



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**FUSIÓN DE VOLÚMENES GENERADOS A
PARTIR DE IMÁGENES MÉDICAS UTILIZANDO
LA TRANSFORMADA HERMITE**

T E S I S

QUE PARA OBTENER EL GRADO DE:

**MAESTRA EN INGENIERÍA
(COMPUTACIÓN)**

P R E S E N T A:

SONIA CRUZ TECHICA

DIRECTOR DE LA TESIS: DR. BORIS ESCALANTE RAMIREZ

MÉXICO, D.F.

2011.



Universidad Nacional
Autónoma de México

Biblioteca Central

Dirección General de Bibliotecas de la UNAM



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.

Dedicatoria

A mis padres y hermanos por todo el apoyo brindado, en especial a mi mamá por su enorme dedicación a lo largo de todos mis estudios. Gracias por su afecto, amor y sostén incondicional.

A Cristian por su apoyo y paciencia en esta etapa, en especial por compartir la realización de uno de mis mayores sueños.

A mis compañeros y amigos de laboratorio de imágenes, Ernesto y Alfonso, gracias por su apoyo y orientación durante el desarrollo de este trabajo, sobre todo por haber hecho más amenas las horas de trabajo y las visitas a los congresos.

Agradecimientos

Al Dr. Boris Escalante Ramírez, por el todo el apoyo que me brindo para el desarrollo de este trabajo, en especial por abrirme las puertas de su laboratorio y permitirme adentrarme en mundo de la investigación en el procesamiento digital de imágenes.

Al Dr. Fernando Arámbula, a la Dra. María Elena Martínez y a la Dra. Lucía Medina Gómez por sus valiosos comentarios y aportaciones en este trabajo.

Al M.C. Sarael Alcauter del Instituto de Nacional de Psiquiatría por sus comentarios y por colaborar en este trabajo de investigación proporcionando algunos casos de estudio empleados en la evaluación de resultados del algoritmo de fusión propuesto.

Al Posgrado de Ciencia e Ingeniería de la Computación, planta docente y administrativa que me permitieron obtener estudios de maestría de calidad.

A la Universidad Nacional Autónoma de México y la Facultad de Ingeniería, a quienes les debo gran parte de mi formación académica.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por el apoyo económico que me proporcionó durante mis estudios de maestría.

Al programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPITT) y al programa de proyectos de la sala Ixtli de la UNAM, ya que una parte de este trabajo fue financiado por ellos con las claves IN106608 e IX100610 respectivamente.

A Cecilia, Lulú, Diana y Amalia por todo el apoyo brindado durante mi estancia en el Posgrado, no sólo en el aspecto administrativo sino también en el motivacional.

A todos, muchas gracias.

Resumen

Las imágenes médicas pueden ser adquiridas mediante diferentes técnicas (modalidades), las cuales tienen sus respectivas áreas de aplicación, algunas de ellas proporcionan información sobre el funcionamiento de algún órgano o sistema mientras que otras sólo contienen información anatómica; en general, en el primer caso las imágenes presentan baja resolución espacial mientras que en el segundo caso se tiene mayor resolución. No obstante, el análisis de imágenes médicas con frecuencia requiere de la valoración de más de una modalidad a fin de que el especialista posea mayor información para la toma de decisiones tanto en el diagnóstico como el tratamiento de enfermedades.

Debido a lo anterior, la fusión de imágenes médicas se ha vuelto una técnica útil para el análisis de imágenes ya que integra la información de imágenes que proceden de diversas modalidades en una sola, simplificando así su interpretación.

En este trabajo se presenta un algoritmo de fusión de volúmenes de imágenes médicas mediante su descomposición en coeficientes de la transformada de Hermite rotada multirresolución en 3D y la aplicación de una regla de fusión de coeficientes basada en el álgebra lineal para evaluar la dependencia lineal de manera local. La elección de dicha transformada como herramienta matemática para el análisis de los datos radica en el hecho de que emplea funciones que han sido propuestas para modelar los perfiles de los campos receptivos presentes en el sistema de visión humano.

Para efectos de comparación se incluye en la evaluación de resultados la fusión mediante la transformada wavelet empleando diversas reglas de fusión y tipos de wavelets. El análisis de dichos resultados demostró que la transformada de Hermite en 3D caracteriza mejor las estructuras importantes de los volúmenes y en consecuencia permite la obtención de volúmenes de mejor calidad perceptiva.

Palabras clave: *fusión de imágenes 3D, imágenes médicas, transformada de Hermite, multirresolución, dependencia lineal.*

Abstract

Medical images can be acquired through different techniques (modalities), which have their own application areas; some of them provide information on the functional activity, while others contain only anatomic information. Usually, in the first case, images have low spatial resolution while in the second case have a higher resolution. However, the analysis of medical images often requires the evaluation of more than one modality; in order provide the specialist with more information for decision making as well as for the diagnosis and the treatment of diseases.

For this reason, medical image fusion has become a useful technique for image analysis. It integrates the information of different imaging modalities in only one, simplifying the interpretation.

In this thesis, a fusion algorithm for medical image volumes which uses a 3D multiresolution directional-oriented Hermite transform decomposition is introduced; this also is based on the linear algebra through a measurement of the linear dependence for the application of a fusion rule to combine the Hermite coefficient. The choice of this transform as a mathematical tool for data analysis lies in the fact that it uses features that have been proposed to model the profiles of the receptive fields present in the human visual system.

For comparison the proposed method is compared with the wavelet transform, using different fusion rules and types of wavelets. The analysis of these results showed that the 3-D Hermite transform best characterized important structures of the volumes and therefore, better perceptual quality volumes are obtained.

Key words: *3D image fusion, medical images, Hermite transform, multiresolution, linear dependence.*

Índice general

Dedicatoria	I
Agradecimientos	III
Resumen	V
Abstract	VII
Índice general.....	IX
Índice de figuras	XIII
Lista de tablas	XIX
1. Introducción.....	1
2. Imágenes médicas	5
2.1. Conceptos básicos.....	5
2.2. Modalidades de adquisición de imágenes médicas.....	8
2.2.1. Radiología convencional	10
2.2.2. Radiología digital	11
2.2.2.1. Tomografía computarizada	12
2.2.3. Medicina Nuclear	13
2.2.3.1. Tomografía por Emisión de Fotón Único (SPECT).....	13
2.2.3.2. Tomografía por Emisión de Positrones (PET)	14
2.2.4. Ecografía.....	15
2.2.5. Resonancia magnética	16
2.2.6. Imagenología óptica	18
2.2.7. La imagen molecular	19
2.3. Futuro de las imágenes médicas	19
3. Fusión de imágenes médicas.....	21
3.1. Fundamentos de fusión de imágenes	21
3.1.1. Conceptos de fusión	21
3.1.2. Propósito de la fusión de imágenes médicas	23
3.2. Técnicas de fusión para el despliegue de imágenes médicas.....	24
3.3. Contexto genérico para esquemas de fusión de imágenes	26
3.4. Niveles de fusión de imágenes.....	27
3.5. Técnicas de fusión de imágenes.....	28
3.5.1. Intensidad-matiz-saturación (IHS)	30
3.5.2. Combinaciones aritméticas.....	31

3.5.2.1.	Combinación ponderada (suma y multiplicación)	31
3.5.2.2.	Promedio	31
3.5.3.	Análisis de componentes principales (PCA)	31
3.5.4.	Métodos basados en teoría de estimación	33
3.5.5.	Técnicas de fusión multirresolución	34
3.5.5.1.	Contexto genérico la fusión multirresolución	34
3.5.5.2.	Descomposición piramidal	36
3.5.5.3.	Descomposición wavelet	38
3.5.6.	Métodos de selección	41
3.6.	Reglas de fusión	43
3.7.	Ventajas y limitantes de los métodos de fusión	44
3.8.	Situación actual de la fusión de imágenes médicas	46
4.	Transformada de Hermite	49
4.1.	Relevancia en Percepción Humana	49
4.2.	Transformada polinomial	50
4.3.	Transformada de Hermite	52
4.4.	Transformada de Hermite Discreta	55
4.5.	Transformada de Hermite 2D	57
4.6.	Transformada de Hermite 2D rotada	58
4.7.	Transformada de Hermite multirresolución 2D	60
4.8.	Transformada de Hermite 3D	62
5.	Principios de evaluación de la fusión de imágenes	65
5.1.	Introducción	65
5.2.	Clasificación de las métricas de calidad	66
5.3.	Métricas de calidad subjetivas	67
5.4.	Métricas de calidad objetivas	68
5.4.1.	Métricas de calidad basadas en el error	69
5.4.2.	Métricas de calidad basadas en la información mutua	70
5.4.3.	Medidas de calidad basadas en la información estructural	71
5.4.4.	Medidas de calidad basadas en coeficientes de correlación	73
5.5.	Comparación de métricas de calidad objetivas	73
6.	Algoritmo de fusión	75
6.1.	Introducción	75
6.2.	Esquema de fusión basado en la transformada de Hermite	76
6.2.1.	Regla de fusión propuesta	78
6.2.1.1.	Selección de coeficientes de frecuencias bajas	78
6.2.1.2.	Selección de coeficientes de frecuencias altas	79
6.2.2.	Reducción de ruido	81
6.3.	Programación del esquema de fusión	82
6.3.1.	Algoritmo rápido para la obtención de la DHT 3D	84
6.3.1.1.	Programación del algoritmo	87
6.3.2.	Algoritmo para la rotación de coeficientes de la DHT	90
6.3.2.1.	Análisis de orientación local de una señal 3D	90
6.3.2.1.	Algoritmo de proyección de los coeficientes en 3D a 1D	91

6.3.2.2.	Algoritmo de proyección de los coeficientes en 1D a 3D.....	92
6.3.3.	Algoritmo rápido para la obtención de la IDHT 3D.....	92
6.3.4.	Implementación de la transformada wavelet 3D	99
7.	Resultados y evaluación de la fusión de imágenes médicas	101
7.1.	Resultados obtenidos de la fusión.....	101
7.1.1.	Comparación con distintas reglas de fusión	107
7.1.2.	Comparación con la transformada wavelet	110
7.2.	Evaluación del método propuesto en presencia de ruido.....	115
7.2.1.	Ruido aditivo normal (Gaussiano).....	115
7.2.2.	Ruido impulsivo (sal y pimienta)	119
7.3.	Comparación con otras técnicas de fusión en 2D	121
7.3.1.	Fusión de imágenes multifocales con referencia.....	122
7.3.2.	Fusión de imágenes médicas (sin referencia).....	128
8.	Conclusiones.....	131
A.	Algoritmos empleados en fusión de imágenes médicas	133
B.	Transformada Curvelet	137
C.	Transformada Contourlet.....	141
D.	Filtros de la HT directa 3D.....	145
E.	Filtros de la HT inversa 3D.....	147
F.	Árbol utilizado en la implementación de la HT 3D	149
	Referencias	151

Índice de figuras

Figura 2.1 Información no deseada en imágenes médicas. (a) La imagen de flourosco피아 presenta mucho ruido, en forma de un punteado aleatorio. (b) La angiografía presenta un artefacto debido al movimiento de la pared abdominal, que produce las líneas blancas que se observan en la imagen del hígado [34].	6
Figura 2.2 Información no deseada en imágenes médicas. (a) Vista 2D de un corte de una imagen donde se muestra el límite entre dos tejidos, el resultado de la intensidad de los voxels se muestra abajo [105]. (b) CT sin medio de contraste. Efecto de volumen parcial dado por techo orbitario (flecha) [81].	7
Figura 2.3 Corrección de la no uniformidad del campo magnético en MR. (a) Imagen original. (b) Estimación de la no uniformidad del campo (utilizando las técnicas basadas en PDE). (c) Imagen reconstruida [55].	7
Figura 2.4 Imagen anatómica y funcional. (a) y (b) Radiografías de tórax. (c) Gammagrafía pulmonar, en la que se aprecia una clara embolia pulmonar, invisible en las radiografías de tórax [113].	8
Figura 2.5 Las tres principales orientaciones de imagen para el cuerpo humano: los planos coronal, sagital y transversal (axial) [34].	10
Figura 2.6 Imágenes de rayos X. (a) Radiografía del cráneo invertida. (b) Radiografía del abdomen utilizando bario como medio de contraste para visualizar el colon [48].	11
Figura 2.7 Imágenes de CT. (a) CT de alta resolución del abdomen. (b) Cortes axiales [48].	12
Figura 2.8 Estudio de SPECT del consumo de anfetamina en un cerebro normal y anormal [48].	14
Figura 2.9 Estudio de PET con FDG asimilado en un cerebro epiléptico [48].	15
Figura 2.10 Imágenes de ultrasonido. (a) Escaneo que muestra el detalle de la cabeza de un feto. (b) Escaneo Doppler de un corazón fetal donde el mapa de colores superpuesto muestra la velocidad del flujo sanguíneo en una región seleccionada [48].	16
Figura 2.11 Imágenes de resonancia magnética. (a) Corte sagital de una MRI del cerebro. (b) MRI de los tejidos de la rodilla [48].	18
Figura 3.1 Datos de múltiples modalidades, registrados en un sistema de referencia común de un paciente, los cuales se combinan para construir una representación integral del paciente [57].	22

Figura 3.2 Aplicación de fusión de imágenes médicas para la localización de lesiones cerebrales. (a) MRI y (b) SPECT del mismo paciente. (c) Imagen co-registrada, MRI + SPECT. La lesión cerebral es más prominente en la imagen compuesta que en las modalidades individuales [34].....	24
Figura 3.3 Cortes axiales de MR y PET presentados con métodos alternativos. (a) Imagen intercalada, por ejemplo, MR en escala de grises y PET en escala de colores hot-body. (b) Composición de color, por ejemplo, MR en 30 tonalidades de grises y PET en 8 tonalidades de la escala de color hot-body con un matiz uniforme. (c) Codificación RG, por ejemplo, MR codificada en rojo lineal y PET en verde lineal [94].....	26
Figura 3.4 Diagrama de bloques de un procedimiento general de fusión de imágenes	27
Figura 3.5 Niveles de fusión de imágenes [88].	28
Figura 3.6 Esquema de la transformación de componentes principales.	32
Figura 3.7 Esquemas de fusión multirresolución de imágenes [136]	34
Figura 3.8 Clasificación de las medidas de nivel de actividad.....	35
Figura 3.9 Estructuras multirresolución y métodos de agrupación de coeficientes de técnicas piramidales y DWT	36
Figura 3.10 Una etapa de la descomposición de una imagen mediante la DWT	39
Figura 4.1 La transformada polinomial directa e inversa.....	52
Figura 4.2 Funciones filtro para $\sigma = 1$	54
Figura 4.3 Descomposición de una imagen con la transformada de Hermite.....	58
Figura 4.4 Descomposición de una imagen con la transformada de Hermite rotada.....	59
Figura 4.5 Expansión multirresolución de la transformada de Hermite	61
Figura 4.6 Distribución de los coeficientes de segundo orden de un voxel	63
Figura 4.7 La transformada de Hermite 3D en un sistema de coordenadas cartesiano y su rotación sobre un volumen.	64
Figura 6.1 Esquema de fusión mediante la HT rotada multirresolución.....	78
Figura 6.2 Construcción de los vectores a partir de una región de 3x3 en el modelo de vector80	
Figura 6.3 Esquema de fusión con reducción de ruido mediante la HT rotada multirresolución.	82
Figura 6.4 Diagrama del flujo de datos general.	83

Figura 6.5 Interfaz gráfica de la aplicación de fusión de volúmenes.	84
Figura 6.6 Matriz de derivadas parciales en tres dimensiones	86
Figura 6.7 Algoritmo rápido para valores grandes de N	87
Figura 6.8 Estructura de registros para la implementación de la DHT 3D con N=2	88
Figura 6.9 Coordenadas esféricas.....	91
Figura 6.10 Obtención de las ponderaciones para la IDHT 3D con N=2 a partir del árbol de derivación.	93
Figura 6.11 Disposición de los datos obtenidos en el programa contra su significado real.....	94
Figura 6.12 Cuatro disposiciones posibles de los ceros que existen entre los datos de las matrices de coeficientes.....	94
Figura 6.13 Representación de un filtro genérico de 3x3x3.....	95
Figura 6.14 Esquema de la transformada wavelet en tres dimensiones.	99
Figura 7.1 Volúmenes de prueba CT/MR. (a) y (d) corresponden a los cortes axiales 45/89, (b) y (e) a los cortes coronales 160/320, (c) y (f) a los cortes sagitales 160/320.	102
Figura 7.2 Volúmenes de prueba MR/PET. (a) y (d) corresponden a los cortes axiales 50/100, (b) y (e) a los cortes coronales 128/256, (c) y (f) a los cortes sagitales 128/256.....	103
Figura 7.3 Volúmenes de prueba MR/fMR. (a) y (d) corresponden a los cortes axiales 22/44, (b) y (e) a los cortes coronales 40/80, (c) y (f) a los cortes sagitales 40/80.....	103
Figura 7.4 Volúmenes de prueba MR-T1/ SPECT (Caso de estudio 1). (a) y (d) corresponden a los cortes axiales 46/91, (b) y (e) a los cortes coronales 55/109, (c) y (f) a los cortes sagitales 46/91.	104
Figura 7.5 Volúmenes de prueba MR-T1/ SPECT (Caso de estudio 2). (a) y (d) corresponden a los cortes axiales 48/160, (b) y (e) a los cortes coronales 128/256, (c) y (f) a los cortes sagitales 128/256.	104
Figura 7.6 Interfaz gráfica para el despliegue del resultado de fusión. (a) Modo 3 vistas donde todas las vistas corresponden al volumen fusionado. (b) Modo cortes axiales donde se contrasta corte a corte el resultado con los datos originales en el plano axial.	105
Figura 7.7 Fusión de CT/MR con utilizando la HT con el esquema propuesto a varios niveles n de resolución. (a) y (b) son los datos originales, (c) es el resultado de la fusión con n=1, (d) es el resultado de la fusión con n=2 y (e) es el resultado de la fusión con n=3.....	106
Figura 7.8 Imágenes multifocales empleadas en una prueba de evaluación de distintas reglas de fusión. (a) Imagen original considerada ‘the ground truth’ para el cálculo de las métricas.	

(b) y (c) Imágenes empleadas en el proceso de fusión donde ciertos objetos han sido emborronados. 108

Figura 7.9 Comparación de fusión en imágenes médicas con distintas reglas de fusión. (a) y (b) corte 68 de los volúmenes de CT y MR respectivamente, (c) resultado empleando máximo absoluto, (d) resultado empleando máximo con verificación por consistencia, (e) resultado empleando saliente y medida de paridad, (f) resultado empleando dependencia lineal. 109

Figura 7.10 Gráfica de la información mutua total del proceso de fusión para los volúmenes CT/MR, MR/PET y Caso 2 para diversos métodos de fusión: bior se refiere a una wavelet biortogonal, coif a una coiflet, db a Daubechies, Hermite a la técnica propuesta y sym a una wavelet de tipo symlet. En el eje 'x' se enlistan todas las técnicas de fusión empleadas en los experimentos y en el eje 'y' los valores de información mutua total entre los volúmenes fusionados y los volúmenes fuente. 111

Figura 7.11 Comparación de fusión en imágenes médicas con distintos métodos de fusión. (a) y (b) corte 50 de los volúmenes de MR y PET respectivamente, (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5, (e) fusión con DWT Daubechies 6, (f) fusión con DHT. 112

Figura 7.11 (Cont.) Comparación de fusión en imágenes médicas con distintos métodos de fusión. (g) fusión con DWT symlet 7. 113

Figura 7.12 Comparación de fusión en imágenes médicas con distintos métodos de fusión. (a) y (b) corte 68 de los volúmenes de CT y MR respectivamente, (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5. 113

Figura 7.12 (Cont.) Comparación de fusión en imágenes médicas con distintos métodos de fusión. (e) fusión con DWT Daubechies 6, (f) fusión con DHT, (g) fusión con DWT symlet 7. 114

Figura 7.13 Comparación de fusión en imágenes médicas degradadas con ruido Gaussiano $\sigma=0.001$ empleando distintos métodos de fusión. (a) y (b) corte 58 de los volúmenes de MR y SPECT (caso de estudio 2) contaminados con ruido, (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5. 116

Figura 7.13 (Cont) Comparación de fusión en imágenes médicas degradadas con ruido Gaussiano $\sigma=0.001$ empleando distintos métodos de fusión. (e) fusión con DWT Daubechies 6, (f) fusión con DHT, (g) fusión con DWT symlet 7. 117

Figura 7.14 Gráfica de los coeficientes de correlación paso altas del resultado de fusión con reducción de ruido Gaussiano con $\sigma=0.001$ de los volúmenes MR-T1/SPECT (caso 2) con MR-T1 y SPECT para diversos métodos de fusión: bior se refiere a una wavelet biortogonal, coif a una coiflet, db a Daubechies, Hermite a la técnica propuesta y sym a una wavelet de tipo symlet. En el eje 'x' se enlistan todas las técnicas de fusión empleadas en los experimentos y en el eje 'y' los valores del coeficiente de correlación paso altas entre el volumen fusionado y las imágenes fuente: CCHA corresponde al coeficiente de correlación entre el volumen

fusionado y el volumen A (MR-T1), y CCHB corresponde al coeficiente de correlación entre el volumen fusionado y el volumen B (SPECT).	118
Figura 7.15 Comparación de fusión en imágenes médicas degradadas con ruido sal y pimienta para una densidad de 0.005 empleando distintos métodos de fusión. (a) y (b) corte 45 de los volúmenes de CT y MR contaminados con ruido.	119
Figura 7.15 (Cont.) Comparación de fusión en imágenes médicas degradadas con ruido sal y pimienta para una densidad de 0.005 empleando distintos métodos de fusión. (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5, (e) fusión con DWT Daubechies 6, (f) fusión con DHT, (g) fusión con DWT symlet 7.	120
Figura 7.16 Aplicación de fusión del método propuesto en 2 dimensiones.	122
Figura 7.17 Imágenes de entrada multifocales para el experimento de Pelota y fondo. (a) Imagen de referencia. (b) Imagen con el fondo desenfocado y la pelota en foco. (c) Imagen con el fondo en foco y la pelota desenfocada.	124
Figura 7.18 Resultados de fusión de imágenes multifocales para el experimento de Pelota y fondo. (a) Imagen obtenida empleando HT, (b) Imagen obtenida empleando WT, (c) Imagen obtenida empleando CT y (d) Imagen obtenida empleando CWT.	125
Figura 7.19 Imágenes de entrada multifocales para el experimento de Líneas. (a) Imagen de referencia. (b) Imagen con la franja central desenfocada. (c) Imagen con las franjas laterales desenfocadas.	126
Figura 7.20 Resultados de fusión de imágenes multifocales para el experimento de Líneas. (a) Imagen obtenida empleando HT, (b) Imagen obtenida empleando WT, (c) Imagen obtenida empleando CT y (d) Imagen obtenida empleando CWT.	127
Figura 7.21 Resultados de fusión de (a) CT y (b) MR. (c) Imagen obtenida empleando HT, (d) Imagen obtenida empleando WT, (e) Imagen obtenida empleando CT y (f) Imagen obtenida empleando CWT.	129
Figura 7.22 Resultados de fusión de (a) CT y (b) MR con ruido Gaussiano $\sigma=0.001$. (c) Imagen obtenida empleando HT, (d) Imagen obtenida empleando WT, (e) Imagen obtenida empleando CT y (f) Imagen obtenida empleando CWT.	130
Figura B.1 Transformada curvlet discreta de una imagen P.	138
Figura B.2 Transformada curvelet a dos niveles de una imagen de resonancia magnética. ...	139
Figura C.1 Diagrama de bloques de la transformada contourlet.	142
Figura C.2 Transformada contourlet a tres niveles de una imagen de resonancia magnética.	143
Figura F.1 Árbol de derivaciones de orden 2 para el cálculo de la HT mediante sumas y diferencias.	150

Lista de tablas

Tabla 2.1 Características de las modalidades de imágenes médicas.	9
Tabla 3.1 Métodos de combinación de imágenes médicas para el despliegue.....	25
Tabla 7.1 Resultados del desempeño de la fusión para diferentes niveles de resolución (Figura 7.7) donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B, MI_{FA} y MI_{FB} es la información mutua entre el volumen fusionado y los volúmenes fuente A y B, MI_{FAB} es la información mutua total y FS la simetría de fusión.	107
Tabla 7.2 Resultados del desempeño de la fusión con las imágenes de la Figura 7.8 empleando la HT con diferentes reglas de fusión. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI_{FR} a la información mutua entre la imagen fusionada y la imagen de referencia.	108
Tabla 7.3 Resultados del desempeño de la fusión con las imágenes de la Figura 7.8 empleando la HT con diferentes reglas de fusión donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B, MI_{FA} y MI_{FB} es la información mutua entre el volumen fusionado y los volúmenes fuente A y B, MI_{FAB} es la información mutua total y FS la simetría de fusión.....	110
Tabla 7.4 Promedio de los resultados obtenidos del proceso de fusión de los 5 volúmenes aplicando distintos métodos de fusión donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B, MI_{FA} y MI_{FB} es la información mutua entre el volumen fusionado y los volúmenes fuente A y B, MI_{FAB} es la información mutua total y FS la simetría de fusión.....	114
Tabla 7.5 Promedio de los resultados obtenidos del proceso de fusión de los 5 volúmenes aplicando distintos métodos de fusión donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B.	118
Tabla 7.6 Resultados del promedio de todos los resultados de los experimentos efectuados con imágenes de referencia. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI a la información mutua entre la imagen fusionada y la imagen de referencia.	123
Tabla 7.7 Resultados del desempeño de la fusión para el experimento Pelota y fondo empleando diferentes métodos de fusión. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI a la información mutua entre la imagen fusionada y la imagen de referencia.	125

Tabla 7.8 Resultados del desempeño de la fusión para el experimento de Líneas empleando diferentes métodos de fusión. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI a la información mutua entre la imagen fusionada y la imagen de referencia..... 127

Tabla 7.9 Resultados del desempeño de la fusión las imágenes correspondientes al corte 70 de los volúmenes de CT y MR, donde MI_{FA} es la información mutua entre la imagen fusionada y la imagen A (CT), MI_{FB} es la información mutua entre la imagen fusionada y la imagen B (MR) y MI_{FAB} es la información mutua total. 128

Tabla A.1 Algoritmos empleados en fusión de imágenes médicas..... 133

Introducción

El procesamiento digital de imágenes, comprende un conjunto de técnicas que operan sobre la representación digital de una imagen, con el propósito de destacar algunos elementos que forman parte de la escena capturada en dicha imagen de manera que se facilite su análisis. En general, el reto es obtener imágenes de buena calidad que nos permitan extraer la mayor cantidad de información posible, para lo cual en ocasiones se requiere realzar o modificar la imagen para mejorar su apariencia, o bien, combinar imágenes o porciones de las mismas para conjuntar la información.

Por otro lado, el campo de las imágenes médicas ha avanzado significativamente en los últimos años convirtiéndose en una herramienta primordial en la práctica de la medicina, y generando a su vez nuevas líneas de investigación científica que abarcan múltiples disciplinas desde la visión por computadora y el procesamiento de imágenes, pasando por técnicas gráficas y de visualización, hasta los entornos inmersivos y dispositivos para la manipulación interactiva de la información.

En la actualidad, los médicos pueden tener a su disposición una mayor variedad de imágenes médicas de distintas modalidades que en general son tridimensionales y con frecuencia deben ser relacionadas de alguna manera, pues debido a las distintas naturalezas de adquisición de cada modalidad, las imágenes médicas poseen características propias que destacan ciertas estructuras o la actividad de algún órgano o sistema, pero no ambas. Así, el problema que se presenta es el manejo eficiente de las imágenes, fundamentalmente en el modo en que son vistas, en cómo son comunicadas y aprendidas y, por último, en la forma de interactuar con la información médica.

De esta forma, la necesidad de la fusión de imágenes médicas parte del hecho de tener información de diversas técnicas de adquisición de una misma región que requiere ser analizada e interpretada. La fusión de imágenes entonces se define como la combinación de imágenes que proceden de diversas modalidades en una sola imagen con mayor cantidad de información, con la finalidad de presentar de manera conjunta e integrada toda la información (anatómica y funcional) para simplificar su interpretación; y dado que las imágenes empleadas en el presente trabajo son tridimensionales, el concepto se extiende a la fusión de volúmenes. En los últimos años se han desarrollado algunas aplicaciones que incorporan la fusión de imágenes médicas y se observa una tendencia en aumento, ya que la fusión de imágenes en este ámbito va más allá de reducir el costo de almacenamiento, pues provee al especialista de una información enriquecida para la toma de decisiones tanto en el diagnóstico como el tratamiento de enfermedades.

Existen varias técnicas de fusión de imágenes en general, pero para aplicaciones médicas han destacado principalmente el análisis multirresolución mediante la transformada wavelet y el análisis geométrico empleando la transformada contourlet. Sin embargo en cuanto a fusión de volúmenes generados de imágenes médicas, existen pocos trabajos desarrollados, lo más común es hacer el procesamiento de todos los cortes que constituyen el volumen para luego integrarlos.

En este trabajo se presenta un algoritmo de fusión de volúmenes de imágenes médicas a partir de la transformada de Hermite (HT: Hermite transform) multirresolución localmente orientada para tres dimensiones. La elección de dicha transformada como herramienta matemática para el análisis de los datos tiene como fundamento el hecho de ser un modelo de representación que incorpora características del sistema de visión humano, como el análisis empleando funciones derivadas de Gaussiana. Estas funciones han sido usadas para modelar los campos receptivos presentes en el ojo humano para las primeras etapas de visión, debido a que se ha demostrado que las derivadas de Gaussianas modelan la respuesta de dichos campos receptivos con mayor exactitud que otras funciones. Por otra parte, es importante señalar que la HT, durante la etapa de análisis, separa la información visual contenida en la imagen en información de baja frecuencia (promedio) y de alta frecuencia, esta última información extraída hace de la HT un buen descriptor de patrones característicos como bordes y texturas, y esto en particular resulta de especial interés para que el proceso de fusión preserve los detalles de los datos originales.

Particularmente, la propiedad de la HT de análisis de orientación local [29, 71] es significativa en el proceso de fusión ya que no sólo permite detectar bordes, sino que también permite el análisis de patrones complejos incluyendo su orientación local, esto hace que en el proceso de fusión se pueda discriminar y seleccionar patrones de acuerdo a su energía y orientación locales. El uso de la HT rotada¹ proporciona un conjunto de coeficientes con alto contenido de energía, de tal manera que se reduce significativamente el número de coeficientes para realizar el proceso de fusión y en consecuencia, se logra una representación poco redundante en información.

El esquema de fusión propuesto se aplica a imágenes médicas tridimensionales (volúmenes) que previamente han sido registradas; así, el problema de fusión se centra en el esquema de extracción de características y la regla de fusión para combinarlas. El proceso general de fusión con la HT es el siguiente: obtener la transformada de Hermite rotada de cada volumen a diferentes escalas, aplicar las reglas de fusión basadas en la dependencia lineal, combinar los coeficientes de la HT multirresolución localmente orientada de cada volumen para generar un sólo conjunto de coeficientes, y finalmente recuperar el volumen compuesto a través de la HT inversa descendiendo así en la pirámide multirresolución hasta obtener una reconstrucción en la escala original.

En resumen, la motivación de este trabajo lo constituye la aplicación médica de la fusión de volúmenes así como el estudio de técnicas de procesamiento inspiradas en modelos biológicos que tienen como referencia al sistema de visión humana. Para este fin, se hace uso de imágenes médicas de referencia, tales como tomografías computarizadas (CT: computed

¹ Los coeficientes se muestran en un sistema de coordenadas rotado por un ángulo θ .

tomography) y resonancias magnéticas (MR: magnetic resonance), así como imágenes de casos de estudio reales proporcionadas por el Instituto Nacional de Psiquiatría, en donde la fusión de imágenes anatómicas y funcionales es primordial para la evaluación de sus pacientes.

Para efectos de comparación se incluye en la evaluación de resultados la fusión mediante la transformada wavelet 3D empleando varios tipos de wavelets con diferentes niveles de descomposición. Asimismo, se programaron otras reglas de fusión y se realizaron pruebas de fusión también en dos dimensiones para poder comparar el esquema completo con otras técnicas de análisis de imágenes como es el caso de la transformada curvelet y la transformada contourlet.

En los siguientes capítulos se profundizará en los temas mencionados, para ello esta tesis se ha organizado en 8 capítulos que a continuación se describen de forma breve. En el capítulo 2 se abordan los conceptos básicos de las imágenes médicas y las diferentes modalidades existentes, así como las técnicas de visualización más utilizadas en ese ámbito. En el capítulo 3 se expondrá una explicación más amplia sobre la definición de fusión, algunas técnicas utilizadas, la aplicación que tiene dentro de la imagenología médica y se concluirá con algunos ejemplos. En el capítulo 4 se detalla la teoría referente a la transformada de Hermite, empezando por la importancia que tiene como modelo inspirado en el sistema de visión humano y algunos conceptos generales de las transformadas polinomiales, enfatizando en el caso particular de la HT; de esta última se mostrarán las expresiones para los filtros de análisis y síntesis hasta tres dimensiones en sus casos continuo y discreto, incluyendo su versión rotada. En el capítulo 5 se habla sobre las métricas de calidad perceptiva y en particular de los criterios de evaluación utilizados para evaluar los resultados. En el capítulo 6 se presenta el método de fusión propuesto a partir de la HT y se describen los algoritmos implementados para llevar a cabo el procesamiento los datos. En el capítulo 7 se exponen los resultados obtenidos y la discusión de los mismos dentro de un marco comparativo con otras técnicas de fusión. En el último capítulo se presentan las conclusiones finales de este trabajo.

Adicionalmente se incluyen algunos apéndices que profundizan en algunos temas como lo es el caso del estado del arte en métodos de fusión aplicados a imágenes médicas, la transformada curvelet y la transformada contourlet. Finalmente, los últimos apéndices tienen la función de complementar la explicación de la implementación del método propuesto expuesta en el capítulo 6.

Imágenes médicas

2.1. Conceptos básicos

Una imagen médica es la representación de puntos en un plano o volumen (representación bidimensional o tridimensional) de la distribución espacial de una o más propiedades físicas o químicas dentro del cuerpo humano. Para obtenerla se requiere irradiar una muestra (o el paciente), la naturaleza de la energía empleada determina el nombre de la modalidad de la imagen [113].

En un sentido más amplio, una imagen digital es un arreglo espacial de intensidades cuya calidad depende del tamaño de los píxeles relativo al tamaño de la imagen y del número de valores de tonos gris disponibles para la descripción del rango de intensidades; la calidad de la imagen es mayor para píxeles pequeños y un gran número de tonos de gris. Por ello, existen parámetros muy importantes dentro del análisis de las imágenes que deben ser considerados, éstos son el contraste, la resolución y la cuantización de niveles de gris.

El contraste determina qué es lo que se ve en la imagen; técnicamente se define como el cambio de la luminancia que existe entre dos puntos de una imagen. Lo más importante, en el contexto que nos ocupa, es tener claro el origen de dicho contraste, esto es, saber qué parámetro físico o químico es el que está siendo representado en forma de intensidad luminosa.

El concepto de resolución está relacionado al grado de detalle percibido en la imagen, puede ser espacial o temporal. La resolución espacial caracteriza la imagen desde el punto de vista de su capacidad para discernir detalles, se define como la distancia mínima que la imagen es capaz de resolver o separar y está relacionada con el tamaño del píxel utilizado para representar la imagen digital. La resolución temporal determina la capacidad del sistema de imagen para “congelar” situaciones en el tiempo, estando muy relacionado con la velocidad de adquisición de las imágenes.

La cuantización de niveles de gris es el número de valores de gris disponibles en la imagen y depende del número de bits usado durante la cuantización, n bits por píxel corresponden a 2^n intensidades de gris. En las imágenes médicas, las diferentes estructuras son representadas por los valores de intensidad disponibles, lo cual es determinado en parte por las propiedades de los tejidos como su grosor, densidad y composición química, adicionalmente a algunos aspectos del proceso de adquisición que pueden ser controlados, como la energía de los fotones de los rayos X. Se requiere de un suficiente número de valores distintos de tonos de gris para que aquello que se desea distinguir sea asignado a diferentes valores de intensidad [34].

No obstante, las imágenes no sólo contienen información de interés, sino también datos erróneos que las distorsionan. Esta información superpuesta no deseada se llama ruido si tiene carácter aleatorio o artefacto en caso contrario (Figura 2.1). Estos últimos suelen depender de problemas en la técnica adquisición de imagen, se podrían considerar errores sistemáticos, frente a los errores aleatorios que introduce el ruido.

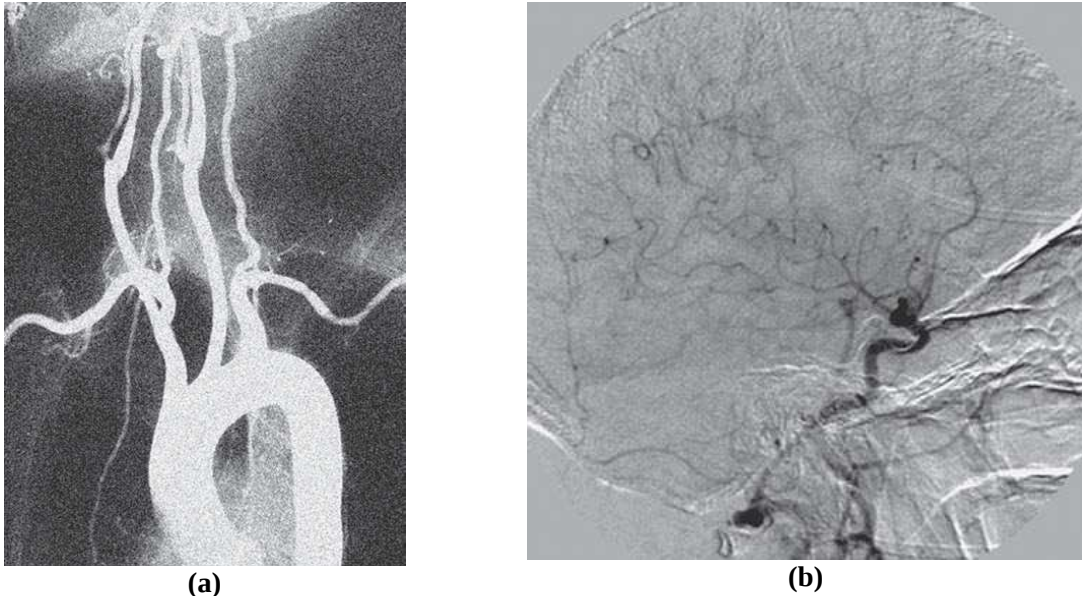


Figura 2.1 Información no deseada en imágenes médicas. (a) La imagen de flouroscofía presenta mucho ruido, en forma de un punteado aleatorio. (b) La angiografía presenta un artefacto debido al movimiento de la pared abdominal, que produce las líneas blancas que se observan en la imagen del hígado [34].

Por otro lado, un volumen de imágenes médicas se obtiene al apilar un conjunto de arreglos bidimensionales que representan los diferentes cortes anatómicos que conforman alguna región de interés, por ejemplo: la tomografía computarizada, la resonancia magnética, el ultrasonido, entre otras. Así, mientras el pixel es el elemento más pequeño de una imagen digital, el voxel lo es dentro del volumen de datos digitales. En el caso de algunas de estas modalidades, que producen imágenes tridimensionales, suelen presentarse otros artefactos durante la etapa de adquisición; el efecto de volumen parcial y la no uniformidad de la intensidad son ejemplo de ello [105].

El efecto de volumen parcial se presenta en las técnicas tomográficas, el cual es un error de cuantización que se introduce al asignar un cierto valor a un voxel que contiene varios tipos de tejido. Esto causa pérdida de definición en los contornos de diferentes órganos, reduciendo la exactitud y confiabilidad de las mediciones realizadas sobre esos contornos (Figura 2.2).

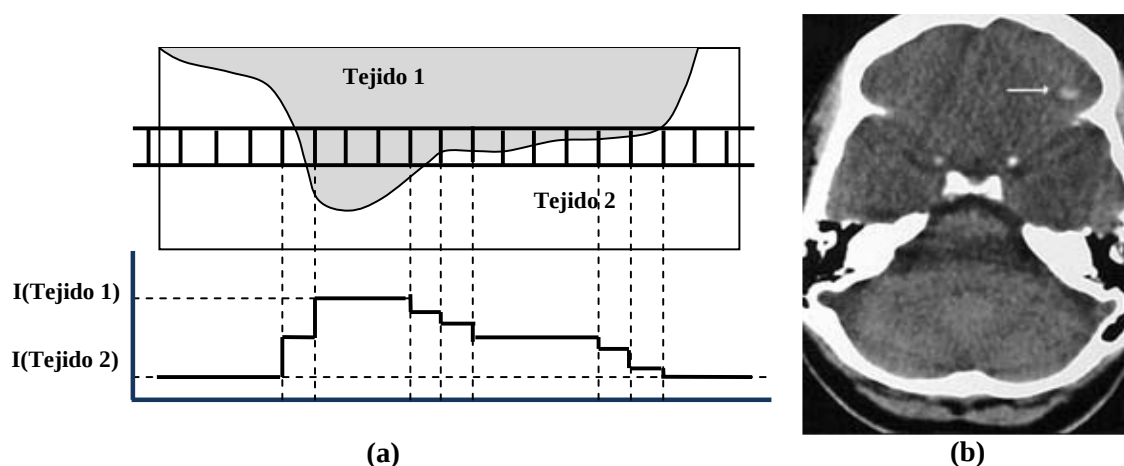


Figura 2.2 Información no deseada en imágenes médicas. (a) Vista 2D de un corte de una imagen donde se muestra el límite entre dos tejidos, el resultado de la intensidad de los voxels se muestra abajo [105]. (b) CT sin medio de contraste. Efecto de volumen parcial dado por techo orbitario (flecha) [81].

La intensidad no uniforme, también conocida como artefacto de sombreado, consiste en cambios locales suaves en la intensidad de la imagen; es un artefacto típico de la adquisición de imágenes de MR que depende de varios factores que incluyen: la forma y las propiedades electromagnéticas del objeto que se explore, la resolución espacial de la antena de radio frecuencia, la respuesta en frecuencia del receptor y las inhomogeneidades del campo magnético principal. Este artefacto puede ser un problema importante para los algoritmos automáticos que utilizan el nivel de gris de los píxeles/voxels, sin embargo, existen modelos multiplicativos de corrección de la no uniformidad, basados en la estimación del campo no uniforme (Figura 2.3).

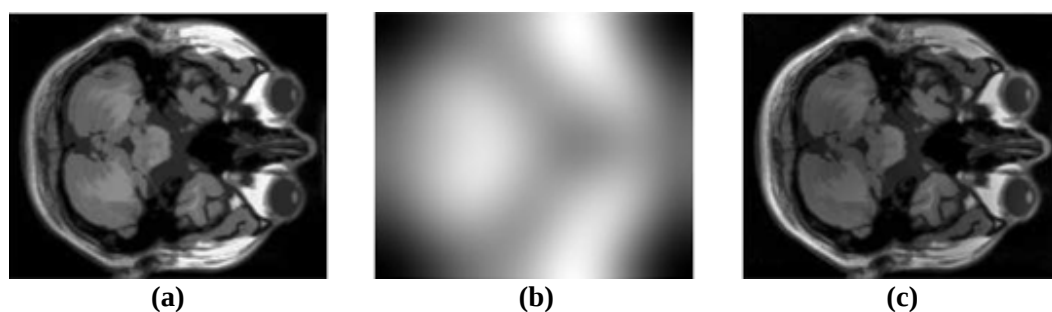


Figura 2.3 Corrección de la no uniformidad del campo magnético en MR. (a) Imagen original. (b) Estimación de la no uniformidad del campo (utilizando las técnicas basadas en PDE²). (c) Imagen reconstruida [55].

² PDE es el acrónimo de ecuaciones diferenciales parciales, las cuales se utilizan para modelar la no uniformidad del campo magnético en las imágenes de Resonancia Magnética.

2.2. Modalidades de adquisición de imágenes médicas

Los sistemas de imágenes médicas detectan diferentes señales físicas del paciente para producir imágenes. Se denomina modalidad a toda técnica de exploración realizada con un equipo de adquisición de datos susceptibles de ser utilizados para realizar un diagnóstico [107]. El elemento básico que define a cada modalidad es el tipo de energía utilizada.

El tipo de energía utilizada determina el tipo de interacción bioquímica que se produce en los tejidos biológicos y, por tanto, en qué medida puede ser nociva para el organismo. Se denominan radiaciones ionizantes aquéllas que por su alta energía son capaces de inducir directamente reacciones químicas, a través fundamentalmente de la ionización de diferentes moléculas. Las radiaciones no ionizantes se limitan a producir calentamiento que, en principio, no presenta efectos biológicos relevantes si es ligero [79].

Un criterio de clasificación de las modalidades es de acuerdo a la naturaleza del contraste de las imágenes, en base a ello se dividen en anatómicas (morfológicas) y funcionales (Figura 2.4). Las primeras se caracterizan por producir imágenes de muy buena resolución, que permiten una representación detallada de la anatomía del paciente. Las segundas, en cambio, se caracterizan por aportar información sobre el funcionamiento de los diferentes órganos o sistemas: algún rasgo de su metabolismo, su perfusión sanguínea, su capacidad para acumular ciertas sustancias, etc. [113].

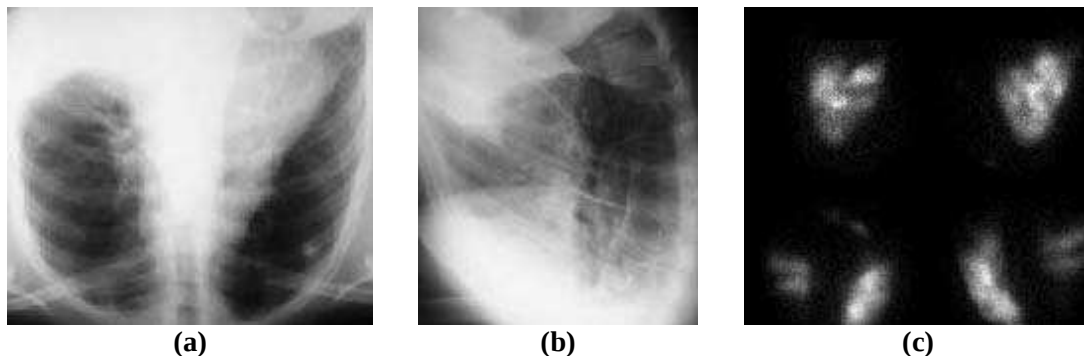


Figura 2.4 Imagen anatómica y funcional. (a) y (b) Radiografías de tórax. (c) Gammagrafía pulmonar, en la que se aprecia una clara embolia pulmonar, invisible en las radiografías de tórax [113].

Asimismo, las modalidades pueden dividirse en imágenes proyectivas y tomográficas según su capacidad para separar objetos a diferentes profundidades. En el primer caso se hace una superposición de todas las estructuras del objeto sobre una superficie bidimensional; mientras que en el segundo caso es posible obtener diferentes planos (Figura 2.5) que facilitan notablemente su interpretación.

Las fuentes de la imagenología médica se diferencian fundamentalmente por el tipo de información que es detectada. Cada modalidad da una representación diferente de los órganos o de su funcionamiento. La tabla 2.1 presenta una clasificación de las modalidades de imagen médica, indicando el tipo de energía utilizada y las características antes mencionadas. En las siguientes secciones se dará una breve descripción de las diferentes modalidades de la imagenología médica.

Tabla 2.1 Características de las modalidades de imágenes médicas.

Modalidad	Energía	Técnica de imagen	Ionizante	Funcional	Tomográfica
Radiología	Rayos X	Radiología convencional	Si	No	No
		Radiología digital	Si	No	No
		Tomografía computarizada (CT: Computed Tomography)	Si	No	Si
Medicina nuclear	Rayos γ	Tomografía computarizada por emisión de fotón gamma único (SPECT: Single Photon Emission Computer Tomography)	Si	Si	Si
		Tomografía por emisión de positrones (PET: Positron Emission Tomography)	Si	Si	Si
Ecografía	Ultrasonido	Ecografía	No	No	Si
Resonancia magnética	Ondas electromagnéticas en espectro de radiofrecuencia	Imagen de resonancia magnética (MRI: Magnetic Resonance Imaging)	No	No	Si
		Imagen de resonancia magnética funcional (fMRI: Functional Magnetic Resonance Imaging)	No	Si	Si
Endoscopía	Luz	Endoscopía	No	No	No

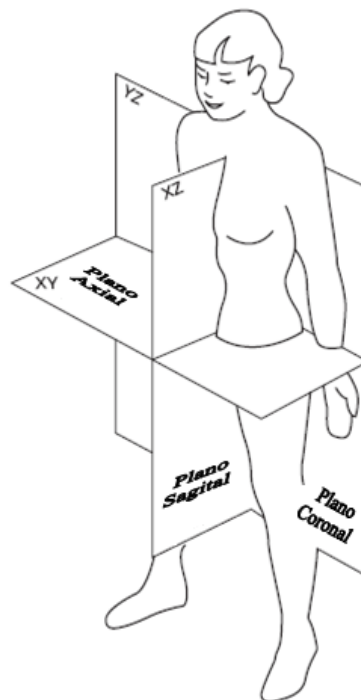


Figura 2.5 Las tres principales orientaciones de imagen para el cuerpo humano: los planos coronal, sagital y transversal (axial) [34].

2.2.1. Radiología convencional

Los rayos X son formas de radiación electromagnética ionizante cuyas propiedades físicas más importantes son su capacidad para atravesar la materia, producir fosforescencia e imprimir películas con emulsiones fotográficas. Los rayos X pueden también producir cambios en los tejidos biológicos y son capaces de producir la ionización de los materiales gaseosos. Debido a estas características, con frecuencia se utilizan para producir imágenes médicas y para desarrollar efectos terapéuticos en algunos pacientes. En la actualidad las modalidades de imágenes médicas basadas en la utilización de rayos X incluyen: radiografía convencional y digital, video-angiografía o fluoroscopia y tomografía.

La radiología convencional permite obtener sobre una película radiográfica la imagen de una parte del cuerpo humano expuesta a un haz de rayos X, en donde la absorción y dispersión que sufre el haz de rayos X depende del espesor de la sustancia atravesada y de los átomos que la constituyen. Dado que diferentes materiales tienen distintos coeficientes de atenuación, la visualización de las diferencias de densidad dentro de los tejidos o los órganos se logra al ubicar una película radiográfica detrás de la parte del cuerpo a estudiar en relación con la fuente de rayos X. Así, el origen del contraste en imágenes radiológicas proviene básicamente de las densidades de los distintos tejidos [73].

Si bien las características de las radiografías son dependientes de los aspectos técnicos relacionados con la potencia y el desplazamiento del sistema fuente detector, hay un conjunto común de atributos característicos que limitan la calidad de la imagen, tales como el hecho de que las imágenes tengan poco contraste debido al bajo valor del coeficiente de atenuación de la mayor parte de tejidos del cuerpo y adicionalmente que puedan estar afectadas por diversas fuentes de ruido difíciles de controlar. Por ello, las mejores imágenes se consiguen en presencia de hueso (aplicación en traumatología) y aire (imagen pulmonar). No obstante, la utilización de “medios de contraste”, que son sustancias densas (radio opacas), administradas mediante distintos tipos de sondas o por vía intravenosa, permiten visualizar tubo digestivo, vesícula, vías biliares y vasos sanguíneos (Figura 2.6). [113].

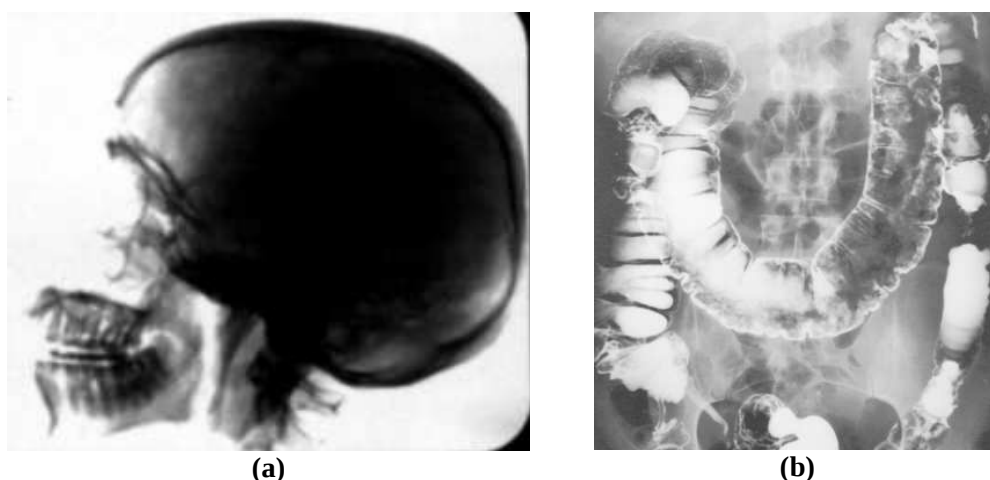


Figura 2.6 Imágenes de rayos X. (a) Radiografía del cráneo invertida. (b) Radiografía del abdomen utilizando bario como medio de contraste para visualizar el colon [48].

2.2.2. Radiología digital

La radiología digital es en esencia similar a la convencional, pero la placa fotográfica se sustituye por un detector electrónico conectado a una computadora. Básicamente consiste en un tubo emisor de Rayos X que estimula una superficie excitable, similar a la fotoestimulación de sales impregnadas en una película, pero por medio de una excitación electrónica. Posteriormente, un lector láser detectará los diferentes niveles de excitación y construirá una imagen radiológica.

Esta modalidad presenta muchas ventajas respecto a la anterior, entre ellas: la reducción de los tiempos de exposición (menos dosis de radiación al paciente para la misma calidad de imagen), la reducción de errores de transmisión de datos (rotulado de la placa, registro de datos del paciente, etc.), la disminución de los contaminantes al sustituir el uso de placas por archivos electrónicos, la posibilidad de compartir las imágenes en tiempo real y almacenarlas en medios informáticos permitiendo su recuperación para futura referencia sobre el mismo paciente [107].

2.2.2.1. Tomografía computarizada

La tomografía computarizada es también conocida por sus acrónimos TAC (Tomografía axial computarizada) y CT ('Computed Tomography') en donde las imágenes del interior del cuerpo se reconstruyen a partir de un conjunto de medidas de proyección, eliminando así el problema de la superposición de objetos en la imagen de radiológica convencional proyectiva [73, 107].

En CT un tubo emisor de rayos X y un detector giran describiendo un círculo alrededor del paciente. La medición de la intensidad de rayos X, que atraviesa un plano del cuerpo humano según un ángulo dado, da lugar a un perfil de proyección (corte o sección axial). La reconstrucción tomográfica se obtiene a partir de los diferentes perfiles de proyección obtenidos utilizando técnicas de reconstrucción de tipo analíticas o algebraicas. La calidad de la imagen depende del número de perfiles considerados (Figura 2.7).

Dado que el mecanismo de producción de contraste en la imagen es el mismo que para la radiología convencional, la CT adolece del mismo problema de falta de contraste entre tejidos blandos. Sin embargo, la imagen de CT presenta notables ventajas respecto a la imagen de rayos X; la imagen generada en este caso está formada por píxeles que, si bien aparecen en la pantalla de manera bidimensional, en realidad representan volúmenes por lo que se pueden considerar de carácter tridimensional. En la radiología convencional, la imagen analógica se obtiene detectando densidades distribuidas sobre la superficie excitada (placa), mientras que en los sistemas basados en CT, esas densidades pueden cuantificarse en una escala de grises mejorando la definición.

Debido a las características del método, se pueden estudiar prácticamente todas las estructuras del cuerpo humano. Además, la CT ha permitido el desarrollo de técnicas terapéuticas complementarias (técnicas intervencionistas) al permitir un abordaje más seguro por ejemplo en punciones y biopsias de abscesos, tumores, quistes, etc.

Dentro de los avances que ha tenido esta modalidad, los más importantes son la utilización de contrastes administrados por diferentes vías y la aparición de la CT de alta resolución, la CT helicoidal y la angio CT para la visualización de imágenes vasculares de alta calidad.

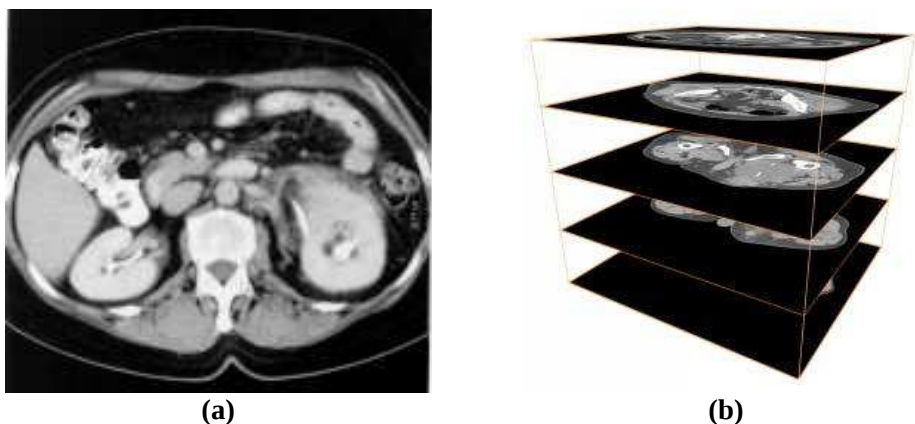


Figura 2.7 Imágenes de CT. (a) CT de alta resolución del abdomen. (b) Cortes axiales [48].

2.2.3. Medicina Nuclear

Se denomina Medicina Nuclear a la especialidad que utiliza imágenes funcionales para el diagnóstico, las cuales se obtienen mediante la detección de la radiación emitida por cada punto de un órgano luego de administrar al paciente una sustancia que incluye trazas de un radioelemento.

El fundamento de la Medicina Nuclear radica en marcar con algún átomo radiactivo determinadas moléculas (denominadas trazadores) cuyo comportamiento biológico se pretende seguir, de modo que su ubicación en el cuerpo pueda ser determinada localizando la radiación gamma emitida, proporcionando así información acerca de la capacidad de tal órgano de asimilar o transformar la sustancia que se ha inyectado; las imágenes obtenidas son, por tanto, un mapa de la distribución de dicho trazador en el organismo. Cada radioelemento es específico para el órgano y la función que se desea estudiar, por ejemplo el isótopo de yodo se utiliza para estudiar la glándula tiroides mientras que el thallium se usa para estudiar los tejidos cardiacos; no obstante, existe un número muy elevado de moléculas utilizadas como trazadores por lo que esta modalidad tiene aplicación clínica en diversas patologías [73].

El contraste de este tipo de imágenes está determinado por la concentración alcanzada por el trazador en los diferentes órganos y sistemas. Una característica importante de estas imágenes es su baja resolución espacial además de ser bastante ruidosas, lo cual se debe a la utilización de dosis muy bajas de radioelementos, que limitan mucho el número de fotones emitidos. Esta limitación es difícil de mejorar sin aumentar las dosis a los pacientes, pero esto implicaría que el método fuera más invasivo cuando precisamente se trata de reducir las reacciones adversas en el organismo al mínimo. Estos factores determinan que el aspecto visual de las imágenes de medicina nuclear sea de menor calidad que el de otras modalidades, sin embargo, esto se ve compensado con el tipo de información que pueden proporcionar sobre el funcionamiento de los sistemas biológicos [113].

En Medicina Nuclear diagnóstica se utilizan principalmente dos tipos de radioisótopos, que dan lugar a dos tecnologías de imagen tomográficas: la emisión de fotones gamma únicos (Single Photon Emission Computed Tomography: SPECT) e imágenes por emisión de positrones (Positron Emission Tomography: PET).

2.2.3.1. Tomografía por Emisión de Fotón Único (SPECT)

En las imágenes de SPECT la fuente de radiación es un emisor gamma producido de forma directa por un isótopo. Para detectar los fotones gamma se utiliza una cámara de rayos gamma típicamente conocida como gammacámara, la cual se hace rotar alrededor del paciente para obtener muchas proyecciones, estas últimas se adquieren en ángulos preestablecidos, típicamente cada 3-6 grados. Repitiendo el proceso para múltiples direcciones es posible determinar la distribución del radioelemento en el interior del cuerpo al utilizar técnicas de reconstrucción tomográfica. En la mayoría de los casos se realiza una rotación completa de 360 grados que permite una reconstrucción tridimensional óptima [113].

Esta técnica se emplea principalmente en diagnósticos de patologías neurológicas y cardiovasculares (Figura 2.8). Las limitaciones de esta técnica son: la necesidad de usar un sistema especial de rotación rápida alrededor del paciente, la pobre resolución espacial (la máxima resolución que se obtiene está en el rango del centímetro), la presencia de artefactos debido a radioisótopos presentes en la nasofaringe (cuando se administra el trazador por vía inhalatoria), y la contraindicación de su uso durante el embarazo.

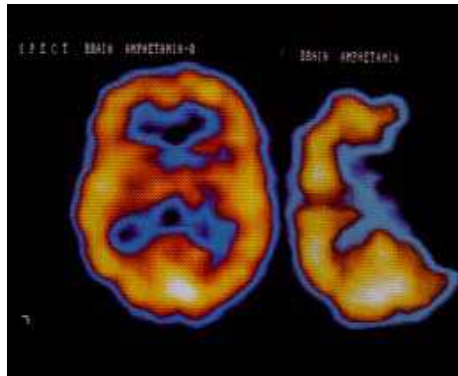


Figura 2.8 Estudio de SPECT del consumo de anfetamina en un cerebro normal y anormal [48].

2.2.3.2. Tomografía por Emisión de Positrones (PET)

En la imagenología por emisión de positrones (PET), el trazador emite positrones que son detectados de manera indirecta, pues cuando un positrón choca con un electrón emite dos fotones gamma que se propagan en direcciones opuestas. La detección casi simultánea de los dos fotones a ambos lados del paciente, permite localizar espacialmente la ubicación del choque de las partículas o aniquilación; dicha detección se realiza en un arreglo en forma de anillo alrededor del paciente constituido por fotodetectores [73].

Las imágenes de PET muestran regiones difusas y contornos de la zona explorada, no planos anatómicos en detalle. Las áreas con mayor intensidad de color corresponden a una mayor captación del radiofármaco administrado, indicando zonas que metabolizan la sustancia en mayor proporción (Figura 2.9). Sus aplicaciones principales se centran en la Oncología, Neurología y Psiquiatría, aunque también se ha probado su utilidad en Cardiología, Traumatología y Endocrinología.

Respecto a las imágenes de SPECT, esta modalidad produce imágenes de mejor resolución dado que el tipo de radiación que se detecta desde el exterior (dos fotones en sentido opuesto) es más exacta que la detección que se realiza en la gammacámara. Por otro lado, la mayor parte de los radiofármacos utilizados son sustancias habituales en la naturaleza, a diferencia de los emisores gamma, que suelen ser átomos pesados; esto permite marcar un mayor número de compuestos biológicos [107].

La principal limitación de esta modalidad consiste en requerir una infraestructura muy costosa que debe incluir un ciclotrón para generar sustancias radioactivas con un periodo de vida muy breve que está entre 5 minutos y media hora aproximadamente. También se requiere

de un laboratorio químico que permita la síntesis de sustancias asimilables por el organismo, las cuales son marcadas con el trazador radioactivo; y adicionalmente, del tomógrafo para realizar la adquisición y de un sistema computacional para procesar las imágenes y realizar la reconstrucción tomográfica [113]. A pesar de ello, en los últimos años se ha conseguido la obtención de imágenes llamadas “híbridas” que proveen información anatómica (CT) y funcional por medio de fusión PET/CT y SPECT/CT en equipos que cuentan con la integración de las dos modalidades en un sólo escáner y hacen posible la obtención de ambos tipos de imagen en una sola sesión, lo cual elimina el problema del registro de imágenes previo al proceso de fusión [112].

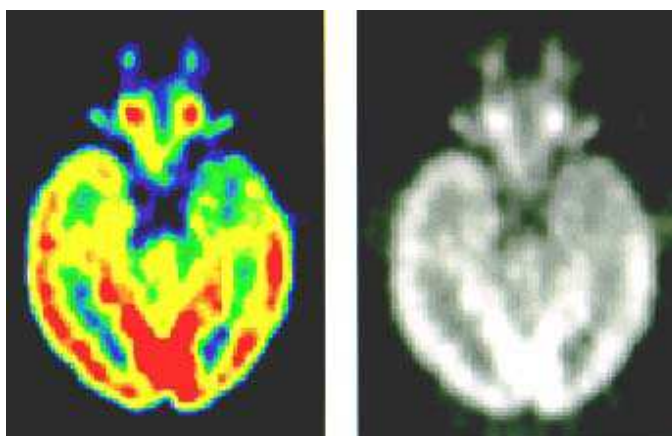


Figura 2.9 Estudio de PET con FDG³ asimilado en un cerebro epiléptico [48].

2.2.4. Ecografía

Esta modalidad de imagen es una técnica de exploración por ultrasonidos. El dispositivo transductor (barra piezoeléctrica que funciona como emisor y receptor) emite una onda ultrasónica de una frecuencia de varios MHz, en una orientación dada. La señal detectada corresponde a la superposición de ecos o reflexiones que se producen debido a los cambios de impedancia acústica en las diferentes fronteras de los órganos.

En sus inicios sólo se podían obtener imágenes en una dimensión (Modo M), más tarde en dos (Modo Bidimensional) y actualmente se pueden llevar a cabo ecos de alta resolución y tridimensionales (volumétricos). La imagen 2D se obtiene por el barrido de un haz según un plano cualquiera. Una de las dimensiones está dada por el barrido del haz, la otra dimensión corresponde al tiempo de retorno del eco. La fuente de contraste de la imagen es la densidad del tejido ya que lo que se representa en la imagen ecográfica son las transiciones entre tejidos de diferente densidad, por ello la ecografía proporciona imágenes especialmente útiles para separar interfases entre estructuras sólidas y líquidas [73]. Las imágenes de ultrasonido tridimensionales se obtienen agregando filas adicionales de elementos de cristal al transductor para permitir el barrido en una dirección perpendicular al plano del escaneo en modo B, así, el movimiento controlado del transductor hace posible combinar diferentes planos de corte [34].

³ La 18-Fluoro-deoxi-glucosa (FDG) es un radiofármaco que permite el estudio del metabolismo celular de la glucosa y es el más utilizado en la modalidad de PET.

Si bien la calidad de la imagen ecográfica es inferior en relación con otras modalidades, sus ventajas se fundamentan en el hecho que la velocidad de adquisición es elevada (15 a 30 imágenes por segundo), lo cual permite la exploración de órganos en movimiento en tiempo real, la naturaleza no ionizante del proceso de adquisición y el precio razonable de los equipos. Estas facilidades han contribuido a extender su uso en prácticamente todas las especialidades médicas para apoyar diversos diagnósticos, tanto en patología de adultos, como en Pediatría, variando la forma y el tamaño de las sondas de exploración de acuerdo al resultado esperado (Figura 2.10-a). Una variante de la técnica, llamada ecografía Doppler, utiliza el efecto físico del mismo nombre para proporcionar información sobre el flujo sanguíneo (Figura 2.10-b), ofreciendo así una nueva fuente de contraste, muy útil en ciertas enfermedades cardiacas.

Sus desventajas son la reducida resolución espacial que proporciona, su escaso contraste, el ruido inherente al proceso de adquisición y cierta dificultad de interpretación. Otro problema es la dificultad de penetración en el tejido (especialmente en el hueso, casi totalmente opaco al ultrasonido).

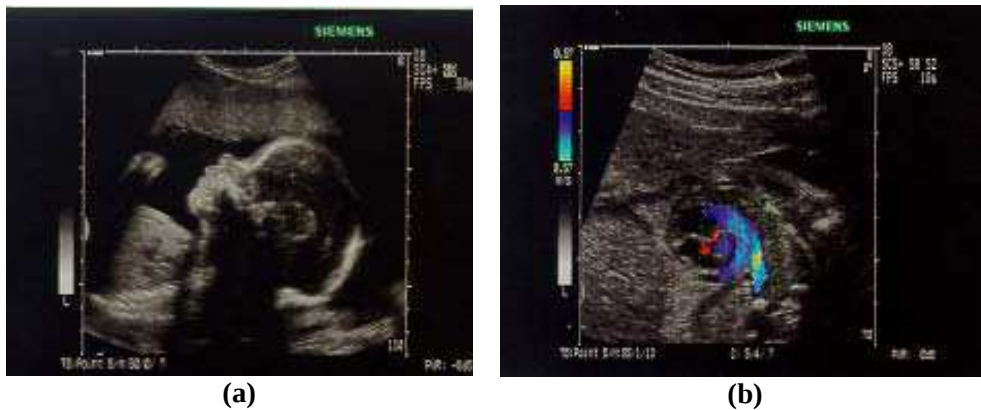


Figura 2.10 Imágenes de ultrasonido. (a) Escaneo que muestra el detalle de la cabeza de un feto. (b) Escaneo Doppler de un corazón fetal donde el mapa de colores superpuesto muestra la velocidad del flujo sanguíneo en una región seleccionada [48].

2.2.5. Resonancia magnética

La imagen de resonancia magnética (MRI) se basa en la capacidad de los electrones y los núcleos atómicos (generalmente núcleos de hidrógeno), con un número impar de protones y/o neutrones, para absorber ondas de radiofrecuencia procedentes de la estimulación de la materia previamente magnetizada por el efecto de un campo magnético. Esta capacidad permite generar una señal que luego de ser tratada en una computadora produce una imagen [107].

Una explicación más amplia del fenómeno físico al que obedece dicha modalidad es la siguiente: un núcleo sometido a un campo magnético H_0 se comporta como vector dipolo magnético que gira alrededor de la dirección de H_0 a una frecuencia conocida como frecuencia de Larmor, la cual es característica del elemento químico considerado y de la intensidad del campo H_0 . En presencia de un campo magnético H_1 , capaz de orientar al dipolo magnético

total en sentido perpendicular a H_0 , se produce un estado excitado de los núcleos atómicos. Al apagar el campo de radio-frecuencia H_1 el estado de excitación va decayendo progresivamente en función de la densidad protónica. El retorno al equilibrio está caracterizado por dos constantes de tiempo de relajación, longitudinal ($T1$) y transversal ($T2$). De manera que cada punto de la imagen es función de la posición en el espacio tridimensional, de la concentración de núcleos atómicos y de las constantes de tiempo de relajación [73].

En resumen, el funcionamiento clínico de estos sistemas consiste en la colocación de la muestra (el paciente) en el seno de un potente campo magnético constante, habitualmente generado mediante un electroimán superconductor. El volumen muestra se excita con impulsos de ondas de radio cuya frecuencia corresponde a la de resonancia del núcleo de hidrógeno, devolviendo después los tejidos esta energía que es captada desde el exterior por una bobina o antena [113].

La MRI tiene la ventaja de no utilizar radiaciones ionizantes, además de mostrarse más sensible que otras técnicas permitiendo una buena discriminación de los tejidos y la adquisición tridimensional de una zona del cuerpo; proporciona buena resolución espacial por la posibilidad de generar cortes más finos y en varios planos (Figura 2.11). Prácticamente no presenta artefactos debidos al aire o a los huesos; los artefactos usualmente se atribuyen a los defectos de los instrumentos, como la no uniformidad del campo y del gradiente, y a los movimientos del paciente durante la adquisición. En general, el contraste de este tipo de imágenes es excelente y muy variado, depende de diversos factores de tipo bioquímico y también de cómo se adquiere la imagen, incluso puede mejorarse adicionalmente mediante la utilización de materiales de contraste paramagnéticos inyectados en el cuerpo [34].

Otra característica de la MRI es que puede proporcionar imágenes anatómicas y también funcionales de diversos tipos (functional MRI, fMRI), y su abanico de indicaciones se amplía de día en día, invadiendo el terreno de otras modalidades. Para el cerebro, por ejemplo, se ha convertido en la modalidad de imagen preferida, por el buen contraste entre tejidos blandos y sus posibilidades de producir distintos tipos de imagen funcional, como es el caso de estudios de perfusión y de actividad cerebral.

Dentro de la evolución de las técnicas de adquisición de imágenes médicas, esta modalidad es la que ha presentado un avance más rápido, por ejemplo, la resolución se ha podido incrementar ya que a diferencia de otras modalidades no hay límite por dosis de radiación ionizante al paciente. La velocidad de adquisición de imagen también ha crecido hasta el punto de obtener imágenes cardíacas en movimiento. Sin embargo, los avances más interesantes en MRI se están produciendo debido a la introducción de nuevos mecanismos de contraste. Para la imagen anatómica se pueden generar diferentes contrastes, capaces de realzar o atenuar diferentes estructuras o tejidos (líquido, grasa, etc.), y capaces también de estudiar el flujo sanguíneo (angiografía por resonancia magnética, MRA), incluso sin utilizar medios de contraste. Otra posibilidad es la realización de análisis químico por resonancia magnética mediante espectroscopia de resonancia magnética (MRS), que abre la posibilidad de realizar biopsias incruentas [113].

Desde un punto de vista práctico, las limitaciones actuales más importantes de la MRI derivan de su elevado precio y en la contraindicación para pacientes con algunos tipos de implantes metálicos.

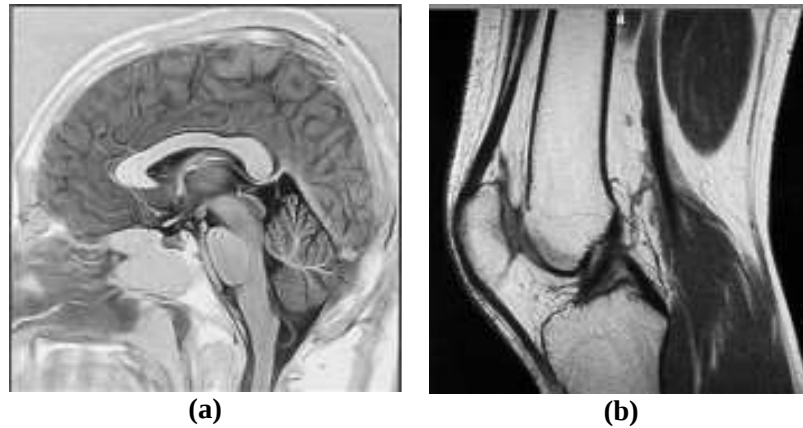


Figura 2.11 Imágenes de resonancia magnética. (a) Corte sagital de una MRI del cerebro. (b) MRI de los tejidos de la rodilla [48].

2.2.6. Imagenología óptica

La imagenología óptica utiliza sondas fluorescentes y bioluminiscentes que emiten radiación en las longitudes de onda visibles o infrarrojas para iluminar la parte del cuerpo en estudio. Los tejidos contienen pequeñas cantidades de sustancias que absorben la luz y consecuentemente emiten una luz de mayor longitud de onda, llamada fluorescencia que es una propiedad natural de los tejidos orgánicos (autofluorescencia). Las enfermedades causan cambios en la concentración de estas sustancias y ello modifica las propiedades fluorescentes de los tejidos, pudiendo ser detectados dichos cambios mediante un método llamado imagen óptica reflectante. Para obtener una mayor fluorescencia tisular se investigan diversas drogas vinculadas con los métodos usados en las imágenes ópticas llamados contrastes fluorescentes o fluorocromos. Con la técnica de la imagen óptica multispectral es posible discriminar los fluorocromos y se cuenta con endoscopios y laparoscopios capaces de detectarlos.

La ventaja que presenta esta modalidad radica en que la gama de colores de la luz permite analizar el contenido de diversas sustancias en la zona de observación, sin necesidad de usar agentes de contraste. No obstante, una gran limitación de las imágenes ópticas es la escasa penetración tisular de la luz, lo que limita el método a exámenes de superficie como la piel, la mucosas por vía endoscópica y la mama.

Hay equipos de imagen óptica diseñados como tomógrafos computarizados, con detectores sensibles a los fotones fluorescentes tisulares, esta técnica se denomina tomografía óptica de coherencia y es capaz de detectar objetos de tan sólo centésimas de milímetro ocultos en el interior de tejidos humanos, lo que es significativamente más pequeño que los detalles que pueden apreciarse con resonancia magnética o ultrasonidos [84].

2.2.7. La imagen molecular

La imagenología molecular se define como la caracterización y medición de los procesos biológicos a nivel celular y molecular en los seres vivos empleando técnicas de imagen mínimamente invasivas. La imagen molecular es el resultado de la exploración de esos procesos y rutas metabólicas en el que se utilizan moléculas marcadas con trazadores como fuente de contraste, empleando la sustancia cuya actividad biológica esté más relacionada, desde el punto de vista bioquímico, con la patología que se quiera analizar. El término implica la convergencia en un nuevo paradigma de imagen entre múltiples modalidades de adquisición, la bioquímica y la informática, a la vez que no denota una modalidad específica, sino que se aplica a cualquier técnica de imagen capaz de seguir con detalle procesos bioquímicos celulares [84].

La medicina nuclear es considerada la modalidad que dio lugar a las imágenes moleculares, pues son suficientemente sensibles como para ser utilizadas a nivel tisular, celular e incluso genético. La MRI por su parte ha sido utilizada como fuente de imagen molecular en estudios de cuantificación espectroscópica, de farmacocinética, de permeabilidad y difusión de agua, o incluso medidas de pH; también se han desarrollado agentes de contraste inteligentes que en presencia de una determinada enzima se activan o se desactivan, cambiando las propiedades paramagnéticas del entorno en el que se encuentran. La CT no es estrictamente un método de imagen molecular, no obstante, su buena resolución espacial en imágenes anatómicas ha propiciado su utilización en combinación con otras modalidades funcionales mediante un proceso de fusión. En cuanto a las imágenes por ultrasonido, los métodos de Doppler ofrecen la información de flujo y la valoración de la angiogénesis tumoral y la respuesta de éstos a las drogas antiangiogénicas, mientras que el ultrasonido de ultra alta frecuencia (más de 200 Mhz) puede servir para determinar propiedades mecánicas celulares.

2.3. Futuro de las imágenes médicas

La imagenología médica continúa evolucionando a pasos agigantados. Nuevos procedimientos diagnósticos ayudarán a identificar el genotipo o el fenotipo de anomalías *in vivo* haciendo posible el diagnóstico de ciertas patologías en etapas tempranas con métodos cada vez menos invasivos. Asimismo, el papel tradicional de la imagen como elemento exclusivamente diagnóstico se ha visto alcanzado por su aplicación como herramienta de auxilio en la terapéutica, a través fundamentalmente de la Radiología Intervencionista⁴ y de la denominada Cirugía Asistida por la Imagen⁵, a tal grado que en unos años dará paso a otro todavía más relevante como es la monitorización de terapias en la farmacogenómica gracias al ordenamiento de las expresiones genéticas y proteicas [113].

⁴ Se denomina Radiología intervencionista a la práctica médica que consiste en realizar intervenciones sobre el paciente mediante catéteres, introducidos generalmente por vía vascular, utilizando la guía de imágenes médicas obtenidas en tiempo real.

⁵ La cirugía asistida por la imagen implica una aplicación muy directa de los métodos de registro o fusión de imagen multimodalidad antes mencionados. Resulta especialmente útil en neurocirugía, donde los sistemas de neuronavegación permiten registrar (hacer coincidir en el espacio) en tiempo real imágenes del paciente (CT o MRI) con el espacio tridimensional real (campo quirúrgico).

En general, es fácil prever una evolución de todas las modalidades de imágenes médicas a fin de obtener cada vez mejor calidad en las imágenes. Los ultrasonidos, la CT y la fluoroscopia digital servirán para guiar microcatéteres en zonas anatómicas objetivo para enviar a su través inhibidores angiogénicos; una vez dentro explotarán y soltarán en el interior su carga terapéutica cuando encuentre un disparo ultrasónico, de forma que la terapia génica pueda evitar los muchos peligros moleculares que de otra forma se encontrarían en su tránsito. La MRI dará excelentes contrastes en el registro de imágenes, suficientes para evaluar imágenes morfológicas y de perfusión, identificando espectroscópicamente en casos de diagnóstico de cáncer y en su monitorización terapéutica [107].

Los avances esperables en los próximos años en la imagen nuclear pasan por el desarrollo de nuevos trazadores, tarea en principio más fácil para radiofármacos emisores de positrones, por las razones antes explicadas. A corto plazo y en términos más cercanos al paciente promedio, el avance más importante que podríamos esperar es el abaratamiento de los estudios con PET, cuyo elevado precio actualmente dificulta su utilización [113]. Por otro lado, la imagenología molecular se considera en la actualidad la tecnología del futuro, ya que permitirá la detección precoz y las características de la enfermedad así como una valoración de la eficacia de los tratamientos.

En cuanto a los equipos de adquisición de imágenes médicas, los fabricantes se están centrando en las modalidades híbridas de imágenes con variaciones en el diseño, usos específicos para algunas aplicaciones y sistemas de utilización simplificados. Ejemplo de ello son los escáneres de tomografía de emisión de positrones/tomografía computarizada (PET-CT) que combinan la imagenología anatómica dada por la CT con la confirmación de la actividad molecular y celular adquirida con la técnica de PET. Su implementación provee información simultánea de la vascularización y del metabolismo celular de la glucosa permitiendo una nueva visión sobre la patofisiología de la isquemia cerebral y la agresión tumoral. Otro sistema híbrido es la SPECT-CT, que combina la sensibilidad funcional de la SPECT con el detalle anatómico de la CT multicorte.

Fusión de imágenes médicas

3.1. Fundamentos de fusión de imágenes

La fusión de datos combina los datos de múltiples sensores para el mejoramiento de aproximaciones e inferencias más específicas, tareas que no podrían llevarse a cabo mediante el uso de una sola fuente de información. Este concepto es aplicado por los seres vivos [67], pues ellos tienen la capacidad de usar sus sentidos para captar información sobre el ambiente, después el cerebro hace la fusión de la información que tiene disponible y toma una decisión.

Los principios de fusión de datos son aplicables a diversos dominios como es el caso del procesamiento de imágenes. En este sentido, la fusión de imágenes puede proveer mejores capacidades de interpretación y resultados más confiables debido a la combinación de datos con diferentes características; como las imágenes varían en resolución espectral, espacial y temporal, mediante la fusión se pretende llegar a una visión completa de los objetos de la escena.

3.1.1. Conceptos de fusión

La fusión de datos abarca un amplio dominio de aplicaciones y métodos por lo que es muy difícil proveer una definición precisa y en ocasiones suele confundirse con otros términos que se refieren al mismo concepto pero en un sentido más general, tal es el caso de la combinación y mezcla de datos [114, 117]. No obstante, existen varias definiciones que pueden ser encontradas en la literatura.

Una de las primeras definiciones de fusión de datos fue enunciada por el Departamento de Defensa de Estados Unidos (1991) [28], estableciendo así que “la fusión de datos es un proceso multinivel y multifacético que trata con la detección automática, asociación, correlación, estimación y combinación de datos de múltiples fuentes”. Este concepto fue redefinido más tarde en [58] como “proceso multinivel y multifacético que trata con la detección automática, asociación, correlación, estimación y combinación de datos *de una o múltiples fuentes*”; definición que es más general con respecto a los tipos de información que pueden ser combinados. Li et al [61] escribió que la fusión se refiere a la combinación de un grupo de sensores con el objetivo de producir una sola señal de mejor calidad y ‘reliability’. De acuerdo con Hall [49] la fusión de datos es un proceso que trata con datos e información de múltiples fuentes que mejora o detalla la información para la toma de decisiones. Mangolini [68] extendió la fusión de los datos a la información en general y también se refiere a la calidad; él define la fusión de datos como “conjunto de métodos, herramientas y medios

usando los datos que se obtienen de varias fuentes de diversa naturaleza, para aumentar la calidad (en un amplio sentido) de la información requerida”. Por su parte, Genderen y Pohl [88] dieron una definición general de fusión de imágenes la cual la describe como la combinación de dos o más imágenes diferentes para formar una nueva imagen mediante un cierto algoritmo.

Tomando como referencia los trabajos de Buchroithner [14] y Wald [116] la siguiente definición fue adoptada en enero de 1998: "la fusión de datos es un marco formal en el cual están expresados los medios y las herramientas para la asociación de los datos originales de fuentes diversas. Tiene como objetivo el obtener información de mayor calidad, donde la definición exacta de 'mayor calidad' dependerá de la aplicación."

Así, en el contexto de las imágenes, la fusión se refiere al proceso de combinar dos o más imágenes de una misma escena con información complementaria o redundante para generar una nueva imagen compuesta, la cual integra la información contenida en las imágenes individuales. El resultado es una imagen que tiene un alto contenido de información y una mejor calidad comparado con cualquiera de las imágenes de entrada. La meta del proceso de fusión es evaluar la información en cada pixel de las imágenes de entrada y retener la información de la imagen que mejor representa el verdadero contenido de la escena o realza la utilidad de la imagen fusionada para una aplicación particular.

La **fusión de imágenes médicas** entonces se define como la combinación de imágenes que proceden de diversas modalidades en una sola imagen con mayor cantidad de información, con la finalidad de presentar de manera conjunta e integrada toda la información (anatómica y funcional) para simplificar su interpretación; y dado que las imágenes empleadas en el presente trabajo son tridimensionales, el concepto se extiende a la fusión de volúmenes (Figura 3.1).

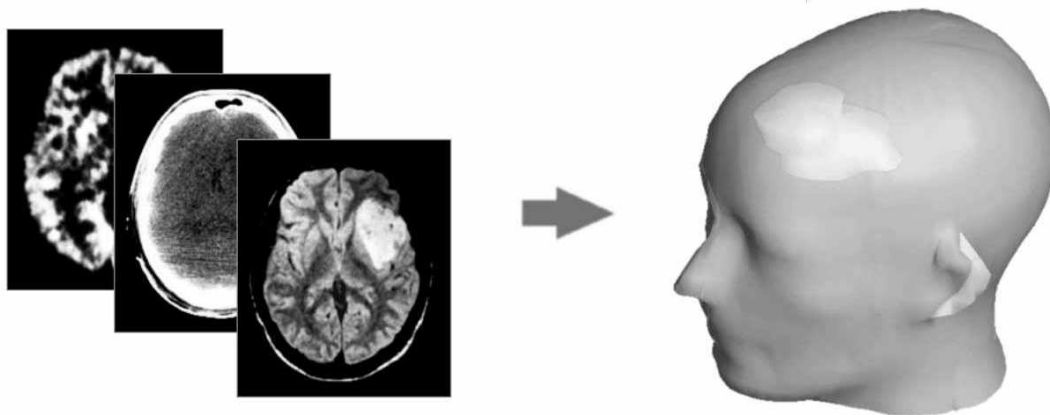


Figura 3.1 Datos de múltiples modalidades, registrados en un sistema de referencia común de un paciente, los cuales se combinan para construir una representación integral del paciente [57].

3.1.2. Propósito de la fusión de imágenes médicas

En general, el propósito del proceso de fusión es integrar diferentes datos con el fin de obtener más información que la que puede ser derivada de cada sensor de datos individualmente. Así, la fusión de imágenes propone la integración de datos dispares y complementarios para mejorar la información que aparece en las imágenes tanto como el incremento de confiabilidad e interpretación, lo cual deriva en una mayor precisión de los datos. También, la fusión de datos provee una tarea operacional robusta, por ejemplo, incremento en la confiabilidad, reducción de ambigüedad, mejora en la credibilidad y en la clasificación.

En la literatura [88] se ha encontrado que la fusión de imágenes es aplicada a la imagenología digital con los siguientes fines:

- Realce de imágenes. La fusión de imágenes puede ser usada como una herramienta para incrementar la resolución espacial.
- Mejorar las correcciones geométricas (precisión del registro). La fusión ofrece la ventaja de una aproximación de rectificación integral que mejora la precisión en el registro ya que la localización de los puntos de control para el registro se facilita si pueden ser ubicados en vistas similares.
- Mejorar ciertas características no visibles en ninguno de los datos por sí solos.
- Complementar conjuntos de datos para ayudar a la clasificación. Si se tienen datos de diferentes fuentes se puede hacer una mejor discriminación de las clases
- Detección de cambios en los datos multitemporales. La combinación de las imágenes temporales mejora la información sobre cambios que pudieron haber ocurrido en el área observada.
- Sustitución de información omitida en una imagen con señales de otro sensor.
- Reemplazo de datos defectuosos.

En imagenología médica, el propósito de la fusión es el mejoramiento de la calidad asistencial ante procesos de diagnóstico y planeación: i) facilitando y adelantando el diagnóstico mediante el uso de técnicas de fusión imagen funcional/anatómica, ii) aumentando la precisión de la planificación y dosimetría en radioterapia, y iii) permitiendo un control de calidad del tratamiento radioterápico basado en mediciones objetivas.

Por ejemplo, MR provee excelente contraste en tejido suave, permitiendo la delineación de tejidos normales y volúmenes que son tumores en diversos sitios. SPECT y PET proveen información metabólica única capaz de resolver ambigüedades en una imagen anatómica y pueden cuantificar la función parcial de un órgano. Las imágenes de ultrasonido pueden proporcionar información volumétrica en tiempo real para la delineación de bordes de órganos tanto en planeación de tratamiento como en la aplicación del mismo. Los datos de estas modalidades ayudan en conjunto a realizar una descripción más completa del paciente. Cada una de esas modalidades por sí misma, sin embargo, no provee toda la información física y geométrica requerida para la planificación del tratamiento terapéutico [57]. Por ello, el objetivo de la fusión de datos médicos es tomar ventaja de la información de cada modalidad integrándolas de tal forma que se reduzcan los errores de interpretación (Figura 3.2).

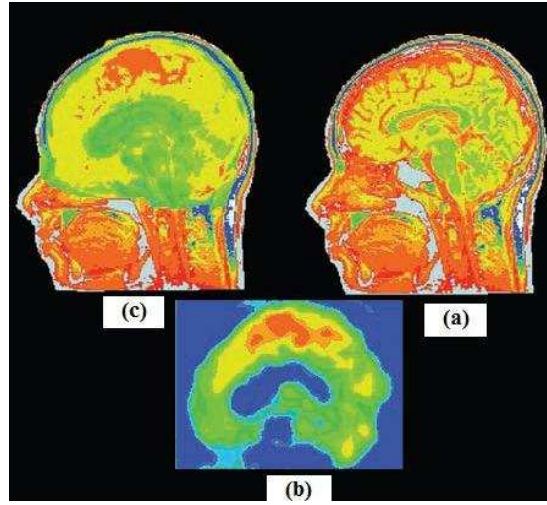


Figura 3.2 Aplicación de fusión de imágenes médicas para la localización de lesiones cerebrales. (a) MRI y (b) SPECT del mismo paciente. (c) Imagen co-registrada, MRI + SPECT. La lesión cerebral es más prominente en la imagen compuesta que en las modalidades individuales [34].

En la actualidad, existen ya algunas aplicaciones para los métodos de fusión de imágenes médicas y se observa una tendencia en aumento, ya que la fusión de imágenes en este ámbito va más allá de reducir el costo de almacenamiento, pues provee de mayor información al médico para la toma de decisiones tanto en el diagnóstico como el tratamiento de enfermedades. Por ejemplo, para la planeación 3D de terapias de radiación [85], mientras que CT y MR son las principales fuentes de imágenes, las modalidades funcionales como PET y SPECT proveen información importante para la identificación de lesiones en un marco estructural anatómico.

3.2. Técnicas de fusión para el despliegue de imágenes médicas

Desde el punto de vista de la visualización, una aplicación que despliega la combinación de imágenes de múltiples modalidades debe permitir una manipulación rápida con una calidad aceptable y fácil de interpretar. Por otro lado, las estaciones de trabajo con capacidades de despliegue limitadas a 8 bits en el caso de imágenes médicas imponen severas restricciones sobre las técnicas que presentan volúmenes fusionados [94]. En este contexto, la combinación de varias modalidades de imágenes médicas se puede llevar a cabo por distintos métodos: adición, despliegue simultáneo, superposición de bordes y estructuras, transparencia y combinación jerárquica [106] (Tabla 3.1). El principio básico de estos métodos (IM: Image Merging por sus siglas en inglés) para dos imágenes A y B es el siguiente:

$$IM = f(A(i, j), B(i, j)) \quad 1 \leq i \leq N; 1 \leq j \leq N \quad (3.1)$$

donde f es la función de combinación y N el tamaño de la imagen.

Tabla 3.1 Métodos de combinación de imágenes médicas para el despliegue

Método	Descripción
Adición	$IM = A(i, j) + B(i, j)$ Un ejemplo de este método es el de <i>Codificación de color: Rojo, Verde y Azul (Red, Green, Blue: RGB)</i> [94]. Esta codificación usa los tres colores primarios para representar cada par o trío de volúmenes. El color del pixel desplegado es una muestra <i>aditiva</i> de dos (o tres) colores. La información del registro espacial está disponible para el usuario, sin embargo, el usuario debe determinar la contribución de pixeles de los volúmenes individuales mediante matiz y brillo. Esto puede hacer que el método sea complicado para la combinación de imágenes anatómicas y funcionales; pues es importante relacionar el nivel de actividad a la localidad espacial.
Despliegue simultaneo o paralelo	$IM = Sim(A(i, j), B(i, j))$ donde Sim es una función de despliegue simultaneo para imágenes con coordenadas comunes. En esta técnica los cortes de dos o más volúmenes se despliegan con escalas de colores compartidas o independientes [94]. Las conexiones espaciales entre cortes correspondientes están hechas de una malla de marcas, puntos simultáneos o despliegue simultaneo de regiones de interés.
Superposición de bordes y estructuras	$IM = A(i, j) + Edg(B(i, j))$ donde Edg es una aproximación para la extracción de bordes.
Transparencia o Composición de color	$IM = (1-t)A(i, j) + (t)B(i, j)$ donde t es el coeficiente de transparencia. La composición de color [94] (también conocida como composición ponderada de opacidad o transparencia) es una técnica de cómputo gráfico usado para componer imágenes de objetos superpuestos de varias opacidades. Este método requiere un nivel adicional de codificación para su representación, cada volumen tiene un nivel de opacidad que varía entre 0 y 1. La composición de color hace uso de la correlación espacial de dos volúmenes de manera que su manipulación permite al usuario cambiar el énfasis, dando a cada volumen los valores extremos para visualizar a cada uno en su forma pura.
Combinación jerárquica	El principio de esta aproximación tiene tres etapas: la transformación de ambas imágenes en un conjunto de patrones primitivos elementales; la combinación de éstos para obtener un solo conjunto y la reconstrucción de una imagen fusionada partiendo de este último conjunto.

Cada uno de estos métodos presenta resultados interesantes [106] ya que todos ellos permiten evaluar la información concerniente al contraste y estructuras de bordes (Figura 3.3). No obstante, el método de adición no siempre puede preservar detalles importantes. El método de superposición depende del algoritmo de extracción de bordes. Los métodos de despliegue simultáneo y transparencia no generan una sola imagen pero a cambio de ello preservan todos los detalles en ambas imágenes. Finalmente, en el método jerárquico las dificultades radican en la regla para combinar ambos conjuntos de patrones primitivos elementales. No obstante, en esta tesis se expone una técnica de fusión que corresponde al rubro de combinación jerárquica, donde el reto ha sido precisamente encontrar una regla de decisión que en conjunto con el método de análisis de datos (la transformada de Hermite) integre la información de las modalidades manteniendo una buena calidad perceptiva sin introducir artefactos; por ello, en las siguientes secciones la fusión de imágenes se limita a este último método.

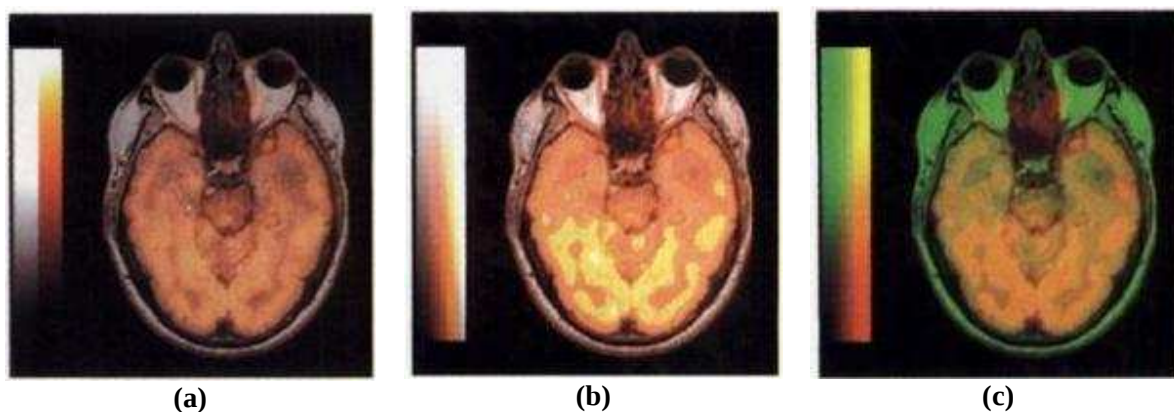


Figura 3.3 Cortes axiales de MR y PET presentados con métodos alternativos. (a) Imagen intercalada, por ejemplo, MR en escala de grises y PET en escala de colores hot-body. (b) Composición de color, por ejemplo, MR en 30 tonalidades de grises y PET en 8 tonalidades de la escala de color hot-body con un matiz uniforme. (c) Codificación RG, por ejemplo, MR codificada en rojo lineal y PET en verde lineal [94].

3.3. Contexto genérico para esquemas de fusión de imágenes

Un sistema típico de fusión de imágenes puede dividirse en seis subsistemas: adquisición de imágenes, registro, pre-procesamiento, fusión, post-procesamiento y despliegue como se muestra en la Figura 3.4. El procedimiento general para la fusión de imágenes [78] es el siguiente: primero, la información reunida de múltiples fuentes de adquisición entra a un proceso de registro, que consiste en asociar los píxeles correspondientes de los mismos puntos físicos del objeto. En esta forma, las imágenes de entrada pueden ser comparadas pixel a pixel. Subsecuentemente, los datos registrados pueden entrar a una etapa de pre-procesamiento (en caso de ser necesario) para la reducción de ruido. Después las imágenes pre-procesadas entran al proceso de fusión; dependiendo del algoritmo empleado para ello, el resultado de esta etapa puede ser una imagen compuesta, un mapa temático, o un gráfico tridimensional. Usualmente, cuando la fusión es implementada a nivel pixel, una etapa de post-procesamiento puede ser aplicada a la imagen fusionada, como clasificación, segmentación o restauración. En esta forma es posible alcanzar resultados cuantitativos, los cuales pueden ser evaluados en la base de métricas predefinidas de acuerdo a los requerimientos de la aplicación. Los resultados

cuantitativos de la evaluación pueden ser usados como retroalimentación para guiar el proceso de fusión dependiendo de la aplicación.

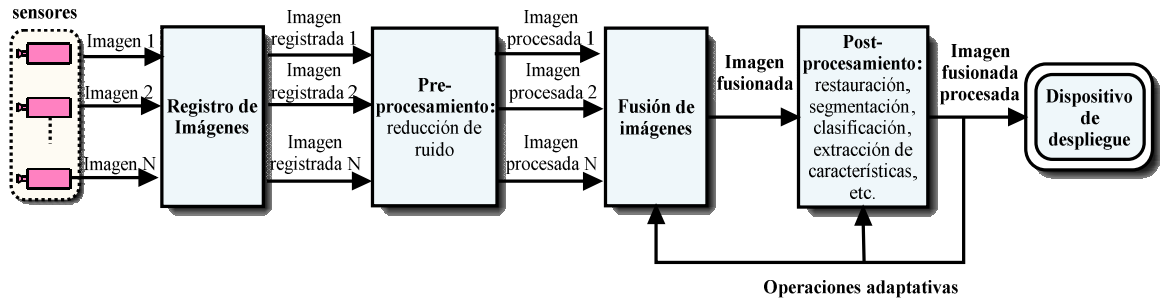


Figura 3.4 Diagrama de bloques de un procedimiento general de fusión de imágenes

Los subsistemas de imágenes están compuestos de múltiples (normalmente dos) sensores de imágenes; en imágenes médicas por ejemplo, estos sensores se refieren a alguna modalidad de adquisición. Dichos sensores, en la mayoría de los sistemas de fusión de imágenes tienen diferentes campos de vista, resoluciones, distorsiones de lente y tasas de muestreo, por lo que resulta de vital importancia alinear las imágenes de entrada apropiadamente, espacial y temporalmente, lo cual implica resolver el problema del registro de imágenes. El registro de imágenes requiere una conversión o transformación de las imágenes; este concepto de alineación asume un “lenguaje común” entre los datos de entrada, por ejemplo: i) estandarización de las unidades de medida; ii) calibración de los sensores; o iii) correcciones para la iluminación y el sombreado [67]. Dada la complejidad del registro de imágenes, el objetivo de este trabajo se enfoca en el subsistema de fusión; en consecuencia, se asume que las imágenes empleadas para la evaluación de la técnica propuesta ya están registradas.

El pre-procesamiento antes o después de la fusión es una etapa generalmente deseable ya que la mayoría de los algoritmos de fusión resaltan los artefactos de los datos originales, obteniendo así resultados indeseables en la última etapa. En cambio, el post-procesamiento dependerá del tipo de despliegue y del propósito del sistema de fusión.

Finalmente, el subsistema de despliegue muestra la imagen fusionada al usuario final. La calidad de esta imagen de salida dependerá en gran parte del proceso de fusión utilizado, del post-procesamiento opcional y del tipo de despliegue.

3.4. Niveles de fusión de imágenes

La fusión de datos tiene lugar en tres niveles: pixel, características, y decisión [88] (Figura 3.5). En la fusión a nivel pixel las imágenes de diferentes fuentes se combinan pixel a pixel, razón por la cual este tipo de fusión representa el nivel mínimo de procesamiento para fusionar parámetros físicamente medidos. En este nivel, existen algunos requerimientos generales [97] impuestos al resultado de la fusión: i) el proceso de fusión debe preservar, hasta donde sea posible, toda la información sobresaliente de las imágenes fuente; ii) el proceso de fusión no debe introducir artefactos o inconsistencias; iii) el proceso de fusión debe ser invariante a la

traslación. En general, los algoritmos empleados en este nivel son más fáciles de implementar y más eficientes desde el punto de vista computacional, por ello casi todos los algoritmos de fusión pertenecen a este nivel [129], y es precisamente en esta categoría en la que se encuentra el algoritmo de fusión tridimensional de esta tesis.

La fusión a nivel de características requiere la extracción de características relevantes en las imágenes de entrada; estas características pueden ser intensidades de píxeles, bordes o texturas dependiendo de la naturaleza de las imágenes. Las características similares son comparadas entre sí y entonces son fusionadas para futuras evaluaciones usando aproximaciones estadísticas o redes neuronales.

Por último, la fusión a nivel de decisión permite la combinación de imágenes en el nivel de abstracción más alto. Las imágenes de entrada son procesadas de forma independiente y sus resultados se combinan utilizando reglas de decisión basadas en lógica difusa, métodos estadísticos, etc. La información obtenida a partir de combinar las imágenes, aplicando reglas de decisión, permite reforzar una interpretación común, resolver diferencias y proporcionar un mejor entendimiento del objeto observado. Dependiendo de qué método se utilice, la fusión se clasifica en fusión dura o fusión suave [59].

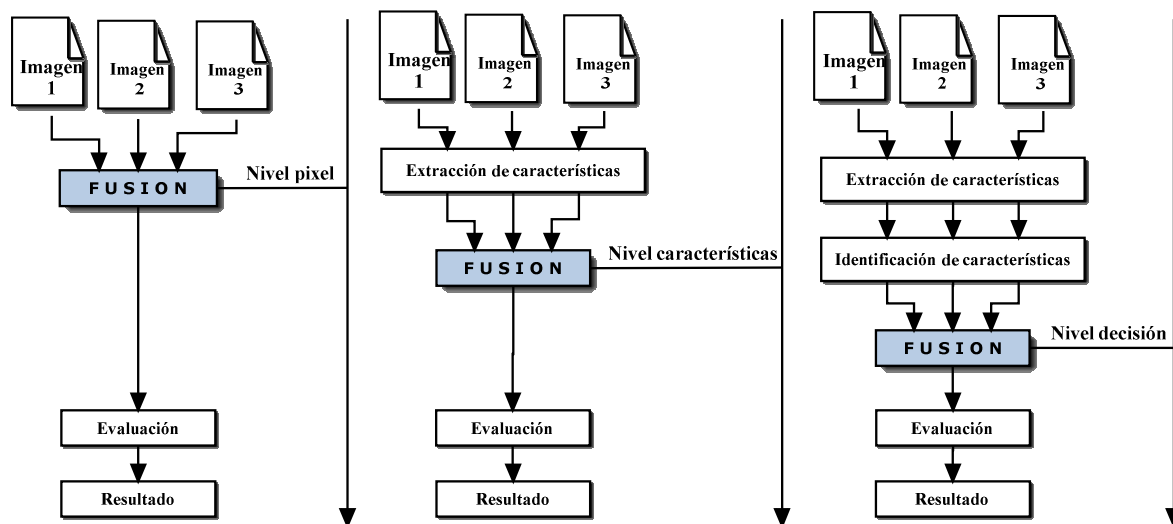


Figura 3.5 Niveles de fusión de imágenes [88].

3.5. Técnicas de fusión de imágenes

Existe una gran cantidad de algoritmos propuestos para resolver el problema de fusión de imágenes, y en especial para hacer la fusión a nivel pixel y con frecuencia con adaptaciones específicas para problemas particulares. Debido a ello es posible encontrar en la literatura diversas categorizaciones de algoritmos de fusión.

- Schowengerdt [98, 93], por ejemplo, los clasificó en técnicas del dominio espectral, técnicas de dominio espacial y técnicas en el espacio de la escala, dependiendo del dominio en el cual es llevado a cabo el proceso de fusión.

- Pohl, et al. [88] los categorizó en dos grandes grupos en relativas al color y en estadísticas/numéricas. Las primeras utilizan los tres canales de imágenes en el espacio de color RGB (p. ej. IHS y HSV). Las técnicas estadísticas están desarrolladas sobre la base de un canal estadístico que incluye correlación y filtros (p. ej. PCA y regresión); mientras que los métodos numéricos realizan operaciones aritméticas como sumas, diferencias, promedios, entre otros.
- Timo Bretschneider y Odej Kao [13] organizó a los diferentes métodos en grupos con respecto a su principal idea de diseño en técnicas por transformación, suma y multiplicación, filtrado, por relaciones inter-banda (correlación) y descomposición wavelet.
- Ranchin y Wald [93] distinguen entre 3 categorías: proyección y sustitución, contribuciones espectrales relativas, y relevantes al concepto de ARSIS (métodos multiescala).
- Zhijun Wang et al. [124] propusieron un método de fusión de imágenes general (GIF por sus siglas en inglés) para categorizar, comparar y evaluar los métodos de fusión existentes.
- Rick S. Blum, Zhiyun Xue, y Zhong Zhang [95] dividen a las técnicas de fusión en dos grandes grupos, métodos basados en descomposición multiescala y métodos basados en la no descomposición multiescala (NMDB por sus siglas en inglés).
- Yang, et al. [132] por su parte, clasifica las diferentes técnicas en modelos del espacio de color, algoritmos estadísticos/numéricos, algoritmos de descomposición multirresolución, y algoritmos radiométricos/espectrales.

A continuación se expondrán las principales técnicas de fusión de imágenes en el siguiente orden:

- Intensidad-Matiz-Saturación (IHS: Intensity Hue Saturation)
- Combinaciones aritméticas {
 - ⇒ Combinación ponderada
 - ⇒ Promedio
- Análisis de componentes principales PCA
- Técnicas multirresolución {
 - ⇒ Pirámide Gaussiana
 - ⇒ Pirámide laplaciana
 - ⇒ Pirámide de contraste
 - ⇒ Pirámide de gradiente
 - ⇒ Pirámide morfológica
 - ⇒ Transformada Wavelet
 - ⇒ Transformada de Hermite
- Métodos de selección

3.5.1. Intensidad-matiz-saturación (IHS)

El modelo de color intensidad-matiz-saturación (IHS: intensity-hue-saturation) [90] es una manera alterna de describir los colores mediante sus componentes RGB. Los componentes de matiz y saturación están íntimamente relacionados con la forma que percibimos el color. El matiz es un atributo cromático que describe un color puro (amarillo puro, rojo puro), mientras que la saturación proporciona una medida del grado en que un color está diluido en luz blanca. La importancia de la componente de intensidad, I , es que está desacoplada de la información cromática contenida en la imagen, lo que permite que pueda ser sustituida por otra información.

Esta transformación separa la información espacial (I) y la espectral (H, S) de una imagen RGB estándar y está relacionada con la percepción humana del color. La ecuación de transformación se expresa en (3.2), (3.3) y (3.4) [89].

$$\begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.2)$$

$$H = \arctan \left\{ \frac{v_2}{v_1} \right\} \quad (3.3)$$

$$S = (v_1^2 + v_2^2)^{1/2} \quad (3.4)$$

El uso de las técnicas IHS de fusión de imágenes se basan en el principio de reemplazar una de las componentes (I, H o S) de un conjunto de datos con el de otra imagen. Lo más común es que el canal de intensidad sea sustituido. La transformación inversa de IHS a RGB convierte los datos en el espacio de la imagen original para obtener la imagen fusionada, esta transformación está definida por (3.5), (3.6) y (3.7).

$$v_1 = S \cos\{H\} \quad (3.5)$$

$$v_2 = S \sin\{H\} \quad (3.6)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & \frac{-\sqrt{6}}{6} & \frac{\sqrt{6}}{2} \\ 1 & \frac{\sqrt{6}}{6} & \frac{-\sqrt{6}}{2} \\ 1 & \frac{\sqrt{6}}{3} & 0 \end{bmatrix} \begin{bmatrix} I \\ v_1 \\ v_2 \end{bmatrix} \quad (3.7)$$

Un ejemplo de la aplicación de este método en imágenes médicas es la técnica denominada ‘codificación en matiz y brillo’, donde el matiz, la saturación y el brillo comúnmente se asumen como descriptores independientes del color. Esta técnica puede resultar difícil para reconocer valores idénticos de matiz cuando está presente una diferencia de brillo. Sin embargo, un efecto potencial más serio ocurre cuando se mezclan datos anatómicos con funcionales [94]. Cuando el brillo es controlado por la imagen anatómica, la actividad funcional está parcialmente restringida a seguir los bordes anatómicos, por ejemplo, con muy poco brillo, un pixel se percibe casi negro sin importar su nivel de actividad funcional. Cuando el rango de brillantez o matiz se comprime a un sólo valor, cada volumen puede ser interpretado de forma independiente, sin embargo, un volumen codificado en matiz es más difícil de interpretar que cuando es codificado en matiz y brillo o sólo brillo.

3.5.2. Combinaciones aritméticas

3.5.2.1. Combinación ponderada (suma y multiplicación)

Para el realce de contraste, la suma y multiplicación de imágenes dan buenos resultados [88]. Un ejemplo de proceso de multiplicación es expresado por las ecuaciones:

$$DN_f = A(w_1 DN_a + w_2 DN_b) + B \quad (3.8)$$

$$DN_f = A DN_a DN_b + B \quad (3.9)$$

A y B son factores de escala y w_1 y w_2 son parámetros de peso. DN_f , DN_a y DN_b se refieren a los valores de la imagen final fusionada y las imágenes de entrada a y b respectivamente.

3.5.2.2. Promedio

Cuando se combinan dos imágenes, una aproximación intuitiva es tomar el promedio de los pixeles de las imágenes fuentes para obtener el valor correspondiente del pixel en la imagen compuesta. Sin embargo, cuando este método se toma directo, el contraste de los objetos presentes en las imágenes se reduce. Para evitar el problema de fusión de promedio, la imagen fusionada puede ser obtenida seleccionando pixeles de sus imágenes fuentes.

3.5.3. Análisis de componentes principales (PCA)

El método de componentes principales (PCA: principal component analysis) es un método estadístico que transforma un conjunto de datos de variables correlacionadas en un nuevo grupo de datos no correlacionados, obtenidos por combinación lineal de las originales [134]. El cálculo de las componentes principales (PC) incluye el uso de la matriz de la covarianza o la matriz de correlación para el cálculo de eigenvalores o eigenvectores y el cálculo de los componentes principales (Figura 3.6).

En términos generales, entre los dos o tres primeros componentes principales se contiene entre el 95% hasta más del 99% de la información contenida en la escena. Por tanto, PCA es

una forma de identificar patrones en los datos y expresarlos en forma distinta para resaltar sus similitudes y diferencias, al mismo tiempo que reduce la dimensión de la imagen sin pérdida significativa de información.

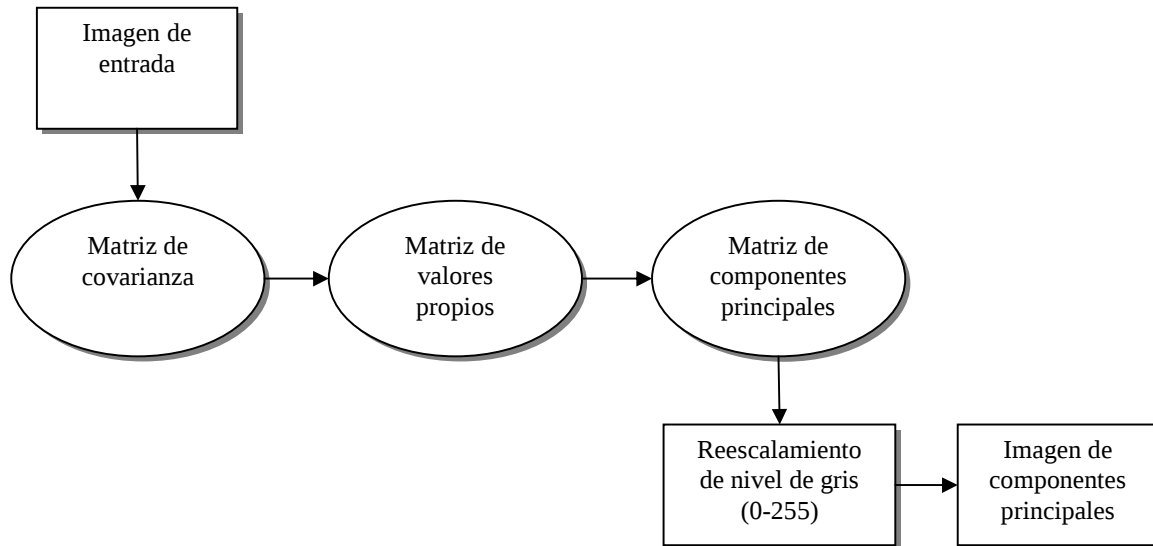


Figura 3.6 Esquema de la transformación de componentes principales.

El análisis por componentes principales (PCA) es una técnica comúnmente usada en imágenes de percepción remota. Ha sido empleada en realce de datos, como una técnica de compresión de imágenes y también para fusión de datos, por ejemplo imágenes radar e imágenes multiespectrales.

Para el proceso de fusión, primero se transforma a las imágenes multiespectrales en un grupo de nuevas bandas que no están correlacionadas, el primer componente PC1 contiene información que es común en todas las bandas multiespectrales, mientras que la información espectral es detallada en los otros componentes. Posteriormente, estos componentes se combinan para obtener un nuevo conjunto de componentes principales, a los cuales se aplica la Transformada inversa de PCA y con ello se obtiene finalmente la imagen fusionada.

La fusión de imágenes mediante PCA tiene dos aproximaciones [88]:

1. Sustitución. Reemplazo del primer componente principal por otra imagen (Principal Component Substitution, PCS).
2. Combinación de componentes principales de todas las imágenes de los distintos sensores.

La primera aproximación sigue la idea de incrementar la resolución espacial de los multicanales de la imagen introduciendo una imagen con alta resolución. El canal que reemplaza al PC1 es normalizado a la varianza de este primer componente y promediado. Un valor alto de resolución de la imagen reemplaza a PC1, el cual es común en todas las bandas mientras la información espectral es única para cada banda. El primer PC1 cuenta con una

varianza máxima la cual puede maximizar el efecto de la alta resolución de datos en la fusión de imágenes.

El segundo procedimiento integra la naturaleza dispar de los datos de entrada del multisensor en una imagen. Las bandas de la imagen de diferentes sensores son combinadas en un archivo y un PCA es calculado desde todas las bandas.

Los métodos de fusión PCA son sensibles a la elección del área de análisis porque la matriz de covarianza y correlación son específicas a la extensión geográfica seleccionada. Por otro lado, el coeficiente de correlación refleja una estrecha relación en muestras homogéneas, sin embargo, los cambios en los valores de las bandas debidos a las diferencias marcadas en los tipos de coberturas también influyen en las correlaciones y particularmente en las varianzas.

3.5.4. Métodos basados en teoría de estimación

La teoría de estimación también ha sido aplicada al problema de fusión de imágenes [130, 99, 77]. Los procedimientos más comunes de estimación incluyen al estimador de máximo a posteriori (MAP: *máximum a posteriori*) y al estimador de máxima verosimilitud (ML: *máximum likelihood*). Sea a el conjunto de imágenes provistas por los diferentes sensores y sea s la escena real a ser estimada; y asumiendo que $p(s|a)$ y $p(a|s)$ denotan las probabilidades condicionales mientras $p(s)$ y $p(a)$ denotan las probabilidades marginales correspondientes. El estimador MAP busca maximizar $p(s|a)$ mientras que el estimador ML busca maximizar $p(a|s)$. La formula de Bayes

$$p(s|a) = \frac{p(a|s)p(s)}{p(a)} \quad (3.10)$$

Indica que los estimadores MAP y ML están estrechamente relacionados. Los métodos de fusión basados en la teoría de estimación con frecuencia son descritos empleando estos componentes primarios conocidos como el modelo de formación de imagen y el modelo a priori. El modelo de formación de imagen representa la relación entre las imágenes de los sensores y la escena verdadera. En [99], se emplea una transformación localmente afín de distorsión cuyos parámetros son variables para la imagen. El modelo a priori representa el conocimiento que se tiene sobre la escena verdadera o la distorsión. Después de elegir formas específicas para la formación del modelo de la imagen y el modelo a priori, diferentes soluciones posibles son asignadas a probabilidades de ser la escena verdadera. Una generalización de esta aproximación fue introducida en [11], la cual permite el uso de distribuciones no Gaussianas. Esta aproximación estima de forma adaptativa las distribuciones empleando un modelo mixto Gaussiano para las distribuciones y una aproximación de estimación iterativa. Otra técnica interesante que usa teoría de estimación son los campos aleatorios de Markov (MRF: *Markov random field*) [126, 6, 56].

3.5.5. Técnicas de fusión multirresolución

Una representación a múltiples escalas proporciona un marco jerárquico para interpretar la información de una imagen o una señal. A diferentes resoluciones espaciales o escalas los detalles en una imagen caracterizan diferentes estructuras físicas de una escena; así, las altas escalas corresponden a una visión global no detallada de la escena, mientras que las bajas escalas corresponden a una vista detallada. Por esta razón, si deseamos describir completamente una imagen debemos analizar un amplio rango de resoluciones.

La idea básica de la fusión multirresolución es efectuar una transformación multiescala (MST: Multi-scale Transform) sobre cada imagen original, luego construir una representación multirresolución. La imagen fusionada se obtiene tomando la transformada multiescala inversa (IMST: Inverse Multi-scale Transform) [136]. Los esquemas basados en la descomposición multirresolución proveen un mejor desempeño debido a que el sistema de visión humano es especialmente sensible a cambios locales de contraste, por ejemplo bordes, y la descomposición multirresolución provee información de la magnitud de los cambios rápidos de contraste en la imagen, además de proporcionar localización tanto en el dominio espacial como en el frecuencial.

3.5.5.1. Contexto genérico la fusión multirresolución

El principio de la fusión multirresolución de imágenes es el dividir los detalles de las imágenes fuente en diferentes niveles de resolución. En el caso de la fusión a nivel pixel, el proceso general puede ser resumido como sigue: en cada posición de muestreo, se hace una decisión sobre las representaciones multirresolución de las imágenes fuente a utilizar. Esta decisión se basa en una medida de nivel de actividad, la cual intenta determinar la calidad de cada imagen fuente. Adicionalmente se emplean algunos métodos de agrupación y combinación de coeficientes para obtener la representación multirresolución de la imagen fusionada. Después puede llevarse a cabo un procedimiento de verificación de consistencia, el cual incorpora la idea de que es poco probable que un coeficiente sea generado de una forma completamente diferente a sus vecinos. Combinaciones diferentes de estas alternativas generan diferentes esquemas de fusión, tal y como se muestra en la Figura 3.7 [136].

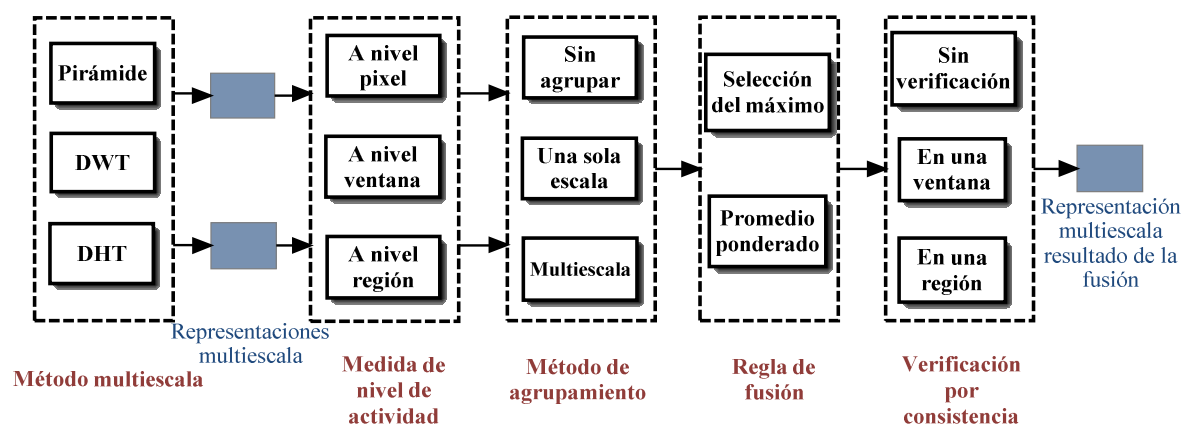


Figura 3.7 Esquemas de fusión multirresolución de imágenes [136]

Los métodos multirresolución más comúnmente usados son la transformada piramidal y la transformada wavelet discreta (DWT). La transformada de Hermite, que es el método propuesto en esta tesis, también puede emplearse como una técnica multirresolución, y existen muchos más como las curvelets, contourlets, wedgelets, ridgelets. Además de la técnica de análisis multirresolución, otro aspecto importante es la forma en la que se combinan los coeficientes, por lo que a continuación se describen algunas alternativas generales para la construcción de la regla de fusión en este tipo de técnicas.

Medida del nivel de actividad

El nivel de actividad de un coeficiente refleja la energía local en el espacio abarcado por el término en la expansión correspondiente a ese coeficiente. Hay tres categorías de métodos para el cálculo del nivel de actividad en la posición p : basado en coeficientes, basado en ventanas y en regiones (Figura 3.8).

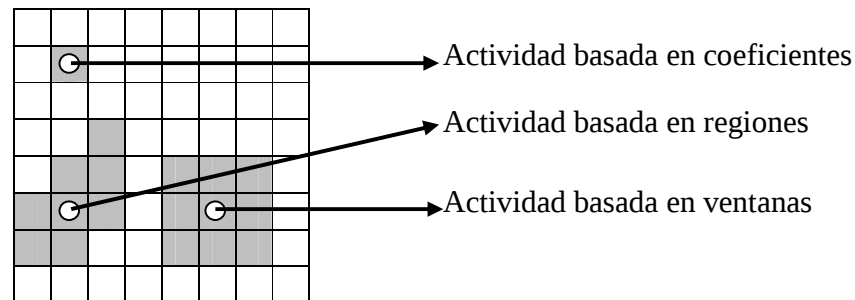


Figura 3.8 Clasificación de las medidas de nivel de actividad

Las medidas de actividad basadas en coeficientes consideran cada coeficiente separadamente. La actividad basada en ventanas se calcula empleando una ventana pequeña (típicamente 3x3 o 5x5) centrada en la posición del coeficiente actual. Para este caso se las reglas más empleadas son el método de promedios ponderados y el método del filtro de rango (rank filter) definido como el i -ésimo valor más largo en el conjunto. Otra opción adicional es medir la actividad mediante el cálculo de propiedades estadísticas como media y desviación estándar. Ambos toman en cuenta texturas de imágenes correspondientes a los coeficientes asociados. Las regiones usadas en la medida de actividad basada en regiones son similares a las ventanas con formas extrañas. La segmentación de regiones puede ser llevada a cabo sobre la misma banda usando la información de los bordes y un algoritmo de etiquetado. La salida es una imagen etiquetada en la cual cada valor diferente representa una región diferente.

Método de agrupación de coeficientes

Después del proceso de descomposición multirresolución cada coeficiente tendrá un conjunto de coeficientes correspondientes en otras bandas de frecuencia y otros niveles de descomposición (Figura 3.9). Para la mayoría de los métodos de fusión los coeficientes de la

descomposición multirresolución no tienen relación entre sí (NG no-grouping schemes), sin embargo podrían relacionarse los coeficientes en una misma escala (SG single-scale grouping) o entre las diferentes escalas que componen la representación de la imagen (MG multiscale grouping).

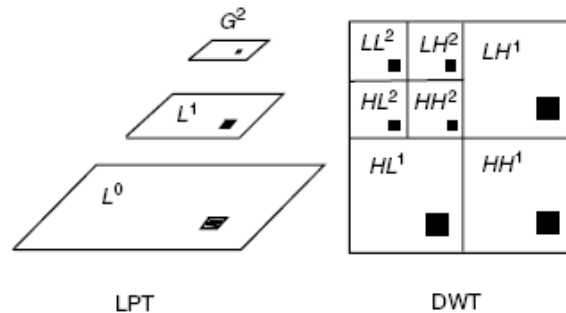


Figura 3.9 Estructuras multirresolución y métodos de agrupación de coeficientes de técnicas piramidales y DWT

Método de combinación de coeficientes y verificación de consistencia

El proceso mediante el cual se combinan los coeficientes es denominado regla de fusión. Existen diversas reglas de fusión por lo que más adelante retomaré este tema, no obstante, cabe destacar que el método más utilizado para combinar los coeficientes provenientes de las bandas de detalle el de selección de máxima nivel de actividad. La verificación de consistencia puede o no aplicarse en el esquema de fusión, el objetivo en ese caso es asegurarse de que el coeficiente de la descomposición multirresolución de la imagen fusionada no provenga de una imagen fuente diferente del que vienen la mayoría de sus vecinos.

A continuación se mencionarán algunas técnicas multirresolución sin profundizar demasiado en ellas.

3.5.5.2. Descomposición piramidal

Una de las primeras aproximaciones multirresolución es la representación piramidal propuesta por Burt, et al. en 1983 [17] y Crowley en 1987 [25]. Una estructura piramidal puede ser descrita como una colección de imágenes a diferentes escalas que juntas representan la imagen original. La pirámide se produce al filtrar la imagen con un filtro paso-bajas y después submuestrear el resultado para generar el siguiente nivel de resolución.

El nivel más bajo o nivel cero $x^{(0)}$ de la pirámide es igual a la imagen original x , esta imagen es filtrada con un filtro paso-bajas w y submuestreada para obtener el siguiente nivel $x^{(1)}$ el cual es filtrado y submuestreado de la misma forma para obtener el nivel $x^{(2)}$, este procedimiento genera los subsecuentes niveles de la pirámide.

El filtro w más utilizado es el filtro Gaussiano y a la pirámide construida se le denomina **pirámide Gaussiana**. Sea $g^{(k)}$ el k -ésimo nivel de la pirámide Gaussiana para una imagen $g^{(0)}$, entonces para $k > 0$ se tiene

$$g^{(k)} = \left[\omega * g^{(k-1)} \right]_{\downarrow 2} \quad (3.11)$$

donde $\downarrow 2$ denota la operación de submuestreo por un factor de 2.

Las operaciones de convolución y el submuestreo producen una pérdida de información de alta frecuencia de un nivel al siguiente. En consecuencia, la **pirámide Laplaciana (LP: Laplacian pyramid)** se define para representar esa pérdida de información entre los niveles adyacentes de la pirámide Gaussiana, pues la diferencia entre dos niveles en dicha pirámide es similar a un filtrado paso altas sobre la imagen. La LP se calcula de la siguiente forma

$$l^{(k)} = g^{(k-1)} - 4\omega * \left[g^{(k)} \right]_{\uparrow 2} \quad (3.12)$$

donde $\uparrow 2$ denota la operación de interpolación por un factor de 2 y el factor de 4 se elige para compensar la pérdida de energía debida al submuestreo. La imagen original puede reconstruirse por su LP y el nivel más alto de la pirámide Gaussiana. En fusión de imágenes, esta estructura fue propuesta en [16].

Existen otras alternativas de descomposición piramidal para la fusión de imágenes, tal es el caso de la pirámide de gradiente [17] y la de radio de paso bajas [111].

La **pirámide de gradiente (GP: gradient pyramid)** de una imagen se obtiene mediante la aplicación de operadores de gradiente a cada nivel de la pirámide Gaussiana. Los operadores de gradiente son usados en la dirección horizontal, vertical y en las 2 direcciones diagonales; se obtienen de la convolución de $g^{(k)}$ con el gradiente d_1 también llamado filtro orientado de segunda derivada.

$$D^{(kl)}(i, j) = d_l * \left[g^{(k)}(i, j) + \omega * g^{(k)}(i, j) \right] \quad (3.13)$$

donde $d_1 = \begin{bmatrix} 1 & -1 \end{bmatrix}$, $d_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}}$, $d_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $d_4 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}}$.

La **pirámide radio de paso bajas (RoLP: Ratio-of-low-pass)** es también conocida como pirámide de contraste [132] y se basa en el contraste de luminancia local. Cada nivel de esta pirámide se calcula como la relación de dos niveles sucesivos de la pirámide Gaussiana. Matemáticamente se define como:

$$R^{(k)}(i, j) = \frac{g^{(k)}(i, j)}{\left[g^{(k+1)}(i, j) \right]_{\uparrow 2}} \quad \text{para } k = n-1, n-2, \dots, 0 \quad (3.14)$$

$$\text{y } R^{(n)} = g^{(n)} \quad (3.15)$$

Las técnicas multirresolución típicamente usan filtros paso bajas o paso banda como parte del proceso. Esas operaciones de filtrado usualmente alteran los detalles de la forma y la localización exacta de los objetos en la imagen. Este problema ha sido tratado mediante el uso de filtros morfológicos para remover los detalles de las imágenes sin efectos adversos. De esta

manera, una **pirámide morfológica (MP: morphological pyramid)** [110] se construye mediante el empleo sucesivo de un elemento estructural y operadores básicos del filtrado morfológico: apertura y cerradura a cada nivel de resolución seguido de un proceso de submuestreo. En fusión de imágenes médicas este método puede encontrarse en [72].

En general, los esquemas de fusión piramidal son muy similares, lo que los hace distintos es manera en la que se construye la estructura piramidal [132]. En una primera etapa se realiza la descomposición piramidal de cada una de las imágenes de entrada; para obtener la pirámide de la imagen compuesta se aplica una regla de fusión entre los niveles correspondientes de las pirámides, y una vez que tiene la pirámide de la imagen fusionada se recupera la imagen aplicando un proceso de reconstrucción que depende del tipo de pirámide.

3.5.5.3. Descomposición wavelet

El análisis basado en la teoría wavelet, permite la descomposición bidimensional de la imagen en diferentes componentes de frecuencia, y diferentes resoluciones. Para ello se utiliza una ventana modulada totalmente escalable que es trasladada a lo largo de la señal y para cada posición se calcula el espectro, repitiendo el procedimiento con una ventana un poco más corta (o larga) para cada nuevo ciclo hasta cubrir toda la señal.

El término wavelet significa “onda pequeña”. La pequeñez se refiere al hecho de que esta función (ventana) es de longitud finita (compactamente soportada) y el término onda se refiere a la condición que esta función es de naturaleza oscilatoria.

El análisis wavelet descrito anteriormente es conocido como transformada wavelet continua (CWT: Continuous Wavelet Transform) y formalmente se escribe como:

$$W_x^\psi(\tau, s) = \langle x, \psi_{s,\tau} \rangle = \frac{1}{\sqrt{|s|}} \int x(t) \psi^*\left(\frac{t-\tau}{s}\right) dt \quad (3.16)$$

donde:

s y τ son parámetros de escala y traslación.

$x(t)$ es la señal a la que se le aplica la transformada.

La función ψ , $\psi(\tau)$, es la función de transformación y es llamada la *wavelet madre*, la cual es una forma de onda oscilante de longitud finita que se utiliza como prototipo para generar las otras funciones ventana (copias escaladas y trasladadas) denominadas “wavelets hijas”.

Para cada escala s y localización τ , los coeficientes de wavelet representan la información contenida en $x(t)$, así la señal original puede ser exactamente reconstruida de los coeficientes de wavelet por:

$$x(t) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty W_x^\psi(s, \tau) \psi_{s,\tau} d\tau \frac{ds}{s^2} \quad (3.17)$$

donde C_ψ es el factor de normalización de la wavelet madre.

Existen diferentes familias de funciones wavelets, las cuales se diferencian en qué tan compactamente son localizadas en el espacio las funciones base y en qué tan suavizadas están. Dentro de una familia de wavelets están las subclases, distinguidas por el número de coeficientes y por el nivel de iteración. Generalmente las subclases se clasifican dentro de una familia por el número de momentos de desvanecimiento. Esto es una relación matemática que deben cumplir los coeficientes y está relacionado directamente con el número de coeficientes.

La aproximación discreta de la transformada de Wavelet (DWT: Discrete Wavelet Transform) puede ser llevada por dos algoritmos, el de Mallat [66] y el de “à trous”.

El algoritmo de Mallat consiste en la aplicación sucesiva de la transformada wavelet mediante el submuestreo de los parámetros de escala y traslación. Si la imagen original tiene C columnas y R filas, a medida que aumentan los niveles de descomposición, se tiene resoluciones más gruesas. En el N ésimo (N th) nivel, la imagen tiene aproximadamente $C/2^N$ columnas y $R/2^N$ filas, que se aplicó la transformada de Wavelet diádica con submuestreo o decimación.

Estas aproximaciones en las imágenes se calculan usando una función de escalamiento para la función Wavelet madre $\psi(x)$. La diferencia entre la información de dos niveles sucesivos del análisis multiresolución, por ejemplo entre la imagen original A_2^j a una resolución 2^j y la imagen aproximada A_2^{j-1} a una resolución 2^{j-1} está dada por la transformada de wavelet. Las matrices de coeficientes resultantes son $C^{H,j-1}_2$, $C^{V,j-1}_2$ y $C^{D,j-1}_2$ con información de detalles horizontal, vertical, y diagonal respectivamente (Figura 3.10).

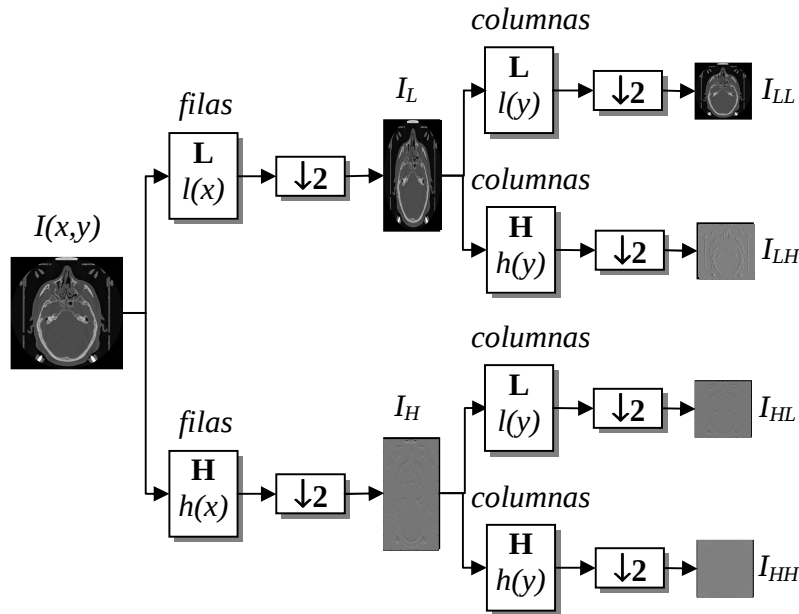


Figura 3.10 Una etapa de la descomposición de una imagen mediante la DWT.

Cuando se aplica la transformada inversa, la imagen original A_2^j puede ser reconstruida a partir de la imagen A_2^{j-1} y las matrices de coeficientes $C^{H,j-1}_2$, $C^{V,j-1}_2$ y $C^{D,j-1}_2$.

Para la realización práctica del algoritmo de Mallat, se usan filtros de cuadratura (quadrature mirror filter) en vez de las funciones de escalamiento y las funciones Wavelet. El filtro L asociado con funciones de escalamiento, es un filtro paso bajo que permite el análisis de datos de baja frecuencia. Mientras el filtro H se asocia con las funciones de Wavelet y es un filtro paso alto que permite el análisis de los componentes de alta frecuencia, por ejemplo, los detalles de la imagen. El número de parámetros de estos filtros y el valor de estos parámetros depende de la función madre Wavelet usada en la parte de análisis.

El algoritmo ‘à trous’ [82, 101] se refiere a insertar ceros en los filtros y fue diseñado para subsanar la falta de invarianza a la traslación de la DWT, esta se logra eliminando el submuestreo y el sobremuestreo en la DWT. Al aplicar esta descomposición todas las imágenes aproximadas tienen el mismo número de columnas y renglones que la imagen original. Esto es una consecuencia del hecho de que el algoritmo de “à trous” es no ortogonal, y una transformada redundante y sobremuestreada [115]. La realización práctica del algoritmo de ‘à trous’, usa un filtro en dos dimensiones asociado a la función de escalamiento.

La transformada wavelet en el contexto de la fusión de imágenes se utiliza para describir diferencias entre imágenes sucesivas provistas por el análisis multirresolución. En general, los métodos de fusión basados en wavelets emplean algoritmos de sustitución y combinación de coeficientes y se fundamentan en las siguientes razones [83]:

- i) Es una aproximación multiescala (multirresolución) muy apropiada para manejar diferentes resoluciones de imágenes.
- ii) La DWT permite la descomposición de imagen en diferentes tipos de coeficientes preservando la información de la imagen.
- iii) Los coeficientes provenientes de diferentes imágenes pueden ser combinados apropiadamente para obtener nuevos coeficientes de manera que la información de las imágenes originales puede ser reunida de forma apropiada.
- iv) Una vez que los coeficientes son mezclados, la imagen final fusionada es reconstruida utilizando la IDWT, donde la información en los coeficientes combinados es también preservada.

Debido a lo anterior, los primeros esquemas de fusión basados en wavelets [61, 23] reportaron resultados superiores a los métodos estándar piramidales, tanto cualitativa como cuantitativamente. No obstante, la DWT no es invariante a la traslación debido al submuestreo empleado en su construcción y esto provoca artefactos de anillos y movimiento en las imágenes. Rockinger et al [97] superó este problema empleando un algoritmo ‘à trous’, sin embargo, este método resultó más costoso computacionalmente debido a la gran cantidad de redundancia, haciéndolo impráctico para grandes cantidades de datos. Hill et al [51] propuso un esquema de fusión introduciendo la transformada wavelet compleja de árbol dual (DTCWT: dual-tree complex wavelet transform), la cual es casi invariante a la traslación y contiene menos redundancia que la DWT ‘à trous’, pero presenta el inconveniente de que los filtros empleados requieren ser diseñados meticulosamente para alcanzar el retraso apropiado de tal forma que el aliasing de la frecuencia de los diferentes árboles pueda ser cancelado; por lo que no resulta fácil encontrar filtros FIR que satisfagan los requerimientos.

3.5.6. Métodos de selección

Los métodos utilizados en selección de fusión por transformada son máximo contraste, verificación de consistencia [61], y medidas de correspondencia – saliente [15]. La idea es formar una imagen compuesta por reglas de selección de fusión de las imágenes A y B , se define como

$$F = FUSE(A(i,j), B(i,j)) \quad (3.18)$$

FUSE es la función que convierte las dos imágenes en la compuesta, usando una técnica de fusión.

Máximo contraste

El método de máximo contraste consiste en seleccionar detalles de máximo contraste. La expresión del contraste de la función FUSE se realiza en una modalidad de pixel a pixel, como sigue:

$$F(i, j) = \begin{cases} A(i, j) & \text{cuando } |A(i, j)| \geq |B(i, j)| \\ B(i, j), \dots \dots \dots \text{otros} & \end{cases} \quad (3.19)$$

donde $F(i, j)$ es el término que define los coeficientes de la imagen fusionada F para toda i, j . A y B representan cada imagen de entrada.

Verificación de consistencia.

Li H. et al [15] usaron el valor máximo absoluto dentro de una ventana como una medida de la actividad asociada con el pixel central. De esta manera, un valor eficiente indica la presencia de un patrón dominante en el área local. Un mapa de decisión binaria, del mismo tamaño de la transformada, es entonces creado para el registro de los resultados seleccionados con base en la regla de máxima selección basada en el contraste.

El contraste de un pixel está definido como el cociente de luminancia del pixel y su luminancia de fondo (background), entre esta última, es decir,

$$\text{Contraste}(i, j) = \frac{L(i, j) - L_b(i, j)}{L_b(i, j)} = \frac{L(i, j)}{L_b(i, j)} - 1 \quad (3.20)$$

donde L , luminancia en (i, j) , o simplemente la intensidad de pixel, y L_b la luminancia de fondo para el área.

Se ilustran dos aspectos, si el mismo objeto aparece con mayor contraste en la imagen A que en la imagen B , en la fusión, el objeto de la imagen A será preservado mientras el objeto de la imagen B será ignorado. De otra forma, suponiendo que los límites externos aparecen más claros en la imagen A mientras que los límites internos del objeto aparecen más claros en

la imagen B , los coeficientes de la transformada de los objetos en la imagen A y B dominarán en diferentes niveles de resolución.

El mapa de decisión binario creado para la aplicación de esta regla está sujeto a una verificación de consistencia. Específicamente, si el valor del pixel central se selecciona de la imagen A mientras la mayoría de los valores de pixeles alrededor se seleccionan de la imagen B , el valor del pixel central es entonces sustituido por el pixel de la imagen B . En la realización, un filtro de mayoría (que produce uno si el número de unos sobrepasa al de ceros, y produce ceros en el otro caso), se aplica para obtener el mapa de decisión binaria. De esta manera, la consistencia va en función de que el valor del pixel se modifique, dependiendo si la mayor intensidad de la suma de los vecinos cercanos pertenece al coeficiente de procedencia del valor indicado o no.

Medidas de correspondencia y saliente

El método de medidas de correspondencia y saliente propuesto por Burt P. J. [15] define dos formas distintas de combinación: correspondencia y saliente.

La medida de correspondencia determina el modo de combinación a cada misma posición (selección o promedio ponderado). La medida de saliente determina cual de los dos coeficientes será copiado a la imagen fusionada (modo de selección) o cual coeficiente será asignado al peso más grande (promedio ponderado). Cuando las dos imágenes son diametralmente diferentes, la imagen compuesta podría seleccionar la componente con mayores salientes. Sin embargo, cuando dos imágenes son similares, la imagen compuesta podría contener la media de dos imágenes fuente. A continuación se describe la aplicación del método en el contexto de la descomposición de imágenes mediante alguna transformación:

- a. Se calculan las varianzas en cada posición espacial (i,j) dentro de una vecindad, σ_A y σ_B (medida de saliente).

- b. De la misma forma, se calcula un coeficiente de correspondencia M

$$M = \frac{2\sigma_{AB}}{\sigma_A^2 + \sigma_B^2} \quad (3.21)$$

- c. Se calculan los coeficientes fusionados utilizando la fórmula $F = W_A A + W_B B$, donde W_A y W_B son pesos asignados para las imágenes fuente A y B respectivamente; y $W_A + W_B = 1$. Dicha fórmula se aplica como sigue:

Si $M > \alpha \rightarrow$ Modo de combinación por promedio ponderado

$$W_{min} = 0.5 (1 - (1-M)/(1-T)) \text{ y } W_{max} = 1 - W_{min}$$

en caso contrario $\rightarrow$ Modo de combinación por selección

$$W_{min} = 0 \text{ y } W_{max} = 1$$

Si $\sigma_A > \sigma_B$

$$W_A = W_{max} \text{ y } W_B = W_{min}$$

en caso contrario

$$W_A = W_{min} \quad y \quad W_B = W_{max}$$

Esta técnica requiere pruebas extensivas para encontrar una α apropiada para cada tipo de imagen usada y no permite cambios graduales entre la selección y promedio. Burt propone un valor de α de 0.75.

3.6. Reglas de fusión

Una vez que los datos a procesar están en el mismo contexto tanto en nivel de resolución como de transformación (técnica de análisis), la forma de fusionarlos depende de la regla de fusión empleada. Sobre este tema existen diversas alternativas que entre otros aspectos incluyen la elección de operadores para llevar a cabo el proceso de fusión; estos operadores determinan la combinación de los grados de credibilidad de la información extraída de los datos originales tomando en cuenta la incertidumbre, imprecisión e incompletitud de la naturaleza de los datos. En la literatura es posible encontrar una gran variedad de formas de combinación, Bloch [10] en particular proporciona una clasificación para los operadores dependiendo de su comportamiento y con ello provee una guía para elegir un operador en un problema dado. Esta elección puede detallarse de acuerdo a las propiedades deseadas para el operador, de su decisividad y de su respuesta a un problema específico; por estas razones será la ésta la clasificación que se menciona a continuación.

Sean a y b dos variables reales que representan los grados de confianza para ser combinados, que pueden tomar valores dentro del intervalo $I [0,1]$. Considerando una función F que actúa sobre ambas variables definiendo una combinación u operador de fusión se tienen los siguientes operadores:

- F es conjuntiva si $F(a,b) \leq \min(a,b)$ (esto corresponde al comportamiento severo)
- F es disyuntiva si $F(a,b) \geq \max(a,b)$ (comportamiento indulgente)
- F se comporta como un compromiso si $a \leq F(a,b) \leq b$ si $a \leq b$ y también si $b \leq F(a,b) \leq a$ (comportamiento precavido)

De acuerdo a lo anterior, se observa que la fusión disyuntiva es una operación con un riesgo mínimo, desde que la información original de ambas fuentes se conserva. La fusión conjuntiva en cambio sólo selecciona información común por lo cual se dice que su comportamiento es muy estricto, incrementa el riesgo de eliminar información importante no presente en cada fuente mientras retiene valores redundantes. El operador de fusión que actúa como un compromiso promedia los datos originales de manera que ni elimina información poco fidedigna ni toma en cuenta información redundante.

La clasificación propuesta por Bloch hace una descripción de los operadores no sólo en términos de su comportamiento sino también respecto a los valores particulares de la información que será combinada, quedando divididos en operadores de comportamiento constante independientes del contexto, operadores de comportamiento variable independientes del contexto y operadores dependientes del contexto.

❖ Operadores de comportamiento constante independientes del contexto (CICB: Context Independent Constant Behavior)

Los operadores CICB son aquellos que tienen el mismo comportamiento independientemente del tipo de datos o los valores de la información a combinar. Estos operadores trabajan sin ayuda contextual o información externa. Algunos ejemplos [92] de este tipo de operadores, para el caso específico de dos imágenes, son los siguientes:

Producto: $F(a,b) = a \times b$.

Mínimo: $F(a,b) = \min \{a,b\}$.

Máximo: $F(a,b) = \max \{a,b\}$.

Promedio aritmético: $F(a,b) = (1/2) \{a+b\}$.

❖ Operadores de comportamiento variable independientes del contexto (CIVB: Context Independent Variable Behavior)

Esta clase de operadores son independientes del contexto, como la clase anterior, pero dependen de los valores numéricos de a y b . Dependiendo de esos valores, estos operadores pueden comportarse de diferentes formas: por ejemplo, conjuntivamente si a y b son menores que 0.5, disyuntivamente si a y b son mayores que 0.5 y precavidamente en caso contrario. Las sumas simétricas asociativas constituyen los operadores difusos CIVB más típicos [22].

❖ Operadores Dependientes del contexto (CD: Context Dependent)

Estos operadores como su nombre lo indica son dependientes del contexto, no sólo dependen de x y y sino también de las especificaciones de las fuentes que se desea fusionar (como conflictos entre fuentes, confiabilidad de las mismas, entre otras cosas). Por ejemplo, es posible construir un operador que se comporte en una forma conjuntiva si las fuentes son consonantes (confiables), en disyuntiva si son disonantes (poco confiables) y como un compromiso si son parcialmente conflictivas.

Estos operadores son en particular interesantes para los problemas de clasificación, y esto es porque sus características adaptativas hacen que ellos sean capaces de combinar información relacionada a una clase de una manera y la información relacionada a otras clases de forma distinta.

3.7. Ventajas y limitantes de los métodos de fusión

Antes de implementar y utilizar un algoritmo de fusión es importante tomar en cuenta el objetivo de la aplicación y el tipo de datos para en base a ello definir la técnica de fusión más adecuada y regla de fusión para llevar a cabo la combinación de los datos [88]. Por ejemplo, en imágenes médicas lo primero que debe definirse son las modalidades médicas con las que se desea trabajar. La siguiente etapa es la elección de un nivel de fusión apropiado ya que las etapas de pre-procesamiento dependen de esto.

Los métodos basados en fusión a nivel pixel tratan con volúmenes de datos muy grandes por lo cual hay una cantidad considerable de tiempo de cómputo. La fusión a nivel de características usa estructuras correspondientes, las cuales hacen de la geometría un asunto menos crítico. Sin embargo, los métodos basados a nivel pixel tienen la ventaja respecto a los de nivel características o decisión de usar los datos en su forma original, lo cual evita pérdidas de información que ocurren durante el proceso de extracción de características.

La pregunta de la técnica a elegir está también relacionada con la definición del criterio de evaluación. Ambos son difíciles de definir y dependen mucho de resultados empíricos. En consecuencia, es muy difícil hacer asunciones generales sobre la calidad de una técnica de fusión. Por otro lado, las características de los datos fusionados dependen mucho del pre-procesamiento aplicado y de la técnica de fusión empleada. Las restricciones están relacionadas con la perturbación del contenido espectral de los datos de entrada y efectos de emborronamiento cuando se introducen imágenes con una baja relación señal a ruido.

En el caso de la fusión de imágenes multimodales, el problema de reconstrucción de las intensidades locales de la imagen fusionada se hace presente debido a que las imágenes de entrada son adquiridas de diferentes sensores multimodales, los cuales poseen diferentes rangos y valores de intensidades [74]. La imagen fusionada es una imagen artificial, que no existe en la naturaleza, y en consecuencia es difícil encontrar un criterio que pueda determinar la forma más eficiente de combinar los rangos de intensidad de los sensores de entrada.

Un problema adicional puede ser la creación de una imagen fusionada a color, como resultado del proceso de fusión. Un primer intento por resolver este problema es el tratar cada canal de color de forma independiente y fusionarlos con los canales de entrada de los otros sensores para crear una representación de imagen fusionada de tres canales, ésta es una técnica simple que no requiere de cómputo intenso. Pero a pesar de que esta técnica parece razonable e incluso produce resultados aceptables en varios casos, ésta no utiliza las dependencias entre los canales de color que podrían beneficiar al contexto de fusión. Otra propuesta ha sido emplear otro espacio de color como el YUV que describe una imagen de color usando una luminancia y dos canales de crominancia, o el IHS que describe un color mediante matiz, saturación e intensidad (luminancia); no obstante, este último método presenta la desventaja de la reducción de detalle espacial respecto a los datos originales.

Los resultados experimentales indican que los algoritmos de fusión que se basan en la descomposición multirresolución muestran un mejor desempeño en comparación con las demás técnicas [132]. Sin embargo, también en ellas encontramos ciertas desventajas, por ejemplo las técnicas basadas en la pirámide Laplaciana incluyen efectos de bloque y son poco flexibles [137], mientras que los métodos basados en las wavelets son sensibles a la elección del nivel de descomposición y a la base de la wavelet.

En cuanto a las reglas de fusión, el esquema más simple es el promedio de las imágenes originales, sin embargo, este método introduce varios efectos indeseables como la reducción del contraste. En conclusión, la selección más adecuada de un algoritmo de fusión a nivel pixel depende de los meritos de cada método, del contexto y de las características de los datos fuente.

3.8. Situación actual de la fusión de imágenes médicas

En la imagenología médica, a pesar de los avances alcanzados, la fusión de imágenes ha tenido algunos obstáculos para convertirse en una herramienta clínica sobresaliente. Entre estos obstáculos se encuentra el registro de imágenes que aún requiere superar imprecisiones producidas por la posición del paciente y la deformación no rígida (un método que aplica deformación para corregir cambios no lineales y movimientos de órganos). Las dificultades con el registro 2D han bloqueado aplicaciones potenciales en cardiología, mientras que la optimización del uso de 3D en fusión de imágenes aún involucra computo más sofisticado. Además, se requiere de la implementación de la conectividad a través de los departamentos de las diferentes modalidades a fin de obtener las imágenes de las modalidades requeridas [12].

No obstante, la fusión de imágenes médicas es un tema que en los últimos años ha sido trabajado utilizando diferentes técnicas destacando principalmente el análisis multirresolución mediante la transformada wavelet [118]. A partir de esta metodología, se ha seguido una línea de investigación en ese sentido con métodos de fusión que presentan algunas variantes; algunos de ellos con la finalidad de evitar la pérdida (inherente a la descomposición wavelet) de algunos detalles en las bandas de alta frecuencia, tal es el caso el uso de la transformada Packet Wavelet; por ejemplo, Anna W., et al en el año 2007 [119] propusieron una técnica que combina dicha transformada con 2v-SVM (Support Vector Machine), lo cual hace posible que los píxeles sean clasificados en la banda de frecuencia correcta como si fueran vectores, evitando así caer en óptimos locales si se empleara una red neuronal; Yang L., et al en el año 2008 [63] también propusieron el uso de la transformada Packet Wavelet con operadores autoadaptativos que ajustan las ponderaciones de los coeficientes de manera automática dependiendo de los datos a fusionar. Otras propuestas se enfocan en mejorar la extracción de características importantes que deben estar presentes en la imagen fusionada, ya sea realizando el proceso de fusión a nivel de regiones [46, 135] en lugar de píxeles, mediante la transformada Wavelet Modulos Maxima [47] o a través del cálculo de la entropía para definir la proporción de información de cada imagen fuente y así conservar detalles de regiones pequeñas [125].

Otros métodos que se encuentran en la literatura son el análisis geométrico multiescala de la transformada Contourlet [131], las redes neuronales [121] y la transformada de Hermite que ha sido aplicada en percepción remota [41, 42] así como en imágenes médicas [39, 79]. Estos métodos presentan mejores resultados que los que se basan en la transformada Wavelet, puesto que se ha demostrado que esta última transformada presenta ciertas limitaciones en el análisis de señales con dimensiones superiores a 1D, ejemplo de ello son los puntos de discontinuidad que no siempre puede detectarlos; otra desventaja consiste en la limitante en la captura de información direccional ya que sólo cuenta con tres bandas pasa altas direccionales: vertical, horizontal y diagonal. Las desventajas mencionadas, ocasionan que haya características de las imágenes originales que no puedan ser conservadas con la precisión deseada e incluso se introduzcan algunos artefactos.

El método de fusión basado en el análisis geométrico de la transformada Contourlet propuesto por Yang L., et al en el año 2008 [131], lleva a cabo una expansión de la imagen expresada por segmentos de contorno que pueden capturar estructuras geométricas 2D de manera más eficiente gracias al análisis en varias direcciones; el algoritmo consiste en la

descomposición de las imágenes originales utilizando la transformada Contourlet, para un primer nivel la imagen se descompone en una banda pasa bajas y varias pasa altas en diferentes direcciones y escalas, la banda de baja frecuencia constituye una aproximación, mientras que las demás preservan los detalles, una vez que se capturan los puntos de discontinuidad se efectúa una descomposición direccional local para sintetizar los puntos dentro de segmentos de contorno independientes, luego se aplican los criterios de fusión para ambos tipos de bandas y se obtiene la imagen fusionada a través de la transformada inversa Contourlet.

El método de fusión de imágenes médicas mediante redes neuronales de funciones base radiales difusas (Fuzzy-RBFNN) fue propuesto por Yang-Ping Wang, et al en el año 2007 [121], el cual ataca el problema de fusionar imágenes médicas emborronadas como resultado de la presencia de ruido inherente a la adquisición de las imágenes en la mayoría de los casos; su método es autoadaptativo y utiliza un algoritmo genético para el entrenamiento de la red, el cual lleva a cabo una búsqueda global utilizando heurísticas para evitar que la red caiga en un óptimo local; así, los píxeles de las imágenes a fusionar son clasificados dentro de grupos característicos de acuerdo a los valores de intensidad de las imágenes y mediante las reglas de operación de la red se van generando los valores de los píxeles que formarán parte de la imagen fusionada.

Finalmente, la fusión de imágenes médicas mediante la transformada de Hermite fue propuesta por Escalante R. en el año 2008 [39], mostrando que la transformada de Hermite permite la identificación de patrones relevantes que son incluidos en la imagen fusionada al mismo tiempo que evita la aparición de artefactos espurios. El algoritmo consiste en obtener la transformada de Hermite rotada de las imágenes originales y aplicar las reglas de selección de coeficientes para generar el conjunto de coeficientes únicos que originarán la imagen fusionada mediante la transformada de Hermite inversa.

En cuanto a fusión de volúmenes generados de imágenes médicas, existen pocos trabajos desarrollados aunque seguramente será parte del trabajo futuro de las investigaciones mencionadas anteriormente, debido a las enormes ventajas que ofrecen las aplicaciones de visualización en 3D para entrenamiento médico y planeación de intervenciones quirúrgicas, donde es de suma importancia tener gran precisión en la localización del problema a tratar. Los primeros trabajos encontrados en la literatura se ocupan del registro de volúmenes ya sea mediante el uso de marcas fiduciarias [127] o por correspondencia de superficies [36], en ninguno de estos artículos se plantea un algoritmo de fusión, constituyen únicamente la base para ello con el proceso de registro y el desarrollo de aplicaciones de visualización de volúmenes donde es posible visualizar más de una modalidad en superposición. Li Bin, et al en el año 2008 [9] realizaron un trabajo de fusión de imágenes médicas en 3D en paralelo para la planeación de tratamientos en radioterapia; el método de fusión propuesto utiliza la transformada wavelet para fusionar cada corte, es decir, en realidad se efectúa la fusión de imágenes en 2D que al final formarán parte del volumen fusionado; lo más interesante de este trabajo es el procesamiento en paralelo que optimiza el tiempo de computo y busca ser una aplicación que pueda ejecutarse en tiempo real en un futuro a corto plazo.

En el apéndice A se exponen algunos de los algoritmos de fusión propuestos para imágenes médicas en las últimas décadas.

Transformada de Hermite

4.1. Relevancia en Percepción Humana

En muchas aplicaciones de codificación de imágenes y visión por computadora, así como en el caso del sistema de percepción humana, se requiere que los datos de la imagen, dados como un arreglo de valores de intensidad, sean interpretados como patrones visualmente significativos mediante algún tipo de procesamiento espacio-temporal de los datos originales. Este procesamiento local se realiza multiplicando la imagen original por una ventana de análisis cuyo tamaño establece el conjunto de puntos que contribuyen en esa etapa básica del procesamiento mientras que la forma determina el peso relativo para cada punto considerado de la imagen. Para describir la imagen en su totalidad es necesario repetir el análisis local un número para un número suficiente de posiciones de la ventana [70].

Por otra parte, en cada posición de la ventana deben realizarse etapas específicas de procesamiento, lo cual implica la búsqueda de patrones visuales que se consideran más relevantes a priori. En general, hacer la mejor elección para la función ventana y el tipo de procesamiento es algo muy complicado, por lo que se toma como referencia el sistema de visión humano (HVS: human vision system). La razón de esto último no sólo es por el hecho de que los procesos visuales que uno desearía automatizar involucran análisis e interpretación de imagen similares a los que lleva a cabo el HVS, sino también porque la mayoría de dichos procesos se llevan a cabo sobre la misma gama de escenarios observados por los seres humanos.

La teoría de Wavelets muestra cómo las señales pueden ser expandidas sobre una familia de funciones, las cuales son versiones de una misma función dilatada y trasladada. Sin embargo, las funciones ventana que satisfacen todas las condiciones necesarias de ortogonalidad tienen dos desventajas: a) usualmente son mucho más grandes que el espaciamiento entre ventanas, y b) frecuentemente son menos suaves que las funciones deducidas a partir de los perfiles de los campos receptivos en la percepción humana. Con el objetivo de lograr funciones más suaves es necesario ignorar la condición de ortogonalidad entre funciones base de ventanas adyacentes. No obstante, no hay necesidad de ignorar dicha condición entre funciones base en una ventana.

Como se mencionó anteriormente, un parámetro importante de la función ventana es su tamaño (o escala) cuya elección representa un problema importante. La ventana tiene que ser lo suficientemente grande para alcanzar una alta reducción de datos, pero la complejidad del análisis dentro de cada ventana se incrementa rápidamente al aumentar su tamaño. La solución al problema puede ser alguna de las siguientes opciones: a) fijar el tamaño de la ventana de análisis de tal manera que todos los patrones visuales de interés sean incluidos, o b) limitar la

complejidad del análisis en cada ventana y en consecuencia determinar el tamaño de ventana necesario para describir la imagen localmente con suficiente precisión. De esta forma, el procesamiento no se limita a una sola escala, pues se puede repetir el proceso a múltiples escalas y después usar la salida de estas etapas para elegir la escala óptima en cada posición. Existen pruebas contundentes que muestran que el HVS usa este principio de *escala-tiempo*.

Frecuentemente, cuando se usan funciones ventana de distintos tamaños el espaciamiento de las funciones ventana es proporcional a su tamaño, por lo que en la mayoría de las estructuras piramidales, el espaciamiento entre ventanas y el tamaño se incrementan en pasos de dos. Se ha demostrado que este incremento exponencial en el tamaño de la ventana y el espaciamiento no sólo es preferible desde un punto de vista de la información sino que además es computacionalmente eficiente.

En cuanto a las técnicas de procesamiento local, está demostrado que el uso de técnicas de pre-procesamiento, esto es, la descripción de la imagen por medio de patrones seleccionados a priori, combinado con el uso de métodos estadísticos provee mejores resultados que el uso directo de métodos de aproximación estadística. Además el uso de patrones con diferentes orientaciones está de acuerdo con el HVS. Generalmente la búsqueda de patrones importantes en imágenes, tales como líneas y bordes, involucra el uso de derivadas de primer y segundo orden en combinación con un filtro paso bajas. Es por eso que la transformada de Hermite utiliza filtros elementales similares a las derivadas de Gaussianas para el procesamiento local, las cuales son un buen modelo de los campos receptivos con respecto al procesamiento espacial de estímulos visuales. Así, el modelo de Hermite, basado en los filtros de análisis de la transformada de Hermite es próximo al modelo de la visión humana en derivadas Gaussianas desarrollado por Richard A. Young [133], ya que puede modelar operaciones de filtrado en el HVS con la misma precisión y con menos parámetros que los ampliamente usados filtros de Gabor [70].

Algunos trabajos recientes han mostrado que esta transformada puede ser usada en la construcción eficiente de aplicaciones tales como reducción de ruido [37], codificación de imagen [38], segmentación de imagen [76], fusión de imágenes [41, 39, 26] entre otras.

4.2. Transformada polinomial

La transformada polinomial es una técnica de descomposición local de señales por medio de polinomios [70].

El proceso de esta técnica consta de dos pasos. En un primer paso, la señal original $L(x)$ es localizada al multiplicarla por una función ventana $V(x)$ en varias posiciones equidistantes, esto con el propósito de lograr una descripción completa de la señal. A partir de la función ventana $V(x)$ es posible definir una función de ponderación periódica T :

$$W(x) = \sum_k V(x - kT) \quad (4.1)$$

El periodo de la función de ponderación es T . Haciendo $W(x)$ diferente de cero para cualquier posición x , se obtiene

$$L(x) = \frac{1}{W(x)} \sum_k L(x) \cdot V(x - kT) \quad (4.2)$$

con lo que se garantiza que las señales localizadas $L(x) \cdot V(x - kT)$ para todas las diferentes posiciones kT contienen suficiente información de la señal original.

El segundo paso consiste en aproximar la información local dentro de cada ventana de análisis en términos de una familia de polinomios ortogonales. Para llevar a cabo la expansión, se toman como funciones base los polinomios $Q_n(x)$, donde n es el grado del polinomio. Estos polinomios están determinados totalmente por la función ventana de tal forma que deben ser ortonormales respecto a $V^2(x)$ es decir

$$\int_{-\infty}^{+\infty} V^2(x) Q_m(x) Q_n(x) dx = \delta_{mn} \quad (4.3)$$

En una transformada polinomial, los coeficientes de una señal $L(x)$ (proyecciones) pueden ser obtenidos como

$$L_n(kT) = \int_{-\infty}^{\infty} L(x) \cdot Q_n(x - kT) \cdot V^2(x - kT) dx \quad (4.4)$$

seguido de un submuestreo en múltiplos de T . Es posible notar que la señal original $L(x)$ se convoluciona con las funciones de análisis (funciones filtro) denotadas por:

$$D_n(x) = Q_n(-x) \cdot V^2(-x) \quad (4.5)$$

es decir:

$$L_n(kT) = \int_{-\infty}^{\infty} L(x) \cdot D_n(x - kT) dx \quad (4.6)$$

El mapeo de la señal original $L(x)$ a los coeficientes $L_n(kT)$ es conocido como transformada polinomial directa. El proceso de reconstrucción de la señal a partir de los coeficientes es conocido como transformada polinomial inversa (ecuación 4.7) y básicamente consiste en interpolar los coeficientes $\{L_n(kT); k \text{ entero}\}$ con las funciones de síntesis (funciones patrón) $P_n(x)$ (ecuación 4.8) y sumar sobre todos los órdenes n . Las transformadas polinomiales directa e inversa se muestran en la Figura 4.1.

$$L(x) = \sum_{n=0}^{\infty} \sum_k L_n(kT) \cdot P_n(x - kT) \quad (4.7)$$

$$P_n(x) = \frac{Q_n(x) \cdot V(x)}{W(x)} \quad (4.8)$$

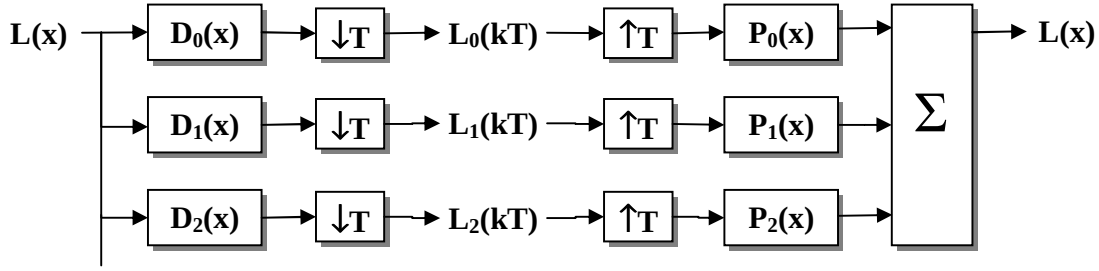


Figura 4.1 La transformada polinomial directa e inversa

4.3. Transformada de Hermite

La transformada de Hermite (HT: Hermite transform), formalmente introducida por Martens [69, 70] al área de procesamiento digital de imágenes, es un caso particular de la transformada polinomial cuando las ventanas de análisis empleadas son funciones gaussianas. La ventana Gaussiana tiene la propiedad de ser isotrópica (invariable con la rotación), separable en coordenadas cartesianas y sus derivadas simulan algunos procesos a nivel retinal o corteza visual del sistema de visión humano; se define de la siguiente manera:

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-\frac{x^2}{2\sigma^2}} \quad (4.9)$$

donde el factor de normalización $1/(\sqrt{\sqrt{\pi}\sigma})$ es tal que $V^2(x)$ tiene energía unitaria.

Los polinomios ortogonales que están asociados a $V^2(x)$ son conocidos como polinomios de Hermite, cuya expresión matemática está dada por la fórmula de Rodrigues [1] (ecuación 4.10), de ahí que el nombre de esta técnica de descomposición local sea el de transformada de Hermite.

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}, \quad n = 0, 1, 2, \dots, \quad (4.10)$$

donde el operador d^n/dx^n se refiere a la n-ésima derivada de la función.

Entre las razones que hay para elegir una ventana Gaussiana están las siguientes:

- La teoría es matemáticamente tratable, por lo que las propiedades de la transformada de Hermite son fácilmente derivadas y evaluadas. Además, la transformada discreta de Hermite es una buena aproximación del caso continuo y sus propiedades están muy relacionadas.
- Las ventanas Gaussianas que están separadas por el doble de la desviación estándar σ son un buen modelo para los campos receptivos.

- La transformada de Hermite involucra funciones filtro que son derivadas de Gaussianas, las cuales han sido ampliamente usadas en visión computacional y en el modelado psicofísico del HVS.
- La ventana Gaussiana minimiza el producto de las incertidumbres en los dominios espacial y frecuencial.

A continuación se describen las propiedades de cada una de las funciones involucradas en la definición de la transformada de Hermite.

Función de ponderación

Como la función de ponderación $W(x)$ es periódica con periodo T , puede ser expandida a través de una serie de Fourier

$$W(x) = \frac{\sqrt{2\sqrt{\pi}\sigma}}{T} w(x) \quad (4.11)$$

donde

$$w(x) = 1 + 2 \sum_{k=1}^{\infty} \exp \left[-\frac{1}{2} \left(k \frac{2\pi\sigma}{T} \right)^2 \right] \cdot \cos \left(k \frac{2\pi x}{T} \right) \quad (4.12)$$

El contraste de esta función de ponderación se determina por el parámetro de muestreo $\tau = T/\sigma$. Como usualmente se quiere limitar el número de descomposiciones locales, es mejor hacer que τ sea tan grande como sea posible. Por otro lado, considerando la ecuación 4.2, se observa que $W(x)$ debe ser aproximadamente constante pues la división entre $W(x)$ podría introducir una sensibilidad a la variación de la escala.

Funciones de análisis

Las funciones de análisis (funciones filtro) determinan qué información se hace explícita en los coeficientes de la transformada de Hermite, por lo que las principales propiedades de esta transformada están determinadas por estas funciones. A partir de la ecuación 4.5 se pueden derivar las funciones filtro específicas

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma\sqrt{\pi}} H_n \left(\frac{x}{\sigma} \right) e^{-\frac{x^2}{\sigma^2}} \quad (4.13)$$

que podemos reescribirla en términos de la definición de derivadas de Gaussiana evaluada en $x=(x/\sigma)$ como

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{\pi}} \cdot \frac{d^n}{dx^n} e^{-\frac{x^2}{\sigma^2}} \quad (4.14)$$

Se puede demostrar que el filtro $D_n(x)$ es igual a la derivada de orden n de una función Gaussiana

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{\pi}} \cdot G_n\left(\frac{x}{\sigma}\right) \quad (4.15)$$

donde $G_n(x) = \frac{d^n e^{-x^2}}{dx^n}$, $n=0,1,2,\dots$,

Es aquí donde se produce la conexión entre los perfiles de los campos receptivos del HVS modelados por las derivadas de Gaussiana, diferentes a éstas por un factor de escala (ecuación 4.16), y los polinomios de Hermite ortonormales respecto a la ventana de análisis.

$$G_n(x) = (-1)^n \cdot H_n(x) e^{-x^2} \quad (4.16)$$

La transformada de Fourier de $D_n(x)$ es

$$d_n(\omega) = \frac{1}{\sqrt{2^n n!}} (j\omega\sigma)^n \cdot e^{-\frac{(\omega\sigma)^2}{4}} \quad (4.17)$$

teniendo un valor extremo cuando $(\omega\sigma)^2 = 2n$, lo cual indica que filtros de órdenes mayores analizan frecuencias sucesivamente más altas de la señal. Sin embargo, para filtros de órdenes muy grandes, los picos en frecuencia se aproximan demasiado, por lo que tales filtros proporcionan poca información adicional. Debido a esto, en la práctica, la transformada de Hermite se limita a pocos términos. Las funciones filtro para $n=0,\dots,4$ se muestran en la Figura 4.2.

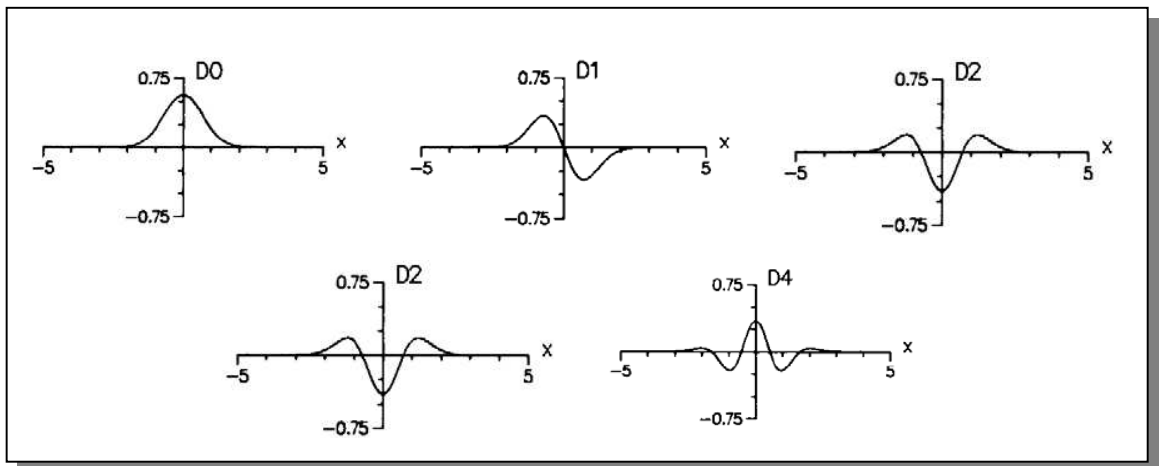


Figura 4.2 Funciones filtro para $\sigma = 1$

Funciones de síntesis

Las funciones de síntesis (funciones patrón) se usan durante la reconstrucción de la señal original que se hace a partir de los coeficientes de la transformada de Hermite. Estas funciones están dadas por la ecuación

$$\begin{aligned} P_n(x) &= H_n\left(\frac{x}{\sigma}\right) V(x) / W(x) \\ &= \frac{T}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{2\pi}} H_n\left(\frac{x}{\sigma}\right) \cdot \frac{e^{-\frac{x^2}{2\sigma^2}}}{w(x)} \end{aligned} \quad (4.18)$$

donde $w(x)$ es la función de ponderación de la ecuación 4.12. Si $w(x)=1$ (para valores de muestreo del parámetro $\tau < 2$), la función patrón es igual a la función Hermite de grado n . La función Hermite tiene la propiedad de ser isomorfa a su transformada de Fourier:

$$P_n(\omega) = \frac{T}{\sqrt{2^n n!}} \cdot (-j)^n H_n(\omega\sigma) e^{-\frac{(\omega\sigma)^2}{2}} \quad (4.19)$$

4.4. Transformada de Hermite Discreta

La transformada de Hermite cuenta con una aproximación discreta (DHT: Discrete Hermite Transform) basada en los polinomios de Krawtchouk. Dado que los filtros de análisis de la transformada de Hermite son similares a las derivadas de Gaussiana excepto por un factor de escala, es posible utilizar la forma discreta de las derivadas de Gaussiana dada por los coeficientes binomiales:

$$C_N^x = \frac{N!}{x!(N-x)!} \quad (4.20)$$

donde N es la longitud de la ventana binomial. Los polinomios ortonormales de Krawtchouk son obtenidos a partir del producto de los polinomios de Krawtchouk por una ventana binomial definida por:

$$V^2(x) = C_N^x / 2^N \quad (4.21)$$

de tal manera que los polinomios ortonormales de Krawtchouk están definidos como:

$$K_n(x) = \frac{1}{\sqrt{C_N^n}} \sum_{k=0}^n (-1)^{n-k} C_{N-x}^{n-k} C_x^k \quad (4.22)$$

para $x=0, \dots, N$ y $n=0, \dots, D_{max}$ donde D_{max} es el máximo orden de la derivada y $D_{max} \leq N$.

Ejemplos de los polinomios de Krawtchouk se presentan a continuación para $N=1$, $N=2$ y $N=4$.

$$K_1 = \frac{1}{\sqrt{2^1}} \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \quad K_2 = \frac{1}{\sqrt{2^2}} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & -2 \\ 1 & -1 & 1 \end{vmatrix} \quad K_4 = \frac{1}{\sqrt{2^4}} \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 4 & 2 & 0 & -2 & -4 \\ 6 & 0 & -2 & 0 & 6 \\ 4 & -2 & 0 & 2 & -4 \\ 1 & -1 & 1 & -1 & 1 \end{vmatrix}$$

La transformada de Hermite discreta de longitud N se aproxima a la transformada de Hermite continua por la siguiente relación $\sigma = \sqrt{N/2}$. Las propiedades de la versión discreta se pueden determinar de forma muy precisa a partir de las propiedades de la versión continua.

Analizando el caso en el que N es par, se tiene que las funciones de análisis y de síntesis pueden ser centradas en el origen desplazando la ventana $N/2$ puntos. Esto lleva a la siguiente definición para las funciones de análisis

$$D_n(x) = K_n\left(\frac{N}{2} - x\right) \cdot V^2\left(\frac{N}{2} - x\right) \quad (4.23)$$

con $x = -(N/2), \dots, (N/2)$. Estas funciones pueden ser expresadas como

$$D_n\left(\frac{N}{2} - x\right) = \frac{(-1)^n}{2^N \sqrt{C_N^n}} \Delta^n [C_N^x \cdot C_x^n] \quad (4.24)$$

donde

$$(-1)^n \Delta^n L(x) = \sum_{k=0}^n (-1)^k C_n^k L(x+k) \quad (4.25)$$

es el operador diferencial. Calculando la transformada Z de esta función filtro se obtiene

$$\begin{aligned} d_n(z) &= \sum_{x=-N/2}^{N/2} D_n(x) z^{-x} \\ &= z^{-N/2} \sqrt{C_N^n} \left(\frac{1-z}{2}\right)^n \left(\frac{1+z}{2}\right)^{N-n} \end{aligned} \quad (4.26)$$

con $n=0, \dots, N$. Estos filtros tienen como ventaja el hecho de que pueden ser realizados aplicando de forma sucesiva (en cascada) una serie de los filtros más sencillos $z^{-1}(1+z)^2$, $z^{-1}(1-z)(1+z)$, $z^{-1}(1-z)^2$, con sus respectivos kérneles $[1 \ 2 \ 1]$, $[-1 \ 0 \ 1]$ y $[1 \ -2 \ 1]$ [50].

Debido a esto, con excepción del término $\sqrt{C_N^n}$, los filtros pueden ser implementados sin multiplicaciones.

4.5. Transformada de Hermite 2D

Un caso especial de la transformada polinomial se presenta cuando la función ventana es separable, es decir $V(x, y) = V(x)V(y)$ y la rejilla de muestreo es cuadrada. Las funciones de análisis y de síntesis son por ende separables y pueden ser implementadas de manera muy eficiente. Como ya se mencionó, en el caso de la transformada de Hermite se trabaja con ventanas Gaussinas, las cuales al ser isotrópicas podemos escribir $\sigma = \sigma_x = \sigma_y$:

$$V(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right) \quad (4.27)$$

Los polinomios de Hermite ortonormales con respecto a $V^2(x, y)$ pueden escribirse entonces como:

$$Q_{n-m,m}(x, y) = \frac{1}{\sqrt{2^n (n-m)! m!}} H_{n-m}\left(\frac{x}{\sigma}\right) H_m\left(\frac{y}{\sigma}\right) \quad (4.28)$$

Los coeficientes polinomiales se calculan convolucionando la imagen localizada $L_{n-m,m}(x, y)$ con las funciones de análisis separables $D_{n-m,m}(x, y) = Q_{n-m,m}(-x, -y)V^2(-x, -y)$ y haciendo enseguida un submuestreo a la señal de salida en las direcciones horizontal y vertical en múltiplos de T; es decir, que para todo $p, q \in S$,

$$L_{n-m,m}(p, q) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} L(x, y) D_{n-m,m}(x - p, y - q) dx dy \quad (4.29)$$

La imagen sintetizada se obtiene por un proceso de interpolar los coeficientes polinomiales con funciones de síntesis $P_{n-m,m}(x, y) = Q_{n-m,m}(x, y)V(x, y)/W(x, y)$ y sumar para todo orden n y m como sigue:

$$L(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{(p,q) \in S} L_{n-m,m}(p, q) P_{n-m,m}(x - p, y - q) \quad (4.30)$$

Los patrones locales de una dimensión tales como bordes y líneas representan un papel importante en la visión temprana, dichos patrones pueden ser encontrados con ayuda de la transformada de Hermite porque dicha transformación emplea como ventana a una función Gaussiana, la cual posee la característica de separabilidad y se sabe que sólo los filtros separables tienen la propiedad de dar curvas de ajuste de orientaciones idénticas para diferentes patrones de una dimensión, tal como bordes, líneas, rayas. En la Figura 4.3 se muestra la descomposición de Hermite, sobre una imagen.

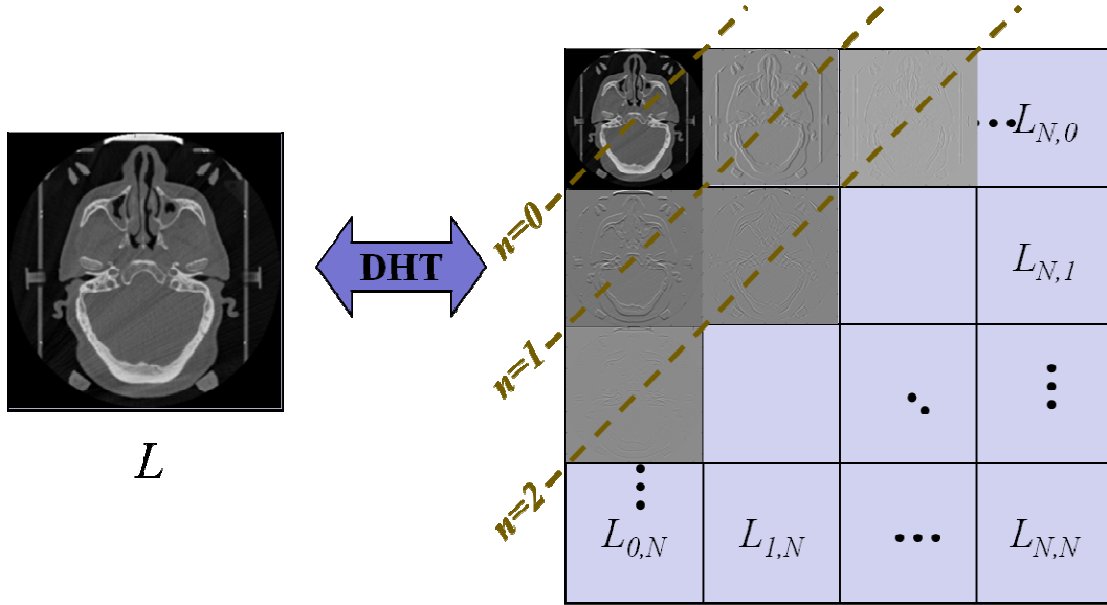


Figura 4.3 Descomposición de una imagen con la transformada de Hermite

4.6. Transformada de Hermite 2D rotada

Freeman y Adelson [44] describen el problema de sintetizar con orientaciones y fases arbitrarias determinando analíticamente su respuesta como una función de orientación. El término de filtro orientable describe una clase de filtros, en la cual un filtro a una orientación arbitraria es sintetizado como una combinación lineal de un conjunto de filtros base.

Con las propiedades de separabilidad y simetría radial de la ventana Gaussiana, es posible incorporar una rotación en los filtros de Hermite que puede ser preestablecida durante el diseño de los filtros [96] o adaptativa [29, 102] de acuerdo a una función de selectividad de orientación de los patrones en la señal 2D. Los $n+1$ filtros de Hermite 2D de *enésimo orden* forman una base de los filtros de orden n sensibles a una orientación θ en particular; en otras palabras, los filtros rotados de orden n pueden ser contruidos a través de combinaciones lineales de los filtros de orden n .

La transformada de Fourier de las funciones de análisis bidimensionales $D_{n-m,m}(x,y)$ expresada en coordenadas polares es:

$$d_{n-m,m}(\omega_x, \omega_y) = \alpha_{n-m,m}(\theta) d_n(\omega) \quad (4.31)$$

donde $\omega_x = \omega \cos \theta$, $\omega_y = \omega \sin \theta$ y $d_n(\omega)$ es la transformada de Fourier de la función filtro de Hermite en una dimensión $D_n(r)$ donde r es la coordenada radial. De esta manera, se observa que la transformada de Fourier de cada función de análisis puede escribirse como un producto que separadamente expresa una preferencia de orientación y una preferencia de frecuencia

radial. La frecuencia radial de las funciones de análisis está dada por la ecuación 4.17, mientras que la selectividad direccional del filtro se expresa por la siguiente función angular:

$$\alpha_{n-m,m}(\theta) = \sqrt{C_n^m} \cos^{n-m} \theta \sin^m \theta \quad (4.32)$$

A partir de la ecuación 4.32 se puede ver que los filtros cuyo orden es el mismo, n , pero con distinto índice direccional m distinguen entre diferentes orientaciones en la imagen.

En base a lo anterior podemos escribir la transformada de Hermite rotada de la siguiente forma:

$$L_{n-m,m}^\theta(x,y) = \sum_{m=0}^n L_{n-m,m} \cdot \alpha_{n-m,m}(\theta) \quad (4.33)$$

donde $L_{n-m,m}^\theta(x,y)$ denota a una versión rotada o proyectada de los coeficientes $L_{n-m,m}(x,y)$ con respecto a un ángulo de preferencia θ .

Para su obtención, en el caso de una rotación adaptativa [29], primero se aplica la transformada de Hermite y los coeficientes de esta transformación son rotados en una orientación local estimada, de acuerdo al criterio de máxima energía de orientación para cada posición de ventana. Esto implica que estos filtros pueden indicar la orientación de un patrón 1D independiente de su estructura interna.

Los coeficientes polinomiales de dos dimensiones se proyectan en coeficientes de una dimensión sobre un eje que hace un ángulo θ con el eje x . Martens [71] demostró que los coeficientes polinomiales de primer orden son una buena aproximación de un detector de bordes óptimo, $\theta = \tan^{-1}(L_{01}/L_{10})$. La Figura 4.4 muestra los pasos para una descomposición direccional de Hermite sobre una imagen, observándose que la energía de los coeficientes se concentra en la dirección horizontal.

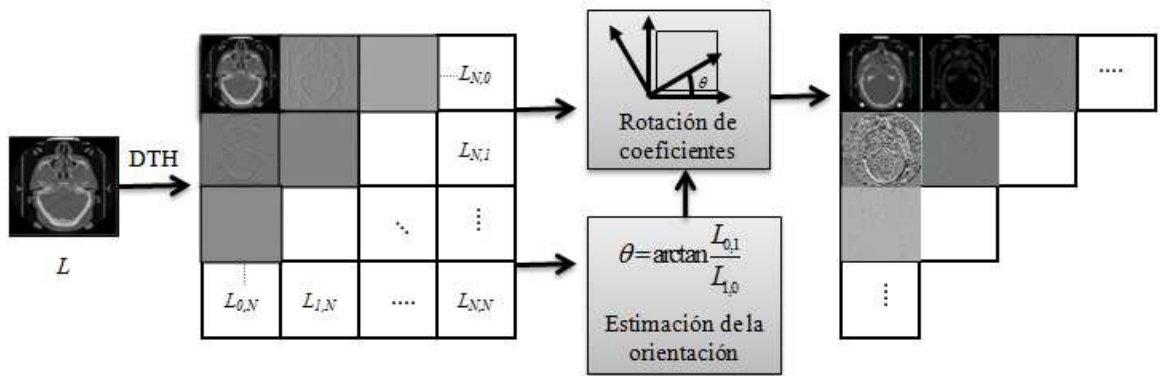


Figura 4.4 Descomposición de una imagen con la transformada de Hermite rotada

Así, para el análisis de orientación, resulta conveniente trabajar con versiones rotadas de la transformada Hermite ya que sus filtros se adaptan al contenido con una orientación local en

cada posición de la imagen; lo cual además presenta la ventaja de compactar la energía en algunos coeficientes.

4.7. Transformada de Hermite multirresolución 2D

En la mayoría de las aplicaciones se desea saber que parte de la información contenida en una imagen debe agregarse con mayor intensidad (por ejemplo, bordes, líneas, esquinas, etc.) y que parte de información debe ser atenuada porque resulta molesta a la vista (por ejemplo, texturas finas o ruido). En tal caso, resulta apropiado llevar a cabo un análisis derivativo para discernir entre las diferentes estructuras de interés [43, 102]. Los cambios locales se descubren mejor con ventanas pequeñas, pero para descubrir representaciones de objetos con baja resolución se necesitan ventanas mayores; no obstante, existe un compromiso en la representación multirresolución para encontrar buenas soluciones.

En representaciones espacio-escala bidimensionales una señal $L(x,y)$ es convolucionada con versiones escaladas de una Gaussiana normalizada para eliminar detalles, es decir, $L(x,y,s)=L(x,y)*G(x,y,s)$ es la representación de la señal a la escala s , donde $G(x,y,s)$ es la Gaussiana con el parámetro de escala s (un medio de la varianza de la Gaussiana). La estructura local de una señal se infiere a partir de las derivadas de las representaciones escaladas, las cuales se pueden obtener convolucionando la función $L(x,y)$ con los operadores derivadas de Gaussiana escalados $G_{n-m,m}(x,y,s)=G_{n-m}(x,s)G_m(y,s)$ para $m=0,1,\dots,n$ y $n=0,1,\dots$, donde n es el orden de derivación total.

Así, para las señales bidimensionales la representación multirresolución se lleva a cabo filtrando la señal en cascada a lo largo de cada coordenada bajo un esquema piramidal, en la cual la imagen se descompone en un número de subimágenes pasobajas o pasobanda, las cuales son submuestreadas en proporción a su resolución obteniéndose la información de las estructuras de la imagen a diferentes escalas (ver Figura 4.5).

Otro factor físico que influye en la representación de los patrones visuales es la inclinación relativa del sistema de visión con respecto a un eje de referencia en la escena. Dicha inclinación hará que las estructuras de la imagen se perciban con una orientación determinada, la cual sólo puede ser vista por operadores sensibles a la orientación. De esta manera, para propósitos de análisis de orientación es conveniente trabajar con las versiones rotadas de estos operadores, es decir, $G_{n-m,m}(x,y,s,\theta)=G_{n-m,m}(xcos\theta+ysen\theta, -xsen\theta+ycos\theta,s)$, donde θ es el ángulo de rotación [103]. En particular, las derivadas direccionales de Gaussiana denotadas por $G_{n,0}(x,y,s,\theta)$ se expresan como

$$G_{n,0}(x,y,s,\theta)=\sum_{m=0}^n \alpha_{n-m,m}(\theta) G_{n,m,m}(x,y,s) \quad (4.34)$$

donde $m=0,1,\dots,n$ y $n=0,1,\dots$, con la función de ángulo de la ecuación 4.32. En este caso, la expansión de las diferencias de Gaussianas (DoG: Difference of Gaussians) se escribe de la siguiente manera:

$$DoG(x, y, s_{k-1}, s_k) = \sum_{n=1}^{\infty} \sum_{j=0}^n \frac{c_n (-\tau s_k)^n}{n!} G_{2n,0}(x, y, s_k, \theta_j) \quad (4.35)$$

donde $\theta_j = \theta_0 + j\pi / (n+1)$ para $j=0, \dots, n$ y

$$c_n = \frac{1}{\sum_{j=1}^n \sin^{2n}\left(\frac{j\pi}{n+1}\right)} = \frac{1}{n+1} \sum_{m=0}^n \frac{(C_n^m)^2}{C_{2n}^{2m}} \quad (4.36)$$

para $n=0, 1, \dots$. Nótese que el orden de derivación n determina la resolución angular y que el ángulo θ_0 es un parámetro libre que puede elegirse según convenga; mientras que el parámetro de espaciamiento entre escalas, τ , generalmente se requiere que sea constante para todo entero k en $\tau = (s_k - s_{k-1})$, lo cual está de acuerdo con los datos psicofísicos que sostienen la existencia en el HVS de un número de canales cuyas frecuencias centrales mantienen una relación de aproximadamente una octava, es decir, $\tau \approx 0.75$.

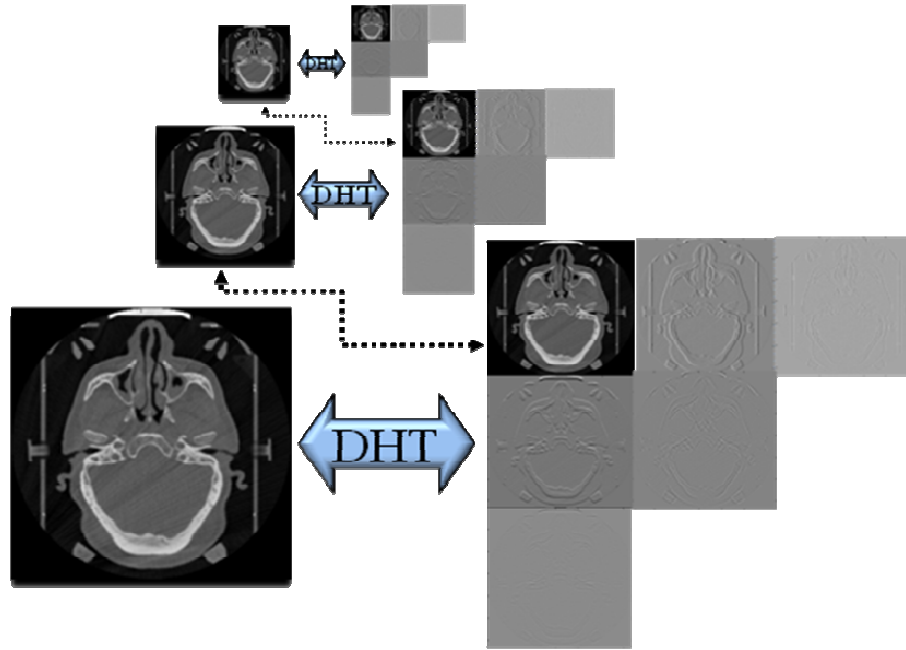


Figura 4.5 Expansión multirresolución de la transformada de Hermite

Partiendo de la expresión 4.35 es posible construir una transformada que resulte más apropiada para la descripción de imágenes, ya que las funciones de análisis que se obtienen incluyen las tres transformaciones geométricas fundamentales: traslación, rotación y escalamiento. El proceso de análisis se formaliza interpretando las operaciones de convolución como producto interno en la siguiente expresión

$$L_n^{(k,j)}(\xi, \eta) = \left\langle L(x, y), G_n^{(k,j)}(x, y, \xi, \eta) \right\rangle_{(x,y)} \quad (4.37)$$

mediante las funciones de análisis

$$G_n^{(k,j)}(x, y, \xi, \eta) = \frac{1}{2s_k} G_{n,0}^* \left(\frac{x \cos \theta_j + y \sin \theta_j}{\sqrt{2s_k}} - \xi, \frac{-x \sin \theta_j + y \cos \theta_j}{\sqrt{2s_k}} - \eta \right) \quad (4.38)$$

donde $\theta_j = \theta_0 + j\pi / (n+1)$ para $j=0, \dots, n$; $n=1, 2, \dots$ y $s_k = (1-\tau)^{-k}$ para todo k entero. Este análisis es referido como **transformada de Hermite multiresolución multidireccional** (MMHT: multidirectional multiresolution Hermite transform) debido a que emplea las derivadas direccionales de la Gaussiana a múltiples escalas y orientaciones y en consecuencia los coeficientes generados $L_n^{(k,j)}$ contienen no sólo información sobre la localización sino también sobre orientación de las estructuras de la imagen a distintas escalas. De manera análoga, al proceso de reconstrucción de la señal a partir de esta representación se le denomina MMHT inversa y se expresa como

$$L(x, y) = \sum_{k=-\infty}^{\infty} \sum_{n=1}^{\infty} \sum_{j=0}^n c_n \tau^n D_n^{(k,j)}(x, y) \quad (4.39)$$

donde las funciones $D_n^{(k,j)}(x, y)$ para $n=1, 2, \dots$ aportan los detalles direccionales de la señal a todas las escalas.

$$D_n^{(k,j)}(x, y) = \left\langle L_n^{(k,j)}(\xi, \eta), G_n^{(k,j)}(x, y, \xi, \eta) \right\rangle_{(\xi, \eta)} \quad (4.40)$$

Para el caso discreto, el filtro DoG puede ser aproximado por una Diferencia de Binomiales (DoB) y dado que las funciones Binomiales son de soporte compacto, su representación mediante las diferencias binomiales se expresa como una suma finita y a partir de esta aproximación se construye la versión discreta de la MMHT [103].

4.8. Transformada de Hermite 3D

En el caso tridimensional [70] los coeficientes de Hermite $L_{l,m-l,n-m}$ se obtienen mediante la convolución de la señal original $L(x, y, z)$ con los filtros de análisis $D_l(x)$, $D_{m-l}(y)$ y $D_{n-m}(z)$ seguido de un submuestreo sobre una malla tridimensional S . Por otro lado, la transformada de Hermite discreta inversa se define como:

$$L(x, y, z) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{l=0}^m \sum_{(p,q,r) \in S} L_{l,m-l,n-m}(p, q, r) P_l(x-p) P_{m-l}(y-q) P_{n-m}(z-r) \quad (4.41)$$

donde $P_l(x)$, $P_{m-l}(y)$ y $P_{n-m}(z)$ son los filtros de síntesis para las direcciones x , y y z respectivamente. La Figura 4.6 muestra la distribución tridimensional de los coeficientes de la transformada de Hermite hasta orden 2 en cada dirección, de manera que en lugar de tener 9 coeficientes por cada pixel en una imagen, en este caso se obtienen 27 coeficientes por cada voxel de un volumen.

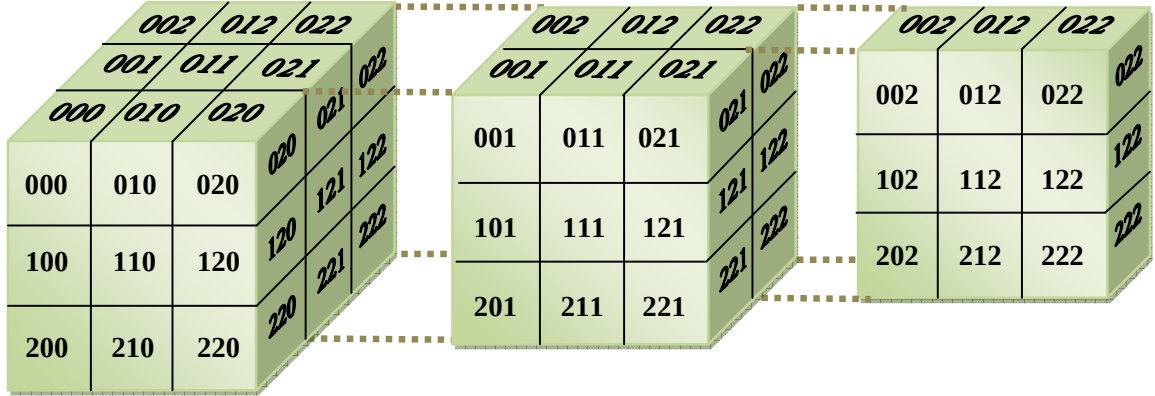


Figura 4.6 Distribución de los coeficientes de segundo orden de un voxel

La transformada de Fourier de los filtros de análisis se puede expresar en coordenadas esféricas de la siguiente manera:

$$d_1(\omega_x) d_{m-l}(\omega_y) d_{n-m}(\omega_z) = \alpha_{l,m-l}(\theta) \alpha_{m,n-m}(\phi) d_n(\omega) \quad (4.42)$$

donde $\omega_x = \omega \cos \theta \cos \phi$, $\omega_y = \omega \sin \theta \cos \phi$, $\omega_z = \omega \sin \phi$ y $d_n(\omega)$ es la transformada de Fourier del filtro de Hermite unidimensional $D_n(r)$, donde r es la coordenada polar y la función α que expresa la selectividad direccional del filtro es la misma que la introducida en la ecuación 4.32.

Por otro lado, el mejor ajuste de la señal original $L(x,y,z)$ mediante un patrón K unidimensional

$$K((x \cos \theta + y \sin \theta) \cos \phi + z \sin \phi) \quad (4.43)$$

se encuentra al maximizar la energía direccional

$$\sum_{n=0}^{\infty} K_{n,\theta,\phi}^2 = \sum_{n=0}^{\infty} \left[\sum_{m=0}^n \sum_{l=0}^m \alpha_{l,m-l}(\theta) \cdot \alpha_{m,n-m}(\phi) L_{l,m-l,n-m} \right] \quad (4.44)$$

para todo (θ, ϕ) .

En consecuencia, la mejor aproximación 1D de los coeficientes de Hermite 3D está dada por

$$\hat{L}_{l,m-l,n-m} = K_{n,\theta,\phi} \alpha_{l,m-l}(\theta) \alpha_{m,n-m}(\phi) \quad (4.45)$$

para $n = 0, \dots, \infty$, $m = 0 \dots n$ y $l = 0 \dots m$, con (θ, ϕ) como los ángulos óptimos, donde $\hat{L}_{l,m-l,n-m}$ son los coeficientes de Hermite rotados. La Figura 4.7 muestra la HT y la HT rotada sobre un

volumen mostrando el cambio de sistema de referencia y la aproximación de los ángulos θ y ϕ de acuerdo a la dirección del gradiente.

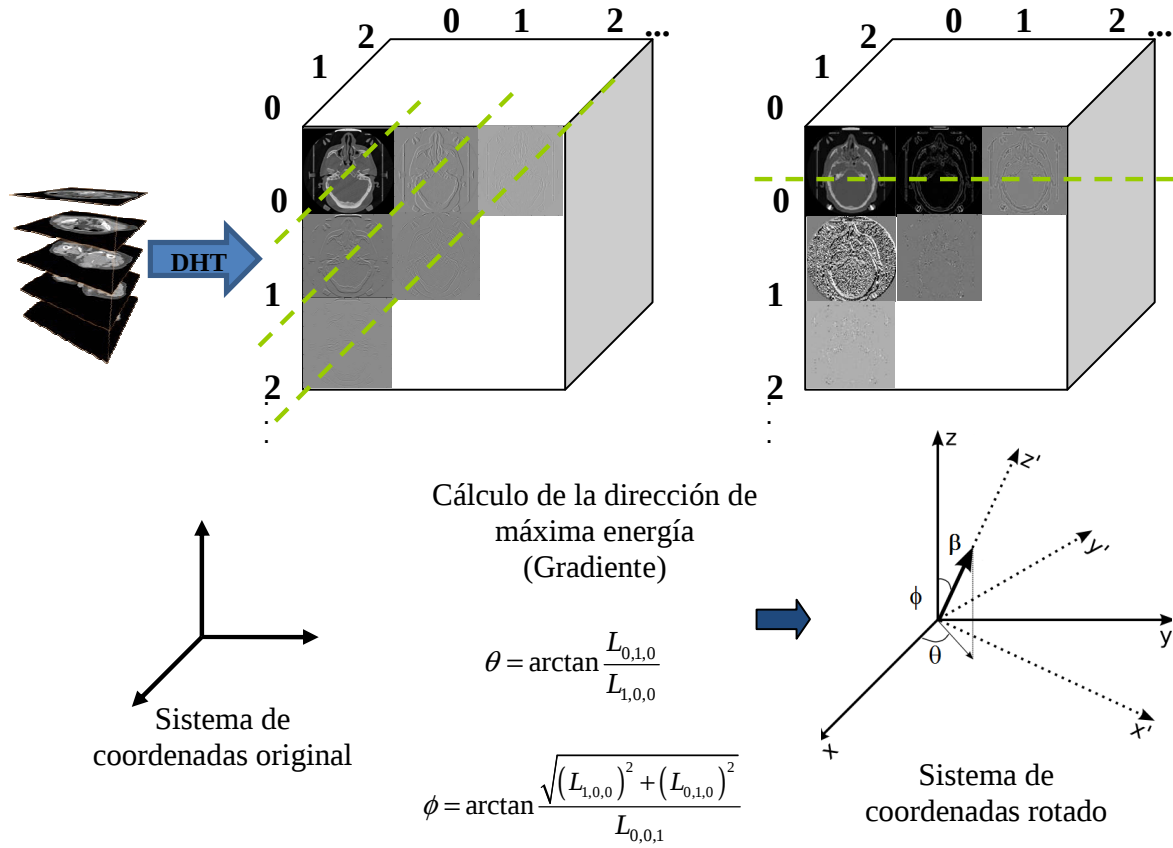


Figura 4.7 La transformada de Hermite 3D en un sistema de coordenadas cartesiano y su rotación sobre un volumen.

Una descomposición multirresolución usando la HT puede ser obtenida a través de un esquema piramidal [43, 102]. En una descomposición piramidal el volumen es descompuesto en un número de subvolúmenes paso banda o paso baja, los cuales son posteriormente submuestreados en proporción a su resolución espacial. En cada subvolumen se transforman los coeficientes de orden cero para obtener una versión escalada de la anterior.

Una vez obtenida la descomposición en coeficientes de Hermite de cada nivel, los coeficientes pueden ser proyectados a una dimensión por su orientación local de máxima energía, de esta forma se obtiene la **transformada de Hermite multirresolución localmente orientada 3D**. Así, este esquema multirresolución de la HT permite obtener información sobre la localización y orientación de la estructura de la imagen a distintas escalas.

Principios de evaluación de la fusión de imágenes

5.1. Introducción

En este capítulo se discuten las métricas de calidad para cuantificar el desempeño de los sistemas de fusión a nivel pixel, cuyo objetivo principal es la preservación de la información visual de las imágenes de entrada. Desde luego, que la información visual puede definirse de varias maneras, y de acuerdo con ello podrían formularse varias métricas. Inicialmente, la información visual que se considera importante está relacionada la información basada en los bordes y las regiones asociadas de la imagen. Este concepto se extiende con la participación del modelado del sistema de visión humano dentro del proceso de formulación de las métricas [128].

El método más confiable y directo para evaluar el desempeño de la fusión de imágenes está basado en pruebas subjetivas que emplean muestras representativas de usuarios. Sin embargo, esta forma de evaluación es costosa en términos de tiempo, esfuerzo y equipo requerido, pues este proceso involucra la participación de varios sujetos que estén dispuestos a evaluar las imágenes bajo condiciones controladas, ya sea comparando varios resultados de fusión o ejecutando específicas tareas visualmente orientadas.

Por ello se prefiere el procesamiento de los datos con el propósito de obtener métricas de desempeño significativas que sean más objetivas, las cuales además dan la oportunidad de utilizar esta cuantificación en el desarrollo de sistemas de fusión de imágenes. De esta forma una métrica de evaluación precisa y objetiva podría utilizarse para proveer información sobre el desempeño de un sistema de fusión y a su vez para guiar el proceso de diseño del algoritmo de fusión como un parámetro de optimización. No obstante, sólo un número limitado de métodos de fusión de imágenes dependen del desempeño de las métricas, pues en general esos métodos están basados en el conocimiento de la aplicación específica.

Asimismo, existe la idea de comparar las imágenes de salida de los algoritmos de fusión con una imagen fusionada ‘ideal’ en orden de tener una estimación del desempeño de fusión. La imagen ideal se obtiene de imágenes con diferentes puntos de foco que son fusionadas manualmente al recortar y pegar para producir una imagen de referencia ideal [128]. Sin embargo, ese método no es aplicable a varias aplicaciones de fusión donde la imagen fusionada ideal no puede ser definida fácilmente ya que no hay forma de obtenerla de forma manual, y este es el caso de las imágenes médicas.

En el contexto de las imágenes médicas, la medida final de la calidad de una imagen reside en lo útil que es la imagen para conseguir un diagnóstico certero y preciso; por ejemplo, la verdadera prueba para un radiólogo es la detección precisa de incluso pequeñas anormalidades [80].

5.2. Clasificación de las métricas de calidad

Se define la calidad de una imagen como: “la capacidad que tiene una imagen de representar el objeto original”. En las métricas de calidad, en vez de tener una imagen y un objeto, la definición se extiende a dos imágenes, donde una se considera la imagen de referencia u original y la otra es la imagen con alguna distorsión. En este caso, la medida de calidad es una relación de semejanza entre la imagen original y la distorsionada.

La definición de calidad de imagen arriba mencionada se conoce como fidelidad de la imagen y también tiene algunos inconvenientes, ya que no es claro que la visibilidad del error esté relacionada con la pérdida de calidad. De hecho, algunas distorsiones que pueden ser claramente visibles, no son a pesar de ello, molestas para el observador [45]. En el contexto de la fusión de imágenes, la imagen fusionada ideal debe contener toda la información visual importante cómo sería percibida por un observador que observa las imágenes de entrada, y al mismo tiempo debe carecer de distorsión o ‘falsa’ información en su contenido.

Así, una medida de desempeño de fusión de imágenes debe ser capaz de: (i) identificar y localizar información visual en imágenes de entrada y en las imágenes fusionadas, (ii) evaluar su importancia perceptiva y (iii) cuantificar la precisión con la cual la información de entrada es representada en la imagen fusionada. Adicionalmente, tal métrica debe distinguir entre la verdadera escena (información de entrada) y los artefactos de fusión, los cuales pueden aparecer en la imagen fusionada [128].

De manera general, las métricas de calidad se dividen en dos grupos principales: las métricas de calidad subjetivas (utilizando observadores) y las métricas de calidad objetiva (medidas matemáticas); de ambas se pueden derivar otras clasificaciones.

En el caso de las métricas de calidad subjetivas es posible identificar dos clasificaciones [86]:

- De acuerdo al rol que toma el sujeto en las pruebas, se clasifican en activas y pasivas, en las primeras, al sujeto se le pide realizar una acción de acuerdo a la información obtenida de la imagen fusionada o secuencia de imágenes, de manera que el desempeño de la fusión se mide a través del desempeño de los sujetos al realizar la tarea requerida. En las pruebas pasivas el usuario sólo expresa una preferencia por alguna imagen de las alternativas que se le presentan, de esta forma la evaluación de la fusión se basa en la percepción de la calidad de la imagen.
- De acuerdo al nivel de control de las condiciones bajo las cuales las pruebas son realizadas, se clasifican en formales e informales. Las pruebas formales requieren el cumplimiento estricto de diversas condiciones: tamaño de despliegue, iluminación, distancia de los sujetos del despliegue, tiempo de exposición o de realización de la

tarea, entre otras. Las pruebas informales emplean un conjunto de condiciones menos estricto, las cuales usualmente varían dentro de un rango razonable.

Las métricas de calidad objetivas cuentan también con diversas clasificaciones:

- De acuerdo a la disponibilidad de la imagen original (libre de distorsión), con la cual la imagen distorsionada es comparada, es posible distinguir tres grupos: métricas de referencia completa, lo que implica la existencia una imagen de referencia⁶ que se conoce en su totalidad, métricas de calidad ‘ciegas’ o sin referencia donde tal imagen de referencia no está disponible y métricas de calidad de referencia limitada donde la imagen de referencia está parcialmente disponible como un conjunto de características extraídas que pueden ser empleadas para la evaluación de calidad [123].
- Dependiendo si las métricas incorporan o no el HVS, los grupos quedan definidos como métricas de error, métricas de calidad perceptible y métricas híbridas [45]. Las métricas de error obtienen una medida de calidad sin tener en cuenta las características del HVS y están orientadas a medir la diferencia entre imágenes. Las métricas de calidad perceptiva son aquellas que incorporan las características del HVS. Finalmente, existe un tercer grupo de métricas híbridas que a pesar de no incluir un modelo de HVS, tratan de obtener valores que se adecuen a la calidad que el observador percibe explotando otras cualidades relativas a la percepción.
- Finalmente, la clasificación puede estar ligada al contexto de la aplicación, por ejemplo, en la fusión de imágenes de detección remota es común dividir las métricas de evaluación de calidad espectral y métricas de calidad espacial [90].

5.3. Métricas de calidad subjetivas

Las pruebas subjetivas se utilizan para obtener una estimación más precisa de la calidad de las imágenes ya que este tipo de métodos son necesarios para comprobar la precisión de los modelos de calidad objetivos, es decir, debe existir una correlación entre los resultados obtenidos mediante el uso de modelos subjetivos y los que se obtienen con los objetivos. Sin embargo, en la práctica resultan muy pesadas en cuanto a la carga temporal que conllevan, además de resultar costosas y difíciles de repetir, puesto que requieren condiciones de visualización especiales [54].

Algunos de los problemas que plantea su utilización son el costo en tiempo y personas para la realización de pruebas ya que lo recomendable es efectuar los test el mayor número de veces posible para conseguir buenos resultados, el efecto del contexto que se refiere al orden e intensidad de las distorsiones que se presentan, la influencia de la escala de puntuación, las condiciones de visibilidad en las que se realizan las pruebas, y la falta de información espacial o temporal de calidad debido a que el valor que proporciona el observador es global y esto ocasiona que no se pueda saber dónde o cuándo se produce este error [45].

⁶ La imagen de referencia es usualmente referida en la literatura como *ground truth*.

Uno de los métodos considerados como el mejor método para la calidad de imagen que ha sido utilizado por varios años es la métrica puntuación de opinión media (MOS: mean opinion score). La métrica MOS es generada promediando los resultados de un conjunto estándar, pruebas subjetivas, donde un número de personas califican la calidad de una serie de imágenes basándose en la recomendación ITU-T J247 [100]. MOS es el promedio aritmético de todas las puntuaciones individuales y tiene un rango de 1 (la calidad más baja) a 5 (la mejor).

Existen además otros procedimientos para la realización de este tipo de evaluaciones y la utilización de uno u otro método depende del problema que se pretende evaluar. No obstante, debido a los inconvenientes que sugiere la evaluación de la calidad en la fusión mediante este tipo de métricas, no se profundizará más en otros tipos de pruebas.

5.4. Métricas de calidad objetivas

Las métricas de calidad objetiva proporcionan un valor numérico de la calidad de una imagen utilizando operaciones matemáticas en las que eventualmente pueden incluirse funcionalidades del HVS [80]. Se pueden utilizar para la optimización de algoritmos en procesamiento de imágenes ya que una de sus principales ventajas frente a los métodos subjetivos es que proporciona una medida casi inmediata basándose completamente en las entradas y la imagen fusionada, y por ello sirven para monitorizar de forma dinámica la calidad de la imagen.

Las métricas objetivas de calidad de imagen se basan en mediciones de características físicas e intentan predecir la calidad perceptiva con precisión y de manera automática. Esto significa que ellas deben predecir la calidad de la imagen que un observador humano promedio podría reportar. Un hecho importante sobre este asunto es la disponibilidad de una “imagen original”, la cual es considerada libre de distorsión o con calidad perfecta. La mayoría de las métricas de calidad objetivas asumen que la imagen de referencia existe e intentan calificar el error entre una imagen distorsionada y una imagen de referencia [40].

Entre las formas disponibles para evaluar objetivamente la calidad de imagen, el error cuadrático medio (MSE: mean square error) y la relación señal pico a ruido (PSNR: peak signal-to-noise ratio) son los más empleados debido a su facilidad de cálculo y a su bajo costo computacional, sin embargo, tales métricas no son necesariamente consistentes con la evaluación de un observador humano [123]. Tanto MSE como PSNR reflejan las propiedades globales de la calidad de imagen pero ellos son ineficientes en la predicción de degradaciones estructurales. Es por ello que en general se recurre al uso de más de una métrica para evaluar el desempeño de un algoritmo de fusión, y con este propósito, a continuación se enumeran algunas métricas objetivas que posteriormente serán empleadas para comparar el resultado de la fusión con otros métodos de fusión.

5.4.1. Métricas de calidad basadas en el error

Estas métricas también se conocen como métricas relacionadas a características estadísticas [120] ya que se basan en funciones matemáticas simples; se basan en el error debido a que obtienen un valor de calidad en términos de variaciones entre la imagen procesada y la original. Generalmente trabajan en el dominio espacial realizando operaciones punto a punto en la imagen. Las más utilizadas son la relación señal a ruido de pico PSNR (5.1) y el error cuadrático medio MSE (5.2).

$$PSNR = 10 \log_{10} \frac{255^2 (MN)}{\sum_{i=1}^M \sum_{j=1}^N [R(i, j) - F(i, j)]^2} \quad (5.1)$$

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [R(i, j) - F(i, j)]^2}{MN} \quad (5.2)$$

donde $F(i, j)$ denota la intensidad del pixel de la imagen fusionada y $R(i, j)$ denota la intensidad del pixel de la imagen original en la posición (i, j) . Si extrapolamos el concepto de estas métricas a la fusión de datos tridimensionales donde $F(i, j, k)$ denote la intensidad del voxel del volumen fusionado y $R(i, j, k)$ denote la intensidad del voxel del volumen original en la posición (i, j, k) , las expresiones quedan de la siguiente forma:

$$PSNR = 10 \log_{10} \frac{255^2 (MNO)}{\sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^O [R(i, j, k) - F(i, j, k)]^2} \quad (5.3)$$

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^O [R(i, j, k) - F(i, j, k)]^2}{MNO} \quad (5.4)$$

La **PSNR** es un radio entre la potencia máxima posible de la imagen reconstruida y la potencia del ruido que afecta la fidelidad de la reconstrucción, por lo que a mayor PSNR, mejor es la calidad de la imagen fusionada. Por el contrario, el **MSE** indica el nivel de error entre la imagen fusionada y la imagen de referencia, así, el menor valor de MSE indica el mejor desempeño del método de fusión.

Cuando se evalúa el desempeño de una técnica de fusión de imágenes empleando alguna de estas métricas se requiere conocer la imagen original. Por esta razón, estas medidas de error sólo pueden ser aplicadas en datos sintéticos (simulados). En tal caso se siguen los siguientes pasos:

- 1) De las imágenes originales se generan las imágenes fuente a ser procesadas, cada una con algún tipo de distorsión dependiendo de las condiciones que se deseen simular. La imagen fusionada ideal es considerada idéntica a la imagen original.

- 2) Las imágenes fuente generadas son fusionadas empleando las técnicas de fusión a evaluar.
- 3) Finalmente se calculan las métricas PSNR y MSE de la imagen fusionada con respecto a la imagen original.

Finalmente, cabe mencionar que estas métricas presentan el inconveniente de proporcionar una idea global sobre la calidad de una imagen, sin embargo en los casos en que la imagen fusionada exhibe artefactos concentrados en un área pequeña, estas medidas todavía pueden producir un valor aceptable, incluso si la imagen es visualmente inaceptable [91]. Una modificación a MSE para reducir la sensibilidad que presenta a cambios globales en la intensidad es el NMSE (normalized mean square error), que normaliza las intensidades haciendo que las imágenes tengan media cero y variancia uno para posteriormente aplicar el MSE [45].

5.4.2. Métricas de calidad basadas en la información mutua

La **información mutua** (MI: mutual information) es un concepto de la teoría de la información que ha sido propuesto como una medida de evaluación de la fusión de imágenes en ausencia de imagen de referencia (ground truth) [120]. Se define como una medida de la dependencia estadística de dos variables aleatorias, cuantificando la cantidad de información que una variable contiene de la otra. La cantidad de información que corresponde a la imagen A contenida en la imagen fusionada F está determinada por

$$MI_{FA}(f, a) = \sum_{f, a} P_{FA}(f, a) \log \left[\frac{P_{FA}(f, a)}{P_F(f) P_A(a)} \right] \quad (5.5)$$

donde f y a denotan los valores de intensidad de imagen de un par de píxeles correspondientes en dos imágenes, consideradas muestras generadas de variables aleatorias F y A , respectivamente. La estimación de las distribuciones conjunta y marginales $p_{FA}(f, a)$, $p_F(f)$ y $p_A(a)$, puede obtenerse mediante la normalización de los histogramas conjunto y marginales de ambas imágenes. Por lo tanto, la información mutua total (5.6) se calcula como la cantidad de información que la imagen fusionada contiene de A y B .

$$MI_F^{AB} = MI_{FA}(f, a) + MI_{FB}(f, b) \quad (5.6)$$

Otra métrica de evaluación es la **simetría de fusión** (FS: fusión symmetry), la cual denota la simetría del proceso de fusión en relación a dos imágenes de entrada [120]. Entre menor sea el FS, mejor es el desempeño del método de fusión. De acuerdo a esta definición la MI es importante cuando una de las dos fuentes es menos importante que la otra, sin embargo, si ambas imágenes de entrada tienen alta calidad, el parámetro FS cobra importancia, y el método con mejor desempeño en ese caso debe ser aquél con el valor más alto de MI y el más bajo en FS.

$$FS = abs \left(\frac{MI_{FA}(f, a)}{MI_{FA}(f, a) + MI_{FB}(f, b)} - 0.5 \right) \quad (5.7)$$

Para el caso de volúmenes estas métricas se calculan de la misma forma, lo que cambia es el cálculo de los histogramas marginales y el histograma conjunto ya que se tiene que considerar una dimensión más.

5.4.3. Medidas de calidad basadas en la información estructural

Las imágenes son señales altamente estructuradas ya que sus píxeles exhiben fuertes dependencias, especialmente cuando éstos están espacialmente próximos, y esas dependencias contienen información importante sobre la estructura de los objetos en la escena visual [123]. Inspirándose en este argumento, se han propuesto medidas de calidad que miden la preservación de la información estructural en la imagen fusionada.

El **índice de similitud estructural** (SSIM: structural similarity index measure) es un modelo propuesto por Wang y Bovik [122, 123] que se basa en la idea de que el HVS extrae información de la estructura de los objetos del campo visual e incorpora este aspecto a la métrica y provee un valor de calidad en el rango de [0,1], siendo 1 el mejor valor de calidad.

Este índice compara los patrones locales de las intensidades de los píxeles que han sido normalizados mediante luminancia y contraste, para ello primero se realiza un cálculo de luminancia que consiste en una media matemática, seguido del cálculo de contraste usando la desviación estándar y finalmente los atributos que representan a la estructura de los objetos en escena se modelan como una correlación, esto se expresa en la siguiente ecuación:

$$Q_0(R, F) = \frac{4\sigma_{RF}\mu_R\mu_F}{(\sigma_R^2 + \sigma_F^2)[(\mu_R)^2 + (\mu_F)^2]} \quad (5.8)$$

el cual puede ser descompuesto como:

$$Q_0(R, F) = \frac{\sigma_{RF}}{\sigma_R\sigma_F} \frac{2\mu_R\mu_F}{(\mu_R)^2 + (\mu_F)^2} \frac{2\sigma_R\sigma_F}{\sigma_R^2 + \sigma_F^2} \quad (5.9)$$

donde μ_R es la media de la imagen original y μ_F la media de la imagen fusionada; σ es la varianza y σ_{RF} es la covarianza entre la imagen original y la imagen fusionada. En la ecuación 5.9 el primer término corresponde a la comparación estructural (correlación entre la imagen original R y la imagen distorsionada F), el segundo término es la distorsión de la luminancia y el tercero evalúa la diferencia de contraste.

Tomando como referencia el SSIM, Piella [87] propuso una métrica de calidad que no requiere imagen de referencia y utiliza medidas locales para estimar que tan bien se representa en la imagen fusionada la información destacada de las imágenes de entrada.

Sea $s(A|w)$ alguna característica de interés de la imagen A dentro de una ventana w , el cual debe reflejar la relevancia local de la imagen a dentro de dicha ventana que puede depender, por ejemplo del contraste, la varianza y la entropía; y $s(B|w)$ la característica de interés de la imagen B en w , se calcula un peso local $\lambda_A(w)$ entre 0 y 1 indicando la importancia relativa de la imagen a comparada con la imagen b . Una elección típica para $\lambda_A(w)$ es

$$\lambda_A(w) = \frac{s(A|w)}{s(A|w) + s(B|w)} \quad (5.10)$$

De manera similar se calcula $\lambda_B(w)$, de manera que la métrica de calidad entre la imagen fusionada F y las imágenes de entrada A y B , $Q(A,B,F)$, se define como

$$Q(A, B, F) = \frac{1}{|W|} \sum_{w \in W} (\lambda_A(w) Q_0(A, F|w) + \lambda_B(w) Q_0(B, F|w)) \quad (5.11)$$

donde $Q_0(A, F|w)$ es el SSIM local calculado para los valores de A y F que están dentro de una ventana w , W es la familia de todas las ventanas y $|W|$ es la cardinalidad de W .

En la ecuación 5.11, los diferentes índices obtenidos en cada ventana son tratados de la misma manera, sin embargo, el HVS da mayor importancia a las regiones visualmente más destacadas en una imagen. Debido a esto, se propone una variante de este índice para dar más peso a aquellas ventanas donde los rasgos sobresalientes de las imágenes de entrada son mayores, que corresponden a las áreas en las cuales hay partes perceptuales más importantes en la escena. La estructura importante global de una ventana están definidas como $C(w) = \max(s(A|w), s(B|w))$. La métrica de calidad de fusión ponderada está entonces definida como

$$Q_W(A, B, F) = \sum_{w \in W} c(w) (\lambda_A(w) Q_0(A, F|w) + \lambda_B(w) Q_0(B, F|w)) \quad (5.12)$$

donde $c(w) = C(w) / (\sum_{w' \in W} C(w'))$.

Una última modificación al índice Q que toma en cuenta la información de bordes es el Q_W , el cual se calcula a partir de imágenes de bordes (por ejemplo, la norma del gradiente) en lugar de las imágenes originales A , B y F . Para ello se combinan $Q_W(A,B,F)$ y $Q_W(A',B',F')$ en un índice de calidad dependiente de los bordes (5.13).

$$Q_E(A, B, F) = Q_W(A, B, F)^{1-\alpha} \cdot Q_W(A', B', F')^\alpha \quad (5.13)$$

donde, el parámetro $\alpha \in [0,1]$ expresa la contribución de las imágenes de bordes comparada con las imágenes originales.

Los últimos tres índices presentados, Q , Q_W y Q_E tienen un rango dinámico de $[-1,1]$. El valor más cercano a 1 denota una mejor calidad en la imagen fusionada.

5.4.4. Medidas de calidad basadas en coeficientes de correlación

El **coeficiente de correlación** es una métrica de calidad espectral y está ligado a la similitud entre dos imágenes [90], por ello esta métrica se emplea principalmente en aplicaciones de fusión de imágenes satelitales donde se realiza la fusión de imágenes multiespectrales con una imagen pancromática que mejor resolución espacial. Éste puede variar entre -1 a +1. Un valor cerca de +1 indica que las dos imágenes son muy similares, mientras que un valor cercano a -1 indica lo contrario. La fórmula para calcular la correlación entre dos imágenes A y B, ambas de MxN píxeles está dada por

$$Corr(A|B) = \frac{\sum_{i=1}^M \sum_{j=1}^N (A(i,j) - \mu_A)(B(i,j) - \mu_B)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N (A(i,j) - \mu_A)^2 \sum_{i=1}^M \sum_{j=1}^N (B(i,j) - \mu_B)^2}} \quad (5.14)$$

donde μ_A es la media de la imagen A y μ_B la media de la imagen B.

Idealmente, la correlación entre la imagen fusionada y la imagen de mejor resolución espacial debe ser la misma que la original con dicha imagen.

Una variante de esta métrica es el **coeficiente de correlación paso altas** (HPCC: high pass correlation coefficient), el cual mide la cantidad de información de bordes transferidos a la imagen fusionada. Una alta correlación entre dos imágenes filtradas con un filtro pasa altas implica que la información espacial ha sido retenida.

Para el caso de las imágenes médicas, podría emplearse el coeficiente de correlación paso altas en entre la imagen fusionada y aquella que tenga mejor calidad espacial, por ejemplo, en la fusión de una imagen anatómica con una funcional, el coeficiente de correlación paso altas puede emplearse para medir la cantidad de bordes de la imagen anatómica transferidos a la imagen fusionada.

5.5. Comparación de métricas de calidad objetivas

De las métricas de calidad objetivas, las más conocidas son las métricas basadas en el error como la PSNR y el MSE; a pesar de requerir una imagen de referencia, se emplean en la mayoría de los casos debido a su simplicidad y velocidad, la cual deriva del hecho de que son sencillas de calcular, operan en el dominio espacial y punto a punto de la imagen. Sin embargo, al no depender de las características del proceso que evalúa (características del monitor, distancia de visualización, etc.) ni del observador, trata todas las distorsiones de la misma forma dándoles la misma importancia, por lo que no predicen de forma correcta la calidad que realmente es apreciada por el observador. Cuando se utilizan sobre diversas escenas o sobre una misma pero con distintos tipos de distorsión, se pone de manifiesto el hecho de que para un mismo valor de PSNR o MSE, se observan calidades diferentes.

La información mutua se planteó como una alternativa para la evaluación de fusión de imágenes donde no se tiene la imagen ideal, pero a pesar de ser ampliamente empleada, el hecho de obtener valores altos no implica una mejor calidad de imagen, lo único que esto nos

indica es que se posee una mayor cantidad de información de las imágenes fuente en el resultado, lo cual no es sinónimo de una mayor calidad visual.

En cuanto a los métodos basados en la similitud estructural, su ventaja radica en que consideran al HVS, sin embargo métricas como el SSIM que es un método sencillo, portable y de poco costo computacional, no puede ser aplicable en todos los casos puesto que también requiere de una imagen de referencia. La alternativa de Piella en este sentido ofrece una ventaja al tomar como modelo la propuesta de Wang y Bovik y adaptar la métrica a más aplicaciones de fusión, sin embargo, es evidente que la complejidad de los métodos aumenta haciendo que su implementación no sea la mejor opción en fusión de datos tridimensionales.

Por otro lado, la imagen fusionada debe contener la información importante de todas las imágenes de entrada, donde la noción de ‘información importante’ depende de la aplicación por lo que es difícil de definir. No existe una medida precisa que nos sirva para determinar qué imagen es mejor que otra en este aspecto, por ejemplo, un HCPP de 0.94 puede ser suficiente para el resultado de fusión de ciertas imágenes, pero en otro proceso de fusión un valor más alto de HCPP no indica que la calidad espacial sea superior en ese caso, ya que este valor es distinto para imágenes diferentes. Por ejemplo, en aplicaciones como la fusión de imágenes médicas, la meta final es combinar perceptualmente elementos de la imagen notables como bordes y regiones de alto contraste. Las técnicas de evaluación de fusión en tal situación pueden ser efectivas solamente basadas en evaluaciones visuales [2, 24]. Además, el problema con las métricas de evaluación es su conexión con la interpretación visual de un observador humano, pues al analizar una imagen el observador no calcula ninguna medida.

En conclusión, las técnicas desarrolladas para evaluar la calidad de la imagen necesariamente dependen del campo de aplicación porque éste determina las características de la tarea de imagenología que se requiere evaluar. De modo que no hay un procedimiento estándar único para medir la calidad de imagen.

Dada la dificultad que representa la utilización de métricas objetivas, en este trabajo sólo se emplearán algunas de las métricas objetivas abordadas en este capítulo. Para los experimentos con volúmenes de imágenes medicas se compararon las métricas que no requieren para su cálculo de una imagen de referencia, tal es el caso de la simetría de fusión, la información mutua y los coeficientes de correlación paso altas del volumen fusionado con cada uno de los volúmenes fuente. Para algunos experimentos con imágenes sintéticas en 2D se utilizaron métricas que comparan si consideran una imagen de referencia como la PSNR, el MSE, el SSIM y la información mutua.

6.1. Introducción

La fusión de imágenes puede ser vista como el proceso de combinar información, obtenida de varias fuentes con la finalidad de construir una imagen artificial que contenga toda la información ‘útil’ que existe en las imágenes de entrada. Así, el principal concepto detrás de los algoritmos de fusión es *detectar las características sobresalientes* en las imágenes de entrada y *fusionar* esos detalles en una imagen sintética (imagen fusionada) [75].

De esta manera, si se tiene un conjunto de N imágenes capturadas de la misma escena donde cada imagen ha sido adquirida usando una técnica diferente de captura, cada imagen presentará características diferentes. Parte de esas diferencias puede aparecer en forma de imprecisión, ambigüedad o incompletitud. Usualmente, esto hace que existan algunas correspondencias erróneas entre varios puntos de la escena observada debido a los diferentes puntos de vista de los sensores de adquisición, para lo cual suele aplicarse un proceso de registro de imágenes. Sin embargo, en este estudio los datos de entrada se asumen libres de problemas de registro, por lo que los objetos presentes en ellos se consideran geométricamente alineados.

Por otro lado, las imágenes de dicho conjunto deben ofrecer información complementaria y redundante sobre la escena observada, y adicionalmente, cada una de esas imágenes debe contener información que puede ser útil para la composición de la imagen fusionada F y que no es provista por las demás. Esto último se lleva a cabo mediante el *proceso de fusión* empleando alguna técnica como las abordadas en el capítulo 3, las cuales pueden ser divididas en técnicas de dominio *espacial* y de *transformada* [51].

En las técnicas de fusión en el dominio espacial el proceso se efectúa con los valores originales de los píxeles. Asumiendo que $g(\bullet)$ representa la regla de fusión, estas técnicas pueden resumirse como sigue:

$$F(x, y) = g(I_1(x, y), \dots, I_N(x, y)) \quad (6.1)$$

La motivación principal detrás del proceso de *fusión en el dominio de una transformación* es el trabajar en un contexto donde las características sobresalientes de las imágenes estén descritas de forma más clara que en el dominio espacial. Para ello, es importante entender la estructura base de la imagen, más que la fusión de píxeles de manera independiente y hacer una buena elección de la transformación [74]. Sea $T\{\bullet\}$ un operador de transformada y $g(\bullet)$ la regla de fusión aplicada, las técnicas de fusión en el dominio de una transformada pueden ser expresadas de la siguiente forma:

$$F(x, y) = T^{-1} \left\{ g \left(T \{ I_1(x, y) \}, \dots, T \{ I_k(x, y) \} \right) \right\} \quad (6.2)$$

En la literatura se encuentran reportados varias técnicas de fusión de imágenes a nivel pixel empleando una transformada, algunas de ellas son el análisis multirresolución mediante la transformada wavelet [47, 119, 125], el análisis geométrico empleando la transformada contourlet [131], la transformada curvelet [65] y la transformada de Hermite [39, 41].

Todas estas transformaciones proyectan las imágenes de entrada en bases localizadas, modelando transiciones bien definidas (bordes) y por consiguiente, describiendo la imágenes empleando una representación más significativa que puede ser empleada para detectar y enfatizar características sobresalientes, las cuales son importantes para la ejecución de tareas en fusión de imágenes. En esencia todas esas transformadas pueden discriminar entre información notable (bordes bien delineadas y texturas) y constante o fondo no texturizado, y pueden también evaluar la calidad de la información sobresaliente provista. Consecuentemente, uno puede seleccionar la información requerida de las imágenes de entrada en el dominio transformado para construir la imagen fusionada [75].

De todos los métodos mencionados, la transformada wavelet ha sido la técnica más empleada para el proceso de fusión, sin embargo se ha demostrado que ésta presenta ciertas limitaciones en el análisis de señales de dos o más dimensiones, ejemplo de ello son los puntos de discontinuidad que no siempre puede detectar, otra desventaja es su limitación para la captura de información direccional. Estas desventajas ocasionan que haya características de las imágenes originales que no puedan ser preservadas e introducción de artefactos. Las transformadas curvelet (apéndice B) y contourlet (apéndice C) por su parte han mostrado mejores resultados que la transformada wavelet debido al análisis multidireccional, pero para ello es preciso hacer una descomposición exhaustiva de las orientaciones en cada nivel de descomposición.

Debido a lo anterior, la transformada de Hermite aporta notables ventajas al proceso de fusión de imágenes ya que este modelo de representación además de incluir algunas propiedades del sistema de visión humano permite un análisis multirresolución de manera que puede describir las estructuras sobresalientes de una imagen con una excelente aproximación en un número reducido de coeficientes al mismo tiempo que reduce el ruido presente sin introducción de artefactos.

6.2. Esquema de fusión basado en la transformada de Hermite

El algoritmo propuesto plantea la fusión de datos tridimensionales mediante el uso de la transformada de Hermite rotada multirresolución y la aplicación de distintas reglas de selección de coeficientes para bajas y altas frecuencias. La Figura. 6.1 muestra una representación simplificada del método. A grandes rasgos, incluye las siguientes etapas:

- i) Obtención de la HT multirresolución de los volúmenes de entrada, a fin de obtener descriptores para texturas y bordes que faciliten su identificación para su preservación durante el proceso de fusión.
- ii) Detección de la orientación de máxima energía para rotar los coeficientes de la HT, obteniéndose así proyecciones en una dimensión para cada nivel de descomposición. Para ello se calculan los ángulos θ y ϕ de acuerdo a la dirección del gradiente y se aplican las funciones de ángulos a los coeficientes a proyectar.
- iii) Selección de coeficientes de la descomposición multirresolución obtenida en el paso i) con la información de los coeficientes rotados del paso ii) para conformar la HT multirresolución del resultado de la fusión. La regla de fusión se conforma de dos métodos de selección dependiendo del tipo de frecuencia.
 - **Selección de coeficientes de baja frecuencia.** Los coeficientes de aproximación en el caso de la HT equivalen a los coeficientes de orden cero (L_{000}), los cuales contienen las características de las zonas homogéneas. Para formar el coeficiente de orden cero del volumen fusionado simplemente se promedian los valores de los coeficientes L_{000} de ambos volúmenes.
 - **Selección de coeficientes de alta frecuencia.** Se toma el coeficiente de primer orden rotado de cada nivel de resolución de los volúmenes de entrada para aplicar una regla de selección de coeficientes paso altas. Se considera sólo ese coeficiente puesto que la información contenida en él es suficiente para describir la información de los bordes del volumen en una determinada localidad espacial. El resultado de esta regla de selección genera un mapa de decisión que se aplicará a los coeficientes de la HT multirresolución del paso i) y con ello se genera el conjunto de coeficientes de detalle del volumen fusionado.
- iv) Obtención del volumen fusionado al aplicar la HT inversa multirresolución donde los coeficientes fusionados en cada nivel pueden reconstruir el conjunto de coeficientes de orden cero del nivel superior.

La elección de las reglas de selección de coeficientes depende del tipo de datos de entrada y las características deseadas en los datos fusionados, por ejemplo, en la elección del método de selección de coeficientes de baja frecuencia, el promedio no siempre resulta la mejor opción puesto que puede reducir notablemente el contraste del resultado, y en estos casos suele ser mejor quedarse con el coeficiente L_{000} de la imagen con mejor contraste. En el caso de los coeficientes de alta frecuencia, existen diversas reglas de selección, pero en la Fig. 6.1 se menciona la *dependencia lineal* ya que el esquema propuesto se aplica con esta regla para compararla con otras reglas tradicionales descritas en el capítulo 3.

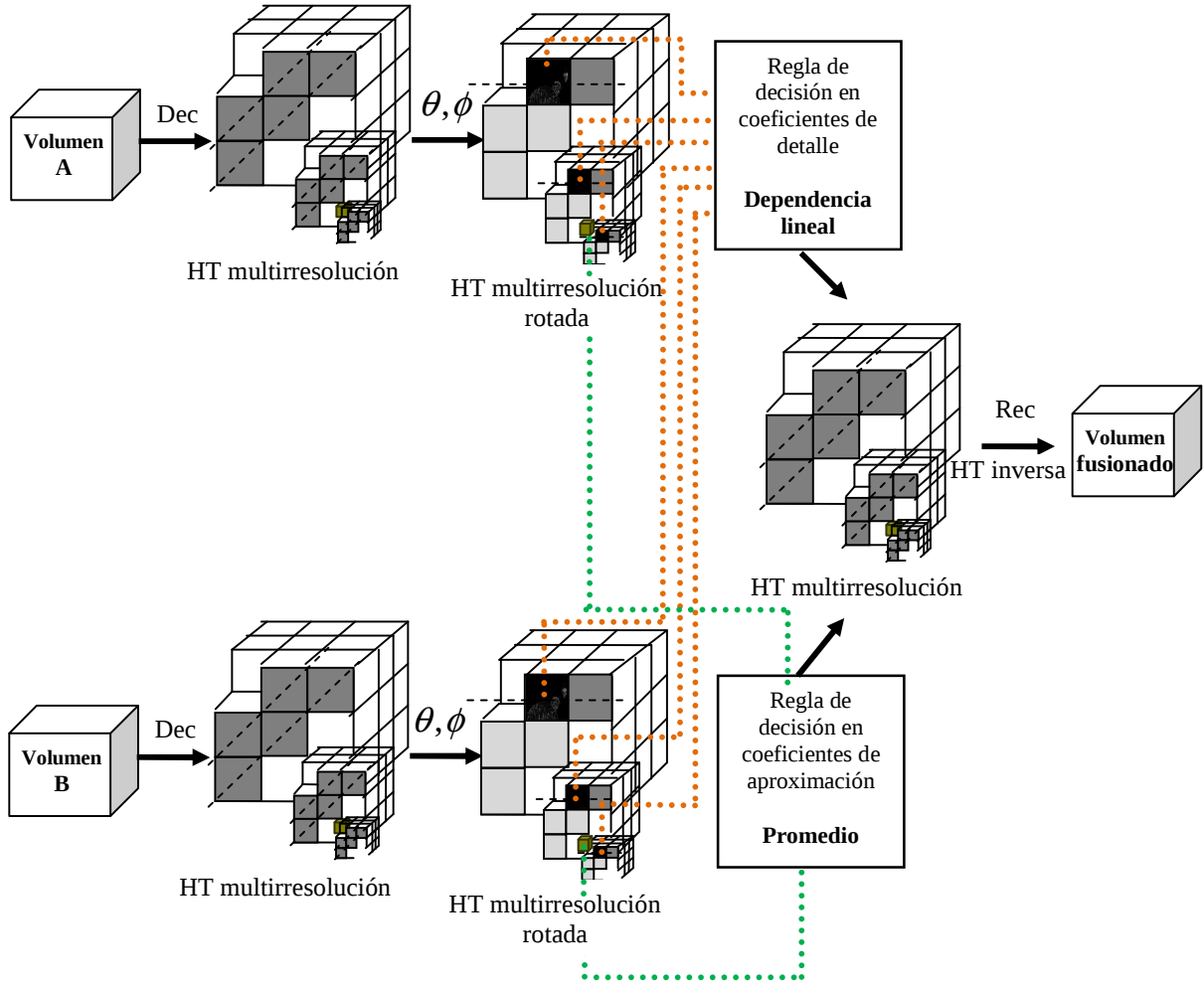


Figura 6.1 Esquema de fusión mediante la HT rotada multiresolución.

6.2.1. Regla de fusión propuesta

La regla de fusión se divide en selección de coeficientes de baja y alta frecuencia. Como ya se mencionó, el coeficiente L_{000} incluye la información general de la imagen tridimensional y el resto de los coeficientes contiene los detalles sobresalientes como bordes, líneas y separación de regiones.

6.2.1.1. Selección de coeficientes de frecuencias bajas

Dado que el coeficiente L_{000} contiene aproximadamente las mismas características que los datos fuente, se aplica alguna de las siguientes reglas de selección:

- Selección del coeficiente del volumen A

$$L_{000_F}(i, j, k) = L_{000_A}(i, j, k) \quad (6.3)$$

- Selección del coeficiente del volumen B

$$L_{000_F}(i, j, k) = L_{000_B}(i, j, k) \quad (6.4)$$

- Selección del promedio de los volúmenes A y B

$$L_{000_F}(i, j, k) = \frac{1}{2} [L_{000_A}(i, j, k) + L_{000_B}(i, j, k)] \quad (6.5)$$

Si se selecciona el coeficiente del volumen A o B, el volumen fusionado tendrá mayor cantidad de características de A o B respectivamente. Ahora si se selecciona el promedio de ambos, el resultado contendrá la información promediada de los coeficientes de ambas imágenes.

6.2.1.2. Selección de coeficientes de frecuencias altas

Los coeficientes de frecuencias altas contienen valores que reflejan las diferencias entre una vecindad de voxels y en consecuencia pueden ser positivos o negativos. En el caso de las imágenes en escala de grises, los valores absolutos de estos coeficientes representan la intensidad de las fluctuaciones de brillo de la escena a una escala dada, por lo que los valores más grandes implican más distinción entre los cambios de brillo que típicamente corresponden a las características sobresalientes de los objetos.

Siguiendo esta lógica, una regla de fusión simple es seleccionar los valores absolutos más grandes de los coeficientes correspondientes de cada volumen fuente, no obstante, hay dos desventajas de este método: la existencia de una alta sensibilidad al ruido y la producción de un efecto de emborronamiento [137]. Para eliminar estos efectos indeseables es menester la aplicación de una regla de fusión que aunque trabaje a nivel pixel considere una vecindad que ayude a detectar si ese pixel pertenece a alguna estructura de interés.

Así, con el propósito de establecer la existencia de un patrón sin realizar una extracción de características previa al proceso de fusión, se optó por una regla que considera como contexto una región pequeña alrededor del voxel de interés para evaluar la dependencia lineal local. Este criterio basado en el álgebra lineal ha sido reportado en la literatura en la detección de cambios en el análisis de imágenes [35, 52] y en esquemas de fusión de imágenes que emplean la transformada wavelet y la transformada curvelet [4, 65].

El fundamento de esta técnica se basa en un modelo empírico de vector de imágenes y la aplicación de los conceptos de dependencia e independencia lineal. Este modelo fue propuesto en [35], donde los conceptos mencionados se aplican para decidir si ha ocurrido un cambio o no entre dos imágenes. Dada una imagen, el modelo de vector se ilustra en la Figura 6.2. Cada pixel con su área de vecindad es la región de soporte de cada vector correspondiente. El pixel central de cada región de soporte es posteriormente reemplazado por el vector correspondiente. La región de soporte puede tener varios tamaños, por ejemplo 3x3, 5x5 o 7x7.

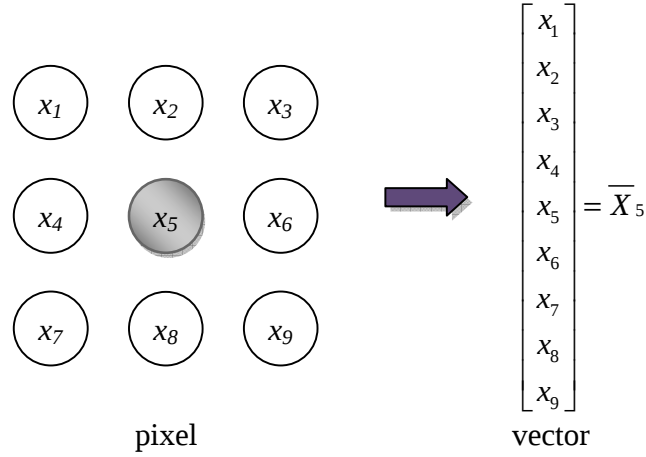


Figura 6.2 Construcción de los vectores a partir de una región de 3x3 en el modelo de vector

Una manera simple y rigurosa para determinar la dependencia lineal es el determinante de Wronskiano, el cual está definido para funciones pero puede aplicarse también a vectores [108].

Para el caso del esquema de fusión propuesto se considera una vecindad de 26 voxeles y en extensión a los esquemas encontrados en [4, 65], la prueba de dependencia lineal consiste en evaluar los voxeles dentro de un cubo pequeño de $w_s \times w_s \times w_s$, si los voxeles son linealmente independientes, no hay estructuras importantes en esa región; por el contrario, si son dependientes esto indica la existencia de un patrón que debe ser conservado en el volumen fusionado. Aplicando el determinante de Wronskiano, la dependencia de un cubo centrado en el voxel (i, j, k) se describe como

$$D_A(i, j, k) = \sum_{m=i-w_s}^{i+w_s} \sum_{n=j-w_s}^{j+w_s} \sum_{o=k-w_s}^{k+w_s} L_A^2(m, n, o) - L_A(m, n, o) \quad (6.7)$$

donde $L_A(m, n, o)$ es el coeficiente de primer orden rotado del volumen fuente A en la posición (m, n, o) . La regla es seleccionar los coeficientes de detalle con la dependencia más alta y se expresa como:

$$L_F(i, j, k) = \begin{cases} L_A(i, j, k) & \text{if } D_A(i, j, k) \geq D_B(i, j, k) \\ L_B(i, j, k) & \text{if } D_A(i, j, k) < D_B(i, j, k) \end{cases} \quad (6.8)$$

6.2.2. Reducción de ruido

En algunos casos no es posible partir de imágenes de gran calidad para la fusión, en ocasiones estas imágenes se encuentran contaminadas por ruido inherente a su proceso de adquisición, tal es el caso de las imágenes de ultrasonido por mencionar un ejemplo en el área médica. Aunque existen varios métodos de reducción de ruido que pueden aplicarse a los datos como parte de un pre-procesamiento antes de aplicar el esquema de fusión, la técnica de representación de las imágenes basada en la HT ha mostrado excelentes resultados en reducción de ruido y conservación de bordes [37, 41], razón por la cual se propone una modificación al esquema de fusión propuesto para realizar la fusión con *reducción de ruido* en un solo algoritmo. A continuación se describen el algoritmo con esta variante cuyas diferencias con el anterior se encuentran a partir del paso iii):

- i) Obtención de la HT multirresolución de los volúmenes de entrada.
- ii) Detección de la orientación de máxima energía (cálculo de θ y ϕ de acuerdo a la dirección del gradiente) para rotar los coeficientes de la HT, obteniéndose así proyecciones en una dimensión para cada nivel de descomposición.
- iii) Selección de coeficientes de la *descomposición multirresolución orientada* obtenida en el paso ii) con la información de esos coeficientes para conformar la HT multirresolución rotada del resultado de la fusión. Para ello se aplican dos métodos de selección, uno para la selección de coeficientes de baja frecuencia y otro para los de alta frecuencia.
- iv) Proyección de los coeficientes unidireccionales de primer y segundo orden de la HT multirresolución rotada obtenida en iii) para obtener los coeficientes de primer y segundo orden en el sistema cartesiano.
- v) Obtención del volumen fusionado al aplicar la HT inversa multirresolución donde los coeficientes fusionados en cada nivel obtenidos en iv) pueden reconstruir el conjunto de coeficientes de orden cero del nivel superior.

De esta manera, si los datos originales tienen ruido, el mapa de decisión de la regla de selección se aplica sobre la HT rotada multirresolución (coeficientes unidireccionales) de los volúmenes de entrada. Esto es posible debido a que al rotar los coeficientes de la HT, sólo se proyectan los coeficientes de primer y segundo orden, eliminando aquellos de órdenes superiores en los cuales se concentran las frecuencias más altas. Así, al antitransformar los datos, esas frecuencias quedan eliminadas quedando como resultado un volumen menos ruidoso que los originales. La Figura 6.3 muestra la estructura del algoritmo de fusión con reducción de ruido.

Cabe destacar que si bien existe una pérdida de información debida a los coeficientes que se discriminan, esta no afecta demasiado a las estructuras importantes ya que como se mencionó en el capítulo 4, las estructuras para el HVS más notables son las líneas y bordes, y estas se encuentran totalmente representadas en los coeficientes que resultan de filtrar los datos con derivadas de primer orden, mientras que la información de texturas también se haya presente en los coeficientes obtenidos de las derivadas de segundo orden.

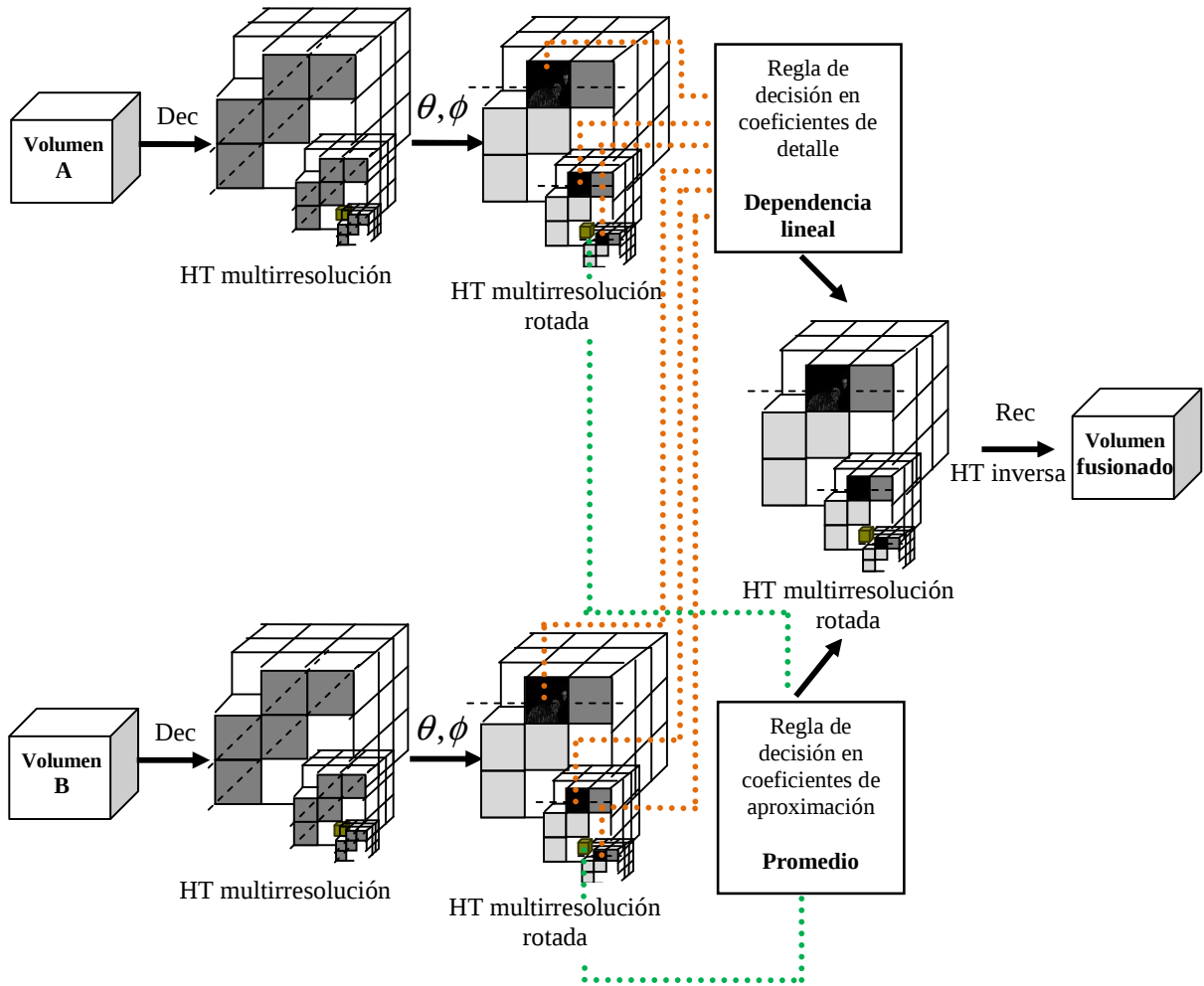


Figura 6.3 Esquema de fusión con reducción de ruido mediante la HT rotada multiresolución.

6.3. Programación del esquema de fusión

La implementación del esquema de fusión propuesto requirió la programación eficiente de la transformada de Hermite multiresolución directa e inversa, así como el mecanismo de orientación de los coeficientes dependiendo de la dirección del gradiente. Para ello se siguió el análisis de Hashimoto y Sklansky para la obtención de aproximaciones discretas de derivadas de Gaussiana [50].

Por otro lado, para la combinación de coeficientes, no sólo se diseñó el algoritmo con la regla de selección basada en la dependencia lineal, sino que se programaron otras reglas para evaluar la robustez del esquema respecto a distintos métodos de selección de coeficientes paso altas. Con este fin se incluyen las reglas de selección del máximo, selección del valor absoluto del máximo, selección del máximo por verificación de consistencia, saliente y medida de paridad [15].

Finalmente, se implementó también la fusión de datos tridimensionales siguiendo el esquema propuesto en el contexto de una transformación pero aplicando la transformada wavelet discreta; de manera que fuera posible comparar los resultados obtenidos para volúmenes en dos dominios distintos. En la Figura 6.4 se muestra el diagrama del flujo de la aplicación diseñada para la fusión de datos mientras que en la Figura 6.5 se muestra una captura de dicha aplicación programada en Matlab. Cabe mencionar que también se realizó una versión de la aplicación en dos dimensiones, en la cual se agregaron dos transformadas más, la transformada curvelet [109] y la transformada contourlet [30]; ambas fueron obtenidas de terceros y sólo se hará referencia a estos resultados, en el siguiente capítulo, como antecedente del presente trabajo por lo que no se describirá a detalle su programación.

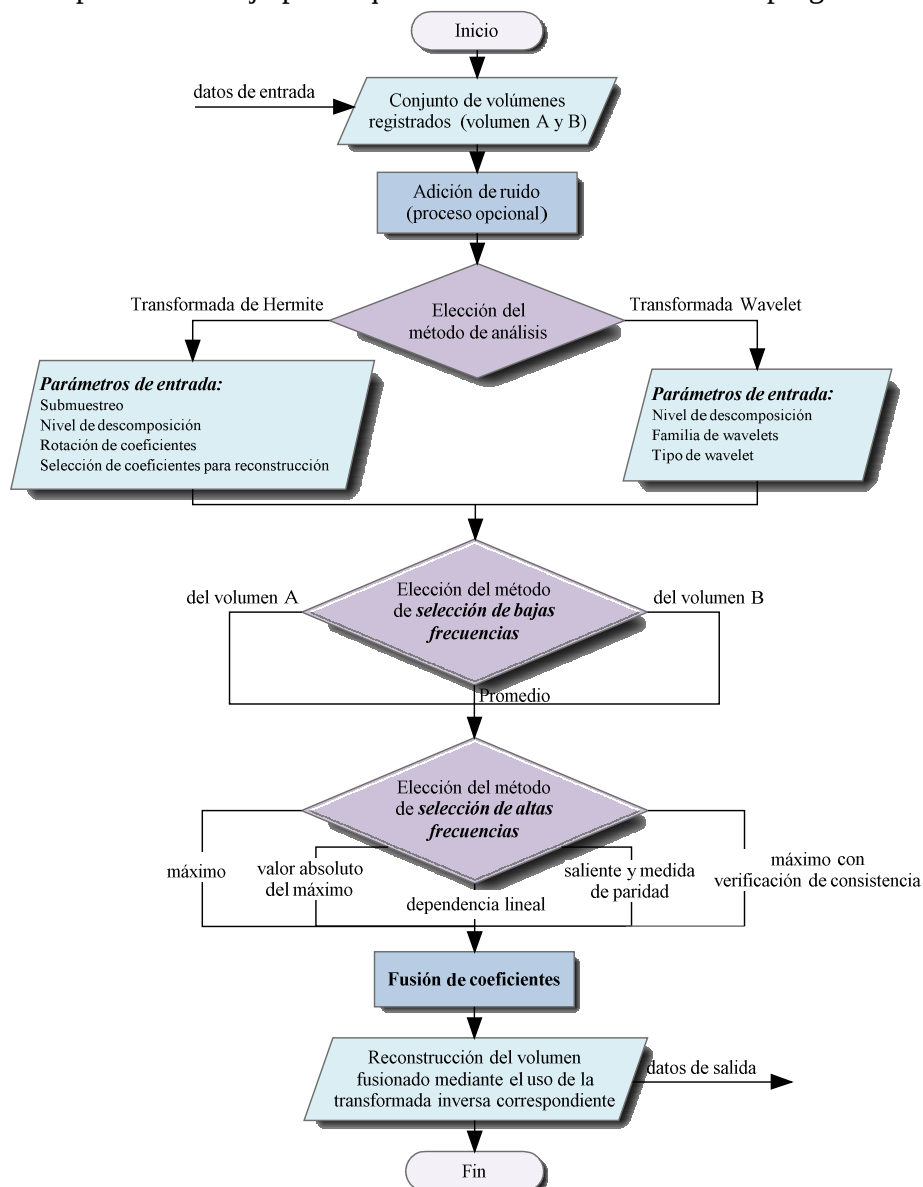


Figura 6.4 Diagrama del flujo de datos general.

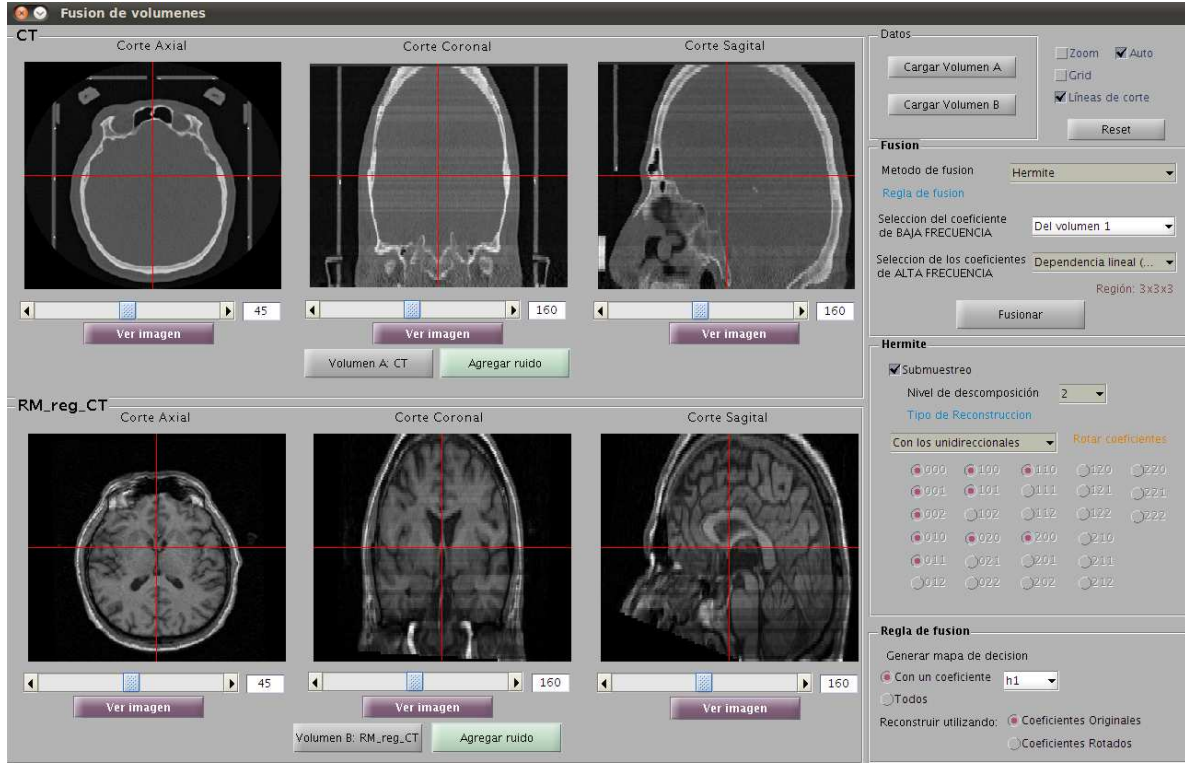


Figura 6.5 Interfaz gráfica de la aplicación de fusión de volúmenes.

6.3.1. Algoritmo rápido para la obtención de la DHT 3D

Como se mencionó en el capítulo 4, la transformada de Hermite utiliza filtros elementales similares a las derivadas de Gaussianas para el procesamiento local, cuya única diferencia radica en un factor de escala (ecuación 4.16); por lo que la eficiencia de la programación de este algoritmo se centra en la obtención de las dichas derivadas como aproximaciones discretas de acuerdo a lo expuesto en [50]. Para explicar el método se partirá del caso unidimensional y posteriormente se extenderá al caso tridimensional.

El procedimiento parte tomando como base un vector definido de la siguiente manera:

$$d_{N,k} = 2^{N-k} \begin{bmatrix} d_{N,k}(N) \\ d_{N,k}(N-1) \\ \vdots \\ d_{N,k}(0) \end{bmatrix} \quad \text{para } k = 0, 1, \dots, N \quad (6.9)$$

En el caso de la transformada en una dimensión, las estimaciones de las derivadas pueden expresarse como

$$\hat{f}^{(k)} = \frac{1}{2^{N-k}} d_{N,k}^T f_0 \quad \text{para } k = 0, 1, \dots, N \quad (6.10)$$

donde f_0 es un vector de datos en una dimensión, de tamaño $(N+1)$ y las derivadas son estimadas en el centro de la ventana. El factor de escala, 2^{N-k} , se elige de tal forma que los vectores tengan las mínimas expresiones enteras. Cada elemento del vector resultante representa un coeficiente de orden k de la transformada de Hermite en un punto dado.

Para determinar el vector $d_{N,k}$, se define una matriz $[P_N]$ de dimensiones $N+1$ por $N+1$ cuyo $k+1$ ésimo renglón es $d_{N,k}^T$. Es así entonces, que el conjunto de ecuaciones (6.10) se puede expresar en forma matricial:

$$\begin{aligned} \mathbf{F} &= (F_0, F_1, \dots, F_N)^T \\ \mathbf{F} &= [P_N] f_0 \end{aligned} \quad (6.11)$$

donde $F_k = 2^{N-k} \hat{f}^{(k)}$. Las matrices $[P_N]$ para $N = 1, 2, \dots, 4$ se muestran a continuación

$$\begin{array}{cccc} \begin{array}{c} \left| \begin{array}{cc} 1 & 1 \\ -1 & 1 \end{array} \right| \\ N = 1 \end{array} & \begin{array}{c} \left| \begin{array}{ccc} 1 & 2 & 1 \\ -1 & 0 & 1 \\ 1 & -2 & 1 \end{array} \right| \\ N = 2 \end{array} & \begin{array}{c} \left| \begin{array}{cccc} 1 & 3 & 3 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & 3 & -3 & 1 \end{array} \right| \\ N = 3 \end{array} & \begin{array}{c} \left| \begin{array}{ccccc} 1 & 4 & 6 & 4 & 1 \\ -1 & -2 & 0 & 2 & 1 \\ 1 & 0 & -2 & 0 & 1 \\ -1 & 2 & 0 & 2 & 1 \\ 1 & -4 & 6 & -4 & 1 \end{array} \right| \\ N = 4 \end{array} \end{array}$$

Procedimiento para generar la matriz $[P_N]$ a partir de $[P_{N-1}]$: El k -ésimo renglón ($k \leq N$) de $[P_N]$ se obtiene mediante la suma de los números adyacentes en el k -ésimo renglón de $[P_{N-1}]$ y para los extremos se consideran ceros, es decir, mediante la convolución de la secuencia $\{1, 1\}$ y la secuencia de números en el k -ésimo renglón de $[P_{N-1}]$. El último renglón de $[P_N]$ se obtiene al restar el número de la derecha del de la izquierda de los números adyacentes del último renglón de $[P_{N-1}]$; lo que equivale a convolucionar la secuencia $\{-1, 1\}$ y la secuencia de números en el último renglón de $[P_{N-1}]$.

En tres dimensiones, las derivadas parciales se estiman mediante el uso de operadores de una dimensión para cada dirección. Podemos escribir una matriz de derivadas parciales $[F]$ como se muestra en la Figura 6.6. Los elementos $F^{(k)}$ representan la derivada parcial estimada con factor de escala $2^{2N-k-l-m}$, donde los elementos de las derivadas parciales que tienen el mismo orden, cuentan también con el mismo factor de escala.

De manera general, la estimación de filtros de derivadas de orden múltiple puede implementarse mediante la aplicación de diversos algoritmos:

- realizando la convolución la imagen tridimensional con un conjunto de filtros de tres dimensiones
- mediante multiplicaciones de matrices en cada ventana de datos,
- convolucionando la imagen con un conjunto de máscaras de una dimensión en cada una de las direcciones de ésta, y
- mediante convoluciones repetidas por dos secuencias $\{1, 1\}$ y $\{1, -1\}$ en las tres direcciones.

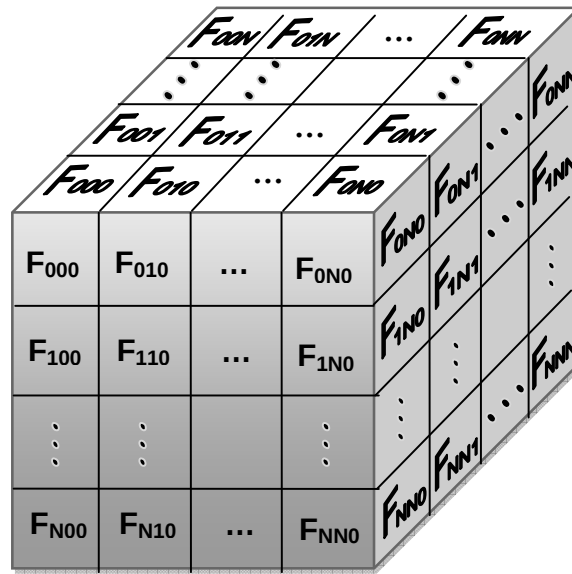


Figura 6.6 Matriz de derivadas parciales en tres dimensiones

De entre todos los algoritmos mencionados, la repetición de convoluciones por dos secuencias puede ser implementada con el menor número de operaciones, requiriendo únicamente de sumas y diferencias. Así, el algoritmo rápido para estimar derivadas de orden múltiple se basa en colocar en paralelo varias configuraciones de filtros de primer orden en cascada.

Dado que las funciones de transferencia de los filtros de estimación de derivadas contienen términos en común, las operaciones redundantes pueden ser eliminadas al compartir los resultados intermedios. Para valores muy grandes de N , es posible construir el algoritmo colocando filtros de primer orden sucesivamente en cascada, tal y como lo muestra la Figura 6.7, en la que se observa que los filtros son implementados con una suma y un retraso, y cada triángulo representa un algoritmo rápido para un tamaño en específico dependiendo si N es número par o impar.

Cuando $N+1$ es una potencia de 2, el número de sumas por muestra a analizar está dado por $(N+1)\log_2(N+1)$ debido a que, si observamos bien, la estructura completa se separa en dos subestructuras y cada una de estas a su vez se separa en otras dos y así sucesivamente hasta un número de veces regido por el factor $\log_2(N+1)$. En caso de que $N+1$ no sea una potencia de 2, el entero más cercano a $(N+1)\log_2(N+1)$ es aquél que nos dará el número exacto de sumas a realizar.

Una propiedad importante de los filtros de estimación de derivadas es su separabilidad, esto es, que las derivadas parciales pueden ser aproximadas al aplicar el algoritmo rápido para una sola dimensión de los datos de entrada y después a las demás del resultado.

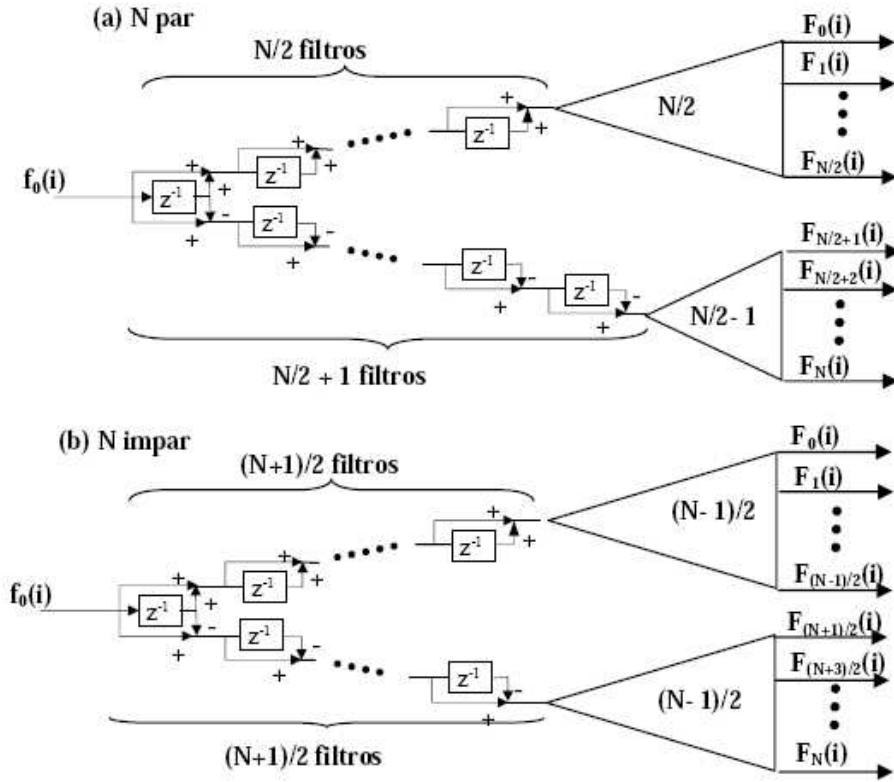


Figura 6.7 Algoritmo rápido para valores grandes de N

6.3.1.1. Programación del algoritmo

En la sección anterior se describió el procedimiento general para la implementación de derivadas de orden N , no obstante, de acuerdo a lo que se mencionó en la sección 4.3, para definir la señal de entrada mediante los coeficientes de la HT unos cuantos términos son suficientes debido a que entre mayor es el orden de dichos coeficientes el contenido que describen corresponde a frecuencias más altas.

Esto representa una gran ventaja desde el punto de vista computacional puesto que se puede obtener una muy buena aproximación de la imagen tridimensional calculando coeficientes de orden muy pequeño, minimizando así el número de operaciones requeridas al mismo tiempo que se reduce el tiempo de procesamiento. Por ello, en el algoritmo de descomposición mediante la HT multirresolución el orden máximo de los coeficientes considerado será $N=2$.

La estructura del algoritmo es la misma que se muestra en la Figura 6.7, sólo que en lugar de obtener las derivaciones hasta segundo orden en cada dirección se aplican en total 6 filtros en cascada que se traducen en operaciones de suma y resta, evitando así las convoluciones que tendrían que aplicarse con filtros tridimensionales de segundo orden (ver apéndice D). Los elementos que componen el algoritmo son:

- Una variable de entrada ‘VOXEL’ que contiene el valor del voxel en turno a procesar.
- Registros para almacenar los retrasos en cada una de las direcciones. Se requieren 18 registros para almacenar los retrasos de columnas (MC), 9 registros por cada columna para los retrasos de renglones (MR) y 36 registros por cada corte para los retrasos entre cortes (MT).
- Las variables auxiliares L1, ..., L62 que contienen los valores de salida de los filtros y que posteriormente son almacenados en los arreglos MR, MC y MT correspondientes.
- Los coeficientes polinomiales de salida, C000, C001, C002, C010, ..., C222.

El volumen se procesa a partir del voxel (1,1,1), ubicado en la esquina superior izquierda del volumen, siguiendo con los voxels de ese mismo renglón tomando columna a columna el voxel correspondiente; posteriormente este valor se almacena en la primera posición del arreglo MR y se tendrá ahora en la entrada al voxel que ocupa la posición (0,0,1) y así sucesivamente. La Figura 6.8 muestra sólo una parte de dicha estructura, la cual está completa en el apéndice F con una descripción más detallada sobre su nomenclatura.

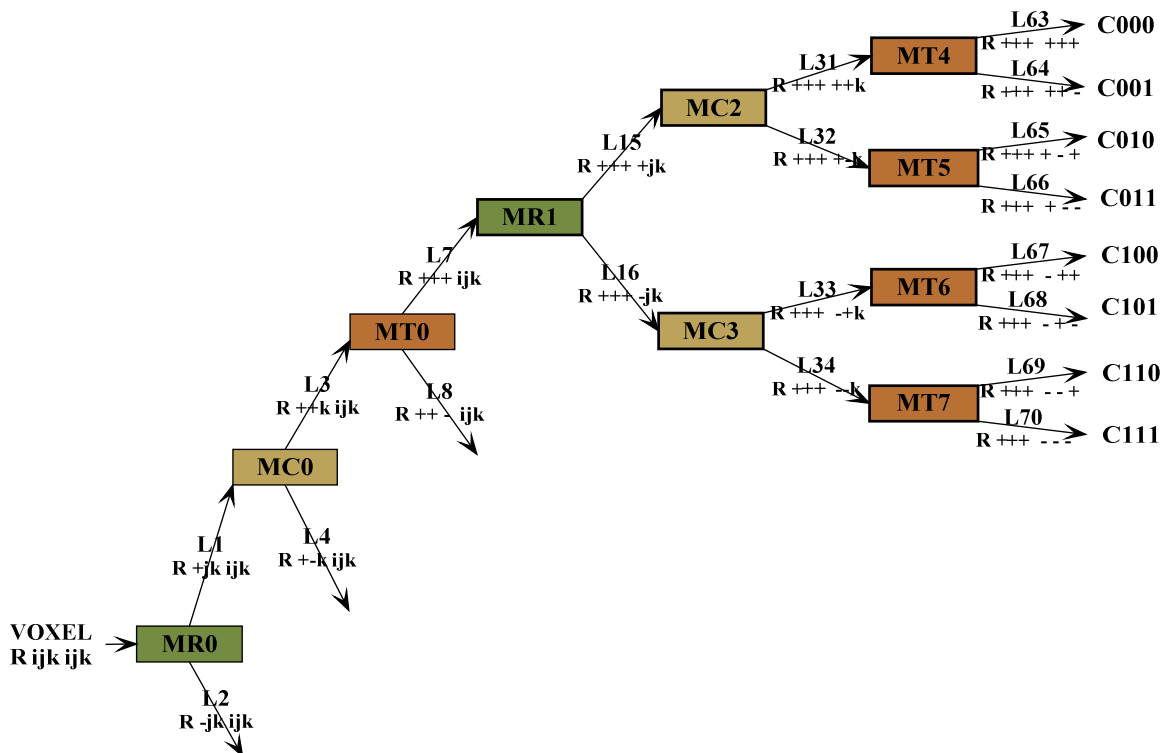


Figura 6.8 Estructura de registros para la implementación de la DHT 3D con N=2

A continuación se muestra el pseudocódigo del algoritmo para la HT directa de segundo orden en tres dimensiones:

```

dht3d(datos, nivel, sub)
Para l=1 hasta nivel {
  Si sub=1 entonces {
    Definir el tamaño de los coeficientes reduciendo a la mitad el
    número de elementos en cada dimensión.
  }
  Para t=1 hasta cortes {
    Para i=1 hasta renglones {
      Para j=1 hasta columnas {

        L(1) = datos(i, j, t) + MR(1, j)
        L(2) = datos(i, j, t) - MR(1, j)
        L(3) = L(1) + MC(1);
        L(4) = L(1) - MC(1);

        ...
        L(61) = L(30) + MC(18, i)
        L(62) = L(30) - MC(18, i)

        Asignar L() a los registros MR, MC y MT correspondientes

        Si sub=1 entonces {
          Tomar solo las posiciones pares de i, j, t
          C000(i, j, t)=L(31) + MT(5, i, j)
          C001(i, j, t)=L(31) - MT(5, i, j)
          ...
          C222(i, j, t)=L(62) - MT(36, i, j)
        }
        Si no{
          C000(i, j, t)=L(31) + MT(5, i, j)
          C001(i, j, t)=L(31) - MT(5, i, j)
          ...
          C222(i, j, t)=L(62) - MT(36, i, j)
        }
      }
    }
  }

  l=l+1
}

```

La función `dht3d` recibe como argumentos de entrada los datos a procesar (`datos`), el número de niveles de descomposición multirresolución (`nivel`) y una bandeara para indicar si se desea aplicar submuestreo de $2 \downarrow$ en cada nivel. El número de iteraciones es de $\text{nivel} \times N$, donde N es la longitud de los datos, es decir, $N = \text{filas} \times \text{columnas} \times \text{cortes}$; en cada iteración se realizan 70 operaciones que comprenden sumas y restas. Como el número de niveles es un número muy pequeño comparado con la longitud de los datos, la complejidad de este algoritmo es de $O(N)$.

6.3.2. Algoritmo para la rotación de coeficientes de la DHT

La rotación de coeficientes se lleva a cabo en dos pasos: i) primero se proyectan los coeficientes de la HT a una dirección, con lo cual se obtienen los ángulos que indican la dirección de máxima energía y dos coeficientes unidireccionales, uno que contiene la información de los coeficientes de primer orden {C001, C010, C100}, y otro con la información de coeficientes de segundo orden {C011, C110, C101, C002, C020, C200}; ii) posteriormente se proyectan nuevamente estos coeficientes unidireccionales para regresar a los coeficientes primer y segundo orden en tres dimensiones.

Entre los pasos i) y ii) descritos se toma la decisión sobre los coeficientes que conformarán el volumen fusionado por lo que la proyección de coeficientes a una dimensión se realiza sobre la HT multiresolución directa y el proceso inverso de pasar de coeficientes unidireccionales a los coeficientes originales se realiza antes de aplicar la antitransformada de los coeficientes seleccionados por la regla de fusión.

6.3.2.1. Análisis de orientación local de una señal 3D

Para la detección de la dirección de máxima energía, una vez calculados los coeficientes de la HT se aplica una función (steering) a los coeficientes, la cual depende de la dirección que se elija. Esta dirección se especifica para cada posición estimando los ángulos que maximizan dicha energía; para el caso tridimensional y tomando coordenadas esféricas, la energía direccional AC es

$$E_{AC}(\theta, \phi) = \sum_{n=1}^{\infty} \{h_n(\theta, \phi)\}^2 \quad (6.12)$$

donde los coeficientes unidireccionales se expresan como:

$$h_n(\theta, \phi) = \sum_{j=0}^n \sum_{l=0}^m \alpha_{l,m-l,n-m}(\theta, \phi) L_{l,m-l,n-m} \quad (6.13)$$

donde $L_{l,m-l,n-m}$ son los coeficientes de la transformada y $\alpha_{l,m-l,n-m}$ son las funciones direccionales que se construyen a partir de los vectores en una esfera unitaria (ver Figura 6.9).

$$\alpha_{l,m,n}(\theta, \phi) = \sqrt{\frac{(l+m+n)!}{l!m!n!}} [\cos(\theta) \sin(\phi)]^l [\sin(\theta) \sin(\phi)]^m [\cos(\phi)]^n \quad (6.14)$$

Por otro lado, si la porción de la imagen $l(x,y,z)$ dentro de la ventana de análisis es efectivamente generada por el desplazamiento del patrón 1D que contiene suficiente información local entonces es posible determinar en forma única los ángulos que maximizan la energía direccional. Bajo estas condiciones el error de aproximación es nulo para la dirección óptima y los coeficientes de la imagen original pueden ser recuperados a partir de los coeficientes rotados unidireccionales mediante:

$$\widehat{L_{l,m-l,n-m}} = \sum_{n=0}^{\infty} h_n(\theta, \phi) \alpha_{l,m-l,n-m}(\theta, \phi) \quad (6.15)$$

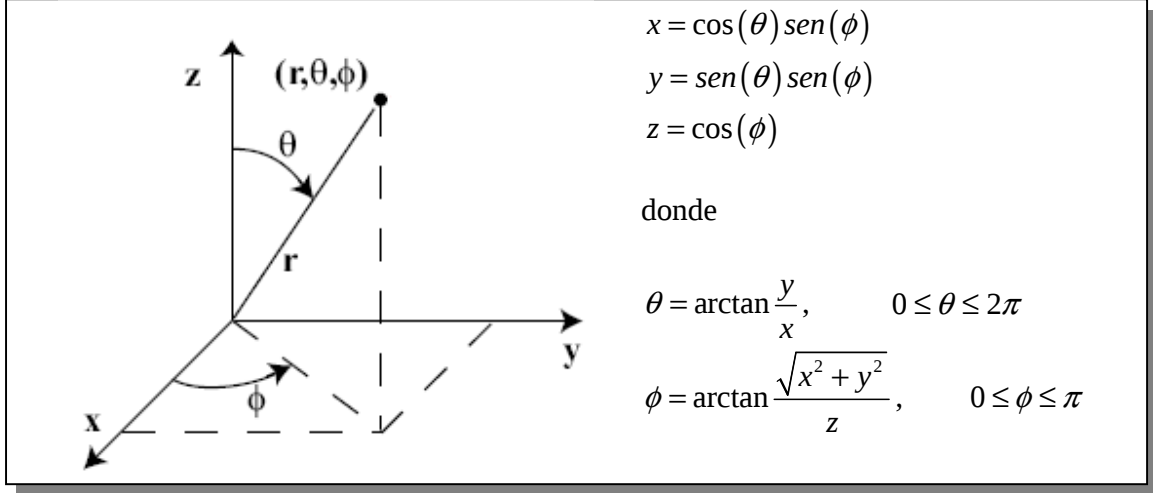


Figura 6.9 Coordenadas esféricas

6.3.2.1. Algoritmo de proyección de los coeficientes en 3D a 1D

Como se mencionó en la sección 6.3.1.1, la expansión polinomial en la representación de datos propuesta se limita a coeficientes de segundo orden, es decir, $L_{i,j,k}$ donde $i + j + k \leq 2$. Así, para el cálculo de las energías direccionales, se tendrá:

$$E(\theta, \phi) = h_1^2(\theta, \phi) + h_2^2(\theta, \phi) \quad (6.16)$$

Desarrollando las expresiones explícitas para h_n , con $n=1,2$:

$$\begin{aligned} h_1(\theta, \phi) &= \alpha_{0,0,1}(\theta, \phi) L_{0,0,1} + \alpha_{0,1,0}(\theta, \phi) L_{0,1,0} + \alpha_{1,0,0}(\theta, \phi) L_{1,0,0} \\ &= \cos(\phi) L_{0,0,1} + \operatorname{sen}(\theta) \operatorname{sen}(\phi) L_{0,1,0} + \cos(\theta) \operatorname{sen}(\phi) L_{1,0,0} \end{aligned}$$

$$\begin{aligned} h_2(\theta, \phi) &= \alpha_{0,0,2}(\theta, \phi) L_{0,0,2} + \alpha_{0,2,0}(\theta, \phi) L_{0,2,0} + \alpha_{2,0,0}(\theta, \phi) L_{2,0,0} + \\ &\quad \alpha_{0,1,1}(\theta, \phi) L_{0,1,1} + \alpha_{1,0,1}(\theta, \phi) L_{1,0,1} + \alpha_{1,1,0}(\theta, \phi) L_{1,1,0} \\ &= \cos^2(\phi) L_{0,0,2} + \operatorname{sen}^2(\theta) \operatorname{sen}^2(\phi) L_{0,2,0} + \cos^2(\theta) \operatorname{sen}^2(\phi) L_{2,0,0} + \\ &\quad \sqrt{2} \operatorname{sen}(\theta) \operatorname{sen}(\phi) \cos(\phi) L_{0,1,1} + \sqrt{2} \cos(\theta) \operatorname{sen}(\phi) \cos(\phi) L_{1,0,1} + \sqrt{2} \cos(\theta) \operatorname{sen}(\theta) \operatorname{sen}^2(\phi) L_{1,1,0} \end{aligned}$$

donde los ángulos θ y ϕ se obtienen de acuerdo a la dirección del gradiente considerando los coeficientes de primer orden en cada dirección, de manera que partiendo de la Figura 6.10 se obtienen las aproximaciones para ambos ángulos:

$$\phi = \arctan \frac{\sqrt{(L_{1,0,0})^2 + (L_{0,1,0})^2}}{L_{0,0,1}} \quad \text{y} \quad \theta = \arctan \frac{L_{0,1,0}}{L_{1,0,0}}$$

A continuación se muestra el pseudocódigo de la función que proyecta los coeficientes en 3D a 1D:

Proy_3D_1D(C)

```

Para t=1 hasta cortes {
  Para i=1 hasta renglones {
    Para j=1 hasta columnas {
      theta(i,j,t)=arctan [ C010(i,j,t) / C100(i,j,t) ]
      phi(i,j,t)=arctan{sqrt[C100(i,j,t)^2+C001(i,j,t)^2] / C001(i,j,t)}
      Calcular las funciones de ángulo α con theta(i,j,t) y phi(i,j,t)
      Obtener h1(i,j,t)
      Obtener h2(i,j,t)
    }
  }
}
```

6.3.2.2. Algoritmo de proyección de los coeficientes en 1D a 3D

El procedimiento inverso al descrito en la sección anterior se basa en la ecuación 6.15, donde cada coeficiente 3D, $L_{i,j,k}$, se recupera al multiplicar el coeficiente unidireccional de orden $i+j+k$ por la función de ángulo correspondiente. A continuación se muestra el pseudocódigo de la función que proyecta los coeficientes en 1D a 3D:

Proy_3D_1D(h, theta, phi)

```

Para t=1 hasta cortes {
  Para i=1 hasta renglones {
    Para j=1 hasta columnas {
      Calcular las funciones de ángulo α con theta(i,j,t) y phi(i,j,t)
      C001(i,j,t)=h1(i,j,t)*α001(i,j,t)
      C010(i,j,t)=h1(i,j,t)*α010(i,j,t)
      C100(i,j,t)=h1(i,j,t)*α100(i,j,t)
      C002(i,j,t)=h2(i,j,t)*α002(i,j,t)
      ...
      C110(i,j,t)=h2(i,j,t)*α110(i,j,t)
    }
  }
}
```

6.3.3. Algoritmo rápido para la obtención de la IDHT 3D

De acuerdo a la teoría vista en el capítulo 4, los filtros de antitransformación necesarios son similares a los filtros de transformación (ver apéndices D y E), sin embargo, este proceso es inverso al anterior ya que en él se hace una síntesis de las matrices de coeficientes de la transformada polinomial para generar una sola matriz que serán los datos recuperados. Al

igual que en el caso de la DHT, para la IDHT se consideraron dos casos, cuando la transformación se realiza sin submuestreo y cuando existe una decimación de $\downarrow 2$

Si *no hay submuestreo*, la reconstrucción se obtiene de sumar los coeficientes obtenidos multiplicados por un factor que indica el número de veces que ese coeficiente aparece en el árbol. Así, para plantear el algoritmo de antitransformación rápida, se tomó el árbol de sumas y diferencias utilizado para la teoría de transformación, pero en este caso, el árbol fue recorrido a la inversa, es decir, de derecha a izquierda, partiendo de que tenemos 27 matrices de coeficientes de la transformada polinomial y queremos llegar al voxel de la imagen recuperada al que corresponde cierta localidad espacial. En la Figura 6.10 se muestra parte del árbol empleado para la obtención de la DHT donde se resalta el hecho de la repetición de coeficientes. Si se observa el árbol completo es posible obtener todos los coeficientes con su factor y notar que estos factores corresponden a los valores del filtro F000 aplicados a los coeficientes de orden 2 con la distribución mostrada en la Figura 4.6.

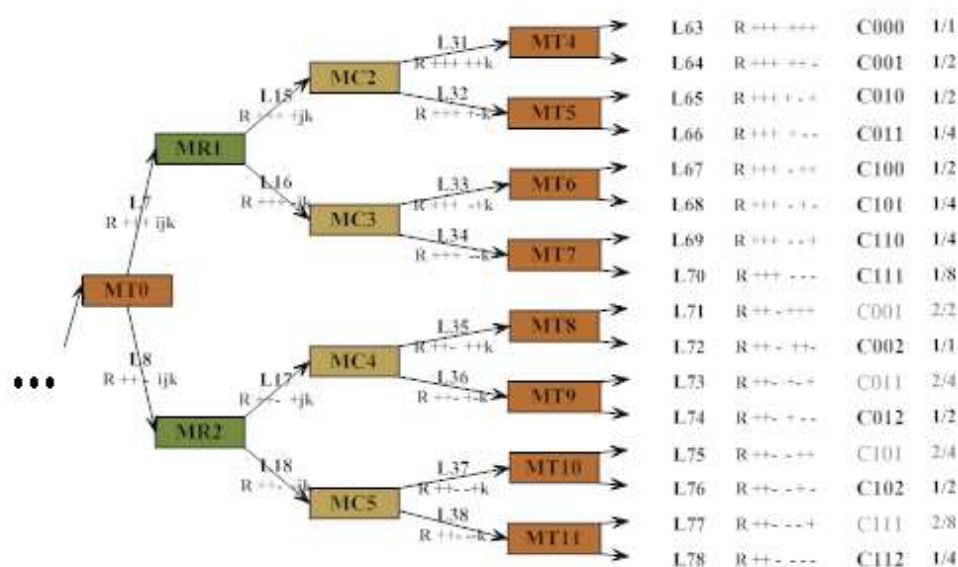


Figura 6.10 Obtención de las ponderaciones para la IDHT 3D con N=2 a partir del árbol de derivación.

Si hay submuestreo, la reconstrucción se vuelve más compleja ya que no es posible utilizar el árbol de manera directa ni los filtros de síntesis del apéndice E, pues se deben considerar los ceros de la ‘interpolación’. Para ello se debe suponer que entre cada voxel de los coeficientes de la transformada existen ceros que rellenan la matriz, los cuales no están realmente almacenados en ésta por ser información redundante. En la Figura 6.11 se muestra un ejemplo de la disposición real de los datos con lo que en realidad representan.

Debido a esta compactación de datos, fue necesario implementar una solución que optimizara el tiempo de procesamiento para no tener que construir la matriz real con los ceros intercalados y así evitar cálculos innecesarios. La solución consistió identificar los casos posibles que se presentan al aplicar el filtro tridimensional en los datos de la matriz no submuestreada, considerando para ello que hay ceros entre cada localidad. Haciendo estas suposiciones, se detectaron ocho casos en los cuales las posiciones de los ceros “imaginarios”

entre los datos tomaban una disposición similar. En la Figura 6.12 se muestran dichos casos, los primeros cuatro corresponden al barrido de columnas y renglones de los cortes impares mientras que los casos restantes se refieren al barrido de los cortes pares (partiendo de la idea de que el primer corte es impar ya que empiezan en 1).

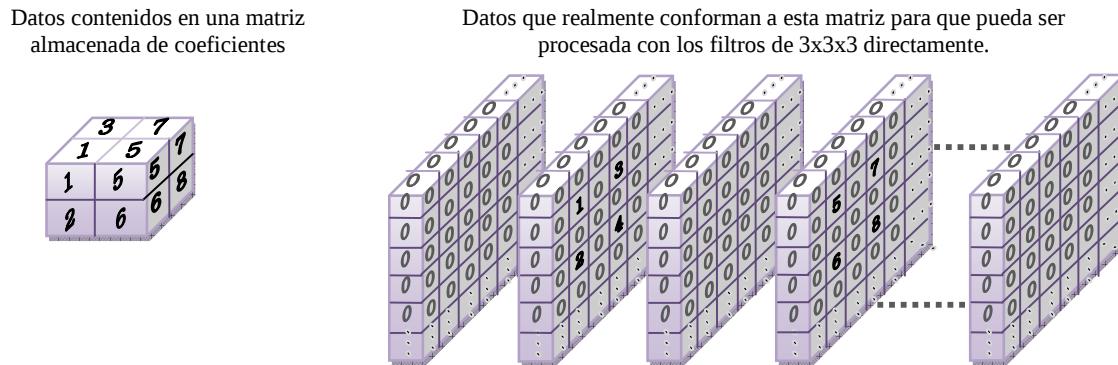


Figura 6.11 Disposición de los datos obtenidos en el programa contra su significado real.

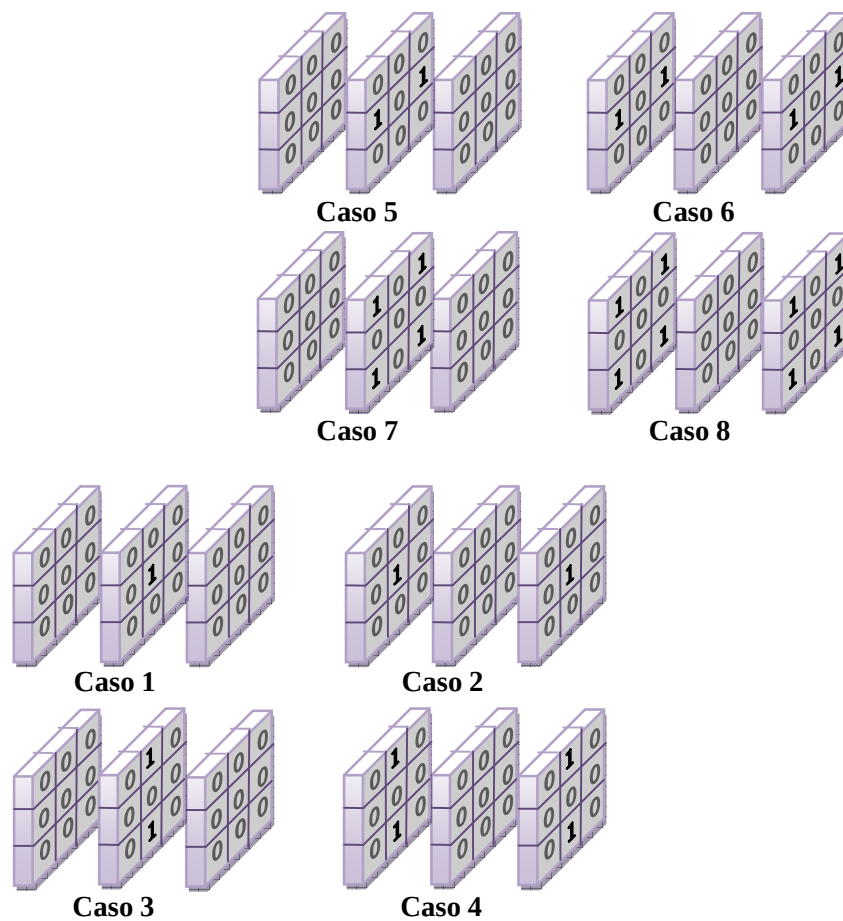


Figura 6.12 Cuatro disposiciones posibles de los ceros que existen entre los datos de las matrices de coeficientes.

Como se puede observar, todos los casos envuelven muchos ceros dentro de la región de filtrado para un voxel en particular, por lo que si se considera un filtro genérico de 3x3x3 de la forma mostrada en la Figura 6.13 y que el barrido se hace sobre la columna 'j' y el Renglón 'i' y el corte 'k', tomando el centro del cubo como el lugar al que apuntan los valores (i,j,k), es posible plantear las siguientes funciones para generar el filtrado para cada matriz de coeficientes, combinando al final estos valores para obtener la matriz recuperada.

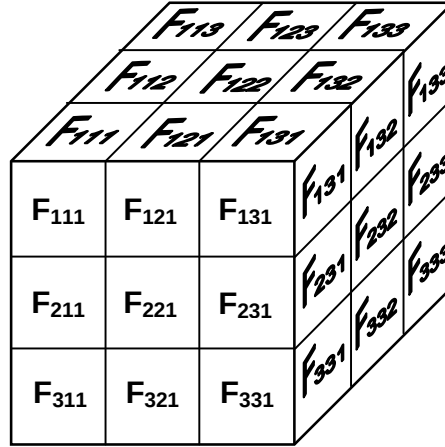


Figura 6.13 Representación de un filtro genérico de 3x3x3.

- Caso 1: $F_{222} \cdot V_{(i,j,k)}$
- Caso 2: $F_{212} \cdot V_{(i,j-1,k)} + F_{232} \cdot V_{(i,j+1,k)}$
- Caso 3: $F_{122} \cdot V_{(i-1,j,k)} + F_{322} \cdot V_{(i+1,j,k)}$
- Caso 4: $F_{112} \cdot V_{(i-1,j-1,k)} + F_{312} \cdot V_{(i+1,j-1,k)} + F_{132} \cdot V_{(i-1,j+1,k)} + F_{332} \cdot V_{(i+1,j+1,k)}$
- Caso 5: $F_{221} \cdot V_{(i,j,k-1)} + F_{223} \cdot V_{(i,j,k+1)}$
- Caso 6: $F_{211} \cdot V_{(i,j-1,k-1)} + F_{213} \cdot V_{(i,j-1,k+1)} + F_{231} \cdot V_{(i,j+1,k-1)} + F_{233} \cdot V_{(i,j+1,k+1)}$
- Caso 7: $F_{121} \cdot V_{(i-1,j,k-1)} + F_{123} \cdot V_{(i-1,j,k+1)} + F_{321} \cdot V_{(i+1,j,k-1)} + F_{323} \cdot V_{(i+1,j,k+1)}$
- Caso 8: $F_{111} \cdot V_{(i-1,j-1,k-1)} + F_{113} \cdot V_{(i-1,j-1,k+1)} + F_{311} \cdot V_{(i+1,j-1,k-1)} + F_{313} \cdot V_{(i+1,j-1,k+1)} +$
 $F_{131} \cdot V_{(i-1,j+1,k-1)} + F_{133} \cdot V_{(i-1,j+1,k+1)} + F_{331} \cdot V_{(i+1,j+1,k-1)} + F_{333} \cdot V_{(i+1,j+1,k+1)}$

De esta manera, los valores de F_{MNO} tomarán los valores adecuados dependiendo del filtro de síntesis que se requiera utilizar, por ejemplo, se está procesando la matriz de coeficientes de orden 000, los valores del filtro F_{MNO} deberán coincidir con los valores para este filtro de antitransformación. Cabe mencionar que en este proceso aún se pueden eliminar algunas sumas, puesto que en ciertos casos F_{MNO} para ciertos valores de 'M', 'N' y 'O', y para determinado orden, se tienen valores nulos que simplifican aún más las sumas de los casos anteriores. Adicionalmente falta considerar un factor de normalización dependiendo del orden del filtro de síntesis, este factor está relacionado con una potencia de 2 y puede ser

considerado al momento de operar con las matrices de coeficientes multiplicándolas por su factor antes de analizar los casos.

Siguiendo este razonamiento, se observa que nuevamente se requieren localidades de memoria que almacenen los valores contenidos en las columnas, renglones y cortes anteriores al voxel procesado. Estos registros son los siguientes:

- MA(18, # columnas) – Almacenan los valores del renglón anterior de las matrices de coeficientes, sólo son 18 porque no todos los filtros de síntesis tienen valores diferentes de cero en la columna central, por ello no se consideran los siguientes filtros: F010, F012, F210, F212, F101, F121, F011, F211 y F111 (ver apéndice E). Para considerar el factor de normalización los valores se multiplican por una constante, para los filtros F000, F002, F020, F200, F220, F202, F022 y F222 el factor es 2, para F001, F021, F201, F221, F100, F102, F120 y F122 el factor es 4 y para F101 y F121 el factor es 8.
- MS(27, # columnas) – Almacenan las sumas y/o diferencias de valores adyacentes del renglón anterior de las matrices de coeficientes. Se considera la suma cuando los signos de los coeficientes de los filtros de síntesis entre las columnas $j-1$ y $j+1$ son iguales, en caso contrario entonces se guarda la diferencia de los valores contiguos de las matrices de coeficientes. Por ejemplo, para la matriz de coeficientes de orden 001 se guardan las sumas de los valores contiguos ya que su filtro de síntesis tiene el mismo signo entre las posiciones 111 y 131, 113 y 133, 311 y 331, 313 y 333. En cambio, la matriz de coeficientes de orden 010 se guardan las diferencias porque en las posiciones mencionadas los signos son contrarios (ver apéndice E). Si los signos son iguales pero negativos, o bien, la diferencia está invertida, por ejemplo -111 y +131, entonces al aplicar alguno de los 8 casos planteados se cambia el signo de ese término en la sumatoria. Para considerar el factor de normalización, algunas de estas sumas y/o diferencias se afectan por una potencia de 2, para los filtros F001, F010, F012, F021, F201, F221, F100, F102, F120, F122, F210 y F212 el factor es 2, para F011, F101, F110, F112, F121 y F211 el factor es 4 y para F111 el factor es 8.
- MZA(18, # renglones, # columnas) – Almacenan los valores del renglón anterior de las matrices de coeficientes que tienen las mismas características descritas para MA.
- MZS(27, # renglones, # columnas) – Almacenan todas las sumas y/o diferencias del corte anterior con las mismas consideraciones que MS.
- Mtmp(27) – Arreglo temporal que almacena ya sea los valores de MA o MS de un renglón anterior para procesar los casos 3, 4, 7 y 8.

A continuación se muestra el pseudocódigo del algoritmo para la IDHT de segundo orden en tres dimensiones considerando sólo un nivel de resolución.

idht3d(C, sub)

```

Si sub=1 entonces {
  p=1, q=1, r=1 → PRIMER CORTE, CORTE IMPAR
  Para j=1 hasta columnas {
    Guardar MA, MS, MZA, MZS
    imagen(p,q,r)=  $\Sigma MA/2^7$  → Caso 1
    q=q+1
    imagen(p,q,r)=  $\Sigma MS/2^7$  → Caso 2
    q=q+1
  }
  p=p+1
  Para i=1 hasta renglones {
    q=1
    Para j=1 hasta columnas {
      Guardar Mtmp como MA
      imagen(p,q,r)=  $(\Sigma MA \Sigma Mtmp)/2^8$  → Caso 3
      q=q+1
      Guardar MA, MZA con los valores de Mtmp
      Guardar Mtmp como MS
      imagen(p,q,r)=  $(\Sigma MS \Sigma Mtmp)/2^8$  → Caso 4
      q=q+1
      Guardar MS, MZS con los valores de Mtmp
    }
    p=p+1, q=1
    Para j=1 hasta columnas {
      imagen(p,q,r)=  $\Sigma MA/2^7$  → Caso 1
      q=q+1
      imagen(p,q,r)=  $\Sigma MS/2^7$  → Caso 2
      q=q+1
    }
    p=p+1
  }
}

r=r+1 → LOS DEMÁS CORTES
Para t=1 hasta cortes {
  p=1, q=1 → CORTE PAR
  Para j=1 hasta columnas {
    Guardar MA, MS
    imagen(p,q,r)=  $(\Sigma MA \Sigma MZA)/2^7$  → Caso 5
    q=q+1
    imagen(p,q,r)=  $(\Sigma MS \Sigma MZS)/2^7$  → Caso 6
    q=q+1
  }
  p=p+1
  Guardar MZA, MZS, MZAtmp, MZStmp con MA y MS

  Para i=1 hasta renglones {
    q=1
    Guardar MATmp y MStmp con MZA y MZS
    Para j=1 hasta columnas {
      Guardar Mtmp como MA
      imagen(p,q,r)=  $(\Sigma MZA \Sigma MZAtmp \Sigma MA \Sigma Mtmp)/2^9$  → Caso 7
      q=q+1
      Guardar MA, MZA con los valores de Mtmp
    }
  }
}

```

```

    Guardar MZAtmp con los valores de MZA

    Guardar Mtmp como MS
    imagen(p,q,r)= (ΣMZS ΣMZStmp ΣMS ΣMtmp)/2^9 → Caso 8
    q=q+1
    Guardar MS, MZS con los valores de Mtmp
    Guardar MZStmp con los valores de MZS
  }
  p=p+1, q=1
  Para j=1 hasta columnas {
    Guardar MA, MS
    imagen(p,q,r)= (ΣMAtmp ΣMA)/2^7 → Caso 5
    q=q+1
    imagen(p,q,r)= (ΣMStmp ΣMS)/2^7 → Caso 6
    q=q+1
  }
  p=p+1
}
r=r+1

p=1, q=1 → CORTE IMPAR
Para j=1 hasta columnas {
  imagen(p,q,r)= ΣMZA/2^7 → Caso 1
  q=q+1
  imagen(p,q,r)= ΣMZS/2^7 → Caso 2
  q=q+1
}
p=p+1
Para i=1 hasta renglones {
  q=1
  Para j=1 hasta columnas {
    imagen(p,q,r)= (ΣMZA ΣMZAi+1)/2^8 → Caso 3
    q=q+1
    imagen(p,q,r)= (ΣMS ΣMZSi+1)/2^8 → Caso 4
    q=q+1
  }
  p=p+1, q=1
  Para j=1 hasta columnas {
    imagen(p,q,r)= ΣMZAi+1/2^7 → Caso 1
    q=q+1
    imagen(p,q,r)= ΣMZSi+1/2^7 → Caso 2
    q=q+1
  }
  p=p+1
}
r=r+1
}

}si no{
  Imagen = Σ(2*C001, 2*C010, 4*C011, 2*C012, 2*C100, 4*C101, 2*C102,...
    2*C021, 4*C110, 8*C111, 4*C112, 2*C201, 2*C120, 4*C121,...
    2*C122, 2*C210, 4*C211, 2*C212, 2*C221)/2^6;
}

```

Aunque en el pseudocódigo sólo se muestra el algoritmo para un solo nivel de resolución, en realidad la función `idht3d` recibe como argumentos de entrada las matrices de coeficientes, el número de niveles de descomposición multiresolución (nivel) y una bandeara para indicar si se considera una interpolación de $2^\uparrow$ en cada nivel, por lo que todo el código se repite tantas veces como niveles de resolución sean indicados. El número de iteraciones es de $\text{nivel} \times N_i$, donde N_i es la longitud de las matrices de coeficientes en cada nivel; en el peor caso que se refiere a la consideración del submuestreo en la DHT, en cada iteración se realizan 160 operaciones que comprenden sumas y restas más 160 operaciones de multiplicación por potencias de 2 las cuales pueden ser implementadas también como corrimientos de bits. Como el número de niveles es un número muy pequeño comparado con la longitud de los datos, la complejidad de este algoritmo es de $O(N)$, siempre y cuando $320 \ll N$.

6.3.4. Implementación de la transformada wavelet 3D

Para implementar la transformada wavelet en 3 dimensiones se aprovechó la separabilidad de sus filtros. De manera análoga al caso bidimensional explicado en la sección 3.5.5.1, para cada tipo de wavelet seleccionada desde la interfaz de la aplicación se obtienen los filtros paso bajo (L) y paso alto (H). En el proceso de análisis, los datos se filtran en una dirección con cada uno de los filtros y esto se sigue de un proceso de decimación de $\downarrow 2$ reduciendo así el tamaño de los datos a la mitad en la dirección filtrada; como en este caso se tienen 3 dimensiones, entonces el proceso se repite tres veces en cada una de las direcciones por cada nivel de descomposición (Figura 6.14).

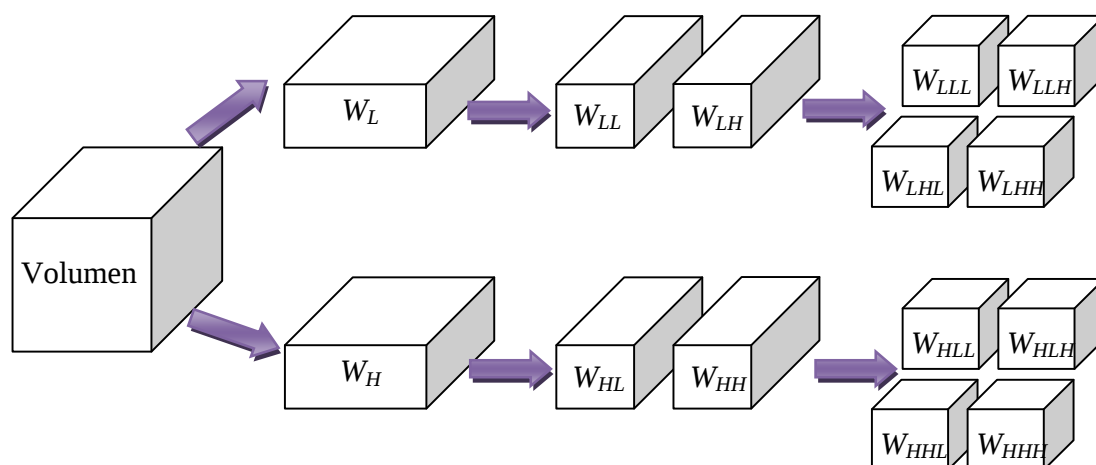


Figura 6.14 Esquema de la transformada wavelet en tres dimensiones.

Del esquema anterior se observa que si el volumen es de tamaño $M1 \times M2 \times M3$, después de aplicar banco de filtros de análisis a los renglones se obtienen dos conjuntos de subbandas de $M1/2 \times M2 \times M3$; luego de aplicar los filtros a las columnas se obtienen cuatro conjuntos de subbandas de $M1/2 \times M2/2 \times M3$; y al aplicar los filtros a la última dimensión se obtienen ocho subbandas de $M1/2 \times M2/2 \times M3/2$.

Finalmente, el banco de filtros de síntesis para la transformada wavelet 3D combina los ocho subbandas para generar el volumen original de $M1 \times M2 \times M3$, para ello, por cada dimensión primero se realiza una interpolación de $2^{\uparrow}$ en cada subbanda y luego se realiza la convolución con los filtros de síntesis cuyos resultados se suman hasta regresar a los datos originales.

Resultados y evaluación de la fusión de imágenes médicas

7.1. Resultados obtenidos de la fusión

En este capítulo se presentan algunos experimentos de fusión de imágenes con el propósito de comparar el desempeño del algoritmo propuesto en esta tesis con otros métodos más tradicionales. Esta comparación se realizó con algunos volúmenes de diferentes modalidades de imágenes médicas para lo cual se incluyen pruebas con la transformada wavelet 3D de familias como daubechies, symlet, coiflet y biortogonal.

Dado que no se cuenta hasta el momento con una base de datos de imágenes médicas registradas, se partió de un conjunto de 5 pares de imágenes tridimensionales previamente registradas, las cuales se describen a continuación, mostrando además tres vistas de cada volumen.



Este par de volúmenes consiste en secuencias de imágenes de tomografía computarizada y resonancia magnética en formato BMP. Cada secuencia consta de 89 imágenes de 320 por 320 de 8 bits por pixel. Las imágenes fueron proporcionadas por el Instituto de Óptica del Consejo Superior de Investigaciones Científicas en España [27] (Figura 7.1).



Este par de volúmenes está conformado también por secuencias de imágenes de resonancia magnética y tomografía por emisión de positrones en formato BMP. Cada secuencia consta de 100 imágenes de 256 por 256 de 8 bits por pixel. Al igual que el caso anterior, este conjunto de imágenes fue proporcionado por el Instituto de Óptica del Consejo Superior de Investigaciones Científicas en España [27] (Figura 7.2).



Ambos volúmenes están en formato analyze (.hdr + .img)⁷. En este caso se tenía un volumen con información anatómica de 288 x 288 x 180 y un volumen con información funcional de 80

⁷ En el formato Analyze cada imagen se compone de un archivo header de información (*.hdr) y de un archivo de datos (*.img). Puede incluir una cuarta dimensión que normalmente corresponde al tiempo (secuencia de imágenes 3D). Es el formato más usado, y es compatible con la mayoría de software.

x 80 x 44, pero para alinearlos el volumen de mayor resolución se redujo a la menor resolución quedando ambos de 80 x 80 x 44 con una resolución de 16 bits por voxel. Fuente desconocida (Figura 7.3).

MR-T1/ SPECT (Caso de estudio 1)

Ambos volúmenes están en formato NIfTI (.nii)⁸. La imagen de resonancia magnética T1 es un estudio del Instituto Neurologico de Montreal (espacio MNI), y el otro es el SPECT de un paciente del Instituto Nacional de Psiquiatría que fue llevado al mismo espacio. Las dimensiones de los datos son 91 x 109 x 91 con 16 bits de profundidad (Figura 7.4).

MR-T1/ SPECT (Caso de estudio 2)

Ambos volúmenes están en formato NIfTI (.nii). Ambas imágenes pertenecen a un paciente del Instituto Nacional de Psiquiatría donde se requería el análisis de la integración de ambos ya que el solo estudio funcional era insuficiente para un diagnóstico correcto. Las dimensiones de los datos son 256 x 256 x 160 con 16 bits de profundidad (Figura 7.5).

Los dos últimos pares de imágenes utilizadas en estos experimentos son cortesía del M.C. Sarael Alcauter del Instituto de Nacional de Psiquiatría.

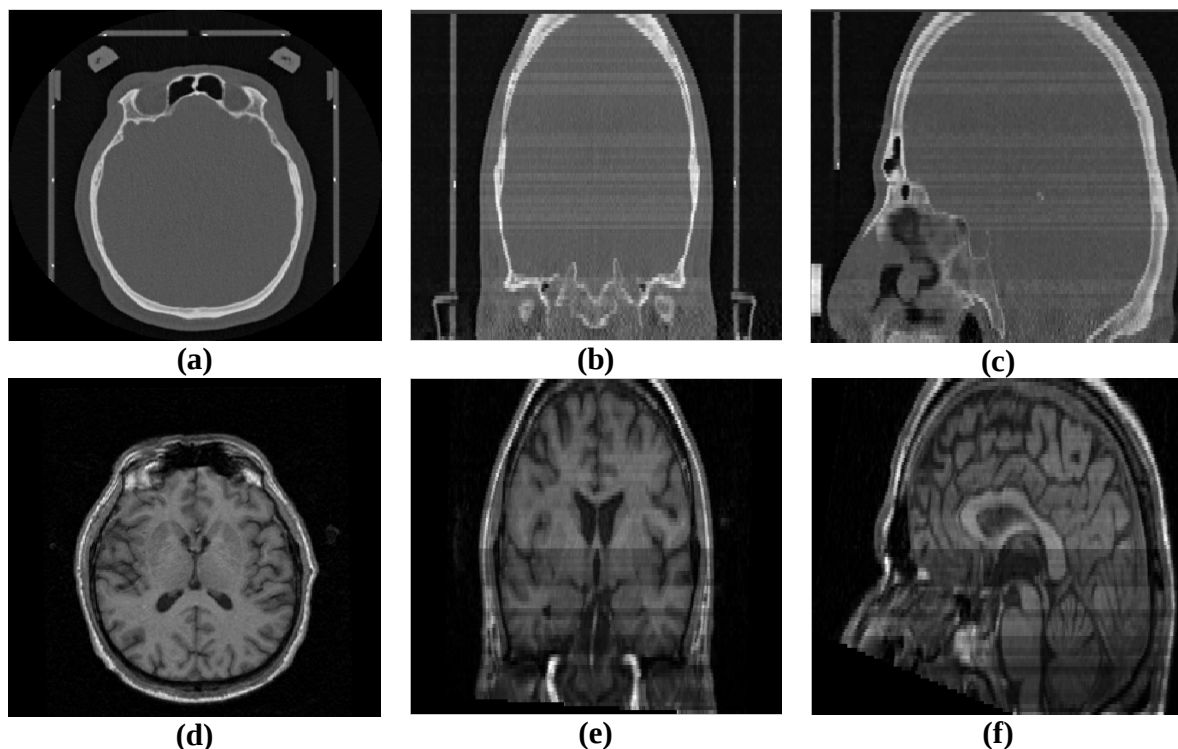


Figura 7.1 Volúmenes de prueba CT/MR. (a) y (d) corresponden a los cortes axiales 45/89, (b) y (e) a los cortes coronales 160/320, (c) y (f) a los cortes sagitales 160/320.

⁸ NIfTI es un formato Analyze mejorado para permitir la compresión de datos y la inclusión de más información en el header.

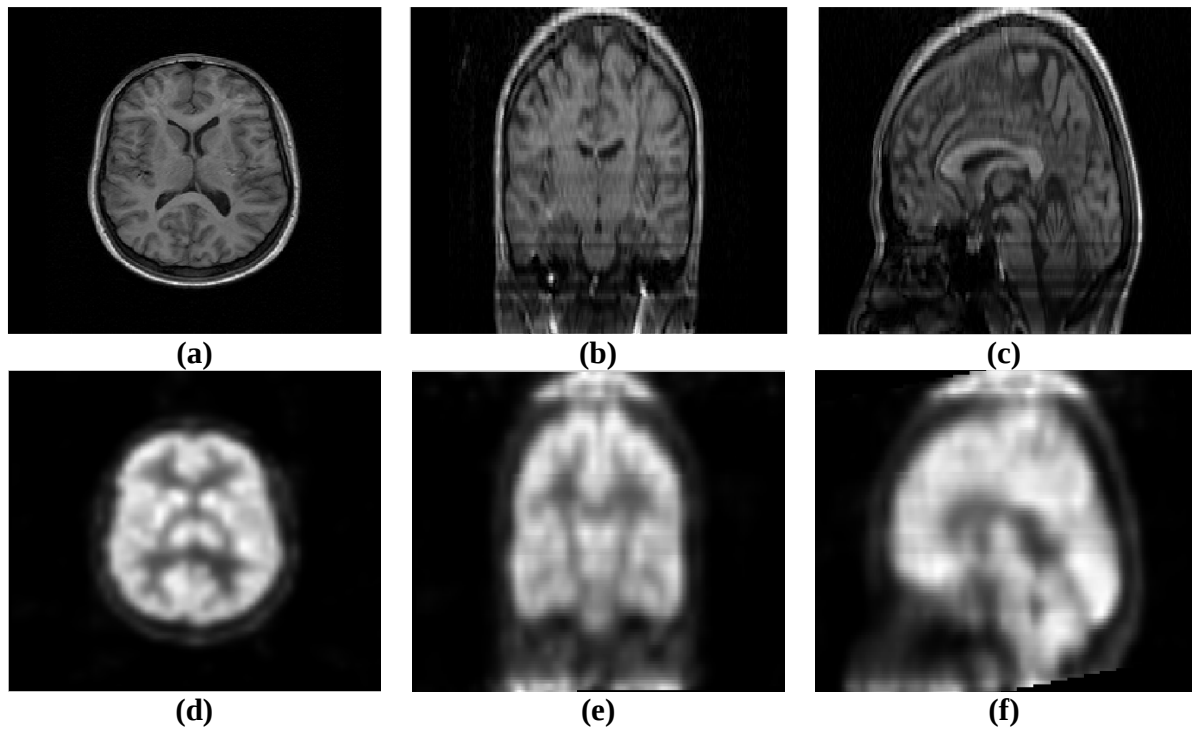


Figura 7.2 Volúmenes de prueba MR/PET. (a) y (d) corresponden a los cortes axiales 50/100, (b) y (e) a los cortes coronales 128/256, (c) y (f) a los cortes sagitales 128/256.

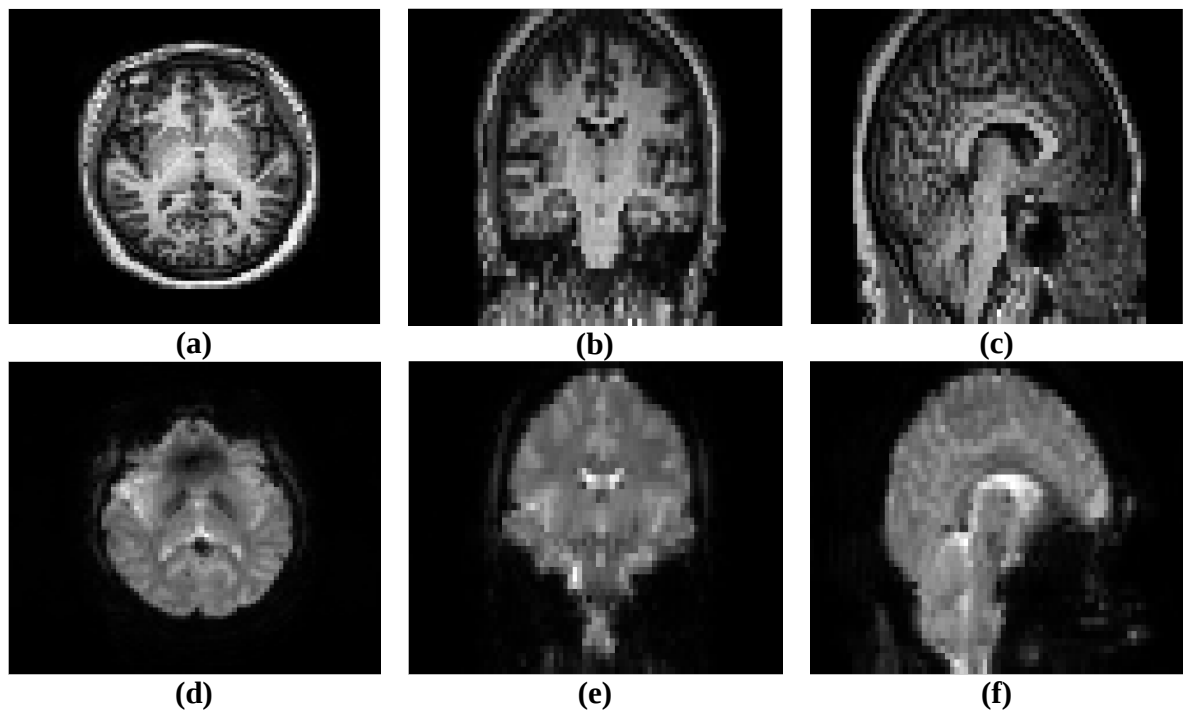


Figura 7.3 Volúmenes de prueba MR/fMR. (a) y (d) corresponden a los cortes axiales 22/44, (b) y (e) a los cortes coronales 40/80, (c) y (f) a los cortes sagitales 40/80.

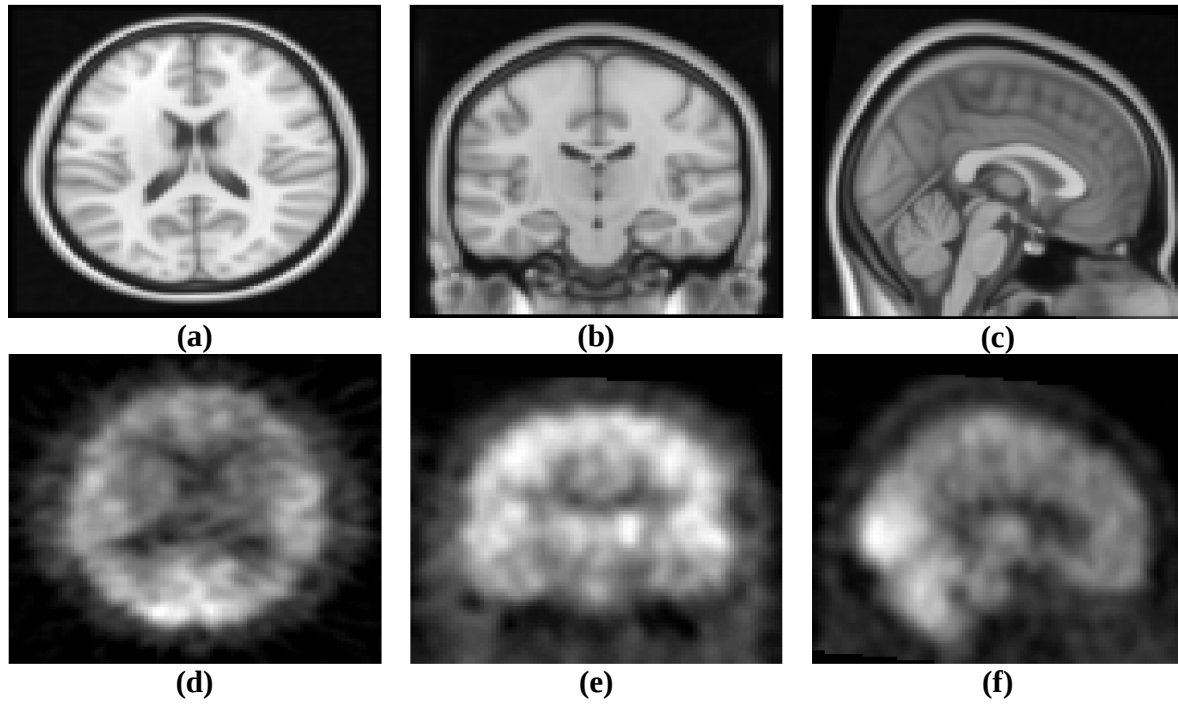


Figura 7.4 Volúmenes de prueba MR-T1/ SPECT (Caso de estudio 1). (a) y (d) corresponden a los cortes axiales 46/91, (b) y (e) a los cortes coronales 55/109, (c) y (f) a los cortes sagitales 46/91.

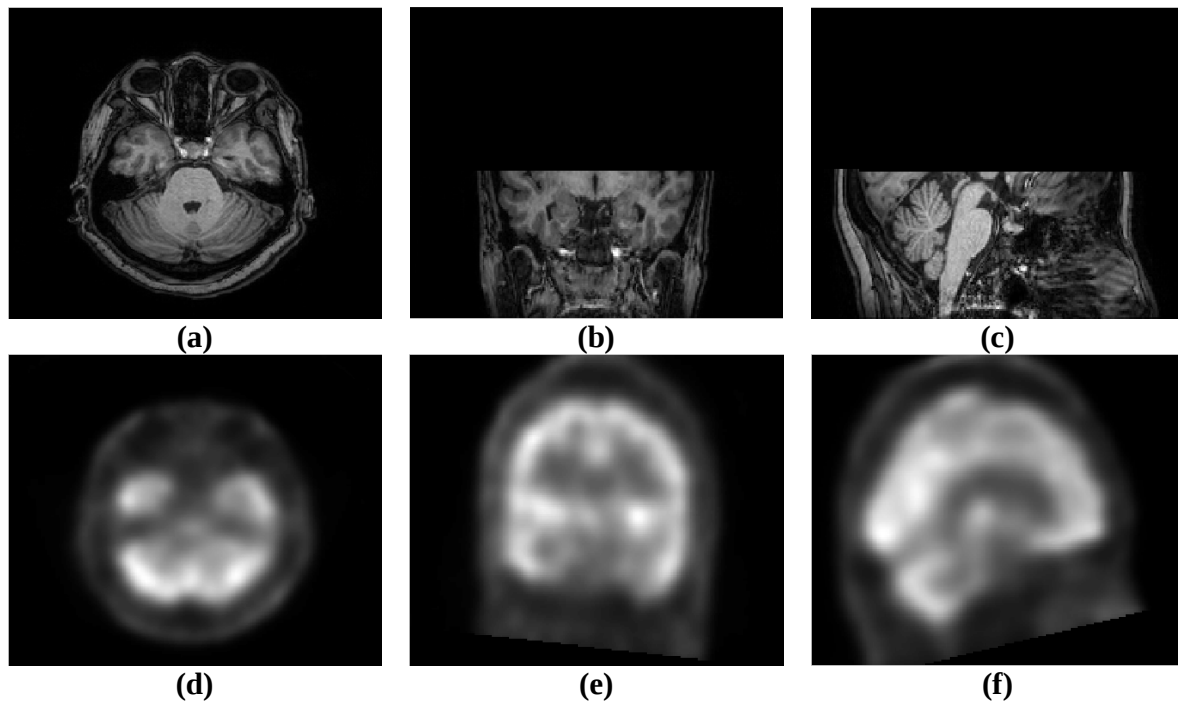


Figura 7.5 Volúmenes de prueba MR-T1/ SPECT (Caso de estudio 2). (a) y (d) corresponden a los cortes axiales 48/160, (b) y (e) a los cortes coronales 128/256, (c) y (f) a los cortes sagitales 128/256.

Aunque el algoritmo de fusión expuesto trabaja con datos tridimensionales, la aplicación de despliegue desarrollada no permite la visualización del volumen en tres dimensiones; como se observa en la Figura 6.5, al cargar los datos de origen sólo se presentan los cortes ortogonales del volumen: axial, coronal y sagital. De la misma manera, el despliegue del resultado de la fusión se realiza por cortes, donde se tiene la opción de mostrar sólo los cortes del resultado o bien comparar el resultado con los datos originales en cualquiera de los tres cortes ya mencionados (ver Figura 7.6).

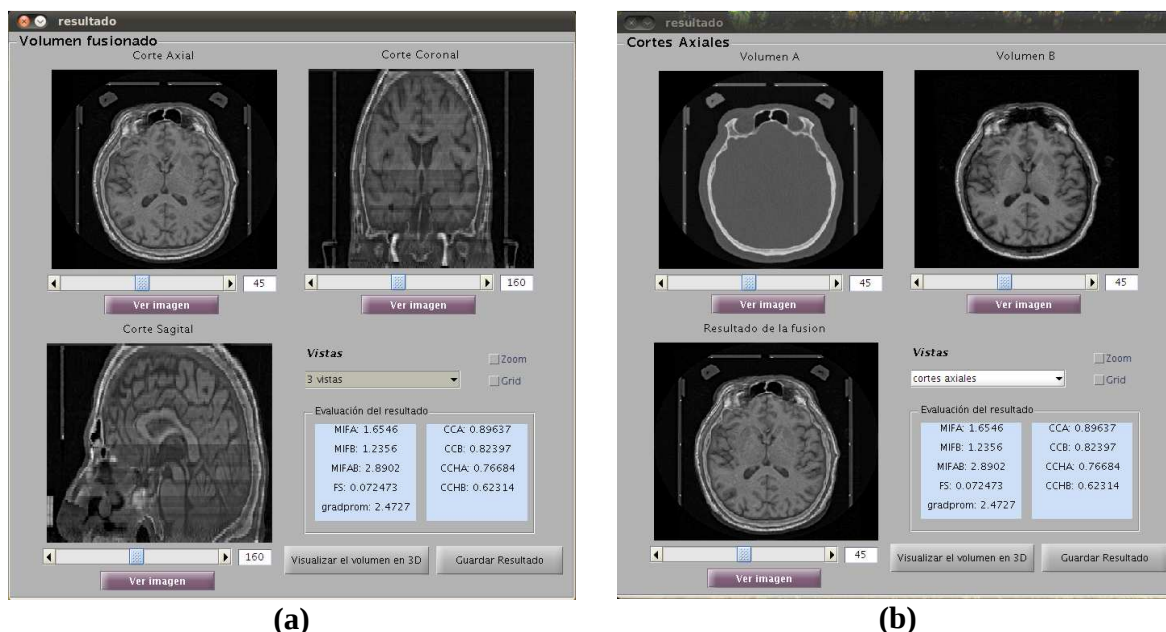


Figura 7.6 Interfaz gráfica para el despliegue del resultado de fusión. (a) Modo 3 vistas donde todas las vistas corresponden al volumen fusionado. (b) Modo cortes axiales donde se contrasta corte a corte el resultado con los datos originales en el plano axial.

Como se puede apreciar en la figura anterior, la vista en cortes axiales es la que presenta mejor calidad visual debido a las inhomogeneidades de captura entre cada corte en algunos de los conjuntos de datos empleados; por ello, en todos los ejemplos mostrados en adelante sólo se considerarán las vistas en cortes axiales y para fines comparativos se seleccionará un solo corte de alguno de los cinco pares de volúmenes para ilustrar los resultados presentados en tablas y gráficas.

Por otra parte, es importante señalar que todas las pruebas realizadas con imágenes médicas utilizan sólo dos niveles de descomposición. Esto es porque se realizaron pruebas con todos los volúmenes y en general se observó que entre los niveles 2 y 3 no existía una diferencia visual significativa en el resultado, no así en el tiempo de ejecución ya que aumentar el número de niveles de resolución implica también mayor costo computacional en proporción $\log_2 N$ donde N es el número de datos. En la Figura 7.7 se ilustra lo anterior con la fusión del primer par de datos (CT/MR), observándose una clara mejoría en el contraste entre los primeros dos niveles de descomposición.

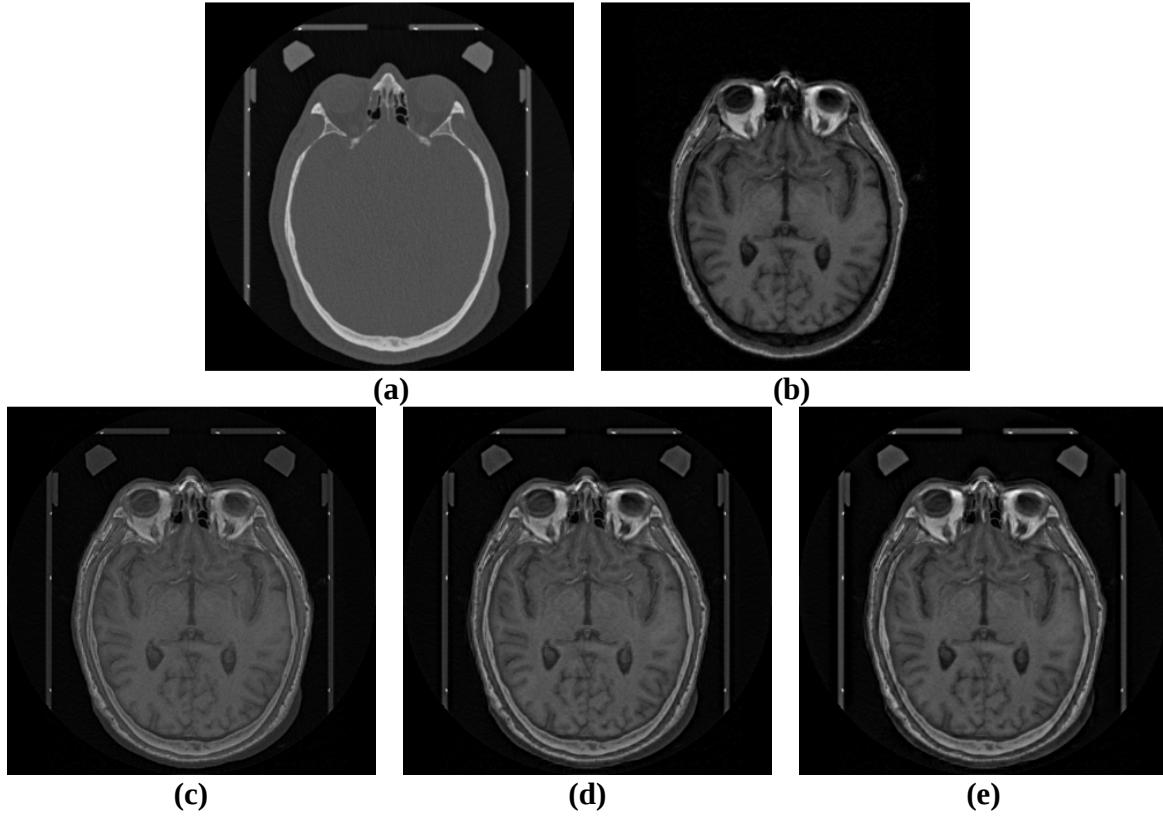


Figura 7.7 Fusión de CT/MR con utilizando la HT con el esquema propuesto a varios niveles n de resolución. (a) y (b) son los datos originales, (c) es el resultado de la fusión con $n=1$, (d) es el resultado de la fusión con $n=2$ y (e) es el resultado de la fusión con $n=3$

Para cada uno de los resultados mostrados en (c), (d) y (e) de la figura anterior se calculó los coeficientes de correlación paso altas, la información mutua y la simetría de fusión (ver Tabla 7.1), sin embargo éstas métricas no necesariamente concuerdan con la calidad visual ya que no consideran al sistema de visión humano. Dada la falta de una imagen de referencia se optó por incluir una métrica más denominada **gradiente promedio**, pues el gradiente de la imagen ha sido utilizado como una medida de nitidez [90], así se observa que el gradiente aumenta a medida que se incrementan el nivel de descomposición de las imágenes en el análisis.

El gradiente en cualquier voxel es la derivada de los valores en una vecindad. Generalmente las imágenes nítidas tienen valores altos de gradiente. De esta manera, el resultado de un proceso de fusión deberá ser una imagen más nítida comparada con la imagen de entrada de menor resolución. El gradiente promedio se define el contraste entre la variación de detalles de patrones sobre la imagen y la claridad de la imagen. El gradiente promedio de una imagen X está dado por

$$\overline{G} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \sqrt{\frac{\Delta I_x^2 + \Delta I_y^2}{2}} \quad (7.1)$$

donde Δx y Δy son los gradientes horizontal y vertical por voxel. En la práctica ambos gradientes fueron calculados mediante la convolución con un filtro sobel.

Tabla 7.1 Resultados del desempeño de la fusión para diferentes niveles de resolución (Figura 7.7) donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B, MI_{FA} y MI_{FB} es la información mutua entre el volumen fusionado y los volúmenes fuente A y B, MI_{FAB} es la información mutua total y FS la simetría de fusión.

N° de niveles	Gradiente promedio	CCHA	CCHB	MI_{FA}	MI_{FB}	MI_{FAB}	FS
1	60.1024	0.8168	0.4739	1.5509	1.5740	3.1249	0.0766
2	69.2615	0.8149	0.4713	1.3376	1.3975	2.7351	0.0712
3	71.4110	0.8185	0.4685	1.1779	1.2285	2.4065	0.0583

7.1.1. Comparación con distintas reglas de fusión

A fin de evaluar el desempeño de la regla de fusión propuesta para los coeficientes de detalle, la dependencia lineal fue comparada con otros métodos de selección abordados en el capítulo 3: selección del valor absoluto del máximo, selección del máximo con verificación de consistencia y saliente y medida de paridad [15]. Para este propósito se realizaron en primera instancia algunas pruebas con imágenes multifocales creadas a partir de una imagen donde todos los objetos están en foco y otras más con los volúmenes de imágenes médicas.

La Figura 7.8 muestra uno de los conjuntos de imágenes multifocales empleados en la evaluación de fusión 2D, este conjunto de imágenes fue generado a partir de la imagen de referencia, cortesía del Dr. Shutao Li [62]. Los resultados obtenidos con las diferentes reglas de fusión no se incluyen ya que son muy parecidos a la imagen de referencia (a), por lo que a simple vista no se logran distinguir las diferencias, como se muestra en la Tabla 7.2 todas las imágenes fusionadas obtuvieron un PSNR mayor a 30 [dB]. En estas pruebas, se realizó la fusión mediante la HT multirresolución con 4 niveles de resolución. El tamaño de la ventana para las reglas de máximo con verificación de consistencia, dependencia lineal y saliente y medida de paridad fue de 3 x 3.

De la tabla 7.2 se observa que el resultado de aplicar la dependencia lineal obtuvo el menor error cuadrático medio (MSE) y en consecuencia la relación señal a ruido (PSNR) más alta. A pesar de no presentar los valores más altos de índice de similitud estructural (SSIM) e información mutua entre la imagen fusionada y la de referencia (MI_{FR}), la regla de dependencia lineal es la segunda mejor evaluada de acuerdo a estas últimas métricas.

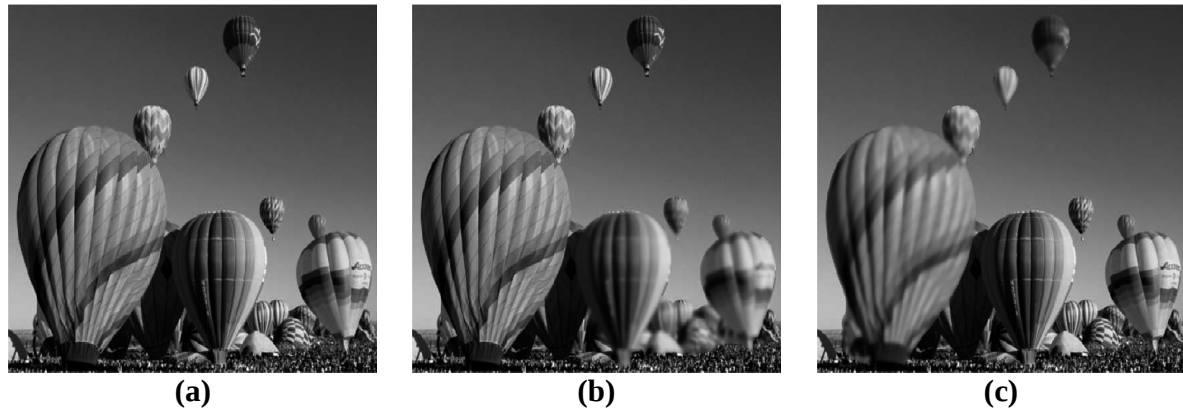


Figura 7.8 Imágenes multifocales empleadas en una prueba de evaluación de distintas reglas de fusión. (a) Imagen original considerada ‘the ground truth’ para el cálculo de las métricas. (b) y (c) Imágenes empleadas en el proceso de fusión donde ciertos objetos han sido emborronados.

Tabla 7.2 Resultados del desempeño de la fusión con las imágenes de la Figura 7.8 empleando la HT con diferentes reglas de fusión. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI_{FR} a la información mutua entre la imagen fusionada y la imagen de referencia.

Regla de fusión	MSE	PSNR [dB]	SSIM	MI_{FR}
Máximo absoluto	4.4293	41.6674	0.997548	5.5351
Máximo con verificación de consistencia	0.4407	51.6886	0.999641	6.5348
Saliente y medida de paridad	4.6604	41.4465	0.996923	5.4942
Dependencia lineal	0.4357	51.7385	0.999625	6.4807

Para ilustrar la comparación entre las reglas de fusión en imágenes médicas 3D, en la Figura 7.9 se muestran los resultados de aplicar estas reglas al conjunto de CT/MR en el corte 68. La fusión se realizó con dos niveles de descomposición de la HT, con una región de $3 \times 3 \times 3$ para la aplicación de todas las reglas de fusión a excepción del máximo absoluto que no requiere una vecindad de voxels.

A diferencia de las imágenes multifocales, en todas las pruebas de imágenes médicas los mejores resultados cuantitativos fueron las últimas dos, con una ligera ventaja de la dependencia lineal sobre el método de Burts.

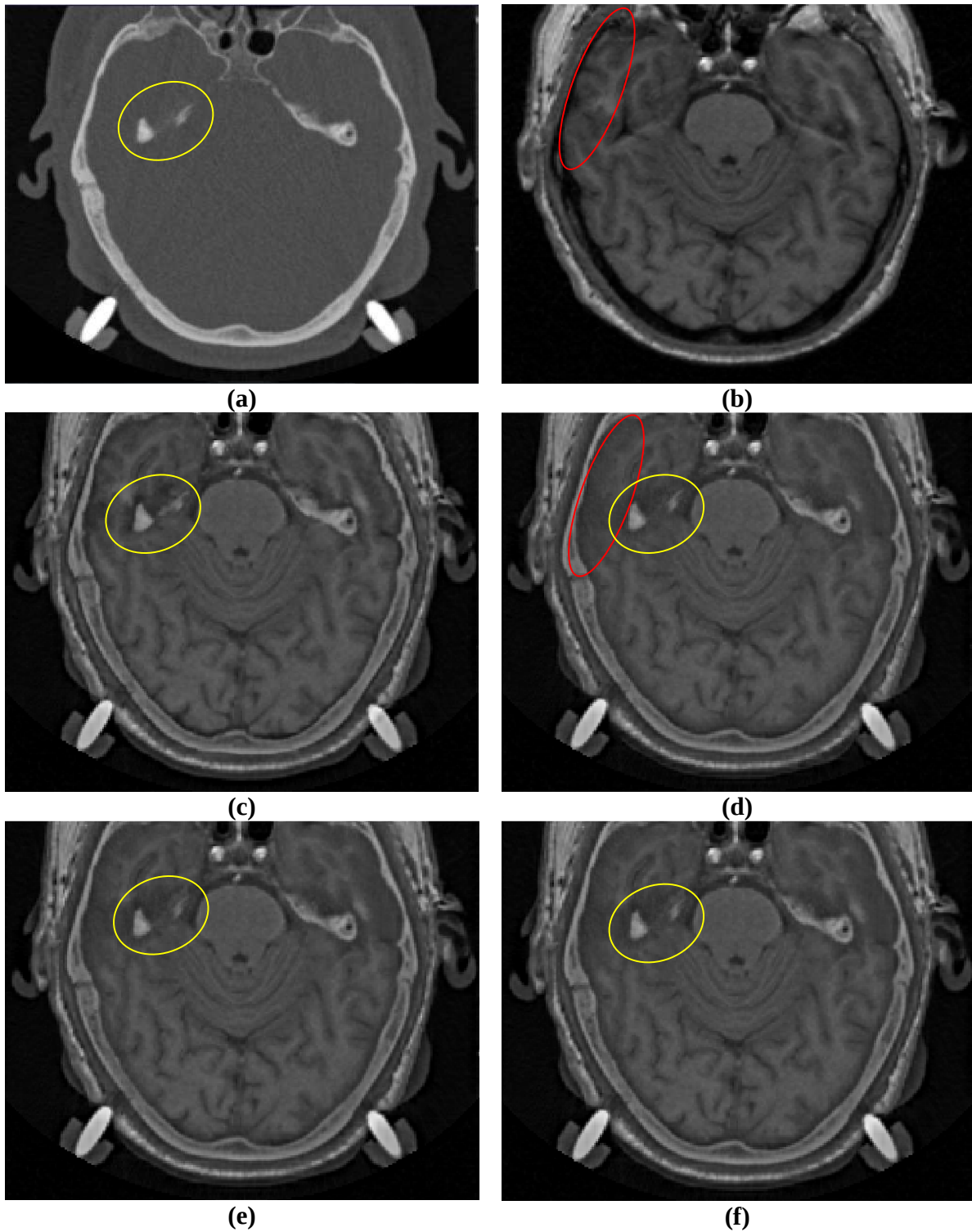


Figura 7.9 Comparación de fusión en imágenes médicas con distintas reglas de fusión. (a) y (b) corte 68 de los volúmenes de CT y MR respectivamente, (c) resultado empleando máximo absoluto, (d) resultado empleando máximo con verificación por consistencia, (e) resultado empleando saliente y medida de paridad, (f) resultado empleando dependencia lineal.

Como en este caso no siempre se da la correlación entre la calidad de las imágenes y la información mutua, se analizaron los detalles en los resultados. En este ejemplo se marcaron dos zonas de interés en cada una de las imágenes originales, en la tomografía se encerró con un círculo amarillo dos regiones de hueso, mientras que en la imagen de resonancia magnética se señaló con rojo una porción de la corteza cerebral. El resultado obtenido de aplicar la regla de máximo absoluto, Figura 7.9 (c), pareciera preservar mejor los detalles de las imágenes originales, no obstante, al no considerar el contexto de cada voxel, es propenso a la introducción de artefactos como lo muestra la zona de marcada en todas imágenes fusionadas. Por otro lado, el resultado obtenido de aplicar la regla de máximo con verificación de consistencia, Figura 7.9 (d) presenta una mejoría respecto al anterior ya que considera a sus vecinos para la selección de coeficientes, sin embargo, la región señalada con rojo de la MR pierde muchos detalles. Finalmente, las reglas de saliente y medida de paridad y dependencia lineal, Figura 7.9 (e) y (f), presentan resultados muy parecidos, por lo que nuevamente se recurrió a una comparación cuantitativa la cual se muestra en la Tabla 7.3, donde se observa una ligera mejora en las métricas de evaluación la conservación de los bordes como el gradiente promedio y los coeficientes de correlación paso altas, mientras que en las métricas basadas en la información mutua se observa que las mejores reglas son el máximo con verificación de consistencia y la medida de saliente y paridad.

Tabla 7.3 Resultados del desempeño de la fusión con las imágenes de la Figura 7.8 empleando la HT con diferentes reglas de fusión donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B, MI_{FA} y MI_{FB} es la información mutua entre el volumen fusionado y los volúmenes fuente A y B, MI_{FAB} es la información mutua total y FS la simetría de fusión.

Regla de fusión	Gradiente promedio	CCHA	CCHB	MI_{FA}	MI_{FB}	MI_{FAB}	FS
Máximo absoluto	69.3731	0.81642	0.473	1.34406	1.39206	2.73612	0.07474
Máximo con verificación de consistencia	68.87646	0.81436	0.4659	1.3399	1.38836	2.72828	0.07112
Saliente y medida de paridad	69.0401	0.81184	0.46924	1.34168	1.40026	2.74196	0.07148
Dependencia lineal	69.2615	0.8149	0.47138	1.33762	1.39758	2.73518	0.07122

7.1.2. Comparación con la transformada wavelet

Como se mencionó en la revisión del estado del arte de las técnicas de fusión, la transformada wavelet es el método más empleado en los algoritmos de fusión, no obstante, es un método que presenta algunas desventajas en la representación de señales de dos o más dimensiones debido en gran parte a su limitación para capturar la información direccional. Para mostrar que la transformada de Hermite supera estas limitaciones se efectuaron pruebas aplicando las mismas reglas de selección de coeficientes a ambas transformadas a los volúmenes descritos al inicio de este capítulo, con las mismas consideraciones de la sección anterior para la regla de dependencia línea (vecindad de $3 \times 3 \times 3$). Este estudio no se limitó a

un solo tipo de wavelet, ya que se seleccionaron las primeras wavelets de 4 familias: biortogonales, de Daubechies, coiflets y symlets.

En las Figura 7.10 se muestra las gráfica con la información mutua total de 3 pares de volúmenes para cada uno de los métodos evaluados. En ella se observa que la transformada de Hermite es una de las técnicas que mejor preserva la información de los datos originales, pues de los tres casos analizados, sólo en el conjunto de datos del caso 2 Hermite se ve superado por el resultado la wavelet biortogonal 2.2.

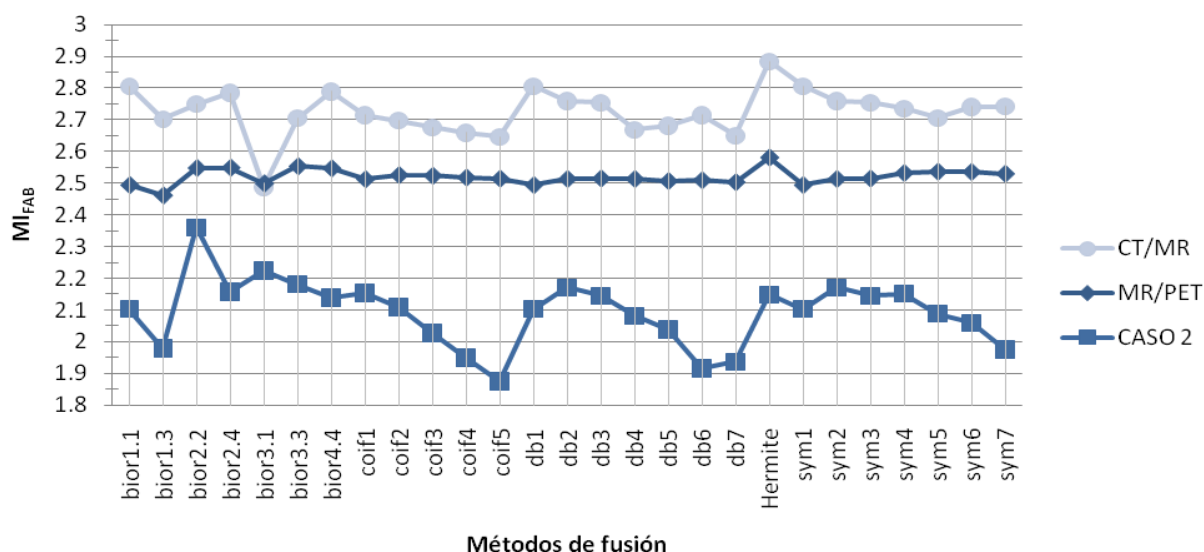


Figura 7.10 Gráfica de la información mutua total del proceso de fusión para los volúmenes CT/MR, MR/PET y Caso 2 para diversos métodos de fusión: bior se refiere a una wavelet biortogonal, coif a una coiflet, db a Daubechies, Hermite a la técnica propuesta y sym a una wavelet de tipo symlet. En el eje 'x' se enlistan todas las técnicas de fusión empleadas en los experimentos y en el eje 'y' los valores de información mutua total entre los volúmenes fusionados y los volúmenes fuente.

A continuación se muestran también dos ejemplos gráficos para comparar el resultado de fusión entre la DHT y la DWT. Para no incluir 27 imágenes se eligió un tipo de wavelet de cada familia, en ambos casos se ilustran los resultados mostrando sólo un corte de cada volumen fusionado, sin embargo cabe destacar que información mutua mostrada en la Figura 7.10 y todas las métricas expuestas en la Tabla 7.4 fueron calculadas a partir de los volúmenes y no sólo de las imágenes que se muestran en las Figuras 7.11 y 7.12.

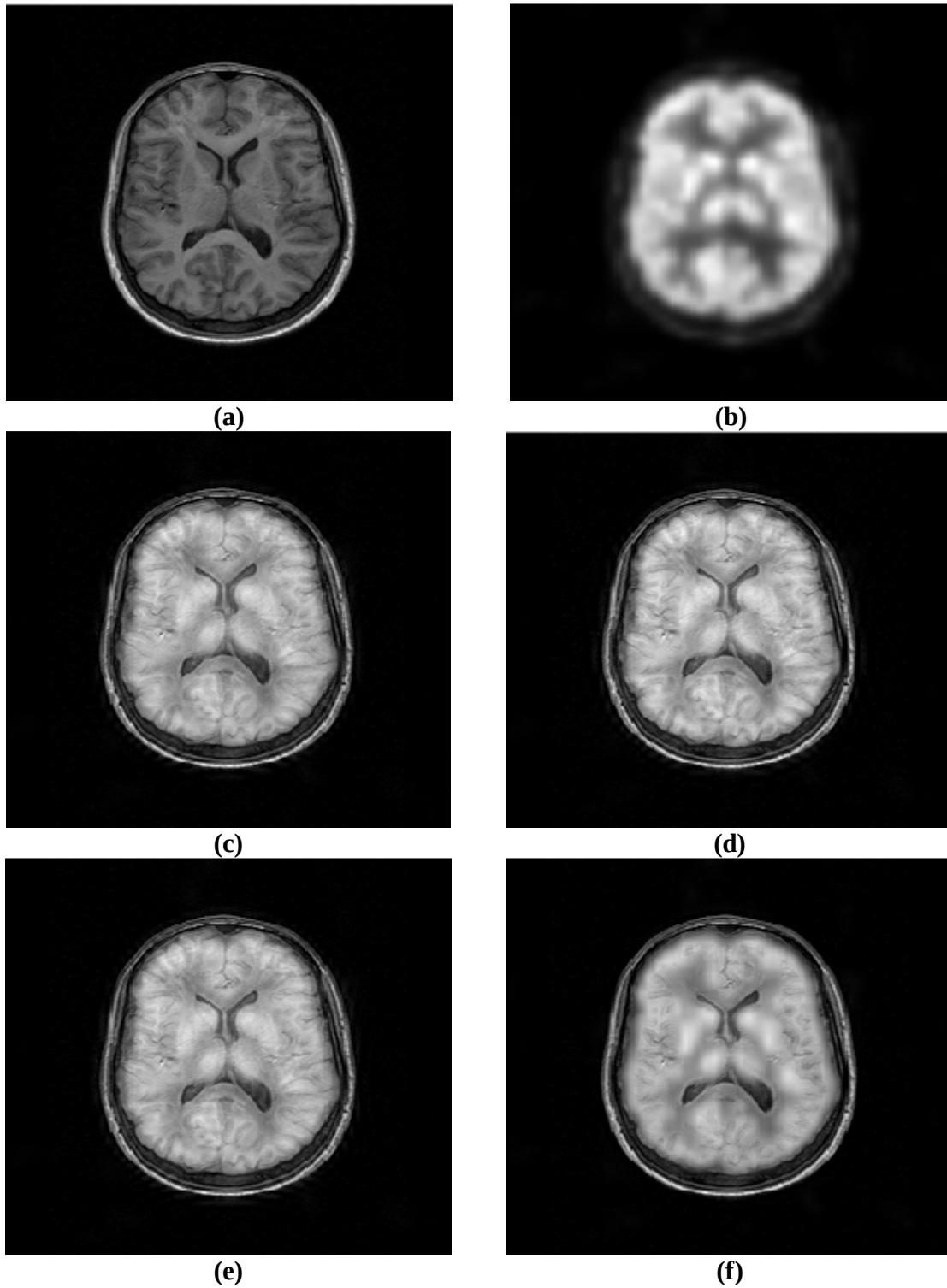
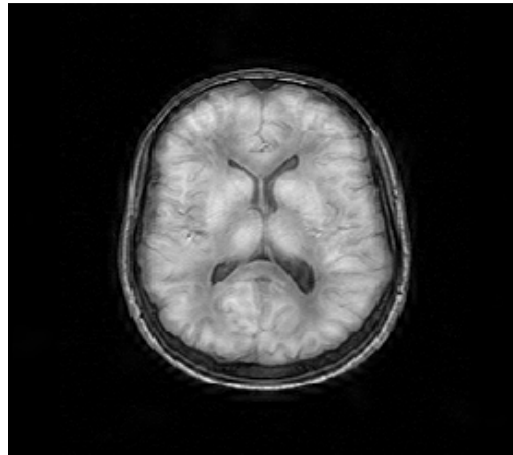
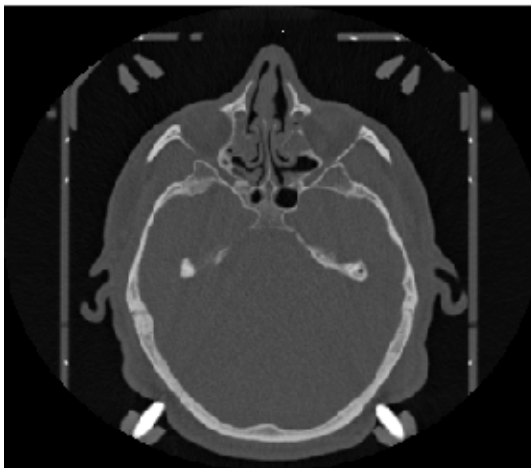


Figura 7.11 Comparación de fusión en imágenes médicas con distintos métodos de fusión. (a) y (b) corte 50 de los volúmenes de MR y PET respectivamente, (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5, (e) fusión con DWT Daubechies 6, (f) fusión con DHT.

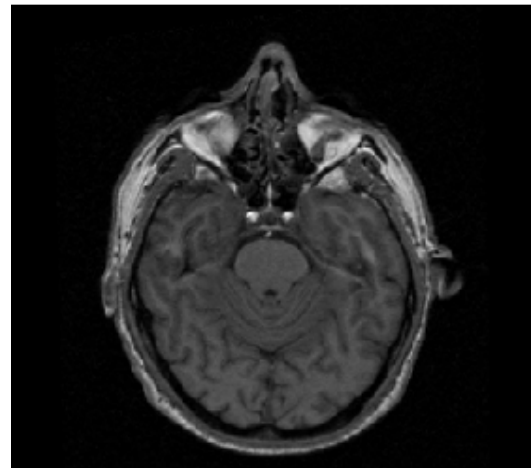


(g)

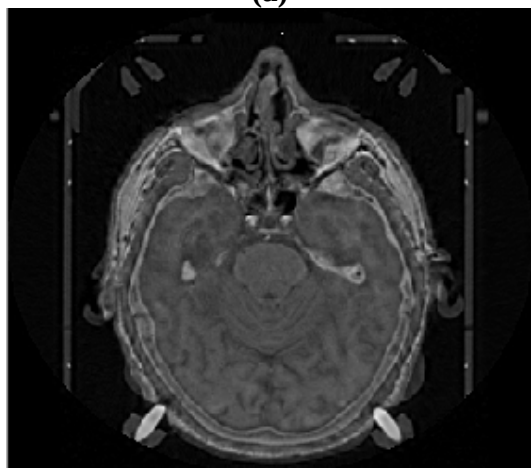
Figura 7.12 (Cont.) Comparación de fusión en imágenes médicas con distintos métodos de fusión. (g) fusión con DWT symlet 7.



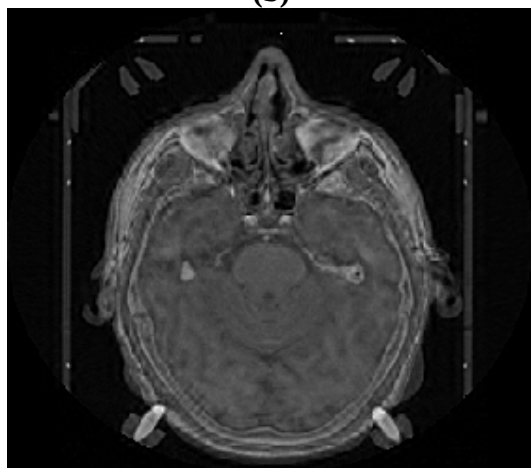
(a)



(b)



(c)



(d)

Figura 7.13 Comparación de fusión en imágenes médicas con distintos métodos de fusión. (a) y (b) corte 68 de los volúmenes de CT y MR respectivamente, (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5.

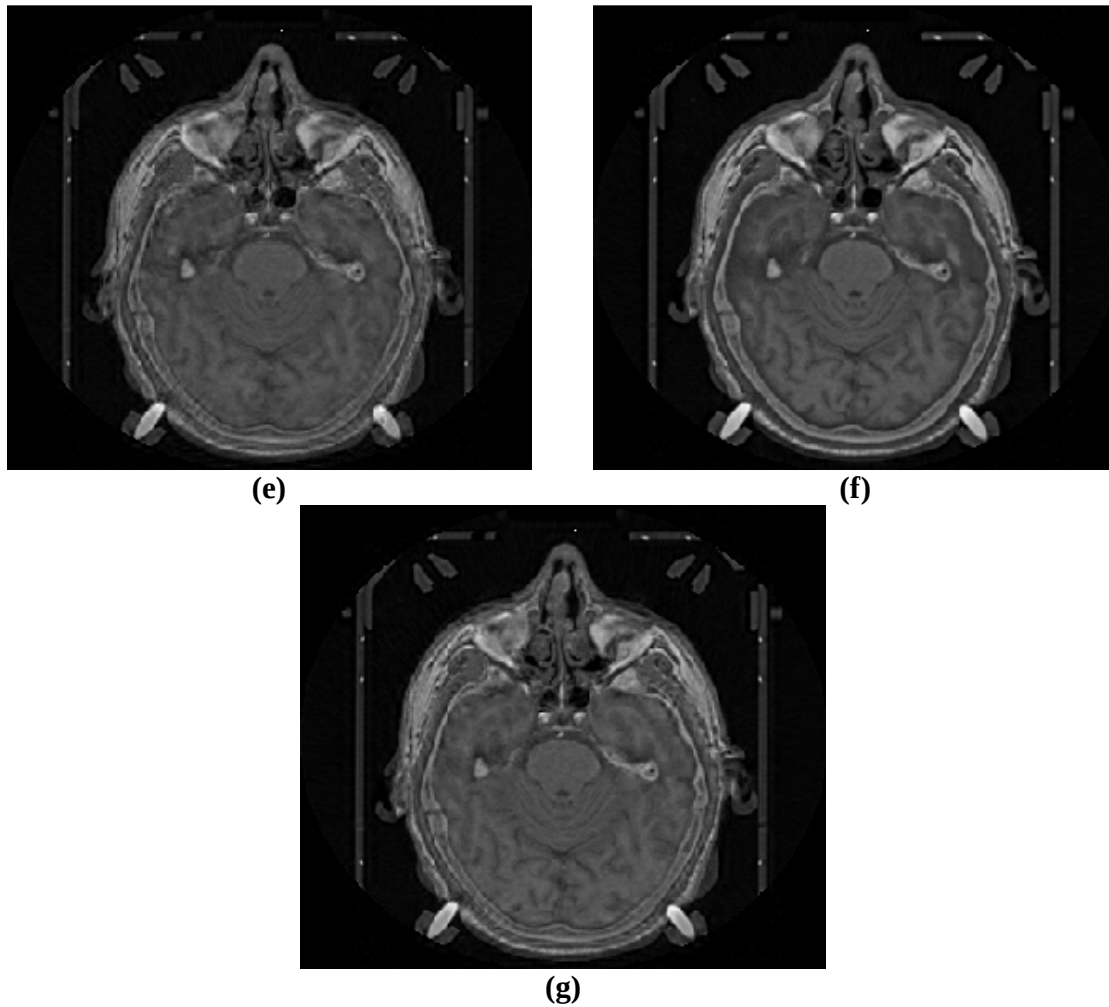


Figura 7.14 (Cont.) Comparación de fusión en imágenes médicas con distintos métodos de fusión. (e) fusión con DWT Daubechies 6, (f) fusión con DHT, (g) fusión con DWT symlet 7.

Tabla 7.4 Promedio de los resultados obtenidos del proceso de fusión de los 5 volúmenes aplicando distintos métodos de fusión donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B, MI_{FA} y MI_{FB} es la información mutua entre el volumen fusionado y los volúmenes fuente A y B, MI_{FAB} es la información mutua total y FS la simetría de fusión.

Método	Gradiente promedio	CCHA	CCHB	MI_{FA}	MI_{FB}	MI_{FAB}	FS
bior3.3	38.4213	0.8768	0.4002	1.2154	1.2642	2.4796	0.09666
coif5	38.3965	0.8752	0.3861	1.1706	1.1749	2.3455	0.07946
db6	38.4312	0.8731	0.3889	1.1855	1.1949	2.3803	0.08546
Hermite	39.2701	0.8920	0.4047	1.1903	1.3472	2.5375	0.09336
sym7	38.3998	0.8727	0.3993	1.1934	1.2224	2.4158	0.08843

De los ejemplos mostrados, en la fusión de CT con MR (Figura 7.12) es más evidente la distorsión del volumen fusionado debido a los artefactos generados por el análisis de las imágenes mediante la transformada wavelet mientras que la transformada de Hermite presenta un resultado integral con mejor calidad visual. Comparando sólo los resultados de obtenidos con wavelets es notorio que las familias biortogonal y symlet son las que mejor se adaptan al contenido de las imágenes y por tanto hay menos artefactos que en los resultados de Daubechies y coiflets.

La Figura 7.11 es más difícil de interpretar debido a que en este caso la fusión es entre una imagen anatómica (MR) y otra funcional (PET), en este caso no necesariamente deben conservarse todos los detalles sobresalientes ya que esto discriminaría casi por completo la información de la actividad cerebral que se desea estudiar. En este sentido se observa que la fusión mediante la DHT permite una mejor integración de la información de ambas fuentes.

7.2. Evaluación del método propuesto en presencia de ruido

Para evaluar el desempeño del método propuesto en imágenes ruidosas se realizaron diversas pruebas con dos tipos de ruido: ruido aditivo normal (Gaussiano) y ruido impulsivo (sal y pimienta) variando los parámetros de ambos tipos, en el primer caso se realizaron pruebas con un rango de varianzas de 0.001 al 0.1 con incrementos de 0.001 y de manera similar para el ruido de sal y pimienta se varió el rango de densidades dentro de estos valores. A continuación se describen las características de cada uno de manera breve así como los resultados obtenidos.

7.2.1. Ruido aditivo normal (Gaussiano)

Este tipo de ruido se caracteriza por tener un espectro de energía constante para todas las frecuencias. Cuando se presenta este problema, el valor exacto de cualquier píxel es diferente cada vez que se captura la misma imagen. Este efecto, suma o resta un determinado valor al nivel de gris real y es independiente de los valores que toma la imagen.

$$G(i, j, k) = I(i, j, k) + n(i, j, k) \quad (7.2)$$

donde $n(i, j, k)$ forma parte de una distribución normal de media μ y variancia σ^2 .

Al aplicarle un filtro pasa bajos el ruido se suaviza bastante, sin embargo esto implica aplicar un pre-procesamiento previo al proceso de fusión por lo que una ventaja adicional de cualquier método de fusión es la reducción de ruido implícita durante el proceso.

Un ejemplo de la efectividad de la transformada de Hermite rotada en la fusión de imágenes degradadas con ruido Gaussiano se encuentra en la Figura 7.13 donde se presentan los resultados de fusionar una imagen de MR y con una de SPECT, notándose claramente la mejora en la calidad de la imagen obtenida respecto a las demás wavelets cuyos resultados en

este caso son muy parecidos. En las Figura 7.14 se muestran las gráficas de los coeficientes de correlación paso altas para este mismo ejemplo, lo cual coincide con las imágenes observadas ya que con el método propuesto se logra una mejor claridad de las estructuras de las imágenes originales. Finalmente, en la tabla 7.5 se muestra los valores de los coeficientes de correlación paso altas de los resultados mostrados en la Figura 7.13.

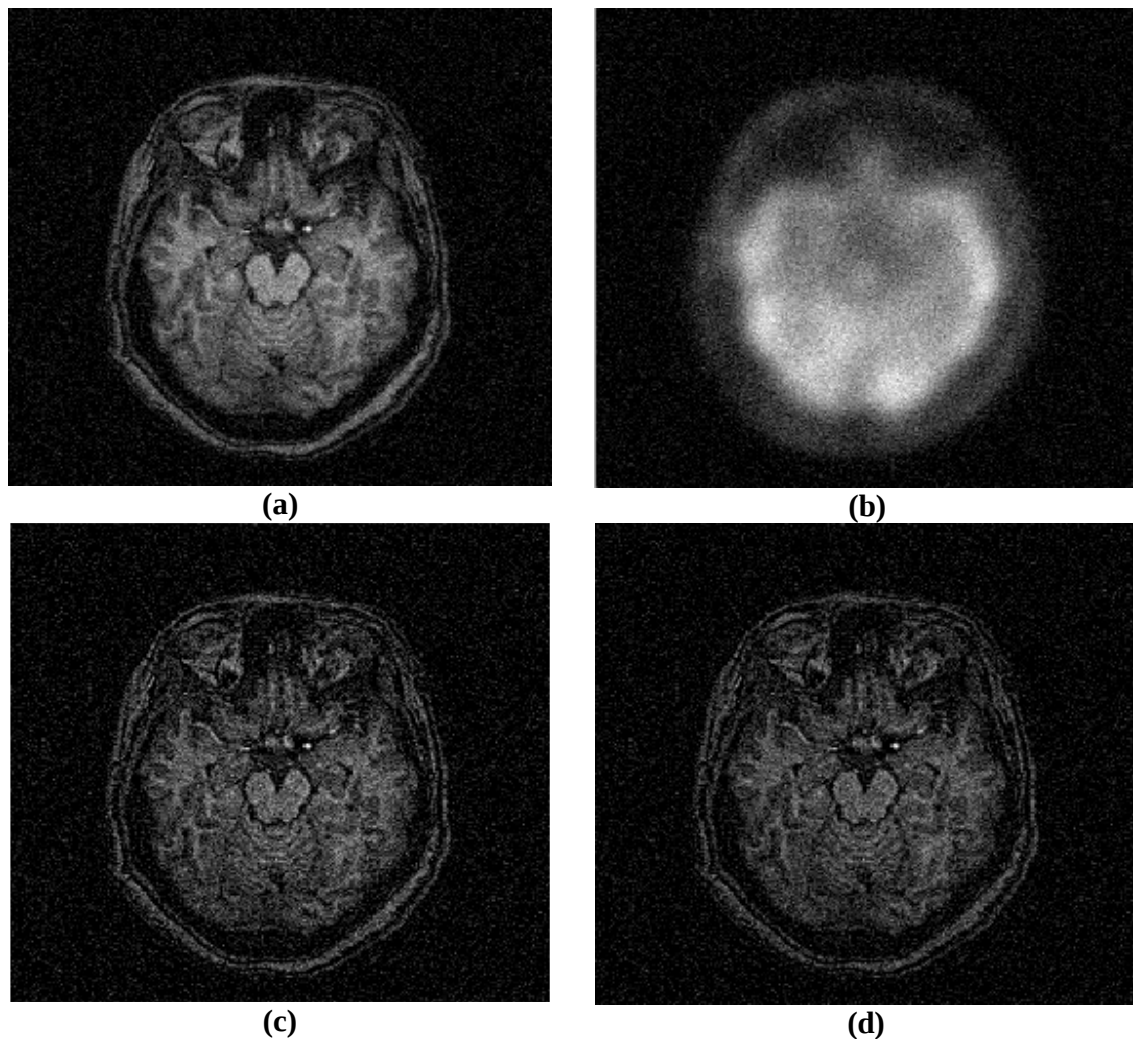


Figura 7.15 Comparación de fusión en imágenes médicas degradadas con ruido Gaussiano $\sigma=0.001$ empleando distintos métodos de fusión. (a) y (b) corte 58 de los volúmenes de MR y SPECT (caso de estudio 2) contaminados con ruido, (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5.

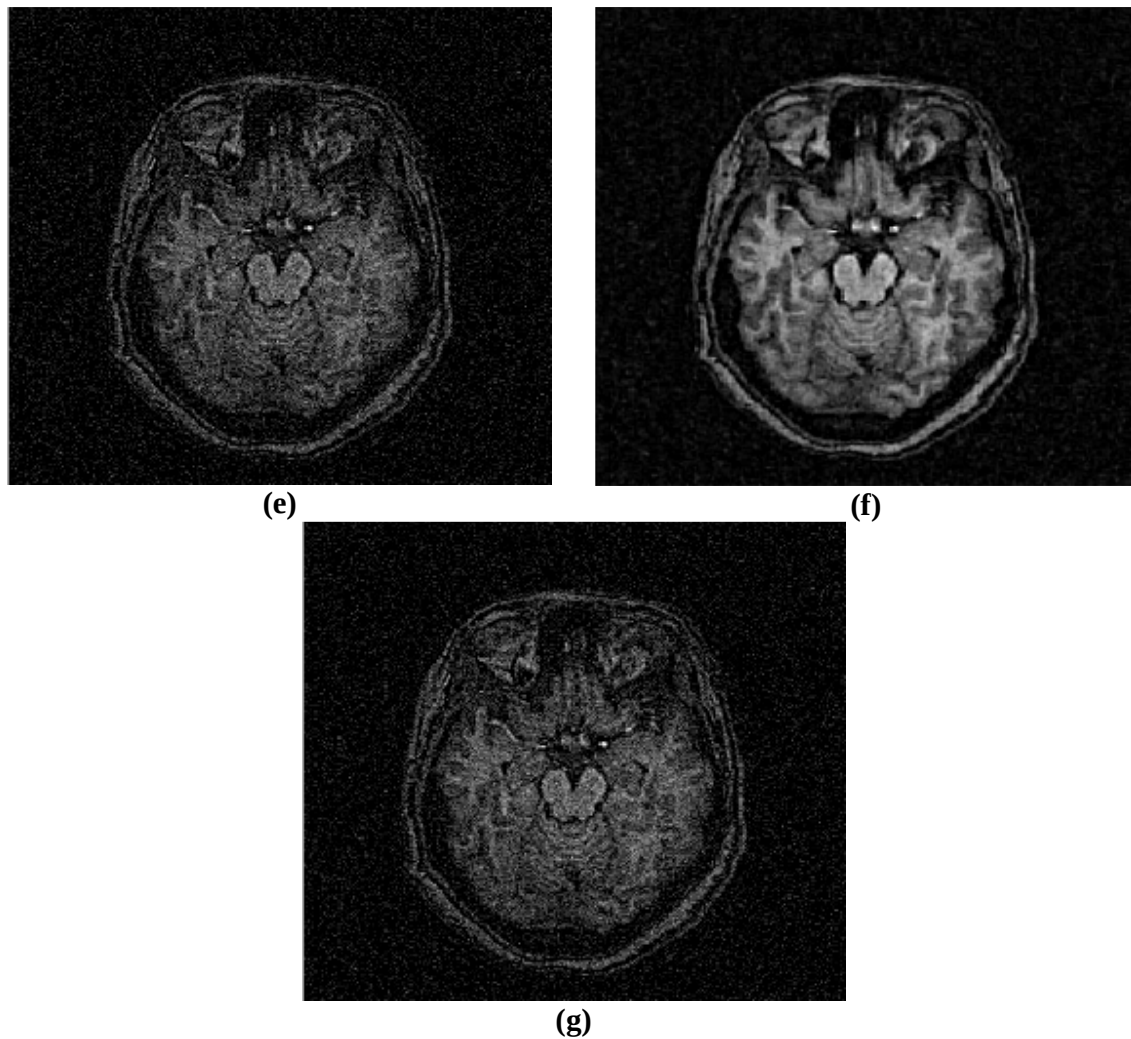


Figura 7.16 (Cont) Comparación de fusión en imágenes médicas degradadas con ruido Gaussiano $\sigma=0.001$ empleando distintos métodos de fusión. (e) fusión con DWT Daubechies 6, (f) fusión con DHT, (g) fusión con DWT symlet 7.

La razón por la cual la transformada de Hermite da mejores resultados respecto a todas las wavelets evaluadas es por el filtrado implícito que resulta al aplicar esta técnica como producto de la rotación de coeficientes, pues esta rotación implica proyectar únicamente los coeficientes hasta segundo orden dejando fuera coeficientes de frecuencias más altas. En cambio, en el caso de las wavelets no existe una forma natural para orientar sus coeficientes ya que todas las wavelets empleadas en estos experimentos carecen de una expresión de analítica que facilite el cambio de sistema de referencias como sucede con Hermite.

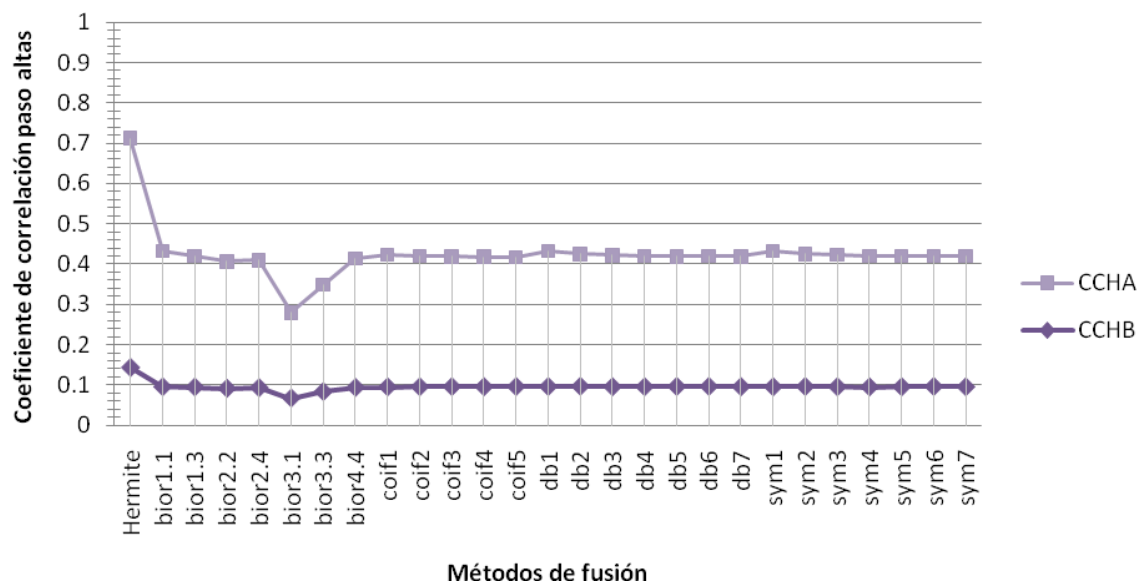


Figura 7.17 Gráfica de los coeficientes de correlación paso altas del resultado de fusión con reducción de ruido Gaussiano con $\sigma=0.001$ de los volúmenes MR-T1/SPECT (caso 2) con MR-T1 y SPECT para diversos métodos de fusión: bior se refiere a una wavelet biortogonal, coif a una coiflet, db a Daubechies, Hermite a la técnica propuesta y sym a una wavelet de tipo symlet. En el eje 'x' se enlistan todas las técnicas de fusión empleadas en los experimentos y en el eje 'y' los valores del coeficiente de correlación paso altas entre el volumen fusionado y las imágenes fuente: CCHA corresponde al coeficiente de correlación entre el volumen fusionado y el volumen A (MR-T1), y CCHB corresponde al coeficiente de correlación entre el volumen fusionado y el volumen B (SPECT).

Tabla 7.5 Promedio de los resultados obtenidos del proceso de fusión de los 5 volúmenes aplicando distintos métodos de fusión donde CCHA y CCHB son los coeficientes de correlación paso altas del volumen fusionado con los volúmenes fuente A y B.

Método	CCHA	CCHB
bior3.3	0.3490	0.0831
coif5	0.4165	0.0957
db6	0.4198	0.0961
Hermite	0.7121	0.1431
sym7	0.4197	0.0950

7.2.2. Ruido impulsivo (sal y pimienta)

Este ruido se caracteriza porque el pixel ruidoso en cuestión no tiene relación alguna con los píxeles circundantes, generalmente afecta a una pequeña cantidad de píxeles de la imagen. Al ver la imagen, encontraremos puntos blancos sobre puntos negros o puntos negros sobre puntos blancos, de ahí el término sal y pimienta. Defectos que contribuyen a este tipo de ruido son las manchas de polvo dentro de las ópticas de la cámara, o bien un CCD defectuoso, que realizará una captura errónea.

$$G(i, j, k) = \begin{cases} I(i, j, k) & \text{si } x < p \\ A & \text{si } p \leq x < q \\ B & \text{si } q \leq x < 1 \end{cases} \quad (7.3)$$

donde $x \sim U[0,1]$, $0 \leq p \leq q \leq 1$, A y B son número fijos, usualmente $A=255$ y $B=0$. Por otro lado, $(q-p)$ es la probabilidad de pasar al valor A (255 y por tanto sal) y $(1-q)$ es la probabilidad de pasar a B (0 y por tanto pimienta), p es la probabilidad de permanecer igual.

Al aplicarle un filtro de mediana, el ruido desaparece casi por completo mientras que con un pasa bajos, cada defecto (pixel erróneo) se ‘lava’, suavizándose, pero al mismo tiempo aumenta de tamaño. Es por ello que en este caso siempre resultará más conveniente aplicar primero el filtro de mediana antes de procesar los datos, ya que como se podrá ver en las pruebas efectuadas, el esquema propuesto empleando la HT rotada el ruido no desaparece, sólo se reduce un poco, en general en mucho menor proporción de lo que puede reducir al ruido Gaussiano (ver Figura 7.15). Por esta razón ninguna de las métricas calculadas aporta algún indicador que refleje la calidad visual en todos los experimentos, pues este ruido deja valores extremos en las imágenes fusionadas que alteran las mediciones.

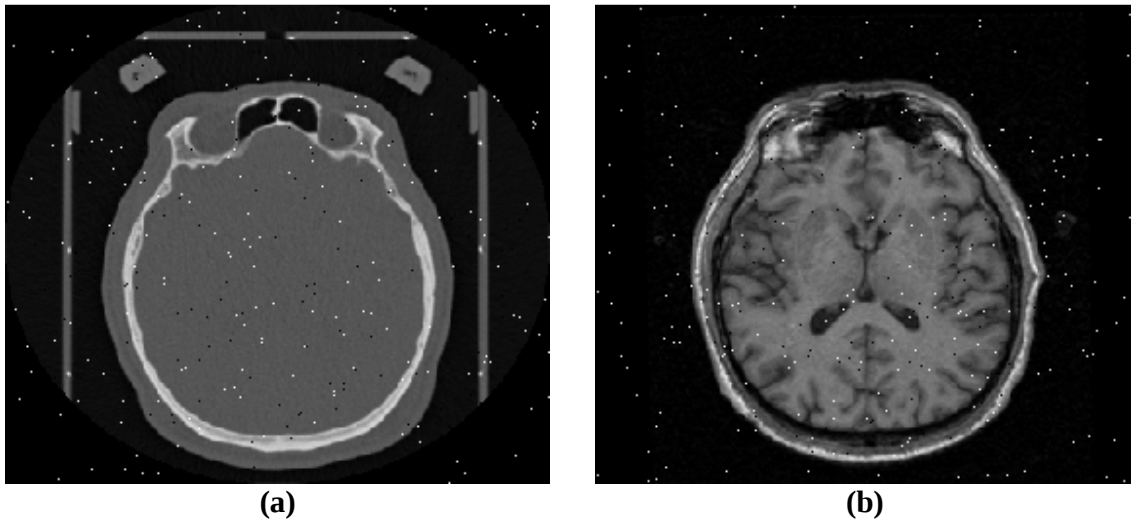


Figura 7.18 Comparación de fusión en imágenes médicas degradadas con ruido sal y pimienta para una densidad de 0.005 empleando distintos métodos de fusión. (a) y (b) corte 45 de los volúmenes de CT y MR contaminados con ruido.

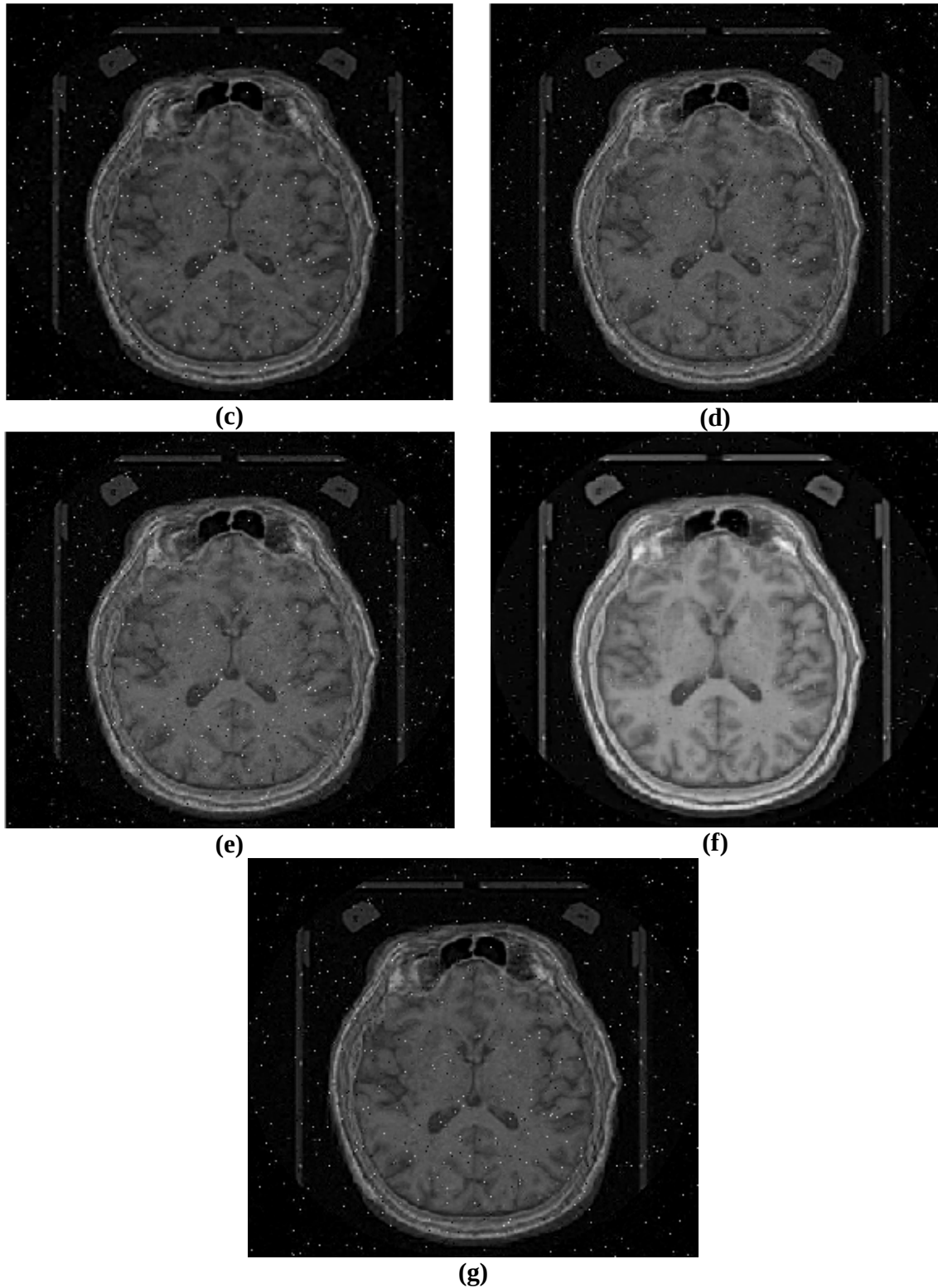


Figura 7.19 (Cont.) Comparación de fusión en imágenes médicas degradadas con ruido sal y pimienta para una densidad de 0.005 empleando distintos métodos de fusión. (c) fusión con DWT biortogonal 3.3, (d) fusión con DWT coiflet 5, (e) fusión con DWT Daubechies 6, (f) fusión con DHT, (g) fusión con DWT symlet 7.

7.3. Comparación con otras técnicas de fusión en 2D

Los sistemas de fusión de imágenes tienen como propósito transferir información ‘interesante’ de las imágenes de entrada a la imagen fusionada. La suposición común para la mayoría de las aproximaciones es la existencia de una imagen de referencia de buena calidad para todas las regiones en las imágenes de entrada [75].

Dada la ausencia de una imagen de referencia en el caso de las imágenes médicas, de manera similar a la fusión de volúmenes, se creó una aplicación de fusión de imágenes 2D que permitiera comparar el esquema propuesto con otras técnicas de representación de imágenes más recientes que la transformada wavelet, donde además pudiera evaluarse con métricas objetivas tomando como referencia una imagen ideal creada sintéticamente. La funcionalidad en cuanto a la posibilidad de aplicar diferentes tipos de ruido a las imágenes es la misma, la importancia de mostrar algunos resultados obtenidos a partir de ella radica en la adición de los algoritmos para realizar fusión mediante la transformada contourlet (CT) [30] y la transformada curvelet (CWT) [109], cada una de las cuales puede ser configurada de diversas formas.

En la Figura 7.16 se muestra la interfaz de dicha aplicación donde las opciones de configuración corresponden al método de fusión de la transformada de Hermite multirresolución, en esta imagen se puede notar que para la selección de reglas de fusión que requieren de la información de los píxeles adyacentes al píxel procesado es posible elegir el tamaño de la vecindad, lo cual no es posible en la aplicación 3D donde todas las pruebas experimentales se realizaron sobre una vecindad tridimensional de 27 vecinos.

Las pruebas expuestas en esta sección se realizaron sobre conjuntos de imágenes multifocales y algunos pares de imágenes médicas, por lo que la evaluación de los experimentos se ha dividido en dos grupos:

- Fusión de imágenes multifocales con referencia (*ground truth*).
- Fusión de imágenes médicas sin referencia.

Con el propósito de comparar cuantitativamente las diversas métricas objetivas, se emplearon varios métodos, incluyendo las métricas más tradicionales así como otras más recientes que poseen una mejor correlación con la evaluación de la percepción humana. Entre las primeras, se consideró la relación señal a ruido de pico (PSNR) y el error cuadrático medio (MSE); para el segundo grupo se empleó la métrica de similitud estructural (SSIM), la información mutua (MI). En experimentos sin imagen de referencia se emplearon únicamente las métricas basadas en la información mutua. Todos estos criterios fueron explicados ampliamente en el capítulo 5.

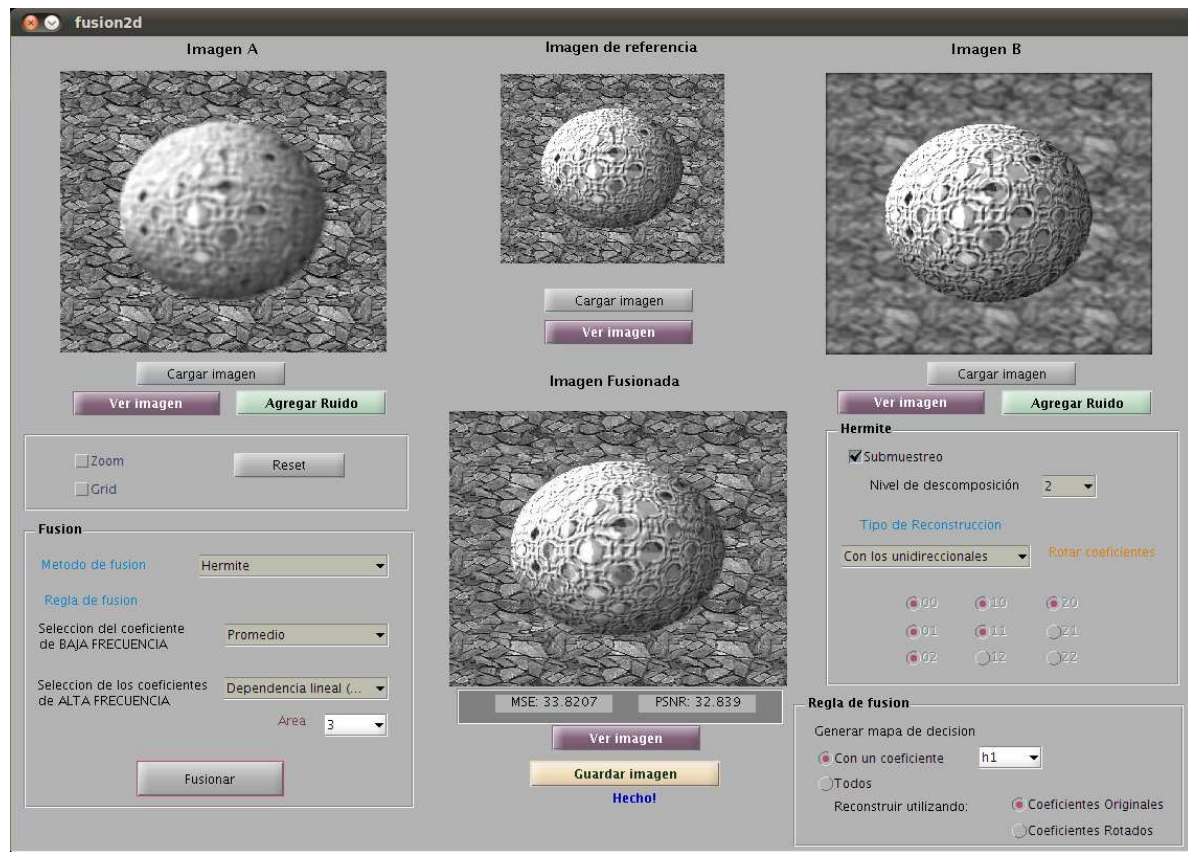


Figura 7.20 Aplicación de fusión del método propuesto en 2 dimensiones.

7.3.1. Fusión de imágenes multifocales con referencia

Las imágenes multifocales son complementarias en diversas formas y una sola no es suficiente en términos de su respectivo contenido de información, por lo que observar una serie de dichas imágenes de forma separada e individualmente no es útil ni conveniente. La ventaja de los datos multifocales puede ser explotada fusionando las regiones más nítidas y no desenfocadas de las diferentes imágenes de entrada para obtener una sola imagen, aprovechando la sobrerepresentación de los datos.

En total se dispuso de 6 conjuntos de imágenes multifocales que fueron creadas a partir de una imagen con todos los objetos enfocados (*ground truth*), a la cual se le emborronaron ciertos objetos, algunos de estos datos fueron proporcionados por el Instituto de Óptica del Consejo Superior de Investigaciones Científicas en España [27] y otras fueron cortesía del Dr. Shutao Li [62]. Para estos datos se realizó la fusión de imágenes aplicando todos los métodos de fusión disponibles en la aplicación con el mismo esquema de fusión ajustando los parámetros de tal forma que las condiciones fueran casi las mismas en todos los casos. En todos los experimentos se tomaron dos niveles de descomposición, las reglas de fusión fueron el promedio para los coeficientes de aproximación y la dependencia lineal para los coeficientes de detalle, esta última considerando una vecindad de 3 x 3. La transformada

wavelet empleada fue Daubechies 4, en el caso de la transformada curvelet se analizaron 8 direcciones, y en relación a la transformada contourlet se empleó la transformada McClellan de filtros 9-7 como filtros direccionales y la wavelet Daubechies 4 como filtro piramidal.

En la Tabla 7.6 se muestra el promedio de los resultados obtenidos de todos los experimentos realizados con este tipo de imágenes. Es importante mencionar que en este caso la información mutua se refiere a la que existe entre la imagen fusionada y la imagen de referencia.

Tabla 7.6 Resultados del promedio de todos los resultados de los experimentos efectuados con imágenes de referencia. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI a la información mutua entre la imagen fusionada y la imagen de referencia.

Método de fusión	MSE	PSNR [dB]	SSIM	MI
Transformada de Hermite	127.0548	36.4254	0.9639	4.2285
Transformada Wavelet	173.6760	31.0859	0.9502	3.4610
Transformada Contourlet	177.4663	30.8361	0.9477	3.3979
Transformada Curvelet	164.2967	31.6076	0.9496	3.4600

De los resultados mostrados en la Tabla 7.6 es posible observar una clara ventaja de la técnica propuesta respecto a otras técnicas de representación de imagen empleadas en la fusión. El MSE y la PSNR reflejan una mayor similitud con la imagen de referencia en términos de error, adicionalmente el SSIM demuestra que no sólo hay se consigue el menor error sino que la calidad de la imagen en cuanto a sus estructuras es superior. Por otro lado, la información mutua también coincide con los resultados de las otras métricas indicando que se logra una mejor conservación de la información de las imágenes originales.

De estos experimentos a continuación se exponen de manera detallada sólo dos ejemplos.

Pelota y fondo.

En este experimento se consideraron dos imágenes en escala de grises de tamaño 512×512 . La Figura 7.17 (a) es la imagen ground truth o de referencia con el fondo y la pelota enfocados proporcionada por el Instituto de Óptica del Consejo Superior de Investigaciones Científicas en España [27]. La Figura 7.17 (b) tiene el fondo desenfocado y la pelota en foco y la Figura 7.17 (c) tiene el fondo en foco pero la pelota esta desenfocada. Los resultados de este experimento se encuentran en la Figura 7.18 y las métricas obtenidas están en la Tabla 7.7 donde se corrobora que el método con el mejor desempeño es el que utiliza la transformada de Hermite.

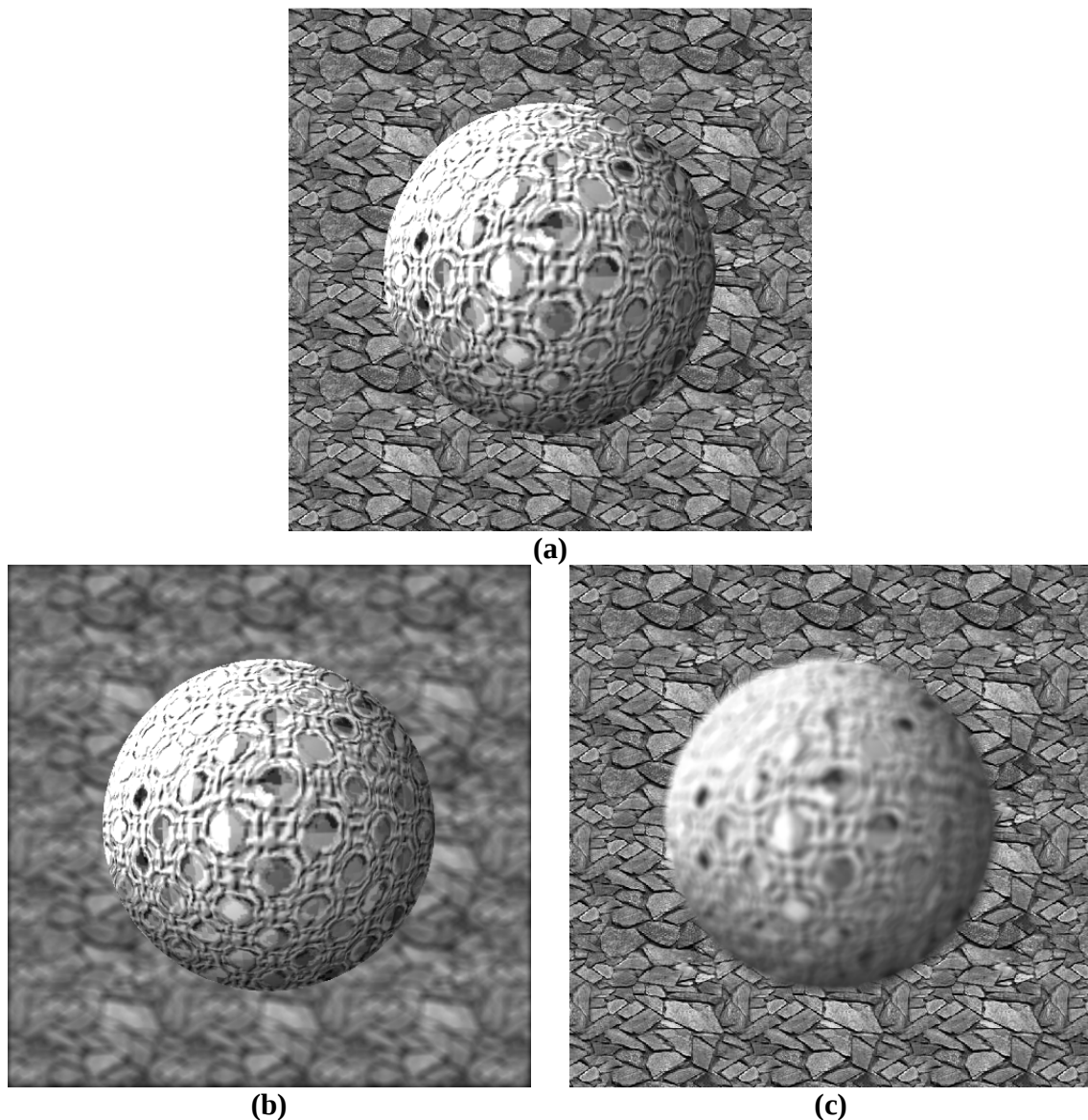


Figura 7.21 Imágenes de entrada multifocales para el experimento de Pelota y fondo. (a) Imagen de referencia. (b) Imagen con el fondo desenfocado y la pelota en foco. (c) Imagen con el fondo en foco y la pelota desenfocada.

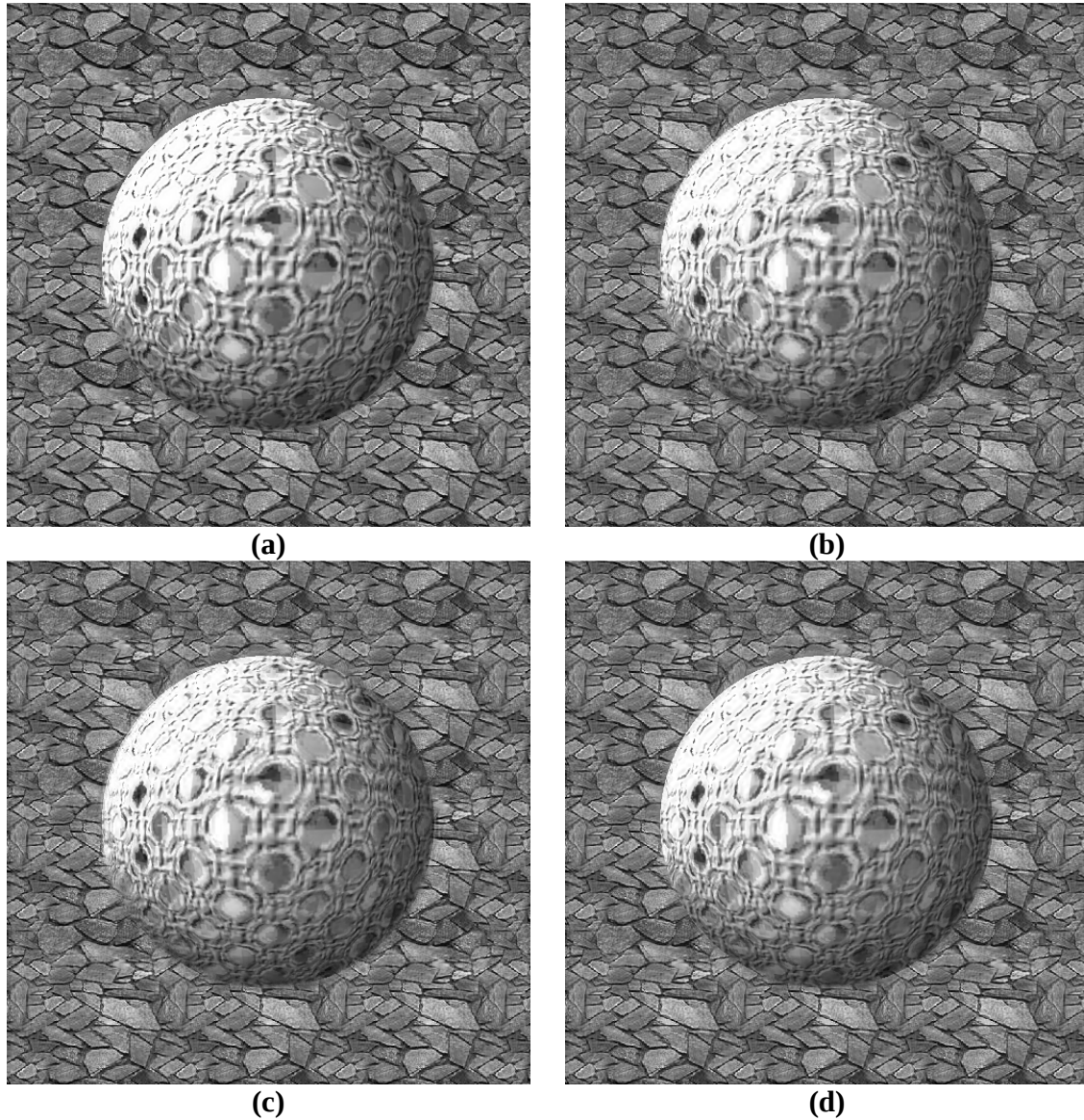


Figura 7.22 Resultados de fusión de imágenes multifocales para el experimento de Pelota y fondo. (a) Imagen obtenida empleando HT, (b) Imagen obtenida empleando WT, (c) Imagen obtenida empleando CT y (d) Imagen obtenida empleando CWT.

Tabla 7.7 Resultados del desempeño de la fusión para el experimento Pelota y fondo empleando diferentes métodos de fusión. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI a la información mutua entre la imagen fusionada y la imagen de referencia.

Método de fusión	MSE	PSNR [dB]	SSIM	MI
Transformada de Hermite	33.8207	32.8389	0.9895	3.6596
Transformada Wavelet	128.5902	27.0387	0.9532	2.5435
Transformada Contourlet	131.6418	26.9368	0.9455	2.5282
Transformada Curvelet	114.9822	27.5244	0.9525	2.5883

Líneas

Para este experimento se consideraron dos imágenes multifocales en escala de grises de tamaño 512×512 . La Figura 7.19 (a) es la imagen de referencia proporcionada por el Instituto de Óptica del Consejo Superior de Investigaciones Científicas en España [27]. Las Figuras 7.19 (b) y 7.19 (c) son las imágenes con distintas regiones desenfocadas. Los resultados de este experimento se muestran en la Figura 7.20 donde es visible la introducción de artefactos en los bordes de las líneas en todos los métodos a excepción del que emplea la transformada de Hermite. En la Tabla 7.8 se condensan las métricas calculadas para estas últimas imágenes donde nuevamente se demuestra la superioridad del método propuesto.

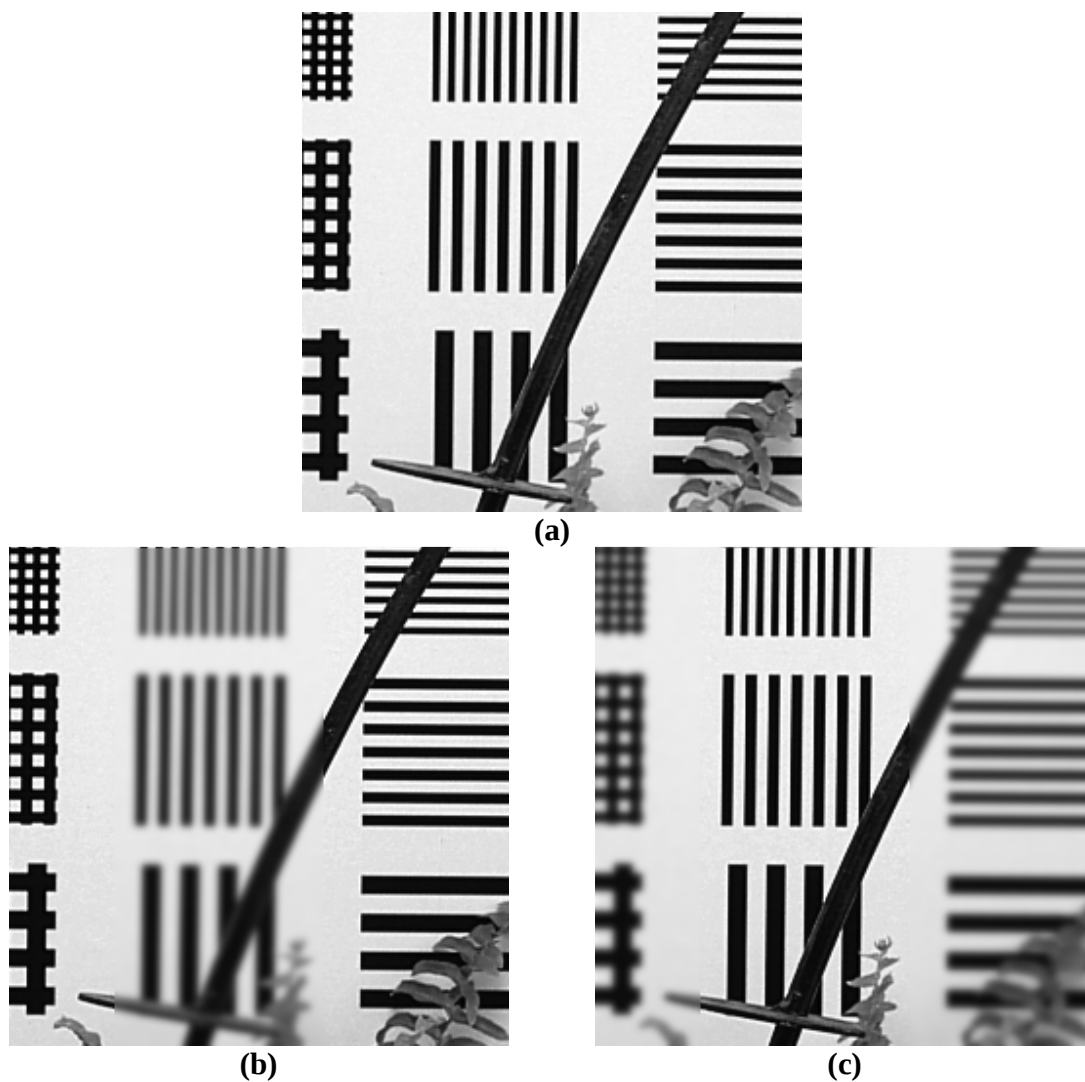


Figura 7.23 Imágenes de entrada multifocales para el experimento de Líneas. (a) Imagen de referencia. (b) Imagen con la franja central desenfocada. (c) Imagen con las franjas laterales desenfocadas.

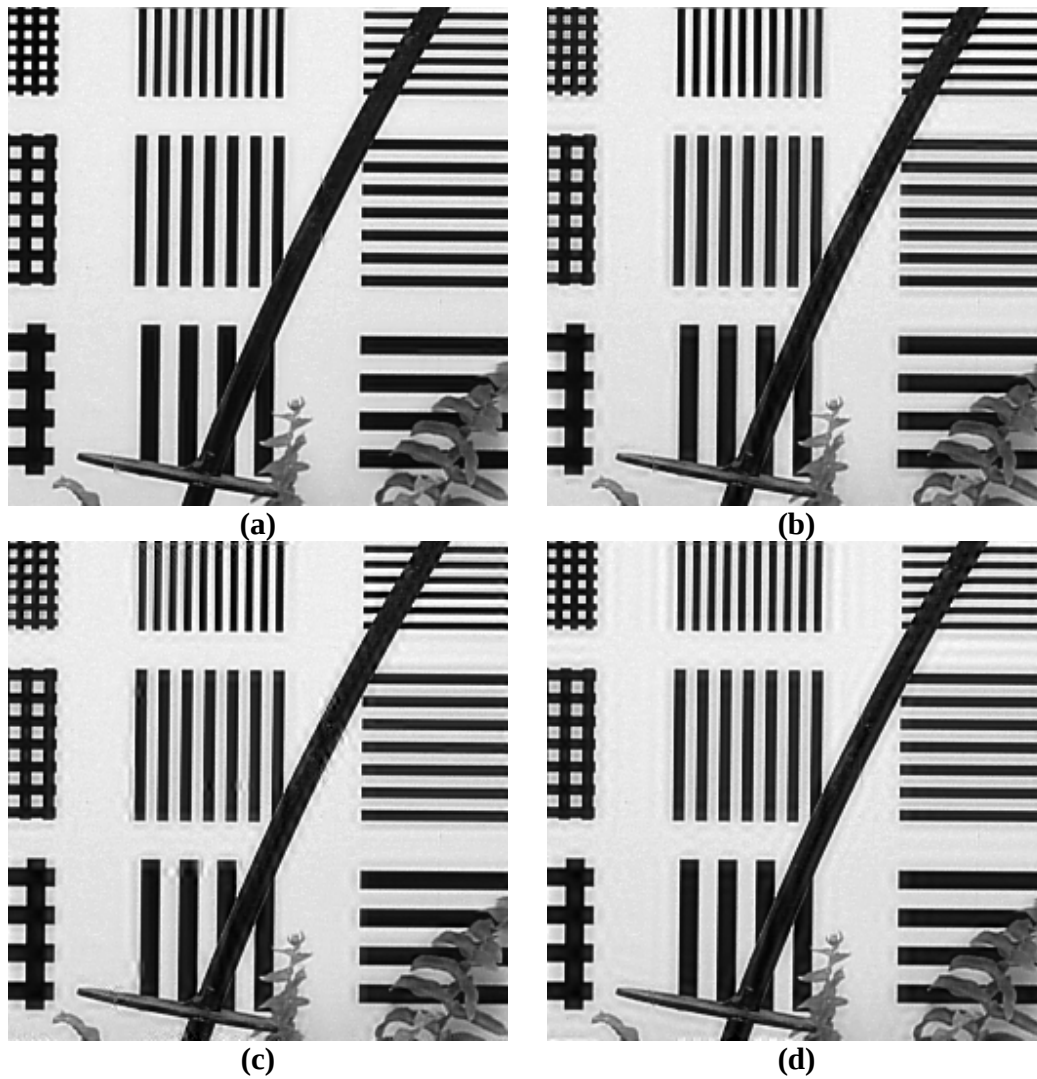


Figura 7.24 Resultados de fusión de imágenes multifocales para el experimento de Líneas. (a) Imagen obtenida empleando HT, (b) Imagen obtenida empleando WT, (c) Imagen obtenida empleando CT y (d) Imagen obtenida empleando CWT.

Tabla 7.8 Resultados del desempeño de la fusión para el experimento de Líneas empleando diferentes métodos de fusión. MSE son las siglas en inglés del error cuadrático medio, PSNR corresponde a la relación señal a ruido, SSIM al índice de similitud estructural y MI a la información mutua entre la imagen fusionada y la imagen de referencia.

Método de fusión	MSE	PSNR [dB]	SSIM	MI
Transformada de Hermite	140.0778	26.6671	0.9765	3.1197
Transformada Wavelet	265.5801	23.8888	0.9433	2.4353
Transformada Contourlet	277.0684	23.7049	0.9398	2.3470
Transformada Curvelet	254.7785	24.0691	0.9364	2.3749

7.3.2. Fusión de imágenes médicas (sin referencia)

Para los experimentos con imágenes médicas se emplearon algunos cortes de los primeros dos volúmenes descritos al inicio de este capítulo y un par de imágenes de CT y MR utilizadas ampliamente en diversos artículos de fusión, las cuales son cortesía del Dr. Oliver Rockinger (ver Figura 7.22). La configuración de cada uno de los métodos de fusión es la misma que en los experimentos realizados en imágenes multifocales por lo que sólo se presentan dos ejemplos de estas pruebas ambos con fusión de modalidades de CT y MR, el primero sin ruido y el segundo con la introducción de ruido Gaussiano con $\sigma=0.001$.

Fusión de imágenes sin ruido

En la Figura 7.21 (a) y 7.21 (b) se muestran los cortes 70 del par de volúmenes de CT/MR empleado en las pruebas de fusión 3D. Las imágenes restantes de esta figura exponen los resultados de aplicar los distintos métodos observándose que a pesar del manejo direccionales de las transformadas contourlet y curvelet, el mejor resultado lo presenta la transformada de Hermite, cuya imagen muestra buen contraste y una excelente conservación de las estructuras sobresalientes de los detalles de encontrados en la corteza cerebral que proporciona la imagen MR. Los demás métodos además de tener un contraste más pobre introducen también artefactos que podrían afectar la correcta interpretación de las imágenes médicas. Asimismo, en la Tabla 7.9 se puede comprobar que la HT conserva mayor la mayor cantidad de información de las imágenes de entrada.

Tabla 7.9 Resultados del desempeño de la fusión las imágenes correspondientes al corte 70 de los volúmenes de CT y MR, donde MI_{FA} es la información mutua entre la imagen fusionada y la imagen A (CT), MI_{FB} es la información mutua entre la imagen fusionada y la imagen B (MR) y MI_{FAB} es la información muta total.

Método	MI_{FA}	MI_{FB}	MI_{FAB}
Transformada de Hermite	1.9378	1.2987	3.2366
Transformada wavelet	1.8213	1.2022	3.0235
Transformada curvelet	1.8279	1.2683	3.0963
Transformada contourlet	1.7910	1.2121	3.0031

Fusión de imágenes ruidosas

En la Figura 7.24 se introdujo ruido gaussiano con $\sigma=0.001$ a las imágenes de entrada de CT y MR, cortesía del Dr. Oliver Rockinger [53], con el propósito de mostrar la eficiencia del método propuesto en presencia de ruido, donde a pesar de que la degradación es pequeña se observa que los métodos con los que se compara el desempeño de la HT son altamente sensibles al ruido preservando ese efecto indeseado en las imagines fusionadas.

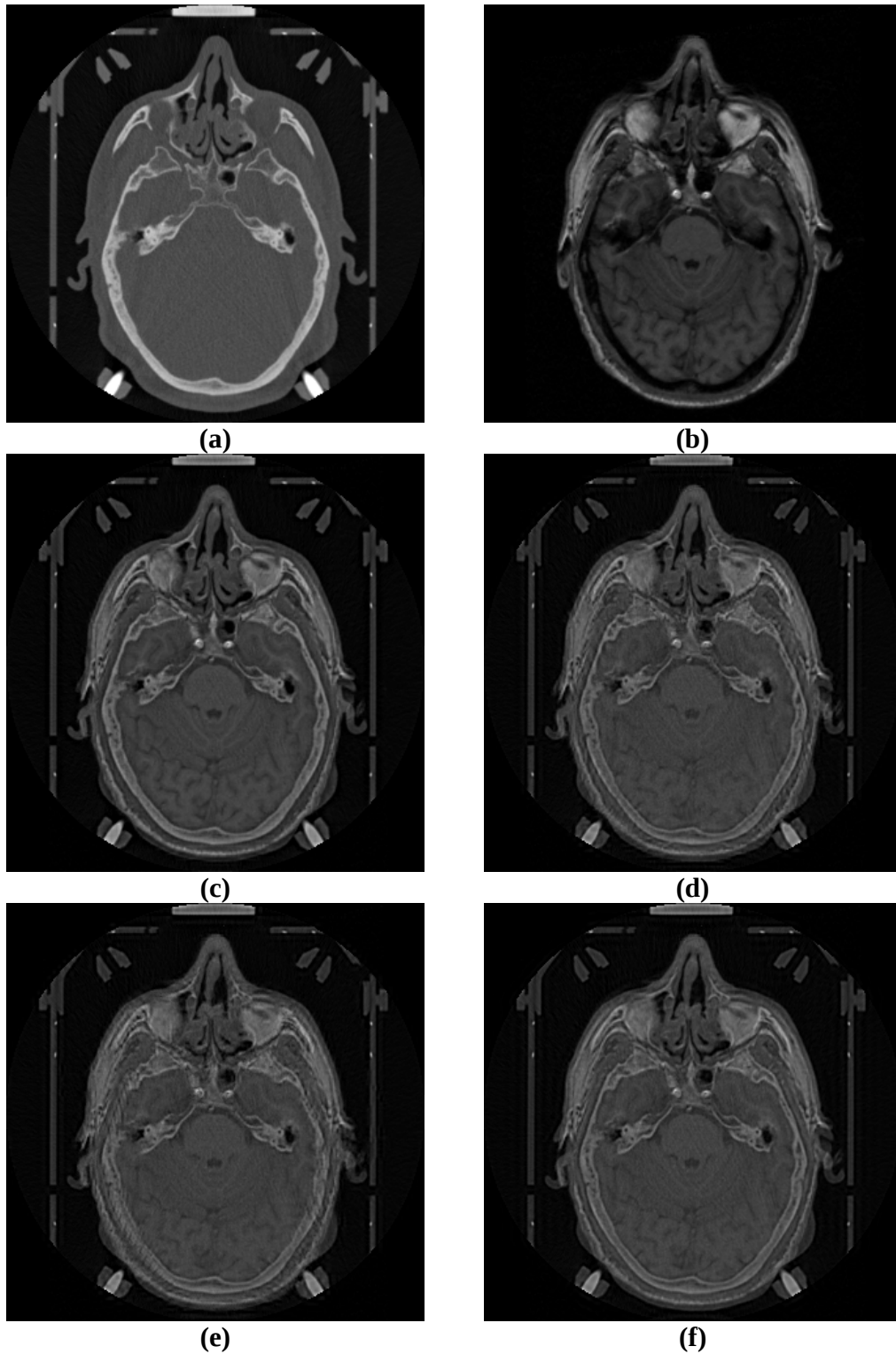


Figura 7.25 Resultados de fusión de (a) CT y (b) MR. (c) Imagen obtenida empleando HT, (d) Imagen obtenida empleando WT, (e) Imagen obtenida empleando CT y (f) Imagen obtenida empleando CWT.

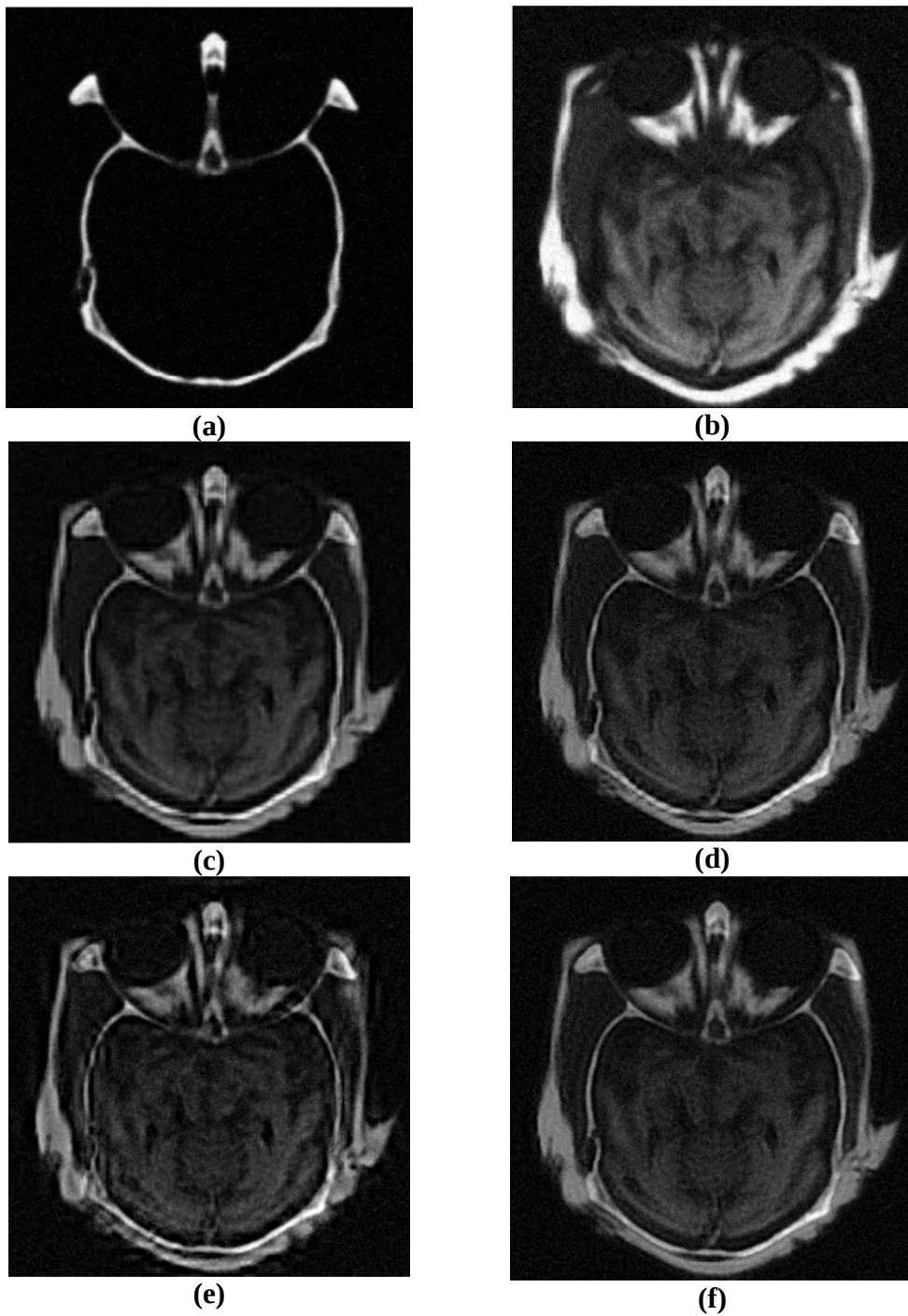


Figura 7.26 Resultados de fusión de (a) CT y (b) MR con ruido Gaussiano $\sigma=0.001$. (c) Imagen obtenida empleando HT, (d) Imagen obtenida empleando WT, (e) Imagen obtenida empleando CT y (f) Imagen obtenida empleando CWT.

En este trabajo se ha presentado un método nuevo de fusión de imágenes tridimensionales que emplea la transformada de Hermite rotada multirresolución en combinación con una regla de fusión basada en la evaluación de la dependencia lineal. Los experimentos realizados han mostrado que esta técnica no sólo presenta resultados aceptables para la integración de información de diversas modalidades como lo son CT/MR, MR/PET y MR/SPECT sino que también reduce la introducción de artefactos en los volúmenes fusionados, incluso en presencia de ruido.

Esta versatilidad del método para la fusión de más de más de dos tipos de modalidades representa una ventaja frente a otros métodos que sólo funcionan para cierta aplicación, pero puede verse también como una desventaja si se considera que en aplicaciones que involucran procesamiento con imágenes médicas se requiere de mucha precisión y usualmente se tiende a la especialización para ciertos tipos de integración de modalidades. En este sentido, se puede decir que los resultados de fusión mediante el esquema propuesto presentan buenos resultados de manera general, pero su área de aplicación sería la fusión de CT/MR porque fue en este tipo de experimentos donde se obtuvieron mejores resultados en cuanto a la conservación y nitidez de las estructuras sobresalientes presentes en las imágenes originales; en el caso de la fusión entre imágenes anatómicas con funcionales no es tan evidente la ventaja que se obtiene al aplicar el método, la interpretación de las imágenes funcionales no es tan sencilla y en este caso es necesaria la evaluación por parte de los médicos que las utilizan.

Las evaluaciones de la fusión de imágenes médicas tridimensionales fueron comparadas con distintos tipos de wavelet a fin de poder generalizar el desempeño entre ambas transformadas en lugar de limitarse a un tipo de wavelet o familia de wavelets específica. A menudo, un argumento a favor de las wavelets es que no existe una transformada wavelet única que resuelva todos los problemas ya que es a partir de la modelación del proceso y de un análisis a priori del tipo de señal tratada y del objetivo que se pretenda (compresión, eliminación del ruido, etc.) que se hace la búsqueda de la familia de wavelets que mejor coincida con las características de la señal a estudiar. Sin embargo en el proceso de fusión en tres dimensiones con o sin ruido, ninguna de las wavelets analizadas pudo superar los resultados alcanzados por la transformada de Hermite, pues esta última ha demostrado ser un modelo eficiente modelo de representación de imágenes debido a que las derivadas de Gaussiana son las funciones base de esta transformada, y estas detectan, representan y reconstruyen de manera optima los patrones relevantes de las imágenes, tales como líneas y bordes.

Por otro lado, el esquema de fusión propuesto sólo se limitó al proceso de fusión, asumiendo la inexistencia de problemas de registro en los datos empleados en todos los experimentos. Esto simplificó el diseño del método pero deja cierta incertidumbre en la evaluación de resultados ya que no se tiene el control del registro y en consecuencia se desconoce el margen de error de cada volumen utilizado en los experimentos. Por esta razón y con el propósito de comparar el método propuesto con otras transformadas discretas que analizan la direccionalidad de los datos, se incluyeron los resultados de la programación en dos dimensiones.

De esta manera, aunque el objetivo principal de la investigación se enfocó a la aplicación de la fusión en imágenes médicas, también se efectuaron algunos experimentos con imágenes multifocales creadas a partir de una imagen con todos los objetos enfocados (*ground truth*), a la cual se le emborraron ciertas áreas, obteniéndose excelentes resultados ya que tanto las evaluaciones subjetivas como las objetivas muestran la superioridad del método respecto a otras técnicas como la transformada wavelet, la transformada contourlet y la transformada curvelet. Esta comparación en dos dimensiones también se realizó con algunas imágenes médicas donde nuevamente la transformada de Hermite mostró una superioridad cuantitativa y cualitativa frente a las demás transformadas.

En las dos aplicaciones de fusión desarrolladas se evaluó también el desempeño del esquema propuesto para fusión con reducción de ruido para lo cual se realizaron experimentos degradando las dos imágenes de entrada con ruido Gaussiano a distintas varianzas y ruido sal y pimienta a distintas densidades. Los resultados mostraron una notable mejora en la calidad visual respecto a todos los métodos empleados, y esto es debido a que la transformada de Hermite rotada como algoritmo en fusión de imágenes, tiene como ventajas la compactación de su energía en algunos coeficientes y su adaptabilidad al contenido local de la imagen, lo cual permite que en el proceso de reconstrucción se ignore gran parte del ruido conservando las características de los bordes en las imágenes.

Además de analizar las diferencias obtenidas en los resultados al variar los métodos de fusión, también se hizo una comparación con otras reglas de fusión como la selección del valor absoluto del máximo, selección del máximo con verificación de consistencia y saliente y medida de paridad [15]. Para este propósito se realizaron en primera instancia algunas pruebas con imágenes multifocales y otras con los volúmenes de imágenes médicas de donde se deduce que si bien no tiene resultados que representen una superioridad apreciable desde el punto de vista perceptivo, si es una regla competitiva respecto a otras reglas ampliamente utilizadas en los algoritmos de fusión ya propuestos.

Algoritmos empleados en fusión de imágenes médicas

En la siguiente tabla se resumen los algoritmos de fusión más novedosos empleados en investigaciones con imágenes médicas.

Tabla A.1 Algoritmos empleados en fusión de imágenes médicas.

Algoritmo	Descripción	Aplicación
Fusión multirresolución morfológica [72] Matsopoulos, G.; Marshall, S. & Brunt, J (1994)	Se construye una pirámide morfológica de cada imagen de entrada y de éstas se obtienen otras mediante diferencias entre los niveles con las cuales se obtiene una pirámide que combina la información de ambas imágenes. Finalmente, para obtener la imagen fusionada se realiza el proceso inverso.	MR y CT de imágenes del cerebro humano.
Fusión por conjuntos difusos y lógica probabilística [24] Colin, A. y Boire, J.-Y (1999)	Se propone el uso de la teoría de probabilidad como un modelo para crear una imagen fusionada entre una modalidad anatómica y una funcional. Para ello se extraen 4 clases de tejidos: materia blanca (WM), materia gris (GM), fluido cerebroespinal + fondo (CSF), y materia gris hipo-perfusionada (HGM), este último intenta caracterizar las áreas patológicas. Cada clase (excepto la HGM) se modela con un mapa difuso, el cual se representa mediante una distribución de probabilidad de los tejidos en la imagen. En dichos mapas se le asigna a cada voxel un valor de pertenencia a cada clase. El mapa para la clase HGM se define durante el proceso de fusión. El proceso de fusión utiliza un operador variable, pues para cada voxel, se calcula el tejido dominante para cada modalidad (esto es, el tejido con la probabilidad más alta). Para cada una de las nueve posibles combinaciones de tejidos dominantes, el operador tiene un comportamiento específico. Una vez que el proceso de fusión se completa, las probabilidades resultantes se utilizan como una base para el etiquetado de imágenes. Para sintetizar la	Imágenes cerebrales 3D funcionales de alta calidad MR/SPECT

	imagen funcional de alta resolución, se multiplica el porcentaje de volumen parcial por el promedio del valor de gris del tejido para una modalidad específica.	
Fusión por conjuntos difusos y lógica probabilística con operadores basados en teoría de la información [7] Barra, V. y Boire, J.-Y (2001)	Esta propuesta se basa en la de Colín y Boire de 1999, la diferencia radica en el operador empleado para la fusión de los mapas difusos de las clases, que en este caso se fundamentan en la teoría de la información. Tal es el caso de los operadores MIM (de información mutua) que de acuerdo al conocimiento de expertos actúa como un operador variable en el proceso de fusión.	Se reportó su aplicación a 2 casos clínicos: el estudio de la Alzheimer mediante fusión de MR/SPECT y el estudio de epilepsia mediante la fusión de MR/SPECT/PET El algoritmo [7] también es utilizado por SEGAMI
Fusión de imágenes médicas volumétricas multimodales [3] Mario Aguilar y Joshua R. New (2002)	La propuesta tiene como motivación los principios neurofisiológicos del sistema de visión humano, el cual está conformado por tres tipos de sensores de luz conocidos como conos que contienen sensibilidades traslapadas (con 3 longitudes de onda distintas). El proceso de fusión consiste en analizar las imágenes de entrada utilizando kernels de convolución 3D para determinar localmente parámetros de fusión relevantes. Luego las bandas resultantes se mapean al espacio de color YIQ y finalmente a RGB	MR y SPECT
Fusión de imágenes médicas multimodales a nivel características en el dominio de la transformada Wavelet Lifting [59] Sudipta Kor y Umashanker Tiwary (2004)	El algoritmo consiste en descomponer las imágenes de entrada usando la transformada wavelet Lifting y luego se calcula el modulo del gradiente de los coeficientes obtenidos. La fusión está basada en la umbralización del valor de dicho gradiente. Finalmente para obtener la imagen fusionada se aplica la transformada wavelet <i>lifting</i> inversa.	MR T1 y T2
Fusión multiescala a nivel pixel [8] Ben Hamza, A.; He, Y.; Krim, H. & Willsky, A (2005)	Se utiliza la medida de actividad de los coeficientes de una descomposición wavelet biortogonal para generar un mapa de decisión aplicando la divergencia Jensen-Renyi basándose en la entropía de la información. El mapa de decisión es segmentado en dos regiones,	Es un método general ya que lo han aplicado a imágenes desenfocadas, CT y MR

A. Algoritmos empleados en fusión de imágenes médicas

	una en donde los pixeles tienen patrones de actividad similares en todas las imágenes de entrada y otra donde esos patrones son diferentes. La regla de fusión es en realidad una solución a un problema de optimización donde lo anterior es tomado en cuenta.	
Fusión mediante transformada Hermite Rotada [39] Boris Escalante (2008)	El algoritmo consiste en obtener la transformada de Hermite rotada de las imágenes originales y aplicar las reglas de selección de coeficientes para generar el conjunto de coeficientes únicos que originarán la imagen fusionada mediante la transformada de Hermite inversa La regla de fusión para los coeficientes de baja frecuencia es el promedio y para los de alta frecuencia es la selección de máximo con verificación de consistencia.	CT/MR MR/PET
Transformada curvelet rápida discreta y transformada wavelet [60] Y. Kiran Kumar (2009)	Se hace la fusión de imágenes a diferentes resoluciones, para lo cual se emplea primero un método de registro en un solo paso luego se aplica una descomposición multiescala de la imagen mediante la transformada curvelet o wavelet La regla de fusión es la selección del máximo absoluto en las escalas finas y el promedio en las escalas gruesas.	PHILIPS HealthCare Se han fusionado MR/ CT y MR/MR de datos pre-operatorios.
Fusión empleando la transformada wavelet Discreta Redundante (RDWT) [104] Richa Singh, Mayank Vatsa, Afzel Noore (2009)	Como primer paso se propone un algoritmo de registro no lineal basado en la maximización de la información mutua normalizada. Una vez que las imágenes han sido registradas, la fusión de imágenes multimodal se lleva a cabo descomponiendo las imágenes en 3 niveles de RDWT. Para preservar las características de ambas imágenes, los coeficientes de aproximación se promedian y para las bandas de detalle, cada banda es dividida en bloques de 3x3 y la entropía de cada bloque se calcula con la siguiente expresión $e_i^{jk} = \ln \sqrt{\frac{\left(\mu_i^{jk} - \sum_{x,y=1}^{3,3} I_i^{jk}(x,y) \right)^2}{\sigma_i^{jk}}} / m^2$ donde j denota las subbandas, $m=3$ el tamaño de cada bloque, k representa el número de bloques, e i ($=1,2$) es usado para diferenciar las dos imágenes multimodales I_1 e I_2. μ_i^{jk} y σ_i^{jk} son la media y la desviación estándar de los coeficientes RDWT del bloque k-ésimo de la subbanda j-ésima de la imagen i-ésima. Con estos valores de entropía, las subbandas de detalle de la imagen fusionada se	MR cerebrales, tipos T1 y T2. Los experimentos fueron realizados en la base de datos BrainWeb.

	generan comparando la entropía entre bloques, si el bloque de la imagen 1 es más grande que el bloque correspondiente a la imagen 2, el primer bloque es seleccionado y viceversa.	
Fusión de imágenes médicas a nivel características empleando ICA [18] Calhoun, V. & Adali, T (2009)	Las imágenes primero son procesadas para extraer ciertas características de interés, para lo cual se considera que una característica es un subconjunto de datos relacionado a una estructura o actividad cerebral seleccionada. Después las características son analizadas mediante jICA, análisis de componentes independientes conjunto. Dados dos conjuntos de datos, X_F y X_G, es posible concatenarlos para formar X_J y escribir la verosimilitud como $L(W) = \prod_{n=1}^N \prod_{v=1}^V p_{J,n}(u_J, v)$ donde $u_J = W_{XJ}$. Visto como un problema de máxima verosimilitud, se estima una matriz W tal que $L(W)$ sea máxima.	Imágenes cerebrales fMRI⁹/sMRI¹⁰ fMRI/ERP¹¹

⁹ fMRI mide la respuesta hemodinámica relacionada a la actividad neuronal del cerebro.

¹⁰ sMRI provee información sobre el tipo de tejido del cerebro.

¹¹ ERP: Even-related potential, es una respuesta electrofisiológica estereotipada a un estímulo interno o externo.

Transformada Curvelet

La transformada curvelet (TCW) es una transformada multiescala direccional desarrollada inicialmente en el dominio continuo por Candès y Donoho en 1999 [19, 33], la cual empleaba un filtrado multiescala seguido de una transformada de bloque ridgelet sobre cada banda de la imagen. Posteriormente, los autores propusieron la segunda generación de transformada curvelet [21], la cual fue definida directamente en el dominio de la frecuencia particionando el espacio sin emplear la transformada ridgelet. Ambas construcciones de wavelet requieren una operación de rotación y corresponden a particiones 2-D en la frecuencia basadas en coordenadas polares. Esto hace que su construcción sea simple en el dominio continuo pero dificulta su implementación en el caso discreto.

La TCW es también considerada una generalización dimensional superior de la transformada wavelet, diseñada para representar imágenes en diferentes escalas y ángulos. Una propiedad importante es que las curvas singulares pueden ser adecuadamente aproximadas con muy pocos coeficientes y en una manera no adaptable por ello su nombre “curvelets”. Además los autores Candès and Donoho han demostrado que la transformada propuesta está diseñada para representar contornos y otras singularidades alrededor de las curvas señaladas mucho más eficientemente que la transformada tradicional de wavelet. Comparada con la transformada wavelet, la transformada curvelet puede representar un contorno liso con muy pocos coeficientes con la misma precisión.

El algoritmo de la transformada curvelet de una imagen P puede resumirse en las siguientes etapas [5] y en el diagrama de la Figura B.1:

- a) La imagen P es dividida en tres subbandas Δ_1 , Δ_2 y P_3 empleando la transformada wavelet aditiva.

El propósito de esta etapa es descomponer la imagen en componentes aditivos, cada uno de los cuales es una subbanda de la imagen. Este proceso aísla los diferentes componentes de frecuencia de la imagen en diferentes planos sin aplicar submuestreo. Para ello se emplea el siguiente algoritmo ‘a trous’:

Dada una imagen P , es posible construir una secuencia de aproximaciones:

$$f_1(P) = P_1, \quad f_2(P_1) = P_2, \quad f_3(P_2) = P_3, \dots, f_n(P_{n-1}) = P_n \quad (1)$$

donde n es un entero preferentemente igual a 3. Para construir esta secuencia, se aplican convoluciones sucesivas con cierto filtro paso bajas. Los planos wavelets se calculan como diferencias entre dos aproximaciones consecutivas P_{l-1} y P_l , por ejemplo:

$$\Delta_l = P_{l-1} - P_l \quad (2)$$

Así, la reconstrucción curvelet está dada por:

$$P = \sum_{l=1}^{n-1} \Delta_l + P_n \quad (3)$$

- b) La partición en regiones traslapadas es realizada en las subbandas Δ_1 y Δ_2 . Estas regiones son pequeñas en dimensiones para transformar líneas curvas en pequeñas líneas rectas dentro de dichas subbandas. Esta parte del proceso es la que produce la mejora en el manejo de los bordes curvilíneos.
- c) La transformada ridgelet es realizada en cada región obtenida del paso anterior de las subbandas Δ_1 y Δ_2 . Dicha transformada pertenece a la familia de las transformadas discretas que emplean funciones base y es en sí misma una herramienta de detección de formas de los objetos en una imagen.

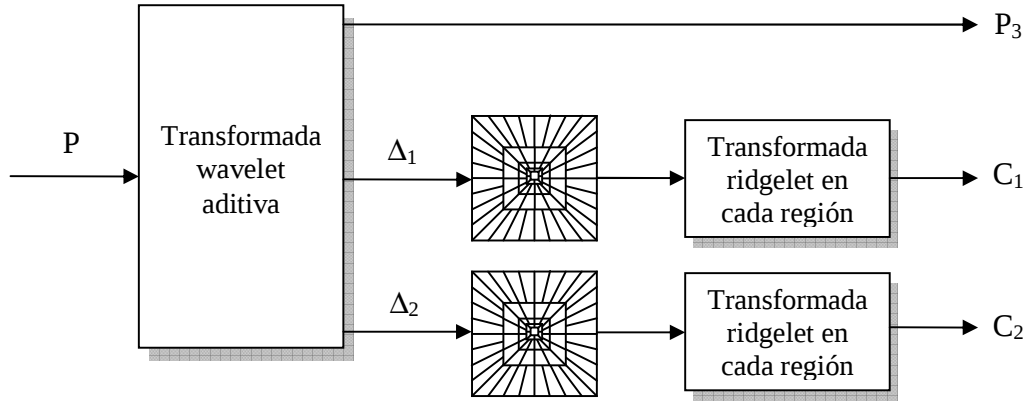


Figura B.1 Transformada curvelet discreta de una imagen P

La Figura B.2 muestra un ejemplo de descomposición de una imagen mediante la transformada curvelet, en ésta se eligieron dos niveles con 8 direcciones de ángulo en la resolución más baja.

Emmanuel Candès, Laurent Demanet, David Donoho y Lexing Ying, describieron en 2005 dos implementaciones digitales de la TCW, denominados la segunda generación de las transformadas curvelet en dos y tres dimensiones [20]. La primera transformación digital se basa en “unequally-spaced fast Fourier transforms (USFFT)”, mientras que el segundo en “wrapping of specially selected Fourier samples”. Las dos implementaciones esencialmente difieren por la selección de la cuadrícula espacial usada para traducir los curvelets en cada escala y ángulo.

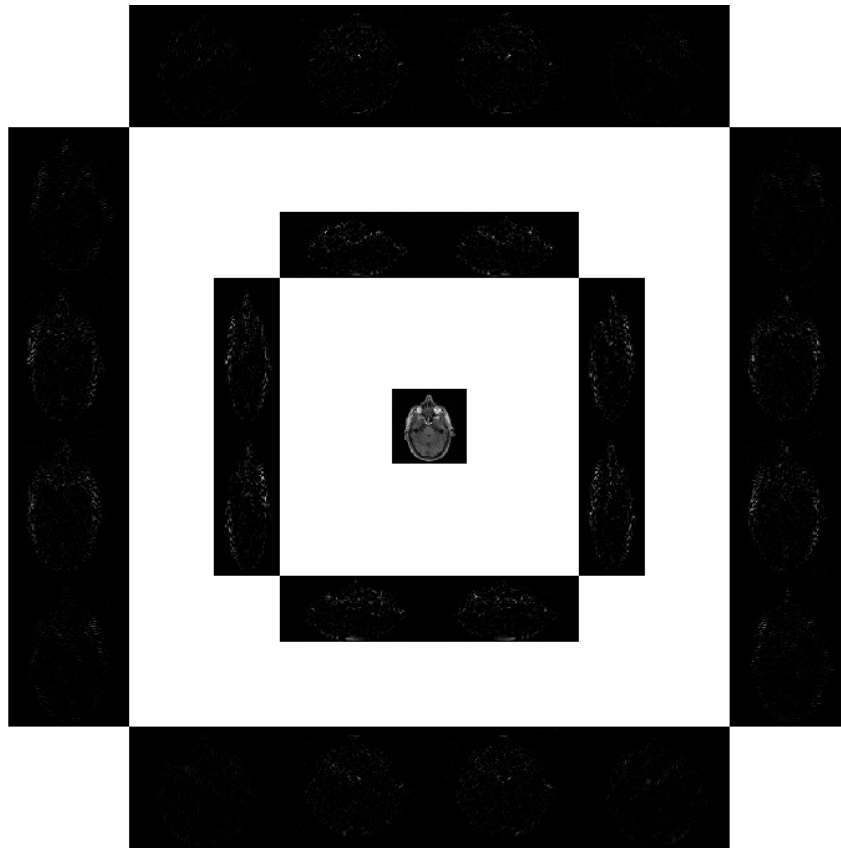


Figura B.2 Transformada curvelet a dos niveles de una imagen de resonancia magnética.

Ambas transformaciones digitales retornan una tabla de coeficientes digitales curvelet indexadas por un parámetro de escala, un parámetro de orientación y un parámetro de localización espacial. Y ambas implementaciones son rápidas ya que se ejecutan en $O(n^2 \log n)$ sobre un arreglo de n por n coeficientes. Dichos algoritmos se encuentran implementados y libres para su estudio en <http://www.curvelet.org> [109].

Esta segunda generación de TCW se ha empleado en diferentes aplicaciones de procesamiento de imágenes, exploración de datos sísmicos, mecánica de fluidos y solución de ecuaciones diferenciales parciales [64]. A pesar de su eficiencia en la representación de contornos, esta transformada aún presenta dos desventajas: 1) no es óptima para aproximaciones de características curvilíneas más allá de singularidades C^2 y, 2) la versión discreta es altamente redundante. Las implementaciones actuales de la transformada curvelet discreta promueven la reducción de redundancia, sin embargo aún no es apropiada para compresión de imágenes.

Transformada Contourlet

La Transformada Contourlet (CT:), creada en el año 2002 por Do y Vetterli, a diferencia de otras transformadas inicialmente definidas en un dominio continuo y que posteriormente se llevaron a un dominio discreto, se definió originalmente en un dominio discreto utilizando una estructura de bancos de filtros no separables, la cual, posteriormente se generalizó a un dominio continuo, mediante un mecanismo de análisis multirresolución [32]. La capacidad de esta transformada para capturar la estructura geométrica intrínseca de las imágenes reside en su propiedad de anisotropía que le permite una extracción mucho más eficaz de los detalles espaciales en diferentes direcciones así como en su auténtica naturaleza 2-D, frente a otras transformadas que se pueden definir como extensiones separables de transformadas 1-D, pues su análisis es multirresolución, local y direccional a través de segmentos de contorno.

La implementación de la TC implica realizar una transformada multirresolución que permita detectar los bordes, seguida de una transformada local direccional para detectar los segmentos de los contornos presentes en la imagen. Para ello se propuso una estructura definida por un doble banco de filtros: la transformada Laplaciana Piramidal (LP) y un banco de filtros direccionales (BFD), por lo cual es también conocida como Banco de Filtros Direccional Piramidal (BFDP) [31]. La LP permite la descomposición subbanda y captura los puntos aislados de la imagen, mientras que el BFD tiene como objetivo unir dichos puntos en estructuras lineales. De esta forma, la TC se puede considerar como una base de funciones con soporte elipsoidal con diferentes escalas, direcciones y relaciones de aspecto.

La descomposición LP en cada nivel genera una versión paso bajas submuestreada de la señal y la diferencia entre la original y la predicción con lo que resulta una imagen paso banda. Dado que el BFD está diseñado para capturar los componentes de alta frecuencia de la imagen de entrada, los cuales permiten representar direccionalidad, las bajas frecuencias de la imagen son eliminadas antes de aplicar el BFD.

La Figura C.1 muestra el banco de filtros de la TC usando una combinación de una LP y un DFB, primero se efectúa una descomposición multiescala en octavas mediante la LP, y posteriormente un BFD es aplicado a cada banda generada. El esquema puede ser iterado sobre la señal paso bajas submuestreada para obtener nuevamente una aproximación y subbandas direccionales de la imagen a otra resolución [32].

Una explicación más explícita de esta técnica de análisis en una imagen $a_0[\mathbf{n}]$ es la siguiente: la salida después de la etapa LP es J imágenes paso banda $b_j[\mathbf{n}]$, $j=1,2,\dots,J$ (en el orden de resoluciones de la más fina a la más gruesa) y una imagen paso bajas $a_J[\mathbf{n}]$. Esto significa que el nivel j -ésimo de la LP descompone la imagen $a_{j-1}[\mathbf{n}]$ en una imagen de mayor

resolución $a_j[\mathbf{n}]$ y una imagen de detalle $b_j[\mathbf{n}]$. Cada imagen paso banda $b_j[\mathbf{n}]$ es además descompuesta por un DFB en 2^{l_j} imágenes paso banda direccionales $c_{j,k}^{(l_j)}[\mathbf{n}]$, $k=0,1,\dots,2^{l_j}$.

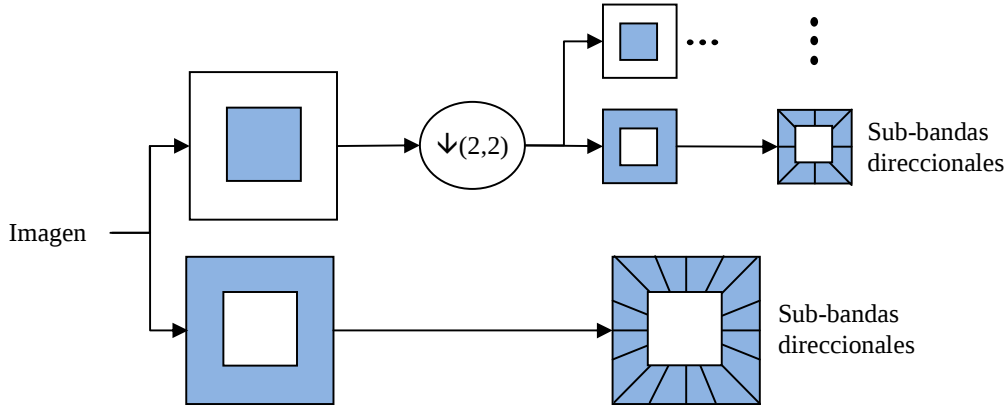


Figura C.1 Diagrama de bloques de la transformada contourlet.

Las principales propiedades de la transformada contourlet discreta están establecidas en el siguiente teorema.

Teorema 1. En un banco de filtros contourlet se cumple lo siguiente:

- 1) Si tanto la LP como el DFB usan filtros de reconstrucción perfecta, entonces la TC también logra una reconstrucción perfecta, lo cual implica que la TC provee un operador estructurado.
- 2) Si la LP y el DFB utilizan filtros ortogonales, la TC provee una estructura acotada, con límites iguales a 1.
- 3) La TC tiene un radio de redundancia menor a $4/3$.
- 4) Si se aplica un DFB de l_j niveles al nivel piramidal j de la LP, las imágenes base de la TC tendrán un soporte de ancho $\approx C2^j$ y largo $\approx C^{2j+l_j-2}$.
- 5) Empleando filtros FIR, la complejidad computacional de la TC es $O(N)$ para una imagen de N píxeles.

La Figura C.2 muestra la descomposición TC de una imagen de resonancia magnetica de 320×320 . Las imágenes de la parte más baja de dicha figura (rectángulos largos, subbandas finas) son el resultado de la descomposición piramidal y ocho direcciones, cada subbanda vertical tiene un tamaño de 160×80 , y cada subbanda horizontal un tamaño de 80×160 . Las imágenes de la parte media representan una descomposición piramidal y cuatro direcciones, cada una con tamaño de 80×80 . La imagen de la parte más alta es el resultado del filtro pasa bajas y tienen un tamaño de 80×80 .

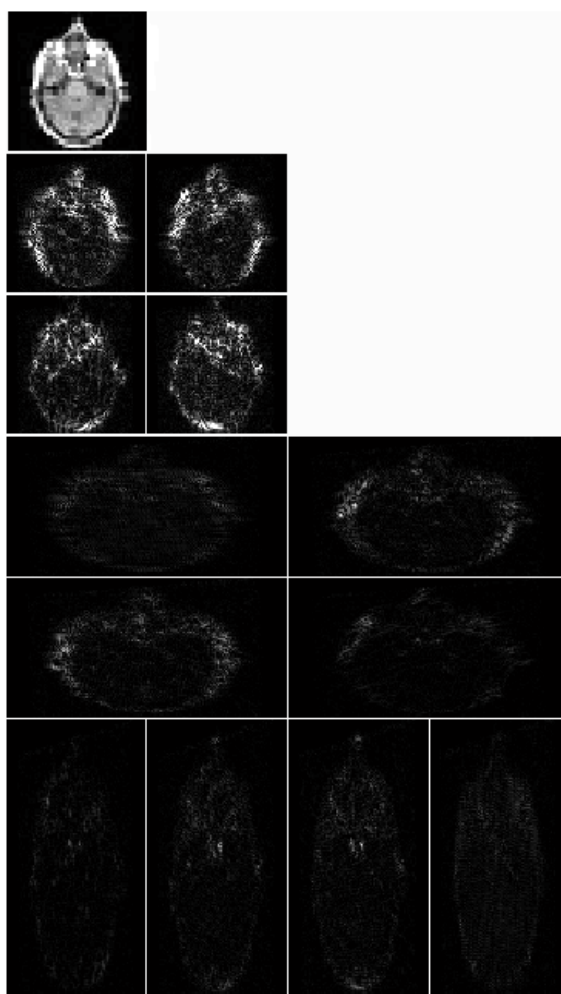
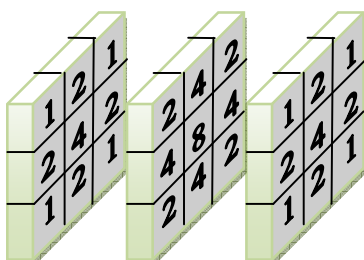


Figura C.2 Transformada contourlet a tres niveles de una imagen de resonancia magnética.

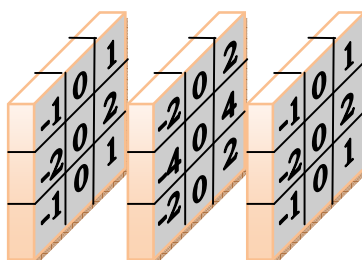
La principal diferencia entre contourlets y curvelets es el hecho de que la transformada contourlet está definida directamente sobre una malla rectangular discreta, sin embargo, las funciones contourlet tienen menos definidas las características geométricas direccionales, lo cual introduce artefactos en aplicaciones de reducción de ruido y compresión [64].

D

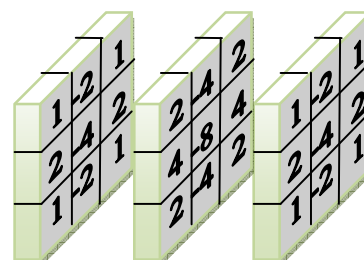
Filtros de la HT directa 3D



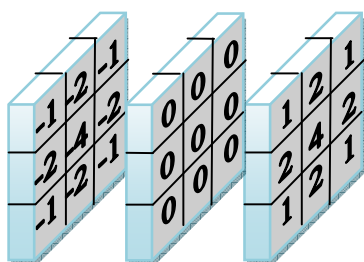
F000



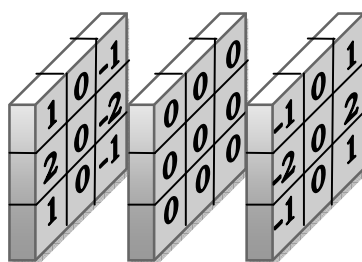
F001



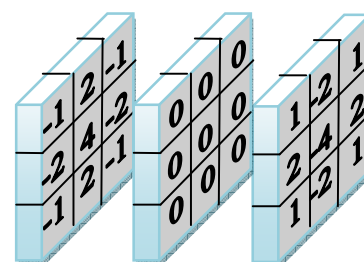
F002



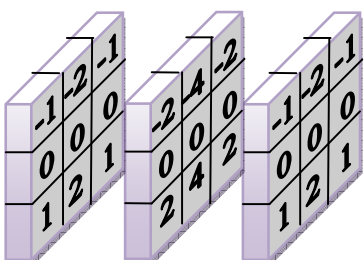
F010



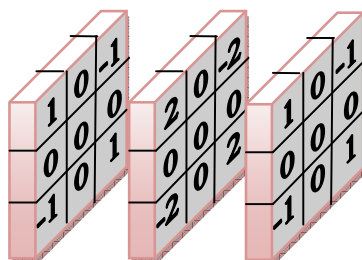
F011



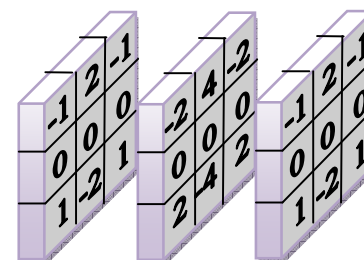
F012



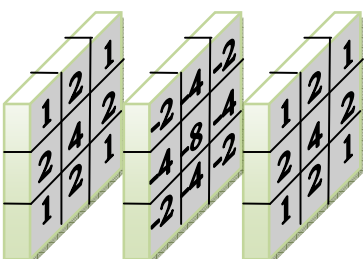
F100



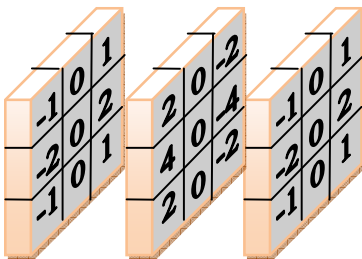
F101



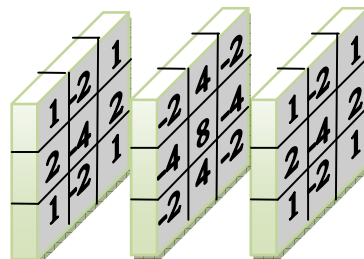
F102



F020

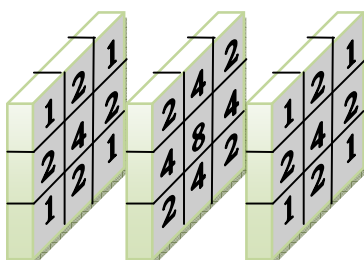


F021

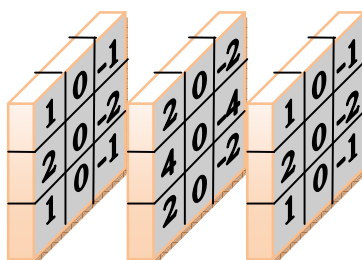


F022

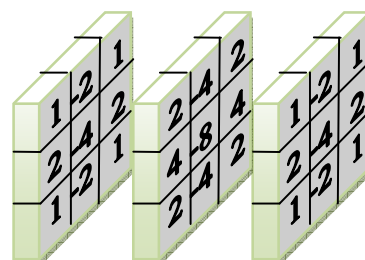
Filtros de la HT inversa 3D



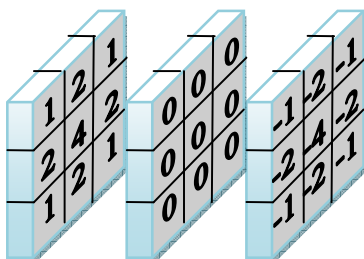
F000



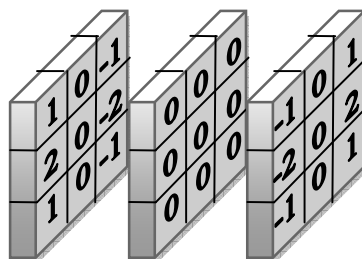
F001



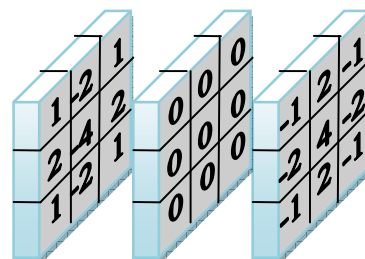
F002



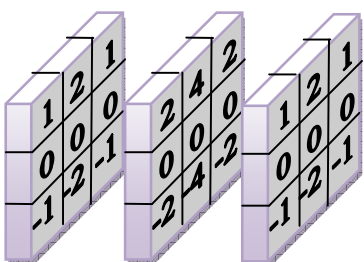
F010



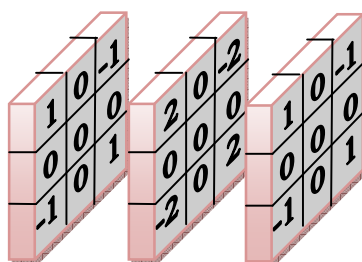
F011



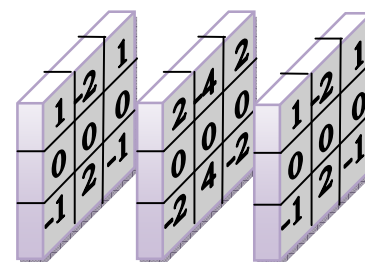
F012



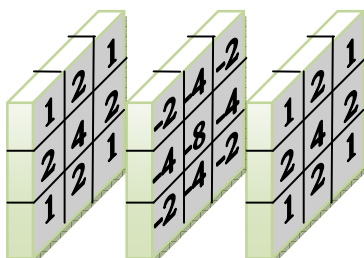
F100



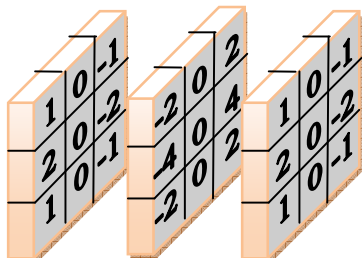
F101



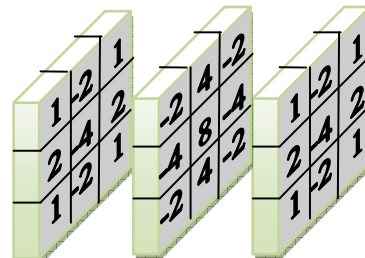
F102



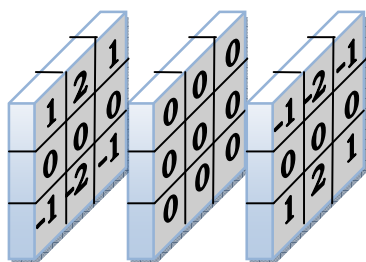
F020



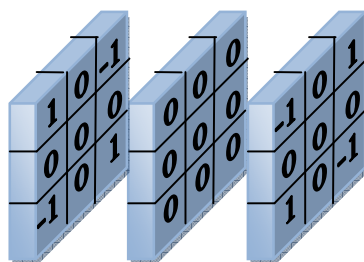
F021



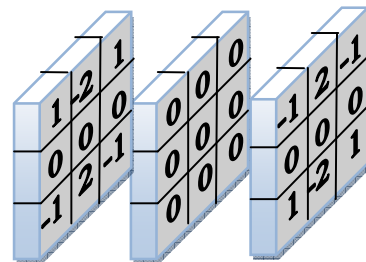
F022



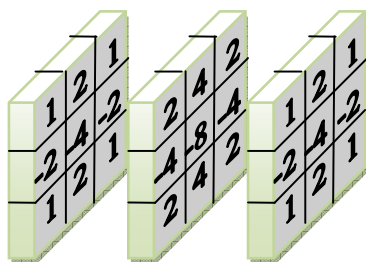
F110



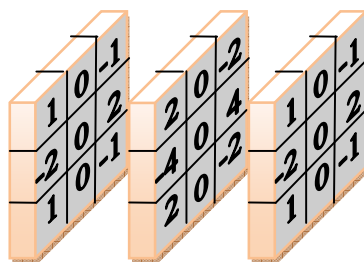
F111



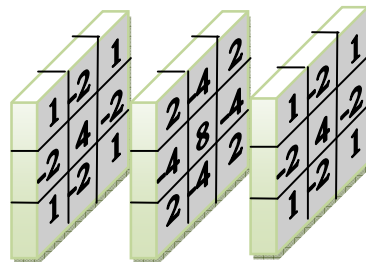
F112



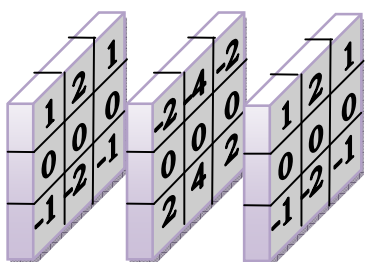
F200



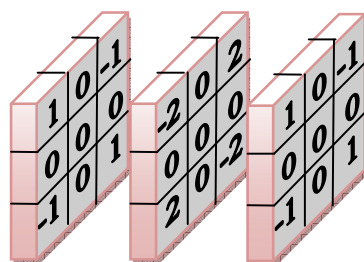
F201



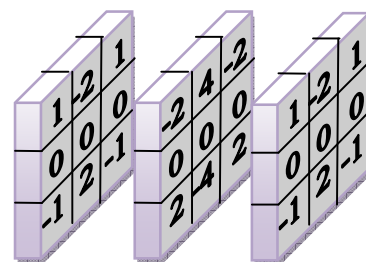
F202



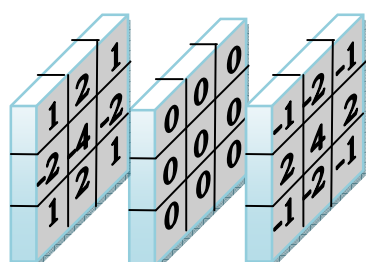
F120



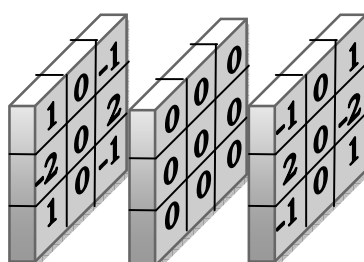
F121



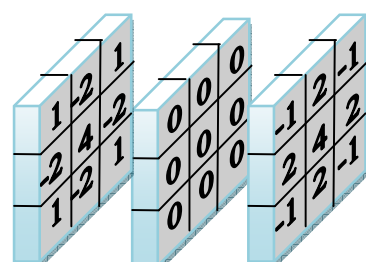
F122



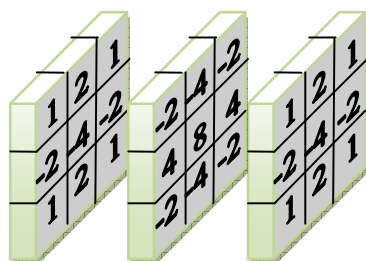
F210



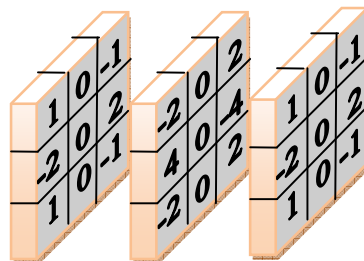
F211



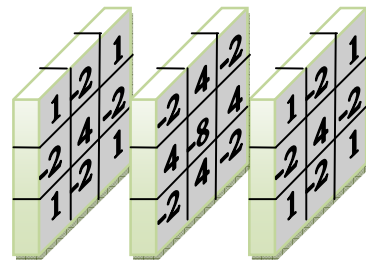
F212



F220



F221



F222

Árbol utilizado en la implementación de la HT 3D

El árbol empleado en el algoritmo rápido de la transformada de Hermite se basa en la aplicación sucesiva de sumas y diferencias para obtener las distintas combinaciones de derivadas en las tres dimensiones x , y y z . Esta estructura se observa en la Figura F.1, cuya nomenclatura es la siguiente:

▪ R ijk ijk:	Representa la acumulación de sumas y diferencias en cada registro involucrado en el proceso, por ejemplo en el voxel de entrada aparece abajo R ijk ijk indicando que no se ha aplicado aún ninguna operación de suma o resta, en L1 aparece R +jk ijk ya que representa el resultado de la primera suma del voxel procesado con el valor del voxel anterior almacenado en MR0.
▪ MR0,...,MR8:	Registros que almacenan el atraso en renglones.
▪ MC0,...,MC17:	Registros que almacenan el atraso en columnas.
▪ MT0,...,MT35:	Registros que almacenan el atraso entre cortes del volumen de datos.
▪ L1,...,L126:	Registros temporales que guardan el resultado de sumar o restar los registros MR, MC o MT con el dato de entrada a procesar.

En la última columna se observan los 27 coeficientes que resultan de obtener las derivadas hasta segundo orden en cada dirección. No obstante, se observa que algunos de ellos pueden obtenerse de varios caminos distintos, por ejemplo el coeficiente C111 se repite 8 veces a lo largo de toda la estructura, y por ello existen registros que no son considerados, pues todos los coeficientes repetidos tienen el mismo valor sin importar el orden en que las derivadas fueron efectuadas para cada dirección. Para resaltar la manera en la que fueron obtenidos todos los coeficientes, las trayectorias consideradas en la programación son las que aparecen como **CXXX** y todos los registros requeridos están en azul mientras que los registros innecesarios aparecen en amarillo.

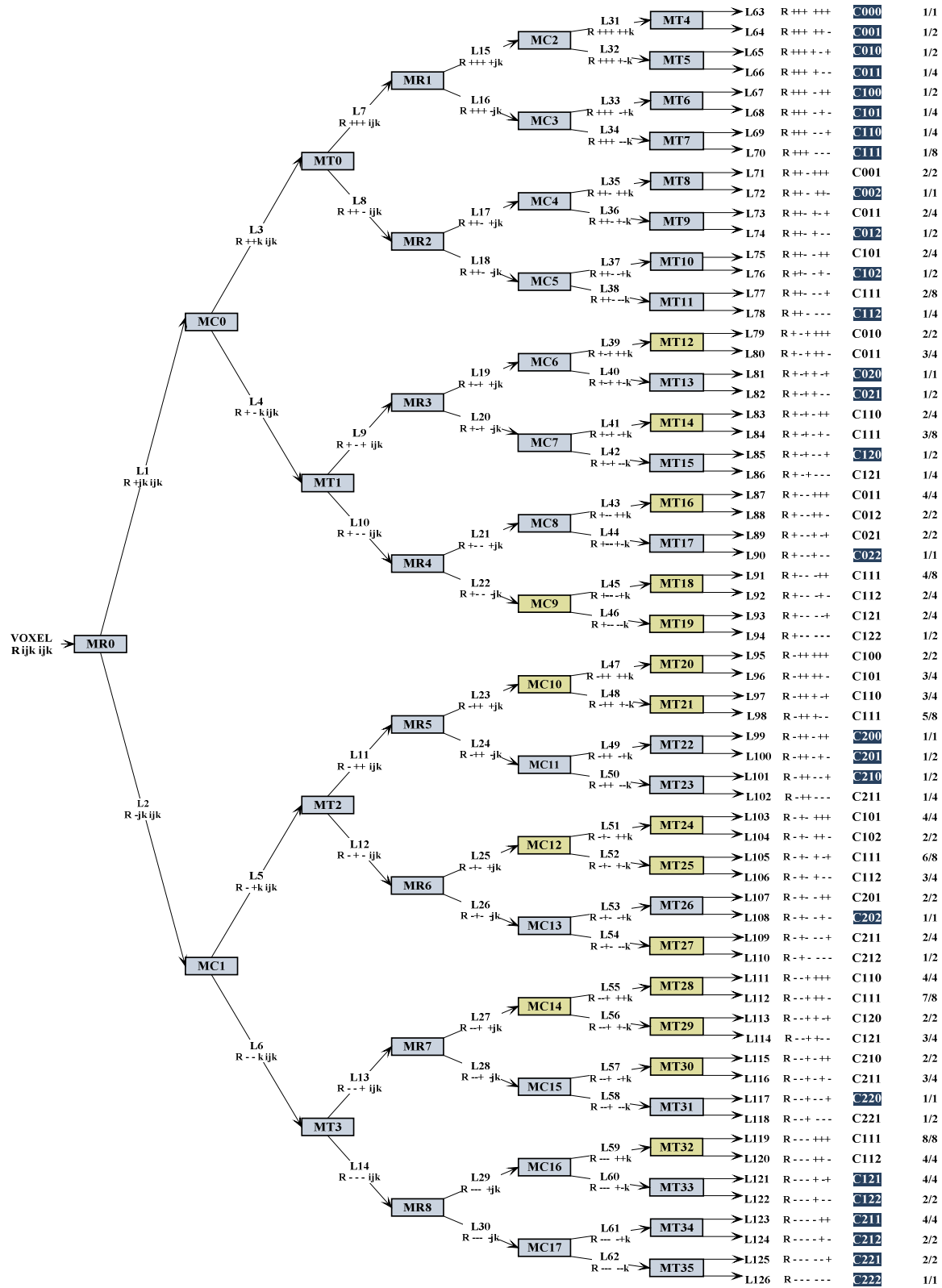


Figura F.1 Árbol de derivaciones de orden 2 para el cálculo de la HT mediante sumas y diferencias.

Referencias

- [1] ABRAMOWITZ, M., AND SETEGUN, I. A. *Handbook of mathematical functions*. Dover, 1965.
- [2] ACHIM, A., LOZA, A., BULL, D., AND CANAGARAJAH, N. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Statistical modelling for wavelet-domain image fusion, pp. 119–138.
- [3] AGUILAR, M., AND NEW, J. Fusion of multi-modality volumetric medical imagery. In *Proceedings of the Fifth International Conference on Information Fusion* (2002), vol. 2, pp. 1206–1212.
- [4] AGUILAR-PONCE, R., TECPANECATL-XIHUITL, J., KUMAR, A., AND BAYOUMI, M. Pixel-level image fusion scheme based on linear algebra. In *IEEE International Symposium on Circuits and Systems. ISCAS 2007*. (2007), pp. 2658–2661.
- [5] ALI, F. E., EL-DOKANY, I. M., SAAD, A. A., AND ABD EL-SAMIE, F. E.-S. Curvelet fusion of mr and ct images. *Progress In Electromagnetics Research* 3 (2008), 215–224.
- [6] AZENCOTT, R., CHALMOND, B., AND COLDEFY, F. Markov fusion of a pair of noisy images to detect intensity valleys. *Int. J. Comput. Vision* 16, 2 (1995), 135–145.
- [7] BARRA, V., AND BOIRE, J.-Y. A general framework for the fusion of anatomical and functional medical images. *NeuroImage* 13, 3 (2001), 410 – 424.
- [8] BEN HAMZA, A., HE, Y., KRIM, H., AND WILLSKY, A. A multiscale approach to pixel-level image fusion. *Integr. Comput.-Aided Eng.* 12, 2 (2005), 135–146.
- [9] BIN, L., LIANFANG, T., YUANYUAN, K., AND XIA, Y. Parallel multimodal medical image fusion in 3d conformal radiotherapy treatment planning. In *The 2nd International Conference on Bioinformatics and Biomedical Engineering. ICBBE 2008*. (2008), pp. 2600–2604.
- [10] BLOCH, I. Information combination operators for data fusion: a comparative review with classification. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 26, 1 (1996), 52–67.
- [11] BLUM, R. S., AND YANG, J. *Advanced Video-Based Surveillance Systems*. Kluwer, 2003, ch. Image Fusion using the Expectation-Maximization Algorithm and a Gaussian Mixture Model.
- [12] BOLAN, C. Multimodality imaging: Fusing imaging for the future. *Imaging Technology News website* (2006).
- [13] BRETSCHNEIDER, T., AND KAO, O. Image fusion in remote sensing. *Proceedings of the 1st Online Symposium of Electronic Engineers*, 2000.
- [14] BUCHROITHNER, M. Geodata interrelations: Inventory and structuring attempt of taxonomic diversity. In *2nd Conf. Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images* (1998), pp. 11–15.

- [15] BURT, P., AND KOLCZYNSKI, R. Enhanced image capture through fusion. In *Proceedings of the Fourth International Conference on Computer Vision* (1993), pp. 173–182.
- [16] BURT, P. J., AND ADELSON, E. H. Merging images through pattern decomposition. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (1985), A. G. Tescher, Ed., vol. 575 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 173.
- [17] BURT, P. J., EDWARD, AND ADELSON, E. H. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications* 31 (1983), 532–540.
- [18] CALHOUN, V., AND ADALI, T. Feature-based fusion of medical imaging data. *IEEE Transactions on Information Technology in Biomedicine* 13, 5 (2009), 711–720.
- [19] CANDÈS, E. J., C, E. J., AND DONOHO, D. L. Curvelets - a surprisingly effective nonadaptive representation for objects with edges, 2000.
- [20] CANDÈS, E., DEMANET, L., DONOHO, D., AND YING, L. Fast discrete curvelet transforms, 2005.
- [21] CANDÈS, E. J., AND DONOHO, D. L. New tight frames of curvelets and optimal representations of objects with piecewise c^2 singularities. *Communications on Pure and Applied Mathematics* 57, 2 (2004), 219–266.
- [22] CHANUSSOT, J., MAURIS, G., AND LAMBERT, P. Improving road detection on sar images using fuzzy fusion methods. In *Proceedings of the 16th IEEE Conference on Instrumentation and Measurement Technology. IMTC/99.* (1999), vol. 3, pp. 1807–1812.
- [23] CHIPMAN, L., ORR, T., AND GRAHAM, L. Wavelets and image fusion. In *International Conference on Image Processing* (1995), vol. 3, pp. 248–251.
- [24] COLIN, A., AND BOIRE, J.-Y. Mri-spect fusion for the synthesis of high resolution 3d functional brain images: a preliminary study. *Computer Methods and Programs in Biomedicine* 60, 2 (1999), 107–116.
- [25] CROWLEY, J. L. A representation for visual information. Tech. Rep. CMU-RI-TR-82-07, Robotics Institute, Pittsburgh, PA, November 1981.
- [26] CRUZ-TECHICA, S., ESCALANTE-RAMIREZ, B., AND ALBOR, E. M. Medical image fusion in 3-d by the multiresolution directional-oriented hermite transform. In *International Seminar on Medical Image Processing and Analysis, SIPAIM 2010* (2010).
- [27] CSIC, S. Instituto de optica. image processing and vision modeling group. <http://www.iv.optica.csic.es/>, (Last modified, June 2011).
- [28] DEPARTMENT OF DEFENCE, U. Data fusion lexicon. Data Fusion Subpanel of the Joint Directors of Laboratories, Technical Panel for C3, 1991.
- [29] DIJK, A. M. V., AND MARTENS, J.-B. Image representation and compression with steered hermite transforms. *Signal Processing* 56, 1 (1997), 1–16.
- [30] DO, M. Contourlet toolbox. <http://www.mathworks.com/matlabcentral/fileexchange/8837>, (Last modified 27 Oct 2005).

- [31] DO, M., AND VETTERLI, M. Contourlets: a directional multiresolution image representation. In *International Conference on Image Processing* (2002), vol. 1, pp. I-357 – I-360.
- [32] DO, M., AND VETTERLI, M. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing* 14, 12 (2005), 2091 –2106.
- [33] DONOHO, D. L., AND DUNCAN, M. R. Digital curvelet transform: Strategy, implementation and experiments. In *Proc. Aerosense 2000, Wavelet Applications VII* (1999), SPIE, pp. 12–29.
- [34] DOUGHERTY, G. *Digital Image Processing for Medical Applications*. Cambridge University Press, 2009.
- [35] DURUCAN, E., AND EBRAHIMI, T. Change detection and background extraction by linear algebra. *Proceedings of the IEEE* 89, 10 (2001), 1368 –1381.
- [36] ELDEIB, A., YAMANY, S., AND FARAG, A. Multi-modal medical volumes fusion by surface matching. In *Proceedings of the Fifth International Symposium on Signal Processing and Its Applications. ISSPA '99*. (1999), vol. 1, pp. 439–442.
- [37] ESCALANTE-RAMIREZ, B., AND MARTENS, J.-B. Noise reduction in computerized tomography images by means of polynomial transforms. *J. Vis. Comun. Image Represent.* 3 (1992), 272–285.
- [38] ESCALANTE-RAMIREZ, B., VENEGAS-MARTINEZ, S., AND GARCIA-UGALDE, F. Pyramidal predictive image coding with polynomial transforms. In *Proceedings SPIE* (1995), E. Still-Image Compression, Majid Rabbani; Edward J. Delp; Sarah A. Rajala, Ed., vol. 2418, Proceedings SPIE, pp. 107–117.
- [39] ESCALANTE-RAMÍREZ, B. The hermite transform as an efficient model for local image analysis: An application to medical image fusion. *Comput. Electr. Eng.* 34, 2 (2008), 99–110.
- [40] ESCALANTE-RAMÍREZ, B., CRUZ-TECHICA, S., NAVA, R., AND CRISTÓBAL, G. *Image Fusion*. INTECH, 2011, ch. A Perceptive-Oriented Approach to Image Fusion, pp. 165–184.
- [41] ESCALANTE-RAMÍREZ, B., AND LÓPEZ-CALOCA, A. *Image Processing for Remote Sensing*. CRC Press, Boca Raton, 2006, ch. The Hermite transform: an efficient tool for noise reduction and image fusion in remote sensing., pp. 539–557.
- [42] ESCALANTE-RAMÍREZ, B., AND LÓPEZ-CALOCA, A. Remote sensing image fusion with a multiresolution directional-oriented image transform based on gaussian derivatives. In *Proceedings of the American Institute of Physics, 5th International Workshop on Information Optics* (2006), pp. 178–187.
- [43] ESCALANTE-RAMÍREZ, B., AND SILVÁN-CÁRDENAS, J. L. Advanced modeling of visual information processing: A multi-resolution directional-oriented image transform based on gaussian derivatives. *Signal Processing: Image Communication* 20, 9-10 (2005), 801 – 812.
- [44] FREEMAN, W., AND ADELSON, E. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 9 (Sept. 1991), 891 –906.
- [45] GALLEGO, A. R. Modelos visuales en el análisis de la calidad de imagen. Master's thesis, Tesis de licenciatura, Universidad Politécnica de Madrid, 2006.
- [46] GARG, S., USHAH KIRAN, K., MOHAN, R., AND TIWARY, U. Multilevel medical image fusion using segmented image by level set evolution with region competition. In *27th Annual*

International Conference of the Engineering in Medicine and Biology Society. IEEE-EMBS 2005. (2005), pp. 7680–7683.

- [47] GUIHONG, Q., DALI, Z., AND PINGFAN, Y. Medical image fusion by wavelet transform modulus maxima. *Opt. Express* 9, 4 (2001), 184–190.
- [48] GUY, C., AND FFYTCH, D. *An Introduction to The Principles of Medical Imaging*. Guy, Chris and Ffytche, Dominic, 2005.
- [49] HALL, D. L. *Mathematical Techniques in Multisensor Data Fusion*. Artech House, Inc., Norwood, MA, USA, 1992.
- [50] HASHIMOTO, M., AND SKLANSKY, J. Multiple-order derivatives for detecting local image characteristics. *Comput. Vision Graph. Image Process.* 39, 1 (1987), 28–55.
- [51] HILL, P., CANAGARAJAH, N., AND BULL, D. Image fusion using complex wavelets. In *Proc. 13th British Machine Vision Conference* (2002), pp. 487–496.
- [52] HUANG, F.-Z., AND SU, J.-B. Face contour detection and tracking with complex backgrounds. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics, 2004.* (2004), vol. 6, pp. 3855 – 3859.
- [53] IMAGEFUSION.ORG. The online resource for research in image fusion. <http://www.imagefusion.org/>, (Last modified, 2008).
- [54] ITU. Methodology for the subjective assessment of the quality of television pictures, 2000. pp. 500-509.
- [55] KADEMI CONSULTING, S. Biomedical image and signal processing. <http://www.khademiconsulting.com>, (Last modified 2008).
- [56] KATARTZIS, A., AND PETROU, M. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Current trends in super-resolution image reconstruction, pp. 1–26.
- [57] KESSLER, M. L., AND LI, K. Image fusion for conformal radiotherapy. Advanced Medical Publishing, 2001.
- [58] KLEIN, L. A. *Sensor and Data Fusion Concepts and Applications*. Society of Photo-Optical Instrumentation Engineers (SPIE), 1993.
- [59] KOR, S., AND TIWARY, U. Feature level fusion of multimodal medical images in lifting wavelet transform domain. In *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEMBS '04.* (2004), vol. 1, pp. 1479–1482.
- [60] KUMAR, Y. K. Comparison of fusion techniques applied to preclinical images: fast discrete curvelet transform using wrapping technique & wavelet transform. *Journal of Theoretical and Applied Information Technology* 5, 6 (2009), 668–673.
- [61] LI, H., MANJUNATH, B., AND MITRA, S. Multi-sensor image fusion using the wavelet transform. In *IEEE International Conference on Image Processing. Proceedings. ICIP-94.*, (1994), vol. 1, pp. 51 –55.
- [62] LI, S., KWOK, J. T., AND WANG, Y. Combination of images with diverse focuses using the spatial frequency. *Information Fusion* 2, 3 (2001), 169 – 176.

- [63] LICAI, Y., XIN, L., AND YUCUI, Y. Medical image fusion based on wavelet packet transform and self-adaptive operator. In *The 2nd International Conference on Bioinformatics and Biomedical Engineering. ICBBE 2008*. (2008), pp. 2647–2650.
- [64] MA, J., AND PLONKA, G. The curvelet transform. *IEEE Signal Processing Magazine* 27, 2 (2010), 118–133.
- [65] MAHYARI, A., AND YAZDI, M. A novel image fusion method using curvelet transform based on linear dependency test. In *International Conference on Digital Image Processing* (March 2009), pp. 351–354.
- [66] MALLAT, S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 7 (July 1989), 674–693.
- [67] MANDINC, D. P., OBRADOVIC, D., KUH, A., ADALI, T., TRUTSCHELL, U., GOLZ, M., BARRIA, J., CONSTANTINIDES, A., AND CHAMBERS, J. Data fusion for modern engineering applications: An overview. In *In Proceedings: ICANN* (2005), vol. 2, pp. 715–721.
- [68] MANGOLINI, M. *Apport de la fusion d’images satellitaires multicapteurs au niveau pixel en télédétection et photo-interprétation*. PhD thesis, University of Nice-Sophia Antipolis, France, November 1994.
- [69] MARTENS, J.-B. The hermite transform-applications. *IEEE Transactions on Acoustics, Speech and Signal Processing* 38, 9 (1990), 1607–1618.
- [70] MARTENS, J.-B. The hermite transform-theory. *IEEE Transactions on Acoustics, Speech and Signal Processing* 38, 9 (1990), 1595–1606.
- [71] MARTENS, J.-B. Local orientation analysis in images by means of the hermite transform. *IEEE Transactions on Image Processing* 6, 8 (1997), 1103–1116.
- [72] MATSOPOULOS, G., MARSHALL, S., AND BRUNT, J. Multiresolution morphological fusion of mr and ct images of the human brain. *IEEE Proceedings on Vision, Image and Signal Processing* 141, 3 (1994), 137–142.
- [73] MEDINA, RUBÉN Y BELLERA, J. Bases del procesamiento de imágenes médicas. Universidad de los Andes, Facultad de Ingeniería, Grupo de Ingeniería Biomédica de la ULA (GIBULA).
- [74] MITIANOUDIS, N., AND STATHAKI, T. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Image fusion schemes using ICA bases, pp. 85–118.
- [75] MITIANOUDIS, N., AND STATHAKI, T. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Enhancement of multiple sensor images using joint image fusion and blind restoration, pp. 299–326.
- [76] MOCTEZUMA, M., ESCALANTE, B., MENDEZ, R., LOPEZ, J., AND GARCIA, F. Coastline detection with polynomial transforms and markovian segmentations. In *Geoscience and Remote Sensing. IGARSS '97. IEEE International Remote Sensing - A Scientific Vision for Sustainable Development* (1997), vol. 1, pp. 38–40.
- [77] MOHAMMAD-DJAFARI, A. Bayesian approach for data and image fusion. In *Bayesian Inference and Maximum Entropy Methods, AIP Conference Proceedings* (2002), C. Williams, Ed., vol. 659, pp. 386–407.

- [78] MORABITO, F., SIMONE, G., AND CACCIOLA, M. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Image fusion techniques for non-destructive testing and remote sensing applications, pp. 367–392.
- [79] MUÑOZ, T. Fusión de imágenes medicas a través de la transformada hermite. Master's thesis, Universidad Nacional Autónoma de México, Facultad de Ingeniería, 2005.
- [80] NAVA, V. Desarrollo de un método de fusión de imágenes basado en energía utilizando la transformada wavelet discreta. Master's thesis, Universidad Nacional Autónoma de México, Maestría en Ingeniería (Computación), 2007.
- [81] ORELLANA P., P. Errores neurorradiologicos frecuentes en tc y rm. *Revista chilena de radiología* 9, 2 (2003), 93–103.
- [82] OTAZU, X. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Theory and implementation of image fusion methods based on the á trous algorithm, pp. 139–156.
- [83] PAJARES, G., AND DE LA CRUZ, J. M. A wavelet-based image fusion tutorial. *Pattern Recognition* 37, 9 (2004), 1855 – 1872.
- [84] PEREIRA, R. H., AND PEREIRA, C. L. Imagenología molecular: Un nuevo paradigma en radiología. *Rev Hum Med* 6, 3 (2006).
- [85] PEREZ, C. A., VIJAYAKUMAR, S., PURDY, J. A., MICHALSKI, J. M., BRADLEY, J., VIJAYAKUMAR, S., PEREZ, C. A., AND LEVITT, S. H. Three-dimensional treatment planning and conformal therapy. In *Technical Basis of Radiation Therapy*, Medical Radiology. 2006, pp. 179–202.
- [86] PETROVIC, V. *Multisensor Pixel-level image fusion*. PhD thesis, Faculty of Science, University of Manchester, 2001.
- [87] PIELLA, G. New quality measures for image fusion. In *Proceedings of the Seventh International Conference on Information Fusion* (2004), P. Svensson and J. Schubert, Eds., vol. I, International Society of Information Fusion, pp. 542–546.
- [88] POHL, C., AND VAN GENDEREN, J. L. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing* 19, 5 (1998), 823–854.
- [89] PRATT, W. *Digital Image Processing*. A John Wiley & Sons, Inc., Publication, 2007.
- [90] PUSHKAR PRADHAM, N. H. Y., AND KING, R. L. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Concepts of image fusion in remote sensing applications, pp. 393–428.
- [91] QIANG WANG, Y. S., AND JIN, J. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Performance evaluation of image fusion techniques, pp. 469–492.
- [92] RAMOS, P. M., CALLAO, M. P., AND RUISÁNCHEZ, I. Data fusion in the wavelet domain by means of fuzzy aggregation connectives. *Analytica Chimica Acta* 584, 2 (2007), 360 – 369.
- [93] RANCHLN, T., AND WALD, L. Fusion of high spatial and spectral resolution images: The arsis concept and its implementation. *Photogrammetric Engineering and Remote Sensing* 66, 1 (2000), 49–61.

- [94] REHM, K., STROTHER, S. C., ANDERSON, J. R., SCHAPER, K. A., AND ROTTENBERG, D. A. Display of merged multimodality brain images using interleaved pixels with independent color scales. *J Nucl Med* 35, 11 (1994), 1815–1821.
- [95] RICK S. BLUM, Z. X., AND ZHANG, Z. *Multi-Sensor Image Fusion and Its Applications*. CRC Press, Boca Raton, 2006, ch. An Overview of Image Fusion, pp. 1–35.
- [96] RIVERO-MORENO, C., AND BRES, S. Texture feature extraction and indexing by hermite filters. In *Proceedings of the 17th International Conference on Pattern Recognition. ICPR 2004*. (2004), vol. 1, pp. 684 – 687.
- [97] ROCKINGER, O., FECHNER, T., AND AG, D. B. Pixel-level image fusion: The case of image sequences. *Proceedings of SPIE*, 1998. Vol. 3374, 1998, pp. 378–388.
- [98] SCHOWEGERDT, R. A. *Remote Sensing: Models and Methods for Image Processing*. Academic Press, 1997.
- [99] SHARMA, R. K., LEEN, T. K., AND PAVEL, M. Bayesian sensor image fusion using local linear generative models. *Optical Engineering* 40, 7 (2001), 1364–1376.
- [100] SHEIKH, H., SABIR, M., AND BOVIK, A. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing* 15, 11 (2006), 3440 – 3451.
- [101] SHENSA, M. The discrete wavelet transform: wedding the a trous and mallat algorithms. *IEEE Transactions on Signal Processing* 40, 10 (Oct. 1992), 2464 –2482.
- [102] SILVAN-CARDENAS, J., AND ESCALANTE-RAMIREZ, B. The multiscale hermite transform for local orientation analysis. *IEEE Transactions on Image Processing* 15, 5 (2006), 1236 –1253.
- [103] SILVAN-CARDENAS, J. L. Compresión de imágenes basada en modelos gaussianos de percepción visual. Master’s thesis, Universidad Nacional Autónoma de México, Facultad de Ingeniería, 2002.
- [104] SINGH, R., VATSA, M., AND NOORE, A. Multimodal medical image fusion using redundant discrete wavelet transform. In *Seventh International Conference on Advances in Pattern Recognition. ICAPR '09*. (Feb. 2009), pp. 232–235.
- [105] SONKA, J. M. F. M. *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*, vol. PM80SC. SPIE Press Book, 2000.
- [106] SOUSSI, N., AND BARAT, J.-L. Merging in medical multimodality imaging. In *Engineering in Medicine and Biology Society. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE* (1996), vol. 5, pp. 2293 –2294.
- [107] TELECOM SALUD, S. Modalidades. <http://www.telecomsalud.com/modalidades.htm>, (Last modified, 2004).
- [108] TENENBAUM, M., AND POLLARD, H. *Ordinary Differential Equations*. Dover, 1963.
- [109] THE CURVELET.ORG TEAM: EMMANUEL CANDES, LAURENT DEMANET, D. D. L. Y. Curvelet.org. <http://www.curvelet.org/software.html>, (Last modified 24 August 2007).

- [110] TOET, A. A morphological pyramidal image decomposition. *Pattern Recognition Letters* 9, 4 (1989), 255 – 261.
- [111] TOET, A. Hierarchical image fusion. *Mach. Vision Appl.* 3, 1 (1990), 1–11.
- [112] TOWNSEND, D. W., AND BEYER, T. A combined pet/ct scanner: the path to true image fusion. *Br J Radiol* 75, 90009 (2002), S24–30.
- [113] VAQUERO, L. J. J., AND DESCO, M. M. Mas de un siglo de imagen medica. *Arbor: Ciencia, pensamiento y cultura* 698 (2004), 337–365.
- [114] VARMA, H., FADAIE, K., HABIBANE, M., AND STOCKHAUSEN, J. Confusion in data fusion. *International Journal of Remote Sensing* 24, 4 (2003), 627–636.
- [115] VETTERLI, M., AND KOVACEVIC, J. *Wavelet and Subband Coding*. Prentice Hall PTR, Englewood Cliffs, 1995.
- [116] WALD, L. Data fusion: A conceptual approach for an efficient exploitation of remote sensing images. In *2nd Conf. Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images* (1998), pp. 17–23.
- [117] WALD, L. Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing* 37, 3 (may. 1999), 1190–1193.
- [118] WANG, A., SUN, H., AND GUAN, Y. The application of wavelet transform to multi-modality medical image fusion. In *Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control. ICNSC '06*. (2006), pp. 270–274.
- [119] WANG, A., WU, J., LI, D., AND CHEN, Y. Research on medical image fusion based on orthogonal wavelet packets transformation combined with 2v-svm. In *IEEE/ICME International Conference on Complex Medical Engineering. CME 2007*. (2007), pp. 670–675.
- [120] WANG, Q., YU, D., AND SHEN, Y. An overview of image fusion metrics. In *IEEE Instrumentation and Measurement Technology Conference. I2MTC '09*. (2009), pp. 918–923.
- [121] WANG, Y.-P., DANG, J.-W., LI, Q., AND LI, S. Multimodal medical image fusion using fuzzy radial basis function neural networks. In *International Conference on Wavelet Analysis and Pattern Recognition. ICWAPR '07*. (Nov. 2007), vol. 2, pp. 778–782.
- [122] WANG, Z., AND BOVIK, A. A universal image quality index. *IEEE Signal Processing Letters* 9, 3 (2002), 81 –84.
- [123] WANG, Z., BOVIK, A., SHEIKH, H., AND SIMONCELLI, E. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (april 2004), 600 –612.
- [124] WANG, Z., ZIOU, D., ARMENAKIS, C., LI, D., AND LI, Q. A comparative analysis of image fusion methods. *IEEE Transactions on Geoscience and Remote Sensing* 43, 6 (June 2005), 1391–1402.
- [125] WENCANG, Z., AND LIN, C. Medical image fusion method based on wavelet multi-resolution and entropy. In *IEEE International Conference on Automation and Logistics. ICAL 2008*. (2008), pp. 2329–2333.

- [126] WRIGHT, W. A. Quick markov random field image fusion. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (1998), I. Kadar, Ed., vol. 3374, pp. 302–308.
- [127] XIE, H., LI, G., NING, H., MENARD, C., COLEMAN, C. N., AND MILLER, R. W. 3d voxel fusion of multi-modality medical images in a clinical treatment planning system. *IEEE Symposium on Computer-Based Medical Systems 0* (2004), 48.
- [128] XYDEASA, C., AND PETROVIC, V. *Image Fusion: Algorithms and Applications*. Academic Press, 2008, ch. Pixel-level image fusion metrics, pp. 429–450.
- [129] YANG, B., JING, Z.-L., AND ZHAO, H.-T. Review of pixel-level image fusion. *Journal of Shanghai Jiaotong University (Science)* 15 (2010), 6–12.
- [130] YANG, J., AND BLUM, R. S. A statistical signal processing approach to image fusion for concealed weapon detection. In *IEEE Intl. Conf. on Image Processing* (2002), Marcel Dekker, pp. 513–516.
- [131] YANG, L., GUO, B., AND NI, W. Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. *Neurocomputing* 72, 1-3 (2008), 203 – 211.
- [132] YANG, Y., HAN, C., KANG, X., AND HAN, D. An overview on pixel-level image fusion in remote sensing. In *IEEE International Conference on Automation and Logistics* (2007), pp. 2339–2344.
- [133] YOUNG, R. The gaussian derivative theory of spatial vision: analysis of cortical cell receptive field line-weighting profiles. Tech. Rep. GMR-4920, General Motors Research, 1986.
- [134] YÉSOU, H., BESNUS, Y., ROLET, J., PION, J., AND AING, A. Merging seasat and spot imagery for the study of geological structures in a temperate agricultural region. *Remote Sensing of Environment* 43, 3 (1993), 265 – 279.
- [135] ZHANG, H., LIU, L., AND LIN, N. A novel wavelet medical image fusion method. In *Proceedings of the 2007 International Conference on Multimedia and Ubiquitous Engineering. MUE '07.* (2007), IEEE Computer Society, pp. 548–553.
- [136] ZHANG, Z., AND BLUM, R. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proceedings of the IEEE* 87, 8 (1999), 1315–1326.
- [137] ZHANG, Z., AND BLUM, R. S. Region-based image fusion scheme for concealed weapon detection. In *Proceedings of the 31st Annual Conference on Information Sciences and Systems* (1997), pp. 168–173.