

01170



**UNIVERSIDAD NACIONAL AUTONOMA DE
MEXICO**

FACULTAD DE INGENIERIA

**FUSION DE IMAGENES MEDICAS A TRAVES DE LA
TRASFORMADA DE HERMITE**

T E S I S

QUE PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERIA ELECTRICA

P R E S E N T A:

JUAN CARLOS MUÑOZ TORRES

DIRECTOR DE TESIS: DR. BORIS ESCALANTE RAMÍREZ

CIUDAD UNIVERSITARIA, 2005

m346573



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

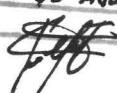
Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

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.

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: MUÑOZ TORRES JUAN CARLOS

FECHA: 4 DE AGOSTO DE 2005

FIRMA: 

Indice

1. Introducción	1
2. Imagen Médica	3
2.1. Conceptos básicos	3
2.2. Modalidades de adquisición de imágenes médicas	7
2.2.1. Radiología convencional	7
2.2.2. Radiología digital	8
2.2.2.1. Tomografía por computadora o tomografía axial computada	9
2.2.3. Medicina nuclear	11
2.2.3.1. Tomografía Computarizada por Emisión de Fotón Único (SPECT)	12
2.2.3.2. Tomografía por Emisión de Positrones (PET)	13
2.2.4. Ecografía	15
2.2.5. Resonancia magnética	16
2.2.6. La imagen molecular	18
2.3.7. Futuro de las imágenes médicas	19
3. Fusión de imágenes médicas	22
3.1. Conceptos de fusión	22
3.2. Tipos de fusión	27
3.3. Métodos representativos de fusión	28
3.3.1. Pirámide laplaciana y gaussiana	29
3.3.1.1. Generación de una pirámide gaussiana	30
3.3.1.2. Generación del kernel	31
3.3.1.3. Funciones de ponderación equivalentes o promedio	32
3.3.1.4. Interpolación de la pirámide gaussiana	34
3.3.2. Pirámide estándar gaussiana y laplaciana	35
3.3.3. Pirámide laplaciana FSD (Filter-Substract-Decimate)	37
3.3.4. Pirámide de cociente de paso bajo	38
3.3.5. Pirámide de gradientes	39
3.3.6. Pirámide de morfología matemática	40
3.3.7. Fusión de imágenes usando la transformada wavelets	43

- 3.3.8. Fusión promedio.....49
- 3.3.9. Máximo contraste.....49
- 3.3.10. Análisis de componentes principales (PCA)51
- 3.4. Ejemplos de métodos de fusión.....53
- 4. Transformada de Hermite.....58
 - 4.1. Introducción58
 - 4.2. Transformada polinomial 1D59
 - 4.3. Transformada de Hermite 1D63
 - 4.4. Transformada Polinomial 2D64
 - 4.4.1. Representación de imágenes con coeficientes polinomiales unidimensionales ...65
 - 4.5. Transformada de Hermite 2D68
 - 4.6. Transformadas Polinomiales Discretas.....69
 - 4.7. Transformada discreta de Hermite 2D71
 - 4.8. Aplicaciones de la Transformada de Hermite74
 - 4.9. Transformada de Hermite Multiescala 1D.....75
 - 4.9.1. Descomposición multiescala.....75
 - 4.9.2. Análisis estructural76
 - 4.9.3. Definición de la Transformada de Hermite Multiescala (THM)77
 - 4.9.4. Acotamiento del rango de escalas.....78
 - 4.9.5. Esquema piramidal continuo79
 - 4.9.6. Predicción entre niveles.....80
 - 4.10. Transformada de Hermite multiescala-multidireccional 2D81
 - 4.10.1. Descomposición multiescala de señales 2D81
 - 4.10.2. Análisis estructural de señales 2D.....81
 - 4.10.3. Definición de la transformada de Hermite multiescala multidireccional (THMM).....82
 - 4.10.4. Filtrado direccional83
 - 4.11. Transformada de Hermite multiescala (THDM) y multiescala multidireccional (THDMM), Caso discreto84
 - 4.11.1. Espacio-escala discreto84
 - 4.11.2. Derivadas discretas85
 - 4.11.3. Definición de la Transformadas de Hermite multiescala discreta (THDM)86
 - 4.11.4. Definición de la transformada Hermite multiescala multidireccional discreta (THDMM).....87

4.12. Clasificación de estructuras en el dominio de la transformada de Hermite discreta rotada (THDR)	87
4.12.1. Descomposición de la imagen	88
4.12.2. Rotación local	89
5. Algoritmo de fusión	91
5.1. Selección del coeficiente de baja frecuencia	98
5.2. Selección del coeficiente de alta frecuencia.....	99
5.2.1. Selección del máximo	99
5.2.2. Selección del saliente y la medida de la paridad	100
5.2.3. Selección del máximo y verificación de consistencia.....	102
6. Resultados de la fusión con imágenes médicas.....	105
6.1. Imágenes de entrada.....	105
6.2. Selección del método del coeficiente de baja frecuencia para la fusión	106
6.3. Selección del método del coeficiente de altas frecuencias para la fusión	108
6.4 Resultados obtenidos de la fusión	113
7. Conclusiones	126
8. Referencias	128

Introducción

El procesamiento digital de imágenes incluye un conjunto de técnicas que operan sobre la representación digital de una imagen, con el objeto de destacar algunos de los elementos que conforman la escena, de modo que se facilite su posterior análisis, ya sea por parte de un usuario (humano) o un sistema de visión computarizado. En general, las técnicas de procesamiento de imágenes son aplicadas cuando resulta necesario realzar o modificar una imagen para mejorar su apariencia o para destacar algún aspecto de la información contenida en la misma, o cuando se requiere, medir, contrastar o clasificar algún elemento de la misma imagen. También se utilizan técnicas de procesamiento, cuando se requiere combinar imágenes o porciones de las mismas o reorganizar su contenido.

La imagenología médica, por su parte, considera un conjunto de modalidades de adquisición de imágenes médicas, las cuales se diferencian en cuanto a la naturaleza de los principios físicos involucrados en el proceso de adquisición. Adicionalmente existen también diferencias en cuanto a la aplicación médica. Las modalidades más comunes de imagenología médica son los rayos X, la tomografía computada, la resonancia magnética nuclear, la imagenología nuclear, la imagenología por ultrasonidos, etc.

La fusión de imágenes es el proceso de combinar dos o más imágenes en una sola que integre información complementaria, tales como orillas o regiones de interés de cada una de las imágenes de entrada para propósitos de percepción visual humana y procesamiento por computadora, por lo que el objetivo de la fusión es hacer que muchas características surjan en la nueva imagen. Dicha fusión es frecuentemente requerida para imágenes obtenidas en diferentes modalidades o técnicas de captura de una misma escena u objetos. Como se discutirá en este trabajo, nos basaremos en la descomposición de las imágenes mediante la transformada de Hermite, debido a que esta transformada tiene un comportamiento como el de la visión humana.

Los objetivos del presente trabajo es obtener una mejor delineación de contornos en imágenes médicas (tomografía computarizada, resonancia magnética y tomografía por

emisión de positrones) para así facilitar el análisis de dichas imágenes. La distribución del trabajo está dividido en: Capítulo I. introducción al presente trabajo. Capítulo 2. Imágenes médicas, se hablará de los tipos de modalidades existentes agrupados en dos grandes bloques: anatómicas (morfológicas o estructurales) y funcionales. Capítulo 3. Fusión de imágenes, daremos una explicación sobre lo que es la fusión de imágenes y algunas técnicas existentes para dicha fusión. En el Capítulo 4, transformada de Hermite, se dará una breve explicación teórica sobre dicha transformada desde el caso continuo y discreto. Capítulo 5. En este capítulo se detallará el funcionamiento del algoritmo empleado para la fusión de imágenes. En el capítulo 6, se detallarán los resultados obtenidos a partir de la transformada de Hermite, se harán comparaciones con otras transformadas y se determinarán las ventajas y desventajas del método empleado. Se culmina la presente tesis con las conclusiones y trabajos futuros a partir de esta investigación.

Imagen Médica

2.1. Conceptos básicos:

Para presentar las nuevas perspectivas que se abren para la imagen médica en el futuro es conveniente establecer algunos conceptos y definiciones, que serán necesarios para estudiar las diferentes técnicas de imagen médica.

Definimos como **imagen médica** a la representación puntual en un plano o volumen (representación bidimensional o tridimensional) de la distribución espacial de una o más propiedades físicas o químicas captadas del cuerpo humano [1]. Dos parámetros de la imagen van a ser de especial interés: el contraste y la resolución.

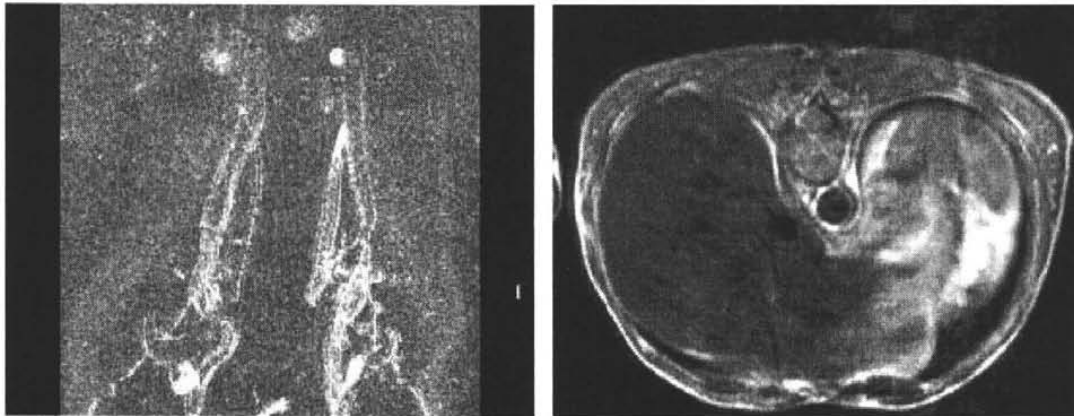
El **contraste** se define como diferencia de intensidad entre dos áreas, medida en valor absoluto o relativo (contraste relativo). Lo más importante, en el contexto que nos ocupa, es tener claro el origen de dicho contraste, esto es, saber qué parámetro físico o químico es el que está siendo representado en forma de intensidad luminosa.

Por otra parte, el concepto de **resolución** nos ayuda a caracterizar la imagen desde el punto de vista de su capacidad para discernir detalles. Es frecuente definir resolución espacial como la distancia mínima que la imagen es capaz de resolver o separar, medida en unidades longitud del mundo físico (milímetros o micras en biología, a lo mejor años luz en imágenes astronómicas). Este modo de medir la resolución es el más intuitivo, pero no el único utilizado: según el tipo de imagen se puede expresar la resolución en puntos por pulgada (dpi, "dots per inch"), pares de líneas por centímetro, etc.

En resumen, el contraste de una imagen determina qué podemos ver en ella, y su resolución espacial nos indica el grado de detalle con que se representará dicha información.

Otro parámetro de interés es la denominada **resolución temporal**, que determina la capacidad del sistema para distinguir en el tiempo dos eventos, estando muy relacionado con la velocidad de adquisición de las imágenes.

Las imágenes contienen no sólo información de interés, que representa fielmente los objetos en que estamos interesados, sino también datos ilegítimos o erróneos que los distorsionan. Esta información superpuesta no deseada se llama **ruido** si tiene carácter aleatorio o **artefacto** en caso contrario (Fig. 2.1). Estos últimos suelen depender de problemas en la técnica de producción de imagen, se podrían considerar errores sistemáticos, frente a los errores aleatorios que introduce el ruido.



A B

Figura 2.1: Información no deseada en imágenes médicas por resonancia magnética. **A.** La angiografía RM presenta mucho ruido, en forma de un punteado aleatorio. **B.** La imagen presenta un artefacto debido al movimiento de la pared abdominal, que produce una falsa textura en el interior del hígado [1].

Últimamente se viene dando en denominar **modalidades de imagen** a las diferentes técnicas de obtención de imagen médica. El elemento básico que define las diferentes modalidades es el tipo de energía utilizada. Como en casi todo proceso de medida, la obtención de imágenes médicas implica irradiar la muestra (el paciente, en este caso) con algún tipo de energía. El carácter de la misma va a definir el contraste de la imagen y también dará nombre a la modalidad correspondiente. Las modalidades fundamentales de imagen médica son: Radiología (radiación electromagnética: rayos X), Ecografía (energía ultrasónica), Medicina Nuclear (radiación electromagnética: radiación gamma) y Resonancia Magnética (radiación electromagnética: ondas de radio). Es interesante resaltar que el tipo de energía utilizada determina el tipo de interacción que se produce en los tejidos biológicos y, por tanto, en qué medida puede ser nociva para el organismo.

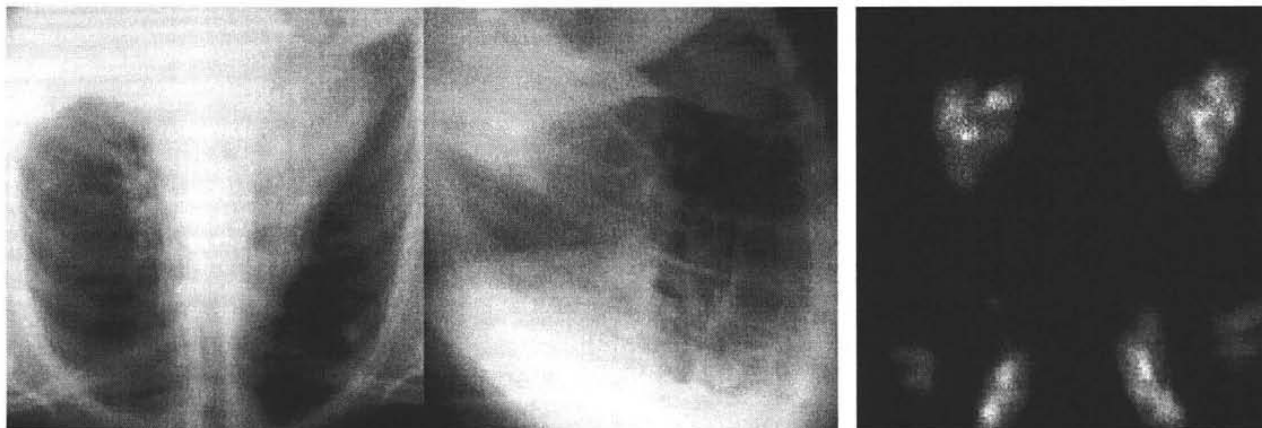
Se denominan **radiaciones ionizantes** aquéllas que por su alta energía son capaces de inducir directamente reacciones químicas, a través fundamentalmente de la ionización de

diferentes moléculas. Las **radiaciones no ionizantes** se limitan a producir calentamiento que, en principio, no presenta efectos biológicos relevantes si la potencia absorbida o depositada es baja.

Es muy frecuente clasificar las modalidades en **anatómicas** (morfológicas o estructurales ya que estas detallan estructuras, texturas, distribución espacial de niveles de gris) y **funcionales** (función del tejido fino, vasos sanguíneos, función celular, metabolismo). Las primeras se caracterizan por producir imágenes de muy buena resolución, que permiten una representación muy 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 (administración de un medicamento), su capacidad para acumular ciertas sustancias, etc.

El origen del contraste es lo que distingue esencialmente las modalidades anatómicas de las funcionales; además es frecuente que la resolución de las primeras, como se dijo, sea mucho mejor.

La Figura 2.2 ilustra esta diferencia: a la izquierda, A, vemos un estudio radiográfico del tórax, modalidad morfológica que representa los pulmones con una resolución inferior al milímetro. A la derecha, B, una gammagrafía pulmonar de perfusión del mismo paciente: es una imagen de Medicina Nuclear que representa el flujo sanguíneo pulmonar, con una resolución que puede ser de uno o dos centímetros (obsérvese el aspecto "borroso" de la imagen). Sin embargo, en esta última podemos observar la existencia de defectos de perfusión que corresponden a una embolia pulmonar, totalmente invisibles en la radiografía torácica. Esta es la razón del uso de las imágenes funcionales: asumimos su falta de resolución a cambio de la interesante información que proporcionan sobre diferentes aspectos del comportamiento de sistemas biológicos.



A **B**

Figura 2.2: **A.** Imagen morfológica: radiografías de tórax, y **B.** Imagen funcional: gammagrafía pulmonar, en la que se aprecia una clara embolia pulmonar, invisible en las radiografías de tórax [1].

Otra característica importante de las imágenes médicas deriva de su capacidad para separar objetos que se hallan a diferentes profundidades. Se llaman **imágenes proyectivas** aquéllas que representan la suma de todas las estructuras del objeto, proyectadas sobre una superficie bidimensional. En cambio, cuando el método de imagen es capaz de separar diferentes planos –“cortar la muestra en rodajas”–, cada uno de los cuales se representa en una imagen bidimensional, se denomina **tomográfico**. La imagen tomográfica crea menos problemas de superposición de objetos, facilitando notablemente su interpretación.

La Tabla 1.1 presenta una clasificación de las modalidades de imagen médica, indicando el tipo de energía utilizada y si la imagen es proyectiva o tomográfica. Los rayos X y γ son energía ionizante, todos los demás no.

Tabla 1.1: Clasificación de las modalidades de imagen médica.

Modalidad	Técnica de Imagen	Energía	Ionizante	Funcional	Tomográfica
Radiología	Radiología Convencional	Rayos X	Si	No	No
	Radiología Digital	Rayos X	Si	No	No
	Tomografía Computarizada (CT: Computed Tomography)	Rayos X	Si	No	Si
Medicina Nuclear	Tomografía Computarizada por emisión de fotón gamma único (SPECT: Single Photon Emission Computer Tomography)	Rayos γ	Si	Si	Si
	Tomografía por Emisión de Positrones (PET: Positron Emission Tomography)	Rayos γ	Si	Si	Si
Ecografía	Ecografía	Ultrasonido	No	No	Si
Resonancia Magnética	Imagen de resonancia Magnética (MRI: Magnetic Resonance Imaging)	Ondas de Radio	No	Si	Si
Endoscopia	Endoscopia	Luz	No	NO	No

2.2. Modalidades de adquisición de imágenes médicas.

Se denomina "modalidad" a toda técnica de exploración realizada con un equipo de adquisición de datos susceptibles de ser utilizados para realizar un diagnóstico [3].

2.2.1. Radiología Convencional

Los rayos X son una clase de radiación electromagnética similar a la luz pero con una longitud de onda más pequeña. Sus 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.

Los rayos X son a menudo utilizados para producir imágenes médicas y para desarrollar efectos terapéuticos en algunos pacientes. En la actualidad, las modalidades de imágenes médica basadas en la utilización de rayos X incluyen: radiografía convencional, videoangiografía o fluoroscopia y la tomografía.

La imagen por rayos X constituyen en la actualidad una de las modalidades más utilizadas en el dominio médico en general, supone del 80% al 90% del total de imágenes que se obtienen [2], permite obtener sobre una película radiográfica, comúnmente denominada "placa", la imagen de una parte del cuerpo humano expuesta a un haz de rayos X. La absorción que sufre el haz de rayos X en su recorrido depende de la densidad de los tejidos que atraviesan: máxima atenuación en el hueso, mínima en el aire e intermedia para tejidos blandos. 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 y de la fuente de rayos X, se podría decir entonces, que la radiografía es la sombra producida por el cuerpo cuando se utiliza como fuente de iluminación un haz de rayos X. En la representación 2D de la imagen obtenida (imagen de proyección), la intensidad de las diferentes zonas de la imagen da una medida de la energía en los órganos o estructuras anatómicas iluminadas, mientras que la apariencia de sus sombras da una indicación de la forma de la estructura y de su posición espacial.

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, las diferencias de densidad entre tejidos blandos no son muy elevadas, por lo que los contrastes máximos se consiguen con hueso (aplicación en traumatología) y aire (imagen pulmonar) y adicionalmente que puedan estar afectadas por diversas fuentes de ruido difíciles de controlar.

2.2.2. Radiología Digital

Continuando con la modalidad radiológica, podemos citar dos avances importantes que se han producido en los últimos años: la radiología digital y la tomografía computarizada.

La radiología digital [1] es en esencia similar a la convencional, la diferencia estriba básicamente en que los rayos X que se emiten desde un tubo, 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.

Esto supone dos importantes ventajas:

Primera, es que se puede reducir la dosis de radiación al paciente, en parte porque el detector es más sensible y sobre todo porque se elimina la repetición de exploraciones gracias a que la computadora es capaz de recuperar información válida aun cuando las condiciones de realización no haya sido las correctas.

Segunda, es que facilita mucho la gestión de información, utilizando medios informáticos para almacenar las imágenes y permitiendo su transmisión electrónica a cualquier punto del hospital donde exista una estación de visualización.

Las ventajas generales que se presenta este método son:

- La reducción de los tiempos empleados en una radiografía convencional (preparado de la película, tiempo de revelado, personal que traslada los estudios, etc.),
- la reducción de errores de transcripción de datos (rotulado de la placa, registro de los datos del paciente, etc.),
- la disminución de contaminantes (tanto los derivados de la película en sí, como los de los líquidos de revelado), y
- el poder compartir las imágenes prácticamente en tiempo real con el médico prescriptor u otros departamentos del hospital.

2.2.2.1. La tomografía por computadora ó tomografía axial computarizada

Una importante modalidad en imágenes médica es la tomografía asistida por computadora (TC) ó tomografía axial computarizada (TAC), disponible desde la década de los 70's [3] y actualmente forma parte de las herramientas habituales de diagnóstico, en donde las imágenes del interior del cuerpo se reconstruyen a partir de un conjunto de medidas de proyección, y constituye un importante avance al eliminar la superposición de objetos en la imagen radiológica convencional proyectiva, facilitando notablemente la interpretación de los resultados [1]. Aunque la fuente de emisión es la misma, se diferencia en varios aspectos significativos, algunos de ellos son:

- El tipo de imagen generada: mientras que en una radiografía se obtiene una imagen analógica, la imagen del TAC es digital con todas las posibilidades que conlleva el tratamiento informático de la misma.
- La forma de adquisición de la imagen: en una radiografía, los rayos X atraviesan el cuerpo e impactan en una película sensible. En el TAC, el paciente descansa sobre una camilla y a su alrededor, gira el tubo que emite la radiación. El tubo gira en un sentido y la placa en el contrario.
- La imagen generada: en la radiología tradicional es plana y no permite discernir sobre diferentes planos anatómicos. La imagen generada en un TAC 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.

La imagen analógica se obtiene detectando intensidad luminosa o densidades distribuidas sobre la superficie excitada (placa), mientras que en los sistemas basados en el análisis informático, como en el caso de la TC, esas densidades pueden cuantificarse en una escala de grises mejorando la definición.

De todos modos ambas técnicas tienen un punto en común muy importante, ambas se basan en la emisión de radiación electromagnética (rayos X), aunque en el caso de la TC se generan dosis más bajas.

Los rayos X se emiten desde diferentes ángulos en forma de abanico, atraviesan el organismo y forman cortes o secciones axiales (perpendiculares al eje mayor longitudinal del cuerpo) generando cortes transversales del sector a estudiar. Las distintas interfases (densidades): hueso, aire, agua y grasa, hacen que el haz tenga otro valor en cada punto midiendo la absorción de rayos X por el tejido atravesado.

Debido a las características del método, se pueden estudiar prácticamente todas las estructuras del cuerpo humano. Además ha permitido el desarrollo de técnicas terapéuticas complementarias al permitir una intervención más segura, por ejemplo en punciones y biopsias de abscesos, tumores, quistes, etc.

El equipo de TC tiene un dispositivo gantry (arco) donde se encuentra el tubo de rayos y los detectores que obtendrán los datos de las diferentes estructuras que atraviesan los rayos X, el correspondiente generador de rayos y la computadora que recibe la información y procesa la imagen.

Los avances tecnológicos más recientes en este terreno son la utilización de contrastes administrados por diferentes vías y la aparición de:

- la TAC de alta resolución,
- la TAC helicoidal, y
- la angio TAC para la visualización de imágenes vasculares de alta calidad.

Entre las ventajas de la TC helicoidal podemos mencionar que reduce el tiempo de exploración, ya que a diferencia del convencional en que el movimiento del tubo de rayos X y el detector se sincroniza con el de la camilla del paciente, de modo que, en vez de obtener un número determinado de cortes, se explora un volumen completo, de modo mucho más rápido; otra ventaja es que disminuye la dosis de contrastes necesarios y aumenta la calidad y precisión diagnóstica con un costo mucho mejor.

Por otra parte, la adquisición volumétrica de la imagen permite visualizar cualquier plano del espacio de los tejidos explorados en relieve, diferenciando los tejidos, lo que añade la posibilidad del color.

Los inconvenientes serían las elevadas dosis de radiación que recibe el paciente. En efecto, la TC es la técnica de imagen médica que suministra mayores dosis de radiación ionizante al paciente (también la helicoidal, aunque es mejor que la convencional también en este aspecto). En términos generales, podemos estimar que un estudio TC de abdomen, por ejemplo, suministra una dosis equivalente a unas 400 radiografías de tórax. Por esta razón se intenta limitar en lo posible el uso de la TC a aquellos casos con indicación clara, en los que el posible riesgo se compensa con un beneficio clínico inmediato [1].

2.2.3. Medicina Nuclear

Se llama Medicina Nuclear a la especialidad que utiliza para el diagnóstico imágenes obtenidas mediante la administración de sustancias radiactivas, es una técnica por la cual las imágenes se obtienen en 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 (o isótopo radiactivo), que es un átomo inestable que sufre desintegración nuclear, emitiendo energía en forma de diferentes tipos de radiación el cual tiene un periodo de vida muy corto y se degrada hasta convertirse en una sustancia inerte. El fundamento radica, por tanto, en marcar algún átomo radiactivo determinadas moléculas cuyo comportamiento biológico

pretendemos seguir ("trazar"), de modo que su ubicación pueda ser determinada localizando la radiación emitida, proporcionando imágenes cuya intensidad cuantifica el funcionamiento del órgano, proporcionando información acerca de la capacidad de tal órgano de asimilar o transformar la sustancia que se ha inyectado. Cada radioelemento es específico para el órgano y la función que se desea estudiar que después, para tener una imagen digital son convertidos a valores de pixel, observando la funcionalidad de la zona explorada.

La adquisición se efectúa midiendo la distribución del radioelemento en el órgano. Sin entrar en detalles sobre este proceso, podemos decir que en el diagnóstico de medicina nuclear se utilizan principalmente dos tipos de radioisótopos, que dan lugar a dos tecnologías de imagen de tipo tomográfico bastante diferentes: tomografía computarizada por emisión de fotón gamma único (Single Photon Emission Computer Tomography: SPECT, también llamado de fotón único o emisores de radiación gamma) y tomografía por emisión de positrones (Positron Emission Tomography: PET) [2].

2.2.3.1. Tomografía Computarizada por Emisión de Fotón (Gamma) Único (SPECT).

En el primer caso (SPECT), la fuente de radiación es un emisor gamma. Los fotones gamma se detectan usando una cámara de cintilación rotativa. La distribución de fotones gamma emitidos según una dirección dada constituye una proyección. Repitiendo el proceso para múltiples direcciones es posible determinar la distribución del radioelemento en el interior del cuerpo al utilizar técnicas de reconstrucción tomográfica. Los emisores gamma sufren una desintegración nuclear que da lugar a la liberación de fotones de alta energía (radiación gamma). Para detectar la emisión gamma de los radioisótopos, el dispositivo más comúnmente empleado (casi exclusivamente en la actualidad) es la denominada "gammacámara". Sin entrar en detalles de funcionamiento, podemos decir que funciona como una cámara fotográfica que en vez de luz "ve" radiación gamma. Si la gammacámara se hace rotar alrededor del paciente, obteniendo muchas imágenes planas o proyecciones (cada 5° de ángulo, por ejemplo), la computadora es capaz de efectuar reconstrucciones tomográficas, al igual que el TC [1]. La máxima resolución geométrica que se obtiene es del orden de centímetro, claramente inferior a la de otros procedimientos de imagen.

2.2.3.2. Tomografía por Emisión de Positrones (PET).

En las imágenes por emisión de positrones, el tomógrafo PET consta de un "gantry" o unidad de exploración (donde se introduce el paciente a ser estudiado) que contiene los fotodetectores. Los radioisótopos emisores de positrones se desintegran emitiendo un positrón (que es la antipartícula del electrón, con carga positiva). Al encontrarse un positrón con un electrón se aniquilan, emitiendo dos fotones que viajan en sentidos opuestos. La aniquilación se produce en un radio muy pequeño del punto de emisión, de unos pocos milímetros [1]. 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. La localización se realiza en un arreglo en forma de anillo alrededor del paciente constituido por detectores del tipo cintilador fotomultiplicador.

Los radioisótopos emisores de positrones presentan diferencias muy importantes con los emisores gamma:

- La mayor parte son elementos muy habituales en sustancias biológicas (flúor, oxígeno, nitrógeno, carbono, etc.), a diferencia de los emisores gamma, que suelen ser átomos pesados, poco habituales en la bioquímica. Esto permite que con emisores de positrones se pueda marcar un número más elevado de sustancias de interés biológico.
- Son de vida media muy breve, pocos minutos en general. Esto determina que deban ser producidos, mediante un ciclotrón, en las proximidades del lugar donde se efectúa el estudio. Esta es una de las razones que explica el elevado precio de estas exploraciones.
- El tipo de radiación que se detecta desde el exterior (dos fotones en sentido opuesto) permite una localización más exacta, produciendo imágenes de mayor resolución que los emisores gamma. Eso sí, la instrumentación requerida para su detección es también más costosa.

La instrumentación utilizada para emisores de positrones es diferente: Se emplean las denominadas cámaras PET que producen directamente cortes tomográficos, detectando las parejas de fotones que produce la aniquilación del positrón. Esto permite obtener mejor resolución geométrica, del orden de unos 5 mm en equipos clínicos actuales y mejores que 2 mm en sistemas experimentales. Por otra parte, los trazadores marcados con radioisótopos emisores de positrones son, en general, más interesantes biológicamente.

Las imágenes de Medicina Nuclear se caracterizan por tener baja resolución y ser además bastante ruidosas, debido a las dosis tan bajas utilizadas que limitan mucho el número de fotones emitidos. Estos factores determinan que su aspecto visual no sea mucho mejor que el de otras modalidades; a pesar de esta "mala calidad" aparente debemos recordar que proporcionan información sobre el funcionamiento de los sistemas biológicos ("imagen funcional") y no sólo sobre su estructura.

Fuera del ámbito clínico, la imagen nuclear tendrá sin duda una aplicación crecientemente interesante en diferentes áreas de investigación biomédica, en conjunción con técnicas avanzadas de tratamiento de imagen. El lado negativo de la imagen nuclear deriva de su escasa resolución y alto ruido, parámetros difíciles de mejorar sin aumentar las dosis a los pacientes. Al igual que ocurre con los Rayos X, las tendencias sociales actuales tienden a disminuir o evitar el uso de radiación ionizante en favor de otras técnicas de imagen.

La aparición de la PET, sin embargo, ha supuesto un reverdecimiento de las esperanzas depositadas en la Medicina Nuclear, ya que ofrece una mejora notable, sobre todo en Oncología, respecto a las exploraciones convencionales mediante SPECT. Los avances esperados 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. Otro avance será la generalización del uso de la PET, sobre todo en indicaciones oncológicas, donde parece tener cada vez más interés clínico.

A corto plazo, y en términos más cercanos al paciente medio, el avance más importante que podríamos esperar es el abaratamiento de los estudios con PET, cuyo elevado precio actual dificulta enormemente su utilización y la obtención de estadísticas fiables sobre posibles indicaciones.

Queremos resaltar finalmente que, pese a la mala prensa del adjetivo "radiactivo" que califica a los trazadores en Medicina Nuclear, las dosis de radiación recibidas por los pacientes son muy pequeñas, inferiores a la de una radiografía de tórax, por ejemplo, y mucho menores que las derivadas del uso del TC [1]. Sin embargo, es innegable que la Medicina Nuclear, como la Radiología, utiliza radiación ionizante. Ante ella no existe un límite de dosis seguro: dosis bajas tienen una probabilidad también baja, pero no nula, de producir efectos indeseables.

Por otra parte, y a diferencia de las modalidades "estructurales", la evolución de la tecnología permite que las imágenes de la MN, sean preferentemente funcionales con la ventaja que ello conlleva, aunque también pueden ser anatómicas o morfológicas o mixtas.

El resultado final, producirá:

- una significativa mejoría del flujo de trabajo del servicio que realiza la prueba,
- una drástica reducción de errores (de transcripción de informes, de datos del paciente, de pérdidas de copias, etc), y
- la mayor satisfacción del médico que prescribe la prueba.

2.2.4. La Ecografía

Es una técnica comúnmente llamada ultrasonido por la utilización de ondas ultrasónicas (energía mecánica) para obtener la imagen. 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 [2].

En sus inicios sólo se podían obtener imágenes en una dimensión, 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. Si bien la resolución geométrica es inferior en relación con otras modalidades (Tomografía computada e imágenes por resonancia magnética nuclear), 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, la radiación utilizada es no ionizante, y por tanto inocua; la imagen es tomográfica ya que sólo se observa un plano de corte; y lo que se representa en la imagen son las transiciones entre tejidos de diferente densidad.

Otro avance reciente, muy llamativo pero de escasa relevancia diagnóstica por el momento, es la ecografía tridimensional, generada mediante el movimiento controlado del transductor, que permite combinar diferentes planos de corte.

La imagen ultrasónica o ecografía es de baja resolución y poco contraste, por lo que desde un punto de vista puramente técnico podría considerarse muy inferior a otras modalidades. Sin embargo, el hecho de no causar daño, sencillez de realización y reducido precio son ventajas que hacen de ella una técnica de primera elección en diversas situaciones (urgencias, embarazo, etc.).

Sus desventajas son la reducida resolución espacial que proporciona, su escaso contraste y cierta dificultad de interpretación. Una variante de la técnica, llamada ecografía doppler, utiliza el efecto físico del mismo nombre para proporcionar información sobre el flujo sanguíneo, ofreciendo así una nueva fuente de contraste, muy útil en ciertas enfermedades cardiacas (valvulares, congénitas).

La ecografía, como todas las modalidades de imagen médica, experimenta continuos avances técnicos. Sin embargo, se encuentra con limitaciones físicas de base, la causa más importante de la limitación en resolución espacial es la longitud de onda del ultrasonido utilizado, del orden de varios milímetros. La disminución de esta longitud de onda (aumento de la frecuencia del ultrasonido) disminuye notablemente su penetración, impidiendo obtener información de objetos profundos.

2.2.5. Imágenes por Resonancia Magnética

La imagen de resonancia magnética (MRI) es una tecnología de imagen, que utiliza un campo magnético y ondas de radiofrecuencia para producir imágenes de órganos y estructuras internas.

Es una técnica que comenzó a ser desarrollada en pleno auge de la aparición clínico-radiológica de la TC, y, en la actualidad, se puede considerar su impacto de diagnóstico, como superior a la misma TC y otras técnicas como las de medicina nuclear y el ultrasonido.

Si bien las primeras investigaciones se iniciaron poco antes de la segunda guerra mundial, los primeros escáneres de MR para estudiar el cerebro humano aparecieron recién a finales de los 70's, seguidos poco tiempo después por los de cuerpo completo [3].

Las dos posibilidades genéricas más importantes de la MR, son:

- Las técnicas de imagen (IRM), y
- Las técnicas de análisis espectrométrico (SRM).

La imagen por resonancia magnética, tiene como equipo necesario, denominado también tomógrafo, un imán, un sistema de antenas y una computadora. El imán crea el campo magnético, las antenas recogen y amplifican la señal y la computadora permite representar la imagen y tratarla, o bien, realizar el análisis espectrométrico. Podemos decir que es un fenómeno donde determinados núcleos atómicos son capaces de absorber y emitir energía electromagnética (ondas de radio) de una frecuencia muy precisa (resonancia) cuando se

someten a un intenso campo magnético. Es decir que pueden absorber selectivamente esa energía electromagnética. Esta capacidad permite generar una señal que, luego de ser tratada en una computadora (como lo hace la TC), produce una imagen.

El funcionamiento de estos sistemas es bastante complejo: El paciente debe colocarse en el hueco de un potente campo magnético constante, habitualmente generado mediante un electroimán superconductor. El volumen muestra se "ilumina" con impulsos de ondas de radio cuya frecuencia corresponde a la de resonancia del núcleo de hidrógeno, aunque hay varios elementos de interés cuyos núcleos presentan el fenómeno de resonancia magnética (hidrógeno, fósforo, sodio, etc.), por el momento en los sistemas para uso clínico se trabaja sólo con el núcleo de hidrógeno, devolviendo después los tejidos a esta energía que es captada desde el exterior por una bobina o antena. La información espacial necesaria para generar imágenes se introduce a base de modificar ligeramente el valor del campo magnético en función de la posición, mediante las llamadas bobinas de gradiente de campo. Una computadora reconstruye las imágenes a partir de esta información.

La conformación y la calidad de la imagen se basan en 3 parámetros fundamentales:

- La densidad protónica o spin,
- el tiempo de relajación longitudinal (T1), y
- el tiempo de relajación transversal (T2)

Lo que permite separar los diferentes tejidos (grasa, músculo, agua, hueso, etc.). La comparación de las imágenes en las secuencias T1 y T2, es la que proporciona la información de la imagen, siendo la densidad protónica, una secuencia intermedia.

La imagen se conforma después de determinar la intensidad de cada píxel en un plano determinado en escala de grises y se almacena en el disco duro de la computadora. Una vez realizado el almacenaje, se puede imprimir en película o guardar en otros medios físicos para recuperarla más adelante.

Este equipo proporciona una buena resolución espacial y un excelente contraste. Puede proporcionar imágenes morfológicas y también funcionales, y su gama de aplicaciones se amplía de día en día, invadiendo el terreno de otras modalidades (TC, angiografía, ecografía, etc.).

El interés por la resonancia magnética se debe al hecho de que el equipo utiliza una radiación no ionizante que permite una buena discriminación de los tejidos y la adquisición tridimensional de una zona del cuerpo. El contraste de los tejidos, puede mejorarse

adicionalmente mediante la utilización de materiales de contraste paramagnéticos inyectados en el cuerpo. El campo de aplicación de esta modalidad es muy amplio y la limitación principal son los costos de los equipos.

Además de sus características físicas, se muestra más sensible que otras técnicas por la posibilidad de generar cortes más finos y en varios planos, no presentar artefactos debidos al aire o a los huesos y permitir visualizar y cuantificar los flujos de manera directa, con la ventaja adicional de no utilizar radiaciones ionizantes.

Debido a sus posibilidades, no presenta las limitaciones propias del tamaño de la región anatómica a explorar y se pueden obtener imágenes volumétricas sin alterar la posición del paciente, pudiendo obtener imágenes múltiples, en múltiples planos.

En contraposición de lo anteriormente dicho, la electrónica de los equipos de MRI es compleja y los imanes para generar el campo magnético de la potencia requerida son muy grandes y bastante costosos. Esto hace que la disponibilidad del método sea todavía relativamente baja, lo que tiende a limitar su uso.

Sin embargo, los avances más interesantes en MRI se están produciendo debido a la introducción de nuevos mecanismos de contraste. Para imagen morfológica o estructural se pueden generar diferentes contrastes, capaces de realzar o atenuar diferentes estructuras o tejidos (líquido, grasa, etc.).

Para tener la capacidad de estudiar el flujo sanguíneo (angiografía por resonancia magnética, MRA) se puede incluso no utilizar medios de contraste y aún más importante es la reciente posibilidad de obtener también imágenes funcionales de diversos tipos, por ejemplo de activación cerebral (Functional MRI, fMRI).

2.2.6. La Imagen Molecular

El concepto de imagen molecular [3], es la acción conjunta de las posibilidades que ofrecen:

- la imagen médica en sí,
- la bioquímica, y
- la informática, a las que últimamente se han sumado
 - la genómica,
 - la proteómica, o
 - la inteligencia artificial.

La acción conjunta de estos recursos otorga un valor añadido a la exploración de las diferentes modalidades por separado o bien formando parte del concepto de "multimodalidad" aprovechando la información conjunta y fusionada de técnicas estructurales, (TAC, MR, etc.) con técnicas funcionales (PET, etc.). Estas ventajas no sólo tienen implicaciones diagnósticas, si no también terapéuticas.

En una amplia definición, se podría decir que se trata del estudio de procesos bioquímicos y posibles rutas metabólicas tanto a nivel anatómico, como a nivel molecular (organismos, órganos, tejidos y células) que pueden provocar alteraciones en la expresión genética de los procesos biológicos normales que realizan los tejidos vivos (replicación, migración, transducción de señal, etc.) y que pueden ser originadas por factores propios (envejecimiento, defectos hereditarios, etc.) o mediados por el entorno.

La imagen molecular será el resultado de la exploración de esos procesos y rutas metabólicas en el que se utilizan moléculas marcadas con trazadores (sonda molecular), empleando la sustancia cuya actividad biológica esté más relacionada del punto de vista bioquímico con la patología que se quiera analizar.

Las técnicas más apropiadas para la elaboración de las imágenes moleculares, son la PET (Tomografía por emisión de positrones) y el SPECT (tomografía por emisión de fotón único). Hasta ahora todo lo más era poder "reconocer" una tumoración pero ahora se trata de poder ver cómo, cuándo y de qué forma se origina y desarrolla su agresividad y su extensión, lo cual conducirá sin duda a un tratamiento más eficaz.

En definitiva, se obtendrán diagnósticos más rápidos y certeros, por medio de técnicas no invasivas, que permitan individualizar los tratamientos y probablemente predecir su resultado.

2.2.7. El Futuro de la Imágenes Médicas

Es innegable la importancia que ha adquirido la imagen médica como herramienta de diagnóstico.

En ocasiones, la combinación de materias aparentemente (al menos hasta hace algunos años) poco relacionadas entre sí, permiten llegar a un nuevo nivel.

Todas las modalidades de diagnóstico por imagen de hoy en día jugarán un papel importante en el futuro. Los ultrasonidos, el Escáner (TAC) y la fluoroscopia digital, servirán para guiar microcatéteres en zonas anatómicas para enviar a través de ellos inhibidores angiogénicos.

Las microburbujas será casi un medio seguro de envío. Una vez dentro del cuerpo 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. Los inhibidores angiogénicos se usarán y la imagen molecular monitorizará su eficacia, proveyendo información sobre la expresión proteica y genética que redefinirán los papeles de la confirmación patológica e histológica.

El papel tradicional de la imagen como elemento exclusivamente diagnóstico, sin disminuir en su importancia, dará paso a otro todavía más relevante como es la monitorización de terapias en la farmacogenómica.

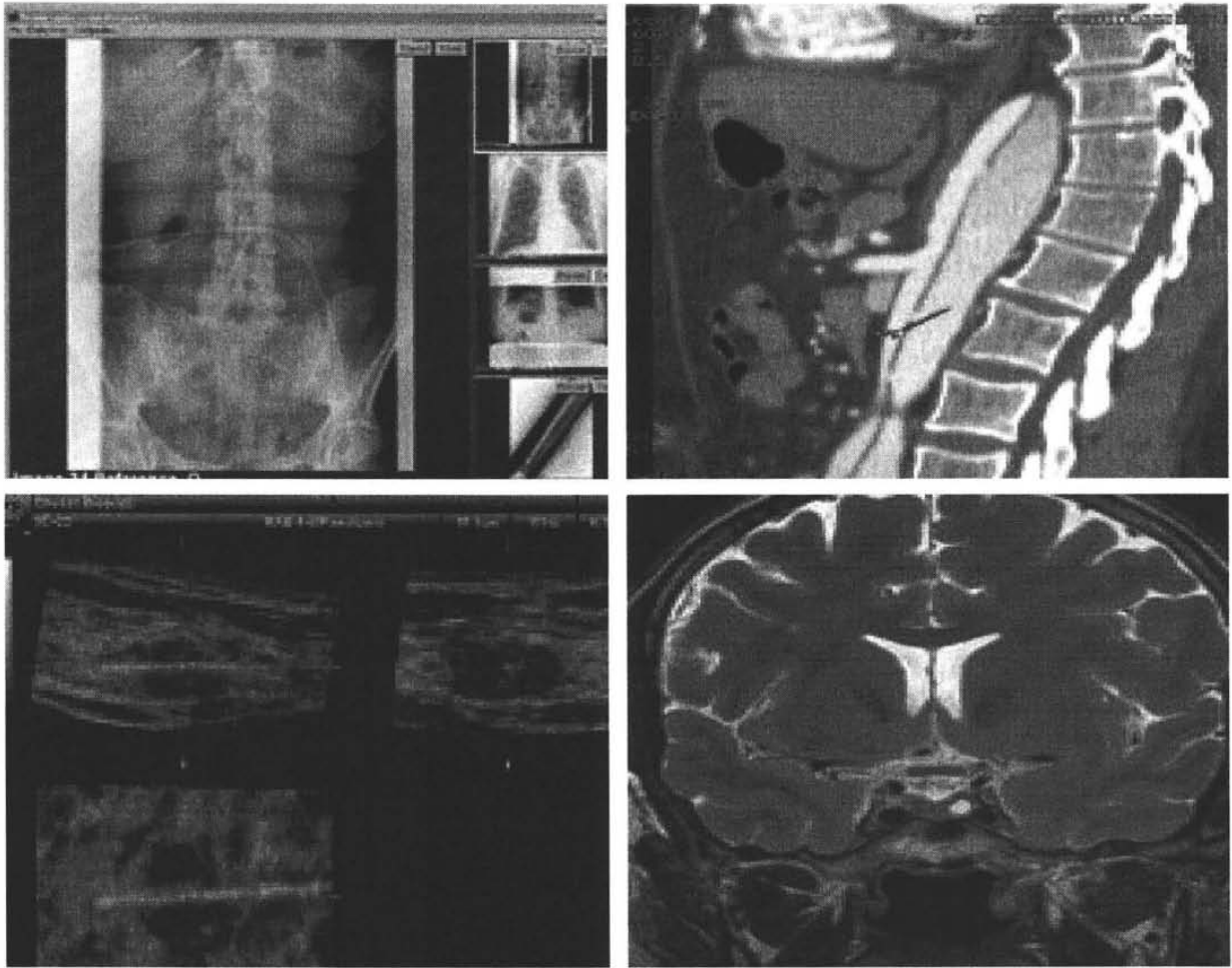
La RM dará excelentes contrastes en las imágenes, suficiente para evaluar imágenes morfológicas y de perfusión, para identificar espectroscópicamente casos con diagnóstico de cáncer y en su monitorización terapéutica.

El gran futuro que le espera a la Imagen en Medicina se basa en que la enfermedad es un proceso genético y que hay un nuevo modelo para radiología, no ligada al diagnóstico sino a las terapias moleculares.

Los métodos a emplear son tan sensibles que la medición de la enfermedad a nivel molecular puede en potencia ser detectada y corregida antes de que aparezcan los primeros síntomas en el paciente.

Mientras la medicina convencional del siglo XX trata los efectos de la enfermedad, en el siglo XXI lo que tratará son sus causas ayudada de manera esencial por las técnicas complementarias de imagen molecular.

La Imagen Molecular permitirá la detección precoz y las características de la enfermedad así como una valoración de la eficacia de los tratamientos.



A	B
C	D

Figura 2.3: Algunas modalidades de imágenes médicas. **A.** Radiología digital. **B.** Tomografía computarizada. **C.** Ecografía. **D.** Resonancia magnética [3].

Fusión de Imágenes Médicas

3.1. Conceptos de fusión

El concepto de fusión de datos es fácil de entender. Sin embargo, su significado exacto varía a partir de un científico a otro. Varias palabras han aparecido, por ejemplo: mezcla, combinación, sinergia, e integración [4]. Todos se refieren más o menos al mismo concepto pero sin embargo se notan diferentes.

Varias definiciones se pueden encontrar en la literatura. Pohl y Van Genderen [5] proponen que "la fusión de imágenes es la combinación de dos o más imágenes diferentes para formar una nueva imagen usando cierto algoritmo". Mangolini [6] amplía la fusión de los datos a la información en general y también se refiere a la calidad. Él define la fusión de los datos como "conjunto de métodos, herramientas y de 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".

Hall y Llinas [7] también se refieren a calidad de la información en su definición pero todavía se centran en los métodos: "las técnicas de fusión de datos combinan los datos de múltiples sensores, y relaciona la información asociadas de las bases de datos, para alcanzar una mejor exactitud y conclusiones más específicas que las que se pueden alcanzar solamente por el uso de un solo sensor".

El departamento de defensa de los Estados Unidos [8] indicó que la "fusión de los datos tiene niveles múltiples, en el proceso multifacético se intercambia con: la detección automática, la asociación, la correlación, la estimación, y la combinación de datos e información de múltiples fuentes".

La calidad se obtiene a través del proceso de refinamiento [9] donde Li, Manjunath y Mitra escriben: "la fusión se refiere a la combinación de un grupo de sensores con el objetivo de producir una sola señal de la mayor calidad y fiabilidad".

Se indica a menudo que la fusión de datos tiene lugar en tres niveles: pixel, atributos, y decisión [5], [6], [10], [11], estos niveles se ilustran en la Figura 3.1.

El nivel de pixel corresponde cuando en imágenes de diferentes fuentes se combinan pixel a pixel. La fusión de imágenes a nivel de pixel representa el nivel mínimo de procesamiento para fusionar parámetros físicamente medidos. En la fusión de imágenes a nivel pixel, el proceso de fusión deberá preservar la información relevante de las imágenes de entrada. En la fusión a nivel de atributo (Figuras o características) requiere la extracción de objetos reconocibles en diversas fuentes de datos, por ejemplo usando un proceso de segmentación. Las Figuras corresponden a características extraídas de las imágenes iniciales, tales que proporcionen forma y vecindad. Estos objetos similares (ejemplo regiones) son asignados entre si y entonces son fusionados para futuras evaluaciones usando aproximaciones estadísticas o redes neuronales [24]. Por último, el nivel de decisión donde una simple imagen fuente es procesada independientemente y sus resultados se combinan usando sus ponderaciones de significancia o exactitud. 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.

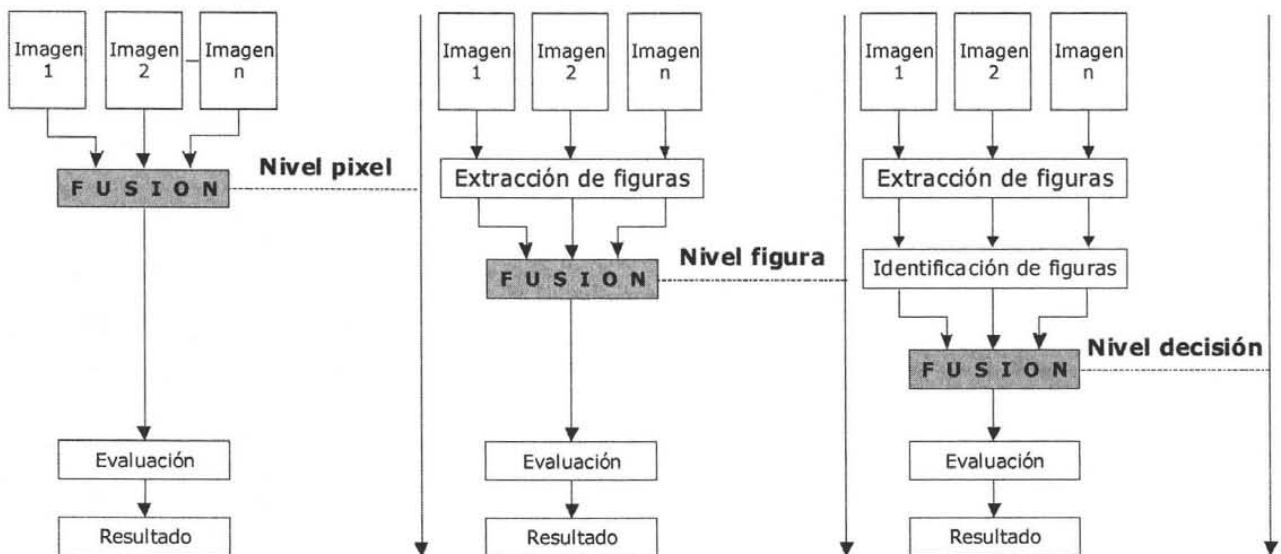


Figura 3.1: Nivel de procesamiento de fusión de imagen.

Basado sobre los trabajos de Buchroithner [12] y Wald [13], la definición siguiente fue adoptada en enero de 1998: "la fusión de datos es un marco formal donde están expresados los medios y las herramientas para la asociación de los datos originales de fuentes diversas.

Tiene como objetivo obtener información de mayor calidad; la definición exacta de 'mayor calidad' dependerá de la aplicación."

Entonces podemos decir que la fusión de imágenes, para nuestro caso, es el proceso de combinar la información procedente de distintas modalidades para producir una única imagen final. El propósito de esta fusión puede ser doble:

- i. Permite presentar de manera conjunta e integrada toda la información (anatómica y funcional) proporcionada por diferentes modalidades de imagen, simplificando su interpretación; y
- ii. Facilita la obtención de datos cuantitativos a partir de informaciones complementarias.

El paso fundamental para integrar imágenes diferentes es el denominado registro. Los algoritmos de registro establecen una referencia geométrica común a partir de las distintas imágenes del objeto, mediante transformaciones de coordenadas, más o menos complejas.

Todos los métodos de registro se basan en la identificación de referencias comunes en las imágenes. Estas referencias pueden ser estructuras propias del objeto (llamadas marcadores anatómicos o 'landmarks'), o bien dispositivos externos adosados al mismo. En este último caso los algoritmos de registro utilizados son sencillos y proporcionan una fiabilidad bien controlada; en el primero, los métodos son experimentales en muchos casos, a veces incluso pueden producir resultados erróneos, lo que hace necesario supervisar cuidadosamente el proceso. Esta es la razón más importante por la que aún no se ha generalizado su uso.

Como es lógico, si se emplean marcadores externos es necesario prepararlos antes de adquirir la imagen, por lo que se habla de registro "a priori". Los métodos de registro "a posteriori" tienen la ventaja de poderse aplicar a imágenes obtenidas con anterioridad, sin marcadores externos. El inconveniente, ya mencionado, es la inestabilidad de los métodos matemáticos involucrados.

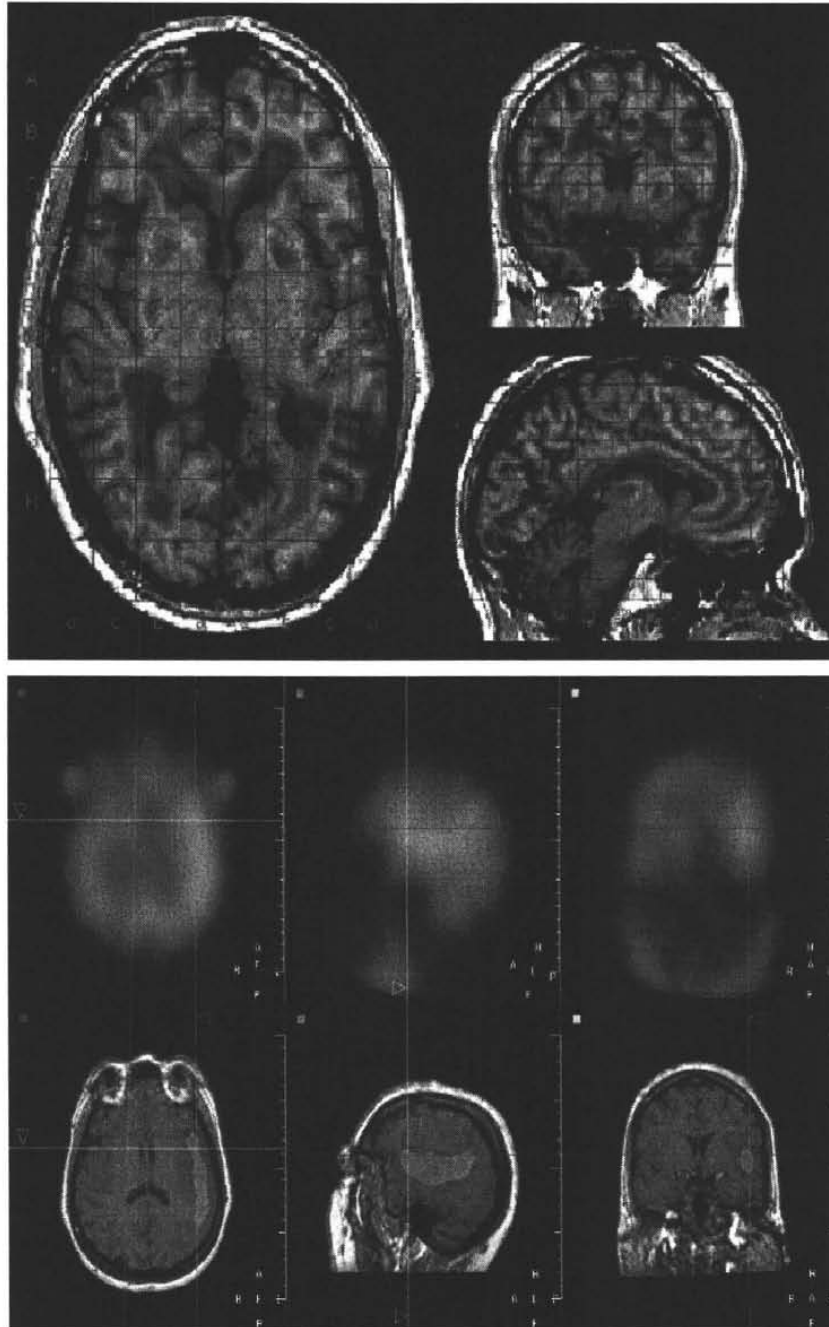
El número de aplicaciones clínicas para los métodos de fusión de imágenes multimodales crece de día en día, existiendo ya algunas muy interesantes y bien consolidadas.

Por el momento, la mayoría de ellas se refieren al sistema nervioso central, ya que la rigidez de la caja craneal facilita la fusión de las imágenes.

En diagnóstico neurorradiológico, la fusión de estudios de imagen cerebral con atlas anatómicos ha sido objeto de múltiples publicaciones, bien como ayuda al diagnóstico o como herramienta docente. Sin embargo, no parece haber alcanzado aún la mayoría de edad en términos de fiabilidad y facilidad de uso: muchos autores rechazan su empleo por no tener en

cuenta correctamente las diferencias anatómicas individuales. Además, no son aplicables cuando existe una importante patología estructural de base, como un tumor, infarto, atrofia, etc., que distorsiona significativamente la anatomía cerebral.

Para el especialista que debe interpretar una imagen funcional (nuclear, habitualmente) la posibilidad de contar con un "patrón" anatómico subyacente de alta calidad (RM, habitualmente) es extraordinariamente útil (Figura 3.2, A). Es cierto que en muchos casos – mal denominados "de rutina"– la interpretación multimodalidad no modificaría el diagnóstico, pero queda un porcentaje significativo que puede beneficiarse enormemente de estas técnicas. Por ejemplo, estudios cerebrales por SPECT y PET, estudios cardíacos para detección de isquemia y, en general, todo análisis comparativo cuantitativo que implica imágenes obtenidas en dos momentos diferentes. En algunos centros se realiza la fusión de imagen de forma habitual en ciertas situaciones clínicas: pacientes epilépticos farmacorresistentes candidatos a cirugía (Figura 3.2, B), para seguir la eficacia del tratamiento en pacientes con sospecha de recidiva y persistencia tumoral, así como en múltiples patologías psiquiátricas y neurológicas con discordancia clínico-radiológica, por ausencia de hallazgos en el CT o la MRI que justifiquen la sintomatología [1].



A

B

Figura 3.2: **A.** Estudio PET fusionado sobre una MRI. El detalle anatómico proviene de la resonancia magnética, el color indica nivel de metabolismo y proviene de la imagen PET. La rejilla superpuesta corresponde al sistema de coordenadas de Talairach, que permite identificar áreas funcionales cerebrales. **B.** Paciente con epilepsia resistente al tratamiento médico. Arriba: La imagen SPECT localiza el foco con poca precisión. Abajo: La fusión de SPECT con MRI permite planificar la cirugía con mayor seguridad [1].

3.2. Tipos de fusión

Existen diferentes tipos de fusión, todo dependiendo del análisis que se desea hacer. A continuación se detallan un conjunto de métodos de fusión con base a la clasificación de los siguientes autores:

Isabelle Bloch [14], establece que la fusión se puede clasificar en función de los valores a combinar, tales como:

- 1) Operadores de funciones constantes independientes del contexto:
 - i) Fusión probabilística y bayesiana, empleando un operador matemático de producto de probabilidad.
 - ii) Conjuntos difusos y teoría de la probabilidad, donde se emplean tres familias de operadores.
 - (a) Norma triangular
 - (b) Co-norma triangular
 - (c) Media
 - iii) Teoría de la evidencia de Dempster-Shafe, empleando el operador de sumas ortogonales.

- 2) Operadores de fusiones variables independientes del contexto; entre los mencionados son:
 - i) Conjuntos difusos y teoría de la probabilidad, haciendo uso del operador de suma simétrica.
 - ii) Sistemas "MYCIN-like", usando el operador de combinación de factores seguros.

- 3) Los operadores dependientes del contexto:
 - i) Operador de dependencia del conflicto.
 - ii) Operador de dependencia en la fiabilidad de la fuente.

Vemos que la autora clasifica la fusión, dependiendo del contexto, es decir, solo si se toma en cuenta el valor que esta estudiando o si se toma en cuenta el entorno.

Tapiador [15], define la fusión de datos como un conjunto de diferentes técnicas que se emplean para mejorar ciertas características en teledetección. La fusión de datos se emplea en observación de la tierra, astronomía y en cualquier ámbito en el que se requiere el procesamiento digital de imágenes. Los métodos mencionados son:

- 1) Intensidad, brillo y saturación.
- 2) Sustitución de componentes principales (PCS, Principal component substitution).
- 3) Contribución espectral relativa. Algoritmo de Brovey o P+XS.
- 4) Cociente Sintético Variable (SVR).
- 5) Cociente Sintético Variable mejorado (SVR).
- 6) Wavelets.
- 7) Fusión de datos mediante entropía (EDF).
- 8) Modelos Markovianos.
- 9) Filtros pasa altas.
- 10) Normalización de imágenes.
- 11) Correlación local.
- 12) Integración contextual.
- 13) Pirámides laplacianas generalizadas.

Oliver Rockinger [16,17], propone otra clasificación basada a nivel píxel-píxel, donde se menciona los siguientes métodos:

- 1) Superposición final.
- 2) Métodos no lineales.
- 3) Aproximación a la optimización.
- 4) Redes neuronales y artificiales.
- 5) Imágenes piramidales.
- 6) Wavelets.

3.3. Métodos Representativos de fusión

Como se mencionó anteriormente y mediante la clasificación expuesta por O. Rockinger [16, 17], López C. [18] describe algunos métodos de fusión más representativos.

Uno de los problemas al combinar las imágenes es conservar los patrones de la misma, es decir, los detalles importantes de los componentes en la imagen se deben preservar en la imagen resultante compuesta. El proceso de fusión no debe introducir patrones que puedan interferir con el análisis siguiente. En este estudio se utilizan diversas técnicas de fusión de imágenes, ocho son los de interés, las pirámides laplaciana y laplaciana FSD (Filter-substrate-decimate), la de cociente, la de contraste, y la de gradiente; la transformadas wavelets y morfológica.

Cuando se hace un acercamiento en una imagen, claramente se ven las subestructuras, aunque se pierda la nitidez en los contornos, contrariamente, cuando se hace un alejamiento para observar toda la imagen, el detalle de la escena se pierde, por lo que el detalle de la imagen puede ser observado sólo a cierta escala, con un cierto intervalo de resolución espacial. Por esta razón, es deseable representar una imagen sobre un intervalo de escalas, dependiendo del contenido estructural.

Una pirámide es una secuencia de imágenes donde cada imagen se filtra y representa una copia submuestreada de sus predecesoras. El término pirámide multiresolución viene de la relación que se establece entre los niveles sucesivos de la pirámide que son copias reducidas de la imagen de entrada.

Las técnicas usadas para generar pirámides pueden clasificarse en tres tipos: lineales, de transformaciones y de morfología matemática. Los tipos lineales se describen a través de las pirámides laplaciana y gaussiana, la FSD, la de cociente, la de contraste y la de gradiente; mientras que la transformada wavelets es una técnica de transformación y por último la pirámide morfológica describe la técnica de morfología matemática.

3.3.1. Pirámide laplaciana y gaussiana.

Cada imagen de una pirámide es obtenida con un filtro pasa bajas y es una copia submuestreada (decimación) de la imagen previa, resultando un grupo de imágenes con resolución espacial diferente por filtración. El filtro lineal pasa bajas es el más usado para la generación piramidal convolucionado con el kernel gaussiano. La pirámide formada usando esta técnica se define como pirámide gaussiana.

El primer paso en la generación de una pirámide laplaciana es la aplicación de un filtro pasa bajas en la imagen original g_0 para obtener la imagen g_1 . La imagen g_1 es una versión

reducida de g_0 en la que tanto la resolución, como la densidad de muestreo disminuyen, el proceso se repite de manera iterativa. En forma similar g_2 es una versión reducida de g_1 . El filtrado se realiza por un proceso equivalente a la convolución con un miembro de una familia de funciones locales simétricas. Un importante miembro de esta familia es la densidad de probabilidad gaussiana. Así la secuencia de imágenes g_0, g_1, \dots, g_n se llama pirámide gaussiana.

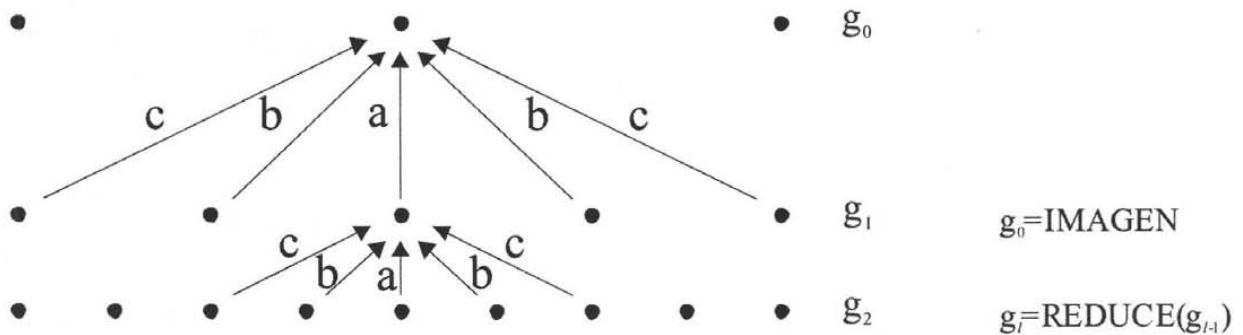


Figura 3.3. Representación gráfica del proceso para generar una pirámide gaussiana (ver texto) [18].

La representación gráfica del proceso para generar una pirámide gaussiana se muestra en la Figura 3.3. Cada punto en la línea representa un nodo con un nivel en la pirámide. El valor de cada nodo en el nivel cero es sólo el nivel de gris de un pixel correspondiente en la imagen. El valor de cada nodo en un nivel más alto son los coeficientes de ponderación promedio de los valores de los nodos en el siguiente nivel menor. El mismo patrón de ponderación es usado para generar todos los niveles.

3.3.1.1. Generación de una pirámide gaussiana.

Suponiendo que la imagen se representa por un arreglo g_0 que contiene C columnas y R líneas de pixeles, cada pixel representa la intensidad de luz en el correspondiente punto de la imagen por un entero I entre 0 y $K-1$. Esta imagen comienza en el nivel inferior o cero de la pirámide gaussiana. El nivel 1 de la pirámide contiene g_1 , el donde es reducido por un filtrado pasa bajas de g_0 . Cada valor con un nivel 1 se calcula con los promedios ponderados de valores en nivel 0, y con una ventana de filtrado de 5×5 . Cada valor en el nivel 2 se

representa como g_2 , y entonces se obtienen los valores del nivel 1 aplicando el mismo patrón de ponderación. La Figura 3.3 representa este proceso en una dimensión. El tamaño de la función de ponderación no es crítico aunque una ventana de 5X5 da un filtrado adecuado.

El proceso de generación de cada nivel a partir del anterior se define por la función

$$g_k = REDUCE(g_{k-1}), \quad (3.1)$$

para niveles $0 < l < N$ y nodos $i, j, 0 \leq i < C_l, 0 \leq j < R_l$

$$g_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) g_{l-1}(2i + m, 2j + n), \quad (3.2)$$

Donde N se refiere al número de niveles de la pirámide, mientras que C_l y R_l son las dimensiones del nivel l th

3.3.1.2. Generación del kernel

El patrón de ponderación de 5X5, w , se usa para generar el arreglo piramidal de sus predecesores. Este patrón de pesos de llama kernel generador, y se elige de acuerdo a ciertas condiciones. La función ponderada $w(m, n)$ es separable, y se define

$$w(m, n) = w'(m)w'(n). \quad (3.3)$$

Los coeficientes de ponderación son siempre reales, se eligen para ser simétricos y normalizados, por ejemplo la función w' normalizada es

$$\sum_{m=-2}^2 w'(m) = 1, \quad (3.4)$$

y la simétrica es

$$w'(i) = w'(-i) \text{ para } i=0,1,2. \quad (3.5)$$

Además, existe una condición adicional de igual contribución, ésta estipula que la suma de todas las contribuciones desde un nodo dado n es independiente de n , es decir, todos los nodos a un nivel dado contribuyen con la misma ponderación total ($=1/4$) para nodos del siguiente nivel. Supóngase que $w'(0) = a$, $w'(-1) = w(1) = b$ y $w'(-2) = w'(2) = c$. En este caso, iguales contribuciones requieren que $a+2c=2b$. Estas tres condiciones se satisfacen cuando w' se define como

$$w'(0) = a, \quad (3.6)$$

$$w'(1) = w'(-1) = \frac{1}{4}, \quad (3.7)$$

$$w'(2) = w'(-2) = \frac{1}{4} - \frac{a}{2}, \quad (3.8)$$

los diferentes valores considerados para a son 0.6, 0.5, 0.4 y 0.3.

3.3.1.3. Funciones de ponderación equivalentes o promedio

La generación de una pirámide iterativa es equivalente a la convolución (caso simétrico) de la imagen g_0 con un grupo de datos de "funciones de ponderación equivalente" h_l

$$g_l = h_l * g_0, \quad (3.9a)$$

o

$$g_l(i, j) = \sum_{m=-M_l}^{M_l} \sum_{n=-M_l}^{M_l} h_l(m, n) g_0(i2^l + m \cdot j2^l + n). \quad (3.9b)$$

El tamaño de M_l de la función de ponderación equivalente se duplica de un nivel al siguiente, al igual que lo hace la distancia entre muestras. El efecto de la convolución de una imagen con una función de ponderación equivalente h_l equivale a generar una imagen aplicando un filtro pasa bajas.

Las funciones de ponderación equivalente de los niveles de la pirámide gaussiana 1, 2, 3 se muestran en la Figura 3.4 para el caso de $a=0.4$. La forma de la función equivalente converge rápidamente a una forma característica conforme se alcanzan niveles sucesivos más altos en la pirámide, de tal forma que sólo cambia su escala. En este caso, para comparar, el eje de escala se ajustó por un factor de dos. El parámetro a del kernel generador es 0.4 y los resultados de la función de ponderación equivalente es muy parecida a la función gaussiana.

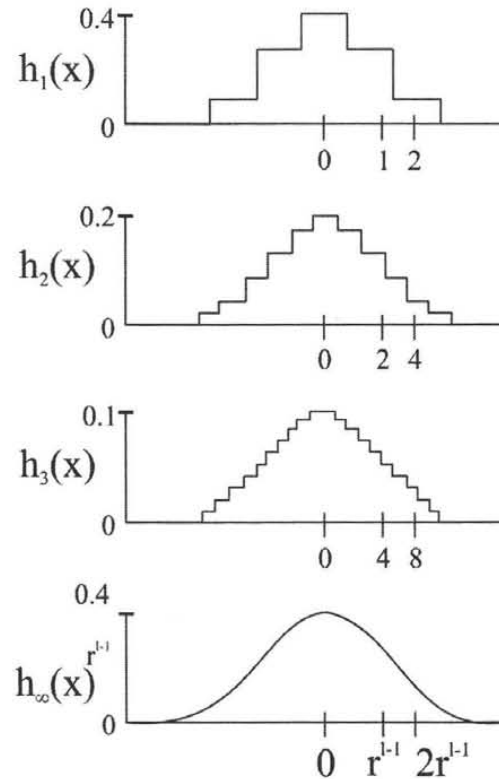


Figura 3.4. Funciones de ponderación equivalentes $h_l(x)$ para nodos en el nivel 1, 2, 3 e infinito de la pirámide gaussiana, el parámetro $a=0.4$ [18].

Las formas características para cuatro elecciones se muestran en la Figura 3.5. Las funciones de ponderación equivalentes son gaussianas, particularmente cuando $a=0.4$. La función es triangular cuando $a=0.5$ y para $a=0.3$ es una gaussiana amplia. Para $a=0.6$ la función es trimodal.

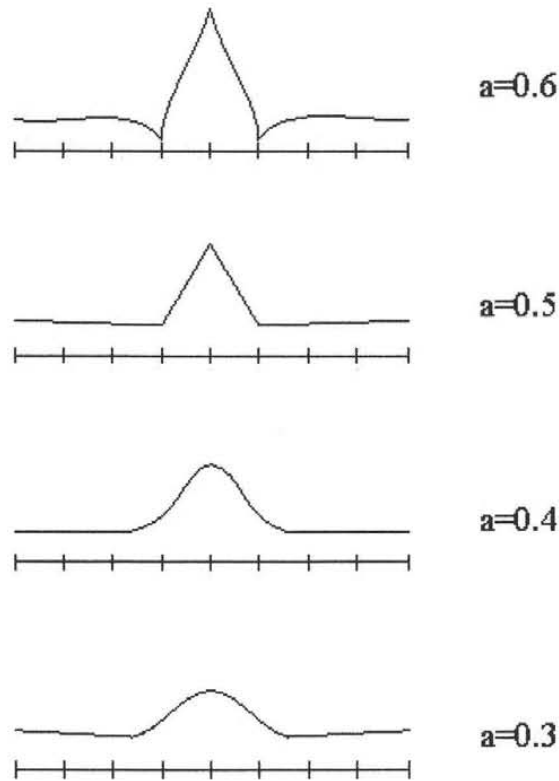


Figura 3.5. La forma de función de ponderación equivalente depende del parámetro α . [18]

3.3.1.4. Interpolación de la pirámide gaussiana

Si se define la función *EXPAND* como la inversa de *REDUCE*, el efecto de *EXPAND* es $(M+1)$ por $(N+1)$ arreglos en $(2M+1)$ por $(2N+1)$ arreglos por interpolación de nuevos nodos de valores entre los valores dados. De aquí, *EXPAND* aplicado al arreglo g_l de la pirámide gaussiana podría dar un arreglo $g_{l,1}$ el cual es del mismo tamaño que g_{l-1} .

Sea g_{l-n} el resultado de la expansión g_l n veces. Entonces

$$g_{l,0} = g_l, \quad (3.10)$$

y

$$g_{l,n} = \text{EXPAND}(g_{l,n-1}), \quad (3.11)$$

EXPAND significa que para niveles $0 < l < N$ y $0 \leq n$ y nodos $i, j, 0 \leq i < C_{l-n}, 0 \leq j < R_{l-n}$

$$g_{l,n}(i, j) = 4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) \cdot g_{l,n-1} \left(\frac{i-m}{2}, \frac{j-n}{2} \right), \quad (3.12)$$

los términos $(i-m)/2$ y $(j-n)/2$ son enteros no incluidos en esta suma.

Si aplicamos la función *EXPAND* l veces a la imagen g_l , entonces se obtiene g_{l-1} , donde es del mismo tamaño de la imagen original g_0

3.3.2. Pirámide estándar gaussiana y laplaciana

Para crear una pirámide gaussiana de una imagen I , dado G_k ($k^{\text{ésimo}}$ nivel de la pirámide gaussiana) hay que aplicar $G_0(i, j) = I(i, j)$ y para $k > 0$

$$G_k = [w * G_{k-1}]_{\downarrow 2}, \quad (3.13)$$

donde w , es el kernel generador y la notación $[...]_{\downarrow n}$ indica que el arreglo de la imagen en el paréntesis es submuestreado (decimación) por n , el símbolo $*$ es el operador convolución. Para todos los elementos, w debe ser la suma igual a la imagen I , por simplicidad, w se define como separable (w) donde $w = \hat{w} * \hat{w}$.

La diferencia entre dos niveles en la pirámide gaussiana es similar a un filtrado pasa alta sobre la imagen. Con el fin de obtener la diferencia entre dos niveles subsecuentes, el nivel de baja resolución deberá ser expandido al mismo tamaño de la imagen de alta resolución. Si esto se hace para cada nivel de la pirámide gaussiana, se crea una pirámide laplaciana, L_k ($k^{\text{ésimo}}$ nivel de la pirámide laplaciana), definido por

$$L_k = G_k - 4w * [G_{k+1}]_{\uparrow 2}, \quad (3.14)$$

usando la notación de convolución, donde $[\dots]_{\uparrow 2}$ indica un sobremuestreo por un factor de dos. Para llevar a cabo el sobremuestreo, se añaden líneas y columnas de ceros a las ya existentes. La convolución con w interpola valores de columnas y líneas de ceros. El kernel w usado en la reconstrucción de la pirámide gaussiana es el filtro 5X5. Siendo \hat{w} el filtro binomial de 3X3.

$$\hat{w} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \frac{1}{16}, \quad (3.15)$$

y

$$w = \hat{w} * \hat{w} = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \frac{1}{256}. \quad (3.16)$$

El valor usado para w se debe a las propiedades similares de una gaussiana, sin embargo, w puede ser definido de diferente manera.

La pirámide laplaciana también puede ser definida usando métodos de promedios ponderados. Dicha pirámide es una secuencia de imágenes de error L_0, L_1, \dots, L_N , donde cada imagen es la diferencia entre dos niveles de la pirámide gaussiana.

De aquí, para $0 \leq l < N$

$$L_l = G_l - EXPAND(G_{l+1}). \quad (3.17)$$

Para una pirámide linealmente filtrada, $EXPAND$ se puede definir como

$$g_l(i, j) = 4 \sum_{m, n=-2}^2 w(m, n) g_{l+1} \left(\frac{i+m}{2}, \frac{j+n}{2} \right), \quad (3.18)$$

donde sólo coordenadas enteras contribuyen a la misma. El nivel superior de la pirámide laplaciana es igual al nivel superior de la gaussiana $L_n = G_n$

El algoritmo de construcción de la pirámide reduce el ancho de banda en un octavo de nivel a nivel. La pirámide laplaciana es equivalente a un filtro pasa banda con ancho de banda igual a la distancia entre octavas (duplicar la frecuencia). Este filtrado sobre una imagen tiene el efecto de pérdida de nitidez. Esto se debe a que los bordes agudos tienen componentes de alta frecuencia que están en la parte menor de la pirámide. Cuando expandimos el tamaño de una imagen menor, la imagen mayor aparece borrosa.

La reconstrucción de una imagen a partir de su pirámide laplaciana es directa, es decir, la imagen original se recupera por la expansión de niveles altos de la pirámide e incorpora el siguiente nivel superior repetidamente. Esto se expresa como

$$\hat{G}_n = L_n, \quad (3.19)$$

y

$$\hat{G}_k = L_k + EXPAND(\hat{G}_{k+1}) \text{ para } k=n-1, n-2, \dots, 0. \quad (3.20)$$

La reconstrucción de la imagen fuente ocurre cuando $\hat{G}_0 = G_0$

3.3.3. Pirámide laplaciana FSD (Filter-Subtract-Decimate)

La pirámide laplaciana FSD se define como la diferencia entre el nivel gaussiano G_k y la copia filtrada del nivel gaussiano anterior a G_k antes de ser submuestreado para el siguiente nivel G_{k+1} . Por esta razón, la función *REDUCE*, eq. 3.1, no trabajará con una laplaciana FSD porque *REDUCE* realiza la operación de filtrado y submuestreo al mismo tiempo. Sea L_k (k^{esimo} nivel de la pirámide laplaciana FSD) definida por

$$L_k = G_k - w * G_k = [I - w] * G_k, \quad (3.21)$$

donde \hat{I} es una matriz de la misma dimensión de w y en que todos los valores son cero excepto el valor central, que es igual a uno. En el caso de la matriz de 3 por 3, la matriz \hat{I} queda definida como

$$\hat{I} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (3.22)$$

Al analizar el proceso de formación de la imagen en la laplaciana L_k , se revela que la pirámide gaussiana G_k , es secuencialmente convolucionada con w , submuestreada, sobremuestreada, convolucionada con w otra vez y entonces restada con ella misma (ecuaciones 3.13 y 3.14). Sin embargo si el paso de remuestreo no se realiza, el resultado es solo ligeramente diferente. Obsérvense las siguientes ecuaciones

$$\begin{aligned} L_k &= G_k - 4w * [[w * G_k]_{\downarrow 2}]_{\uparrow 2} \approx G_k - w * w * G_k \\ &= [1 - w * w] * G_k \\ &= [1 + w] * [1 - w] * G_k. \end{aligned} \quad (3.23)$$

De aquí, la pirámide laplaciana FSD puede ser convertida a pirámide laplaciana a través de la conversión

$$L_k \approx [1 + w] * L_k. \quad (3.24)$$

sin embargo, la pirámide laplaciana reconstruida no es un réplica exacta de la original.

3.3.4. Pirámides de cociente de paso bajo (Ratio of Low Pass Pyramids)

La pirámide de cociente bajo (RoLP) adquiere su nombre a partir de la relación que existe entre niveles sucesivos de la pirámide. Las pirámides RoLP son muy similares a la laplaciana, en vez de tomar la diferencia entre el nivel de una pirámide gaussiana, la RoLP toma la relación entre los niveles de una gaussiana. La pirámide RoLP, R_k , se define matemáticamente como

$$R_k = \frac{G_k}{EXPAND(G_{k+1})} \text{ para } k=n-1, n-2, \dots, 0, \quad (3.25)$$

y

$$R_n = G_n. \quad (3.26)$$

Cada nivel en la pirámide RoLP es la relación de dos niveles sucesivos en la pirámide gaussiana.

Sea \hat{G}_0 la imagen reconstruida de la pirámide RoLP. El proceso de reconstrucción es el inverso del proceso de construcción

$$\hat{G}_n = R_n, \quad (3.27)$$

y

$$G_k = R_k EXPAND(G_{k+1}) \text{ para } k=n-1, n-2, \dots, 0. \quad (3.28)$$

El proceso de reconstrucción para la pirámide RoLP es exacto, en otras palabras, $\hat{G}_0 = G_0$.

3.3.5. Pirámide de gradientes

El término de gradiente no es muy propio ya que una pirámide de gradientes es una colección de cuatro pirámides. Dado que D_{kl} , representa $k^{\text{ésimo}}$ nivel y $l^{\text{ésimo}}$ la orientación de la pirámide de gradientes para una imagen I , D_{kl} se obtiene de la convolución de G_k con el gradiente d_l también llamado filtro orientado de segunda derivada [23].

$$D_{kl} = d_l * [G_k + w * G_k], \quad (3.29)$$

donde

$$d_1 = [1 \quad -1], \quad (3.30)$$

$$d_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}}, \quad (3.31)$$

$$d_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad (3.32)$$

$$d_4 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}}. \quad (3.33)$$

La reconstrucción de una imagen a partir de una pirámide de gradiente requiere una pirámide laplaciana y una pirámide laplaciana FSD como paso intermedio. Cada nivel de la pirámide de gradiente D_{kl} es convertido a su correspondiente de segunda derivada en la pirámide o nivel orientado laplaciano L_{kl}

$$L_{kl} = -\frac{1}{8} d_1 * D_{kl}. \quad (3.34)$$

Una pirámide laplaciana FSD, L_k se forma entonces por la suma de sus pirámides orientadas

$$L_k = \sum_{l=1}^4 L_{kl}. \quad (3.35)$$

3.3.6. Pirámide de morfología matemática

La morfología matemática es una herramienta para extraer componentes de una imagen que sean útiles en la representación de la forma de una región. El lenguaje de la morfología matemática es la teoría de conjuntos, estos representan las formas de los objetos de una imagen. La localización de contornos se estima con filtros lineales aunque alteran la luminosidad del objeto. Por otro lado, los filtros morfológicos remueven detalles sin añadir influencias en las escalas de grises y son apropiados para extraer formas.

Como antes indicamos, la resolución jerárquica, llamada piramidal, es producida por un filtro pasa bajas y es entonces remuestreada para generar el siguiente nivel de resolución menor. Las bases de la pirámide morfológica requieren del teorema del remuestreo morfológico el

definido como una aproximación de filtrado relacionada con la imagen muestreada. Esto manifiesta qué clase de formas son preservadas y cuales son eliminadas.

Ahora bien, una imagen binaria discreta se representa en un espacio bidimensional entero (2D) en un plano discreto Z^2 , donde cada elemento de un conjunto es una dupla (vector bidimensional) cuyas coordenadas son (x,y) del pixel de una imagen. Sea $f(x,y)$ y $A(x,y)$ donde $f(x,y)$ es la imagen de entrada y $A(x,y)$ es el elemento estructural.

Los constructores de bloques de morfología son los operadores de dilatación (se denotada por el símbolo \oplus) y erosión (se denotada por el símbolo \ominus). A partir de ellos podemos componer las operaciones de apertura, ($f \circ A = (f \ominus B) \oplus B$, erosión seguida de dilatación) denotada por el símbolo \circ y cerradura ($f \bullet A = (f \oplus B) \ominus B$, dilatación seguida de erosión) denotada por el símbolo \bullet . Son estas dos operaciones las que tienen relación con la representación de la forma, la descomposición y la extracción de primitivas.

Usualmente se emplea la erosión y la dilatación por pares, ya sea la dilatación seguida de la erosión o viceversa. El resultado de esta aplicación sucesiva de erosiones (contracción del objeto) y dilataciones (expansión del objeto) es una eliminación de los detalles menores, sin distorsionar la forma global del objeto.

Los filtros morfológicos son secuencias de operaciones morfológicas que tienen propiedades especiales con respecto a las formas de la imagen. Algunos filtros morfológicos son idempotentes y crecientes. Idempotentes significa que las aplicaciones sucesivas del filtro deja el resultado sin cambios después que el filtro has sido aplicado por primera vez, similarmente al efecto de aplicar un filtro pasa banda lineal sobre una señal. Creciente significa que la operación mantiene relaciones de inclusión en las imágenes que transforma, o bien, si A es un subconjunto de B , entonces A filtrada por F es un subconjunto de B filtrado por F .

Los filtros morfológicos simples son transformaciones de apertura y cerradura. La apertura es una erosión seguida por una dilatación y generalmente suaviza el contorno de una imagen, eliminando protuberancias delgadas. La cerradura es una dilatación seguida por una erosión y éste tiende también a suavizar secciones de contorno pero, en oposición a la apertura y generalmente fusiona separaciones estrechas y profundas, eliminando pequeños huecos y rellenando agujeros de contornos.

Hasta aquí se han definido los filtros morfológicos, por lo que se puede describir la generación de pirámides morfológicas de manera similar a la generación de pirámides gaussianas. Para generar la pirámide, la imagen actual es filtrada y submuestreada. Sea I la imagen original, la

base de la pirámide M_0 , necesita ser una copia morfológicamente filtrada de I , el filtro usado para generar M_0 determinará que filtro se debe usar para reconstruir la pirámide. La función *REDUCE* se puede definir nuevamente, en este caso como filtro morfológico de la imagen seguida de un submuestreo.

Para generar el siguiente nivel en la pirámide, *REDUCE* será aplicado a los niveles actuales. Sea M_0 la base de la pirámide y

$$M_k = REDUCE(M_{k-1}) \text{ para } 1 \leq k \leq n, \quad (3.36)$$

donde n es la profundidad de la pirámide y $REDUCE F(M_{k-1})_{\downarrow 2}$. De aquí, es una pirámide de diferencias similar a la pirámide laplaciana lineal que puede ser construida

$$D_n = M_n, \quad (3.37)$$

y

$$D_k = M_k - EXPAND(M_{k+1}) \text{ para } k=n-1, n-2, \dots, 0. \quad (3.38)$$

Aquí, la función *EXPAND* se define como un submuestreo seguido por una cerradura. *REDUCE* y *EXPAND* pueden utilizar cualquier filtro, sin embargo, se debe de asegurar que el filtro usado en la función *EXPAND* sea complementaria al filtro empleado en la función *REDUCE*.

Si el filtrado de la imagen fuente I es una cerradura, entonces se debe de llevar a cabo una reconstrucción mínima y la función *EXPAND* deberá usar el filtro de la erosión.

Si el filtrado inicial de la imagen fuente es una apertura, entonces se debe de llevar a cabo una reconstrucción máxima y la función de *EXPAND* deberá usar el filtro de dilatación.

La reconstrucción de la imagen \hat{M}_0 a través de la pirámide puede ser obtenida

$$\hat{M}_n = D_n, \quad (3.39)$$

y

$$\hat{M}_k = D_k + EXPAND(\hat{M}_{k+1}) \text{ para } k=n-1, n-2, \dots, 0, \quad (3.40)$$

\hat{M}_n es una reconstrucción exacta de la pirámide.

3.3.7. Fusión de imágenes usando la transformada wavelets

La teoría de wavelets (ondeletas) se refiere a la representación de una función en términos de una familia biparamétrica de dilataciones y traslaciones de una función fija que, en general no es senoidal, por ejemplo

$$f(x) = \int_{\mathbb{R}} |a|^{-1/2} \psi\left(\frac{x-b}{a}\right) W_{\psi} f(a,b) da db ,$$

donde $W_{\psi} f$ es una transformada de f adecuadamente definida.

Sea $\psi \in L^2(\mathbb{R})$. La función dilatada y trasladada se define como:

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right), \dots\dots\dots 0 \neq a \in \mathbb{R}, b \in \mathbb{R} . \tag{3.41}$$

La reconstrucción de la función original se hace a partir de sus wavelets de descomposición. Para realizar la descomposición y la construcción de la función existe la condición de que ψ tenga una atenuación suficiente, esto es

$$\int \frac{|\Psi(w)|}{w} < \infty , \tag{3.42}$$

donde $\Psi(w)$ es la transformada de Fourier de $\psi(t)$. Para propósitos prácticos se puede utilizar la versión discreta de la transformada wavelets. Considerando una señal a resoluciones sucesivas separadas por un factor de 2, la transformada discreta de wavelets puede ser definida para $a = 2^m$ y $b = n2^m$, donde m y n son enteros. De modo alternativo, se imagina un desarrollo en serie

$$f(t) = \sum_{m,n} c_{m,n} \psi_{m,n}(t), \quad (3.43)$$

donde $\psi_{m,n}(t) = 2^{-m/2} \psi[2^{-m}t - n]$, y $\psi_{m,n}(t)$ constituye una base ortonormal, así que los coeficientes de la transformada de wavelets pueden ser obtenidos por el producto interno

$$c_{m,n} = \langle f, \psi_{m,n} \rangle = \int \psi_{m,n}(t) f(t) dt. \quad (3.44)$$

Para construir un análisis multiresolución, es necesario, que la función de escalamiento ϕ , así como sus versiones de dilatación y traslación, cumplan con $\phi_{m,n}(t) = 2^{-m/2} \phi[2^{-m}t - n]$. Para una m dada, las $\phi_{m,n}$ son ortonormales. Sea V_m un espacio extendido sobre las $\phi_{m,n}$ con una resolución de 2^m . Un análisis multiresolución es una serie de subespacios sucesivos anidados que se generan $\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \dots$. Para cada m los $\phi_{m,n}$ constituyen un espacio W_m que es exactamente un complemento ortonormal de V_m en V_{m-1} . Matemáticamente se tiene $V_m \perp W_m$ y $V_{m-1} = V_m \oplus W_m$. El análisis multiresolución requiere un grupo de subespacios anidados como se ilustra en la Figura 3.6. La representación en el espacio de frecuencia de la Figura 3.6 se muestra de manera intuitiva en la Figura 3.7.

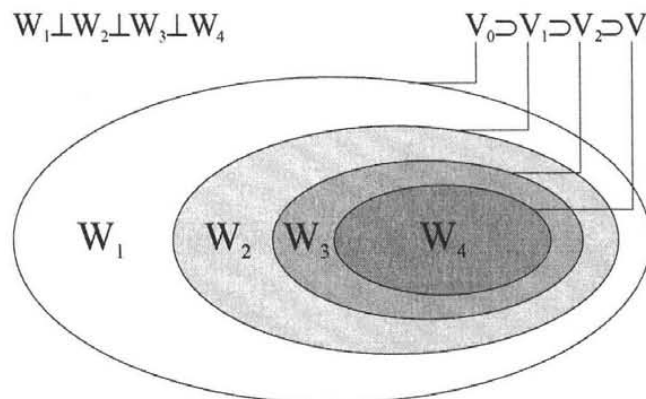


Figura 3.6. Espacios de multiresolución anidados. El espacio V_0 puede ser descompuesto en un subespacio de resolución más bajo V_1 , la diferencia entre V_0 y V_1 puede ser representada por el subespacio complementario W_1 . Similarmente, se descompone V_1 en V_2 y W_2 [18].

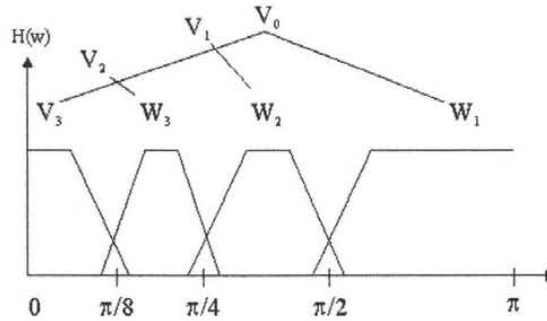


Figura 3.7. Representación en el espacio de frecuencia de la Figura 3.6 [18].

Los coeficientes $a_{m-1,n} = \langle f, \phi_{m-1,n} \rangle$ y $a_{m,n} = \langle f, \phi_{m,n} \rangle$ describen las aproximaciones de la función f a las resoluciones 2^m y 2^{m-1} respectivamente, mientras que los coeficientes $c_{m,n} = \langle f, \psi_{m,n} \rangle$ describen la información perdida cuando se va de una aproximación a otra.

Dado que $h_n = 2^{1/2} \int \phi(t-n)\phi(2t)$ y $g_l = (-1)^l h_{1-l}$, se puede demostrar que existen relaciones recursivas entre coeficientes $c_{m,n}$ y $a_{m,n}$, y son

$$C_{m,n} = \sum_k g_{2n-k} a_{m-1,k} \tag{3.45}$$

$$a_{m,n} = \sum_k h_{2n-k} a_{m-1,k} \tag{3.46}$$

Estas ecuaciones representan operaciones de filtrado usando un filtro pasa baja h y un pasa alta g . Si la función f está dada en forma muestreada, entonces se puede usar estas muestras para una aproximación de coeficientes de resolución de orden alto $a_{0,n}$. La aproximación de las funciones se calcula recursivamente combinando una operación de submuestreo con la operación de filtrado. Debido a sus asociaciones con bases de wavelets ortonormales, estos filtros proporcionan una reconstrucción perfecta como en el caso de los filtros

$$a_{m-1,l}(f) = \sum_n [h_{2n-l} a_{m,n}(f) + g_{2n-l} c_{m,n}(f)]. \tag{3.47}$$

Se han construido filtros de respuesta a impulsos finitos (FIR) para la transformación de wavelets y se han utilizado tanto para la reconstrucción como para la descomposición. Sin embargo, estos filtros no son simétricos y son de fase no lineal.

Desafortunadamente, no existen filtros FIR ortonormales no triviales con fase lineal y propiedades perfectas. Una solución es relajar la condición de utilizar el mismo filtro para usar la reconstrucción y la descomposición y adoptar las llamadas bases biortonormales, así las propiedades de fase lineal se pueden preservar. En este esquema, la descomposición se sigue haciendo usando las ecuaciones 3.45 y 3.46. La reconstrucción se acompaña por un grupo diferente de filtros

$$a_{m-1,l}(f) = \sum_n [a_{m,n}(f) + \tilde{g}_{2n-l} c_{m,n}(f)]. \quad (3.48)$$

Las siguientes condiciones se imponen para asegurar una reconstrucción perfecta

$$\tilde{g}_n = (-1)^n h_{1-n}, \quad (3.49)$$

$$g_n = (-1)^n \tilde{h}_{1-n}, \quad (3.50)$$

$$\sum_n h_n \tilde{h}_{n+2k} = \delta_{k,0}. \quad (3.51)$$

Las relaciones entre los filtros y las funciones wavelets y de escalamiento están dadas por las ecuaciones

$$\phi(t) = \sum_n h_n \phi(2t-n), \quad (3.52)$$

$$\tilde{\phi}(t) = \sum_n \tilde{h}_n \tilde{\phi}(2t-n), \quad (3.53)$$

$$\psi(t) = \sum_n g_n \phi(2t-n), \quad (3.54)$$

$$\tilde{\psi}(t) = \sum_n \tilde{g}_n \tilde{\phi}(2t-n). \quad (3.55)$$

La estructura de funciones del banco de filtros con su wavelets asociada y funciones de escalamiento se presentan en la Figura 3.8.

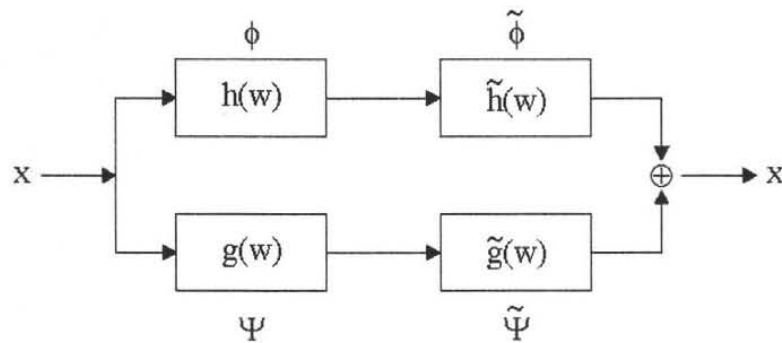


Figura 3.8. Estructura de un banco de filtros y sus correspondientes wavelets [18].

Como en el caso de la transformada 2D de Fourier, la transformada de wavelets bidimensional se considera como una extensión sencilla del caso en una dimensión. La Figura 3.9 muestra la descomposición piramidal multiescala de una imagen a_j . La operación de filtrado con los filtros h y g se aplican a lo largo de la dirección horizontal (líneas), seguido por una operación de submuestreo a lo largo de las líneas. Las dos imágenes resultantes son filtradas a lo largo de la dirección vertical (columnas), seguido por la operación de submuestreo a lo largo de la columna. De esta forma, en el primer nivel obtenemos cuatro nuevas imágenes correspondientes a a_1, d_1^1, d_1^2 , y d_1^3 , que son cuatro veces más pequeñas que la inicial. Estas subimágenes corresponden a las bandas de salidas de bajo-bajo (LL), bajo-alto (LH), alto-bajo (HL) y alto-alto (HH). Si se continúa procesando y se van permutando los resultados tras cada etapa del filtrado, llegaríamos a una distribución como se muestra en la Figura 3.10. El esquema de reconstrucción de la imagen se muestra en la Figura 3.11.

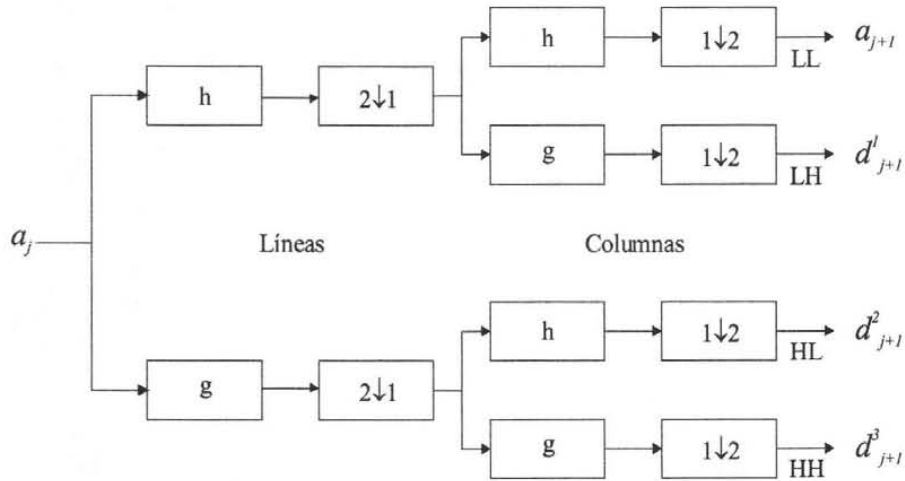


Figura 3.9. Descomposición de una imagen utilizando la transformada wavelets packets. $2 \downarrow 1$ submuestreo de líneas y $1 \downarrow 2$ submuestreo de columnas [18].

LL^3	LH^3	LH^2	LH^1
HL^3	HH^3		
HL^2	HH^2		
HL^1		HH^1	

Figura 3.10. Ejemplo de una señal en 2D DWT con tres niveles de descomposición: niveles de descomposición, 1, 2, 3, H bandas de frecuencia alta, L bandas de frecuencia baja. Después de un nivel de descomposición se tienen 4 bandas de frecuencias bajo-bajo (LL), bajo-alto (LH), alto-bajo (HL) y alto-alto (HH). El siguiente nivel de descomposición se aplica en la banda bajo-bajo (LL), del estado de descomposición actual, el cual sigue el procedimiento recursivo. De aquí, un nivel N de descomposición será $3N+1$ bandas de frecuencias diferentes octavas de una DWT bidimensional de tres niveles [18].

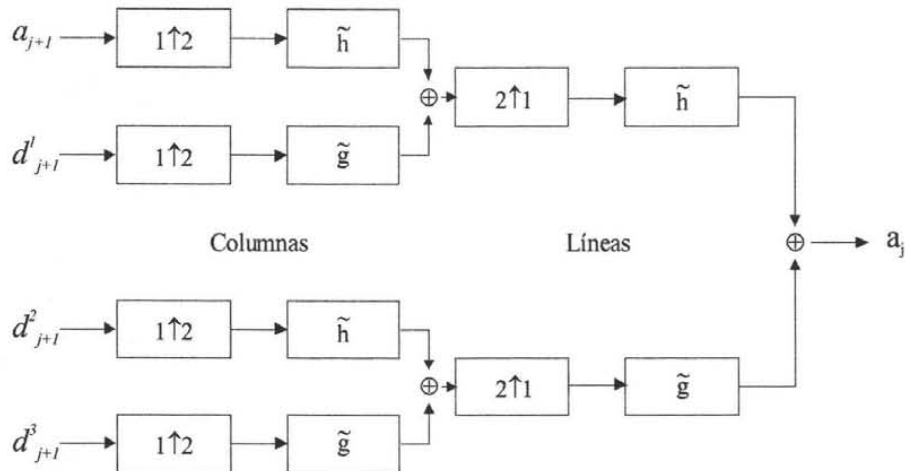


Figura 3.11. Reconstrucción de una imagen utilizando la transformada de wavelets discreta inversa bidimensional. $2 \uparrow 1$ sobremuestreo de líneas y $1 \uparrow 2$ sobremuestreo de columnas [18].

3.3.8. Fusión promedio

El método más simple de fusión de imágenes es tomar el promedio de las dos imágenes pixel por pixel. Sin embargo, cuando este método se toma directo, el contraste de figuras presentes en las imágenes se reduce.

Cuando se combinan dos imágenes, una aproximación intuitiva puede ser el promedio de los píxeles de las imágenes fuentes para obtener el valor correspondiente del pixel en la imagen compuesta. Esta aproximación no es útil porque las Figuras que aparecen en una imagen y en otras no, dan una composición con un contraste reducido o aparecerían sobrepuestas en Figuras desde las otras imágenes. Para evitar el problema de fusión de promedio, la imagen compuesta puede ser obtenida seleccionando píxeles de sus imágenes fuentes.

3.3.9. Máximo contraste

La lógica detrás de este método es que al seleccionar detalles de máximo contraste, la fusión de imágenes dará un mejor detalle para el análisis humano. La expresión del contraste de la función $FUSE$ ($C_k = FUSE(A_k, B_k)$ para $K=n, n-1, n-2, \dots, 0$; n es el número de niveles y $FUSE$ es la función que convierte las dos imágenes en una compuesta) se implementa en una modalidad de pixel a pixel, como sigue:

$$C(i, j) = \begin{cases} A(i, j) & \text{cuando } |A(i, j) - 1| \geq |B(i, j) - 1| \\ B(i, j) & \text{otro caso} \end{cases}, \quad (3.56)$$

A y B representan las imágenes de entrada.

$$C(i, j) = \begin{cases} A(i, j) & \text{cuando } |A(i, j)| \geq |B(i, j)| \\ B(i, j) & \text{otro caso} \end{cases}. \quad (3.57)$$

Sin embargo, se puede implementar el contraste de fusión de manera diferente como:

$$C(i, j) = \begin{cases} A(i, j) & \text{cuando } |A(i, j) - 1| \geq |B(i, j)| \\ B(i, j) & \text{otro caso} \end{cases}, \quad (3.58)$$

donde $C(i, j)$ es el término que define los coeficientes de la imagen fusionada C para todo i, j .

En el caso de la pirámide RoLP la selección del pixel utiliza un área local de contraste para determinar los pixeles que van a ser seleccionados. El contraste de un pixel esta definido como el cociente de luminosidad de pixel y su luminancia de fondo entre esta última, es decir

$$\text{Contraste } (i, j) = \frac{L(i, j) - L_b(i, j)}{L_b(i, j)} = \frac{L(i, j)}{L_b(i, j)} - 1, \quad (3.59)$$

donde L , es la luminosidad en (i, j) , o simplemente la intensidad del pixel y L_b es la luminancia de fondo para el área. Nótese que el cociente entre L y L_b es el valor de la pirámide RoLP, para conseguir un contraste para un cierto valor de pixel, simplemente se resta un valor de la RoLP.

El ojo humano detecta contrastes muy bien. Cuando hablamos de ventajas en reglas de fusión, los píxeles de máximo contraste son seleccionados para cada imagen fuente de la forma compuesta.

3.3.10. Análisis de componentes principales (principle components análisis, PCA)

El método de componentes principales es útil en aplicaciones como codificación de imágenes, compresión de datos de imagen, realce de imagen, detección de cambios, dimensionalidad multitemporal y fusión de imágenes. PCA es un método que transforma datos multibandas intercorrelacionadas en un nuevo grupo de datos no correlacionados, obtenidos por combinación lineal de las originales.

El cálculo de los 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, (ver Figura 3.12). En el caso de imágenes multiespectrales las tres primeras componentes principales contienen el 98-99 % de la información.

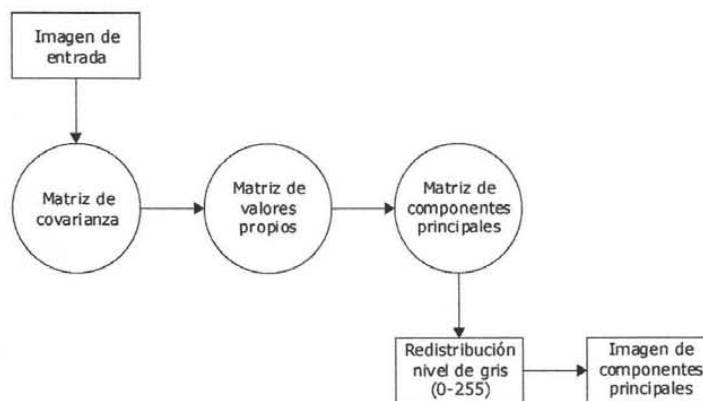


Figura 3.12. Esquema de la transformación de componentes principales [18].

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.

La imagen multiespectral digital puede ser considerada como una variable aleatoria vectorial. Cada banda supone un vector. La función de distribución multivariable constituye el

histograma multiespectral de la imagen. Si suponemos que la función de distribución univariable, es decir, el histograma de cada banda, se asemeja a una distribución normal, entonces cada banda se caracteriza completamente por la media y la varianza de sus datos, y la relación entre las bandas, se describe mediante los correspondientes coeficientes de correlación entre bandas. La imagen multiespectral será: $z = (z_1, z_2, \dots, z_n)$, donde n representa el número de bandas de la imagen.

La fusión de imágenes mediante PCA tiene dos aproximaciones:

1. Sustitución. Reemplazo del primer componente principal por otra imagen (Principal Component Substitution, PCS), Figura 3.13.
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 canales 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, común en todas las bandas mientras la información espectral es única para cada banda. El primer PC1 cuenta con una varianza máxima que maximizar el efecto de la alta resolución de datos en la fusión de imágenes.

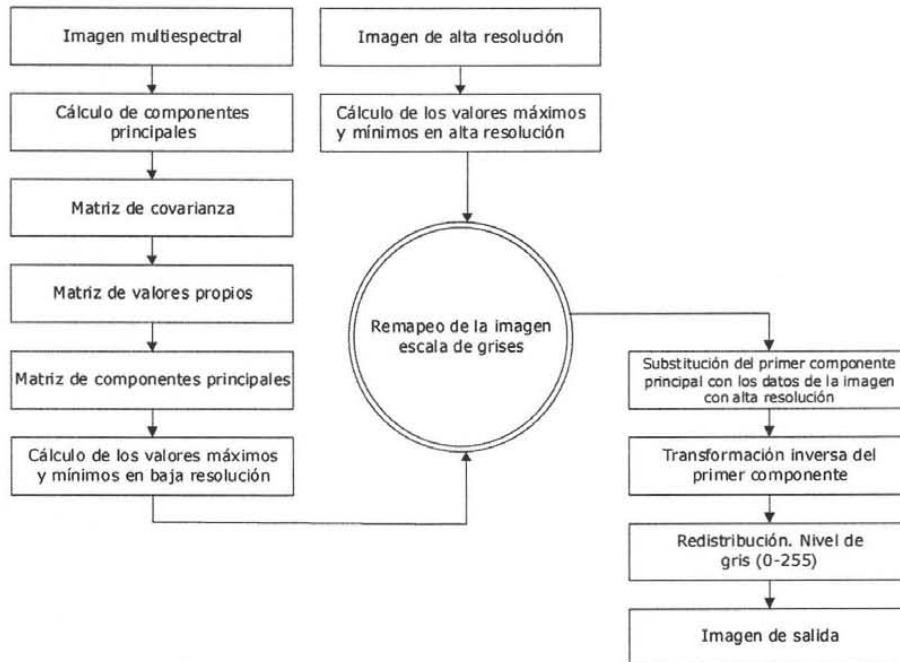


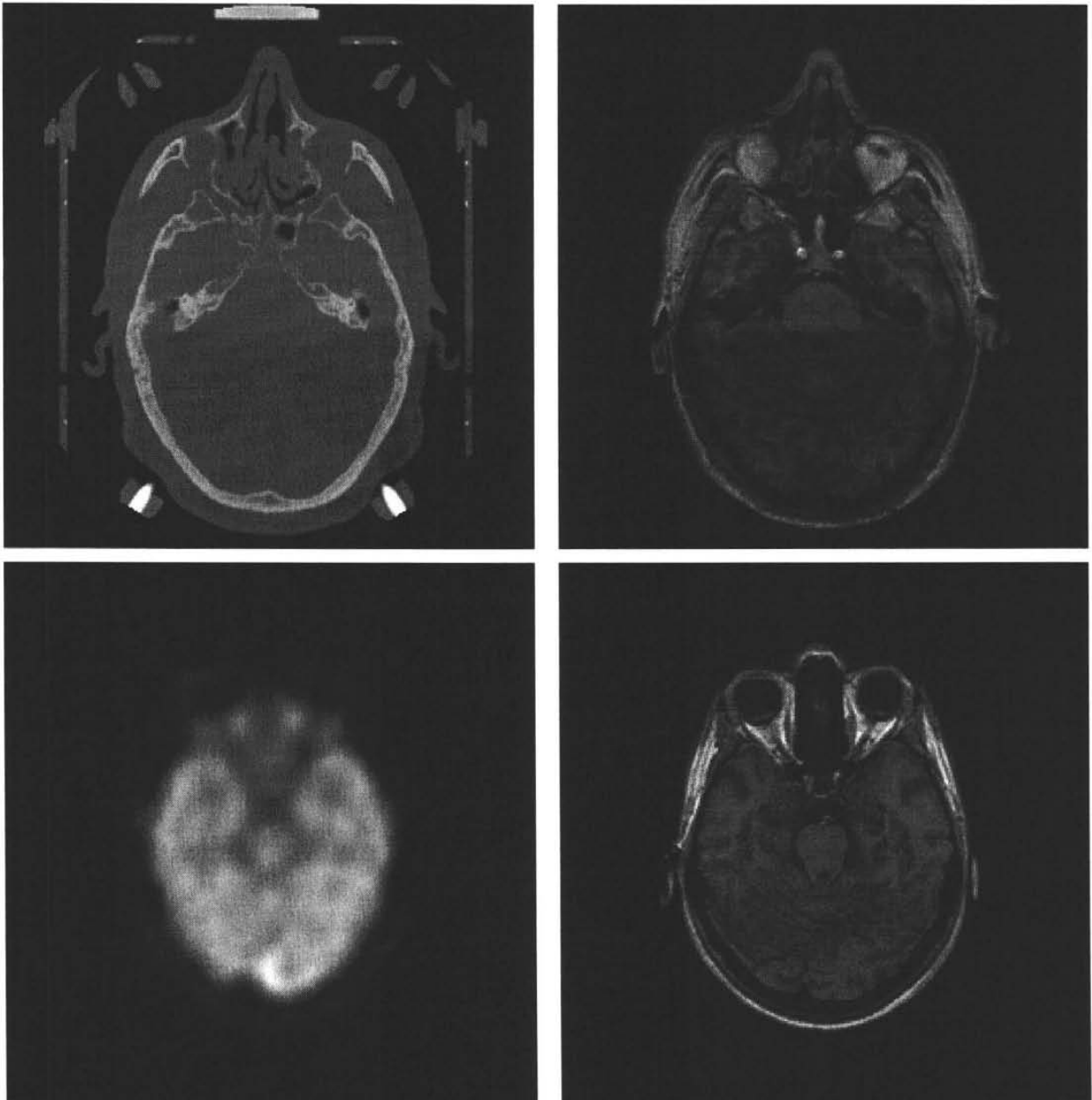
Figura 3.13. Fusión de componentes principales por medio de sustitución [18].

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. 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.

3.4. Ejemplos de métodos de fusión

A continuación se muestran algunos ejemplos de los métodos de fusión teniendo como entrada las imágenes de la Figura 3.14. En las Figuras 3.15, 3.16 y 3.17 se visualizan los resultados.



A	B
C	D

Figura 3.14: Imágenes de Entrada. **A.** Tomografía computarizada. **B.** Resonancia magnética registrada con la imagen de tomografía computarizada. **C.** Tomografía por emisión de positrones registrada con la imagen de resonancia magnética. **D.** Resonancia magnética. [27]

La Figura 3.15 muestra algunos ejemplos de fusión con los métodos de: Promedio, Método PCA, Selección del Máximo y Selección del Mínimo.

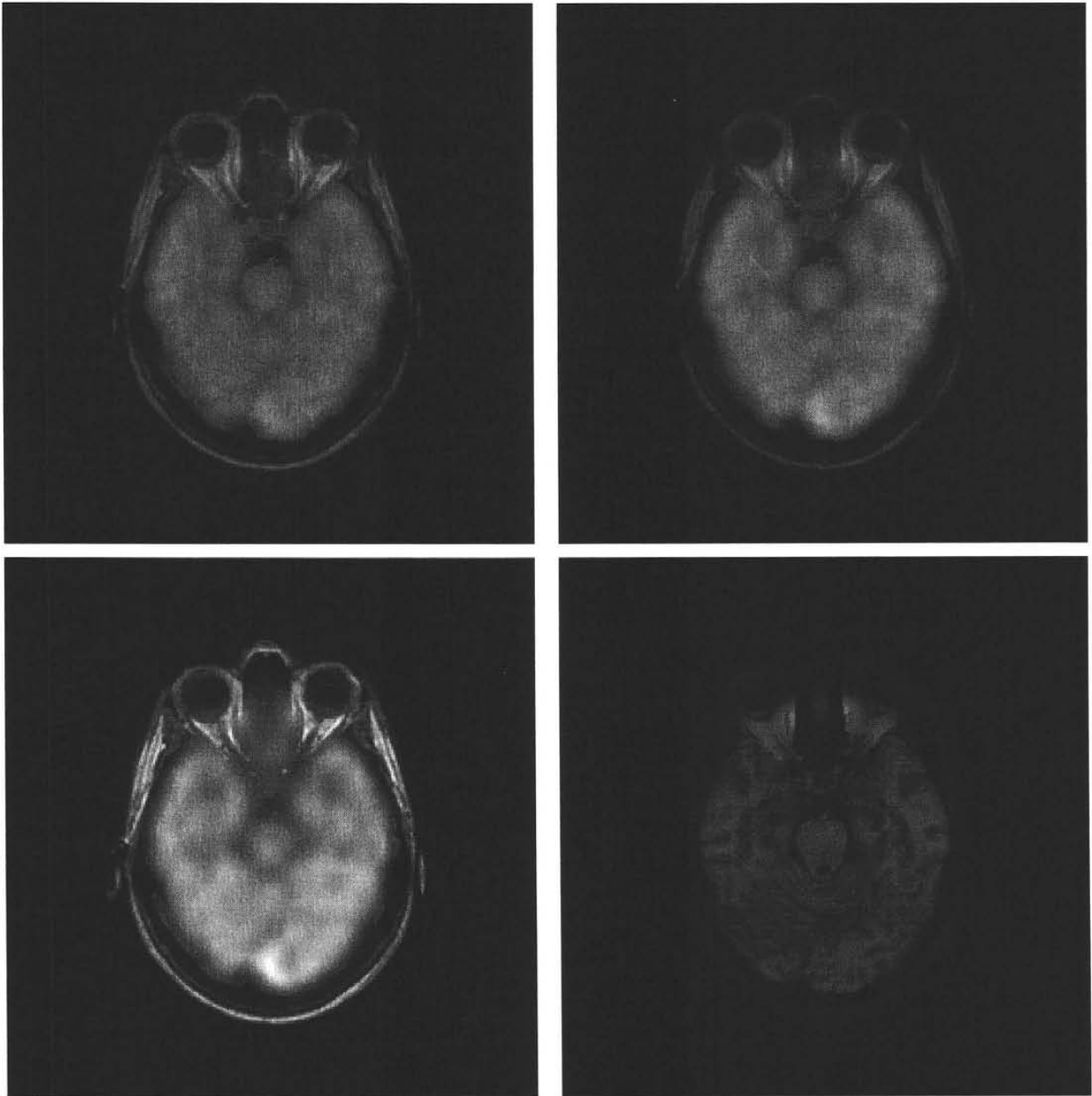
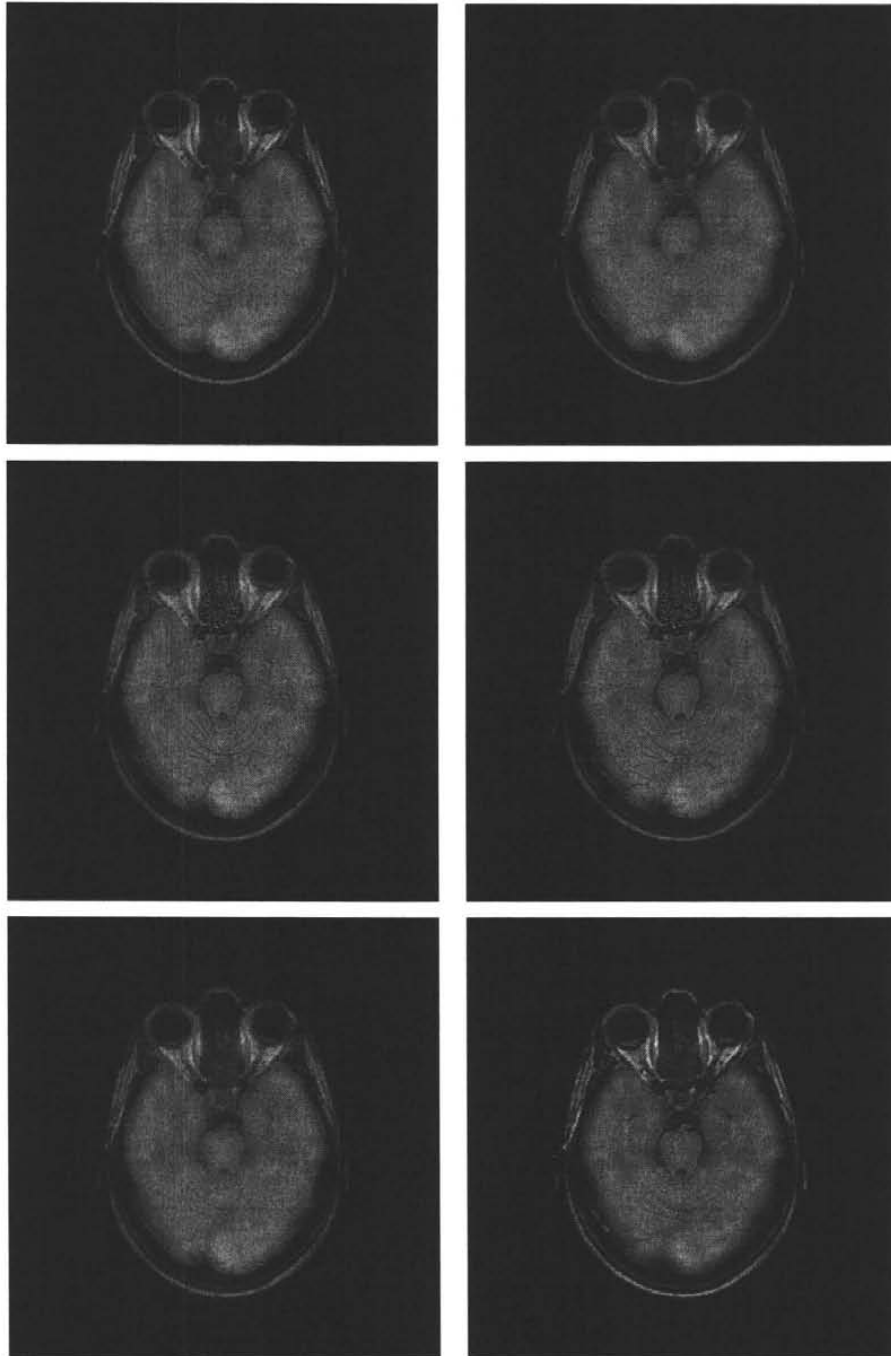


Figura 3.15: Fusión con los métodos: **A.** Promedio; **B.** Método PCA; **C.** Selección del Máximo y **D.** Selección de Mínimo.

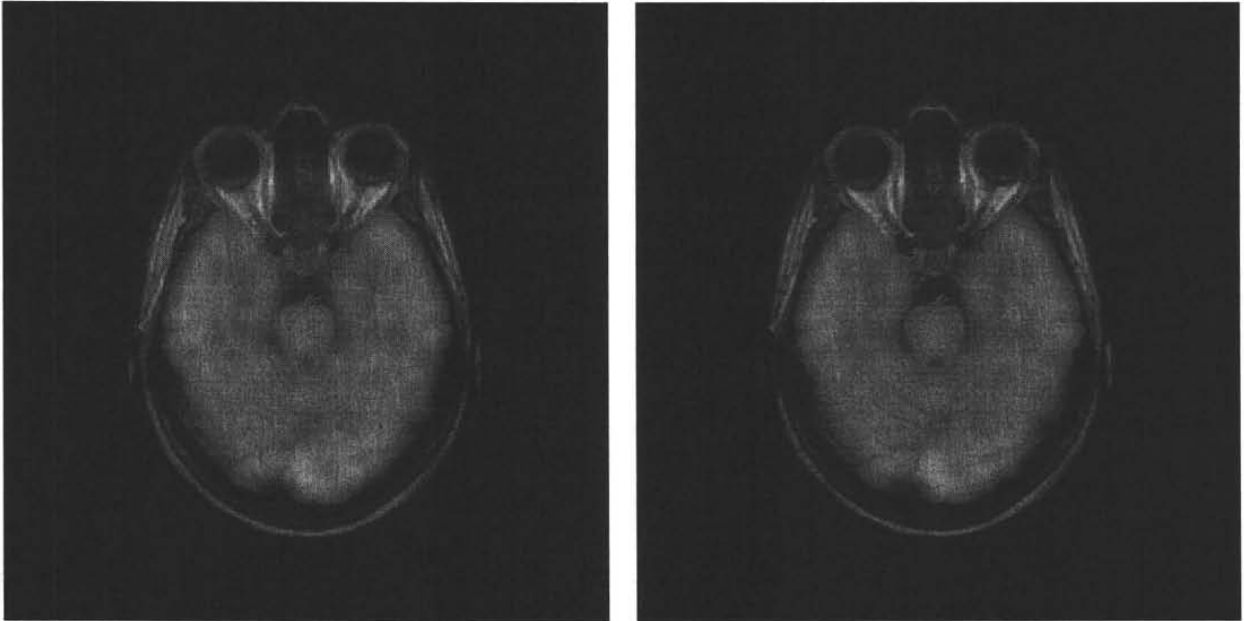
La siguiente Figura (3.16) muestra la fusión de imágenes por métodos piramidales.



A	B
C	D
E	F

Figura 3.16: Fusión con métodos piramidales. **A.** Pirámide Laplaciana; **B.** Pirámide FSD; **C.** Pirámide del Cociente; **D.** Pirámide del Contraste; **E.** Pirámide del Gradiente y **F.** Pirámide morfológica.

En la Figura 3.17 se muestran los ejemplos de fusión a través de transformadas.



A **B**

Figura 3.17: Fusión con métodos de transformada. **A.** Transformada Wavelets Daubechies (2,2) [17] (Daubechies Spline -DBSS-) y **B.** Transformada Wavelets Haar [17] (shift invariant discrete wavelet transform -SIDWT-).

Transformada de Hermite

4.1. Introducción

Las imágenes digitales están especificadas por arreglos de datos que corresponden a valores de intensidad. Para interpretar las imágenes se necesita hacer explícita la información importante. Usualmente esto implica determinar las relaciones espacio-temporales entre las intensidades para proponer la manera de procesar localmente los datos. Este tipo de procesamiento involucra dos decisiones importantes:

Primero, para efectuar el procesamiento local, usualmente las imágenes se multiplican por una función ventana. Este procesamiento local debe repetirse un número suficiente de veces, para describir a la imagen completamente. Algunas características que se eligen para la ventana son: el tamaño, la forma y el espaciamiento. El tamaño, establece el conjunto de píxeles de la imagen que contribuyen en un paso específico de procesamiento. La forma determina el peso relativo de cada píxel de la imagen que está contribuyendo. El espaciamiento entre ventanas se detallará más adelante.

En segundo lugar, para cada posición de la ventana, deben proponerse pasos específicos de procesamiento para ser aplicados. Estos pasos están determinados por el tipo de patrones que se estén buscando y que deben definirse a priori. En general, es muy difícil escoger la función ventana y los pasos de procesamiento basándose en argumentos teóricos puramente. Es por ello que un número considerable de aplicaciones se utiliza como referencia al Sistema de Visión Humano, HVS (Human Visual System) [19].

Entonces para la función ventana, cabe mencionar que el problema de la determinación de su tamaño se puede resolver limitando la complejidad del análisis que se realiza en cada ventana y determinando posteriormente el tamaño de ventana necesario para describir localmente a la imagen con precisión suficiente. De esta manera no se restringe el procesamiento a un solo tamaño de ventana (escala), sino que se repite el mismo procesamiento a diferentes tamaños y subsecuentemente se utilizan las salidas en esta etapa

del procesamiento para seleccionar el tamaño óptimo para cada posición. Este mismo principio es utilizado por el HVS. Ahora bien, si se utilizan ventanas de diferentes tamaños, el espaciamiento entre ellas es generalmente proporcional a su tamaño, de hecho para estructuras piramidales, el espaciamiento y tamaño se incrementa en pasos de 2. La transformada de Hermite involucra estas consideraciones en su definición y desempeño.

Ahora bien, en cuanto a las técnicas de procesamiento local, está demostrado que el uso de técnicas de preprocesamiento, esto es, describir a 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. La transformada de Hermite utiliza derivadas de funciones gaussianas para el procesamiento local, estas pueden modelar operaciones de filtrado en Visión Humana con la misma precisión y con menos parámetros que los ampliamente usados filtros de Gabor.

4.2. Transformada polinomial 1D

La transformada polinomial consiste en una técnica que descompone una señal y la aproxima localmente por polinomios.

Existen dos pasos que involucran el análisis en dicha transformada. Primero, la señal original $L(x)$ se localiza, multiplicándola por una función ventana $V(x)$, para tener una descripción completa de la señal implica que el proceso de localización se repita en un número suficiente de veces para diferentes posiciones de la ventana. Para este caso, siendo el más simple, se debe considerar un espaciamiento equidistante entre ventanas. A partir de la función $V(x)$, se puede construir una función de peso por repetición periódica.

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

la función de peso es periódica con período T .

Si $W(x)$ es diferente de cero para todo x , tenemos

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

de esta forma se garantiza que las señales localizadas en $L(x) \cdot V(x - kT)$ tienen suficiente información en todas las diferentes posiciones de kT de la señal original.

En segundo lugar tenemos la aproximación de una sección de la señal dentro de la ventana $V(x - kT)$ por polinomios. Para llevar a cabo esta expansión se utilizan como funciones base los polinomios $G_n(x)$, de grado n , ortonormales respecto a $V^2(x)$:

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

Estando únicamente determinados estos polinomios por $V^2(x)$.

Los polinomios ortonormales para una función ventana $V^2(x)$ arbitraria están dados por

$$G(x) = \frac{1}{\sqrt{M_{n-1} M_n}} \begin{vmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \\ 1 & x & \cdots & x^2 \end{vmatrix}, \quad (4.4)$$

donde el determinante M_n está definido como:

$$M_n = |c_{i+j}|_{i,j=0,\dots,n}, M_{-1} = 1, \quad (4.5)$$

y

$$C_n = \int_{-\infty}^{+\infty} x^n V^2(x) dx, \quad (4.6)$$

es el momento de orden n .

si $V^2(x)$ es una función par, podemos obtener las siguientes expresiones:

$$\begin{aligned} G_0(x) &= \sqrt{c_0} \\ G_1(x) &= x/\sqrt{c_2} \\ G_2(x) &= (c_0x^2 - c_2)/\sqrt{c_0(c_0c_4 - c_2^2)} \\ G_3(x) &= (c_2x^3 - c_4x)/\sqrt{c_2(c_2c_6 - c_4^2)} \end{aligned}$$

son los polinomios ortonormales de hasta orden 3.

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

$$V(x-kT) \left[L(x) - \sum_{n=0}^{\infty} L_n(kT) \cdot G_n(X-kT) \right] = 0, \quad (4.7)$$

con

$$L_n(kT) = \int_{-\infty}^{+\infty} L(x) \cdot G_n(x-kT) V^2(x-kT) dx. \quad (4.8)$$

Para este caso, si $L(x)$ es analítica y finita para toda x , es suficiente para garantizar la convergencia de la expansión en series en la ecuación 4.7 para la mayoría de funciones de ventana. Así, el error de aproximación entre la señal y un polinomio puede hacerse arbitrariamente pequeño eligiendo un orden para la expansión polinomial suficientemente grande, como en el caso de las expansiones de Taylor. Esto implica que la descripción de las señales localizadas $L(x) \cdot V(x-kT)$, puede reducirse especificando un conjunto finito de coeficientes polinomiales $L_n(kT)$. La energía de la señal dentro de la ventana puede expresarse en términos de los coeficientes de la expansión de acuerdo con la generalización del teorema de Parseval para polinomios ortonormales [19],

$$\int_{-\infty}^{+\infty} L^2(x) V^2(x - kT) dx = \sum_{n=0}^{\infty} L_n^2(kT). \quad (4.9)$$

Combinando las ecuaciones 4.2 y 4.7, se tiene la expansión para la señal completa

$$L(x) = \sum_{n=0}^{\infty} \sum_k L_n(kT) \cdot P_n(x - kT), \quad (4.10)$$

donde

$$P_n(x) = G_n(x) V(x) / W(x). \quad (4.11)$$

La ecuación 4.8 implica que los coeficientes $L_n(kT)$, pueden derivarse a partir de la convolución entre la señal $L(x)$, y los filtros

$$D_n(x) = G_n(-x) V^2(-x), \quad (4.12)$$

seguido de un muestreo en múltiplo de T . Este mapeo a partir de la señal original $L(x)$, a los coeficientes polinomiales $L_n(kT)$, se conoce como **Transformada Polinomial Directa**. La reconstrucción de la señal a partir de estos coeficientes, puede lograrse de acuerdo con la ecuación 4.10 y se conoce como **Transformada Polinomial Inversa** y consiste en interpolar los coeficientes $L_n(kT)$, para k entero, con la función patrón $P_n(x)$, sumando los n órdenes. La Transformada Polinomial Directa e Inversa se ilustra en la Figura 4.1

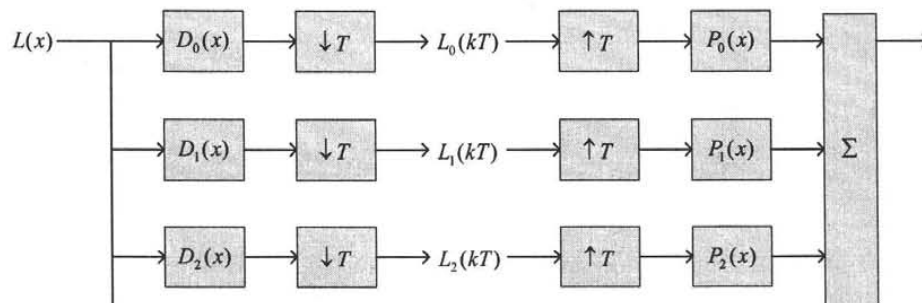


Figura 4.1: Ilustración de la transformada polinomial directa e inversa.

4.3. Transformada de Hermite 1D

Para ilustrar las propiedades de la transformada polinomial, nos concentraremos en el caso especial cuando la función ventana local es gaussiana, es decir

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} e^{-x^2/2\sigma^2}, \quad (4.13)$$

los polinomios ortogonales asociados con $V^2(x)$ son conocidos como polinomios de Hermite, y por tanto, nos referimos a su resultado en la técnica de descomposición local, como **Transformada de Hermite**.

Debido a que la función de peso $W(x)$, es periódica con período T y puede expandirse en una serie de Fourier

$$W(x) = \frac{\sqrt{2\sqrt{\pi}\sigma}}{T} w(x), \quad (4.14)$$

donde

$$w(x) = 1 + 2 \sum_{k=1}^{\infty} \exp\left[-\frac{1}{2}\left(k \frac{2\pi\sigma}{T}\right)^2\right] \cdot \cos\left(k \frac{2\pi x}{T}\right). \quad (4.15)$$

Las funciones filtro determinan qué información es la que se hace explícita en los coeficientes de la Transformada de Hermite. Las propiedades principales de la Transformada de Hermite son determinadas por estas funciones. De la expresión general 4.12 se tiene que:

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma\sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) e^{x^2/\sigma^2}, \quad (4.16)$$

dado que los polinomios de Hermite $\{H_n(x/\sigma); n = 0, 1, \dots\}$, son ortogonales sobre la ventana gaussiana $V^2(x)$, está demostrado que las funciones filtro $D_n(x)$ son iguales a la derivada de orden n de una Gaussiana, es decir

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{d^n}{d\left(\frac{x}{\sigma}\right)^n} \left[\frac{1}{\sigma\sqrt{\pi}} e^{-x^2/\sigma^2} \right]. \quad (4.17)$$

Las funciones patrón $P_n(x)$, se requieren para efectuar la síntesis de la señal original a partir de los coeficientes de la Transformada de Hermite y están dados por la siguiente expresión

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

donde $w(x)$ es la función peso de la ecuación 4.15.

4.4. Transformada Polinomial 2D

Si se proporciona una función ventana $V(x,y)$, los polinomios ortonormales $G_{m,n-m}(x,y)$, donde m y $n-m$, son los grados con respecto a x y y , respectivamente, determinados por

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} V^2(x,y) G_{m,n-m}(x,y) G_{j,i-j}(x,y) dx dy = \delta_{ni} \delta_{mj}, \quad (4.19)$$

para $n, i = 0, 1, \dots, \infty$; $m = 0, 1, \dots, n$; y $j = 0, 1, \dots, i$. La descomposición de señales 2D en polinomios localizadas se da por

$$L(x,y) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{(p,q) \in S} L_{m,n-m}(p,q) \cdot P_{m,n-m}(x-p, y-q), \quad (4.20)$$

donde (p, q) son todas las coordenadas en una retícula S de muestreo 2D. La única condición para la retícula de muestreo es que la función peso

$$W(x, y) = \sum_{(p, q) \in S} V(x - p, y - q), \quad (4.21)$$

sea distinta de cero para todas las coordenadas (x, y) .

Los coeficientes polinomiales $L_{m, n-m}(p, q)$ se derivan del ventaneo de la imagen con los filtros

$$D_{m, n-m}(x, y) = G_{m, n-m}(-x, -y)V^2(-x, -y), \quad (4.22)$$

y muestreando la salida en $(p, q) \in S$. Las funciones patrón que se usan para interpolar los coeficientes polinomiales están definidas como

$$P_{m, n-m}(x, y) = G_{m, n-m}(x, y)V(x, y)/W(x, y), \quad (4.23)$$

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

4.4.1. Representación de imágenes con coeficientes polinomiales unidimensionales

El análisis de imágenes debe tratar de descomponer una señal en patrones que sean perceptualmente importantes. Se sabe que en el marco de la visión y la percepción visual, los patrones 1D (una dimensión) como líneas o bordes juegan un rol importante en visión. Ahora, se establecerá que el mejor ajuste en 1D para una imagen puede ser hallado con la ayuda de las transformadas polinomiales. Esto resulta equivalente a encontrar los coeficientes de la derivadas de las funciones gaussianas y separarlas en aquellos de orden cero (o filtro de suavizado), de primer orden (o detector de líneas), segundo orden (o detector de texturas) y el ángulo correspondiente θ que nos indica la dirección donde se obtiene la mayor energía.

Utilizando un criterio de cuadrado ponderado, se minimiza:

$$E^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [K(x \cos \theta + y \operatorname{sen} \theta) - L(x, y)]^2 V^2(x, y) dx dy, \quad (4.24)$$

sobre todos los patrones k en 1D y ángulos θ . Se define la función ventana 1D

$$V_{\theta}^2(u) = \int_{-\infty}^{+\infty} V^2(u \cos \theta - v \operatorname{sen} \theta, u \operatorname{sen} \theta + v \cos \theta) dv, \quad (4.25)$$

proyectando la función en dos dimensiones $V^2(x, y)$ en un eje que haga un ángulo θ con el eje x . Esta función ventana es independiente de la orientación si $V^2(x, y)$ es rotacionalmente simétrica. Es posible expandir el patrón $K(u)$ en la base $\{F_{n,\theta}(u); n = 0, 1, \dots\}$ de polinomios ortonormales sobre $V_{\theta}^2(u)$, esto es

$$V_{\theta}(u) \left[K(u) - \sum_{n=0}^{\infty} K_{n,\theta} \cdot F_{n,\theta}(u) \right] = 0. \quad (4.26)$$

Sustituyendo las expansiones polinomiales 2D y 1D para $L(x, y)$ y $K(u)$, respectivamente, en la ecuación 4.24, y tomando la derivada parcial con respecto a $K_{n,\theta}$, se encuentra la solución óptima siguiente

$$K_{n,\theta} = \sum_{k=0}^n \sum_{l=0}^k L_{l,k-l} \cdot h_{n,\theta}(l, k-l), \quad (4.27)$$

para los coeficientes de patrones 1D, donde

$$h_{n,\theta}(l, k-l) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_{n,\theta}(x \cos \theta + y \operatorname{sen} \theta) \cdot G_{l,k-l}(x, y) V^2(x, y) dx dy, \quad (4.28)$$

es una función de ángulos que es completamente determinada por $V^2(x, y)$.

Los polinomios ortogonales $F_{n,\theta}(u)$ y la función ángulo pueden determinarse sin necesidad de conocer explícitamente a $V_\theta(u)$. De hecho, la ecuación 4.4 implica que únicamente los momentos

$$c_{n,\theta} = \int_{-\infty}^{+\infty} u^n V_\theta^2(u) du, \quad (4.29)$$

son necesarios para especificar completamente a los polinomios ortogonales. Sin embargo, el cálculo de estos momentos puede basarse directamente en $V^2(x, y)$, ya que

$$c_{n,\theta} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \cos \theta + y \sin \theta)^n V^2(x, y) dx dy. \quad (4.30)$$

A partir de la ortogonalidad de los polinomios $F_{n,\theta}(u)$, es posible derivar las propiedades siguientes:

$$h_{n,\theta}(l, k-l) = 0 \text{ si } k > n, \quad (4.31)$$

y

$$\sum_{k=0}^{\infty} \sum_{l=0}^k h_{m,\theta}(l, k-l) h_{n,\theta}(l, k-l) = \delta_{mn}, \quad (4.32)$$

para la función ángulo.

El error de aproximación en 1D

$$E^2 = \sum_{k=0}^{\infty} \sum_{l=0}^k L_{l,k-l}^2 - \sum_{n=0}^{\infty} K_{n,\theta}^2, \quad (4.33)$$

puede ser minimizado sobre el ángulo θ maximizando la energía direccional

$$\sum_{n=0}^{\infty} K_{n,\theta}^2, \quad (4.34)$$

donde $K_{n,\theta}$ se determina por medio de los coeficientes polinomiales 2D a través de la ecuación 4.27. En la práctica, sólo algunos de los primeros términos de esta energía direccional son suficientes para calcular un estimado de la dirección óptima. Cabe mencionar que la energía direccional se encuentra a través de una simple combinación de los coeficientes de los polinomios en 2D. Esto es computacionalmente más eficiente que el hecho de usar distintos filtros para calcular la energía en cada dirección, especialmente si se quieren revisar un gran número de direcciones.

Si la imagen original $L(x,y)$, es localmente unidimensional, entonces el error de estimación debe ser cero para el ángulo óptimo θ , por lo cual los coeficientes polinomiales 2D deben satisfacer

$$L_{l,k-l} = \sum_{n=k}^{\infty} K_{n,\theta} \cdot h_{n,\theta}(l, k-l). \quad (4.35)$$

Si la imagen $L(x,y)$, es localmente de 1D entonces la ecuación 4.35 puede utilizarse como una aproximación óptima para coeficientes polinomiales 2D.

4.5. Transformada de Hermite 2D

La Transformada de Hermite se obtiene cuando en la transformada Polinomial 2D, la función ventana es una función Gaussiana. Una ventaja importante de las funciones ventana Gaussianas en dos dimensiones, es la propiedad que tienen de ser espacialmente separables y rotacionalmente simétricas. El hecho de que la función ventana sea separable permite que las funciones filtro y patrón sean también separables y como consecuencia que puedan ser implementadas de manera sencilla. Por ejemplo, los coeficientes polinomiales se encuentran mediante la convolución de la imagen con funciones filtro $D_m(x)D_{n-m}(y)$ donde $D_m(x)$ es la función filtro 1D para la ventana $V(x)$, seguida de un muestreo de la salida en las direcciones vertical y horizontal en múltiplos del espacio de muestreo T . Las propiedades de las funciones

filtro incluyen el hecho de que son separables espacial y polarmente. La transformada de Fourier de $D_m(x)D_{n-m}(y)$, expresada en coordenadas polares $w_x = w \cos \theta$ y $w_y = w \sin \theta$ es

$$d_m(w_x)d_{n-m}(w_y) = g_{m,n-m}(\theta) \cdot d_n(w), \quad (4.36)$$

con w_x, w_y frecuencias espaciales en x y y , donde $d_n(w)$ es la transformada de Fourier de la función filtro de Hermite 1D $D_n(r)$, considerando a r como coordenada radial y

$$g_{m,n-m}(\theta) = \sqrt{\frac{n!}{m!(n-m)!}} \cos^m \theta \cdot \sin^{n-m} \theta, \quad (4.37)$$

expresa la selectividad direccional del filtro. De esta manera los filtros de orden mayores a n analizan sucesivamente frecuencias radiales mayores, esto es, resoluciones espaciales más grandes como en el caso 1D. Los filtros del mismo orden n e índice m (direccional) diferente distinguen entre diferentes orientaciones en la imagen. La relación con la función de ángulo general presentada anteriormente es

$$h_{n,\theta}(l, k-l) = \delta_{nk} \cdot g_{l,k-l}(\theta), \quad (4.38)$$

resultando en una simplificación sustancial del caso general. Está demostrado por Daugman [25] la importancia de filtros polarmente separables, esto es, filtros que pueden expresarse como el producto de una función de frecuencia espacial y una función de orientación, ambas sintonizadoras. De manera más específica, sólo los filtros separables tienen la propiedad de arrojar curvas sintonizadoras de orientación idénticas para diferentes patrones 1D, como líneas y bordes. Esto implica que los filtros pueden detectar la orientación de un patrón 1D independientemente de su estructura interna.

4.6. Transformadas Polinomiales Discretas

Hasta ahora las señales y filtros se han asumido como continuos, pero las aplicaciones prácticas de las transformaciones polinomiales requieren ser formuladas para señales

discretas. Para ello, podemos definir a las transformadas polinomiales directamente sobre señales discretas, esto es, sin la necesidad de una liga explícita entre señales continuas y discretas. En el caso de las transformadas polinomiales para 1D, las expresiones para las funciones peso, los filtros y las funciones patrón siguen siendo válidos, mientras se reemplace la variable continua x por una discreta y todas las expresiones con integrales tendrán que utilizar en su lugar sumatorias. Por ejemplo, la ecuación 4.6 que calcula el momento de orden n tendrá que cambiar por

$$c_n = \sum_x x^n V^2(x), \quad (4.39)$$

para $n = 0, \dots, N$. Si la ventana discreta es finita, esto es, $V(x)=0$ para $x < N_1$ y $x > N_2$ entonces la transformada Polinomial tendrá un orden finito $N=N_2-N_1$. Los coeficientes polinomiales de hasta orden N , serán suficientes para reconstruir perfectamente cualquier señal discreta. La razón es que la señal discreta dentro de la ventana $V(x)$ tiene únicamente $N+1$ grados de libertad.

Para el caso de las transformadas polinomiales 2D, la mayoría de los resultados para imágenes analógicas pueden ajustarse de manera directa a imágenes discretas. Sin embargo, existe una complicación que se debe cuando se proyecta la función $V^2(x, y)$ sobre un eje que hace un ángulo θ con el eje x , esto es

$$V_\theta^2(u) = \sum_v V^2(u \cos \theta - v \sin \theta, u \sin \theta + v \cos \theta), \quad (4.40)$$

sólo se consideran los puntos muestreados en la proyección. De esta forma, u y v asumen únicamente los valores

$$u = x \cos \theta + y \sin \theta, \quad v = -x \sin \theta + y \cos \theta, \quad (4.41)$$

donde x y y toman todos los valores (enteros) para los cuales $V(x, y)$ es diferente de cero. Las funciones de suavizado de la imagen se obtienen usualmente seleccionando el ángulo θ donde las líneas de proyección pasan a través de un número mayor de puntos posibles (ver

Figura 4.2). La aplicación de la técnica de aproximación 1D, requiere conocer la función ángulo

$$h_{n,\theta}(l, k-l) = \sum_{x,y} F_{n,\theta}(x \cos \theta + y \sin \theta) \cdot G_{l,k-l}(x,y) V^2(x,y), \quad (4.42)$$

donde $F_{n,\theta}(u)$ es el polinomio ortonormal de orden n sobre la ventana 1D $V_\theta^2(u)$.

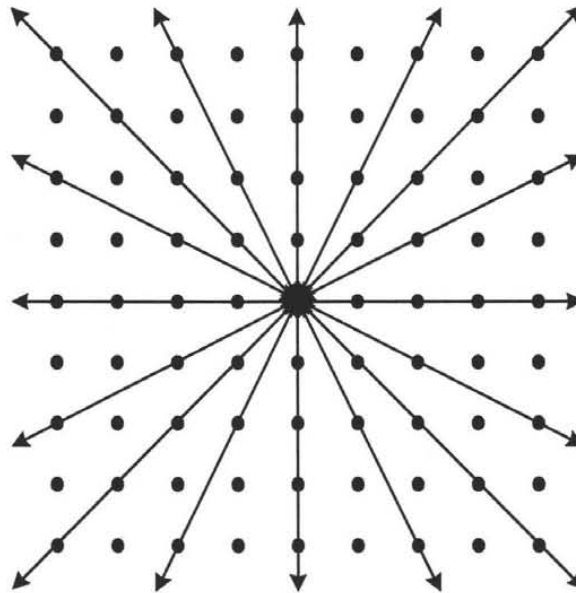


Figura 4.2: Angulos discretos en una malla de muestreo cuadrada 2D

4.7. Transformada discreta de Hermite 2D

El paralelo discreto de una ventana Gaussiana es una ventana binomial dada por

$$V^2(x) = \frac{1}{2^M} C_M^x, \quad (4.43)$$

para $x=0,\dots,M$. Los polinomios ortonormales discretos que están asociados con este tipo de ventanas, se conocen como polinomios de Krawtchouk [26]

$$G_n(x) = \frac{1}{\sqrt{C_M^n}} \sum_{k=0}^n (-1)^{n-k} C_{M-x}^{n-k} \cdot C_x^k, \quad (4.44)$$

para $x, n=0, \dots, M$.

Para valores grandes de M , la ventana binomial se reduce a una ventana Gaussiana; de manera más específica

$$\lim_{M \rightarrow \infty} \frac{1}{2^M} C_M^{x+(M/2)} = \frac{1}{\sqrt{\pi} \sqrt{\frac{M}{2}}} \exp \left[- \left(\frac{x}{\sqrt{\frac{M}{2}}} \right)^2 \right], \quad (4.45)$$

para $x=-(M/2), \dots, M/2$. También puede demostrarse que el mismo límite convierte a un polinomio de Krawtchouk en un polinomio de Hermite [19], es decir

$$\lim_{M \rightarrow \infty} G_n \left(x + \frac{M}{2} \right) = \frac{1}{\sqrt{2^n n!}} H_n \left(\frac{x}{\sqrt{\frac{M}{2}}} \right). \quad (4.46)$$

De esta manera, la transformada discreta de Hermite de longitud M aproxima a la transformada de Hermite analógica de extensión $\sigma = (M/2)^{1/2}$. Las propiedades de la transformada discreta de Hermite pueden ser adecuadas a partir de las correspondientes propiedades de la transformada de Hermite analógica.

Si nos concentramos en el caso en M par, las funciones filtro y patrón pueden ser concentradas en el origen mediante el movimiento de la ventana binomial sobre $M/2$. Esto permite la definición de las funciones filtro para la transformada discreta de Hermite

$$D_n(x) = G_n \left(\frac{M}{2} - x \right) \cdot V^2 \left(\frac{M}{2} - x \right), \quad (4.47)$$

para $x=-(M/2), \dots, M/2$. Estas funciones también pueden expresarse como

$$D_n\left(\frac{M}{2} - x\right) = \frac{(-1)^n}{2^M \sqrt{C_M^n}} \Delta^n [C_M^x \cdot C_x^n], \quad (4.48)$$

donde

$$(-1)^n \Delta^n L(x) = \sum_{k=0}^n (-1)^k C_n^k L(x+k), \quad (4.49)$$

es el operador diferencia de n -ésimo orden. Calculando la transformada z de esta función filtro se tiene

$$d_n(z) = \sum_{x=-M/2}^{M/2} D_n(x) z^{-x} = z^{-M/2} \sqrt{C_M^n} \left(\frac{1-z}{2}\right)^n \left(\frac{1+z}{2}\right)^{M-n}, \quad (4.50)$$

si se expresa en frecuencias angulares

$$d_n(e^{-jw}) = \sqrt{C_M^n} \left(j \operatorname{sen} \frac{w}{2}\right)^n \left(\cos \frac{w}{2}\right)^{M-n}, \quad (4.51)$$

para $n=0, \dots, M$. Es obvio que para w pequeñas este filtro se reduce a una operación derivativa de orden n , como en el caso analógico.

Los filtros anteriores tienen la ventaja importante de que pueden implementarse como filtros $z^{-1}(1-z)^2$, $z^{-1}(1-z)(1+z)$, $z^{-1}(1+z)^2$, cuyos kernels son $[1 \ 2 \ 1]$, $[-1 \ 0 \ 1]$ y $[1 \ -2 \ 1]$, respectivamente. De esta manera con excepción del factor de amplificación $(C_M^n)^{1/2}$ estos filtros pueden realizarse sin la obtención de multiplicaciones generales.

El cálculo de la función ángulo $h_{n,\theta}$ es una aplicación directa de la ecuación 4.41, aunque los cálculos pueden ser más largos para valores mayores de n . Algunas desviaciones sustanciales, con respecto a la función ángulo de la Transformada analógica de Hermite dada en la ecuación 4.37, ocurrirán sólo si M es pequeño. En la mayoría de las aplicaciones solo

interesaran las funciones ángulo $h_{n,\theta}$ para valores de n pequeños. Las expresiones explícitas para $n=0,1,2$ se muestran a continuación

$$\begin{aligned}
 h_{0,\theta}(0,0) &= 1 \\
 h_{1,\theta}(0,0) &= 0 \\
 h_{1,\theta}(1,0) &= \cos\theta \\
 h_{1,\theta}(0,1) &= \text{sen}\theta \\
 h_{2,\theta}(0,0) &= 0 \\
 h_{2,\theta}(1,0) &= 0 \\
 h_{2,\theta}(0,1) &= 2 \\
 h_{2,\theta}(2,0) &= \cos^2\theta \cdot \sqrt{1-\frac{1}{M}} \sqrt{1-\frac{1}{M}(\cos^4\theta + \text{sen}^4\theta)} \\
 h_{2,\theta}(1,1) &= \sqrt{2} \cos\theta \text{sen}\theta \sqrt{1-\frac{1}{M}(\cos^4\theta + \text{sen}^4\theta)} \\
 h_{2,\theta}(2,0) &= \text{sen}^2\theta \cdot \sqrt{1-\frac{1}{M}} \sqrt{1-\frac{1}{M}(\cos^4\theta + \text{sen}^4\theta)}.
 \end{aligned} \tag{4.52}$$

Cabe mencionar que $h_{0,\theta}$ y $h_{1,\theta}$ son los mismos que para el caso analógico y que la diferencia máxima para $h_{2,\theta}$, entre los casos analógico y el discreto, es muy pequeña usualmente, por ejemplo, para $M=6$ es menor al 5 %.

4.8. Aplicaciones de la Transformada de Hermite

Martens [20] propone la codificación jerárquica como una de las aplicaciones de la Transformada de Hermite. Un codificador jerárquico compara las descripciones de la imagen en niveles sucesivos de detalle. Las representaciones de imágenes propuestas por Martens [19] son descripciones polinomiales locales y finitas en 2D, en éste propone la utilización de polinomios locales 1D motivado por el hecho de que en las imágenes naturales la mayoría de las descripciones locales 2D pueden ser aproximadas correctamente por descripciones 1D, de esta manera, el poder descriptivo de la aproximación 1D es mucho más parecido o adecuado que en el caso de las descripciones 2D. Esto refleja el hecho de que las estructuras locales 1D como los bordes representan una mayoría de las regiones no constantes. La importancia de las estructuras 1D ha sido reconocida ampliamente en visión por computadora, donde muchos algoritmos han sido diseñados específicamente para detectar y localizar estas estructuras. En la codificación por transformada polinomial 1D, el análisis se efectúa en cada

nivel del codificador y no se restringe a la transformación polinomial directa 2D en coeficientes $L_{m,n-m}$, sino que estos coeficientes son mapeados subsecuentemente en un ángulo óptimo θ y los coeficientes polinomiales 1D

$$K_{k,\theta} = \sum_{n=0}^k \sum_{m=0}^n L_{m,n-m} \cdot h_{k,\theta}(m, n-m), \quad (4.53)$$

para $k=0, 1, \dots, N$, utilizando el procedimiento descrito en la sección 4.4.1 representación de imágenes con coeficientes polinomiales unidimensionales, donde también se define la función ángulo $h_{k,\theta}$. El ángulo óptimo se encuentra maximizando la energía direccional $\sum_{k=1}^N K_{k,\theta}^2$, o equivalentemente, el coeficiente de contraste

$$C_{\theta} = \log \left[1 + \sum_{k=1}^N K_{k,\theta}^2 \right], \quad (4.54)$$

sobre todos los ángulos. Como resultado se tiene que para cada posición de la ventana en el nivel i , la imagen original se mapea en coeficientes 1D $h_{k,\theta}$, y un ángulo óptimo θ .

4.9. Transformada de Hermite Multiescala 1D

4.9.1. Descomposición multiescala

Para entender la noción de escala se debe ver a la imagen como la representación de una escena física donde por ejemplo, el efecto de la distancia de observación se puede traducir en un escalamiento de la distribución de intensidades de la representación, esto significa que al alejarse del objeto observado aumenta la escala de la representación y por lo tanto se pierden detalles, mientras que al acercarse se reduce la escala y en consecuencia se perciben más detalles. La representación espacio-escala emplea las versiones escaladas de la Gaussiana de integral unitaria $G(x) = \exp(-x^2) / \sqrt{(\pi)}$ para eliminar detalles de la señal de luminosidad $L(x)$ mediante la operación de convolución, es decir que $L(x,s) = L(x) * G(x,s)$ es la

representación de la señal a la escala s , donde $G(x, s) = G(x/\sqrt{4s})/\sqrt{4s}$ es la gaussiana a la escala $s > 0$. La variable normalizada $\xi = x/\sqrt{4s}$ es referida como coordenada natural. Para llegar a un resultado que ponga en evidencia la relación existente entre las estructuras de la imagen percibidas a diferentes escalas, se necesita tener una idea sobre la forma en la que el sistema de visión humana podría llevar a cabo dicho proceso [21]. Los resultados psicofísicos evidencian la existencia de por lo menos cuatro canales en cada punto del campo visual a nivel de la retina (específicamente en las células ganglionares). Dichos canales exhiben una respuesta aproximada a una diferencia de Gaussianas (DoG del inglés Difference of Gaussians) cuya respuesta al impulso se escribe como $DoG(x, s_1, s_2) = G(x, s_1) - G(x, s_2)$ para $s_2 > s_1$. La respuesta del filtro DoG contiene la información que debe ser agregada a la representación a la escala s_2 para obtener la representación a la escala menor s_1 . Así, la señal se puede expresar como una suma de canales DoG:

$$L(x) = \sum_{k=-\infty}^{\infty} L(x) * DoG(x, s_{k-1}, s_k), \quad (4.55)$$

con $s_k > s_{k-1}$. Requiriendo que el parámetro de espaciamento $\tau = (s_k - s_{k-1})$ sea constante para todo entero k , de acuerdo con datos psicofísicos que soportan la existencia en el HVS (Sistema de Visión Humano) de un número de canales cuyas frecuencias centrales mantienen una relación de aproximadamente una octava, es decir $\tau \approx 0.75$.

4.9.2. Análisis estructural

En la mayoría de las aplicaciones se desea saber qué parte de la información contenida en cada canal, hablando en términos perceptivos, debe agregarse con mayor intensidad (por ejemplo, bordes, líneas, esquinas, etc.) y qué 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. Así, la estructura local de una señal se infiere a partir de las derivadas de las representaciones escaladas. Por otra parte, las derivadas de $L(x, s)$ se obtienen convolucionando la función $L(x)$ con los operadores derivadas de Gaussiana (DG), $G_n(x, s) = \partial_x^n G(x, s)$. Por ende, el siguiente paso es expresar al filtro DoG en función de las

DG, y esto se consigue tomando la expansión en serie de Taylor de la Gaussiana a la escala s_{k-1} alrededor de la escala s_k , es decir,

$$DoG(x, s_{k-1}, s_k) = \sum_{n=1}^{\infty} \frac{(-\tau s_k)^n}{n!} G_{2n}(x, s_k). \quad (4.56)$$

En la práctica se debe limitar el número de términos de esta serie a un número pequeño. Estudios neurofisiológicos indican que el sistema visual trabaja con derivadas hasta de cuarto orden. En este estudio matemático, sin embargo, se consideran todos los términos de la serie ya que las aproximaciones se pueden estudiar como casos particulares.

4.9.3. Definición de la Transformada de Hermite Multiescala

La ecuación 4.56 muestra explícitamente que las derivadas de orden impar no contribuyen en la descripción de la imagen. Por otro lado, es deseable que, al menos en la etapa de análisis del modelo, aparezcan todas las derivadas con el fin de tener una descripción más rica de la estructura de la señal. Por lo tanto, es necesario distinguir en el modelo la fase de análisis de la fase de síntesis. Esto se logra factorizando las DG que aparecen en la serie 4.56 utilizando la propiedad de cerradura bajo convolución. Más aún, si se interpretan las operaciones de convolución como producto interno, se puede escribir las expresiones para la reconstrucción de la señal como:

$$L(x) = \sum_{k=-\infty}^{\infty} \sum_{n=1}^{\infty} \tau^n D_n^{(k)}(x), \quad (4.57)$$

donde las funciones

$$D_n^{(k)}(x) = \langle L_n^{(k)}(\xi), G_n^{(k)}(x, \xi) \rangle_{\xi}, \quad (4.58)$$

para $n=1,2,\dots$ aportan los detalles de la señal a todas las escalas y están determinadas por las derivadas escaladas,

$$L_n^{(k)}(x, \xi) = \left\langle L(x), G_n^{(k)}(x, \xi) \right\rangle_x, \quad (4.59)$$

determinados proyectando la señal de entrada sobre las funciones de análisis

$$G_n^{(k)}(x, \xi) = \frac{1}{\sqrt{2k_s}} G_n^* \left(\frac{x}{\sqrt{2s_k}} - \xi \right), \quad (4.60)$$

para $n=1,2,\dots$ y $s_k = (1-\tau)^{-k} s_0$ para todo k entero. Al análisis de la señal expresado en la ecuación 4.59 le conocemos como THM [21] (**Transformada de Hermite Multiescala**) directa, mientras que la reconstrucción de la señal mediante la ecuación 4.57 y las funciones de detalle en la ecuación 4.58 se conocen como THM inversa.

4.9.4. Acotamiento del intervalo de escalas

Considerando un número finito K de canales DoG y suponiendo que la escala menor es suficientemente pequeña, la descomposición multicanal se reduce a

$$L(x, s_0) = L(k, s_K) + \sum_{k=1}^K \sum_{n=1}^{\infty} \tau^n D_n^{(k)}(x), \quad (4.61)$$

donde s_0 y s_K determinan las resoluciones límite del análisis. El residuo paso bajas $L(x, s_K)$ no puede ser despreciado para valores de K en un intervalo práctico, pero su variación es, en general, menor que la de la señal original. En cambio, el residuo paso altas $L(x) - L(x, s_0)$ es despreciado en virtud de que s_0 es suficientemente pequeño. Para casos en los cuales el residuo paso altas contiene información relevante, se puede agregar un canal definido entre las escalas 0 y s_0 ($\tau=1$), ya que la señal es la representación a la escala cero. Por lo tanto, generalizando las definiciones de las ecuaciones 4.58, 4.59 y 4.60 para $n=0$, se puede escribir la THM como

$$L(x) = D_0^{(k)}(x) + \sum_{n=1}^{\infty} D_n^{(0)}(x) + \sum_{k=1}^K \sum_{n=1}^{\infty} \tau^n D_n^{(k)}(x), \quad (4.62)$$

donde los dos primeros términos del miembro derecho son justamente el residuo paso bajas y el residuo paso altas respectivamente. El caso cuando $K=0$ se conoce como **transformada de Hermite de una sola escala** (o simplemente TH). En tal caso no aparece el tercer término de la ecuación 4.62.

4.9.5. Esquema piramidal continuo

A un esquema de procesamiento en el cual la imagen se descompone en un número de subimágenes pasa bajas o pasa banda, las cuales son submuestreadas en proporción a su resolución, se le conoce como esquema piramidal. Cada nivel de la pirámide guarda una relación lineal con los niveles adyacentes y dicha relación es invariante con la escala. Esta definición está colocada en el contexto de señales discretas, pero puede ser extrapolada al caso continuo si se reemplaza el remuestreo por un reescalamiento de la coordenada espacial. En tal caso, todos los niveles de la pirámide quedan definidos en un dominio continuo. La relación piramidal de la THM se deriva del hecho de que las funciones de análisis de un orden dado forman una pirámide Gaussiana dada por la relación lineal

$$\sqrt{\tau^n} G_n^{(k)}(x, \xi) = \int_{-\infty}^{\infty} G_0^{(k-1)}(x, \eta) P_n(T\xi - \eta) d\eta, \quad (4.63)$$

donde los coeficientes de la pirámide $P_n(x) = G_n^*(-x/\sqrt{4\sigma})/\sqrt{4\sigma}$ son derivadas de Gaussiana a la escala $\sigma = \tau T^2/4$, la cual se mantiene constante para todos los niveles. El factor $T = 1/\sqrt{1-\tau}$ determina la proporción entre las coordenadas de un nivel dado y del nivel anterior. Para $\tau = 0.75$ se tiene $T=2$. Aplicando sucesivamente la propiedad 4.63 en la expresión 4.58 se obtiene la relación piramidal ascendente

$$\tau^n D_n^{(k)} = \left\{ [(\sqrt{\tau^n} L_n^{(k)})^{\uparrow T} * P_n]^{\uparrow T} * \underbrace{P_0^{\uparrow T} \dots * P_0}_{k-1 \text{ factores}} \right\} * P_0^{(0)}, \quad (4.64)$$

para $k=1, \dots, K-1$ y $P_n^{(0)}(x) = G_n^*(-x/\sqrt{4s_0})/\sqrt{4s_0}$. Por otra parte, si se aplica la misma propiedad a la fase de análisis de la ecuación 4.59, se obtiene la relación piramidal descendente

$$\sqrt{\tau^n} L_n^{(k)} = \left\{ \left[(L * P_0^{(0)} * \underbrace{P_0^{\downarrow T} \dots P_0^{\downarrow T}}_{k-1 \text{ factores}}) * (-1)^n P_n \right]^{\downarrow T} \right\}. \quad (4.65)$$

Los símbolos $\uparrow T$ y $\downarrow T$ significan que las coordenadas se multiplican por T y T^{-1} respectivamente. De las ecuaciones 4.64 y 4.65 se obtiene el esquema piramidal mostrado en la Figura 4.3 donde, por simplicidad, se escribió $L_n^{(k)}$ en lugar de $\sqrt{\tau^n} L_n^{(k)}$. Como se puede observar en este diagrama, el esquema piramidal tiene la característica de que la reconstrucción se realiza por etapas agregando sucesivamente las contribuciones de cada nivel de la pirámide comenzando con el residuo pasa bajas.

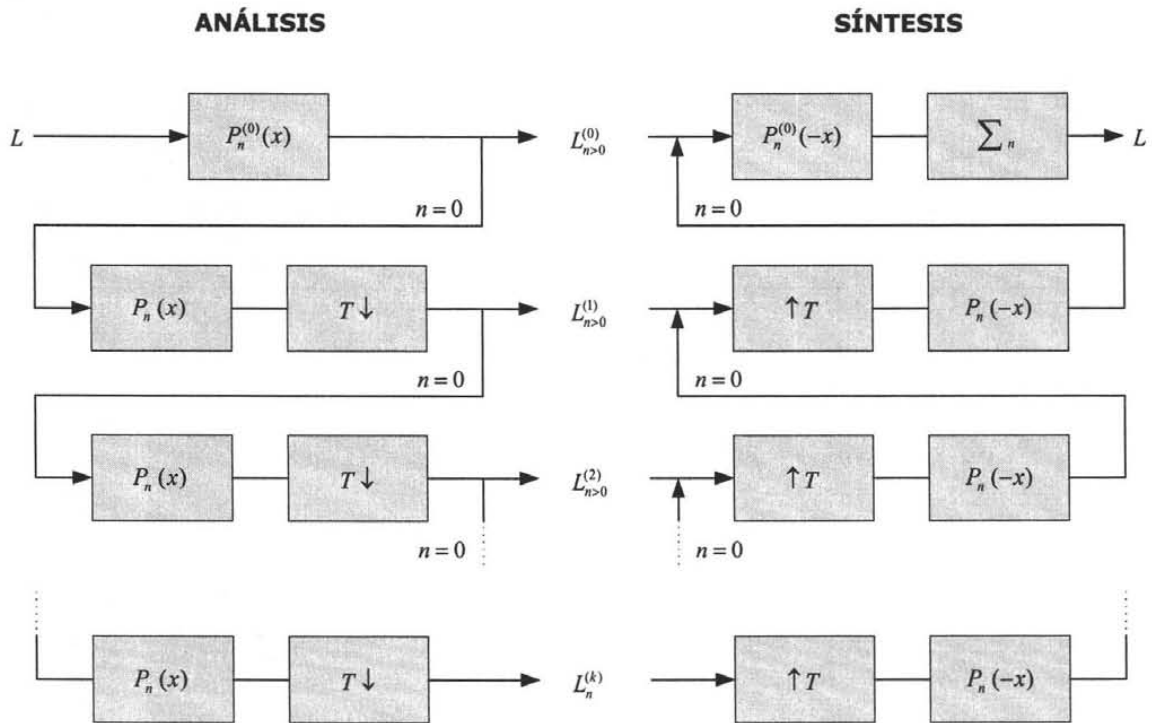


Figura 4.3: Esquema piramidal de análisis y síntesis de la THM.

4.9.6. Predicción entre niveles

Una desventaja de la THM es que incrementa las dimensiones del espacio de representación de la señal (de posición a posición-escala-orden de derivación) que en la práctica se traduce

en un incremento en la redundancia. Parte de esa redundancia se debe a la dependencia lineal entre las funciones de análisis de distintas escalas. En particular, se puede demostrar que las funciones del nivel $k-1$ se relacionan con las del nivel k mediante

$$\sqrt{\tau^n} G_n^{(k-1)}(x, \xi) = \sum_{m=n}^{\infty} c_{n,m} \int_{-\infty}^{\infty} \sqrt{\tau^m} G_m^{(k)}(x, \eta) P_{m-n}(\eta - T^{-1}\xi) d\eta, \quad (4.66)$$

con las constantes $c_{n,m} = \sqrt{C_m^n T^{-m} [(1+T^2)/2]^{m-n}}$. Esta dependencia lineal entre las funciones de análisis permite insertar la THM en un esquema de predicción en el cual cada nivel se estima a partir del siguiente nivel.

4.10. Transformada de Hermite multiescala-multidireccional 2D

4.10.1. Descomposición multiescala de señales 2D

Para el caso de señales multidimensionales la representación multiescala se lleva a cabo filtrando la señal en cascada a lo largo de cada coordenada ya que la gaussiana multidimensional se expresa como un producto de gaussianas 1D de cada coordenada. En particular, para el caso bidimensional se definen los operadores de derivadas escaladas como $G_{m,n-m}(x, y, s) = G_m(x, s)G_{n-m}(y, s)$ para $m=0,1,\dots,n$ y $n=0,1,\dots$, donde n es el orden de derivación total. Por lo tanto, la descomposición de señales bidimensionales en canales DoG es análoga al caso de señales 1D.

4.10.2. Análisis estructural de señales 2D

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. La ventaja de usar derivadas de Gaussiana se hace evidente cuando se requiere representar estructuras orientadas ya que la función DoG es isotrópica. Se puede demostrar que el conjunto formado por las $n+1$ derivadas bidimensionales de orden n permite representar

patrones de hasta n orientaciones simultáneas. Para propósitos de análisis de orientación es conveniente trabajar con las versiones rotadas de estos operadores. Sean los operadores $G_{m,n-m}(x, y, s, \theta) = G_{m,n-m}(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, s)$, donde θ es el ángulo de rotación, se puede demostrar que estas funciones están relacionadas linealmente con las funciones referidas al sistema original. Las derivadas direccionales de Gaussiana (DDG), denotadas por $G_{m,0}(x, y, s, \theta)$ se han empleado eficazmente en la detección y discriminación de orientaciones. Estos operadores son ideales para la descripción de estructuras orientadas porque satisfacen la llamada propiedad de direccionamiento (steering), a partir de la cual se demuestra que sólo se requiere conocer la respuesta de la DDG de orden n en $n+1$ direcciones para poder conocer la respuesta en cualquier dirección. La expansión en serie de Taylor de la DoG bidimensional puede escribirse ya sea en función de las DG bidimensionales o bien en función de las DDG. La primera expansión produce la THM en dos dimensiones, que no es otra cosa que la transformación definida en la sección anterior pero aplicada sobre ambas coordenadas espaciales. El segundo caso es más interesante y vale la pena describirlo con mayor detalle. En este caso la expansión de la DoG se escribe como

$$DoG(x, y, s_{k-1}, s_k) = \sum_{n=1}^{\infty} \sum_{j=0}^n \frac{c_n (-s_k)^n}{n!} G_{2n,0}(x, y, s_k, \theta_j), \quad (4.67)$$

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

$$c_n = \frac{1}{\sum_{j=1}^n \text{sen}^{2n} \left(\frac{j\pi}{n+1} \right)} = \frac{1}{n+1} \sum_{m=0}^n \frac{(C_n^m)^2}{C_{2n}^{2m}}, \quad (4.68)$$

para $n=1, 2, \dots$. Note que el orden de derivación n determina la resolución angular y que el ángulo θ_0 es un parámetro libre que puede ser elegido según convenga.

4.10.3. Definición de la Transformada de Hermite Multiescala Multidireccional (THMM)

Partiendo de la expansión 4.67 y procediendo como en el caso 1D, se construye una transformada que resulta 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

$$L_n^{(k,j)}(\xi, \eta) = \langle L(x, y) - G_n^{(k,j)}(x, y, \xi, \eta) \rangle_{(x,y)}, \quad (4.69)$$

mediante las funciones de análisis

$$G_n^{(k,j)}(x, y, \xi, \eta) = \frac{1}{2s_k} G_{n,0}^* \left(\frac{x \cos \theta_j + y \sin \theta_j}{\sqrt{2s_k}} - \xi, \frac{-x \sin \theta_j + y \cos \theta_j}{\sqrt{2s_k}} - \eta \right), \quad (70)$$

con $\theta_j = \theta_0 + j/(n+1)$ para $j=0, \dots, n$; $n=1, 2, \dots$ y $s_k = (1-\tau)^{-k}$ para todo k entero, es referido como **transformada de Hermite multiescala multidireccional** (THMM) [21] porque emplea las DDG a múltiples escalas y orientaciones y por tanto, al proceso de reconstrucción de la señal a partir de esta representación se le denomina THMM inversa. En este caso, la reconstrucción de la señal

$$L(x, y) = \sum_{k=-\infty}^{\infty} \sum_{n=1}^{\infty} \sum_{j=0}^n c_n \tau^n D_n^{(k,j)}(x, y), \quad (4.71)$$

se lleva a cabo sumando las contribuciones de las funciones de detalle direccional

$$D_n^{(k,j)}(x, y) = \langle L_n^{(k,j)}(\xi, \eta) G_n^{(k,j)}(x, y, \xi, \eta) \rangle_{(\xi, \eta)}, \quad (4.72)$$

sobre el conjunto de escalas y orientaciones previamente definido.

4.10.4. Filtrado direccional

La representación de imágenes mediante la THMM tiene la desventaja de que las funciones de análisis de la ecuación 4.70 son no-separables y, por lo tanto, su implantación puede resultar costosa en cuanto a técnicas computacionales. Sin embargo, ese costo puede ser reducido si

se calcula la transformada del caso separable y posteriormente se transforman los coeficientes del mismo orden total a través de las funciones de ángulo

$$\alpha_{m,n-m}(\theta) = \sqrt{C_n^m} \cos^m \theta \operatorname{sen}^{n-m} \theta, \quad (4.73)$$

para $m=0,1,\dots,n$, ya que las funciones de análisis en el punto $(\xi, \eta) = (0,0)$ están relacionadas linealmente como

$$G_n^{(k,j)}(x, y) = \sum_{m=0}^n \alpha_{m,n-m}(\theta_j) G_m^{(k)}(x) G_{n-m}^{(k)}(y), \quad (4.74)$$

para $n=1,2,\dots$. La proyección de la señal sobre las funciones de análisis equivale a una convolución con los filtros direccionales que se obtienen reflejando las funciones de la ecuación 4.74 respecto a x y y , más la normalización de coordenadas respecto a escala y orientación. El filtrado direccional es ampliamente citado en la literatura como mecanismo para hacer explícita la información relevante de una imagen. Por ejemplo, considérese una estructura 1D orientada con un ángulo ϕ medido respecto al eje x . La THMM que se obtiene para esta estructura tiene la forma $L_n^{(k,j)} = \cos^n(\theta_j - \phi) H_n^{(k)}$ donde $H_n^{(k)}$ es la THM del perfil 1D de la estructura. Estos coeficientes 1D contienen información sobre contraste y nivel de luminosidad, mientras que los coeficientes $L_n^{(k,j)}$ contienen además información sobre la localización y orientación de la estructura a distintas escalas.

4.11. Transformada de Hermite multiescala (THDM) y multiescala multidireccional (THDMM), Caso discreto.

4.11.1. Espacio-escala discreto

En la práctica las imágenes obtenidas son representadas digitalmente para su manipulación y almacenamiento por lo que es necesaria una formulación discreta del modelo. Un problema concerniente a la representación espacio-escala de señales discretas es cómo discretizar la teoría espacio-escala y al mismo tiempo mantener las propiedades del caso continuo. Para

señales unidimensionales se puede formular una teoría discreta completa siguiendo el mismo axioma de causalidad del caso continuo [21]; $H(\cdot, s)$ es un kernel discreto de la representación espacio-escala si, y solo si, el número de extremos locales de $L * H$ no excede el número de extremos de L . El único kernel de soporte compacto que satisface este axioma es la función binomial generalizada. Y si la definición se combina con el requerimiento de que la familia de transformaciones de suavizado debe obedecer la propiedad de semi-grupo, $H(\cdot, s_2) * H(\cdot, s_1) = H(\cdot, s_2 + s_1)$, entonces la única función que puede ser candidata para kernel del espacio-escala discreta es la función binomial simétrica $B[x, N] = 2^{-2N} C_{2N}^{x+N}$ para $N \in \mathbb{Z}^+, x = -N, \dots, N$ y $B[x, N] = 0$ para cualquier otro valor de x . Este kernel satisface varias propiedades en el dominio discreto que son similares a las del kernel Gaussiano en el dominio continuo; por ejemplo, tiende a la función delta discreta cuando $N \rightarrow 0$, mientras que para valores grandes de N se aproxima a la gaussiana. A pesar de que esta formulación de la teoría espacio-escala es completa y análoga al caso continuo, no puede ser extrapolada a espacios de más de una dimensión. Se ha demostrado que usando argumentos similares al caso continuo no existe ningún kernel trivial en dos o más dimensiones que garantice no introducir nuevos extremos locales. Sin embargo, la causalidad de los extremos locales puede ser expresada alternativamente: Si para un nivel de escala s_0 , un punto x_0 es un máximo (mínimo) local a ese nivel, entonces su valor no debe incrementar (decrementar) al incrementar la escala. Intuitivamente, este requerimiento previene que los extremos locales sean realizados o aparezcan de la nada. Basado en este axioma, se puede generalizar el kernel binomial a dos dimensiones como $B[x, y, N] = B[x, N]B[y, N]$. En este caso, aunque la escala es la misma en las orientaciones vertical y horizontal, no es así para las orientaciones oblicuas; lo cual supone una limitación para el análisis de orientaciones.

4.11.2. Derivadas discretas

El equivalente discreto de las derivadas escaladas se escribe en términos de diferencias finitas del kernel binomial, es decir, $B_n[x, N] = \delta_x^n B[x - N, N - n/2]$ para $n=0, \dots, 2N$. En lo sucesivo se hará referencia a estas funciones como Diferencias de Binomial (DB). Se puede demostrar que las DB tienden asintóticamente a las DG al crecer la escala, específicamente, $B_n[x, N] \sim G_n(x, N/4), N \rightarrow \infty$. De ahí que la mayoría de sus propiedades puedan predecirse con bastante precisión a partir de las propiedades correspondientes del caso continuo

haciendo $N=4s$. La generalización de estos operadores a 2D es inmediata por la separabilidad del kernel binomial, es decir, $B_{n,m}[x, y, N] = B_n[x, N]B_m[y, N]$. Por lo que el filtro de derivada direccional se puede definir convenientemente como

$$B_n[x, y, N, \theta_j] = \sum_{m=0}^n C_m^n \cos^m(\theta_j) \text{sen}^{n-m}(\theta_j) B_m[x, N] B_{n-m}[y, N], \quad (4.75)$$

donde el parámetro de orientación está discretizado como en el caso continuo: $\theta_j = \theta_0 + j\pi/(n+1)$ para $j=0, \dots, n$ y un valor θ_0 arbitrario.

4.11.3. Definición de la transformada de Hermite multiescala discreta (THDM)

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 DB se expresa como una suma finita. A partir de esta aproximación se construye la versión discreta de la THM a la que denominamos THDM. La descomposición de una señal discreta se representa como una suma finita de las funciones de detalle discretas $D_n^{(k)} = (L_n^{(k)})^{\uparrow T_k} * T_k B_n^{(k)}$, esto es,

$$L[x] = D_0^{(k)}[x] + \sum_{n=1}^{N_0} D_n^{(0)}[x] + \sum_{k=1}^K \sum_{n=1}^{\tau N_k} \Gamma_n^{(k)}(\tau) D_n^{(k)}[x], \quad (4.76)$$

donde, como en el caso continuo, la fase de análisis $L_n^{(k)} = (L * (-1)^n B_n^{(k)})^{\downarrow T_k}$ puede verse como la proyección de la señal de entrada sobre las funciones de análisis

$$B_n^{(k)}[x, \xi] = B_n^*[x - T_k \xi, N_k / 2]. \quad (4.77)$$

Note la similitud de estas expresiones con las obtenidas para el caso continuo con número finito de canales DoG. De hecho, los coeficientes $\tau_n^{(k)}$ tienden a τ^n a medida que el índice de escala k crece ($N_k \rightarrow \infty$). El primer término de la ecuación 4.76 corresponde al residuo pasa bajas, el segundo, al canal DoB entre las escalas cero y N_0 , y el tercero, a los canales

espaciados logarítmicamente con parámetro de espaciamento $\tau = (N_k - N_{k-1})$. El factor de normalización $T_k = \sqrt{N_k / 2}$ debe ser entero. Evidentemente, el valor $T_k = 2^k$ genera una secuencia de escalas $N_k = 2^{2k+1}$ cuya constante de espaciamento es $\tau=0.75$.

4.11.4. Definición de la transformada de Hermite multiescala multidireccional (THDMM)

La THDM se generaliza directamente a dos o más dimensiones en forma similar al caso continuo, mientras que la versión discreta de la THMM se calcula a partir de ésta mediante la definición de la ecuación 4.75. Es decir, para un punto genérico se tiene que

$$L_n^{(k,j)} = \sum_{m=0}^n \alpha_{m,n-m}(\theta_j) L_{m,n-m}^{(k)} \quad (4.78)$$

para el análisis de la señal, mientras que para la síntesis de la señal se efectúa la THDM inversa con los coeficientes

$$H_{m,n-m}^{(k)} = c_n \sum_{j=0}^n \alpha_{m,n-m}(\theta_j) L_n^{(k,j)} \quad (4.79)$$

Es de esperarse que, debido al submuestreo, los coeficientes $L_{m,n-m}^{(k)}$ y $H_{m,n-m}^{(k)}$ reproduzcan imágenes ligeramente diferentes. Por lo general suele elegirse $\theta_0 = \arctan(L_{0,1}^{(k)} / L_{1,0}^{(k)})$ en todos los niveles, lo que ocasiona que $L_1^{(k,0)} = 0$. La ventaja de la THDMM es que tiende a compactar más la energía a lo largo de una coordenada.

4.12. Clasificación de estructuras en el dominio de la transformada de Hermite discreta rotada (THDR)

En esta sección se presenta un esquema de representación de imágenes basado en una clasificación perceptual de los coeficientes de la **transformada de Hermite discreta rotada** (THDR) sobre una retícula de muestreo aproximadamente hexagonal. Las funciones de

análisis de la THD corresponden a funciones binomiales que aproximan a las derivadas de Gaussiana con la ventaja de que pueden ser calculadas mediante un algoritmo rápido. El muestreo hexagonal es muy eficiente en la transferencia de información espacial y además tiene gran relevancia en el HVS. Los coeficientes de un mismo orden se mapean mediante una transformación unitaria especificada localmente. La transformación se basa en funciones binomiales generalizadas, y como resultado se obtienen los coeficientes de la THD referidos a un sistema de coordenadas rotado de tal forma que la energía de los coeficientes se concentra a lo largo de una coordenada. Dicha representación permite un proceso de clasificación perceptual en el cual se define un umbral de error máximo tolerado para la aproximación del bloque por una constante (patrón 0D) o por una estructura orientada (patrón ID).

4.12.1. Descomposición de la imagen

El primer paso consiste en el mapeo de los datos de la imagen (mediante la THD a una sola escala) en un conjunto de coeficientes espectrales a partir de los cuales es posible obtener una reconstrucción perfecta de la imagen. La THD se basa en los filtros binomiales

$$B_{n,k} = (-1)^n 2^{-N} \sqrt{C_N^n} \Delta_k^n \{C_{N-n}^{k-n}\}, \quad (4.80)$$

para $n, k=0, \dots, N$ donde n es el orden de derivación y $N+1$ es la longitud del filtro. Para las pruebas se emplearon filtros con N igual a una potencia de 2. Por otro lado, se definen las funciones binomiales

$$A_{n,k} = \sqrt{2^{-N}} \Delta_k^n \{C_{N-n}^{k-n}\}. \quad (4.81)$$

para $n, k=0, \dots, N$, las cuales están relacionadas con los filtros binomiales mediante $B_{n,k} = W_k A_{n,n-k}$, donde

$$W_k = \sqrt{2^{-N} C_N^k}, \quad (4.82)$$

es la ventana binomial. Las funciones binomiales satisfacen un número de propiedades importantes las cuales pueden ser estudiadas usando la representación matricial $A = [A_{n,k}]_{n,k=0,\dots,N}$. Por ejemplo, para $N=8$ la matriz se escribe explícitamente como

$$\frac{1}{\sqrt{2^8}} \begin{pmatrix} 1 & 8 & 28 & 56 & 70 & 56 & 28 & 8 & 1 \\ 1 & 6 & 14 & 14 & 0 & -14 & -14 & -6 & -1 \\ 1 & 4 & 4 & -4 & -10 & -4 & 4 & 4 & 1 \\ 1 & 2 & -2 & -6 & 0 & 6 & 2 & -2 & -1 \\ 1 & 0 & -4 & 0 & 6 & 0 & -4 & 0 & 1 \\ 1 & -2 & -2 & 6 & 0 & -6 & 2 & 2 & -1 \\ 1 & -4 & 4 & 4 & -10 & 4 & 4 & -4 & 1 \\ 1 & -6 & 14 & -14 & 0 & 14 & -14 & 6 & 1 \\ 1 & -8 & 28 & -56 & 70 & -56 & 28 & -8 & 1 \end{pmatrix}, \quad (4.83)$$

donde cada renglón representa una máscara de análisis. Las funciones binomiales satisfacen la propiedad de ortogonalidad $AA=I$, donde I es la matriz identidad. Por lo tanto, la transformación definida por la matriz A lleva a cabo las mismas operaciones tanto para el análisis como para la síntesis. La salida de los filtros binomiales en una posición genérica se puede obtener mediante la transformación definida por la matriz $B=WA$, donde $W = \text{diag}(W_k)_{k=0,\dots,N}$ es una matriz diagonal que representa la ventana binomial. De la propiedad de ortogonalidad de la matriz A se puede demostrar que $BB^T=W^2$. Si F es una matriz cuadrada que contiene los datos de un bloque de $(N+1) \times (N+1)$ de la imagen de entrada, entonces $G=BF B^T$ son coeficientes de la THD. Se dice que la THD es una expansión local de los datos de la imagen ya que para cada punto sobre una retícula de muestreo se puede obtener la representación local $W^2 F W^2 = B^T G B$.

4.12.2. Rotación local

La rotación local en el dominio de la transformada puede verse como un mapeo de los coeficientes espectrales a través de una transformación unitaria. Esta se lleva a cabo sobre los coeficientes de un mismo orden y puede ser representada matricialmente como se describe a continuación. Sean g_n para $n=0,\dots,2N$, los vectores columna formados por las

antidiagonales de la matriz de coeficientes G y sean las funciones binomiales generalizadas definidas como

$$A_{n,k}^{(\theta)} = c^{-k} s^k \Delta^n \{ C_{N-n}^{k-n} c^{2k-n} s^{N-2k+n} \}, \quad (4.84)$$

con $c = \cos \theta$ y $s = \sin \theta$ para $0 \leq \theta < \pi$. Se puede ver que si el parámetro de sesgo θ se hace igual a $\pi/4$ se obtiene el caso simétrico, es decir, $A_{n,k} = A_{n,k}^{(\pi/4)}$. Usando la notación matricial se denotará la matriz de funciones binomiales generalizadas por $A^{(\theta)}$. Esta matriz satisface la misma propiedad de ortogonalidad del caso simétrico. Las primeras tres matrices (excepto el caso $N=0$, el cual se reduce a la unidad) tienen la siguiente expresión explícita

$$\begin{pmatrix} s & c \\ c & -s \end{pmatrix}, \begin{pmatrix} s^2 & 2sc & c^2 \\ sc & c^2 - s^2 & -sc \\ c^2 & -2sc & s^2 \end{pmatrix}, \begin{pmatrix} s^3 & 3s^2c & 3sc^2 & c^3 \\ s^2c & 2sc^2 - s^3 & c^3 - 2s^2c & -s^2c \\ s^3 & -3s^2c & 3sc^2 & -c^3 \end{pmatrix}. \quad (4.85)$$

La rotación de los vectores de coeficientes para un ángulo θ dado se obtienen a partir de la transformación definida por la matriz $R^{(\theta)} = WA^{(\theta)}W^{-1}$, es decir que el vector $g_n^{(\theta)} = R^{(\theta)}g_n$ representa los coeficientes de orden n referido a los ejes coordenados rotados por un ángulo θ medido en sentido inverso a las manecillas del reloj, mientras que la rotación inversa se consigue aplicando la matriz de rotación transpuesta a los vectores de coeficientes rotados. En lo subsecuente se hará referencia a los coeficientes rotados a través de la matriz $G^{(\theta)}$ cuyas antidiagonales son justamente los vectores $g_n^{(\theta)}$. En las pruebas presentadas aquí, se eligió el ángulo del gradiente $\theta = \arctan(G_{0,1} / G_{1,0})$.

Algoritmo de fusión

El algoritmo diseñado en el laboratorio de procesamiento digital de imágenes esta basado en los esquemas de fusión por transformada [16, 17, 22].

Podemos describir este algoritmo de fusión en 3 fases o etapas:

1. Empleo de una transformación, para nuestro caso la transformada de Hermite rotada que se ha discutido en capítulo 4, sección 4.12, que descompone las señales (imágenes) y fusionarlas (las Figuras 5.2 y 5.3 muestran los coeficientes de Hermite no rotados y los coeficientes rotados se muestran en las Figuras 5.4 y 5.5).

a. Rotación Local (Figura 5.1):

- i. Operador derivativo gaussiano 2D

$$D_{n,m} = (x, y; s) = D_n(x; s)D_m(y; s)$$

- ii. Operador derivativo gaussiano rotado 2D

$$D_{n,m} = (x, y; s, \theta) = D_{n,m}(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta; s)$$

$$D_{n-m,m} = (x, y; s, \theta) = \sum_{k=0}^n \alpha_{k,n-k;n-m,m}(\theta) D_{k,n-k}(x, y; s)$$

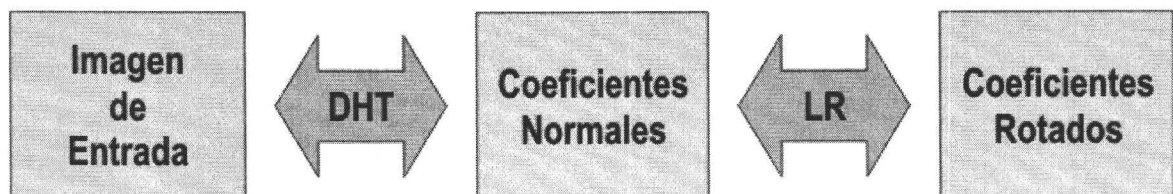
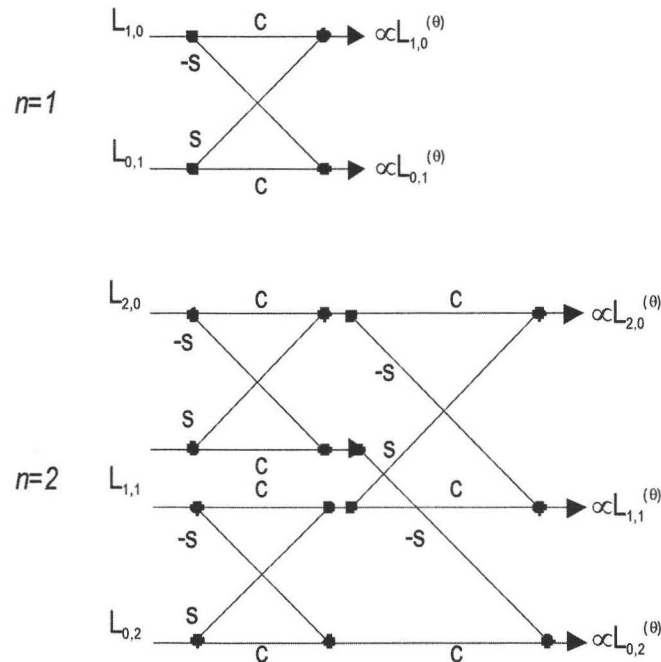


Figura 5.1: Rotación local.

b. Algoritmo de Rotación

i. Mariposas de binomiales generalizados con $s = \text{sen}\theta, c = \text{cos}\theta$

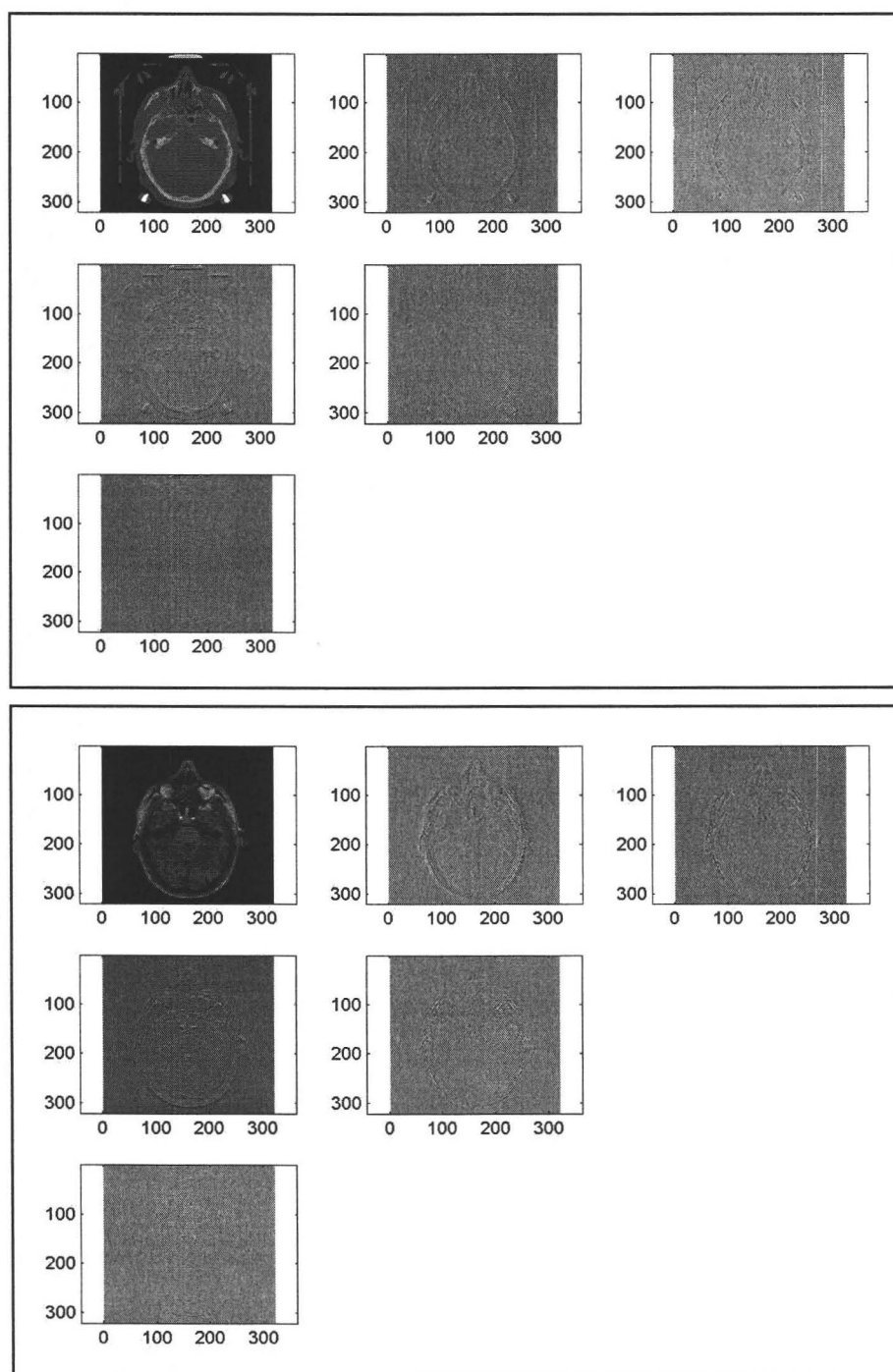


$$\theta = \arctan \frac{L_{0,1}}{L_{1,0}}$$

2. Se hace uso de algunos métodos de selección descritos en este capítulo; se realizan las selecciones de los componentes de baja y alta frecuencia.
3. Se emplea la transformada de Hermite rotada para reconstruir la imagen y así obtener el proceso de fusión sobre esta imagen final.

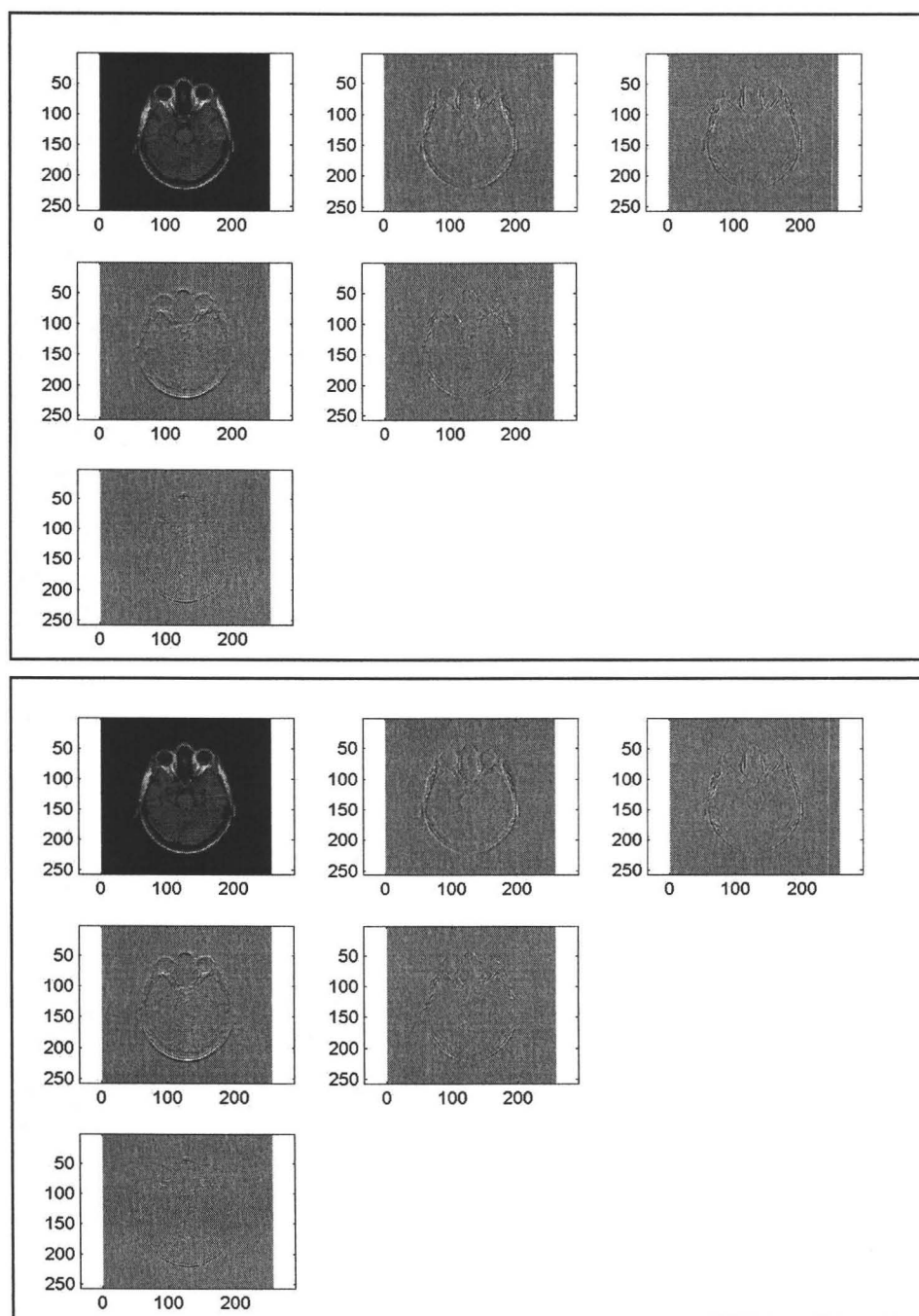
Si tenemos las siguientes imágenes, Figura 3.14 y hacemos uso de las herramientas matemáticas sobre la transformada de Hermite y aplicamos el algoritmo de fusión, obtenemos los coeficientes de baja y alta frecuencia para cada una de las imágenes de la Figura 3.14.

Las Figuras 5.2 y 5.3 muestran los coeficientes por medio de la transformada de Hermite y empleando la rotación de esta transformada tenemos los coeficientes de las imágenes en las Figuras 5.4 y 5.5.



A
B

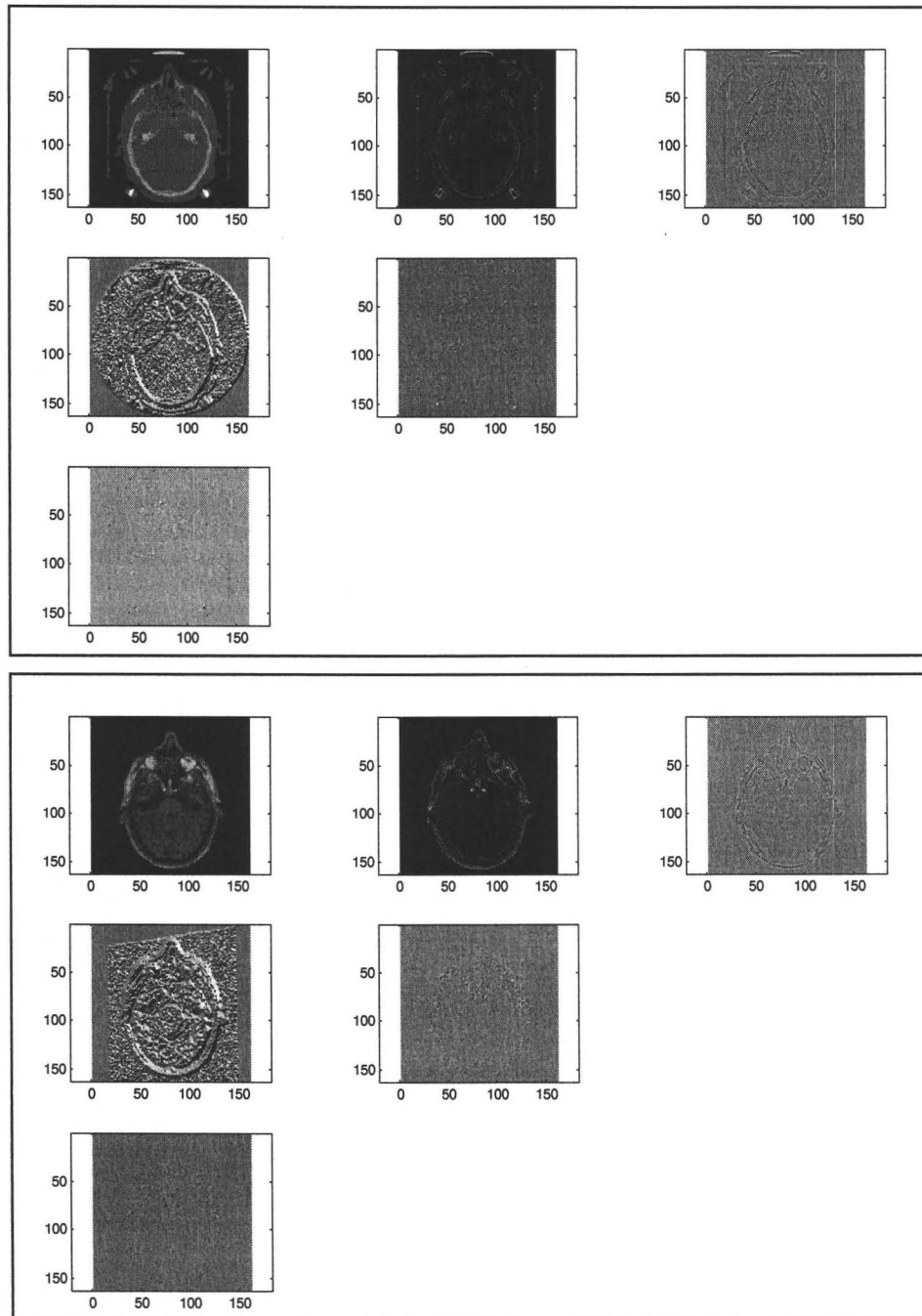
Figura 5.2: Coeficientes de las imágenes de entrada por medio de la **Transformada de Hermite**. **A.** Coeficientes de la imagen de la Figura 3.14 **A.** **B.** Coeficientes de la imagen de la Figura 3.14 **B.**



A

B

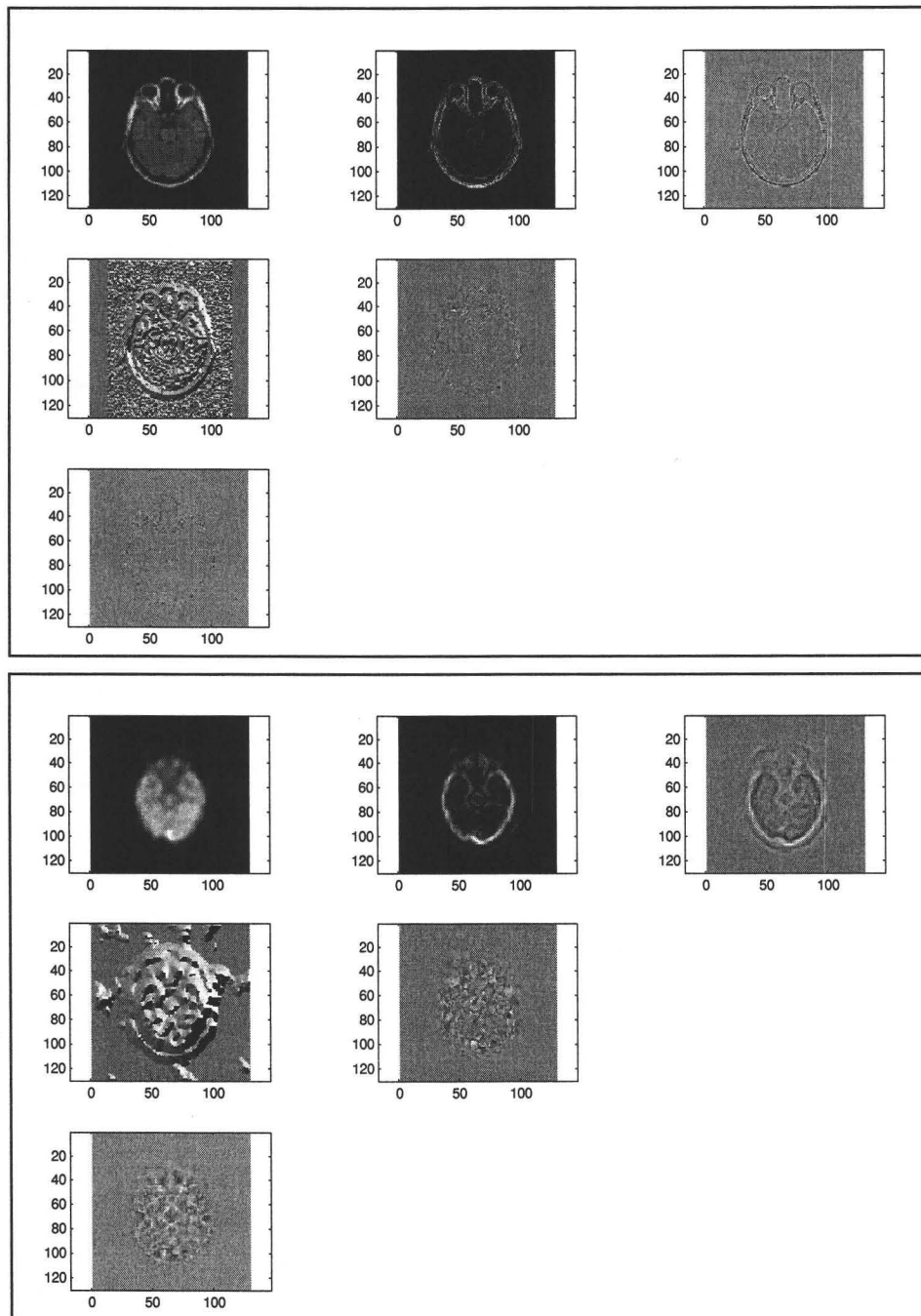
Figura 5.3: Coeficientes de las imágenes de entrada por medio de la **Transformada de Hermite**. **A.** Coeficientes de la imagen de la Figura 3.14 C. **B.** Coeficientes de la imagen de la Figura 3.14 D.



A

B

Figura 5.4: Coeficientes de las imágenes de entrada por medio de la **Transformada de Hermite Rotada**. **A.** Coeficientes de la imagen de la Figura 3.14 A. **B.** Coeficientes de la imagen de la Figura 3.14 B.



A

B

Figura 5.5: Coeficientes de las imágenes de entrada por medio de la **Transformada de Hermite Rotada**. **A.** Coeficientes de la imagen de la Figura 3.14 C. **B.** Coeficientes de la imagen de la Figura 3.14 D.

Para realizar el proceso de fusión se une la información de los dos conjuntos de coeficientes, siendo necesario hacer uso de dos tipos de métodos, uno para los componentes de baja frecuencia y otro para los de alta frecuencia.

Para comprender el proceso de descomposición de las imágenes por medio de la transformada de Hermite observemos la Figura 5.6, donde se explicará la selección de coeficientes en el proceso de dicha fusión y su reconstrucción de la imagen a partir de los coeficientes seleccionados.

Se toma el coeficiente que contiene las características de las zonas homogéneas (coeficiente de baja frecuencia), éste se selecciona por el método llamado selección de pasa bajas [16, 17], para extraer las características principales de dicho coeficiente. Para los coeficientes, los de alta frecuencia, y que contienen características de bordes en las imágenes, se utiliza por el método llamado selección de pasa altas [16, 17].

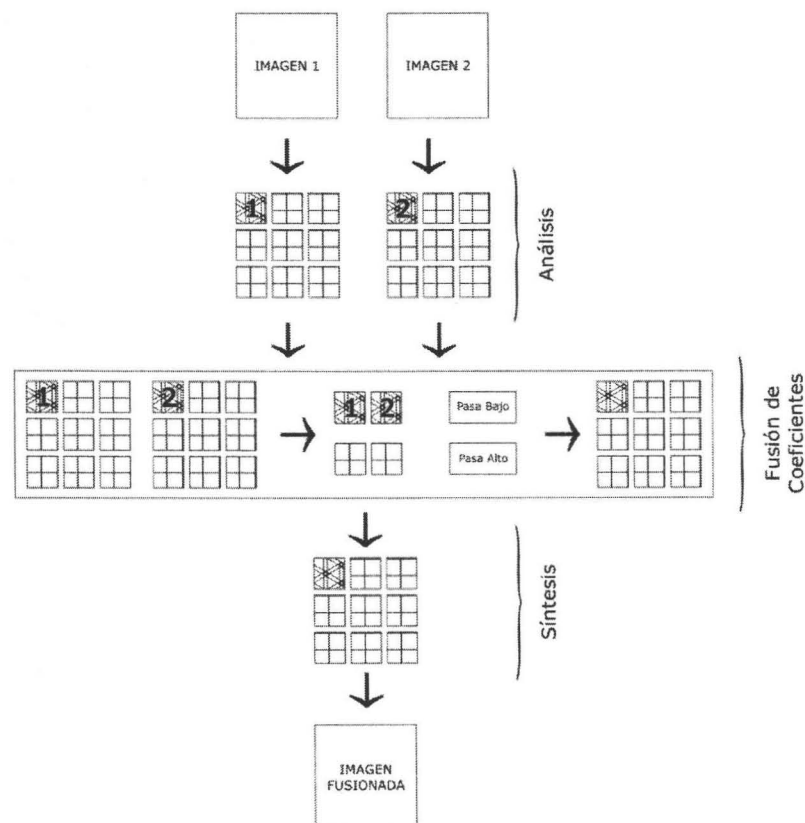


Figura 5.6: Proceso de descomposición y fusión de las imágenes.

5.1. Selección del Coeficiente de Baja Frecuencia.

Los métodos que se emplean para seleccionar los componentes de baja frecuencia son:

- Selección del coeficiente de la imagen 1.
- Selección del coeficiente de la imagen 2.
- Selección del promedio de los coeficientes de las imágenes 1 y 2

Si se selecciona el coeficiente de la imagen 1, la imagen fusionada tendrá mayor cantidad de características de dicha imagen seleccionada; si se selecciona el coeficiente de la imagen 2 sucede exactamente lo descrito anteriormente pero para la imagen 2. Ahora si se selecciona el promedio de las imágenes 1 y 2, la imagen fusionada, contendrá la información promediada de los coeficientes de ambas imágenes. En la Figura 5.7 podemos observar esquemáticamente dicho proceso.

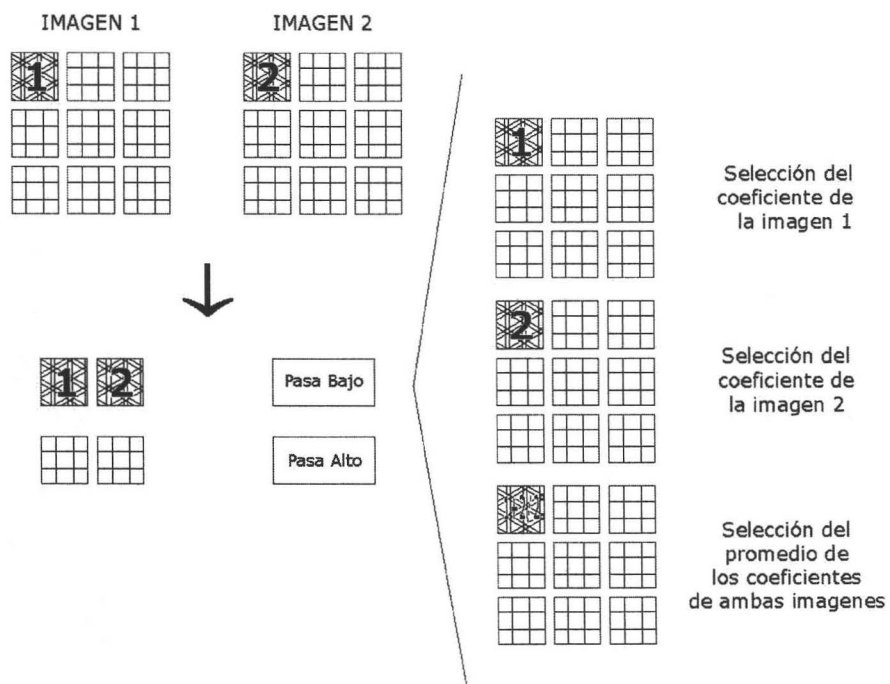


Imagen 5.7: Selección del coeficiente de baja frecuencia.

5.2. Selección del Coeficiente de Alta Frecuencia.

Para este caso la selección de los coeficientes se emplea los siguientes métodos:

1. Selección del máximo.
2. Selección del saliente y medida de paridad
3. Selección del máximo y verificación de consistencia

A continuación se detallan cada uno de los métodos.

5.2.1. Selección del máximo [16, 17]

Para esta técnica, de los pares de valores de los coeficientes que se tienen de las imágenes, se selecciona el valor que tenga mayor magnitud, así la imagen resultante (fusionada) contendrá la información de las dos imágenes. Para visualizar esta técnica consideremos los valores de los coeficientes de alta frecuencia de las imágenes mostradas en la Figura 5.8.

15	25	68	42	7	65	98	32	1	54
8	9	54	8	2	25	47	89	32	2
23	12	3	56	8	1	5	65	21	9
19	65	95	32	54	21	5	2	48	3
58	48	36	25	9	94	48	75	21	6

65	98	68	42	54
25	47	89	32	2
23	12	65	56	9
21	65	95	48	54
94	48	75	25	9



Figura 5.8: Selección del máximo. **A.** Coeficiente de A; **B.** Coeficiente de B y **C.** Coeficientes fusionados.

La rutina para esta selección se basa en:

- Datos de entrada: Coeficientes de A y B.
- Comparación punto a punto de los valores de cada coeficiente.
- Selección del máximo valor de la comparación previa, punto a punto.
- Construcción de la imagen resultante que contiene los valores máximos de los datos de entrada.

Se puede observar claramente que los coeficientes de las imágenes de entrada, únicamente se toman de ellos los valores de mayor peso numérico y son los que se encuentran contenidos en la imagen resultante.

5.2.2. Selección del saliente y la medida de la paridad [16, 17].

Para este método se calcula una matriz de valores salientes a partir de los valores de entrada (Figura 5.7 (A) y (B)), basada en autocorrelaciones que se obtienen de cada coeficiente. Para obtener el cálculo de la paridad se calcula la correlación cruzada entre los coeficientes de entrada, la cual es sopesada por las matrices de saliente antes calculadas. Dado un umbral predeterminado, obtenido anteriormente por los trabajos de O. Rockinger [16, 17], se verifica si los valores de paridad exceden o no a este umbral y se obtiene el valor correspondiente en la selección. Para observar como trabaja esta técnica tenemos la Figura 5.9.

0.9375	8.3125	9.9375	0.4375
3.3750	21.4375	26.6250	1.1875
7.4375	45.5000	40.2500	7.8125
3.6250	11.8750	5.9375	0.5625

S1=1.0e+003*

4.0625	18.3125	5.5000	3.3750
7.2500	49.0625	32.1250	4.1875
8.5625	25.4375	27.5625	1.3125
5.8750	16.5625	7.6875	0.3750

S2=1.0e+003*

0.0009	0.0691	0.988	0.0002
0.0114	0.4596	0.7089	0.0014
0.0553	2.0703	1.6201	0.0610
0.0131	0.1410	0.0353	0.0003

0.0165	0.3353	0.302	0.0114
0.0526	2.4071	1.0320	0.0175
0.0733	0.6471	0.7597	0.0017
0.0345	0.2743	0.0591	0.0001

0.4382	0.7527	0.8474	0.2550
0.7652	0.7338	0.9826	0.5249
0.9902	0.8519	0.9324	0.3268
0.8938	0.9471	0.9675	0.9231

4.0625	16.8425	6.1489	2.0728
6.9388	43.1958	31.0109	2.3511
8.4406	30.0831	30.3743	5.2343
4.1400	13.2313	6.4129	0.4976

A	B
C	D
E	F

Figura 5.9: Cálculo del saliente y medida de la paridad. **A.** Coeficiente de A; **B.** Coeficiente de B; **C.** Matriz saliente S1; **D.** Matriz saliente S2; **E.** Matriz de pares; **F.** Coeficientes fusionados.

La rutina empleada para esta selección se expresa como:

- Datos de entrada: Coeficientes de A y B.
- Cálculo de los Salientes:
 - Cálculo de autocorrelaciones para cada coeficiente.

- Cálculo de la paridad:
 - Se calcula la correlación de los coeficientes de entrada.
 - Se pondera el resultado de la correlación con la suma de los salientes calculados anteriormente.
- Selección de los valores:
 - Dado el umbral predeterminado, se compara con los valores de la matriz de paridad.
 - Se hace una comparación entre sí de los valores salientes.
 - El mayor valor resultante entre las comparaciones anteriores se emplean para reconstruir los valores del coeficiente de salida.

5.2.3. Selección del máximo y verificación de consistencia [16, 17]

Este es igual que el método de selección del máximo; sólo que después de los máximos coeficientes, en ésta, para cada posición y tomando una vecindad, por ejemplo de conectividad 8, se determina a qué coeficiente corresponde cada vecino y sumando después los pesos (según el coeficiente al que correspondan), dicha suma de mayor valor, va a indicar el coeficiente al que pertenece esa posición; si el valor inicial corresponde a la suma de mayor peso, no se modifica y se continúa evaluando el valor de la siguiente posición, de no ser así, se cambia por el valor en esa posición del coeficiente de mayor peso. Para verificar la consistencia o solidez, tomaremos como ejemplo una vecindad de conectividad 8, pudiendo observar así como trabaja esta técnica mostrada en la Figura 5.10.

0.9375	8.3125	9.9375	0.4375
3.3750	21.4375	26.6250	1.1875
7.4375	45.5000	40.2500	7.8125
3.6250	11.8750	5.9375	0.5625

4.0625	18.3125	5.5000	3.3750
7.2500	49.0625	32.1250	4.1875
8.5625	25.4375	27.5625	1.3125
5.8750	16.5625	7.6875	0.3750

4.0625	18.3125	9.9375	3.3750
7.2500	49.0625	32.1250	4.1875
8.5625	45.5000	40.2500	7.8125
5.8750	16.5625	7.6875	0.5625

4.0625	18.3125	5.5000	3.3750
7.2500	49.0625	32.1250	4.1875
8.5625	25.4375	40.2500	7.8125
3.6250	11.8750	5.9375	0.5625



Figura 5.10: Selección del máximo y verificación de consistencia. **A.** Coeficientes de A. **B.** Coeficientes de B. **C.** Coeficientes de los máximos. **D.** Coeficientes fusionados.

Las Figuras 5.10 (A) y (B) representan los coeficientes de entrada. La Figura 5.10 (C) nos muestra la selección de los máximos de los coeficientes iniciales, mientras que la Figura 5.10 (D) representa la selección final del coeficiente, donde se ha verificado si el valor seleccionado, con respecto a sus vecinos, es el indicado o se cambia por el del otro coeficiente.

Para verificar esta técnica y tomando una vecindad con conectividad 8, evaluamos la intensidad total de ambas imágenes respecto a un valor determinado. Por ejemplo, observemos el valor de la posición (3,2) en la Figura 5.10 (C), 45.5000, este pertenece al coeficiente de A; observemos sus vecinos cercanos en dicha conectividad, y a qué coeficiente

pertenecen: 49.0625 (B), 32.1250 (B), 40.2500 (A), 7.6875 (B), 16.5625 (B), 5.8750 (B), 8.5625 (B). Sumando las intensidades correspondientes para cada coeficiente tenemos 40.2500 del coeficiente A y 127.1250 del coeficiente B. Observamos que la mayor intensidad corresponde al coeficiente B (para este píxel en particular), por lo que, se cambia el valor de 45.5000 del coeficiente A, por el correspondiente en el coeficiente B (25.4375).

Podemos ver otro ejemplo, la posición (4,1), en la Figura 5.10 (C), tiene el valor de 5.8750 que pertenece al coeficiente de B. Si se observa la vecindad de éste (conectividad 8) tenemos: 8.5625 (B), 45.5000 (A), 16.5625 (B). Al sumar las intensidades respectivas a cada coeficiente tenemos: 45.5000 de A y 25.1250 de B; por lo que la intensidad mayor es la correspondiente en A se cambia el valor 5.8750 (B) por 3.6250 (A).

De esta manera, la consistencia o solidez va en función de que el valor del píxel se modifique, dependiendo si la mayor intensidad de la suma de los vecinos cercanos pertenece al coeficiente de procedencia del valor indicado o no.

El algoritmo para este método consiste en:

- Datos de entrada: Coeficientes de A y B.
- Comparación punto a punto de los valores de cada coeficiente.
- Selección del valor máximo de la comparación previa, punto a punto.
- Construcción de la imagen resultante con los valores máximos.
- De la imagen resultante con los máximos, y con una vecindad de conectividad 8, se suman las intensidades de los valores correspondientes a los coeficientes A y B respectivamente.
- Se comparan las sumas de las intensidades, se verifica si el valor del punto que se está evaluando corresponde al coeficiente de mayor valor de intensidad, de ser así, se deja el valor asignado, en caso contrario, se cambia por el valor correspondiente en posición al coeficiente con mayor intensidad.
- Se obtiene el coeficiente de salida.

Resultados de la fusión con imágenes médicas

La fusión de imágenes médicas se realizó empleando 3 etapas mediante la transformada polinomial de Hermite Rotada:

1. Análisis: En esta se realiza la descomposición en coeficientes de cada una de las imágenes a fusionar,
2. Selección de coeficientes: Consiste en hacer una selección entre los coeficientes de cada una de las imágenes de entrada para obtener uno solo conjunto de estos, y
3. Síntesis: Esta se lleva a cabo para construir el conjunto de coeficientes seleccionados y así obtener la imagen de salida (fusionada).

6.1 Imágenes de entrada.

Las imágenes de entrada (Figura 3.14), que se desean fusionar pertenecen a una secuencia de imágenes obtenidas por métodos tomográficos, la fusión se realizará a partir de las imágenes de: tomografía computarizada (Figura 3.14 A) y resonancia magnética (Figura 3.14 B), y tomografía por emisión de positrones (Figura 3.14 C) y resonancia magnética (Figura 3.14 D).

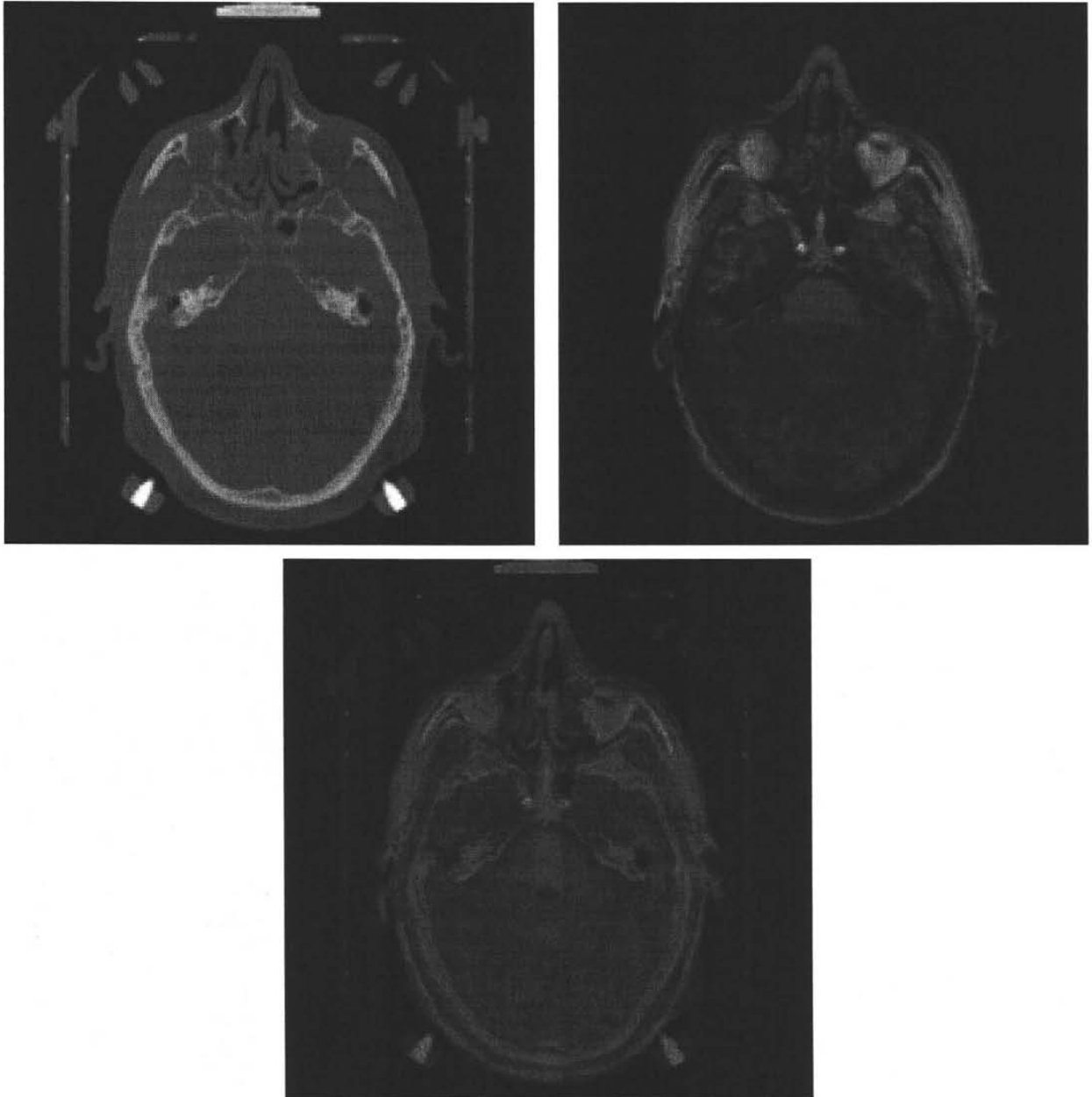
6.2 Selección del método del coeficiente de baja frecuencia para la fusión.

Con nuestras imágenes de entrada procederemos a evaluar los métodos de pasa bajos, ya discutido en capítulo 5 sección 5.1, para determinar cómo funciona y cuales son los efectos en la fusión. Hacemos mención que para llevar a cabo esta prueba se selecciona para el método pasa altas el de máximo por verificación de consistencia.

Los métodos del coeficiente de bajas frecuencias consisten en seleccionar el coeficiente de una de las dos imágenes o el promedio de éstas.

La Figura 6.1 muestra la fusión obtenida después de seleccionar los coeficientes de baja frecuencia de las imágenes de tomografía computarizada (Figura 3.14 A) y resonancia magnética (Figura 3.14 B).

En la Figura 6.1 A, podemos observar que existe mayor información de la tomografía computarizada (Figura 3.14 A), ya que en la fusión se tomó en cuenta el coeficiente de esta imagen. En la Figura 6.1 B, se observa ahora mayor información de la imagen de resonancia magnética (Figura 3.14 B), pues la fusión tomó el coeficiente de baja frecuencia de la resonancia magnética y por último en la Figura 6.1 C podemos ver la fusión tomando el promedio de los coeficientes de baja frecuencia observando evidencia de información de las dos imágenes de entrada.



A B
C

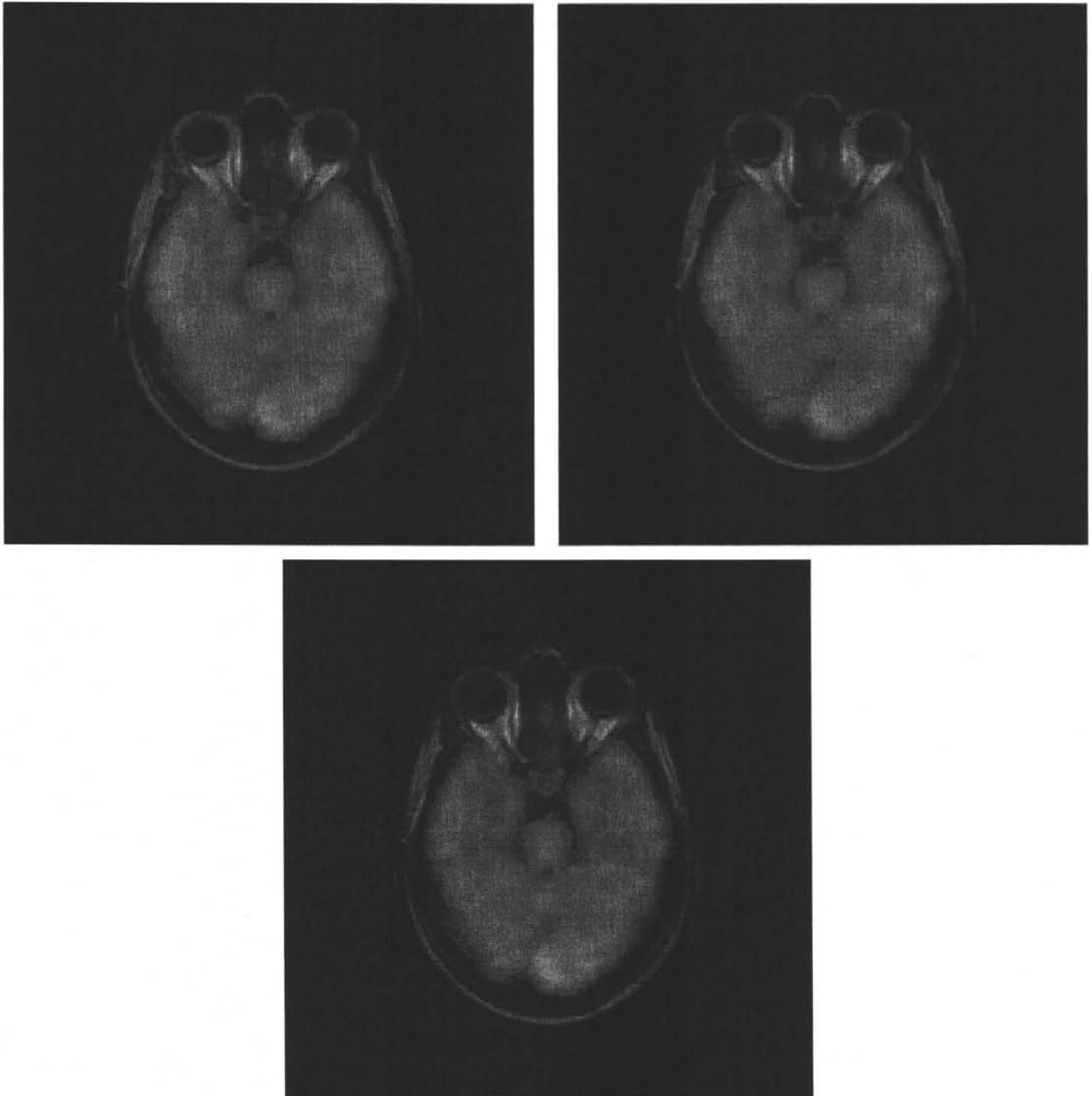
Figura 6.1: Selección de los métodos pasa bajas. **A.** Fusión tomando el coeficiente de la tomografía computarizada. **B.** Fusión tomando el coeficiente de baja frecuencia resonancia magnética y **C.** Fusión tomando el promedio de los coeficientes de baja frecuencia de la tomografía computarizada y la resonancia magnética.

6.3 Selección del método del coeficiente de altas frecuencias para la fusión.

Una vez visto el método de selección de coeficientes de bajas frecuencia, ahora evaluaremos el método de selección de coeficientes alta frecuencia. Para determinar cómo trabaja cada uno de los métodos empleados y saber cual nos da mejores resultados (vistos en el capítulo 5, sección 5.2) fusionaremos las imágenes de entrada de la Figura 3.14 C y D.

En la Figura 6.2, en (A) tenemos la fusión por medio de la selección del máximo, en (B) se muestra la fusión por selección del cálculo del saliente y medida de la paridad y en (C) se observa la fusión por selección del máximo y verificación de la consistencia. Hacemos mención que se utilizó el método para la selección del coeficiente de baja frecuencia el del promedio de ambas imágenes.

Para apreciar algunas diferencias cualitativas del resultado de la fusión (Figura 6.2) por los métodos empleados de selección de los coeficientes de alta frecuencia presentamos un método analítico (Figuras 6.3, 6.4 y 6.5) por medio de perfiles para dichos métodos y en la Figura 6.6 se encuentran superpuestas las columnas y las filas de cada uno de los tres métodos empleados.



A **B**
C

Figura 6.2: Selección de los métodos pasa altas. **A.** Fusión por medio de la selección del máximo. **B.** Fusión por selección del cálculo del saliente y medida de la paridad y **C.** Fusión por selección del máximo y verificación de la consistencia.

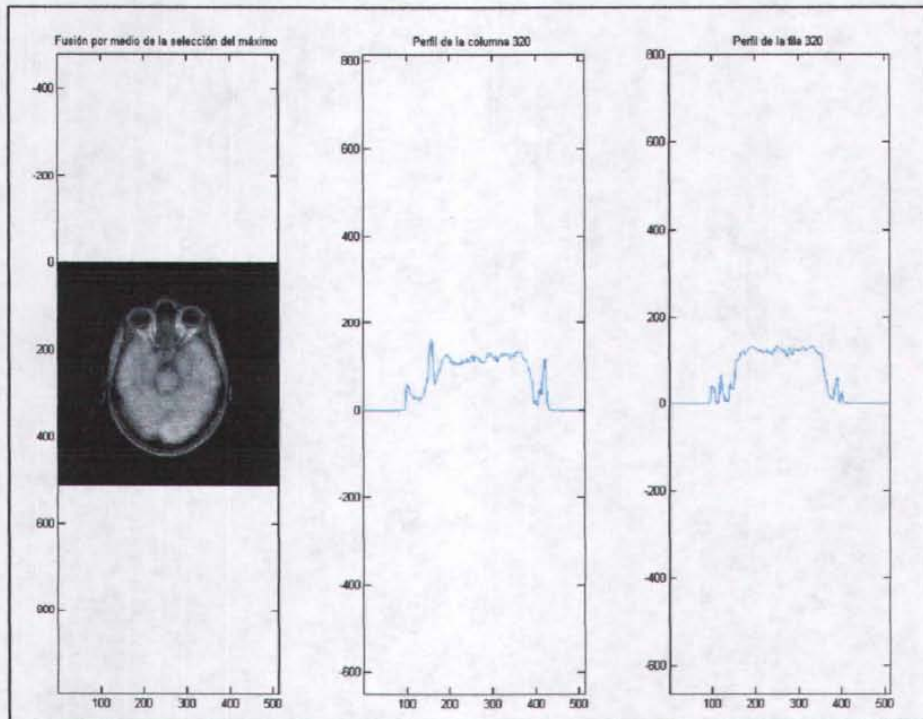


Figura 6.3: Perfil obtenido por la selección del coeficiente de alta frecuencia de la fila y columna 320 método de selección del máximo.

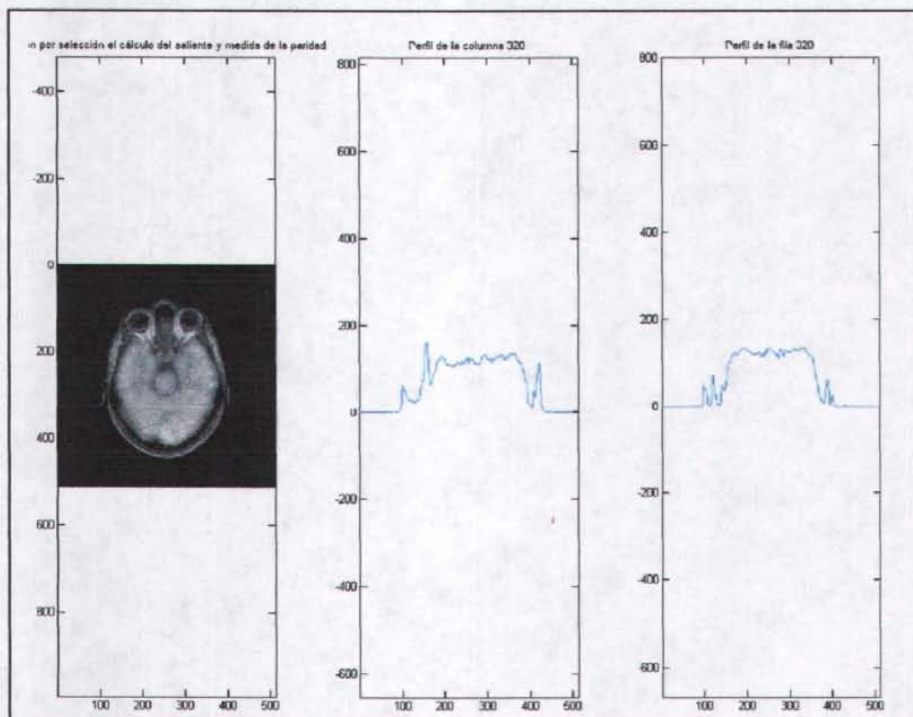


Figura 6.4: Perfil obtenido por la selección del coeficiente de alta frecuencia de la fila y columna 320 por el método de selección del cálculo del saliente y medida de la paridad.

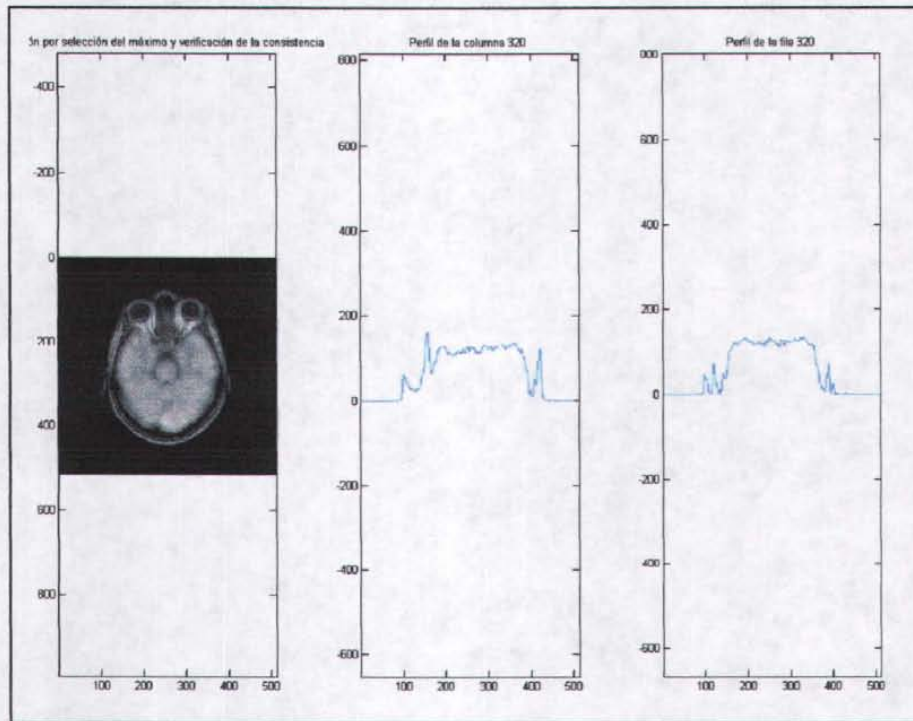
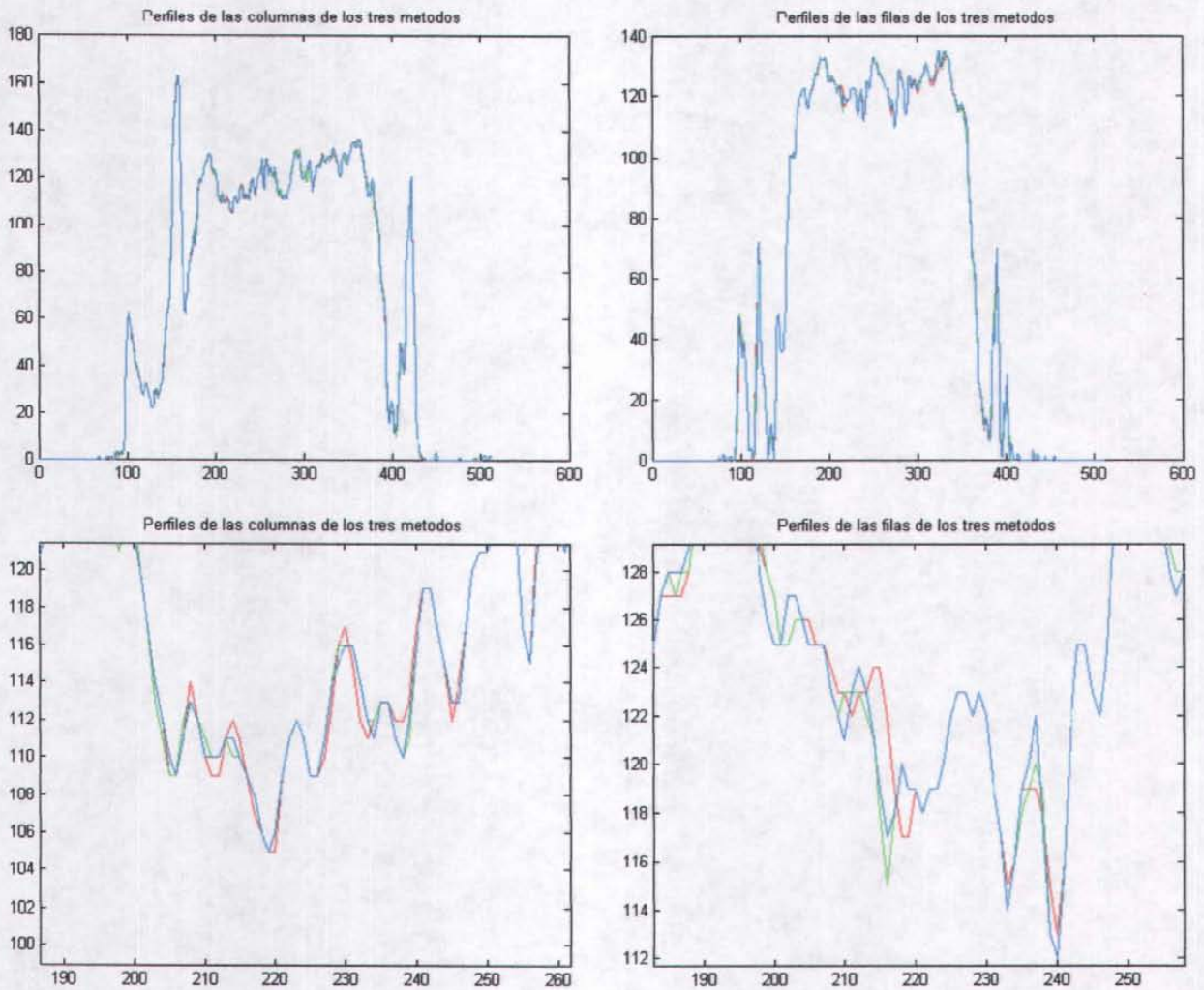


Figura 6.5: Perfil obtenido por la selección del coeficiente de alta frecuencia de la fila y columna 320 por el método de selección del máximo y verificación de la consistencia.



A	B
C	D

Figura 6.6: Perfiles de los 3 métodos pasa altas. **A.** Perfiles de las columnas. **B.** Perfiles de las filas. **C** y **D.** Acercamientos de A y B correspondientemente. Rojo, selección del máximo; verde, selección del cálculo del saliente y medida de la paridad y azul, selección del máximo y verificación de la consistencia.

En las Figuras 6.6 A y B es difícil apreciar las diferencias entre cada uno de los métodos pero al hacer un acercamiento, Figura 6.6 C y D se nota un desfase entre el método de selección del máximo (rojo) y los otros dos.

Ahora, para seleccionar cual de los dos métodos restantes sería útil, ya que ambos muestran un perfil muy similar para la fusión de las imágenes, se optó por seleccionarlo en cuanto a su

tiempo de cómputo empleado tomando el método de selección del máximo y verificación de consistencia, ya que éste método emplea un tiempo de cálculo menor (en las imágenes de la figura 3.17 son del orden de centésimas de segundo).

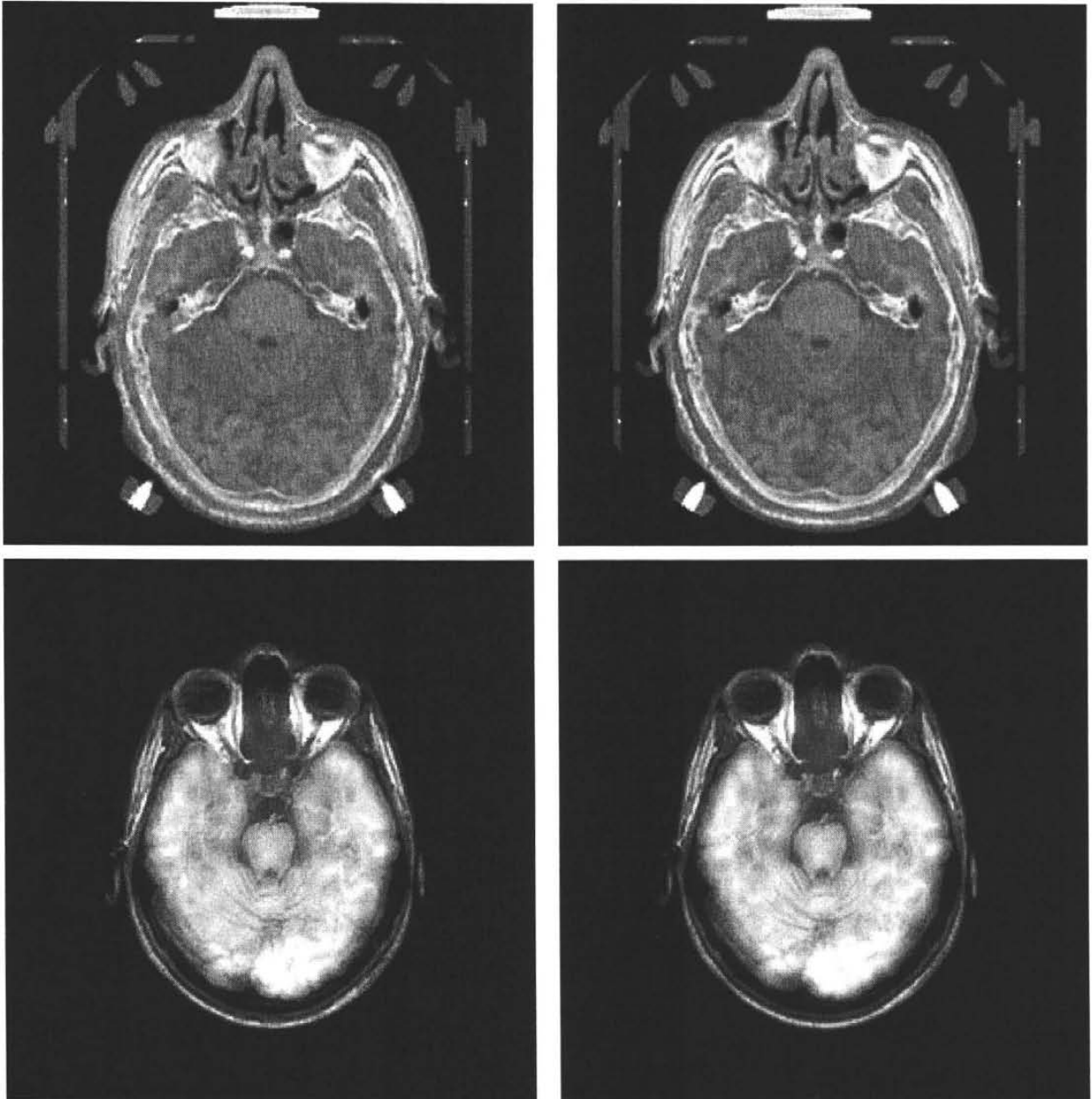
6.4 Resultados obtenidos de la fusión.

Las imágenes que se fusionaron pertenecen a una secuencia de imágenes obtenidas por métodos tomográficos:

1. Tomografía computarizada (Figura 3.14 A) con un tamaño de 640 X 640 pixeles,
2. Resonancia magnética (Figura 3.14 B), que previamente ha sido registrada con la tomografía computarizada (Figura 3.14 A) con un tamaño de 640 X 640 pixeles,
3. Tomografía por emisión de positrones (Figura 3.14 C), que al igual ha sido registrada con la imagen de resonancia magnética (Figura 3.14 D) y un tamaño de 512 X 512 pixeles, y
4. Resonancia magnética (Figura 3.14 D) con un tamaño de 512 X 512 pixeles.

Las imágenes fusionadas fueron 1 (tomografía computarizada) con 2 (resonancia magnética) y 3 (tomografía por emisión de positrones) con 4 (resonancia magnética), empleando para la selección de los coeficientes de baja frecuencia el promedio de las imágenes a fusionar y para la selección de los coeficientes de alta frecuencia la selección del máximo y verificación de consistencia, por lo dicho en las secciones 6.2 y 6.3.

En las imágenes de la Figura 6.7 se muestran los resultados de fusionar por medio de la transformada de Hermite Rotada (en adelante la llamaremos transformada de Hermite únicamente pero se refiere al caso de coeficientes rotados) y la transformada Wavelet Haar ya que esta última esta implementada en el algoritmo de fusión [16, 17, 22].

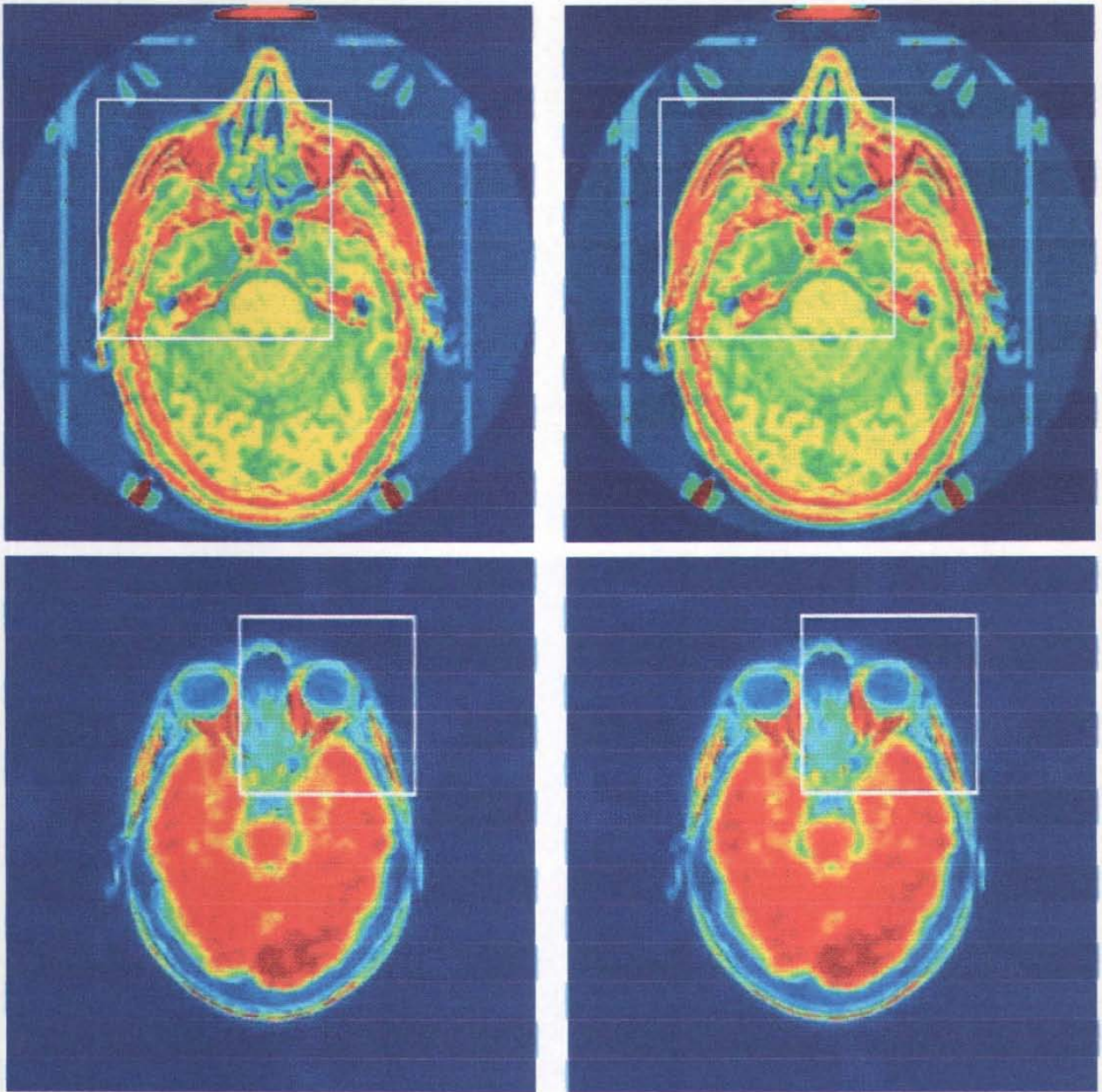


A	B
C	D

Figura 6.7: Fusión por medio de transformadas sobre las imágenes de entrada de la Figura 3.17. **A.** y **C.** Transformada de Hermite. **B.** y **D.** Transformada Wavelet Haar.

Las imágenes vistas (Figura 6.7) se desplegaron en niveles de gris. Para poder hacer una mejor discriminación sobre los efectos de las transformadas empleadas, la Figura 6.8

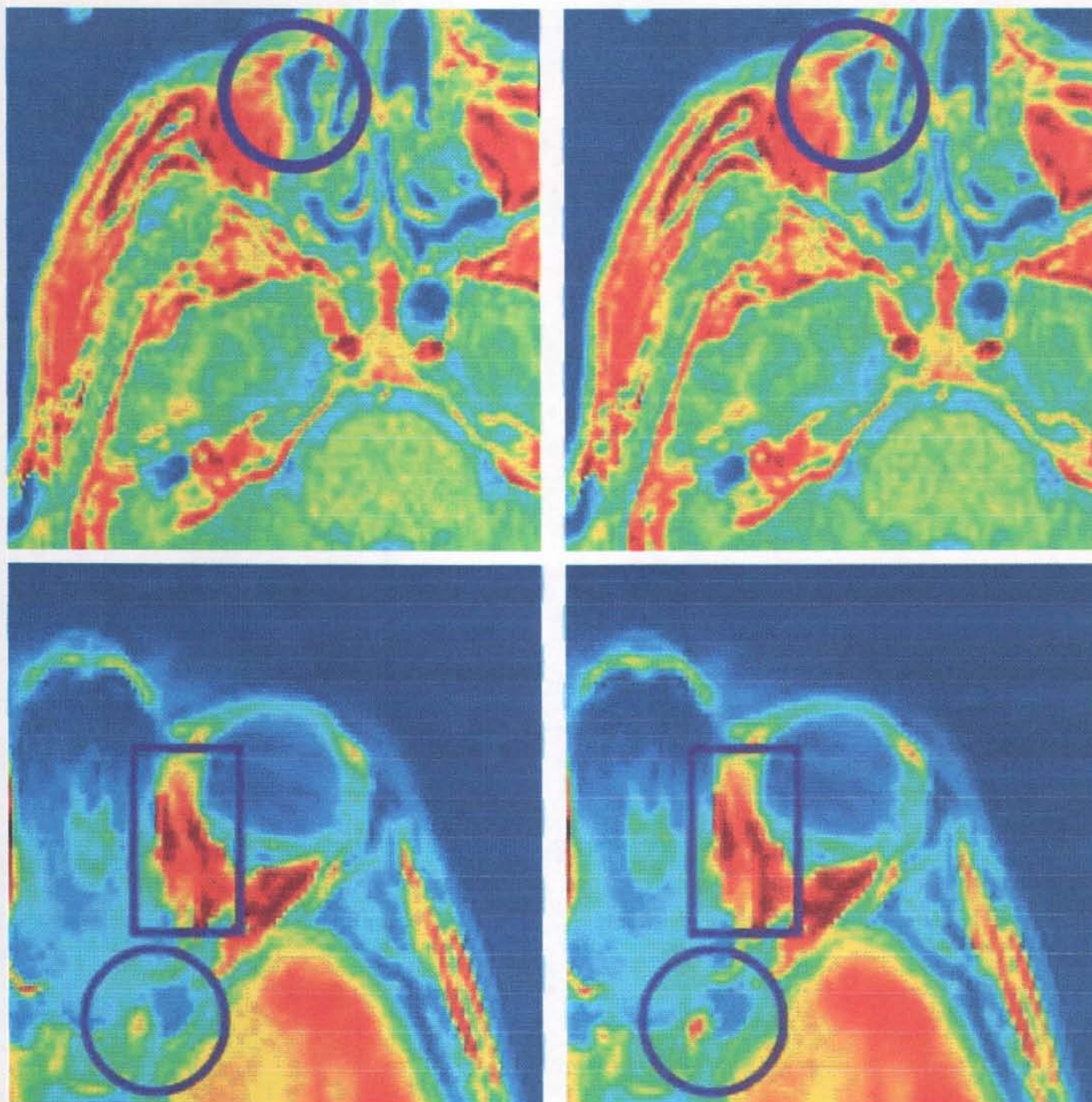
despliega estas mismas imágenes pero empleando el modelo de color HSV. (Matiz (*Hue*), Saturación (*Saturation*), Valor (*Value*)).



A	B
C	D

Figura 6.8: Fusiones visualizadas con el espacio de color HSV sobre las imágenes de entrada de la Figura 3.17. Transformada de Hermite (**A** y **C**). Transformada Wavelet Haar (**B** y **D**).

Para acentuar algunas diferencias sobre estas últimas imágenes hacemos un acercamiento de las mismas (recuadro) y marcando una región de interés con un círculo (Figura 6.9).

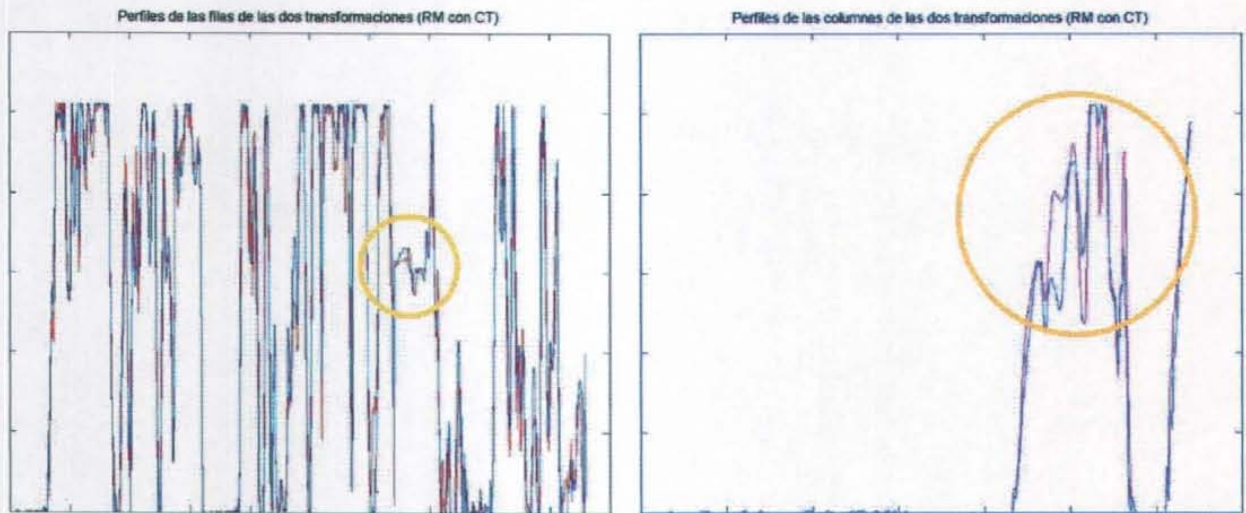


A	B
C	D

Figura 6.9: Acercamientos de la Figura 6.6. Transformada de Hermite (A y C) y Transformada Wavelet Haar (B y D).

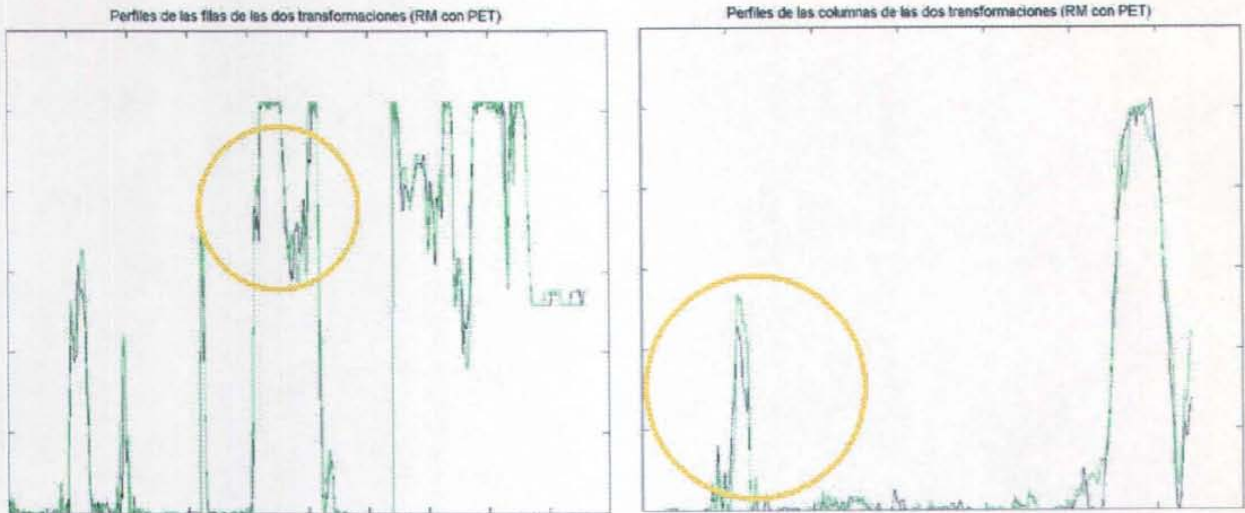
En la Figura 6.9 A y B se puede visualizar cómo la transformada Wavelet Haar difumina más la región marcada con el círculo en contraposición con la transformada de Hermite en la misma región. En la Figura 6.10, C y D marcadas con el rectángulo, mantienen la misma difuminación para la transformada Wavelet Haar, no así en la región marcada con el círculo cuya apariencia manifiesta una difuminación mayor hecha por la transformada de Hermite.

Para tener una apreciación más cuantitativa las Figuras 6.10 y 6.11 muestran los perfiles de los puntos de interés donde se observan los cambios en los niveles de intensidad de algunas líneas y columnas en las regiones marcadas en la Figura 6.9.



A **B**

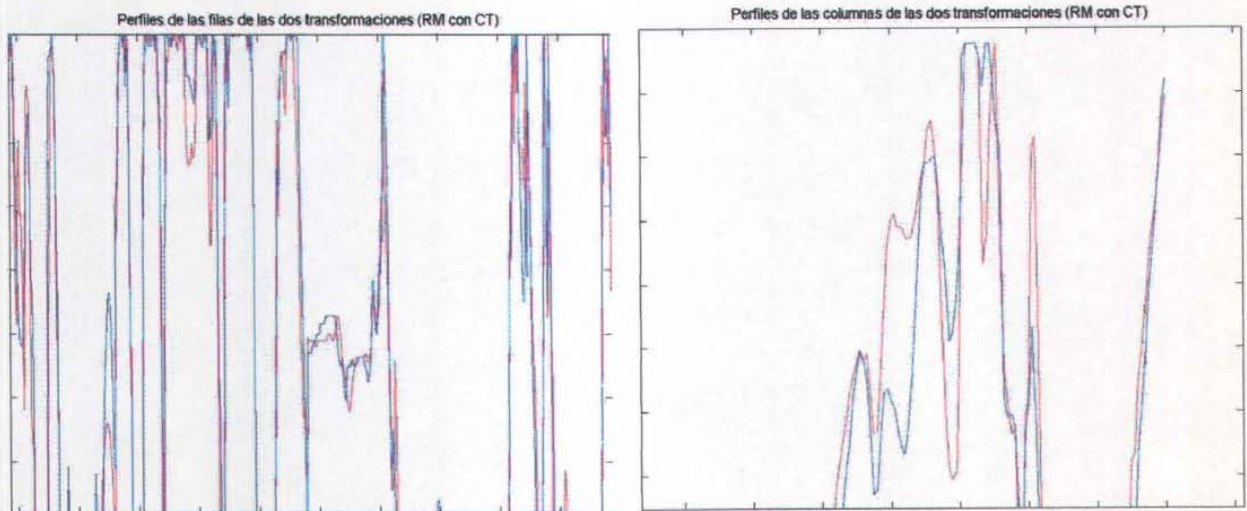
Figura 6.10: Perfiles (líneas y columnas) de las fusiones de la Figura 6.9 A y B. **A.** Líneas de las imágenes fusionadas con transformada de Hermite (roja) y con la transformada de Wavelet Haar (azul). **B.** Columnas de las imágenes fusionada con transformada de Hermite (roja) y con la transformada de Wavelet Haar (azul).



A B

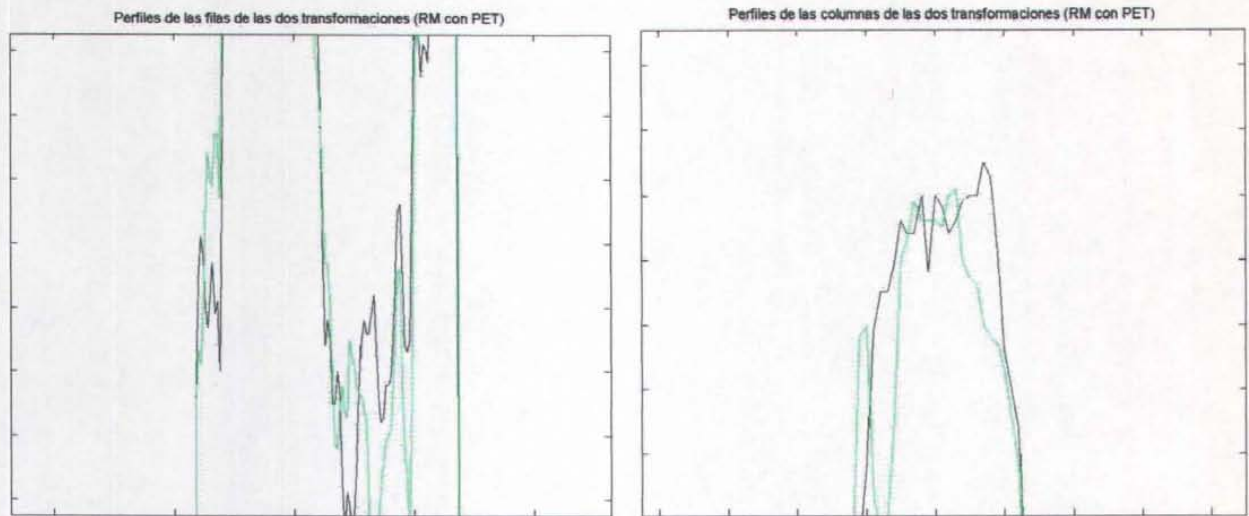
Figura 6.11: Perfiles (líneas y columnas) de las fusiones de la Figura 6.9 C y D. **A.** Líneas de las imágenes fusionada con transformada de Hermite (negra) y con la transformada de Wavelet Haar (verde). **B.** Columnas de las imágenes fusionadas con transformada de Hermite (negra) y con la transformada de Wavelet Haar (verde).

Para poder notar algunas diferencias acercamos las imágenes de la Figura 6.10 y 6.11 en las regiones marcadas por los círculos y se muestran en la Figura 6.12 y 6.13 respectivamente.



A B

Figura 6.12: Perfiles aumentados (filas y columnas) de la Figura 6.10. **A.** Líneas, transformada de Hermite (roja) y transformada de Wavelet Haar (azul). **B.** Columnas, transformada de Hermite (roja) y transformada de Wavelet Haar (azul).



A B

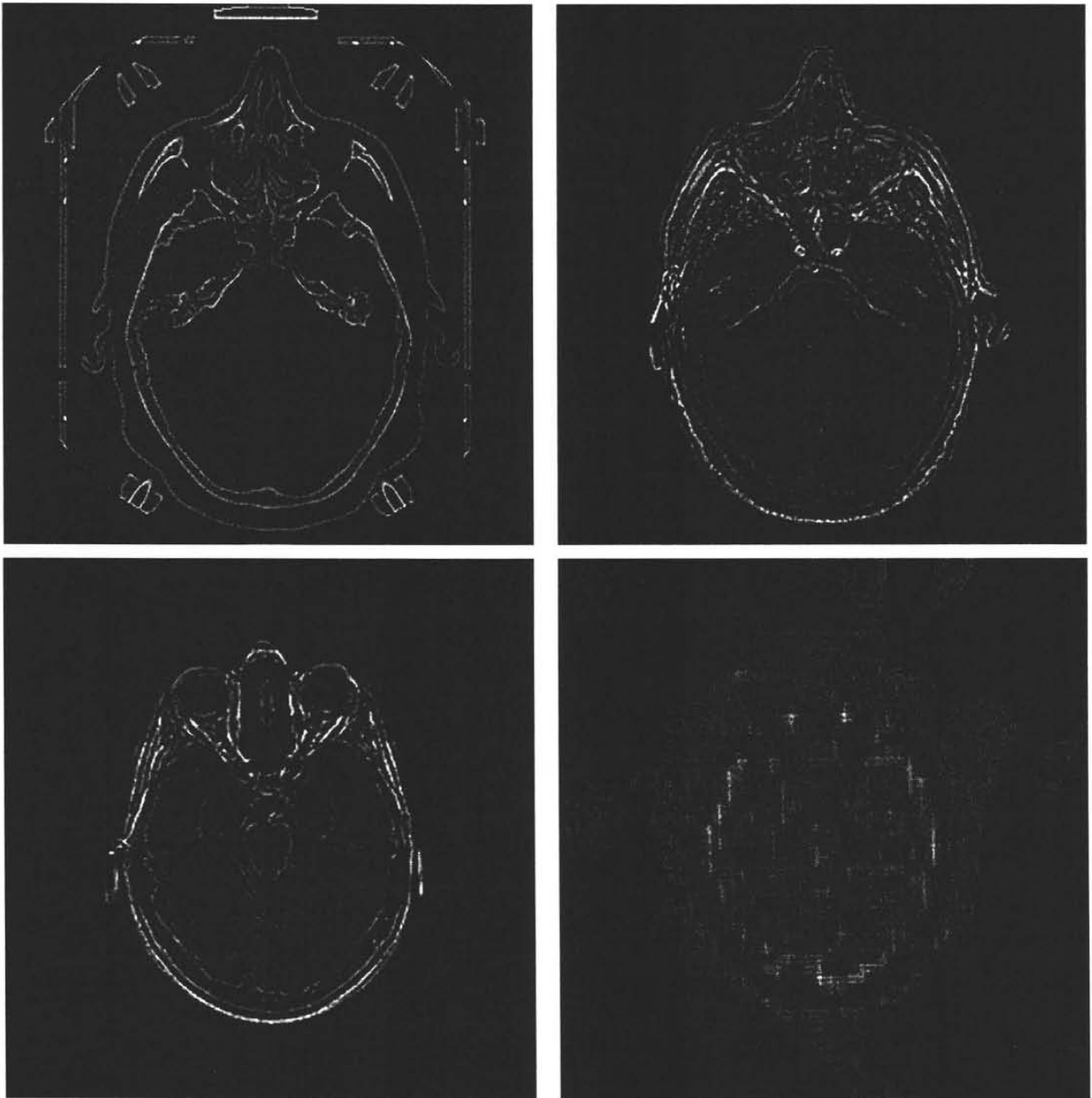
Figura 6.13: Perfiles aumentados (filas y columnas) de las fusiones de la Figura 6.11. **A.** Líneas, transformada de Hermite (negra) y transformada de Wavelet Haar (verde). **B.** Columnas, transformada de Hermite (negra) transformada de Wavelet Haar (verde).

En estas Figuras se pueden apreciar algunas diferencias sobre la transición de los niveles de intensidad de ambas transformadas. En la mayoría de los puntos la transformada de Hermite mantiene menos saturados los niveles de intensidad como consecuencia la transformada Wavelet Haar satura más los contornos de las imágenes (Figura 6.9).

La detección de bordes es parte fundamental en la mayoría de los sistemas de visión, la transformada de Hermite nos ofrece mejores resultados en la detección de bordes o contornos, para poder ver estos efectos sobre detección de contornos se aplicó un filtro pasa altas con un núcleo de la siguiente forma:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

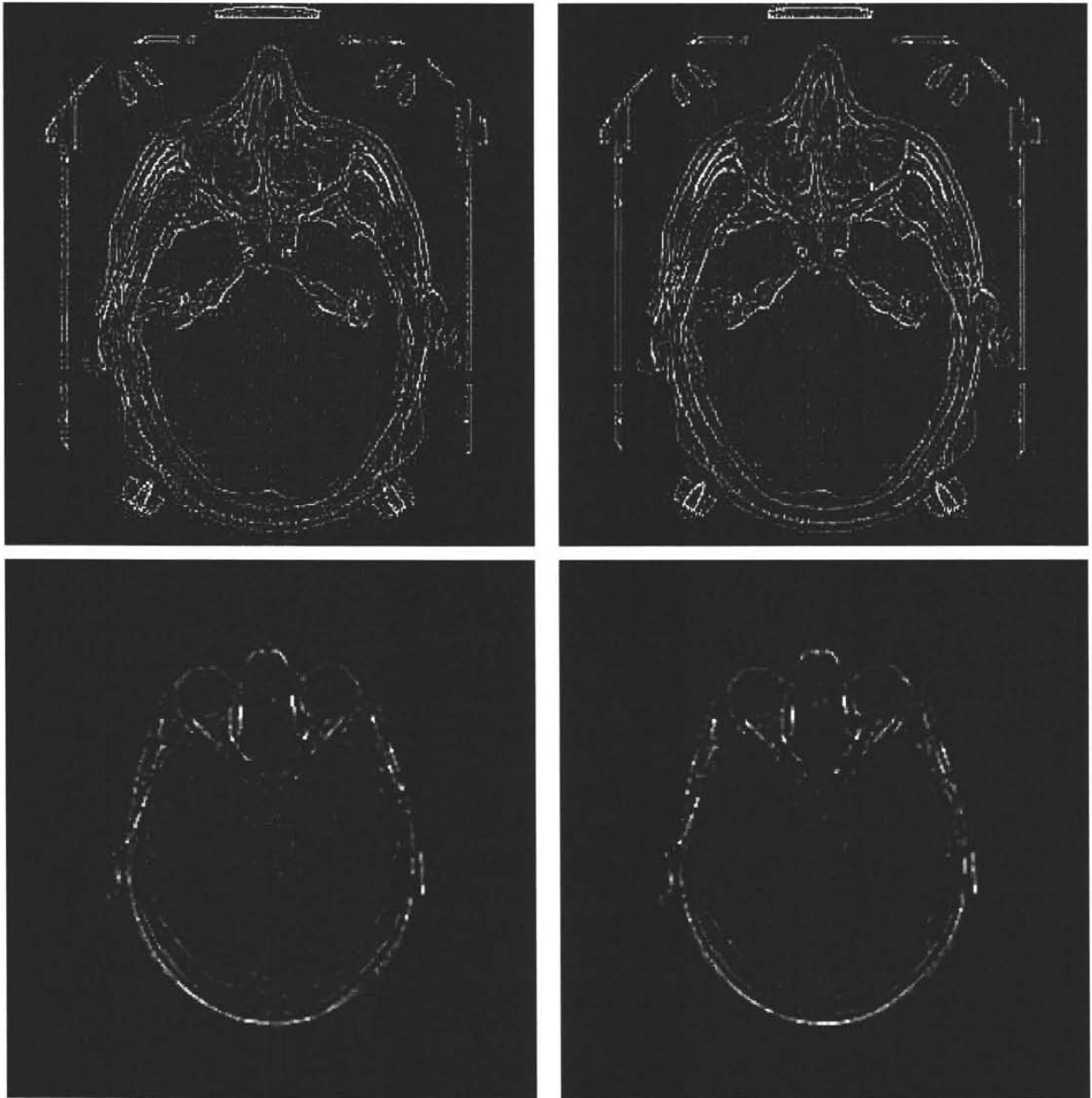
Con el propósito de resaltar y así evaluar los contornos derivados mediante las transformadas se aplicó el filtro pasa alta para las imágenes de entrada (Figura 3.14) para así tener un patrón de comparación, los resultados de dicho filtrado se muestran en la Figura 6.14.



A	B
C	D

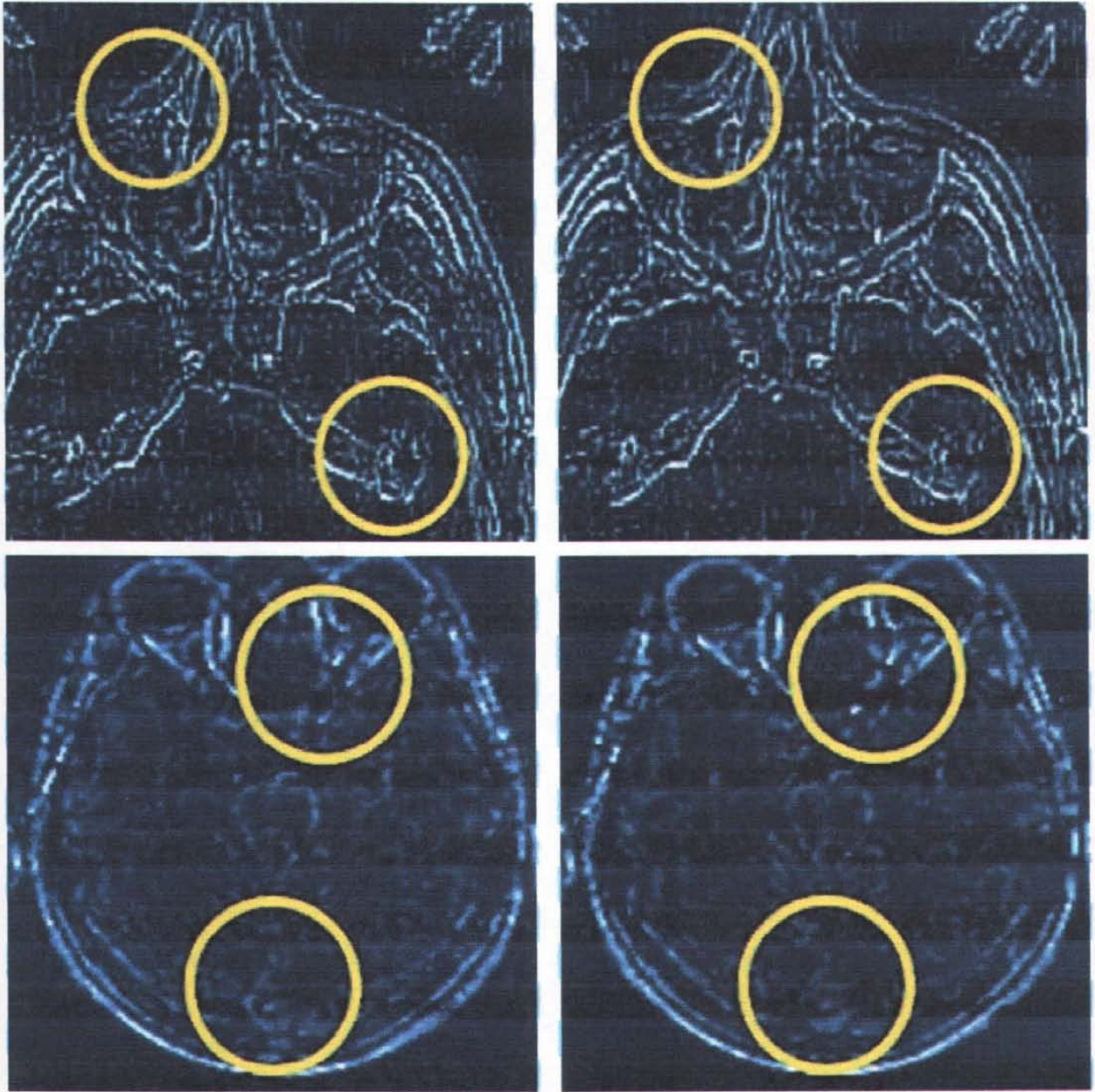
Figura 6.14: Detección de bordes utilizando un filtro pasa alta de las imágenes de entrada de la Figura 3.14. **A.** Tomografía computarizada **B.** Resonancia magnética registrada con la imagen de tomografía computarizada **C.** Tomografía por emisión de positrones registrada con la imagen de resonancia magnética. **D.** Resonancia magnética.

En la Figura 6.15 observamos los resultados de aplicar el mismo filtro pasa altas a las imágenes fusionadas (Figura 6.7) y se hace un acercamiento para notar algunas diferencias entre ambas transformadas (Figura 6.16).



A	B
C	D

Figura 6.15: Filtrado sobre las imágenes fusionadas de la Figura 6.5. **A.** y **C.** Transformada de Hermite. **B.** y **D.** Transformada Wavelet Haar.

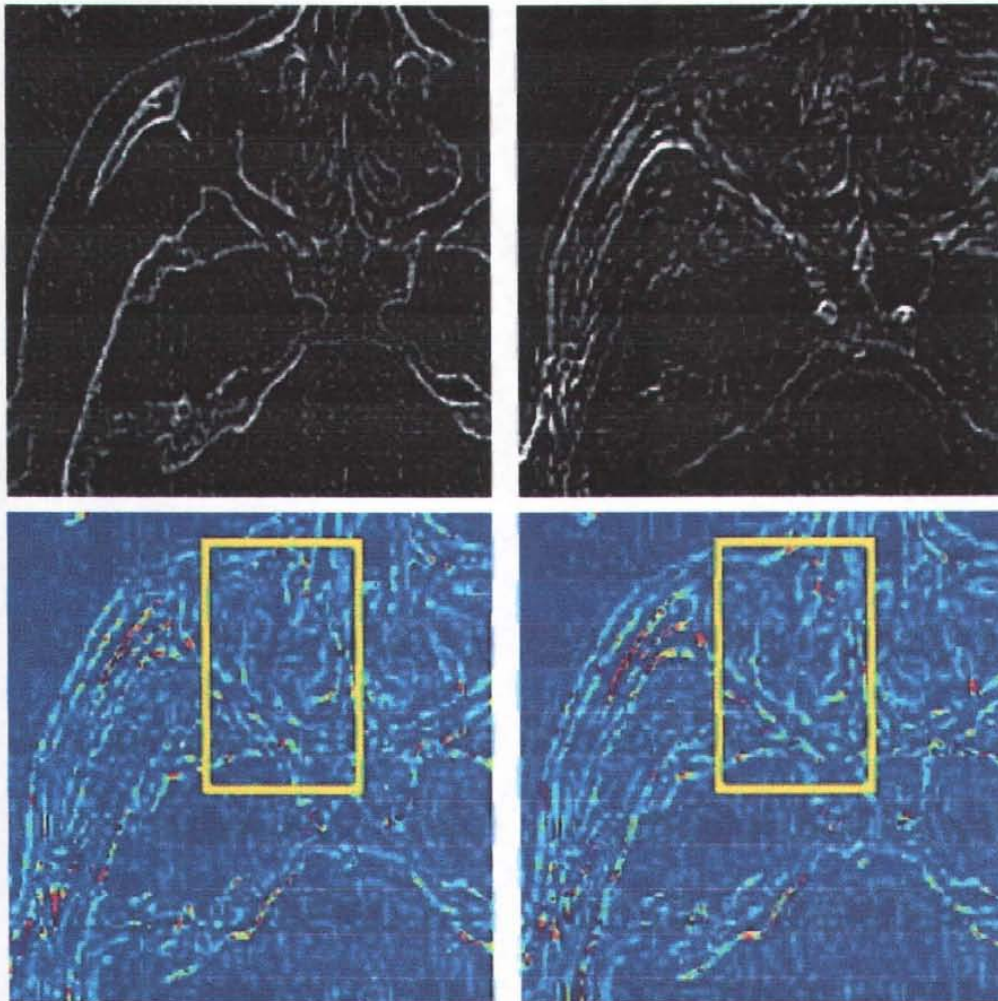


A	B
C	D

Figura 6.16: Acercamientos de la Figura 6.15. Transformada de Hermite (**A** y **C**) y Transformada Wavelet Haar (**B** y **D**).

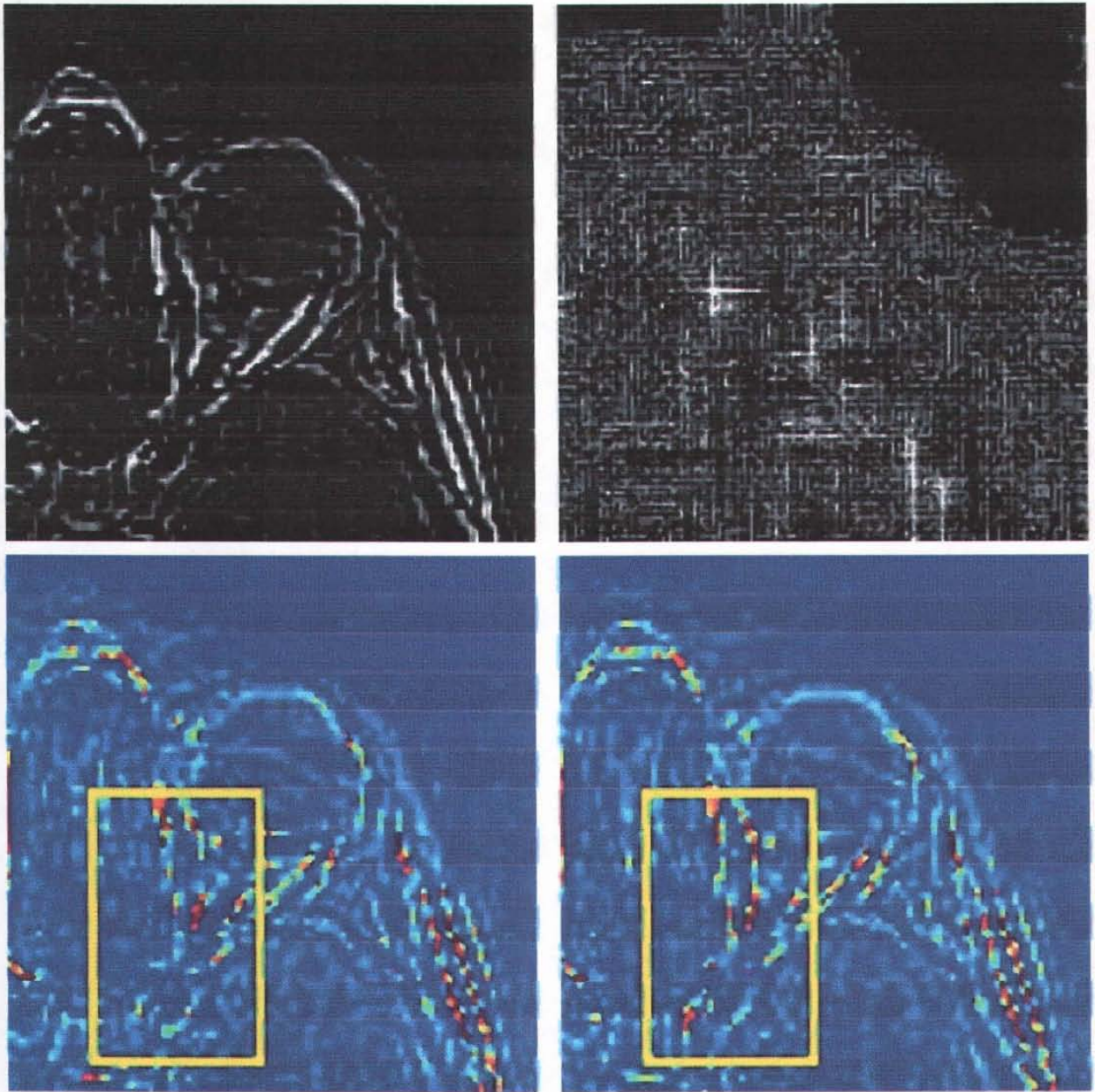
En las imágenes de la Figura 6.16 en las regiones marcadas por los círculos se observan mejor definidos los contornos a través de la transformada de Hermite que con la transformada Wavelet Haar

En las Figuras (6.17 y 6.18) se presentan las imágenes fusiones que han sido filtradas mostradas bajo otro modelo de color y acercadas (C y D) con el propósito de apreciar mas detalle se compararon con las imágenes fusionadas (A y B).



A	B
C	D

Figura 6.17: Comparación entre las imágenes fusionadas (RM con CT) y las fusionadas filtradas. **A.** Tomografía computarizada **B.** Resonancia magnética, **C.** Imagen fusionada y filtrada (pasa alta) por medio de la transformada de Hermite y **D.** Imagen fusionada y filtrada (pasa alta) por medio de la transformada Wavelet Haar.



A	B
C	D

Figura 6.18: Comparación entre las imágenes fusionadas (RM con PET) y las fusionadas filtradas. **A.** Resonancia magnética **B.** Tomografía por emisión de positrones, **C.** Imagen fusionada y filtrada (pasa alta) por medio de la transformada de Hermite y **D.** Imagen fusionada y filtrada (pasa alta) por medio de la transformada Wavelet Haar.

En estas últimas Figuras se observa que la transformada de Hermite preserva mejor los contornos para el caso de fusionar una imagen de tomografía computarizada con una de resonancia magnética (Figura 6.17 C y D) que utilizando la transformada wavelet Haar y en la Figura 6.18 se preservan mejor los bordes utilizando la transformada wavelet Haar después de haber fusionado imágenes de resonancia magnética con tomografía por emisión de positrones.

Del capítulo 3 sección 3.3; donde se analizaron los diferentes métodos de fusión más representativos, se deriva la conclusión que el mejor método a emplear depende del tipo de imágenes a fusionar, sin embargo se puede decir que en términos generales los métodos de fusión a través de transformadas nos dan buenos resultados, pues como se vio, la fusión utilizando un tipo de transformada nos ayuda a reconstruir "perfectamente" la imagen de entrada, es por ello que la fusión por medió de la transformada de Hermite se comparó con la transformada wavelet Haar.

CONCLUSIONES

La fusión de imágenes médicas, como se mencionó en nuestro caso, es el proceso de combinar la información procedente de distintas modalidades (fuentes) para producir una única imagen final. Por una parte permite presentar de manera conjunta e integrada toda la información anatómica y funcional proporcionada por diferentes modalidades de imagen; por otra parte, puede facilitar la obtención de datos cuantitativos a partir de informaciones complementarias.

Mediante este trabajo de investigación se comprobó por un lado, la posibilidad de aplicar la transformada de Hermite rotada a imágenes médicas (tomografía computarizada, resonancia magnética y tomografía por emisión de positrones para nuestra cuestión), con el propósito de obtener una mejor delineación sobre los contornos. Esta herramienta es importante en el ámbito del procesamiento digital de imágenes pues busca aumentar la información existente en la imagen y facilitar el análisis que en el caso de imágenes de médicas éste resulta particularmente útil en los estudios que utilizan técnicas como la Detección de Bordes, Segmentación y Clasificación.

La transformada de Hermite se emplea en dichos procesos de señales por la forma en que se modela, tomando en cuenta características fundamentales con relación a las etapas primarias de procesamiento del Sistema Visual Humano (filtros Gaussianos) y su descomposición (rotada) lleva a cabo un análisis local de los datos para extraer información adicional sobre clasificación de patrones, lo que simplifica considerablemente el proceso de reconstrucción. Acorde con Martens [20] los bordes representan la mayoría de las regiones no constantes en imágenes y su importancia ha sido reconocida ampliamente en visión computacional, donde muchos algoritmos han sido diseñados específicamente para detectar y localizar estas estructuras. Hacer uso de esta transformada en este trabajo resultó de gran importancia ya que al hacer la descomposición de cada una de las imágenes, en estas se resaltan las

características más importantes, ya que en el proceso de selección se toman estas características y el resultado contiene la mayor información de ambas imágenes a fusionar.

Entre las diferentes técnicas de selección de coeficientes y en el caso de los de baja frecuencia, va a depender de la información que se necesite resaltar como resultado final. En nuestro caso se requiere resaltar la información de ambas imágenes, por lo que para lograr una fusión con estas características tomamos el promedio de ambos coeficientes.

Para la selección de los coeficientes de alta frecuencia, los métodos empleados nos arrojan buenos resultados pero para este trabajo se seleccionó el de selección del máximo y verificación de la consistencia, ya que es un método que toma en cuenta los valores de los vecinos cercanos y es de rápido cálculo computacional.

Los resultados obtenidos a través de la fusión con la transformada de Hermite fueron comparados con la técnica de transformación wavelets Haar y podemos concluir que la transformada de Hermite (rotada) como se vio en el capítulo 6 sección 6.4 tiene ventajas sobre esta.

Como trabajos futuros se encuentran el de diseñar nuevos esquemas de selección de coeficientes orientados a objetos y no solo a energías y magnitudes como lo hacen los métodos actuales como también la eliminación del ruido contenido en las imágenes médicas antes de hacer fusión.

REFERENCIAS

- [1] M. Desco, "Técnicas avanzadas de imagen en Medicina", Medicina experimental, Hospital general universitario G. Marañón, pp. 1-14.
- [2] R. Medina y J. Bellera, "Bases del procesamiento de imágenes médicas", Universidad de los Andes, Facultad de Ingeniería, Grupo de Ingeniería Biomédica de la ULA (GIBULA), Venezuela, pp. 26-32.
- [3] Telecom Salud, Home site, disponible en: <http://www.telecomsalud.com/modalidades.htm>
- [4] Lucien Wald, "Some Terms of Reference in Data Fusion", IEEE Transactions on Geoscience and Remote Sensing, Vol. 37, No. 3, pp. 1190-1193, May 1999.
- [5] C. Pohl y J. L. Van Genderen, "Multisensor image fusion in remote sensing: Concepts, methods and applications," *Int. J. Remote Sensing*, vol. 19, no. 5, pp. 823-854, 1998.
- [6] M. Mangolini, "Apport de la fusion d'images satellitaires multicateurs au niveau pixel en télédétection et photo-interprétation," Thèse de doctorat, Univ. Nice-Sophia Antipolis, France, p. 174, 1994.
- [7] D. L. Hall y J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, pp. 6-23, Jan. 1997.
- [8] U.S. Department of Defence, Data Fusion Subpanel of the Joint Directors of Laboratories, Technical Panel for C3, "Data fusion lexicon," 1991.
- [9] H. Li, B. S. Manjunath, y S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Comput. Vis., Graph., Image Process.: Graph. Models Image Process.*, vol. 57, pp. 235-245, 1993.
- [10] G. Brandstätter y A. Sharov, "Environmental monitoring in the high Arctic using different types of high-resolution satellite imagery," *Int. Arch. Photogramm. Remote Sensing*, pt. 7, vol. 32, pp. 201-216, 1998.
- [11] B. Csathó y T. Schenk, "Multisensor data fusion for automatic scene interpretation," *Int. Arch. Photogramm. Remote Sensing*, pt. 7, vol. 32, pp. 336-341, 1998.

- [12] M. Buchroithner, "Geodata interrelations: Inventory and structuring attempt of taxonomic diversity," in *Proc. 2nd Conf. Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*, pp. 11-15, 1998.
- [13] L. Wald, "Data fusion: A conceptual approach for an efficient exploitation of remote sensing images," in *Proc. 2nd Conf. Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*, pp. 17-23, 1998.
- [14] I. Bloch, "Information combination operators for data fusion: A comparative review whit classification", *System, man and cybernetics, Part A, IEEE. Transaction on*, Vol. 26, ISSUE: 1, pp. 52-67, January 1996.
- [15] Tapiador, F.J. y Casanova J.L., "La fusión de datos en teledetección", revista de teledetección, 2001, No. 15, junio 2001, pp 1-6. Disponible en: telenet.uva.es/promotores/revista/revista_15/AET15_1.pdf.
- [16] O. Rockinger, "Pixel level fusion of image sequences using wavelet frames", in: *Proc. 16th Leeds Annual Statistical Research Workshop*, Leeds University Press, 1996, pp. 149-154.
- [17] O. Rockinger, "Image sequence fusion using a shift invariant wavelet transform", in: *Proc. IEEE Intl. Conference on Image Processing*, 1997, pp. III-288-291.
- [18] A. Lopez, A. A., "Fusión de imágenes satelitales de la Ciudad de México mediante métodos multiresolución", Tesis de Maestría en Ingeniería Eléctrica, Universidad Nacional Autónoma de México, División de Estudios de Posgrado de la Facultad de Ingeniería, capítulos 3 y 4, 2003.
- [19] J. B. Martens, "The Hermite Transform-Theory", *IEEE transactions on Acoustics, Speech and Signal Processing*; vol. 38 no. 9, pp 1595-1606, 1990.
- [20] J. B. Martens, "The Hermite Transform-Applications", *IEEE transactions on Acoustics, Speech and Signal Processing*; vol. 38 no. 9, pp 1607-1618, 1990.
- [21] J. Silván, "Compresión de imágenes basada en modelos gaussianos de percepción visual", Tesis de Maestría en Ingeniería Eléctrica, Universidad Nacional Autónoma de México, División de Estudios de Posgrado de la Facultad de Ingeniería, capítulos 2 y 3, 2002.
- [22] C. Zambrano, "Fusión de imágenes de radar aerotransportado AeS-1 y LANSAT TM5, mediante la Transformada Polinomial de Hermite", Tesis de Maestría en Ingeniería Eléctrica, Universidad Nacional Autónoma de México, División de Estudios de Posgrado de la Facultad de Ingeniería, capítulos 2 y 3, 2004.

- [23] R. C. Gonzalez y R. E. Woods, "Digital image processing", Prentice Hall, second edition, 2002. Páginas 567-585.
- [24] T. Baidyk, E. Kussul, O. Makeyev, A. Caballero, L. Ruiz, G. Carrera y G. Velasco, "Flat image recognition in the process of microdevice assembly", in Elsevier B. V., Pattern Recognition Letters 25, pp. 107-118, 2004.
- [25] J. Daugman, "Six formal properties of two-dimensional anisotropic visual filters: Structural principles and frequency/orientation selectivity", IEEE Trans. Syst., Man, Cybern., vol. 13, pp. 882-887, sept./Oct. 1993.
- [26] G. Szegő, Orthogonal Polynomials (American Mathematical Society). Colloquium Publication, 1959.
- [27] Interfile -Grupo Imagen UMCE HGGM- Generado por el programa de Jesús, adquiridas el: 24/09/2002 (CT y RM_reg_CT), 19/09/2002 (PET_reg_RM) y 06/06/2001 (RM).