud y orientación de un borde [Viv97]; por último, la detección de bordes y texturas también se utilizan para la extracción de parámetros característicos.

En segmentación, las técnicas más comunes son los umbrales en amplitud, donde una vez identificado el valor de dicha amplitud se pueden discriminar componentes no necesarios para el análisis. El etiquetado de componentes es un método de segmentación de imágenes binarias que examina la conectividad de los pixels con sus vecinos identificando y etiquetando los que pertenecen a un mismo conjunto. La aproximación basada en fronteras de discontinuidad, que segmenta objetos en base a sus perfiles o contornos, de tal forma que algunas técnicas como seguimiento de contornos, conectividad, eslabonamiento de bordes y transformada Hough son aplicables a segmentación de imágenes [Gon87]. Las aproximaciones basadas en regiones y agrupamiento, cuya idea principal es identificar varias regiones con las mismas características. La comparación por templates como su nombre lo indica se refiere a la definición de un template que se compara con la imagen para definir regiones que se le parezcan. La segmentación por textura, muy útil en la segmentación de objetos que tienen un fondo con textura rugosa, donde la segmentación por bordes tiene algunos errores.

Las técnicas de clasificación más utilizadas son muy parecidas a las de segmentación pues estas dos tareas del análisis de imágenes están muy ligadas. Se tiene la clasificación supervisada con dos vertientes, los métodos de distribución libre, donde no se requiere conocimiento a priori de funciones de probabilidad y se basan en el razonamiento y la heurística, dentro de esta categoría se encuentran los árboles de decisión. Las técnicas estadísticas se basan en modelos de distribución de probabilidad y pueden ser paramétricos (como la función gaussiana) o no paramétricos.

## 2.3 Procesamiento digital de imágenes

Por *procesamiento digital de imágenes* se entiende “la generación de imágenes a partir de otras ya existentes con el objetivo de hacer más fácil su interpretación, resaltando o discriminando características importantes para el sistema de percepción humana” [Hor86]. Generalmente las técnicas de procesamiento de imágenes provienen de la teoría de sistemas lineales. El resultado del procesamiento, puede ser una imagen cuyo ruido haya desaparecido, con perfilación realizada o con acentuación de bordes y esquinas. Se puede decir que “el procesamiento digital de imágenes tiene sus orígenes en la percepción remota” [Jai89], ya que las imágenes que se obtenían del sistema de captación de imágenes de los aviones de reconocimiento generalmente tenían algún tipo de degradación como perfilación, distorsiones geométricas y ruido de fondo, razón por la cual, era necesario realizar el procesamiento adecuado para recuperar de la mejor manera la imagen original. Hasta la fecha todas las imágenes que se obtienen de satélites y aviones de reconocimiento son procesadas antes de publicarlas.

### 2.3.1 ¿Qué áreas componen al procesamiento digital de imágenes?

Dentro del procesamiento digital de imágenes, encontramos las siguientes áreas:

- **Realce:** Para mejorar la percepción de la imagen con base en características importantes para el observador humano, se parte de que una imagen se necesita realzar para una mejor apreciación incluso si ésta no sufre degradación alguna.
- **Restauración:** Es necesaria en la reducción o eliminación de alguna degradación producida en la imagen. Aunque de cierta manera se lleva a cabo un realce de la imagen la diferencia es que se parte de que la imagen se encuentra degradada.
- **Codificación:** Se trata de la representación de la imagen con el menor número posible de bits, es decir, se reduce la cantidad de información, conservando una calidad visual de la imagen tal que los observadores puedan apreciar sus características sin problema alguno. Actualmente, esta área está muy ligada a las

dos anteriores pues con realce o restauración se eliminan algunas características que son poco importantes perceptualmente.

- **Comprensión**<sup>2</sup>: Se representa simbólicamente el contenido de una imagen, a diferencia de los anteriores, en esta área la entrada es una imagen y la salida es una representación simbólica, tiene múltiples aplicaciones en visión computarizada, robótica, etc...

Las aplicaciones potenciales en procesamiento de imágenes son ilimitadas, ejemplo de ello son: percepción remota y administración de recursos terrestres, pronóstico de temperatura, sistemas de comunicación y almacenamiento de imágenes, además de sistemas multimedia, sistemas de información geográfica, administración de bases de datos, sistemas médicos, etcétera [Gud95]. Algunas técnicas de procesamiento de imágenes son útiles para entender las limitaciones en el sistema de formación y para el diseño de sistemas de preprocesamiento. Tomando en cuenta que el sentido de la vista es de particular importancia para la interacción de cualquier ser humano con el mundo exterior, se puede deducir que el procesamiento de imágenes tiene todavía un gran potencial de aplicaciones.

## 2.4 Codificación y compresión de imágenes.

Codificación significa un cambio de representación de la información que contiene una imagen con el objetivo de hacer explícitas ciertas entidades y entenderlas mejor. Dicho de otra forma, se sabe que la información se representa por medio de datos que generalmente en los sistemas de despliegue visual se presentan como un conjunto finito de niveles de intensidad. Dicha representación, depende de la aplicación que se requiera; por ejemplo, si lo que se desea es visualizar la información en una computadora es la representación numérica espacial la que da idea de los diferentes niveles de intensidad; pero si deseamos conocer sus componentes frecuenciales, entonces por medio de la transformada de Fourier lo podemos hacer; si lo deseable es realizar una representación con el menor número de datos, entonces las técnicas de compresión son muy útiles.

En un sistema de comunicaciones básico como el presentado en la figura 1.2. se observan los bloques que corresponden a la codificación de fuente y canal. La codificación de fuente forma representaciones eficientes de la imagen tratando de minimizar la cantidad de información. En este caso se utiliza la *transformada polinomial* y codificación de Huffman para realizar la codificación de fuente. La codificación de canal representa un papel importante cuando la fuente es propensa a errores y ruido.

---

<sup>2</sup>En inglés: "Image Understanding"

## 2.4.1 ¿Cuándo se efectúa compresión de datos?

Cuando la cantidad necesaria de datos que se usan para representar la imagen es menor que la cantidad que se utiliza para su visualización se dice que se está realizando compresión de datos. El proceso de compresión consiste en eliminar datos redundantes dentro de la representación. La cantidad de compresión lograda se mide con la tasa de compresión

$$T.C. = \frac{\text{bits\_de\_imagen\_original}}{\text{bits\_de\_imagen\_comprimida}} \quad (2.1)$$

que se ve afectada por los siguientes factores: la tasa de muestreo, cuyo incremento conlleva una compresión mayor; los niveles de cuantización, el aumento en el número de niveles disminuye la tasa de compresión; finalmente, el ruido que reduce la correlación entre los píxeles limitando la cantidad de compresión que se puede lograr.

### 2.4.2.1. Redundancia

Los métodos por los cuales se forman imágenes digitales generan una gran cantidad de datos para representar la información, generalmente la representación de esa información contiene datos que carecen de información, a ellos se les llama datos redundantes; se pueden identificar dos clases de redundancia:

- **Redundancia estadística**, que se refiere a la relación estadística que guardan los datos entre sí, la manera de eliminar este tipo de redundancia es decorrelacionando los datos que componen a la imagen. Dentro de este grupo encontramos cuatro tipos de redundancia,
  - *Redundancia espacial*, es aquella que se debe a la correlación de datos, es decir la relación que guarda el valor de un píxel con el de sus vecinos.
  - *Redundancia espectral*, es la correlación que existe entre varias bandas de la imagen, por ejemplo en imágenes multiespectrales de color (Rojo, Verde, Azul).
  - *Redundancia temporal*, es la correlación entre imágenes de una secuencia, por ejemplo en video digital.
  - *Redundancia del código*, se refiere a la correlación que existe entre los símbolos del alfabeto utilizado [Gon87].
- **Redundancia subjetiva**, que se refiere a aquella información que no es importante a la vista.

La redundancia estadística puede ser removida y posteriormente con la operación inversa generalmente se puede recuperar, no así la redundancia subjetiva que ya no se regenera, sin embargo puede ser aproximada.

## 2.4.2 ¿Porqué existe la necesidad de codificar y comprimir?

Las razones por las que se realiza codificación y compresión de imágenes se listan a continuación

- **Almacenamiento de información.** La gran cantidad de información visual que se genera día con día y las aplicaciones que necesitan de una reducción de datos eficiente han dado lugar a una demanda de algoritmos de compresión. Por ejemplo, si se tiene una imagen astronómica cuyas dimensiones en pantalla son 1024 pixeles por cada lado y el valor de cada pixel es de dos bytes, entonces se necesitan dos megabytes de espacio aproximadamente. El problema surge cuando las imágenes generadas tienen que almacenarse para su posterior visualización [Gud95].
- **Capacidad de canal** Las aplicaciones de los algoritmos de compresión de imágenes en secuencias son una buena razón para comprimir. Como ejemplo tenemos que una imagen de video a color (3 colores) de baja resolución, con calidad de TV, cuyas dimensiones son de 512 pixels por lado y 8 bits por pixel, necesita aproximadamente de  $6 \times 10^6$  bits para ser representada; el problema se agrava aun más cuando se trata de enviar video por línea telefónica usando un modem de 36600 bits por segundo, pues la transmisión toma un tiempo aproximado de 6 minutos para una sola imagen de una secuencia de video [Rab91].
- **Seguridad en la transmisión.** El cambio de representación o codificación es muy utilizado para evitar que la información sea interpretada por algun decodificador no autorizado, en otras palabras se logra el envío de información dirigida a los usuarios deseados, por ejemplo un proveedor de señales puede cancelar el suministro de una de ellas a algun sistema de televisión cuando alguno de los dos lo requiera [Sk188].
- **Transmisión eficiente.** La codificación es muy útil para preservar los datos que contienen la información que es básicamente el objetivo de la codificación de canal, es decir se asegura que los bits que se envían sean los que recibe el decodificador [Vit79].

## 2.4.3 Repaso a la teoría de la información

Para determinar la fiabilidad de un método de codificación, se ha dado especial importancia a la medida subjetiva de la visión, sin embargo existen sistemas de comunicación visual donde el daño causado a la imagen es imperceptible pero es necesario saber que cantidad de información se ha perdido; las medidas que aquí se utilizan para determinar el desempeño de un método de codificación son : entropía y relación señal a ruido, que logran respectivamente una medida cuantitativa de



la información y el error de la imagen para codificación con pérdidas. Para las medidas mencionadas es necesario conocer primero conceptos como información y error cuadrático medio relacionados con la teoría de la información.

## Información

La noticia del clima en alguna ciudad contiene cierta cantidad de información, pero la noticia de un terremoto devastador en ese mismo lugar contiene mayor información ya que se trata de un evento inesperado. Para definir la información de manera cuantitativa contenida en tales eventos se tienen que cumplir las siguientes tres propiedades: (1) La información contenida en los eventos se define como una medida de incertidumbre de los eventos. (2) Los eventos menos probables deben contener mayor información que los eventos más probables. Por último, (3) la información de eventos no relacionados tomados como un solo evento debe ser igual a la suma de información de los eventos.

Se sabe que una medida de incertidumbre es la probabilidad de ocurrencia del evento, a partir de ahí se puede lograr una medida de información de un evento  $\alpha$  en base a su probabilidad  $P(\alpha)$ :

$$I(\alpha) = -\log P(\alpha) \quad (2.2)$$

donde  $I(\alpha)$  es la información, la base del logaritmo determina las unidades. La base dos corresponde a bits [Vit79] [Pap91].

## Fuente discreta sin memoria

La manera más simple de una fuente de información es la fuente discreta sin memoria, que significa que los símbolos sucesivos de salida de la fuente son estadísticamente independientes. Sea un conjunto  $U = \{a_1, a_2, \dots, a_A\}$  con probabilidades  $P(a_k)$  donde  $k = 1, 2, \dots, A$  y cada variable emitida es independiente de las anteriores y posteriores. De acuerdo a la definición de información, si cada salida de la fuente es el evento  $\alpha_k$  entonces la información de cada evento se denota por

$$I(\alpha_k) = -\log P(\alpha_k) \quad (2.3)$$

## Entropía

Si tomamos el promedio de la información en todos los símbolos posibles de la fuente obtenemos la información promedio por símbolo de la fuente  $H(U)$  también conocida como entropía

$$H(U) = \sum_{k=1}^A P(a_k) I(\alpha_k) = -\sum_u P(u) \log_2 P(u) \quad (2.4)$$

donde  $u$  es cada evento del conjunto  $U$ , y el resultado de  $H(U)$  son bits por símbolo. La entropía se entiende como la información por símbolo de la fuente y que es necesaria para determinar un símbolo desconocido [Pee93] [Sk188].

### Relación Señal a Ruido

Para evaluar la calidad de la estimación de las imágenes se proponen el error cuadrático medio (RMSE<sup>3</sup>) y la relación señal a ruido de reconstrucción pico (PSNR<sup>4</sup>) de la imagen como medidas de error.

El error cuadrático medio en este caso nos da la información del promedio de error de la imagen, como se puede observar en la siguiente ecuación:

$$RMSE = \sqrt{\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N [f(i, j) - \hat{f}(i, j)]^2} \quad (2.5)$$

que representa la desviación estándar del error de la imagen [Rab91]. A partir de ahí, se puede obtener la relación señal a ruido pico (PSNR en dB) utilizando

$$PSNR = 20 \log_{10} \left( \frac{255}{RMSE} \right) \quad (2.6)$$

Para una imagen de 8 bits, cuyos niveles de intensidad son 256 (0-256). Sin embargo, la forma más eficiente de evaluar la calidad de una imagen sigue siendo la apreciación subjetiva de la imagen. Se sabe que un RMSE alto o PSNR bajo no siempre implica una calidad subjetiva deplorable. Aunque se establece como regla que una imagen con una PSNR menor a 30 dB presenta degradaciones visibles, es decir la calidad visual empieza a decrecer rápidamente.

#### 2.4.3.1 Teoremas fundamentales de codificación.

Los tres teoremas básicos relacionados con la codificación o representación de información se describen en esta sección.

##### Teorema de codificación sin ruido

Se considera un canal sin ruido cuando la señal que entrega la codificación de fuente no sufre alteración alguna antes de su llegada a la decodificación de fuente, por lo tanto la tarea de la codificación de fuente es entonces realizar la mayor compresión de los datos.

<sup>3</sup>En inglés: "Root Mean Square Error"

<sup>4</sup>En inglés: "Peak Signal to Noise Ratio"

El teorema de codificación sin ruido, también conocido como el *primer teorema de Shannon*<sup>5</sup>, define el promedio mínimo de la longitud de palabra del código por símbolo de fuente que se puede obtener; también establece que se puede codificar cualquier fuente sin pérdida, con un código cuyo número promedio de bits por símbolo de fuente sea arbitrariamente cercano, pero no menor, a la entropía  $H(U)$  de la fuente en bits.

Si  $U$  es una fuente ergódica con un alfabeto de tamaño  $n$  y una entropía  $H(U)$ . Considerese la codificación de bloques de  $N$  símbolos fuente a la vez en palabras de código binarias. Para cualquier  $\delta > 0$ , es posible que eligiendo  $N$  suficientemente grande, se pueda construir un código de tal manera que el número promedio de bits por símbolo fuente original  $\bar{L}$  satisfaga

$$H(U) \leq \bar{L} < H(U) + \delta \quad (2.7)$$

Si  $H(U)$  es el límite inferior de  $\bar{L}$ , entonces podemos calcular la eficiencia  $\eta$  de un esquema de codificación

$$\eta = \frac{H(U)}{\bar{L}} \quad (2.8)$$

### Teorema de codificación con ruido

Se ha mencionado ya que si un sistema de comunicaciones presenta alteraciones que den como consecuencia ruido en la transmisión de la señal, entonces la importancia de la codificación se enfoca en la fiabilidad de la comunicación. Se puede determinar la fiabilidad por medio del segundo teorema de Shannon, conocido como el *teorema de codificación con ruido*. Dice que: siendo  $C$  la capacidad del canal sin memoria y  $R$  la tasa de transmisión; "para cualquier  $R < C$  se tienen un entero  $r$ , un código de longitud de bloque  $r$  y tasa  $R$  tales que la probabilidad de que se produzca un error de decodificación del bloque es menor o igual que  $\epsilon$ , para toda  $\epsilon > 0$ . De tal manera que se puede tener una probabilidad de error arbitrariamente pequeña siempre y cuando la tasa del mensaje codificado sea menor que la capacidad del canal"[Gon87].

### Teorema de codificación de fuentes o teoría de la tasa de distorsión.

En esta parte se tiene la teoría desarrollada cuando se presentan degradaciones a la información a pesar de tener un canal libre de errores. Para ello es necesario conocer la teoría de tasa de distorsión.

La tasa de distorsión se refiere a la mínima cantidad de bits que se requieren para codificar una fuente siempre y cuando la degradación se mantenga en un nivel aceptable. La teoría de la tasa de distorsión establece límites de desempeño para

---

<sup>5</sup>Shannon (1948)

compresión de datos cuando implica pérdidas acorde a un criterio de fidelidad. La función de la tasa de distorsión  $R(D)$  tiene las siguientes propiedades:

- Se puede encontrar un sistema de codificación con distorsión promedio  $D$  a una tasa arbitrariamente cercana a  $R(D)$ .
- No se puede encontrar un sistema de codificación que reconstruya la imagen con distorsión  $D$  a una tasa menor de  $R(D)$ .
- La función  $R(D)$  es convexa, continua y estrictamente decreciente.

## 2.5 Conclusiones

Se tiene en este capítulo un panorama general de las áreas que se encuentran involucradas en el desarrollo del proyecto. Se realizó una breve descripción del sistema visual humano, y se mencionaron las investigaciones efectuadas desde tiempos antiguos.

Además, se presentó una descripción del análisis y procesamiento de imágenes, ya que son estas, las áreas donde principalmente se desarrolla el trabajo. Al final se presentan conceptos de teoría de la información, con ellos se tiene de una manera concreta el alcance de los algoritmos que se proponen más adelante.

Los objetivos de este trabajo son aquellos relacionados con el uso de una técnica específica de análisis de imágenes para realizar codificación. Por esta razón, una presentación exhaustiva de las técnicas de compresión y codificación esta fuera del alcance de este proyecto. Sin embargo, al final del texto se propone una bibliografía correspondiente al tema. Los esquemas de codificación se trataran en los capitulos 4 y 5 que corresponden a la descripción y los resultados obtenidos.

# Capítulo 3

## Transformada Polinomial

Se han utilizado típicamente técnicas como la transformada de Fourier para realizar representación y análisis, aprovechando que la energía tiene una distribución desigual en el dominio de la frecuencia. Sin embargo, la evolución en las herramientas de procesamiento de imágenes ha permitido realizar análisis más complejo, ejemplo de ello son las técnicas de análisis adaptivo, como el análisis tiempo frecuencia [Cor96], la transformada ondícula [Mey94], la piramide gaussiana [Aka92] y la transformada polinomial [Mar90b].

En este capítulo se describe la transformada polinomial mencionando sus características, la caracterización de señales e imágenes por medio de la localización con filtros polinomiales y finalmente la técnica en el plano discreto.

### 3.1 Propiedades

La transformada polinomial tiene su origen en el procesamiento escala-espacio, que generalmente se utiliza para la detección de extremos (caso unidimensional) y en la representación de patrones visuales significativos de una imagen (caso bidimensional). Este procesamiento se realiza por medio de dos pasos:

1. La señal se multiplica por una función ventana cuya forma tamaño y posición se determinan para conocer la contribución de cada punto de la señal, el conjunto de puntos que contribuyen a un paso de procesamiento y el espaciado de la localización respectivamente. Para que la descripción sea completa se requiere que la ventana tome un cierto número de posiciones, se considera que el espacio entre ventanas es equidistante.
2. La selección del procesamiento implica una determinación a priori de lo que se busca. En este caso la parte de la señal que se encuentra localizada por la función ventana se aproxima por una suma de polinomios ponderados.

Las funciones ventana que se han utilizado han dado lugar a numerosas técnicas de procesamiento. Con el uso de una función ventana Gaussiana se tiene que la transformada polinomial recibe el nombre de "*Transformada de Hermite*", cuyo parámetro de escala es la desviación estándar, que indica el tamaño de la ventana. Un problema que surge es la escala o el tamaño óptimo para realizar el análisis, para cuestiones de codificación se utiliza una ventana grande y de esta manera es sensible a objetos grandes pero se pierde detalle; además, la implantación es más compleja. En el otro extremo se encuentra un tamaño de ventana pequeño que generalmente no se utiliza en casos de codificación. La solución al problema de escala es entonces, realizar análisis con una escala fija que involucre a una gran parte de las estructuras visualmente importantes que se extraerán y usar múltiples escalas de análisis tratando de abarcar un número aún mayor de estructuras contenidas en la imagen. De esta manera se tiene la teoría escala-espacio[Wit84], que es una aproximación del funcionamiento del sistema de visión humano.

Las principales propiedades de la "*transformada polinomial de Hermite*" se indican a continuación:

- Analiza la imagen por medio de funciones ventana traslapadas, proyectando la señal en bases obtenidas a partir de polinomios ortogonales.
- La ventana que se utiliza es una Gaussiana, se encontró que los filtros de análisis obtenidos a partir de la función ventana son equivalentes a derivadas de Gaussiana.
- En señales bidimensionales imita algunas características del sistema de visión humano, como la detección de bordes.

### 3.2 Transformada polinomial continua unidimensional.

La transformada polinomial se describe aquí de manera detallada para señales continuas unidimensionales; se presentan los conceptos básicos. El análisis de una señal por medio de esta técnica se muestra a continuación:

Tenemos que la señal original  $f(x)$  se analiza por medio de transformaciones polinomiales, y sus coeficientes polinomiales se obtienen con

$$f_n(kT) = \int_{-\infty}^{+\infty} f(x) \cdot D_n(kT - x) dx \quad (3.1)$$

y se recupera la señal de manera completa a partir de los coeficientes generados con

$$f(x) = \sum_{n=0}^{\infty} \sum_k f_n(kT) \cdot P_n(x - kT) \quad (3.2)$$

si los coeficientes de la expansión se multiplican por constantes fijas  $t_n$ , entonces para el caso de una transformada finita la señal de salida sera:

$$\hat{f}(x) = \sum_{n=0}^{\infty} \sum_k t_n \cdot f_n(kT) \cdot P_n(x - kT) \quad (3.3)$$

donde:

- $f(x)$  es la señal original
- $\hat{f}(x)$  es la señal aproximada de  $f(x)$
- $f_n(kT)$  es el coeficiente de orden  $n$ , en la posición  $k$  con periodo  $T$ .
- $D_n(x)$  es el filtro de análisis
- $P_n(x)$  es el filtro de síntesis
- $n$  es el orden de el coeficiente, su rango es de  $0$  a  $N$ .
- $T$  es el periodo de localización de la señal
- $t_n = 1$  para  $0 \leq n \leq N$  y  $t_n = 0$  para  $n > N$ . En el caso de una transformación finita

La ecuación de análisis de la señal que se presenta en la ecuación 3.1, indica un mapeo de la señal original por medio de los filtros  $D_n(x)$  para obtener los coeficientes  $L_n(kT)$ , que recibe el nombre de *transformada polinomial directa*. La señal completa se obtiene por medio de la ecuación 3.2, a partir de los filtros de síntesis  $P_n(x)$ , que recibe el nombre de *transformada polinomial inversa*. Cuando los coeficientes se multiplican por una constante  $t_n$ , como en la ecuación 3.3, la reconstrucción de la señal se ve afectada, esta constante se usa para determinar el conjunto finito de coeficientes que se utilizará en la reconstrucción.

### 3.2.1 Filtros de análisis y síntesis

Los filtros de análisis se generan a partir de la multiplicación de un polinomio con la función ventana, como se presenta a continuación

$$D_n(x) = G_n(-x)V^2(-x) \quad (3.4)$$

donde:  $G_n(x)$  es el polinomio de orden  $n$ ; y  $V(x)$  es la función ventana. Estos filtros determinan que información se va a hacer explícita en los coeficientes polinomiales. La señal se recupera por medio de la multiplicación del coeficiente de orden  $n$  con el filtro de síntesis, este se genera con

$$P_n(x) = \frac{G_n(x)V(x)}{W(x)} \quad (3.5)$$

donde:  $W(x)$  es la función de peso o ponderación. En la figura 3.1 se observa la transformación polinomial directa e inversa.

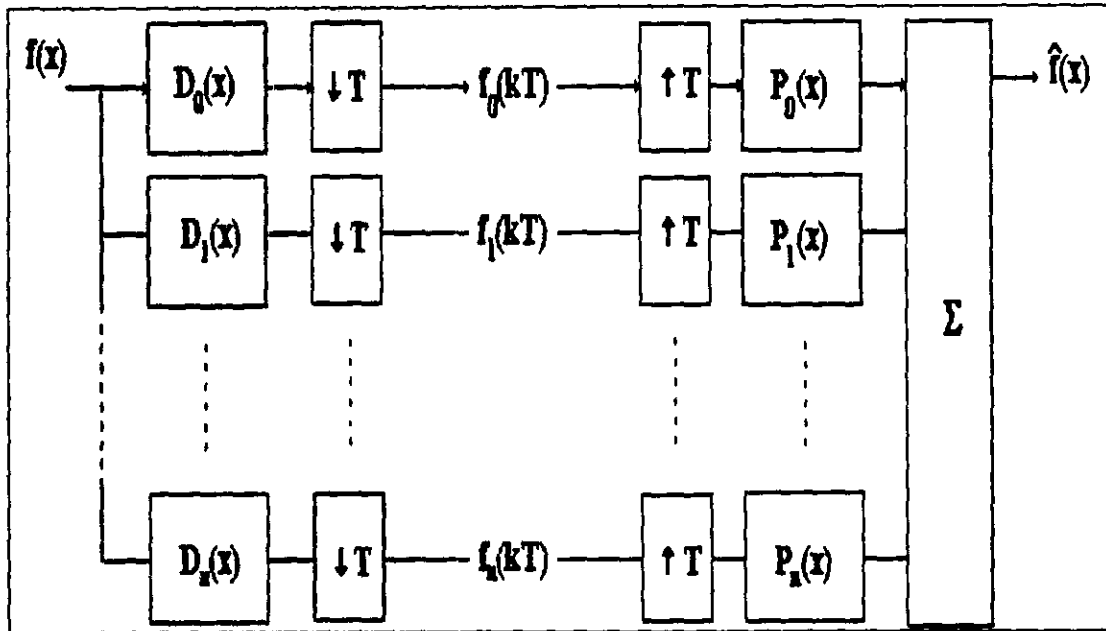


Figura 3.1: Transformada polinomial directa e inversa,  $f(x)$  es la señal original localizada por medio de los filtros de análisis  $D_n$  que submuestreada por un factor  $T$  da como resultado los coeficientes  $f_n$ . La señal reconstruida  $\hat{f}(x)$  se obtiene de la sumatoria de los coeficientes interpolados por el mismo factor  $T$  y ponderados por los filtros  $P_n$ .

### 3.2.2 Función de peso o ponderación

La función de peso se construye a partir de la función ventana por medio de repeticiones periódicas, esto hace que la función peso sea por si misma periódica con periodo  $T$ .

$$W(x) = \sum_k V(x - kT) \quad (3.6)$$



dado que la función peso es no cero para toda  $x$ , tenemos

$$f(x) = \frac{1}{W(x)} \sum_k f(x) \cdot V(x - kT) \quad (3.7)$$

así se garantiza que las señales localizadas  $f(x) \cdot V(x - kT)$  para todas las posiciones  $kT$  de la ventana contienen suficiente información de la señal  $f(x)$ .

### 3.2.3 ¿Cómo se obtienen los polinomios?

Los polinomios  $G_n(x)$  de orden  $n$  utilizados como funciones base para aproximar la parte de la señal que se encuentra dentro de la función ventana, se generan en concordancia con ella y se determinan únicamente por la función  $V^2(x)$ . Cumplen con una condición de ortogonalidad:

$$\int_{-\infty}^{\infty} V^2(x) G_m(x) G_n(x) dx = \delta_{mn}. \quad (3.8)$$

Los polinomios  $G_n(x)$  ortonormales a la función ventana  $V(x)$ , elegida arbitrariamente, se determinan por medio de

$$G_n(x) = \frac{1}{\sqrt{M_{n-1} M_n}} \begin{vmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} \quad (3.9)$$

donde  $M_n$  es el determinante que se define por

$$M_n = |c_{i+j}|_{i,j=0,\dots,n}, M_{-1} = 1 \quad (3.10)$$

y

$$c_n = \int_{-\infty}^{\infty} x^n V^2(x) dx \quad (3.11)$$

que es similar al momento de orden  $n$  -ésimo. Suponiendo que la función ventana  $V^2(x)$  sea par, los polinomios ortonormales que se generan hasta orden 3 son

$$\begin{aligned} G_0(x) &= \frac{1}{\sqrt{c_0}} \\ G_1(x) &= \frac{x}{\sqrt{c_2}} \\ G_2(x) &= \frac{(c_0 x^2 - c_2)}{\sqrt{c_0(c_0 c_4 - c_2^2)}} \\ G_3(x) &= \frac{(c_0 x^3 - c_4 x)}{\sqrt{c_2(c_2 c_6 - c_4^2)}} \end{aligned} \quad (3.12)$$

Bajo condiciones muy generales, para la señal original  $f(x)$ , se tiene que

$$V(x - kT) \left[ f(x) - \sum_{n=0}^{\infty} f_n(kT) \cdot G_n(x - kT) \right] = 0 \quad (3.13)$$

con la función de análisis de la ecuación 3.1. Para garantizar la convergencia de la expansión en series para muchas funciones ventana se requiere que  $f(x)$  sea analítica y finita para toda  $x$ .

El error de aproximación que se genera con la descripción de la señal dentro de la ventana de análisis se puede hacer arbitrariamente pequeño tomando el grado de la expansión suficientemente alto, como se sabe de la expansión de Taylor. De tal manera que la señal localizada  $f(x) \cdot V(x - kT)$  se puede describir con un error de aproximación pequeño por un conjunto finito de coeficientes  $f_n(kT)$ .

La energía de la señal dentro de la ventana se expresa en términos de los coeficientes de la expansión:

$$\int_{-\infty}^{+\infty} f^2(x) V^2(x - kT) dx = \sum_{n=0}^{\infty} f_n^2(kT) \quad (3.14)$$

y esta es la generalización del teorema de parseval para polinomios ortonormales.

### 3.2.4 ¿Qué función ventana utilizar?

La parte más importante de la *transformada polinomial* es la elección de la función ventana local que analiza la señal, generalmente se trata de un tipo de función que cumpla con ciertas características deseables, en este caso se utiliza la función Gaussiana.

La función Gaussiana es útil para la suavización de una señal y sus derivadas se usan para la detección de extremos (máximos y mínimos en una señal), ha dado lugar al desarrollo de la teoría escala-espacio [Wit84]; los polinomios asociados a una Gaussiana reciben el nombre de "*polinomios de Hermite*" la técnica de análisis derivada de ellos se conoce como "*transformada polinomial de Hermite*". Este tipo de función se elige debido a las propiedades que se mencionan a continuación.

- Las propiedades de la función Gaussiana se pueden derivar y evaluar de una manera sencilla ya que el tratamiento de la teoría matemática correspondiente es más fácil.
- Su factor de normalización es tal que  $V^2(x)$  tiene energía unitaria. La expresión matemática de la función ventana Gaussiana local es la siguiente

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (3.15)$$

con desviación estándar  $\sigma$ .

- Las ventanas gaussianas separables en 2D y con desviación estándar  $\sigma$  son un buen modelo de los campos receptivos traslapados en la retina encontrados en experimentos fisiológicos.
- La transformada polinomial Hermite genera funciones filtro que son derivadas de Gaussiana, este tipo de funciones tiene un amplio uso en procesamiento de imágenes y en el modelado del sistema de visión humano.
- La ventana gaussiana minimiza el producto de incertidumbres en el dominio del espacio y la frecuencia, propiedad interesante en el análisis de imágenes.

La “transformada polinomial de Hermite” solo se limita a unos pocos términos en la práctica, debido a que los filtros de ordenes altos dan muy poca información adicional. A continuación se derivan las expresiones para las diferentes funciones involucradas en la “transformada polinomial de Hermite”.

### Función peso:

Como ya se mencionó la función peso  $W(x)$  es periódica con periodo  $T$ ,

$$W(x) = \frac{\sqrt{2\sqrt{\pi}\sigma}}{T} w(x) \quad (3.16)$$

se puede expandir en series de Fourier [Mar90b] con

$$w(x) = 1 + 2 \sum_{k=1}^{\infty} \exp \left[ -\frac{1}{2} \left( k \frac{2\pi\sigma}{T} \right)^2 \right] \cdot \cos \left( k \frac{2\pi x}{T} \right) \quad (3.17)$$

Como generalmente lo deseable es que se reduzca el número de descomposiciones locales, para casos de codificación se puede hacer el parámetro de muestreo  $\tau = T/\sigma$  tan grande como sea posible <sup>1</sup>, de otra manera  $W(x)$  es aproximadamente constante. De lo anterior se tiene que si

$$f(x) = \frac{1}{W(x)} \sum_k f(x) \cdot V(x - kT) \quad (3.18)$$

y  $\tau \leq 2$ , entonces  $W(x)$  se aproxima a una constante y el error de aproximación disminuye.

---

<sup>1</sup>El límite superior de  $\tau$ , se conoce a partir del tamaño de la función, ya que se tiene que dar peso a la mayoría de los puntos de la señal, para que la reconstrucción no presente degradaciones considerables.

## Función de análisis

La *transformada de Hermite* tiene como principal descriptor de sus propiedades a las funciones de análisis que a partir de la ecuación 3.4 resultan en

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) e^{-x^2/\sigma^2} \quad (3.19)$$

pues los *polinomios de Hermite*  $\{H_n(x/\sigma); n = 0, 1, \dots\}$  son ortogonales sobre la ventana Gaussiana  $V^2(x)$ . De lo anterior se ha demostrado [Abr65] que las funciones de análisis son iguales a derivadas de orden  $n$  -ésimo de una Gaussiana, esto es

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{d^n}{d\left(\frac{x}{\sigma}\right)^n} \left[ \frac{1}{\sigma \sqrt{\pi}} e^{-x^2/\sigma^2} \right] \quad (3.20)$$

y cuya transformada de Fourier es

$$d_n(\omega) = \frac{1}{\sqrt{2^n n!}} \cdot (j\omega\sigma)^n e^{-(\omega\sigma)^2/4} \quad (3.21)$$

El filtro tiene un valor extremo para  $(\omega\sigma)^2 = 2n$ , de ahí que el orden de los filtros este directamente relacionado con la frecuencia de la señal de análisis pues conforme aumenta el orden de la transformación, la información adicional que generan los filtros es escasa, motivo por el cual la *transformada polinomial de Hermite* se limita a unos pocos términos, esto es más notorio en el plano discreto.

**Funciones de síntesis** Las funciones  $P_n(x)$  que se requieren para recuperar la señal a partir de los coeficientes de la transformada polinomial Hermite son

$$P_n(x) = \frac{T}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{2\pi}} H_n\left(\frac{x}{\sigma}\right) e^{-x^2/2\sigma^2} / w(x) \quad (3.22)$$

donde  $w(x)$  es la función peso de la ecuación 3.17. Si  $w(x) = 1$  como ya se había mencionado antes para valores  $\tau < 2$ , entonces  $P_n(x)$  es igual a la *función de Hermite* de orden  $n$ , que tiene la propiedad de ser isomórfica a su transformada de Fourier,

$$p_n(\omega) = \frac{T}{\sqrt{2^n n!}} \cdot (-j)^n H_n(\omega\sigma) e^{-(\omega\sigma)^2/2} \quad (3.23)$$

### 3.3 Transformada polinomial para señales bidimensionales

De lo anterior se puede generalizar fácilmente la *transformada polinomial* a dos dimensiones<sup>2</sup>. La ventana de análisis local es  $V(x, y)$ . Los polinomios ortonormales asociados  $G_{m,n-m}(x, y)$ , donde  $m$  y  $n - m$  son los grados con respecto a  $x$  y  $y$  respectivamente, se determinan únicamente por la condición de ortogonalidad:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} V^2(x, y) G_{m,n-m}(x, y) G_{j,i-j}(x, y) dx dy = \delta_{ni} \delta_{mj} \quad (3.24)$$

para  $n = 0, 1, \dots, \infty$ ;  $m = 0, \dots, n$ ;  $i = 0, 1, \dots, \infty$  y  $j = 0, \dots, i$ .

Los coeficientes polinomiales se obtienen convolucionando la señal original con las funciones filtro

$$f_{m,n-m}(p, q) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) D_{m,n-m}(p - x, q - y) dx dy \quad (3.25)$$

donde  $(p, q)$  cubren todas las coordenadas de una rejilla de muestreo bidimensional  $S$ . La única condición para la rejilla de muestreo es que la función peso

$$W(x, y) = \sum_{(p,q) \in S} V(x - p, y - q) \quad (3.26)$$

sea diferente de cero para todas las coordenadas  $(x, y)$ . Las funciones filtro se derivan de la siguiente expresión

$$D_{m,n-m}(x, y) = G_{m,n-m}(-x, -y) V^2(-x, -y) \quad (3.27)$$

La señal se recupera con

$$f(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{(p,q) \in S} f_{m,n-m}(p, q) \cdot P_{m,n-m}(x - p, y - q) \quad (3.28)$$

donde las funciones de síntesis  $P_{m,n-m}(x, y)$  están descritas por

$$P_{m,n-m}(x, y) = G_{m,n-m}(x, y) V(x, y) / W(x, y) \quad (3.29)$$

para  $n = 0, 1, \dots, \infty$  y  $m = 0, \dots, n$ .

Si tenemos que la función ventana es separable  $V(x, y) = V(x) \cdot V(y)$  y la rejilla de muestreo es cuadrada, entonces surge un caso especial de la transformada polinomial, donde las funciones de análisis y síntesis son también separables y los coeficientes se pueden obtener por medio de la convolución de la señal con los filtros  $D_m(x) D_{n-m}(y)$  que son funciones filtro de  $V(x)$  y  $V(y)$  respectivamente.

<sup>2</sup>Se considera que las imágenes son señales bidimensionales.

### 3.3.1 Transformada polinomial de Hermite bidimensional

Si la función ventana es una función Gaussiana, entonces se obtiene la *transformada polinomial de Hermite bidimensional*; la función Gaussiana es espacialmente separable y rotacionalmente simétrica, estas propiedades son heredadas a las funciones filtro. La ventaja de tener filtros separables es la capacidad de detectar orientaciones unidimensionales independientemente de su estructura interna.

## 3.4 Transformada polinomial discreta

La implantación de la *transformada polinomial* se lleva a cabo dentro del caso discreto. Una descripción discreta de la técnica se puede lograr a partir de dos enfoques que se mencionan a continuación:

- El primer enfoque consiste de realizar aproximaciones de la señal continua  $f(x)$  en su versión discreta produciendo un número determinado de coeficientes  $f_q$  a partir de la interpolación con  $I(x)$ , como se observa

$$f(x) = \sum_q f_q \cdot I(x - q\Delta) \quad (3.30)$$

si aplicamos una *transformada polinomial directa* a esta señal, se obtienen los coeficientes

$$f_n(pT) = \sum_q f_q [D_n(x) * I(x)]_{x=pT-q\Delta} \quad (3.31)$$

- De otra manera se puede definir la *transformada polinomial* directamente en señales discretas sin necesidad realizar la aproximación a señales continuas. Las expresiones para todas las funciones se realizan con la sustitución de las integrales por sumatorias, por ejemplo,

$$c_n = \sum_x x^n V^2(x) \quad (3.32)$$

es el momento de orden  $n$  -ésimo para  $n = 0, \dots, N$ . Para propósitos de implantación es útil tomar una ventana finita de tal manera que  $V(x) = 0$  para  $x < N_1$  y  $x > N_2$ . Sin embargo, se puede realizar una reconstrucción perfecta para cualquier señal discreta pues dentro de la ventana  $V(x)$  se tienen  $N + 1$  grados de libertad.

### 3.4.1 Transformada polinomial de Hermite discreta.

Cuando se elige una función binomial como función ventana, la cual es una aproximación discreta de la función Gaussiana, entonces surge la *transformada polinomial de Hermite discreta*, también llamada *Transformada binomial*, la expresión que define a  $V(x)$  es

$$V^2(x) = \frac{1}{2^M} C_M^x \quad (3.33)$$

para  $x = 0, \dots, M$ .  $M$  es el tamaño de la ventana binomial. Los polinomios ortonormales que se asocian a esta función ventana son los "*polinomios de Krawtchouk*"

$$G_n(x) = \frac{1}{\sqrt{C_M^n}} \sum_{k=0}^n (-1)^{n-k} C_{M-x}^{n-k} \cdot C_x^k \quad (3.34)$$

para  $x, n = 0, \dots, M$ . Para valores de  $M$  muy grandes la ventana binomial tiende a una Gaussiana y los polinomios de Krawtchouk a polinomios Hermite. La *transformada polinomial de Hermite discreta* de longitud  $M$  se aproxima a la analógica de  $\sigma = \sqrt{M/2}$ . Sus propiedades se pueden deducir a partir de la *transformada polinomial de Hermite continua*.

## 3.5 Conclusiones

Se describió en este capítulo la técnica de análisis de imágenes conocida como la *transformada polinomial*. Se presenta el análisis a señales unidimensionales, y posteriormente se hace la generalización a señales bidimensionales, dicho de otra forma son imágenes.

Para la descripción de la transformada polinomial se utilizan las características de una función ventana Gaussiana, sin embargo, la implantación computacional se efectúa con una función ventana binomial.

Lo visto en este capítulo solo se circunscribe al análisis que se efectúa en una señal a una escala. La transformada polinomial para realizar una mejor localización de estructuras importantes necesita variar el valor de escala, esto la hace una técnica adaptiva. La variación de escalas se menciona más adelante en la propuesta de un algoritmo para resolver el problema de codificación.

## Capítulo 4

# Codificación de imágenes a partir de la representación polinomial

Por lo antes mencionado, se sabe que el análisis de imágenes por medio de la transformada polinomial modela algunas propiedades del sistema visual humano; análisis que además, es útil para cuestiones de realce, restauración y reducción de ruido [Esc92b]. Debido a la generación de información redundante, la transformada polinomial presenta problemas para la codificación y compresión de imágenes, a pesar de ello se han propuesto algoritmos de codificación en esquema piramidal predictivo obteniendo buenos resultados [Esc95] [Kay94]. Se logra eliminar la redundancia en la expansión bidimensional de la transformada polinomial siguiendo los tres pasos que se mencionan a continuación: (1) La proyección de los coeficientes bidimensionales en un espacio unidimensional, asumiendo que existe una dirección óptima para cada posición de la ventana. (2) La búsqueda de estructuras primarias -en este caso bordes- y su caracterización en parámetros. (3) Utilización de un esquema a múltiples resoluciones, con el cual se espera una reducción importante de datos. Se proponen algoritmos de codificación usando los pasos antes mencionados.

### 4.1 Representación de imágenes con coeficientes polinomiales unidimensionales

Los filtros de análisis y síntesis son diferentes debido a la condición de ortogonalidad de las funciones polinomiales, por esta razón la información obtenida es redundante. La redundancia de tipo estadístico que se genera puede reducirse de manera notable si realizamos una proyección de la descomposición bidimensional en todas las orientaciones posibles, tratando de encontrar aquella donde el contraste sea máximo y conservando los valores de orientación y magnitud del coeficiente en esa dirección; de esa manera, se logra una representación donde se tienen algunas de las estructuras



más importantes para la percepción como son los bordes y líneas [Lop96]. Esta es una representación en coeficientes polinomiales unidimensionales y se obtiene como se indica a continuación.

Sea  $f(x, y)$  la imagen original y  $f_{m,n-m}(p, q)$  el coeficiente de orden  $(m, n-m)$  que se obtiene a partir de la ecuación 3.25. Sea  $E^2$  un criterio de error cuadrado ponderado que se desea minimizar

$$E^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [K(x\cos\theta + y\sin\theta) - f(x, y)]^2 V^2(x, y) dx dy \quad (4.1)$$

para todos los coeficientes unidimensionales  $K$  y ángulos  $\theta$ . Donde  $V(x, y)$  es la ventana gaussiana bidimensional. Si ahora proyectamos la función ventana bidimensional  $V^2(x, y)$  sobre un eje que hace un ángulo  $\theta$  con el eje  $x$  se tiene

$$V_\theta^2(u) = \int_{-\infty}^{+\infty} V^2(u\cos\theta - v\sin\theta, u\sin\theta + v\cos\theta) dv. \quad (4.2)$$

Si esta ventana es rotacionalmente simétrica, entonces es independiente de la orientación. En este caso se utiliza una función gaussiana que cumple con esa propiedad. Podemos realizar la expansión de los coeficientes unidimensionales  $K(u)$  en las bases  $\{F_{n,\theta}(u); n = 0, 1, \dots\}$  de polinomios ortonormales sobre  $V(u)$ , esto es

$$V_\theta(u) \left[ K(u) - \sum_{n=0}^{\infty} K_{n,\theta} \cdot F_{n,\theta}(u) \right] = 0. \quad (4.3)$$

Sustituyendo la expansión polinomial en  $f(x, y)$  y la expansión unidimensional en  $K(u)$  en la ecuación 4.1, además de tomar la derivada parcial con respecto a  $K_{n,\theta}$  se obtiene

$$K_{n,\theta} = \sum_{k=0}^n \sum_{l=0}^k f_{l,k-l} h_{n,\theta}(l, k-l) \quad (4.4)$$

que es el coeficiente unidimensional  $K_n$  en la dirección  $\theta$ . Donde la función ángulo que proyecta los coeficientes en el ángulo  $\theta$ , se expresa por medio de

$$h_{n,\theta}(l, k-l) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_{n,\theta}(x\cos\theta + y\sin\theta) \cdot G_{l,k-l}(x, y) V^2(x, y) dx dy \quad (4.5)$$

Los polinomios ortonormales  $F_{n,\theta}(u)$  y la función ángulo se pueden calcular conociendo los momentos necesarios para especificar totalmente los polinomios

$$c_{n,\theta} = \int_{-\infty}^{+\infty} u^n V_{\theta}^2(u) du \quad (4.6)$$

Estos momentos se pueden encontrar utilizando la ventana bidimensional  $V(x, y)$  como se muestra

$$c_{n,\theta} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \cos \theta + y \sin \theta)^n V^2(x, y) dx dy \quad (4.7)$$

#### 4.1.1 Propiedades de la función ángulo

La función ángulo tiene las siguientes propiedades derivadas de la ortogonalidad de los polinomios  $F_{n,\theta}(u)$

1. Acotación de los coeficientes bidimensionales que se toman para realizar la proyección en un ángulo  $\theta$

$$h_{n,\theta}(l, k-l) = 0 \quad \text{si } k > n, \quad (4.8)$$

2. Ortogonalidad de las funciones ángulo para diferente orden y un ángulo  $\theta$

$$\sum_{k=0}^{\infty} \sum_{l=0}^k h_{m,\theta}(l, k-l) h_{n,\theta}(l, k-l) = \delta_{mn} \quad (4.9)$$

#### 4.1.2 Expansión unidimensional en coeficientes óptimos

Una vez realizada la proyección en un número suficiente de direcciones se tiene una expansión unidimensional aún más numerosa que la bidimensional. Sin embargo, lo que se desea es reducir la cantidad de datos, razón por la cual se busca aquella dirección donde el contraste sea máximo. Inicialmente se tiene el siguiente error

$$E^2 = \sum_{k=0}^{\infty} \sum_{l=0}^k f_{l,k-l}^2 - \left[ \sum_{n=0}^{\infty} K_{n,\theta}^2 \quad \text{para toda } \theta \right] \quad (4.10)$$

que se puede minimizar sobre el ángulo  $\theta$  maximizando la energía direccional

$$\sum_{n=0}^{\infty} K_{n,\theta}^2. \quad (4.11)$$

Debido a que los coeficientes unidimensionales  $K$  están acotados, encontramos la dirección óptima maximizando la energía direccional  $\sum_{k=1}^N K_{k,\theta}^2$ , o el coeficiente de contraste

$$C_\theta = \log \left[ 1 + \sum_{k=1}^N K_{k,\theta}^2 \right] \quad (4.12)$$

sobre todos los coeficientes unidimensionales  $K$  y los ángulos  $\theta$ . De tal manera que verificando el valor del coeficiente de contraste  $C_\theta$ , se puede determinar cual es el ángulo óptimo  $\theta$  y el valor del coeficiente unidimensional  $K_n$ . Después de la elección de los valores de  $K_n$  y las orientaciones  $\theta$  se tiene una representación unidimensional óptima.

### 4.1.3 Reconstrucción de la expansión bidimensional

La reconstrucción de la expansión bidimensional se logra a través de la representación unidimensional orientada óptimamente

$$f_{l,k-l} = \sum_{n=k}^{\infty} K_{n,\theta} \cdot h_{n,\theta}(l, k-l) \quad (4.13)$$

para el caso  $n > N$ , los coeficientes unidimensionales  $K_{n,\theta}$  son desconocidos y se asumen como igual a cero. Esto indica que la expansión bidimensional se reconstruirá como una matriz triangular superior de coeficientes bidimensionales implicando la pérdida de información. Sin embargo, se considera que no es relevante para el sistema de visión humano.

## 4.2 Representación de imágenes con parámetros de bordes

Se ha demostrado que los cambios de intensidad en una imagen son una de las estructuras más importantes para la percepción humana. Se les llama bordes, contornos o fronteras de regiones; su identificación es sumamente útil para realizar procesamiento. En esta sección se buscan los bordes con la finalidad de obtener parámetros, con los cuales sea posible reconstruir un borde, llamados "*parámetros característicos de un borde*". Los valores de amplitud del borde, su orientación en

FALTA PAGINA

No. **47**

donde  $f(x)$  es la función borde gaussiano unidimensional,  $A_b$  es el valor medio del borde y  $x_b$  es la posición central del borde en el eje  $x$ . El borde ideal es la función escalón que se define por

$$f(x) = \begin{cases} A_b - \frac{\Delta f}{2}, & x < x_b \\ A_b + \frac{\Delta f}{2}, & x \geq x_b \end{cases} \quad (4.15)$$

donde  $\Delta f$  es la magnitud del borde. Ambas funciones se observan en la gráfica de la figura 4.1. Una función escalón bidimensional en los ejes  $x, y$  se puede observar en la gráfica de la figura 4.2.

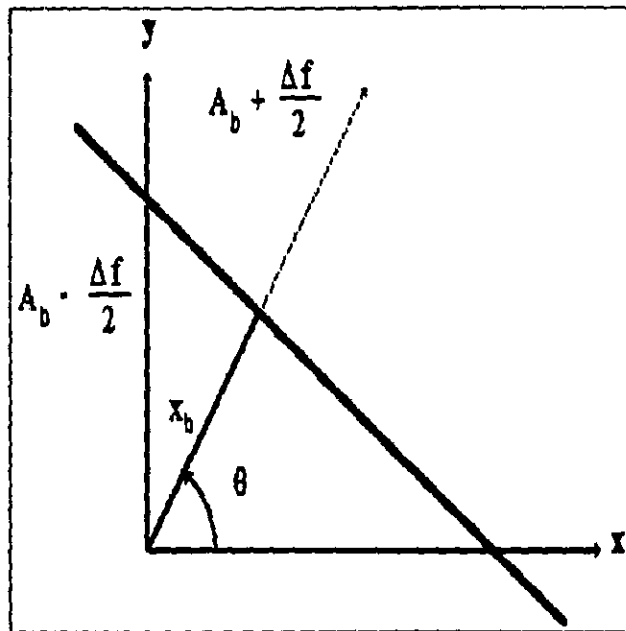


Figura 4.2: Diagrama de un borde con sus parámetros en el plano  $(x, y)$

Se sabe que las derivadas de primero y segundo orden de una función Gaussiana, son detectores eficientes de extremos (máximos y mínimos) y cruces por cero respectivamente [Mar82] [Hen93] [Lim90]. También se ha mencionado que los filtros de análisis de orden  $n$  de la transformada polinomial de Hermite corresponden a derivadas de Gaussiana de orden  $n$  [Abr65]. De lo anterior se deduce entonces que en los coeficientes polinomiales bidimensionales de primero y segundo orden, se encuentra contenida gran parte de la información que corresponde a estructuras primarias -en este caso bordes-. Por lo tanto se pueden modelar dichas estructuras si conocemos los coeficientes mencionados. Para un borde dado la respuesta a sus operadores diferenciales es

$$f_0(x) = A_b + \Delta f \cdot E_0(x_b) \quad (4.16)$$

y

$$f_n(x) = \Delta f \cdot E_n(x_b) \quad (4.17)$$

donde

$$E_n(x_b) = \frac{1}{2} \left[ \int_{x_b}^{+\infty} G_n(x) V^2(x) dx - \int_{-\infty}^{x_b} G_n(x) V^2(x) dx \right] \quad (4.18)$$

para  $n = 0, \dots, \infty$ , donde  $G_n(x)$  es el polinomio ortonormal de grado  $n$  sobre la ventana  $V^2(x)$ . La respuesta de un borde ideal y uno gaussiano a los operadores diferenciales de primero y segundo orden se puede observar en la gráfica de la figura 4.3.

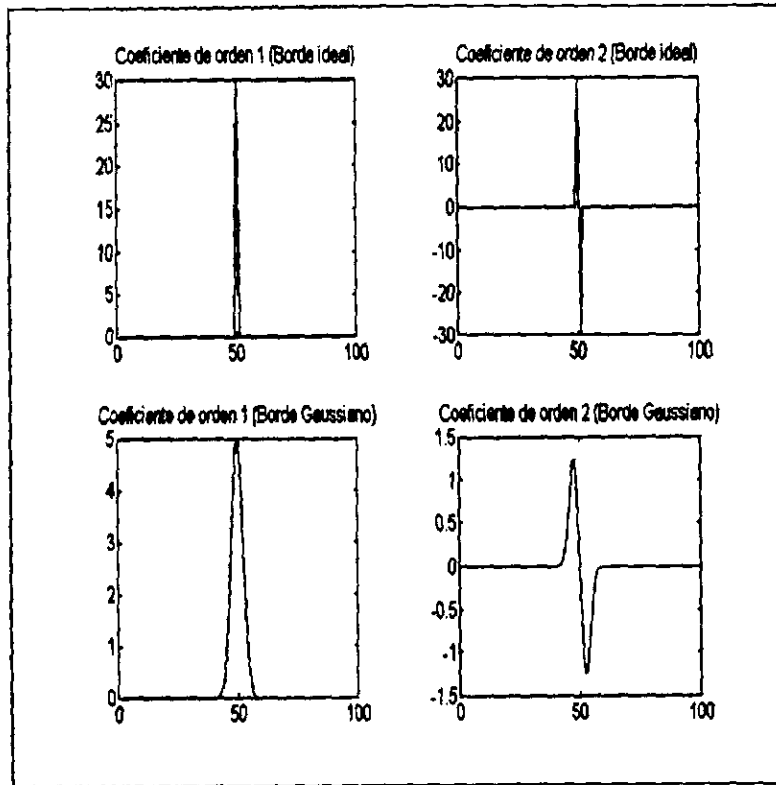


Figura 4.3: Respuesta de bordes Gaussiano e ideal a los operadores diferenciales de primero y segundo orden. Arriba a la izquierda se presenta la respuesta del borde ideal al operador diferencial de primer orden, a la derecha se tiene la respuesta al operador de segundo orden. Abajo se tienen las respuestas del borde gaussiano al operador diferencial de primero y segundo orden (izquierda y derecha respectivamente).

La curva de respuesta de un borde cuando se hace la gráfica del coeficiente de primer orden contra el coeficiente de segundo orden se observa en la figura 4.4. Los coeficientes Hermite para una ventana gaussiana de varianza  $\sigma$  centrada en el origen satisface la siguiente expresión

$$f_n = \frac{\Delta f \cdot \sigma}{\sqrt{2n}} D_{n-1}(-x_b, \sigma) \quad (4.19)$$

para  $n = 1, \dots, \infty$ , donde  $D_n$  es la derivada de orden  $n$  -ésimo de una Gaussiana, como se definió en el capítulo anterior.

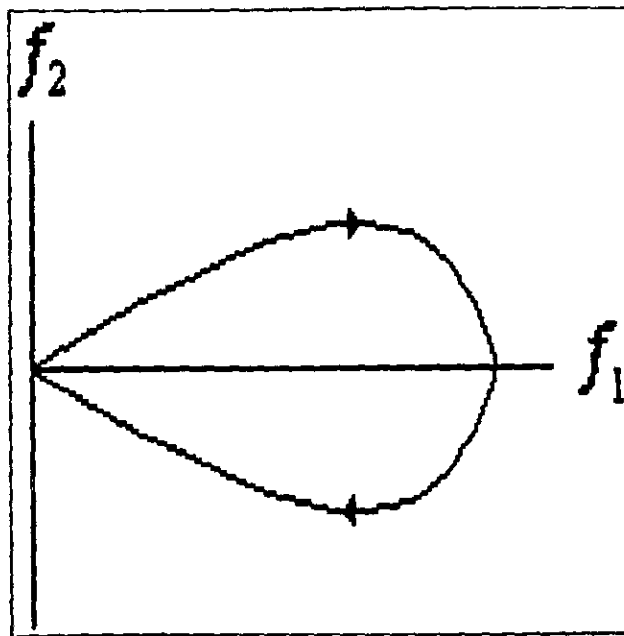


Figura 4.4: Curva típica de respuesta a los operadores diferenciales de un borde, en el plano  $(f_1, f_2)$ . Siguiendo los valores que toman los operadores diferenciales en la figura 4.3 es posible imaginar una gráfica de esta forma.

#### 4.2.2 Extracción de parámetros característicos de un borde.

La búsqueda de estructuras primitivas en una imagen tiene varios objetivos, entre otros delimitar regiones, realzar imágenes, restaurarlas, etcétera; en este caso se

trata de representar una imagen con parámetros de los bordes que están contenidos en ella. Tales parámetros son :

- Amplitud ( $\Delta f$ )
- Posición ( $x_b$ )
- Orientación ( $\theta$ )
- Valor medio ( $A_b$ )

La orientación es un parámetro que se obtiene de la representación unidimensional e indica la dirección en la que el máximo contraste se encuentra. Los coeficientes de primer y segundo orden de la descomposición polinomial nos ayudan a determinar los valores de posición y amplitud.

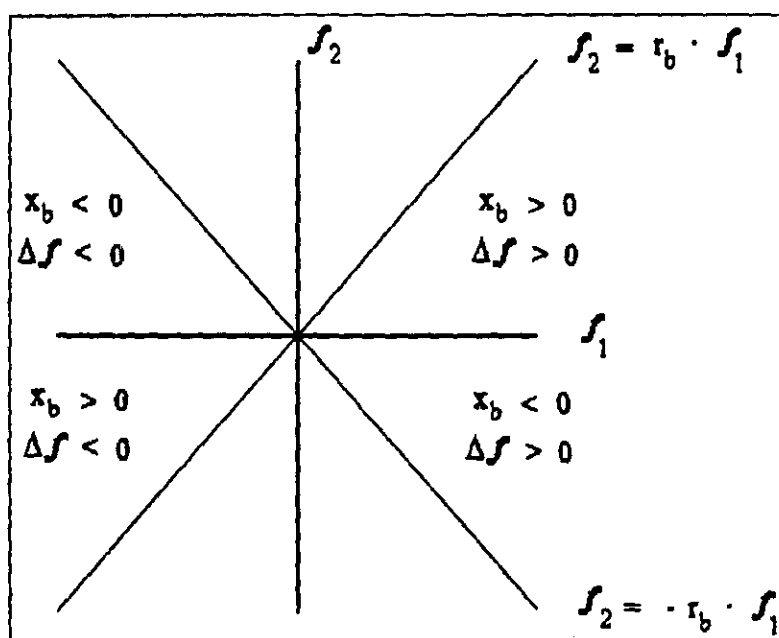


Figura 4.5: Partición del plano ( $f_1, f_2$ ). Umbrales para posición  $x_b$

De la definición de  $f_n$  se desprende la estimación de la posición del borde que se observa en la figura 4.5.

$$\frac{x_b}{\sigma} = \frac{f_2}{f_1} \quad (4.20)$$

Los valores de amplitud ( $\Delta f$ ) se estiman a partir de  $f_1$  en la ecuación 4.19. En la figura 4.5 se puede apreciar la partición del plano ( $f_1, f_2$ ) donde dado un valor



de  $x_b$  se determina si un pixel corresponde a un borde contenido en la ventana. Si se excede el umbral, entonces el pixel pertenece a una estructura no contenida completamente en la ventana que puede ser una esquina.

Para el caso discreto se tiene que

$$\frac{x_b}{\sqrt{M/2}} = \sqrt{1 - \frac{1}{M}} \cdot \frac{f_2}{f_1} \quad (4.21)$$

aquí  $x_b$  asume valores discretos  $-(M-1)/2, \dots, (M-1)/2$ . De tal manera que se puede lograr cuantización alta. El parámetro de magnitud  $\Delta f$  se obtiene de la misma manera que para el caso continuo a partir del coeficiente de primer orden.

### 4.3 Algoritmos de codificación propuestos

Se ha visto hasta aquí el análisis de imágenes con la transformada polinomial y la eliminación de la redundancia que trae implícita. Esta sección corresponde a los algoritmos que se proponen para efectos de codificación, tomando en cuenta que ya se ha logrado una reducción importante en los datos a través de las diferentes representaciones de la imagen. A manera de resumen, es conveniente recordar que las representaciones que se han repasado son: *bidimensional*, *unidimensional* y en *parámetros característicos*. La representación bidimensional se utiliza para cuestiones de comparación. Los esquemas que se proponen son básicamente dos: *el procesamiento en una escala y en múltiples escalas*. Se presentan los diagramas correspondientes a cada esquema.

#### 4.3.1 Codificación en una escala

En la figura 4.6, se presenta el diagrama que corresponde a la codificación realizada en una escala. Se puede observar que es básicamente un cambio de representación lo que conforma al algoritmo de codificación con escala fija.

De manera general el funcionamiento es el siguiente: se obtienen los coeficientes bidimensionales (2D) con una ventana Binomial  $V(x)$  cuyo tamaño  $M$  determina su varianza. A continuación y como se indica en la ecuación 4.4 se realiza la proyección en coeficientes unidimensionales (1D) y de acuerdo con la ecuación 4.12, se buscan las orientaciones óptimas; es decir, aquellas donde el contraste sea mayor para un punto dado. Después se realiza la detección de bordes y la extracción de los parámetros característicos. Finalmente la representación se submuestra para reducir datos y la representación de la imagen sea eficiente, la representación que se obtiene se compone de los siguientes elementos:

- Promedio local, que son los valores promedio de los bordes encontrados en cada imagen. Aquí se guarda la mayor parte de la información.

- Orientaciones. Para cada punto de la imagen se encuentra la orientación óptima de la estructura a la que pertenece.
- Posición del borde, con respecto al origen de la ventana.
- Magnitud del borde, al que pertenece el punto que se esta analizando.

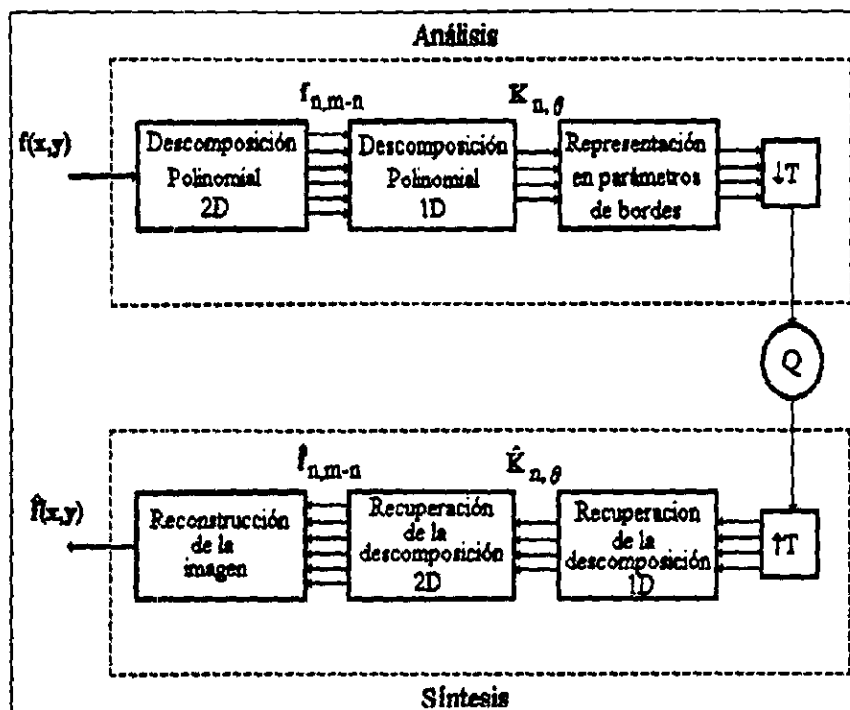


Figura 4.6: Esquema de codificación en una escala.

A partir del esquema de la figura 4.6 se obtienen los tres tipos de representaciones, en cada una de ellas se realiza la variación del tamaño de la ventana y el submuestreo. El tamaño de la ventana se hace suficientemente grande para garantizar un submuestreo alto sin gran pérdida de información. El submuestreo se ve limitado por la calidad de la reconstrucción de la imagen, de tal forma que se establece un límite en el cual la función peso de reconstrucción debe ser aproximadamente una constante.

### Representación bidimensional

Se sabe que el conjunto de coeficientes que se obtiene después de realizar las transformaciones polinomiales a la imagen es la *representación bidimensional*. Esta representación esta formada por aquellos coeficientes  $f_{m,n-m}$  donde  $n = 0, 1, \dots, N$ ,  $m = 0, 1, \dots, n$  y  $N$  es el máximo orden de la expansión. De esta manera se obtiene una matriz triangular superior de coeficientes<sup>1</sup>. Sin embargo, también es posible

<sup>1</sup>Suponiendo que en una matriz de coeficientes polinomiales  $m$  sean los renglones y  $n - m$  las columnas

obtener una *representación bidimensional* cuya matriz de coeficientes sea completa, ya que los filtros son separables y el orden máximo del filtro es  $N$ , por lo tanto el coeficiente de mayor orden en esta expansión completa es  $f_{N,N}$ . La representación bidimensional cuya matriz de coeficientes es completa contiene algunos elementos que no son perceptualmente importantes y son aquellos cuyo orden sea mayor a  $N$ . En cuanto a cuestiones de codificación, esta representación solo logra cierta eficiencia cuando se realiza submuestreo. Esto se debe a la naturaleza de la transformación, pues se trata de una expansión, lo que implica un aumento en la cantidad de datos y generación de redundancia. La *representación bidimensional* se utiliza como comparación pues se ha probado que tiene una buena calidad de reconstrucción [Esc92a] [Esc95] [Mar90b].

### Representación unidimensional

De la proyección en direcciones óptimas, surge la *representación unidimensional*, que tiene la característica de eliminar parte de la redundancia estadística existente en la representación bidimensional. En esta representación, se tienen coeficientes polinomiales por cada orden de la expansión y una imagen de orientaciones óptimas. Los coeficientes  $K_n$  donde  $n = 0, 1, \dots, N$ , contienen ahora la mayor parte de la información. Se ha observado que estos coeficientes representan también, una compactación de la energía de la imagen. Aquí las estructuras unidimensionales de la imagen, principalmente bordes, se pueden identificar fácilmente.

Únicamente con el cambio de representación logrado a partir de las proyecciones unidimensionales, se puede decir que se logra una cierta compresión de los datos de la imagen. Las imágenes que se pueden cuantizar son el coeficiente de orden cero y la imagen de orientaciones.

### Representación en parámetros característicos

A partir de la identificación de estructuras unidimensionales, estas se pueden modelar por medio de sus parámetros. La representación en parámetros característicos, consiste de la imagen de baja frecuencia, las orientaciones, magnitud y posición de un borde.

Entre las propiedades de los parámetros esta la cantidad de muestreo que se puede efectuar en ellos, que es como máximo hasta  $N$ , donde  $N$  es el tamaño de la ventana. El potencial de cuantización también es significativo, ya que la imagen de posiciones por ejemplo, toma valores cuyo rango es  $-\frac{M-1}{2}, \dots, \frac{M-1}{2}$ . Por lo tanto si la ventana es de tamaño  $M = 2$ , el rango es  $-0.5, \dots, 0.5$ . El mayor problema para realizar la cuantización de las imágenes es aquella que corresponde a las magnitudes de los bordes contenidos en la imagen, ya que su rango es muy amplio y su alta cuantización conlleva degradaciones considerables en la imagen.

### 4.3.2 Codificación en múltiples escalas

La transformada polinomial se engloba dentro de las técnicas de análisis escala-espacio, por lo cual, el algoritmo de codificación es más eficiente si se toma en cuenta este principio. El análisis a múltiples escalas presenta la ventaja de dar más detalles de la imagen, ya que se parte de que la imagen esta formada por estructuras cuya varianza es diferente. Por ejemplo, con una varianza grande, la función ventana detecta estructuras muy grandes, entonces los parámetros que se obtienen solo se refieren a esas estructuras. Sin embargo, si se localiza una imagen con una ventana cuya varianza sea muy pequeña, las altas frecuencias se detectan muy bien, pero necesita de un mayor número de localizaciones para conseguir una buena reconstrucción. El algoritmo que se propone se observa en la figura 4.7.

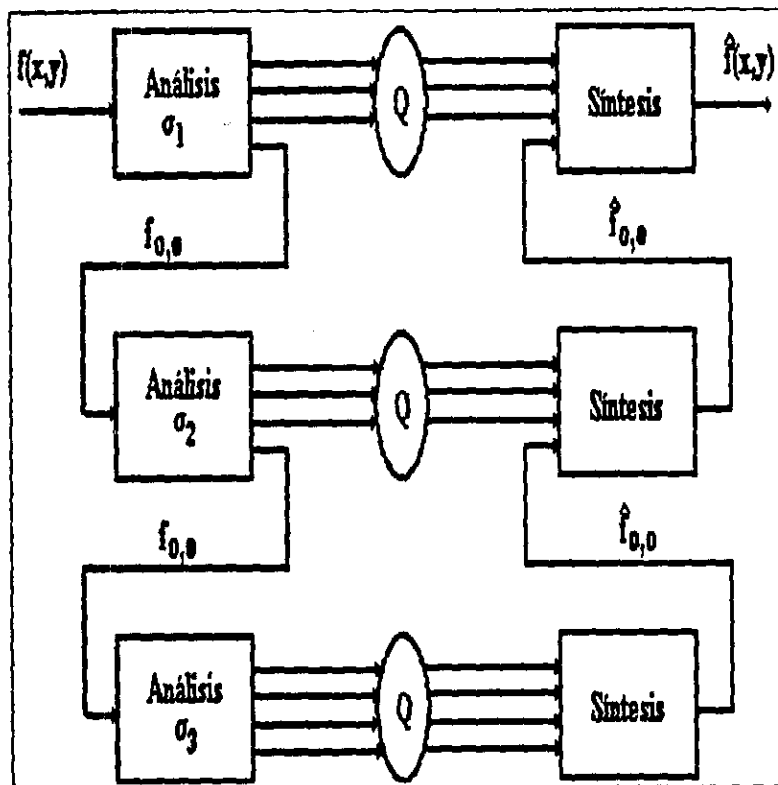


Figura 4.7: Esquema de codificación a múltiples escalas.

La localización de la imagen se efectúa de la siguiente manera: Se realiza el análisis de la imagen a una resolución dada, cuya varianza de la ventana es  $\sigma_1$ . En el siguiente nivel se toma el coeficiente de orden cero, que guarda la información de baja frecuencia de la imagen original y se le aplica nuevamente una transformación polinomial cuya varianza de la ventana ahora es  $\sigma_2$ , y así sucesivamente.

La varianza de la imagen es diferente en cada nivel pues aunque  $\sigma_n$ , donde  $n$  es el nivel, tenga el mismo valor, la imagen se va suavizando cada vez más; si a ello le

aunamos un muestreo, entonces la compresión de los datos es mayor.

La reconstrucción de la imagen se lleva a cabo en el sentido inverso, pues del nivel mas bajo se interpola el coeficiente de orden cero de la expansión superior, de tal manera que en la salida del primer nivel se tiene una recuperación de la imagen.

Las representaciones que se proponen en esta parte son esencialmente las mismas que las de una escala: representación bidimensional, representación unidimensional y representación en parámetros característicos

De la misma manera que en el caso anterior, la *representación bidimensional*, se usa para efectos de comparación y a pesar de ser una transformada cuya expansión es redundante, se ha comprobado que con otros esquemas de codificación se puede utilizar para realizar codificación piramidal predictiva [Esc95].

La *representación unidimensional*, a pesar de no tomar la expansión completa en coeficientes polinomiales, presenta generalmente buenos resultados.

En el caso de la *representación en parámetros característicos*, la pérdida de información entre niveles, debe ser mayor. Pues algunos detalles muy finos se perderan.

## 4.4 Conclusiones

La transformada polinomial se ha utilizado comunmente para casos de realce y restauración de imágenes, sin embargo, en este capítulo se han propuesto métodos de codificación de imágenes, que involucran el uso de la transformada polinomial como técnica de análisis de imágenes.

Se presenta la proyección en coeficientes óptimos unidimensionales y a partir de ellos, la extracción de los parámetros característicos de los bordes contenidos en la imagen.

Se sabe que los bordes son una de las estructuras primarias más importantes para la percepción que se encuentran contenidas en una imagen.

Los algoritmos propuestos consisten en buscar representaciones que consigan realizar compresión. Para ello, se obtienen las representaciones bidimensional, unidimensional y en parámetros característicos en una escala y a múltiples escalas. Los resultados que se obtienen con estos algoritmos se presentan en el proximo capítulo.

# Capítulo 5

## Resultados

En este capítulo se presenta el funcionamiento de los algoritmos de codificación propuestos. Para ello se utilizan las imágenes LENA y CASA, cuyas características se describen. Se mencionan las medidas de fiabilidad que se utilizan para conocer el desempeño de los algoritmos. Codificación en una escala y en múltiples escalas, son las dos categorías en las que se divide el conjunto de representaciones logradas a partir de los algoritmos propuestos. En cada una de las categorías se tiene la variación del tamaño de la función ventana ( $M$ ) y el muestreo ( $T$ ). Al final de cada categoría se presentan gráficas de desempeño del algoritmo con los valores de entropía, relación señal a ruido y tasa de compresión como medidas de fiabilidad.

### 5.1 Características generales

#### 5.1.1 De las imágenes utilizadas

Las imágenes que se utilizan para la presentación de resultados son Lena y Casa cuyas dimensiones<sup>1</sup>, entropía y datos estadísticos como media y desviación estándar, se indican en la tabla 5.1.

donde se aprecian las dimensiones, datos estadísticos y los valores de entropía.

Nombre Imagen	Dimensiones por lado	Media	Desviación Estándar	Entropía
Lena	256 pixels	119.6597	57.5659	7.638083
Casa	256 pixels	133.7711	55.6163	6.517494

Tabla 5.1 Datos estadísticos de las imágenes.

En la figura 5.1a se observa la imagen LENA y la figura 5.1b corresponde a su histograma, esta imagen presenta una gran cantidad de contornos que delimitan diferentes regiones en la imagen, esto se debe a la iluminación con la que cuenta la

<sup>1</sup>Las dimensiones de la imagen impresa en las figuras están directamente relacionadas con el despliegue en pantalla, un pixel en la pantalla, equivale a 25 milímetros en la figura.

escena, sin embargo se observan estructuras como bordes o textura de alta frecuencia muy bien definidas. El histograma presenta una distribución con picos localizados en diferentes valores de la imagen, esto hace que sea casi uniforme y por lo tanto tenga una entropía alta.

En la figura 5.1c se tiene a la imagen CASA que a diferencia de la anterior esta compuesta principalmente por bordes que delimitan claramente zonas de contraste alto, dentro de esas regiones se encuentran frecuencias que no llegan a ser muy altas; la figura 5.1d presenta el histograma donde se observa que valores se encuentran contenidos en esta imagen.

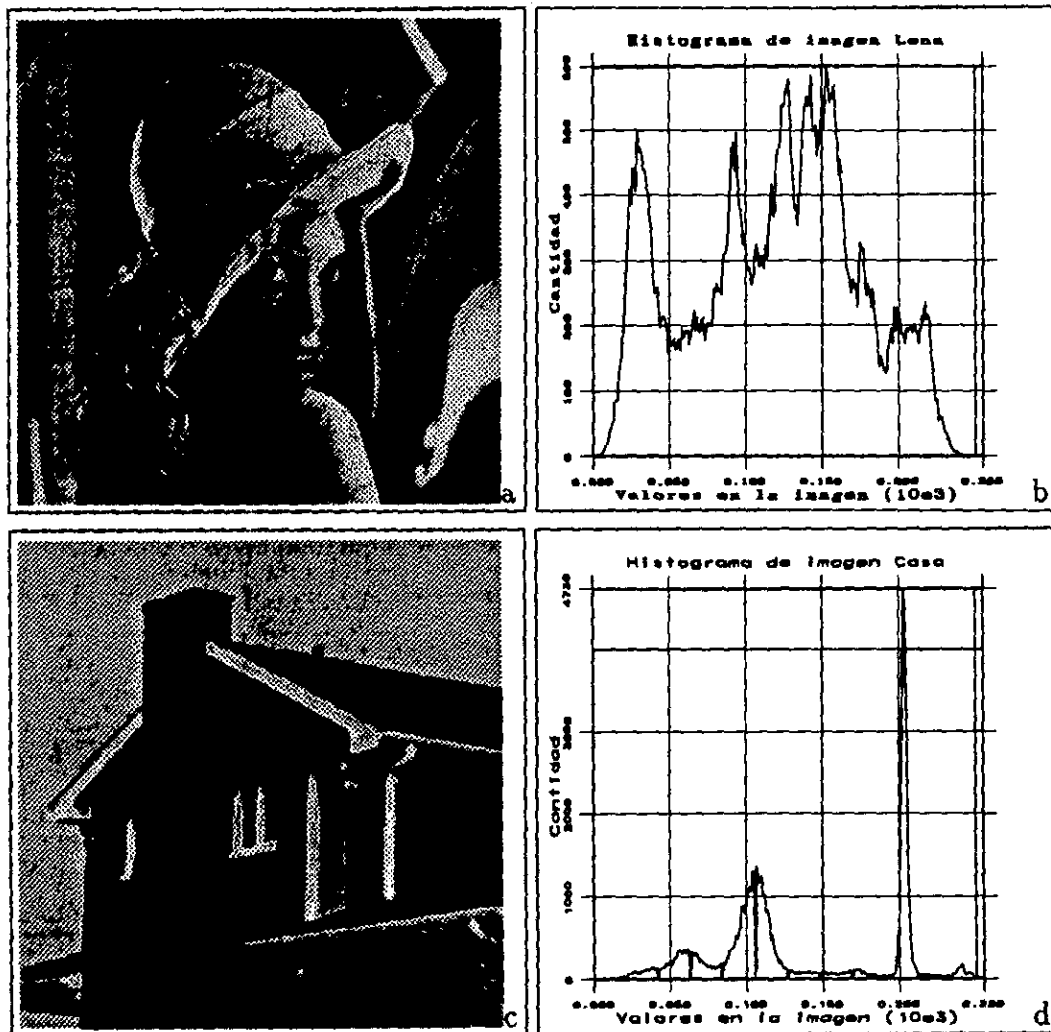


Figura 5.1: Arriba a la izquierda, se observa (a) la imagen Lena, a la derecha se encuentra (b) el histograma de valores contenidos en la imagen. Abajo a la izquierda, se observa (c) la imagen Casa, a la derecha (d) su histograma de valores.

### 5.1.2 De los algoritmos de codificación propuestos

El análisis y codificación por medio de la transformada polinomial a las imágenes mencionadas se efectúa siguiendo los algoritmos propuestos. Para efectos de codificación se realizan variaciones al tamaño de la ventana ( $M$ ) y a la cantidad de muestreo ( $T$ ).

Los algoritmos propuestos se dividen en dos categorías:

- Codificación en una escala
- Codificación en múltiples escalas

En cada una de ellas se obtienen las representaciones propuestas que son: bidimensionales, unidimensionales y en parámetros característicos de estructuras unidimensionales. La representación bidimensional que se obtiene es una matriz triangular superior de coeficientes. Es decir, solo aquellos coeficientes cuyo orden sea menor o igual al orden indicado por el tamaño de la ventana de análisis se tomarán en cuenta. Se ha demostrado [Mar90b], que los coeficientes mayores al orden máximo indicado por la ventana no son relevantes para la percepción. La energía de la imagen se encuentra concentrada principalmente en los coeficientes de orden menor, disminuyendo conforme aumenta el orden del coeficiente.

La representación unidimensional es la proyección en la dirección óptima de los coeficientes bidimensionales. Se representa por un conjunto de coeficientes unidimensionales  $K_n$ , que se encuentran en una dirección óptima  $\theta_{optima}$ . El orden de los coeficientes unidimensionales es  $n = 0, \dots, M$ .

En la representación en parámetros característicos se presentan la imagen de distancias de bordes al centro de la ventana  $x_b$ , la de magnitudes de los bordes  $\Delta f$ , orientaciones  $\theta_{optima}$  y la imagen de baja frecuencia o valor medio  $A_b$ . Es en esta representación donde los niveles de compresión son los más altos. Se obtiene esta representación a partir de los coeficientes unidimensionales por lo cual los valores de ventana  $M$  y muestreo  $T$  son los mismos.

La función ventana que se utilizó es una *distribución binomial*, cuyos valores usados son  $M = 2, 4, 6, 8$ . El muestreo en cada representación se toma  $T = M$ , porque  $M$  es la varianza de la función ventana, por lo tanto es la mayor tasa de muestreo que se puede alcanzar para una imagen suavizada con una ventana binomial de tamaño  $M$ . En las representaciones a múltiples escalas cuando  $M = 6, 8$ , se utilizan solo dos niveles; debido a que último nivel de la pirámide en estos casos tiene imágenes muy pequeñas.

## 5.2 Medidas de fiabilidad

Para medir el desempeño de los algoritmos se utilizan la entropía y la relación señal a ruido pico (PSNR), que se describieron con las ecuaciones 2.4 y 2.6 respectivamente. Otra manera de evaluar la calidad de la reconstrucción de la imagen es por medio de la apreciación subjetiva, para ello se presentan las reconstrucciones que se obtienen.



### 5.2.1 Entropia

Se ha visto ya el calculo y concepto de la entropia en el capitulo 2 de este texto. Sin embargo, para realizar el calculo de la entropia total en un esquema de codificación como el propuesto en este trabajo, la formula es

$$H_{total} = \sum_{j=1}^{N+1} \left( \frac{H_j}{M^j} \right) \quad (5.1)$$

donde:

- $H_{total}$  es la entropia total de la representación de la imagen.
- $H_j$  es la entropia en el nivel  $j$ .
- $N$  es el número de niveles, cuyo rango es  $(0, \dots, M)$ .
- $M$  es el tamaño de la ventana<sup>2</sup>.

### 5.2.2 Relación señal a ruido

La relación señal a ruido, se calcula en base al error que se genera entre la imagen reconstruida y la imagen original. Por esta razón, este calculo es independiente del esquema de codificación utilizado. Se utilizan las ecuaciones 2.5 y 2.6.

Sirve como una medida de referencia para conocer la calidad de la reconstrucción de la imagen. Generalmente una imagen que tiene una relación señal a ruido mayor a 35dB, presenta degradaciones imperceptibles. Las degradaciones son visibles cuando el valor de la relación señal a ruido es menor a 30 dB.

### 5.2.3 Tasa de compresión

La tasa de compresión se calcula a partir de conocer el número de bytes que se obtienen después de realizar la cuantización y codificación de Huffmann [Rab91], la fórmula utilizada es la ecuación 2.1 como se muestra a continuación

$$TC = \frac{\text{bytes\_imagen\_original}}{\text{bytes\_codigo}} \quad (5.2)$$

donde:  $TC$  es la tasa de compresión,  $\text{bytes\_imagen\_original}$  es el número de bytes de la imagen original, que se calcula conociendo las dimensiones; y  $\text{bytes\_codigo}$  es el número total de bytes de la representación.

Se presentan gráficas de desempeño para cada categoría, donde se presenta el aumento del tamaño de la ventana y los valores de las medidas de fiabilidad propuestas.

---

<sup>2</sup>Se asume que  $T = M$  en la codificación propuesta.

## 5.3 Algoritmos de codificación en una escala

En esta sección se presenta el funcionamiento del algoritmo de codificación en una escala. Se tienen las tres representaciones: bidimensional, unidimensional y en parámetros característicos. El tamaño de la ventana tiene un rango  $M = 2, 4, 6, 8$ . La varianza de la ventana es fija, esto tiene como consecuencia que estructuras con varianza diferente no sean localizadas por el análisis y por lo tanto no se modelen. Sin embargo dichas estructuras se encuentran contenidas en la señal de baja frecuencia. La tasa de compresión es directamente proporcional al tamaño de la ventana, por esta razón, cuando la ventana incrementa su tamaño se obtiene una tasa de compresión alta. Por lo tanto, la entropía de la representación disminuye con el incremento del tamaño de la ventana.

### 5.3.1 Representación bidimensional

En la siguiente serie de figuras, se observan las representaciones polinomiales bidimensionales de las imágenes Lena y Casa. Se obtuvieron variando el tamaño de ventana  $M$  y el valor del muestreo  $T$ . Debido a que se trata de una escala de análisis, la mayor parte de la información queda concentrada en el coeficiente de orden cero ( $f_{0,0}$ ). Por ello la entropía es muy alta, sin embargo se logra realizar codificación eficiente, es decir valores modestos en las tasas de compresión. La representación bidimensional se presenta para efectos de comparación.

En la figura 5.2a y 5.2b se aprecia la representación bidimensional de las imágenes Lena y Casa respectivamente. El tamaño de ventana binomial empleado es  $M = 2$  y el valor de muestreo  $T = 2$ . La entropía de los conjuntos es de 6.63 bits por pixel (bpp) para la imagen Lena y 5.43 bpp para la Casa. La tasa de compresión que se alcanza después de un codificador de Huffman es de 1.19 veces para la imagen Lena y 1.46 veces para la Casa.

En la figura 5.2, los coeficientes se encuentran dispuestos de la misma manera en (a) y en (b), en forma de matriz de coeficientes. En la fila superior se tienen de izquierda a derecha, los coeficientes  $f_{0,0}, f_{0,1}, f_{0,2}$ . Los coeficientes de abajo corresponden en el mismo orden a  $f_{1,0}, f_{1,1}$ . El coeficiente que queda en la parte inferior de las dos figuras corresponde a  $f_{2,0}$ .

Como se puede observar en la figura 5.2, los bordes son resaltados en los coeficientes de primer orden y las estructuras de orden 2 se observan mejor en los coeficientes de segundo orden. Aquí la causa por la cual se está logrando una representación eficiente de la imagen, es el muestreo que se efectúa en el conjunto de coeficientes. Las reconstrucciones que se observan en la figura 5.3 de (a) Lena y (b) Casa, tienen degradaciones prácticamente imperceptibles que seguramente se deben al tipo de expansión en coeficientes. Si la expansión es completa, la reconstrucción es perfecta, es decir sin error numérico; ya que la propiedad de traslape entre ventanas permite generar mayor información de la que se tenía inicialmente. La relación señal a ruido es 39.3741 y 45.6788 dB para las reconstrucciones de Lena y Casa respectivamente.



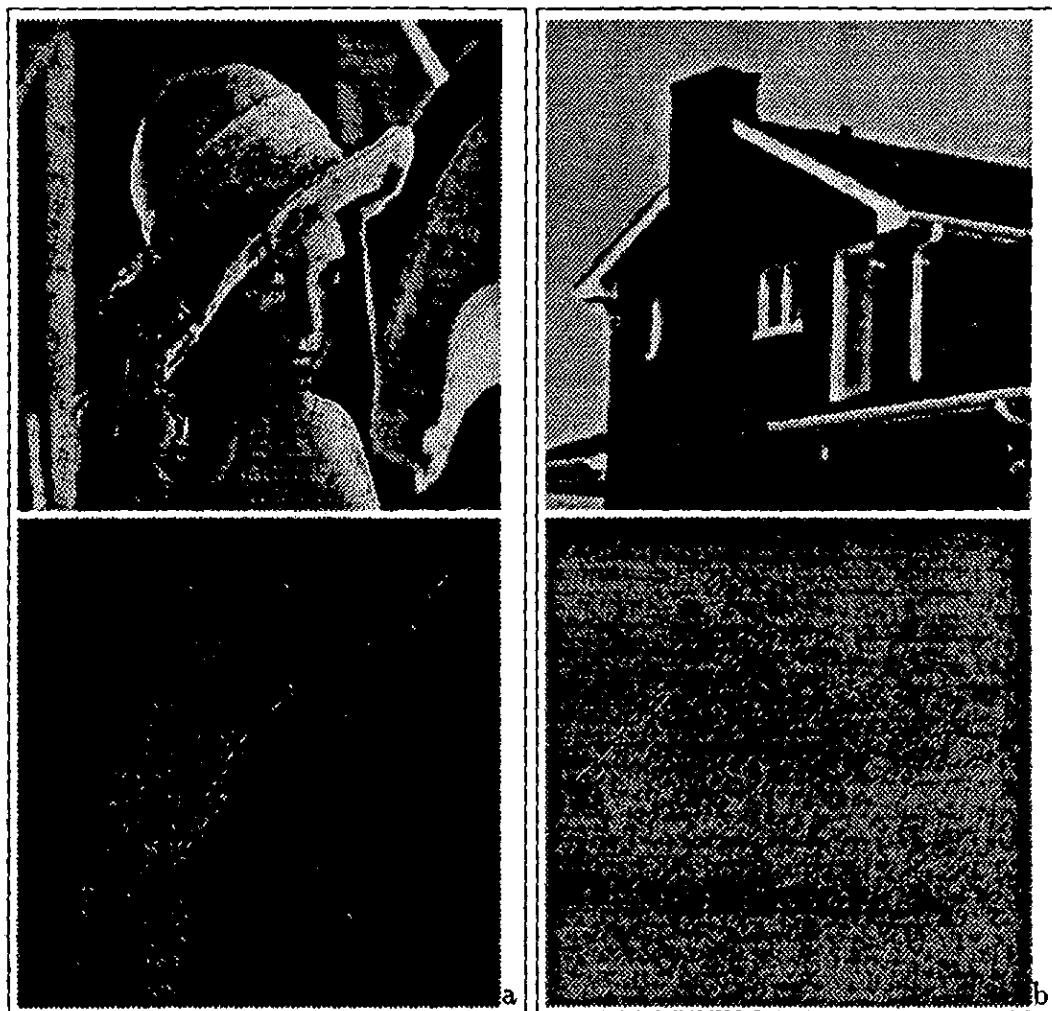


Figura 5.3: Reconstrucción de imágenes a partir de su representación bidimensional ( $M = 2$  y  $T = 2$ ). Arriba a la izquierda (a), se tiene la imagen Lena reconstruida, abajo de ella se encuentra la imagen de error con respecto a la original. Arriba a la derecha (b), se encuentra la imagen de la Casa reconstruida, abajo de ella se encuentra la imagen de error con respecto a la original.

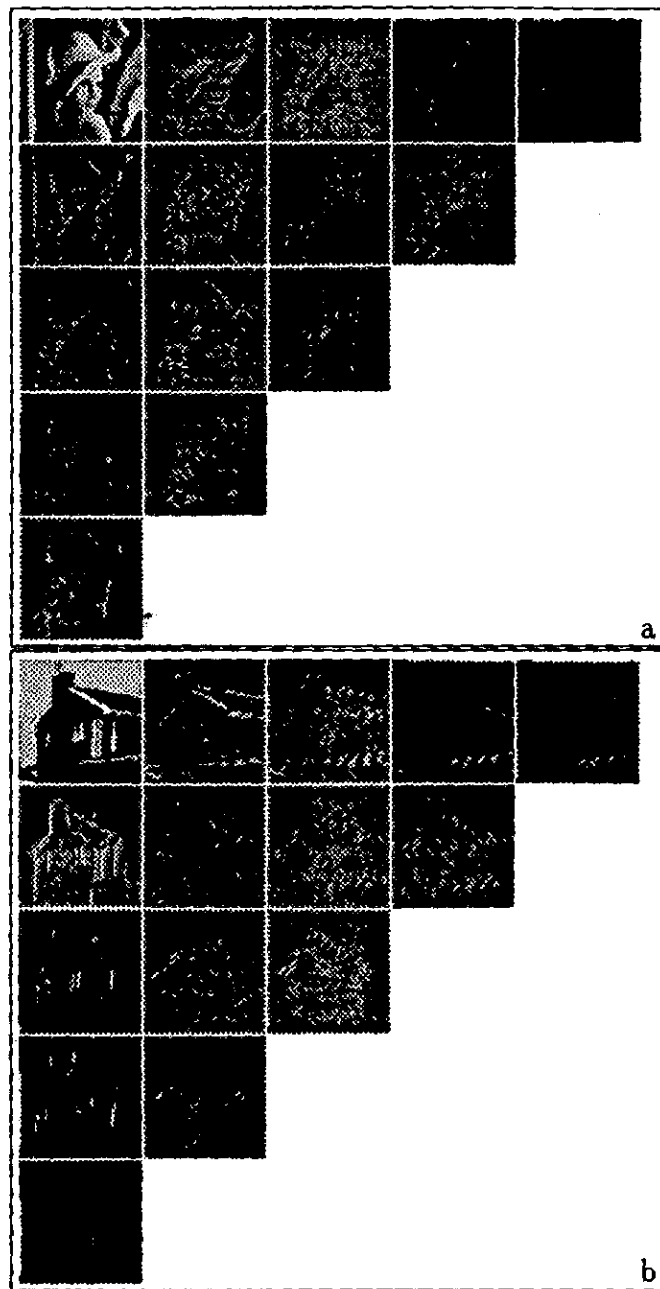


Figura 5.4: La representación bidimensional de la imagen Lena se presenta arriba (a), la correspondiente a la imagen Casa abajo (b), con tamaño de ventana  $M = 4$  y submuestreo  $T = 4$ . Los coeficientes se encuentran dispuestos (de la misma manera en las dos representaciones) de izquierda a derecha y empezando por la fila superior:  $f_{0,0}, f_{0,1}, f_{0,2}, f_{0,3}, f_{0,4}$ ; en la siguiente fila se encuentran los coeficientes  $f_{1,0}, f_{1,1}, f_{1,2}, f_{1,3}$ . La siguiente fila se compone de los coeficientes  $f_{2,0}, f_{2,1}, f_{2,2}$ . Los coeficientes  $f_{3,0}, f_{3,1}$  forman la penúltima fila y  $f_{4,0}$  es la última fila.

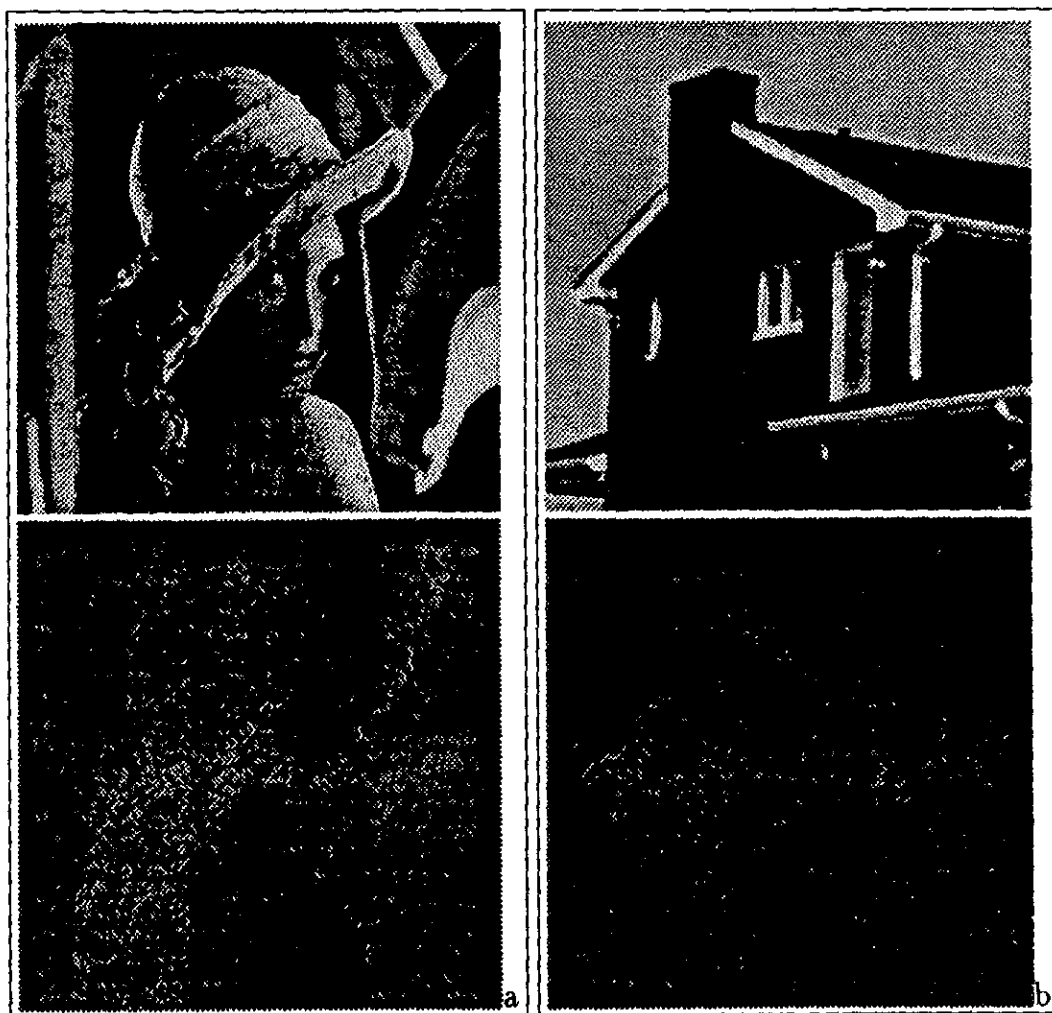


Figura 5.5: Las imágenes reconstruidas a partir de sus respectivas representaciones bidimensionales ( $M = 4$  y  $T = 4$ ), se observan en esta figura. Arriba a la izquierda (a), se tiene la imagen Lena reconstruida. Arriba a la derecha (b), se encuentra la imagen de la Casa reconstruida. Abajo de cada una de ellas se encuentran las imágenes de error con respecto a la original.

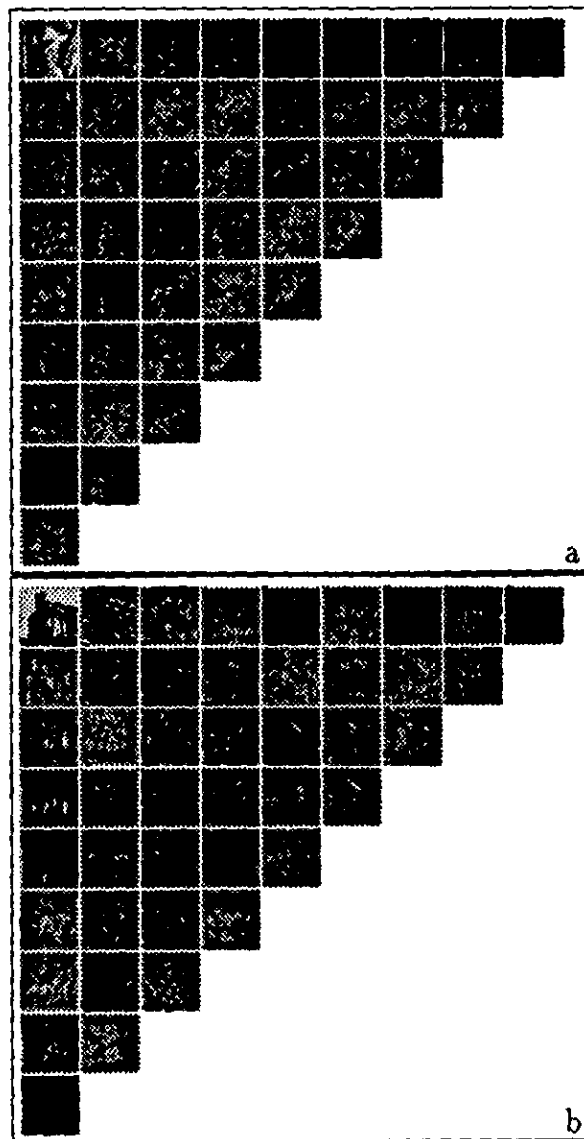


Figura 5.6: Representación bidimensional de las imágenes (a) Lena y (b) Casa, con tamaño de ventana  $M = 8$  y muestreo  $T = 8$ . Los coeficientes  $f_{R,C}$  donde: Los renglones  $R$  tienen un rango que es  $(0, \dots, M)$  y las columnas  $C$  tienen un rango  $(0, N)$  siempre y cuando  $R + C \leq M$ . Si por ejemplo se quiere saber cual es el coeficiente  $f_{4,3}$ , simplemente se busca en el renglon 4 y la columna 3. El resultado es la matriz triangular superior para cada representación.

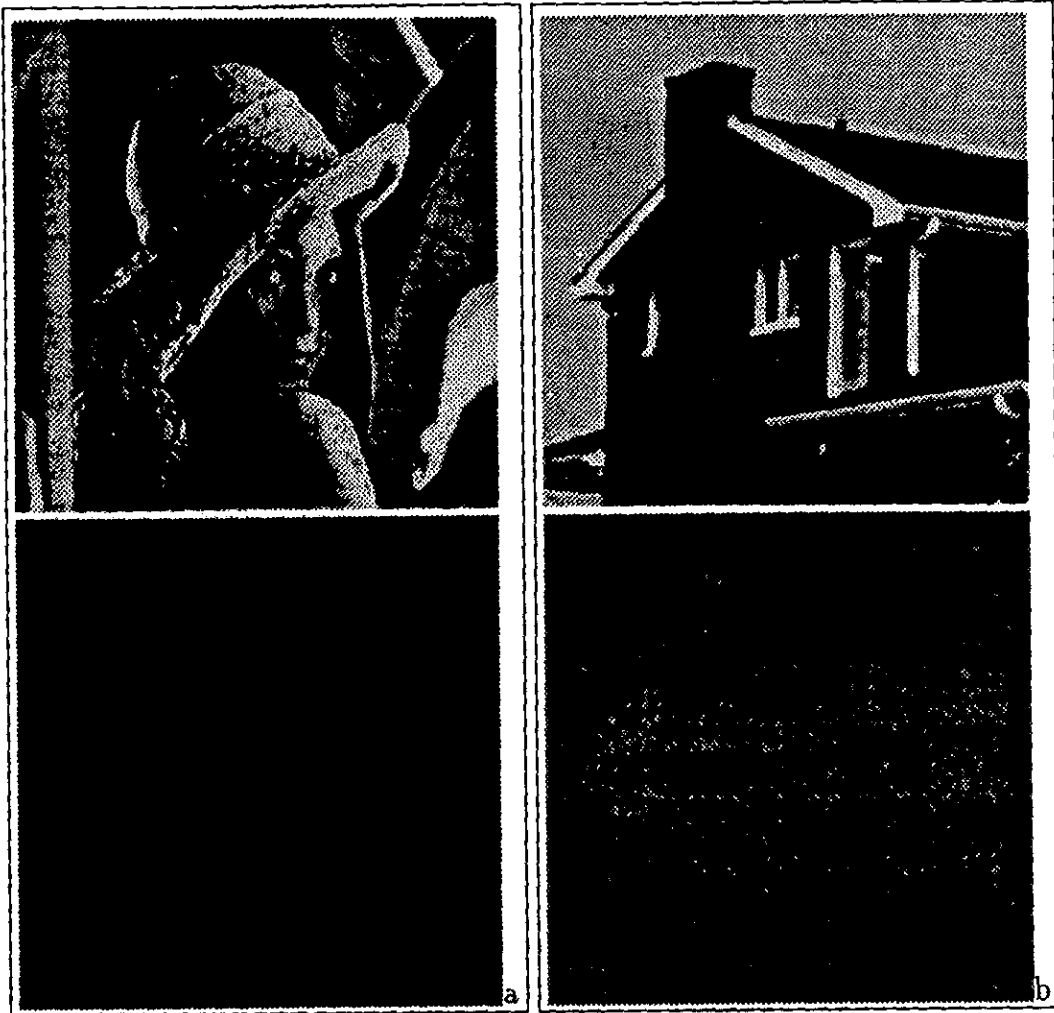


Figura 5.7: Las imágenes reconstruidas a partir de la representación bidimensional con parámetros  $M = 8$  y  $T = 8$ , (a) Lena y (b) Casa. Sus respectivos errores se observan abajo de cada una de ellas.



En la representación bidimensional de las imágenes Lena y Casa, con tamaño de ventana  $M = 4$  y muestreo  $T = 4$  (figuras 5.4a y 5.4b respectivamente), se puede observar una pequeña compactación de datos. Los coeficientes tienen ahora localizadas estructuras de frecuencias más altas en los coeficientes mayores a orden 2.

En la figura 5.5, se presentan las imágenes reconstruidas a partir de las representaciones anteriores de (a) Lena y (b) Casa. El valor de la relación señal a ruido de las imágenes es 36.44 y 42.17 dB para Lena y Casa respectivamente. Esto indica que la reconstrucción es buena y se comprueba en la apreciación visual de la imagen. De la misma manera que en el caso anterior, el error se debe al tipo de expansión que se efectuó, matriz de coeficientes triangular superior y no completa.

Como se ha observado en el desempeño del algoritmo, si el muestreo  $T$  se hace cada vez mayor, se pueden obtener tasas de compresión altas. Para ello se realiza el análisis de la imagen con una ventana de tamaño  $M = 8$  y muestreo  $T = 8$ . Con ello se logra comprimir la imagen de manera significativa, pues el conjunto de coeficientes que se genera tiene un valor de entropía muy bajo 2.03 bpp para la imagen Lena y 1.77 bpp para la imagen Casa. Los coeficientes se encuentran dispuestos en una matriz triangular superior de imágenes, como se indica en la figura 5.6, (a) representación bidimensional de la imagen Lena y (b) de la imagen Casa.

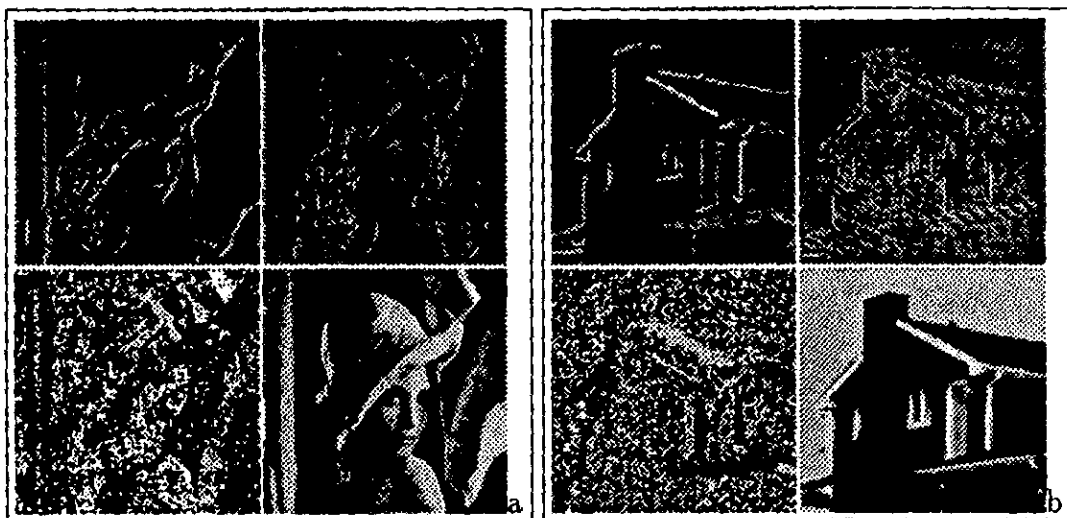


Figura 5.8: Representación polinomial unidimensional de las imágenes (a) Lena y (b) Casa, con  $M = 2$  y  $T = 2$ . En las dos figuras se encuentran dispuestos los coeficientes  $K_n$  de la misma forma: de izquierda a derecha en la parte superior se observan los coeficientes de orden  $K_1$  y  $K_2$ , abajo se encuentra la imagen de orientaciones óptimas  $\theta_{optima}$  y la señal de baja frecuencia o coeficiente de orden  $K_0$ .

La reconstrucción de las imágenes a partir de las representaciones de la figura 5.6, se

observa en las figuras 5.7a y 5.7b, Lena y Casa respectivamente. Aquí aparecen unos artefactos que son similares al fenómeno de Gibbs [Mal92b]; estos artefactos están principalmente localizados cerca de lugares donde hay bordes de alto contraste. La imagen Lena tiene mayor contenido de artefactos debido a que su frecuencia espacial es más alta que la imagen Casa. Los errores de otro tipo como pérdida de estructuras no son perceptualmente importantes.

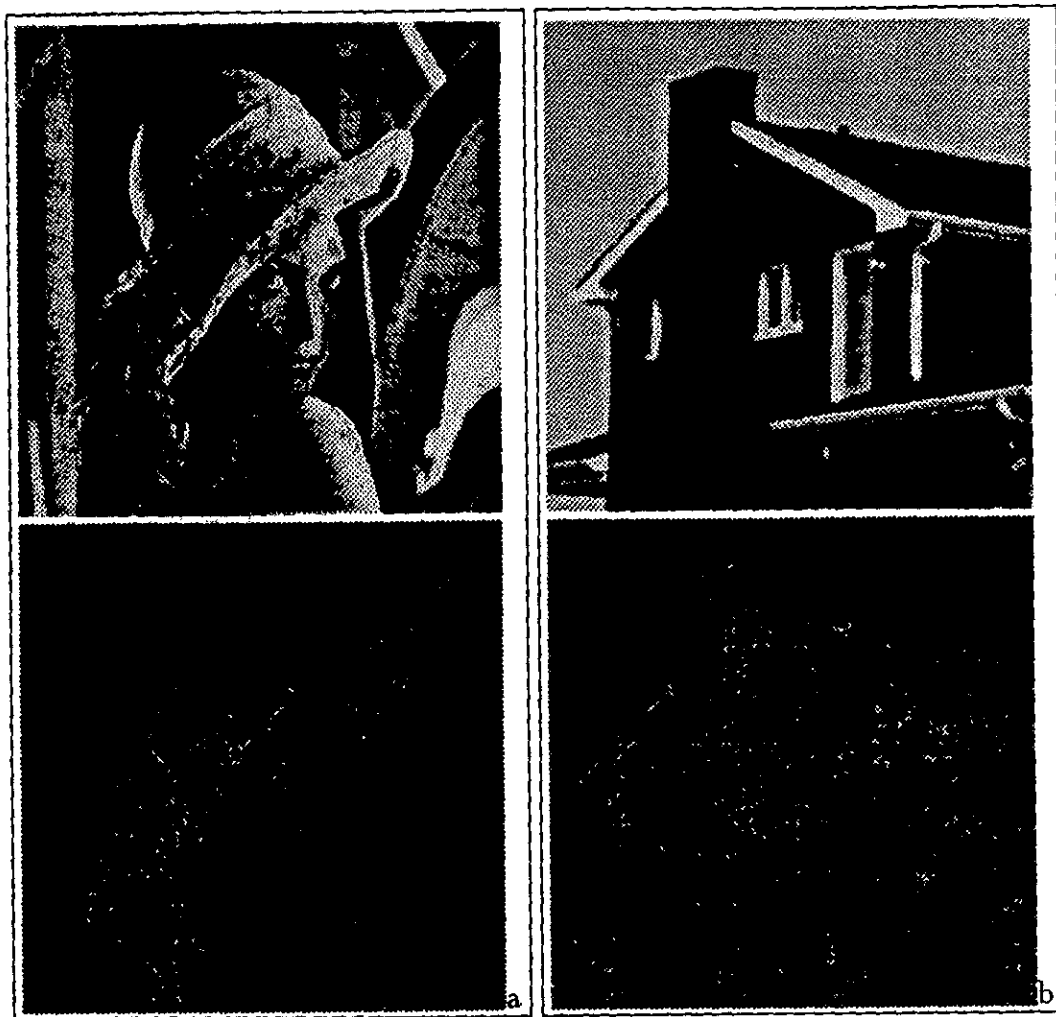


Figura 5.9: Reconstrucción a partir de la representación unidimensional de la figura 5.8 . Arriba se observan las reconstrucciones de las imágenes (a) Lena y (b) Casa. Abajo de ellas, su respectivo error de reconstrucción.

### 5.3.2 Representación unidimensional

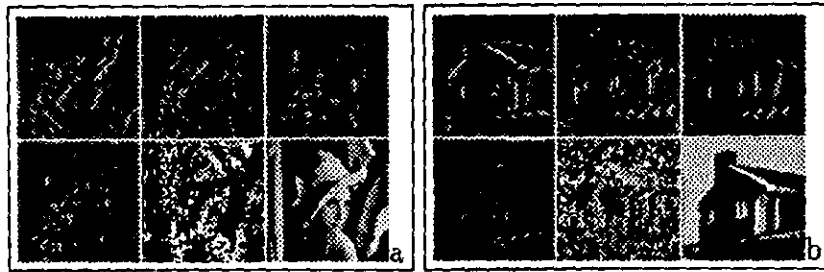


Figura 5.10: En las representaciones unidimensionales de las imágenes (a) Lena y (b) Casa, con  $M = 4$  y  $T = 4$ , se encuentran los coeficientes dispuestos de la siguiente manera: arriba y de izquierda a derecha, se tienen los coeficientes  $K_1$ ,  $K_2$  y  $K_3$ . Abajo y de izquierda a derecha se observa el coeficiente  $K_4$ , la imagen de orientaciones óptimas  $\theta_{optima}$  y el coeficiente de orden 0,  $K_0$ .

A partir de las proyección en una dirección óptima, se obtienen los coeficientes que forman la representación polinomial unidimensional, se utilizan las representaciones antes presentadas para obtener las que se ilustran en la figura 5.8, donde se observa la representación unidimensional de las imágenes (a) Lena y (b) Casa. El coeficiente de orden 0,  $K_0$ , es el que contiene la mayor cantidad de información y corresponde a la suavización de la imagen original por la función ventana binomial. En la parte superior de las imágenes se observan los coeficientes  $K_0$  y  $K_1$ , que contienen la información de estructuras primarias de la imagen. La imagen de orientaciones óptimas  $\theta_{optima}$  se encuentra abajo a la izquierda y es un control que indica la dirección donde el contraste de la imagen es mayor en un punto dado. El tamaño de ventana utilizado es  $M = 2$  y el muestreo  $T = 2$ .

La reconstrucción de las imágenes aparece en la figura 5.9, (a) la imagen Lena y (b) la imagen Casa. El error de reconstrucción se presenta abajo de las reconstrucciones correspondientes. La relación señal a ruido de las imágenes es 37.18 y 41.58 dB para Lena y Casa respectivamente. La calidad subjetiva de la imagen es muy buena pues no se aprecian degradaciones.

Con la representación unidimensional de las imágenes Lena en la figura 5.10a y Casa en la figura 5.10b, se logra una tasa de compresión de 4.5 y 5 veces respectivamente, después de haber pasado por un codificador de Huffman. Esto quiere decir que se han compactado los datos conservando importantes estructuras de la imagen. Se utiliza una ventana de tamaño  $M = 4$  y  $T = 4$ .

En las figuras 5.11a y 5.11b se aprecia la reconstrucción de las imágenes Lena y Casa respectivamente, abajo de ellas su correspondiente imagen de error. Aquí se puede observar que las degradaciones son mas notorias, principalmente en la imagen Lena que contiene información más compleja que la imagen Casa.

La representación unidimensional con tamaño de ventana  $M = 8$  y  $T = 8$ , que se

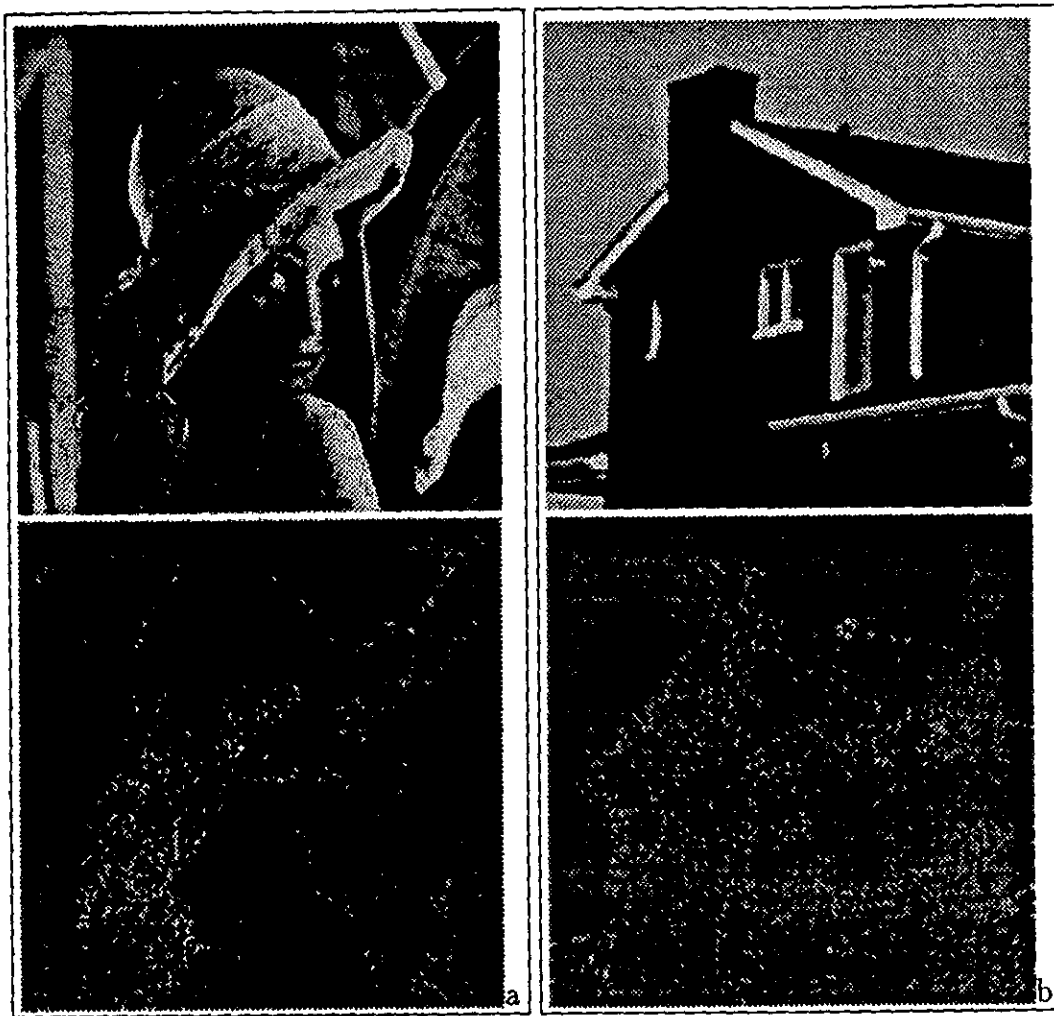


Figura 5.11: A partir de las representaciones de la figura 5.10, se recuperan las imágenes (a) Lena y (b) Casa. Abajo de ellas la imagen de error correspondiente.

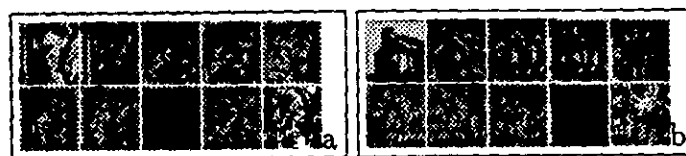


Figura 5.12: Representación unidimensional con  $M = 8$  y  $T = 8$  de (a) Lena y (b) Casa. En las dos figuras se encuentran los coeficientes dispuestos de la misma manera: de arriba a abajo y de izquierda a derecha se encuentran los coeficientes  $K_0, K_1, K_2, K_3, K_4$  y  $K_5, K_6, K_7, K_8$  y orientaciones  $\theta_{optima}$ .

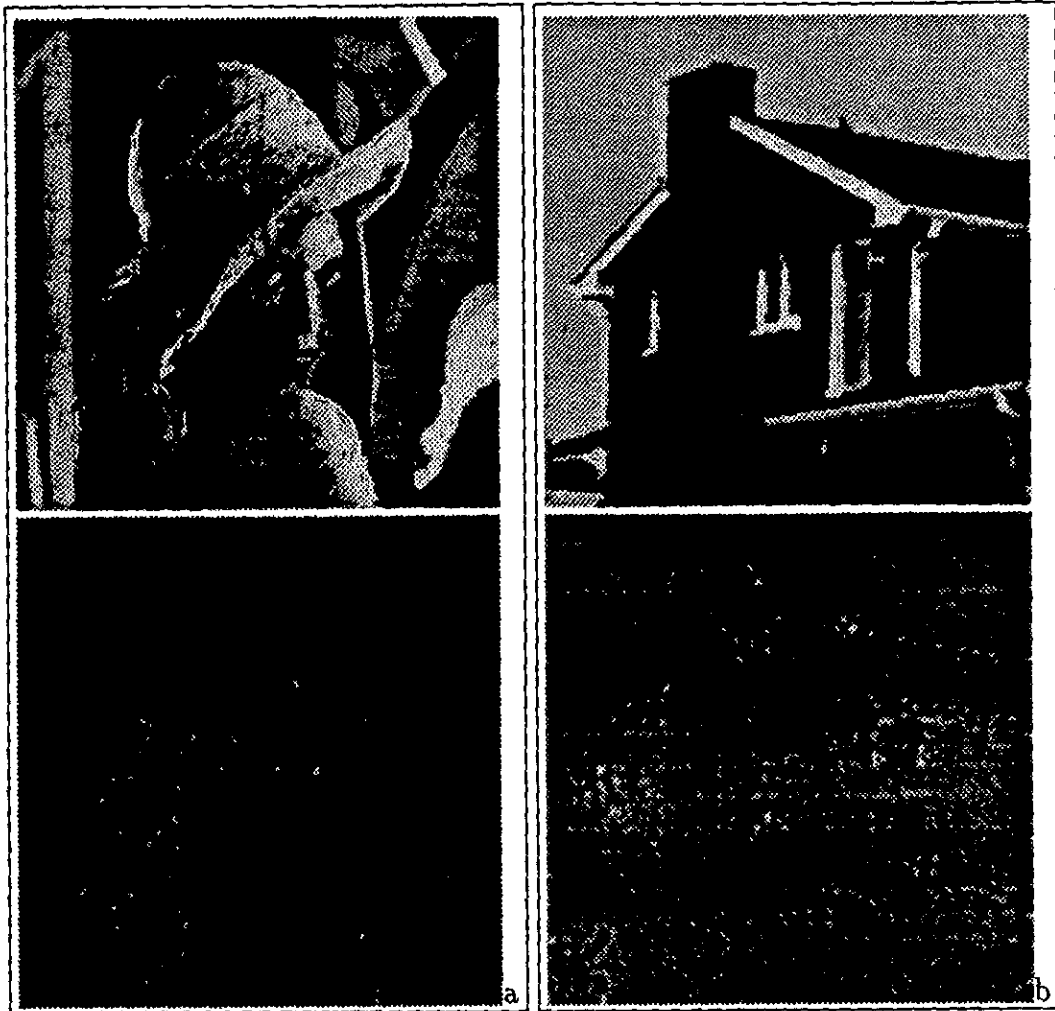


Figura 5.13: Reconstrucción a partir de la representación  $M = 8$  y  $T = 8$ , de las imágenes (a) Lena y (b) Casa. Abajo de cada reconstrucción se encuentran las imágenes de error correspondientes.



Figura 5.14: Acercamiento a la reconstrucción de la imagen Lena: (a) con  $M = 2, T = 2$ , (b) con  $M = 4, T = 4$ , (c) con  $M = 8, T = 8$ . Se observa una parte de la cara, aumentada 2 veces.

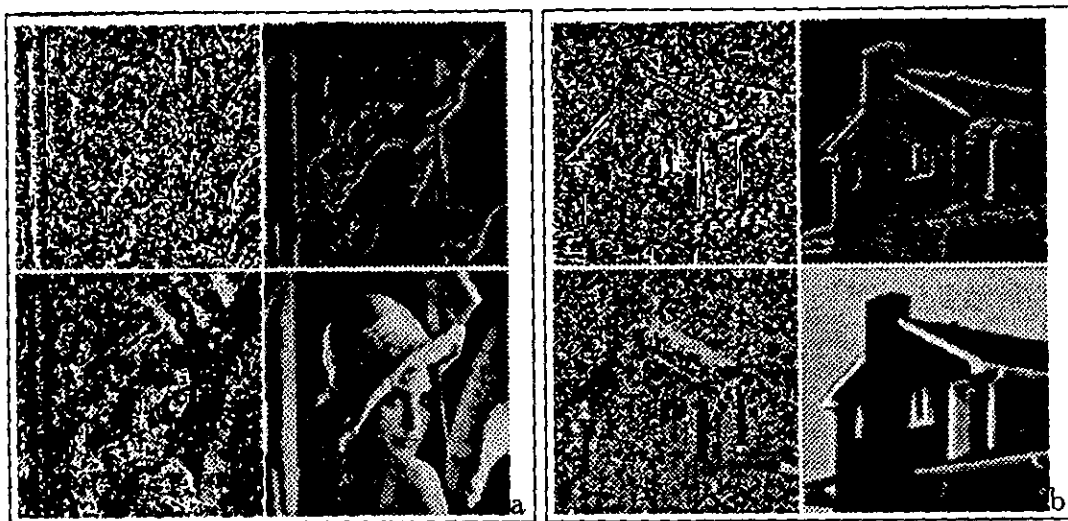


Figura 5.15: Representación en parámetros característicos de bordes, con  $M = 2$  y  $T = 2$  de la imágenes (a) Lena y (b) Casa. En la parte superior de las dos figuras se encuentra la imagen con valores de posición  $x_b$  y a su derecha la imagen de magnitud de los bordes contenidos en la imagen  $\Delta f$ . Abajo se tienen las imágenes de orientaciones  $\theta_{optima}$  y a la derecha el valor medio  $A_b$ .

observa en la figura 5.12, corresponde a las imágenes (a) Lena y (b) Casa. En las imágenes se aprecian el coeficiente de orden cero  $K_0$  en la parte superior izquierda y la imagen de orientaciones óptimas  $\theta_{\text{óptima}}$  en la parte inferior derecha. Los coeficientes  $K_n$  donde  $n = 1, \dots, 8$ , se encuentran de izquierda a derecha y de arriba a abajo después del coeficiente  $K_0$ . La entropía de la representación de las figuras es de 0.7125 y 0.6519 bpp para Lena y Casa respectivamente. Con esto se obtiene una compresión importante.

La reconstrucción de las imágenes Lena y Casa, se muestra en las figuras 5.13a y 5.13b respectivamente. Se observa un efecto de bloques debido principalmente al muestreo tan alto al que se somete la imagen, este es más notorio en la imagen Lena que en la imagen Casa. La relación señal a ruido que se obtiene es de 26.57 dB para la imagen Lena y 29.71 dB para la imagen Casa. La imagen Casa se encuentra más cerca del umbral de percepción antes mencionado, por lo cual tiene una mayor calidad subjetiva.

A manera de presentación de los efectos que tiene el procesamiento efectuado en la imagen Lena, se presenta un acercamiento a las reconstrucciones logradas, estas tienen un factor de 2. La figura 5.14a corresponde a la lograda a partir de la representación unidimensional con  $M = 2$  y  $T = 2$ , la figura 5.14b a la reconstrucción a partir de la representación unidimensional con  $M = 4$  y  $T = 4$  y finalmente la figura 5.14c a la reconstrucción que se logra a partir de la representación unidimensional con  $M = 8$  y  $T = 8$ . En ellas se puede observar el aumento de la degradación a partir del uso de una ventana más grande y un submuestreo mayor.

### 5.3.3 Representación en parámetros característicos

La parte final del algoritmo es la búsqueda de parámetros que caracterizan estructuras primarias contenidas en la imagen, en este caso bordes. Los parámetros son la distancia del borde al centro de la ventana  $x_b$ , la magnitud del borde  $\Delta f$ , la orientación óptima del borde  $\theta_{\text{óptima}}$  y el valor medio de ese borde  $A_b$ . Es en esta parte donde se esperan las mayores tasas de compresión. Como se está haciendo una selección de estructuras perceptualmente importantes, se realiza entonces cierta eliminación de la redundancia psicovisual.

En la figura 5.15, se tiene la representación en parámetros característicos de las imágenes (a) Lena y (b) Casa, con  $M = 2$  y  $T = 2$  para una escala de análisis. Se observa en la imagen superior izquierda de los dos conjuntos a la imagen de valores de distancia del borde  $x_b$ , esta imagen es altamente cuantizable ya que el tamaño de la ventana  $M$ , limita el número de valores. A la derecha se encuentra la imagen de magnitud de los bordes contenidos en la imagen, en la imagen Lena esta imagen tiene un mayor contenido de información que en la imagen Casa, esto se debe a la cantidad de bordes que se encuentran en cada una de ellas. Abajo se observan las imágenes de orientaciones  $\theta_{\text{óptima}}$  y la de valor medio  $A_b$ . Se logra una reducción importante en la entropía del conjunto con valor de 4.37 y 3.95 bpp para la imagen Lena y Casa respectivamente.

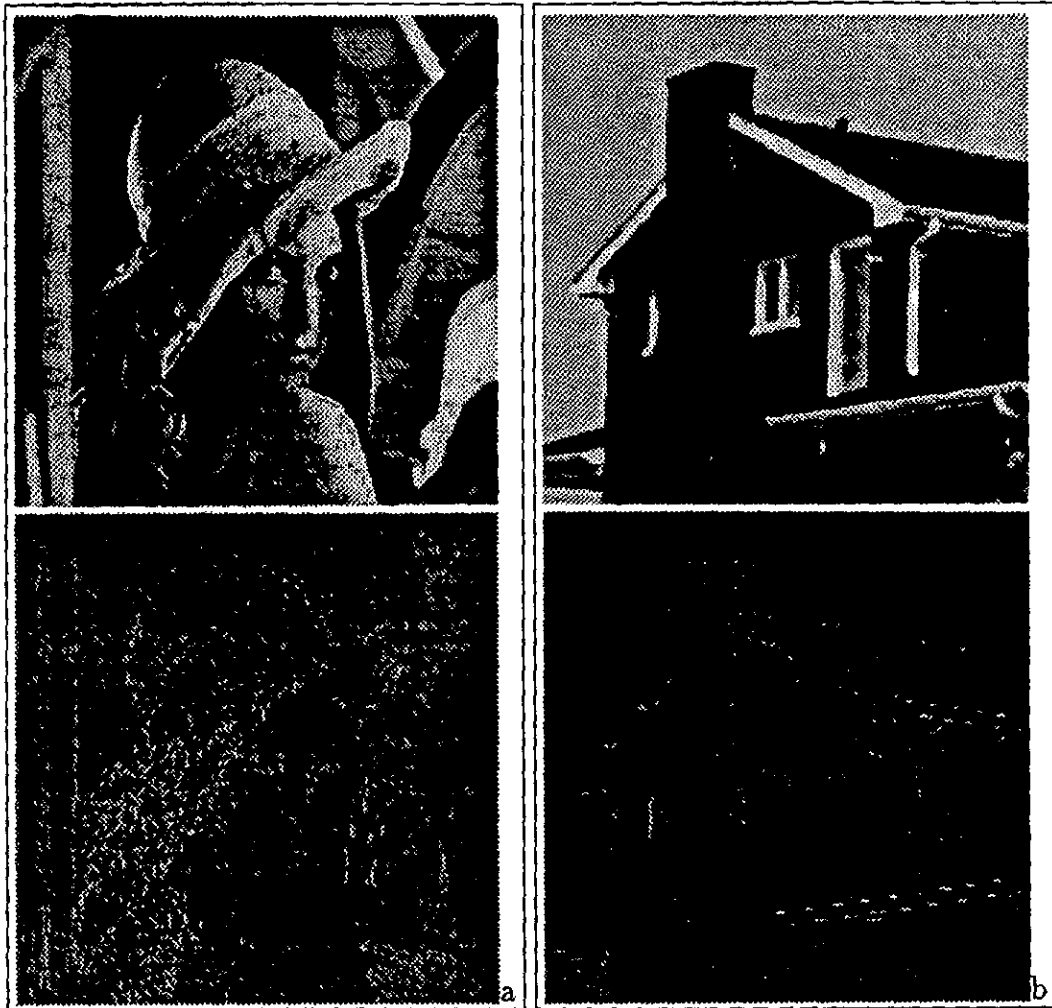


Figura 5.16: Reconstrucción a partir de la representación en parámetros característicos que se observa en la figura 5.15. Se tiene en la parte superior la reconstrucción y abajo de ella su correspondiente error respecto de la imagen original. (a) Imagen Lena y (b) imagen Casa



Esta reducción de datos tiene su efecto en la reconstrucción, en la figura 5.16 donde se aprecian las imágenes reconstruidas (a) Lena y (b) Casa. La degradación no es muy evidente aún, y prueba de ello es el valor de relación señal a ruido que se tiene 31.99 y 34.89 dB para las imágenes Lena y Casa respectivamente. Las imágenes reconstruidas presentan una ligera segmentación que es más clara en la imagen Lena. Con tamaño de ventana  $M = 4$  y  $T = 4$ , se obtiene la representación en parámetros característicos que se ilustra en la figura 5.17, para (a) imagen Lena y (b) imagen Casa. La diferencia notable con la representación presentada anteriormente, es el muestreo  $T$ . Aquí la compresión es mayor, por lo cual la entropía es de 1.08 y 0.98 bpp, lo que resulta en una tasa de compresión de 7 y 8 veces para Lena y Casa respectivamente, después de pasar por un codificador Huffman.

La reconstrucción tiene mayor degradación, y se empiezan a notar artefactos debidos al submuestreo que se efectúa en la imagen. Se ilustran las reconstrucciones en la figura 5.18, (a) imagen Lena y (b) imagen Casa. El efecto de bloques empieza a hacerse más notorio en estas reconstrucciones, afectando principalmente a la imagen Lena. La relación señal a ruido es 27.12 y 29.91 dB para Lena y Casa respectivamente.

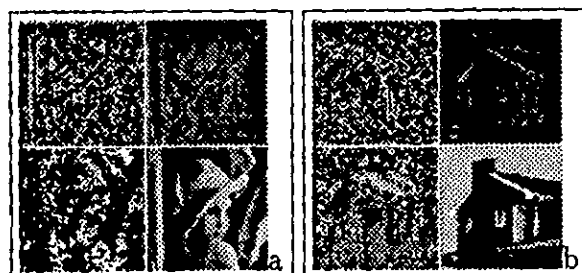


Figura 5.17: Representación en parámetros característicos de bordes, con  $M = 4$  y  $T = 4$  de las imágenes (a) Lena y (b) Casa. En la parte superior de las dos figuras se encuentra la imagen con valores de posición  $x_b$  y a su derecha la imagen de magnitud de los bordes contenidos en la imagen  $\Delta f$ . Abajo se tienen las imágenes de orientaciones  $\theta_{optima}$  y a la derecha el valor medio  $A_b$ .

En la figura 5.19 se observa la representación en parámetros característicos de las imágenes (a) Lena y (b) Casa, con tamaño de ventana  $M = 8$  y muestreo  $T = 8$ . La cantidad de muestreo que se aplica en esta representación es lo suficiente para lograr una compresión alta. La entropía del conjunto es de 0.27 y 0.24 bpp.

La reconstrucción de las imágenes que se presenta en la figura 5.20, (a) Lena y (b) Casa, tiene una relación señal a ruido 21.08 y 23.78 dB respectivamente. Es decir que la calidad subjetiva de la imagen deja mucho que desear. esto es evidente en las figuras donde se observa claramente un efecto de bloques en la imagen Lena y en la imagen Casa se presenta en los bordes diagonales principalmente.

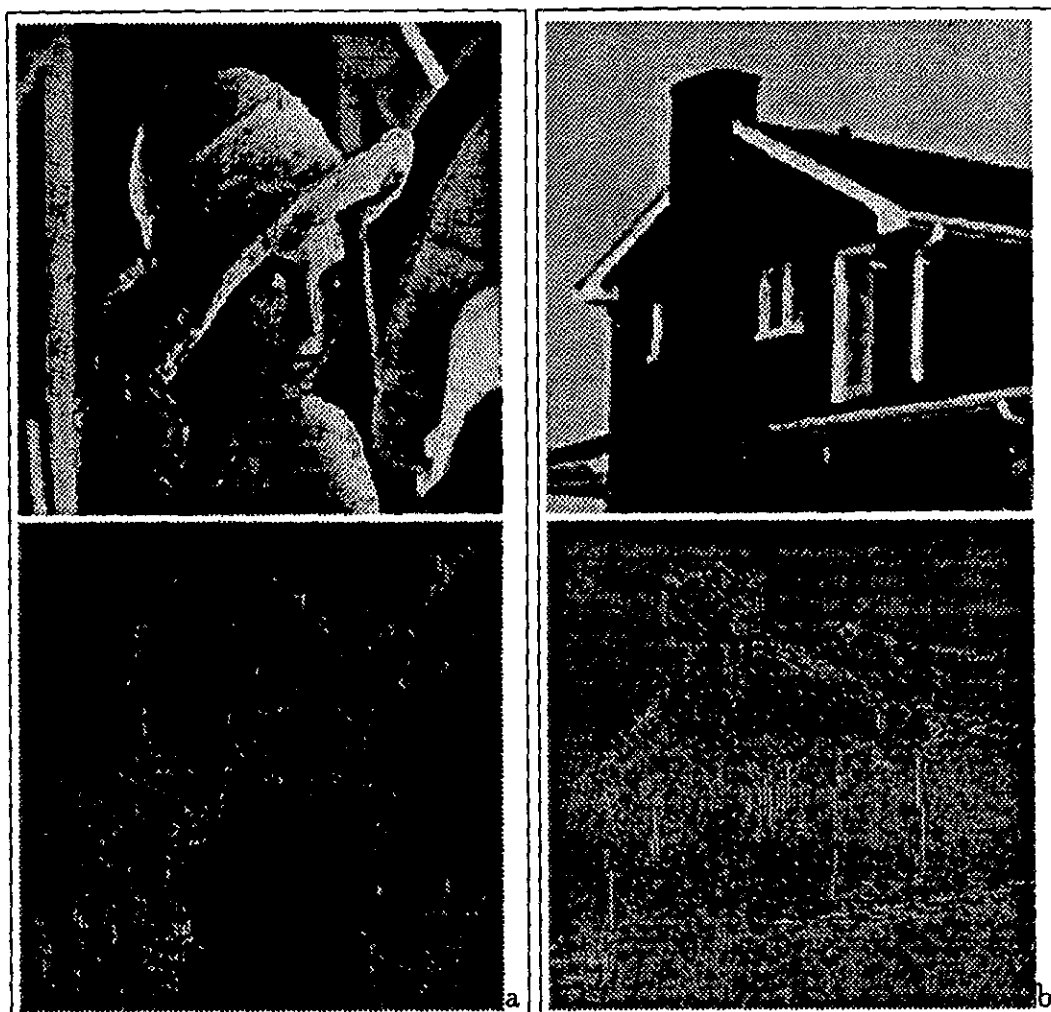


Figura 5.18: Reconstrucción de las imágenes (a) Lena y (b) Casa, a partir de sus respectivas representaciones que se muestran en la figura 5.17. Arriba se tiene la reconstrucción y abajo su correspondiente imagen de error.

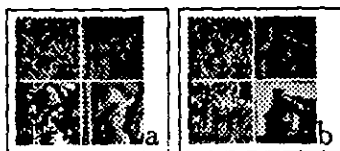


Figura 5.19: Representación en parámetros característicos de bordes, con  $M = 8$  y  $T = 8$  de las imágenes (a) Lena y (b) Casa. En la parte superior de las dos figuras se encuentra la imagen con valores de posición  $x_b$  y a su derecha la imagen de magnitud de los bordes contenidos en la imagen  $\Delta f$ . Abajo se tienen las imágenes de orientaciones  $\theta_{optima}$  y a la derecha el valor medio  $A_b$ .

Los acercamientos a las reconstrucciones de la imagen Lena que se muestran en la figura 5.21, tienen un factor de aumento 2. Con ellos se puede observar el efecto del incremento en la tasa de muestreo en el análisis de la imagen. De los tres acercamientos el que presenta la mejor calidad subjetiva es (a) cuando  $M = 2$  y  $T = 2$ . Cuando el tamaño de la ventana aumenta a  $M = 4$  y  $T = 4$ , empiezan a aparecer artefactos en la imagen. La mayor degradación se observa en la reconstrucción (c), que se obtiene de un tamaño de ventana  $M = 8$  y  $T = 8$ , aquí es notorio el efecto de bloqueo y la aparición de artefactos es mayor.

### 5.3.4 Gráficas de desempeño, para el algoritmo en una escala

Las siguientes figuras, corresponden a gráficas que se presentan a manera de resumen del desempeño de los algoritmos propuestos para una escala de análisis. Se observa que para cada curva existe un identificador que denota cada una de las representaciones logradas: **2D** es la representación bidimensional, **1D** es la representación unidimensional y **PC** la representación en parámetros característicos.

En la figura 5.22 se tienen gráficas de desempeño para las imágenes (a) Lena y (b) Casa. En ellas se puede observar una medida de la calidad de la reconstrucción con la relación señal a ruido pico (PSNR) en cada variación del tamaño de la ventana  $M$ , recordando que en cada  $M$ , el valor del muestreo es  $T = M$ . Se observa una mayor calidad en la reconstrucción de la imagen Casa, esto se debe al contenido de estructuras principalmente bordes en la imagen.

Las gráficas que se observan en la figura 5.23, corresponden a los valores de entropía para cada representación con diferentes valores de  $M$  en una escala; (a) corresponde a la imagen Lena y (b) a la imagen Casa. Se observa que las curvas son continuamente decrecientes, esto quiere decir que se presenta una disminución de la entropía o de la cantidad de información contenida en el conjunto, cuando existe variación en el

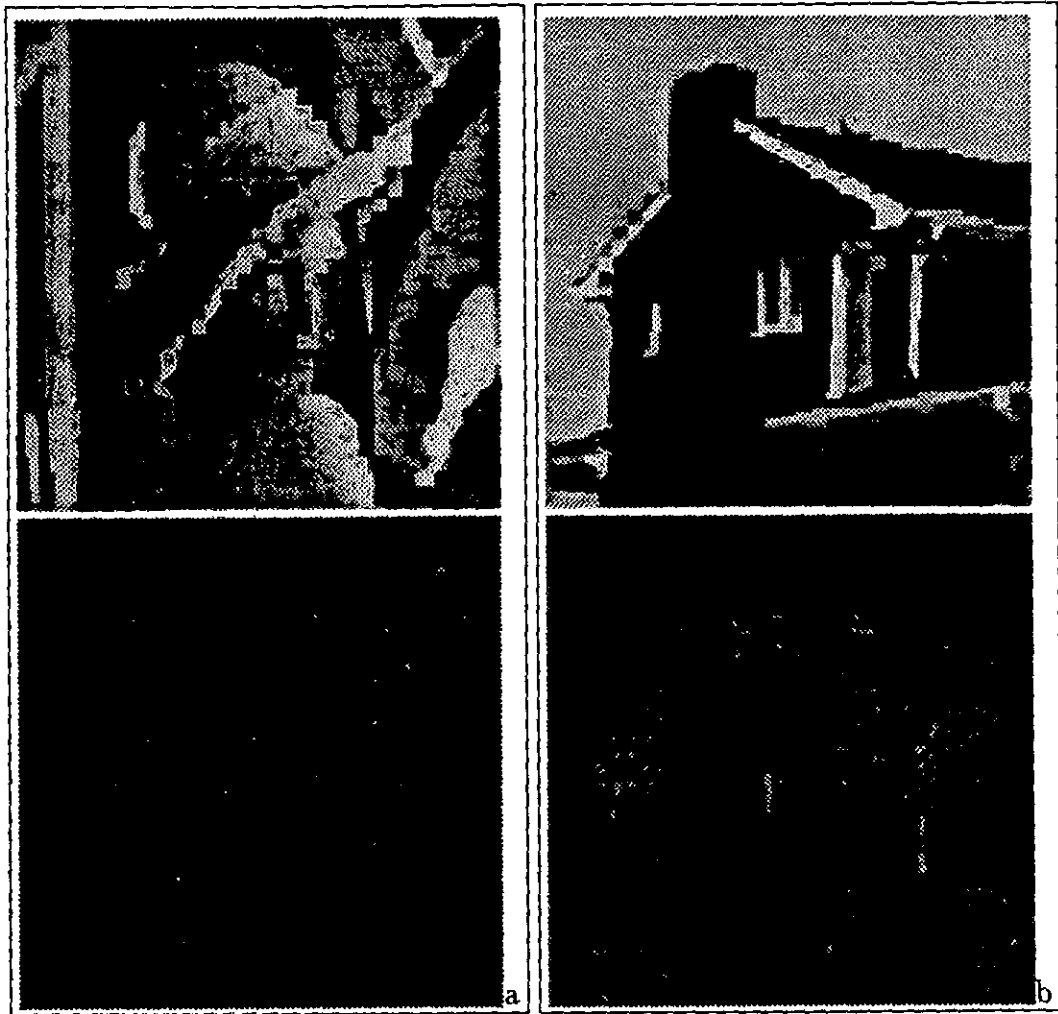


Figura 5.20: Reconstrucción de las imágenes (a) Lena y (b) Casa, con su respectiva imagen de error. Se recuperaron a partir de las representaciones mostradas en la figura 5.19.



Figura 5.21: Acercamiento a una parte de la imagen Lena reconstruida de las representaciones en parámetros característicos con (a)  $M = 2, T = 2$ , (b)  $M = 4, T = 4$  y (c)  $M = 8, T = 8$ . Aumentada 2 veces.

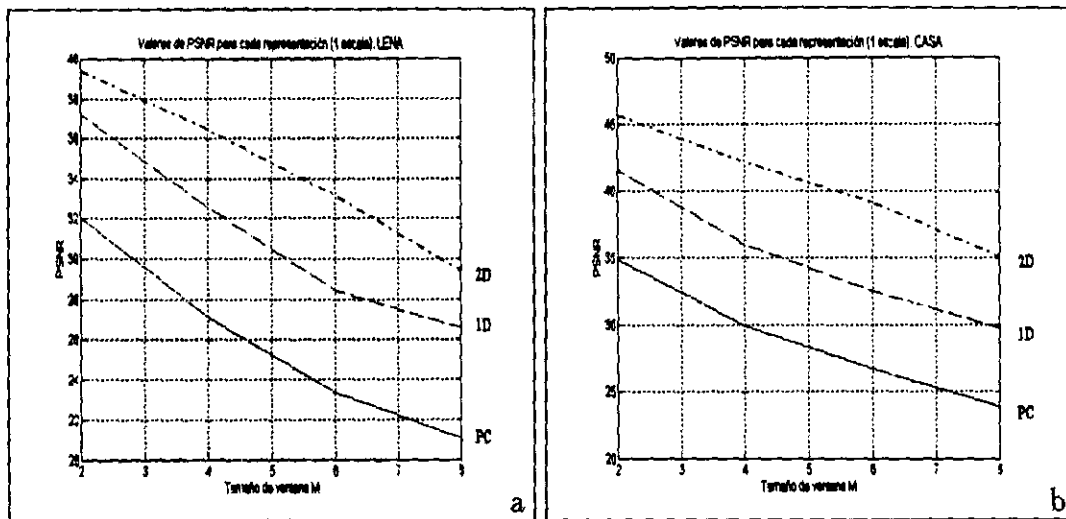


Figura 5.22: Gráficas que corresponden a los valores de relación señal a ruido (PSNR) en cada representación de las imágenes (a) Lena y (b) Casa, para diferentes tamaños de ventana  $M$ . Análisis realizado en una escala en las representaciones bidimensional, unidimensional y en parámetros característicos (2D, 1D y PC respectivamente).

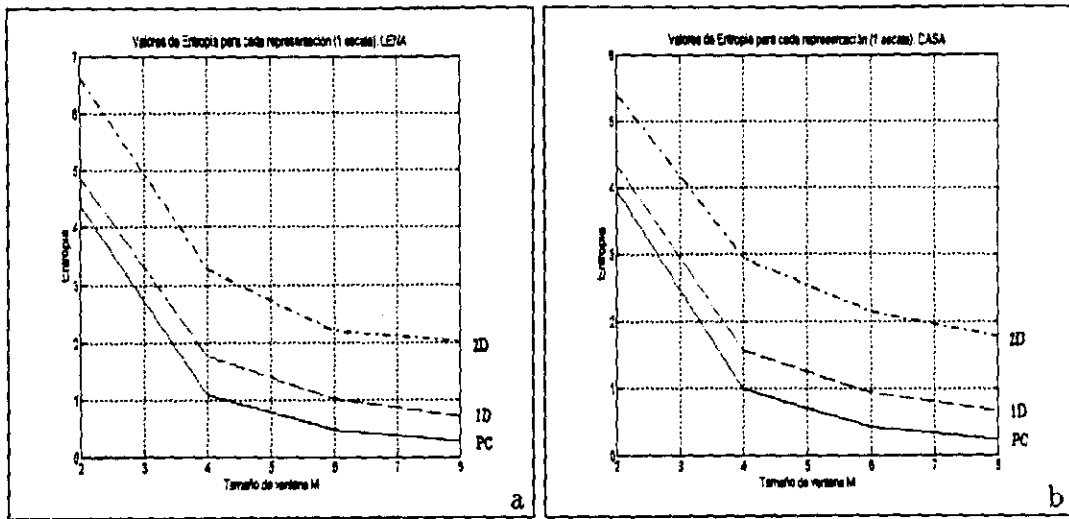


Figura 5.23: Valores de Entropía en cada representación de las imágenes (a) Lena y (b) Casa con diferentes tamaños de ventana  $M$ . El análisis se realizó en una escala para las representaciones 2D, 1D y PC.

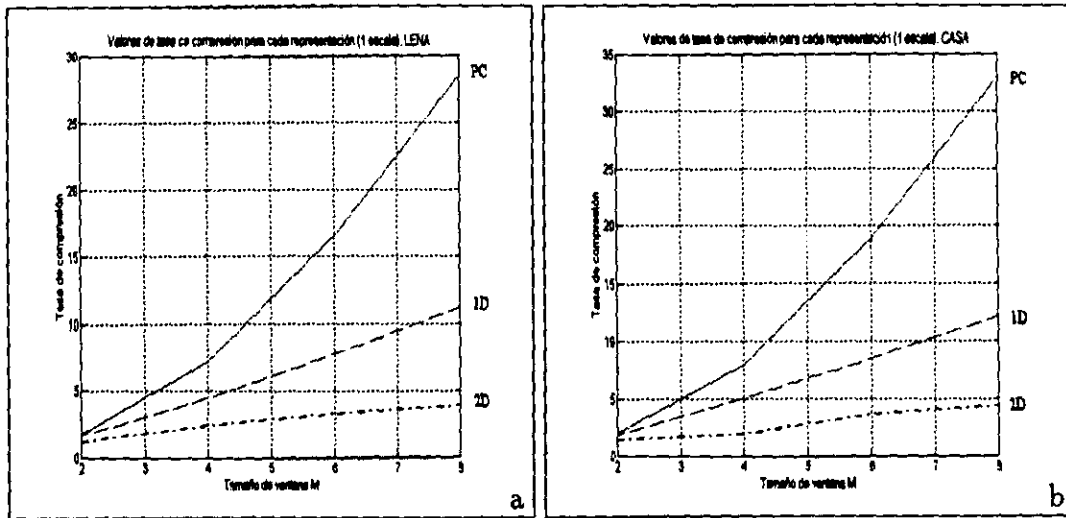


Figura 5.24: Tasa de compresión en cada representación de las imágenes (a) Lena y (b) Casa, con diferentes tamaños de ventana  $M$ . El análisis se realizó en una escala para las representaciones 2D, 1D y PC. La tasa de compresión se obtiene después de realizar codificación de Huffman.

tamaño de ventana  $M$  y por consiguiente en el parámetro de muestreo  $T$ .

La tasa de compresión calculada como se indicó antes, presenta una curva continuamente creciente. Esto se deduce por el valor de entropía pues aunque la tasa de compresión se calcula a partir del número de bytes que se entregan al final de la codificación de Huffman [Rab91], esta se acerca a los valores obtenidos por la entropía. Se presentan los valores para las imágenes Lena en la figura 5.24a y Casa en la figura 5.24b. La mayor tasa de compresión se alcanza con una ventana muy grande. Sin embargo y regresando a la figura 5.23, se observa la menor calidad de reconstrucción.

## 5.4 Algoritmos de codificación en múltiples escalas

La transformada polinomial es una técnica escala-espacio, que utiliza ventanas de análisis de diferente tamaño para realizar la localización de la imagen.

Si se utiliza una ventana binomial, entonces en la variación del tamaño  $M$  va implícito un cambio de la varianza de la función. Sin embargo, esto es poco práctico para la implantación en computadora. De tal manera, que lo adecuado es variar el tamaño de la señal mediante muestreo  $T$ , y así reducir el área de análisis. Con ello se abarcan estructuras cuya varianza sea diferente. Dicho en otras palabras, si se disminuye la señal de análisis, equivale a aumentar el tamaño de la ventana. De esta manera se detectan estructuras muy suaves, es decir con una varianza muy grande. Se presentan los resultados para representaciones con dos y tres niveles de análisis. Al final un estudio comparativo de las dos representaciones. El análisis en tres niveles se efectúa solo con ventanas de tamaño  $M = 2, 4$ , debido a que el muestreo  $T$  es el máximo permitido por la función ventana.

### 5.4.1 Representación bidimensional

La representación bidimensional en múltiples escalas permite realizar un análisis más completo del contenido de estructuras. Se utilizan dos y tres niveles para el análisis.

#### Dos niveles de análisis

En la figura 5.25 se observa la representación bidimensional en dos escalas de la imagen Lena, cuyo valor de entropía es de 6.59 bpp. Se utilizó una ventana de tamaño  $M = 2$  y muestreo  $T = 2$ . Con el conjunto de coeficientes del segundo nivel, se reconstruye el coeficiente de orden cero en el primer nivel, queda entonces la representación bidimensional completa para ese nivel y se puede reconstruir la imagen original.

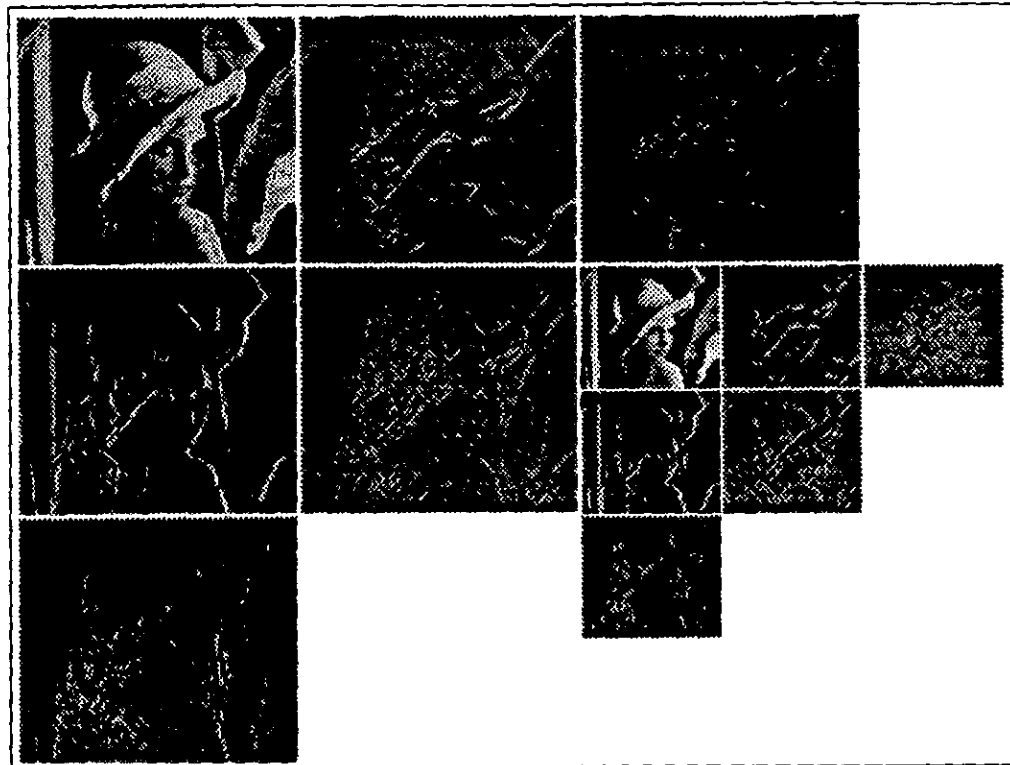


Figura 5.25: Representación polinomial bidimensional en dos escalas de la imagen Lena con tamaño de ventana  $M = 2$  y  $T = 2$ . Los coeficientes se encuentran dispuestos de la siguiente manera: en la fila superior  $f_{0,0,0}$ ,  $f_{0,0,1}$ ,  $f_{0,0,2}$ , en la siguiente fila  $f_{0,1,0}$ ,  $f_{0,1,1}$  y el conjunto de imágenes que pertenece a los coeficientes del siguiente nivel, en el mismo orden  $f_{1,0,0}$ ,  $f_{1,0,1}$ ,  $f_{1,0,2}$ , abajo de ellos  $f_{1,1,0}$ ,  $f_{1,1,1}$ ; en la fila inferior se encuentran los coeficientes  $f_{0,2,0}$ ,  $f_{1,2,0}$ .



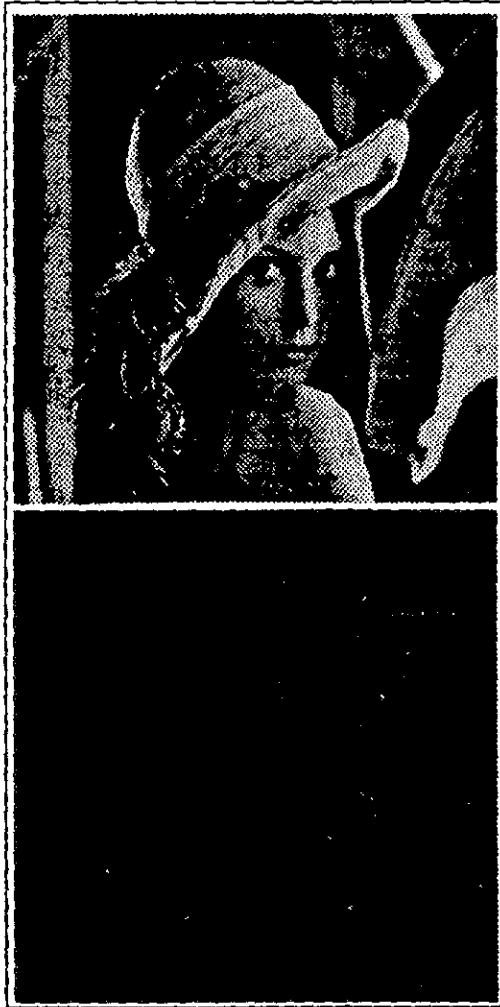


Figura 5.26: Reconstrucción de la imagen Lena a partir de la representación bidimensional de la figura 5.25. Arriba se tiene la imagen reconstruida y abajo de ella el error de reconstrucción.

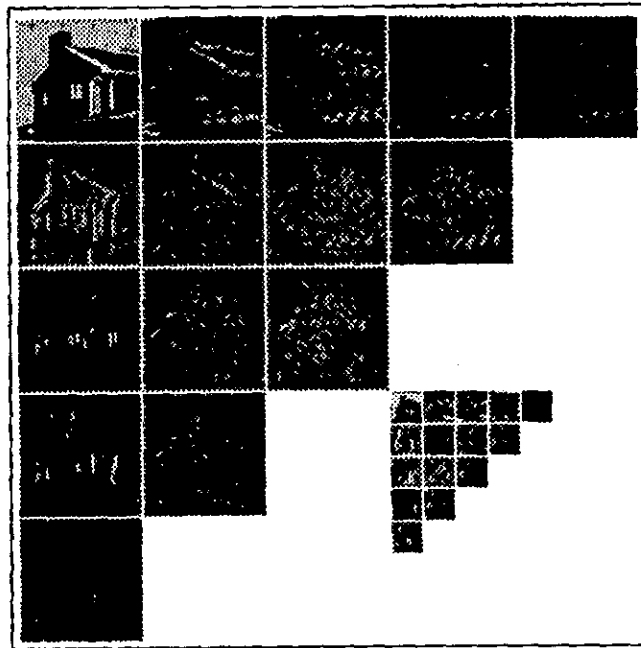


Figura 5.27: Se observan dos conjuntos de imágenes que corresponden a la representación bidimensional de la imagen Casa en dos escalas con tamaño de ventana  $M = 4$  y  $T = 4$ . Los coeficientes se encuentran dispuestos de la misma manera para los dos conjuntos, formando una matriz triangular superior.

La calidad de la reconstrucción de la imagen es muy buena. Perceptualmente hablando no se perciben errores, como se comprueba en la figura 5.26. La relación señal a ruido que se logra es muy alta, 100 dB.

En la representación bidimensional de la figura 5.27, se puede observar la disminución notoria del segundo conjunto de imágenes que forman el siguiente nivel. Se utilizó una ventana de tamaño  $M = 4$  y  $T = 4$ . Con el conjunto de imágenes que corresponde al segundo nivel se reconstruye el coeficiente de orden cero del primer nivel que se observa en la parte superior izquierda de la figura. La entropía para esta representación es de 2.6 bpp.

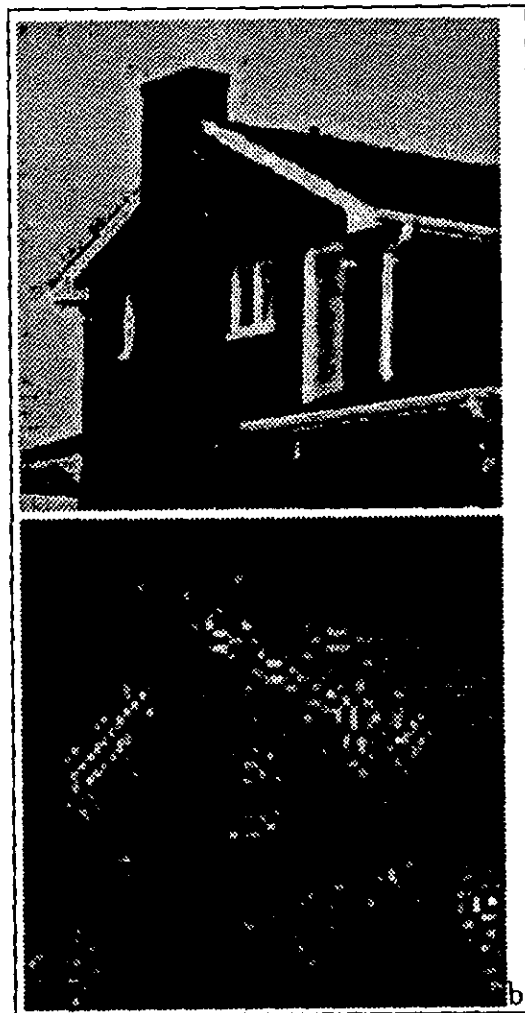


Figura 5.28: Reconstrucción de la imagen Casa a partir de la representación bidimensional de la figura 5.27. Arriba se tiene la imagen reconstruida y abajo de ella el error de reconstrucción.

La reconstrucción de la imagen Casa a partir de la representación bidimensional

de la figura 5.27 se observa en la figura 5.28, abajo el error de reconstrucción. Se presentan errores en la imagen principalmente en lugares donde la textura es muy suave. Se generan pérdidas a través de los dos niveles de análisis.

### Tres niveles de análisis

La figura 5.29 muestra la representación bidimensional en los tres niveles de análisis de la imagen Lena. El nivel más bajo presenta al coeficiente con el mayor contenido de energía, aún así la entropía es muy alta, 6.6 bpp.

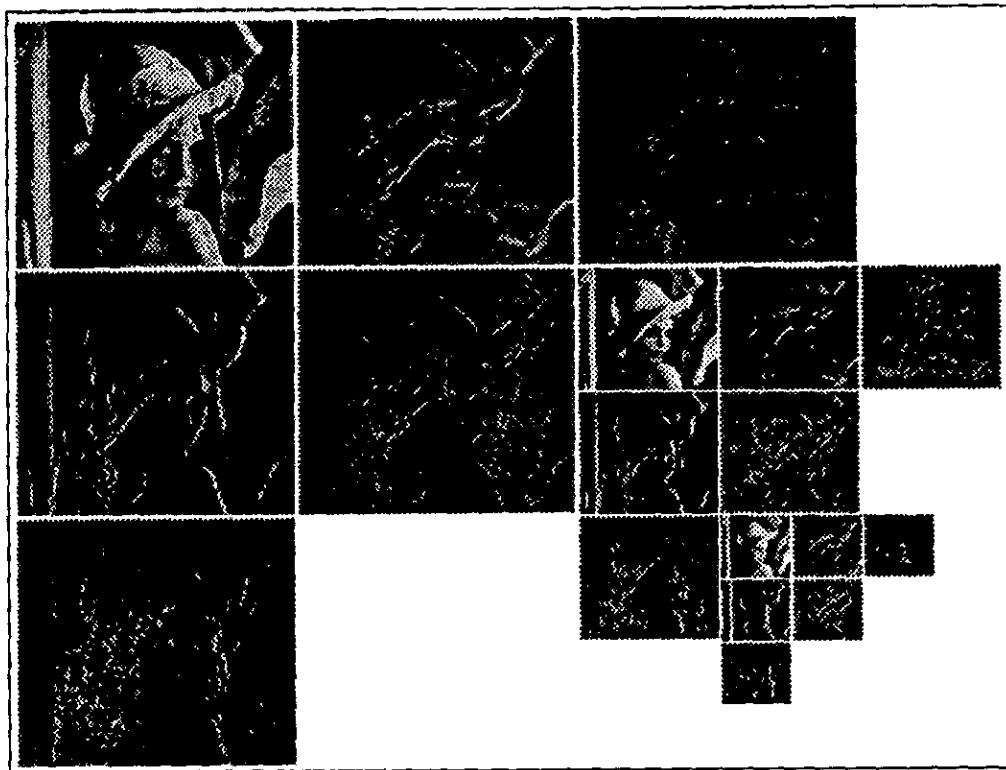


Figura 5.29: Representación bidimensional de la imagen Lena para tres escalas con  $M = 2$  y  $T = 2$ . Los coeficientes se encuentran dispuestos de la misma manera en los tres conjuntos:  $f_{0,0}$ ,  $f_{0,1}$ ,  $f_{0,2}$  en la primera fila,  $f_{1,0}$ ,  $f_{1,1}$  en la segunda fila y  $f_{2,0}$  al final.

La representación bidimensional que se muestra en la figura 5.30 corresponde a la imagen Casa, y se observan claramente los niveles de análisis. Se tiene en el tercer nivel al coeficiente de orden cero que tiene un mayor contenido de energía y es el que se conserva para la recuperación sucesiva de coeficientes de orden cero en los siguientes niveles.

En la figura 5.31, se observan las reconstrucciones de las imágenes (a) Lena a partir de la representación mostrada en la figura 5.29 y (b) Casa a partir de la repre-

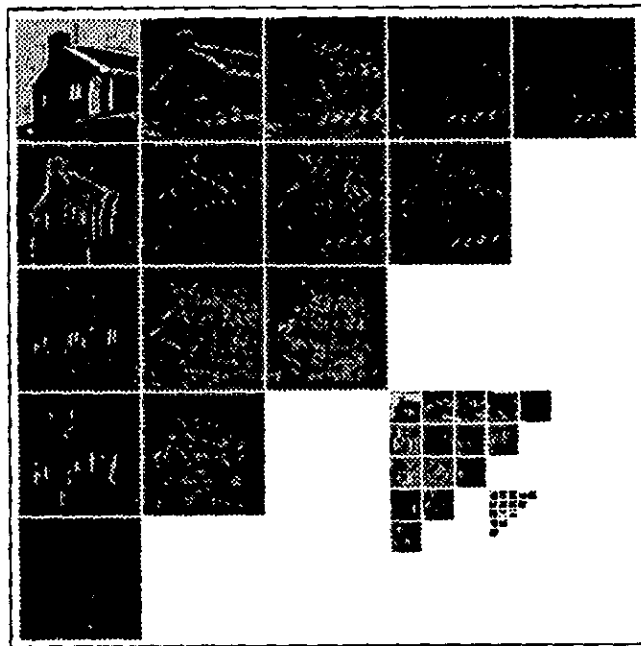


Figura 5.30: Representación bidimensional de la imagen Casa para tres escalas de análisis con  $M = 4$  y  $T = 4$ . Los coeficientes se encuentran ordenados de la misma forma para los tres conjuntos, en la primera fila  $f_{0,0}, f_{0,1}, f_{0,2}, f_{0,3}, f_{0,4}$ , la segunda fila  $f_{1,0}, f_{1,1}, f_{1,2}, f_{1,3}$ , la siguiente fila  $f_{2,0}, f_{2,1}, f_{2,2}$ , la penúltima fila  $f_{3,0}, f_{3,1}$  y la última fila  $f_{4,0}$ .

sentación mostrada en la figura 5.30. Es muy notorio el efecto de bloques que tiene la imagen Casa reconstruida, esto se debe principalmente al muestreo de la imagen y a la pérdida que existe entre los niveles.



Figura 5.31: Reconstrucción de las imágenes (a) Lena y (b) Casa a partir de las representaciones bidimensionales en tres escalas. Abajo de las imágenes se encuentra su imagen de error correspondiente.

#### 5.4.2 Representación unidimensional

A partir de la representación bidimensional se obtiene la proyección en coeficientes en una dimensión, que forman la representación unidimensional. Con ella se pretende reducir la cantidad de información redundante contenida en la representación anterior.

## Dos niveles de análisis

En dos niveles de análisis, la representación unidimensional de la imagen Lena se observa como en la figura 5.32. Los coeficientes de orden 1 y 2 se encuentran arriba y abajo la imagen de orientaciones y el siguiente nivel que tiene los elementos de la representación dispuestos de la misma forma para ese nivel. El coeficiente de orden cero se localiza en la parte inferior derecha de la figura. La entropía que se logra en esta representación es 4.43 bpp. El tamaño de ventana es  $M = 2$ .

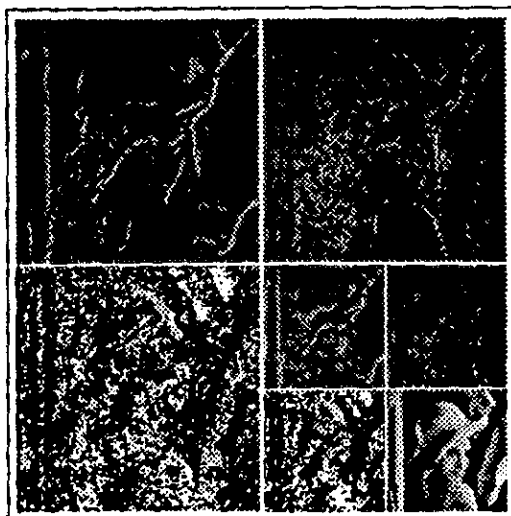


Figura 5.32: Representación polinomial unidimensional de la imagen Lena con  $M = 2$  y  $T = 2$ . En la parte superior se tienen los coeficientes  $K_{0,1}, K_{0,2}$ , abajo a la izquierda se encuentra la imagen de orientaciones  $\theta_0$ , a la derecha se tiene el conjunto de imágenes que corresponde al segundo nivel, los coeficientes  $K_{1,1}, K_{1,2}$  se encuentran en la parte superior y  $\theta_1$  y el valor medio  $A_b$  en la parte inferior.

La representación unidimensional para la imagen Casa con  $M = 4$  y  $T = 4$  se observa en la figura 5.33. De la misma manera que en la anterior aparecen los conjuntos de imágenes por nivel que forman la representación unidimensional.

A manera de demostración se presenta en la figura 5.34, la representación unidimensional de la imagen Lena con  $M = 8$  y  $T = 8$ . Se observa en lugar del coeficiente de orden cero del primer nivel, el conjunto de coeficientes que forman el segundo nivel de la representación. Con esto se ilustra una compresión de datos significativa.

La figura 5.35 contiene la reconstrucción de las imágenes Lena y Casa con su respectiva imagen de error. La figura 5.35a corresponde a la imagen Lena reconstruida de la representación con  $M = 2$  y  $T = 2$ . La figura 5.35b corresponde a la imagen Casa reconstruida de la representación con  $M = 4$  y  $T = 4$ . Es más notoria la degradación en la imagen Casa debido principalmente al submuestreo de la imagen.

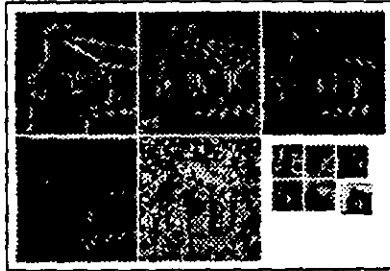


Figura 5.33: Representación polinomial unidimensional en dos escalas de análisis de la imagen Casa con tamaño de ventana  $M = 4$  y  $T = 4$ . Los coeficientes se encuentran dispuestos de la misma forma en los dos conjuntos. En la fila superior se tienen los coeficientes  $K_{0,1}$ ,  $K_{0,2}$ ,  $K_{0,3}$ , en la fila inferior se muestran  $K_{0,4}$ ,  $\theta_0$  y el conjunto de coeficientes del siguiente nivel  $K_{1,1}$ ,  $K_{1,2}$ ,  $K_{1,3}$ ,  $K_{1,4}$ ,  $\theta_1$ ,  $K_{1,0}$ .

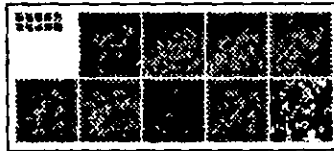


Figura 5.34: Representación polinomial unidimensional de la imagen Lena con  $M = 8$  y  $T = 8$ . El conjunto de imágenes del segundo nivel se observa en la parte superior izquierda.





Figura 5.35: Imágenes reconstruidas (a) Lena y (b) Casa, a partir de las representaciones unidimensionales mostradas en las figuras 5.32 y 5.33 respectivamente. Abajo se encuentra la imagen de error correspondiente.

### Tres niveles de análisis

En la figura 5.36 se observa la representación unidimensional de la imagen Lena con tamaño de ventana  $M = 2$  y muestreo  $T = 2$ , en tres escalas de análisis. La tasa de compresión lograda en esta representación es de 1.9 veces. La imagen se reconstruye a partir de las estimaciones del coeficiente de orden 0,  $K_{n,0}$ , para cada nivel partiendo del tercer nivel.

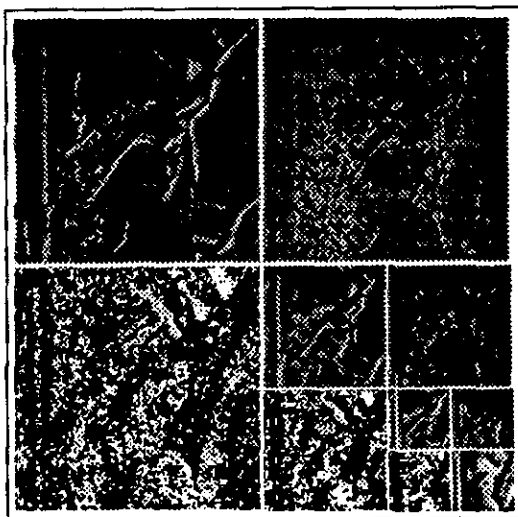


Figura 5.36: Representación polinomial unidimensional en tres escalas de la imagen Lena con  $M = 2$  y  $T = 2$ . Se observan tres conjuntos de imágenes, en la parte superior los coeficientes  $K_{0,1}$ ,  $K_{0,2}$ . En la parte inferior,  $\theta_0$  y los coeficientes que corresponden a los siguientes niveles. En la parte inferior derecha se muestra el coeficiente de orden 0 en el último nivel,  $K_{2,0}$ .

La representación unidimensional para la imagen Casa se observa en la figura 5.37, se usó una ventana de tamaño  $M = 4$  y muestreo  $T = 4$ . Se aprecia la diferencia en el tamaño de las imágenes de los tres conjuntos. La entropía que existe en la representación es de 1.28 bpp, esto habla de una pequeña compresión de la imagen. La reconstrucción de las imágenes, se observa en la figura 5.38. (a) Corresponde a la imagen Lena a partir de la representación mostrada en la figura 5.36 con valores  $M = 2$  y  $T = 2$ , abajo se observa la imagen de error con mayor cantidad de errores. Sin embargo, la imagen reconstruida se percibe con un efecto de suavizado. En (b) se presenta la reconstrucción de la imagen Casa a partir de la representación mostrada en la figura 5.37, con  $M = 4$  y  $T = 4$ ; contiene un efecto de bloques muy notorio, esto hace que la imagen de error también sea significativa. La relación señal a ruido de la reconstrucción es de 32.63 dB para la imagen Lena y 23.44 db para la imagen Casa.

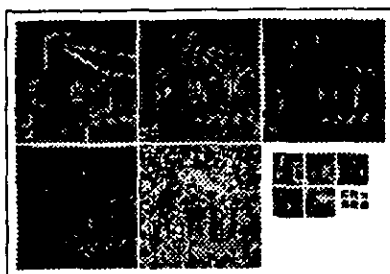


Figura 5.37: Representación polinomial unidimensional en tres escalas de la imagen Casa con  $M = 4$  y  $T = 4$ . Se observan los coeficientes que corresponden al primer nivel  $K_{0,1}$ ,  $K_{0,2}$ ,  $K_{0,3}$ , abajo  $K_{0,4}$  y  $\theta_0$ . En la siguiente imagen se observan los conjuntos de coeficientes que corresponden a los siguientes niveles de análisis.

### 5.4.3 Representación en parámetros característicos

En la representación unidimensional se tienen las estructuras primarias que componen a la imagen. Por lo tanto, se hace la extracción de parámetros que permitan conocer características de esas estructuras primarias. A esto se le llama representación en parámetros característicos.

#### Dos niveles de análisis

La figura 5.39, ilustra la representación en parámetros característicos en dos niveles de análisis de la imagen Lena. Se observa que la imagen de distancias  $x_b$  tiene pocos valores, por lo cual es altamente cuantizable.

La representación en parámetros característicos de la imagen Casa en la figura 5.40, tiene como principal característica el muestreo  $T = 4$ . La entropía de este conjunto es 0.637 bpp, lo que representa una compresión deseable.

La figura 5.41 muestra la representación en parámetros característicos de la imagen Lena, con un tamaño de ventana  $M \approx 8$  y  $T = 8$ . Con esta representación se logra un valor de entropía de 0.167 bpp, que equivale a una tasa de compresión de 47 veces la cantidad de bytes en la imagen. Sin embargo la calidad de la reconstrucción para este conjunto deja mucho que desear, tiene una relación señal a ruido de 14.74 dB.

En la figura 5.42a, se muestra la reconstrucción de la imagen Lena a partir de la representación en parámetros característicos de la figura 5.40. El error se aprecia abajo de la reconstrucción, la relación señal a ruido de esta imagen es 28.84 dB, esto quiere decir que a pesar de presentar cierta degradación la imagen tiene una aceptable calidad subjetiva.

En la figura 5.42b, se observa la reconstrucción de la imagen Casa a partir de la representación mostrada en la figura 5.41, se puede apreciar el efecto de bloques suavizados en la imagen. La relación señal a ruido es 23.76 dB.



Figura 5.38: Reconstrucción de las imágenes (a) Lena y (b) Casa a partir de las representaciones mostradas en las figuras 5.36 y 5.37 respectivamente. Abajo de ellas se encuentra su correspondiente imagen de error.

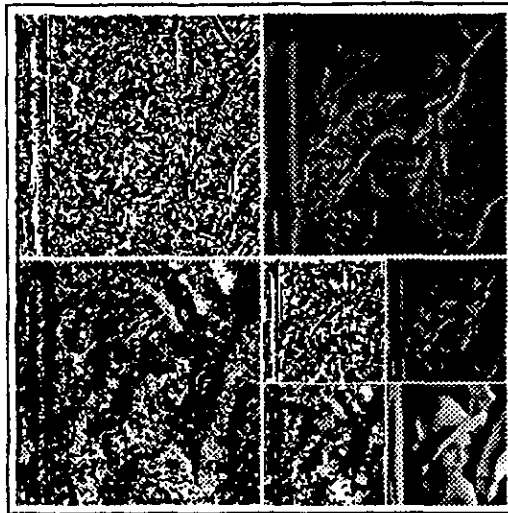


Figura 5.39: Representación en parámetros característicos de la imagen Lena en dos escalas de análisis con  $M = 2$  y  $T = 2$ . En la parte superior se observan las imágenes de distancia  $x_b$  y magnitud del borde  $\Delta f$ . Abajo se encuentra la imagen de orientaciones y el conjunto de imágenes que corresponden al siguiente nivel de análisis con la imagen de baja frecuencia  $A_b$ .

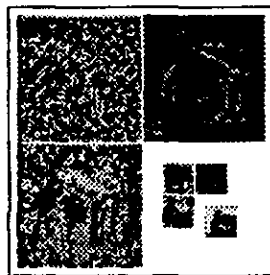


Figura 5.40: Representación en parámetros característicos de la imagen Casa en dos escalas de análisis con  $M = 4$  y  $T = 4$ . En la parte superior se observan las imágenes de distancia  $x_b$  y magnitud del borde  $\Delta f$ . Abajo se encuentra la imagen de orientaciones y el conjunto de imágenes que corresponden al siguiente nivel de análisis con la imagen de baja frecuencia  $A_b$ .



Figura 5.41: Representación en parámetros característicos de la imagen Lena en dos escalas de análisis con  $M = 8$  y  $T = 8$ . En la parte superior se observan las imágenes de distancia  $x_b$  y magnitud del borde  $\Delta f$ . Abajo se encuentra la imagen de orientaciones y el conjunto de imágenes que corresponden al siguiente nivel de análisis con la imagen de baja frecuencia  $A_b$ .

### Tres niveles de análisis

La representación en parámetros característicos que se observa en la figura 5.43, corresponde a la imagen Lena. Se realizó el análisis en tres niveles. La entropía total del conjunto es de 3.43 bpp.

La figura 5.44, ilustra la representación en parámetros característicos de la imagen Casa, en tres niveles con tamaño de ventana  $M = 4$  y muestreo  $T = 4$ . Se puede observar que la compactación de datos es mayor, y que el coeficiente de orden 0 es muy pequeño.

En la figura 5.45a se muestra la reconstrucción de la imagen Lena a partir de la representación de la figura 5.43. Tiene una buena calidad subjetiva aún cuando el valor de relación señal a ruido sea de 26.33 dB. De la misma manera que en la reconstrucción en dos niveles se observa segmentación de la imagen de manera muy escasa.

En la figura 5.45b se tiene la reconstrucción de la imagen Casa, a partir de la representación que se muestra en la figura 5.44. El efecto de bloques es más visible, sin embargo la relación señal a ruido es de 20.20dB.

### 5.4.4 Gráficas de desempeño, para el algoritmo en dos y tres escalas

A manera de resumen se presentan las gráficas de desempeño de los algoritmos a múltiples escalas, en dos y tres niveles de análisis. En las gráficas se pretende hacer un estudio comparativo que sirva como referencia en la evaluación del algoritmo. La manera de identificar cada curva se mencionó antes. Se tienen gráficas para las tres representaciones de las dos imágenes Lena y Casa.

#### Dos niveles de análisis

En la figura 5.46, se observan las gráficas de los valores de relación señal a ruido (PSNR). (a) Corresponde a las reconstrucciones de la imagen Lena, (b) a las de la



Figura 5.42: Reconstrucción de las imágenes (a) Lena y (b) Casa a partir de las representaciones mostradas en las figuras 5.39 y 5.40 respectivamente. Abajo se muestra la imagen de error.

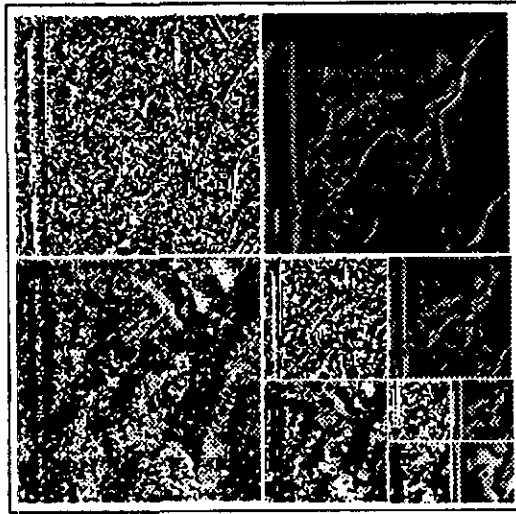


Figura 5.43: Representación en parámetros característicos de la imagen Lena en tres escalas con  $M = 2$  y  $T = 2$ . Se observan tres conjuntos de imágenes, en la parte superior las imágenes de distancia  $x_b$  y magnitud del borde  $\Delta L$ . En la parte inferior,  $\theta_0$  y las magnitudes y distancias de bordes en los siguientes niveles se encuentran a la derecha. En la parte inferior derecha se muestra la señal de baja frecuencia  $A_b$  en el último nivel.

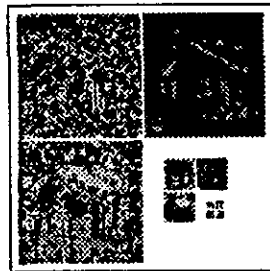


Figura 5.44: Representación en parámetros característicos de la imagen Casa en tres niveles para  $M = 4$  y  $T = 4$ . Se observan tres conjuntos de imágenes, en la parte superior las imágenes de distancia  $x_b$  y magnitud del borde  $\Delta L$ . En la parte inferior,  $\theta_0$  y las magnitudes y distancias de bordes en los siguientes niveles se encuentran a la derecha. En la parte inferior derecha se muestra la señal de baja frecuencia  $A_b$  en el último nivel.





Figura 5.45: Reconstrucciones de las imágenes (a) Lena y (b) Casa con sus respectivas imágenes de error debajo de ellas. Se obtienen a partir de las representaciones mostradas en las figuras 5.43 y 5.44, con valores de  $M = 2, 4$  y  $T = M$ .

imagen Casa. La curva que indica la representación bidimensional (2D), se encuentra recortada en las dos gráficas, esto es por el valor tan alto que alcanza cuando  $M=2$ . En las dos gráficas se observa que la mayor PSNR lograda es en la representación bidimensional y que se obtiene mejor calidad en la reconstrucción con la imagen Casa que con Lena.

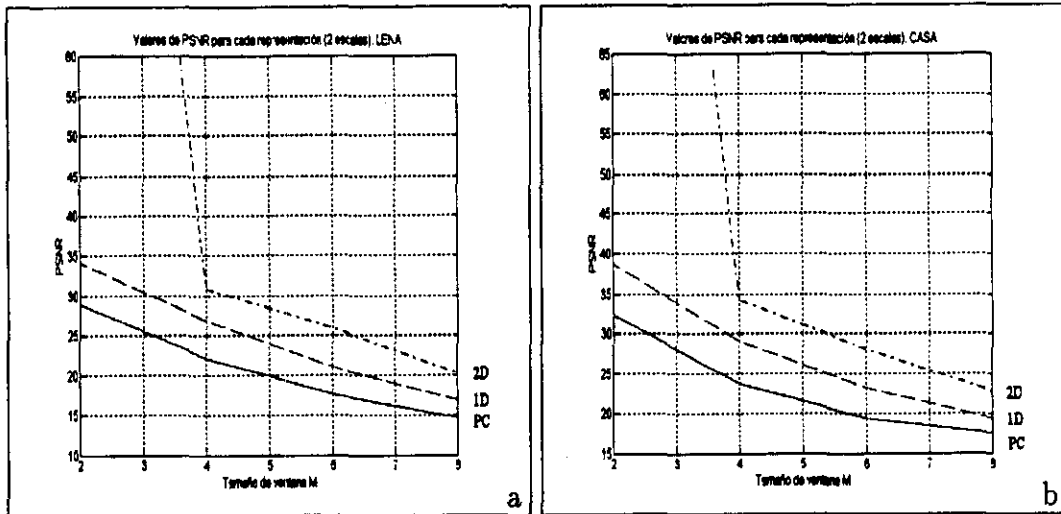


Figura 5.46: Gráficas que corresponden a los valores de relación señal a ruido (PSNR) en cada representación de las imágenes (a) Lena y (b) Casa, para diferentes tamaños de ventana  $M$ . Análisis realizado en dos escalas en las representaciones bidimensional, unidimensional y en parámetros característicos (2D, 1D y PC respectivamente).

La entropía disminuye de manera significativa con el aumento del tamaño de la ventana, esto se muestra en la figura 5.47, donde se presentan las gráficas de valores de entropía en cada representación para las representaciones de (a) Lena y (b) Casa en dos escalas. Los valores de entropía más bajos los tiene la representación en parámetros característicos (PC).

Como consecuencia de los cambios de representación se obtiene una cierta tasa de compresión. Esta se ilustra en las gráficas que se presentan en la figura 5.48. Corresponden a representaciones de (a) Lena y (b) Casa. Se observa en las dos gráficas un comportamiento similar para las tres curvas. La curva correspondiente a la representación en parámetros característicos, tiene unos valores de compresión muy altos, estos se incrementan con el tamaño de la ventana de análisis. Las representaciones bidimensional (2D) y unidimensional (1D), aunque también aumentan, no logran niveles de compresión muy altos.

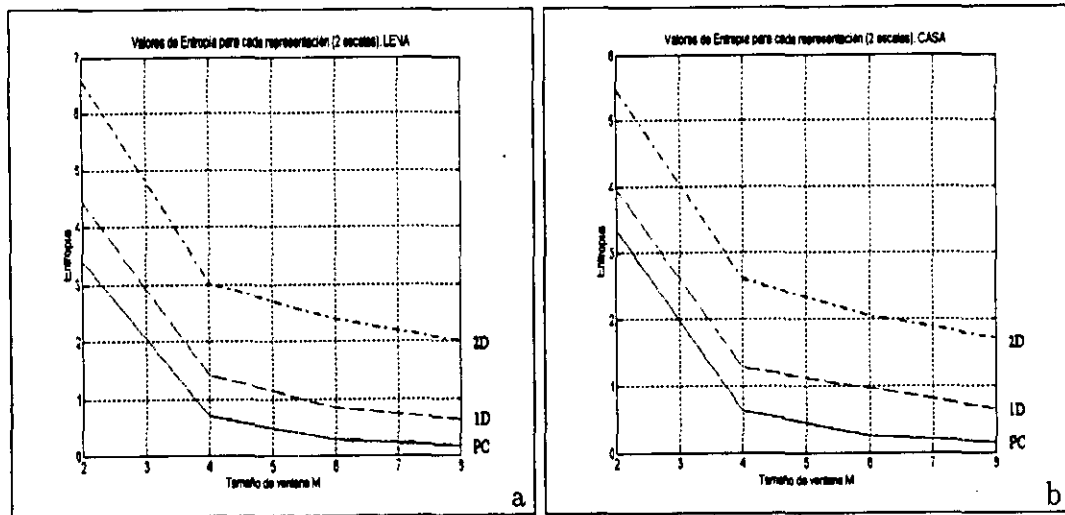


Figura 5.47: Valores de Entropía en cada representación de las imágenes (a) Lena y (b) Casa con diferentes tamaños de ventana  $M$ . El análisis se realizó en dos escalas para las representaciones 2D, 1D y PC.

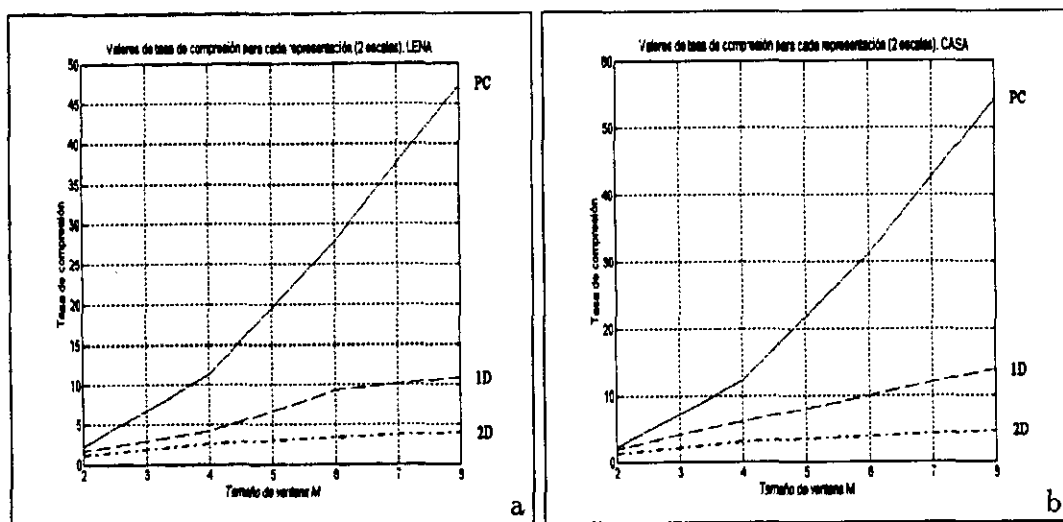


Figura 5.48: Tasa de compresión en cada representación de las imágenes (a) Lena y (b) Casa, con diferentes tamaños de ventana  $M$ . El análisis se realizó en dos escalas para las representaciones 2D, 1D y PC. La tasa de compresión se obtiene después de realizar codificación de Huffman.

### Tres niveles de análisis

Aquí se realizó la representación de las imágenes, solo en dos tamaños de ventana,  $M = 2, 4$ . Esto se debe al muestreo aplicado,  $T = M$ . De tal manera que se presentan gráficas con una línea para cada representación.

En la figura 5.49a, se tienen los valores de PSNR para cada representación de la imagen Lena en tres niveles y se observa recortada la línea en la representación bidimensional (2D), esto se debe al valor que alcanza la reconstrucción de la imagen, más de 100 dB. Sin embargo, las otras representaciones presentan una disminución significativa en la calidad de la reconstrucción de la imagen. En la figura 5.49b, se observan las curvas de PSNR correspondientes a la imagen Casa, en ellas se observa una disminución muy grande cuando existe un cambio en el tamaño de la ventana  $M$ . Sin embargo, la calidad no es muy mala, pues se ubica arriba de los 20 dB.

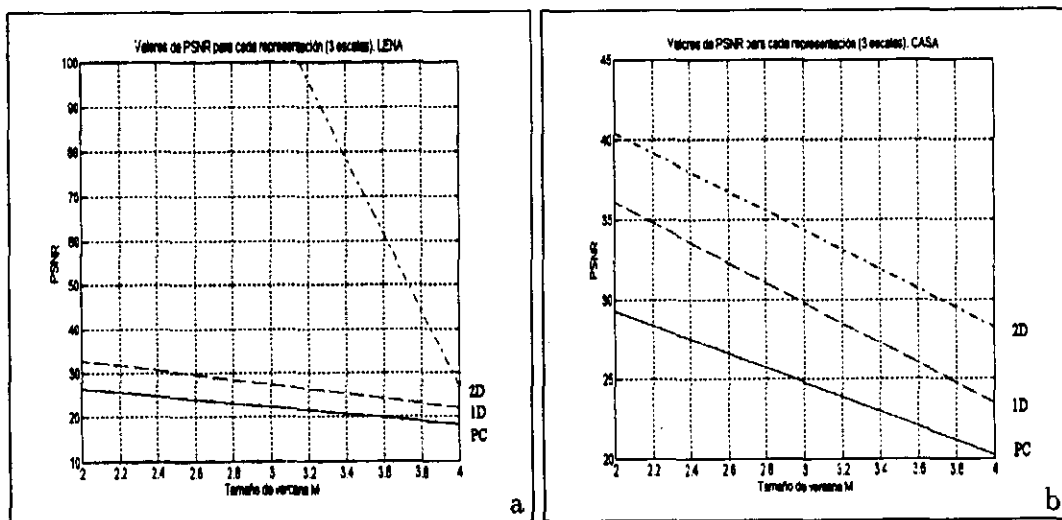


Figura 5.49: Gráficas que corresponden a los valores de relación señal a ruido (PSNR) en cada representación de las imágenes (a) Lena y (b) Casa, para diferentes tamaños de ventana  $M$ . Análisis realizado en tres escalas en las representaciones bidimensional, unidimensional y en parámetros característicos (2D, 1D y PC respectivamente).

En la figura 5.50, se presentan los valores de entropía para las representaciones en tres niveles de las imágenes (a) Lena y (b) Casa. En las dos se presenta una disminución importante de la entropía cuando se aumenta el tamaño de la ventana  $M$ .

En las gráficas de la figura 5.51, aunque parecen significativos los valores de la tasa de muestreo, son muy escasos para la calidad de reconstrucción que se obtiene. En (a), se observa que la representación en parámetros característicos (PC) es la que

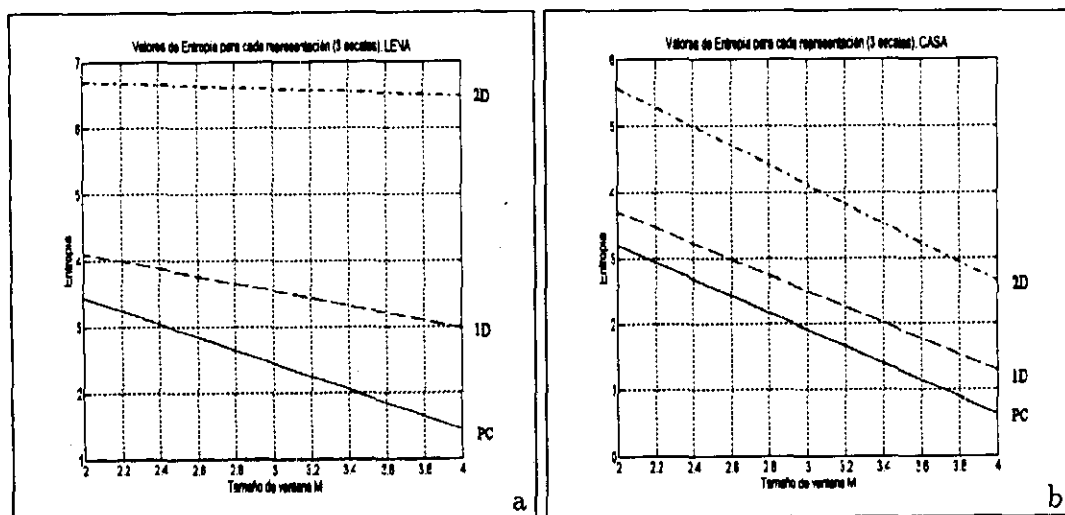


Figura 5.50: Valores de Entropía en cada representación de las imágenes (a) Lena y (b) Casa con diferentes tamaños de ventana  $M$ . El análisis se realizó en tres escalas para las representaciones 2D, 1D y PC.

mayor tasa de compresión alcanza. En (b), la gráfica se dispara en forma más significativa.

## 5.5 Comparación con JPEG

Ya se ha visto que las representaciones obtenidas a lo largo del proyecto presentan medidas de fiabilidad que son competitivas. Para demostrar esto se presenta una serie de imágenes que se obtuvieron a partir de la compresión con JPEG<sup>3</sup>. Se utilizó un paquete computacional llamado XVIEW, en Solaris de SUN.

El estándar JPEG para compresión de imágenes, utiliza la transformada coseno discreta (DCT) para realizar el análisis de imágenes [Rab91]. Los coeficientes que se generan, se cuantizan por medio de un algoritmo con valores fijos, y se determina de acuerdo a la calidad requerida de la imagen, cuales de esos coeficientes se usaran para la reconstrucción.

En las figuras 5.52 y 5.53, se observa la imagen Lena después de comprimirla con JPEG. El parámetro utilizado para efectuar la compresión es la calidad de reconstrucción de la imagen, que se indica en cada una de las figuras.

<sup>3</sup>En inglés: "Joint Photographic Experts Group".

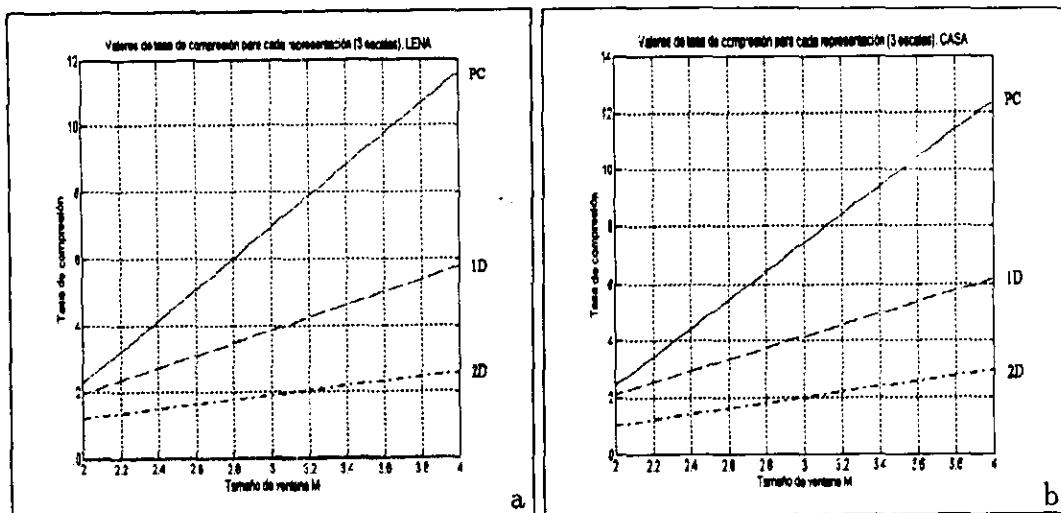


Figura 5.51: Tasa de compresión en cada representación de las imágenes (a) Lena y (b) Casa, con diferentes tamaños de ventana  $M$ . El análisis se realizó en tres escalas para las representaciones 2D, 1D y PC. La tasa de compresión se obtiene después de realizar codificación de Huffman.

Se puede observar que la calidad de reconstrucción es mala cuando la tasa de compresión aumenta. Para la imagen de la figura 5.53d, se tiene una tasa de compresión de 39.5 veces y la relación señal a ruido es de 22.142 dB. Si se compara este resultado con el obtenido por el algoritmo de parámetros característicos de bordes que se presenta en la figura 5.20a, se observa que los valores no están muy alejados de los obtenidos con JPEG, ya que el valor de relación señal a ruido es de 21.089 dB y la tasa de compresión es de 28.5 veces.

Si bien, ciertamente los valores de las medidas de fiabilidad en los algoritmos propuestos son bajos comparados con el JPEG, también es cierto que al conjunto de algoritmos propuestos les faltan detalles para lograr ser totalmente óptimos.

En la figura 5.54, se puede apreciar las gráficas que se obtuvieron después de realizar la compresión con JPEG, para las distintas calidades de reconstrucción. Los valores que se muestran corresponden a relación señal a ruido pico (PSNR) y la tasa de compresión.



Figura 5.52: Imagen Lena recuperada después de la compresión con JPEG con diferente calidad de reconstrucción: (a) 100 %, (b) 75 %, (c) 50 % y (d) 25 %.



Figura 5.53: Imagen Lena recuperada después de la compresión con JPEG con diferente calidad de reconstrucción: (a) 10 %, (b) 5 %, (c) 3 % y (d) 1 %.



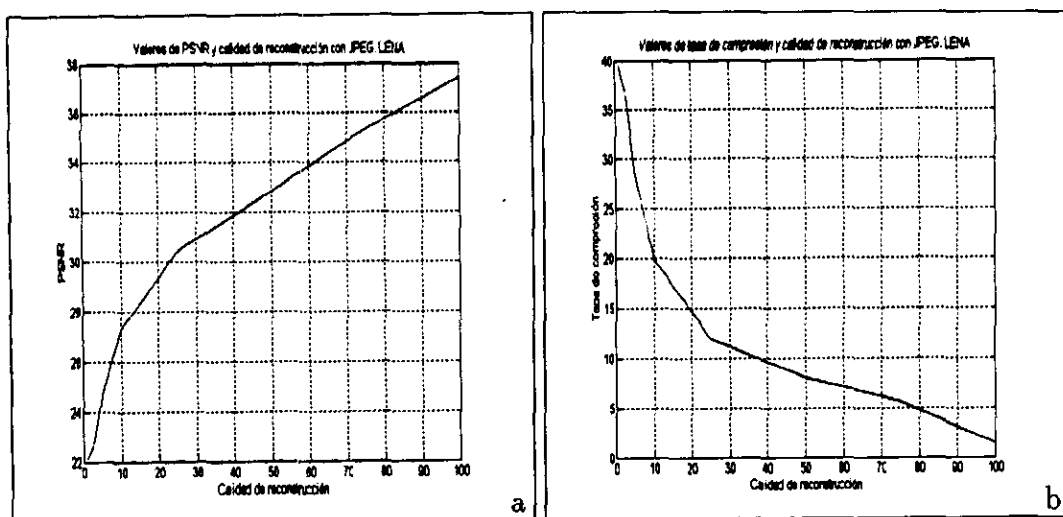


Figura 5.54: Gráficas de desempeño del JPEG aplicado sobre la imagen Lena. (a) Corresponde a los valores de PSNR, (b) a los valores de tasa de compresión.

## 5.6 Conclusiones

Se presentaron los resultados que se obtienen al aplicar el algoritmo de codificación propuesto a las imágenes Lena y Casa. Con las variaciones de algunos valores como el tamaño de la ventana y tomando el máximo muestreo posible  $T = M$ . También se realizó la variación en la cantidad de escalas, desde una hasta tres escalas de análisis.

Las características de las imágenes permiten realizar una evaluación subjetiva de las reconstrucciones logradas y determinar en que condiciones el algoritmo tiene un mejor comportamiento. Se eligieron las imágenes Casa y Lena por las características que presentan, gran cantidad de estructuras perceptualmente importantes y frecuencias altas, principalmente en la imagen Lena.

Se utilizaron como medidas de fiabilidad, la entropía, la relación señal a ruido y la tasa de muestreo. Con ellas se da un panorama general del funcionamiento del algoritmo con una imagen específica. A partir de dicho panorama en forma de gráficas, se puede realizar una evaluación. Se observa en las gráficas que cuando la entropía decrece, la calidad de la reconstrucción decrece también.

La representación bidimensional se presenta para fines de comparación, ya que no es la codificación el objetivo principal. Se usa principalmente en casos de restauración, realce y reducción de ruido. Sin embargo, se lograron resultados con cierto nivel de

compresión y las mejores reconstrucciones de la imagen.

La representación unidimensional, se usa principalmente en restauración de imágenes, sin embargo, se obtienen tasas de compresión y calidad de reconstrucción que es aceptable, excepto cuando se trata de un muestreo muy alto.

La representación en parámetros característicos de estructuras unidimensionales tiene como principal objetivo la codificación, por medio de la eliminación de redundancia estadística y psicovisual. Si a ello aunamos una tasa de muestreo alta, entonces tenemos tasas de compresión altas también. El problema que surge aquí es la calidad perceptual de la reconstrucción de la imagen, que es menor conforme aumenta la tasa de compresión.

La tasa de compresión a la cual el algoritmo funciona mejor, depende en gran medida del contenido de estructuras de la propia imagen. Si la imagen esta formada principalmente de estructuras bidimensionales, entonces la tasa de compresión no puede ser muy alta, de otra manera se degrada de manera importante la imagen.

En las pruebas realizadas con algoritmos en múltiples resoluciones, se observan degradaciones muy importantes cuando la tasa de muestreo y el tamaño de la ventana aumentan, esto se produce debido a las pérdidas que existen en la recuperación de la señal de baja frecuencia en cada nivel. Para resolverlo se proponen algunas ideas que se mencionan en el siguiente capítulo.

Se presentó una serie de imágenes comprimidas con JPEG, para comparar el comportamiento del algoritmo, se deduce que la propuesta tiende a ser un eficiente algoritmo de compresión siempre y cuando se logren realizar ciertas tareas como la cuantización de manera optima.

# Capítulo 6

## Perspectivas y conclusiones generales

### 6.1 Perspectivas

Como ya se mencionó en los primeros capítulos de este trabajo, los métodos de codificación y compresión de imágenes están en constante evolución para satisfacer las necesidades de almacenamiento, transmisión, capacidad de canal, etcétera. Razón por la cual, el diseño e implantación de este algoritmo tiene la posibilidad de continuar mejorando si se toman en cuenta tres aspectos importantes:

1. La compresión de datos.
2. La calidad de la reconstrucción.
3. La robustez del algoritmo.

Y ¿cómo se puede mejorar la tasa de compresión de datos?, la reducción de datos se puede lograr de varias formas con el uso de la transformada polinomial, como se describe a continuación:

- Realizando muestreos altos con grandes ventanas de análisis. Se ha demostrado que así se logra comprimir con una tasa mayor.
- Utilizando los algoritmos a múltiples escalas. Tratando además, de reducir el error de recuperación que se presenta en cada escala y que se acumula a través de los niveles de la pirámide.
- Eliminando la redundancia que existe en una representación a múltiples niveles mediante un algoritmo semejante al laplaciano. Esto es, transmitiendo sólo los errores de la representación en los niveles más altos. Si suponemos que los niveles más bajos son prácticamente un muestreo de los superiores, entonces el error es muy pequeño y la tasa de compresión muy alta.

- Identificando estructuras a través de una representación a múltiples niveles. Es decir, dado que se cuenta con una representación en parámetros de estructuras primarias (bordes y líneas), se puede realizar un seguimiento de ellas en cada nivel de la pirámide. Se trata de agrupar los pixels de cada estructura, conservando los parámetros que son similares y determinando donde se encuentran en cada nivel. De tal manera, que exista una especie de liga entre ellos [Mal92a]. Esto es posible, ya que la teoría escala espacio muestra que las estructuras desaparecen después de un cierto número de niveles y se pueden unir a otras [Wit84].
- Con algoritmos de cuantización que permitan un comportamiento adaptivo. Un ejemplo es la cuantización vectorial, ya que las representaciones que se obtienen se pueden utilizar para tales fines, pues los vectores pueden estar formados por los parámetros o los coeficientes de la representación.
- Finalmente el uso de orientaciones de una manera adaptiva. El número de orientaciones utilizado en los resultados expuestos fue de 8 para todas las representaciones. Sin embargo, se puede determinar el número de orientaciones conociendo el tamaño de la ventana de análisis. Por ejemplo, para una ventana de tamaño  $M = 2$ , se puede utilizar 4 orientaciones; ya que con tal número quedan localizados los pixels adyacentes. Para una ventana cuyo tamaño sea  $M = 4$ , son necesarias 8 orientaciones. Por lo tanto, se puede calcular el número de orientaciones con respecto al tamaño de la ventana de análisis  $M$ . Reduciendo la información contenida en la imagen de orientaciones óptimas cuando  $M$  es pequeña.

El problema de lo anteriormente expuesto es la calidad de la reconstrucción que ya presenta graves degradaciones cuando el algoritmo tiene una razón de compresión mayor a 30 veces. Sin embargo, se puede resolver de la siguiente manera:

- Con la búsqueda más eficiente de parámetros, que caracterizan a las estructuras primarias (bordes y líneas) y determinando los parámetros de otras estructuras como esquinas o puntos, que aunque no son tan frecuentes como las anteriores existen y es posible caracterizarlas.
- Obteniendo una imagen de error de reconstrucción, ya sea en una escala o en múltiples escalas. Esto evidentemente aumenta la cantidad de información.

El otro aspecto en el que se pueden realizar mejoras a los algoritmos propuestos es la robustez. Esta es una cuestión interesante cuando de codificar se trata. Como se puede observar en los resultados obtenidos, la imagen Lena consigue menores tasas

de compresión que la imagen Casa, esto se debe a las estructuras que se encuentran en cada una de las imágenes. Se puede lograr un cierto grado de robustez que permita lograr la reconstrucción de estructuras que no sean unidimensionales. Ya que si se presenta ruido en la imagen, este puede causar serios daños a la recuperación de la imagen.

Existen también perspectivas en cuanto a las aplicaciones de esta técnica de análisis como son:

- **Realce.** Si se tienen parámetros que caracterizan a las estructuras primarias, entonces se puede mediante unas modificaciones, realizar una reconstrucción diferente. De tal manera, que se de prioridad a los bordes, como en una perfilación y se elimine el ruido presente, o se eliminen la estructuras unidimensionales y se mantengan solo las estructuras bidimensionales.
- **Segmentación.** De lo anterior se tiene que la identificación de parámetros es en estructuras que principalmente indican cambios de intensidad en la imagen, por lo tanto delimitan regiones. Es así como se puede aprovechar para realizar una segmentación que incluso puede ser más eficiente, ya que se tienen características de las fronteras. De hecho, en el desarrollo del algoritmo se puede decir que se realiza segmentación de estructuras primarias.
- **Secuencias de imágenes.** El modelado de estructuras primarias que se obtiene a partir del análisis efectuado con la transformada polinomial, es útil para mejorar la segmentación o estimación de movimiento contando con una buena descripción de los bordes de la imagen; se puede también reducir el costo computacional si se usan algoritmos a múltiples escalas. Asimismo, conforme se obtengan mejores tasas de compresión se puede utilizar para comprimir las imágenes de error que se obtienen entre los cuadros reconstruidos de la secuencia de imágenes [Gat96].
- **Optimización para implantación.** Se puede buscar la optimización del algoritmo de la transformada polinomial para lograr su implantación en arquitectura electrónica, por medio de la obtención de algoritmos rápidos.

## 6.2 Conclusiones generales

Los algoritmos propuestos cumplen con los objetivos planteados inicialmente que son: (1) el cambio de representación de una imagen de manera eficiente, (2) el uso de la transformada polinomial para tal efecto y (3) obtener altas tasas de compresión.

La codificación y compresión de imágenes y de manera general el procesamiento de imágenes, van en dirección de un mejor entendimiento del sistema visual humano para la realización de sus tareas. Con esto, se expresa la necesidad de modelar cada vez más y mejor las estructuras importantes para la percepción. De lo anterior se desprende, que la técnica utilizada para el análisis, la transformada polinomial, es de vanguardia pues es un modelo del sistema de visión humano. Con ella se puede continuar con el sentido que lleva el área en general.

La transformada polinomial se ha probado con éxito en otras áreas del procesamiento de imágenes, como son realce y restauración. Por lo cual, se puede vislumbrar su uso en algoritmos de múltiple propósito, por ejemplo codificación eliminando ruido gaussiano.

El estudio comparativo realizado que se presenta en el último capítulo, muestra claramente como se pierde calidad conforme la razón de compresión se va haciendo mayor. Se proponen mejoras para incrementar la calidad en la reconstrucción. Es importante notar que la aplicación de los algoritmos a las dos imágenes usadas generó resultados diferentes, esto se debe a las características propias de la transformada polinomial, que es un modelo del sistema de visión humano; y al contenido de estructuras en las dos imágenes.

Se puede deducir lo siguiente: En la transmisión de imágenes es conveniente tomar en cuenta el contenido de la imagen. Si esta tiene un mayor número de estructuras unidimensionales que estructuras bidimensionales, entonces la calidad de la reconstrucción es muy buena aún con tasas de compresión altas. Ejemplo de ello son los dibujos animados o caricaturas, en el caso de secuencia de imágenes.

En las gráficas de comportamiento se observa que los algoritmos propuestos, tienen valores muy competitivos. Se esperan resultados aún mejores a partir de las perspectivas.

Los algoritmos propuestos, sobretodo el que corresponde a parámetros característicos tiene un desempeño óptimo, aún cuando existen ya algoritmos de compresión de imágenes que funcionan de manera eficiente, como el JPEG mostrado en el capítulo 5. Sin embargo, se tiene que tomar en cuenta que el JPEG tiene como técnica de análisis a la DCT que no genera una expansión de coeficientes, además de ello, tiene un algoritmo para cuantización de los coeficientes que permite valores fijos.

# Bibliografía

- [Abr65] Abramowitz, M. and Stegun, I. *Handbook of Mathematical Functions*. Dover, New York, 1965.
- [Aka92] Akansu, Ali N. and Haddad, Richard A. *Multiresolution Signal Decomposition*. Academic Press Inc. (Harcourt Brace Jovanovich, Publishers), 1992.
- [Asi84] Asimov, Isaac. *Historia de la ciencia*. Planeta, 1984.
- [Bal82] Ballard, Dana H. and Brown, Christopher M. *Computer Vision*. Prentice Hall, Inc., 1982.
- [Cor96] Corlay Trujillo, Monica. *Estudio comparativo e implementación de métodos tiempo frecuencia*. Tesis de maestría, DEPFI-UNAM, marzo 1996.
- [Esc92a] Escalante Ramírez, Boris. *Perceptually-Assesed Digital Processing of Medical Images*. PhD thesis, Technische Universiteit Eindhoven, september 1992.
- [Esc92b] Escalante-Ramirez, Boris and Martens, Jean Bernard. Noise reduction in computerized tomography images by means of polinomial transforms. *Journal of Visual Communication and Image Representation*, 3(3):272-285, september 1992.
- [Esc95] Escalante Ramírez, Boris; Venegas Martínez, Santiago and García Ugalde, Francisco. Pyramidal predictive image coding with polynomial transforms. *Still-Image Compression, SPIE*, 2418:107-117, february 1995.
- [Gat96] Gatica Pérez, Daniel. *Estimación de movimiento y segmentación basada en movimiento en secuencias de imágenes*. Tesis de maestría, DEPFI-UNAM, junio 1996.
- [Gon87] González, Rafael C. y Woods, Richard E. *Tratamiento digital de imágenes*. Addison-Wesley/Díaz de Santos, 1987.
- [Gud95] Gudivada, Venkat N. and Raghavan, Vijay V. Content-based image retrieval systems. *IEEE Computer Magazine*, 28(9):18-22, september 1995.

- [Hen93] Hendee, William R. and Wells, Peter. *The Perception of Visual Information*. Springer-Verlag New York, Inc., 1993.
- [Hor86] Horn, Berthold Klaus Paul. *Robot Vision*. MIT Press, Mc.Graw-Hill Book Co., 1986.
- [Jai89] Jain, Anil K. *Fundamentals of Digital Image Processing*. Prentice Hall, Inc., 1989.
- [Kay94] Kayargadde, Vishwakumara and Martens, Jean Bernard. Estimation of edge parameters and image blur using polynomial transforms. *CVGIP: Graphical Models and Image Processing*, 56(6):442-461, november 1994.
- [Lim90] Lim, Jae S. *Two-dimensional Signal and Image Processing*. Prentice Hall, Inc., 1990.
- [Lir87] Lira Chavez, Jorge. *La percepción remota: Nuestros ojos desde el cielo*. Fondo de Cultura Económica, 1987.
- [Lop96] Lopez Miranda, Juan Roman. *Realce y restauración de imágenes por medio de transformaciones polinomiales*. Tesis de maestria, DEPMI-UNAM, agosto 1996.
- [Mal92a] Mallat, Stephane and Liang, Hwang Wen. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2):617-643, july 1992.
- [Mal92b] Mallat, Stephane and Zhong, Sifen. Characterization of signals from multiescale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7):710-732, july 1992.
- [Mar82] Marr, David. *Vision*. W.H.Freeman and Company New York, 1982.
- [Mar90a] Martens, Jean-Bernard. The hermite-transform applications. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(9):1607-1618, september 1990.
- [Mar90b] Martens, Jean-Bernard. The hermite-transform theory. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(9):1595-1606, september 1990.
- [Mey94] Meyer, Yves. *Wavelets, Algorithms and Applications*. SIAM, 1994.
- [Net89] Netravali, Arun N. and Haskell, Barry G. *Digital Pictures. Representation and Compression*. Plenum Press, june 1989.
- [Pap91] Papoulis, Athanasios. *Probability, Random Variables and Stochastic Processes*. Mc.Graw Hill Book Co., 1991.



- [Pee93] Peebles, Peyton Z. *Probability, Random Variables and Random Signal Principles*. Mc.Graw Hill Book Co., 1993.
- [Pro88] Proakis, J.G., Manolakis, D.G.. *Introduction to Digital Signal Processing*. Macmillan Publishing Company, N.Y., 1988.
- [Rab91] Rabbani, Majid y Jones Paul. *Técnicas de compresión de imágenes*. SPIE, 1991.
- [Sk188] Sklar, Bernard. *Digital Communications, Fundamentals and Applications*. Prentice Hall, 1988.
- [Vit79] Viterbi, Andrew J. and Omura, Jim K. *Principles of Digital Communication and Coding*. Mc.Graw Hill Book Co., 1979.
- [Viv96] Viveros Cancino, Oscar y Escalante Ramírez, Boris. Codificación unidimensional de imágenes mediante la transformada polinomial. *Conferencia de Ingeniería Eléctrica, CINVESTAV, IPN, Robotica y Visión*, septiembre 1996.
- [Viv97] Viveros Cancino, Oscar y Escalante Ramírez, Boris. Parametric image coding by means of polinomial transforms. In *Visual Communications and Image Processing, SPIE*, volume 3024, february 1997.
- [Wit84] Witkin, Andrew P. Scale space filtering: a new approach to multi-scale description. In *Image Understanding 1984*, volume ch.3, pages 79-95, 1984.
- [You91] Young, Richard A. Oh say, can you see? the physiology of vision. Technical report, Computer Science Department, General Motors Research Laboratories, Warren, Michigan USA, 1991.