

UNIVERSIDAD POLITÉCNICA DE CARTAGENA

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA DE TELECOMUNICACIÓN



PROYECTO FIN DE CARRERA

Ingeniero de Telecomunicación

APLICACIÓN DE MEDIDAS ESTADÍSTICAS DE SIMILITUD AL REGISTRO DE IMAGEN MULTIMODO

Autor: Luis Eduardo García Capel

Director: Juan Morales Sánchez

Departamento: Departamento de Tecnologías
de la Información y las
Comunicaciones

Fecha: febrero de 2007

AGRADECIMIENTOS

Quisiera agradecer en primer lugar a toda mi familia el haberme ofrecido el apoyo necesario para darme la posibilidad de estudiar lo que realmente deseaba. Sin ellos nada de esto habría sido posible.

Mención especial merecen los amigos y compañeros de clase que han estado conmigo durante todos estos años. De una u otra forma han hecho posible que el tiempo se haya hecho mucho más ameno y que los momentos difíciles hayan dejado de serlo.

También me gustaría expresar mi agradecimiento al director de este proyecto, Juan Morales Sánchez, por haberme brindado la oportunidad de realizar este proyecto, por haber tenido siempre la paciencia necesaria conmigo, y por compartir su conocimiento, experiencia y dedicación para hacer realidad este trabajo.

RESUMEN

Este proyecto final de carrera tiene como objetivo analizar el comportamiento de varias medidas estadísticas de similitud cuando éstas se aplican al campo del registro de imágenes de distintos tipos. El resultado de estos procesos de registro depende de la calidad de las imágenes, el nivel de gris, la naturaleza de las mismas y el contenido presente en ellas. Hay multitud de medidas que han sido propuestas para el registro de imagen, pero no hay una sola medida que produzca siempre un mejor resultado que otras. Las cuatro medidas de similitud que se presentan en este proyecto son cuatro: el coeficiente normalizado de correlación cruzada (NCC), la información mutua (MI), el ratio de correlación (CR) y el criterio de Woods (WC). Algunos de estos métodos se derivan de otros añadiendo diversas variaciones, que tienen como objetivo mejorar el resultado del registro bajo unas ciertas condiciones y para un tipo específico de imágenes.

El algoritmo de registro utilizado se divide en varias etapas claramente identificables de forma funcional. Estas etapas resultan esenciales para el buen funcionamiento del sistema como conjunto y pueden implementarse de muy distintas maneras, aunque éste no es el objetivo fundamental del trabajo aquí presente. Por lo tanto, este proyecto se basa en el cálculo de las medidas de similitud de los niveles de gris de las imágenes referencia y objetivo dadas, para posteriormente utilizarlas en la detección del movimiento o deformación que se haya llevado a cabo entre las dos fracciones de imágenes seleccionadas en cada iteración del proceso. Para cada una de estas iteraciones se extrae una porción de la imagen referencia y se aplica una de las medidas de similitud en una zona amplia de la imagen objetivo para crear un mapa de similitud. El conjunto de mapas calculados nos permite determinar el lugar hacia donde se ha deformado cada parte de la imagen. Este mapa de deformaciones, traducido en un conjunto de vectores de desplazamiento, debe ser continuo y suave para garantizar la coherencia, por lo que el resultado del registro dependerá del conjunto de desplazamientos individuales de los conjuntos de píxeles de las imágenes comparadas durante todo el proceso.

Para la consecución de un resultado lo más fiel posible a su realidad, en el proceso de registro se hace uso de la diferencia cuadrática media normalizada. Esta medida nos facilita la tarea de localización de bloques y nos facilita la tarea de realizar un ajuste paramétrico en 2D, con el que se demuestran resultados mucho mejores que con la técnica de *Block-matching*, que únicamente utiliza modelos de correspondencia *punto* frente a los tres que utiliza la técnica anterior: *punto*, *curva* y *punto+curva*. Para contrastar estos hechos, se realiza una comparativa entre ambos métodos tanto en el cálculo de los mapas de similitud como para el registro de una

imagen completa, en la que se observa la ineficacia del algoritmo de *Block-matching* para la obtención de buenos resultados de registro.

Con el objetivo de analizar claramente los resultados de estas cuatro medidas, también se muestran los resultados obtenidos para registros reales de varios pares de imágenes. Dichas imágenes empleadas son de igual y distinta naturaleza, y nos proporcionan un sentido amplio de los campos de acción de las diversas medidas. Estos resultados demuestran la extensa aplicación de los métodos analizados, así como nos permiten apreciar algunos aspectos y consecuencias concretas de estas determinadas técnicas, como son por ejemplo el tiempo de procesamiento, la sensibilidad frente a distintos tipos de imágenes, la suavidad del registro o las limitaciones existentes.

Para la cuantificación del grado de similitud entre la imagen referencia y la imagen registrada a partir de la imagen objetivo, se hace uso de una medida simple como es el *Peak Signal Noise Ratio* (PSNR). Ésta, aun no siendo una medida muy elaborada, nos proporciona mejor efecto que otras medidas más complicadas como la información mutua, que no produce un resultado coherente con la realidad del resultado para imágenes monomodo. Para esta tarea de determinación de la lealtad del resultado obtenido con el registro, también se calcula el error cuadrático medio del mapa de vectores de desplazamiento (o deformación) calculado frente al real, es decir, el que en realidad se ha aplicado sobre la imagen referencia para obtener la imagen objetivo. Esta medida sigue en todo momento al PSNR, complementándose perfectamente. Asimismo, los mapas de desplazamiento son analizados por su rigidez y elasticidad, o sea, mediante el cálculo de las derivadas primera y segunda de los vectores calculados, lo que nos permite apreciar las diferencias producidas en los diferentes registros con las distintas medidas de similitud.

Finalmente, tras toda esta serie de medidas, comparativas y conclusiones, los resultados ilustran que la conveniencia de una u otra medida dependen de la naturaleza de las imágenes u otros requisitos y limitaciones que se nos planteen. El tiempo de procesamiento será un aspecto muy a tener en cuenta en algunos casos, mientras que en otros tendrá más importancia el hecho de que dos imágenes sean de un determinado tipo. De este modo se puede comprobar la alta versatilidad de estas medidas y su grado de fiabilidad bajo ciertas condiciones.

ÍNDICE

1. INTRODUCCIÓN.....	1
1.1. PLANTEAMIENTO INICIAL DEL PROYECTO	1
1.2. OBJETIVOS Y MOTIVACIÓN DE ESTE TRABAJO	1
1.3. PUNTO DE PARTIDA	2
1.4. DESCRIPCIÓN DEL RESTO DE CAPÍTULOS.....	2
2. CONOCIMIENTOS PREVIOS	5
2.1. INTRODUCCIÓN	5
2.2. EL REGISTRO DE IMAGEN.....	5
2.2.1. Naturaleza de las imágenes.....	6
2.2.2. Tipos de transformaciones.....	7
2.2.2.1 Transformaciones paramétricas.....	8
2.2.2.2 Transformaciones no paramétricas	9
2.3. CLASIFICACIÓN DE LOS ALGORITMOS DE REGISTRO DE IMÁGENES.....	9
2.3.1. Emparejamiento geométrico.....	10
2.3.2. Registro basado en intensidades.....	11
2.3.2.1 Registro basado en intensidad estándar (SIB).....	11
2.3.2.2 Registro de emparejamiento y suavizado (P&S)	12
2.4. MEDIDAS DE SIMILITUD	12
2.4.1. Medidas basadas en el nivel de intensidad.....	13
2.4.1.1 Medidas basadas en promedios	14
2.4.1.2 Medidas basadas en la estadística de la imagen	14
3. MEDIDAS DE SIMILITUD	15
3.1. INTRODUCCIÓN	15
3.2. MÁSCARAS DE PASO Y PONDERACIÓN.....	15
3.3. EL COEFICIENTE DE CORRELACIÓN CRUZADA.....	16
3.3.1. Expresión matemática del NCC	17
3.3.2. Algoritmo de implementación	20

3.3.3. Limitaciones.....	22
3.3.4. Ejemplos	23
3.4. EL COEFICIENTE DE INFORMACIÓN MUTUA	26
3.4.1. Expresión matemática del NMI.....	26
3.4.2. Algoritmo de implementación.....	29
3.4.3. Limitaciones.....	30
3.4.4. Ejemplos	30
3.5. EL RATIO DE CORRELACIÓN	33
3.5.1. Expresión matemática del CR.....	33
3.5.2. Algoritmo de implementación.....	35
3.5.3. Limitaciones.....	36
3.5.4. Ejemplos	36
3.6. EL CRITERIO DE WOODS	39
3.6.1. Expresión matemática del WC.....	39
3.6.2. Algoritmo de implementación.....	40
3.6.3. Limitaciones.....	41
3.6.4. Ejemplos	42
3.7. COMPARATIVA DE LOS MÉTODOS DE SIMILITUD EXAMINADOS.....	45
3.7.1. Influencia del contenido del bloque	45
3.7.2. Influencia del contraste, iluminación y offset.....	48
3.7.3. Influencia del tipo de imagen.....	51
3.7.4. Tiempo de procesamiento.....	52
4. DESCRIPCIÓN DEL PROCEDIMIENTO DE REGISTRO.....	57
4.1. INTRODUCCIÓN.....	57
4.2. DEFINICIONES INICIALES	57
4.3. CORRESPONDENCIA ESPACIAL BIDIMENSIONAL.....	58
4.4. ESTIMACIÓN DEL MOVIMIENTO EN 2D	59
4.4.1. Cálculo de la correspondencia paramétrica en 2D.....	60
4.4.1.1 Cálculo del coeficiente de similitud entre dos conjuntos de datos 2D	62
4.4.1.2 Ajuste del modelo paramétrico en 2D.....	64
4.4.2. Cálculo de los vectores de correspondencia en 2D.....	67
4.4.2.1 Filtrado de interpolación 2D	68

4.4.2.2	Aplicación de restricciones paramétricas en 2D	68
5.	RESULTADOS	71
5.1.	INTRODUCCIÓN	71
5.2.	EVALUACIÓN DEL RESULTADO DEL REGISTRO	71
5.2.1.	El error cuadrático medio (ECM)	71
5.2.2.	El PSNR (Peak Signal-to-Noise Ratio)	72
5.2.3.	Cálculo de la energía del VCD	72
5.3.	RESULTADOS OBTENIDOS	73
5.3.1.	Imágenes sintéticas	73
5.3.1.1	Caso 1. Bola con estrellas	73
5.3.1.2	Caso 2. Mosaico	75
5.3.1.3	Caso 3. Conjunto de figuras geométricas	77
5.3.2.	Imágenes fotográficas.....	78
5.3.2.1	Caso 1. Mujer de espaldas	78
5.3.2.2	Caso 2. Planta química	81
5.3.3.	Imágenes aéreas	83
5.3.3.1	Caso 1. Centro residencial	83
5.3.3.2	Caso 2. Costa de San Diego (EEUU)	85
5.3.4.	Imágenes médicas.....	88
5.3.4.1	Caso 1. Imagen MRI cerebral.....	88
5.3.4.2	Caso 2. Imagen CT del tórax.....	90
5.3.5.	Imágenes multimodo	93
5.3.5.1	Caso 1. Tórax CT-PET	93
5.3.5.2	Caso 2. Cerebro CT-PET	96
5.3.5.3	Caso 3. Imagen sagital CT-PET	99
5.3.5.4	Caso 4. Tórax CT-PET	101
5.4.	DISCUSIÓN DE LOS RESULTADOS	103
6.	CONCLUSIONES Y LÍNEAS FUTURAS	105
6.1.	CONCLUSIONES	105

6.1.1. Esencia del proyecto.....	105
6.1.2. Conocimientos adquiridos	105
6.1.3. Dificultades encontradas.....	105
6.1.4. Valoración	108
6.2. FUTURO.....	108
6.2.1. Posibles mejoras al procedimiento de registro.....	108
6.2.1.1 Mejora de las medidas de similitud.....	108
6.2.1.2 Reducción del tiempo de procesamiento.....	108
6.2.1.3 Portabilidad	109
6.2.2. El futuro: hacia dónde se dirige el registro de imágenes	109
7. BIBLIOGRAFÍA.....	111

1. INTRODUCCIÓN

1.1. PLANTEAMIENTO INICIAL DEL PROYECTO

El *registro de imágenes* es un tema de actualidad en multitud de campos, especialmente en medicina, donde la alineación de dos imágenes puede ser vital para el diagnóstico o la determinación de la evolución de una cierta enfermedad. Este hecho hace que el estudio del registro de imágenes resulte un tema muy interesante a tratar, siendo por tanto éste el centro del estudio en este trabajo, por medio del análisis de algunas de las distintas medidas de similitud con las que se cuenta hoy día para llevar a cabo dicho cometido. Este capítulo nos servirá para analizar el estado actual de la técnica y definir lo que serán los objetivos a conseguir con la elaboración de este proyecto.

1.2. OBJETIVOS Y MOTIVACIÓN DE ESTE TRABAJO

El campo del registro de imágenes es muy amplio. Las técnicas de registro se utilizan hoy día para múltiples aplicaciones, desde la visión por ordenador hasta el registro de imágenes médicas, pasando por campos tan incipientes como el de seguimiento de objetos. Sin embargo, el objetivo de este trabajo se concentra en analizar a través del uso de la herramienta de procesamiento Matlab un conjunto de técnicas de cálculo de similitud para su aplicación al registro de imágenes bidimensionales. Nos centraremos fundamentalmente en las imágenes de naturaleza multimodo, uno de los campos sobre los que más se está trabajando y que hoy en día tiene un mayor peso en los métodos de registro de imágenes médicas. Sin embargo, un análisis previo de imágenes de la misma modalidad (imágenes monomodo) nos permitirá analizar de manera correcta las distintas medidas de similitud empleadas, así como evaluar su funcionamiento bajo diferentes condiciones de operación.

Así pues, para la consecución de dicho objetivo fundamental será necesario llevar a cabo una serie de objetivos menores, que agrupados entre sí contribuirán a lograr la meta deseada. Estos objetivos menores serán la investigación sobre los trabajos de medidas de similitud efectuados

hasta el momento, la implementación adecuada de las funciones de similitud necesarias en el programa de procesamiento Matlab y la configuración de un proceso de registro adecuado para llevar a cabo el análisis de las distintas medidas.

Como se ha comentado, la motivación de este trabajo se centra en la adquisición de una serie de conocimientos sobre un tema muy concreto y de tanta relevancia en la actualidad, el registro de imágenes. Por otra parte, no se puede obviar un primer acercamiento al campo del procesamiento de la señal en su condición más práctica, favoreciendo el aprendizaje de una metodología de trabajo propia, más acorde con el trabajo de investigación y científico.

1.3. PUNTO DE PARTIDA

Antes de todo es necesario realizar un estudio del estado de la técnica hasta la fecha. Son muchos los trabajos y publicaciones realizados sobre el registro de imagen en la actualidad. Muchos de ellos sugieren nuevos métodos para realizar el registro entre un par de imágenes de una naturaleza específica, mientras que otros tantos se centran en una o varias medidas de similitud para analizar su potencia y aplicaciones bajo unas ciertas condiciones de trabajo, con el objetivo de obtener la medida más adecuada para cada caso en concreto.

Las cuatro técnicas de similitud analizadas a lo largo de este trabajo han sido ampliamente estudiadas por multitud de investigadores, y es por ello que resulta tremendamente interesante evaluar su comportamiento comparativo y conjunto en el campo del registro de imágenes multimodo. Es además en el mundo de la medicina donde podemos encontrar más aplicaciones de las medidas estudiadas, y por lo tanto donde las publicaciones editadas se encuentran en mayor abundancia.

El hecho de que se haya trabajado tanto sobre este tema no debe suponer que no quede nada por hacer. El aspecto novedoso que presenta este proyecto es el de recopilar las cuatro medidas de similitud empleadas en un solo trabajo, analizando su comportamiento para imágenes monomodo y multimodo y constatando los resultados ya presentes en otros trabajos. Además, este trabajo intenta añadir más información si cabe sobre estas medidas a la ya existente, sugiriendo asimismo todo tipo de ideas que contribuyan al progreso de la ciencia y la técnica en este determinado campo.

1.4. DESCRIPCIÓN DEL RESTO DE CAPÍTULOS

En este primer capítulo hemos tratado los puntos básicos que motivan la realización del proyecto. En primer lugar se han definido los objetivos generales y la motivación desde el punto de vista científico y de enriquecimiento personal. Posteriormente se ha pasado a hablar de la situación actual del campo de acción, así como de lo que somos capaces de aportar mediante la realización de este proyecto.

En el Capítulo 2 abarcaremos los aspectos fundamentales del registro de imágenes, que nos permitan la comprensión de la terminología utilizada en capítulos posteriores. Veremos en qué consiste esta técnica y qué es lo que se pretende con ello, clasificación de los algoritmos, etc. Todo

ello nos permitirá comprender todas las fases del proyecto, que serán abordadas en los subsiguientes capítulos del proyecto.

El Capítulo 3 se dedicará a la descripción detallada de cada una de las cuatro medidas de similitud a tratar: el coeficiente de correlación cruzada, la información mutua, el ratio de correlación y el criterio de Woods. En este apartado introduciremos estas medidas, basándonos en su formulación matemática y en los aspectos que es necesario tener presentes para su implementación. Como culminación del capítulo se analizarán conjuntamente dichas medidas para determinar sus ventajas e inconvenientes entre sí.

Una vez se conocen los fundamentos básicos tratados en el Capítulo 2, seremos capaces de realizar una descripción detallada de los pasos a seguir para el registro de un par de imágenes dado. El Capítulo 4 recogerá las etapas necesarias para la obtención de una imagen correctamente registrada, describiendo de forma sencilla e ilustrativa cada uno de estos bloques vitales, para facilitar el seguimiento del tema abordado.

A continuación se llegará al capítulo central del proyecto, el Capítulo 5. En esta sección se recogerán los resultados obtenidos mediante el registro para imágenes completas, así como las ideas más relevantes, siempre tratando de analizar todo lo visto previamente de la forma más ilustrativa e intuitiva posible. Finalmente se discutirán los resultados obtenidos por los cuatro métodos empleados.

Por último, en el Capítulo 6 se expondrán las conclusiones que nos reporta el trabajo, así como todos los problemas encontrados durante el proceso de elaboración de este proyecto. Como colofón se estudiará el posible futuro del registro de imágenes y los trabajos en los que sería interesante embarcarse en los próximos años.

En el Capítulo 7 se recoge la bibliografía utilizada y las fuentes de donde se han extraído las imágenes utilizadas.

2. CONOCIMIENTOS PREVIOS

2.1. INTRODUCCIÓN

En este capítulo abordamos los conocimientos previos necesarios para la comprensión de los subsiguientes apartados. En un primer lugar comenzamos definiendo el *registro de imágenes* y todo lo relacionado con esta materia, para posteriormente entrar un poco más en cada una de las variables de esta técnica. Estos conceptos nos ayudarán y servirán como punto de partida en lo que se refiere a una mejor comprensión del trabajo realizado.

2.2. EL REGISTRO DE IMAGEN

Se denomina *registro de imagen* al proceso de encontrar la mejor superposición de dos o más imágenes de una misma escena tomadas en diferentes instantes de tiempo, desde diferentes puntos de vista, y/o tomadas por diferentes sensores. Geométricamente esto supone alinear una de las imágenes, denominada *imagen objetivo*, con otra imagen, denominada *imagen de referencia*. Este tema tiene multitud de aplicaciones, incluyendo la superresolución, detección de rostros, codificación de vídeo, imágenes médicas, predicción del tiempo, codificación de bases de datos, etc. Las diferencias entre las dos imágenes pueden ser de diversa naturaleza: desplazamientos, giros, deformaciones, distinto contraste, etc.; así se puede comprobar cómo a veces este tipo de cálculos llega a ser realmente complejo. Un ejemplo de imágenes para el registro tomadas desde distintos puntos de vista se muestra en la Figura 2.1.

El registro de dos imágenes se considera no trivial a causa de la posibilidad de grandes diferencias en el contenido de las imágenes. Por ejemplo, las diferencias en medicina pueden ser debidas a variaciones biológicas naturales, diferentes tipos de información biológica en las dos imágenes, posicionamiento del paciente en los escáneres y cambios relacionados con el paciente a lo largo del tiempo, consecuencia directa de la edad, el proceso de enfermedad o los propios efectos de la terapia. Para las imágenes médicas, el registro a menudo lidia con una deformación elástica (o no rígida) para hacer frente a las deformaciones producidas por el cuerpo capturado.

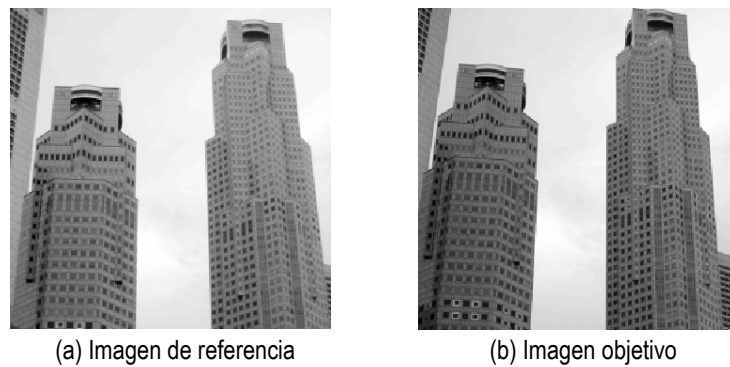


Figura 2.1. Ejemplo de dos imágenes para el registro captadas desde diferentes puntos de vista.

En general, el mal funcionamiento del procedimiento de registro de imágenes aparece cuando encontramos ausencia de información compartida entre las imágenes utilizadas. A veces ocurre el fenómeno conocido por *oclusión* de objetos, lo que hace estos objetos indetectables en la imagen, y a su vez produce la aparición de otros objetos que tampoco están presentes en la otra imagen.

A veces la solución del registro no es única. Es por ello que si no se dispone de una información adicional, a priori todas las soluciones son idénticamente válidas. Cuanto más general es el planteamiento del problema, más grande se hace la ambigüedad.

2.2.1. Naturaleza de las imágenes

Un aspecto muy importante a tener en cuenta en el procedimiento de registro es la naturaleza de las imágenes de las que se dispone. Podemos encontrarnos en dos posibles grupos: imágenes monomodo e imágenes multimodo.

Los algoritmos de registro de imágenes monomodo son aquellos que se utilizan para las imágenes adquiridas por la misma fuente, mientras que los algoritmos dedicados a imágenes multimodo se encargan de aquellas que han sido tomadas por distintos medios de adquisición.

El registro de imágenes monomodo en medicina se utiliza mayoritariamente para detectar cambios físicos o funcionales en un mismo paciente (por ejemplo, antes y después de una intervención quirúrgica para observar la evolución de tumores) o para comparar pacientes entre sí.

En el campo de la medicina podemos encontrar multitud de ejemplos sobre el registro de imágenes multimodo. Entre ellos destacan los de registro de imágenes CT (*Computed Tomography*)/MRI (*Magnetic Resonance Imaging*) del cerebro o imágenes PET (*Positron Emission Tomography*)/CT de todo el cuerpo para la localización de tumores. Otro uso extendido es el de registro de imágenes CT para la segmentación de partes específicas de la anatomía y el registro de ultrasonidos e imágenes CT para la localización de la próstata en radioterapia.

La alineación geométrica o registro de imágenes multimodo es una tarea fundamental en numerosas aplicaciones de procesamiento de imágenes médicas. El diagnóstico médico, por

ejemplo, a menudo se beneficia de la complementariedad de la información en imágenes de distinta modalidad. En planificación de radioterapia, el cálculo de la dosis se basa en datos CT, mientras que el contorno de los tumores es a menudo mejor observado en un escaneo MRI. Para análisis de función cerebral, las imágenes MRI proporcionan información anatómica mientras que la información funcional puede ser obtenida a través de imágenes PET.

Una técnica muy empleada para la representación del resultado en el registro de imágenes de diferente modalidad es la técnica de fusión de imágenes, que proporciona una visualización conjunta (superpuesta) de la imagen de referencia y de la imagen registrada, aplicando para ello el arte de la transparencia. De este modo, el canal de transparencia se selecciona de acuerdo con la siguiente ecuación:

$$[r, g, b]_{fusion} = \alpha[r, g, b]_{REF} + (1 - \alpha)[r, g, b]_{REG}, \quad (2.1)$$

donde r , g y b son los canales de rojo, verde y azul respectivamente de la imagen fusionada. El valor del factor α puede estar entre 0 y 1. Si $\alpha = 0$, solamente se visualiza la imagen registrada, mientras que si $\alpha = 1$, solamente se representa la imagen de referencia.

2.2.2. Tipos de transformaciones

La etapa de transformación es la encargada de mapear el sistema de coordenadas de una imagen al sistema de coordenadas de otra, y puede ser global o local, dependiendo de los datos de la imagen que se utilizan como soporte. Un modelo de transformación se denomina global cuando un cambio en cualquiera de los parámetros influye en la transformación de la imagen como un todo. En la transformación local, tal cambio influye solamente en una parte de la imagen, es decir, se trata a la imagen como un mosaico bien dividido, de forma que los parámetros de transformación son diferentes para cada uno de estos fragmentos. Debe observarse que la transformación que se produce no afecta a la intensidad del píxel, sino única y exclusivamente a sus posiciones espaciales. Por tanto, los niveles de intensidad asociados a las posiciones del píxel quedan inalterados.

Con la transformación se desea que el registro efectuado establezca en último término una correspondencia biunívoca entre cada par de imágenes. Este emparejamiento habitualmente se representa mediante un conjunto de vectores de correspondencia (véase la Figura 2.2), que definen esta determinada transformación a aplicar sobre una de las imágenes para obtener la otra.

El modelo con el que se relaciona la transformación detectada entre la imagen objetivo y la de referencia se puede agrupar además en otros dos tipos: modelos de transformación paramétricos y no paramétricos.

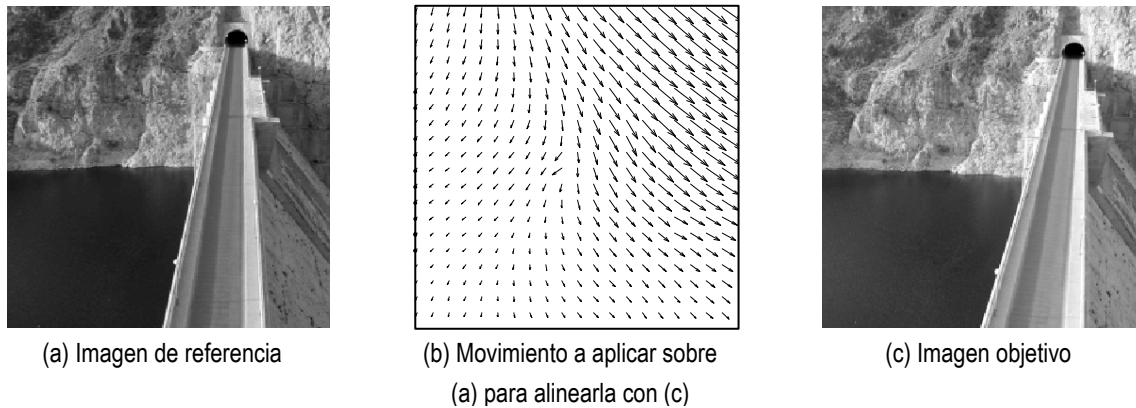


Figura 2.2. Ejemplo de vectores de correspondencia.

2.2.2.1 Transformaciones paramétricas

Las transformaciones paramétricas son llamadas también transformaciones *lineales*, ya que son una combinación de translación, reflexión, rotación y transformación afín (escalado) o proyectiva (distinta perspectiva). Este tipo de transformaciones son globales en general, por lo que no son capaces de modelar deformaciones locales, producidas en distintas partes de la imagen y de distinta forma. Todas las transformaciones locales pueden introducir solapamientos o huecos. Las transformaciones afín, proyectiva y curva se pueden restringir de tal modo que los solapamientos no ocurran (véase la Figura 2.3).

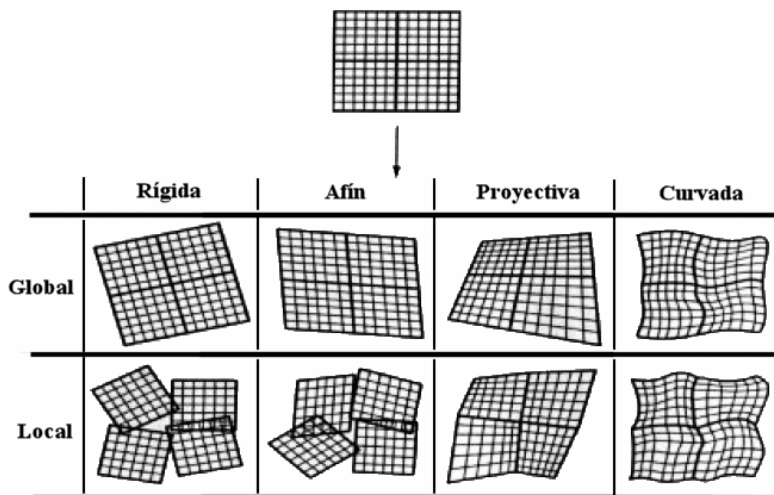


Figura 2.3. Ejemplos de varias transformaciones 2D.

A menudo se suele hablar de transformación rígida cuando la distancia entre dos puntos cualesquiera en la primera imagen se preserva cuando esos dos puntos son mapeados en la segunda imagen. La transformación rígida está limitada a la rotación, la translación y la reflexión.

La translación suele ser la más presente en todas las aplicaciones descritas, por lo que a veces se utiliza para ello una primera fase en el proceso de registro con el fin de estimar este desplazamiento.

En un gran número de aplicaciones actuales, los métodos de registro rígido de imágenes no proporcionan una solución adecuada, por lo que se utilizan complejas transformaciones *no lineales* para detectar, emparejar o corregir las diferencias no locales existentes. El registro de imágenes médicas pasa por ser la aplicación donde más claramente se requiere un registro no rígido. En estos casos puede hablarse de registro *deformable*.

2.2.2.2 Transformaciones no paramétricas

Este tipo de modelos de transformación son conocidos también como *elásticos* y en este caso sí que permiten deformaciones locales. Las aproximaciones no paramétricas incluyen *warping* polinomial o interpolación de funciones de base suave entre otros, como las RBFs.

Las funciones de base radial o RBFs son funciones de mapeo global, aunque son capaces de hacer frente a distorsiones geométricas variantes o locales. La función de transformación resultante es una combinación lineal de funciones radialmente simétricas trasladadas, más un polinomio de pequeño grado.

2.3. CLASIFICACIÓN DE LOS ALGORITMOS DE REGISTRO DE IMÁGENES

El registro no rígido ha sido un campo de trabajo muy importante durante las últimas décadas. Hoy en día las técnicas que podemos encontrar son muy diversas, aunque el debate todavía sigue abierto para ciertos aspectos y aplicaciones. En este apartado se ha optado por utilizar la clasificación de los algoritmos para el registro no rígido propuesta por Pascal Cachier y Nicholas Ayache [13]. Ésta supone la división de la mayoría de los métodos de registro no rígido en dos categorías principales: en primer lugar los algoritmos geométricos; y por otra parte los métodos basados en intensidad, que a su vez se encuentran divididos en estándares (*Standard Intensity-Based, SIB*) y de emparejamiento y suavizado (*Pair and Smooth, P&S*).

Con la transformación elegida y el método de evaluación de la transformación definidos, el registro de imágenes se convierte en una tarea de optimización para encontrar los mejores parámetros de transformación. Actualmente, el uso de la información de intensidad parece estar mucho más extendido que el uso de las características geométricas para el registro no rígido. Este hecho se debe al éxito de las medidas de similitud de intensidad en el caso rígido, especialmente para el registro multimodo, y también a la dificultad de extraer automáticamente regiones conocidas que permanezcan lo suficientemente estables bajo movimiento no rígido y variación de intensidad.

Sin embargo, hoy en día la estrategia de mayor aceptación es desarrollar técnicas mixtas que contemplen las ventajas de un emparejamiento basado en la información de intensidad con las ventajas proporcionadas por un emparejamiento basado en la geometría.

2.3.1. Emparejamiento geométrico

En este tipo de técnica se lleva a cabo la detección y extracción de las estructuras marcadas en las imágenes, a las que comúnmente se las denomina *características* o *rasgos* de la imagen. Estas características van desde bosques, lagos o campos, a caminos o ríos. Estos rasgos deben estar presentes en ambas imágenes y a ser posible dispersos sobre toda la imagen, de modo que el número de ellos sea lo más alto posible. Estas características generalmente se pueden dividir en tres grupos principales:

1. **Características tipo región:** Estos objetos suelen ser áreas cerradas de alto contraste y un tamaño aproximado, lagos, edificios, bosques, etc. Estas regiones se describen mediante sus centros de masa, que es invariante a la rotación y al escalado, y estable frente al ruido y la iluminación no uniforme.
2. **Características tipo línea:** Este tipo de características pueden ser representaciones de línea genéricas, contornos, líneas costeras, caminos, etc. Las correspondencias entre líneas se representa generalmente mediante los pares de extremos de línea o sus puntos intermedios.
3. **Características tipo punto:** Pueden representar intersecciones entre líneas, cruces de carreteras, centros de regiones muy concretos, puntos de alta varianza o discontinuidades locales de curvatura. Las esquinas constituyen un caso especial de puntos, ya que se pueden considerar como puntos de alta curvatura que delimitan una región.

El proceso de emparejamiento geométrico consiste en extraer en primer lugar el conjunto de características detectadas en la imagen, manual o automáticamente. En los primeros métodos geométricos las características eran objetos seleccionados manualmente por expertos facultativos en ambas imágenes, por lo que se les daba el nombre de métodos *basados en marcadores*. En la evolución propia de estos tipos de emparejamientos, este paso ha ido automatizándose hasta el punto de ir prescindiéndose de la intervención humana, dando lugar a los llamados métodos *basados en segmentación de imagen*.

Una vez extraídas las características, se emparejan juntas mediante una transformación suave no rígida, de tipo global. El problema de encontrar una fórmula precisa para la transformación que interpole o aproxime el conjunto de pares no es trivial, incluso con características de tipo punto y transformaciones rígidas. Aproximaciones alternativas incluyen la resolución de elementos finitos, ajustes multidimensionales con *splines* (véase apartado 4.4.1.1), difusión de lugares conocidos a píxeles vecinos u otras técnicas de extrapolación.

Debido al elevado coste computacional que conlleva el emparejamiento de características, sobre todo para imágenes de tamaños elevados, se han llevado a cabo numerosas investigaciones a lo largo de los años cuyo objetivo ha sido la reducción de dicho coste de cálculo. La mayoría de ellas hacen uso de técnicas piramidales, también denominadas jerárquicas, multirresolución o multiescala.

2.3.2. Registro basado en intensidades

Los métodos basados en intensidades integran en una sola fase la detección y el emparejamiento de características. Estos métodos tratan las imágenes sin una detección previa de rasgos significativos ni segmentación. Para la estimación de la correspondencia entre imágenes se utilizan ventanas de tamaño predefinido tomadas de las imágenes, o algunas veces incluso imágenes completas.

El uso de ventanas rectangulares resulta adecuado sólo cuando nos encontramos con imágenes que difieren por efectos de translación. Si las imágenes han sufrido transformaciones más complejas, este tipo de ventanas pueden no cubrir apropiadamente la misma zona del conjunto de datos de la imagen de referencia y la imagen objetivo. Para solucionar los problemas presentes con las ventanas cuadradas, se utilizan las ventanas circulares, que en cierto modo ayudan a evitar estas limitaciones, pero para deformaciones geométricas tales como los cambios de perspectiva encontramos que esta nueva solución tampoco es suficiente.

Otra ventaja de estos métodos está directamente relacionada con las zonas donde existe un área casi o totalmente homogénea. Este tipo de áreas, al no tener detalles característicos, pueden ser emparejadas con ciertas partes similares de la otra imagen, por lo que se debe de tratar de buscar características en zonas bien definidas y contrastadas de la imagen. Estos métodos representan la aproximación más simple al problema del registro de imagen, al hacer casi exclusivamente uso de los valores de intensidad de imagen, prescindiendo del reconocimiento previo de formas presentes en las imágenes. El factor que diferencia a un algoritmo basado en intensidad de otro suele ser la medida de similitud empleada por cada uno de ellos para establecer la correspondencia adecuada entre pares de imagen.

En el caso de imágenes médicas, que no son ricas en detalles, la formación distintiva está contenida en los niveles de gris principalmente. Es por ello que para este tipo de aplicaciones se hace uso preferentemente de estos algoritmos. La condición necesaria para poder aplicar este tipo de algoritmos a las imágenes médicas es que ambas imágenes tengan funciones de intensidad parecidas.

Como ya se ha comentado, este tipo de algoritmos se divide en dos grupos: los algoritmos basados en intensidad de tipo estándar (SIB) y los de emparejamiento y suavizado (P&S). Éstos se describen a continuación.

2.3.2.1 *Registro basado en intensidad estándar (SIB)*

En estos algoritmos las imágenes son registradas usando una medida de similitud de intensidad, mediante una transformación T no rígida. Esta transformación tiene dos metas opuestas: ser lo más suave posible y maximizar la medida de similitud. La principal característica de este tipo de técnicas, al contrario que los algoritmos P&S, es que está presente la misma transformación en la medida de similitud de intensidad y en la limitación de la suavidad de la transformación.

Podemos dividir esta categoría de registro en tres subclases:

1. **Paramétrica:** La transformación T se encuentra limitada para minimizar la medida de similitud de intensidad en un espacio de pocas dimensiones y de transformación no rígida.
2. **Competitiva:** La transformación T no tiene forma a priori, pero minimiza una suma ponderada de la energía de similitud de intensidad y de regularización. Esta aproximación a veces se justifica mediante la formulación bayesiana.
3. **Fluida:** La transformación T no tiene forma a priori, pero su evolución hacia el mínimo de la energía de similitud está limitada para ser suave. Una "distancia fluida" entre dos transformaciones se puede definir como el camino más corto presente entre ellas, es decir, se puede diseñar un algoritmo "fluido" que minimice la suma ponderada de la energía de similitud y de la distancia "fluida" entre las transformaciones.

Como podemos ver, el primer tipo de regularización es de tipo *paramétrico*, mientras que los otros dos son *no paramétricos*. Sin embargo, estas tres subclases no son exclusivas, por lo que encontraremos algunos algoritmos que combinen varias de ellas.

2.3.2.2 Registro de emparejamiento y suavizado (P&S)

Esta categoría es intermedia entre las dos vistas anteriormente. Los algoritmos de emparejamiento y suavizado también están basados en intensidad, ya que usan una energía de similitud de la intensidad para cuantificar la calidad de la correspondencia entre las imágenes. Sin embargo, estos algoritmos operan en dos pasos, que pueden alternarse o no. En primer lugar se usa una medida de similitud de intensidad para encontrar un conjunto relativamente denso de puntos o características entre ambas imágenes. Estas correspondencias se aproximan entonces mediante una transformación suave no rígida.

La principal característica de estos algoritmos, en contraposición al registro SIB, es que usan dos transformaciones (a menudo implícitamente): una primera transformación C para calcular la medida de similitud; y una segunda transformación T en la que la limitación de suavidad se mantiene. Es por ello que en la práctica la suavidad de las soluciones encontradas es mucho más uniforme y controlable que en otros métodos. En esta categoría encontramos los algoritmos de *Block-matching*, para registro rígido y no rígido y los algoritmos basados en flujo óptico, por ejemplo.

En el capítulo 4 se describirá el procedimiento utilizado para realizar el registro de las imágenes empleadas en este proyecto, utilizando para ello un algoritmo de tipo P&S.

2.4. MEDIDAS DE SIMILITUD

En el registro de imágenes la precisión de los resultados depende en gran medida de la medida de similitud empleada. Para el registro de imágenes se emplean muchas medidas de similitud. Estas medidas pueden ser calculadas sobre imágenes completas o con bloques de ellas.

Tradicionalmente, la limitada información sobre el comportamiento de una medida de similitud era obtenida estudiando la calidad del registro final o dibujando gráficas de similitud de los valores de medida obtenidos por rotación o translación de la imagen. Hoy día estamos asistiendo a procedimientos de optimización mediante la evaluación estadística de las medidas de similitud. Estos tipos de técnicas estiman entre otras cosas el rango de captura, el número y localización de un óptimo local y la precisión de un óptimo global, para obtener el mejor resultado. No hay una única medida de similitud que pueda producir un mejor resultado que otra para todo tipo de imágenes, de tal modo que la búsqueda de factores de similitud constituye un área de investigación relevante en el mundo del procesado de la señal.

A grandes rasgos, las medidas de similitud están divididas en dos tipos: medidas basadas en métodos geométricos, que utilizan para ello la extracción de distintas características, y medidas basadas en niveles de intensidad. Éstas últimas a su vez también se encuentran divididas en dos tipos: las de promedios y las de tipo estadístico. La división efectuada está directamente relacionada con el tipo de algoritmo de registro llevado a cabo, por lo que a veces en la bibliografía el tipo de emparejamiento y las medidas de similitud son consideradas como una misma cosa.

En el pasado, los métodos de registro se han basado mayoritariamente en emparejar las características correspondientes de las imágenes a registrar, pero más recientemente, el interés se ha dirigido a las medidas de la correspondencia global obtenida directamente a partir de las intensidades de la imagen.

Los registros de imágenes basados en medidas de similitud ajustan los parámetros de una transformación espacial apropiada, hasta que la medida en cuestión alcanza un óptimo. Dadas las imágenes a ser registradas y el modelo de transformación espacial, el resultado de la tarea de registro principalmente depende de la medida de similitud y del método de optimización.

Cada una de las distintas medidas de similitud introducidas hasta la fecha tiene diferentes propiedades, y son por lo tanto diferentemente sensibles a la modalidad, contenido y diferencia de la imagen en cuanto a muestreo, interpolación, solapamiento parcial y degradación (por ejemplo ruido). Para poder elegir la medida de similitud y su implementación de la manera más apropiada, así como el método de optimización, es a menudo deseable disponer de alguna información *a priori* sobre el comportamiento de la función de similitud con respecto a los factores mencionados. Sin embargo, incluso en el caso simple de registro rígido de dos imágenes, el parámetro espacio es demasiado grande para analizar la función de similitud en cada punto de este espacio. A partir de la precisión y robustez del propio registro es posible obtenerse, no obstante, una información limitada de la medida de similitud empleada.

2.4.1. Medidas basadas en el nivel de intensidad

Las medidas basadas en nivel de intensidad utilizan los valores de los *píxeles* directamente, usando para ello una medida basada en promedios o en la estadística de la imagen. Este tipo de medidas son por lo habitual más lentas que las basadas en métodos geométricos, que utilizan una menor cantidad de datos, pero requieren poca o ninguna segmentación. Los métodos de similitud

por nivel de intensidad tratan de encontrar un máximo (o mínimo) valor donde las dos imágenes estén perfectamente alineadas.

Los métodos de registro basados en este tipo de medidas por nivel de intensidad han capturado actualmente mucho interés para el registro de imágenes multimodo. A continuación se describen las dos clases principales de medidas basadas en el nivel de intensidad.

2.4.1.1 *Medidas basadas en promedios*

Si las funciones de intensidad entre las imágenes son muy similares, entonces se pueden aplicar métodos sencillos basados en promedios. Este tipo de medidas utilizan operaciones de tipo no estadístico para el cálculo de la similitud, como la diferencia, el cociente y el sumatorio. Ejemplos de este tipo de métodos son la suma de las diferencias absolutas (*Mean Absolute Difference, MAD*), el coeficiente normalizado de correlación cruzada (*Normalized Cross-Correlation Coefficient, NCC*) y medidas basadas en distancias geométricas, como la diferencia cuadrática media (*Mean Squared Difference, MSD*) o la raíz cuadrática media (*Root Mean Square, RMS*).

2.4.1.2 *Medidas basadas en la estadística de la imagen*

Si las funciones de intensidad son al menos estadísticamente dependientes (típicamente en el registro multimodo), se pueden usar medidas basadas en la estadística de la imagen, como el ratio de correlación (*Correlation Ratio, CR*), el criterio de Woods (*Woods Criterion, WC*), la información mutua (*Mutual Information, MI*) o heurísticos.

Dentro de las medidas de tipo estadístico, podemos encontrar otro grupo distintivo, las medidas basadas en la teoría de la información. La correcta alineación geométrica entre imágenes es crítica para la obtención de resultados precisos, y una de las formas de llevar a cabo esto es mediante la aplicación de estas técnicas. Las medidas basadas en la teoría de la información hacen uso de la teoría matemática de la información, fundamentada en la teoría de la probabilidad y la estadística. Este tipo de medidas hacen uso de la entropía entre otras importantes medidas de información. La información mutua se define como la cantidad de información en común entre dos variables, y es el ejemplo más claro de las medidas de similitud de este tipo.

3. MEDIDAS DE SIMILITUD

3.1. INTRODUCCIÓN

Las medidas de similitud son una de las bases fundamentales del proceso de registro de imágenes, ya que son éstas las que nos permiten detectar las parte de la imagen objetivo que tienen correspondencia con otras de la imagen de referencia. Son muchas las medidas que se han estudiado y de las que podemos encontrar multitud de trabajos. Algunos ejemplos de dichas medidas son la correlación cruzada [1], la entropía de la diferencia entre imágenes, la correlación y la diferencia del gradiente [4], la energía del histograma [5], la información mutua [7], el ratio de correlación y el criterio de Woods [12].

En este capítulo vamos a analizar cada una de las cuatro medidas más utilizadas en la práctica, que se han implementado y evaluado a lo largo del proyecto. Estas cuatro medidas son el coeficiente normalizado de correlación cruzada (NCC), la información mutua (MI), el ratio de correlación (CR) y el criterio de Woods (WC). Estas tres últimas son de tipo estadístico y nos permiten registrar imágenes de distinta modalidad.

3.2. MÁSCARAS DE PASO Y PONDERACIÓN

Para el cálculo de los diversos promediados sobre bloques no rectangulares, y de distribuciones geométricas arbitrarias, se hace uso de una ventana especial llamada *máscara de datos*. Este tipo de máscaras son tales que al multiplicar el conjunto de datos por ellas sólo permanece el conjunto de datos que coincide con las posiciones de la máscara donde el valor es igual a 1, mientras que los valores no deseados quedan anulados al ser multiplicados por un valor igual a 0 de la máscara de paso. En el caso de imágenes, este tipo de máscaras tienen carácter bidimensional. La Figura 3.1 muestra el resultado de aplicar una máscara de paso circular a una imagen fotográfica.

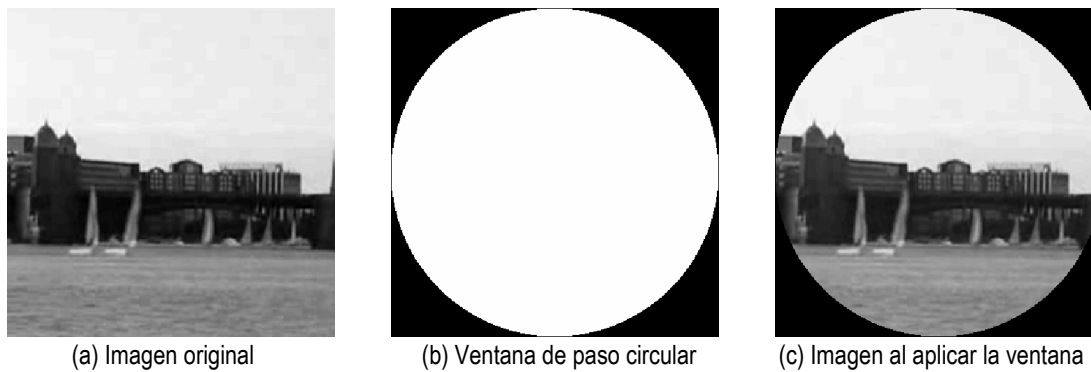


Figura 3.1. Ejemplo de máscara de paso circular aplicada a una imagen.

Del mismo modo que se aplica una máscara de paso, se puede aplicar una *máscara de ponderación*, que nos puede ser de utilidad para ponderar el peso de cada uno de los datos implicados en el cálculo de un determinado estadístico, los denominados *momentos*. El peso de la máscara puede venir dado por la posición geométrica, la intensidad, etc.

3.3. EL COEFICIENTE DE CORRELACIÓN CRUZADA

La medida de similitud más frecuente es el coeficiente normalizado de correlación cruzada, o *Normalized Cross Correlation (NCC)*. La correlación se calcula entre parejas de bloques pertenecientes a la imagen de referencia y a la imagen objetivo, buscando el máximo en dicha medida. Es el bloque con el que se consigue el máximo el que determina la correspondencia finalmente establecida.

El registro basado en el coeficiente normalizado de correlación cruzada nos permite en principio alinear con precisión únicamente imágenes con translaciones entre sí, aunque también es posible la aplicación de esta medida en el caso de rotaciones leves o escalado.

Las dos principales desventajas de los métodos de correlación son las siguientes:

1. La excesiva planicidad de los máximos de similitud detectados debido a la autosimilitud de las imágenes.
2. La elevada carga computacional, aunque inferior a otras medidas.

Para solucionar el primer punto se suele resaltar el máximo aplicando un preprocesado a la imagen o utilizando una correlación de bordes o de vectores. De este modo la medida de similitud es menos vulnerable a las diferencias de intensidad entre las imágenes de referencia y objetivo.

Recientemente ha surgido un gran interés por el registro multimodo, por lo que la principal atención se ha dirigido a métodos basados en el ratio de correlación, que veremos más adelante a lo largo de este tema. Sin embargo, la correlación como método de registro todavía sigue en uso y se utiliza en gran número de aplicaciones gracias a su sencilla implementación *hardware*, lo que la hace muy extendida para aplicaciones en tiempo real.

3.3.1. Expresión matemática del NCC

Sean dos conjuntos de datos $I(\mathbf{n})$ y $J(\mathbf{n})$ de cualquier dimensión, donde \mathbf{n} representa la posición de un determinado dato en el conjunto, y sean $I_{\mathbf{v}}(\mathbf{n})$ y $J_{\mathbf{v}+\mathbf{x}}(\mathbf{n})$ los subconjuntos de datos extraídos de las posiciones de $I(\mathbf{n})$ y $J(\mathbf{n})$ indexadas por las matrices de índices \mathbf{v} y $\mathbf{v}+\mathbf{x}$ respectivamente. El término \mathbf{x} representa un cierto desplazamiento respecto a los índices \mathbf{v} , por lo que \mathbf{x} recibe el nombre de *matriz de desplazamientos*. Se define el *coeficiente de correlación* de la siguiente forma:

$$r_{\mathbf{v}}(\mathbf{x}) = \frac{\widehat{I_{\mathbf{v}} J_{\mathbf{v}+\mathbf{x}}}}{\sqrt{\widehat{I_{\mathbf{v}}^2} \widehat{J_{\mathbf{v}+\mathbf{x}}^2}}} = \frac{\overline{I_{\mathbf{v}} J_{\mathbf{v}+\mathbf{x}}} - \bar{I}_{\mathbf{v}} \bar{J}_{\mathbf{v}+\mathbf{x}}}{\sqrt{\left(\overline{I_{\mathbf{v}}^2} - \bar{I}_{\mathbf{v}}^2\right) \left(\overline{J_{\mathbf{v}+\mathbf{x}}^2} - \bar{J}_{\mathbf{v}+\mathbf{x}}^2\right)}}, \quad (3.1)$$

donde se definen

$$\bar{W} = \frac{1}{N_{\eta}} \sum_{x=-L_x}^{L_x} \sum_{y=-L_y}^{L_y} W(x, y) \eta(x, y), \quad (3.2)$$

$$\overline{W^2} = \frac{1}{N_{\eta}} \sum_{x=-L_x}^{L_x} \sum_{y=-L_y}^{L_y} W^2(x, y) \eta(x, y), \quad (3.3)$$

$$\widehat{W^2} = \frac{1}{N_{\eta}} \sum_{x=-L_x}^{L_x} \sum_{y=-L_y}^{L_y} (W(x, y) - \bar{W})^2 \eta(x, y), \quad (3.4)$$

como la media, el valor cuadrático medio y la varianza del conjunto de datos definido por una imagen bidimensional arbitraria $W(x, y)$ y la máscara de paso $\eta(x, y)$, donde N_{η} representa el número de puntos a considerar dentro de una área de dimensiones $N_x \times N_y$, con $N_x = 2L_x + 1$ y $N_y = 2L_y + 1$, correspondiente a la máscara de paso, dada por:

$$N_{\eta} = \sum_{x=-L_x}^{L_x} \sum_{y=-L_y}^{L_y} \eta(x, y), \quad (3.5)$$

Para dos imágenes $W(x, y)$ y $Z(x, y)$ también definimos

$$\overline{WZ} = \frac{1}{N_\eta} \sum_{x=-L_x}^{L_x} \sum_{y=-L_y}^{L_y} W(x, y)Z(x, y)\eta(x, y), \quad (3.6)$$

$$\hat{WZ} = \frac{1}{N_\eta} \sum_{x=-L_x}^{L_x} \sum_{y=-L_y}^{L_y} (W(x, y) - \overline{W})(Z(x, y) - \overline{Z})\eta(x, y), \quad (3.7)$$

$$\overline{(W - Z)^2} = \frac{1}{N_\eta} \sum_{x=-L_x}^{L_x} \sum_{y=-L_y}^{L_y} (W(x, y) - Z(x, y))^2 \eta(x, y), \quad (3.8)$$

respectivamente como la correlación, la covarianza y la diferencia cuadrática media entre los dos conjuntos de datos.

También definimos el *coeficiente de diferencia cuadrática* de la siguiente forma:

$$d_v(x) = \frac{\overline{(I_v - J_{v+x})^2}}{\hat{I}_v^2} = \frac{\overline{I}_v + \overline{J}_{v+x} - 2\overline{I}_v\overline{J}_{v+x}}{\overline{I}_v^2 - \overline{I}_v^2}. \quad (3.9)$$

El coeficiente de diferencia cuadrática media (también llamado MSD) se calcula paralelamente al del coeficiente de correlación, y se usa debido a la poca fiabilidad de esta medida en bloques de bajo contraste o textura, es decir, de baja varianza y similar nivel medio de intensidad. Con un uso conjunto del coeficiente de correlación, que es sensible a las transiciones de niveles de gris, y de la diferencia cuadrática, que es sensible a los propios niveles de gris, se puede conseguir una mejora en la capacidad de detección de patrones, en lo referente a discriminación, selectividad o directividad en ángulo de giro. Por lo tanto, la medida de similitud resultante de esta combinación tiene la siguiente expresión:

$$\rho = \frac{r}{1 + k \cdot d}, \quad (3.10)$$

siendo r el coeficiente de correlación, d la diferencia cuadrática y k un parámetro que nos permite ponderar la importancia de la medida de la diferencia cuadrática frente a la correlación.

En la práctica, un valor aproximado de k que suele dar buenos resultados es $k = 3$, al proporcionar un compromiso perfecto entre sensibilidad al tipo de patrón y los niveles de intensidad y contraste. Los valores resultantes de p suelen umbralizarse a un valor cercano a 0.25-0.3, por lo que valores menores quedan automáticamente reasignados a un nivel igual a cero.

En la Figura 3.2 podemos observar el resultado de un mapa de similitud para el caso del coeficiente de correlación y la combinación de éste con la diferencia cuadrática. Un mapa de similitud se genera calculando el valor de la similitud entre dos bloques determinados, en las posiciones donde ambos pueden superponerse por tener el mismo tamaño. Cuando calculamos la medida de similitud de dos bloques de igual tamaño, el resultado que obtendremos será un escalar que de cierto modo indicará el grado de interdependencia entre ambos bloques. Si por el contrario nos disponemos a comparar dos bloques de distinto tamaño (J mayor que I), el cálculo del mapa debe hacerse por iteraciones, es decir, el bloque I (Figura 3.2(a)) se debe comparar con cada posible combinación de las divisiones de J (Figura 3.2(b)) en subbloques del tamaño de I . Por ejemplo, si el bloque I tiene un tamaño de 2×2 y el bloque J uno de 3×3 , tendremos cuatro iteraciones:

$$I = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \quad J = \begin{pmatrix} p & q & r \\ s & t & v \\ w & x & y \end{pmatrix}$$

- Iteración 1: Se calcula la similitud de $I = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y $J' = \begin{pmatrix} p & q \\ s & t \end{pmatrix}$, resultando M_1 .
- Iteración 2: Se calcula la similitud de $I = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y $J' = \begin{pmatrix} q & r \\ t & v \end{pmatrix}$, resultando M_2 .
- Iteración 3: Se calcula la similitud de $I = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y $J' = \begin{pmatrix} s & t \\ w & x \end{pmatrix}$, resultando M_3 .
- Iteración 4: Se calcula la similitud de $I = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y $J' = \begin{pmatrix} t & v \\ x & y \end{pmatrix}$, resultando M_4 .

El resultado es un mapa de similitud de la forma $MAPA = \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}$.

A la hora de calcular el mapa de similitud también hay que tener en cuenta la presencia o no de una máscara de paso. Si esto es así, sólo se calcularán los valores del mapa pertenecientes a los lugares donde esa máscara de paso sea 1, mientras que el resto de valores serán asignados directamente a 0.

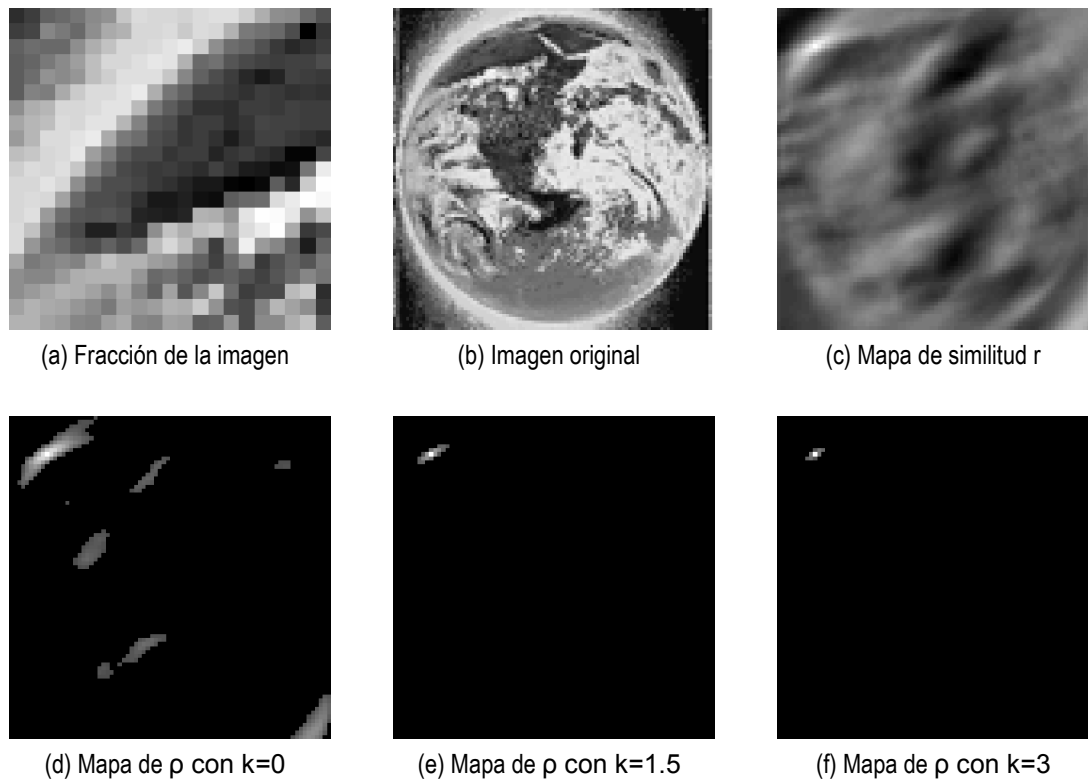


Figura 3.2. Influencia del parámetro de ponderado k en los mapas de similitud.

Como podemos observar, la Figura 3.2(a) es un fragmento de la esquina superior izquierda de la Figura 3.2(b), y se va calculando el valor de la medida de similitud para cada una de las posibles combinaciones sobre la imagen (b). Obviamente, el valor mayor de este mapa (en blanco) coincide con la posición de la que se ha extraído la imagen (a). En la Figura 3.2(d) vemos el resultado de aplicar una umbralización de 0.3 a la Figura 3.2(c) sin tener en cuenta la diferencia cuadrática; en la Figura 3.2(e) vemos el resultado de una umbralización del mismo tipo, pero esta vez con un factor de ponderado de $k = 1.5$; por último, en la Figura 3.2(f) vemos el resultado para el valor típico de $k = 3$.

3.3.2. Algoritmo de implementación

A continuación se describen los distintos pasos del algoritmo que calcula el coeficiente de correlación de dos bloques de datos. Para ello se hace uso de la convolución, que nos permite desarrollar un procedimiento más rápido y eficiente, a partir de la ecuación (3.1). El hecho de que la convolución consiga un ahorro del tiempo de procesamiento importante se debe a la idea de que se produce una reducción en el número de operaciones requeridas, haciendo para ello uso de la simetría de las máscaras de convolución.

La correlación cruzada es similar en naturaleza a la convolución de dos funciones. Mientras que la convolución implica invertir una señal, y luego desplazarla y multiplicarla por otra señal, la correlación sólo implica desplazarla y multiplicarla (no invertirla).

Para ilustrar la afirmación anterior supongamos la correlación cruzada de dos funciones complejas $f(m)$ y $g(m)$ de una variable discreta m , denotada como $f \bullet g$, y que se puede expresar como

$$(f \bullet g)(m) = \sum_n \bar{f}(n)g(m+n), \quad (3.11)$$

donde $\bar{f}(m)$ es el complejo conjugado de $f(m)$. Ya que la convolución entre dos funciones discretas (representada por un $*$) está definida como

$$(f * g)(m) = \sum_n f(n)g(m-n), \quad (3.12)$$

se puede expresar la relación existente entre ambas mediante

$$f \bullet g \equiv \bar{f}(-m) * g(m). \quad (3.13)$$

La expresión (3.13) viene a decir exactamente lo ya comentado anteriormente. Así pues, invirtiendo el orden de los datos utilizados en la convolución, seremos capaces de llevar a cabo las operaciones para calcular la correlación cruzada con una mayor velocidad que la conseguiríamos utilizando todas las operaciones implícitas en las fórmulas introducidas en este apartado.

La ecuación (3.1) será por lo tanto la utilizada como base para el cálculo de la similitud entre las imágenes. Ésta la podemos describir como $r_v(x) = \frac{cfactor1 - A \cdot Jm \cdot Im}{\sqrt{Ci \cdot Cj}}$, reduciéndose todo a

obtener estos términos por separado:

- 1) Se calcula el área de la máscara de paso utilizada (sumatorio del número de unos que contiene) y se asigna a la variable **A**.
- 2) Se aplica la ventana reflejada al primer bloque (bloque **I**), también reflejado, y se asigna a la variable **IMaskReverse**.
- 3) El valor promedio de los píxeles de la máscara se calcula entonces como el sumatorio de $\frac{Imaskreverse}{A}$, y se asigna a la variable **Im**.

- 4) Se realiza la convolución 2D del segundo bloque (bloque **J**) con la máscara sin reflejar dividida por **A**. Todo esto se asigna a la variable **Jm**.
- 5) Asignamos a **Ci1** el sumatorio de **IMaskReverse** al cuadrado.
- 6) Se calcula **Ci** como **Ci = Ci1 - A·Jm·Jm**, produciendo un valor escalar.
- 7) Se lleva a cabo la convolución 2D del bloque **J** al cuadrado con la máscara de paso sin reflejar, y se almacena en **Cj1**.
- 8) Se calcula **Cj** como **Cj = Cj1 - A·Jm·Jm**, produciendo un resultado escalar.
- 9) De nuevo calculamos la convolución 2D, pero esta vez entre el bloque **J** e **IMaskReverse**, asignando el resultado a **cfactor1**.
- 10) Calculamos **cfactor** como **cfactor = cfactor1 - A·Jm·Jm**.
- 11) Por último, el resultado final del coeficiente de correlación lo proporciona el cociente $\frac{cfactor}{\sqrt{Ci \cdot Cj}}$.

Para el cálculo de la diferencia cuadrática media se sigue un algoritmo muy similar, ya que se puede describir (3.9) como $d_v(x) = \frac{Ci1 + Cj1 - 2cfactor1}{Ci}$:

- 1) Se calcula **Ci**, **Ci1**, **Cj**, **Cj1** y **cfactor1** como en el algoritmo anterior.
- 2) Se obtiene la variable **dfactor** como **dfactor = Ci1 + Cj1 - 2·cfactor1**.
- 3) El resultado final lo proporciona el cociente $\frac{dfactor}{Ci}$.

Un aspecto muy a tener en cuenta en estos algoritmos es el hecho de que pueden producirse divisiones por cero debido a que el producto entre **Ci** y **Cj** puede llegar a ser nulo. Para estos casos, es conveniente asignarse un valor igual a 0 para el coeficiente de correlación y un valor alto para la diferencia cuadrática (por ejemplo 100), de modo que se indique un parecido casi nulo entre los conjuntos de datos. Además, por la misma razón, para las zonas del mapa de similitud resultante que se encuentran fuera de las zonas de paso de la máscara, también se asigna un valor nulo al coeficiente de correlación y un valor alto a la diferencia cuadrática.

3.3.3. Limitaciones

La correlación de datos es capaz de detectar formas o distribuciones geométricas de datos bien definidas, como los presentes en bordes, contornos o transiciones. Además, este parámetro se muestra más efectivo cuanto más contrastado es el patrón buscado, o dicho de otro modo, cuanto

mayor es su varianza. El coeficiente de correlación es por definición independiente de los valores medios de los bloques comparados, y por consiguiente de los niveles medios de iluminación de las porciones de imagen. La contribución a partir de un píxel dado es fuertemente dependiente de la intensidad de dicho píxel, por lo que algunas grandes diferencias en intensidad pueden causar un efecto substancial sobre la medida de similitud.

Sin embargo, aunque la correlación resulta estupenda para detectar movimiento rígido, ésta presenta numerosas limitaciones para el omnipresente caso de movimiento deformable, debido a su gran sensibilidad a las deformaciones y a las rotaciones de bloque. En casos de considerable deformación o rotación entre los bloques, la correlación se muestra poco robusta cuanto mayor es el tamaño de bloque utilizado. Por tanto, si cogemos bloques de tamaños menores, estaremos en cierto modo evitando el problema de las deformaciones o rotaciones, pero por otra parte se puede llegar a una pérdida drástica de la capacidad de discriminación de patrones. Cada tipo de aplicación determinará el tamaño de los bloques a utilizar para el cálculo del mapa de similitud.

3.3.4. Ejemplos

A continuación mostramos una serie de ejemplos para advertir el funcionamiento del coeficiente normalizado de correlación y contrastar lo afirmado en el apartado anterior. En la Figura 3.3 se observa una imagen tomada a partir de la columna vertebral de un ser humano. De esta imagen se extrae un bloque de la zona indicada por el cuadro blanco, que se utiliza para calcular el mapa de similitud en toda la imagen original. La Figura 3.3(b), Figura 3.3(c) y Figura 3.3(d) muestran el resultado para un tamaño de bloque de 6 x 6 píxeles, mientras que la Figura 3.3(e), Figura 3.3(f) y Figura 3.3 (g) lo hacen para un tamaño de 11 x 11, y la Figura 3.3(h), Figura 3.3(i) y Figura 3.3(j) para un bloque de 16 x 16. La Figura 3.4 analiza este mismo comportamiento para el caso en que la imagen con la que comparar haya sufrido una rotación.

Al analizar ambos conjuntos de figuras, podemos sacar las siguientes conclusiones:

- Un tamaño de bloque menor lleva consigo un mapa de similitud más definido y por lo tanto localiza más zonas donde el parecido es alto con el bloque tomado.
- Conforme aumentamos el tamaño del bloque, la suavidad del mapa de similitud aumenta y las zonas con un alto grado de similitud disminuyen, perdiendo capacidad de detección para grandes deformaciones.
- El hecho de comparar un bloque con una imagen deformada de la primera provoca que las zonas de detección se vean alteradas. Cuanto más pequeño sea el tamaño del bloque, más fácil será detectar la deformación producida, mientras que un mayor tamaño de bloque aumentará las limitaciones del algoritmo frente a las deformaciones no rígidas.

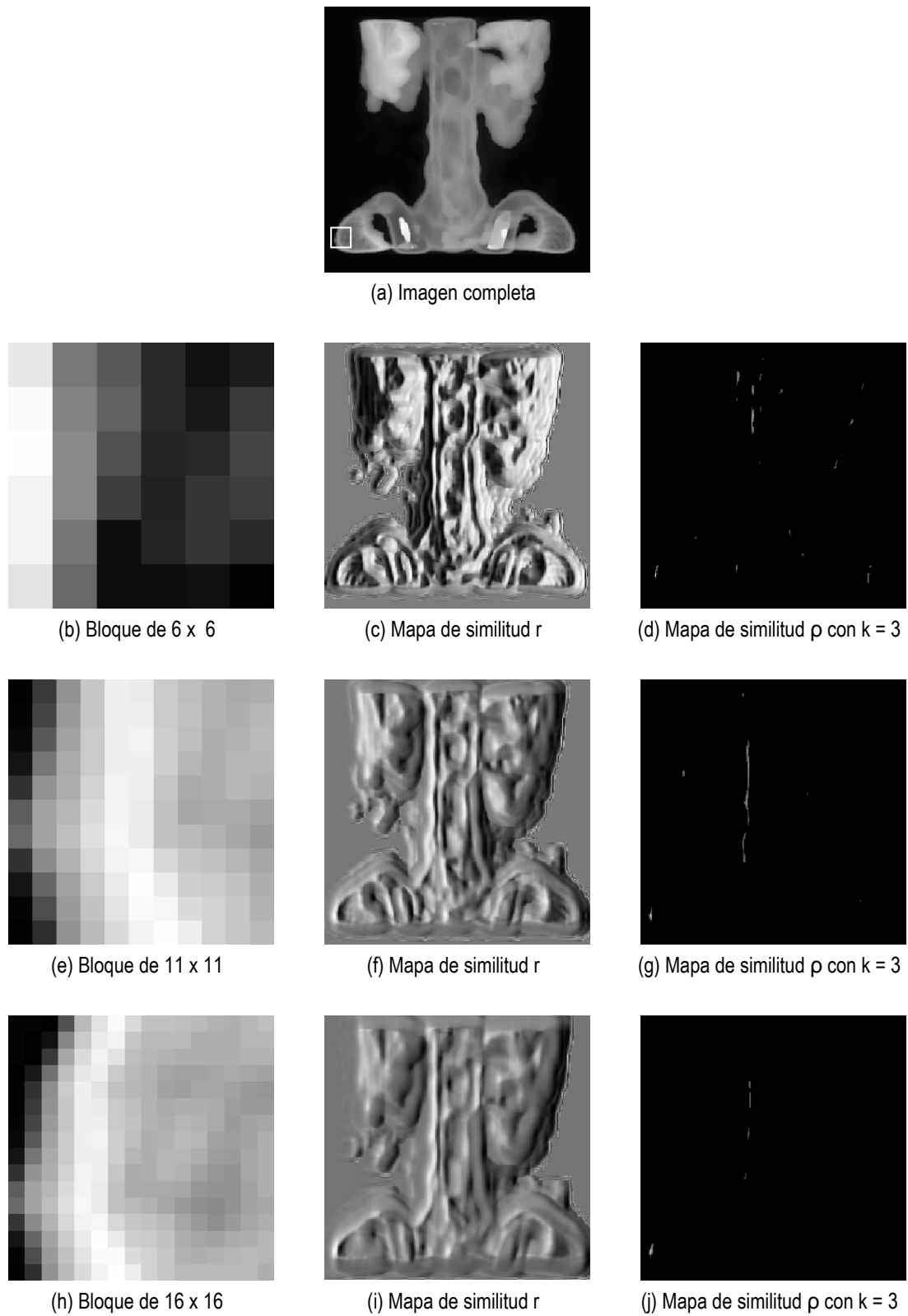


Figura 3.3. Ejemplo de la repercusión del tamaño del bloque sobre el mapa de similitud en una imagen sin rotación.

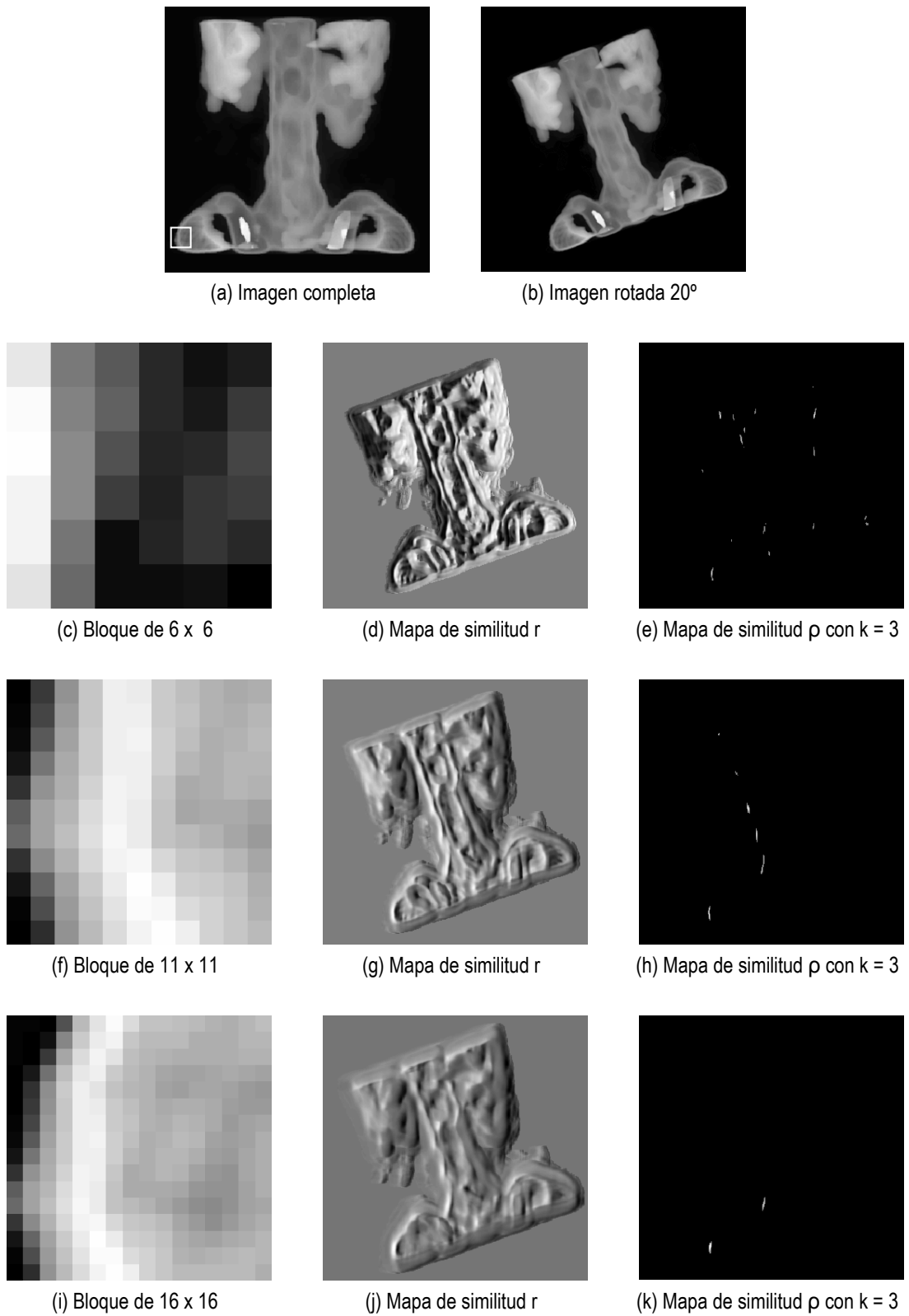


Figura 3.4. Ejemplo de la repercusión del tamaño del bloque sobre el mapa de similitud en una imagen con rotación.

3.4. EL COEFICIENTE DE INFORMACIÓN MUTUA

En teoría de probabilidad y de la información, la información mutua (MI) de dos variables aleatorias es una cantidad que mide la dependencia estadística de las dos variables (o conjuntos de datos). La unidad más común de medida de la información mutua es el *bit*, cuando para ello se hace uso de logaritmos de base 2.

En gran cantidad de aplicaciones, se trata de encontrar el máximo de la información mutua, lo que implica incrementar la dependencia entre dos variables. A menudo, maximizar el coeficiente de información mutua es equivalente a minimizar la entropía condicional. Algunos ejemplos de las aplicaciones de este tipo de medida de similitud incluyen el cálculo de la capacidad del canal en teoría de la información, selección de características en aprendizaje máquina o procesamiento de imágenes médicas para registro de imagen. En este último caso, la imagen es deformada hasta que la información mutua entre la imagen referencia y la registrada se maximiza.

Los métodos basados en la información mutua son los métodos basados en intensidad de píxel más recientes, siendo también los más populares para el registro multimodo. El registro de imágenes multimodo es un problema complejo que se encuentra muy habitualmente en el registro de imágenes médicas, por ejemplo para la comparación simultánea de imágenes anatómicas y funcionales que permiten un diagnóstico más fiable. La información mutua no asume una relación lineal entre los valores de los píxeles de las dos imágenes, pero en su lugar asume que la ocurrencia de los valores más probables en las dos imágenes se maximiza en el proceso de registro. Como podrá apreciarse más adelante, el hecho de que la información mutua sea una medida de tipo estadístico resulta especialmente apropiado para estas aplicaciones de registro multimodo.

Los criterios de registro MI muestran que la información mutua de los valores de intensidad de los correspondientes pares de píxeles es máxima si las imágenes están geoméricamente alineadas. Ya que no se impone ningún tipo de limitaciones a la naturaleza de la relación entre las intensidades de las imágenes a ser registradas, y no se hacen presunciones sobre el contenido de las imágenes de las modalidades involucradas, el criterio de información mutua es muy general y potente. Esto nos permite su aplicación para registros completamente automatizados sin segmentación previa o preprocesado.

Si combinamos la medida de la información mutua normalizada (*Normalized Mutual Information, NMI*), con estrategias piramidales o jerárquicas (con resolución creciente), se puede conseguir una aceleración importante del proceso de registro. De forma similar a la NMI, y también proveniente de la teoría de la información, se puede emplear otra medida de similitud basada en la entropía cruzada, el denominado coeficiente de correlación de entropía (*Entropy Correlation Coefficient, ECC*).

3.4.1. Expresión matemática del NMI

Sean de nuevo dos imágenes $I(\mathbf{n})$ y $J(\mathbf{n})$ bidimensionales. Sean también $I_{\mathbf{v}}(\mathbf{n})$ y $J_{\mathbf{v}+\mathbf{x}}(\mathbf{n})$ las subimágenes extraídas de las posiciones de $I(\mathbf{n})$ y $J(\mathbf{n})$ indexadas por las matrices de índices \mathbf{v} y

$v+x$ respectivamente. El término x representa el desplazamiento respecto a los índices v , recibiendo el nombre de *matriz de desplazamientos*. Se define el *coeficiente de información mutua normalizado (NMI)* de la siguiente forma:

$$NMI(I_v, J_{v+x}) = \frac{H(I_v) + H(J_{v+x})}{H(I_v, J_{v+x})} = 1 + \frac{I(I_v, J_{v+x})}{H(I_v, J_{v+x})}, \quad (3.14)$$

donde $H(\cdot)$ representa la entropía de la imagen escogida, que es posible calcular a partir del histograma de niveles de intensidad en la imagen, es decir, de la probabilidad de aparición de cada uno de los 256 niveles de gris, p_i e $I(\cdot)$ es la información mutua. Ambas están dadas por las expresiones

$$H(x) = -\sum_{i=0}^{255} p_i(x) \log p_i(x), \quad (3.15)$$

$$H(x, y) = -\sum_{i=0}^{255} \sum_{j=0}^{255} p_{ij}(x, y) \log p_{ij}(x, y), \quad (3.16)$$

$$I(x, y) = \sum_{i=0}^{255} \sum_{j=0}^{255} p_{ij}(x, y) \log \frac{p_{ij}(x, y)}{p_i(x) p_j(y)}, \quad (3.17)$$

$$I(x, y) = H(x) - H(x|y) = H(y) - H(y|x) = H(x) + H(y) - H(x, y), \quad (3.18)$$

donde $H(x)$ y $H(y)$ son las entropías marginales, $H(x|y)$ y $H(y|x)$ son las entropías condicionales, $H(x, y)$ es la entropía conjunta de x e y , $p(x, y)$ es la función de distribución de probabilidad conjunta de x e y , y $p(x)$ y $p(y)$ son las funciones de distribución de probabilidad marginal de x e y respectivamente.

Para obtener los valores de las distribuciones de probabilidad se suelen utilizar los histogramas, una herramienta que nos proporciona una aproximación de dichas funciones estadísticas. Un histograma es simplemente una matriz que cuenta el número de observaciones que caen en cada uno de los intervalos definidos en ella, denominados *bins*. Estos intervalos se especifican de modo que no se solapen entre sí, y deben ser adyacentes. Dependiendo del conjunto de datos del que dispongamos, el histograma a calcular será de una dimensión o de dos dimensiones (aunque más dimensiones son posibles). En el primer caso se calcula el número de coincidencias de los datos de una variable dentro de un conjunto de *bins* determinado, mientras que en el segundo se evalúan dichas coincidencias cuando se dispone simultáneamente de dos

variables distintas, resultando una matriz rectangular cuyo tamaño viene determinado por el tamaño de los *bins* escogidos para cada uno de los conjuntos de datos presentes.

La Figura 3.5 muestra el resultado de distintos histogramas cuando disponemos de uno o de dos conjuntos de datos para su cálculo. Para la representación se han utilizado tamaños de *bin* de 32 en cada eje.

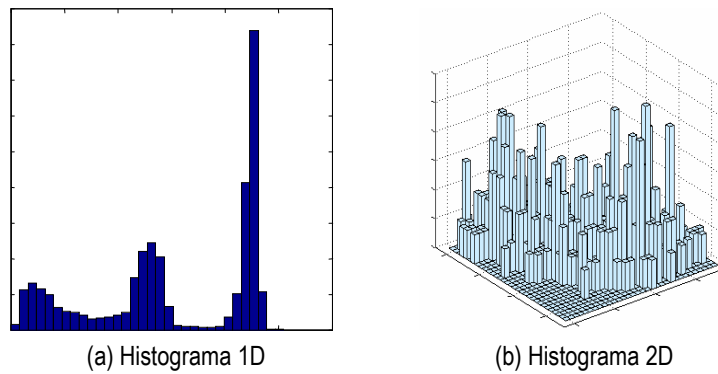


Figura 3.5. Ejemplos de distintos tipos de histogramas.

Intuitivamente, si la entropía $H(x)$ es tomada como la medida de la incertidumbre sobre un conjunto de datos, entonces $H(x|y)$ puede ser vista como la cantidad de incertidumbre que queda sobre x tras conocer y , y por lo tanto la primera de estas igualdades puede ser interpretada como la incertidumbre en x menos la incertidumbre en x cuando se conoce y , que es equivalente a decir que es la incertidumbre en x que se elimina al conocer y .

A partir de las expresiones (3.15), (3.16), (3.17) y (3.18) se pueden deducir las siguientes conclusiones sobre la información mutua:

1. La información mutua es no negativa $\rightarrow I(x, y) \geq 0$ (ya que $H(x) \geq H(x|y)$).
2. La información que contiene una imagen acerca de ella es su propia información $\rightarrow I(x, x) = H(x)$.
3. La información mutua es simétrica $\rightarrow I(x, y) = I(y, x)$.
4. Si $I(x, y) = 0$, las imágenes x e y son independientes entre sí.

Dado que el NMI se encuentra comprendido entre 1 y 2, generalmente se prefiere el uso del denominado coeficiente de correlación de entropía (ECC), que sí que toma valores comprendidos entre 0 y 1, y se puede calcular de la siguiente forma a partir del NMI:

$$ECC(I_v, J_{v+x}) = 2 - \frac{2}{NMI(I_v, J_{v+x})}. \quad (3.19)$$

3.4.2. Algoritmo de implementación

Para la implementación del algoritmo que calcula la información mutua normalizada de un par de conjuntos de datos o imágenes se parte de las expresiones (3.14) y (3.19), de modo que el algoritmo queda:

- 1) Se calcula el histograma del primer bloque (imagen I) y se normaliza dividiendo por el número de muestras procesadas, lo que da la probabilidad de cada uno de los niveles de gris. Este valor se asigna a la variable P_i .
- 2) Se calcula la entropía del bloque I mediante la expresión (3.15), obteniendo la variable H_I .
- 3) Se coge el segundo bloque (imagen J), que debe ser de las mismas dimensiones que el primero, y se calcula la probabilidad de ocurrencia de cada nivel de gris tal y como se hizo en 1), asignándolo todo a la variable P_j .
- 4) Obtenemos el valor de la entropía H_J a partir de la fórmula (3.15) a partir de la matriz de probabilidades P_j .
- 5) Se crea una nueva matriz combinación de los bloques I y J , asignando en la primera columna el valor de los niveles de gris de un bloque y en la segunda los de otro. Esta nueva variable se denomina IJ .
- 6) Se calcula la probabilidad de ocurrencia de cada nivel de gris de IJ a través del histograma normalizado 2D, obteniendo la variable P_{ij} .
- 7) La entropía de la matriz combinación de I y J se calcula a partir de P_{ij} usando ahora la expresión (3.16). Este valor se asigna a H_{IJ} .
- 8) Obtenemos el NMI mediante el cociente
$$NMI = \frac{H_I + H_J}{H_{IJ}} .$$
- 9) El ECC se obtiene a partir del valor anterior a través de la relación
$$ECC = 2 - \frac{2}{NMI} .$$

Como vemos, la manera de calcular los mapas de similitud mediante este algoritmo es la misma que la descrita en el apartado 3.3.1, de tal manera que se va calculando la similitud iterativamente para cada posición. Según la implementación efectuada, los aspectos más críticos a la hora de realizar el cálculo del NMI y el ECC son dos: en primer lugar el tamaño de los bloques para cada cálculo debe ser el mismo; y por otro lado hay que resolver el caso en que la entropía conjunta sea igual a cero.

Un aspecto muy importante a tener en cuenta durante la ejecución ocurre cuando la entropía conjunta de los bloques I y J es nula, lo que provoca una división por cero. Para remediarlo, se ha optado por asignar un valor de ECC igual a 0 que indique que no hay similitud

siempre que la entropía conjunta H_D sea nula. Para evitar calcular los logaritmos de valores nulos, otra medida a tomar será la de coger sólo las probabilidades cuyo valor sea distinto de cero.

3.4.3. Limitaciones

La principal desventaja de los métodos que utilizan información estadística como la entropía es principalmente el aumento considerable del coste computacional, al compararlos con medidas como la correlación o la diferencia cuadrática media. Otro de los principales peligros o escollos de la información mutua es el de tratar los valores de intensidad de la imagen de una manera *puramente cualitativa*, sin considerar ninguna noción de proximidad en el espacio de intensidad.

Debido a que los valores de intensidad de las imágenes médicas son muy distintos, los *bins* de los histogramas son a menudo muy grandes. Para calcular el histograma más eficientemente, los valores de intensidad de la imagen son reescalados antes del procedimiento de registro. Los valores de nivel de gris más utilizados son 256 o 64 para ambas imágenes, aunque los resultados se vuelven peores con la reducción de estos niveles de gris.

3.4.4. Ejemplos

A continuación, vamos a observar el funcionamiento de la medida de similitud ECC para algunas variaciones, a partir de las cuales podemos analizar su utilidad ante desplazamiento rígido y ante rotaciones. Un análisis más profundo de ésta y otras medidas de similitud lo encontraremos en el apartado 3.7.

En la Figura 3.6 podemos ver el resultado de aplicar la medida de similitud ECC en una imagen de la que se ha extraído un fragmento determinado. Este fragmento queda indicado por un cuadro blanco en la Figura 3.6(a). En la Figura 3.6(b) se muestra un bloque de tamaño 6 x 6 que se aplica a la imagen de la Figura 3.6(a) para obtener el mapa de similitud representado en la Figura 3.6(c). También se ha calculado el mismo mapa de similitud (Figura 3.6(d)) cuando hacemos uso de la diferencia cuadrática media y de una umbralización de 0.3, ya vista en el apartado 3.3.1. Ahora esa expresión (3.10) nos ha quedado de la siguiente manera:

$$\rho = \frac{ECC}{1 + k \cdot d}, \quad (3.20)$$

donde k de nuevo es el parámetro de ponderado de la diferencia cuadrática d .

En la Figura 3.6(e), Figura 3.6(f) y Figura 3.6(g) podemos ver estos mismos mapas de similitud para el caso de un tamaño de bloque mayor. Todo ello nos permite reafirmar, tal y como vimos para el caso de la correlación cruzada, que un tamaño de bloque mayor resulta más selectivo, pero tiene el inconveniente de tener peores prestaciones ante imágenes con deformaciones complejas, es decir, para casos de deformaciones no rígidas. También vemos que la

adición de la diferencia cuadrática y una umbralización del mapa nos permiten obtener más claramente los lugares en donde se produce una mejor estima de la posición del bloque. Esto, como ya veremos más adelante, nos ayudará a distinguir el tipo de modelo (punto, curva o mixto) que describe el mapa de similitud.

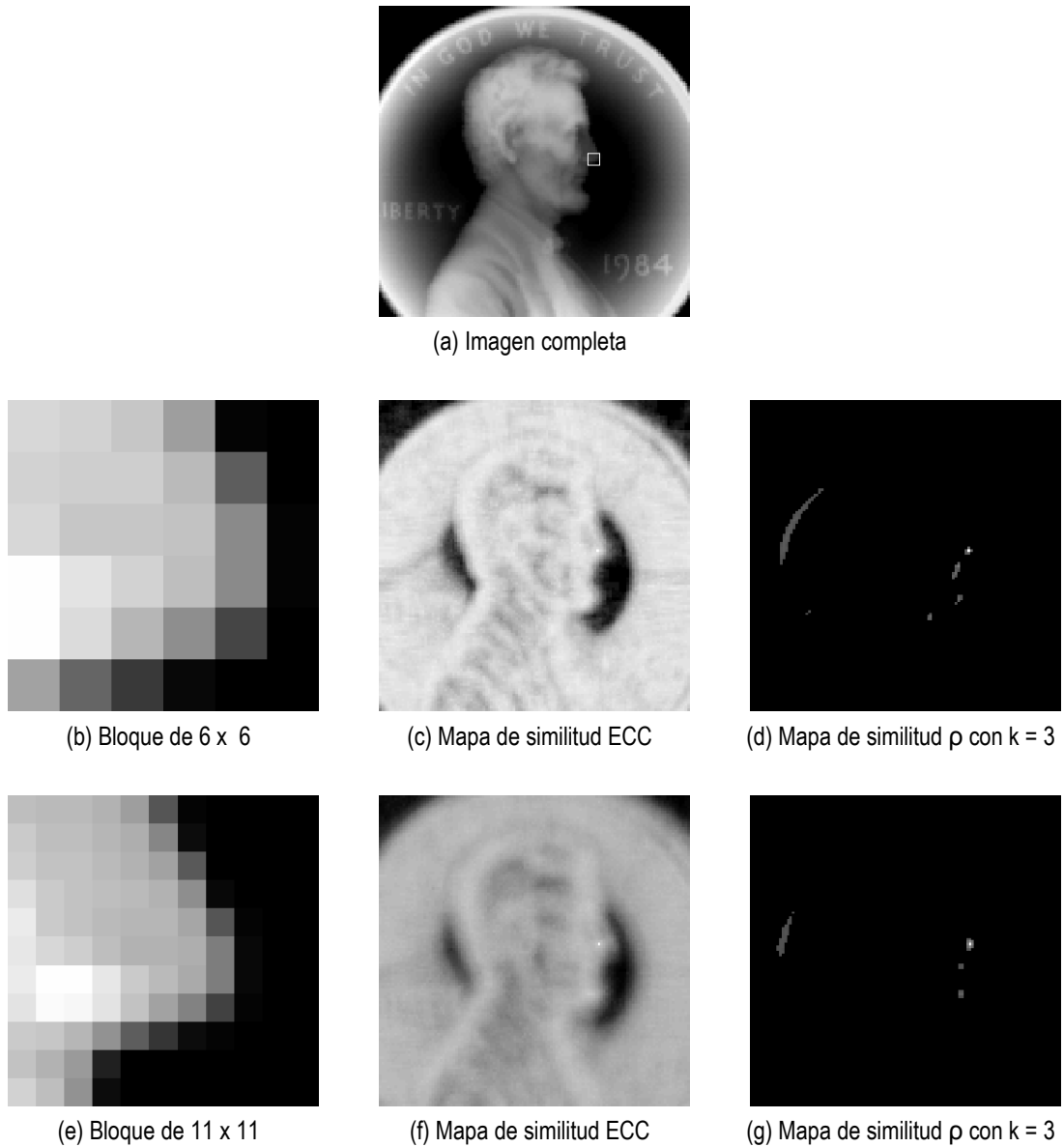


Figura 3.6. Ejemplo de mapas de similitud ECC ante distintos tamaños de bloque.

En la Figura 3.7 se muestra el resultado de los mapas de similitud ante distintas rotaciones de la imagen, para lo que se emplea un bloque de 11 x 11 píxeles. En la Figura 3.7(c) se ve la imagen tras una rotación de 10°, lo que proporciona al aplicar la medida de similitud con el bloque dispuesto en la Figura 3.7(b) los mapas mostrados en la Figura 3.7(d) y Figura 3.7(e). En la Figura

3.7(f), Figura 3.7(g) y Figura 3.7(h) observamos estos mismos resultados para una rotación de 20°.

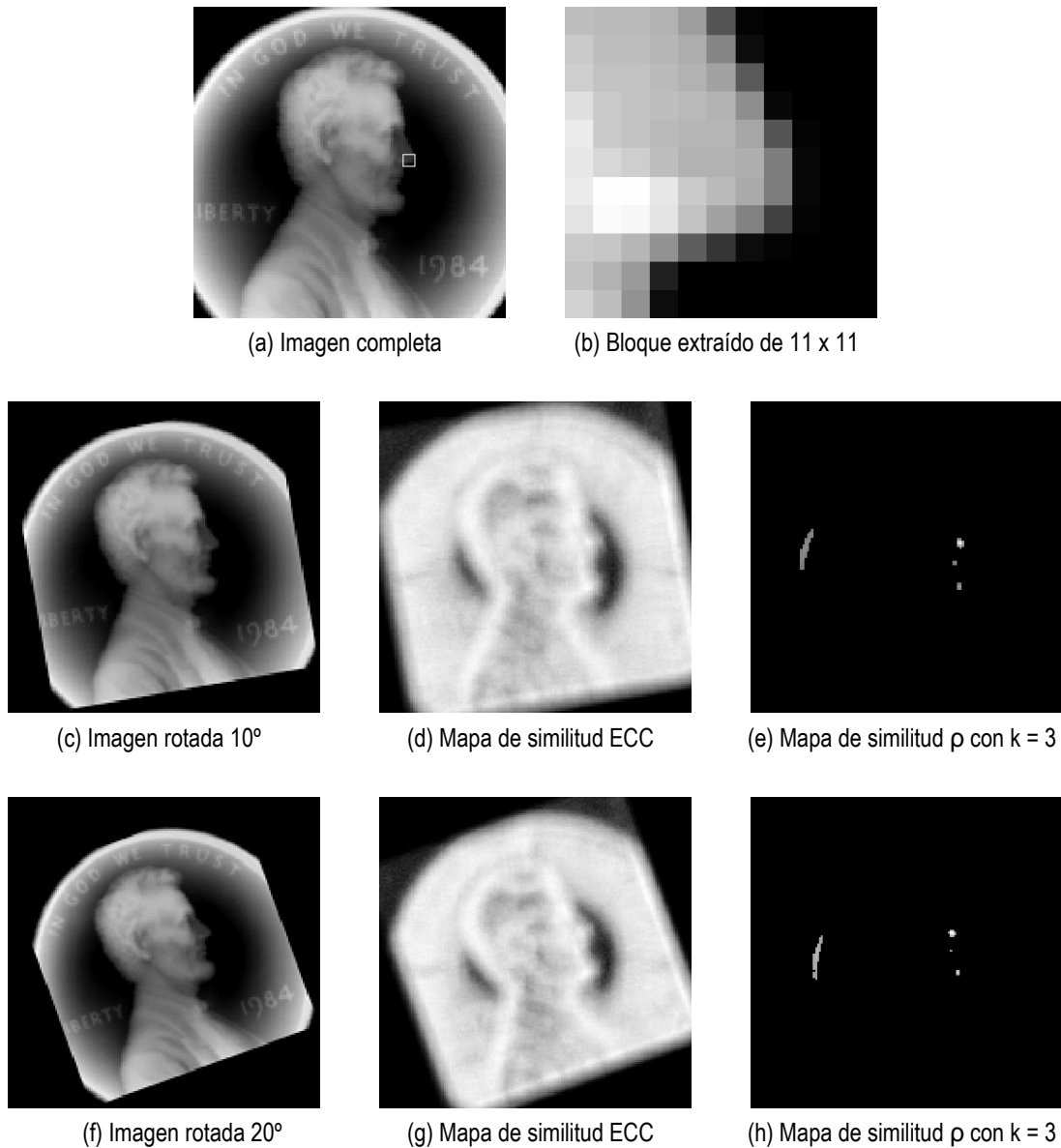


Figura 3.7. Ejemplo de mapas de similitud ECC ante distintas rotaciones.

Los resultados obtenidos reflejan que incluso para rotaciones moderadas de la imagen original, como puede ser una rotación de 20°, el mapa de similitud indica para fracciones de bloque con ciertas características bien definidas (tal y como puede ser la nariz escogida en el bloque de la Figura 3.7(b)) la posición aproximada desde donde se ha desplazado ese determinado bloque. Para rotaciones o deformaciones mucho más acentuadas, la medida de similitud comienza a perder efectividad, hasta el punto de llegar incluso a producir errores críticos de detección en el

momento de llevar a cabo el registro de dos imágenes con este tipo de deformación. Como ya se ha comentado anteriormente, un procedimiento más exhaustivo de análisis tendrá lugar en el último punto del reciente capítulo.

3.5. EL RATIO DE CORRELACIÓN

En los últimos años, las nuevas medidas de similitud basadas en píxeles han permitido un progreso importante en el registro de imágenes multimodo, notablemente debido al incremento de las medidas de similitud basadas en la teoría de la información. Su amplio éxito ha conducido al abandono progresivo de medidas que usan la estadística estándar de imagen, como la media y la varianza. Hasta hace poco, tales medidas han estado esencialmente basadas en heurísticos. En este punto tratamos el *ratio de correlación (CR)*, una medida basada en la teoría de probabilidad, y que se define como una medida de la relación entre la dispersión estadística dentro de categorías individuales, y la dispersión en la totalidad de la muestra. Para la definición de este concepto asumiremos la dependencia funcional entre las intensidades de las imágenes a registrar.

Como ya se ha visto en el capítulo 2, el principio general del registro basado en la intensidad de los píxeles consiste en cuantificar la calidad de emparejamiento con respecto a la medida de similitud del solapamiento de los píxeles de las imágenes. Ya que esta medida se asume máxima cuando las imágenes están correctamente alineadas, estas aproximaciones se implementan usando un esquema de optimización.

Las medidas simples de correlación han sido usadas extensamente en el campo de las imágenes médicas, aunque básicamente asumen una relación lineal entre las intensidades de las imágenes. Tal hipótesis es generalmente demasiado ingenua en el registro multimodo, donde medidas como la información mutua, el ratio de correlación o el criterio de Woods proporcionan un resultado más acorde a la naturaleza multimodo de las imágenes. Al contrario que el NCC clásico, el ratio de correlación es una medida de similitud de carácter estadístico y es transparente a los diferentes significados de la intensidad captada en cada imagen.

3.5.1. Expresión matemática del CR

Supongamos que tenemos dos imágenes $I(\mathbf{n})$ y $J(\mathbf{n})$ bidimensionales que han sido registradas. Si aleatoriamente seleccionamos píxeles en la región de solapamiento de estas imágenes, observaremos que los pares de intensidad que obtenemos son estadísticamente consistentes, es decir, todos los píxeles que tienen cierta intensidad i en $I(\mathbf{n})$ pueden tener intensidades agrupadas en $J(\mathbf{n})$, posiblemente muy diferentes de i . Dependiendo del tipo de imágenes, cualquier isoconjunto $I(\mathbf{n}) = i$ podría proyectarse a uno o varios de dichos grupos (*clusters*). En el caso de un solo *cluster* por isoconjunto, la intensidad en $J(\mathbf{n})$ podría predecirse aproximadamente a partir de la intensidad en $I(\mathbf{n})$, aplicando una función simple. Este razonamiento es válido sólo si las imágenes están correctamente registradas. Por lo tanto, podríamos usar el grado de dependencia funcional entre $I(\mathbf{n})$ y $J(\mathbf{n})$ como criterio de emparejamiento.

Para llevar a cabo el cálculo del ratio de correlación entre dos imágenes, tenemos que ser capaces de definir las como variables aleatorias, es decir, determinar sus funciones de densidad de probabilidad (fdp) marginal y conjunta. Una técnica común consiste en normalizar el histograma 2D del par de imágenes. Entonces, las imágenes pueden ser vistas como variables aleatorias discretas. Escogiendo esta aproximación discreta, no hay necesidad de manipular explícitamente el histograma 2D de las imágenes, tal y como ocurría para la información mutua. En su lugar, el ratio de correlación puede calcularse recursivamente mediante la acumulación de operaciones locales.

Sean $I_v(\mathbf{n})$ y $J_{v+x}(\mathbf{n})$ las subimágenes extraídas de las posiciones de dos imágenes sin registrar $I(\mathbf{n})$ y $J(\mathbf{n})$, indexadas por las matrices de índices \mathbf{v} y $\mathbf{v+x}$ respectivamente. Podemos representar la región de solapamiento de ambos subconjuntos de datos como Ω , y a $N = \text{Card}(\Omega)$ como el número total de píxeles que esta región contiene. Consideramos los isoconjuntos de $I_v(\mathbf{n})$, $\Omega_i = \{\omega \in \Omega, I_v(\omega) = i\}$ y sus cardinales $N_i = \text{Card}(\Omega_i)$. Los momentos totales y condicionales (media y varianza) de $J_{v+x}(\mathbf{n})$ son entonces

$$m = \frac{1}{N} \sum_{\omega \in \Omega} J_{v+x}(\omega), \quad (3.21)$$

$$m_i = \frac{1}{N_i} \sum_{\omega \in \Omega_i} J_{v+x}(\omega), \quad (3.22)$$

$$\sigma^2 = \frac{1}{N} \sum_{\omega \in \Omega} J_{v+x}(\omega)^2 - m^2, \quad (3.23)$$

$$\sigma_i^2 = \frac{1}{N_i} \sum_{\omega \in \Omega_i} J_{v+x}(\omega)^2 - m_i^2. \quad (3.24)$$

A partir de las relaciones anteriores es posible obtener una expresión simple del ratio de correlación, dada por

$$\eta(J_{v+x}|I_v) = 1 - \frac{1}{N\sigma^2} \sum_i N_i \sigma_i^2. \quad (3.25)$$

La medida anterior no requiere el cálculo del histograma 2D de las imágenes. Esta es la gran diferencia respecto a la información mutua. Los algoritmos clásicos para el cálculo de la información mutua tienen una complejidad de orden $n_I n_J$, con n_I y n_J siendo el número de niveles de intensidad en las imágenes I_v y J_{v+x} respectivamente. El cómputo del ratio de correlación anterior tiene solamente una complejidad de orden n_I , y es independiente de n_J .

El ratio de correlación, como vemos, mide la dependencia funcional entre los subconjuntos de imágenes $I_v(\mathbf{n})$ y $J_{v+x}(\mathbf{n})$. Toma valores entre 0 (sin dependencia funcional) y 1 (dependencia puramente determinista). Por el uso de un cociente en lugar de una diferencia, $\eta(J_{v+x}|I_v)$ es invariante a cambios multiplicativos en $J_{v+x}(\mathbf{n})$. También hay que apuntar que el ratio de correlación es asimétrico por naturaleza ya que las dos variables fundamentalmente no juegan el mismo papel en la relación funcional. En general, $\eta(J_{v+x}|I_v) \neq \eta(I_v|J_{v+x})$.

3.5.2. Algoritmo de implementación

El procedimiento para el cálculo del ratio de correlación consiste en calcular cada uno de los términos presentes en (3.25), para lo que se divide el cálculo en varios pasos. El hecho de no hacer uso del histograma 2D explícitamente nos permite ahorrar gran cantidad de recursos de cálculo. A continuación se enumeran los pasos a seguir:

- 1) Se calcula el número de elementos de la región de solapamiento (número de píxeles) sumando el número de elementos de cualquiera de las dos imágenes (\mathbf{I} o \mathbf{J}), por tener ambas el mismo tamaño. El valor calculado se asigna a la variable \mathbf{N} .
- 2) Calculamos la varianza del bloque \mathbf{J} y la almacenamos en una variable llamada $\mathbf{var_j}$.
- 3) Pasamos a calcular la varianza de \mathbf{J} condicionada a \mathbf{I} mediante un pequeño algoritmo:
 - a. Inicializamos el valor de la variable llamada **sumatorio** a 0.
 - b. Se recorre iterativamente el bloque \mathbf{I} mediante los niveles de gris presentes en él, escogiendo en cada iteración las posiciones de los píxeles con una misma intensidad.
 - c. Para cada una de las iteraciones, se calcula el número de coincidencias del nivel de gris dado por el número de iteración (iteración 1, nivel de gris 0; iteración 2, nivel de gris 1; etc.) y se almacena en \mathbf{N}_i .
 - d. Si \mathbf{N}_i no es cero, lo que implica que al menos hay un píxel con dicho nivel de gris, se calcula la varianza condicional como la varianza de los píxeles del bloque \mathbf{J} en las posiciones dadas por los píxeles de misma intensidad detectados en la iteración donde esté. Este valor se almacena en $\mathbf{var_ji}$.
 - e. Se calcula el producto $\mathbf{N}_i \cdot \mathbf{var_ji}$ y se suma a la variable **sumatorio**.
 - f. Si se han recorrido todos los niveles de gris, se termina el algoritmo. Si no, volvemos al punto b.
- 4) Se calcula el valor del ratio de correlación a partir de la fórmula
$$CR = 1 - \frac{I}{N \cdot var_j} \cdot \text{sumatorio}$$
, derivada de (3.25).

Un aspecto crucial de este algoritmo es el caso en el que se produce un valor de la varianza del bloque J igual a 0. En este caso, y al ser este factor el divisor en la expresión final, se asigna al valor del CR un valor final de 0, de modo que indique una similitud nula.

Otra de las cosas que hay que tener presentes es que para calcular el ratio de correlación es necesario que el tamaño de las subimágenes empleadas sea el mismo. Para ello, se puede hacer uso de un algoritmo iterativo que vaya recorriendo cada una de las posibles posiciones de la imagen más grande, tal y como se vio para el caso de la información mutua. Al recorrer todas las coordenadas de la imagen, obtendremos el mapa de similitud completo entre las subimágenes presentes en el bloque I y J .

En el caso de que se haga uso de máscaras de paso o de búsqueda, sólo procesaremos aquellos píxeles que deje pasar la máscara de paso, mientras que en la imagen mayor sólo buscaremos donde nos permita la máscara de búsqueda.

3.5.3. Limitaciones

En la derivación del ratio de correlación, hemos tomado como hipótesis la dependencia funcional entre las intensidades de las imágenes. Aunque tal hipótesis no es tan general como se desea, nos permite modelar más fácilmente el proceso a seguir. Cuando las imágenes o conjuntos de datos a ser comparados por el método que nos proporciona el ratio de correlación son adquiridos por procedimientos muy diferentes, la situación se vuelve compleja. Entonces hay dos asuntos a tratar: primero, la medida de similitud debe ser capaz de determinar la correspondencia entre imágenes de distinto tipo; y segundo, la medida tiene que ser sensible a las diferencias locales de la imagen. El problema se hace difícil porque las dos propiedades deseadas para nuestra medida de similitud son contrapuestas.

Un aspecto muy a tener en cuenta a la hora de calcular la similitud de un par de imágenes mediante el ratio de correlación, es la propiedad de asimetría que ésta presenta. Así, encontraremos que comparar I con J no proporcionará el mismo resultado que comparar J con I . Habrá por tanto que decidir correctamente el orden en que estas imágenes serán comparadas, mediante un estudio de tipo experimental por ejemplo.

3.5.4. Ejemplos

En este apartado vamos a observar el comportamiento del ratio de correlación para algunos casos concretos de imágenes reales. En primer lugar analizaremos la actuación ante distintos tamaños de bloque, mientras que en un segundo ejemplo veremos el resultado del mapa de similitud ante rotaciones simples sin deformación. Para el cálculo de estos mapas de similitud con el uso de la diferencia, de nuevo utilizamos la fórmula vista en (3.10), que ahora nos queda

$$\rho = \frac{\eta}{I + k \cdot d}. \quad (3.26)$$

En la Figura 3.8 se puede ver el resultado de aplicar una fracción de la imagen, mostrada en la Figura 3.8(a), a la totalidad de la imagen. Estos fragmentos de distintos tamaños se cogen de la parte procedente del ojo de la izquierda de la imagen, y están representados para distintos tamaños en la Figura 3.8(b) y Figura 3.8(e). Al lado de dichas figuras podemos observar el resultado del mapa de similitud del ratio sin umbralizar y el ratio umbralizado a 0.3 usando la diferencia cuadrática (imágenes del centro y derecha). A partir de los resultados se aprecia una mayor suavidad del mapa para tamaños del bloque mayores, mientras que el número de puntos tras la umbralización ha disminuido, reduciendo el problema de la correspondencia. Sin embargo, esto puede dejar un número de zonas detectadas insuficientes para nuestro propósito.

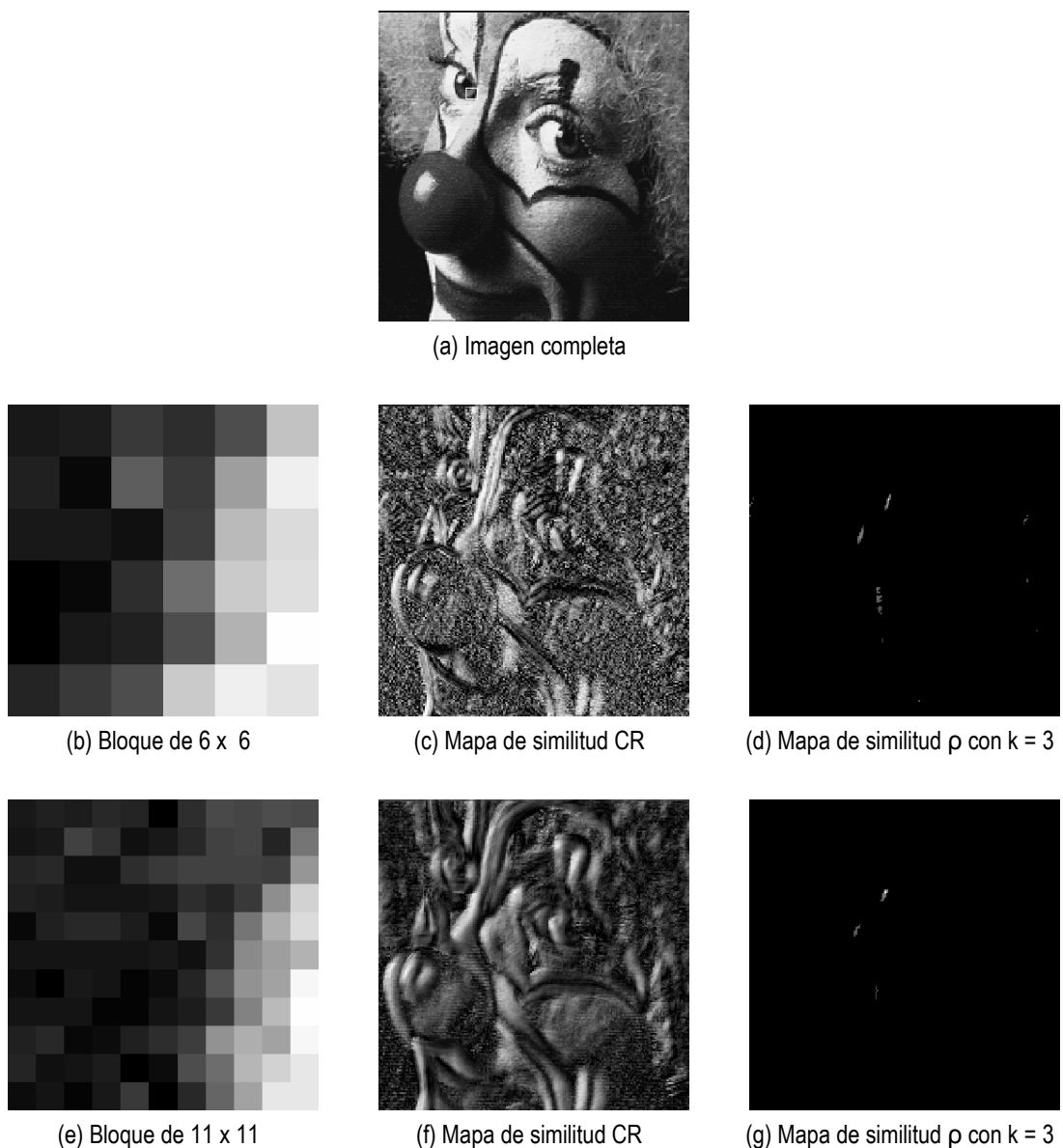


Figura 3.8. Ejemplo de mapas de similitud CR ante distintos tamaños de bloque.

En la Figura 3.9 podemos ver la sensibilidad de esta medida de similitud frente a las rotaciones. El hecho de que se detecten ligeramente menos zonas de alta correspondencia se debe al propio algoritmo de cálculo del ratio de correlación, ya que éste hace uso de las coordenadas del primer bloque que tienen un mismo nivel de gris, y éstas coordenadas son las tomadas en el segundo bloque de datos para analizar la similitud de los conjuntos comparados. La consecuencia de rotar los bloques unos determinados grados conlleva que al tomar los píxeles de la segunda imagen, éstos no tengan una varianza constante como deberían tener si no hubiesen sido rotados. En la Figura 3.9(c), Figura 3.9(d) y Figura 3.9(e) se puede ver el caso para una rotación de 10° , mientras que una rotación de 20° se recoge en la Figura 3.9(f), Figura 3.9(g) y Figura 3.9(h).

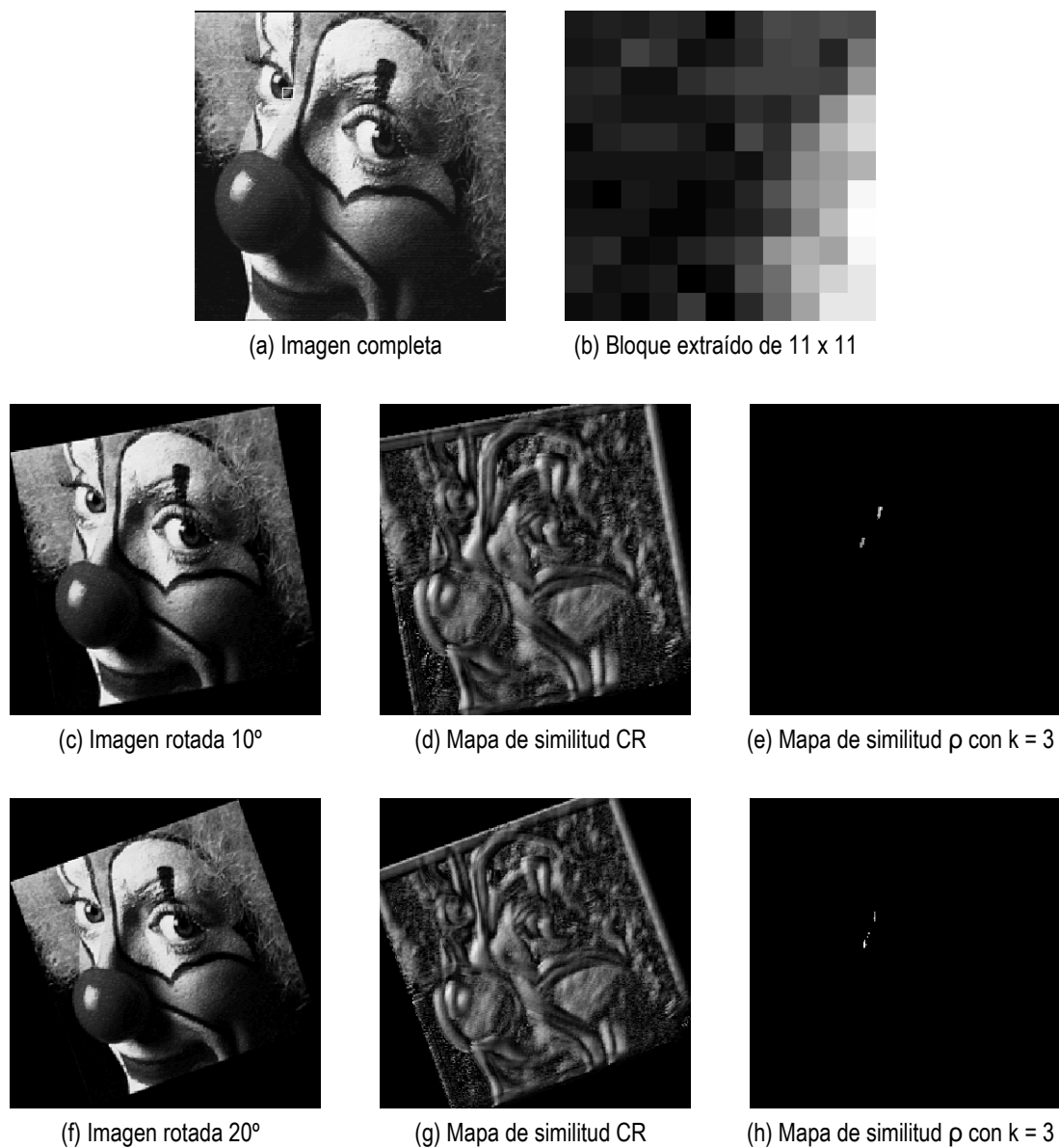


Figura 3.9. Ejemplo de mapas de similitud CR ante distintas rotaciones.

3.6. EL CRITERIO DE WOODS

El método heurístico derivado por Woods y otros, fue originalmente diseñado para registro multimodo PET/MR, pero también ha sido usado con otras modalidades. Este método es muy similar al del ratio de correlación y opera directamente sobre los valores del nivel de gris de la imagen sin necesidad de una segmentación previa. Se trata de un método estadístico que hace uso implícitamente del histograma 2D, lo que nos permite una mejora en cuanto a tiempo de procesado frente a la información mutua.

El criterio de Woods (WC) se basa en premisas relativas a los niveles de gris de las diferentes modalidades, lo que a menudo reduce su aplicabilidad a algunas combinaciones multimodo muy específicas. Woods obtuvo un ratio para usar la uniformidad de los valores de intensidad de las imágenes como una medida de similitud. La asunción de Woods fue que si dos imágenes están alineadas de forma precisa, entonces el valor de cualquier píxel en una imagen está relacionado con el valor del píxel correspondiente en la otra imagen por un simple factor multiplicativo R . Si las imágenes están desalineadas, este factor multiplicativo no es constante, sino que varía de un píxel a otro a través de la imagen. El algoritmo de registro mediante esta técnica básicamente mueve la imagen objetivo en relación a la de referencia hasta que la variación píxel a píxel se minimiza.

A lo largo de los últimos años se han presentado numerosas versiones del criterio de Woods para cada caso concreto: registro PET/PET, PET/MR, MRI/fMRI, etc. Cada una de ellas introduce cambios respecto de las otras para ajustarse a las necesidades de cada problema en particular.

3.6.1. Expresión matemática del WC

Como se ha comentado, el criterio de Woods cambia en función de la aplicación. Aquí se presenta la formulación original de esta medida de similitud para el caso de registro multimodo PET/MR.

Sean dos imágenes dadas por $I(\mathbf{n})$ y $J(\mathbf{n})$, siendo $I(\mathbf{n})$ la imagen de referencia y $J(\mathbf{n})$ la imagen objetivo o de test. Sean también $I_{\mathbf{v}}(\mathbf{n})$ y $J_{\mathbf{v}+\mathbf{x}}(\mathbf{n})$ los subconjuntos de datos extraídos de las posiciones de $I(\mathbf{n})$ y $J(\mathbf{n})$ indexadas por las matrices de índices \mathbf{v} y $\mathbf{v}+\mathbf{x}$ respectivamente, donde el término \mathbf{x} representa el desplazamiento respecto a los índices \mathbf{v} , y recibe el nombre de *matriz de desplazamientos*. De acuerdo con la notación introducida en la sección 3.5.1, el criterio de Woods se puede escribir de la siguiente forma:

$$W(J_{\mathbf{v}+\mathbf{x}}|I_{\mathbf{v}}) = \frac{I}{N} \sum_i N_i \frac{\sigma_i}{m_i}, \quad (3.27)$$

donde la notación $W(J_{v+x} | I_v)$ se utiliza para enfatizar el hecho de que el criterio de Woods es asimétrico, tal y como ocurre con el ratio de correlación. En la expresión anterior, un máximo de similitud se presenta con un valor de $W(J_{v+x} | I_v) = 0$, por lo que a veces es interesante hacer que el resultado se encuentre entre 0 y 1, ocurriendo el 0 cuando no hay dependencia entre las imágenes y el 1 cuando las imágenes están completamente relacionadas. Para ello se puede reescribir (3.27) como

$$W(J_{v+x} | I_v) = 1 - \frac{1}{N} \sum_i N_i \frac{\sigma_i}{m_i}. \quad (3.28)$$

Aunque el ratio de correlación (3.25) y el criterio de Woods son diferentes, expresan la misma idea básica. Así, podemos identificar dos diferencias: primero, el ratio de correlación suma varianzas σ_i^2 , mientras que el criterio de Woods suma desviaciones típicas normalizadas, $\frac{\sigma_i}{m_i}$; segundo, la propiedad de invarianza multiplicativa se logra en el ratio de correlación mediante la división global por σ^2 , mientras que en el criterio de Woods cada término de la suma se divide por una media condicional m_i .

Como vemos, para valor de píxel de la imagen de referencia, se calculan los valores de la media y la desviación típica de los píxeles de la imagen objetivo. La desviación típica normalizada se calcula dividiendo la desviación típica por la media correspondiente. La media ponderada de la desviación típica es entonces proporcional al número total de píxeles que tienen un determinado valor de intensidad. La función W será la función de coste a ser minimizada. Éste es un procedimiento razonable, ya que si la desviación típica es pequeña quiere decir que hay sólo una pequeña variación de los niveles de gris, y esto ocurrirá cuando ambas imágenes estén espacialmente alineadas.

3.6.2. Algoritmo de implementación

El algoritmo de implementación es muy similar al visto en el apartado 3.5.2 para el caso del ratio de correlación. En este caso también se van cogiendo las coordenadas de los niveles de gris iguales, y se van calculando las variables necesarias para completar la fórmula vista en (3.28). Los pasos del algoritmo se muestran a continuación:

- 1) En primer lugar se calcula el número de píxeles que corresponden a la zona de solapamiento de ambos conjuntos de datos. Al ser ambos conjuntos del mismo tamaño, basta con tomar el número de elementos de uno de ellos y asignárselo a una variable que llamaremos N .
- 2) Se calcula la desviación típica normalizada del segundo bloque (imagen J) condicionada al primer bloque (imagen I) mediante el algoritmo:

- a. Se inicializa el valor de la variable llamada **sumatorio** a 0.
 - b. Se recorre iterativamente el bloque **I** mediante los niveles de gris presentes en él, escogiendo en cada iteración las posiciones de los píxeles con una misma intensidad.
 - c. Para la iteración actual, se calcula el número de coincidencias del nivel de gris dado por el número de iteración (iteración 1, nivel de gris 0; iteración 2, nivel de gris 1; etc.) y se almacena en **N_i**.
 - d. Si **N_i** no es cero, lo que implica que al menos hay un píxel con dicho nivel de gris, se calcula la desviación típica condicional como la desviación típica de los píxeles del bloque **J** en las posiciones dadas por los píxeles de misma intensidad detectados en la iteración actual. Este valor se almacena en **std_{ji}**.
 - e. De la misma forma que se calcula **std_{ji}**, también calculamos la media de los píxeles en dichas posiciones y la guardamos en la variable **media**.
 - f. Se calcula el producto $N_i \frac{std_{ji}}{media}$ y se suma a la variable **sumatorio**.
 - g. Si se han recorrido todos los niveles de gris, se termina el algoritmo. Si no, volvemos al punto b.
- 3) Se calcula el valor del criterio de Woods a partir de la fórmula

$$W = 1 - \frac{I}{N} * sumatorio .$$

Los principales problemas a tener en cuenta a la hora de implementar este algoritmo se encuentran al evaluar bloques de datos cuyo nivel de gris es constante para la segunda imagen (**J**). En estos casos, la desviación típica normalizada siempre será igual a cero, produciendo un valor de similitud máximo incluso aunque la primera imagen (**I**) sea completamente diferente. Para solucionar esto, y aun en perjuicio del caso en el que ambos bloques de datos sean constantes, se hace necesario asignar al resultado de la medida de similitud un 0, siempre y cuando la imagen **J** sea de valor constante.

Otro caso a tener muy en cuenta se produce cuando la media de la imagen **J** es igual a cero, lo que provoca una indeterminación en la desviación típica normalizada. Para estos casos también se puede adoptar la solución de asignar un cero al valor de esta medida normalizada, de modo que sea tomada como una zona de baja similitud.

3.6.3. Limitaciones

Hoy en día, está generalmente admitido que el método de Woods es el mejor para algunos casos específicos de registro multimodo. De algún modo, los resultados corroboran esta

observación, sugiriendo que tener en cuenta intensidades cercanas en las imágenes (PET por ejemplo) podría ser crucial para el buen resultado del registro. Sin embargo, la información mutua se muestra mejor adaptada a imágenes de bajo ruido, y las discrepancias encontradas entre el ratio de correlación y el criterio de Woods a veces llegan a ser sorprendentes a pesar de que estas medidas están formalmente basadas en consideraciones similares. De hecho el ratio de correlación nos proporciona a parte de una justificación teórica del criterio de Woods, una serie de mejoras prácticas perceptibles.

Una de las limitaciones más importantes del criterio de Woods tiene que ver con la media de las imágenes tratadas. Si utilizamos imágenes con zonas de nivel de gris con media muy pequeña, el resultado de la desviación típica normalizada se irá a valores altos; si por el contrario la media de la imagen es grande, esta medida normalizada será por lo general bastante baja, produciendo valores del criterio de Woods muy cercanos a 1.

3.6.4. Ejemplos

Vamos a proceder a estudiar algunos resultados que proporciona el criterio de Woods para distintos casos procedentes de una misma imagen, de la que se extrae un bloque cuadrado de un tamaño determinado, y que se utiliza para calcular el mapa de similitud. En concreto veremos la relevancia del tamaño del bloque escogido, así de la sensibilidad de dicha medida de similitud frente a pequeñas rotaciones de la imagen.

La Figura 3.10 muestra una imagen a la que se le calculan los mapas de similitud a partir de un fragmento extraído de la zona delimitada por el cuadro blanco (Figura 3.10(a)). Para el cálculo de estos mapas se hace uso del criterio de Woods dado por (3.28) y de la diferencia cuadrática, lo que nos permite combinar estas dos medidas para formar

$$\rho = \frac{W}{1 + k \cdot d}, \quad (3.29)$$

donde como siempre k es un factor que mide la importancia de la diferencia cuadrática en la expresión, representada por d .

A partir de los resultados obtenidos se pueden establecer las siguientes conclusiones:

1. Los valores de similitud nulos que se observan en la Figura 3.10(c) son debidos a que durante el cálculo se han encontrado bloques de varianza 0, lo que implica que el bloque es constante y por lo tanto deben asignarse a 0 tal y como se vio en el apartado 3.6.2.
2. Las zonas de la imagen con un nivel de gris más alto (más claro) producen valores de similitud más altos, por el hecho de tener una media más elevada.

3. Cuanto mayor es el tamaño del bloque, más posibilidades hay de no encontrar un grupo de píxeles con varianza igual a cero, lo que a su vez permite una mejor definición de las zonas de correspondencia detectadas.
4. El hecho de emplear la diferencia cuadrática junto con una umbralización nos permite eliminar aquellas zonas que han dado una correspondencia alta por el hecho de tener un valor alto del nivel de gris.

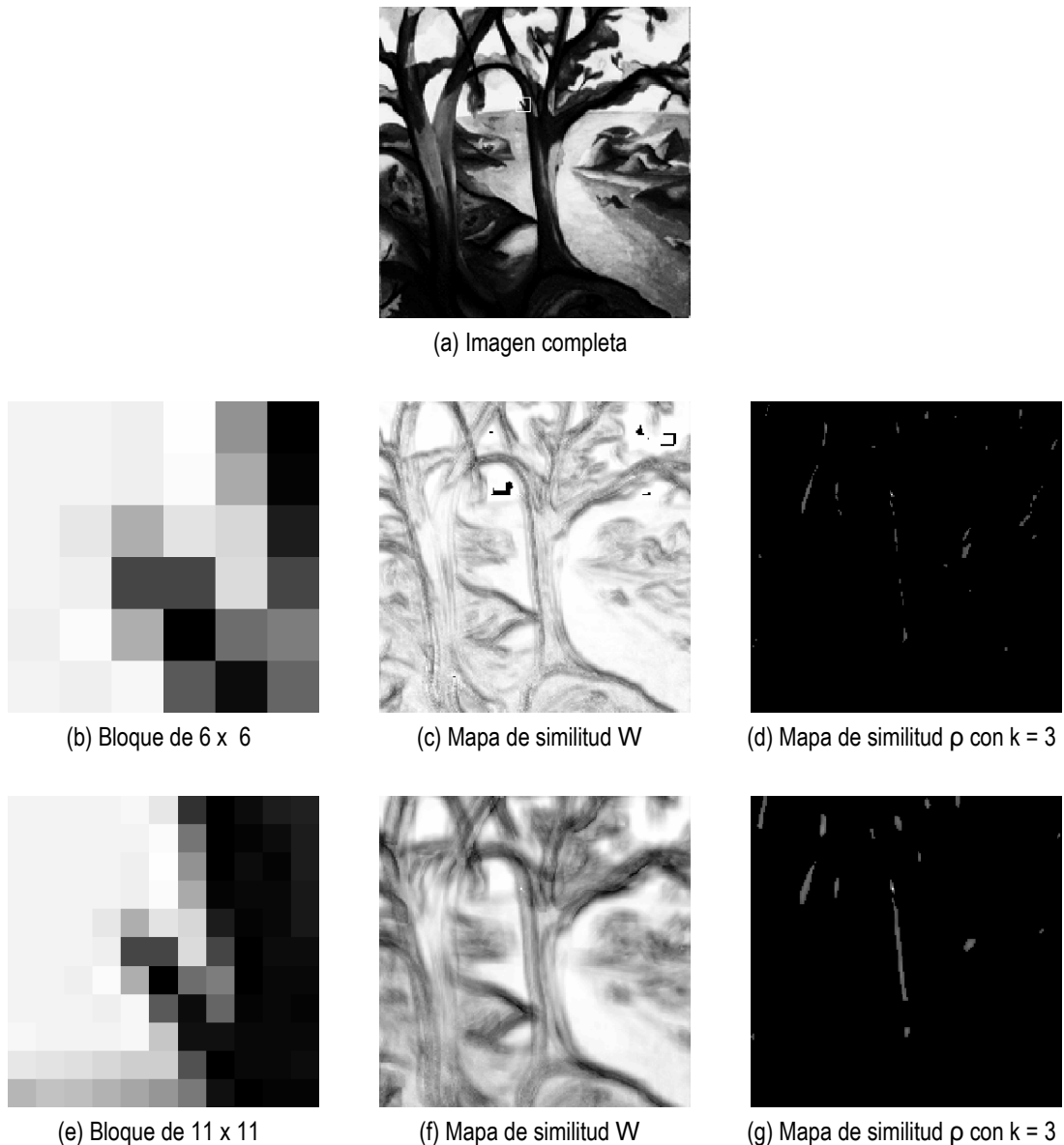


Figura 3.10. Ejemplo de mapas de similitud W ante distintos tamaños de bloque.

En la Figura 3.11 se muestra el resultado de calcular un mapa de similitud antes dos diferentes rotaciones de la imagen. Para el caso de la imagen mostrada en la Figura 3.11(a), una

rotación implica que las hojas se orienten de tal forma que aumenta el número de correspondencias posibles. Es fácil determinar, después de haber visto los ejemplos de este capítulo, cómo para determinadas imágenes una rotación tiene más repercusión que otra, pero igual de importante es el bloque del que se está buscando la correspondencia, en este caso representado en la Figura 3.11(b). En la Figura 3.11(d) y Figura 3.11(g) se puede ver el mapa de similitud para el criterio de Woods sin umbralizar para dos rotaciones de 10° y 20° respectivamente. Por su parte, en la Figura 3.11(e) y Figura 3.11(h) podemos apreciar el mapa de similitud combinando la medida de la diferencia cuadrática y una umbralización del coeficiente de similitud.



Figura 3.11. Ejemplo de mapas de similitud W ante distintas rotaciones.

3.7. COMPARATIVA DE LOS MÉTODOS DE SIMILITUD EXAMINADOS

Los métodos vistos en apartados anteriores son diferentes entre sí, pero a su vez proporcionan resultados muy similares ante determinados casos. En esta sección analizamos la influencia de dichos métodos de similitud ante ciertos casos particulares, comparando tales resultados de tal modo que nos permitan deducir algunas conclusiones importantes. Algunos de estos aspectos a analizar son la influencia del contenido del bloque, la eficacia frente a imágenes multimodo o el tiempo de procesamiento.

3.7.1. Influencia del contenido del bloque

Un aspecto muy a tener en cuenta a la hora de calcular el mapa de similitud entre dos bloques escogidos, como pueden ser una fracción de la imagen referencia y la imagen objetivo completa, es el conjunto de datos que dichos bloques contienen. Si el primer bloque contiene una serie de niveles de gris que definen una zona fácilmente reconocible, como un borde, una esquina o una transición, la medida de similitud por lo general reducirá las correspondencias en el mapa de similitud a una serie de zonas donde dicha área característica se pueda detectar. Si por el contrario el primer bloque no tiene píxeles que definan una zona fácilmente reconocible, como puede ser una zona de nivel de gris constante, el resultado del mapa de similitud podrá dar como resultado una serie de puntos con alta correspondencia pero muy alejados entre sí, como resultado precisamente de que la imagen original tenga varias zonas donde el nivel de gris sea constante o casi constante.

En la siguiente figura (Figura 3.12(a)) se recoge la imagen que se ha usado para el cálculo de los mapas de similitud. Esta imagen es muy adecuada debido a la presencia de zonas fácilmente detectables como bordes o zonas de nivel constante. En la Figura 3.12(b) se muestra el fragmento extraído de uno de los bordes del aparato de la Figura 3.12(a), marcado en esta última con un recuadro blanco.

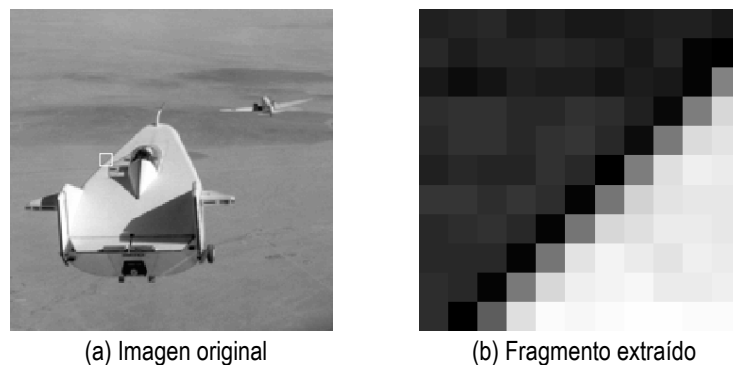


Figura 3.12. Imagen original y fragmento de la misma usados para el cálculo del mapa de similitud (caso 1).

En la Figura 3.13 podemos ver el resultado de aplicar las medidas de similitud tratadas a lo largo de este capítulo (coeficiente de correlación, información mutua, ratio de correlación y criterio de Woods) para el cálculo de la correspondencia entre el bloque de la Figura 3.12(b) y la imagen de donde se ha extraído (Figura 3.12(a)). En esta figura se puede apreciar claramente la disparidad de cada uno de los métodos de similitud, muy obvia al comparar los mapas de similitud calculados en la Figura 3.13(a), Figura 3.13(b), Figura 3.13(c) y Figura 3.13(d). En esta última se ven los valores de similitud altos obtenidos por la presencia de un nivel alto de los niveles de gris. En la Figura 3.13(e), Figura 3.13(f), Figura 3.13(g) y Figura 3.13(h) se muestran también los mapas de similitud calculados, pero esta vez haciendo uso de la diferencia cuadrática ($k=3$) y de una umbralización posterior de 0.3. Los resultados que nos da esta medida son muy similares para las cuatro medidas utilizadas, lo que proporciona una buena detección de las posibles zonas de correspondencia, situadas en la parte izquierda del artefacto mostrado en la imagen.

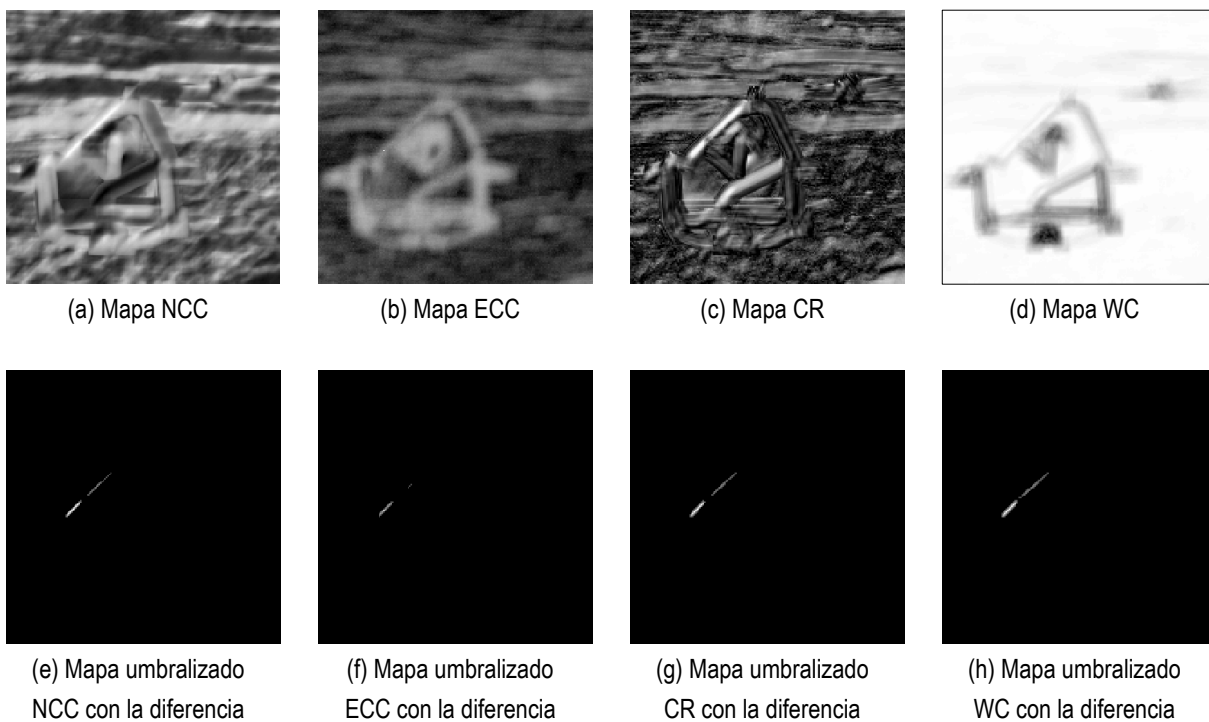


Figura 3.13. Resultado de los mapas de similitud para las cuatro medidas utilizadas (caso 1).

En el siguiente experimento (caso 2) se escoge una zona de nivel de gris constante para ver el comportamiento ante bloques con características poco definidas. En la Figura 3.14(a) se puede ver el recuadro desde donde se coge el bloque mostrado en la Figura 3.14(b).

En la Figura 3.15 se muestran los resultados de los mapas de similitud correspondientes a las cuatro medidas empleadas, con y sin el uso de la diferencia cuadrática ($k=3$) y la umbralización (0.3). De nuevo vemos que los mapas de la Figura 3.15(a), Figura 3.15(b), Figura 3.15(c) y Figura 3.15(d) tienen diferencias significativas, pero el uso de la diferencia cuadrática ayuda en cierto

modo a igualar estos comportamientos, como se puede ver en la Figura 3.15(e), Figura 3.15(f), Figura 3.15(g) y Figura 3.15(h).

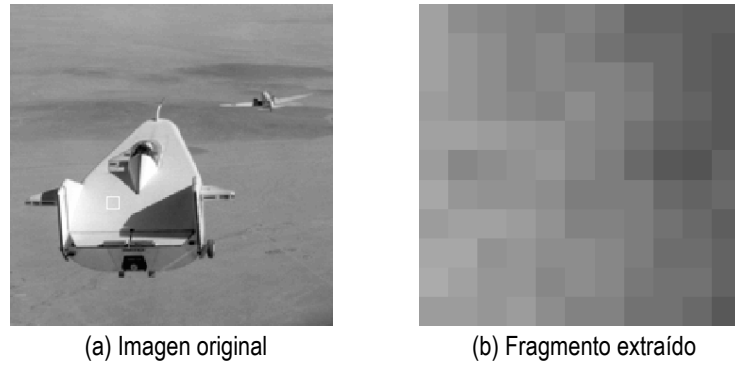


Figura 3.14. Imagen original y fragmento escogido de la misma con nivel de intensidad casi constante (caso 2).

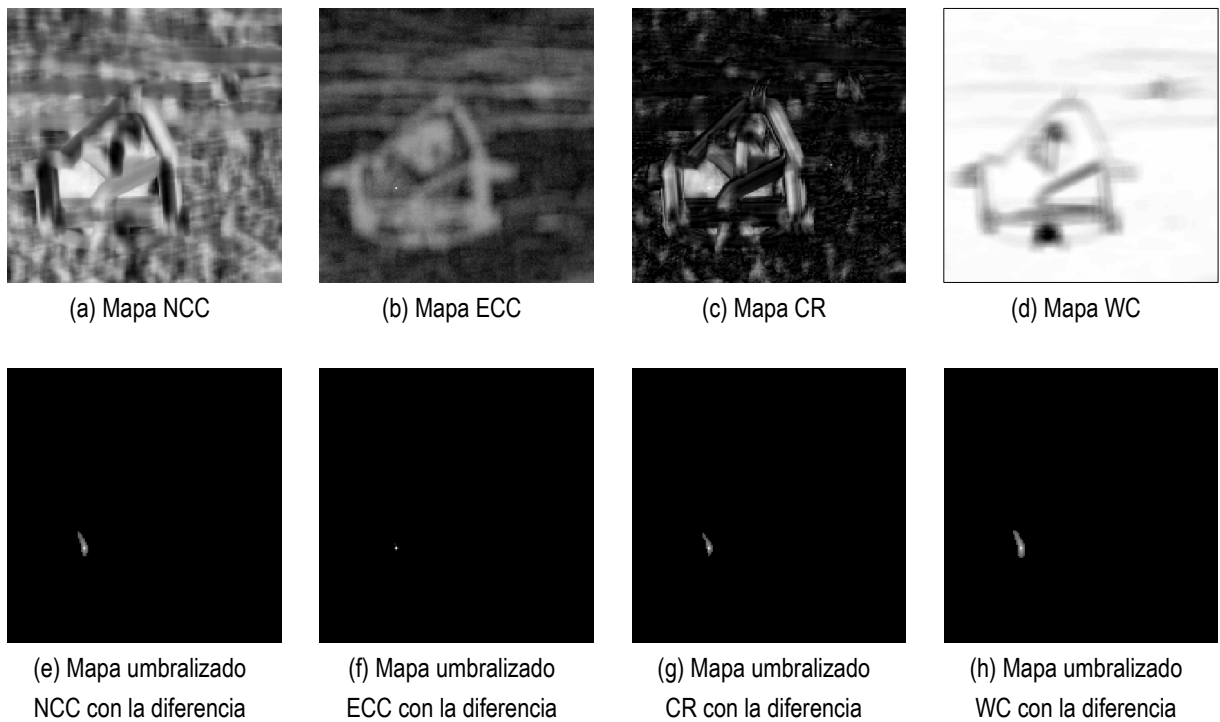


Figura 3.15. Resultado de los mapas de similitud para las cuatro medidas utilizadas y un bloque casi constante (caso 2).

A partir de los dos casos vistos, es fácil observar que la MI proporciona un valor de similitud por lo general más bajo y suave, al tratarse de una medida más cualitativa que cuantitativa, mientras que por ejemplo el WC tiende a dar mapas de similitud con valores altos cuando las zonas calculadas poseen una media de nivel de gris más elevada.

3.7.2. Influencia del contraste, iluminación y offset

Tan importante como conocer el funcionamiento de una medida de similitud frente al tipo de información escogida para calcular la correspondencia, es la sensibilidad de ésta ante cambios en la iluminación, contraste o el *offset*. Si existe *offset*, las imágenes deben normalizarse antes de ser procesadas. Para los otros dos casos vamos a observar la influencia en un ejemplo, cuyas imágenes se muestran en la Figura 3.16.

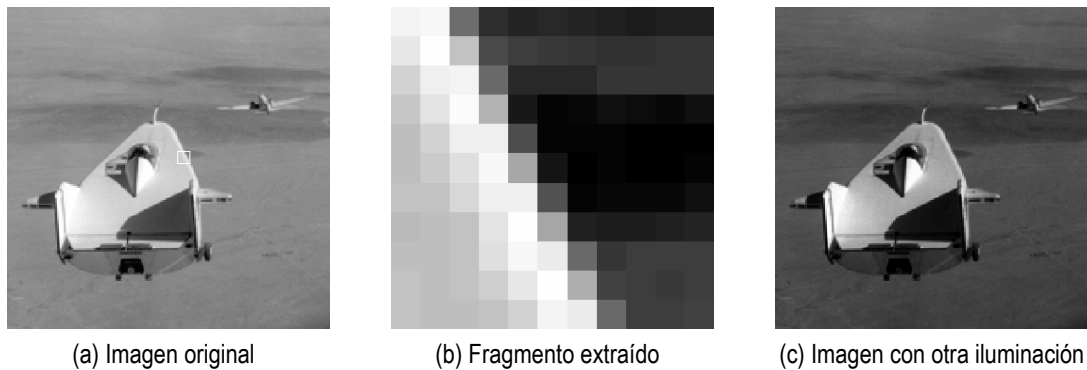


Figura 3.16. Imagen original, fragmento extraído de ésta y la misma imagen con un cambio de iluminación.

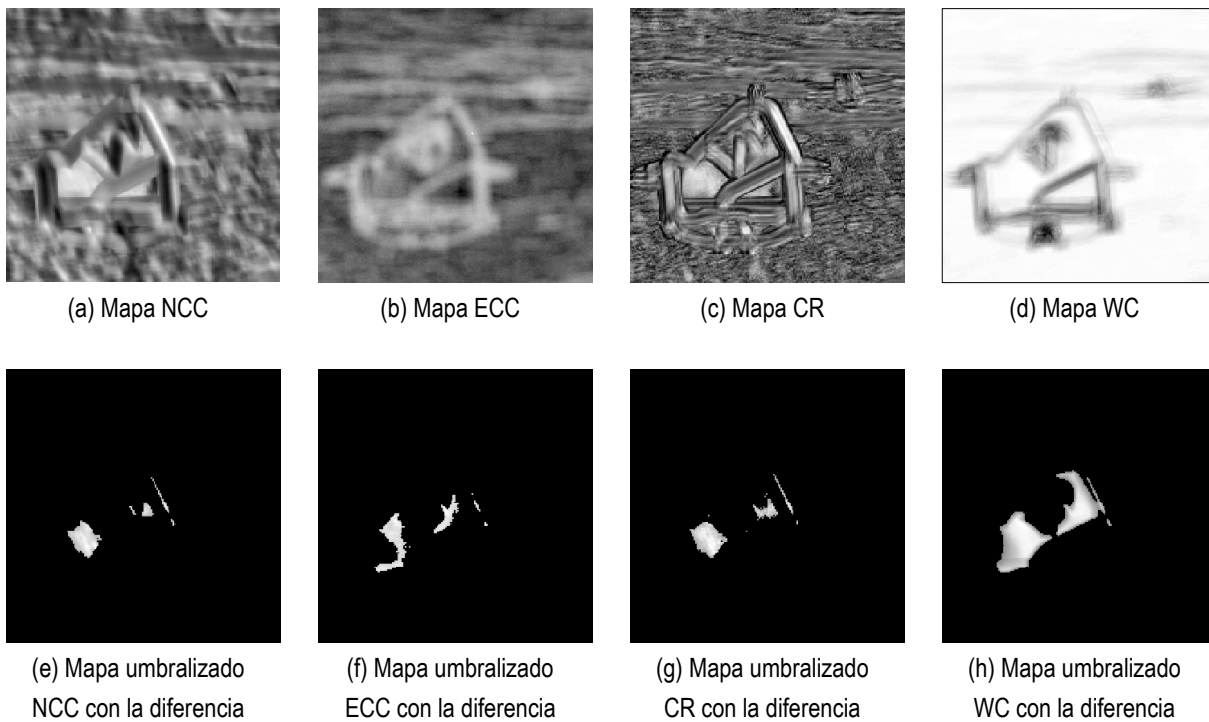


Figura 3.17. Resultado de los mapas de similitud para el caso de un cambio de iluminación.

Como se puede ver en la figura anterior, cuando tenemos dos imágenes con niveles distintos de iluminación o contraste, el uso de la diferencia cuadrática no se muestra eficiente para discriminar zonas donde el resultado de la similitud es alto, tal y como muestran la Figura 3.17(e), Figura 3.17(f), Figura 3.17(g) y Figura 3.17(h). Para estos casos puede bastar con umbralizar frente a un valor lo suficientemente elevado de la medida de similitud, para eliminar las grandes zonas de coincidencia. Un ejemplo se muestra en la siguiente figura (Figura 3.18), para un umbral de 0.85.

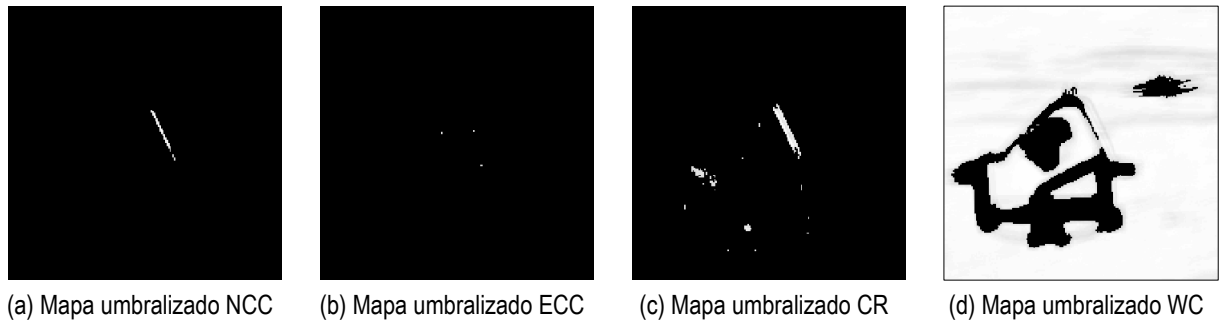


Figura 3.18. Mapas de similitud sin aplicar la diferencia cuadrática y con un umbral de 0.85.

En aplicaciones de registro de imágenes, encontraremos que un bloque extraído de una de las imágenes (la de referencia por ejemplo) no se utilizará para comparar con la totalidad de la otra imagen (en este caso la imagen objetivo), sino que se elige una determinada zona de vecindad sobre la que se le calcula el mapa de similitud. A partir de ese mapa de similitud se puede establecer una correspondencia aproximada del bloque en esta zona de búsqueda. Esto en cierto modo elimina los problemas encontrados por la presencia de varios lugares con alta similitud separados entre sí, tal y como se puede ver claramente en la Figura 3.18(c).

Sin embargo, otra técnica que puede emplearse para resolver en cierto modo este cambio entre los niveles de grises de dos imágenes conocidas, es llevar a cabo una corrección precisamente de estos niveles mediante un estudio estadístico de la imagen. Estas técnicas calculan la relación entre las intensidades de píxel de las imágenes a registrar mediante un método de estimación determinado y basado en el histograma conjunto de los niveles de gris presentes en las imágenes. Una vez realizada la corrección, es posible realizar el registro como si las imágenes fueran de una modalidad parecida o tuviesen escaso cambio en la iluminación.

En la Figura 3.19 se muestra un ejemplo de dos imágenes cuyos niveles de gris son diferentes entre sí, siguiendo un comportamiento cuadrático. La Figura 3.20(a) por su parte refleja el resultado del histograma conjunto entre las dos imágenes, lo que da una idea de la transformación que ha sufrido un cierto nivel de gris entre ambas imágenes. Una vez conocido este histograma, hay que tratar de ajustar este comportamiento mediante una curva paramétrica. La curva resultante de ajustar los puntos del histograma mostrado en la Figura 3.20(a) se ilustra en la Figura 3.20(b), donde se puede ver un gran ajuste a la transformación que existe entre los niveles

de gris. El resultado final de la corrección se incluye en la Figura 3.21, asignando a cada nivel de gris original de la imagen objetivo el nuevo nivel de gris dado por la curva de corrección calculada.

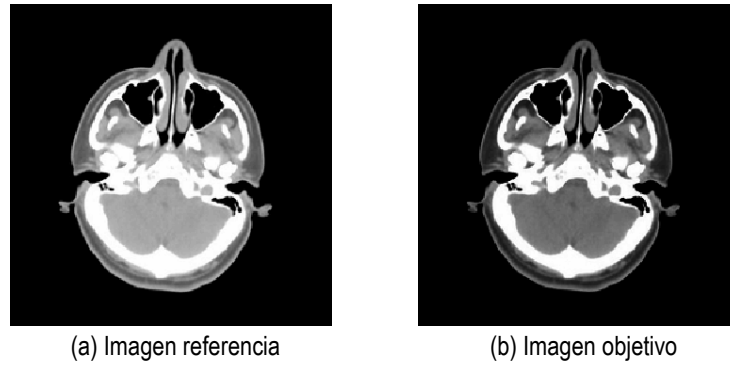


Figura 3.19. Ejemplo de imagen referencia y objetivo con variación entre sus niveles de gris.

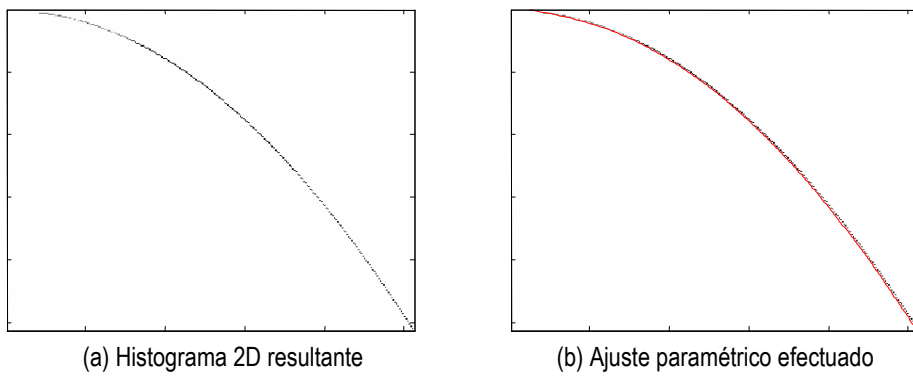


Figura 3.20. Histograma calculado y curva de ajuste al modelo.



Figura 3.21. Imagen corregida mediante la curva de corrección hallada.

La conclusión que podemos sacar a partir de la figura anterior es que si el modelo de transformación es fácilmente ajustable (como en este caso de transformación cuadrática), la corrección que se puede efectuar es lo suficientemente buena como para asegurar un registro entre las imágenes satisfactorio, tras una etapa previa de corrección.

3.7.3. Influencia del tipo de imagen

Como ya se trató en el apartado 2.2.1, las imágenes para el registro pueden ser monomodo o multimodo. A lo largo de las secciones anteriores se ha ilustrado el comportamiento de imágenes de tipo monomodo, por lo que ahora vamos a ver el comportamiento de las cuatro medidas de similitud estudiadas frente a una imagen multimodo. Las imágenes vistas en el apartado 3.7.2, donde se han tratado imágenes con diferentes niveles de iluminación o contraste, también podrían considerarse como un caso particular de registro multimodo.

En la figura siguiente (Figura 3.22) se muestra el ejemplo de un cerebro a partir del que se han obtenido dos imágenes: una estructural (Figura 3.22(a)) y otra funcional (Figura 3.22(c)). Además, en esta misma figura se puede ver el fragmento escogido de la primera imagen para el cálculo de la correspondencia en la segunda. El fin de la búsqueda de estas correspondencias para pares de imágenes multimodo es el de alinear las imágenes, de modo que se permita una visión mucho más completa del conjunto, localizando para ello en una imagen determinada las zonas significativas de la otra por mera fusión de los resultados.

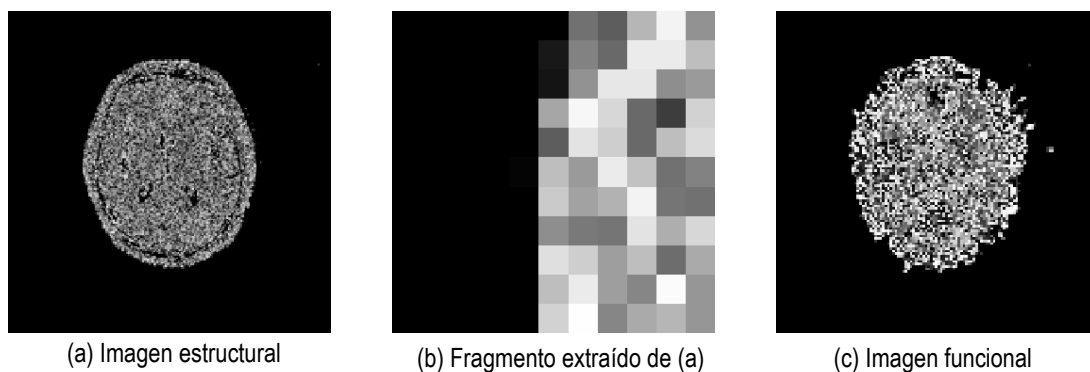


Figura 3.22. Pares de imágenes multimodo y bloque extraído para el cálculo de los mapas de similitud.

En la Figura 3.23 vemos los mapas de similitud calculados para cada una de las medidas sin haber llevado a cabo ningún tipo de corrección de gris previa, ya que éstos se encuentran muy cercanos entre sí. En la fila superior de la figura se muestran los mapas de similitud haciendo uso exclusivamente de las fórmulas estudiadas a lo largo de este capítulo, mientras que en la parte inferior se pueden observar los resultados obtenidos al aplicar la diferencia cuadrática al cálculo de los mismos. Para estos últimos cálculos se ha hecho uso de un factor de ponderado de la diferencia de $k = 2$ y de un umbral posterior de 0.25. El hecho de que el criterio de Woods ahora produzca un valor del mapa de similitud satisfactorio se debe a la propia hipótesis en la formulación del

algoritmo. Se trata por tanto de un método indicado para este tipo de aplicaciones, lo que claramente limitaba su funcionamiento para otro tipo de imágenes, tal y como se ha podido ver en ejemplos anteriores.

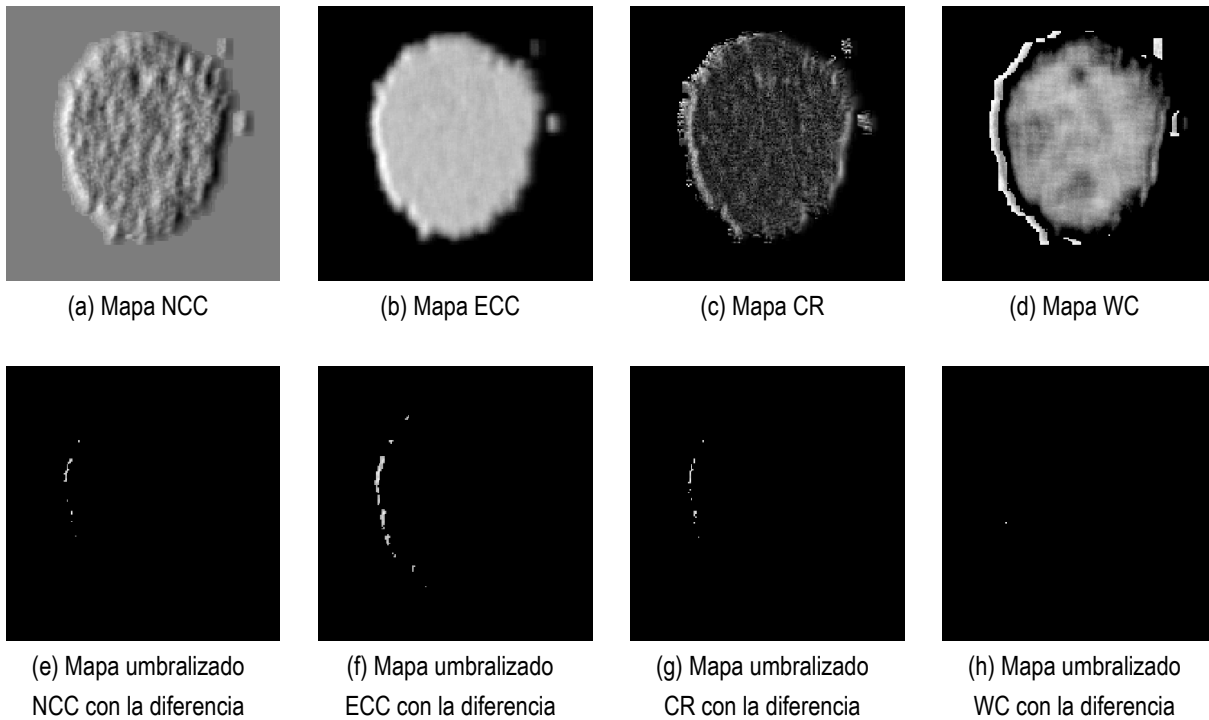


Figura 3.23. Conjunto de mapas de similitud con y sin el uso de la diferencia cuadrática y la umbralización.

3.7.4. Tiempo de procesamiento

En aplicaciones de registro de imágenes reales, un aspecto muy a tener en cuenta es el tiempo de procesamiento requerido. Cada medida de similitud hace uso de un algoritmo diferente que le confiere unas características intrínsecamente relacionadas con la capacidad de cálculo necesaria para llevarlas a cabo. En aplicaciones de tiempo real, donde el tiempo de cálculo es esencial, existe un compromiso importante entre la resolución de las imágenes procesadas y el tiempo de procesamiento. La duración del cálculo de cada una de las medidas de similitud es el tema a tratar en este apartado.

El tiempo de cálculo de un mapa de similitud está fundamentalmente relacionado con tres aspectos principales:

1. El tamaño del bloque que se usa para comparar en una determinada zona de la otra imagen.
2. La zona de búsqueda que determina el tamaño del mapa de similitud.

3. Las características concretas de ambos bloques de imagen.

El punto 1 se debe principalmente al hecho de que cuanto mayor es el tamaño del bloque escogido, más cantidad hay que procesar en la convolución 2D, más cantidad de *píxeles* hay que tener en cuenta para el cálculo de la varianza, desviación típica o media, y el histograma 2D tiene que procesar mayor número de elementos. Sin embargo, un mayor tamaño del bloque conlleva también un menor número de puntos en el mapa de similitud calculado, al encontrar menos zonas en la segunda imagen con las que se puede superponer dicho bloque, lo que a veces también implica una mejora en el tiempo de procesamiento.

El punto 2 determina que la zona en la que se calcula el mapa de similitud es de vital importancia, por el hecho de que el bloque escogido tiene que ir moviéndose sobre ella, calculando el valor de la similitud para todas las posibles posiciones del bloque sobre la zona de búsqueda. En el coeficiente de correlación cruzada normalizada (NCC), al no hacerse una división de la zona de búsqueda en tamaños del bloque escogido por el uso de la convolución, la influencia es menos apreciable a la hora de observar el tiempo que ha llevado el procesamiento. Para los otros tres casos, el tamaño de dicha zona de búsqueda resulta muy importante por tener precisamente que realizarse una segmentación previa en tamaños de los del bloque seleccionado, a modo de poder calcular el valor de similitud en cada uno de los segmentos dados.

Aunque a primera vista pueda parecer que el contenido del bloque escogido o de la zona de búsqueda no tenga importancia, también está ligeramente relacionado con el tiempo de cálculo del mapa de similitud. En el caso del NCC, y al realizarse una convolución 2D de todos los datos, las características de los bloques no afectan en gran medida. Para el caso del ECC, el CR o el WC, dependientes claramente del contenido de los bloques procesados, estas características no son despreciables. En el ECC se hace uso del histograma y del cálculo de las entropías, por lo que una serie de valores constantes en los bloques nos proporcionarían sólo unos pocos valores de la probabilidad no nulos, acelerando este proceso. En el caso del CR y el WC, al tener un algoritmo muy similar, el contenido de los bloques determina el número de iteraciones en las que se calculan las medidas condicionadas, de manera que si el primer bloque tiene valores de gris muy diversos, habrá que recorrer más iteraciones, mientras que para valores más reducidos estas iteraciones se minimizarán a la cantidad de grises encontrados.

Las pruebas recogidas en este apartado se han llevado a cabo con la versión 7.0.0.19920 (R14) de Matlab, con un procesador AMD de 3400 MHz y 1024 MB de RAM. En la página siguiente se muestra la Tabla 3.1, que recoge el tiempo de procesamiento para diversos tipos de imágenes y distintos tamaños del bloque y de la zona de búsqueda.

Caso	Tamaño del bloque	Zona de búsqueda	NCC	ECC	CR	WC
1. Imagen PET con varios niveles de gris en el bloque	11x11	200x200	0.031 s	395.828 s	44.016 s	63.703 s
2. Imagen PET con nivel de gris constante en el bloque	11x11	200x200	0.015 s	338.031 s	2.797 s	3.235 s
3. Imagen aérea	11x11	200x200	0.063 s	411.734 s	127.703 s	195.906 s
4. Imagen aérea	11x11	100x100	0.016 s	90.453 s	29.468 s	41.641 s
5. Imagen aérea	16x16	200x200	0.062 s	599.672 s	285.422 s	421.485 s
6. Imagen PET completa	200x200	200x200	5.953 s	6.750 s	6.156 s	6.235 s
7. Imagen PET completa	100x100	100x100	0.406 s	0.706 s	0.500 s	0.516 s

Tabla 3.1. Tiempo de procesamiento para diversos tamaños del bloque escogido y de la zona de búsqueda.

A partir de los resultados obtenidos para los distintos métodos de similitud podemos sacar varias conclusiones:

1. El tipo de imágenes que se comparan tiene influencia directa sobre el tiempo de procesamiento requerido (casos 1 y 3).
2. Para un mismo par de tamaño de bloque y zona de búsqueda, el tipo de bloque escogido interfiere en el tiempo de procesamiento (casos 1 y 2).
3. El hecho de reducir el tamaño de la zona de búsqueda conlleva una reducción drástica del tiempo de cálculo, tal y como muestran los casos 3 y 4.
4. Si se incrementa el tamaño del bloque, manteniendo el de la zona de búsqueda, se produce un aumento en el tiempo de procesamiento muy considerable para la ECC, el CR y el WC (casos 3 y 5).
5. Para el cálculo de la similitud entre bloques del mismo tamaño (casos 6 y 7), los tiempos de cálculo se igualan más para los cuatro métodos, y éstos a su vez se decrecen bastante al reducirse las dimensiones de los bloques a comparar.

6. Podemos ordenar las cuatro medidas de similitud en orden creciente según su tiempo de procesamiento, que en la mayoría de los casos se respeta: NCC, CR, WC y ECC.

Como hemos visto, el NCC constituye una medida mucho más rápida que el resto, y de hecho muestra un comportamiento ideal para aplicaciones de tiempo real, casi independientemente del tamaño de los bloques y zonas usadas. El hecho de que en el CR y el WC se haga uso implícitamente del histograma 2D de los niveles de gris de las imágenes nos proporciona una serie de mejoras de procesamiento frente a la información mutua, que muestra ser el procedimiento más lento y menos dependiente del contenido de los bloques. Sin embargo, vemos también que dependientemente del tipo de imágenes utilizadas o de los tamaños de bloque, las medidas se vuelven más o menos pausadas.

En la aplicación de estas medidas para el registro de imágenes nos encontraremos con un compromiso entre resolución y tiempo de procesamiento. Para reducir el tiempo de procesamiento manteniendo la resolución en las imágenes hay que intentar hacer uso de bloques extraídos de la imagen de referencia lo suficientemente pequeños para que ese tiempo sea menor, pero lo suficiente grandes también como para que no se produzcan demasiadas zonas de alta similitud en lugares donde realmente no haya un gran parecido. La zona de búsqueda, lejos de ser la totalidad de la imagen objetivo, generalmente se opta por tomarla en un área cercana a la posición de la que se ha tomado el bloque en la otra imagen. De esta forma se consigue reducir considerablemente el tiempo de procesamiento en los procedimientos de registro.

4. DESCRIPCIÓN DEL PROCEDIMIENTO DE REGISTRO

4.1. INTRODUCCIÓN

Sean I y J dos imágenes con cierta relación entre sí a partir de las cuales se desea calcular una cierta relación de dependencia entre ellas, atendiendo a las restricciones propias del movimiento deformable. Esta relación se representará en forma de vectores de deformación (véase Figura 2.2(b)), y representa el movimiento de los píxeles de una imagen respecto de la otra, de forma que esta transición sea suave, continua y conexa.

Según el tipo de los datos que estemos tratando, la clase de correspondencia será de una u otra manera, ya que nos encontraremos ante *estimación de movimiento* si los conjuntos de datos proceden de la misma fuente pero fueron tomados en distinto instante de tiempo; de deformación espacial si los conjuntos fueron tomados de la misma fuente pero en distintas posiciones espaciales; o bien de registro multimodo si ambos conjuntos no proceden de fuentes de la misma naturaleza. Todos estos casos son tratados de forma muy similar en este trabajo, obviando el origen de los datos comparados. En algunos casos, dependiendo de las características de las imágenes, será necesario seguir importantes cambios en los procesos llevados a cabo durante el algoritmo de registro de imágenes.

4.2. DEFINICIONES INICIALES

La ambigüedad en el cálculo de la correspondencia entre dos imágenes $I(\mathbf{x})$ y $J(\mathbf{x})$ es enorme, sobre todo en zonas de bajo contraste o texturas. Para la estimación de dicha correspondencia se suele hacer uso de una serie de restricciones o condiciones de contorno que nos permitan alcanzar al menos una solución cercana a la correcta. En este caso, la restricción principal vendrá dada por la suposición de un movimiento deformable de la imagen.

A través de una medida de similitud como las vistas en el capítulo 3, es posible calcular para cada píxel de la imagen I lo que se llama un *Mapa Planar de Similitud (MPS)*, haciendo para ello

uso de sucesivos pares de subimágenes, cada uno de los cuales denominamos *Conjunto Planar de Datos (CPD)*, extraídos de I y J . A partir de la información contenida en el MPS, se recoge la información referente al modelo paramétrico de estos mapas en el *Vector de Parámetros de Correspondencia (VPC)*. Por último, para cada píxel de la imagen I se transforman dichos modelos paramétricos en *Vectores de Correspondencia de Datos (VCD)* o *Vectores de Campo de Movimiento (VCM)*, que relacionan espacialmente ambas imágenes de forma explícita, mediante el uso de una serie de proyecciones y de restricciones de movimiento (regularización). En la Figura 4.1 puede verse de forma simplificada la secuencia básica utilizada para la estimación de la deformación.

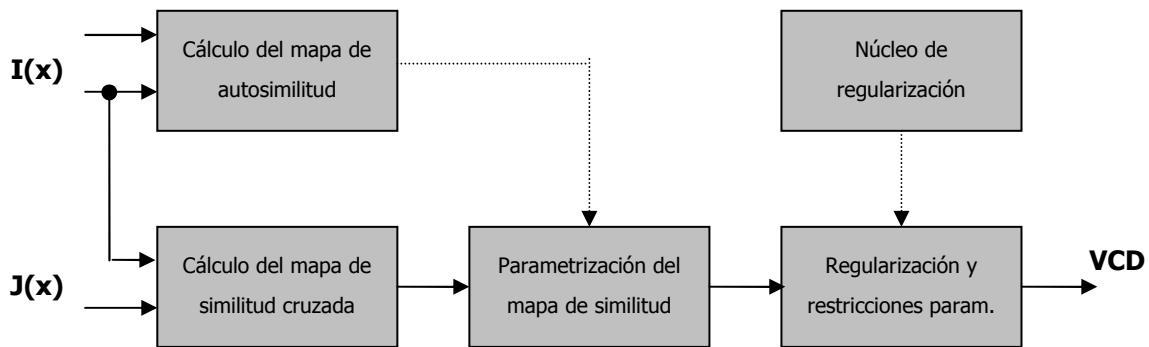


Figura 4.1. Esquema básico de la correspondencia de datos 2D.

Para el cálculo final de la correspondencia entre las imágenes se puede utilizar una aproximación jerárquica a la solución. Este tipo de técnicas (multirresolución o multiescala) se basan en la obtención sucesiva de aproximaciones a la correspondencia de mayor precisión, lo que permite que a partir de la resolución de peor calidad, la solución obtenida en cada resolución sirva de base para la de la resolución siguiente, acelerando con ello la convergencia a una solución estable.

4.3. CORRESPONDENCIA ESPACIAL BIDIMENSIONAL

Supongamos que para cada píxel $\mathbf{x} \equiv (x, y)$ de una imagen $I(\mathbf{x})$ se conoce un modelo paramétrico $\Omega(\mathbf{x})$ que define las regiones de $J(\mathbf{x})$ de correspondencia más probable. A partir de estas hipótesis se puede definir el funcional

$$\varepsilon(\mathbf{v}) = \sum_{\mathbf{x} \in \Phi} V\{\Omega(\mathbf{x}), \mathbf{v}(\mathbf{x})\} + \lambda \sum_{\mathbf{x}} \|\nabla^{\sigma} \mathbf{v}(\mathbf{x})\|^2. \quad (4.1)$$

Este funcional describe un compromiso entre la fidelidad del emparejamiento paramétrico dado por $\Omega(\mathbf{x})$, y la suavidad de la transformación estimada, $\mathbf{v}(\mathbf{x})$. En el primer término de dicha expresión, Φ indica el dominio en el que se han detectado posibles correspondencias

(características) $\Omega(\mathbf{x})$, y $V\{\Omega(\mathbf{x}), \mathbf{v}(\mathbf{x})\}$ representa una función de pérdidas que nos da una medida del coste que supone asumir que $\mathbf{v}(\mathbf{x})$ es el vector de correspondencia para el punto \mathbf{x} , sabiendo que la región de emparejamiento probable es la dada por el modelo $\Omega(\mathbf{x})$. En el segundo término se puede observar la expresión de *regularización de Tikhonov*, dada por la energía de la derivada espacial de orden σ de la transformación estimada. El parámetro λ es una constante positiva que pondera en este caso el peso de la regularización efectuada.

Aunque las imágenes $I(\mathbf{x})$ y $J(\mathbf{x})$ no aparecen explícitamente en la fórmula anterior, hay que decir que su contenido sólo se halla de forma implícita en los modelos paramétricos $\Omega(\mathbf{x})$, para cuyo cálculo sí que se hace uso de las imágenes definidas.

El mapeo $\mathbf{v}(\mathbf{x})$ que minimiza la expresión (4.1) se considera la mejor estimación posible de la correspondencia entre ambas imágenes, o lo que es lo mismo,

$$\mathbf{v}_o = \underset{\mathbf{v}}{\operatorname{argmín}} \varepsilon(\mathbf{v}). \quad (4.2)$$

En el caso de emparejamiento sin ambigüedad en el dominio Φ , es decir, ($\Omega(\mathbf{x}) = (v_x, v_y)$), la solución se puede obtener mediante un procedimiento directo. En un caso general de emparejamiento, el lugar del plano con el mapeo más probable se suele representar como una función paramétrica, mientras que la solución \mathbf{v}_o se puede obtener mediante un procedimiento iterativo de cálculo.

El hecho de que la convexidad de $\Omega(\mathbf{x})$ sea constante es una característica que se puede imponer a priori, y no representa ningún tipo de problema puesto que esta propiedad es compatible con todas las formas típicas que las regiones de emparejamiento adoptan. Sin embargo, es posible también hacer uso de una serie de áreas disjuntas mediante una modelación de varias funciones convexas disjuntas. Esta posibilidad nos permite alcanzar el mínimo global mediante una técnica iterativa conocida como *Project Onto Convex Sets (POCS)*.

Básicamente, el proceso de obtención de la correspondencia bidimensional entre dos imágenes se puede resumir en dos etapas separadas y secuenciales:

1. Se calculan los modelos paramétricos $\Omega(\mathbf{x})$ a partir de dos imágenes I y J .
2. A partir de los modelos paramétricos calculados en el paso anterior, se calcula la transformación \mathbf{v}_o que minimiza la función $\varepsilon(\mathbf{v})$.

4.4. ESTIMACIÓN DEL MOVIMIENTO EN 2D

Para la labor de estimar el movimiento entre dos imágenes se va a hacer uso de varias etapas, siendo cada una de estas etapas un subsistema del sistema total de procesamiento utilizado para la estimación del movimiento. Como una primera aproximación, este sistema podría

dividirse en dos etapas principales: una de cálculo de la correspondencia paramétrica (VPC) a partir de los dos CPDs iniciales de las imágenes; y otra de cálculo tras la correspondencia paramétrica previa, de los VCD que relacionan ambos CPDs de forma explícita a partir de los VPC resultantes en la etapa anterior. En la siguiente figura se describen los dos bloques citados.

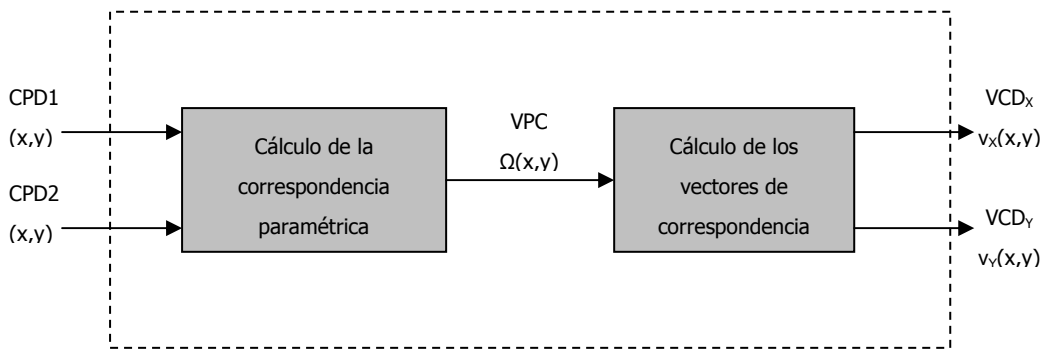


Figura 4.2. Diagrama de bloques básico para la estimación de movimiento en 2D.

Tal y como muestra la figura anterior, las entradas al sistema son dos CPDs, I y J , ambos de dimensiones $N_x \times N_y$ píxeles, que representan niveles de gris. La salida del primer bloque es otro CPD con celdas de $N_x \times N_y$ píxeles, donde cada una de estas celdas contiene un número variable de parámetros numéricos, que en conjunto representan el llamado VPC. Dichos VPCs determinan para cada píxel de la imagen I la región del espacio donde la correspondencia en J es mayor, según una determinada medida de similitud. El emparejamiento resultante se puede considerar como difuso, ya que se asigna toda una posible región de correspondencia parametrizada de J para cada píxel de I , lo que provoca cierta ambigüedad.

La salida procedente del primer bloque se introduce en el que se encuentra a continuación, con el objetivo de hallar los vectores de correspondencia. Las salidas de este segundo bloque son dos CPDs de dimensiones $N_x \times N_y$ píxeles cada uno. Cada posición de estos nuevos CPDs contiene un escalar que representa una de las dos componentes espaciales de los VCD, de modo que $\mathbf{v}(x, y) \equiv (v_x(x, y), v_y(x, y))$. El conjunto VCD calculado se ajustará a las restricciones de suavidad y continuidad en el entorno de cada uno de los píxeles escogidos.

4.4.1. Cálculo de la correspondencia paramétrica en 2D

El proceso de cálculo de la correspondencia paramétrica necesita de varios componentes, y además está dividido en un determinado número de pasos. Para empezar, es necesario almacenar las diferentes variables a procesar, como pueden ser por ejemplo los CPDs entrantes a dicho bloque. También son necesarios una serie de índices de direccionamiento de píxeles, que nos permitan determinar un vecindario alrededor de un determinado píxel escogido $\mathbf{x} \equiv (x, y)$, y una serie de procesos que nos permitan calcular la medida de similitud entre determinados bloques, para posteriormente realizar un ajuste paramétrico que nos proporcione la correspondencia paramétrica deseada.

En cada acceso de lectura, cada una de las memorias que contienen los CPDs devuelve un CPD de tamaño reducido y constante. La memoria que contiene el CPD1 devuelve una subimagen centrada en el píxel $(x + \Delta x, y + \Delta y)$, siendo Δx y Δy los desplazamientos del bloque, mientras que la memoria que contiene el CPD2 devuelve un bloque centrado en el píxel $(x + x_0 + \Delta x, y + y_0 + \Delta y)$. El tamaño de los bloques elegidos debe ser muy pequeño, para que así nos permita detectar el movimiento deformable. Este tamaño suele ser típicamente de 5×5 o 7×7 píxeles.

Para la lectura de los bloques en las dos memorias iniciales, se deben generar secuencialmente los índices para el acceso a un determinado bloque centrado en el píxel correspondiente. Así pues, durante el algoritmo iterativo que conlleva este proceso, se generan todos los posibles índices $(N_x \times N_y)$ para obtener las posiciones $(x + \Delta x, y + \Delta y)$ y $(x + x_0 + \Delta x, y + y_0 + \Delta y)$ que nos permitan leer los dos CPDs. De este modo, por cada uno de los índices recorridos en (x, y) se generará una serie de posiciones de vecindario $(\Delta x, \Delta y)$ que describirán un área rectangular centrada en el píxel correspondiente. El desplazamiento (x_0, y_0) procede de la estimación de movimiento obtenida en la resolución inmediatamente anterior, utilizando el esquema jerárquico de procesamiento. Este desplazamiento constituye el punto de partida para la búsqueda del emparejamiento entre bloques, y además será nulo en la escala inicial de la estimación de movimiento.

La secuencia de direcciones generada se obtiene mediante un barrido directo de todos los píxeles, tal y como sucede en las televisiones con cañón de rayos catódicos. Este procedimiento consiste en seleccionar un bloque centrado en la posición (x, y) de la primera imagen y escoger un bloque centrado en $(x + \Delta x, y + \Delta y)$ en la segunda imagen, de tal modo que para cada bloque de la primera imagen haya una serie de posibles bloques en la segunda, también centrados en la misma posición pero con un desplazamiento dado por la vecindad proporcionada por $(\Delta x, \Delta y)$. En la siguiente figura (Figura 4.3) se puede observar este concepto.

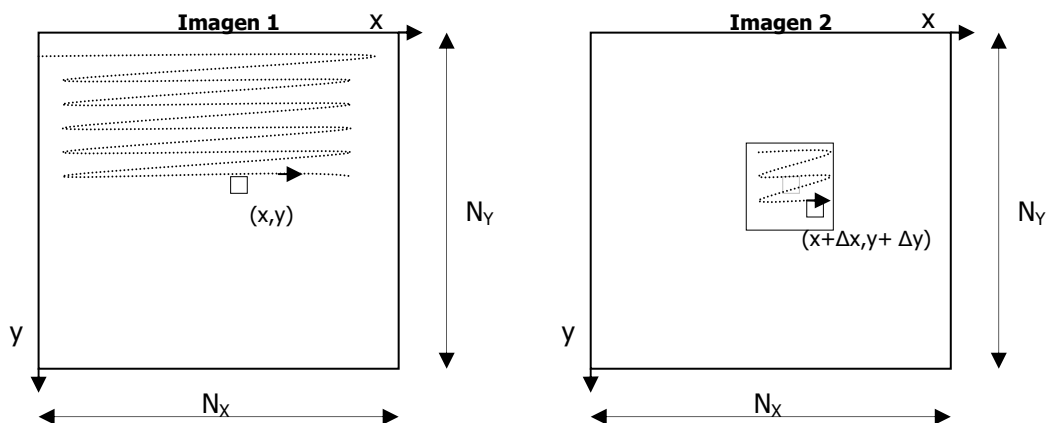


Figura 4.3. Secuencia de generación de direcciones para cada imagen.

Después de que los dos bloques de las imágenes hayan sido seleccionados en cada iteración, se calcula la similitud entre dichas subimágenes, y que como ya se vio en el capítulo 3, cuanto mayor es este coeficiente (generalmente el máximo es 1), más alto es el parecido entre los dos CPDs. En un primer instante se calcula la autosimilitud de la subimagen obtenida de la memoria que contiene el CPD inicial de I , almacenando el resultado en el *Mapa Planar de Autosimilitud (MPAS)*. Por otro lado se calcula la similitud entre las dos subimágenes, cada una de las cuales pertenecientes a los CPDs I y J , obteniéndose en este caso el *Mapa Planar de Similitud (MPS)*.

El siguiente bloque ajusta un modelo paramétrico simple polinómico al MPS de su entrada, determinando el orden y los parámetros del modelo que mejor se ajusta al MPS, mientras que el MPAS se utiliza para validar la coherencia del ajuste paramétrico efectuado. En último lugar, una memoria se encarga de almacenar los VPCs procedentes del bloque anterior. Esta memoria está constituida por $N_x \times N_y$ celdas vectoriales (o matriciales si hay varios modelos para un mismo MPS) de dimensión variable, en función del número de parámetros del modelo en cada píxel.

En la siguiente figura (Figura 4.4) se representa de forma básica la secuencia de bloques comentada hasta ahora.

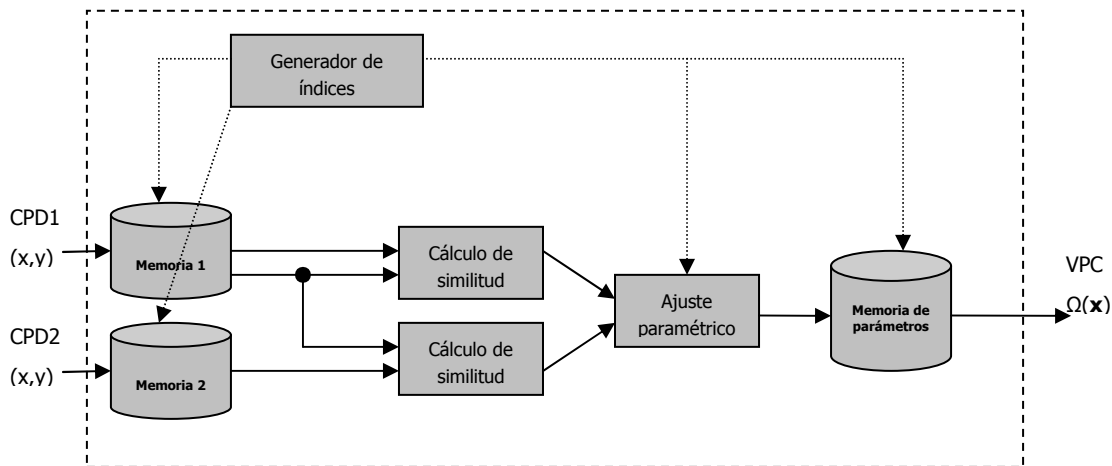


Figura 4.4. Esquema básico para el cálculo de la correspondencia paramétrica 2D.

4.4.1.1 Cálculo del coeficiente de similitud entre dos conjuntos de datos 2D

El cálculo del grado de parecido entre dos determinados bloques se hace a través de las medidas de similitud, ya tratadas ampliamente en el capítulo 3 de este trabajo. La misión de este sistema es encontrar la similitud entre los dos CPDs situados en su entrada, tanto para el cálculo de la autosimilitud como para su cálculo entre dos bloques de distinta fuente. La Figura 4.5 muestra el diagrama de bloques interno de la etapa de cálculo de similitud vista en la Figura 4.4. Hay que señalar que la salida $\rho_c(\mathbf{x})$ de dicho sistema es un escalar, ya que se produce la comparación de dos bloques del mismo tamaño.

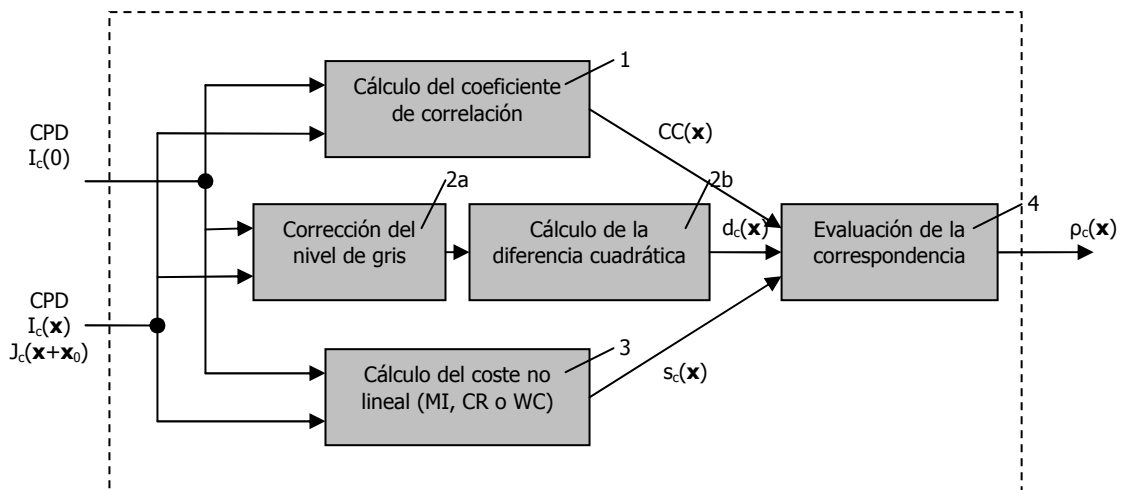


Figura 4.5. Cálculo de la similitud entre bloques 2D.

Básicamente, y a partir de los bloques CPD que entran en el subsistema de la Figura 4.5, se calculan tres medidas (la de similitud, la diferencia cuadrática y una de coste no lineal) que posteriormente son evaluadas conjuntamente mediante la información que todas ellas poseen, obteniéndose de esta forma un coeficiente mixto de similitud. El bloque 1 de la Figura 4.5 mide el parecido de los datos mediante la técnica lineal del coeficiente de correlación. Debido al hecho de que algunas medidas resultan poco fiables en bloques de bajo contraste, se hace uso de la diferencia cuadrática media normalizada en el bloque 2b. Esta tarea complementa a la desarrollada en el bloque o subsistema 1 para la localización de bloques, al resultar una medida muy efectiva en zonas de baja varianza y similar nivel medio de intensidad.

Dado que la fuente o naturaleza de las imágenes a registrar es a veces diferente, se hace necesario conocer la relación entre las intensidades de píxel para poder encontrar la correspondencia espacial entre ambas imágenes. Para obtener esta relación se efectúa un ajuste no lineal del histograma conjunto de los niveles de gris, lo que nos permite calcular la diferencia cuadrática incluso para estos casos donde los niveles de gris son diferentes. Este sistema, que ya se introdujo en el apartado 3.7.2, constituye el bloque 2a y se utiliza exclusivamente para el caso en que las imágenes tengan niveles de gris bien diferenciados, actuando como un cortocircuito para el caso de registro monomodo.

En el bloque de corrección de nivel de gris, el ajuste se realiza mediante interpolación por *splines*, siendo un *spline* una curva definida a trozos mediante polinomios. La interpolación mediante *splines* se utiliza porque da lugar a resultados similares a interpolaciones de alto orden, requiriendo solamente el uso de polinomios de bajo grado, a la vez que se evitan las oscilaciones, que en la mayoría de las aplicaciones resultan indeseables. Estas oscilaciones aparecen al interpolar mediante polinomios de grado elevado, y la facilidad de cómputo de los *splines* los hace populares para estos casos. El primer paso consiste por tanto en obtener el histograma 2D conjunto de niveles de gris, contabilizando para ello el número de ocurrencias de cada pareja de niveles de gris. Este histograma se puede umbralizar posteriormente, teniendo como objetivo de este procesado el eliminar las combinaciones con un número escaso de ocurrencias, que no

proporcionan información significativa, para así, tener en cuenta sólo los niveles de gris de mayor probabilidad. Después, se procede a promediar los valores en vertical (imagen a convertir) del histograma para cada uno de los valores del eje horizontal (imagen de referencia), obteniendo una curva única y equivalente al histograma conjunto procesado. Para obtener la expresión matemática de intensidades de píxel, se aplica el método proporcionado por la interpolación mediante *splines*. Así, se obtiene una nueva imagen con la intensidad de píxel calculada frente a la que se ha tomado como referencia. Esta nueva imagen corresponde a la imagen con las intensidades de píxel evaluadas de acuerdo a la relación de intensidades existente.

A veces es útil realizar un promediado de las medidas efectuadas para distintas rotaciones relativas entre los bloques de entrada, de tal modo que se eviten las limitaciones de dichas medidas frente al movimiento no rígido. Sin embargo, dependiendo de la aplicación, esta fase no es crítica y se puede obviar para rebajar el tiempo de cálculo global.

El bloque 3 de la Figura 4.5 es un bloque que reemplaza al bloque 1 cuando se hace necesario efectuar una serie de operaciones no lineales, dependientes en gran medida de la naturaleza de los datos; estas operaciones podrían ser por ejemplo alguna de las otras tres medidas de similitud vistas en el capítulo 3: el ratio de correlación, la información mutua y el criterio de Woods. En general, este sistema es prescindible cuando ambas imágenes proceden de la misma fuente, pero para aplicaciones de *fusión de datos* (registro multimodo) sí resulta necesario. Esta conclusión se deriva de que si la relación entre los niveles de gris de las imágenes es diferente, los métodos anteriores (correlación y diferencia) pueden resultar ineficaces, por lo que hará falta una función que nos permita maximizar el valor para dos puntos que sean tomados de la misma zona de las imágenes.

Cuando se utiliza la técnica de *Block-matching*, los bloques 2a y 2b puede despreciarse, ya que en este caso sólo se hace uso de una de las medidas de similitud (bloque 1 o 3), y de donde se escoge únicamente la zona donde dicho coeficiente es mayor.

4.4.1.2 *Ajuste del modelo paramétrico en 2D*

Después de calcular el grado de similitud de las imágenes, el siguiente paso es parametrizar mediante un determinado modelo el resultado obtenido en la fase anterior. Hay varios tipos de modelos a los que se puede producir un ajuste: modelo *punto*, *curva*, *mixto* e *indeterminado*. En un modelo *punto*, los píxeles no nulos están concentrados en torno a un lugar, definiendo una nube de valores. La dimensión del modelo punto es 0. En un modelo *curva* los píxeles determinan una forma básica de curva en el plano. Para este caso la dimensión se considera igual a 1. Cuando los píxeles no tienen una distribución bien definida donde sea aplicable uno de los dos casos, nos encontramos ante un modelo *indeterminado*, al que se le asigna dimensión 2. Por último, cuando hay zonas bien diferenciadas y aisladas podemos hablar de modelo *múltiple* o *mixto*, y cada una de esas zonas suele tratarse independientemente. En la Figura 4.6 se ilustra el concepto de los distintos modelos paramétricos.

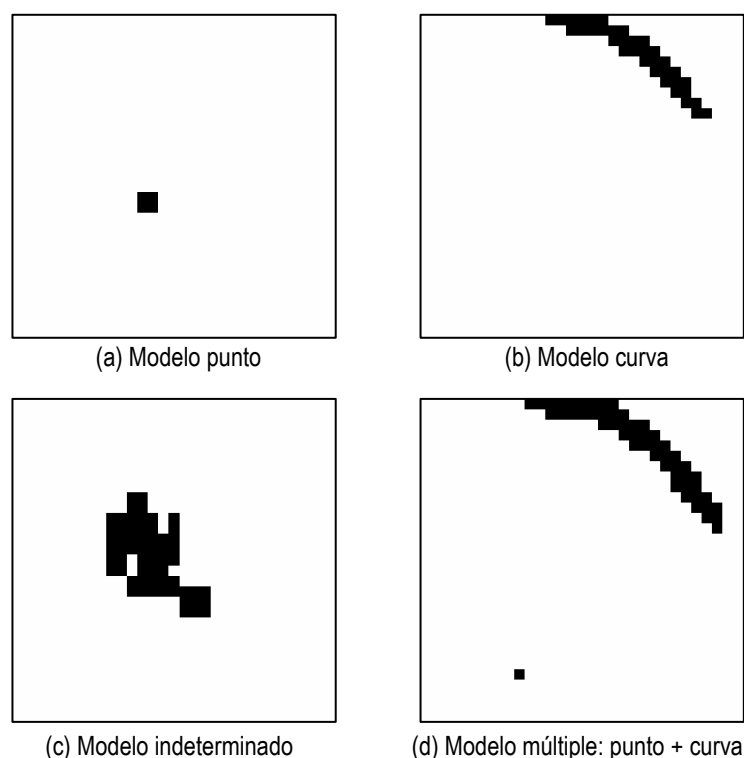


Figura 4.6. Modelos de ajuste paramétrico 2D.

Generalmente, es suficiente con utilizar funciones polinómicas para ajustar las curvas del modelo, que habitualmente tienen una curvatura leve y nos permiten hacer uso únicamente de funciones polinómicas cuadráticas para la caracterización. Así pues, las entradas al bloque de determinación del modelo paramétrico realizan una primera detección de la presencia de zonas inconexas dentro de un mismo MPS, y en su caso realizan una separación en submapas de similitud. Sólo se considerarán modelos dobles, por lo que si el número de zonas detectadas es mayor, todas ellas se descartarán por implicar demasiada incertidumbre. Una vez que se han separado las distintas zonas, llega el momento de ajustar los modelos a tipo punto y tipo curva. Una elección final determina el modelo que mejor se ajusta a cada subMPS mediante la comparación del error de ajuste normalizado, proporcionando sus parámetros a la salida en el VPC.

Un caso particular en el que cada una de las posibles correspondencias paramétricas es de tipo punto es la técnica denominada *Block-matching*. Como ya se ha comentado, esta técnica sólo deja pasar el píxel de mayor valor que define el MPS, lo que provoca que el ajuste sea siempre de este tipo.

4.4.1.2.1 Ajuste de modelos punto

Los parámetros de ajuste a modelos punto vienen dados por la posición de origen de $\rho(\mathbf{x})$ (o del submapa $\rho_i(\mathbf{x})$, con i igual a 1 o 2), es decir, (x_0, y_0) , y por las coordenadas del centro de masas, que pueden ser calculadas mediante los momentos de primer orden (m_x, m_y) . Así pues,

los parámetros del modelo a su salida estarán dados por $\Omega_p = \{m_x + x_0, m_y + y_0\}$ y por el error de ajuste E_p .

4.4.1.2.2 Ajuste de modelos curva

Para ajustar el conjunto de datos a un modelo curva, se calcula la línea curva paramétrica en dos dimensiones que mejor se ajusta a la distribución existente en el MPS $\rho(\mathbf{x})$ (o bien del submapa $\rho_i(\mathbf{x})$), mediante un criterio de error cuadrático medio mínimo. Los factores que determinan dicha curva son los dos siguientes:

- **Rotación:** La orientación fundamental de los datos en el espacio 2D dado por los ejes de coordenadas $(x + x_0, y + y_0)$ viene dada por un ángulo θ_c . Para el cálculo de este ángulo se hace uso de un algoritmo iterativo que nos proporciona unos nuevos ejes de coordenadas (\hat{x}, \hat{y}) facilitados por la ecuación

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} \cos\theta_c & \text{sen}\theta_c \\ -\text{sen}\theta_c & \cos\theta_c \end{bmatrix} \begin{bmatrix} x + x_0 \\ y + y_0 \end{bmatrix}. \quad (4.3)$$

- **Forma:** La función de forma viene dada por la ecuación $\hat{y} = f(\mathbf{a}, \hat{x})$, siendo \mathbf{a} los respectivos parámetros, y \hat{x} la variable dependiente. En la práctica se toma una función polinómica para dicha función, por lo que en este la caso dicha función se particulariza como

$$\hat{y} = a\hat{x}^2 + b\hat{x} + c. \quad (4.4)$$

De esta manera, la curva 2D queda definida por un total de cuatro parámetros: el ángulo θ_c y los tres coeficientes constantes de \mathbf{a} . A la salida del sistema nos quedaría un conjunto de parámetros dados por $\Omega_c = \{\theta_c, a, b, c\}$ y el error de ajuste normalizado E_c , que nos permitirá estimar el modelo que mejor se ajusta al conjunto de puntos dado.

4.4.1.2.3 Decisión del orden del modelo

Tras analizar los errores de ajuste a los modelos del MPS $\rho(\mathbf{x})$ dados por E_p y E_c , el sistema debe tomar uno de los cuatro posibles valores del conjunto $\{0, 1, 2, 3\}$ a su salida, correspondientes a los ajustes paramétricos en *punto*, *curva*, *indeterminado* y *múltiple* (doble). El modelo de orden dos sólo se elige cuando no se considera aplicable ni el modelo de orden 0 ni de orden 1 en ninguno de los submapas de similitud.

4.4.2. Cálculo de los vectores de correspondencia en 2D

En la Figura 4.2 vimos que los dos pasos fundamentales para el cálculo del movimiento (o deformación) en un conjunto de datos son el cálculo de la correspondencia paramétrica y el cálculo de los vectores de correspondencia. Una vez que tenemos los correspondientes modelos paramétricos en los VPCs, $\Omega(x, y)$, hay que llevar a cabo una conversión de estos datos en vectores de emparejamiento o VCD que relacionen de forma explícita los CPDs I y J . En la Figura 4.7 se puede observar el proceso de la conversión de datos comentado.

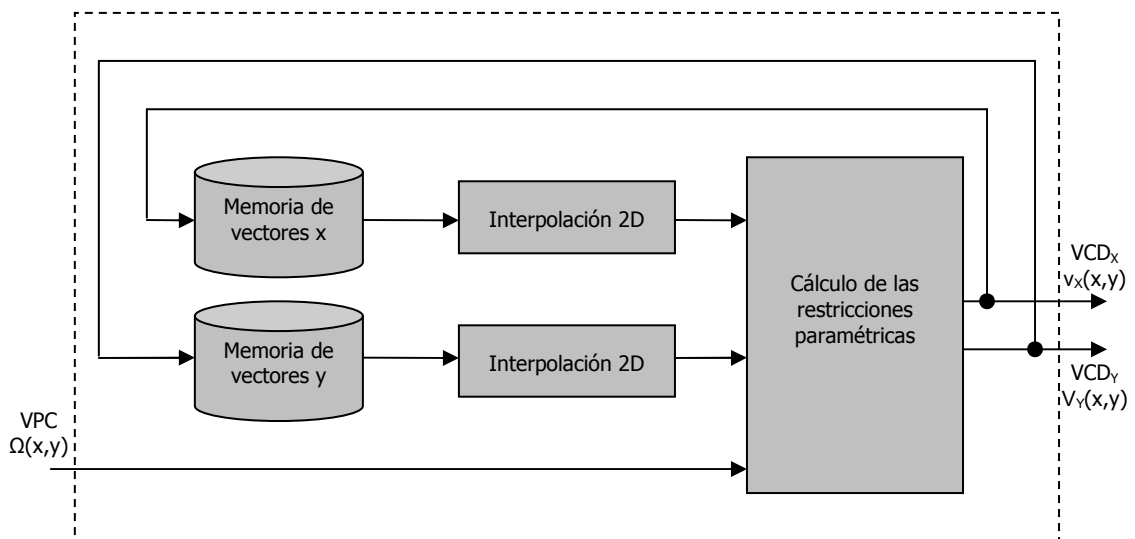


Figura 4.7. Sistema para el cálculo de la correspondencia en 2D.

Tal y como muestra la figura anterior, las realimentaciones de la salida sugieren un proceso iterativo que consta de dos fases bien diferenciadas:

- **Aplicación de condicionantes a priori (regularización de Tikhonov):** Este proceso se lleva a cabo mediante una interpolación o filtrado paso bajo de los vectores de emparejamiento parcialmente calculados. Este proceso es crítico para el sistema y se aplica de forma global a todo el conjunto de VCD, siendo necesario que el filtrado varíe su comportamiento con el tiempo para adaptarse rápidamente a todas las situaciones.
- **Aplicación de restricciones paramétricas espaciales:** Para ello se hace uso de una proyección de los vectores parciales de emparejamiento sobre los modelos paramétricos correspondientes. Esta fase tiene como objetivo forzar el cumplimiento de las restricciones paramétricas de todos los modelos punto, curva o doble (los indeterminados no implican restricciones) ya que la fase anterior puede haber alejado ciertos vectores de dichos modelos.

Todo este proceso se repite iterativamente, de modo que los vectores de emparejamiento se vayan aproximando a la solución óptima. El resultado final de estos vectores estará restringido a la información presente en la correspondencia paramétrica inicial, y además estará suavizado en el espacio 2D. Matemáticamente, este proceso iterativo se resume mediante la minimización del funcional ε ya visto en la expresión (4.1). Puesto que el objetivo es la minimización de la energía de curvatura de los VCD obtenidos al final, se buscará la minimización de la segunda derivada de dichos vectores. Debido a la naturaleza convexa de los modelos elegidos (polinomios de segundo orden), el método de regularización propuesto alcanza la solución óptima, lo que implica la presencia de un único mínimo.

4.4.2.1 *Filtrado de interpolación 2D*

Para el cálculo de la interpolación se hace uso de la *transformada rápida de Fourier (FFT)* en 2D de los datos planares que se introducen en el bloque de interpolación. Dentro de dicho bloque, también hay una memoria de solo lectura que contiene los coeficientes de un filtro de interpolación en el dominio de la frecuencia. A partir de los datos de entrada pasados al dominio de la frecuencia mediante la FFT, se aplica este filtrado mediante un multiplicador. Finalmente, y antes de dar la salida, se aplica la transformada inversa rápida de Fourier (IFFT) en 2D para pasar los datos de nuevo al dominio del tiempo. Aquí de nuevo se hace uso de un sistema de direccionamiento de píxeles similar al visto en el apartado 4.4.1. El objetivo del filtrado en el dominio de la frecuencia no es otro que el de reducir considerablemente el coste computacional.

El diseño del filtro de interpolación es un aspecto decisivo para el funcionamiento global del sistema. El objetivo fundamental es el de garantizar una solución espacialmente suave en los vectores de correspondencia, lo que implica que los vectores de correspondencia pertenecientes a píxeles cercanos tengan un valor similar en cuanto a distancia euclídea. Este filtrado es equivalente a la minimización de la energía de curvatura entre los vectores de correspondencia, mediante la minimización de la segunda derivada de los vectores de correspondencia.

4.4.2.2 *Aplicación de restricciones paramétricas en 2D*

Tal y como se muestra en la Figura 4.7, las entradas del sistema de aplicación de restricciones paramétricas son dos CPDs (las componentes v_x y v_y de los VCD parciales) y los VPCs. Por su parte, las salidas son los dos CPDs proyectados sobre los modelos paramétricos, para lo que se hace uso de una unidad de cálculo que proyecta un determinado vector sobre un modelo paramétrico. En cada iteración, el sistema direcciona un píxel distinto (x, y) y se lee un elemento de los VCD y un VPC de las memorias que contienen dichos valores. El correspondiente vector de emparejamiento se proyecta sobre la función definida por los parámetros de correspondencia y el vector proyectado resultante se almacena en la misma posición (x, y) de la iteración pertinente, para dar lugar a la salida del sistema.

4.4.2.2.1 Proyección del vector de correspondencia sobre el modelo

La Figura 4.8 ilustra el concepto de proyección de un punto sobre un modelo curva. Un punto P , al proyectarse nos devuelve un punto P' , que cumple la condición de ser el punto más cercano a la curva en cuanto a distancia euclídea. En el caso particular de modelos de tipo punto, la proyección de P es simplemente un remplazado de valores, de tal manera que el valor de P' se asigna siempre en coincidencia con la posición almacenada en el modelo.

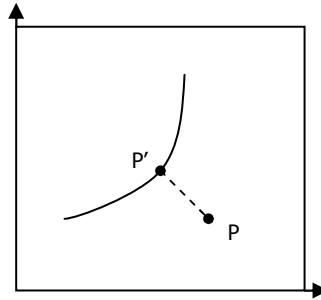


Figura 4.8. Proyección 2D sobre un modelo de curva.

El hecho de que se proyecte sobre modelos conlleva un aumento de la suavidad en los VCDs, pues éstos deben ajustarse a unos modelos que intrínsecamente describen transiciones suaves.

Para el cálculo de la proyección 2D se parte de las dos coordenadas del vector VCD y del conjunto de parámetros de correspondencia del modelo. En el caso de modelos curva simples, el sistema proyecta el vector de entrada sobre la función paramétrica correspondiente, mientras que si nos encontramos ante un caso de modelo doble, se efectúa una doble proyección. La salida, dividida en dos componentes, nos proporciona la proyección anterior si el modelo es tipo curva, el vector procedente (las coordenadas) del VPC si se trata de un modelo punto o el mismo vector sin proyectar para modelos indeterminados.

Para llevar a cabo la proyección de un punto cualquiera (x, y) sobre un modelo polinómico de segundo orden (tipo parábola), se utiliza un algoritmo eficiente de descenso de gradiente, *Least Mean Square (LMS)*, que minimiza la distancia euclídea del punto a la parábola. Este proceso de proyección se realiza en tres etapas fundamentales:

1. **Rotación inversa:** En esta fase se centra y gira el punto a proyectar para alinearlo en ángulo con el modelo, cuya orientación y origen vienen dados respectivamente por el ángulo θ_c y el par (x_0, y_0) . Para ello se rota un ángulo $-\theta_c$ el vector a proyectar, mediante la transformación

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} \cos \theta_c & -\text{sen} \theta_c \\ \text{sen} \theta_c & \cos \theta_c \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}. \quad (4.5)$$

2. **Proyección:** Para efectuar la proyección del punto rotado a la parábola se puede hacer uso de una función de error o de coste, minimizable mediante el algoritmo de descenso de gradiente, que a través de varias iteraciones va aproximando la posición de la proyección.
3. **Rotación:** En esta etapa se invierte el proceso llevado a cabo en la primera etapa por un ángulo θ_c , mediante la siguiente expresión:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta_c & \text{sen } \theta_c \\ -\text{sen } \theta_c & \cos \theta_c \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \quad (4.6)$$

donde (\tilde{x}, \tilde{y}) son los puntos proyectados en la etapa anterior, tras completar todas las iteraciones.

Todas las operaciones anteriores se repiten para cada uno de los vectores existentes, de tal modo que se obtengan todas las parejas (x, y) y (x', y') .

5. RESULTADOS

5.1. INTRODUCCIÓN

Una vez que se han asentado las bases del proceso de registro de imágenes, es el turno de analizar los resultados para casos específicos de imágenes reales. Para dicho análisis se va a observar el resultado obtenido por cada una de las técnicas de similitud estudiadas, lo que nos permitirá ver la diferencia entre la imagen referencia y la registrada en imágenes de distinta naturaleza, y las limitaciones de estos métodos ante grandes deformaciones. Cuatro serán las posibles fuentes de imágenes a estudiar: imágenes de tipo médico, imágenes fotográficas, imágenes aéreas e imágenes sintéticas. En primer lugar veremos el comportamiento para imágenes monomodo, mientras que después pasaremos a las de tipo multimodo, el último objetivo de este trabajo.

5.2. EVALUACIÓN DEL RESULTADO DEL REGISTRO

Para determinar si el resultado de dos imágenes a registrar ha sido satisfactorio, a veces no basta con observar la imagen de referencia y la imagen transformada a simple vista, pues este resultado a menudo es engañoso. Para ello, es útil hacer uso de algunas medidas que nos permitan evaluar la fidelidad de los resultados obtenidos frente a la deformación inicial que la imagen objetivo poseía. Algunas de estas medidas son el PSNR (*Peak Signal-to-Noise Ratio*), el error cuadrático medio o el cálculo de la energía de los vectores de transformación estimados (VCD).

5.2.1. El error cuadrático medio (ECM)

El error cuadrático medio (ECM), o *Mean Squared Error (MSE)*, es una medida del cuadrado del error entre dos imágenes I y J . Este error es la cantidad que una de las imágenes difiere con respecto de la otra, y viene dado por la expresión siguiente:

$$ECM = \frac{1}{N} \sum_{i=1}^N (I_i - J_i)^2. \quad (5.1)$$

En registro de imágenes, es habitual hacer uso de esta medida para cuantificar el error que se produce entre el vector de movimiento real de la imagen objetivo y referencia, y el vector de movimiento estimado en el procedimiento de registro. La presencia de este error se debe a que a veces la estimación no se calcula con una precisión suficiente.

Otra medida directamente derivada del ECM es la raíz cuadrática media, o *Root Mean Square (RMS)*, calculada como la raíz cuadrada del ECM, de la siguiente manera:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (I_i - J_i)^2} = \sqrt{ECM}. \quad (5.2)$$

5.2.2. El PSNR (*Peak Signal-to-Noise Ratio*)

El PSNR calcula el parecido entre dos imágenes I y J , de valores de intensidad comprendidos entre 0 y 1 (ó 0 y 255), produciendo un resultado en decibelios (dB) que nos da una idea del grado de similitud entre ambas imágenes. Hay varias versiones de SNR, pero el PSNR es el más común en el procesamiento de imágenes, probablemente por el hecho de proporcionar números más prácticos que en el resto de medidas. Esta medida se puede definir como

$$PSNR = 20 \log \frac{b}{rms}, \quad (5.3)$$

donde b es el valor mayor posible de la señal, y rms es la raíz cuadrática media vista en (5.2). El PSNR se da en unidades de decibelios, que miden el ratio entre el pico de la señal y la diferencia entre dos imágenes. Un incremento de 20 dB corresponde a un decrecimiento de una décima parte en la diferencia rms entre las dos imágenes.

5.2.3. Cálculo de la energía del VCD

Una forma de ver de qué manera se ha detectado la deformación de una imagen es mediante el cálculo de las *energías de difusión y curvatura* de los vectores VCD. Estas medidas reciben también el nombre de *rigidez y elasticidad* respectivamente, y están calculadas a partir de

la primera y segunda derivada de los vectores de movimiento, lo que permite hacernos una idea de la suavidad de la transformación global llevada a cabo al final del registro.

5.3. RESULTADOS OBTENIDOS

5.3.1. Imágenes sintéticas

5.3.1.1 Caso 1. Bola con estrellas

En la Figura 5.1 podemos ver el primer caso de registro tratado para una imagen sintética. Este tipo de imágenes se caracterizan por ser generadas artificialmente, y son de gran interés al permitirnos crear la imagen objetivo a partir de una imagen de referencia con sólo aplicar una transformación conocida. Posteriormente, esto nos permitirá evaluar la transformación detectada en el registro frente a la aplicada.

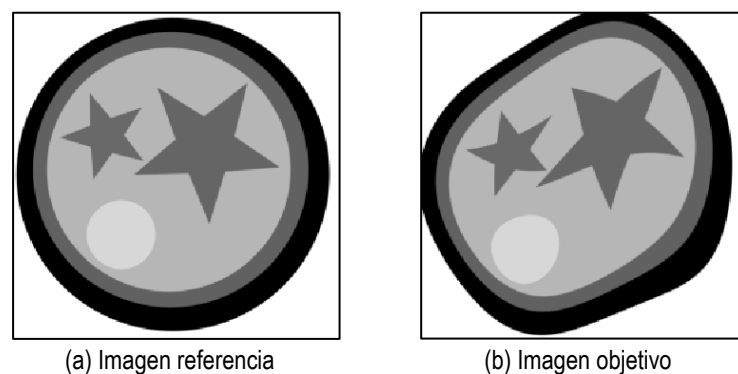


Figura 5.1. Par de imágenes sintéticas para el registro.

En la Figura 5.2 se ilustra el efecto de realizar el registro de las imágenes utilizando la técnica de ajuste paramétrico usual y la de *Block-matching (BM)*. En la Figura 5.2(a) se puede ver la imagen registrada mediante la medida de similitud NCC, mientras que los vectores de movimiento detectados se muestran en la Figura 5.2(b). En la Figura 5.2(c) y Figura 5.2(d) se ve el efecto de aplicar la transformación inversa dada por los vectores de movimiento a una rejilla rectangular, y en la Figura 5.2(e), Figura 5.2(f) y Figura 5.2(g) se observan los resultados anteriores pero para la técnica BM.

Con sólo observar las imágenes reflejadas en la Figura 5.2, podemos demostrar que el modelo de ajuste paramétrico seguido es mucho más robusto que el que nos proporciona la técnica de *Block-matching*. *Block-matching* resulta un procedimiento poco preciso a la hora de realizar el registro de dos imágenes, pecando de las desventajas asociadas al no uso de la diferencia cuadrática y al uso exclusivo de modelos tipo *punto* para solventar el nivel de similitud en zonas de baja varianza. A partir de ahora sólo se hará uso de técnicas de ajuste paramétrico,

sin limitar su comportamiento a correspondencias de tipo *punto*. De este modo, con el uso conjunto de la diferencia cuadrática, se obtendrán resultados más satisfactorios.

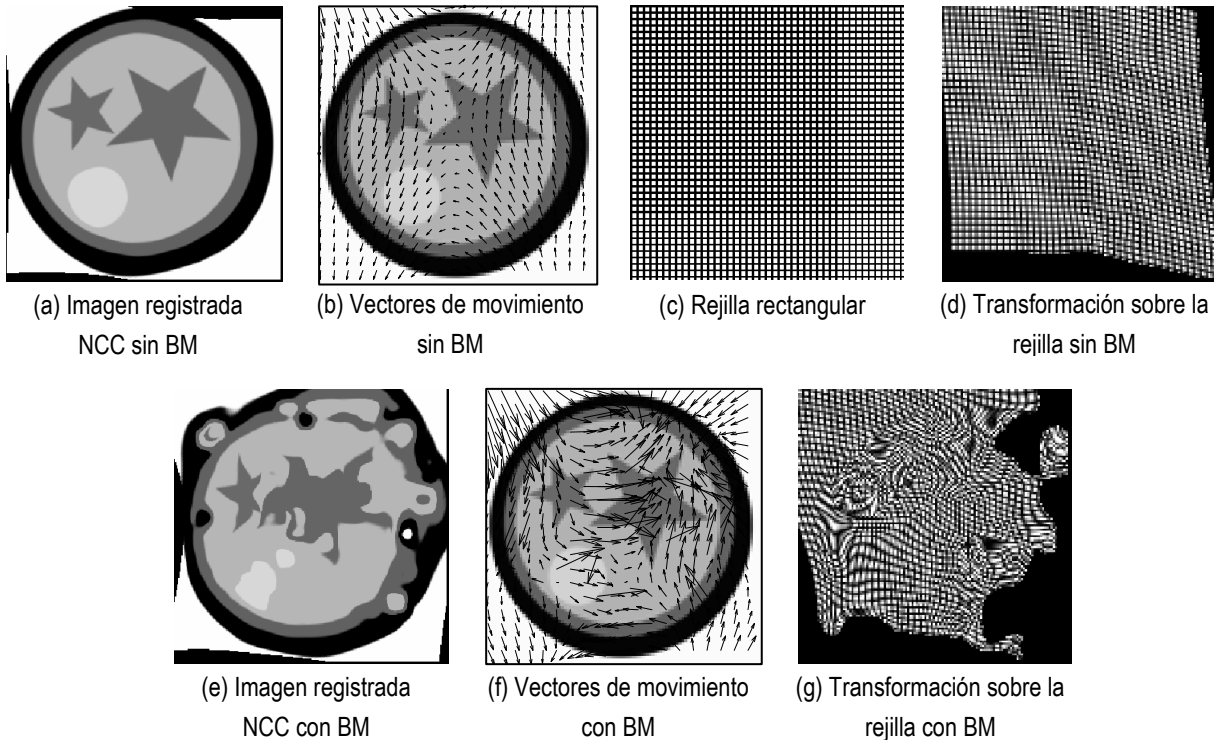


Figura 5.2. Resultado del registro con y sin *Block-matching*.

En la Figura 5.3 podemos ver el resultado del registro para las otras tres medidas de similitud analizadas, haciendo uso de nuevo del ajuste paramétrico descrito en el capítulo anterior. Estos resultados consiguen a simple vista un alineamiento muy bueno frente a la imagen de referencia, aunque aparecen unos bordes negros debido a que su valor se ha desplazado dentro.

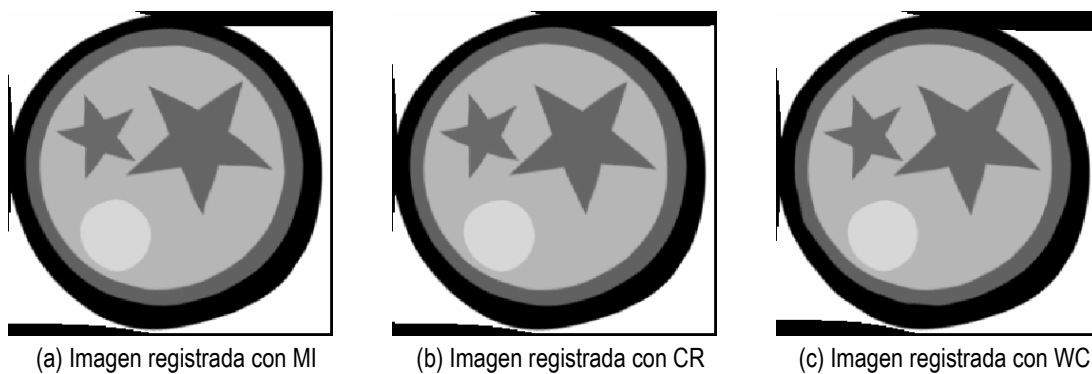


Figura 5.3. Resultado del registro para MI (ECC), CR y WC.

En la tabla siguiente (Tabla 5.1) se resumen las medidas que nos cuantifican los resultados del registro. El PSNR indica el parecido de la imagen de referencia con la registrada, mientras que la rigidez y la elasticidad nos indican las características de los vectores de transformación calculados. Como vemos, el mejor resultado lo produce la información mutua seguida del CR. El criterio de Woods muestra ser el menos eficaz, al estar basado en un tipo específico de imágenes.

Medida	PSNR	RIGIDEZ			ELASTICIDAD		
		Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	12.31 dB	0.0588	0.0321	0.0036	0.0632	0.0058	0.1111
CR	12.39 dB	0.0581	0.0317	0.0034	0.0286	0.0052	0.0056
WC	11.67 dB	0.0791	0.0448	0.0084	0.1595	0.0087	0.7266
MI	12.41 dB	0.0648	0.0396	0.0044	0.0348	0.0050	0.0144

Tabla 5.1. Resumen de los resultados obtenidos para el caso 1 de imagen sintética.

5.3.1.2 Caso 2. Mosaico

Para este caso, la imagen objetivo se generará a partir de la imagen referencia aplicando una transformación conocida. Esto nos permitirá evaluar el error cuadrático medio que se produce en el proceso de registro. Dichas imágenes se representan junto a los vectores de movimiento reales en la siguiente figura (Figura 5.4).

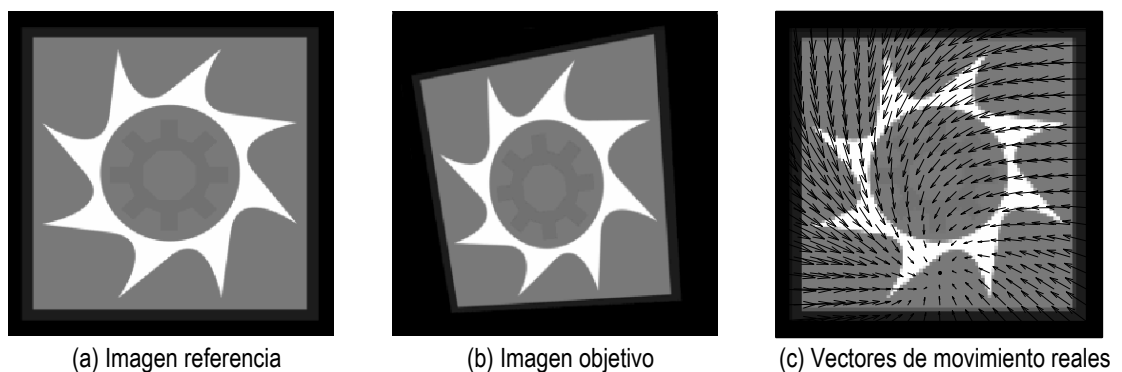


Figura 5.4. Imágenes de partida para el registro del caso 2, junto con los vectores de movimiento reales.

La Figura 5.5 muestra el resultado del registro para las distintas medidas de similitud. En este caso la gran deformación existente en el borde izquierdo de la imagen hace imposible que el registro se alinee perfectamente. La Tabla 5.2 recoge los resultados del proceso de registro.

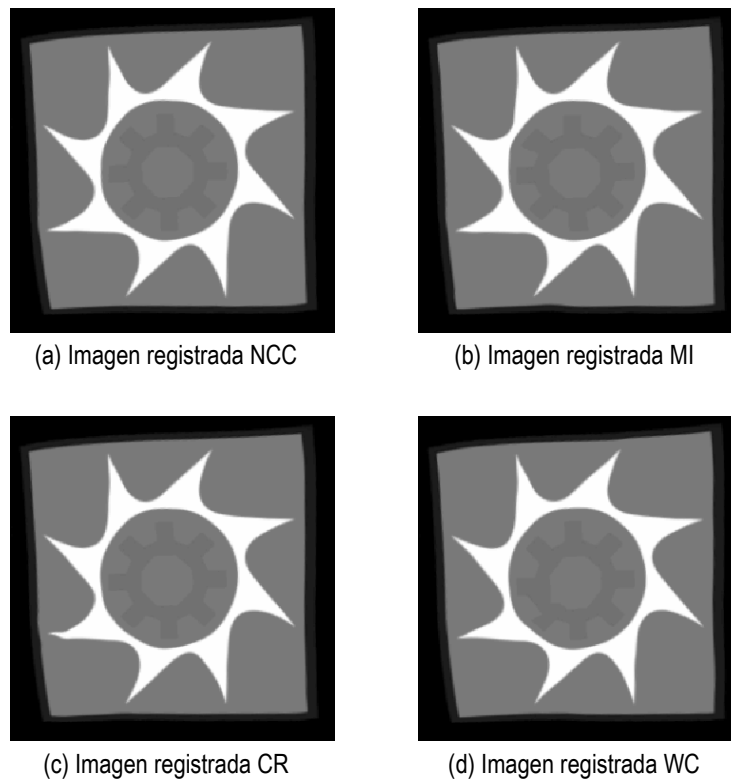


Figura 5.5. Resultado del registro para todas las medidas de similitud.

Medida	PSNR	ECM	RIGIDEZ			ELASTICIDAD		
			Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	18.82 dB	0.3232	0.0925	0.0898	0.0028	0.0168	0.0044	0.0014
CR	18.80 dB	0.3225	0.0926	0.0890	0.0027	0.0206	0.0046	0.0028
WC	18.61 dB	0.4100	0.0984	0.0856	0.0061	0.0512	0.0082	0.1002
MI	18.71 dB	0.3326	0.0950	0.0901	0.0034	0.0300	0.0045	0.0150

Tabla 5.2. Resumen de los resultados obtenidos para el caso 2 de imágenes sintéticas.

A partir de los datos anteriores, es fácil ver cómo el PSNR se encuentra estrechamente ligado con el ECM, ya que al crecer el primero el segundo decrece. Básicamente, esto se deduce teniendo en cuenta los vectores de transformación calculados, que al ser más cercanos a los reales producen un mejor resultado de registro, cuantificado por el PSNR, y a su vez reducen su diferencia con los vectores reales, lo que implica que el error disminuya y se acerque a la transformación real entre las imágenes.

5.3.1.3 Caso 3. Conjunto de figuras geométricas

La Figura 5.6 muestra el par de imágenes utilizadas para el registro en este caso. Ahora contamos con una serie de figuras que han sido deformadas de una u otra manera, por lo que en cierto modo el registro tiende a estar localizado en cuatro partes diferentes. La Figura 5.7, por su parte, ilustra el resultado obtenido mediante las cuatro técnicas de cálculo de similitud.

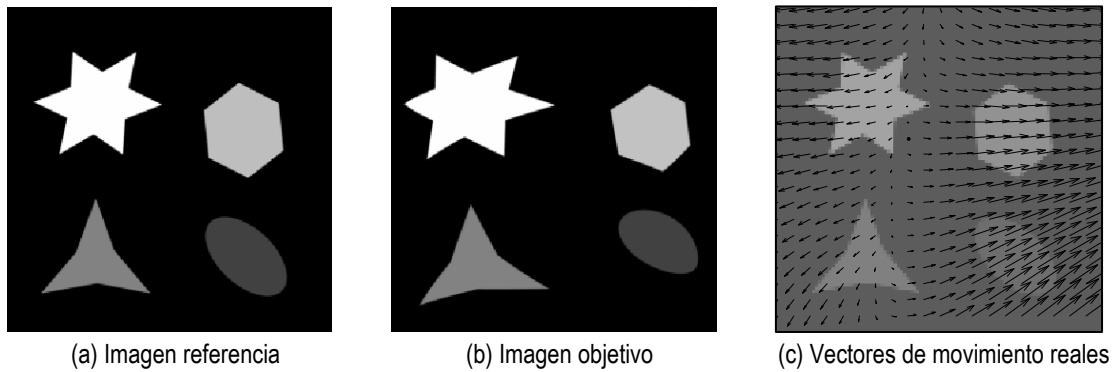


Figura 5.6. Imágenes para el registro del caso 3, junto con los vectores de movimiento reales.

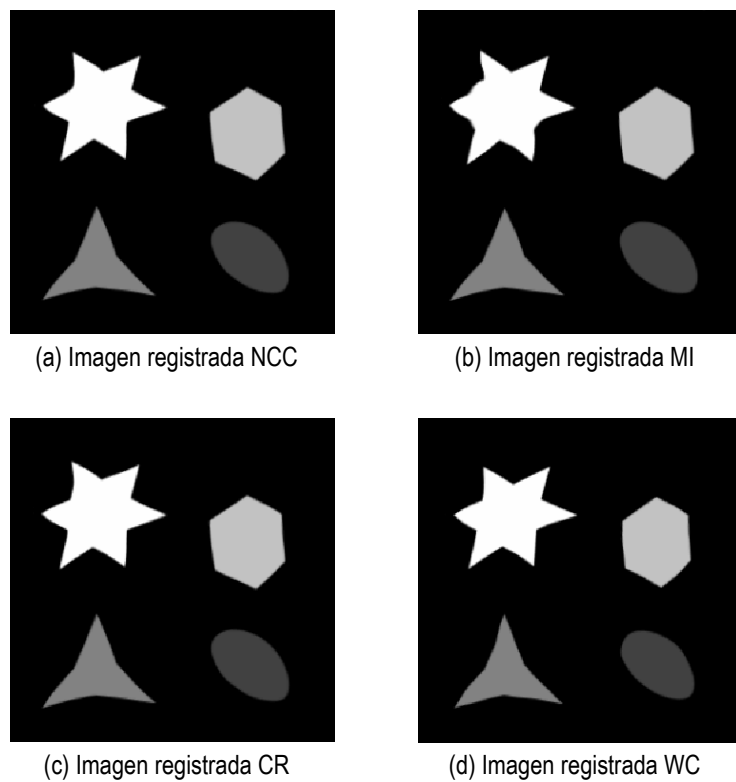


Figura 5.7. Resultado del registro para todas las medidas de similitud.

Como observamos, los distintos métodos proporcionan resultados muy similares, aunque con algunas particularidades. Por ejemplo, la MI y el CR crean ciertas fluctuaciones en los bordes de la figura en forma de estrella que el resto de medidas no producen. La Tabla 5.3 resume los resultados obtenidos:

Medida	PSNR	ECM	RIGIDEZ			ELASTICIDAD		
			Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	23.07 dB	0.0595	0.0706	0.0371	0.0042	0.0184	0.0011	0.0102
CR	22.92 dB	0.0650	0.0661	0.0355	0.0037	0.0131	0.0010	0.0024
WC	20.72 dB	0.1834	0.0726	0.0568	0.0031	0.0142	0.0007	0.0023
MI	21.12 dB	0.1617	0.0765	0.0394	0.0078	0.0613	0.0011	0.1505

Tabla 5.3. Resumen de los resultados obtenidos para el caso 3 de imágenes sintéticas.

En este caso, al tratarse de una transformación más sencilla, es posible reducir el error entre la transformación detectada y la real. Esta conclusión se refleja en el resultado obtenido por el ECM, que esta vez es menor y de nuevo vuelve a seguir al PSNR. Ahora es el NCC es que nos proporciona una mayor alineación entre las imágenes, mientras que el WC vuelve a ser el que peor resultados obtiene. La mayor rigidez y elasticidad nos las proporcionan de nuevo los métodos con peor PSNR, por lo que interesa que para este determinado caso estos valores sean menores.

5.3.2. Imágenes fotográficas

Las imágenes fotográficas son aquellas obtenidas mediante una máquina fotográfica, como una cámara de fotos. El registro de estas imágenes nos sirve para corregir las deformaciones producidas por la adquisición de dichas imágenes en un instante de tiempo diferente o desde una perspectiva distinta. Para estas imágenes, el conjunto de vectores de movimiento es desconocido, por lo que el objetivo del registro será corregir la desviación existente entre ambas imágenes.

5.3.2.1 Caso 1. Mujer de espaldas

La Figura 5.8 ilustra el par de imágenes de las que se parte en este primer caso. Se trata de un cambio de perspectiva en la toma de la fotografía, por lo que habrá que intentar aplicar una transformación lo suficientemente buena como para que parezca que en realidad la imagen registrada se ha adquirido desde la misma posición que la imagen de referencia.

En la Figura 5.9 podemos ver el resultado de aplicar al procedimiento de registro cada una de las cuatro medidas de similitud estudiadas, que nos proporcionan un resultado muy similar a simple vista. La acción principal que se realiza sobre la imagen objetivo es la de desplazar la zona

del lado izquierdo de la espalda de la mujer un poco más a su izquierda, levantando también ligeramente la zona del hombro, tal y como se aprecia en los resultados.

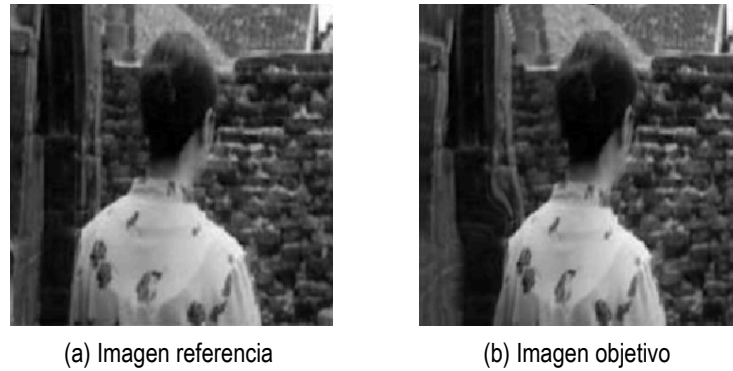


Figura 5.8. Par de imágenes fotográficas utilizadas para el registro del caso 1.



Figura 5.9. Resultado del registro para las cuatro medidas de similitud.

El conjunto de vectores detectados en cada caso se representa en la Figura 5.10. En ella también se muestra el resultado de aplicar estos vectores (recordemos que se aplican en sentido

opuesto a los mostrados) a una rejilla rectangular del tamaño de la imagen. Es fácil ver que cada método proporciona un movimiento peculiar, pero a veces estas diferencias no son apreciables a simple vista (véase la Figura 5.9).

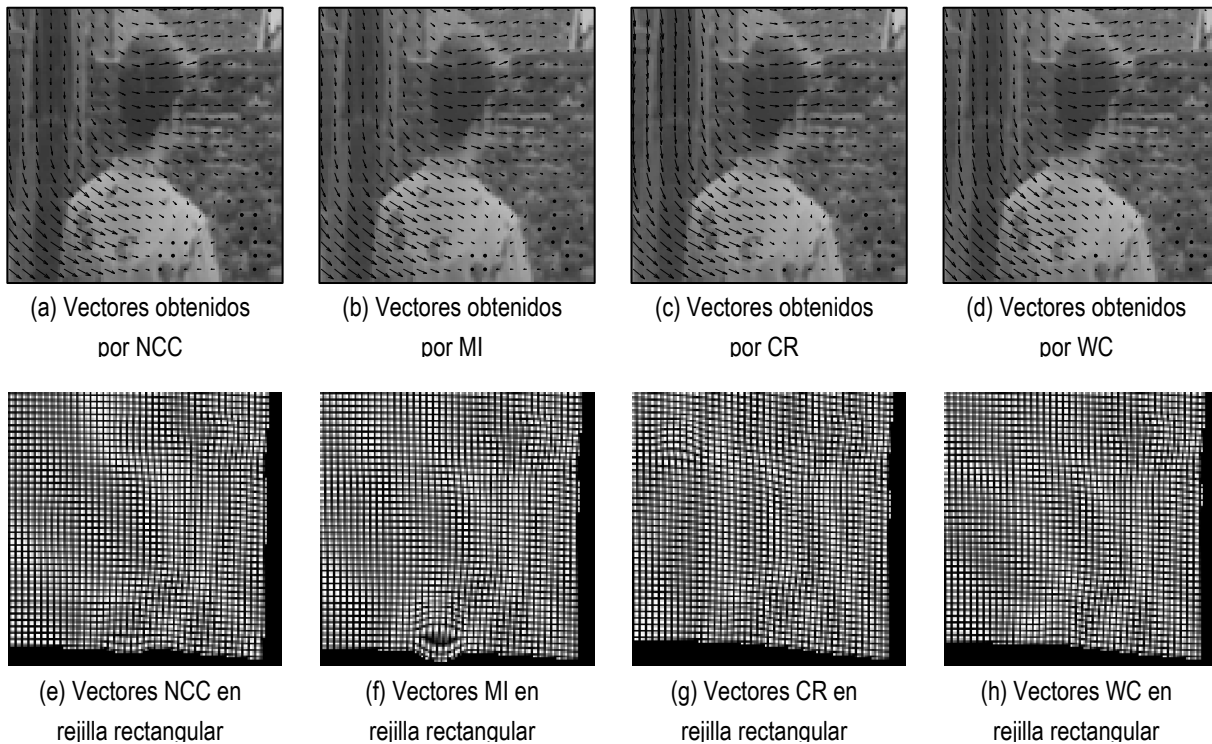


Figura 5.10. Vectores de movimiento detectados y su aplicación a una rejilla rectangular.

La siguiente tabla (Tabla 5.4) resume los resultados obtenidos para cada uno de los métodos empleados:

Medida	PSNR	RIGIDEZ			ELASTICIDAD		
		Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	22.94 dB	0.0314	0.0222	0.0010	0.0306	0.0062	0.0054
CR	23.01 dB	0.0371	0.0290	0.0011	0.0397	0.0073	0.0104
WC	22.67 dB	0.0320	0.0217	0.0010	0.0306	0.0051	0.0066
MI	22.84 dB	0.0333	0.0216	0.0017	0.0445	0.0060	0.0331

Tabla 5.4. Resumen de los resultados obtenidos para el caso 1 de imágenes fotográficas.

Como vemos, el registro que produce un mejor alineamiento entre la imagen referencia y la registrada es el ratio de correlación, seguido por el coeficiente de correlación, la información mutua y el criterio de Woods. Este último método sigue demostrando ser el que peor resultado nos proporciona en cuanto a similitud entre la imagen registrada y la de referencia.

5.3.2.2 Caso 2. Planta química

El segundo caso que vamos a tratar en el registro de imágenes fotográficas es el de una planta química de la que se han tomado fotografías desde dos diferentes perspectivas. Ambas imágenes se muestran en la Figura 5.11.

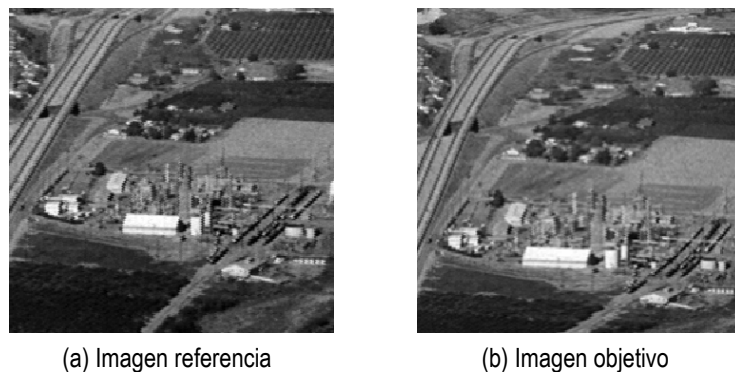


Figura 5.11. Par de imágenes fotográficas utilizadas para el registro del caso 2.

Como vemos, la imagen objetivo se ha adquirido desde una posición inferior a la de la imagen de referencia. Además, la imagen se encuentra ensanchada en su zona central, por lo que básicamente el procedimiento de registro deberá tratar de corregir este desplazamiento vertical y conseguir a su vez que la parte central se comprima horizontalmente. Los resultados de este registro para las cuatro técnicas de cálculo de la similitud se exponen en la Figura 5.12.

Tal y como muestran las imágenes de la Figura 5.12, la presencia de un borde negro inferior sugiere que los datos antes presentes en esa posición se han desplazado hacia otro lugar. Ese lugar en este caso es el centro de la imagen. También se ha producido el efecto contrario en el borde superior, donde los datos han desaparecido al ser sustituidos por los provenientes de píxeles de posiciones inferiores. En los bordes laterales se aprecia el efecto de la compresión llevada a cabo para solventar el ensanchamiento que había sufrido la imagen objetivo frente a la de referencia.

El movimiento detectado por el registro se observa mejor en la Figura 5.13, donde también se han aplicado los vectores calculados inversos a una rejilla rectangular del tamaño de las imágenes originales.

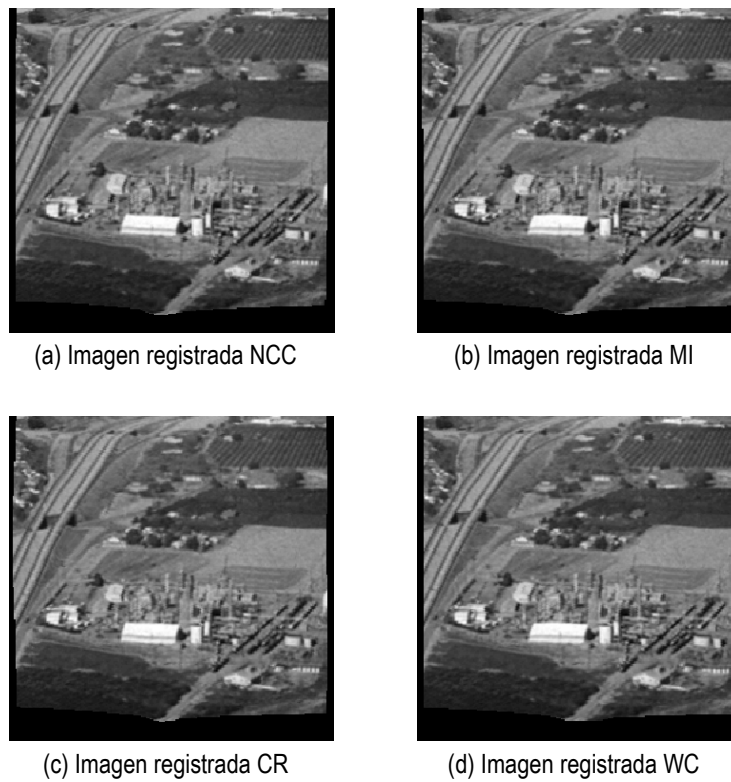


Figura 5.12. Resultado del registro para las cuatro medidas de similitud.

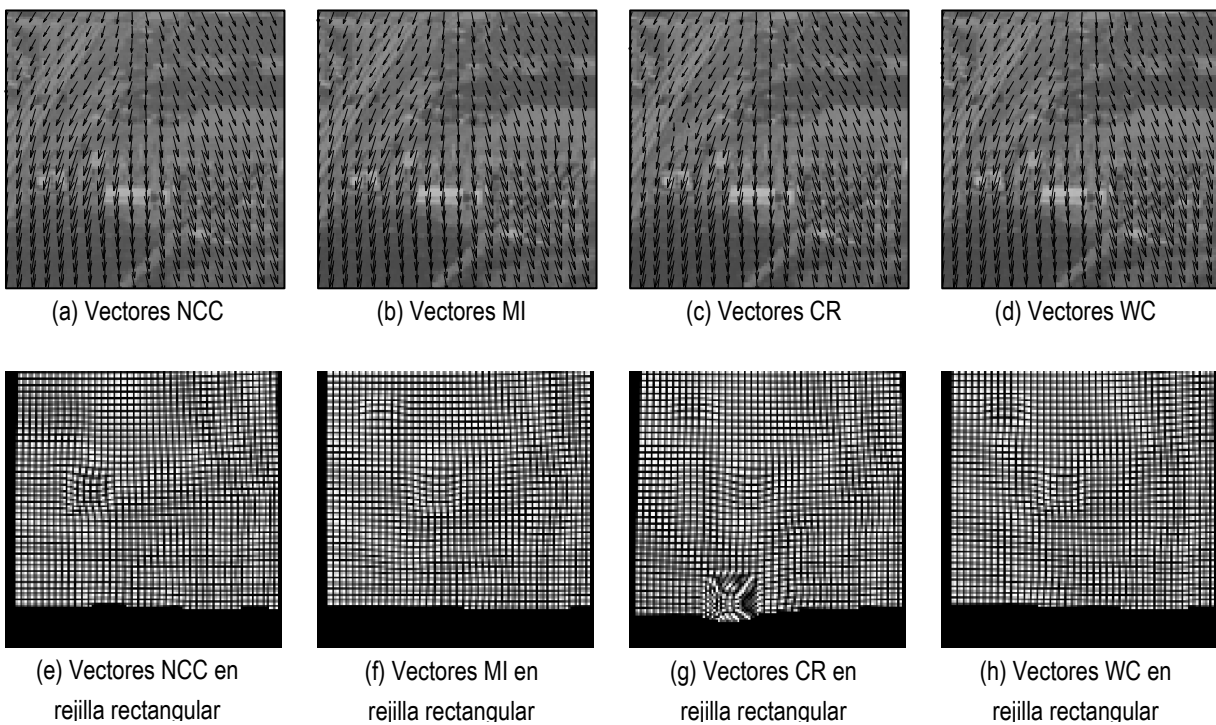


Figura 5.13. Vectores de movimiento detectados y su aplicación a una rejilla rectangular.

La Tabla 5.5 sintetiza los parámetros más relevantes del registro obtenidos para este caso. De nuevo conseguimos valores del PSNR muy cercanos entre todos los métodos utilizados, lo que hace que por simple visualización de las imágenes sea muy complicado detectar cualquier fallo de correspondencia. El método que nos proporciona una mayor rigidez y elasticidad en su transformación es el ratio de correlación, debido precisamente a esos cambios poco suaves que se observan a la izquierda de la Figura 5.13(c). Por el contrario, el criterio de Woods obtiene los vectores con mayor suavidad entre todas las medidas de similitud.

Medida	PSNR	RIGIDEZ			ELASTICIDAD		
		Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	17.19 dB	0.0445	0.0296	0.0024	0.0696	0.0067	0.3183
CR	17.48 dB	0.0600	0.0305	0.0220	0.1225	0.0075	0.7137
WC	17.44 dB	0.0428	0.0289	0.0019	0.0554	0.0072	0.0436
MI	17.64 dB	0.0469	0.0469	0.0026	0.0752	0.0075	0.2291

Tabla 5.5. Resumen de los resultados obtenidos para el caso 2 de imágenes fotográficas.

5.3.3. Imágenes aéreas

Las imágenes aéreas son aquellas imágenes tomadas desde el aire mediante algún tipo de dispositivo especial, y sus aplicaciones son diversas, yendo desde la localización de blancos militares a la detección de construcciones para una zona determinada tras un cierto intervalo de tiempo.

5.3.3.1 Caso 1. Centro residencial

En este primer caso partimos de dos imágenes tomadas desde la misma fuente, por lo que el hecho de que ambas sean parecidas nos permite que los métodos para la realización del registro sean los mismos que los vistos en ejemplos anteriores. Las imágenes de referencia y objetivo se muestran en la Figura 5.14, en las cuales se pueden distinguir dos imágenes aéreas pertenecientes a una zona residencial, donde las características más importantes son las intersecciones entre carreteras y algunas edificaciones. Estas características nos permiten una mejor distinción del movimiento entre ellas.

La Figura 5.15 ilustra por su parte el resultado de la imagen registrada, al haber aplicado durante el proceso de registro cada una de las medidas de similitud. Al no haber un gran movimiento entre ambas imágenes, éstas sufren solamente ligeras transformaciones, lo que tiene como consecuencia la aparición de bordes negros de zonas que se han desplazado hacia el interior de la imagen.

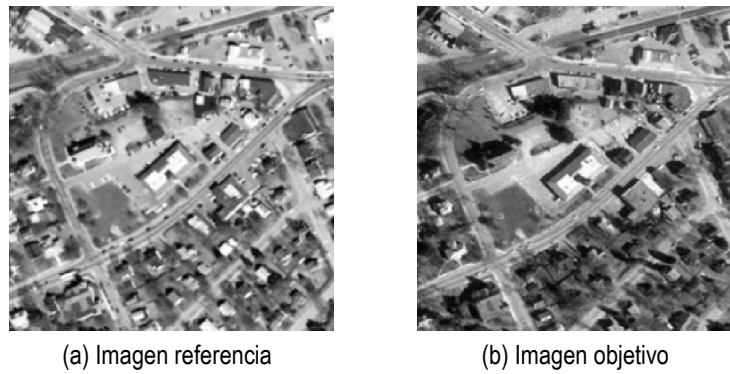


Figura 5.14. Imágenes aéreas de un centro residencial utilizadas para el registro del caso 1.

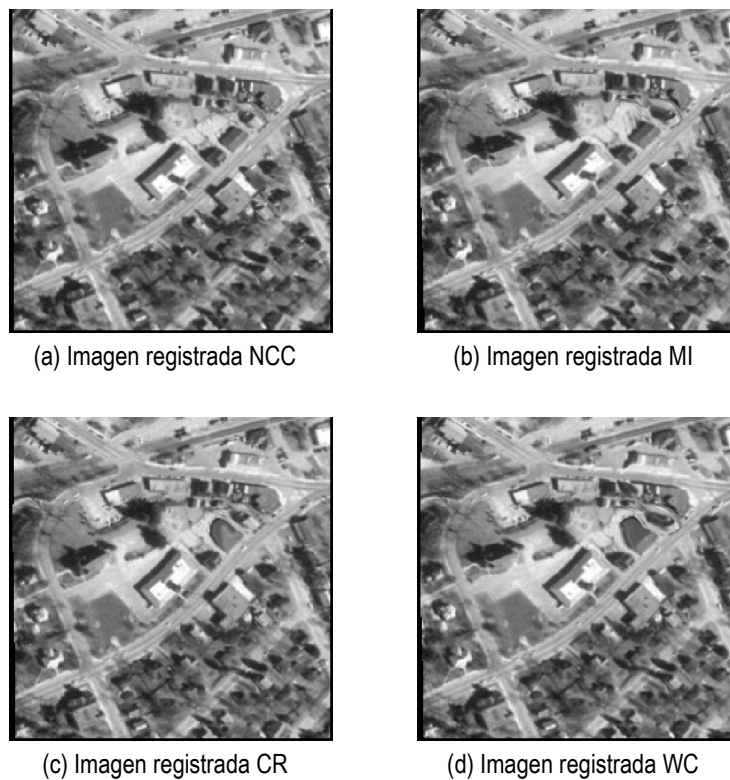


Figura 5.15. Resultado del registro para las cuatro medidas de similitud.

Los vectores de transformación obtenidos pueden visualizarse en la Figura 5.16. En ellos es posible ver que el NCC realiza una detección completamente suave, mientras que el resto de los casos detectan zonas donde el movimiento es mayor. Esto es debido a la propia naturaleza de estas tres medidas, que nos permiten un alto valor de la similitud incluso para imágenes con distintos niveles de gris.

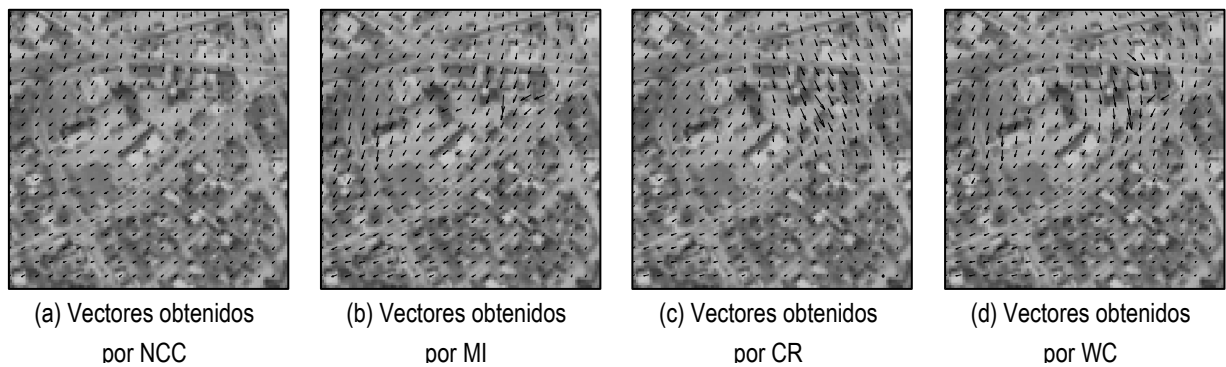


Figura 5.16. Vectores de movimiento detectados.

La tabla siguiente (Tabla 5.6) resume los datos concernientes al PSNR, la elasticidad y rigidez de los vectores de movimiento, donde podemos reafirmar lo dicho anteriormente, y es que el NCC realiza una transformación completamente suave, mientras que el CR y el WC llevan a cabo un estiramiento de una de las naves de la figura, provocando un PSNR menor que en los otros dos casos.

Medida	PSNR	RIGIDEZ			ELASTICIDAD		
		Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	14.37 dB	0.0013	0.6081e-3	0.0136e-3	0.0024	0.0087e-3	0.2448e-3
CR	13.95 dB	0.0219	0.0032	0.0099	0.0572	0.0002	0.3187
WC	14.02 dB	0.0457	0.0049	0.0382	0.1731	0.0010	1.3669
MI	14.13 dB	0.0264	0.0031	0.0120	0.1131	0.0005	0.4563

Tabla 5.6. Resumen de resultados para el registro de un área residencial.

5.3.3.2 Caso 2. Costa de San Diego (EEUU)

Para este caso se analiza de nuevo una imagen aérea, pero esta vez donde hay presente una cierta rotación entre la imagen referencia y objetivo, así como varias deformaciones locales existentes. Dichas imágenes corresponden a parte de la costa de San Diego, en EEUU, y pueden apreciarse en la Figura 5.17.

Las transiciones entre tierra y mar constituyen las características más importantes que nos permiten detectar la deformación existente entre las imágenes de la Figura 5.17. El nivel de gris en ellas es lo suficientemente particular como para no obviar las zonas donde se produce un alto emparejamiento, lo que originará que los resultados obtenidos para todos los casos sean de bastante calidad.

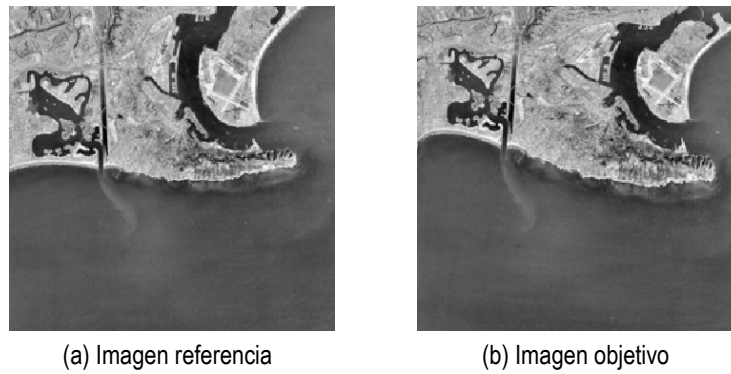


Figura 5.17. Imágenes aéreas de la costa de San Diego utilizadas para el registro del caso 2.

La Figura 5.18 muestra el resultado del registro de imagen para las imágenes de partida, donde se observa claramente la rotación aplicada a la imagen. La Figura 5.19, por su parte, ilustra los vectores de transformación detectados en la imagen objetivo. También se incluye el resultado de aplicar estos vectores a una rejilla rectangular, que nos permita observar mejor la transformación aplicada a la imagen de la Figura 5.17(b) con el fin de llevar a cabo el registro.

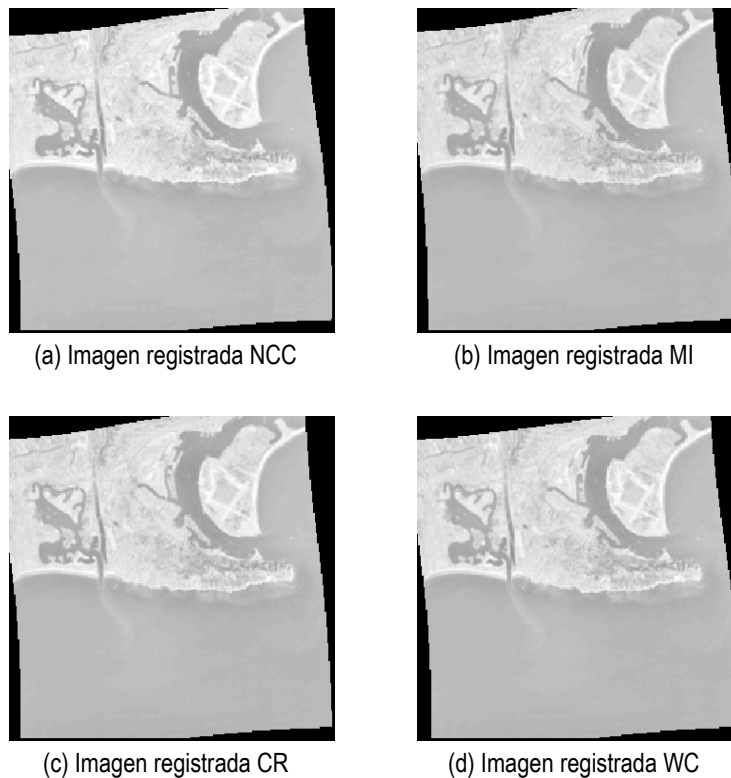


Figura 5.18. Resultado del registro para las cuatro medidas de similitud.

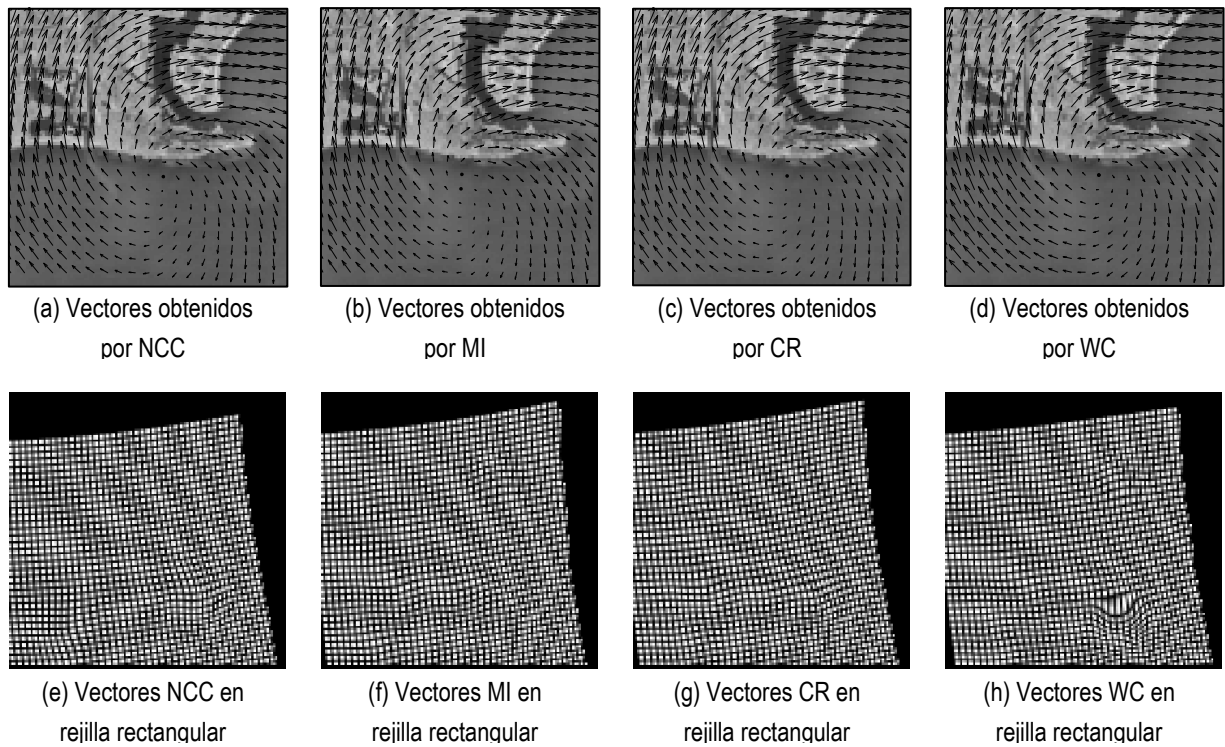


Figura 5.19. Vectores de movimiento detectados y su aplicación a una rejilla rectangular.

Como se puede apreciar, el hecho de poseer información presente en una de las imágenes pero no presente en la otra hace que tras el registro haya zonas que están en la imagen registrada pero que en la imagen referencia no aparecen. Sin embargo, la imagen se reorienta para conseguir la mayor coincidencia de puntos significativos, como en este caso podría ser toda la zona de transición entre tierra y agua, muy sencilla de identificar para los cuatro métodos de similitud.

La Tabla 5.7 muestra el grado de similitud y las características de los vectores hallados en los cuatro casos.

Medida	PSNR	RIGIDEZ			ELASTICIDAD		
		Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	8.53 dB	0.0875	0.0622	0.0062	0.0243	0.0008	0.0059
CR	8.66 dB	0.0946	0.0591	0.0104	0.0510	0.0012	0.0800
WC	8.59 dB	0.0990	0.0630	0.0108	0.0558	0.0014	0.0857
MI	8.67 dB	0.0912	0.0607	0.0088	0.0521	0.0013	0.0460

Tabla 5.7. Resumen de resultados para el registro de una fotografía aérea de la costa de San Diego.

Como se ve, la transformación llevada a cabo por los cuatro métodos es de tipo circular en los extremos y casi nula en el centro, aunque con transiciones suaves entre ellas. Este comportamiento se refleja claramente en los resultados de la rigidez y la elasticidad. De nuevo el PSNR tiende a ser muy similar para todos ellos, con una ligera ventaja de la MI y el CR.

5.3.4. Imágenes médicas

La aplicación del registro de imágenes a la medicina es un campo de gran actualidad hoy en día. Estas imágenes nos permiten conocer la evolución de tumores, el movimiento de ciertos órganos o la comparación de cuerpos de distintos pacientes por medio del registro, lo que resulta de gran ayuda para el diagnóstico y prevención de enfermedades.

5.3.4.1 Caso 1. Imagen MRI cerebral

La Figura 5.20 muestra un ejemplo de registro para dos imágenes de tipo MRI, donde la imagen objetivo está desplazada y deformada frente a la de referencia. El MRI (*Magnetic Resonance Imaging*) es un método utilizado para la visualización en el interior de organismos vivos. Se usa principalmente para demostrar alteraciones de tipo patológico o fisiológico y es una forma común de imagen médica. Así pues, el objetivo del registro podría ser alinear dichas imágenes para su almacenamiento en una base de datos, o simplemente comparar la evolución del cerebro tras un cierto periodo de tiempo.

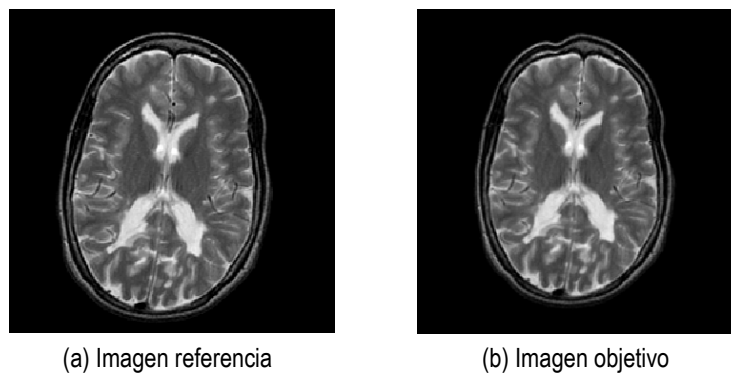


Figura 5.20. Imágenes cerebrales MRI utilizadas para el registro del caso 1.

El resultado del registro de las imágenes anteriores mediante las cuatro medidas de similitud estudiadas se ilustra en la Figura 5.21.

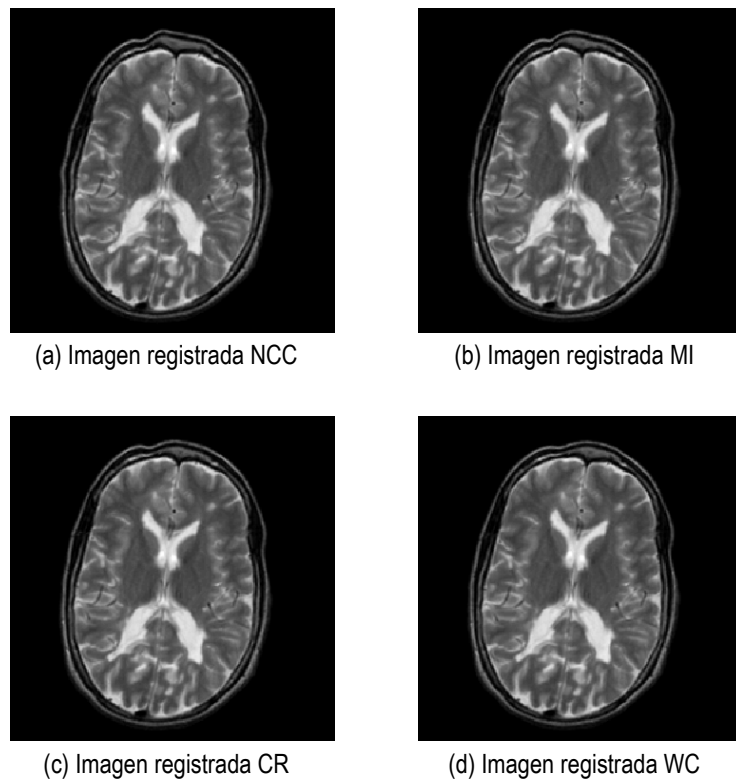


Figura 5.21. Resultado del registro para las cuatro medidas de similitud.

Como observamos, las imágenes registradas han sufrido un aumento de tamaño que la imagen objetivo necesitaba frente a la imagen referencia, y además se ha producido un pequeño giro hacia la derecha para conseguir el alineamiento. La hendidura de la parte frontal izquierda también se ha disminuido un poco, pero no lo suficiente como para igualar a la imagen de referencia.

La Figura 5.22 representa la rotación, la expansión y los movimientos comentados, a través de los vectores de transformación:

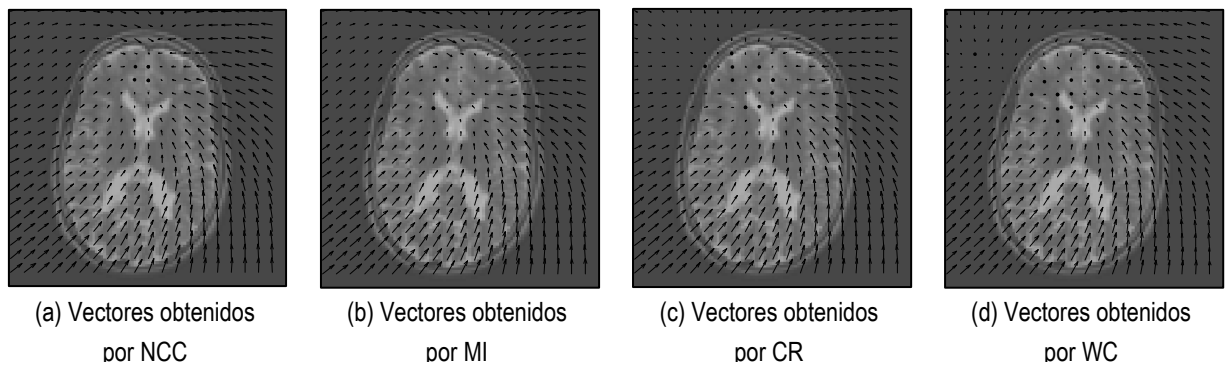


Figura 5.22. Vectores de movimiento detectados.

La Tabla 5.8 resume los resultados concernientes a la similitud de la imagen registrada y la imagen referencia mediante el PSNR, y las características de los vectores de movimiento. En ella se puede observar un alto grado de emparejamiento entre las imágenes (muchos dBs), pero esto es consecuencia directa del hecho de que el fondo de la imagen sea negro, el mismo valor que se inserta cuando los píxeles de una posición se desplazan a otro lugar. El criterio de Woods se muestra como el peor resultado, mientras que es el NCC el que obtiene un mejor alineamiento entre las imágenes.

Medida	PSNR	RIGIDEZ			ELASTICIDAD		
		Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	26.86 dB	0.0596	0.0275	0.0152	0.0026	0.0433	0.0023
CR	26.02 dB	0.1817	0.0204	0.0132	0.0004	0.0146	0.0011
WC	24.62 dB	0.3252	0.0264	0.0148	0.0011	0.0457	0.0038
MI	26.58 dB	0.1019	0.0274	0.0162	0.0014	0.0320	0.0015

Tabla 5.8. Resumen de resultados para el registro de una imagen médica MRI.

5.3.4.2 Caso 2. Imagen CT del tórax

La tomografía computerizada (CT) es un tipo de imagen médica que se emplea para generar una imagen tridimensional del interior de un objeto, a través de una serie de imágenes de rayos X bidimensionales alrededor de un solo eje de rotación. Un par de estas imágenes serán utilizadas para llevar a cabo el registro de este segundo caso. Dichas imágenes se pueden visualizar en la siguiente figura (Figura 5.23).

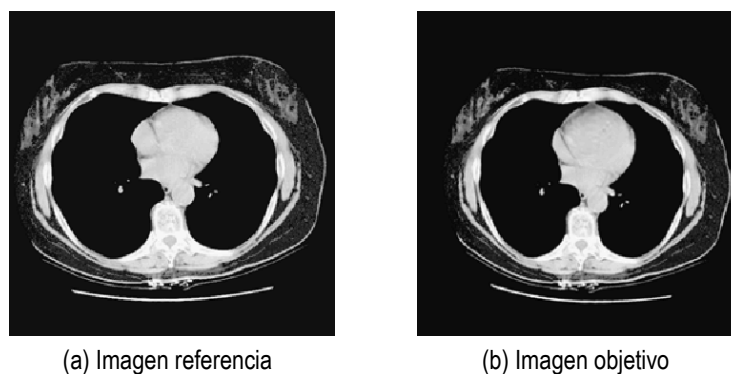


Figura 5.23. Imágenes CT de un tórax utilizadas para el registro del caso 2.

En la Figura 5.23(b) se puede observar que el tórax está ligeramente rotado hacia la derecha con respecto a la imagen de referencia, y además parece que el tamaño de éste es menor. También se hace presente una ligera deformación en la parte central de la imagen CT, que aparece más estrecha que en la imagen de referencia. Los resultados de dichas correcciones se muestran en las imágenes registradas de la Figura 5.24.

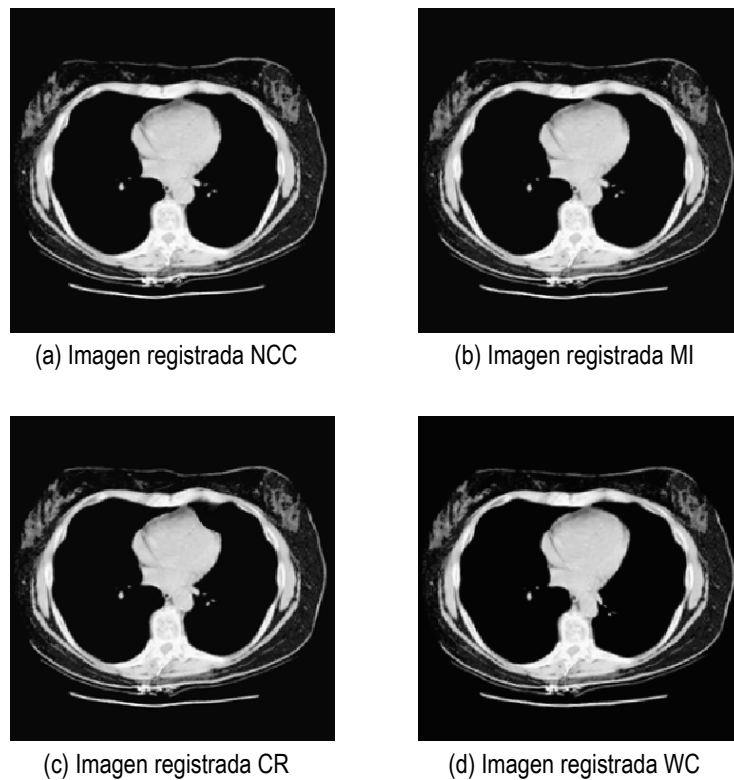


Figura 5.24. Resultado del registro para las cuatro medidas de similitud.

Como vemos en las imágenes anteriores, parece que todas las medidas de similitud empleadas consiguen un registro del contorno satisfactorio, pero en el CR por ejemplo se produce una deformación en la parte central de la imagen que el resto no hace. Esta idea se observa mejor mediante los vectores de movimiento y su aplicación a una rejilla rectangular, ambos representados en la Figura 5.25.

A partir de las imágenes de la Figura 5.25 es fácil apreciar las diferencias entre los distintos métodos de similitud, ya que cada uno de ellos genera deformaciones distintamente al resto. Estas deformaciones se traducen en zonas donde los movimientos resultan ser menos suaves, que son además las causantes de estas diferencias apreciables entre métodos.

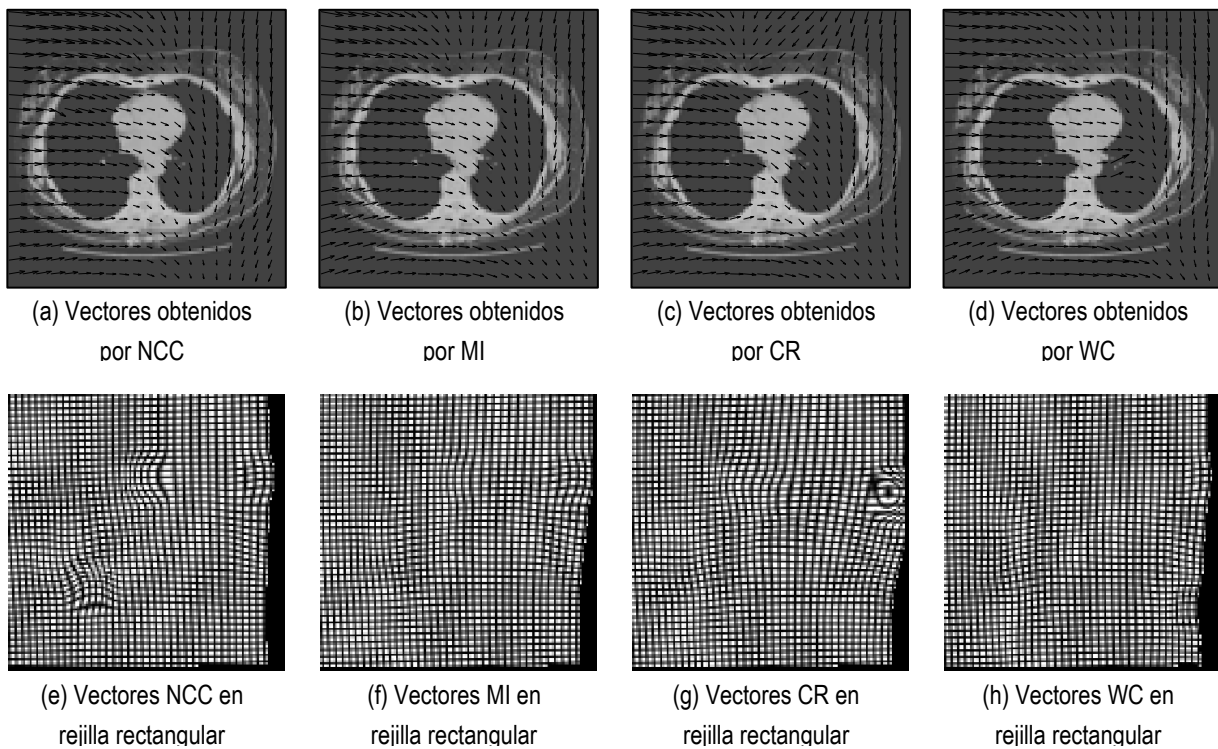


Figura 5.25. Vectores de movimiento detectados y su aplicación a una rejilla rectangular.

La Tabla 5.9 resume las características del registro para los distintos casos de medidas de similitud:

Medida	PSNR	RIGIDEZ			ELASTICIDAD		
		Media	Mediana	Varianza	Media	Mediana	Varianza
NCC	19.37 dB	0.0457	0.0288	0.0028	0.0607	0.0037	0.0415
CR	18.84 dB	0.0727	0.0292	0.0288	0.1430	0.0059	0.7087
WC	18.01 dB	0.0640	0.0332	0.0183	0.1254	0.0072	0.5512
MI	18.91 dB	0.0383	0.0251	0.0018	0.0503	0.0034	0.0275

Tabla 5.9. Resumen de resultados para el registro de una imagen médica CT.

Los resultados contenidos en la tabla anterior reflejan el buen funcionamiento del NCC en cuanto al registro del par de imágenes anterior, dejando de nuevo al WC como la peor alternativa de medida de similitud a emplear. Las variaciones poco suaves en la transformación, presentes en el CR y el WC principalmente, contribuyen a un peor registro de las imágenes, principalmente reflejado en el PSNR.

5.3.5. Imágenes multimodo

Una vez que se ha comprobado el funcionamiento de las cuatro medidas de similitud estudiadas para registros de naturaleza monomodo, llega el turno de conocer el comportamiento cuando la procedencia de las imágenes es de distintas fuentes. Aunque el problema del registro multimodo podría ser solucionado en algunos casos mediante una transformación de las imágenes a un mundo pseudo-monomodo, por ejemplo con el uso de una corrección de los niveles de gris como la estudiada en apartados anteriores, procuraremos utilizar esta técnica exclusivamente para el cálculo de la diferencia cuadrática, de tal modo que se aprecie la funcionalidad de los métodos para el cálculo de la similitud en imágenes multimodo. Así pues, el NCC se convierte en una medida ineficiente y no capaz de funcionar correctamente para el cálculo de la similitud, lo que limita nuestra aplicación al ECC, CR y WC.

El método para la realización del registro será el mismo utilizado para los casos de registro monomodo vistos en los apartados anteriores, con la única salvedad de que habrá que realizar una transformación de niveles de gris previa al cálculo de la diferencia cuadrática (apartado 3.7.2 y 4.4.1.1). Esto conlleva también un cambio en los coeficientes de la expresión que relaciona la similitud y la diferencia cuadrática con la umbralización. Otra posibilidad es el uso exclusivo de la medida de similitud sin hacer uso de la diferencia, pero esta opción adolece de requerir de un mayor ajuste del parámetro de umbralización para la obtención de un resultado satisfactorio.

5.3.5.1 Caso 1. Tórax CT-PET

En este primer caso se utiliza el método descrito para realizar el registro de dos imágenes CT y PET, correspondientes a un tórax humano. Así, la primera imagen es de tipo CT y la segunda es una imagen de tipo PET, que se representan en la Figura 5.26.

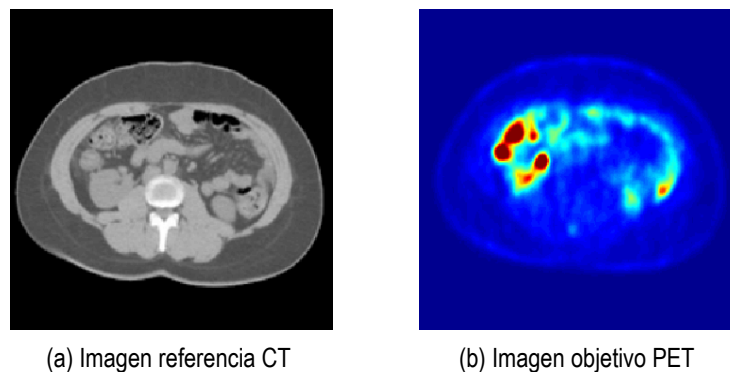


Figura 5.26. Imágenes multimodo de un tórax utilizadas para el registro del caso 1.

Al no ser el NCC válido para el registro multimodo por su comportamiento tan dependiente del nivel de gris, el resultado del registro para los tres casos restantes se ilustra en la Figura 5.27, junto con los vectores de movimiento calculados.

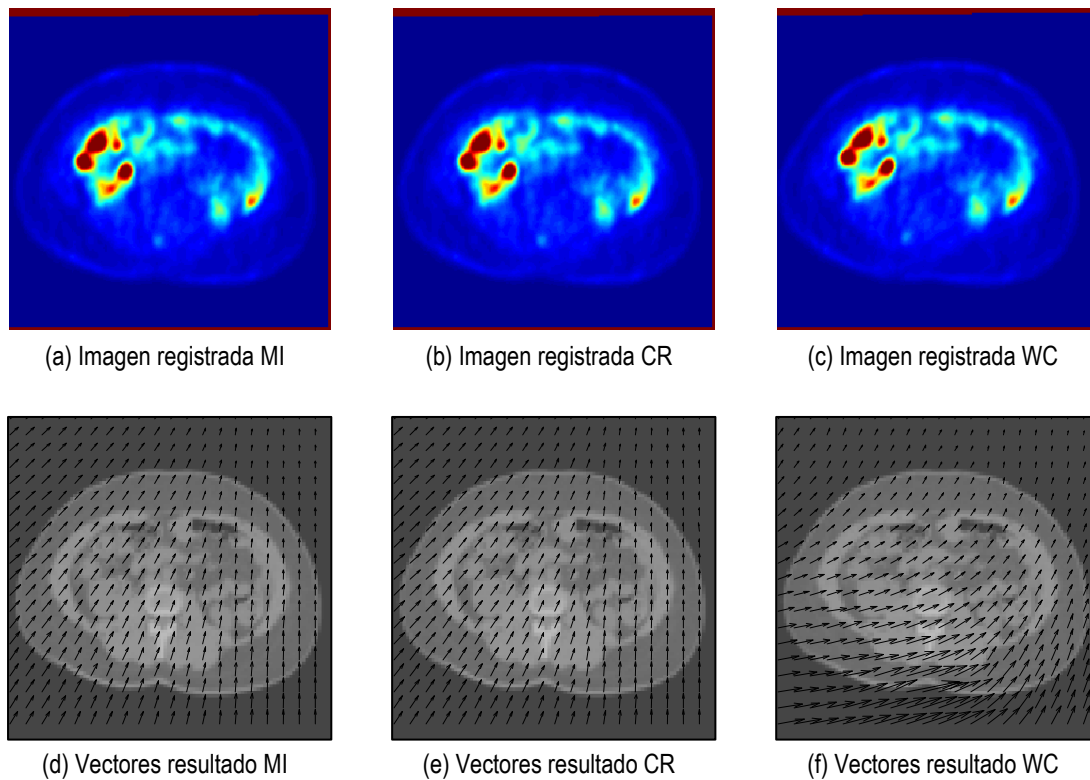


Figura 5.27. Resultado del registro para MI, CR y WC.

Como vemos en la figura anterior, las dos primeras medidas de similitud proporcionan un valor muy similar, mientras que el criterio de Woods detecta zonas donde la deformación es mayor, provocando un ligero movimiento no presente en la imagen PET frente a la CT. Sin embargo, el movimiento general de la totalidad de la imagen, que es el que en realidad nos interesa para conseguir el alineamiento, sí que logra corregirse para los tres métodos empleados. Para el cálculo de los mapas de similitud de las imágenes anteriores se ha hecho uso de un factor de ponderación de la diferencia cuadrática de $k=3.5$ y de un umbral de aproximadamente 0.3.

Para el registro de las imágenes anteriores, como se ha comentado, se ha hecho uso de una corrección de niveles de gris en lo que se refiere al cálculo de la diferencia cuadrática. Para que este resultado sea mejor, se ha decidido realizar una corrección para cada una de las imágenes, haciendo en un primer paso la corrección de la imagen PET hacia la CT, y en un segundo la corrección de la imagen de referencia con la nueva imagen PET corregida. El resultado de las dos imágenes corregidas se muestra en la siguiente figura (Figura 5.28) junto con las curvas de ajuste (para este caso se ha elegido un polinomio de quinto orden) de los histogramas correspondientes, donde no se ha realizado ninguna umbralización previa.

Como podemos ver en dicha figura, se consigue que el fondo de la imagen tipo PET se iguale con la de tipo CT, y además, el nivel de la sección del tórax de ambas imágenes se aproxima en cuanto a magnitud, lo que para el cálculo de la diferencia cuadrática supone una mejora sustancial. Los bordes de las imágenes quedan ahora mucho más parecidos, haciendo que sean

estos puntos los que permitan hallar la transformación necesaria para registrar la imagen PET a la posición de la imagen CT.

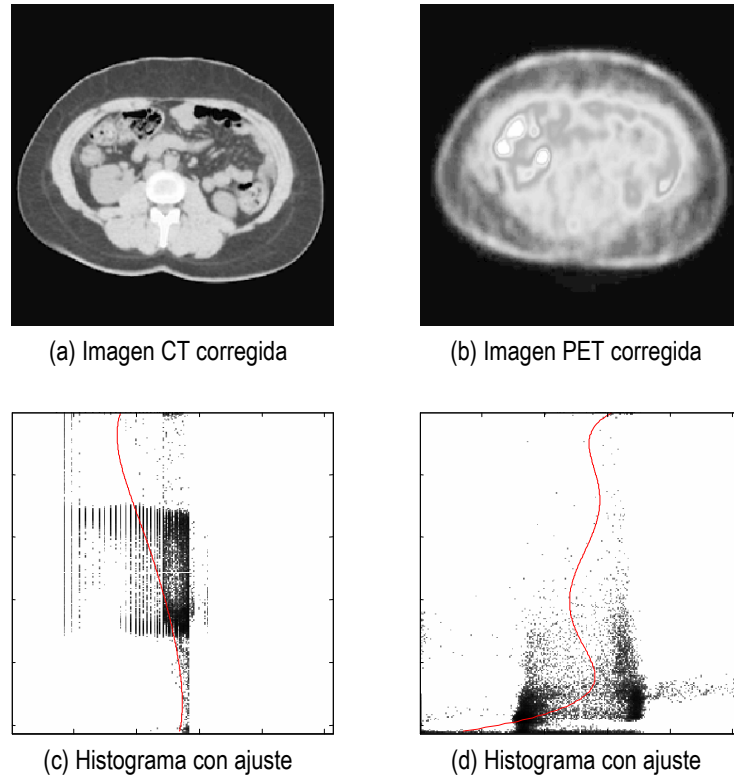


Figura 5.28. Corrección de grises efectuada en el caso 1.

Una manera espléndida de representar los resultados es mediante la técnica de fusión de imágenes, ya introducida en el capítulo 2, que proporciona una visualización conjunta aplicando el arte de la transparencia. Para el caso particular de un registro de tipo CT/PET, podemos describir la expresión (2.1) de la siguiente manera:

$$[r,g,b]_{fusion} = \alpha[r,g,b]_{CT} + (1-\alpha)[r,g,b]_{PET}, \quad (5.4)$$

donde r , g y b son los canales de rojo, verde y azul respectivamente de la imagen fusionada, y el valor del factor α puede estar entre 0 y 1. Si $\alpha = 0$, solamente se visualizará la imagen de PET y si $\alpha = 1$, solamente se representará la imagen de CT.

En la Figura 5.29 se representa la fusión para los tres métodos anteriores con un factor de transparencia de $\alpha = 0.3$. Ahora es posible integrar la información funcional (PET) con la estructural (CT) en una misma representación.

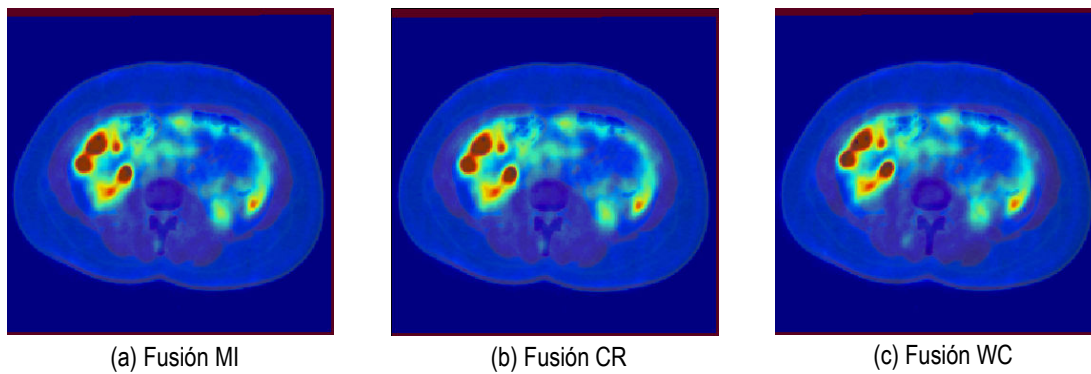


Figura 5.29. Resultado de la fusi3n de im3genes PET y CT para un $\alpha = 0.3$.

En la siguiente tabla (Tabla 5.10) se representan los valores estadísticos para los vectores de desplazamiento calculados. En este caso ya no resulta posible utilizar el PSNR para la medida de la correspondencia entre im3genes, al estar ahora los niveles de gris diferentemente distribuidos por tratarse de distintas fuentes. Como vemos, los vectores de movimiento del WC presentan valores superiores, tal y como se podía observar en los vectores de movimiento calculados para este m3todo.

Medida	RIGIDEZ			ELASTICIDAD		
	Media	Mediana	Varianza	Media	Mediana	Varianza
CR	0.0044	0.0037	0.0285e-3	0.0042	0.0001	0.3316e-3
WC	0.0663	0.0164	0.0681	0.1872	0.0001	2.5708
MI	0.0037	0.0035	0.0072e-3	0.0025	0	0.1224e-3

Tabla 5.10. Resumen de resultados para el registro multimodo de un t3rax.

5.3.5.2 Caso 2. Cerebro CT-PET

En este segundo caso vamos a proceder a realizar el registro de una secci3n cerebral, en lugar de un t3rax como en el ejemplo anterior. El procedimiento para el c3lculo de los vectores de movimiento a aplicar seguir3 las mismas etapas comentadas anteriormente, y de nuevo evaluaremos las tres t3cnicas multimodo para comparar el resultado que 3stas nos proporcionan.

En la Figura 5.30 representamos las dos figuras mencionadas: una de tipo estructural (CT) y otra de tipo funcional (PET). El objetivo ser3 por tanto integrar ambas tipos de informaci3n en una sola imagen que permita detectar un malfuncionamiento o cualquier tipo de anomalía presente en la actividad cerebral.

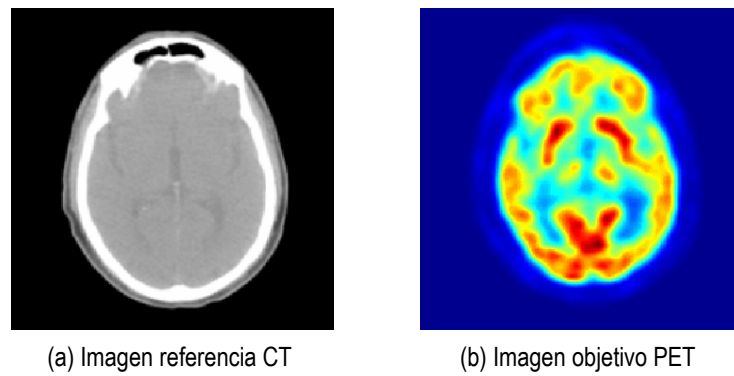


Figura 5.30. Imágenes multimodo de un cerebro, utilizadas para el registro del caso 2.

Para la realización del registro se ha hecho uso de un factor de ponderado de la diferencia de $k=3.5$, mientras que la umbralización se ha fijado en torno a 0.3 de nuevo. El resultado de la fusión con un $\alpha = 0.3$ para los tres métodos, junto con los vectores de movimiento detectados, se ilustra en la Figura 5.31.

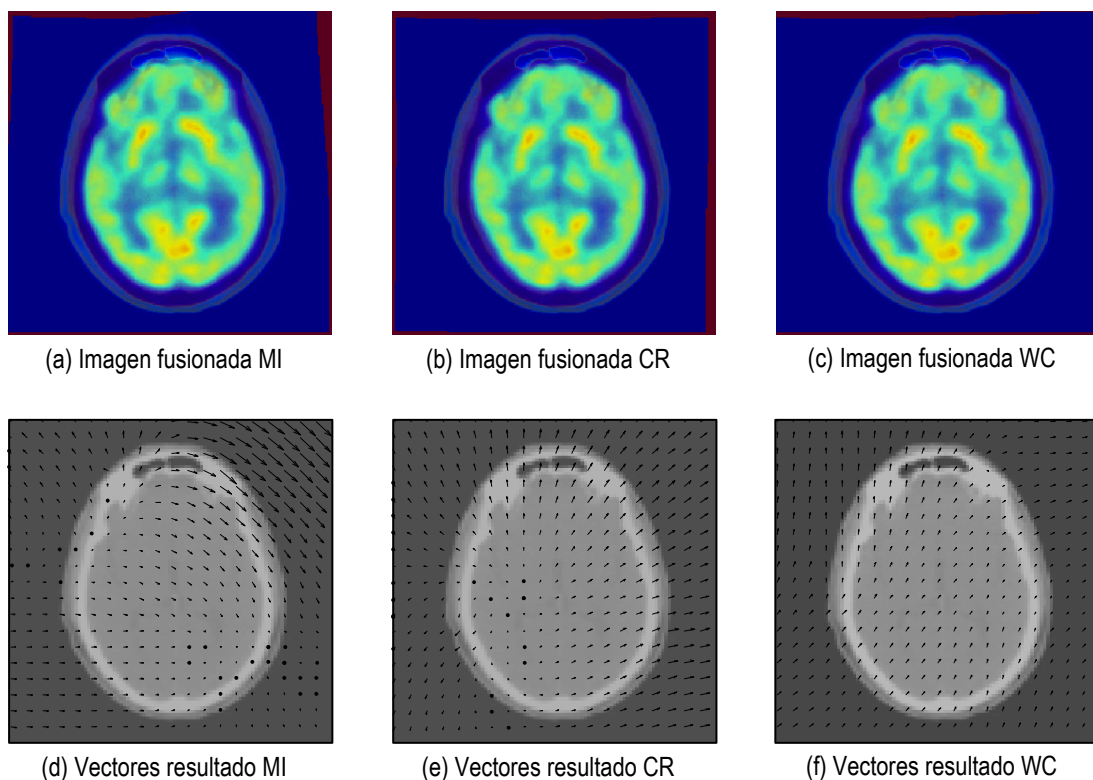


Figura 5.31. Resultados obtenidos para MI, CR y WC.

Como podemos apreciar en la figura anterior, el CR y el WC proporcionan un resultado muy similar, mientras que la MI realiza una detección más fuerte en la esquina superior derecha de la

imagen. Así pues vemos que se obtiene la alineación deseada entre las imágenes, lo que nos proporciona una correcta fusión CT-PET.

La corrección de grises llevada a cabo para las dos imágenes se muestra en la Figura 5.32.

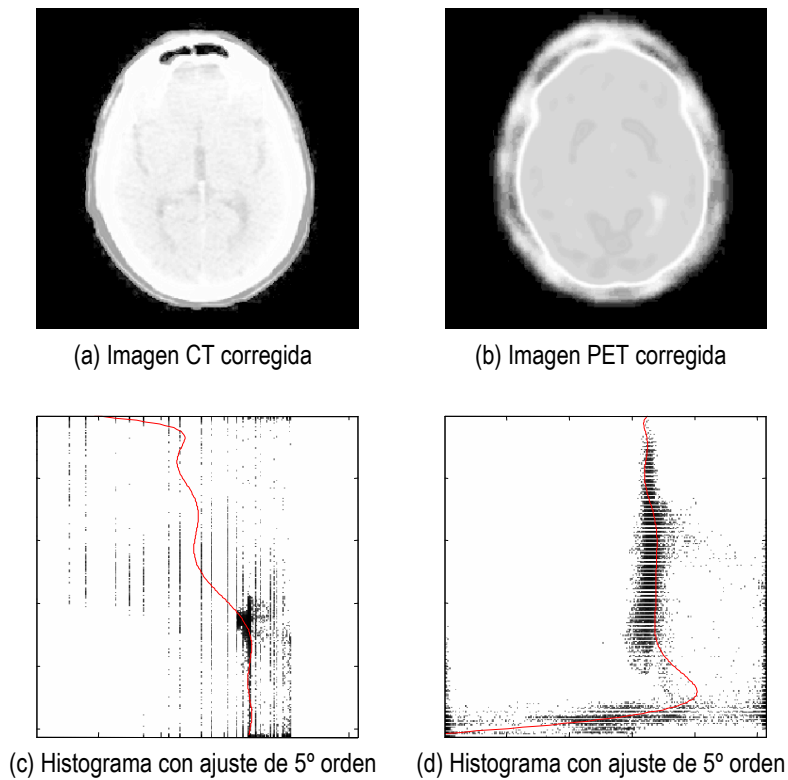


Figura 5.32. Corrección de grises efectuada en el caso 2.

La Tabla 5.11 resume los valores estadísticos para la rigidez y elasticidad de los vectores de movimiento detectados, aplicados posteriormente para la realización del registro multimodo. En ella podemos apreciar el efecto de los vectores para la MI comentados anteriormente, que producen valores mayores en la tabla que en los otros dos casos.

Medida	RIGIDEZ			ELASTICIDAD		
	Media	Mediana	Varianza	Media	Mediana	Varianza
CR	0.0085	0.0062	0.0001	0.0070	0.0001	0.0016
WC	0.0033	0.8088e-3	0.0001	0.0089	0.0342e-3	0.0025
MI	0.0221	0.0077	0.0026	0.0331	0.0001	0.0498

Tabla 5.11. Resumen de resultados para el registro multimodo de una sección cerebral.

5.3.5.3 Caso 3. Imagen sagital CT-PET

Tal y como se ha visto en los dos casos anteriores, los resultados producidos por los tres métodos de similitud poseen ciertas diferencias apreciables en una u otra medida, aunque siempre se encuentra la deformación como conjunto. En este tercer caso realizamos el registro de un par de imágenes sagitales CT-PET, haciendo uso exclusivamente del CR por ofrecer el mejor compromiso entre resultado y tiempo de computación. El resto de métodos de similitud son igualmente aplicables a éste y otros casos, tal y como se ha visto en los dos ejemplos anteriores.

En la Figura 5.33 se representan las imágenes utilizadas para el registro, donde se puede observar un ligero desplazamiento entre las posiciones de cada una de las imágenes.

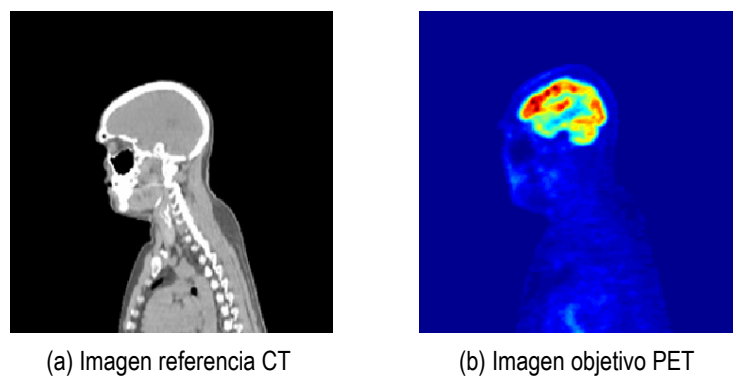


Figura 5.33. Imágenes multimodo de una sección humana sagital utilizadas para el registro del caso 3.

La Figura 5.34 representa el resultado del registro para la medida de similitud utilizada, el ratio de correlación, utilizando un factor $k=3$ y un umbral de 0.25. En la Figura 5.34(a) se muestra el resultado del registro final, mientras que en la Figura 5.34(b) se observan los vectores de desplazamiento detectados entre las imágenes. Por último, en la Figura 5.34(c) se ve el resultado de aplicar estos vectores a una rejilla rectangular.

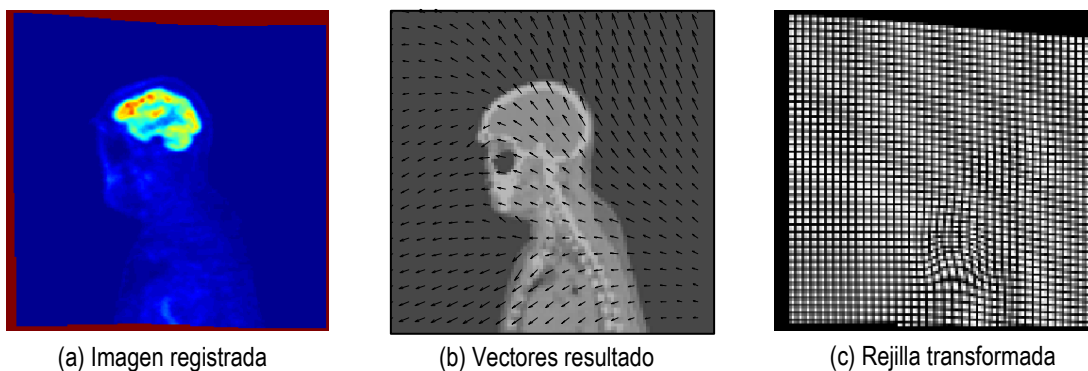


Figura 5.34. Resultados obtenidos para el CR.

En la siguiente figura (Figura 5.35) se puede observar la corrección de grises efectuada para el cálculo de la diferencia cuadrática. Aunque lejos de obtenerse una corrección perfecta, por el simple hecho de que la aproximación del histograma por una curva continua (de quinto orden en este caso) es muy complicada para imágenes multimodo, los bordes de la imagen quedan con un nivel de gris muy similar, ideal para el cálculo del movimiento de precisamente estas transiciones.

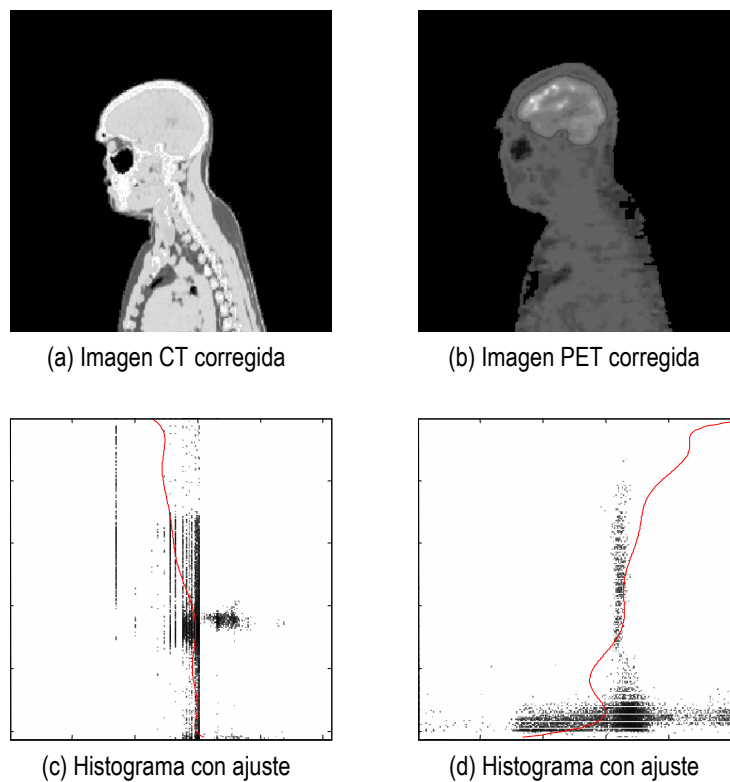


Figura 5.35. Corrección de grises efectuada en el caso 3.

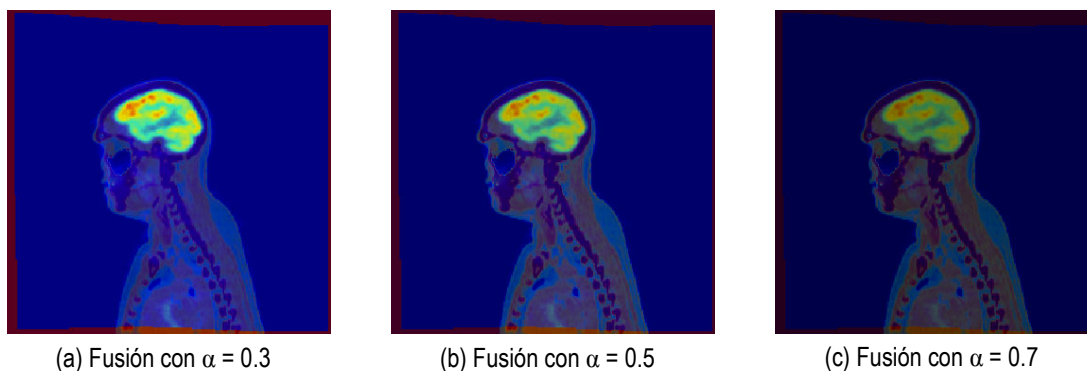


Figura 5.36. Fusión para distintos valores de α .

En la figura anterior (Figura 5.36) es posible ver el resultado de la fusión entre la imagen de referencia CT con la PET registrada, para diferentes valores del factor de transparencia α . Cuanto menor es α , mayor es el peso otorgado a la información funcional, mientras que si éste es mayor, la información estructural adquiere más peso. Esto se puede deducir directamente a partir de la expresión (5.4).

Por último, un resumen de los aspectos característicos de los vectores de movimiento obtenidos para el CR se muestra en la Tabla 5.12.

Medida	RIGIDEZ			ELASTICIDAD		
	Media	Mediana	Varianza	Media	Mediana	Varianza
CR	0.0245	0.0095	0.0046	0.0505	0.0002	0.1374

Tabla 5.12. Resumen de resultados para el registro multimodo de una sección sagital.

5.3.5.4 Caso 4. Tórax CT-PET

En este último caso se procede a realizar el registro de una nueva sección de un tórax humano correspondiente a una fractura, de nuevo exclusivamente para la medida de similitud CR, aunque como ya se ha comentado, tanto MI como WC son perfectamente aplicables.

El par de imágenes utilizado se muestra en la figura siguiente (Figura 5.37), donde se representa la imagen referencia CT y la de objetivo PET, con un cierto desplazamiento entre ambas adquisiciones.

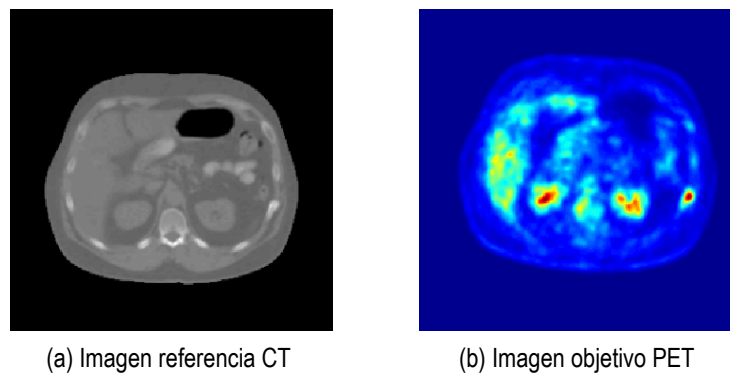


Figura 5.37. Imágenes multimodo de un tórax humano utilizadas para el registro del caso 4.

Para el cálculo del registro se hace uso de un parámetro de ponderado $k=3.5$ y de un umbral igual a 0.45, lo que da como resultado los efectos mostrados en la Figura 5.38. En esta figura se puede observar la imagen registrada (Figura 5.38(a)), los vectores de desplazamiento

calculados (Figura 5.38(b)) y el resultado de aplicar estos vectores a una rejilla rectangular (Figura 5.38(c)).

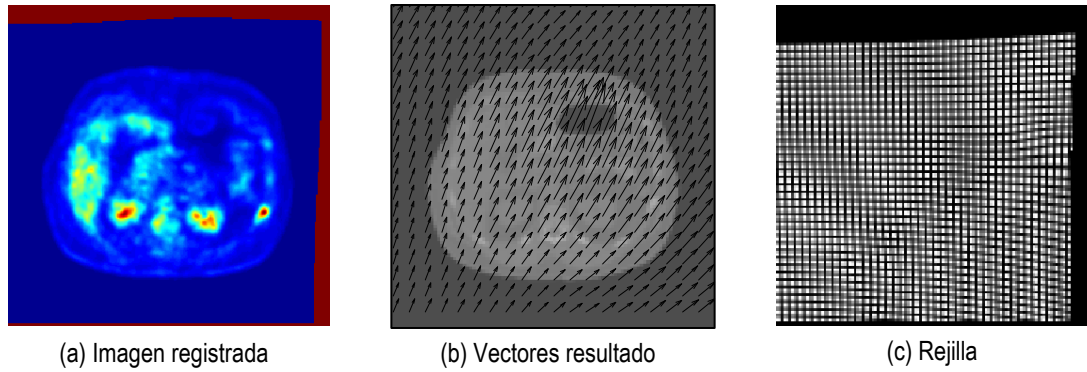


Figura 5.38. Resultados obtenidos para el CR.

La Figura 5.39 muestra las correcciones de nivel de gris llevadas a cabo para el cálculo de la diferencia cuadrática media. Como se ve, de nuevo se consigue una aproximación importante entre las imágenes, que nos permite detectar principalmente el movimiento de los bordes.

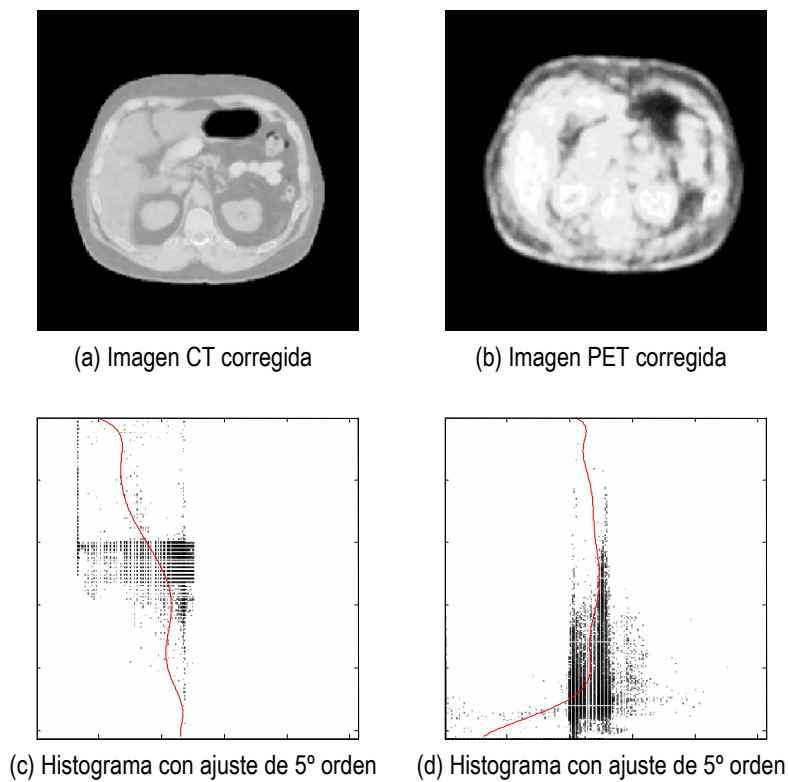


Figura 5.39. Corrección de grises efectuada en el caso 4.

Una comparativa de la fusión entre la imagen de referencia y la registrada para distintos valores α de transparencia se ilustra en la Figura 5.40.

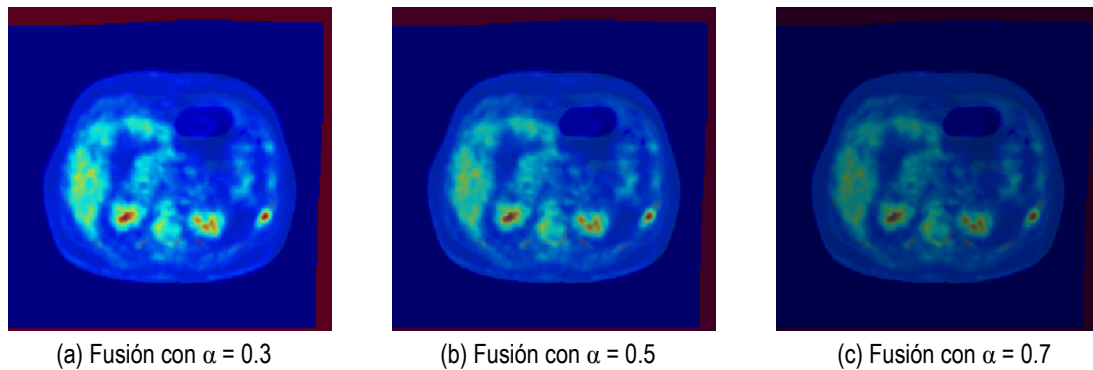


Figura 5.40. Fusión para distintos valores de α .

En la Tabla 5.13 se recogen los datos concernientes a la rigidez y elasticidad de los vectores de movimiento hallados entre las dos imágenes. Como se ha visto en la Figura 5.38(b), este movimiento es muy acentuado en una misma dirección, lo que se traduce en unos valores relativamente altos de media y varianza.

Medida	RIGIDEZ			ELASTICIDAD		
	Media	Mediana	Varianza	Media	Mediana	Varianza
CR	0.0639	0.0134	0.0943	0.1353	0.0004	3.1201

Tabla 5.13. Resumen de resultados para el registro multimodo de un tórax (caso 4).

5.4. DISCUSIÓN DE LOS RESULTADOS

Las pruebas llevadas a cabo sobre imágenes de distintos tipos nos confirman que no existe una medida de similitud especialmente idónea para todo tipo de aplicaciones. Incluso en imágenes del mismo tipo, uno u otro método puede producir un mejor o peor resultado dependiendo de las características intrínsecas de cada imagen. El criterio de Woods se muestra quizás como la medida de similitud más endeble y que produce en general peores resultados para registro monomodo, pero constituye una afirmación lógica si se tienen en cuenta las hipótesis sobre las que está basado este procedimiento.

Los píxeles a los que se les detecta la correspondencia entre las imágenes, se desplazan siempre hacia el punto donde la similitud es mayor. El desplazamiento de estos puntos puede producir zonas vacías sobre las que no se desplaza ningún píxel, provocando la aparición de bordes

negros. Estos bordes, si el fondo de la imagen no es de este color, generan valores del PSNR entre imágenes menores, además de obtener una imagen registrada un tanto antiestética. En imágenes médicas, donde este fondo es en muchos casos negro, el resultado ofrece una realidad cercana a la imagen de referencia, sea cual sea la deformación detectada.

No obstante, a la hora de llevar a cabo el registro entre un par de imágenes, además del tipo de los datos también habrá que tener en cuenta el tiempo de procesamiento que requiere cada método. De esta manera se establece un compromiso fundamental que es necesario afrontar, optando por la medida más adecuada si las restricciones de cómputo son menores; o por la medida más rápida que ofrezca un resultado suficientemente bueno para aplicaciones en tiempo real.

Por su parte, para registro multimodo son muchos los factores a tener en cuenta para la obtención de un resultado satisfactorio. Por una parte, es posible generar un par de imágenes pseudo-monomodo a partir de un par multimodo mediante la técnica de la corrección de niveles de gris. Aunque esta técnica reduciría el problema a un caso típico de registro monomodo, es necesaria una buena corrección de la intensidad para que esto sea posible. Es por ello que a veces se hace más interesante emplear exclusivamente medidas de similitud capaces de detectar correspondencia entre imágenes de distinta naturaleza, como el CR, el WC o la MI. En estos casos, es viable la consecución del registro mediante un ajuste preciso del coeficiente de umbralización, en el cálculo de los mapas de similitud. Para evitar esta laboriosa tarea, se hace uso de una corrección de niveles de gris para llevar a cabo solamente el cálculo de la diferencia cuadrática entre imágenes. El grado de calidad del registro vendrá dado esencialmente por la corrección efectuada en ambas imágenes, y que para imágenes de distinta naturaleza el ajuste del histograma por una curva resulta a menudo muy complicado.

En última instancia, la fusión entre imágenes de distinta modalidad nos permite la superposición de dos imágenes para la consecución de un todo con más información de la que tendría cada una de las imágenes por separado. Para adecuar la importancia de las imágenes de uno u otro tipo, se hace uso de un coeficiente de transparencia que pondera las imágenes en la fusión final.

6. CONCLUSIONES Y LÍNEAS FUTURAS

6.1. CONCLUSIONES

6.1.1. Esencia del proyecto

Analizar distintas medidas de similitud de tipo estadístico para el laborioso proceso del registro de imágenes, comprobando sus ventajas y aplicaciones ante diferentes tipos de imágenes y medios de adquisición.

6.1.2. Conocimientos adquiridos

Familiarización con el campo del registro de imágenes, principalmente en cuanto al procedimiento llevado a cabo, tanto para el registro de imágenes procedentes de distintas fuentes como de distinta naturaleza.

Establecimiento de las bases sobre las distintas medidas de similitud que hay en la actualidad para utilizarlas al registro, así como las posibles aplicaciones que nos permite el uso del registro de imágenes.

Profundización en la herramienta de trabajo Matlab, especialmente en los útiles de procesado de la imagen, representación de la información y de cálculo de estadísticos de la imagen.

6.1.3. Dificultades encontradas

Aunque Matlab nos proporciona una herramienta de gran potencia, es a veces muy complicado obtener el máximo rendimiento para el procedimiento de registro entre dos imágenes. Son muchas las operaciones a realizar, por lo que a menudo el alcance del óptimo se antoja ciertamente difícil.

Los detalles que han complicado en cierto modo la realización de algunas etapas del proyecto se resumen en los siguientes puntos:

- Tiempo de procesamiento elevado:

Resulta totalmente lógico pensar que a partir de los resultados vistos en el capítulo 3, el proceso de registro está asociado a una gran carga computacional. Aunque la NCC nos permite obtener resultados en un tiempo relativamente pequeño, éste es un problema del que adolecen el resto de casos. En estas circunstancias hay que optar por reducir la resolución de las imágenes para conseguir una mejora sustancial en el procesamiento, de modo que nos permita realizar más operaciones en el mismo periodo de tiempo.

Para conseguir una mejora de estas prestaciones puede bastar con reducir el tamaño de la imagen. Se ha demostrado que con tamaños de 256 x 256 píxeles se obtienen imágenes de resolución suficiente, pero a veces el procesamiento de estos bloques tan grandes no es lo suficientemente rápido, para lo que se emplean resoluciones menores que nos permitan obtener unos vectores de movimiento aplicables a las imágenes del tamaño deseado, mediante un remuestreo por ejemplo. A esta técnica se la conoce como resolución *subpíxel* y nos permite resolver el problema anterior aun a costa de un peor resultado.

Otras maneras de aumentar la velocidad de cálculo es mediante la elección del tamaño de los bloques a comparar, tal y como muestra la Tabla 3.1. Este parámetro está estrechamente relacionado con el tiempo de procesamiento, aunque habrá que fijar también una zona de búsqueda adecuada para que dicho cálculo obtenga una mejora real.

- Peculiaridades de las medidas de similitud:

El proceso de implementación de las medidas de similitud ha sido realmente laborioso por la minuciosidad con la que ha habido que tratar estas medidas para la obtención de un resultado satisfactorio. Para la comprobación de dichas medidas, se ha intentado abordar todos los casos posibles que nos permitieran un comportamiento adecuado ante cualquier adversidad. A veces, el obtener una medida idónea para un caso particular implica obtener una no tan buena para otro caso diferente, lo que a fin de cuentas hace que se trate de un compromiso que nos proporcione la mejor solución global.

Los problemas debidos a medidas fuera de rango han estado también presentes durante la implementación de las medidas de similitud estudiadas. Estos casos aislados se han debido básicamente a problemas en la discretización de las imágenes, valores nulos de las entropías para la MI o a problemas relacionados con las medidas estadísticas de valor nulo para el CR y el WC. Se ha optado siempre por asignar los valores de la similitud resultante a cero como medida de precaución ante estos valores que resultaban indeterminados, implicando esta medida la adjudicación de un parecido nulo entre las imágenes.

- Correspondencia entre bloques muy alejados:

A veces el resultado del registro para una imagen determinada no resulta tal y como se espera debido la homogeneidad (varias zonas parecidas dentro de la imagen) de las muestras. Esto conlleva consigo que una zona característica de la imagen puede tener varias

correspondencias de alto nivel de similitud en la otra imagen, lo que tiene como consecuencia unos vectores de movimiento detectados muy dispersos y poco suaves.

Para resolver el problema anterior se ha optado por reducir la zona de vecindad hasta un margen que determinara un resultado satisfactorio en cada caso. El hecho de reducir este rango tiene como consecuencia la disminución de la capacidad de detección ante grandes deformaciones, pero resulta la mejor opción en cuanto a tiempo de procesamiento y a resultado final. Sin embargo, hay que tener presente que el óptimo de este rango de búsqueda depende exclusivamente de las imágenes tratadas, siendo claramente diferentes entre sí para dos imágenes con deformaciones visualmente diferentes.

- Método de evaluación del emparejamiento de la imagen registrada:

En un primer lugar, para el cálculo del grado de afinidad entre la imagen de referencia y la obtenida mediante el proceso de registro, se hizo uso de la información mutua entre ambas imágenes. Esta medida, sin embargo, no resultaba muy fiable para imágenes de naturaleza multimodo. La razón se deriva del hecho de que no detecta el sentido de los bordes, es decir, que si se compara una transición de 0 a 1 con una de 1 a 0 por ejemplo, la MI dirá que ambos bloques son idénticos, lo cual es obviamente un error. Por esa razón esta medida no proporciona un resultado totalmente adecuado para la comparación de imágenes monomodo.

Ante las restricciones que se observan con la medida anterior, se ha optado por el uso del PSNR para el cálculo de la similitud en imágenes monomodo. Esta medida se basa en la diferencia entre los niveles de gris de la imagen referencia y la registrada, por lo que para estos casos produce un mejor resultado, acorde además con el error cuadrático medio de los vectores de transformación. Además, el hecho de que el resultado de esta medida sea en decibelios nos reporta unos datos de carácter más práctico.

- Registro de imágenes multimodo:

En un principio se intentó adoptar para la realización del registro entre imágenes de distinta naturaleza el uso exclusivo de las medidas de similitud estudiadas, sin hacer uso de la diferencia cuadrática, por otra parte inútil al no existir una relación entre los niveles de gris. Esta medida requería de un ajuste muy grande, lo que llevó a optar por requerir de esta medida mediante una conversión previa de los niveles de gris entre las imágenes.

La conversión entre niveles de gris también ha necesitado de un cierto estudio previo de las imágenes. Este hecho se deriva de que para el ajuste del histograma hay que fijar parámetros clave como la umbralización o el coeficiente del polinomio con el que se aproxima el modelo. Un histograma cuyos puntos tienen un comportamiento más abrupto requiere un ajuste de orden mayor, mientras que si los puntos son más suaves interesa un orden menor. Además, si este ajuste es de un orden muy alto, se puede originar un *sobreajuste*, lo que provoca a veces que el resultado sea peor que en un ajuste más general (de menor orden).

6.1.4. Valoración

Tomando como punto de partida las funciones implementadas relativas a las medidas de similitud abordadas, el trabajo se ha centrado en la consecución de una serie de resultados y conclusiones para estas medidas. Con esto se pretende evaluar las medidas de tipo estadístico para su aplicación en el registro multimodo.

La búsqueda de unas medidas de similitud que funcionen correctamente para obtener el registro entre un par de imágenes se traduce en largas horas de trabajo centradas en la delicada labor de recopilación, selección y estudio detallado de documentación, que nos permitan una implementación lo más eficaz posible.

No se obvia tampoco la gran vistosidad de este trabajo, que cuenta con la presencia de gran cantidad de imágenes que lo hacen más llevadero para su realización, pero que también pretende serlo para su lectura. El hecho de trabajar con imágenes nos proporciona una representación mucho más visual de los resultados alcanzados, lo que permite que con sólo un vistazo seamos capaces de determinar si el registro entre dos imágenes ha sido satisfactorio o no.

Habría que resaltar también que el trabajo llevado a cabo supone solamente un grano de arena en el desierto, pues son muchas las personas que han analizado éstas y muchas más medidas de similitud para su aplicación al registro de imagen. A pesar de que no se abarcan más que cuatro medidas, este trabajo pretende ser una síntesis de las características principales para estos métodos de similitud.

6.2. FUTURO

6.2.1. Posibles mejoras al procedimiento de registro

6.2.1.1 Mejora de las medidas de similitud

Como se ha observado, cada medida de similitud tiene una serie de características que la hace única. Una combinación de dichas medidas nos podría proporcionar un método para determinar con mayor exactitud la correspondencia entre dos bloques de imágenes. Para ello habría que hacer uso de alguna expresión que relacionara dichas medidas, o mediante un selector que dependiendo de las características propias del bloque elegido fuera capaz de seleccionar correctamente la medida de similitud más apropiada.

6.2.1.2 Reducción del tiempo de procesamiento

Las herramientas de procesamiento que nos proporciona Matlab son especialmente potentes, pero tienen la desventaja de ser a veces demasiado lentas para aplicaciones de tiempo real. Sería ideal para la mejora sustancial de la necesidad de cálculo que dichas funciones se optimizaran más aún si cabe, para garantizar un resultado más rápido sin necesidad de reducir la calidad de las imágenes.

La evolución de los ordenadores personales y de los procesadores para aplicaciones industriales van a ayudar especialmente a solucionar el problema del tiempo de procesamiento, pero quizás en un futuro las necesidades de procesamiento sean también mayores por la existencia de imágenes de distinto tipo. Tenemos sin embargo que estar tranquilos, pues la propia evolución tecnológica nos permitirá solventar el problema de la necesidad de cálculo.

6.2.1.3 Portabilidad

Esta es una cualidad muy interesante para toda aplicación tipo software, como la utilizada en este proyecto. La portabilidad nos permitiría realizar el procedimiento de registro en todo tipo de unidad de procesamiento, permitiendo también al usuario corriente el contar con esta clase de herramientas para su uso personal.

6.2.2. El futuro: hacia dónde se dirige el registro de imágenes

El registro de imágenes es un proceso complejo, tanto por la cantidad de etapas que son necesarias como por el hecho de que un cambio en la implementación de una de estas etapas puede variar completamente el resultado obtenido en la imagen registrada.

Hoy en día no conocemos el futuro en el campo del registro de imágenes, pero no es muy atrevido plantearse una evolución del registro de imágenes 2D en el campo de las imágenes médicas, pasando a un enfoque casi exclusivo sobre las imágenes 3D, que precisarán de una mayor capacidad de procesamiento y un cambio sustancial en el planteamiento de todo el proceso visto en este trabajo, con el fin no sólo de obtener un resultado satisfactorio, sino que este resultado se obtenga en el menor tiempo posible.

Teniendo en cuenta todo esto, no es difícil aventurar que vamos a poder aprovechar todos los conocimientos y avances que se han realizado hasta la fecha, para la mejora y la adaptación frente a los nuevos retos que se nos van a presentar en el campo del registro de imágenes en el futuro.

7. BIBLIOGRAFÍA

□ Referencias

- [1] Juan Morales Sánchez. Estimación de Movimiento no Rígido mediante Modelos Deformables. Tesis Doctoral, 2005.
- [2] Ian T. Young, Jan J. Gerbrands, Lucas J. van Vliet. Fundamentals of Image Processing. Delft University of Technology.
- [3] Francisco Javier Rodrigo Saura, Juan Morales Sánchez, José Luis Sancho Gómez. Medida de Similitud Combinada para el Registro y Fusión de Imágenes. Universidad Politécnica de Cartagena.
- [4] Graeme P. Penney, Jürgen Weese, John A. Little, Paul Desmedt, Derek L. G. Hill, David J. Hawkes. A Comparison of Similarity Measures for Use in 2D-3D Medical Image Registration. IEEE Transactions on Medical Imaging, vol. 17, nº 4, Agosto 1998.
- [5] Darko Skerl, Bostjan Likar, Franjo Pernus. A Protocol for Evaluation of Similarity Measures for Rigid Registration. IEEE Transactions on Medical Imaging, vol. 25, nº 6, Junio 2006.
- [6] Federico J. Bonsignore Caro. Aplicación de las Técnicas de la Teoría de la Información en el Registro de Imágenes Médicas. XIII Seminario de Ingeniería Biomédica, 2004. Facultades de Medicina e Ingeniería, Univ. de la República Oriental del Uruguay.
- [7] Frederik Maes, André Collignon, Dirk Vandermeulen, Guy Marchal, Paul Suetens. Multimodality Image Registration by Maximization of Mutual Information. KU Leuven, Laboratory for Medical Imaging Research. Proceedings of MMBIA '96.
- [8] Juan Morales, Luis Weruaga, Rafael Verdú. Non-Rigid Image Registration with Parametric Landmarks and Tikhonov Regularization. IEEE Transactions on Image Processing.
- [9] Peter Rogelj, Stanislav Kovacic. Similarity Measures for Non-Rigid Registration. University of Ljubljana.
- [10] G. S. Cox. Template Matching and Measures in Image Processing. University of Cape Town. 1995.

- [11] Zhiyong Gao, Jiarui Lin, Bangquan Xu. The Affection of Grey Levels on Mutual Information Based Medical Image Registration. Huazhong University of Science and Technology, P.R. China.
 - [12] Alexis Roche, Grégoire Malandain, Xavier Pennec, Nicholas Ayache. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. INRIA Sophia Antipolis, EPIDAURE project, France. Proceedings MICCAI '98, vol. 1496 of LNCS.
 - [13] Pascal Cachier, Nicholas Ayache. Regularization in Image Non-Rigid Registration: I. Trade-off between Smoothness and Intensity Similarity. INRIA. Informe de investigación nº 4188, Mayo 2001.
 - [14] Pascal Cachier. How to Trade-off between Regularization and Image Similarity Regularization in Non-Rigid Registration. INRIA Sophia Antipolis, EPIDAURE project, France. Octubre 2001.
 - [15] Pascal Cachier, Eric Bardinet, Didier Dormont, Xavier Pennec, Nicholas Ayache. Iconic Feature Based Non-Rigid Registration: the PASHA Algorithm. Computer Vision and Image Understanding 89 (2003) 272–298.
 - [16] Lilla Zöllei. A Unified Information Theoretic Framework for Pair- and Group-wise Registration of Medical Images. Tesis doctoral. Department of Electrical Engineering and Computer Science. Massachusetts Institute of Technology. Enero 2006.
 - [17] Tomasz Lukasiak. MRI-fMRI Image Registration Using the Woods Criterion.
 - [18] Wikipedia. La enciclopedia libre: <http://www.wikipedia.org/>.
- Imágenes utilizadas
- [19] Ejemplos incluidos en el software Matlab.
 - [20] The Whole Brain Atlas: <http://www.med.harvard.edu/AANLIB/home.html>.
 - [21] Normal and Benign Pathologic Findings in 18FDG-PET and PET/CT: <http://www.med.harvard.edu/JPNM/chetan/popup.html>.
 - [22] Thorax Radiology Images: http://www.upstate.edu/cdb/grossanat/Thorax_Radiology.shtml.
 - [23] The USC-SIPI Image Database: <http://sipi.usc.edu/database/index.html>.
 - [24] Computer Vision Test Images: <http://www.cs.cmu.edu/~cil/v-images.html>.