

Algunas contribuciones sobre la corrección de
inhomogeneidades de iluminación en
imagenes digitales.

Elsa Fernandez

Resumen

El problema de las distorsiones introducidas por la fuente de iluminación está presente en todas las aplicaciones que utilizan imágenes digitales para el reconocimiento y detección en su más amplia acepción. En este trabajo nos hemos concentrado en las imágenes de Resonancia Magnética Nuclear, dado su creciente uso diagnóstico. Además, en ellas este problema se puede formular de una forma más asequible para su solución que en el caso general, debido a que las imágenes presentan una clara estructura como funciones constantes por trozos. Hemos implementado y contrastado diversas aproximaciones, que incluyen nuestras proposiciones basadas en la minimización de una función objetivo definida sobre el error de clasificación y el bias de iluminación. Además de las imágenes de resonancia magnética, hemos adaptado nuestro algoritmo de corrección de la iluminación para aplicaciones en el reconocimiento de caras, con resultados positivos.

ÍNDICE GENERAL

1	Introducción	1
1.1	Motivación y descripción general de los objetivos de la tesis . . .	2
1.2	Estructura de la memoria	9
1.3	Publicaciones realizadas durante el desarrollo de la tesis	10
2	El problema de la corrección de iluminación	13
2.1	Planteamiento general	14
2.1.1	Características específicas del problema para IRM	15
2.1.2	Ilustración de los efectos de la iluminación	15
2.2	El filtrado homomórfico	22
2.3	Clasificación versus restauración	25
3	Algoritmos de corrección del bias en IRM	29
3.1	Algoritmos basados en métodos bayesianos.	30
3.2	Agrupamiento borroso	32
3.3	Métodos no paramétricos	33
3.3.1	BMAP	35
3.3.2	BFCM	43
3.4	Métodos paramétricos	47
3.4.1	Algoritmo de Wells	48
3.4.2	PABIC	53
3.5	Nuestras proposiciones	60
3.5.1	El gradiente	60
3.5.2	GradPABIC	63
3.5.3	GradClassLeg	63
4	Resultados experimentales en IRM	65
4.1	Consideraciones metodológicas	65
4.2	Experimentos realizados sobre imágenes de tablero de ajedrez	67

4.2.1	Sobre el filtrado homomorfo	69
4.2.2	Comparación entre algoritmos evolutivos	69
4.2.3	Comparación con algoritmos no evolutivos	76
4.2.4	Análisis de sensibilidad BMAP	80
4.2.5	Análisis de sensibilidad del algoritmo de Wells	82
4.2.6	Conclusiones sobre las imágenes de tablero de ajedrez	83
4.3	Experimentos sobre las imágenes de cerebro sintéticas.	84
4.3.1	Resultados de los algoritmos evolutivos	85
4.3.2	Resultados con el algoritmo de Wells, BMAP y BFCM	90
4.3.3	Conclusiones sobre las imágenes de cerebro	102
4.4	Resultados sobre imágenes del IBSR	102
4.4.1	Resultados básicos	102
4.4.2	Estudio de sensibilidad de GradClassLeg	105
4.4.3	Estudio de sensibilidad del algoritmo de Wells	108
4.4.4	Estudio de sensibilidad de BMAP	112
4.5	Discusión y conclusiones	113
5	La corrección de la iluminación en reconocimiento de caras	119
5.1	Antecedentes	119
5.2	Planteamiento del problema	121
5.3	Resultados experimentales en reconocimiento de caras	123
5.4	Discusión	127
6	Conclusiones y trabajo futuro	131
7	Apéndice 1: Fundamentos de la imagen de resonancia magnética nuclear	135
7.1	Nociones físicas básicas	135
7.2	Espectroscopía de RMN	138
7.3	Componentes hardware	140
7.4	Principios de visualización	143
7.5	Visualización mediante la TF	146
7.6	Presentación de la imagen	148
7.7	Artefactos en la imagen	148
7.7.1	El ruido en las imágenes de RMN	150
8	Apéndice 2: Polinomios de Legendre	153
8.1	Ortogonalidad y productos de polinomios	154

9	Apéndice 3: Estrategias Evolutivas y Algoritmos Meméticos	157
9.1	Estrategias evolutivas	157
9.1.1	Estrategias evolutivas para la corrección de la iluminación	158
9.2	Algoritmos Meméticos	159
9.2.1	Algoritmos meméticos instantáneos	159

ÍNDICE DE FIGURAS

2.1	Histograma de una imagen de tablero de ajedrez con dos niveles de gris, sin corrupciones de ningun tipo.	16
2.2	Imagen de tablero de ajedrez original	16
2.3	Histograma de una una imagen de tablero de ajedrez corrompida con ruido gaussiano aditivo.	17
2.4	Imagen del tablero de ajedrez corrompida por ruido gaussiano aditivo.	17
2.5	Ejemplo de campo de iluminación no uniforme.	18
2.6	Ejemplo de imagen corrompida con un campo de iluminación.	18
2.7	Histograma de la imagen de tablero de ajedrez corrompida por un campo de iluminación no uniforme.	19
2.8	Resultados de clasificación versus SNR en el caso (a) de ruido gaussiano y (b) de campo de iluminación inhomogeneo.	20
2.9	Clasificación obtenida aplicando el algoritmo k-medias a una imagen corrompida con un campo de iluminación inhomogeneo.	21
2.10	Imagen de la clasificación obtenida por el algoritmo de las k-medias sobre una imagen corrompida por ruido gaussiano aditivo.	21
2.11	Resultados del cálculo de los bordes en el caso (a) de la imagen original en (c) y (d) de la corrompida con un campo de iluminación inhomogeneo	23
2.12	Resultados del filtrado homomórfico sobre una imagen de tablero de ajedrez corrupta (a) reflectancia, (b) campo de iluminación estimado.	26
2.13	Resultados de la aplicacion del filtro homomórfico sobre una rodaja de un volumen de cerebro (a) imagen corrupta, (b) reflectancia y (c) campo de iluminación.	26
3.1	Ilustracion del efecto del sesgo debido al campo de iluminacion y del ruido aditivo en una linea de una imagen, o en una funcion unidimensional.	54

4.1	Imagen de tablero de ajedrez original.	68
4.2	Ejemplo de campo de iluminación empleado para corromper la imagen de tablero de ajedrez.	68
4.3	Ejemplo de imagen de ajedrez corrompida por un campo de iluminación.	68
4.4	Correlación con las originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos: filtrado homomórfico, GradPABIC, EE y PABIC	70
4.5	Varianza de las correlaciones con los originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos GradPABIC, EE y PABIC	70
4.6	Evolución de la función de ajuste en sendas instancias del algoritmo PABIC (a) y la EE (b).	73
4.7	Resultados de la recuperación de la imagen con algoritmos evolutivos (a) Imagen corrupta, (b) filtrado homomórfico, (c) PABIC, (d) EE, (e) GradPABIC, (f) BMAP	74
4.8	Resultados visuales de la estimación del sesgo del campo de iluminación (a) sesgo original, (b) filtrado homomórfico, (c) PABIC, (d) EE, (e) GradPABIC, (f) BMAP.	75
4.9	Correlaciones con los originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos GradClassLeg, EE y PABIC	76
4.10	Varianzas de las correlaciones con los originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos GradClassLeg, EE y PABIC	77
4.11	Correlaciones de la imagen corregida, (a) para los algoritmos GradPABIC, Wells, (b) GradClassLeg y Wells	78
4.12	Aciertos en la clasificación de las imágenes corregidas comparados con la clasificación nominal.	79
4.13	Correlaciones de las imágenes corregidas y de la imagen original, (a) algoritmos GradPABIC, BFCM, BMAP, (b) algoritmos GradClassLeg, BFCM, BMAP	79
4.14	Sensibilidad del BMAP respecto del tamaño del filtro de suavización sobre una imagen de tablero de ajedrez, con radio 2 de vecindario en el MRF que modela las probabilidades a priori.	80
4.15	Sensibilidad del BMAP respecto del tamaño del vecindario en el MRF de las probabilidades a priori.	81

4.16	Influencia del diametro de la mascara de suavización sobre el resultado del algoritmo de Wells.	83
4.17	Imágenes de cortes del volumen sintético de cerebro usados como originales (a) corte axial, (b) corte coronal, (c) corte sagital . . .	86
4.18	Imágenes de cortes del volumen sintético de cerebro corrompidas con un sesgo de iluminacion (a) corte axial, (b) corte coronal, (c) corte sagital	86
4.19	Clasificación de las imágenes de cortes del cerebro en líquido cerebro-espinal, materia gris y materia blanca. Cortes (a) axial, (b) coronal y (c) sagital	87
4.20	Correlaciones de las imágenes corregidas (a,c,e) y de los sesgos estimados (b,d,f) para los algoritmos evolutivos, sobre las imagenes de los cortes axial, coronal y sagital	88
4.21	Correlaciones de las imágenes corregidas (a,c,e) y de los sesgos estimados (b,d,f) para los algoritmos evolutivos y GradClassLeg sobre las imágenes de los cortes axial, coronal y sagital	89
4.22	Resultados visuales de las estrategias evolutivas sobre los cortes del volumen sintético.	91
4.23	Resultados visuales del algoritmo PABIC sobre las imágenes de los cortes del volumen sintético. (a, b, c) imágenes corregidas, (d, e, f) sesgos estimados.	92
4.24	Resultados visuales del algoritmo GradPABIC sobre los cortes del volumen sintético	93
4.25	Resultados visuales de la aplicación del algoritmo GradClassLeg, estado inicial arbitrario, sin estimar las intensidades medias de las clases. (a,b,c) imágenes corregidas, (d,e,f) sesgos de iluminación estimados.	94
4.26	Resultados visuales de aplicar el algoritmo GradClassLeg sobre las imágenes sintéticas corruptas, estimando las medias de las intensidades de las clases. (a,b,c) imágenes corregidas, (d,e,f) sesgos estimados.	95
4.27	Correlación de las imagenes restauradas y la original sobre los cortes del volumen sintético de cerebro, GradPABIC versus el algoritmo de Wells, BFCM y BMAP	96
4.28	Correlación de las imágenes restauradas y la original sobre los cortes del volumne sintético de cerbro, GradClassLeg versus el algoritmo de Wells, BFCM y BMAP	98

4.29	Resultado de la aplicación del algoritmo de Wells. (a,b,c) Imágenes corregidas(d,e,f) Sesgos estimados.	99
4.30	Resultado de la aplicación del algoritmo BFCM a las distintas rodajas del volumen sintético de cerebro corrompido con un campo sintético. (a, b, c) imágenes corregidas, (d, e, f) sesgos estimados.	100
4.31	Resultados de la aplicación del algoritmo BMAP sobre las imágenes de la figura. (a, b, c) Imágenes corregidas. (d, e, f) Sesgos estimados.	101
4.32	Resultados de aciertos de la clasificación de los píxeles en materia gris, materia blanca, líquido cerebro-espinal obtenidos por GradClassLeg, Wells y BMAP sobre cada rodaja.	104
4.33	Una rodaja de (#32) del volumen empleado en los experimentos, (a) la imagen original con las materias no cerebrales removidas. (b) imagen de la segmentación manual donde cada píxel tiene el valor de intensidad de la clase asignada.	106
4.34	Bias detectados en la rodaja del volumen en la figura (a) GradClassLeg, (b) BMAP, (c) Wells	107
4.35	Imagen corregida de la rodaja #32 del volumen obtenido del IBSR. (a) GradClassLeg, (b) BMAP, (c) Wells.	107
4.36	Imágenes de la clasificación obtenidas por los algoritmos con los valores nominales de intensidad correspondiente a cada clase.	108
4.37	Sensibilidad del GradClassLeg a variaciones en la ganancia del gradiente, aciertos en las distintas rodajas del volumen.	109
4.38	Sensibilidad de GradClassLeg para distintos números de iteraciones.	110
4.39	Sensibilidad de GradClassLeg a las variaciones en el orden de los productos 2D de polinomios de Legendre.	111
4.40	Efecto del tamaño del filtro pasabaja en los resultados obtenidos con el algoritmo de Wells.	112
4.41	Efecto de la variación del parámetro de potencial en el algoritmo BMAP para la imagen del IBSR:	113
4.42	Estudio del efecto del cambio, para una imagen del IBSR, del radio del vecindario en BMAP.	114
4.43	Sensibilidad del BMAP al tamaño de la máscara de convolución de la estimación del campo de iluminación.	115
5.1	Imágenes originales de uno de los sujetos de la colección de Yale	125
5.2	Diferencias respecto de la cara de referencia para un individuo.	125

5.3	Cara de referencia utilizada en la corrección de la iluminación . .	126
5.4	Error de restauracion versus la velocidad de descenso de gradiente.	127
5.5	Error final de la corrección versus el orden maximo de los polinomios de Legendre.	128
5.6	Imágenes del sujeto tras la corrección de la iluminación.	129
5.7	Resultados de clasificación sobre las imágenes de caras del conjunto de Yale, antes y después de la corrección	130
7.1	Ecuaciones de Bloch	138
7.2	Diagrama temporal de la secuencia espin-echo.	140
7.3	Esquema de un sistema de visualización de IRMN	141
7.4	Esquema de la disposición de las bobinas de gradiente para un sistema de imagen de resonancia magnética de apertura horizontal	142
7.5	Algunas bobinas de RF típicas.	142
7.6	Ilustración de la obtención de la imagen mediante retroproyección.	144
7.7	Pulsos para la obtención de la imagen de una rodaja mediante retroproyección.	145
7.8	Secuencia de pulsos para la visualización mediante TF.	146
8.1	Los primeros 5 polinomios de Legendre en el intervalo (-1,1). . .	154
8.2	Productos de polinomios de Legendre de hasta orden 4.	156
9.1	Algoritmo general de una estrategia evolutiva	159
9.2	Algoritmo Memético general	160
9.3	Algoritmo Memético instantáneo	161
9.4	Algoritmo Memético instantáneo de individuo único	162

1. INTRODUCCIÓN

En este capítulo procederemos a la descripción de la motivación y objetivos generales de la memoria de la tesis. Los trabajos de desarrollo de la presente tesis han sido similares a un viaje de exploración sobre un tema de interés. Al comenzar este viaje no existían objetivos definidos salvo una inquietud sobre el papel del efecto de la iluminación en los procesos de reconocimiento sobre la imagen digital. En este viaje hemos estudiado dos casos de aplicación muy distintos: la segmentación de imágenes de Resonancia Magnética (IRM) nuclear y el reconocimiento de caras en imágenes frontales en niveles de gris. El interés sobre la imagen de resonancia magnética nuclear proviene de la colaboración con el Instituto de Resonancia Magnética de la Universidad Complutense de Madrid.

La IRM es diferente de la mayor parte de las otras técnicas de visualización médica porque no es el resultado de una transformación de tomografía¹. IRM permite visualizar con gran contraste los tejidos blandos y ha revolucionado la capacidad de diagnosticar las patologías que los afectan. La visualización de resonancia magnética está basada en el fenómeno conocido como resonancia magnética nuclear (RMN) y las IRM son mediciones agregadas de la composición de los tejidos al nivel molecular. El hecho de que la estructura molecular permanece constante dentro del tejido y varía entre distintos tejidos determina la efectividad de la IRM. La IRM tiene una alta resolución espacial y proporciona mucha información sobre la estructura anatómica, permitiendo estudios patológicos o clínicos cuantitativos, la derivación de atlas anatómicos digitalizados y también la guía antes y durante la intervención terapéutica. Las aplicaciones iniciales de la IRM requirieron principalmente que la imagen fuera lo suficientemente clara y libre de artefactos como para soportar las tareas de diagnóstico cualitativo de

¹Tomografía es el problema inverso a la proyección. Consiste en la obtención de un volumen 3D a partir de un conjunto de imágenes 2D obtenidas mediante proyecciones desde distintos ángulos. Esta es la técnica que permite obtener las imágenes volumétricas a partir de imágenes de rayos X.

patologías. Muchas de las tareas que se requieren de los radiólogos son repetitivas, lo que justifica la introducción de métodos automatizados de proceso de la imagen. Por ello existe una necesidad creciente de resultados cuantitativos basados en el análisis automatizado de la imagen. El proceso automatizado puede ir desde la eliminación de ruido impulsivo, mediante filtros frecuenciales y otros como los filtros anisotrópicos, hasta las operaciones de registro y segmentación. El registro de las imágenes consiste en el alineamiento de diferentes imágenes para obtener una visión más completa. Se aplica en la mezcla de imágenes de múltiples modalidades y la comparación de los pacientes con atlas anatómicos. La segmentación de la imagen consiste en su descomposición en regiones. Los criterios de identificación de regiones varían de aplicación en aplicación. La segmentación de la imagen es crítica en aplicaciones como la diagnosis de la esquizofrenia, la detección de tumores y la cirugía basada en realidad aumentada. Los métodos de segmentación deben ser fiables y reproducibles. Por ejemplo, en estudios de enfermedades degenerativas cerebrales como la esquizofrenia, el mal de Alzheimer [106], esclerosis múltiple, es necesario medir de forma precisa la cantidad de materia gris, materia blanca, lesiones en la materia blanca, líquido cerebro-espinal y sus distribuciones espaciales y cambios temporales.

A continuación detallaremos la motivación y objetivos de la tesis. Completan este capítulo introductorio una excursión por la estructura de la memoria de la tesis y la enumeración de las publicaciones producidas durante el desarrollo de los trabajos, donde no incluimos los artículos en preparación.

1.1. Motivación y descripción general de los objetivos de la tesis

Los trabajos de la presente tesis doctoral tienen como eje central el estudio de las soluciones a uno de los problemas centrales en el tratamiento de la imagen digital y en los procesos subsecuentes de reconocimiento automático y humano: la corrección de los efectos nocivos de una iluminación no controlada e inhomogénea.

Formación de la imagen

La imagen se genera cuando un mecanismo sensor recibe la energía que, proveniente de un sistema emisor de iluminación, le llega reflejada desde la escena que está siendo visualizada. De una forma u otra, el sistema de iluminación define

un campo de iluminación en el que las cosas son visibles para el sensor. Si ese campo es constante en su intensidad la imagen obtenida será un reflejo fiel de la distribución de cosas en la escena visualizada². Sin embargo, si el campo de iluminación no es constante, encontraremos que se producen, en la imagen que observamos variaciones de la intensidad de la imagen que no se corresponden con objetos en la escena.

En las cámaras digitales convencionales que registran la luz visible, el sistema de iluminación es el sol o alguna fuente artificial de luz. La fuente de iluminación emite fotones que llegan hasta la escena, son reflejados y llegan hasta la cámara. En situaciones de iluminación controlada suele ser preferible realizar un campo de iluminación homogéneo tanto en intensidad como en color, lo que facilita los procesos posteriores.

En los sistemas de generación de la imagen que se emplean en imagen médica existe una variedad de sistemas de iluminación, captación y transformación de esa información en una imagen que pueda ser interpretada por un ser humano o procesada automáticamente. En el caso que nos ocupa, la IRM tiene como sistema de iluminación un sistema emisor de radio-frecuencia que perturba el estado de reposo de los spines en su alineación con el campo magnético dominante establecido por el magneto, como se explica en el Apéndice I.

Las inhomogeneidades

La existencia de inhomogeneidades puede ser inherente al sistema, por ejemplo, cuando se utilizan bobinas de superficie en la obtención de IRM. Estas bobinas tienen una penetración en los tejidos que decrece con la distancia a la superficie y las imágenes resultantes ofrecen necesariamente una disminución de la intensidad en el sentido que se aleja de la superficie. En otras configuraciones de obtención de IRM, como los magnetos de apertura lateral, las inhomogeneidades son debidas o bien a un funcionamiento defectuoso de las bobinas de radio-frecuencia, o bien a la aparición de materiales ferromagnéticos en el objeto que está siendo visualizado.

En el caso del funcionamiento defectuoso de las bobinas de radio-frecuencia las variaciones del campo de iluminación son suaves y generales, se extienden a toda la imagen. Al igual que ocurre con las bobinas de superficie. Cuando

²Resulta difícil traducir las palabras inglesas derivadas de *image*. Por ejemplo, la palabra *imaged* que podría usarse en el contexto de este párrafo se refiere a algo de lo que se ha construido una imagen. La palabra más cercana a este significado en castellano sería visualizado.

aparecen materiales ferromagnéticos, el campo de iluminación presenta inhomogeneidades localizadas, donde la imagen se satura en un entorno de la posición del objeto. Aunque las imágenes RM pueden parecer visualmente uniformes, las inhomogeneidades perturbando la clasificación automática basada en las intensidades. Por ejemplo, en el caso ideal, la diferenciación entre la materia gris y la materia blanca en el cerebro se puede realizar fácilmente, debido a las diferencias en la señal que exhiben estos tejidos. Sin embargo, en la práctica las inhomogeneidades pueden ser de tal magnitud que causan el solapamiento de las distribuciones de la intensidad de señal asociadas a los distintos tejidos [120]. En [70] se muestra cómo la inhomogeneidad de la iluminación elonga los clusters en el espacio de características, llegando a fundir, por ejemplo, los clusters que representan el cerebro y el líquido cefalorraquídeo en el caso de IRM del cráneo. También se ha mostrado [113] cómo la estructura de las bobinas de radio-frecuencia y la propia estructura geométrica del objeto visualizado pueden inducir inhomogeneidades de iluminación que son frecuentes en las imágenes de cerebro.

En las imágenes de cámaras convencionales, la inhomogeneidad del campo de iluminación se debe a la iluminación artificial, a la presencia de ventanas o a posiciones extremas relativas del objeto visualizado y la fuente de iluminación (el sol). Si consideramos el caso de las imágenes de caras, la dirección de la que procede la iluminación puede crear sombras en la cara que modifican sustancialmente la función de intensidad de la imagen. La iluminación frontal, esto es, desde la posición de la cámara cuando el sujeto está mirando directamente a la cámara, es la ideal puesto que no genera sombras en los ojos, la nariz, mentón, etc. Las iluminaciones desde otras direcciones introducen sombras, además de variar suavemente la intensidad de los píxeles de la piel. La cara es una superficie suave, pero que varía de forma muy compleja. Esto significa que pequeñas variaciones de la dirección de la iluminación pueden producir fuertes variaciones de la iluminación reflejada por partes aparentemente suaves de la cara, debido a las variaciones del ángulo de la normal a la superficie y la dirección de la iluminación.

Además de las variaciones en intensidad, las imágenes de color obtenidas por cámaras convencionales presentan sensibilidades espectrales que deben ser calibradas según la aplicación. La fuente de iluminación tiene habitualmente su propio espectro, lo que hace que la imagen de color de una cara pueda ser muy variable, puesto que es el resultado del producto de los espectros de la fuente de iluminación y de la piel de la cara. Muchas de las dificultades encontradas en

sistemas de seguimiento basados en color están relacionadas a esta variabilidad espectral.

Efectos en el proceso de las imágenes

Puesto que los procesos de análisis y reconocimiento en las imágenes se basan en las intensidades observadas, es fácil deducir que el efecto de un campo de iluminación inhomogeneo puede ser muy perjudicial tanto para los sistemas automatizados como para el análisis por inspección humana.

Las aplicaciones avanzadas que usan el contenido morfológico de la IRM usualmente requieren segmentación del volumen en tipos de tejidos. Aunque existen algunos trabajos que aplican filtros y transformaciones lineales [114] para la detección de tejidos, la mayor parte de las técnicas que aparecen en la literatura son técnicas de clasificación estadística o aprendizaje automático [11]. En las IRM se aplican tanto la segmentación no supervisada como clasificación supervisada de los píxeles (voxels³) [11]. La segmentación no supervisada se puede realizar mediante la aplicación de algoritmos de agrupamiento (*clustering*) sobre las intensidades de los píxeles en las diferentes modalidades, de forma conjunta o independiente por modalidad, ya sea utilizando redes neuronales artificiales o métodos borrosos (*fuzzy clustering*). Estas técnicas no supervisadas tratan de detectar las regiones en la imagen en base al descubrimiento de agrupamientos en los espacios de características. Estas regiones se corresponderán con los tejidos biológicos. Las variaciones en el campo de iluminación pueden alterar de forma significativa la distribución de la intensidad en la imagen, haciendo que no sean distinguibles las agrupaciones que deberían serlo en la imagen “limpia”.

Si bien existen muchos trabajos sobre el efecto de las inhomogeneidades de iluminación en los algoritmos de segmentación basados en aprendizaje automático, no existe mucha información sobre su efecto en otros tipos de algoritmos de segmentación no supervisada, como los basados en la manipulación de los conjuntos de nivel (*level sets*) [118] mediante esquemas de evolución de ecuaciones diferenciales. La propia definición de los conjuntos de nivel está basada en la función intensidad y las deformaciones que la inhomogeneidad de iluminación introduce hacen que los conjuntos de nivel pierdan su significado, puesto que no se

³En IRMN es frecuente que se denominen voxels a las unidades de visualización en la imagen, puesto que las imágenes suelen corresponder a un volumen y los puntos en la imagen corresponden a una discretización de dicho volumen. A lo largo de la memoria utilizaremos preferentemente la denominación pixel.

garantiza su constancia en la imagen. También alteran el cálculo del gradiente, pudiendo hacer inefectivos los algoritmos basados en la detección de bordes. Por ejemplo, métodos como el propuesto en [109], que realizan una aproximación al contorno de un objeto de poco contraste (el hipocampo en la base del cerebro) mediante contornos activos guiados por interacción manual, son muy sensibles a las variaciones que el campo de iluminación puede introducir en el cálculo del gradiente de la imagen. En [4] se presenta un método de segmentación del cerebro en las rodajas de IRM, basado en contornos activos, que afirman poseer robustez a inhomogeneidades de la iluminación y que utiliza información a priori para segmentar algunos componentes como los ojos.

También se han aplicado técnicas supervisadas de clasificación basadas en redes neuronales, árboles de decisión y otras. Algunos autores han aplicado redes neuronales artificiales de entrenamiento supervisado, como en [123] para la diagnosis de cancer de mama. En otros casos se proponen técnicas mixtas que tienen una primera fase no supervisada de identificación de agrupamientos en los datos, y una segunda fase supervisada de construcción del clasificador sobre la representación inducida por las funciones de pertenencia a los agrupamientos. Por ejemplo, [54], [105] utilizan el mapa autoorganizativo (SOM: *Self Organizing Map*) y un perceptron multicapa para realizar la segmentación supervisada del cerebro en IRM. La clasificación supervisada de los píxeles precisa de un conjunto de muestras etiquetadas (conjunto de entrenamiento) de los píxeles de la imagen. Las variaciones en los campos de iluminación entre las imágenes de entrenamiento y la imagen sobre la que se prueba el algoritmo pueden ser catastróficas en términos de los resultados de la clasificación.

Finalmente, en algoritmos de segmentación como el propuesto en [87] que utiliza la intensidad directamente para generar, mediante líneas de proyección, modelos geométricos de las características anatómicas de interés, las variaciones de intensidad debidas a la inhomogeneidad del campo de iluminación pueden introducir fuertes deformaciones en el modelo geométrico que se persigue.

El reconocimiento de caras se considera en general un problema resuelto en circunstancias controladas de captura de la imagen, cuando la vista es frontal y la iluminación procede del punto de vista de la cámara. La técnica más difundida de construcción del clasificador consiste en dos fases: la primera es la reducción de dimensiones mediante el análisis en componentes principales (*Principal Component Analysis: PCA*), la segunda es la clasificación propiamente dicha. Es bien conocida la sensibilidad del PCA a grandes variaciones en las imágenes. En esencia, PCA define un hiperplano de proyección que ajusta

por mínimo error cuadrático los datos. Si los datos están muy dispersos, la aproximación típica consiste en definir varias transformaciones PCA ajustadas a regiones específicas del espacio de datos. Los campos de iluminación inhomógenos y las variaciones de pose producen fuertes cambios en la apariencia de la cara, los cuales introducen dificultades muy fuertes en el desarrollo de sistemas robustos en condiciones generales. También las variaciones en la dirección de la iluminación introducen sombras y deformaciones en la función intensidad. Esto es, la iluminación es una de las posibles causas de ineffectividad del esquema de reconocimiento antedicho. Pero no sólo los métodos basados en la apariencia global se ven afectados, también los métodos basados en características geométricas de la cara son sensibles a la iluminación debido a las dificultades que ésta introduce al extraer los puntos característicos de la cara.

En los sistemas de detección y seguimiento de caras, el color es una de las características más eficientes para la detección en tiempo real [64], debido a que se realiza como una decisión independiente para cada pixel, en base a características definidas en un espacio de baja dimensión. La aproximación habitual consiste en estudiar el agrupamiento de los pixeles en el espacio de color elegido. Las variaciones espectrales debidas a la cámara o a la fuente de iluminación pueden desplazar dramáticamente la posición en la que se encuentran estas agrupaciones de pixeles, invalidando todo el trabajo de estimación de parámetros de agrupación realizado.

Objetivos

El objetivo global de los trabajos de esta tesis, y de la presente memoria, podemos resumirlo en la siguiente frase: “Estudiar métodos de corrección de iluminación que tengan aplicación general en diversos campos en los que se utilizan imágenes digitales”. Estudiamos el problema en el contexto de dos aplicaciones: la segmentación de la IRM y el reconocimiento de caras en imágenes monocromas (de niveles de gris).

El problema de la inhomogeneidad de iluminación en IRM ha recibido mucha atención y se han propuesto una gran cantidad de algoritmos. Nuestro objetivo es reproducir los resultados encontrados en la literatura y proponer mejoras en términos de la restauración de la imagen sin efectos de iluminación y de la clasificación de los píxeles de la imagen para la segmentación en regiones de la imagen. En esta memoria de tesis no consideramos los métodos basados en el gradiente de la imagen.

El problema de inhomogeneidad de iluminación o de las variaciones introducidas en la imagen debidas a las variaciones de iluminación sobre el reconocimiento de caras, se plantea en esta tesis como un resultado adicional. Esto es, proponemos la aplicación de un modelado y estimación del campo de iluminación basado en un modelo paramétrico que ha sido probado con anterioridad sobre las imágenes de resonancia magnética. El objetivo en este caso es probar la utilidad de esta aproximación en un contexto computacional completamente distinto.

En la presente tesis doctoral no hemos considerado el problema de la detección de la cara, por lo que no consideramos los métodos de corrección espectral, que se plantean para la definición de algoritmos robustos de identificación de los píxeles de cara.

Algoritmos propuestos y su relación con la literatura

Consideraremos en primer lugar el caso de la segmentación de IRM mediante algoritmos de agrupamiento. En esta memoria nos hemos concentrado en aquellos algoritmos que realizan esta operación considerando explícitamente la posibilidad de un campo de iluminación inhomogeneo. No hemos considerado algoritmos basados en texturas, operadores morfológicos o modelos estadísticos espaciales. En esta clase de algoritmos de segmentación se distinguen, como ya se comentará en el capítulo 3, dos aproximaciones esencialmente distintas: los métodos paramétricos y los no paramétricos. Nuestro interés se centra en los métodos paramétricos, si bien hemos implementado algunos métodos no paramétricos para contrastar experimentalmente los resultados de los algoritmos. Creemos que el interés de los métodos paramétricos radica en que introducen restricciones globales sobre la corrección de la imagen a través del modelo del campo de iluminación. Estas restricciones globales sirven de término de regularización implícito en el proceso de minimización, análogamente a los términos de restricciones de suavidad que se introducen en otros métodos, como las distribuciones a priori en los métodos bayesianos de proceso de la imagen [49], [50]. Los métodos no paramétricos generalmente necesitan considerar vecindarios u otras formas de especificar dependencias espaciales que puedan regularizar la solución. En ocasiones estas restricciones espaciales se introducen por medio de la suavización del campo de iluminación estimado por un filtro pasabaja [103]. En concreto, en nuestros trabajos hemos estudiado la mejora de eficiencia que se consigue con un modelo de campo de iluminación dado por una combinación lineal de productos 2D de polinómios de Legendre.

Los métodos que se han propuesto para conseguir sistemas de reconocimiento de caras robustos frente a cambios de iluminación y/o pose consisten en generar las vistas a partir de modelos 3D o modelos lineales creados a partir de un conjunto de imágenes que modelan las condiciones de iluminación y pose. Estos métodos necesitan conocer las imágenes que se obtienen bajo las distintas condiciones de iluminación y pose, para poder construir la transformación de generación (*re-rendering*) de la imagen. En este trabajo apuntamos una solución alternativa que pasa por realizar la corrección de la imagen previamente a la realización de los procesos computacionales de reconocimiento. Este método no puede corregir las variaciones debidas a los cambios grandes de pose o al movimiento de la cámara. Por contra, no necesita todas las vistas de la cara, sino solamente una imagen de referencia que se puede obtener de un conjunto de imágenes “normales” de las caras.

Varios

Por supuesto que al profundizar en el tema de la tesis quedan aspectos del procesado de la imagen y de la visión por computador que, estando relacionados, se alejan del eje central de la tesis y no podemos hacer nada más que reconocer su existencia. Este es el caso del problema de la constancia del color [115] que tiene importancia en aplicaciones de reconocimiento de objetos y seguimiento de objetos en escenas. Si bien ha recibido especial atención en el campo de la búsqueda basada en contenidos en bases de datos de imágenes [53], la posibilidad de descomponer el color en las contribuciones debidas a la fuente de iluminación y al propio objeto es de interés para aplicaciones como la detección de caras en condiciones generales de iluminación, puesto que se ha comprobado la fuerte influencia que la cromaticidad de la fuente de luz tiene sobre la caracterización de los colores de piel.

1.2. Estructura de la memoria

En el capítulo inicial, capítulo 2 hacemos una introducción general del problema de la corrección de las inhomogeneidades de iluminación, con mención especial en las características de las imágenes de resonancia magnética nuclear (IRM).

En el capítulo 3 abordamos los antecedentes de los trabajos relativos a la corrección de las IMRN. Además de una introducción general, damos una descripción detallada de los algoritmos que han sido implementados para contrastar con

las proposiciones propias de esta tesis doctoral. En este capítulo presentamos también nuestras aportaciones concretas.

En el capítulo 4 presentamos resultados empíricos comparativos de los algoritmos descritos en el capítulo 3 sobre un conjunto de imágenes sintéticas del cerebro e imágenes obtenidas del repositorio IBSR [65].

En el capítulo 5 introducimos el problema de la corrección de la iluminación en el reconocimiento de caras y adaptamos nuestras proposiciones para este problema. El objetivo es mostrar la generalidad de la aproximación que seguimos y su potencialidad en esa área de investigación.

Como es habitual, en el capítulo 6 se resumen las conclusiones de la presente tesis doctoral y se indican las líneas de trabajo futuro abiertas a nuestro parecer.

Hemo incluido tres apéndices en la memoria que recojen algunos aspectos de interés que no tienen un espacio natural en el discurso de la memoria pero que son necesarios para hacerla sino autocontenida, al menos si completa. El primer apéndice recoge una descripción somera de los principios de la generación de la imagen de resonancia magnética nuclear. Este material está orientado a hacer inteligible el problema de la iluminación en IRM y la importancia de los algoritmos que incluyen el modelado de la iluminación en sus presupuestos. El segundo apéndice recoge la definición y algunas propiedades matemáticas de los polinomios de Legendre, con el objeto, esta vez si, de hacer autocontenida la memoria. El tercer y último apéndice presenta los algoritmos meméticos y las estrategias evolutivas. Este apéndice podría formar parte del capítulo 3 pero creemos que allí rompería la fluidez del discurso.

1.3.Publicaciones realizadas durante el desarrollo de la tesis

Los trabajos de la tesis doctoral han dado lugar a las siguientes publicaciones, sin incluir artículos en preparación.

El desarrollo y aplicación de estrategias evolutivas se ha descrito en las siguientes publicaciones

- Fernández E., M. Graña, J. Ruiz-Cabello, On a Gradient Based Evolution Strategy for Parametric Illumination Correction, *Electronic Letters*, 2004, 40 (9) , p.531-32 , introduce de forma breve nuestra aproximación.
- Fernández E., M. Graña, J. Ruiz-Cabello, On a Gradient Based Evolution Strategy for Parametric Illumination Correction, en *Information Process-*

ing with Evolutionary Algorithms, M. Graña, R. Duro, A. d'Anjou, P.P. Wang (eds) Springer Verlag, London, pp.61-72. Introduce de forma más extensa nuestra aproximación evolutiva a la corrección paramétrica de la iluminación

- Fernández E., M. Graña, J. Ruiz Cabello, (2004) On Evolution Strategies for Illumination Correction, Information Sciences, in press. Se presentan los experimentos más exhaustivos que demuestran la potencia de la utilización del gradiente en un algoritmo de estrategia evolutiva para la corrección paramétrica de la iluminación.
- Fernández E., M. Graña, J. Ruiz-Cabello, An Instantaneous Memetic Algorithm for Illumination Correction, Proc. 2004 Int. Congress of Evolutionary Computation (CEC2004). Introduce nuestro trabajo en el contexto de los algoritmos meméticos.

El desarrollo de algoritmos basados en el descenso de gradiente se han presentado en

- Fernández E., M. Graña, J. Ruiz Cabello, (2004) Simultaneous adaptive segmentation and parametric correction of illumination inhomogeneities in MRI, MEDSIP2004. Se introduce el descenso combinado de gradiente respecto de los parámetros del modelo de bias y de las medias de las intensidades en las clases. Se consigue la estimación simultánea del bias de la iluminación y de las intensidades medias de las clases.

2. EL PROBLEMA DE LA CORRECCIÓN DE ILUMINACIÓN

El problema de la corrección de la inhomogeneidad del campo de la iluminación (denominado Retinex en algunos contextos [30], “ânclaje de la luz” [62], o constancia del color en otros [53], [115]) aparece en prácticamente todas las áreas de aplicación del procesado de la imagen digital. En esta tesis doctoral nos hemos concentrado en el estudio de este problema en el contexto de la segmentación de la imagen de resonancia magnética y, con menos énfasis, en el contexto del reconocimiento de caras.

Existe muchas otras aplicaciones del proceso de la imagen digital en las que se presenta el problema de la inhomogeneidad de la iluminación y que caen fuera del ámbito de la tesis. Citaremos algunos para mostrar la generalidad de este problema.

El primer ejemplo es el proceso de las imágenes de reconocimiento remoto, en particular las imágenes hiperespectrales. Las imágenes hiperespectrales se obtienen mediante sensores que muestrean de forma fina el espectro luminoso visible e infrarrojo. Estas imágenes son particularmente sensibles a la inhomogeneidad en el campo de iluminación, puesto que cada pixel es un vector de alta dimensión que se ve afectado en sus componentes de muy diversas formas por los cambios de iluminación. En [42] se estudia el efecto que la corrección de la iluminación (usando información topográfica adicional) puede tener en la estimación de los espectros extremos (denominados *endmembers*) y en el subsecuente análisis de la imagen mediante algoritmos de descomposición espectral (*spectral unmixing*).

En otro ámbito más cercano a los problemas de reconocimiento, [48] estudia la corrección de los campos de iluminación en imágenes de color y su efecto en los sistemas de búsqueda en bases de datos de imágenes basadas en contenido (*Content Based Image Retrieval*), o [60] estudia la influencia de la iluminación en la clasificación visual de objetos. También en los sistemas de seguimiento (tracking) de objetos la iluminación es reconocida como un factor crítico.

Usualmente se consideran tan sólo variaciones globales (suma o resta de una constante a toda la imagen) para demostrar la robusted de los sistemas. Sin embargo, las variaciones direccionales de la iluminación que se manifiestan como inhomogeneidades del campo de iluminación todavía no han sido modeladas y estudiadas, ni consideradas experimentalmente [66], excepto en contextos muy restringidos como el futbol robotizado de la *robo-cup* [69].

Otro de los problemas centrales en visión por computador es el cálculo del flujo óptico. Recientemente se está teniendo en consideración la influencia de la iluminación inhomogénea y variable en este problema [76]. No sólo es de interés la corrección de iluminación en problemas de reconocimiento y segmentación, en [31] se describe un método de estabilización de la imagen (corrección de desplazamientos indeseados de la cámara) con precisión subpixel que pretende ser invariante a las variaciones de iluminación y su efecto en la eficiencia de la compresión de video en el estándar MPEG-2.

Muchas de estas aplicaciones comparten la formulación del problema que se da en las secciones siguientes. Sin embargo, en algunos casos el modelo de la influencia de la iluminación es bastante distinto y las soluciones que se presentan en esta tesis no son de aplicación para ellos.

En este capítulo presentamos en la sección 2.1 el problema en su formulación formal más general. La sección 2.2 presenta el filtrado homomórfico como algoritmo básico de restauración de la imagen, que puede asumirse como referente para los demás. La sección 2.3 presenta una discusión sobre el papel de los algoritmos de corrección en el proceso de la imagen.

2.1. Planteamiento general

Podemos tomar un planteamiento abstracto y general del problema de la iluminación partiendo de un modelo simple de formación de la imagen digital, que es común a todas las formas de sensorización. Siguiendo la notación del libro [55] y considerando imágenes con soporte bidimensional (no volúmenes), este modelo se expresa en la siguiente ecuación:

$$f(x, y) = b(x, y) r(x, y) + \eta(x, y), \quad (2.1)$$

donde $r(x, y)$ se corresponde con la imagen original no corrompida, o con el mapa de reflectancia de la escena; el bias de la iluminación $b(x, y)$ modela el efecto de la distribución de energía que ilumina la escena y nos permite percibirla;

$\eta(x, y)$ modela el ruido aditivo (usualmente Gaussiano) que corresponde a efectos inevitables en el sistema de sensorización.

El factor $b(x, y)$ modela tanto el efecto de las fuentes de luz en los sistemas basados en sensores ópticos, como el efecto de la distribución de la energía del pulso de radio frecuencia en las imágenes de resonancia magnética nuclear y de la homogeneidad del campo magnético externo B_0 . En general el objetivo es obtener una estimación de $r(x, y)$ que nos permita obtener información exacta sobre la escena y los objetos en ella, ya sea para realizar tareas de reconocimiento y clasificación, ya sea para realizar tareas de metrología.

2.1.1. Características específicas del problema para IRM

Se asume en general que la función de intensidad de la imagen real es constante por trozos, esto es, la señal de RM detectada en un tejido es constante y su valor teórico puede ser deducido si se conoce la composición química del tejido. Esto es, los posibles valores de intensidad son un conjunto discreto:

$$f(x, y) \in I, \quad (2.2)$$

donde $I = \{\mu_1, \dots, \mu_K\}$, y K es el número de clases de tejidos en la imagen.

En general los algoritmos de corrección de la inhomogeneidad de la iluminación encontrados en la literatura asumen esta estructura de la intensidad de la imagen de resonancia magnética nuclear, y de hecho ésta es clave para su formulación y desarrollo, con excepciones como [112], donde no asumen ningún modelo de distribución de la intensidad en clases y operan directamente sobre el histograma de la imagen.

2.1.2. Ilustración de los efectos de la iluminación

Para ilustrar los efectos de los distintos elementos en la ecuación 2.1 consideremos el histograma de la figura 2.1 que corresponde a una imagen de un tablero de ajedrez con dos únicos niveles de gris, que se muestra en la figura 2.2. Esta imagen es la que nos sirve de base para algunos de los experimentos que se describen más adelante. Uno de los procesos más comunes que se realizan con las imágenes es la segmentación en regiones. En el caso de la figura del tablero de ajedrez esta segmentación es trivial y puede realizarse en términos de los niveles de gris directamente: se asocia cada una de las dos regiones con un nivel de gris. Los sistemas automatizados tratarían en este caso de descubrir cuáles son los

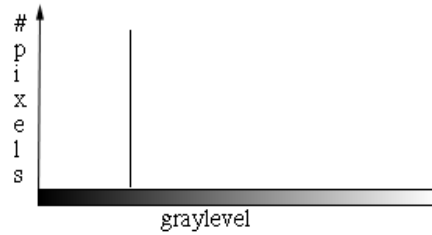


Figura 2.1: Histograma de una imagen de tablero de ajedrez con dos niveles de gris, sin corrupciones de ningún tipo.

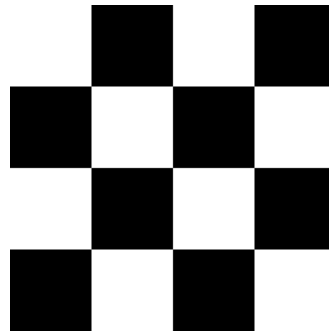


Figura 2.2: Imagen de tablero de ajedrez original

niveles de gris que se asocian con cada clase a partir del histograma, lo que en el caso de la figura 2.1 es trivial.

Cuando se añade ruido aditivo gaussiano a la imagen, que corresponde al término $\eta(x, y)$ en la ecuación 2.1 obtenemos imágenes como la de la figura 2.4 cuyos histogramas son análogos al de la figura 2.3. Se aprecia visualmente en el histograma que se mantienen diferenciadas dos masas de probabilidad que corresponden *grosso modo* con las dos clases de píxeles que constituyen la imagen. Los procesos automáticos de segmentación realizan normalmente procesos de agrupamiento (*clustering*) sobre el histograma. Dado el histograma en la figura 2.3 no resulta difícil para estos algoritmos descubrir las dos clases de píxeles y realizar la partición de la imagen, si bien normalmente no descubrirá los valores exactos de los niveles de gris de las clases en la imagen limpia.

En la figura 2.5 se muestra un campo de iluminación no homogéneo, que corresponde al factor $b(x, y)$ en la ecuación 2.1. Multiplicando este campo a la

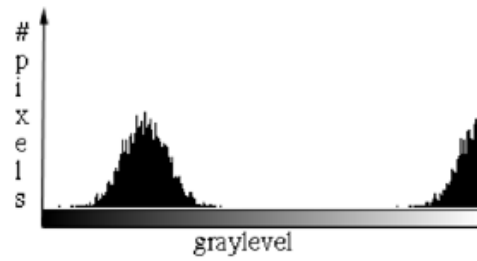


Figura 2.3: Histograma de una una imagen de tablero de ajedrez corrompida con ruido gaussiano aditivo.

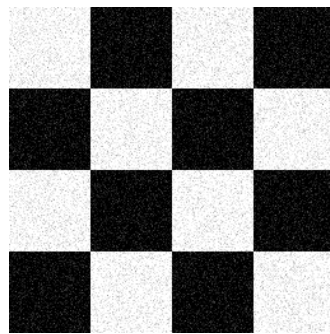


Figura 2.4: Imagen del tablero de ajedrez corrompida por ruido gaussiano aditivo.

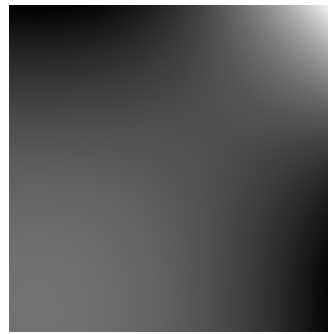


Figura 2.5: Ejemplo de campo de iluminación no uniforme.

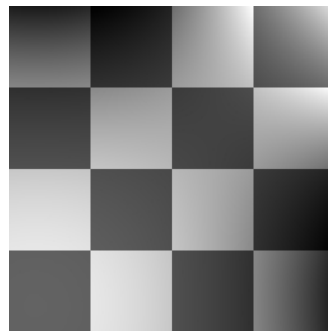


Figura 2.6: Ejemplo de imagen corrompida con un campo de iluminación.

imagen original del tablero de ajedrez, obtenemos la imagen corrupta sintética presentada en la figura 2.6. El histograma que se obtiene de la imagen corrupta se presenta en la figura 2.7. En este caso es muy difícil que un algoritmo de agrupamiento sea capaz de descubrir las dos clases originales, y mucho más que sea capaz de identificar los niveles de gris en la imagen original. Es por esta razón que se necesitan algoritmos que incluyan explícitamente la estimación del campo de iluminación cuando este es un efecto no despreciable.

Efecto sobre la clasificación Para tratar de mostrar de forma cuantitativa el efecto relativo de ambas formas de corrupción de la imagen hemos realizado el siguiente experimento:

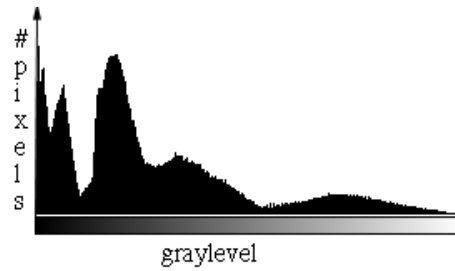


Figura 2.7: Histograma de la imagen de tablero de ajedrez corrompida por un campo de iluminación no uniforme.

1. Corrompemos la imagen original de la tabla de ajedrez en un caso con ruido gaussiano de media cero y amplitud creciente, y en otro con un campo inhomogeneo que también está amplificado por un factor de ganancia creciente.
2. Calculamos los representantes de los agrupamientos aplicando el algoritmo de las k-medias.
3. Calculamos la clasificación de los pixeles en clases correspondientes según estos representantes, obteniendo una segmentación de la imagen.
4. Calculamos la tasa de aciertos en la clasificación no supervisada como la correlación de la imagen clasificada con la distribución original de clases.

¹

En la figura 2.8(a) presentamos la evolución de los aciertos en función de la relación señal ruido (SNR) de la imagen corrompida con ruido aditivo gaussiano de amplitud creciente. Se observa que la clasificación obtenida es perfecta hasta valores de SNR muy bajos.

En la figura 2.8(b) se muestra el resultado del mismo experimento realizando la corrupción de la imagen añadiéndole un bias de iluminación multiplicado por

¹Al hacer esto estamos asumiendo que la clasificación no supervisada es equivalente a una clasificación supervisada. Esto es aceptable habida cuenta que solo hay dos clases, y no tenemos peligro de que se produzcan confusiones en la identificación de las clases: la clase 1 es la de menor intensidad y la 2 la de mayor intensidad. Obviamente esto no ocurre cuando tenemos más clases y el espacio de las características es multidimensional.

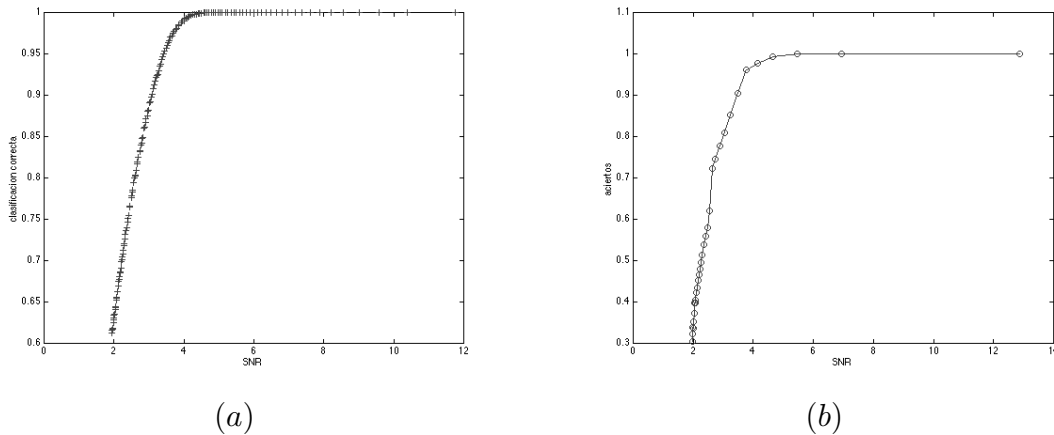


Figura 2.8: Resultados de clasificación versus SNR en el caso (a) de ruido gaussiano y (b) de campo de iluminación inhomogeneo.

un parámetro de ganancia creciente, de forma que obtenemos una secuencia de imágenes con SNR decreciente (distorsión creciente).

Se puede apreciar comparando las gráficas de las imágenes que para valores de SNR altos se obtienen una clasificación perfecta. Sin embargo cuando decrece el SNR la tasa de aciertos se degrada mucho más rápidamente en el caso de las imágenes corrompidas por el campo de iluminación. Para un SNR de 2 se obtiene una tasa de aciertos del 30% frente al 60% que se obtiene para un SNR similar en el caso de las imágenes corrompidas por ruido gaussiano aditivo. Para evidenciar completamente el efecto de la iluminación mostramos los resultados de clasificación en dos casos de SNR similar. En la figura 2.9 se muestra la clasificación obtenida por el algoritmo de las *k*-medias sobre la imagen presentada en la figura 2.6, mientras que en la figura 2.10 muestra la clasificación sobre la imagen de la figura 2.4. Se aprecia claramente el efecto catastrófico del campo de iluminación inhomogeneo.

Efecto sobre el gradiente En muchos de los algoritmos de segmentación se utilizan los gradientes de la imagen, para aproximar los contornos de los objetos o bien para limitar la detección de las regiones en la imagen. Si consideramos, por ejemplo, la derivada en x de la imagen, sustituyendo la ecuación 2.1:

$$\frac{\partial}{\partial x} f(x, y) = \frac{\partial}{\partial x} [i(x, y) \cdot r(x, y)] + \frac{\partial}{\partial x} \eta(x, y). \quad (2.3)$$

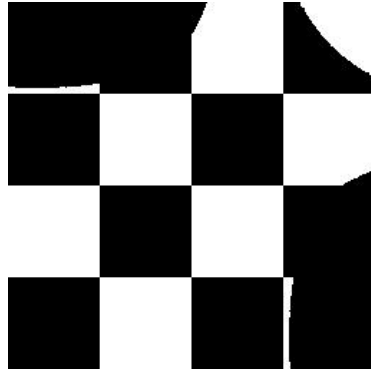


Figura 2.9: Clasificación obtenida aplicando el algoritmo k-medias a una imagen corrompida con un campo de iluminación inhomogeneo.

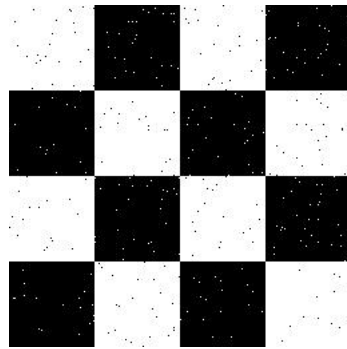


Figura 2.10: Imagen de la clasificación obtenida por el algoritmo de las k-medias sobre una imagen corrompida por ruido gaussiano aditivo.

El efecto del ruido aditivo puede eliminarse de la imagen de gradiente realizando un filtrado de suavización previo al cálculo del gradiente. Sin embargo, es fácil ver que el efecto de la inhomogeneidad del campo de iluminación no se elimina fácilmente. Si consideramos el término correspondiente tenemos:

$$\frac{\partial}{\partial x} [i(x, y) \cdot r(x, y)] = r(x, y) \frac{\partial}{\partial x} i(x, y) + i(x, y) \frac{\partial}{\partial x} r(x, y). \quad (2.4)$$

El cálculo de la derivada tiene dos componentes que corresponden a

- la derivada del campo de iluminación, que suele ser reducida debido a que el campo de iluminación es una función suave, multiplicada por la reflectancia. Esto puede hacer que este término tenga influencia considerable en el gradiente calculado en las regiones con valores altos de reflectancia.
- la derivada de la reflectancia que aparece atenuada o amplificada por el campo de iluminación. En algunos casos, esta atenuación puede ser muy importante.

En la figura 2.11 se muestran un corte axial de una imagen sintética de cerebro y los bordes detectados en la imagen. La imagen 2.11(a) presenta la imagen original y la figura 2.11(b) presenta la imagen corrompida por un campo de iluminación. La imagen 2.11(c) es la imagen de bordes obtenida de la imagen 2.11(a), mientras que la imagen 2.11(d) es la imagen de bordes de la imagen 2.11(b). Se aprecia claramente la atenuación que produce el campo de iluminación en la imagen de bordes.

2.2.El filtrado homomórfico

El filtrado lineal no puede directamente separar y eliminar los efectos de un campo de iluminación inhomogeneo [55], [15], [100]. Sin embargo si se puede plantear un esquema de filtrado lineal sobre manipulaciones de la imagen y bajo ciertas asunciones. Si en la ecuación 2.1 descartamos el ruido aditivo y consideramos la ecuación:

$$f(x, y) = b(x, y) r(x, y). \quad (2.5)$$

Podemos considerar la transformación logarítmica de esta ecuación:

$$\ln f(x, y) = \ln b(x, y) + \ln r(x, y). \quad (2.6)$$

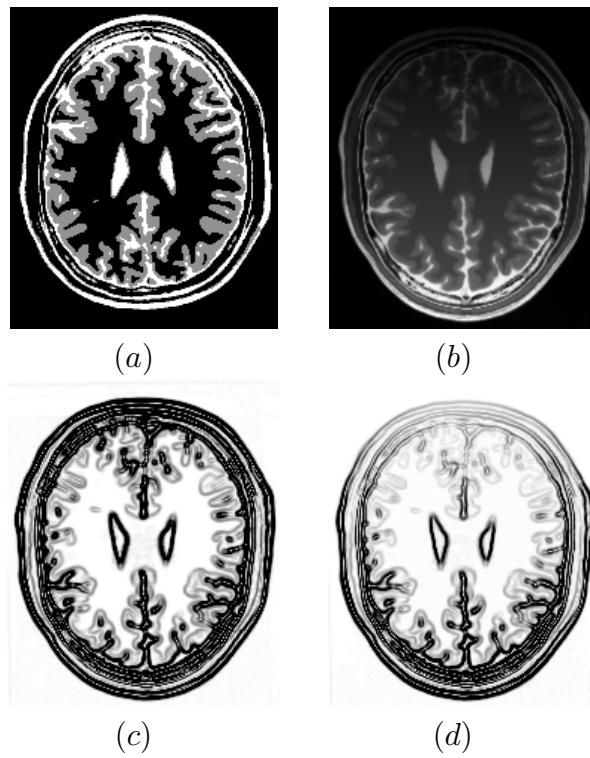


Figura 2.11: Resultados del cálculo de los bordes en el caso (a) de la imagen original en (c) y (d) de la corrompida con un campo de iluminación inhomogeneo

De esta forma el modelo se convierte en un modelo lineal en la iluminación y la reflectancia. Si consideramos la transformada de Fourier de esta imagen logarítmica tenemos:

$$F(u, v) = B(u, v) + R(u, v), \quad (2.7)$$

donde $F(u, v)$, $B(u, v)$ y $R(u, v)$ son, respectivamente, los pares transformados de $\ln f(x, y)$, $\ln b(x, y)$ y $\ln r(x, y)$. El filtrado lineal de esta imagen en el espacio transformado:

$$F(u, v) H(u, v) = B(u, v) H(u, v) + R(u, v) H(u, v), \quad (2.8)$$

donde $H(u, v)$ es la función de transferencia de un filtro lineal, nos permite eliminar uno de los componentes de la ecuación, siempre y cuando estos tengan dominios disjuntos en el espacio. Esto es, sea $D_B = \{(u, v) \mid |B(u, v)| \neq 0\}$, $D_R = \{(u, v) \mid |R(u, v)| \neq 0\}$ los dominios en los que no se anulan las transformadas de los logarimos del bias de iluminación y la reflectancia, si tenemos que

$$D_B \cap D_R = \emptyset, \quad (2.9)$$

entonces es posible diseñar un filtro que elimine por completo los efectos de la iluminación preservando la reflectancia, por ejemplo haciendo que $H_B(u, v)$ se anule en D_B y sea la identidad para el resto del dominio frecuencial, y en especial en D_R :

$$H_B(u, v) = \begin{cases} 0 & (u, v) \in D_B \\ 1 & (u, v) \notin D_B \end{cases}. \quad (2.10)$$

Entonces, tendríamos que

$$\begin{aligned} F(u, v) H_B(u, v) &= B(u, v) H_B(u, v) + R(u, v) H_B(u, v) \\ &= R(u, v), \end{aligned} \quad (2.11)$$

y podríamos recuperar la reflectancia calculando

$$r(x, y) = \exp(\mathcal{F}^{-1}(F(u, v) H_B(u, v))), \quad (2.12)$$

donde \mathcal{F}^{-1} denota la inversa de la transformada de Fourier. Esta es la aplicación usual del filtrado homomorfo, que se puede compactar en la siguiente formulación:

$$\hat{r}(x, y) = \exp(\mathcal{F}^{-1}(\mathcal{F}(f(x, y)) H_B(u, v))), \quad (2.13)$$

donde \mathcal{F} y \mathcal{F}^{-1} denotan, respectivamente, la transformada de Fourier directa e inversa.

En general, la condición 2.9 no se cumple para imágenes usuales y, en particular, para las imágenes de resonancia magnética. Usualmente se asocian los efectos de iluminación con funciones suaves y, por tanto, regiones de bajas frecuencias. Por ello, los filtros utilizados son filtros pasa-alta que eliminarían los componentes de la transformada en estas frecuencias. Sin embargo, lo habitual es que la imagen de reflectancia contenga componentes de baja frecuencia significativos y que no se pueden descartar sin incurrir en errores. Por ejemplo, en las imágenes de resonancia magnética es de esperar que grandes regiones tengan un valor constante correspondiente a un tejido y que, por tanto, los componentes de baja frecuencia de la imagen de reflectancia sean no despreciables.

Para ilustrar las dificultades que el filtro homomorfo tiene para extraer los componentes de iluminación y reflectancia en imágenes como las que nos interesan, presentamos los resultados que se obtienen sobre la imagen 2.6. La imagen de reflectancia de esta imagen, la imagen “limpia” tiene grandes regiones constantes, por lo que es un caso extremo de los discutidos arriba. La figura 2.12 presenta (a) la reflectancia y (b) el campo de iluminación, obtenidos utilizando un filtro frecuencial construido a partir de una función gaussiana de varianza 5. Se aprecia que la mayor parte de la información queda en la imagen del campo de iluminación y que la reflectancia se reduce a una imagen de bordes. Además el rango de valores de la reflectancia está considerablemente comprimido y es preciso amplificarlo considerablemente para obtener la imagen de la figura 2.12a.

En la figura 2.13 presentamos los resultados con el mismo filtro homomorfo sobre una rodaja de un volumen simulado de cerebro obtenido de [14] y corrompido sintéticamente por un campo de iluminación no uniforme. La figura 2.13a presenta la rodaja corrupta. La figura 2.13b presenta la imagen de reflectancia calculada. También en este caso el rango de valores está muy comprimido y precisa amplificación para obtener la visualización. Por último, la figura 2.13c presenta el campo de iluminación estimado. También en este caso el campo tiene mucha información estructural de las regiones de la imagen original, que obviamente son regiones suaves.

2.3. Clasificación versus restauración

El problema de la iluminación en las imágenes puede abordarse desde dos perspectivas que son relativamente complementarias:

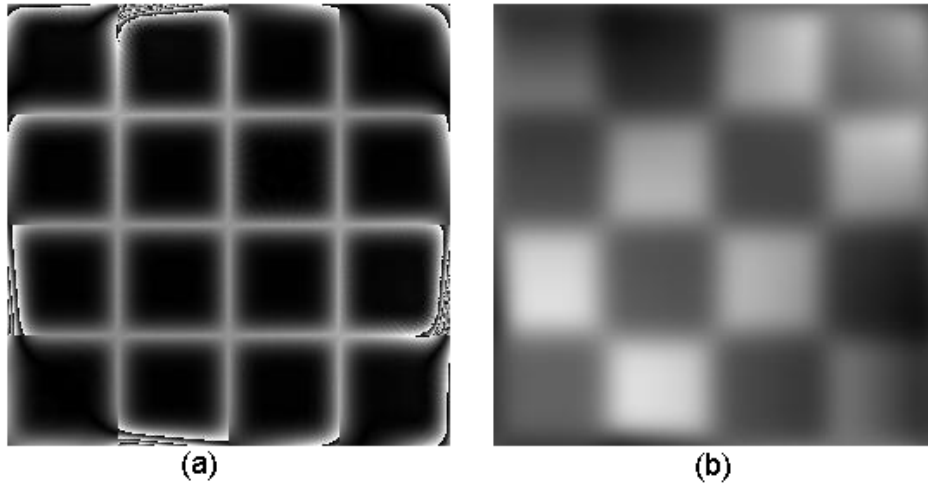


Figura 2.12: Resultados del filtrado homomórfico sobre una imagen de tablero de ajedrez corrupta (a) reflectancia, (b) campo de iluminación estimado.

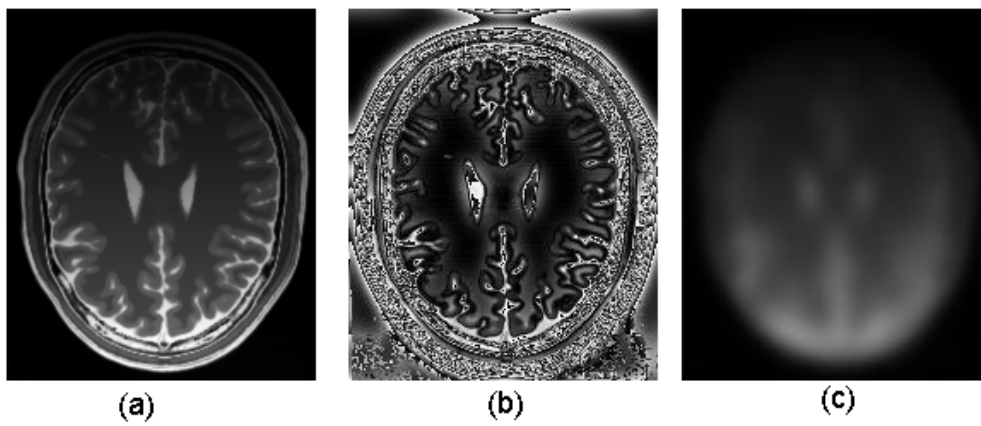


Figura 2.13: Resultados de la aplicación del filtro homomórfico sobre una rodaja de un volumen de cerebro (a) imagen corrupta, (b) reflectancia y (c) campo de iluminación.

- Clasificación: se trata de identificar regiones en la imagen con una cierta significación física, para ello se clasifica cada pixel en alguna de las clases. El problema puede incluir la identificación de los parametros asociados a las clases (p.e.: la media y varianza de la intensidad). Dentro del proceso de clasificación pueden incluirse parámetros espaciales.
- Restauración: se trata de restaurar la imagen que correspondería a la escena bajo unas condiciones de iluminación ideales. No se impone ninguna restricción a la imagen limpia.

Se aprecia en la literatura que para algoritmos cuyo objetivo es la clasificación de los pixeles y la identificación de regiones no es de interés obtener una buena restauración de la imagen original, por lo que posiblemente los resultados de los algoritmos diseñados con este objetivo no incluirán una buena corrección de la imagen. El caso de los algoritmos de restauración es similar, si bien es de esperar que la clasificación sobre una imagen restaurada será de mejor calidad que sobre una imagen corrupta.

La distinción fundamental entre estas dos aproximaciones se produce considerando como se evaluarán los resultados de los algoritmos. En el caso de los algoritmos de clasificación, esta evaluación se realizará contra la verdad del terreno dada por clasificaciones de la imagen obtenidas de expertos, dando los porcentajes de error y aciertos. En el caso de los algoritmos de restauración, la evaluación de los resultados se realizará contra las imagenes no corruptas, si se conocen, en términos de la distorsión, relación señal-ruido o correlación.

El algoritmo de filtrado homomórfico presentado en la sección 2.2 es un algoritmo de restauración, que no nos proporciona información alguna sobre la clasificación de los pixeles, la descomposición de la imagen en regiones u otros procesos de reconocimiento. El algoritmo de las k-medias es un algoritmo típico de agrupamiento, que puede aplicarse para la estimación de los representantes de la intensidad de las clases en la imagen. Este algoritmo nos proporcionará una segmentación de la imagen e información sobre las clases, pero no imagen restaurada.

Es más, la clasificación presenta problemas específicos, como son la aparición en IRM de clases desconocidas correspondientes a anomalías biológicas (tumores) que exigen su detección. Una forma de atacar el problema es, por ejemplo, usar información geométrica sobre los organos visualizados para poder detectar anomalías mediante el registro de la imagen sobre la información geométrica [94].

3. ALGORITMOS DE CORRECCIÓN DEL BIAS EN IRM

En este capítulo presentamos la descripción formal de los algoritmos de corrección del bias y clasificación de los píxeles de la imagen de RM, tanto nuestras proposiciones como los algoritmos de comparación. En el capítulo 4 presentamos los resultados empíricos sobre imágenes simuladas e imágenes reales de RM.

Debido a su importancia, distinguimos dos aproximaciones generales que han producido numerosas aportaciones en la literatura y les dedicamos sendas secciones introductorias. En la sección 3.1 delineamos la descripción general de los métodos bayesianos de proceso de la imagen que se originan a partir del trabajo de Geman [49]. En la sección 3.2 presentamos los algoritmos de agrupamiento borroso, que tienen su raíz en los trabajos de Bezdek (*fuzzy clustering*) [9], [12].

En la sección 3.3 presentamos los métodos no paramétricos, incluyendo la descripción detallada de BMAP y BFCM como representantes, respectivamente, de la formulación bayesiana [49] y de la basada en algoritmos de agrupamiento borroso.. En la sección 3.4 presentamos los métodos paramétricos, incluyendo la descripción detallada de los algoritmos PABIC y el algoritmo adaptativo de Wells, que también puede incluirse dentro de los métodos bayesianos. Finalmente, en la sección 3.5 presentamos los algoritmos que son nuestra aportación a la segmentación y corrección de la iluminación en IRM.

En la descripción de los algoritmos incluimos la formulación del modelo de imagen que es la base de su desarrollo. Si bien estos modelos se ajustan en general al modelo dado en la sección 2.1, existen pequeñas variaciones y detalles que reflejan la idiosincrasia de cada autor y que merecen ser tenidas en cuenta para una mejor comprensión de los algoritmos.

3.1. Algoritmos basados en métodos bayesianos.

La formulación bayesiana [49], [50], [122] del procesado de las imágenes digitales, incluyendo la corrección de bias de iluminación, así como la restauración ante diversos tipos de ruido y la segmentación en regiones o contornos, es un área muy activa de investigación. También se ha aplicado extensamente a las imágenes de resonancia magnética [33], [61], [68], [89], [91], [103], [104], [126]. Merece una consideración especial esta clase de métodos que pueden dar lugar tanto a métodos paramétricos como no paramétricos.

En esencia el proceso bayesiano consiste en formular el problema como uno de estimación del objeto de máxima probabilidad *a posteriori*. Utilizando el teorema de Bayes, la probabilidad *a posteriori* se descompone en las probabilidades condicionales y *a priori*. Usualmente la probabilidad condicional modela el ruido que afecta la percepción de la imagen y es asumido como gaussiano. La probabilidad *a priori* de la imagen se modela habitualmente como un campo markoviano. De esta forma se obtiene un medio elegante de especificar restricciones espaciales en las imágenes a través de los potenciales de los cliques en el campo markoviano. De esta formulación resulta que la distribución de probabilidad *a posteriori* tiene forma exponencial y su exponente es una función de energía. El problema de la estimación de máxima probabilidad *a posteriori* (MAP) se convierte en esta formulación en un problema de minimización de la función de energía que se puede realizar mediante métodos codiciosos (*greedy*), que son en realidad métodos locales, o mediante métodos de búsqueda globales, como el enfriamiento simulado [1], [74]. La eficiencia se inclina habitualmente por los métodos locales. En la sección 3.3 expondremos con detalle la aproximación propuesta por [104], a continuación comentamos las peculiaridades de algunas de las aproximaciones bayesianas encontradas en la literatura, que reflejan la variedad de formulaciones que pueden realizarse en este marco teórico.

En [61] la probabilidad condicional se modela mediante ventanas de Parzen [28] para tomar en cuenta las variaciones debidas a las inhomogeneidades en los tejidos y los ruidos de captura. De esta forma consigue modelar distribuciones condicionales no gaussianas, a cambio de un considerable costo computacional. Las probabilidades *a priori* modelan la relación entre tejidos vecinos, penalizando las clasificaciones distintas entre vecinos. También introduce las inhomogeneidades debidas a las bobinas de radio frecuencia como una probabilidad *a priori* adicional. La inhomogeneidad se modela como un campo markoviano.

Los autores prueban la implementación de los modos de iteración condicionales (*Iterated Conditional Modes: ICM*) y del enfriamiento estadístico para el cálculo de la estimación MAP del campo de iluminación y de la segmentación de la imagen.

En [91] se realiza mediante métodos bayesianos la correspondencia entre un atlas anatómico del craneo-cerebro con el volumen de IRM. El proceso es, por tanto, un proceso mixto de reconocimiento y registro geométrico, el cual, además de la clasificación de los voxels en tejidos, proporciona una transformación no rígida entre el atlas y la imagen del sujeto que facilita la segmentación fiable del craneo y el cerebro en la IRM, así como la estimación precisa de las probabilidades *a priori* en cada voxel. El objetivo es la detección precisa de anomalías que pueden corresponder a situaciones patológicas. En este trabajo se modelan los campos de iluminación mediante combinaciones de splines pesados por las probabilidades *a priori* de las clases en cada pixel. Cada clase de tejido tiene su propio campo global de iluminación, lo que responde a la observación de la interacción entre la composición química de los tejidos y la señales de radio frecuencia. Una aproximación similar para la extracción del cerebro se halla en [89], donde se combina la detección basada en la clasificación de los píxeles con la segmentación basada en contornos dinámicos.

Aunque cae fuera del ámbito de esta tesis, una de las modalidades de IRM más modernas es la imagen funcional IRMf (*fMRI*), que registra la evolución temporal de la señal en cada voxel de la imagen. Este tipo de imágenes tiene aplicación e interés en el estudio de fenómenos dinámicos, como el flujo sanguíneo en el cerebro. En [68] se presenta la aplicación del modelado bayesiano a este tipo de imágenes, con el objetivo de segmentar regiones de especial actividad. La formulación bayesiana se complica en este caso por el modelo de mezclas lineales subyacente. El problema en este caso se centra en inferir a partir de la distribución *a posteriori* si un pixel está activo o no, esto es, si refleja actividad o no.

En [126] se realiza una extensión del modelo probabilístico *a priori*, los campos markovianos: se presentan y fomulan los campos markovianos ocultos (*HMM-RF: hidden markov random field*), de forma análoga a las cadenas de markov ocultas que se utilizan en síntesis y reconocimiento de voz. En este modelo el campo markoviano oculto representa las clases de tejidos, mientras que las intensidades observables se modelan como una mezcla lineal de las intensidades de las clases de tejidos.

Finalmente, en [33] se presenta un método variacional no paramétrico de

estimación del campo de iluminación y de la segmentación MAP de la imagen de resonancia.

3.2. Agrupamiento borroso

La aplicación de métodos basados en técnicas de agrupamiento borroso (*fuzzy clustering* [10]) a la segmentación de imágenes de resonancia magnética ha sido extensa [3], [7], [8], [11], [16], [17], [32], [57], [58], [99], [127], [77]. El problema se considera como la clasificación no supervisada de los píxeles de la imagen en clases. La aproximación borrosa consiste en asociar etiquetas de pertenencia a las clases que no son exclusivas, esto es, un píxel puede tener posibilidad no nula de pertenencia a más de una clase. Por contra, los métodos no borrosos usualmente asocian a cada píxel un coeficiente pertenencia rígido (*crisp*) que es cero para todas las clases excepto una. La diferencia entre los coeficientes de pertenencia borrosos y una distribución de probabilidad de pertenencia no está clara, sobre todo cuando se exige que los coeficientes borrosos sumen uno para cada píxel. El proceso de clasificación se convierte en un proceso de minimización de una función objetivo que asocia a cada píxel la distorsión respecto de todas las clases ponderada por los coeficientes de pertenencia elevados a una potencia que modula el grado de borrosidad de los agrupamientos. En esta función objetivo puede también incluirse el campo de iluminación mediante una sencilla sustitución. Formulada la función objetivo, la minimización puede realizarse por medio de alguna de las múltiples estrategias de optimización, desde los métodos de descenso de gradiente hasta los algoritmos evolutivos [58].

Los trabajos clásicos [7], [11], [16] se dedicaron a la segmentación de la imagen de resonancia, en especial a las imágenes multimodales, donde cada píxel se representa por un vector de características que corresponden a los valores de cada modalidad, usualmente imágenes pesadas en T1, T2 y de densidad de protones. De esta forma se conseguía obtener segmentaciones de calidad, dado que en algunos casos, los tejidos se discriminan bastante bien en este espacio 3D. En algunos casos se realizaron comparaciones directas con redes neuronales artificiales supervisadas [57] entrenadas con regiones seleccionadas de la imagen.

En [8] se introduce la noción de validación de los agrupamientos para guiar el proceso de reagrupamiento (*reclustering*). El proceso se convierte en un proceso iterativo, en el que se mejora la calidad del agrupamiento aproximándose a los resultados de los algoritmos supervisados. La validación utilizada es una medida de separación y compactación de los agrupamientos, y el proceso de rea-

grupamiento es un proceso de corte y mezcla guiado por la medida de validación.

En [17] el agrupamiento borroso se utiliza para realizar una segmentación inicial de la imagen multimodal de la cabeza, que se pasa a un sistema de razonamiento basado en conocimiento para la extracción del craneo. Posteriormente, se utilizan histogramas multiespectrales para la separación de los tejidos tumorales de los tejidos sanos. Esta línea de trabajos comienza con [77], donde se formula una estrategia de aplicación del algoritmo de agrupamiento borroso similar.

También en [127] se propone un método que combina el agrupamiento borroso y métodos de representación del conocimiento y razonamiento basados en reglas para la segmentación y proceso de imágenes. El agrupamiento borroso se utiliza para la extracción de características comunes al tipo de imágenes que se van a procesar y se entiende, por tanto, como un proceso de aprendizaje de las distribuciones de estas características. El método se propone como un método general aplicable, entre otros, a las imágenes de IRM.

En [99] se introduce un factor de inhomogeneidad en la función objetivo del agrupamiento borroso que modela el campo de iluminación. Para conseguir un algoritmo adaptativo se deriva la función objetivo respecto de los parámetros, incluido el campo de iluminación, y se iteran las ecuaciones obtenidas igualando estas derivadas a cero. Para evitar la solución trivial se introducen términos de regularización que imponen una estructura suave al campo de iluminación estimado.

En [19] se aplican tanto el agrupamiento borroso como el mapa autoorganizado (*SOM: Self Organizing Map*) a las imágenes funcionales de resonancia magnética, en un proceso en cascada, para la obtención de un mapa de actividad. El SOM se aplica como un preproceso que identifica agrupamientos secundarios que son luego refinados en agrupamientos primarios mediante el agrupamiento borroso, los que finalmente proporcionan la segmentación de la imagen.

3.3. Métodos no paramétricos

Los métodos no paramétricos incluyen explícitamente la estimación del campo de iluminación, pero no proporcionan ningún modelo espacial global de dicho campo. Se estima el campo en cada pixel de forma independiente y, en ocasiones, se filtra el campo estimado mediante un filtro pasabaja o una máscara de suavización para evitar la solución trivial en la que el campo estimado coincide con la diferencia en cada pixel respecto del valor nominal de la clase

asociada con él.

Algunos de los métodos son heurísticos, producto de la experiencia y la intuición de los diseñadores. En [24] la corrección de la iluminación se calcula localmente como la diferencia entre las medianas del histograma de las intensidades en un vecindario y la mediana del histograma global para la clase más representativa en el vecindario. En [21] el proceso que se realiza consiste en seleccionar los píxeles que corresponden al aire, que tienen intensidad casi nula, sustituyendo su intensidad por el promedio de las intensidades no nulas. La imagen resultante se suaviza. La imagen corregida se calcula como el ratio de la imagen original a la imagen suavizada multiplicado por la intensidad media de la imagen. Esta aproximación es equivalente a considerar que la imagen suavizada es la imagen de referencia para la corrección de la intensidad.

Estamos interesados en algoritmos desarrollados a partir de principios generales. Una de estas aproximaciones es el proceso de la imagen basada en el modelado bayesiano del problema [49], [50]. Se trata de obtener una estimación de máxima probabilidad *a posteriori* (MAP) de la clasificación de los píxeles, sujeto a la estimación MAP del campo de iluminación.

El primero de los algoritmos no paramétricos que implementamos y estudiamos es el algoritmo bayesiano *Biased MAP* (BMAP) propuesto por [104]. En este algoritmo se itera la estimación de la clasificación MAP de los píxeles y la estimación de máxima verosimilitud de los parámetros del modelo (el campo de iluminación, la media y varianza de las clases de intensidad). La probabilidad condicional de la intensidad respecto de la clase se modela como una distribución gaussiana, mientras que la probabilidad a priori se modela como un Campo Markoviano (*MRF: Markov Random Field*) con vecindario de orden 2 cuyos potenciales se estiman al comenzar el proceso. Los potenciales se estiman a partir de una clasificación inicial que puede realizarse con el algoritmo de las *k*-medias u otro similar. BMAP utiliza la iteración de modos condicionales (*ICM: iterated conditional modes*) para realizar la estimación MAP de la clasificación de los píxeles. La estimación de los parámetros de la probabilidad condicional se realiza de forma directa dada una clasificación de los píxeles de la imagen. La estimación del campo de iluminación también es directa.

Otra categoría de aproximaciones son los métodos de agrupamiento, en particular consideramos aquí los métodos borrosos. En estos algoritmos se trata de obtener, conocida una estimación del campo de iluminación, una clasificación no supervisada de los píxeles en agrupamientos que corresponden a los tejidos de la imagen. La estimación del campo de iluminación viene dada usualmente como

un cociente o diferencia entre el valor predicho y el valor actual de la intensidad en la imagen. Los algoritmos de agrupamiento borroso usualmente se definen como la minimización de una función criterio que es una variante borrosa de la distorsión [9], [10], [12]. En el algoritmo *Biased Fuzzy C Means* (BFCM) se introduce un término de vecindario en esta función objetivo para forzar condiciones de continuidad en la clasificación de los píxeles. La estimación de los parámetros se realiza de forma directa resolviendo las ecuaciones diferenciales de la función objetivo respecto de los parámetros.

3.3.1.BMAP

En nuestros trabajos hemos implementado con fidelidad la proposición del BMAP [104] como ejemplo de las aproximaciones bayesianas a la clasificación y corrección de IRM. En esta sección vamos a describir este algoritmo ciñendonos a la nomenclatura y notación usada en [103]. Como todos los algoritmos bayesianos, la descripción abarca la formulación de las probabilidades condicionales y la probabilidad *a priori* que nos dan la forma de la función de energía a minimizar para obtener la estimación MAP deseada. El algoritmo de minimización ideal sería un algoritmo global, como el simulated annealing [1], [74], pero las restricciones computacionales obligan a la aplicación de algoritmos de búsqueda local, en este caso el ICM. Comenzamos introduciendo la notación utilizada por los autores.

Se denota el lugar de un píxel en la malla como $i \in I \subset \mathbb{N}^2$ (en caso de imágenes 2D). La imagen observada se denota

$$\mathbf{y} = (y_i; i \in I).$$

donde y_i es la intensidad observada en el píxel i . La segmentación de una imagen expresada como la clasificación de cada píxel en alguna de las clases predefinidas se denota:

$$\mathbf{x} = (x_i; i \in I, x_i \in \Omega).$$

donde $\Omega = \{\omega_1, \dots, \omega_K\}$ es un conjunto de clases de objetos a los que se asocian los píxeles. Cada clase corresponde a un valor de intensidad en el proceso de generación de la imagen ideal (sin ruido). La ecuación de generación de la imagen se formula como:

$$y_i = (1 + \beta_i) \mu_{x_i} + \eta_i,$$

donde η_i es una variable aleatoria que modela el ruido aditivo, β_i es el sesgo en la intensidad debido al campo de iluminación en el píxel, y μ_{x_i} es la intensidad

correspondiente a la clase del pixel. El ruido se considera en [104] que sigue una distribución gaussiana de varianza dependiente de la clase. Esto es, si $x_i = \omega_k$ entonces $\eta_i \approx N(0, \sigma_k)$.

La densidad de probabilidad condicional

La densidad de probabilidad condicional de la intensidad observada en un pixel respecto de la clase del pixel es de la forma:

$$p(y_i | x_i = \omega_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - (1 + \beta_i) \mu_{x_i}}{\sigma_{x_i}} \right)^2 \right\}. \quad (3.1)$$

Esta densidad de probabilidad modela el efecto del ruido aditivo en la generación de la imagen como la desviación respecto de la intensidad media de la clase modulada por el sesgo β_i debido al campo de iluminación.

La densidad de probabilidad *a priori*

La distribución de probabilidad a priori de las clases de los pixeles viene dada por un Campo Markoviano :

$$p(\mathbf{x}) = \frac{1}{Z} \exp \left\{ -\sum_{c \in C} V_c(\mathbf{x}) \right\}, \quad (3.2)$$

donde Z es la constante de normalización y la suma de los potenciales $V_c(\mathbf{x})$ se extiende a todo el conjunto de cliques C definidos sobre la malla de los sitios de los pixeles. Los potenciales modelan las relaciones espaciales entre las intensidades de los pixeles y permiten introducir restricciones de suavidad u otras restricciones globales. Los Campos Markovianos son una forma elegante de especificar las probabilidades *a priori* de las imágenes a través de las propiedades espaciales deseadas en la imagen. Estas propiedades se expresan mediante los potenciales asociados con los cliques de sitios de pixeles. Potenciales positivos y elevados corresponden con configuraciones indeseadas, potenciales pequeños y negativos modelan configuraciones deseables. Si consideramos sólo los cliques formados por uno y dos pixeles, tenemos que el potencial puede expresarse como:

$$\sum_c V_c(\mathbf{x}) = \sum_i \{V_1(x_i) + V_2(x_i)\}. \quad (3.3)$$

Los potenciales locales se pueden expresar como:

$$V_1(x_i) = \sum_{\omega \in \Omega} -\alpha_{0\omega} \delta(x_i = \omega). \quad (3.4)$$

$$V_2(x_i) = \sum_{i,j \in c_l(i)} [\alpha_l (-\delta(x_i = x_j) + \delta(x_i \neq x_j))]. \quad (3.5)$$

En estas expresiones

$$\delta(c) = \begin{cases} 1 & \text{si } c \\ 0 & \text{sino} \end{cases}. \quad (3.6)$$

Se denota por $c_l(i)$ los lugares de los píxeles en un vecindario de radio l en torno al pixel i de referencia. Un vecindario de radio l en torno al pixel i puede expresarse como $c_l(i) = \{j : j \in I, |i - j| < \sqrt{l}, i \neq j\}$. El potencial $V_1(x_i)$ corresponde a las probabilidades *a priori* de las clases de píxeles. El potencial $V_2(x_i)$ especifica una restricción de suavidad en la clasificación, penalizando las clasificaciones de píxeles en el vecindario distintas de la clasificación del pixel central, y premiando las que son iguales.

En [104] los vecindarios considerados son de radio 1 y 2, por motivos computacionales y, por tanto, los parámetros de la densidad de probabilidad *a priori* a estimar son:

$$\alpha = (\alpha_{\omega}, \omega \in \Omega; \alpha_1, \alpha_2).$$

Esto es, necesita estimar las probabilidades *a priori* α_{ω} de las clases de píxeles y los pesos en los vecindarios de orden 1 y 2 de las coincidencias en la clase del pixel.

Estimación MAP de la segmentación de la imagen

La segmentación de la imagen viene dada por una clasificación de los píxeles. El objetivo del BMAP es la estimación de MAP de la segmentación de la imagen dada una observación:

$$\mathbf{x}^* = \arg \max_{\mathbf{x}} p(\mathbf{x} | \mathbf{y}).$$

Según el teorema de Bayes, y habida cuenta la independencia entre la imagen observada y la clasificación, la estimación MAP de la clasificación de los píxeles puede escribirse como:

$$\mathbf{x}^* = \arg \max_{\mathbf{x}} p(\mathbf{y} | \mathbf{x}, \beta^*, \theta^*) p(\mathbf{x} | \alpha^*).$$

En esta ecuación se hacen explícitas las dependencias de los parámetros de los modelos de las densidades de probabilidad condicionales y *a priori*. Puesto que la probabilidad *a posteriori* puede escribirse en términos de la función de energía $u(\mathbf{x}|\mathbf{y})$

$$p(\mathbf{x}|\mathbf{y}) \propto \exp(-u(\mathbf{x}|\mathbf{y}))$$

El problema de la maximización de la probabilidad a posteriori se convierte en un problema de minimización de dicha energía.

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} u(\mathbf{y}|\mathbf{x}, \beta^*, \theta^*, \alpha^*).$$

La minimización de la energía se puede realizar de diversas maneras, aplicando métodos de búsqueda de mínimos globales como el enfriamiento simulado [1], [74] o métodos de búsqueda local. En [104] se realiza mediante la aplicación de ICM, que es un algoritmo voraz (*greedy*) de búsqueda local. ICM no garantiza el óptimo global pero sí soluciones subóptimas de calidad obtenidas en un tiempo de cálculo razonable. ICM asume que los píxeles son independientes, esto es:

$$p(\mathbf{x}|\mathbf{y}) = \prod_i p(x_i|y_i),$$

de donde la función de energía puede expresarse como la suma de las energías individuales en cada pixel

$$u(\mathbf{x}|\mathbf{y}) = \sum_i u(x_i|y_i).$$

En términos de las expresiones de las probabilidades a priori y condicionales, la energía local de cada pixel puede escribirse:

$$u(x_i|y_i) = -\frac{1}{2} \left(\frac{y_i - (1 + \beta_i) \mu_{x_i}}{\sigma_{x_i}} \right)^2 + \log(\sigma_{x_i}) + V_1(x_i) + V_2(x_i)$$

El ICM parte de una estimación inicial $\mathbf{x}(0)$ de la clasificación e itera la siguiente secuencia:

1. Para cada pixel i
 1. Evalua la función de energía para todos posibles estados de la clasificación del pixel en función del valor observado y las clasificaciones previas de los pixeles vecinos

$$\{u(x_i|y_i, x_j(t); j \neq i); x_i \in \Omega\}.$$

2. Selecciona el estado de mínima energía como estado del pixel en la siguiente iteración:

$$x_i(t+1) = \arg \min \{u(x_i | y_i, x_j(t); j \neq i); x_i \in \Omega\}$$

3. Se contabiliza un "flip" cada vez que varía el estado de un pixel

$$flips = |\{x_i(t+1) \neq x_i(t)\}|.$$

2. Si el número de flips es cero o menor que un valor de umbral de convergencia, se detiene el proceso.

El proceso de cálculo de ICM es un proceso síncrono, en el que todos los píxeles cambian simultáneamente de estado basándose en la información del estado de sus vecinos en el instante anterior de tiempo. Esta asunción es la que lo convierte en un algoritmo local.

Estimación de los parámetros de la d.d.p. condicional

El proceso completo del BMAP incluye la estimación inicial de los parámetros de la d.d.p. *a priori* y la reestimación en cada iteración de los parámetros de la d.d.p. condicional. Esto es, dada una clasificación MAP de los pixels \mathbf{x}_k en base a los valores $\{(\beta, \theta)_k, \alpha\}$ de los parámetros, se reestiman $(\beta, \theta)_{k+1}$. El proceso converge hacia un punto fijo en el que los parámetros no varían. Por otra parte, los parámetros α de la d.d.p. *a priori* se estima a partir de una clasificación inicial \mathbf{x}_0 obtenida por otros medios, como el algoritmo de las k-medias.

La estimación de los parámetros condicionales se hace maximizando la verosimilitud de la muestra

$$\left(\hat{\beta}, \hat{\theta}\right) = \arg \max_{(\beta, \theta)} p(\mathbf{y} | \mathbf{x}, \beta, \theta), \quad (3.7)$$

o equivalentemente, la log-verosimilitud de la muestra la cual toma la siguiente forma, supuesto que los pixeles son estadísticamente independientes:

$$\log p(\mathbf{y} | \hat{\mathbf{x}}) = \sum_{i \in I} \left\{ -\frac{1}{2} \left(\frac{y_i - (1 + \beta_i) \mu_{\hat{x}_i}}{\sigma_{\hat{x}_i}} \right)^2 + \log(\sigma_{\hat{x}_i}) \right\}, \quad (3.8)$$

donde $\hat{\mathbf{x}}$ es la clasificación de los pixeles en un paso de la iteración. Calculando la derivada respecto de los distintos parámetros, igualando a cero y despejando

obtenemos las siguientes relaciones:

$$\hat{\mu}_\omega = \frac{\sum_{i \in R_\omega} y_i (1 + \beta_i)}{\sum_{i \in R_\omega} (1 + \beta_i)^2}, \quad (3.9)$$

$$\hat{\sigma}_\omega^2 = \frac{\sum_{i \in R_\omega} (y_i - (1 + \beta_i) \mu_\omega)^2}{|R_\omega|}, \quad (3.10)$$

$$\hat{\beta}_i = \frac{y_i}{\mu_{x_i}} - 1. \quad (3.11)$$

Donde R_ω denota el conjunto de píxeles (regiones de la imagen) que son clasificados como de clase ω . En estas expresiones los nuevos valores se denotan por $(\hat{\mu}_\omega, \hat{\sigma}_\omega^2, \hat{\beta}_i)$ y los antiguos por $(\mu_\omega, \sigma_\omega^2, \beta_i)$.

Estimación de los potenciales *a priori*

La estimación de los parámetros de la probabilidad *a priori* se obtendría maximizando la densidad *a posteriori* de los parámetros para una clasificación dada:

$$\hat{\alpha} = \arg \max_{\alpha} p(\alpha | \mathbf{x}). \quad (3.12)$$

Sin embargo en [104] se sigue un proceso un poco más directo. Los potenciales de los cliques unitarios están relacionados con las probabilidades *a priori* de las clases, por lo que una expresión adecuada para ellos, dada una clasificación inicial de los píxeles, es:

$$\alpha_{0\omega} = \log \frac{|R_\omega|}{\sum_{\omega \in \Omega} |R_\omega|}. \quad (3.13)$$

Para la estimación de los potenciales de los cliques de segundo orden se considera la probabilidad conjunta de un pixel y su vecindario. Sea ω la clase del pixel y t la configuración de los píxeles vecinos. La probabilidad conjunta será

$$p(\omega, t) = \frac{1}{Z'} \exp \{ \alpha_{0\omega} + 2\alpha_1 (f_\omega - 3) + 2\alpha_2 (s_\omega - 6) \}, \quad (3.14)$$

donde f_ω y s_ω representan el número de píxeles de clase ω en vecindarios N_1 y $N_2 - N_1$, respectivamente. (N_m denota el vecindario de radio m). El log-ratio de verosimilitud de dos clases distintas con el mismo vecindario es como sigue

$$\log \frac{p(\omega, t)}{p(\omega', t)} = \alpha_{0\omega} - \alpha_{0\omega'} + 2\alpha_1 (f_\omega - f_{\omega'}) + 2\alpha_2 (s_\omega - s_{\omega'}), \quad (3.15)$$

donde podemos estimar el ratio

$$\frac{\widehat{p}(\omega, t)}{\widehat{p}(\omega', t)} = \frac{n_{\mathbf{x}}(\omega, t)}{n_{\mathbf{x}}(\omega', t)}, \quad (3.16)$$

donde $n_{\mathbf{x}}(\omega, t)$ es el numero de veces que se encuentra esta configuración en la clasificación de la imagen. Las estimaciones de los potenciales (α_1, α_2) se obtiene resolviendo por minimos cuadrados el conjunto de ecuaciones que se plantean al considerar todos los pares $n_{\mathbf{x}}(\omega, t) > 0$, $n_{\mathbf{x}}(\omega', t) > 0$.

Inicialización

Se puede utilizar el algoritmo de las k-medias para estimar valores iniciales de las medias y varianzas en la distribución condicional y asumir un valor cero para el sesgo debido al campo de iluminación y los parámetros de la distribución *a priori*. La energía total inicial es

$$u^0(\mathbf{x}|\mathbf{y}) = \sum_{\omega} \sum_{i=R} (y_i - \mu_{x_i})^2. \quad (3.17)$$

El algoritmo completo

El algoritmo BMAP en resumen consta de los siguientes pasos

1. Estimación de una clasificación inicial $\widehat{\mathbf{x}}(0)$ basada en los representantes de intensidad obtenidas mediante el algoritmo k-medias u otro similar
2. Estimación de los parámetros de las densidades *a priori* y condicional: $(\widehat{\alpha}, \widehat{\theta}(0))$. El sesgo inicial es nulo $\widehat{\beta}(0) = 0$.
3. Repetir hasta que convergen los estimadores
 1. Estimación MAP de $\widehat{\mathbf{x}}(t+1)$ mediante ICM a partir de $(\widehat{\alpha}, \widehat{\theta}(t), \widehat{\beta}(t), \widehat{\mathbf{x}}(t))$.
 2. Estimación del sesgo $\widehat{\beta}(t+1)$ a partir de $\widehat{\mathbf{x}}(t+1)$. Suavización del sesgo estimado aplicando un filtro lineal.
 3. Estimación de los parámetros de la densidad condicional $\widehat{\theta}(t+1)$ a partir de $\widehat{\mathbf{x}}(t+1)$, y $\widehat{\beta}(t+1)$ suavizado.

Discusión

Hemos implementado el algoritmo siendo fieles a la descripción dada en [104] y recojida en las secciones anteriores. Los resultados experimentales se obtuvieron aplicando los ajustes indicados en el artículo [104] en lo referente al tamaño del vecindario y del filtro de suavizado aplicado al bias.

Los potenciales expresan el conocimiento previo en dos aspectos:

- Distribución a priori de las clases: debería ser especificada a partir de conocimiento previo sobre los objetos físicos. Por ejemplo, en el caso de las imágenes de cerebro, la relación entre la materia gris, materia blanca y líquido cerebro-espinal es bien conocida y puede introducirse.
- Las restricciones de suavidad: los potenciales de los cliques de orden 2 especifican la probabilidad de que pixeles vecinos sean de la misma clase, lo que equivale a especificar los tamaños de regiones con intensidad constante.

La estimación de estos potenciales a partir de la segmentación inicial que se obtiene mediante la clasificación por el vecino más cercano usando los representantes obtenidos mediante el k-medias u otro algoritmo similar resulta muy poco fiable, a menos que los autores consideren que esta clasificación no va a ser modificada sustancialmente en la aplicación del resto del algoritmo. Otros algoritmo bayesianos, por ejemplo [98], fijan los potenciales de los cliques sin utilizar información directa de la imagen. El algoritmo que se presenta en [98] es un algoritmo de segmentación basado también en el modelado bayesiano de la imagen usando vecindarios de tamaño decreciente, siendo el vecindario inicial el que incluye toda la imagen.

Además, es bien conocido el carácter local del k-medias como algoritmo de minimización, por lo que no se garantiza que las distribuciones de las clases correspondan a la realidad física del problema, ni tan siquiera correspondan a la segmentación óptima en términos de distorsión. Cuando el número de clases aumenta, es de esperar que las regiones detectadas por el k-medias sean bastante erráticas y las estimaciones de los potenciales de segundo orden pueden ser inconsistentes. Por último, el papel que juega el sesgo puede hacer que la clasificación varíe fuertemente. Al no tener en cuenta este factor en la estimación de los potenciales se puede estar incurriendo en un error considerable.

3.3.2.BFCM

Hemos seleccionado como representante de las técnicas basadas en agrupamiento borroso la presentada en [3] debido a que es una de las pocas referencias que considera explícitamente la inhomogeneidad de la iluminación. El modelo de la imagen que considera este trabajo es logarítmico:

$$y_k = x_k + \beta_k; \forall k \in \{1, \dots, N\}, \quad (3.18)$$

donde k denota la localización del pixel, y_k, x_k son, respectivamente, los logaritmos de la intensidad observada y la real; β_k es el sesgo debido al campo de iluminación en el pixel k .

La función objetivo

La función objetivo estandar para el algoritmo borroso de C-medias (FCM: Fuzzy C Means [10], [12]) para la partición de una muestra $\{x_k\}_{k=1}^N$ en c agrupamientos viene dada por:

$$J = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|x_k - v_i\|^2, \quad (3.19)$$

donde $\{v_i\}_{i=1}^c$ son los prototipos de los agrupamientos y la matriz $[u_{ik}] = U$ representa la matriz de partición, $U \in \mathcal{U}$, definidas por las siguientes restricciones

$$\mathcal{U} = \left\{ u_{ik} \in [0, 1] \left| \sum_{i=1}^c u_{ik} = 1, \forall k \text{ and } 0 < \sum_{k=1}^N u_{ik} < N, \forall i \right. \right\}. \quad (3.20)$$

Los coeficientes de pertenencia u_{ik} especifican la posibilidad de que un elemento de la muestra pertenezca a una clase. Si sólo uno de ellos es 1 para un elemento de la muestra, la partición es dura (*crisp*), sino es borrosa (*fuzzy*). En este último caso, los coeficientes de pertenencia se *parecen* a una distribución de probabilidad, cuando los tomamos para cada elemento de la muestra. El algoritmo FCM se propuso inicialmente en [29], y la demostración de su convergencia apareció en [9]. Basicamente, la función objetivo 3.19 se minimiza cuando los coeficientes de pertenencia son altos para los elementos de la muestra cercanos al representante de su clase y son bajos para los que están lejos respecto de los representantes.

En [3] se propone una modificación de la función objetivo 3.19 que tiene en cuenta los sesgos debidos a inhomogeneidades de iluminación. Para ello se consideran los vecindarios de los pixeles:

$$J_m = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|x_k - v_i\|^2 + \frac{\alpha}{N_R} \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \left(\sum_{x_r \in \mathcal{N}_k} \|x_r - v_i\|^2 \right). \quad (3.21)$$

En esta expresión \mathcal{N}_k denota el vecindario del pixel k de cardinalidad N_R . El efecto del vecindario en la función objetivo se controla por medio del parámetro α . La sustitución del modelo de la imagen en la ecuación nos da la expresión en términos de la imagen observada y el sesgo de iluminación:

$$J_m = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|y_k - \beta_k - v_i\|^2 + \frac{\alpha}{N_R} \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \left(\sum_{x_r \in \mathcal{N}_k} \|y_r - \beta_r - v_i\|^2 \right). \quad (3.22)$$

Formalmente, el proceso de agrupamiento es un problema de minimización con restricciones

$$\min_{U, \{v_i\}_{i=1}^c, \{\beta_k\}_{k=1}^N} J_m \text{ sujeto a } U \in \mathcal{U}. \quad (3.23)$$

Estimación de los parámetros

La estimación de los parámetros se realiza por minimización directa de la función 3.22 incorporando las restricción $\sum_{i=1}^c u_{ik} = 1$ mediante un multiplicador de Lagrange:

$$F_m = \sum_{i=1}^c \sum_{k=1}^N \left(u_{ik}^p D_{ik} + \frac{\alpha}{N_R} u_{ik}^p \gamma_i \right) + \lambda \left(1 - \sum_{i=1}^c u_{ik} \right), \quad (3.24)$$

donde $D_{ik} = \|y_k - \beta_k - v_i\|^2$ y $\gamma_i = \sum_{x_r \in \mathcal{N}_k} \|y_r - \beta_r - v_i\|^2$. Para calcular el valor de los parámetros, se iguala a cero la derivada de la función 3.24 respecto del parámetro y se despeja el parámetro buscado.

Para estimar los coeficientes de pertenencia, derivamos respecto de u_{ik} y tenemos (para $p > 1$):

$$\left[\frac{\partial F_m}{\partial u_{ik}} = p u_{ik}^{p-1} D_{ik} + \frac{\alpha p}{N_R} u_{ik}^{p-1} \gamma_i - \lambda \right]_{u_{ik} = u_{ik}^*} = 0. \quad (3.25)$$

Resolviendo para u_{ik}^* obtenemos:

$$u_{ik}^* = \left(\frac{\lambda}{p \left(D_{ik} + \frac{\alpha}{N_R} \gamma_i \right)} \right)^{1/(p-1)}, \quad (3.26)$$

considerando la condición $\sum_{i=1}^c u_{ik}^* = 1$, podemos calcular el valor del multiplicador de Lagrange:

$$\lambda = \frac{p}{\left(\sum_{i=1}^c \left(\frac{1}{D_{ik} + \frac{\alpha}{N_R} \gamma_i} \right)^{1/(p-1)} \right)^{(p-1)}}. \quad (3.27)$$

Sustituyendo, llegamos a una expresión para los coeficientes de pertenencia:

$$u_{ik}^* = \frac{1}{\sum_{j=1}^c \left(\frac{D_{ik} + \frac{\alpha}{N_R} \gamma_i}{D_{jk} + \frac{\alpha}{N_R} \gamma_j} \right)^{1/(p-1)}}. \quad (3.28)$$

Para calcular los prototipos de los agrupamientos, derivamos respecto de ellos:

$$\left[\sum_{k=1}^N u_{ik}^p (y_k - \beta_k - v_i) + \frac{\alpha}{N_R} \sum_{k=1}^N u_{ik}^p \left(\sum_{x_r \in \mathcal{N}_k} (y_r - \beta_r - v_i) \right) \right]_{v_i=v_i^*} = 0. \quad (3.29)$$

Despejando v_i^* obtenemos:

$$v_i^* = \frac{\sum_{k=1}^N u_{ik}^p \left((y_k - \beta_k) + \frac{\alpha}{N_R} \left(\sum_{x_r \in \mathcal{N}_k} (y_r - \beta_r) \right) \right)}{(1 + \alpha) \sum_{k=1}^N u_{ik}^p}. \quad (3.30)$$

Los prototipos resultan ser una especie de promedio de las intensidades de los pixeles corregidos restando el sesgo estimado en el pixel mas la influencia de sus vecinos, también corregidos. Los coeficientes de pertenencia normalizados se convierten en una especie de distribución de probabilidad.

Finalmente, para estimar el sesgo de la iluminación hacemos el cálculo similar a los anteriores consistente en igualar a cero la derivada respecto del sesgo β_k debido al campo de iluminación:

$$\left[\sum_{i=1}^c \frac{\partial}{\partial \beta_k} \sum_{k=1}^N u_{ik}^p \|y_k - \beta_k - v_i\|^2 \right]_{\beta_k=\beta_k^*} = 0. \quad (3.31)$$

Despejando β_k obtenemos:

$$\beta_k^* = y_k - \frac{\sum_{k=1}^N u_{ik}^p v_i}{\sum_{k=1}^N u_{ik}^p}. \quad (3.32)$$

El sesgo resulta ser la diferencia entre el valor de intensidad observado y el predicho a partir de los coeficientes de pertenencia del pixel y las medias actuales de intensidad correspondiente a cada clase en la imagen.

El algoritmo completo

El BFCM completo consiste en los siguientes pasos

1. Selecciona los prototipos iniciales $\{v_i\}_{i=1}^c$. Fija un sesgo inicial $\{\beta_k\}_{k=1}^N$ con valores idénticos para todos los pixeles y muy pequeños.
2. Repite hasta que se produzca la convergencia de los representantes de clases $\|\mathbf{v}_{\text{nuevo}} - \mathbf{v}_{\text{viejo}}\| < \varepsilon$
 1. Estima la matriz de partición
 2. Estima los centros de las clases
 3. Estima el sesgo

Es importante notar que el sesgo del campo de iluminación no se suaviza ni se procesa de ninguna otra manera. Esto implica que la estimación es completamente local a cada pixel

Discusion

El proceso de estimación asume que la función a minimizar es algo tan sencillo como una parábola, con un único mínimo global. Esta asunción no parece razonable, dado que este mismo tipo de algoritmos precisan de procesos de minimización más sofisticados incluso cuando no incluyen la información relativa al vecindario.

La necesidad de filtrar el estimador del sesgo para que sea una función suave no aparece en el artículo [3], la estimación del sesgo es un proceso local a cada pixel y el resultado evidente que encontramos más adelante es que el sesgo calculado es trivialmente una imagen de diferencias de los pixeles respecto de las

intensidades de referencia de las clases. Este “pequeño defecto” hace que los resultados presentados por los autores en su artículo sean poco creíbles.

No se dan indicaciones sobre el parámetro p del algoritmo, el exponente que controla el grado de “borrosidad” del algoritmo. En la descripción de los experimentos se especifica el valor trivial $p = 2$ que reduce muchas de las características del algoritmo al caso del agrupamiento basado en la distancia euclídea.

3.4. Métodos paramétricos

Los métodos paramétricos imponen un modelo del bias cuyos parámetros son estimados a partir de la imagen. Los dos algoritmos incluidos en esta sección introducen algún modelo espacial del campo de iluminación, pero varían sustancialmente en la forma de resolver el problema.

Uno de los primeros trabajos que manejan un modelo de la inhomogeneidad de la iluminación es el de [93] que sigue un proceso iterativo de identificación de regiones constantes en la imagen mediante el ajuste polinomial a las diversas regiones segmentadas en la imagen. El modelo polinómico permite modelar las inhomogeneidades locales de los tejidos.

En las descripciones de los algoritmos en las siguientes subsecciones, se asume que el histograma de la imagen original limpia está formado por impulsos correspondientes a las intensidades nominales medias de las clases de tejidos, similar al de la figura 2.1. En [81] se propone un método basado en la observación de que la degradación de la imagen producida por un campo de iluminación inhomogéneo aumenta la cantidad de información en la imagen. Propone un método que estima el campo de iluminación en un proceso que trata de minimizar la información en la imagen corregida, sujeto a la restricción de preservar las estadísticas globales de la imagen. El campo de iluminación se modela como una combinación lineal de polinomios modulados para que preserven las estadísticas globales de la imagen. En [78], [79] la imagen se modela como un proceso estocástico y la estimación MAP de la clasificación y el campo de iluminación se realiza mediante un algoritmo EM. El campo de iluminación se asume como una función suave que es combinación lineal de una base de funciones, en este caso también se trata de funciones polinomiales.

En [111], [112] se propone un algoritmo basado en la idea de que el efecto del campo de iluminación es una convolución de los histogramas (densidades de probabilidad en su nomenclatura) de la imagen y del campo de iluminación. Esto es, el campo actúa como una suavización del histograma de la imagen

“limpia”. La estrategia de corrección consiste en tratar de maximizar el contenido frecuencial de la imagen corregida, buscando un campo suave y que varíe lentamente. Para hacer la búsqueda tratable, se restringe a campos con distribución gaussiana de media cero. Además, dado que una función gaussiana se puede expresar como convolución de funciones gaussianas de radio menor, el proceso de búsqueda se convierte en un proceso iterativo de suavización. Aunque no está descrito explícitamente, el campo de iluminación está parametrizado a través de su distribución de probabilidad. Una de las características de este algoritmo es que no asume ningún modelo previo sobre la clasificación de los píxeles de las imágenes en clases de tejidos.

El primer algoritmo que presentamos en detalle en esta sección es el algoritmo de Wells, el cual modela el campo de iluminación como un vector aleatorio de tamaño el de la imagen con distribución de probabilidad gaussiana multivaluada. El problema que trata de resolver es el de la estimación MAP del campo de iluminación de la imagen, conocidos los valores medios de las clases de intensidad correspondientes a los tejidos en la imagen. La clasificación de los píxeles es desconocida y para resolver simultáneamente ambos problemas se realiza un proceso de maximización de la esperanza (EM). Este algoritmo es local y tiene buenas propiedades de convergencia. Sin embargo, la complejidad computacional de tratar con la matriz de covarianza del campo de iluminación obliga a realizar aproximaciones, lo que hace que el método final sea más parecido a un método no paramétrico. La estimación del sesgo debido al campo de iluminación se convierte en la práctica en una suavización de la clasificación MAP de los píxeles.

El segundo algoritmo propone un modelo del campo de iluminación construido como combinación lineal de productos de polinomios de Legendre. El problema de la corrección de la iluminación se convierte en la estimación de los parámetros de la combinación lineal. Para ello los autores proponen un algoritmo aleatorio que se enmarca en la familia de algoritmos conocidos como Estrategias Evolutivas.

3.4.1. Algoritmo de Wells

En [120] se propone un método de restauración de la imagen basado en el modelado bayesiano de la imagen, que ha sido mejorado en [56] teniendo en cuenta la clase que modela todos los tejidos no interesantes. El método se aplica a

imagenes multimodales. Se considera la imagen logaritmo

$$\begin{aligned} Y_i &= g(X_i) \\ &= [\ln([X_i]_1), \ln([X_i]_2), \dots, \ln([X_i]_m)]^T \end{aligned} \quad (3.33)$$

donde X_i es la señal de resonancia magnética nuclear observada en el voxel i -ésimo y m es la dimensión de la señal. La probabilidad condicionada de la imagen sigue el habitual modelo gaussiano centrado en la intensidad media de la clase asociada al pixel, incluyendo el sesgo debido a la inhomogeneidad de la iluminación:

$$p(Y_i | \Gamma_i, \beta_i) = G_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i), \quad (3.34)$$

donde

$$G_{\psi_{\Gamma_i}}(x) = (2\pi)^{-m/2} |\psi_{\Gamma_i}|^{-1/2} \exp\left(-\frac{1}{2}x^T \psi_{\Gamma_i}^{-1}x\right) \quad (3.35)$$

es la distribución gaussiana m -dimensional con matriz de varianza ψ_{Γ_i} , Y_i es el logaritmo de la intensidad (vector $m \times 1$) en el voxel i , Γ_i es la clase de tejido en el voxel i , $\mu(x)$ es (vector $m \times 1$) la media de la intensidad de los voxels de clase x , ψ_x es la matriz $m \times m$ de covarianza de las intensidades de los voxels de clase x , β_i es el sesgo (vector $m \times 1$) producido por el campo de iluminación en el voxel i . Observese que el ruido gaussiano aditivo está incorporado en este modelo de la densidad de probabilidad condicional.

La distribución de probabilidad *a priori* de las imagenes se denota

$$p(\Gamma_i). \quad (3.36)$$

Si esta probabilidad es uniforme sobre el conjunto de clases de tejidos, el método proporciona una estimación de máxima verosimilitud de la clasificación del tejido. Una distribución *a priori* dependiente de la posición espacial del voxel como la que se describe, por ejemplo, en [67] precisa de información anatómica detallada. El sesgo de iluminación se denota $\beta = (\beta_0, \beta_1, \dots, \beta_{n-1})^T$, donde n es el número de voxels de los datos. El campo de iluminación también se modela como una variable aleatoria multidimensional con distribución gaussiana:

$$p(\beta) = G_{\psi_\beta}(\beta), \quad (3.37)$$

donde ψ_β es la matriz $n \times n$ de covarianza de todo el campo, para poder formular estimadores tratables esta matriz se considera que es una matriz por bandas.

Se asume que las clases de los tejidos y el campo de iluminación son independientes estadísticamente. Usando la definición de probabilidad condicional, se puede obtener la expresión para la probabilidad de la imagen y la clase condicionada al campo de iluminación:

$$p(Y_i, \Gamma_i | \beta_i) = p(Y_i | \Gamma_i, \beta_i) p(\Gamma_i) \quad (3.38)$$

y la probabilidad condicional de la intensidad, calculando el marginal sobre las clases de los tejidos

$$p(Y_i | \beta_i) = \sum_{\Gamma_i} p(Y_i, \Gamma_i | \beta_i) = \sum_{\Gamma_i} p(Y_i | \Gamma_i, \beta_i) p(\Gamma_i). \quad (3.39)$$

Asumiendo la independencia estadística de los voxels, se puede escribir la densidad de probabilidad *a priori* de la imagen completa como:

$$p(Y | \beta) = \prod_i p(Y_i | \beta_i). \quad (3.40)$$

Derivación del estimador para el bias

La regla de Bayes permite obtener la probabilidad *a posteriori* del campo de iluminación, dada la intensidad observada:

$$p(\beta | Y) = \frac{p(Y | \beta) p(\beta)}{p(Y)},$$

donde $p(Y)$ es una constante de normalización. La estimación del campo de iluminación se puede realizar como un proceso de estimación MAP:

$$\hat{\beta} = \arg \max_{\beta} p(\beta | Y). \quad (3.41)$$

La condición necesaria para el máximo de la probabilidad a posteriori de β es que su gradiente respecto de β sea nulo:

$$\left[\frac{\partial}{\partial [\beta_i]_k} \ln p(\beta | Y) \right]_{\beta=\hat{\beta}} = 0; \forall i, k \quad (3.42)$$

donde $[\beta_i]_k$ es el k -ésimo componente del bias en el voxel i . Sustituyendo tenemos:

$$\left[\frac{\partial}{\partial [\beta_i]_k} \left(\sum_j \ln p(Y_j | \beta_j) + \ln p(\beta) \right) \right]_{\beta=\hat{\beta}} = 0; \forall i, k. \quad (3.43)$$

Diferenciando los logaritmos y teniendo en cuenta que sólo el i -ésimo término depende de β_i , tenemos:

$$\left[\frac{\sum_{\Gamma_i} p(\Gamma_i) \frac{\partial}{\partial[\beta_i]_k} G_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i)}{\sum_{\Gamma_i} p(\Gamma_i) G_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i)} + \frac{\frac{\partial}{\partial[\beta_i]_k} p(\beta)}{p(\beta)} \right]_{\beta=\hat{\beta}} = 0; \forall i, k. \quad (3.44)$$

Esta expresión se escribe de forma compacta como sigue:

$$\left[\sum_j W_{ij} [\psi_j^{-1}(Y_i - \mu_j - \beta_i)]_k + \frac{\frac{\partial}{\partial[\beta_i]_k} p(\beta)}{p(\beta)} \right]_{\beta=\hat{\beta}} = 0; \forall i, k \quad (3.45)$$

donde los pesos W_{ij} son de la forma:

$$W_{ij} = \frac{\left[p(\Gamma_i) G_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i) \right]_{\Gamma_i=\text{tejido-clase-}j}}{\sum_{\Gamma_i} p(\Gamma_i) G_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i)}. \quad (3.46)$$

Estos pesos W_{ij} tienen un aspecto similar a la probabilidad *a posteriori* de la clase respecto de la intensidad observada en el pixel. Observese en este sentido que el denominador es un término de normalización que hace que la expresión se pueda interpretar como una distribución de probabilidad. La relación 3.45 se puede reescribir en términos de residuales de la intensidad y el bias como

$$\left[[\bar{R}_i]_k - [\psi^{-1}_{ii}\beta_i]_k + \frac{\frac{\partial}{\partial[\beta_i]_k} p(\beta)}{p(\beta)} \right]_{\beta=\hat{\beta}} = 0; \forall i, k, \quad (3.47)$$

donde el residual es de la forma

$$\bar{R}_i \equiv \sum_j W_{ij} \psi_j^{-1}(Y_i - \mu_j),$$

y la covarianza media inversa es de la forma:

$$\overline{\psi^{-1}}_{ik} = \begin{cases} \sum_j W_{ij} \psi_j^{-1} & \text{si } i = k \\ 0 & \text{sino.} \end{cases}$$

Diferenciando el último término de la ecuación 3.47 y escribiendo en notación matricial obtenemos:

$$\bar{R} - \overline{\psi^{-1}}\beta + \psi_{\beta}^{-1}\hat{\beta} = 0. \quad (3.48)$$

En resumen, la condición de gradiente nulo para la estimación MAP del campo de iluminación se puede escribir como:

$$\widehat{\beta} = H\bar{R} \quad (3.49)$$

donde el operador lineal H está definido por:

$$H \equiv \left[\overline{\psi^{-1}} + \psi_{\beta}^{-1} \right]^{-1}. \quad (3.50)$$

El estimador del sesgo se obtiene aplicando un operador lineal H a los residuales promedio de los voxeles, y el operador H está determinado por las matrices de covarianza de las clases y del propio sesgo de iluminación. A pesar de las apariencias, de todas maneras el estimador en la ecuación 3.49 no es lineal en Y , puesto que tanto \bar{R} como H son no lineales en Y .

Método de expectación-maximización (EM)

El proceso de estimación se propone como un proceso de expectación maximización (EM) [25]. En este caso, los datos desconocidos son las clases de tejido asociadas a los voxeles.

La estimación se resume en los dos siguientes pasos:

$$W_{ij} \leftarrow \frac{\left[p(\Gamma_i) G_{\psi_{\Gamma_i}} (Y_i - \mu(\Gamma_i) - \beta_i) \right]_{\Gamma_i = \text{tejido-clase-}j}}{\sum_{\Gamma_i} p(\Gamma_i) G_{\psi_{\Gamma_i}} (Y_i - \mu(\Gamma_i) - \beta_i)} \quad (3.51)$$

$$\widehat{\beta} \leftarrow H\bar{R} \quad (3.52)$$

cuya interpretación es la siguiente:

- En el paso de expectación (paso E) en la ecuación 3.51 se estiman las probabilidades a posteriori de las clases de tejidos en los voxels, a partir de un campo de iluminación previo.
- En el paso de maximización (paso M) en la ecuación 3.52 se calculan los residuales en base a las probabilidades a posteriori y se calcula el estimador MAP del bias.

Es sabido que el algoritmo EM tiene buenas propiedades de convergencia, por lo que el esquema dado por las ecuaciones 3.51 y 3.52 convergerá a un mínimo,

si bien no hay garantía de que sea un mínimo local. El paso M (ec. 3.52) implica el conocimiento y la manipulación de la covarianza del campo de iluminación, una matriz de $n \times n$. En [120] se muestra que H puede ser convenientemente aproximado por un filtro lineal pasabaja. Las condiciones de convergencia del algoritmo EM, aseguran que el proceso no divergerá a pesar de la inexactitud en H .

Discusión

El algoritmo de Wells parte de una fundamentación rigurosa, pero se encuentra con el escollo computacional de la estimación y manipulación de la matriz de covarianza del bias. La solución que se aporta es muy aproximada y descansa en la robusted del algoritmo EM. En realidad lo único que difrencia al algoritmo de Wells del algoritmo de filtrado homomorfo es la fase de clasificación MAP de los pixeles. Desde ese punto de vista, podría considerarse como un algoritmo no paramétrico. Se trata en realidad de una de las primeras intentonas de formulación bayesiana del problema, que luego ha sido continuado por trabajos como el BMAP y los referenciados anteriormente. A pesar de la simplicidad de la formulación final, el algoritmo es muy efectivo, sobre todo en las imagenes reales.

3.4.2.PABIC

Uno de los escasos algoritmos con modelos paramétricos del sesgo debido al campo de iluminación se propone en [116]. En él se asume el modelo de imagen siguiente, que concuerda con el presentado en la sección 2.1.1:

$$s(\mathbf{x}) = (o(\mathbf{x}) + n_{bio}(\mathbf{x})) \otimes h(\mathbf{x}) + b(\mathbf{x}) + n_{MR}(\mathbf{x}) \quad (3.53)$$

donde \mathbf{x} denota el sitio de un elemento de la imagen (pixel o voxel). La imagen original toma el valor $o(\mathbf{x}) = \mu_{k(\mathbf{x})}$ cuando la clase del pixel (voxel) es $k(\mathbf{x})$. Esto es, se asume un conjunto finito de clases de tejidos caracterizados por valores de intensidad constantes. El término $n_{bio}(\mathbf{x})$ denota el efecto de las estructuras internas de los tejidos. La imagen original $o(\mathbf{x})$ tendría un histograma consistuido por impulsos, similar al de la figura 2.1, los efectos de mezcla producidos por los volúmenes parciales en las imagenes se modelan en [116] por la convolución con un kernel de suavizado $h(\mathbf{x})$. Finalmente, los efectos del sistema de captura de la imagen se reflejan en los términos $b(\mathbf{x})$ y $n_{MR}(\mathbf{x})$. El primero

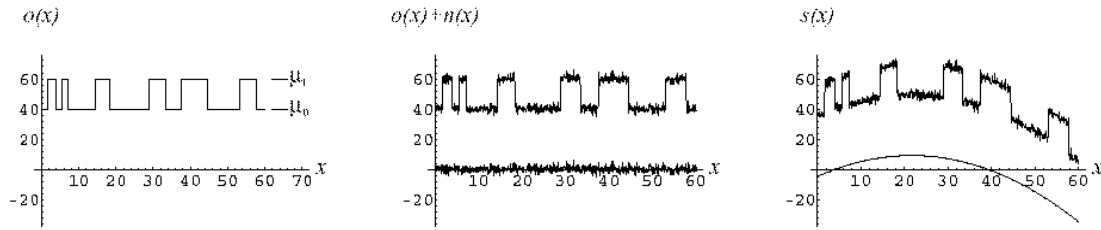


Figura 3.1: Ilustración del efecto del sesgo debido al campo de iluminación y del ruido aditivo en una línea de una imagen, o en una función unidimensional.

corresponde a los posibles sesgos sistemáticos debidos a inhomogeneidades en la señal de radiofrecuencia, el segundo corresponde al ruido térmico y a los errores de redondeo en los procesos computacionales de recuperación de la imagen. En realidad, formalmente en [116] se ignora $h(\mathbf{x})$. Y los términos $n_{MR}(\mathbf{x})$ y $n_{bio}(\mathbf{x})$ se funden en un único término $n(\mathbf{x})$. La figura 3.4.2 muestra el efecto de $b(\mathbf{x})$ y $n(\mathbf{x})$ sobre la imagen original representada por una función 1D que podría corresponder a una línea en la imagen. La imagen original es constante a trozos, alternando entre los dos valores de intensidad que corresponden a las clases. El ruido aditivo perturba esta señal pero mantiene la apariencia de las dos clases separadas. El bias (también aditivo) deforma la señal haciendo que los valores nominales de las clases dejen de tener valor para la clasificación automática.

Hacemos énfasis en el modelo de la imagen presentado en [116] puesto que, aún siendo similar a la ecuación 2.1, el efecto de la iluminación se presenta como un término aditivo, no multiplicativo. Esto es, la ecuación 3.53 podría equipararse a la imagen logaritmo sobre la que se realiza el filtrado homomorfo clásico. La corrección de la iluminación, dada una estimación $\hat{b}(\mathbf{x})$ del sesgo debido al campo de iluminación, se realiza de la siguiente manera:

$$\tilde{s}(\mathbf{x}) = s(\mathbf{x}) - \hat{b}(\mathbf{x}), \quad (3.54)$$

si asumimos que la imagen que procesamos es el logaritmo de la imagen capturada $s(\mathbf{x}) = \log(f(\mathbf{x}))$, con $f(\mathbf{x}) = o(\mathbf{x}) \cdot \exp(b(\mathbf{x}) + n(\mathbf{x}))$ entonces la corrección es equivalente a la siguiente ecuación:

$$s(\mathbf{x}) - \hat{b}(\mathbf{x}) = \log \left(o(\mathbf{x}) + \frac{n(\mathbf{x})}{\exp(\hat{b}(\mathbf{x}))} \right). \quad (3.55)$$

Esta ecuación muestra que, cuando realizamos la corrección del sesgo, existe un efecto sobre el componente de ruido aditivo en la imagen. En general la contribución de este ruido es mayor en las zonas de menor intensidad en la imagen, por lo que en esas zonas se producirá, en general, un descenso de la relación señal ruido (SNR). Por ello es recomendable seleccionar las regiones de baja intensidad, corregir su ruido antes de la corrección del sesgo y/o evitar la corrección de sesgo para esos píxeles.

Modelo paramétrico del campo de iluminación

El modelo del campo de iluminación $b(\mathbf{x})$ propuesto en [116] es una combinación lineal de productos 2D de polinomios de Legendre (vease apéndice 8). En concreto, para 2D

$$b(\mathbf{x}, \mathbf{p}) = \sum_{i=0}^l \sum_{j=0}^{l-i} p_{ij} P_i(x) P_j(y). \quad (3.56)$$

Donde $\mathbf{x} = (x, y)$ y $P_i(\cdot)$ es el ponomio de Legendre de orden i . El número de parámetros que componen \mathbf{p} es $m = (l + 1) \frac{l+2}{2}$. Para 3D el bias se modela por

$$b(\mathbf{x}, \mathbf{p}) = \sum_{i=0}^l \sum_{j=0}^{l-i} \sum_{k=0}^{l-i-j} p_{ijk} P_i(x) P_j(y) P_k(z). \quad (3.57)$$

Donde $\mathbf{x} = (x, y, z)$. el número de parámetros que componen \mathbf{p} es $m = (l + 1) \left(\frac{l+2}{2}\right) \left(\frac{l+3}{3}\right)$. Esto es, podemos construir modelos de alta dimensión mediante productos externos de los polinomios de legendre. Esta consturcción no preserva muchas de las propiedades de los polinomios de Legendre, por lo que las funciones resultantes no pudedn llamarse propiamente polinomios de alta dimensión.

La función de ajuste de la estrategia evolutiva

La estragía evolutiva que asume como espacio de búsqueda los parámetros del modelo del campo de iluminación. Por lo tanto los individuos en las poblaciones nos darán hipótesis sobre el vector de parámetros de la combinación lineal.

El objetivo es la minimización del error de ajuste de la imagen corregida a la hipotesis de que la imagen es constante a trozos, con distintas clases de tejidos

correspondiendo a valores de intensidad distintos, esto es, para una única clase con intensidad media μ tendríamos un error de ajuste

$$e(\mathbf{p}) = \sum_{\mathbf{x}} (\widehat{r}(\mathbf{x}, \mathbf{p}) - \mu)^2, \quad (3.58)$$

donde $\widehat{r}(\mathbf{x}, \mathbf{p}) = f(\mathbf{x})/\widehat{b}(\mathbf{x}, \mathbf{p})$ son los pixeles (voxeles) de la imagen corregida removiendo el bias de iluminación, sustituyendo obtenemos:

$$e(\mathbf{p}) = \sum_{\mathbf{x}} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu \right)^2. \quad (3.59)$$

En [116] se propone la utilización de una función *valley*(d) = $d^2/(d^2 + 3\sigma^2)$ en lugar de la expresión cuadrática en la ecuación 3.59, esto es:

$$e(\mathbf{p}) = \sum_{\mathbf{x}} \text{valley} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu \right). \quad (3.60)$$

Se asegura que esta formulación es más robusta a errores y desviaciones del modelo (outliers). Sin embargo en nuestros experimentos no hemos encontrado una especial sensibilidad del algoritmo a esta función y hemos prescindido de ella.

Cuando tenemos varias clases de tejidos, la ecuación de error de ajuste se generaliza en [116] como sigue:

$$e(\mathbf{p}) = \sum_{\mathbf{x}} \prod_{k=1}^c \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_k \right)^2. \quad (3.61)$$

Donde $\{\mu_k; k = 1, \dots, c\}$ son los valores nominales de intensidad de las clases de tejidos en la imagen. La estructura de la ecuación de error 3.61 está pensada para evitar una clasificación explícita de los pixeles. Si consideramos la ecuación alternativa

$$e(\mathbf{p}) = \sum_{\mathbf{x}} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_{k(\mathbf{x})} \right)^2, \quad (3.62)$$

donde $k(\mathbf{x})$ es la clase a la que se asigna el pixel \mathbf{x} , usualmente la que corresponde a la intensidad nominal más cercana

$$k(\mathbf{x}) = \arg \min_k \left\{ \frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_k \right\}, \quad (3.63)$$

veremos que la diferencia entre las dos formulaciones es que para cada pixel en la ecuación 3.61 tenemos un producto mientras que en 3.62 sólo se considera el error respecto de la clase asignada al pixel. El producto en 3.61 será cero si alguno de los factores es cero: pixeles que se ajustan a alguna de las clases tendrán error muy pequeño, pixeles que no se ajustan bien a ninguna de las clases tendrán su término de error exponenciado al número de clases. La idea detrás de estas formulación pudiera ser considerar funciones de pertenencia suaves o, al menos, no nulas, de forma similar a las funciones de pertenencia de los algoritmos borrosos[11]. En general, el rango de valores de la ecuación 3.62 será mucho menor, varios órdenes de magnitud menor, que el que toma la función en 3.61.

El algoritmo PABIC

Para ajustar la definición del algoritmo a las expresiones originales en [116], consideraremos que la población¹ en la generación t se denota:

$$\mathbf{x}_t = [\mathbf{p}], \quad (3.64)$$

esto es, la población consta de un único individuo que consiste en el vector de parámetros de la combinación lineal de productos 2D de Legendre que forma la estimación actual del campo de iluminación. También se introduce la notación para una muestra de una variable aleatoria normal multivariada como \mathbf{r}_t . Con estas precisiones el algoritmo PABIC se presenta con las siguientes ecuaciones:

$$\begin{aligned} \mathbf{r}_t &\approx N(0, \mathbf{I}), \\ \mathbf{x}_{t+1} &= \mathbf{x}_t + \mathbf{A}_t \mathbf{r}_t, \\ \mathbf{A}_{t+1} &= \begin{cases} \mathbf{A}_t \left(\mathbf{I} + (c_g - 1) \frac{\mathbf{r}_t \mathbf{r}_t'}{\mathbf{r}_t' \mathbf{r}_t} \right) & \text{si } e(\mathbf{x}_{t+1}) < e(\mathbf{x}_{opt}) \\ \mathbf{A}_t \left(\mathbf{I} + (c_s - 1) \frac{\mathbf{r}_t \mathbf{r}_t'}{\mathbf{r}_t' \mathbf{r}_t} \right) & \text{en otro caso} \end{cases} \\ \mathbf{x}_{opt} &= \begin{cases} \mathbf{x}_{t+1} & \text{si } e(\mathbf{x}_{t+1}) < e(\mathbf{x}_{opt}) \\ \mathbf{x}_{opt} & \text{en otro caso} \end{cases} \end{aligned} \quad (3.65)$$

El algoritmo es estrictamente elitista: sólo se acepta una nueva solución si mejora la mejor encontrada hasta el momento \mathbf{x}_{opt} . Es un algoritmo de individuo único, similar al *simulated annealing* [74] pero con importantes diferencias, que discutiremos más adelante. La mecánica del algoritmo es simple: se genera un

¹De aquí en adelante no existe conflicto entre el sitio de un pixel y la población de la estrategia evolutiva. Por ello mantenemos la notación original en la que \mathbf{x} es la población del algoritmo evolutivo.

vector aleatorio según una distribución normal multivariante de varianza unitaria. La solución actual se modifica añadiéndole una perturbación aleatoria que corresponde a una muestra de un vector aleatorio con distribución normal multivariada de matriz de covarianza \mathbf{A}_t . Si la nueva solución mejora la existente, se almacena como tal. La matriz de covarianza se perturba aleatoriamente con tendencia a encogerse si se ha encontrado un óptimo o a expandirse en otro caso. La expansión contribuye a la búsqueda en regiones más amplias y el encogimiento está pensado para concentrar la búsqueda en una región local, en la creencia de que existe una cierta continuidad en el espacio de búsqueda y la optimización local puede llevarnos al óptimo global.

Discusión

El primer aspecto a discusión es la relación entre PABIC y el Simulated Annealing (SA) [1], [74]. Ambos métodos son algoritmos aleatorios de búsqueda en las que se proponen soluciones y son aceptadas. En el caso del SA, los problemas de optimización son discretos, esto es, los espacios de las variables involucradas son discretos. Se define una topología en este espacio discreto que permite generar aleatoriamente vecinos a una solución dada. Se define $N(x)$ como el conjunto de las soluciones vecinas a una dada. Partiendo de una solución aleatoria x_0 , el esquema del algoritmo consiste en la iteración de los siguientes pasos

1. Generar una solución alternativa x_{t+1}^h muestreando aleatoriamente $N(x_t)$.
2. Evaluar la función objetivo sobre la solución alternativa $C(x_{t+1}^h)$ y el incremento de la función objetivo obtenido al pasar de la solución previa a la nueva solución $\Delta C(x_t, x_{t+1}^h)$.
3. Aceptar la solución según una decisión probabilística que garantiza que la distribución de probabilidad de los estados converge hacia una distribución de Boltzmann. La decisión probabilística tiene el aspecto:

$$P_a(x_t \rightarrow x_{t+1}^h) = \exp\left(-\frac{\Delta C(x_t, x_{t+1}^h)}{\beta}\right),$$

donde β es el parámetro de temperatura que decae hasta cero durante el proceso de optimización.

Este esquema es similar al PABIC presentado en 3.65, sin embargo se aprecian las siguientes importantes diferencias:

- (a) El espacio de búsqueda en PABIC no es discreto, por tanto no es necesario definir vecindarios o topologías específicas para el problema. Para generar soluciones alternativas basta con calcular perturbaciones sobre la solución actual.
- (b) En PABIC, la decisión de aceptar una solución como la óptima es independiente del estado actual. Tal como está formulado en [116], el proceso de generación de soluciones es un paseo aleatorio no guiado. No parte de la solución óptima actual x_{opt} para generar la nueva alternativa x_{t+1} sino del estado interno x_t que evoluciona independientemente: x_{opt} sirve tan sólo como una “memoria”. Esta estrategia da pésimos resultados en los experimentos, por la que hemos modificado en nuestras implementaciones: las nuevas soluciones se generan a partir del último óptimo encontrado:

$$\mathbf{x}_{t+1} = \mathbf{x}_{opt} + \mathbf{A}_t \mathbf{r}_t. \quad (3.66)$$

- (c) La convergencia hacia el óptimo está garantizada en el PABIC debido a que (1) la perturbación aleatoria tiene probabilidad no nula de alcanzar cualquier punto del espacio de búsqueda y (2) seleccionamos el óptimo encontrado, ya sea que utilizemos la versión original del algoritmo o nuestra variante 3.66. En el SA, la convergencia teórica está garantizada por la distribución estacionaria a la que converge el algoritmo, que es la distribución de Boltzmann, y los ratios de convergencia han sido establecidos en la literatura [1], [74]. No existe nada similar para el PABIC. De hecho, si bien nuestra variante 3.66 es más rápida en la convergencia hacia un óptimo local, es más parecida a un algoritmo de búsqueda local y más susceptible de caer en óptimos locales.

El PABIC se presenta, por otro lado, como un algoritmo evolutivo, en concreto como una estrategia $(1 + 1) - ES$, esto es, con la particularidad de que la población sólo tiene un individuo. Es bien conocida la sensibilidad de las estrategias evolutivas, y los algoritmos evolutivos en general, al tamaño de la población [5], [6]. Grandes poblaciones permiten una exploración más exhaustiva del espacio, mientras que poblaciones pequeñas tienden a converger a óptimos locales y a quedarse en ellos por falta de variedad en los individuos. Por ello la proclamación de las buenas propiedades de convergencia no coincide con el acervo teórico y experimental sobre las estrategias evolutivas. Otra de las particularidades de las estrategias evolutivas es la varianza de mutación autoadaptable,

que en general se expresa de la siguiente manera:

$$\begin{aligned}\sigma' &= \sigma \cdot \exp(\tau \cdot N(0, 1)), \\ x_{i(t+1)} &= x_{i(t)} + N_i(0, \sigma').\end{aligned}$$

Esto es, en la estrategia evolutiva se muta la varianza de mutación y esa varianza mutada se usa para generar un individuo mutado, al que se asocia. Si ese individuo es seleccionado para la siguiente generación, su varianza de mutación también permanece en la nueva generación. En 3.65 el papel de esta mutación autodaptable lo toma la matriz \mathbf{A}_t y los parámetros c_g y c_s que modelan, respectivamente, la matriz de covarianza de la perturbación, la tasa de crecimiento y decrecimiento de la matriz correspondientes a la expansión y contracción del radio de la búsqueda cuando se producen, respectivamente, fracasos y éxitos al evaluar nuevas soluciones. La matriz siempre se perturba aleatoriamente en uno u otro sentido (expansión o contracción)

3.5. Nuestras proposiciones

La aportación más importante de esta tesis doctoral es la proposición de algoritmos de corrección de la iluminación basados en el modelado del bias como combinación lineal de los productos 2D de polinomios de Legendre (Apendice 2). Para presentarlas en primer lugar vamos a presentar el gradiente de la función de error en la que se basan. A continuación introducimos los dos algoritmos GradPABIC y GradClassLeg que son variantes de la minimización de la función de error y utilizan este gradiente.

3.5.1. El gradiente

En primer lugar consideramos una versión simplificada de la función de error 3.61 propuesta en [116], basada en el modelado del bias como combinación lineal de productos de polinomios de Legendre:

$$e(\mathbf{p}, I) = \sum_{\mathbf{x}} \prod_k \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_k \right)^2, \quad (3.67)$$

donde $I = \{\mu_k; k = 1, \dots, c\}$ es el conjunto de intensidades correspondientes a las clases en la imagen, \mathbf{x} es el sitio de un pixel (voxel) en la imagen. Consideramos en primer lugar el gradiente de esta función 3.67 respecto de los parámetros de

la combinación lineal de productos de polinomios de Legendre, en el caso 2D en las que $\mathbf{x} = (x, y)$:

$$\nabla_{\mathbf{p}} e(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial p_{ij}} e(\mathbf{p}, I) \right\},$$

donde:

$$\frac{\partial}{\partial p_{ij}} e(\mathbf{p}, I) = \sum_{\mathbf{x}} \sum_m \prod_{k \neq m} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_k \right)^2 \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_j \right) \frac{-f(\mathbf{x}) P_i(x) P_j(y)}{\widehat{b}^2(\mathbf{x}, \mathbf{p})}. \quad (3.68)$$

Si las imagenes son 3D en las que $\mathbf{i} = (x, y, z)$:

$$\nabla_{\mathbf{p}} e(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial p_{ijl}} e(\mathbf{p}, I) \right\},$$

donde

$$\frac{\partial}{\partial p_{ijl}} e(\mathbf{p}, I) = \sum_{\mathbf{x}} \sum_m \prod_{k \neq m} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_k \right)^2 \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_j \right) \frac{-f(\mathbf{x}) P_i(x) P_j(y) P_l(z)}{\widehat{b}^2(\mathbf{x}, \mathbf{p})}.$$

Si consideramos el gradiente de la función de error 3.67 respecto de las intensidades medias obtenemos, tanto para imagenes 2D como 3D:

$$\nabla_I e(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial \mu_j} e(\mathbf{p}, I) \right\},$$

donde

$$\frac{\partial}{\partial \mu_j} e(\mathbf{p}, I) = \sum_{\mathbf{x}} -\frac{1}{2} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_j \right) \prod_{k \neq j} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_k \right)^2. \quad (3.69)$$

El gradiente total será la reunión de estos dos gradientes parciales:

$$\nabla e(\mathbf{p}, I) = \{ \nabla_{\mathbf{p}} e(\mathbf{p}, I), \nabla_I e(\mathbf{p}, I) \}.$$

Si consideramos la ecuación alternativa

$$e_c(\mathbf{p}, I) = \sum_{\mathbf{x}} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_{k(\mathbf{x})} \right)^2, \quad (3.70)$$

donde $k(\mathbf{i})$ es la clase a la que se asigna el pixel \mathbf{i} , usualmente la que corresponde a la intensidad nominal más cercana

$$k(\mathbf{x}) = \arg \min_k \left\{ \frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_k \right\}, \quad (3.71)$$

podemos formular también las correspondientes alternativas a los gradiente de esta función 3.70 respecto de los parámetros de la combinación lineal de productos de polinomios de Legendre. En el caso 2D en las que $\mathbf{x} = (x, y)$:

$$\nabla_{\mathbf{p}} e_c(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial p_{ij}} e_c(\mathbf{p}, I) \right\}, \quad (3.72)$$

donde:

$$\frac{\partial}{\partial p_{ij}} e_c(\mathbf{p}, I) = \sum_{\mathbf{i}} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_{k(\mathbf{x})} \right) \frac{-f(\mathbf{x}) P_i(x) P_j(y)}{\widehat{b}^2(\mathbf{x}, \mathbf{p})}. \quad (3.73)$$

Si las imagenes son 3D en las que $\mathbf{x} = (x, y, z)$:

$$\nabla_{\mathbf{p}} e_c(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial p_{ijl}} e_c(\mathbf{p}, I) \right\}, \quad (3.74)$$

donde

$$\frac{\partial}{\partial p_{ijl}} e_c(\mathbf{p}, I) = \sum_{\mathbf{i}} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_{k(\mathbf{i})} \right) \frac{-f(\mathbf{x}) P_i(x) P_j(y) P_l(z)}{\widehat{b}^2(\mathbf{x}, \mathbf{p})}. \quad (3.75)$$

Si consideramos el gradiente de la función de error 3.67 respecto de las intensidades medias obtenemos, tanto para imagenes 2D como 3D:

$$\nabla_I e_c(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial \mu_j} e_c(\mathbf{p}, I) \right\}, \quad (3.76)$$

donde

$$\frac{\partial}{\partial \mu_j} e_c(\mathbf{p}, I) = \sum_{\mathbf{i}} -\frac{1}{2} \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p})} - \mu_{(\mathbf{i})} \right) \delta_{j, k(\mathbf{i})}. \quad (3.77)$$

El gradiente total será nuevamente la reunión de estos dos gradientes parciales:

$$\nabla e_c(\mathbf{p}, I) = \{ \nabla_{\mathbf{p}} e_c(\mathbf{p}, I), \nabla_I e_c(\mathbf{p}, I) \}. \quad (3.78)$$

Esta es una formulación más cercana a las aproximaciones basadas en agrupamiento borroso.

3.5.2. GradPABIC

Este algoritmo cae en la categoría general de los algoritmos meméticos [95], [96]. La presentación de ideas relevantes sobre estrategias evolutivas y algoritmos meméticos se ha recogido en el Apéndice 3.

El algoritmo que proponemos es precisamente un algoritmo memético de individuo único, cuyas mutaciones son generadas como perturbaciones en la dirección del gradiente de la función de ajuste.

$$\begin{aligned} \mathbf{r}_t &\approx N(0, \mathbf{I}), \\ \mathbf{x}_{t+1} &= \mathbf{x}_{opt} + \nabla e(\mathbf{p}, I) \mathbf{r}_t, \\ \mathbf{x}_{opt} &= \begin{cases} \mathbf{x}_{t+1} & \text{si } e(\mathbf{x}_{t+1}) < e(\mathbf{x}_{opt}) \\ \mathbf{x}_{opt} & \text{en otro caso} \end{cases} \end{aligned} \quad (3.79)$$

Este algoritmo puede aplicarse tanto a la estimación del bias como a la estimación de las medias, variando la definición de los individuos (cromosomas en terminología evolutiva) correspondientemente. Si los individuos están constuidos por los parámetros del bias

$$\mathbf{x}_t = [\mathbf{p}], \quad (3.80)$$

donde \mathbf{p} es el vector de coeficientes de la combinación lineal de productos de polinomios de Legendre, entonces el gradiente aplicado es

$$\nabla_{\mathbf{p}} e(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial p_{ij}} e(\mathbf{p}, I) \right\}, \quad (3.81)$$

si los individuos están constituidos por los parámetros del bias y las medias de intensidad de las clases

$$\mathbf{x}_t = \{\mathbf{p}, I\}, \quad (3.82)$$

donde I es el conjunto de intensidades medias de las clases en la imagen, entonces el gradiente aplicado es

$$\nabla e(\mathbf{p}, I) = \{\nabla_{\mathbf{p}} e(\mathbf{p}, I), \nabla_I e(\mathbf{p}, I)\}. \quad (3.83)$$

3.5.3. GradClassLeg

La segunda variante del algoritmo que proponemos es un algoritmo de descenso de gradiente [121] de la función de ajuste.

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \alpha_t \nabla e(\mathbf{p}, I), \quad (3.84)$$

Este algoritmo puede aplicarse tanto a la estimación del campo de iluminación como a la estimación de las medias, variando la definición de las variables de estado correspondientemente. Si los individuos están constuidos por los parámetros del bias

$$\mathbf{x}_t = [\mathbf{p}], \quad (3.85)$$

donde \mathbf{p} es el vector de coeficientes de la combinación lineal de productos de polinomios de Legendre, entonces el gradiente aplicado es

$$\nabla_{\mathbf{p}} e(\mathbf{p}, I) = \left\{ \frac{\partial}{\partial p_{ij}} e(\mathbf{p}, I) \right\}, \quad (3.86)$$

si los individuos están constituidos por los parámetros del bias y las medias de intensidad de las clases

$$\mathbf{x}_t = \{\mathbf{p}, I\}, \quad (3.87)$$

donde I es el conjunto de intensidades medias de las clases en la imagen, entonces el gradiente aplicado es

$$\nabla e(\mathbf{p}, I) = \{\nabla_{\mathbf{p}} e(\mathbf{p}, I), \nabla_I e(\mathbf{p}, I)\}. \quad (3.88)$$

4. RESULTADOS EXPERIMENTALES EN IRM

En este capítulo presentamos los resultados experimentales de los algoritmos de corrección del campo de iluminación en IRM realizados sobre varios conjuntos de imágenes. En la sección 4.1 introducimos el método de trabajo y presentación de resultados adoptado. En la sección 4.2 presentamos los resultados sobre un conjunto de imágenes sintéticas construidas a partir de un tablero de ajedrez. En las secciones 4.3 y 4.4 presentamos los resultados sobre un conjunto de imágenes sintéticas de cerebro, obtenidas de la Brain Web Database [14] y una las extraídas del *International Brain Segmentation Repository (IBSR)* [65].

4.1. Consideraciones metodológicas

La validación de los resultados de los algoritmos de corrección de las imágenes de RM necesita de datos de contraste para poder cuantificar y valorar numéricamente la calidad de los algoritmos. Los resultados visuales pueden ser ilustrativos pero raramente validan apropiadamente un algoritmo. Una aproximación habitual en la validación de algoritmos de inteligencia artificial y de visión por computador es aplicar estos algoritmos a datos sintéticos de los que conocemos los resultados deseables. De esta forma es posible cuantificar con exactitud la bondad de la aproximación a los resultados deseables. El inconveniente de esta metodología es que las simulaciones son simplificaciones de la realidad y rara vez se corresponden exactamente con datos reales. Si se aplican los algoritmos sobre datos reales, solo es posible realizar comparaciones con los resultados proporcionados por un experto humano o tomar medidas indirectas¹ de la bondad del resultado. El inconveniente de los resultados del experto humano es la inconsistencia entre expertos y la imprecisión o errores del resultado. La ventaja es que estamos validando los algoritmos sobre situaciones reales.

¹Algunos autores [81] refieren como medida de evaluación el coeficiente de variación en los tejidos antes y después de la corrección.

En los experimentos computacionales que se refieren en este capítulo hemos utilizado los siguientes conjuntos de imágenes:

- Imágenes corrompidas a partir de la imagen de una imagen de tablero de ajedrez.
- Imágenes de cerebro simuladas obtenidas del sitio web [14] corrompidas con bias de iluminación sintetizados. Debido a las limitaciones computacionales² hemos trabajado sobre cortes específicos 2D del volumen, variando el campo de iluminación. Las imágenes de partida son rodajas de un volumen pesado en T1.
- Imágenes reales obtenidas del repositorio de imágenes de segmentación de cerebro [65]. Estas imágenes incluyen la clasificación manual proporcionada por un experto.

Las imágenes sintéticas se han obtenido añadiendo a las imágenes limpias un bias conocido construido como combinación lineal de productos 2D de polinomios de Legendre. Los campos de iluminación sintéticos utilizados tienen una variación mucho más fuerte que aquellos que se encuentran en los casos reales. Nuestros experimentos sintéticos se pueden considerar como un “caso peor” para los algoritmos de corrección y clasificación como se verá al contrastar los resultados sobre los casos reales del IBSR con los resultados de los casos sintéticos.

Concretamente, en el caso de las imágenes sintéticas, el conocimiento de la imagen original y de los campos de iluminación que los distorsionan nos permiten calcular las siguientes medidas que nos proporcionan una evaluación cuantitativa de los algoritmos.

- La correlación de la imagen restaurada con la imagen original.
- La correlación del campo de iluminación estimado con el campo utilizado para corromper la imagen.

Los resultados de los algoritmos sobre las imágenes sintéticas se presentan tanto en forma cuantitativa como en forma cualitativa. Esta última consiste en mostrar imágenes resultantes de los procesos. La presentación cuantitativa

²La realización de experimentos exhaustivo consume mucho tiempo de cálculo. Algunos autores [116] realizan el submuestreo de la imagen para acelerar el proceso. Nosotros no hemos reducido las imágenes de tamaño y trabajamos sobre la imagen en su resolución original.

consiste en gráficas que muestran las correlaciones de los resultados con las imágenes originales versus la distorsión de la imagen corrupta. Esta distorsión se calcula como la inversa de la correlación de la imagen corrupta con la imagen limpia.

Sobre la imagen original obtenida de [65] calculamos los resultados como los aciertos de la clasificación de los tejidos en cada una de las rodajas del volumen, respecto de la clasificación manual. Presentamos la gráfica de estos aciertos para cada rodaja del volumen. También presentamos algunas imágenes de ejemplo.

4.2. Experimentos realizados sobre imágenes de tablero de ajedrez

El primer experimento computacional se ha realizado sobre un conjunto de imágenes obtenidas a partir de una imagen de tablero de ajedrez (figura 4.1) mediante la adición de un campo de iluminación generado de forma artificial. Un ejemplo de campo de iluminación se presenta en la figura 4.2, mientras que un ejemplo de imagen corrupta por el campo de iluminación se presenta en la figura 4.3. El objetivo del experimento es evaluar las calidades de los distintos algoritmos sobre una imagen muy simple. Los experimentos sobre imágenes sintéticas sencillas son práctica común para mostrar las propiedades de este tipo de algoritmos.

Para probar la variabilidad de los resultados de los algoritmos aleatorios (las Estrategias Evolutivas, PABIC y GradPABIC) hemos repetido las ejecuciones 30 veces para cada imagen y los resultados proporcionados corresponden a la media y varianza de los resultados de los algoritmos sobre las 30 repeticiones. La media nos da la información de precisión y la varianza la información de fiabilidad del algoritmo. Grandes varianzas corresponden a resultados con poca posibilidad de ser reproducidos en repeticiones distintas de los experimentos.

Los resultados cuantitativos se muestran como gráficas cuyo eje de ordenadas es la inversa de la correlación de la imagen corrupta con la imagen original limpia. De esta manera, las imágenes experimentales se han ordenado por distorsión creciente para mostrar la dependencia, si existe, de los resultados respecto de esta distorsión. El eje de abscisas corresponde con la correlación entre la imagen corregida y la imagen limpia, o bien entre el campo de iluminación estimado y el campo sintético utilizado para corromper la imagen.

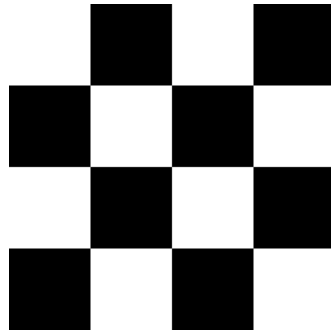


Figura 4.1: Imagen de tablero de ajedrez original.



Figura 4.2: Ejemplo de campo de iluminación empleado para corromper la imagen de tablero de ajedrez.

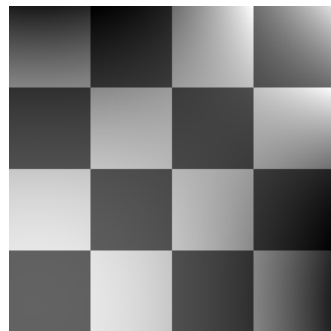


Figura 4.3: Ejemplo de imagen de ajedrez corrompida por un campo de iluminación.

Parametros de los algoritmos

Algunos de los experimentos referidos a continuación son experimentos de análisis de sensibilidad de los algoritmos en los que observamos la variación de la respuesta del algoritmo a variaciones en alguno de sus parámetros. En los restantes experimentos hemos utilizado los siguientes valores de los parámetros:

- PABIC: los dados en la referencia [116], 9000 iteraciones en cada repetición del proceso de optimización para cada imagen
- Estrategias Evolutivas: población 100 individuos, 90 generaciones,
- GradPABIC: 900 iteraciones.
- Filtrado homomorfo: filtrado con una máscara de convolución de suavizado de tamaño 30×30 .
- BMAP: el radio del vecindario es 2 y la máscara de convolución es de tamaño 13×13 como se sugiere en la referencia original.
- GradClassLeg: $\alpha = 0.01$, numero de pasos 900.

4.2.1. Sobre el filtrado homomorfo

No se realizan en general experimentos computacionales aplicando el filtrado homomorfo. La razón es simple: todos los algoritmos propuestos mejoran al filtrado homomorfo sistemáticamente y la comparación del filtrado homomorfo y otros algoritmos no es un tema de interés en esta tesis. Un resultado típico es el que aparece en las gráficas de las figuras 4.4a y 4.9a. En ellas el filtrado homomorfo es mejorado sistemáticamente por todos los demás algoritmos probados. Si bien este es un algoritmo de comparación básico, no lo hemos considerado en adelante dado que, en general es más difícil de ajustar que los restantes algoritmos y produce peores resultados.

4.2.2. Comparación entre algoritmos evolutivos

Tanto las Estrategias Evolutivas, como PABIC, como GradPABIC son algoritmos evolutivos y, por tanto, relativos competidores cercanos. En primer lugar consideramos la comparación de rendimientos entre el algoritmo PABIC y las

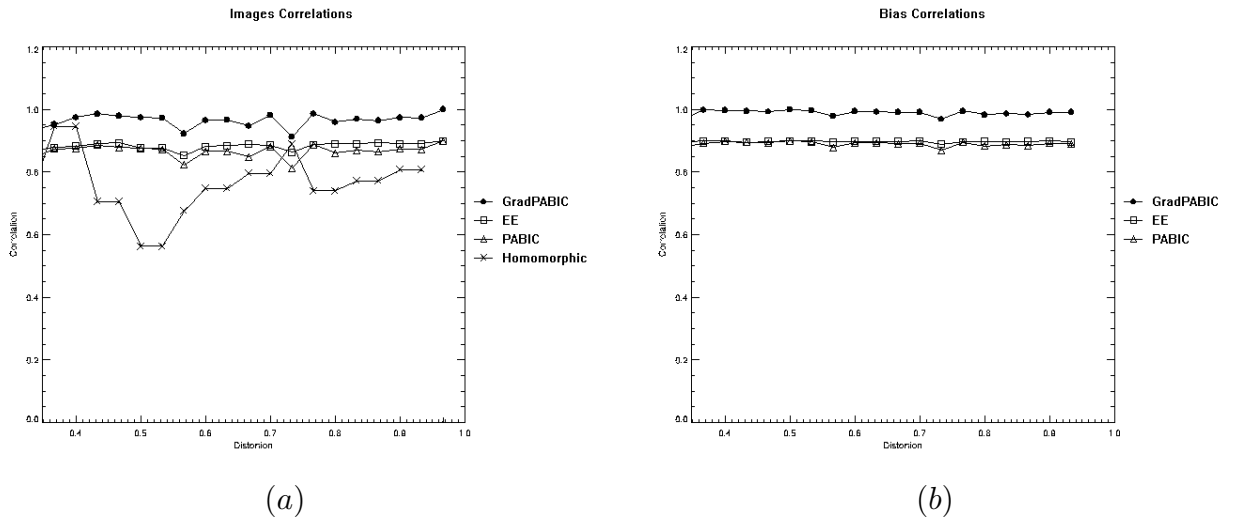


Figura 4.4: Correlación con las originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos: filtrado homomórfico, GradPABIC, EE y PABIC

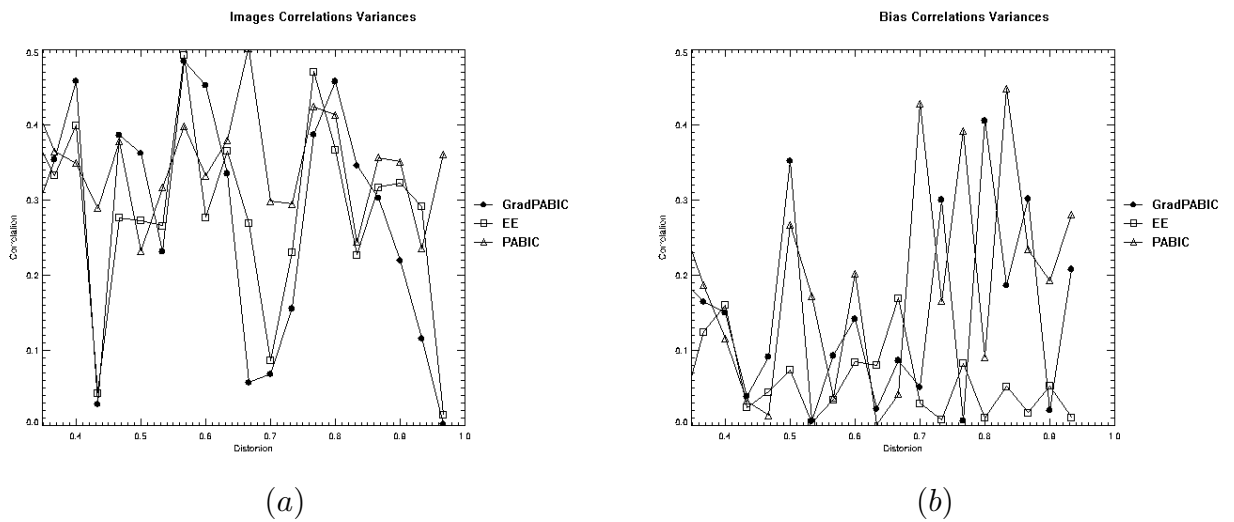


Figura 4.5: Varianza de las correlaciones con los originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos GradPABIC, EE y PABIC

Estrategias Evolutivas diseñadas específicamente para el problema de la corrección de iluminación. Las comparamos a continuación con nuestras proposiciones GradPABIC y GradClassLeg.

La figura 4.4 muestra, para las imágenes corruptas ordenadas por distorsión creciente, la gráfica de la correlación de la imagen recuperada tras la corrección respecto a la imagen original (a) y del sesgo del campo de iluminación estimado respecto de los originales utilizados para corromper la imagen original (b). Comparando los resultados de la bondad de las imágenes corregidas con la bondad de los sesgos de los campos de iluminación estimados, se observa que estas últimas son mucho más cercanas a 1. Esto se debe a la suavidad de las imágenes de bias, lo que hace que la correlación sea en general más alta.

La figura 4.5 muestra las varianzas de los resultados de los algoritmos, para las imágenes corruptas ordenadas por distorsión creciente, en la recuperación de la imagen original (a) y del sesgo del campo de iluminación (b).

Estrategias Evolutivas versus PABIC

Una de las razones de este experimento es la confrontación del algoritmo PABIC contra las Estrategias Evolutivas (EE) convencionales: PABIC es un algoritmo evolutivo de población de individuo único, los autores [116] lo definen como una estrategia evolutiva (1+1), similar al algoritmo de enfriamiento estadístico en la medida en que en cada iteración se genera una solución potencial que es aceptada si mejora las soluciones previas, sin que se haya formulado una estructura de vecindarios. Sin embargo en el PABIC no se aplica un test de aceptación que garantice una distribución de probabilidad estacionaria con propiedades específicas, como es la distribución de Boltzman para el enfriamiento estadístico. Puesto que PABIC cae en el ámbito de las EE, debe esperarse que presente las sensibilidades propias de estos algoritmos. En concreto la sensibilidad al tamaño de la población. Usualmente, mayor población implica una mejor exploración del espacio de búsqueda. Es de esperar, por tanto, que los resultados de la implementación de una versión de las Estrategias Evolutivas dedicada al problema (con poblaciones más grandes) puedan mejorar los resultados del PABIC a un costo computacional equivalente. Para poder realizar una comparación no sesgada por la cantidad de cálculo asignado a cada algoritmo, hemos ejecutado los algoritmos permitiendo que realicen un número similar, sino exactamente idéntico, de evaluaciones de la función fitness. La evaluación de la función fitness se asume como la unidad computacional para la cuantificación de la carga com-

putacional. La EE aplicada a este problema consiste en asociar a cada individuo un vector de coeficientes de la combinación lineal de polinomios de Legendre y se precisa en la sección 9.1.1 del Apéndice 1.

En los resultados mostrados en las gráficas 4.4a y 4.4b se aprecia que las EE genéricas proporcionan un resultado sistemáticamente mejor al PABIC; esto es, EE mejora a PABIC excepto en alguna instancia especial. En la figura 4.5 no se aprecia una mejora tan sistemática de la EE respecto del algoritmo PABIC, si bien la varianza de la primera es menor que la de la segunda en la mayor parte de los casos. Esta mejora es menos apreciable en el caso de la recuperación del bias.

Además de estos resultados de rendimiento, si consideramos el potencial para la mejora que todavía tienen los algoritmos en el momento de terminar la ejecución, las estrategias evolutivas tienen mayor potencial que el algoritmo PABIC, debido a que la mayor población implica la posibilidad de generar nuevas soluciones. En la figura 4.6a se muestra la evolución de la función de error para una instancia del algoritmo PABIC. Puede observarse claramente que el algoritmo está encallado en un mínimo local desde una época temprana de la evolución. Esto demuestra el carácter local del algoritmo. En la figura 4.6b se muestra, para una instancia de la ejecución de la EE, la evolución de la función de ajuste del mejor individuo (línea continua) y la media de las funciones de ajuste de los individuos de cada generación. Una medida de la uniformidad de la población es la diferencia entre estas dos curvas. Mientras la población no sea uniforme la variedad de los individuos proporciona material adicional para la mejora, respecto del algoritmo de individuo único. Se aprecia en la figura 4.6b que el algoritmo se termina antes de que se agote esta variedad, y que todavía son posibles mejoras significativas de la función de ajuste. En conclusión, la EE mantiene potencial de mejora al terminar su ejecución, mientras que el PABIC permanece estancado en un mínimo local desde épocas tempranas de su evolución.

Comparación de GradPABIC con PABIC, EE

Como se ha descrito en secciones anteriores, GradPABIC es un algoritmo evolutivo de individuo único, basado en el algoritmo PABIC. La información de gradiente que se añade en la mutación aporta previsiblemente más rápida convergencia, sin modificar las características de localidad de la búsqueda del óptimo. Es más, es de esperar que el algoritmo sea todavía más local que PABIC. Los

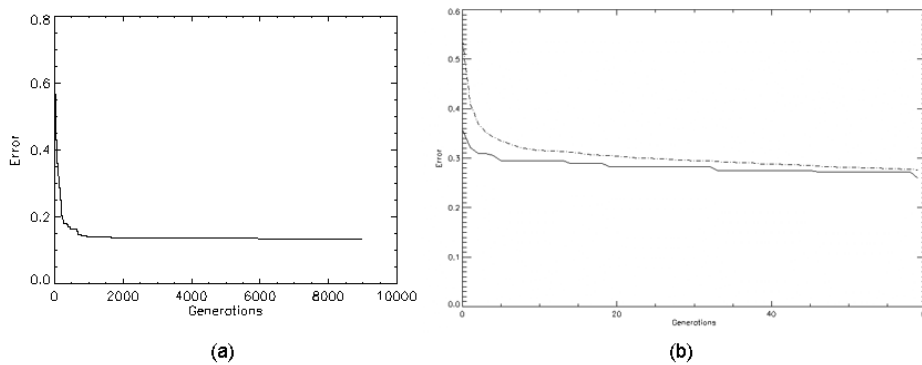


Figura 4.6: Evolución de la función de ajuste en sendas instancias del algoritmo PABIC (a) y la EE (b).

experimentos computacionales tratan de evaluar si la información de gradiente introduce alguna mejora sistemática sobre PABIC.

La correlación entre las imágenes restauradas y la imagen original, para las imágenes del tablero de ajedrez ordenadas por distorsión de la imagen corrupta respecto del original, se muestra en 4.4. Se puede apreciar que GradPABIC mejora a los otros dos algoritmos en todos los casos. También se aprecia que no se produce degradación de la respuesta de los algoritmos conforme aumenta la distorsión de la imagen corrupta. La figura 4.4b presenta las correlaciones de los bias estimados por los algoritmos evolutivos respecto del bias original sintético utilizado para generar las imágenes experimentales. Nuevamente, GradPABIC mejora sistemáticamente a PABIC y a la EE y la respuesta no se degrada en ningún algoritmo conforme aumenta la distorsión

La varianza de los resultados en las repeticiones de los algoritmos son una indicación de la fiabilidad de los algoritmos y la confianza en que los resultados de las figuras 4.4a y 4.4b son representativos de la respuesta que nos proporcionarían los algoritmos. En la figura 4.5a se muestran las varianzas de los resultados obtenidos en términos de correlación entre la imagen original y la imagen restaurada. Los resultados son inconclusivos, en términos generales GradPABIC tiene una varianza menor pero los resultados no son definitivos. La figura 4.5b presenta las varianzas de los resultados en términos de la correlación entre los sesgos estimados y los originales. Tampoco se puede asegurar que GradPABIC tenga inferior varianza, aunque tiene una tendencia mejor que

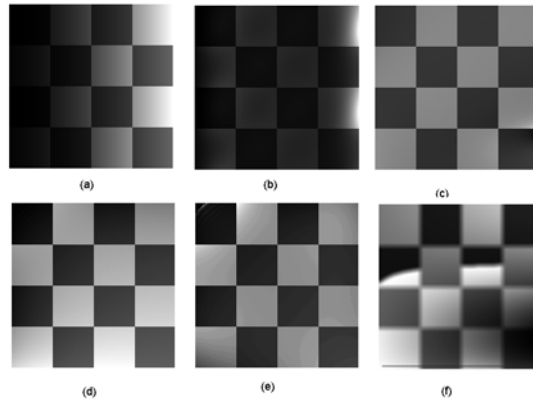


Figura 4.7: Resultados de la recuperación de la imagen con algoritmos evolutivos (a) Imagen corrupta, (b) filtrado homomórfico, (c) PABIC, (d) EE, (e) GradPABIC, (f) BMAP

los otros dos algoritmos.

En la figura 4.7 se muestran los resultados visuales de la recuperación de la imagen original para un caso concreto. La imagen corrupta es la de la figura 4.7a. El resultado del filtrado homomórfico es el de la figura 4.7b que puede apreciarse como muy deficiente. La figura 4.7c muestra el resultado obtenido con el algoritmo PABIC y la figura 4.7d muestra el resultado de la EE. Ambos son resultados aceptables y que no se distinguen mucho a simple vista del resultado del GradPABIC en la figura 4.7e.

En la figura 4.8 se muestra los resultados visuales de la estimación del sesgo de iluminación para un caso concreto. La figura 4.8a presenta el sesgo original utilizado para corromper la imagen. La figura 4.8b presenta el sesgo estimado por el filtrado homomorfo. Las figuras 4.8c, 4.8d y 4.8e presentan los sesgos estimados, respectivamente, por PABIC, la EE y GradPABIC. Otra vez, resulta difícil distinguir los resultados de estos últimos algoritmos a simple vista. También en la estimación de los sesgos el filtrado homomorfo da resultados muy inferiores a los de los otros algoritmos.

GradClasLeg versus algoritmos evolutivos

GradClassLeg no es un algoritmo evolutivo, sino un algoritmo de descenso de gradiente. Por ello la única variabilidad que presenta es la debida a las condi-

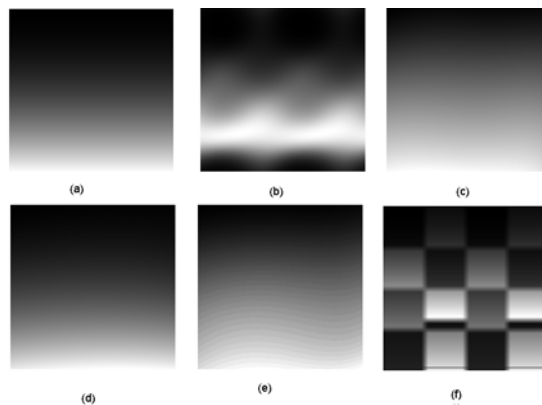


Figura 4.8: Resultados visuales de la estimación del sesgo del campo de iluminación (a) sesgo original, (b) filtrado homomorfo, (c) PABIC, (d) EE, (e) GradPABIC, (f) BMAP.

ciones iniciales. Hemos de indicar aquí que el algoritmo GradClassLeg es muy sensible al ajuste de sus parámetros, en especial la ganancia, lo que puede provocar comportamientos divergentes en ocasiones.

La correlación entre las imágenes restauradas y la imagen original, para las imágenes del tablero de ajedrez ordenadas por distorsión de la imagen corrupta respecto del original, se muestra en la figura 4.9a para los algoritmos GradClasPabic, PABIC y EE. Se muestra la media de 30 repeticiones, que en el caso de GradClasPabic se producen variando los estados iniciales aleatoriamente. Se confirma que GradClassLeg mejora tanto la EE como a PABIC. La figura 4.9b muestra las correlaciones de los sesgos de iluminación estimados y los originales. Nuevamente, GradClasPabic mejora ambos algoritmos evolutivos.

La figura 4.10a muestra la varianza de las correlaciones obtenidas entre las imágenes restauradas y la imagen original. Al igual que ocurría con el GradPABIC, los resultados no son definitivos, pero parecen indicar que GradClassLeg tiene menor varianza. Esta disminución de la varianza de GradClassLeg se aprecia mucho más en los resultados relativos a la correlación de los sesgos estimados que se presenta en la figura 4.10b.

Dado que los resultados visuales son muy similares a los ofrecidos en las figuras 4.7 y 4.8, no presentamos nuevos resultados visuales para GradClassLeg. Una somera comparación entre las figuras 4.4 y 4.9 muestra que el algoritmo

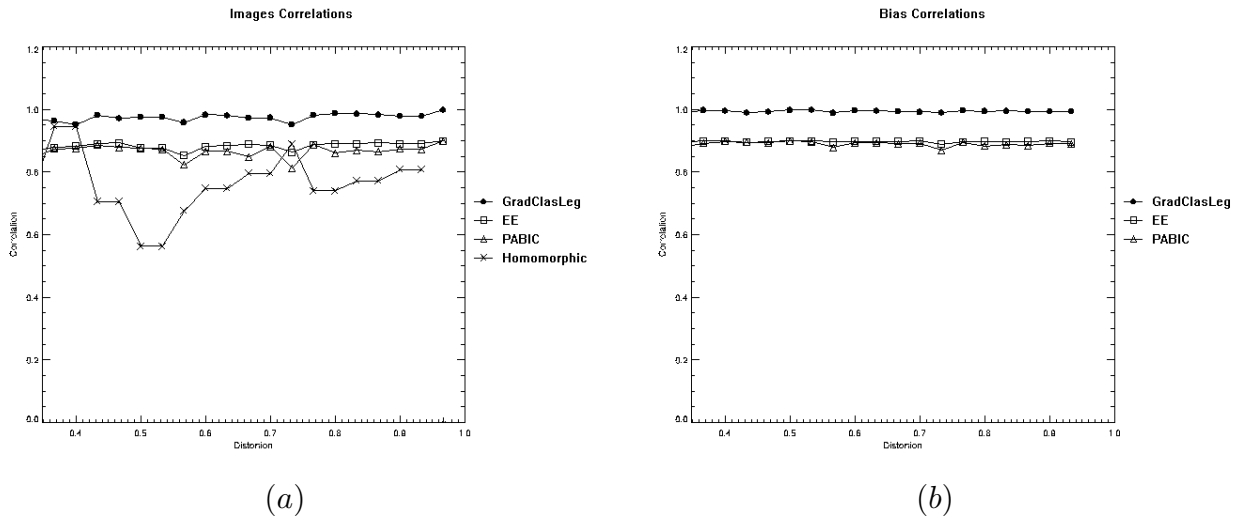


Figura 4.9: Correlaciones con los originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos GradClassLeg, EE y PABIC

GradClassLeg mejora al algoritmo memético GradPABIC ligeramente. Este resultado es bastante inesperado puesto que es de esperar que el algoritmo aleatorio sea capaz de escapar de mínimos locales mejor que el algoritmo determinista. Sin embargo puede entenderse si consideramos que las repeticiones realizadas para estimar el comportamiento medio conllevan un proceso de remuestreo que introduce aleatoriedad en el conjunto total del experimento. Esto es, el experimento completo con las 30 repeticiones es, en sí, un algoritmo aleatorio de minimización.

4.2.3. Comparación con algoritmos no evolutivos

En esta sección revisamos los resultados de la comparación de los algoritmos propuestos con los algoritmos de corrección que no son de índole evolutiva, el algoritmo de Wells, el algoritmo bayesiano BMAP, el algoritmo borroso BFCM. Puesto que las correlaciones con los sesgos del campo de iluminación no aportan información adicional como hemos comprobado en la sección anterior, prescindimos de presentar las gráficas correspondientes en esta comparación.

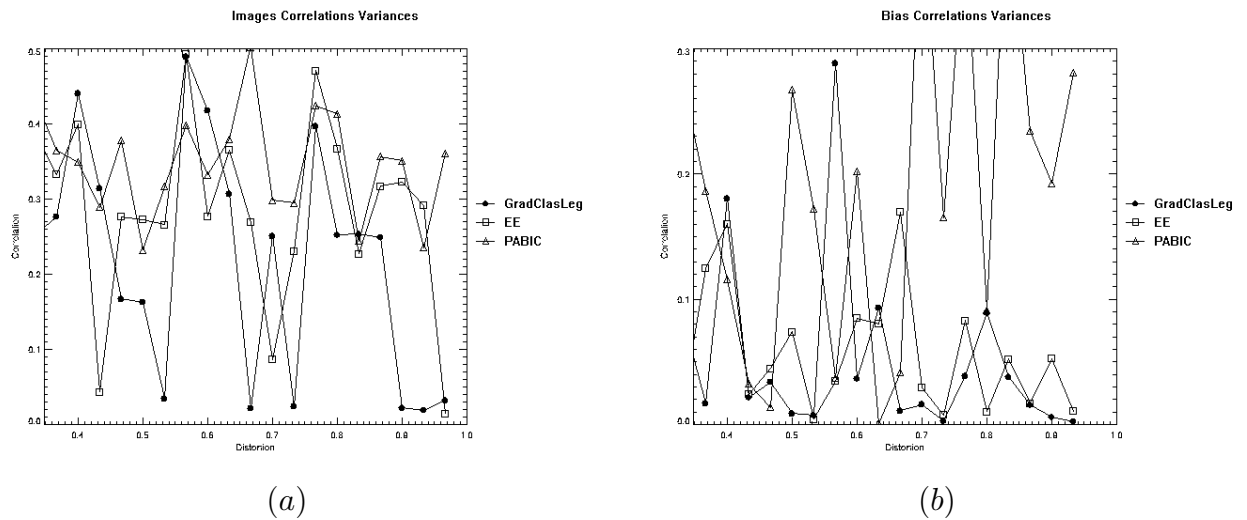


Figura 4.10: Varianzas de las correlaciones con los originales de la imagen corregida (a) y del sesgo estimado (b), para los algoritmos GradClassLeg, EE y PABIC

Algoritmo adaptativo de Wells.

La correlación entre las imágenes restauradas y la imagen original, para las imágenes del tablero de ajedrez ordenadas por distorsión de la imagen corrupta respecto del original, se muestra en la figura 4.11a comparando los resultados obtenidos con GradPABIC y con el algoritmo adaptativo de Wells. En la figura 4.11b presentamos los mismos resultados comparando con GradClasLeg. Ambos algoritmos mejoran sustancialmente al algoritmo de Wells. Como el algoritmo de Wells es básicamente un algoritmo de clasificación más bien que de restauración, hemos considerado las gráficas de aciertos en las clasificaciones obtenidas con el algoritmo de Wells, tomando como referencia la clasificación de la imagen original usando los valores nominales de la intensidad de las clases. Para GradPABIC hemos realizado la clasificación de los píxeles de la imagen restaurada usando también los valores nominales de las clases. Estas gráficas de aciertos en la clasificación son proporcionales a las tasas de aciertos en una clasificación supervisada. Se muestran estos resultados en la figura 4.12. Se observa que GradClassPabic mejora al algoritmo de Wells. También se observa que los resultados se degradan sistemáticamente al aumentar la distorsión de la imagen corrupta, algo que no ocurriría al considerar la restauración de la imagen.

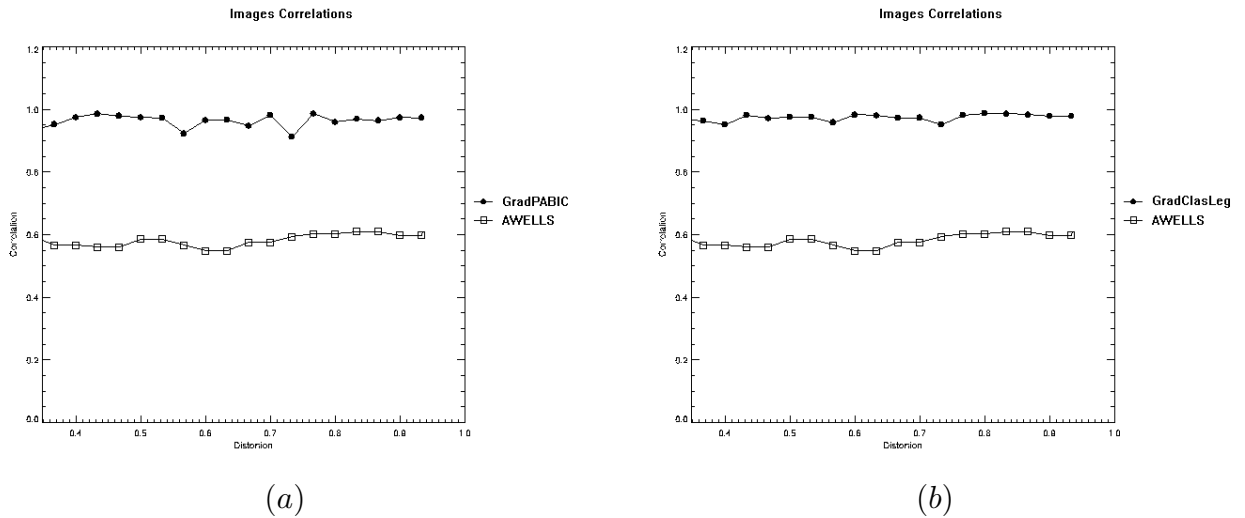


Figura 4.11: Correlaciones de la imagen corregida, (a) para los algoritmos GradPABIC, Wells, (b) GradClassLeg y Wells

Hay que hacer notar que estamos utilizando los valores de los parámetros (tamaño de la máscara de suavización) recomendados en el artículo original [120]. Parece ser que estos valores no se ajustan muy bien a los datos con los que realizamos estos experimentos, en el sentido de que la variación debida al campo de iluminación es muy global y suave, frente a los casos estudiados en [120] que tenían fenómenos muy localizados. Por esta razón realizamos el experimento de análisis de sensibilidad que se detalla más adelante, en la sección 4.2.5.

BMAP y BFCM

La correlación entre las imágenes restauradas y la imagen original, para las imágenes del tablero de ajedrez ordenadas por distorsión de la imagen corrupta respecto del original, se muestra en la figura 4.13a para los algoritmos GradPABIC, BFCM y BMAP. El algoritmo BMAP proporciona unos resultados realmente malos comparado con los restantes algoritmos. Creemos que se deben a que la definición de los parámetros del algoritmo son más adecuados para la corrección de inhomogeneidades locales y de pequeña magnitud. Por ello realizamos un experimento de análisis de sensibilidad que comentamos más adelante en la sección 4.2.4. En nuestras imágenes experimentales sintéticas se producen grandes variaciones en el gradiente tanto en magnitud como en dis-

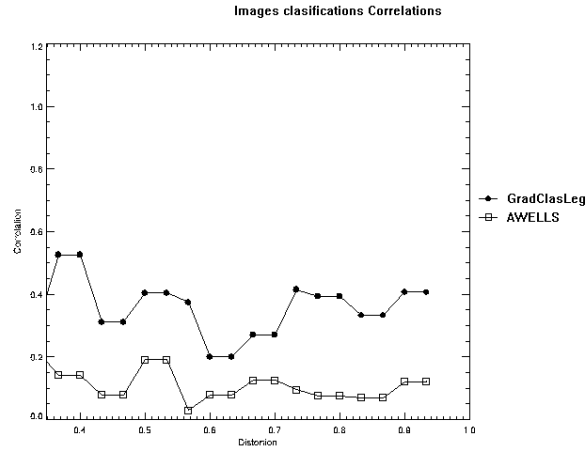


Figura 4.12: Aciertos en la clasificación de las imágenes corregidas comparados con la clasificación nominal.

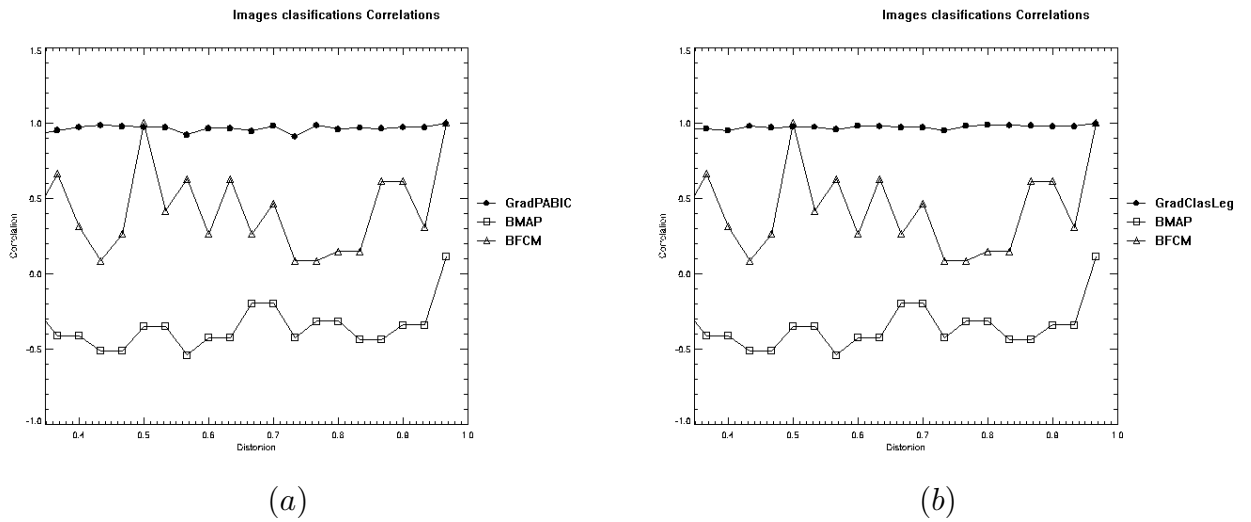


Figura 4.13: Correlaciones de las imágenes corregidas y de la imagen original, (a) algoritmos GradPABIC, BFCM, BMAP, (b) algoritmos GradClassLeg, BFCM, BMAP

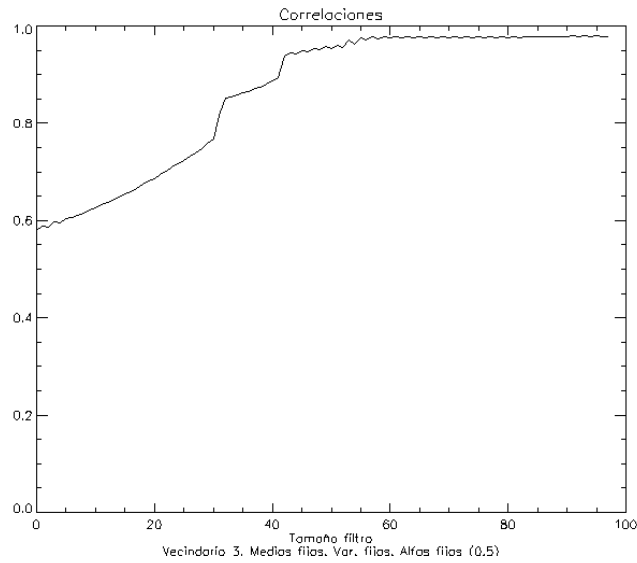


Figura 4.14: Sensibilidad del BMAP respecto del tamaño del filtro de suavización sobre una imagen de tablero de ajedrez, con radio 2 de vecindario en el MRF que modela las probabilidades a priori.

tribución espacial en la imagen. El algoritmo BFCM mejora al BMAP, y el GradPABIC mejora a ambos. La figura 4.13b muestra la comparación de resultados de los mismos algoritmos con GradClassLeg. También, se observa que GradClassLeg mejora notablemente respecto de BMAP y BFCM.

Aunque el BMAP está basado en un modelo probabilístico, la realización es determinista y por tanto no tiene sentido realizar las repeticiones del algoritmo para observar la varianza en su respuesta. Lo mismo ocurre con el algoritmo BFCM. Las únicas dependencias aleatorias son las condiciones iniciales y para tenerlas en cuenta lo que hacemos es iniciar todos los algoritmos en las mismas condiciones. Por todo esto, no presentamos gráficas de varianza en los resultados de estos algoritmos.

4.2.4. Analisis de sensibilidad BMAP

En los experimentos comentados en secciones anteriores, se han aplicado los valores de los parametros de los algoritmos que se sugieren en las referencias originales. Sin embargo, los resultados obtenidos nos hacen dudar de que estos

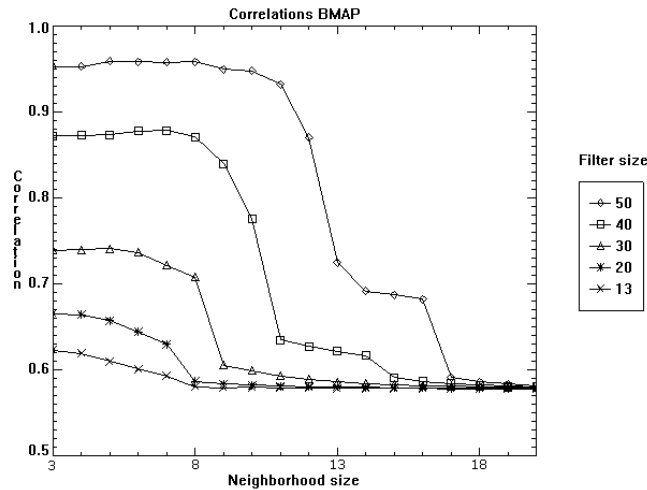


Figura 4.15: Sensibilidad del BMAP respecto del tamaño del vecindario en el MRF de las probabilidades a priori.

valores sean apropiados para el conjunto de imágenes que estamos probando. Para verificarlo, el caso del BMAP realizamos el siguiente experimento de análisis de sensibilidad de los dos parámetros libres: la potencia del filtro pasabaja (que consiste en la convolución con una máscara de promediado) realizado sobre del estimador del sesgo y el tamaño del vecindario en el campo markoviano que modela las probabilidades *a priori* de las imágenes.

Para evaluar experimentalmente la influencia del tamaño de la máscara de suavización que se aplica sobre el sesgo del campo de iluminación estimado para regularizarlo, hemos realizado la restauración de una de las imágenes de tablero de ajedrez corrupta que son la base experimental en la sección 4.2 variando dicho tamaño. Por supuesto que los resultados de este experimento no tienen validez general en cuanto a la decisión de fijar este parámetro del algoritmo, sin embargo si que muestran la fuerte sensibilidad que este algoritmo puede tener respecto de este parámetro. En la figura 4.14 se muestra como varía la correlación de la imagen restaurada respecto de la imagen original al variar el diámetro de la máscara. Se aprecia una fuerte sensibilidad hasta un valor de diámetro de 60, a partir del cual los resultados son estables y cercanos al valor óptimo de correlación 1. Los valores recomendados en la referencia original [103] son muy

inferiores a éste. Esta sensibilidad se produce en los algoritmo no paramétricos en general y son una fuente de incertidumbre respecto de la optimalidad del resultado.

El radio del vecindario en el campo markoviano es otro parámetro a fijar en el BMAP para el que el valor sugerido es 2. Este radio determina la suavidad de las imágenes y la extensión de las regiones constantes en la imagen. Intuitivamente, el resultado esperado al variar este parámetro es que, para nuestra imagen de referencia que tiene grandes trozos constantes, se mejorasen los resultados del algoritmo.

En la figura 4.15 se muestra un estudio conjunto de sensibilidad del orden del vecindario y del diámetro de la máscara de suavización al aplicar BMAP sobre una imagen de tablero de ajedrez corrupta. Obviamente, por limitaciones computacionales, para vecindarios de orden superior a 2 no se realizan las estimaciones de los potenciales de la distribución *a priori* a partir de la imagen segmentada inicialmente. Se observa que el rendimiento del algoritmo disminuye al aumentar el tamaño del vecindario, un resultado inesperado puesto que se espera que las restricciones de suavidad más fuertes que se implementan con un vecindario más grande fueren la definición de regiones más grandes consistentes con las regiones de la imagen original. Por contra, el factor más influyente en la mejora del rendimiento de BMAP resulta ser el tamaño de la máscara pasabaja que suaviza el estimador del sesgo de la iluminación. Se aprecia que para valores grandes de este tamaño de máscara, los resultados se mantienen aceptables a pesar del crecimiento del tamaño del vecindario. Sin embargo, en todos los casos el crecimiento del radio del vecindario llega a provocar un descenso catastrófico del rendimiento. El valor crítico del radio de vecindario depende del tamaño de la máscara de suavización. Lo que se observa, al contemplar las imágenes producidas por la ejecución del algoritmo, es que una de las dos clases tiende a fagocitar a la otra y a dominar la imagen, de ahí el descenso en rendimiento.

4.2.5. Análisis de sensibilidad del algoritmo de Wells

De forma análoga a la exploración de comportamiento realizada para el algoritmo BMAP, realizamos, sobre una de las imágenes de tablero de ajedrez corruptas, un experimento de sensibilidad al tamaño de la máscara de suavización que se aplica en el algoritmo de Wells sobre la clasificación MAP. En la figura 4.16 se aprecia con claridad la dependencia respecto de este parámetro para el algoritmo de Wells. Conforme aumenta el radio de la máscara de suavización,

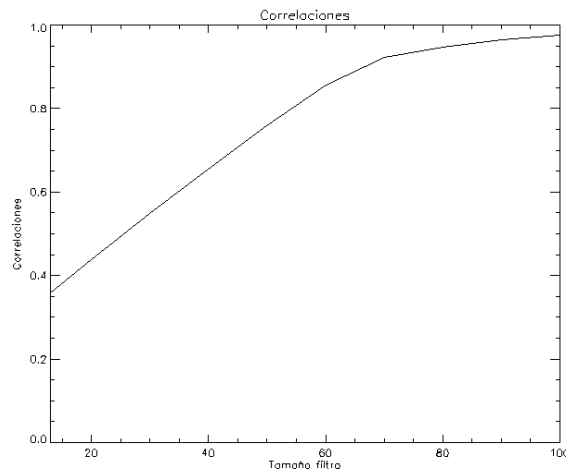


Figura 4.16: Influencia del diametro de la mascara de suavización sobre el resultado del algoritmo de Wells.

se mejoran los resultados del algoritmo de forma lineal. Nuevamente, el escenario de la experimentación es muy desfavorable para el algoritmo, que está más adaptado a inhomogeneidades más localizadas y de menor magnitud que las que se presentan en la imagen experimental.

4.2.6. Conclusiones sobre las imágenes de tablero de ajedrez

En primer lugar constatamos que el algoritmo PABIC sufre de las deficiencias esperables en un algoritmo de individuo único frente a un algoritmo evolutivo de población mayor que 1 individuo. El algoritmo PABIC se encalla en mínimos locales con mayor facilidad que una EE con el mismo tiempo de cálculo y posee menor esperanza de mejora según se desprende de la observación de las gráficas de la evolución de las curvas de error. Además, la EE mejora sistemáticamente al algoritmo PABIC dado el mismo tiempo de cálculo medido en llamadas a la función de ajuste.

Hemos realizado experimentos análisis de sensibilidad de los algoritmos de Wells y BMAP para intentar determinar si los parámetros recomendados en la literatura son los más apropiados para el conjunto de imágenes experimentales. Encontramos que aumentando el tamaño de la máscara de suavización del algoritmo de Wells mejoran considerablemente los resultados, lo que indica que

los valores recomendados son más apropiados para situaciones localizadas. Para el BMAP encontramos que también el aumento de tamaño de la máscara de convolución tiene un efecto positivo en nuestro experimento, mientras que, sorpresivamente, el aumento del vecindario del campo markoviano tiene un efecto negativo, a pesar de que la imagen original está constituida por grandes regiones constantes.

En suma los experimentos sobre este conjunto de imágenes demuestran que los algoritmos propuestos GradPABIC y GradClassLeg mejoran, en términos de la correlación de la imagen restaurada y la imagen original, a los algoritmos que se han implementado como algoritmos competidores. La varianza de estos resultados no es en general menor para nuestros algoritmos. Un resultado indeseable, puesto que quisieramos que fueran fiables en su mejoría respecto de los demás algoritmos.

4.3.Experimentos sobre las imagenes de cerebro sintéticas.

Las imagenes experimentales se obtienen como rodajas de un volumen simulado de una imagen pesada en T1 de un modelo de cerebro que hemos obtenido de la pagina [14]. Los principios de la sintesis del volumen de señal IRM que han servido para la realización de estas simulaciones se presentan en [20], [22], [72], [73]. Se trata de un modelo de craneo sano cuya anatomia se ha obtenido de un individuo sano. En la misma página web se proporcionan imágenes con ruido aditivo y campos de iluminación no uniformes simulados. Sin embargo, hemos encontrado que las perturbaciones de iluminación son muy reducidas en comparación con las que hemos venido probando. Por ello hemos decidido añadir nuestros propios campos de iluminación sobre las imágenes.

En la figura 4.17 se muestran las imagenes originales en el experimento. Debido a la gran cantidad de cómputo necesario, no hemos aplicado los algoritmos a los volúmenes 3D sino a cortes representativos de las estructuras cerebrales: axial, coronal y sagital.

Por otra parte, la figura 4.18 muestra un caso de corrupción de la imagen con un campo de iluminación sintético, cuyos parámetros son, por tanto, conocidos y podemos aplicar las mismas estrategias de evaluación cuantitativa de la sección anterior. Esto es, los resultados se presentarán como gráficas cuyo eje de ordenadas es la inversa de la correlación de la imagen corrupta con la imagen original limpia. De esta manera, las imágenes experimentales se han ordenado por distorsión creciente para mostrar la dependencia, si existe, de los resultados

respecto de esta distorsión. El eje de abscisas corresponde con la correlación entre la imagen corregida y la imagen limpia, o bien entre el campo de iluminación estimado y el campo sintético utilizado para corromper la imagen.

En relación con las imágenes del tablero de ajedrez, estas imágenes tienen mayor complejidad y son más cercanas a las aplicaciones reales. Tienen tres clases de tejidos de interés: materia gris, materia blanca y líquido cerebrospinal, además de las regiones sin tejidos, el aire, donde los datos de la imagen corresponden sólo al ruido aditivo. En la figura 4.19 se muestran las imágenes de la clasificación de las imágenes originales en tres clases cuyas intensidades medias se han obtenido aplicando el algoritmo de k-medias sobre las imágenes originales. Los valores medios se han utilizado en el capítulo experimental 4 como valores de las medias de intensidad de las clases.

Los valores de los parámetros de los algoritmos son los siguientes:

- PABIC: número de generaciones:900, $c_{\text{grow}} = 1.01$, $c_{\text{shrink}} = 0.99$
- GradClasLeg: grado máximo de los polinomios 6, $\alpha = 0.001$
- GradPABIC: grado máximo de los polinomios 6,
- EE: grado máximo de los polinomios 6, número de individuos 10, número de mutaciones 150, número de generaciones 60
- Algoritmo de Wells: Diámetro de la máscara de suavización: 25
- BMAP: $\alpha_1 = 0.56$, $\alpha_2 = -0.5$, vecindario=3, diámetro de la máscara de suavización=13
- BFCM: alfa: 0.01, radio del vecindario: 1

4.3.1. Resultados de los algoritmos evolutivos

La figura 4.20 muestra para cada una de las rodajas originales correspondientes a los cortes axial, coronal y sagital, las gráficas de las correlaciones entre la imagen original y las imágenes corregidas, así como las correlaciones entre los sesgos originales y los estimados, para las rodajas corruptas ordenadas por distorsión creciente, cuando se aplican los algoritmos GradPABIC, PABIC y EE. La misma información relativa a GradClassLeg se muestra en la figura 4.21. Las figuras (a,c,e) presentan las correlaciones de la imagen original y las figuras (b,d,f) las

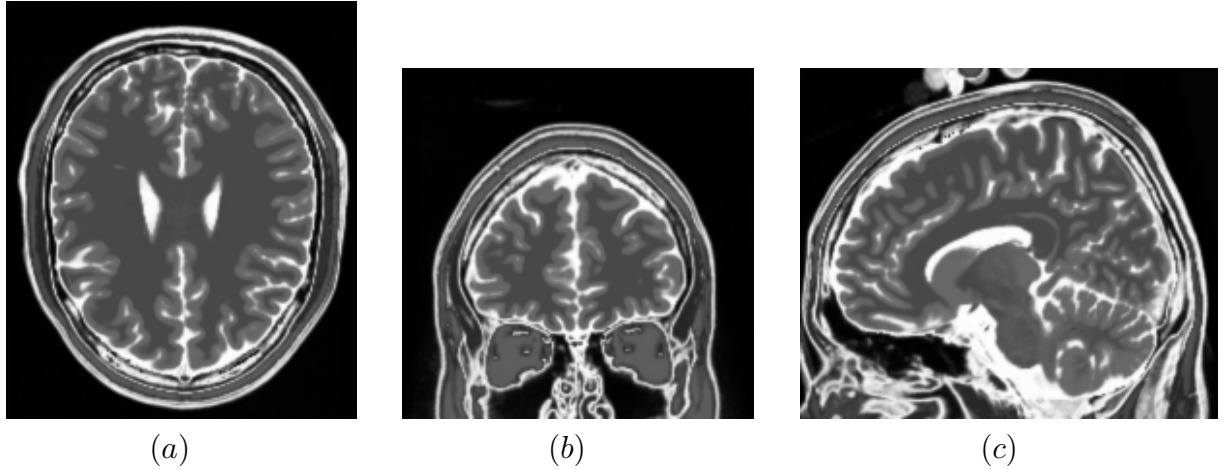


Figura 4.17: Imágenes de cortes del volumen sintético de cerebro usados como originales (a) corte axial, (b) corte coronal, (c) corte sagital

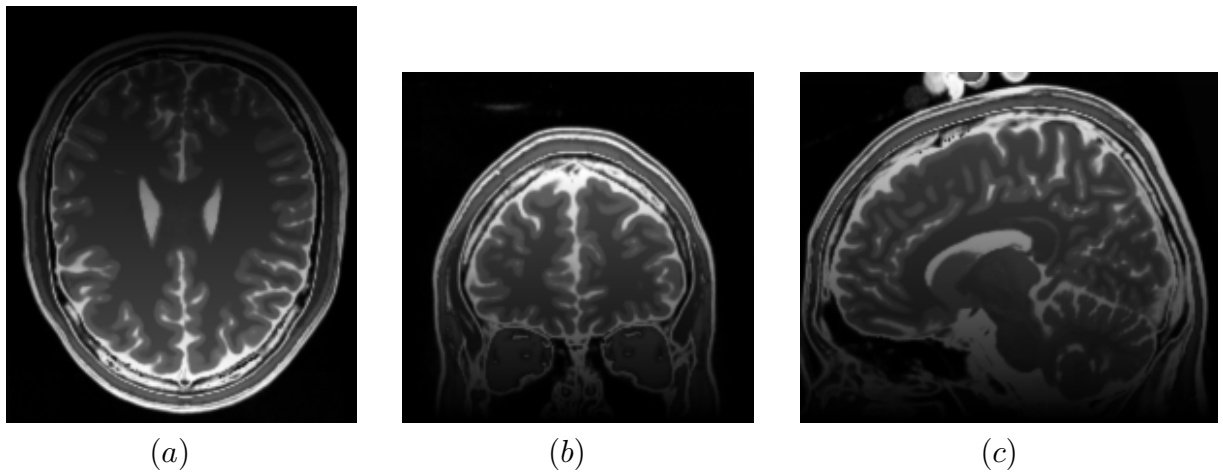


Figura 4.18: Imágenes de cortes del volumen sintético de cerebro corrompidas con un sesgo de iluminación (a) corte axial, (b) corte coronal, (c) corte sagital

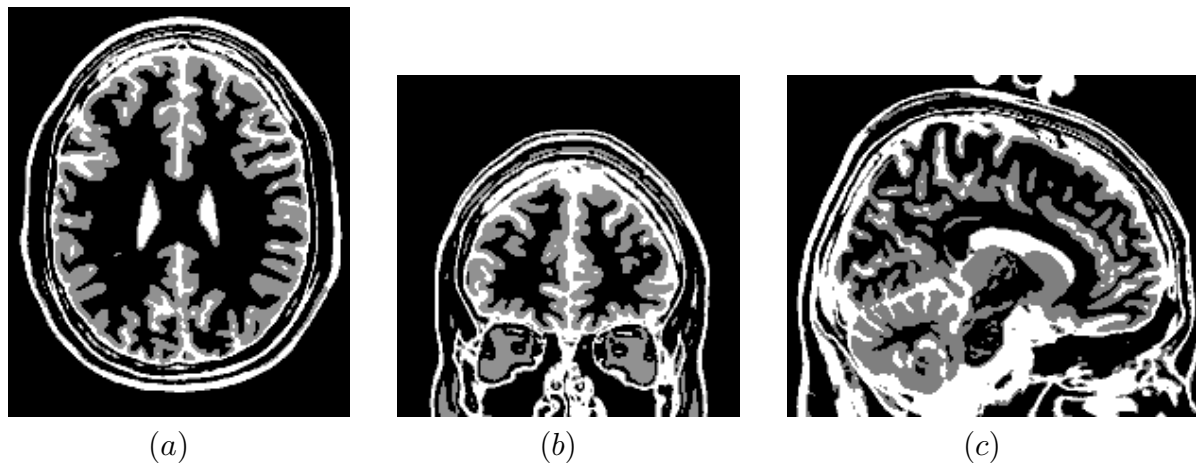


Figura 4.19: Clasificación de las imágenes de cortes del cerebro en líquido cerebro-espinal, materia gris y materia blanca. Cortes (a) axial, (b) coronal y (c) sagital

correlaciones de los sesgos. Se aprecia fácilmente que tanto GradPABIC como GradClassLeg mejoran a EE y al PABIC en términos de calidad de la imagen restaurada. Sin embargo en el caso de la calidad del sesgo estimado no se aprecia una diferencia significativa entre los algoritmos que permita asegurar la superioridad de un algoritmo respecto de los demás.

Es de apreciar que no se produce la mejora de la EE sobre el PABIC como en las imágenes de tableros de ajedrez. Esto puede ser debido al aumento del número de clases que lleva consigo un aumento en la complejidad de la superficie de la función de ajuste. Las estrategias evolutivas genéricas estarían todavía en una etapa temprana cuando son forzadas a terminar por la limitación computacional impuesta, mientras que el PABIC ha tenido una convergencia más rápida hacia un mínimo local.

La variabilidad en el sesgo estimado es muy superior al caso de las imágenes de ajedrez, suponemos que debido a la mayor complejidad de las imágenes se producen interacciones entre los estimadores del sesgo y la estructura de la imagen. El alto orden de los polinomios de Legendre utilizados puede dar lugar a este tipo de fenómenos que pueden considerarse como consecuencia de un sobreajuste.

En las figuras 4.22, 4.23, 4.24, 4.25, 4.26 se muestran resultados visuales de la aplicación de los algoritmos sobre las rodajas del volumen sintético presentado

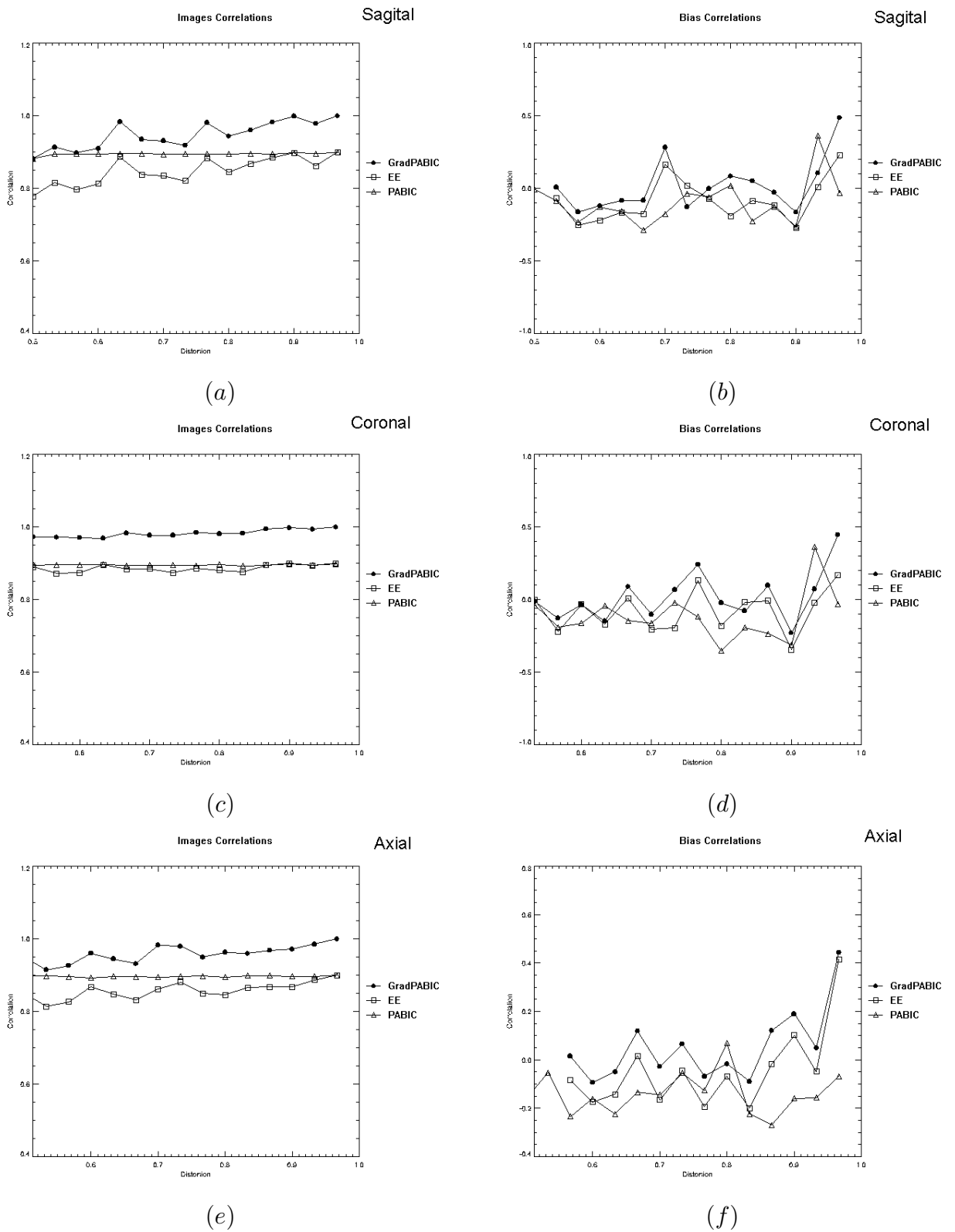


Figura 4.20: Correlaciones de las imágenes corregidas (a,c,e) y de los sesgos estimados (b,d,f) para los algoritmos evolutivos, sobre las imágenes de los cortes axial, coronal y sagital

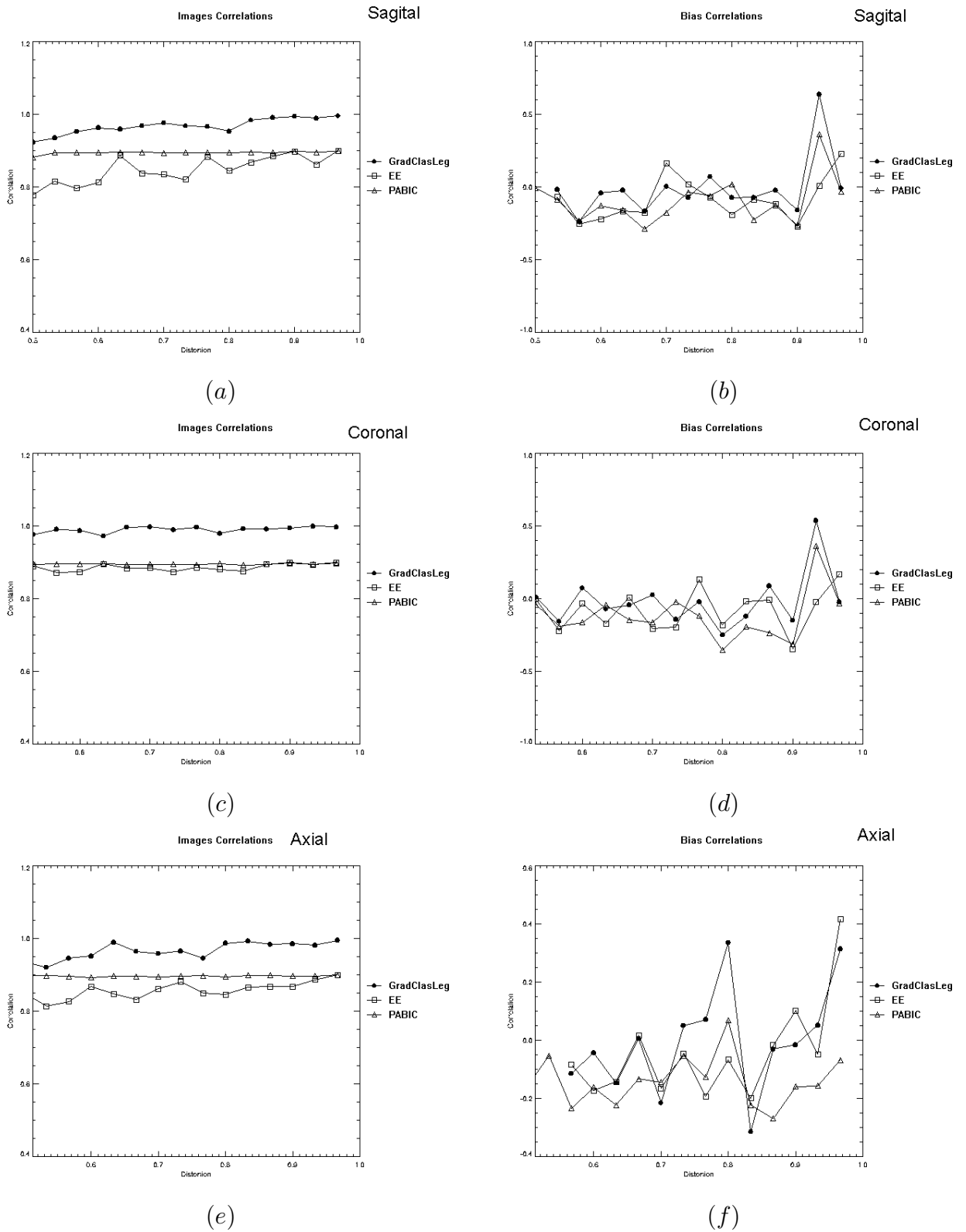


Figura 4.21: Correlaciones de las imágenes corregidas (a,c,e) y de los sesgos estimados (b,d,f) para los algoritmos evolutivos y GradClassLeg sobre las imágenes de los cortes axial, coronal y sagital

en la figura 4.18. En estas figuras se muestran tanto las imágenes corregidas como las estimaciones del sesgo del campo de iluminación.

Las diferencias entre las imágenes corregidas son difícilmente apreciables para la percepción visual humana, aunque pueden tener como sabemos efectos importantes en las realizaciones de algoritmos automáticos de segmentación. En este ejemplo concreto, se puede apreciar que las imágenes restauradas con EE, PABIC y gradPABIC (figuras 4.22 y 4.23, 4.24) son un poco más claras que las restauradas con GradClassLeg (figuras 4.25, 4.26). La diferencia entre las imágenes en las figuras 4.25 y 4.26 consiste en que las primeras se obtuvieron fijando el valor de las medias de intensidad correspondientes a las clases en los valores estándar utilizados por los restantes algoritmos, mientras que las segundas se obtuvieron permitiendo que el algoritmo GradClassLeg reestimara dichas medias en el proceso de descenso de gradiente de la función de error. Un resultado interesante es que las imágenes corregidas cuando GradClassLeg estima las medias son más claras que las obtenidas fijando estos valores a los nominales utilizados por los otros algoritmos.

Las diferencias en las visualizaciones de los sesgos estimados son más apreciables. En concreto, los estimados con EE y GradPABIC son bastante similares. También lo son los estimados con GradClassLeg (con estimación de las medias de intensidad) y PABIC. Mientras que los estimados por GradClassLeg con las medias de intensidad fijadas son bastante distintos. En todo caso se trata, como era de esperar, de funciones muy suaves.

4.3.2. Resultados con el algoritmo de Wells, BMAP y BFCM

En la figura 4.27 se presentan las gráficas de resultados comparativos entre el algoritmo de Wells, BMAP, BFCM y GradPABIC sobre las imágenes sintéticas de cerebro corruptas con diversos campos de iluminación de distorsión creciente. En este caso no presentamos los resultados de correlación de los sesgos de iluminación estimados. Se aprecia en la comparación de GradPABIC con el algoritmo adaptativo de Wells unos resultados muy similares y poco variables respecto de la distorsión. Los algoritmos BMAP y BFCM proporcionan resultados sistemáticamente inferiores a GradPABIC. Se aprecia también una fuerte variabilidad de los resultados de BMAP y BFCM. Los pésimos resultados de BMAP son debidos probablemente al tamaño del filtro pasabaja que se aplica en la estimación del bias, de acuerdo con los resultados del análisis de sensibilidad presentado en la sección 4.2.4.

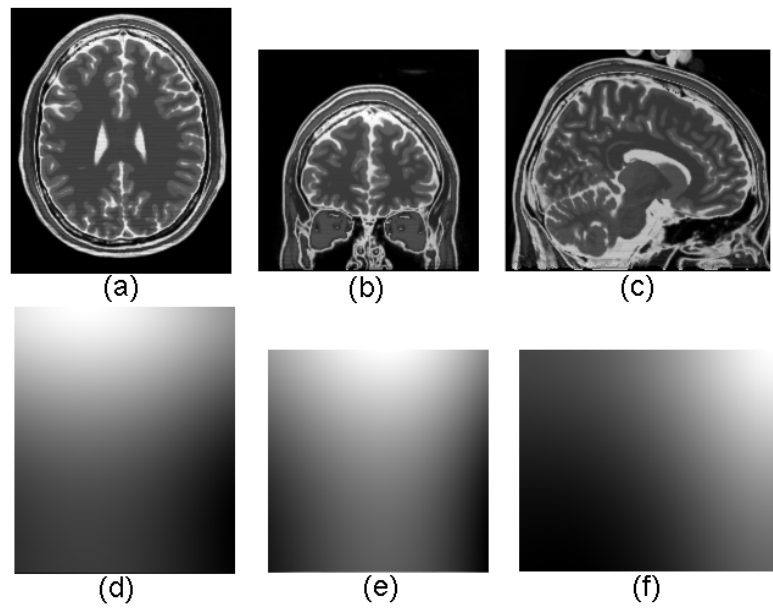


Figura 4.22: Resultados visuales de las estrategias evolutivas sobre los cortes del volumen sintético.

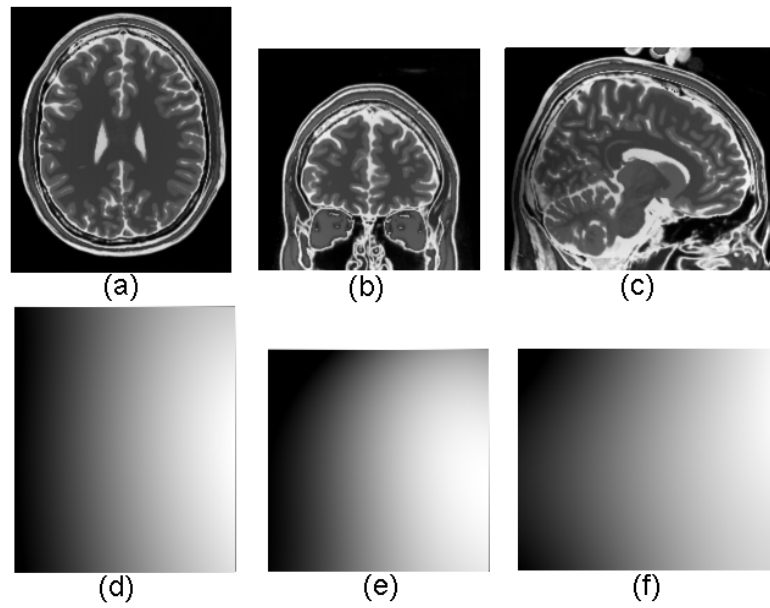


Figura 4.23: Resultados visuales del algoritmo PABIC sobre las imágenes de los cortes del volumen sintético. (a, b, c) imágenes corregidas, (d, e, f) sesgos estimados.

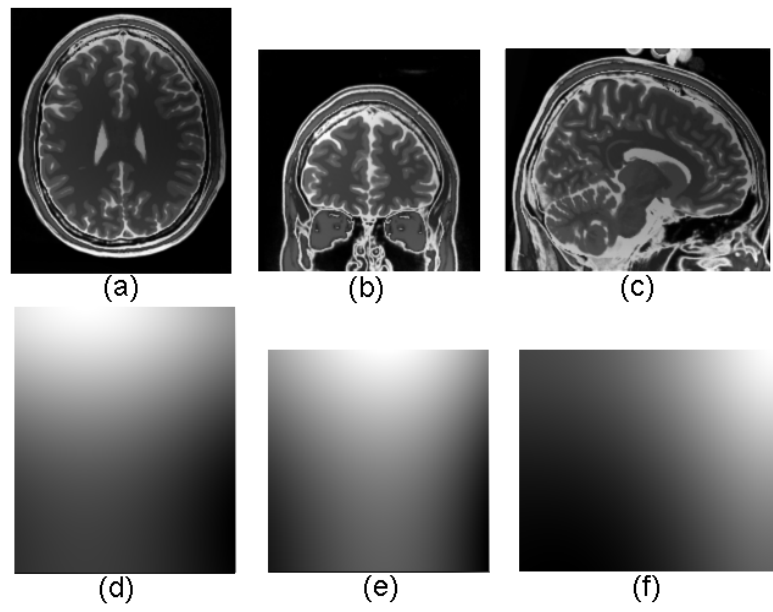


Figura 4.24: Resultados visuales del algoritmo GradPABIC sobre los cortes del volumen sintético

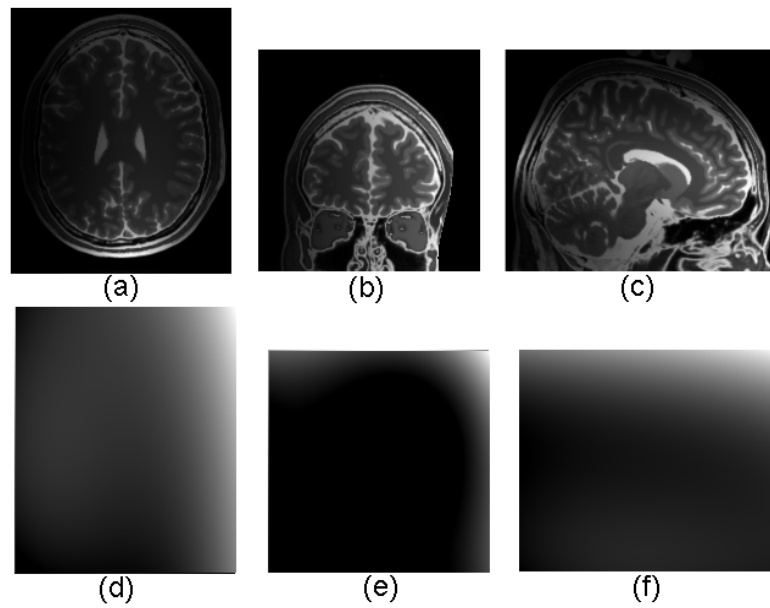


Figura 4.25: Resultados visuales de la aplicación del algoritmo GradClassLeg, estado inicial arbitrario, sin estimar las intensidades medias de las clases. (a,b,c) imágenes corregidas, (d,e,f) sesgos de iluminación estimados.

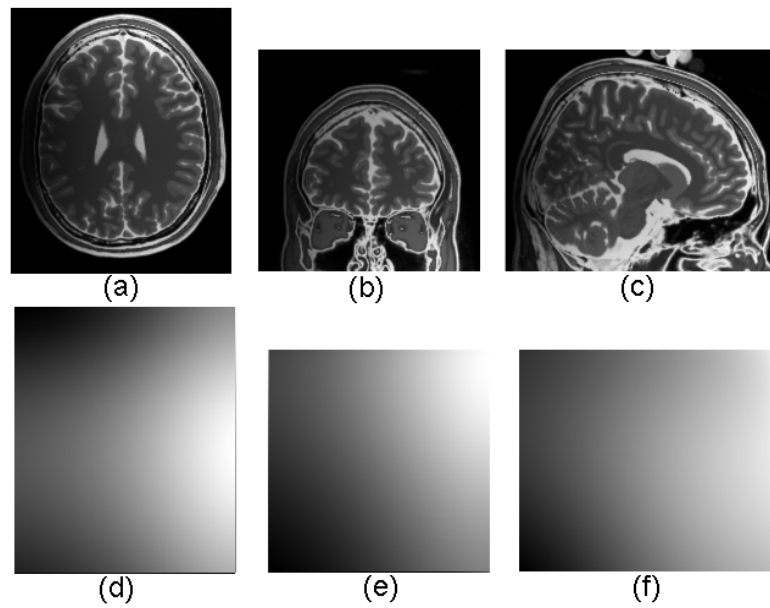


Figura 4.26: Resultados visuales de aplicar el algoritmo GradClassLeg sobre las imágenes sintéticas corruptas, estimando las medias de las intensidades de las clases. (a,b,c) imágenes corregidas, (d,e,f) sesgos estimados.

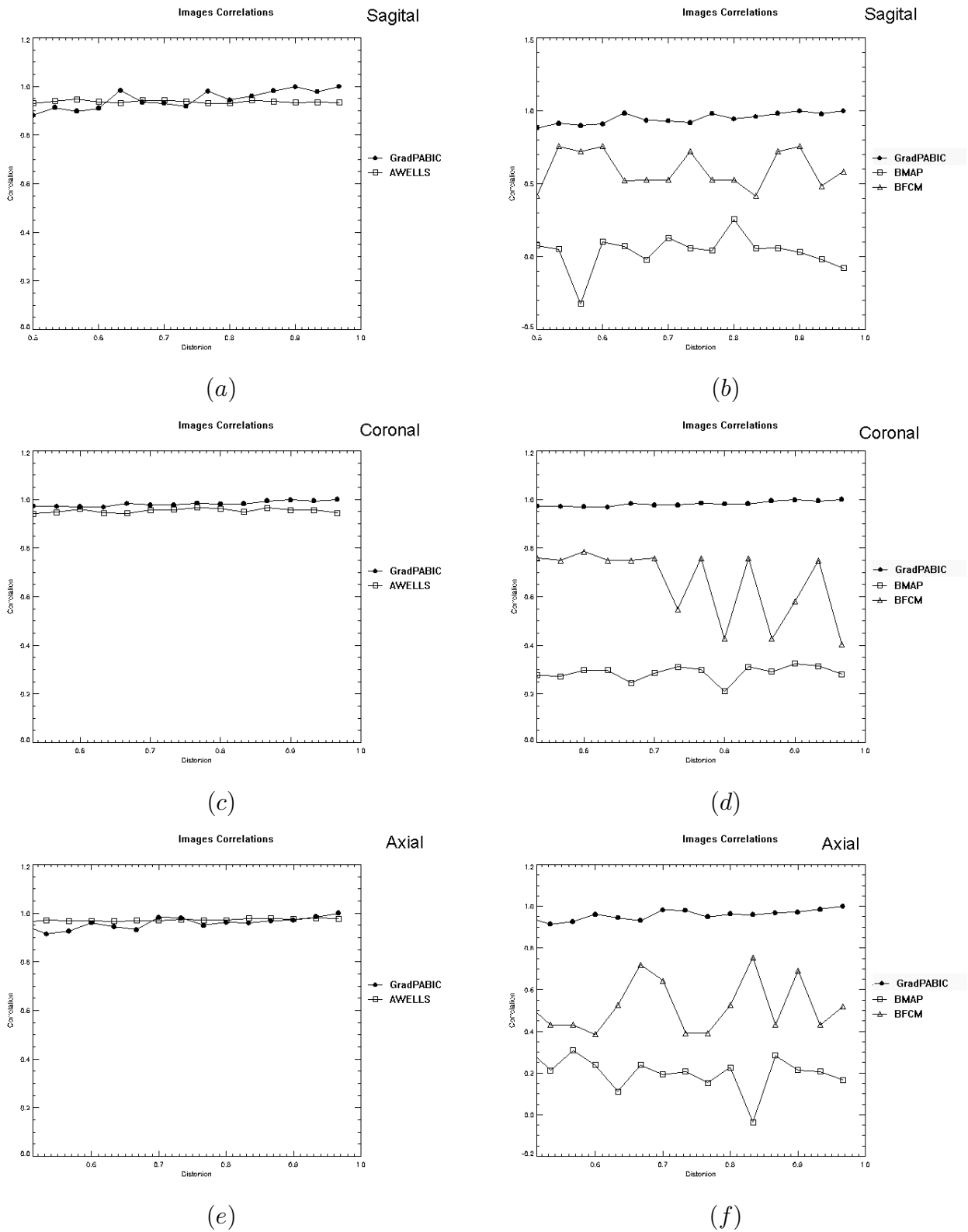


Figura 4.27: Correlación de las imágenes restauradas y la original sobre los cortes del volumen sintético de cerebro, GradPABIC versus el algoritmo de Wells, BFCM y BMAP .

En la figura 4.28 se presentan los resultados comparativos del algoritmo GradClassLeg frente al algoritmo adaptativo de Wells, BFCM y BMAP. GradClassLeg proporciona resultados sistemáticamente superiores a los restantes algoritmos, si bien la diferencia con el algoritmo de Wells es pequeña. La diferencia de rendimiento de los algoritmos sobre los distintos cortes no es significativa para los algoritmos GradPABIC y GradClassLeg, ni para el algoritmo de Wells. Sin embargo si se aprecian respuestas muy distintas en los diversos cortes para los algoritmos BFCM y BMAP. En ningún caso se incluye la estimación de la intensidad media de las clases en la ejecución de los algoritmos, siendo el objetivo de los experimentos estimar la imagen corregida para su clasificación posterior, u obtener la clasificación de los píxeles directamente.

En las figuras 4.29,4.30, 4.31 se muestran resultados visuales de la aplicación de los algoritmos sobre las rodajas del volumen sintético presentado en la figura 4.18. Las imágenes corregidas por el algoritmo de wells (figura 4.29) son de buena calidad, comparables a las obtenidas en la sección anterior. Los sesgos estimados mantienen parte de la estructura de la imagen, producto del proceso que consiste en la suavización de la estimación puntual del sesgo atribuible al campo de iluminación. De no aplicarse esta suavización, el sesgo estimado y la corrección serían triviales.

Esto es lo que ocurre con el algoritmo BFCM, cuyos resultados visuales se presentan en la figura 4.30. Aparentemente las imágenes que presentamos como sesgos (figuras 4.30d,e,f) son más parecidas a lo que entendemos como imagen corregida, mientras que las imágenes corregidas (figuras 4.30a,b,c) son más parecidas a campos de iluminación cuantizados o discretizados. Pensamos que este intercambio de roles se debe a la ausencia de un paso de suavizado en la estimación del sesgo. El sesgo se convierte en una estimación puntual que degenera en una especie de imagen corregida, mientras que las imágenes corregidas están restringidas por condiciones de suavidad impuestas en la función objetivo, lo que degenera en una función suave discretizada.

El algoritmo BMAP, cuyos resultados visuales aparecen en la figura 4.31 tiene su propia idiosincrasia. La estimación del campo de iluminación se convierte en una suavización de la imagen, mostrando las regiones de mayor concentración de los píxeles con valores significativos. Por contra, la imagen corregida tiene un aspecto poco natural, parece haber sido sometida a alguna forma de reducción del contraste. En la imagen corregida aparecen efectos sobre regiones de la imagen que no corresponden a tejidos y que son debidas a que el sesgo estimado se extiende más allá de la región de tejidos en la imagen.

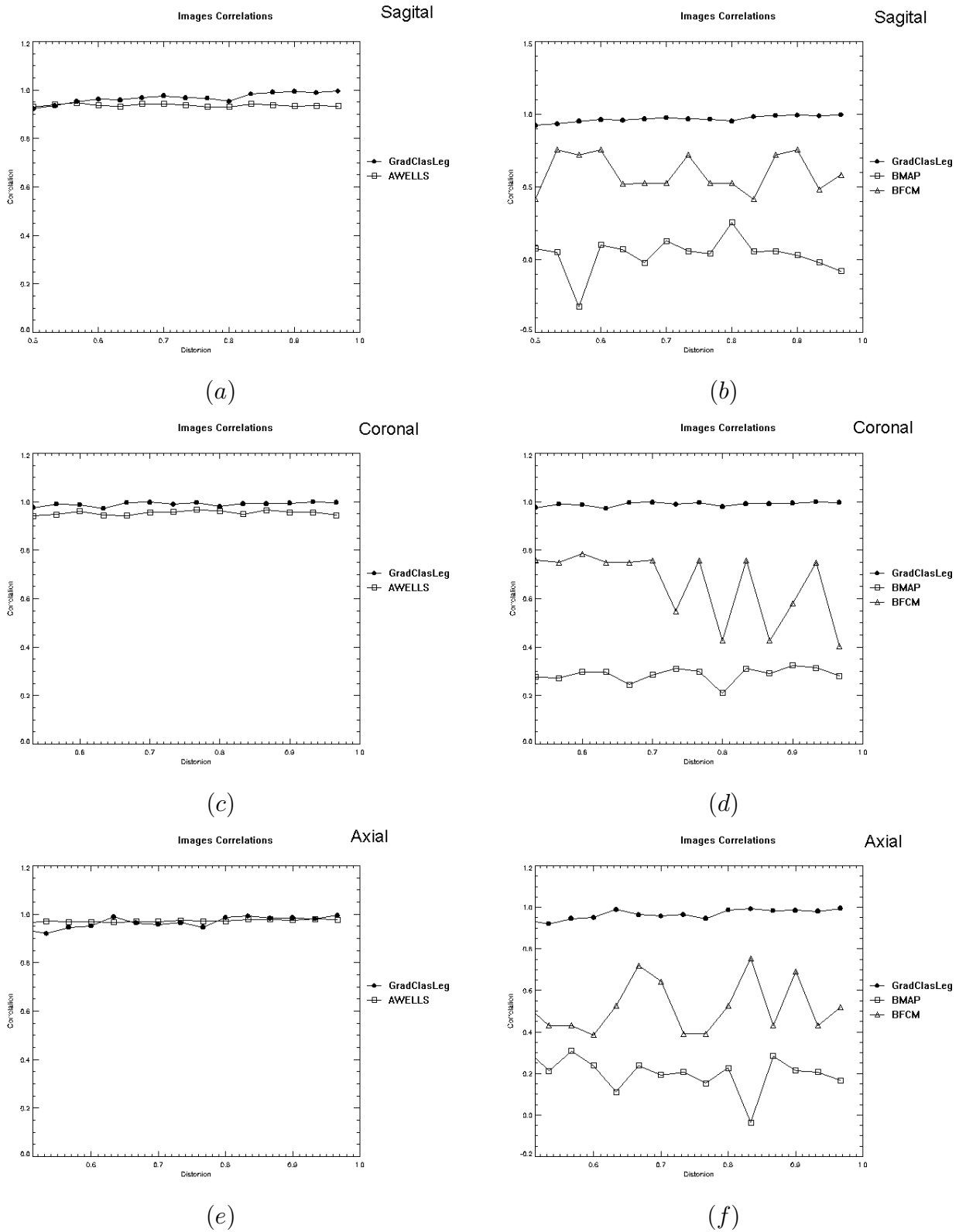


Figura 4.28: Correlación de las imágenes restauradas y la original sobre los cortes del volumen sintético de cerebro, GradClassLeg versus el algoritmo de Wells, BFCM y BMAP

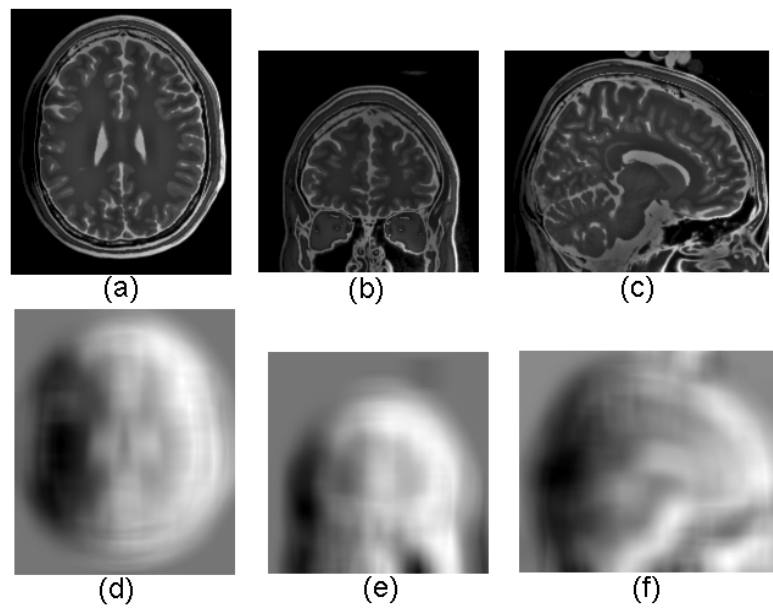


Figura 4.29: Resultado de la aplicación del algoritmo de Wells. (a,b,c) Imagenes corregidas(d,e,f) Sesgos estimados.

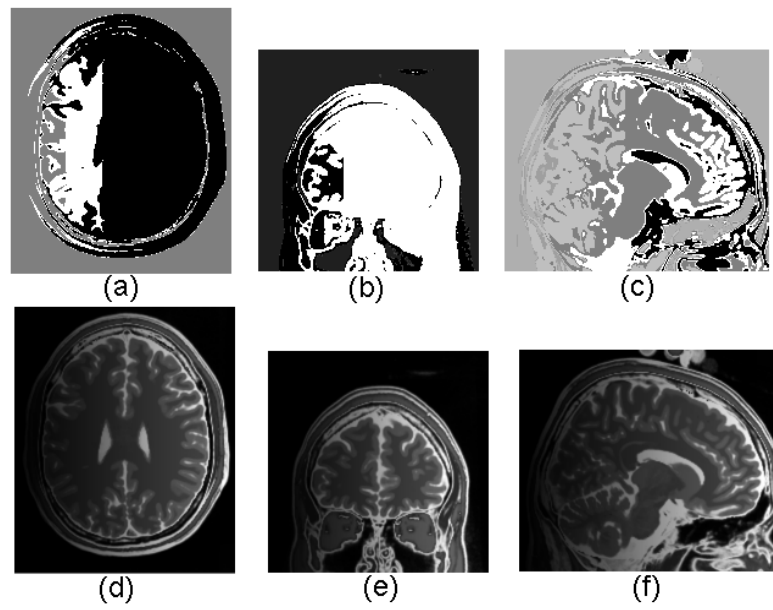


Figura 4.30: Resultado de la aplicación del algoritmo BFCM a las distintos rodajas del volumen sintético de cerebro corrompido con un campo sintético. (a, b, c) imágenes corregidas, (d, e, f) sesgos estimados.

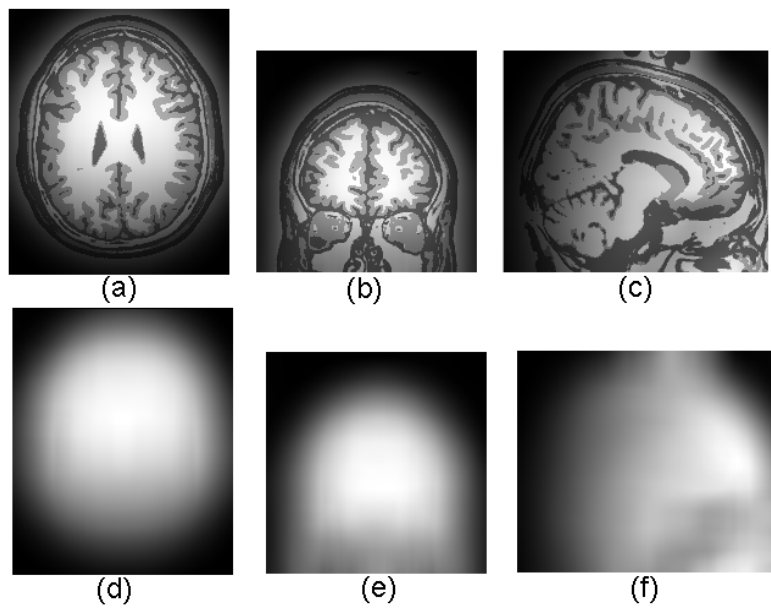


Figura 4.31: Resultados de la aplicación del algoritmo BMAP sobre las imágenes de la figura. (a, b, c) Imágenes corregidas. (d, e, f) Sesgos estimados.

4.3.3. Conclusiones sobre las imágenes de cerebro

Estas imágenes son más complejas que los tableros de ajedrez y producen algunos resultados contradictorios con los obtenidos en la sección anterior. Uno de ellos es la variabilidad de la correlación de los sesgos estimados con los originales. Además nos encontramos que la EE no mejoran ya al PABIC, posiblemente debido a que para las EE la restricción computacional impuesta fuerza una terminación prematura. Por otro lado, el algoritmo de Wells mejora considerablemente su rendimiento, acercándose al de los algoritmos GradPABIC y GradClassLeg. Estos últimos algoritmos mejoran sistemáticamente la calidad de los resultados obtenidos con los restantes algoritmos. Los algoritmos BFCM y BMAP presentan resultados bastante pobres, posiblemente debido a la forma en la que se estiman los sesgos de iluminación.

4.4. Resultados sobre imágenes del IBSR

Las imágenes que se han utilizado en experimento de validación extraídas del IBSR mantenido por el Massachusetts General Hospital, Center for Morphometric Analysis son imágenes pesadas en T1 de sujetos normales en las que se han aislado manualmente las regiones correspondientes a la materia gris, materia blanca y líquido cerebro espinal. El cráneo y resto del espacio ha sido eliminado de la imagen. Las imágenes se proporcionan como rodajas de 256×256 píxeles, 8 bits/píxel correspondientes a cortes coronales que se agrupan en volúmenes. El número de rodajas varía de una imagen a otra. Las segmentaciones se proporcionan como volúmenes en los que los valores de los píxeles son: no clase=0, csf=128 gray=192 white=254. Las inhomogeneidades de iluminación, si existen son desconocidas.

En esta sección además de proporcionar los resultados de los algoritmos, introducimos unos análisis de sensibilidad de los algoritmos, para intentar fijar la bondad de los resultados y su fiabilidad.

4.4.1. Resultados básicos

La única forma de validar los resultados de los algoritmos, en este caso, es la comparación de los resultados de la segmentación obtenidos por los algoritmos con las segmentaciones proporcionadas. Hay que tener en cuenta que estas segmentaciones son manuales y, por tanto, pueden contener errores.

Hemos probado GradClassLeg, algoritmo de Wells y BMAP sobre estas imágenes aplicandolos independientemente para cada rodaja, sin reducción de tamaño. Los algoritmos no han realizado reestimación de las medias de intensidad, han asumido los valores nominales de las imagenes de clasificación como los valores de las medias de las clases de intensidad.

Aunque no realizamos un proceso 3D propiamente dicho, sino una sucesión de procesos 2D independientes, en esta sección pasamos a tratar un volumen completo, lo cual supone un aumento considerable en la carga de cálculo necesario. Por ello resulta inviable realizar experimentos repetitivos y descartamos los algoritmos evolutivos debido a su caracter aleatorio y la baja repetibilidad de sus resultados. Esta es la razón por la que aplicamos GradClassLeg y no GradPABIC a este conjunto de datos. La razón para aplicar el algoritmo de Wells radica en los buenos resultados obtenidos con las imágenes de cerebro. Por último, aplicamos BMAP debido a que esperamos que el caracter local de las inhomogeneidades le favorezca.

En la figura 4.32 se muestran los resultados de los algoritmos sobre una de las imagenes del IBSR. Los resultados se presentan como promedio de aciertos en la clasificación de los pixeles en cada rodaja del volumen. Se advierte que existen fuertes variaciones, comunes a los tres algoritmos, de la tasa de aciertos de clasificación entre algunas rodajas vecinas. Esto es un indicio de una baja calidad de la segmentación manual. En todo caso es de resaltar que el GradClassLeg mejora en casi todas las rodajas a los otros dos algoritmos. Los parametros de los algoritmos utilizados son los siguientes:

- GradClassLeg: $\alpha = 0.01$, 25 pasos de iteracion, orden mínimo 1, orden máximo 5.
- BMAP: vecindario 2, potenciales del MRF que modela la probabilidad a priori $\alpha_1 = \alpha_2 = 0.5$, diametro de la máscara de suavización del sesgo 20,
- BFCM: vecindario de regularización de radio 2.
- Algoritmo de Wells: máscara de suavizado de diámetro 25.

Los resultados visuales sobre una de las rodajas (la número 32) se presentan en las figuras 4.33, 4.34, 4.35, 4.36. La figura 4.33a presenta la imagen de cerebro original, la figura 4.33b presenta la imagen de la segmentación manual de esta rodaja.

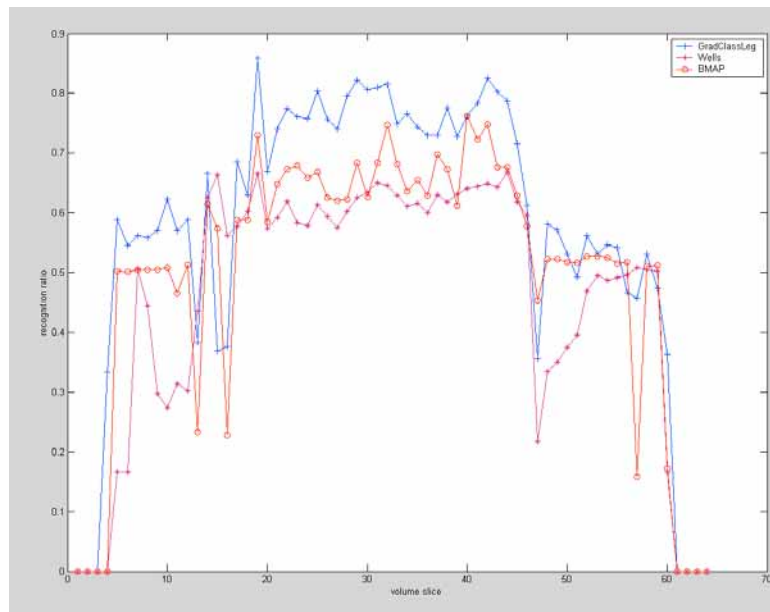


Figura 4.32: Resultados de aciertos de la clasificación de los píxeles en materia gris, materia blanca, líquido cerebro-espinal obtenidos por GradClassLeg, Wells y BMAP sobre cada rodaja.

En la imagen 4.34 se presentan los sesgos de iluminación estimados por los algoritmos. El sesgo estimado por GradClassLeg se presenta en 4.34a y tiene el caracter suave que es de esperar. Las figuras 4.34b y 4.34c presentan, respectivamente, los sesgos estimados por BMAP y el algoritmo de Wells. En ambos casos hemos invertido el sesgo para su visualización puesto que los valores eran inferiores a cero, de ahí que el fondo aparezca como blanco. Se aprecia en estas imágenes que el sesgo estimado por BMAP es muy localizado, mientras que el estimado por el algoritmo de Wells es mucho más difuso.

La imagen 4.35 presenta, en el mismo orden que la anterior, las imágenes corregidas por los algoritmos. Se aprecia un mayor contraste en la imagen recuperada con GradClassLeg respecto a los otros algoritmos.

La imagen 4.36 presenta, en el mismo orden que las anteriores, las imágenes de clasificación de la imagen corregida utilizando los valores nominales de las clases de intensidad. Se aprecia que la clasificación del GradClassLeg es la más consistente con la segmentación manual. Una de las clases, la primera, es muy poco frecuente y la fuente de la mayor parte de nuestros problemas respecto de la clasificación porque se pierde casi siempre.

4.4.2. Estudio de sensibilidad de GradClassLeg

Las condiciones en las que se prueban los algoritmos varían tremendamente al pasar de considerar imágenes sintéticas a imágenes reales con segmentaciones manuales como único referente de validez. Para asegurar que los valores presentados en la sección anterior son fiables y repetitivos, incluimos un pequeño estudio de la sensibilidad de los algoritmos.

Hemos realizado un estudio de sensibilidad del algoritmo GradClassLeg sobre una la imagen del IBSR empleada en el estudio comparativo. En la figura 4.37 se muestran los resultados de reconocimiento para las rodajas en el volumen cuando se varia $\alpha \in \{0.5, 0.1, 0.01\}$. No se aprecian cambios de importancia. Esto es, el algoritmo se comporta de forma similar para ganancias en un rango amplio de valores.

En la figura 4.38 se muestran los resultados al variar el numero de iteraciones que se realizan. Tampoco se aprecian cambios significativos, excepto en las rodajas de los extremos en las que se obtienen resultados muy distintos para 25 pasos. Estas rodajas son poco significativas, porque tienen pocos pixeles de interés en los que opera el algoritmo. Además una de las clases (líquido encefaloraquídeo) es muy poco frecuente o inexistente en ellas. De todas maneras, el

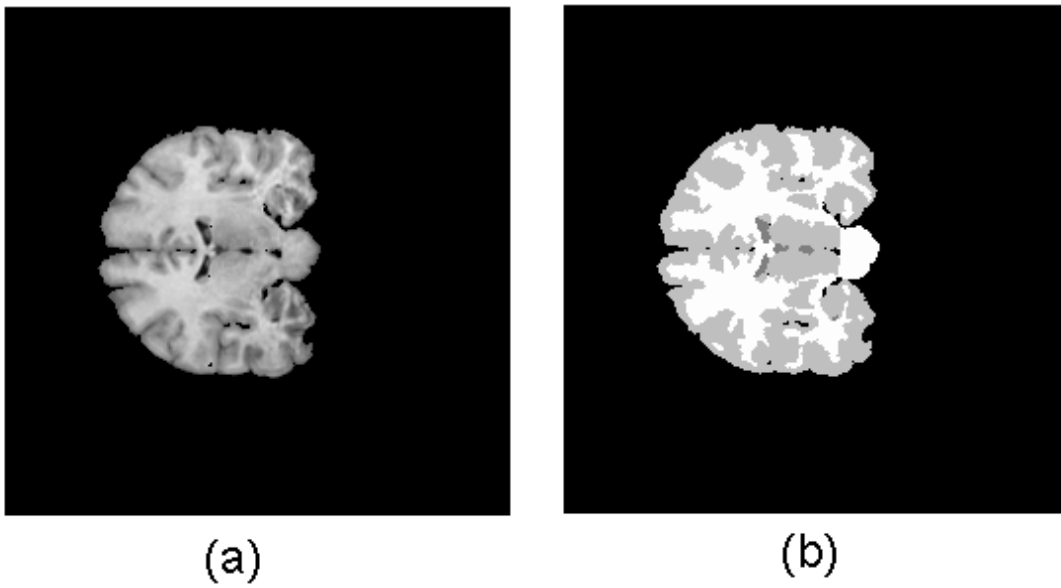


Figura 4.33: Una rodaja de (#32) del volumen empleado en los experimentos, (a) la imagen original con las materias no cerebrales removidas. (b) imagen de la segmentación manual donde cada pixel tiene el valor de intensidad de la clase asignada.

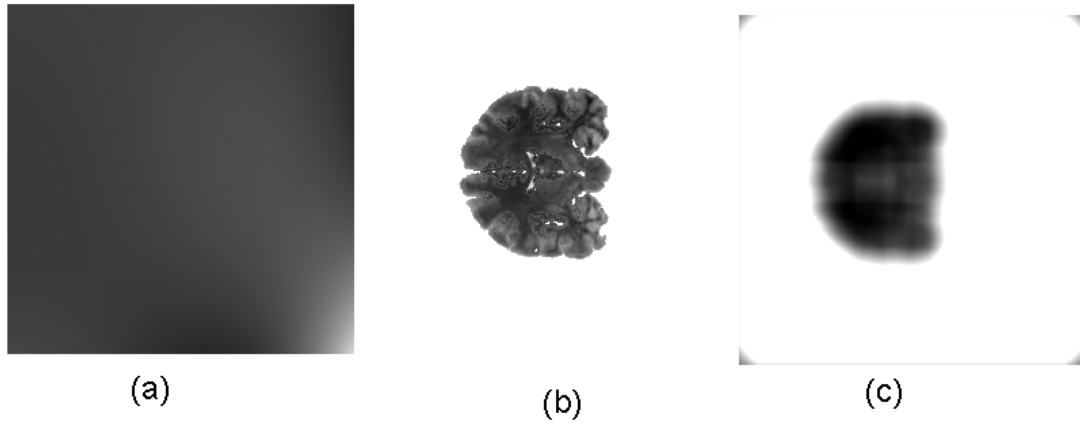


Figura 4.34: Bias detectados en la rodaja del volumen en la figura (a) Grad-ClassLeg, (b) BMAP, (c) Wells

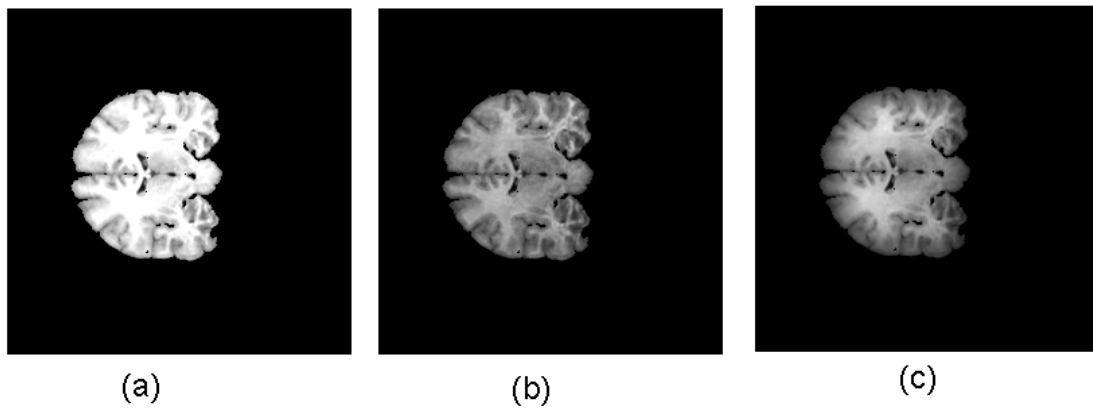


Figura 4.35: Imagen corregida de la rodaja #32 del volumen obtenido del IBSR. (a) GradClassLeg, (b) BMAP, (c) Wells.

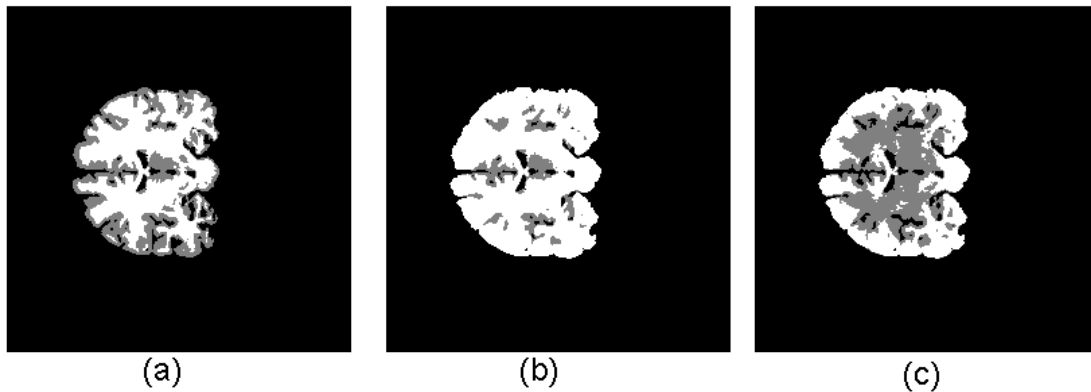


Figura 4.36: Imágenes de la clasificación obtenidas por los algoritmos con los valores nominales de intensidad correspondiente a cada clase.

resultado se puede interpretar en el sentido de que el número de pasos más adecuado es 25. El algoritmo degenera a partir de ese número de pasos, en ocasiones divergiendo en el error.

En la figura 4.39 se muestra el efecto que tiene sobre el reconocimiento la variación en el rango de orden que se utilizan para contruir los productos de polinomios de Legendre. En las rodajas centrales, que son las que contienen más información de interés, los órdenes de 1-4 y de 1-5, son los que dan mejor resultado. Tanto los ordenes bajos, 1-2 como los órdenes altos, 1-8 dan peores resultados, lo que se puede interpretar como que las variaciones del campo no son de tan suaves como algunos autores pretenden, ni tan bruscas como se sugiere en otros trabajos basados en la suavización del estimador del campo de iluminación.

4.4.3. Estudio de sensibilidad del algoritmo de Wells

En la figura 4.40 se muestra la sensibilidad del algoritmo de Wells al tamaño de la máscara de suavización sobre la imagen experimental. Se aprecia claramente que el aumento del tamaño de la máscara tiene un efecto definido y mejora el resultado del algoritmo. El tamaño que hemos asumido como standard (25×25) da buenos resultados, y las mejoras que produce un aumento de tamaño no justifican el aumento del costo de cálculo. En concreto, las mejoras más

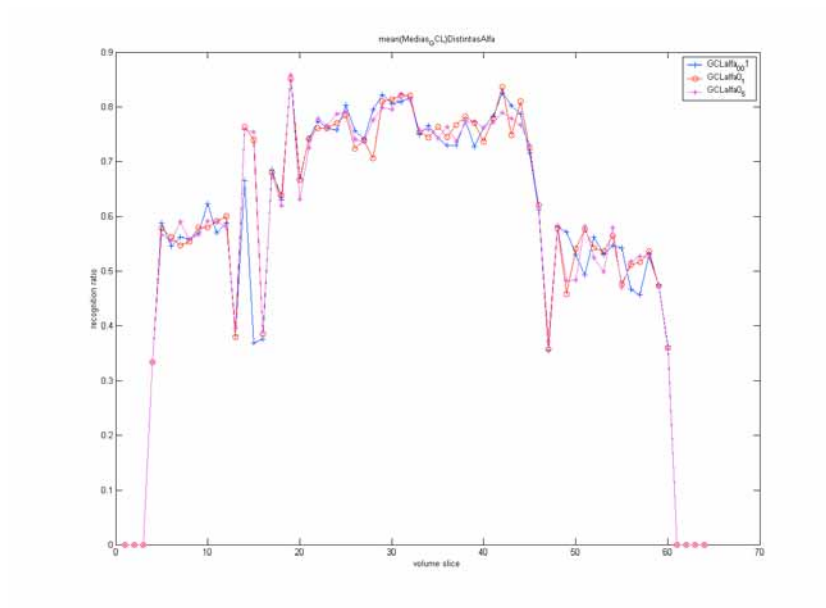


Figura 4.37: Sensibilidad del GradClassLeg a variaciones en la ganancia del gradiente, aciertos en las distintas rodajas del volumen.

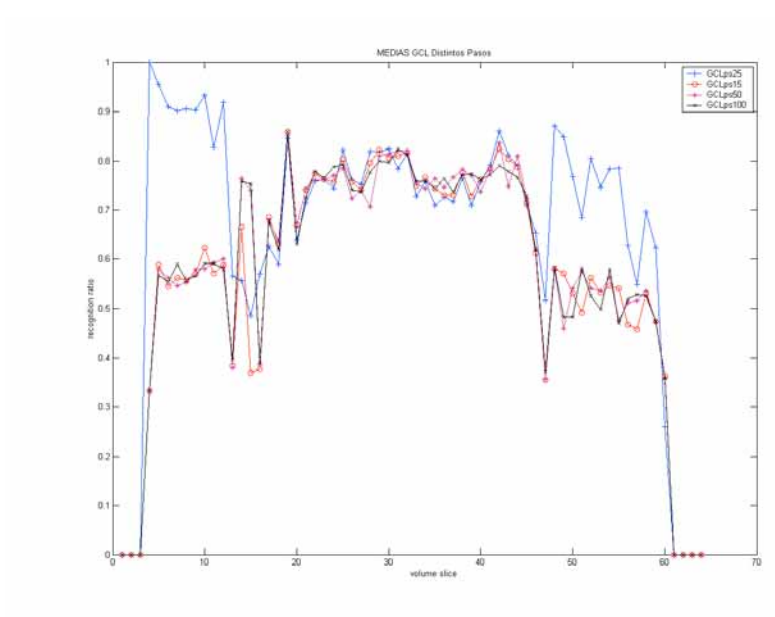


Figura 4.38: Sensibilidad de GradClassLeg para distintos números de iteraciones.

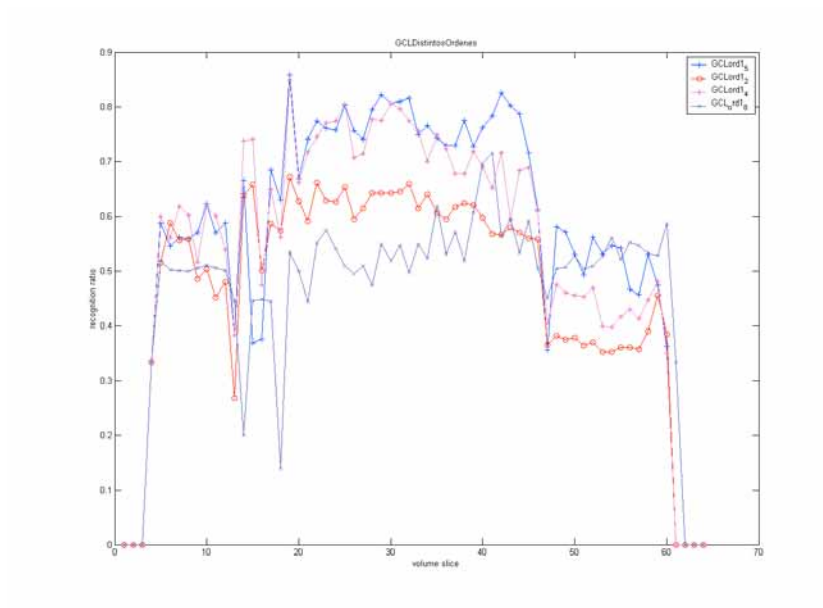


Figura 4.39: Sensibilidad de GradClassLeg a las variaciones en el orden de los productos 2D de polinomios de Legendre.

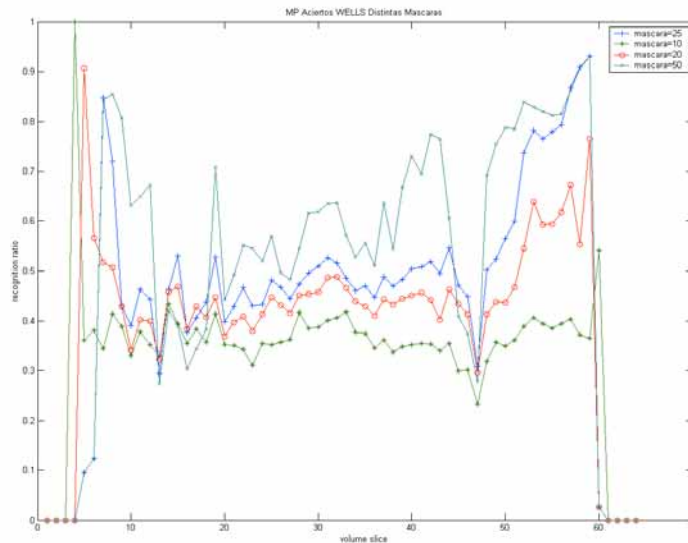


Figura 4.40: Efecto del tamaño del filtro pasabaja en los resultados obtenidos con el algoritmo de Wells.

espectaculares son en las rodajas de los extremos que tienen muy pocos píxeles de materia intracraneana.

4.4.4. Estudio de sensibilidad de BMAP

El algoritmo BMAP tiene como parámetros que pueden influir en su eficiencia el tamaño del vecindario, el tamaño de la máscara pasabaja empleada para regularizar el bias y el tamaño del vecindario. En esta sección vamos a comentar los resultados de este algoritmo cuando varían por separado los parámetros del algoritmo. En la figura 4.41 se muestra el efecto de variar el parámetro alfa, encontrándose que no es significativo. En la figura 4.42 se muestra el efecto del tamaño de vecindario, que no es tampoco un parámetro crítico. El aumento de vecindario produce algunas mejoras menores en algunas rodajas, pero es bastante inefectivo en general. Por último en la figura 4.43 se muestra el efecto de la variación del tamaño de la máscara de suavización aplicada a la estimación del campo de iluminación. Se observa que no es tan influyente como puede

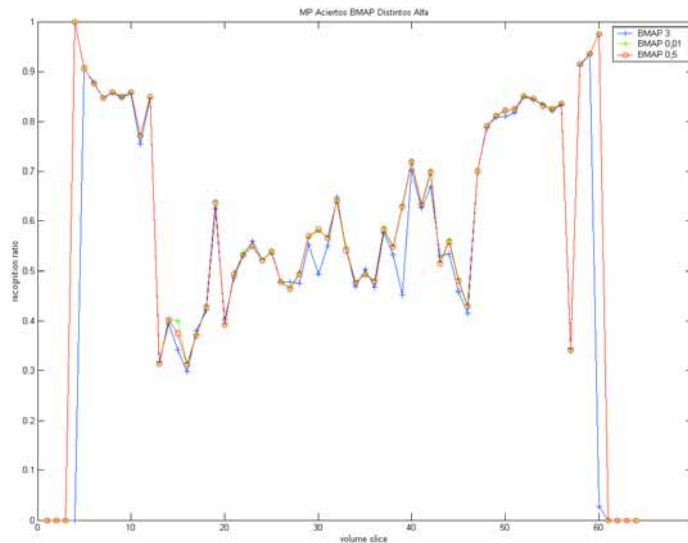


Figura 4.41: Efecto de la variación del parámetro de potencial en el algoritmo BMAP para la imagen del IBSR:

deducirse de los resultados presentados en la sección 4.2.4.

4.5. Discusión y conclusiones

En este capítulo hemos presentado resultados de experimentos computacionales de aplicación de los algoritmos descritos en el capítulo 3 han sido probados en un par de conjuntos de imágenes sintéticas, corrompidas por campos de iluminación también sintéticos conocidos. Este marco experimental es, a nuestro juicio el más adecuado para la evaluación de algoritmos de restauración de imágenes, y más concretamente para los algoritmos de corrección de la iluminación, puesto que permite una evaluación cuantitativa del algoritmo, mediante el cálculo de las correlaciones entre la imagen restaurada y la imagen original, así como las correlaciones entre los campos de iluminación estimados y los que actualmente se han aplicado a la imagen.

Considerando en conjunto los dos tipos de imágenes sintéticas que han servido de base a los experimentos podemos concluir lo siguiente:

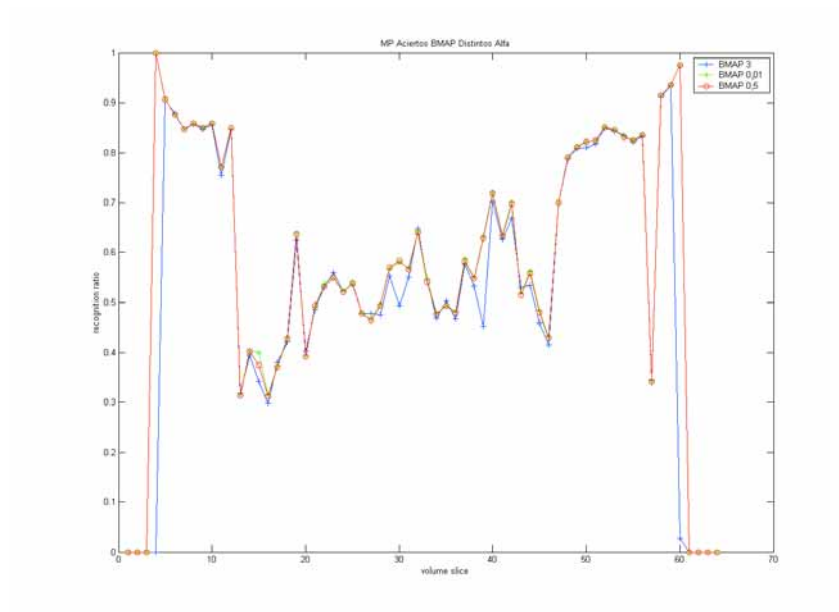


Figura 4.42: Estudio del efecto del cambio, para una imagen del IBSR, del radio del vecindario en BMAP.

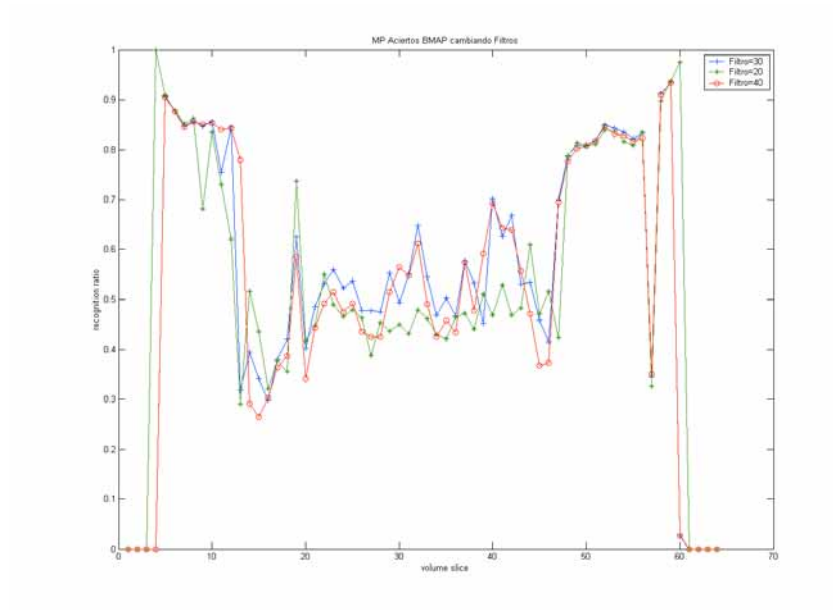


Figura 4.43: Sensibilidad del BMAP al tamaño de la máscara de convolución de la estimación del campo de iluminación.

- Los algoritmos estudiados mejoran en términos generales al filtrado homomórfico, que proporciona la “bottom-line” de los resultados elementales a superar por cualquier solución más sofisticada del problema. En esta medida, los algoritmos pueden considerarse como no triviales.
- Nuestras proposiciones, tanto GradPABIC como GradClassLeg, mejoran prácticamente en todos los casos a los restantes algoritmos extraídos de la literatura.
- Se produce una fuerte diferencia de comportamiento del algoritmo de Wells en las imágenes de cerebro respecto de las imágenes del tablero de ajedrez. Pensamos que esto se debe a que los restantes algoritmos se degradan más al considerar un número mayor de clases de intensidad.
- El algoritmo BMAP presenta unos pésimos resultados. Son precisamente estos resultados obtenidos con los ajustes de parámetros sugeridos en la referencia original los que nos indujeron a realizar los experimentos de sensibilidad que se comentan en la sección 4.2.4. Estos experimentos de sensibilidad apuntan a que la razón de la respuesta del BMAP sobre las imágenes experimentales sintéticas radica en el pequeño diámetro de la máscara de suavización del estimador del sesgo de iluminación.
- El algoritmo BFCM se comporta de forma poco aceptable, no sólo es sistemáticamente mejorado por nuestras proposiciones, sino que los resultados son malos tanto desde el punto de vista de la clasificación como de la corrección de la imagen. Posiblemente se debe a la ausencia de un filtro de suavización del estimador del sesgo de iluminación.
- El algoritmo PABIC adolece de los problemas que son de esperar en las estrategias evolutivas con tamaño reducido de población. Las estrategias evolutivas clásicas presentan resultados similares o mejores que PABIC con el mismo coste computacional y sus evoluciones indican que las EE todavía tienen posibilidad de mejorar los resultados al terminar la ejecución forzadas por la restricción computacional, mientras que PABIC se estanca rápidamente en mínimos locales.
- En la aplicación sobre imágenes reales extraídas del ISBR también GradClassLeg produce mejores resultados que el algoritmo de Wells y BMAP, si consideramos el resultado promedio sobre las rodajas.

- Sobre las imágenes de IBSR el comportamiento de BMAP es más aceptable que sobre las imágenes sintéticas, debido posiblemente a que las variaciones de iluminación son más locales.

5. LA CORRECCIÓN DE LA ILUMINACIÓN EN RECONOCIMIENTO DE CARAS

El problema de la iluminación también se presenta en dominios de procesado de la imagen lejanos a las imágenes médicas. En el caso del reconocimiento de caras para identificación biométrica, el efecto de la iluminación es de importancia para el rendimiento de los sistemas de clasificación. En este capítulo referimos la forma en que la aproximación desarrollada para las imágenes de resonancia magnética nuclear (IRMFN) puede extenderse a la mejora de sistemas de reconocimiento de caras. Los trabajos que presentamos están en una fase menos madura que en el caso de las IRM, pero los resultados son positivos y animan a la profundización en esta línea de investigación.

Comenzaremos con una revisión de algunos trabajos previos encontrados en la literatura, para continuar exponiendo la aproximación que proponemos y los resultados experimentales obtenidos hasta el momento.

5.1. Antecedentes

En tiempos recientes, las nuevas bases de datos experimentales diseñadas para testar y desarrollar nuevos algoritmos de reconocimiento de caras, tienen en cuenta las dificultades que introducen las variaciones de iluminación, así como las poses y expresiones faciales. Una de las primeras versiones de bases de datos experimentales con estas restricciones en mente es la de primera de Yale [124], a la que siguió una base de datos mucho más extensa y detallada [125], [51] construida con 10 sujetos en 9 poses y bajo 64 condiciones de iluminación conseguidas con un complejo aparato de luces y cámaras. Similar diseño tiene la base de datos de la CMU [110] entre otras. Este interés proviene del reconocimiento del hecho de que las variaciones en pose e iluminación pueden introducir variaciones en la imagen superiores a las variaciones entre sujetos. La aproximación clásica a la construcción de métodos y algoritmos de reconocimiento de caras

(ejemplos de algoritmos recientes son [13], [26],[71],[82] , [83], [84], [85], [86], [88]) no tienen en cuenta la iluminación en la construcción del clasificador, simplemente esperan que el algoritmo propuesto sea suficientemente robusto como para que no le afecten demasiado los cambios de iluminación. Es posible que esta forma de pensar sea incorrecta fundamentalmente. En [2] se mostraba como una variedad de aproximaciones a la extracción de características (i.e. filtros de Gabor) para el reconocimiento de caras supuestamente robustas frente a cambios de iluminación resultan estrepitosamente frágiles.

Los métodos de reconocimiento de caras basados en la geometría de las características de la cara son potencialmente insensibles a cambios de iluminación, siempre que la detección de esas características en la imagen sea un proceso robusto. Sin embargo, este proceso es muy poco robusto en general y sensible a su vez a las variaciones que introducen los cambios de iluminación. Por estas razones ha ganado aceptación general los métodos basados en la apariencia, esto es, métodos basados en representaciones de baja dimensión de la imagen obtenidas mediante transformaciones (muchas veces lineales) de la imagen. Un inconveniente de estos métodos basados en la apariencia es que para que se produzca el reconocimiento de forma fiable de un objeto bajo una iluminación precisa es preciso que haya sido visto en circunstancias similares.

En [51] se propone un método generativo basado en apariencias que se basa en las siguientes observaciones. El conjunto de imágenes de un objeto en una pose fija bajo todas las posibles condiciones de iluminación es un cono convexo en el espacio de las imágenes: el cono de iluminación. Cuando la reflectancia es aproximadamente Lambertiana, el cono de iluminación se puede construir a partir de un conjunto reducido de imágenes. El cono de iluminación se puede aproximar bien por un subespacio lineal. Bajo pose e iluminación variables, el conjunto de imágenes se puede caracterizar por una familia de conos de iluminación parametrizados por la pose. A partir del cono de iluminación se pueden sintetizar las imágenes de las caras bajo vistas no conocidas previamente. El cono de iluminación se calcula mediante un proceso de estéreo fotométrico a partir de un conjunto de imágenes con distintas iluminaciones. El proceso de reconocimiento lleva consigo la síntesis mediante rendering de las posibles vistas para su comparación con la imagen a clasificar.

En [108] se propone un cociente de imágenes como la forma de obtener una representación invariante a cambios de iluminación del objeto. El proceso necesita un conjunto bootstrap de imágenes que sirve de base de representación de los posibles albedos de las superficies. Se realiza un ajuste de mínimos cuadra-

dos para extraer los coeficientes que representan la iluminación respecto de las tres fuentes de iluminación que forman la base de un espacio de vectores de iluminación. A partir de estos coeficientes se obtiene la imagen cociente de la imagen dada respecto de las imágenes bootstrap que sirve de representación invariante para los procesos de reconocimiento. En [128] se propone un método simplificado basado en la idea de que el ratio entre una imagen y la representante canónica del individuo es una caracterización de la superficie de la cara. A partir de esta idea se propone un método basado en los cocientes de imágenes para generar vistas sintéticas de un cara con distintas iluminaciones.

En [75] se realiza el seguimiento robusto (*tracking*) de una cara utilizando un modelo cilíndrico sobre el que se mapea la textura de una cara, el seguimiento se hace por mínimo error cuadrático entre la imagen real y la imagen renderizada del modelo variando su pose. Para tomar en cuenta las diferencias de iluminación se estima la dirección de la iluminación en la imagen y se simula en el modelo antes de calcular el error de seguimiento. En [92] se estudia el problema de la invarianza de color para el tracking de caras en base a la detección del color, proponiendo una restricción de color invariante a la iluminación y robusta frente a la calibración incorrecta de la cámara. Es digno de mención por su originalidad el trabajo de [97] se propone una técnica de reconstrucción 3D de caras a partir de imágenes 2D mediante el reconocimiento de partes de la cara y su registro mediante técnicas de análisis en componentes principales. El reconocimiento invariante a iluminación de las partes de la cara se consigue mediante una colección de plantillas (*templates*). La aproximación se denomina forma a partir del aprendizaje (*shape from learning*).

5.2.Planteamiento del problema

La aproximación que tomamos en este capítulo consiste en tratar de estimar el campo de iluminación y obtener una imagen corregida de la cara sobre la que realizar el proceso de reconocimiento. En nuestra aproximación no necesitamos un conjunto de imágenes de bootstrap para estimar la dirección de la iluminación. Sea $\Omega = \{\omega_1, \dots, \omega_c\}$ el conjunto de individuos a reconocer de los cuales disponemos de imágenes de referencia (entrenamiento) $R = \{r_{ij}(x, y); i = 1, \dots, c; j = 1, \dots, n_c\}$ donde r_{ij} es la j -ésima imagen del i -ésimo individuo. Si r_{i0} es la imagen obtenida con la iluminación de referencia, la corrección de la iluminación para cualquier nueva imagen $f(x, y)$ obtenida para

este individuo puede formularse como el siguiente problema de minimización:

$$\min_{\beta} e(f, \beta, r_{i0}) \quad (5.1)$$

donde β corresponde al campo de iluminación presente en la imagen f producto de la diferente iluminación respecto de la imagen de referencia r_{i0} . El problema que plantea esta formulación es la necesidad de determinar la clase a la que pertenece la imagen previamente a la realización de la corrección de la iluminación. Para evitar este problema podemos considerar una imagen de referencia genérica independiente del individuo. La elección más natural para esta imagen es el promedio de las imágenes de referencia de los individuos:

$$r^*(x, y) = \frac{1}{c} \sum_{i=1}^c r_{i0}. \quad (5.2)$$

La imagen de referencia $r^*(x, y)$ nos da la distribución de niveles de gris que modela la superficie del objeto genérico bajo la iluminación estándar (usualmente frontal). El problema de la corrección de la iluminación se formula, por tanto, como la minimización

$$\min_{\beta} e(f, \beta, r^*) \quad (5.3)$$

donde β corresponde al campo de iluminación presente en la imagen f que hace que se aleje de la imagen de referencia $r^*(x, y)$. Si el bias de iluminación β tiene una forma paramétrica, como la combinación lineal de productos de polinomios de Legendre, la función de error dependerá explícitamente del vector \mathbf{p} de parámetros del bias de iluminación. Si mantenemos el modelo multiplicativo de formación de la imagen dado por la ecuación 2.1 la función de error toma la forma:

$$e(f, \mathbf{p}, r^*) = \sum_{x,y} \left(\frac{f(x, y)}{\widehat{b}(x, y)} - r^*(x, y) \right)^2. \quad (5.4)$$

La resolución del problema dado por la ecuación 5.3 la realizaremos siguiendo el algoritmo de descenso de gradiente [121] para la estimación de los parámetros de la combinación lineal de productos de polinomios de Legendre. Esto es, nuestro algoritmo tiene la forma

$$\mathbf{p}(t) = \mathbf{p}(t-1) + \Delta_{\mathbf{p}} e, \quad (5.5)$$

donde $\Delta_{\mathbf{p}}e$ es el gradiente del error de corrección de la iluminación respecto de los parámetros del modelo del bias.

$$\Delta_{\mathbf{p}}e(\mathbf{p}, r^*(x, y)) = \left\{ \frac{\partial}{\partial p_{ij}} e(\mathbf{p}, r^*(x, y)) \right\}, \quad (5.6)$$

donde

$$\frac{\partial}{\partial p_{ij}} e(\mathbf{p}, r^*(x, y)) = 2 \sum_{x,y} \left(\frac{f(x, y)}{\widehat{b}(x, y)} - r^*(x, y) \right) \frac{-f(x, y) P_i(x) P_j(x)}{\widehat{b}^2(x, y)}. \quad (5.7)$$

5.3. Resultados experimentales en reconocimiento de caras

La evaluación de los sistemas de reconocimiento está habitualmente circunscrita a su aplicación a conjuntos de entrenamiento y validación (test) extraídos de bases de datos disponibles públicamente (p.e.: [35], [36], [37], [38], [39], [40], [41], [124], [125]). Muchas de estas bases de datos no están sistemáticamente organizadas para probar la invarianza a iluminación, pose o expresión. La base de datos de Yale inicial [124] tomaba en cuenta explícitamente la variación de iluminación, así como variaciones elementales de expresión. En nuestros trabajos en esta memoria nos hemos concentrado en esta base de datos por que su pequeño tamaño permite la realización rápida de experimentos exhaustivos.

El algoritmo de reconocimiento de caras que aplicamos sigue el siguiente esquema:

1. Calculamos a partir de las imágenes en la base de datos la transformación en componentes principales. El número de autovectores (caras propias o eigenfaces) que consideramos es el máximo posible, esto es, tantos como imágenes en la base de datos.
2. Realizamos la transformación de todas las imágenes de la base de datos en sus coeficientes de componentes principales.
3. Seleccionamos un conjunto de entrenamiento con un muestreo aleatorio idéntico para cada individuo.
4. Como clasificador para los experimentos de reconocimiento utilizamos la técnica del vecino más cercano (*Nearest Neighbor, NN*).

La razón de utilizar la transformación en componentes principales y la búsqueda por el vecino más cercano es que son, respectivamente, el método de extracción de características más extendido y el algoritmo de clasificación más sencillo y de fácil comprensión. Puesto que nuestro interés está en evaluar el impacto sobre el sistema de reconocimiento de la corrección previa de la imagen, escogemos los mecanismos más simples para evitar la confusión de efectos y causas en el rendimiento final. El experimento completo consiste en aplicar el esquema computacional y de cálculo precedente tanto a las imágenes originales como a las imágenes después de la corrección de la iluminación.

La figura 5.1 muestra las imágenes de uno de los sujetos utilizados en el experimento antes de la corrección. La imagen con expresión neutra e iluminación frontal es la cara de referencia para este sujeto (tercera en la fila inferior). Entre las restantes imágenes dos muestran iluminaciones extremas desde la izquierda y la derecha del sujeto, sin embargo se pueden apreciar variaciones de intensidad en las restantes imágenes del individuo que pueden deberse a movimientos incontrolados del sujeto o a momentos distintos de captura de la imagen. Para evitar introducir diferencias debidas únicamente al fondo de la imagen, se ha realizado un recorte de la imagen original, sin atender con precisión al alineamiento de las imágenes. En la figura 5.2 se muestran las diferencias de las distintas imágenes del individuo en la figura 5.1 respecto de la cara de referencia, con la iluminación frontal y expresión neutra. Se aprecian algunas diferencias debidas a la imprecisión de la alineación de la cara en las distintas imágenes, sobre todo en los contornos de la cabeza. Pero también se aprecian, diferencias significativas en muchas de las imágenes en la zona de los pómulos que son debidas a variaciones de la iluminación incontroladas. Muchas de las diferencias son también debidas a los cambios de expresión. Las diferencias más notables, sin duda, se encuentran en las imágenes con iluminaciones extremas.

Como ya hemos indicado, la corrección respecto de la imagen de referencia de un individuo implica el reconocimiento previo de este individuo, algo que no encaja con nuestro propósito de mejorar precisamente el sistema de reconocimiento mediante la corrección de la iluminación. La imagen de referencia es la media de las imágenes de referencia de los individuos y se muestra en la figura 5.3. Esta figura viene a ser como la caracterización de la superficie de una cara a partir de un conjunto de referencia.

Para decidir el valor de los parámetros involucrados en el algoritmo de corrección, hemos realizado un cuidadoso estudio experimental sobre el efecto de los parámetros del algoritmo en los resultados de la corrección y la velocidad

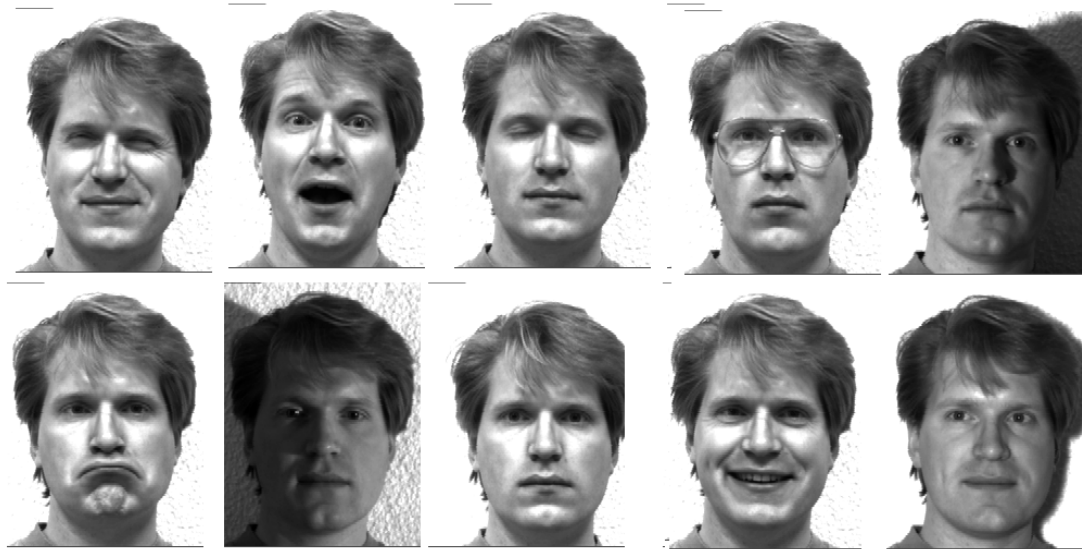


Figura 5.1: Imágenes originales de uno de los sujetos de la colección de Yale



Figura 5.2: Diferencias respecto de la cara de referencia para un individuo.



Figura 5.3: Cara de referencia utilizada en la corrección de la iluminación

de convergencia del algoritmo. En la figura 5.4 se muestra como afecta el valor del parámetro de ganancia en el descenso de gradiente al error final promedio obtenido de la iteración. A partir de los datos de estos experimentos se fijó un valor $\alpha = 0.013$ como valor óptimo de la ganancia del algoritmo de entrenamiento. Así mismo, se estudió el número de iteraciones más adecuada, encontrando que basta con 12 iteraciones para alcanzar el estado estable del error. En la figura 5.5 se presenta la evolución del error promedio final en función del orden máximo de los polinomios de Legendre. Un orden máximo 11 parece adecuado para el problema.

La figura 5.6 muestra las imágenes de las caras después de la corrección mediante la división por los bias estimados. El algoritmo utiliza productos de polinomios de Legendre hasta orden 9, con 100 iteraciones, $\alpha = 0.013$.

Una vez decididos los parámetros del algoritmo de corrección, realizada la corrección y transformadas las imágenes de la base de datos en sus coeficientes principales sobre las caras propias calculadas sobre las imágenes corregidas, hemos realizado el experimento computacional de reconocimiento variando el tamaño del conjunto de entrenamiento, seleccionando de forma aleatoria, manteniendo la distribución *a priori* de los individuos, imágenes para entrenamiento. Cada experimento de entrenamiento se ha repetido 100 veces para obtener un resultado suave. El resultado experimental conseguido se sintetiza en la figura 5.7. En ella se muestra la evolución del reconocimiento sobre la base de datos de Yale básica [124] conforme se incrementa el conjunto de entrenamiento. Se observa

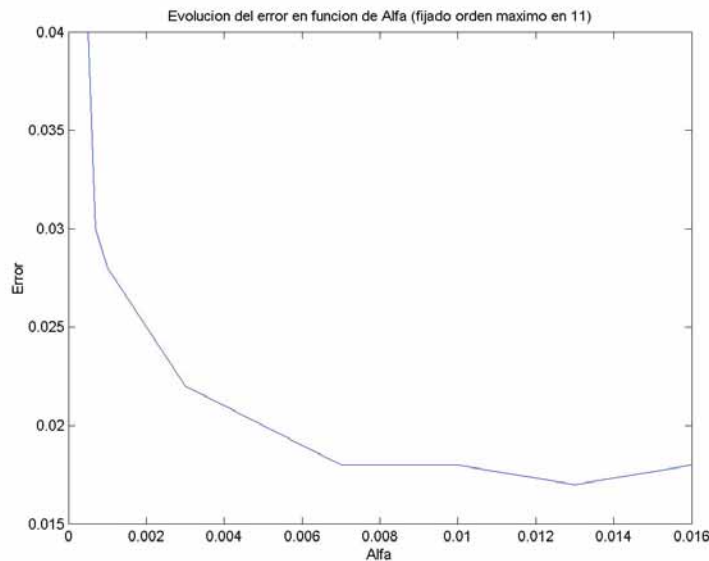


Figura 5.4: Error de restauracion versus la velocidad de descenso de gradiente.

una mejora sistemática de los resultados cuando el experimento de reconcimiento se realiza sobre las imagenes corregidas por el algoritmo descrito en la seccion 5.2.

5.4.Discusion

El trabajo relatado en este capítulo representa una fase muy preliminar de los trabajo que pueden realizarse en el reconocimiento automático. En esta fase, hemos realizado una extrapolación del método desarrollado en la sección 3.5.3 para la corrección de la iluminación en imagenes de caras y la hemos testado sobre la base de datos [124]. Variando el conjunto de entrenamiento controlamos la cantidad de información de que dispone el sistema de reconocimiento. Los resultados son positivos, ya que se obtiene una mejora sistemática del 5% aproximadamente a pesar de que no se tiene en cuenta la variabilidad en la imagen debida a cambios de expresión.

Entre los trabajos futuros consideramos de interés desarrollar métodos más eficientes de cálculo de las estimación del bias, así como métodos capaces de

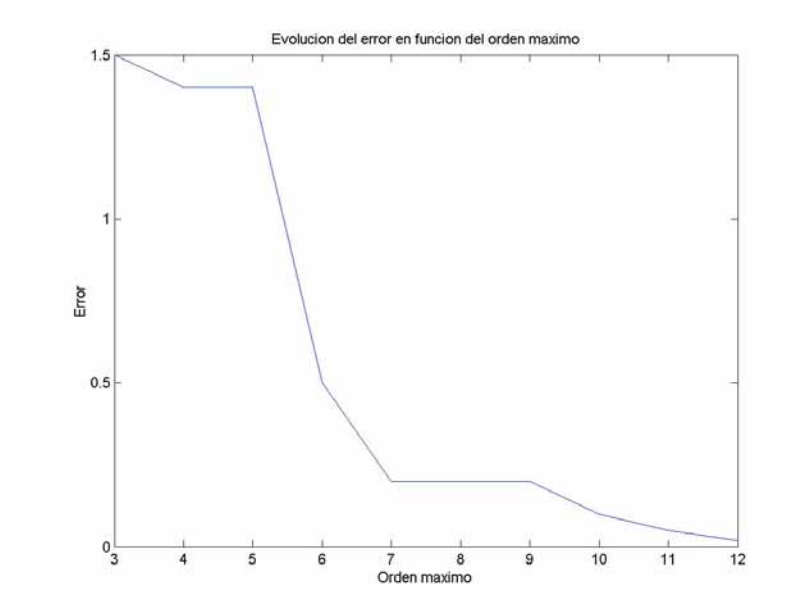


Figura 5.5: Error final de la corrección versus el orden maximo de los polinomios de Legendre.

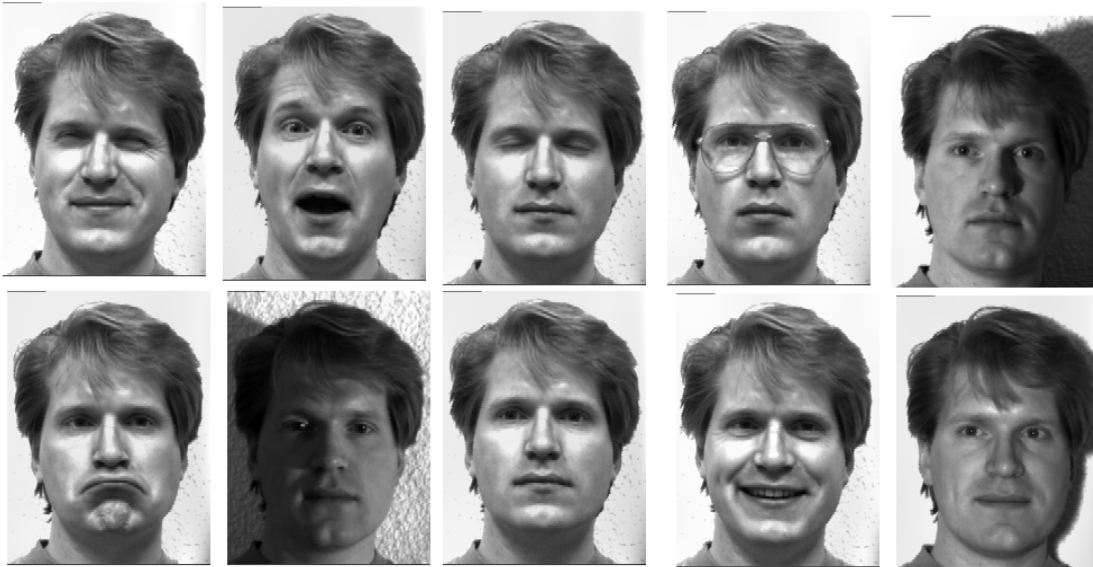


Figura 5.6: Imágenes del sujeto tras la corrección de la iluminación.

determinar el orden de los polinomios de forma automatizada, podando algunos coeficientes de la combinación lineal para aumentar la precisión y aliviar la carga computacional. Parece evidente que la fuerza de la corrección no puede ser la misma en todas las imágenes. En este sentido hemos encontrado que el nivel de error al terminar la ejecución para las imágenes con iluminaciones extremas (que son las que motivan el trabajo) es muy alto en relación al obtenido en otras imágenes, lo que es indicativo de la posibilidad de mejorar los resultados si se continúa la estimación para estas imágenes, o de la necesidad de introducir polinomios de más alto orden.

Los trabajos no tienen en cuenta la pose o la expresión. Consideramos que se pueden incluir o encastrar nuestro método en los algoritmos de reconocimiento de expresión o de pose, con el objeto de contribuir a mejorar sus resultados.

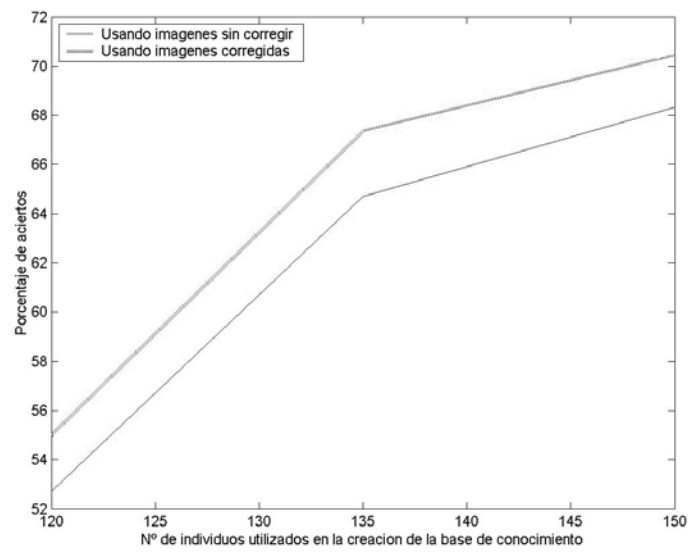


Figura 5.7: Resultados de clasificación sobre las imágenes de caras del conjunto de Yale, antes y después de la corrección

6. CONCLUSIONES Y TRABAJO FUTURO

Presentamos un breve resumen del contenido de la memoria, las contribuciones más relevantes y las líneas de trabajo futuro que podrían derivarse de la presente tesis doctoral.

Resumen

En la presente memoria de tesis doctoral hemos explorado el problema de la corrección de las inhomogeneidades de iluminación en las imágenes digitales. Hemos centrado nuestra atención en dos campos de aplicación: la imagen de resonancia magnética y el reconocimiento de caras.

Existe una extensa literatura sobre la corrección de iluminación en IMR, por lo que hemos realizado un extenso estudio del estado del arte, concentrados en cuatro algoritmos que hemos implementado y probado. Incluimos una detallada descripción de estos algoritmos. Estos algoritmos caen dentro de las dos categorías principales: métodos paramétricos y métodos no paramétricos. Nos hemos decantado por los métodos paramétricos a la hora de realizar nuevas proposiciones de algoritmos.

Hemos realizado un conjunto exhaustivo de experimentos computacionales para determinar la validez de los algoritmos propuestos, contrastándolos con los resultados obtenidos con las implementaciones de los algoritmos seleccionados entre los descritos en la literatura.

Los resultados positivos obtenidos sobre las imágenes de resonancia magnética nos han impulsado a probar la extensión de los métodos computacionales en otro paradigma radicalmente distinto, el reconocimiento de caras, para el cual hemos adaptado con relativo éxito los métodos de corrección paramétrica propuestos.

Contribuciones

Hemos implementado varios algoritmos encontrados en la literatura para la segmentación de la imagen de resonancia magnética basada en la clasificación de sus píxeles, que incluyen explícitamente la estimación del campo de iluminación para contrastar sus resultados con los obtenidos por nuestras proposiciones de algoritmos paramétricos.

Hemos realizado un análisis de la sensibilidad del algoritmo BMAP, mostrando su sensibilidad relativa al filtro de suavización aplicado sobre el estimador del sesgo de iluminación.

Hemos propuesto y aplicado un algoritmo de descenso de gradiente para la estimación tanto de las medias de las clases de intensidad, como del sesgo del campo de iluminación.

Hemos propuesto un algoritmo memético instantáneo para la estimación del campo de iluminación mediante la minimización del error de clasificación y restauración.

Hemos realizado un exhaustivo estudio empírico de los algoritmos sobre imágenes sintéticas e imágenes obtenidas del repositorio ISBR.

Las imágenes sintéticas nos han permitido realizar una cuidadosa evaluación cuantitativa de los algoritmos en base a la correlación entre las imágenes restauradas y las imágenes “limpias”. Las imágenes sintéticas representan el caso peor debido a la gran variación de intensidad que introducen los bias sintéticos.

Hemos propuesto un algoritmo paramétrico de corrección de la iluminación para imágenes de caras que consigue mejorar los resultados de clasificación de las caras de forma sistemática.

Conclusiones

A partir de la experimentación con imágenes sintéticas se deduce que nuestras proposiciones tienen igual o mejor respuesta que los algoritmos en la literatura. La experimentación con volúmenes reales extraídos del IBSR confirma nuestra confianza en el algoritmo de descenso de gradiente GradClassPabic.

En relación al reconocimiento de caras hemos aplicado nuestra aproximación consiguiendo una mejora apreciable de los resultados de validación. Dadas las diferencias entre las imágenes de resonancia magnética y las de reconocimiento de caras es necesario adaptar la formulación del problema, introduciendo una imagen de referencia, en lugar de las medias de las clases de intensidad en la imagen.

Líneas de trabajo futuro

Uno de los problemas típicos en todos los sistemas automáticos, ya sean de clasificación o de aproximación, es determinar el grado del sistema. En concreto, en nuestros trabajos la decisión de los órdenes de los polinomios de Legendre utilizados es una decisión tomada a priori y que no se ve modificada durante el proceso de estimación del campo de iluminación. Sería deseable desarrollar técnicas capaces de detectar si el orden de los polinomios es el adecuado y si hay órdenes redundantes que no aportan nada a la solución buscada. El proceso sería uno similar a los de poda realizados sobre sistemas de entrenamiento automático (redes neuronales, árboles de clasificación, etc) pero aplicado a la estructura de nuestro estimador del campo de iluminación.

Sería deseable también adaptar el paso de descenso de gradiente de acuerdo a la evolución del error, para acelerar la convergencia y evitar oscilaciones indeseadas. En nuestras aplicaciones experimentales hemos optado por realizar pocas iteraciones con un paso de descenso relativamente grande, sobre todo por consideraciones computacionales, pero un método fundamentado de ajuste del paso y control de parada podría mejorar los resultados obtenidos.

Los trabajos se han realizado sobre todo con imágenes de cerebro, sería deseable realizar estos experimentos sobre imágenes de otras partes del cuerpo que pueden presentar distintas inhomogeneidades y mayor número de clases de tejidos.

El problema del número de clases de tejidos es un problema similar al del orden de los polinomios, nosotros hemos fijado a priori este número en todos los casos, pero sería deseable que el algoritmo detectase el número real de clases de intensidad. Este es un problema que se ha aproximado desde muy diversas perspectivas en los algoritmos de agrupamiento y que no está resuelto de forma universal. En la IRM toma una forma especial, dado que en general se conocen los tejidos que están presentes en la imagen, salvo las patologías que precisamente son las que interesa detectar en muchos casos. Dado que las diferencias entre las respuestas de RM de los tejidos patógenos y los sanos pueden ser muy similares, el problema de distinguir nuevos tejidos inesperados de sesgos de iluminación no es un problema trivial.

Lo que muestran los trabajos experimentales en esta tesis es que la adopción de un modelo paramétrico del campo de iluminación puede favorecer considerablemente a los algoritmos y procesos de estimación. En realidad, los algoritmos propuestos son una combinación de métodos de agrupamiento y de estimación del sesgo de iluminación. Pensamos que esta aproximación puede combinarse

con otras formas de segmentación como el agrupamiento borroso o los métodos bayesianos, dando lugar a algoritmos con nuevas potencialidades y propiedades.

En lo que concierne al reconocimiento de caras invariante a cambios de iluminación, quedan todavía por realizar experimentos con base de datos más extensas que la utilizada en los trabajos descritos en esta memoria. Esperamos que estos trabajos, confirmen la utilidad de la aproximación propuesta. Pensamos que los métodos de ajuste automático de los órdenes de los polinomios utilizados pueden ser de gran utilidad en este problema concreto, dado que las sombras provocadas por los cambios de orientación de la iluminación corresponden en general a polinomios de alto orden y es difícil afinar a priori en el ajuste de estos órdenes.

7. APENDICE 1: FUNDAMENTOS DE LA IMAGEN DE RESONANCIA MAGNÉTICA NUCLEAR

Las fuente más inmediata de la que hemos extraído los fundamentos expuestos en este apéndice es [63], pero pueden encontrarse en otros libros de consulta como [27], [34], [59], [80], [101], [107]. En este apéndice recogemos la recapitulación resumida de los principios y los aparatos que permiten la adquisición de las imágenes de resonancia magnética nuclear. Esta recopilación está dirigida a comprender la fuente de las inhomogeneidades en la iluminación que se pretende corregir.

7.1. Nociones físicas básicas

La visualización de resonancia magnética nuclear (RMN) es una modalidad de visualización tomográfica, que produce una imagen de una rodaja de un volumen de un objeto. Los elementos de resolución del volumen se denominan voxels, mientras que los elementos de la imagen discreta producida se denominan pixels. La visualización se basa en la absorción y emisión de energía en el rango del espectro que corresponde a la radio-frecuencia. Dado que no se pueden visualizar objetos de tamaño menor que la longitud de onda de la energía utilizada para la visualización, la RMN es posible mediante variaciones espaciales de fase y frecuencia de la radio frecuencia absorbida y emitida por el objeto visualizado. El cuerpo humano contiene un 63% de átomos de hidrógeno (grasa y agua) cuya señal de RMN es la que se visualiza. El único protón en el núcleo del átomo de hidrogeno tiene espín, lo que permite conceptuarlo como un imán y causa que el núcleo genere una señal de RMN. El spin es una propiedad de la naturaleza, como la carga o la masa, viene en múltiplos de $1/2$ y puede ser positivo o negativo. Los electrones, protones y neutrones tienen spin de valor $1/2$. Partículas de spin opuesto pueden cancelarse para que el efecto exterior

sea la ausencia de spin. En este sentido son las partículas no emparejadas las que producen los efectos de RMN. Cuando se coloca una partícula en un campo magnético de intensidad B , una partícula con spin neto puede absorber un fotón de frecuencia $\nu = \gamma B$, donde, para el hidrógeno, $\gamma = 42.58 \text{ MHz/T}$.

Para comprender como se comportan las partículas con espín en un campo magnético, consideremos un protón. Este protón tiene una propiedad llamada espín, que puede conceptuarse como un momento magnético, que causa que el protón se comporte como un minúsculo imán con polos norte y sur. Cuando el protón se sitúa en un campo magnético externo, el vector espín de la partícula se alinea con el campo exterior. Hay una configuración de baja energía (N-S-N-S) y otra de alta energía (N-N-S-S). La partícula puede cambiar de configuración por la absorción de un fotón pasando de la configuración de baja energía a la de alta. La energía del fotón debe ser idéntica a la de la diferencia entre la de los estados, que es proporcional a la frecuencia $E = h\nu = h\gamma B$. La frecuencia se denomina de Larmor.

Cuando un conjunto de espines se coloca en un campo magnético externo, cada uno se coloca en una de las dos configuraciones posibles. La distribución de probabilidad de las poblaciones de espines en cada uno de los estados N^+ y N^- sigue la distribución de Boltzmann. La señal en RMN resulta de la diferencia entre la energía absorbida por los espines que hacen una transición del estado de menor energía al de mayor energía, y la energía emitida por los espines que simultáneamente hacen una transición del estado de mayor energía al de menor. Por tanto la señal es proporcional a la diferencia de poblaciones en estos estados. RMN es capaz de detectar esas pequeñas diferencias en población a frecuencias específicas. Para una descripción más sucinta se adoptan los paquetes de espines que están caracterizados por un vector que describe el espín resultante del balance de poblaciones dentro del paquete. La magnetización neta es el balance total de los vectores de todos los paquetes de espines. Se adopta la convención de que el campo magnético exterior y la magnetización neta se hallan sobre el eje Z espacial.

En equilibrio, la magnetización neta coincide con la magnetización en el eje Z , $M_0 = M_Z$, y no existen las magnetizaciones transversales M_X, M_Y . Es posible cambiar la magnetización neta exponiendo al sistema de espines a energía de frecuencia igual a la diferencia entre estados espine. Si se introduce suficiente energía es posible saturarlo y hacer $M_Z = 0$. La constante temporal que describe como M_Z vuelve al estado de equilibrio se denomina tiempo de relajación del

latice de espines T_1 :

$$M_Z = M_0 (1 - e^{-t/T_1}).$$

La precesión se produce cuando el vector de la magnetización neta está en el plano XY . Entonces girará en torno al eje Z con la frecuencia Larmor, que es la del fotón que puede cambiar de estado al espín.

Además de la rotación, el paquete de espines también sufre un proceso de desfase, debido a que cada espín rota independientemente a su frecuencia Larmor. El desfase aumenta con el tiempo. La constante temporal que describe el retorno al equilibrio de la magnetización transversal M_{XY} se denomina tiempo de relajación espín-espín T_2 .

$$M_{XY} = M_{XY_0} e^{-t/T_2}$$

T_2 es siempre menor o igual que T_1 . En el proceso de relajación, la magnetización en el plano XY tiende a cero mientras que la magnetización longitudinal crece hasta que la magnetización neta se alinea con el eje Z . Ambos procesos son simultáneos.

Una bobina de cable colocada alrededor del eje X producirá un campo magnético a lo largo del eje X cuando la corriente pasa por la bobina. Una corriente alterna producirá un campo magnético que alternará en dirección. Si tomamos un marco de referencia que rota alrededor del eje Z , (X', Y', Z') , el campo magnético en el eje X' será constante. En RMN el campo magnético creado por la corriente alterna a la frecuencia de Larmor pasado por bobina se denomina B_1 . Cuando esta corriente se activa y desactiva, crea un campo magnético pulsante B_1 en el eje X' . Los espines responden a este pulso de forma que la magnetización neta rota en la dirección del campo B_1 . el ángulo de rotación depende del tiempo transcurrido τ (mucho menor que T_1 y T_2) y del campo B_1 :

$$\theta = 2\pi\gamma\tau B_1.$$

Campos magnéticos variables en el tiempo a la frecuencia Larmor causan transiciones entre estados espín y cambios en M_Z . Existe una distribución de frecuencias en una muestra de moléculas, sólo las frecuencias iguales a la frecuencia Larmor afectan T_1 . Puesto que la frecuencia Larmor depende de B_0 , T_1 dependerá también de la intensidad del campo B_0 , en general es inversamente proporcional a la densidad de movimientos moleculares en la frecuencia Larmor. También dependerá de la temperatura y la viscosidad de la solución.

$$\begin{aligned}\frac{dM_{x'}}{dt} &= (\omega_o - \omega)M_{y'} - \frac{M_{x'}}{T_2} \\ \frac{dM_{y'}}{dt} &= -(\omega_o - \omega)M_{x'} + 2\pi\gamma B_1 M_z - \frac{M_{y'}}{T_2} \\ \frac{dM_z}{dt} &= -2\pi\gamma B_1 M_{y'} - \frac{(M_z - M_{z_o})}{T_1}\end{aligned}$$

Figura 7.1: Ecuaciones de Bloch

En la figura 7.1 recojemos las ecuaciones de Bloch que modelan el comportamiento del vector de magnetización en cualquier circunstancia. Su integración proporcionan los componentes (X' , Y' , Z) de la magnetización como funciones del tiempo. Estas ecuaciones se han utilizado para la síntesis de imágenes de resonancia magnética [14], [22], [73] para la experimentación con algoritmos de segmentación.

7.2. Espectroscopía de RMN

La rotación transversal alrededor del eje Z induce una corriente en una bobina localizada entorno al eje X . La gráfica de la corriente en función del tiempo es una onda seno que decae con una constante temporal T_2^* debido al desfase de los paquetes de espines. Esta señal se denomina inducción libre decadente (FID). Los vectores de magnetización transversal que rotan más rápido que el marco rotatorio de referencia se dice que tienen una frecuencia relativa positiva (+). Los que rotan más despacio se dice que tienen una frecuencia relativa negativa (-).

Un conjunto de pulsos de RF aplicados a una muestra para producir una forma específica de señal de MRN se denomina secuencia de pulsos. En una secuencia 90-FID la magnetización neta se rota en el plano $X'Y'$ con un pulso 90° . EL vector de magnetización neta tiene precesión en torno al eje $+Z$. La magnitud del vector decae también con el tiempo. Las secuencias de pulsos se presentan como diagramas temporales de multiples ejes ordenados. Un diagrama temporal para la secuencia 90-FID es un gráfico temporal de la energía RF junto con el gráfico temporal de otra señal del sistema. La figura 7.7 muestra un diagrama temporal. Cuando una secuencia se repite, por ejemplo para mejorar

la relación señal ruido (SNR), la amplitud de la señal tras la transformación de Fourier S dependerá de T_1 y el tiempo entre repeticiones de la secuencia, llamado tiempo de repetición TR .

$$S = k\rho (1 - e^{-TR/T_1}),$$

donde k es una constante proporcional y ρ es la densidad de espines en la muestra.

Una secuencia de pulsos comunmente usada es la secuencia *espin-eco*. En ella se aplica un pulso de 90° primeramente, que produce la rotación de la magnetización neta en el plano $X'Y'$. Después la magnetización comienza a desfasarse. En un momento posterior, se aplica un pulso 180° que produce la rotación de 180° alrededor del eje X' . Este pulso produce la recuperación de la fase parcialmente al menos y produce una señal denominada eco. La figura 7.2 presenta el diagrama temporal de la secuencia espin-eco. La ecuación para la repetición de la secuencia espin eco en función del tiempo de repetición TR y el tiempo de eco TE (definido como el tiempo entre el pulso 90° y la máxima amplitud del eco es la siguiente:

$$S = k\rho (1 - e^{-TR/T_1}) e^{-TE/T_2}.$$

Una secuencia de pulsos de recuperación de la inversión consiste en un pulso 180° que rota la magnetización neta al eje $-Z$. La magnetización pasa por la relajación espín lattice y vuelve a su posición de equilibrio en el eje $+Z$. ANtes de que alcance el equilibrio, se aplica un pulso de 90° , que rota la magnetización longitudinal en el plano XY . Cuando la magnetización está presente en el plano XY , rota alrededor del eje Z , se desafasa y produce un FID. La figura muesra el diagrama temporal de esta secuencia de pulsos. La señal cuando no hay repetición sigue la ecuacion

$$S = k\rho (1 - 2e^{-TI/T_1}).$$

Cuando se producen repeticiones cada TR segundos, la señal es

$$S = k\rho (1 - 2e^{-TI/T_1} + e^{-TR/T_1}).$$

Cuando los átomos se colocan en un campo magnetico se produce un pequeño flujo de los electrones en el sentido del campo, esto produce un pequeño campo magnético que se opone al campo exterior. De esta forma el campo en el nucleo es menor que el campo exterior $B = B_0(1 - \sigma')$. Dado que la composición

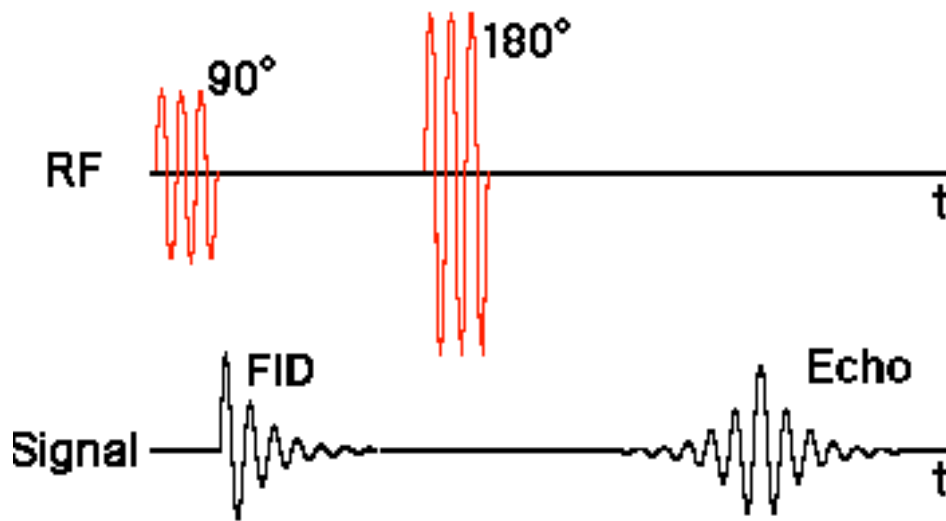


Figura 7.2: Diagrama temporal de la secuencia espin-echo.

electrónica varía de un elemento a otro, también variará el campo magnético en el núcleo para distintos elementos químicos, lo que dificulta que puedan compararse espectros obtenidos en distintos aparatos con intensidades de campos distintos. La forma de evitar el problema es calcular el desplazamiento químico respecto de un elemento estandar y utilizarla para normalizar las frecuencias Larmor que dependen del campo.

7.3. Componentes hardware

En la figura 7.3 se encuentra un diagrama de los componentes hardware del sistema de imagen de resonancia magnética. El magneto produce el campo magnético B_0 necesario para la generación de las imágenes. La mayor parte de los magnetos son superconductores que tienen resistencia nula cuando la bobina se sumerge en helio líquido para mantenerla a una temperatura de cero grados absolutos. Dentro del magneto se encuentran las bobinas de gradiente que permiten generar gradientes en las direcciones X , Y , Z en el campo magnético B_0 . Dentro de las bobinas de gradiente se sitúa la bobina de radiofrecuencia (RF) que genera el campo B_1 necesario para rotar los espines 90° o 180° , y que detecta la señal emitida por los espines. El paciente se posiciona dentro del magneto

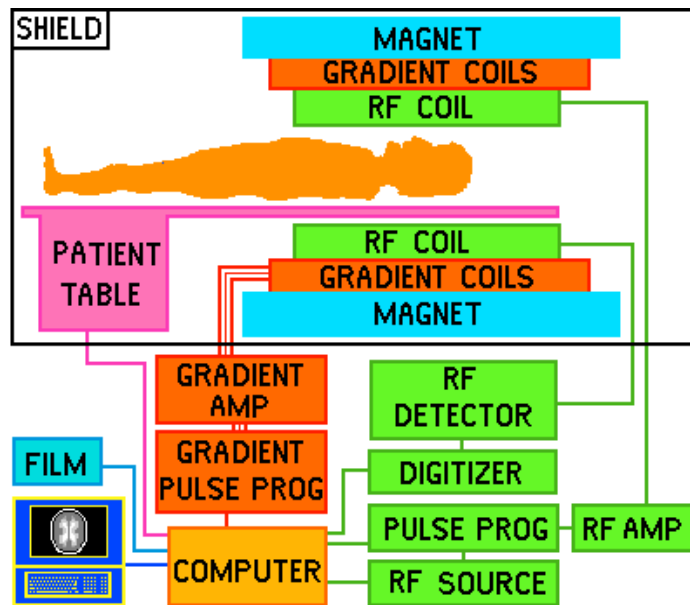


Figura 7.3: Esquema de un sistema de visualización de IRMN

por medio de una mesa controlada por computador, con una precisión de 1 milímetro en equipos estándar. La habitación está blindada para evitar tanto la interferencia desde el exterior como la propagación de las potentes señales de RF generadas por el sistema hacia el exterior. Además del posicionamiento del paciente, el computador controla la fuente de radio frecuencia, el programador de pulsos y el programador de gradientes. En las instalaciones modernas, el cálculo de la transformada de Fourier discreta se realiza en procesadores vectoriales dedicados. El amplificador de RF eleva la potencia de la señal desde milivatios hasta kilovatios, así como el amplificador de gradiente asegura que la señal de gradiente tendrá potencia suficiente para dirigir las bobinas de gradiente.

Las bobinas de gradiente están a temperatura ambiente y se disponen por pares en los que fluyen corrientes en direcciones opuestas, que permiten generar gradientes en B_0 en las direcciones positiva y negativa de cada eje. En la figura 7.4 se muestra la configuración típica de las bobinas de gradiente para un sistema de apertura horizontal.

Las bobinas de RF producen el campo B_1 que rota la magnetización neta en una secuencia de pulsos. También detectan la magnetización transversal en su

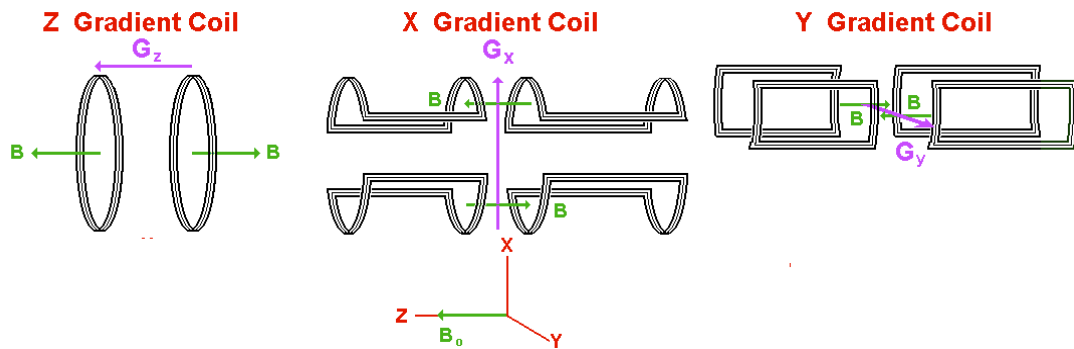


Figura 7.4: Esquema de la disposición de las bobinas de gradiente para un sistema de imagen de resonancia magnética de apertura horizontal

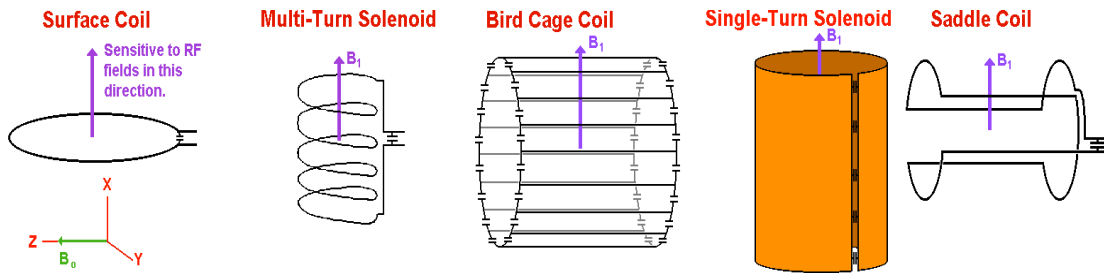


Figura 7.5: Algunas bobinas de RF típicas.

predecesión en el plano XY . Las bobinas pueden ser de transmisión y recepción, de transmisión o recepción únicamente. En el segundo caso, bobinas separadas se combinan para realizar la excitación y detección de la señal emitida por los espines del objeto visualizado. La bobina de RF es el equivalente a la lente en los sistemas ópticos. Distintas bobinas pueden ser útiles para distintas aplicaciones. El campo B_1 debe ser perpendicular al campo B_0 . La bobina de RF debe resonar a la frecuencia Larmor. Todas las bobinas de visualización se componen de un inductor y un conjunto de elementos capacitivos. La frecuencia de resonancia está determinada por la inductancia (L) y la capacitancia (C): $n = (2\pi\sqrt{LC})^{-1}$. Estos elementos se pueden ajustar para una mejor visualización dependiendo del individuo. La figura 7.5 muestra algunas bobinas de RF típicas, todas son emisor/receptor excepto la de superficie.

El detector de cuadratura es un dispositivo que extrae las señales $M_{x'}$ y $M_{y'}$ de la señal detectada por la bobina RF. Su corazón es el denominado "mezclador doble-balanceado" (DBM), que tiene dos entradas y una salida. Si las entradas son $\cos(A)$ y $\cos(B)$, la salida es $1/2 \cos(A+B)$ y $1/2 \cos(A-B)$. Por esta razón, al DBM a menudo se le denomina detector de productos. El detector de cuadratura contiene típicamente dos DBM, dos filtros, dos amplificadores y un cambiador de fase de 90° . Las dos entradas son las frecuencias n y n_0 , las dos salidas son los componentes M_X y M_Y de la magnetización transversal.

7.4. Principios de visualización

El principio básico es la ecuación de resonancia, $n_0 = gB_0$, dice que el espín de un objeto resuena a una frecuencia proporcional a la magnitud del campo B_0 . Donde g es la constante giromagnética. Cuando una región de espín está experimentando el mismo campo magnético, sólo hay un pico en el espectro de RMN. Si cada región de espín experimenta un campo magnético único, podríamos visualizar sus posiciones. Un gradiente de un campo magnético es una variación del campo magnético respecto de la posición. El gradiente unidimensional es una variación respecto de una dirección. Un gradiente a lo largo del eje X indica que el campo magnético B_0 aumenta en esa dirección

El isocentro del imán $(x, y, z) = (0, 0, 0)$, donde la magnitud del campo es B_0 y la frecuencia de resonancia es n_0 . Si un gradiente lineal del campo se aplica al objeto las regiones con spin experimentarán distintos campos, lo que resulta en más de una señal. La amplitud de la señal es proporcional al número de espines en un plano perpendicular al gradiente. Este procedimiento es la codificación de frecuencia y hace que la frecuencia de resonancia sea proporcional a la posición espacial del espín.

$$\begin{aligned} n &= g(B_0 + xG_x) = n_0 + gxG_x, \\ X &= (n - n_0) / (gG_x). \end{aligned}$$

Este principio es la base tras toda visualización de resonancia magnética. La imagen se genera a partir del espectro RMN mediante el método de proyección inversa (Backprojection), también conocido como transformada de Radon inversa [23]. Esta es una de las primeras técnicas de visualización de RMN que tuvo éxito. Se coloca el objeto en el campo B_0 y se le somete a varias visualizaciones variando el ángulo del gradiente unidimensional, guardándose los diferentes espectros obtenidos. La retroproyección de estos espectros proporciona la imagen

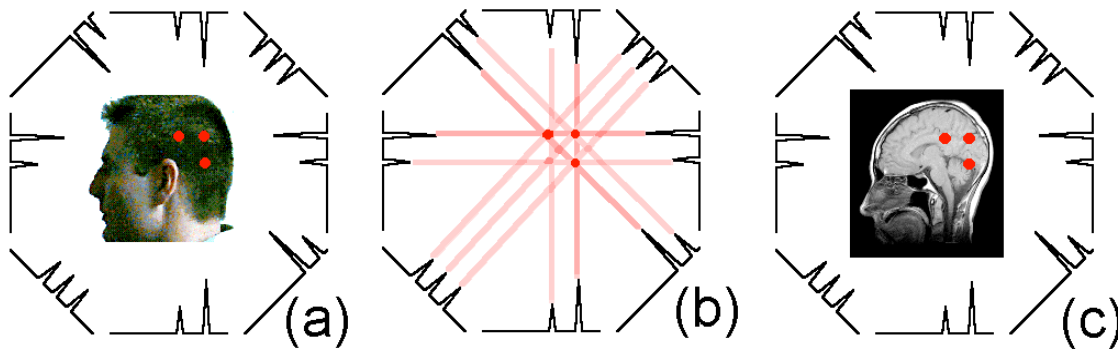


Figura 7.6: Ilustración de la obtención de la imagen mediante retroproyección.

de tomografía. En la figura 7.6 se ilustra la retroproyección que obtiene las imágenes de tomografía.

La variación del ángulo q del gradiente se consigue por la aplicación de combinaciones lineales de dos gradientes. Aquí se aplican los gradientes en X e Y en las siguientes proporciones para obtener el gradiente deseado G_f :

$$\begin{aligned} G_y &= G_f \sin q \\ G_x &= G_f \cos q \end{aligned}$$

La selección de la rodaja a visualizar se consigue aplicando un gradiente unidimensional lineal durante el periodo en que el pulso de RF se aplica. Un pulso de 90° aplicado en conjunción con un gradiente del campo magnético rotará algunos o muchos de los espines que están localizados en un plano a través del objeto. La razón de que no se roten todos los espines estriba en algunas de las frecuencias tienen un B_1 menor que el requerido para la rotación de 90° . Como consecuencia los espines seleccionados no constituyen una rodaja exacta del objeto. Una posible solución es la adopción de un pulso con perfil de función *sinc* para obtener la rotación de los espines. La figura muestra una secuencia de pulsos que permitiría la obtención de la imagen de RMN mediante retroproyección.

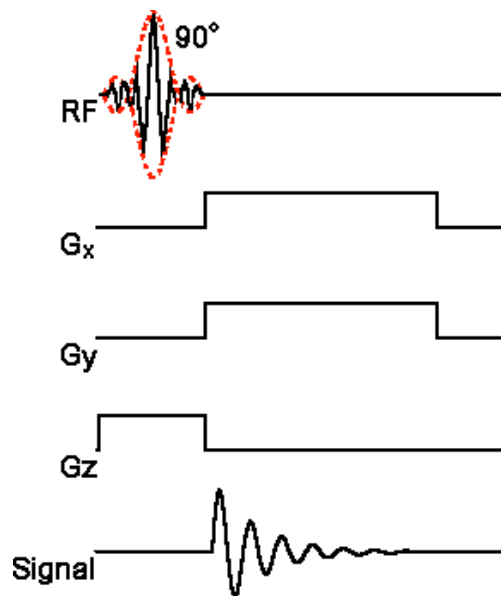


Figura 7.7: Pulsos para la obtención de la imagen de una rodaja mediante retroproyección.

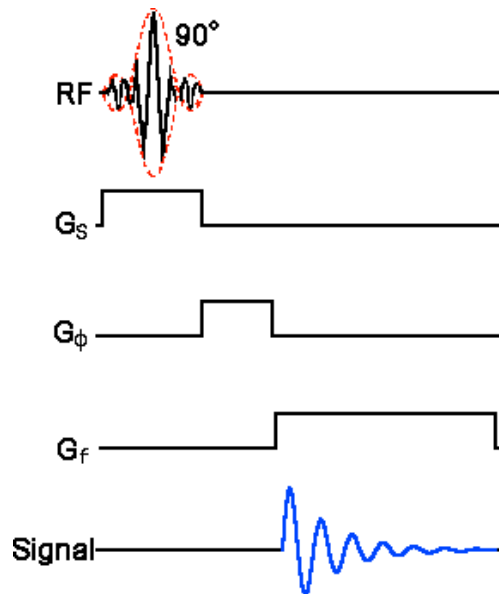


Figura 7.8: Secuencia de pulsos para la visualización mediante TF.

7.5. Visualización mediante la TF

El gradiente de codificación de la fase es un gradiente en el campo magnético B_0 . Este gradiente se utiliza para impartir un ángulo de fase específico al ángulo de magnetización transversal. El ángulo de fase específico depende de la localización del vector de magnetización transversal. Si G_x es el gradiente en el campo magnético a lo largo del eje X , los vectores de magnetización de regiones con frecuencias Larmor idénticas tendrán una precesión en la dirección del campo magnético a una frecuencia dada por la ecuación de resonancia:

$$n = g(B_0 + xG_x) = n_0 + gxG_x.$$

Mientras el gradiente de codificación de fase está activo, cada vector de magnetización transversal tiene una frecuencia de Larmor única. Si el gradiente en X se desactiva, las frecuencias Larmor vuelven a ser idénticas, sin embargo el ángulo de fase (el ángulo respecto de un eje de referencia cuando el gradiente ha sido desactivado) no es idéntico. Si podemos medir esta fase, podemos asignar posiciones en el eje X .

Los diagramas de tiempo para las secuencias de pulsos de visualización tienen entradas para la RF, los pulsos de gradiente del campo magnético (selección de la rodaja, codificación de fase, codificación de frecuencia) y las señal. La figura 7.8 muestra la secuencia más sencilla para la visualización mediante FT. Los pulsos de los tres gradientes representan la magnitud y duración de los gradientes del campo magnético. La secuencia de sucesos es la siguiente:

1. Se activan simultáneamente el gradiente de selección de rodaja y el pulso de RF. AL finalizar el pulso RF se desactiva el gradiente de selección de rodaja.
2. Se activa el gradiente de codificación de fase.
3. Se activa el gradiente de codificación en frecuencia y se graba la señal captada (FID)

Esta secuencia se repite 128 o 256 veces para obtener los datos para una imagen. Cada vez se modifica la magnitud del gradiente de codificación de fase, en pasos iguales. Los gradientes son ortogonales entre si.

Si imaginamos un cubo de espines colocado en un campo magnético. El cubo está compuesto de varios elementos de volumen, cada uno con su propio vector de magnetización neto. Supongamos que queremos visualizar una rodaja en el plano XY . El campo B_0 está a lo largo del eje Z , el gradiente de selección de rodajas se aplica a lo largo del eje Z . El pulso RF rota sólo los paquetes de espines que satisfacen la condición de resonancia (en el plano XY en este ejemplo). La localización del plano respecto del isocentro en el eje Z viene dada por $Z = \Delta n / gG_s$, donde $\Delta n = (n - n_0)$, G_s es la magnitud del gradiente selector de rodajas y g es el ratio giromagnético. Tras la aplicación del gradiente de fase, los espines en el plano seleccionado giran a la misma frecuencia pero desfasados a lo largo del eje X (si es el seleccionado para esta tarea). La aplicación del pulso de RF hace que los espines giren a frecuencias que dependen de su posición en Y . Por tanto cada posición en XY está caracterizada por un par (fase, frecuencia). La localización se consigue aplicando la transformada de Fourier a los datos de los FID registrados tomados como una función 2D, cuyos ejes son la codificación frecuencial y la codificación en fase. Los picos de esta señal nos dan la localización de los espines en el plano XY .

De hecho se utilizan secuencias de pulsos más complejas que la de la figura 7.8. En las figuras se muestran secuencias espin eco y de recuperación de la inversión que incluyen el gradiente de codificación en fase.

7.6. Presentación de la imagen

Los datos brutos, o datos en el espacio-k es una matriz (típicamente 256×256) de datos complejos cuyas partes reales e imaginarias corresponden a M_x y M_y respectivamente. Los datos tienen típicamente 16 bits de resolución. Antes de realizar la transformada de Fourier se suavizan los datos convolucionandolos con una función Lorentziana de forma lineal (equivalente a multiplicar en el dominio espacial por un cono exponencial). Después de la transformación de Fourier la imagen visualizada es la magnitud de los valores complejos, por ello nunca habrá valores negativos en la imagen. Estos datos pueden ser expandidos a una matriz de 512×512 mediante interpolación o repetición. La reducción de los 15 bits de resolución de los niveles de gris hasta los 8 bits típicos de las pantallas se puede hacer de diversas maneras que pueden enfatizar algunos aspectos de la imagen. Típicamente se hace mediante una función de transferencia que recorta los niveles indeseados y es lineal en el intervalo seleccionado. Este recorte puede tener efecto dramáticos en la imagen visualizada.

Los aparatos clínicos no usan el sistema de coordenadas del sistema de resonancia magnética XYZ para la colección y presentación de las imagenes. EN su lugar se utiliza el sistema de coordenadas anatomico. Los tres ejes son izquierda-derecha (L/R), superior-inferior (S/I) y anterior-posterior (A/P). Similarmente, en los aparatos clínicos los planos de visualización son:

- Plano axial: plano de visualización perpendicular al eje más largo del cuerpo. Los lados de este plano son L/R y A/P. (eje Z)
- Plano coronal: plano que biseca el frente del cuerpo respecto de la espalda. Los lados del plano son L/R y S/I. (eje Y)
- Plano sagital: plano que biseca los lados izquierdo y derecho del cuerpo. Los lados del plano son S/I y A/P. (eje X)

7.7. Artefactos en la imagen

Se denominan artefactos a los defectos o anomalías en las imagenes obtenidas que son debidas a circunstancias especiales de los objetos visualizados o al mal funcionamiento de algún componente hardware. La relación entre las causas y los artefactos presentes en la imagen no es biunívoca, diversas causas pueden

producir un mismo artefacto en la imagen. Una enumeración no exhaustiva de los artefactos más comunes es lo que sigue:

- Cuadratura RF. Está causada por defectos en el circuito de recepción de RF, típicamente en el detector de cuadratura. El efecto más habitual es la aparición de ruido coherente en forma de ondas espaciales, puesto que se corrompe los datos en el espacio transformado de Fourier.
- Inhomogeneidad de B_0 . La inhomogeneidad del campo B_0 puede producir distorsiones espaciales y de intensidad en la imagen. Las distorsiones de intensidad se deben a que la homogeneidad del campo tiene un valor mayor o menor en una cierta región del espacio. La respuesta de la señal T_2^* es entonces mayor o menor en esa región. Las distorsiones espaciales se producen cuando existe un gradiente en el campo B_0 , que causa que los espines resuenen a frecuencias Larmor distintas de las prescritas por una secuencia de visualización.
- Gradiente. El mal funcionamiento de las bobinas de gradiente puede producir distorsiones espaciales.
- Inhomogeneidad de RF. Se trata de una variación de la intensidad a lo largo de la imagen. La causa puede ser tanto la inhomogeneidad del campo B_1 como la inhomogeneidad de la sensibilidad del receptor. Las bobinas de superficie presentan naturalmente este artefacto. En otros casos, la presencia del artefacto revela bien defectos en las bobinas de RF, o bien la presencia de metales ferromagnéticos en el objeto visualizado, que aparecen como regiones de señal vacía.
- Movimiento. El movimiento del objeto visualizado resulta en el emborronamiento de toda la imagen con imágenes fantasma en la dirección de la codificación de fase. Movimientos de porciones del objeto resultan en el emborronamiento de esa porción a lo largo de la secuencia de imágenes. La respiración o el corazón son fuentes de este artefacto. La solución consiste en sincronizar la toma de las imágenes con el movimiento rítmico biológico.
- Flujo. El flujo de los fluidos hace que no se obtenga respuesta a pesar de que los fluidos contengan muchos átomos de hidrógeno. La razón es que el fluido excitado por los pulsos de RF deja de estar en el corte visualizado antes de que se registre la señal, perdiéndose su respuesta.

- Desplazamiento químico. Las diferencias de composición de los tejidos hacen que la respuesta de los tejidos (p.e.: de grasa) sufran un retraso respecto de los restantes átomos, lo que se convierte en desplazamientos espaciales.
- Volumen parcial. Se debe a que la resolución de la visualización no permite resolver los distintos tejidos en un voxel.
- Circularidad (wrap around). Es la ocurrencia de una parte de la anatomía visualizada, que está fuera del campo de vista, en el interior de dicho campo. Está causada porque el campo de vista es menor que el tamaño del objeto visualizado. La solución computacional se basa en el sobremuestreo de la imagen, filtrado digital y decimación.

El objeto de esta tesis es la corrección de los artefactos que disminuyen la intensidad que surgen de las inhomogeneidades en los campos magnéticos, tanto de los campos magnéticos B_0 y B_1 como de la bobina de RF.

7.7.1.El ruido en las imágenes de RMN

En las imágenes de resonancia magnética nuclear, el ruido tiene una distribución de Rician en las zonas de baja intensidad, mientras que en las zonas de alta intensidad la distribución se separa poco de la normal. Debido a las características específicas de la imagen y al tratamiento que habitualmente se hace de ellas, han sido objeto especial de investigación las técnicas no lineales de eliminación de ruido. Entre ellas destaca el filtrado anisotrópico que preserva los bordes entre regiones de intensidad constante salvo ruido [90], [117].

El trabajo de [52] merece destacarse por el cuidadoso análisis que realiza de las fuentes de ruido en la creación de la imagen. En concreto llega a la siguiente expresión del ratio señal ruido (SNR) :

$$\begin{aligned} \text{SNR} &= \frac{M_s}{\sigma_n} = \left[\frac{2x\sqrt{\rho}}{\gamma\mu_0\sqrt{KT}\pi} \right] \left[\frac{1}{r_0^2\sqrt{I}} \right] \left[\omega_o V_h \sqrt{T} \right] \\ &= Cf(Obj) f(Im), \end{aligned} \quad (7.1)$$

donde C corresponde a las constantes físicas que no pueden ser controladas ni manipuladas en el primer paréntesis de la expresión, $f(Obj)$ representan las dimensiones del objeto y $f(Im)$ corresponde al último parentesis y contiene los parámetros de visualización que están bajo el control del operador: la frecuencia

del campo magnético (ω_o), el volumen del voxel (V_h), el tiempo de relajación (T). Que el ruido varie dependiendo de las dimensiones y forma del objeto es una propiedad muy específica de la imagen de resonancia magnetica nuclear. Que el ruido en la imagen dependa de parámetros de la visualización que son ajustados por el operador, implica que protocolos de adquisición cuidadosos pueden reducir la presencia del ruido.

8. APENDICE 2: POLINOMIOS DE LEGENDRE

Los polinomios de Legendre [102], [119], también llamados funciones de Legendre del primer tipo, coeficientes de Legendre o armónicos zonales son soluciones de la ecuación diferencial de Legendre. Si l es entero son polinomios. La figura 8.1 muestra los polinomios de Legendre $P_n(x)$ para $n = 1, 2, \dots, 5$ y $x \in [-1, 1]$. El polinomio de Legendre $P_n(x)$ se puede definir por la integral circular:

$$P_n(x) = \frac{1}{2\pi i} \oint (1 - 2tx - t^2)^{1/2} dt, \quad (8.1)$$

donde el contorno incluye el origen y es atravesado en dirección contraria a las agujas del reloj.

Los primeros polinomios de Legendre son:

$$P_0(x) = 1 \quad (8.2)$$

$$P_1(x) = x \quad (8.3)$$

$$P_2(x) = \frac{1}{2} (3x^2 - 1) \quad (8.4)$$

$$P_3(x) = \frac{1}{2} (5x^3 - 3x) \quad (8.5)$$

$$P_4(x) = \frac{1}{8} (35x^4 - 30x^2 + 3) \quad (8.6)$$

$$P_5(x) = \frac{1}{8} (63x^5 - 70x^3 + 15x) \quad (8.7)$$

Los polinomios de Legendre pueden generarse también mediante la ortonormalización de Gram-Schmidt en el intervalo abierto $(-1, 1)$.

Los polinomios de Legendre son un caso especial de funciones ultrasféricas con $\alpha = 1/2$, un caso especial de los polinomios de Jacobi $P_n^{(\alpha, \beta)}$ con $\alpha = \beta = 0$, y pueden ser escritos como funciones hipergeométricas usando la fórmula de Murphy:

$$P_n(x) = P_n^{(0,0)} = {}_2F_1 \left(-n, n+1; 1; \frac{1}{2}(1-x) \right). \quad (8.8)$$

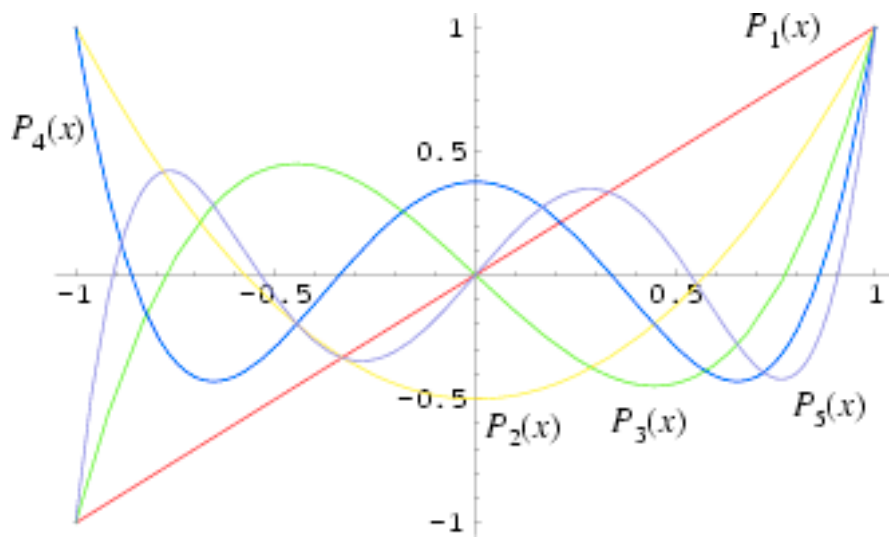


Figura 8.1: Los primeros 5 polinomios de Legendre en el intervalo $(-1,1)$.

La representación de Rodrigues da la siguiente fórmula:

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l. \quad (8.9)$$

Realizando la expansión se obtiene:

$$P_l(x) = \frac{1}{2^l} \sum_{k=0}^{\lfloor l/2 \rfloor} (-1)^k \binom{l}{k} \binom{2l-2k}{l} x^{l-2k}. \quad (8.10)$$

8.1. Ortogonalidad y productos de polinomios

Una de las propiedades de los polinomios de Legendre que más nos interesan para nuestro trabajo, es su ortogonalidad, que nos permite asumir un conjunto de estos polinomios como una base de funciones para la representación de otras funciones.

Los polinomios de Legendre son ortogonales en el intervalo $(-1, 1)$ con función de peso 1, y cumplen:

$$\int_{-1}^1 P_n(x) P_m(x) dx = \frac{2}{2n+1} \delta_{n,m}, \quad (8.11)$$

donde $\delta_{n,m}$ es la función delta de Kronecker.

Podemos definir funciones suaves de soporte multidimensional como productos externos de los polinomios 1D:

$$P_{ij}(x, y) = P_i(x) P_j(y), \quad (8.12)$$

$$P_{ijk}(x, y, z) = P_i(x) P_j(y) P_k(z). \quad (8.13)$$

Lo que nos interesa es que la ortogonalidad de estos polinomios se mantiene en su extensión multidimensional:

$$\begin{aligned} & \int_{-1}^1 \int_{-1}^1 P_{n,m}(x, y) P_{k,l}(x, y) dx dy \quad (8.14) \\ &= \int_{-1}^1 \int_{-1}^1 P_n(x) P_k(x) P_m(y) P_l(y) dx dy \\ &= \left(\frac{2}{2n+1} \right)^2 \delta_{n,k} \delta_{m,l}, \end{aligned}$$

de forma que podemos seguir considerando estos polinomios como bases de funciones que nos permiten representar funciones suaves como combinaciones lineales. En la figura se muestra los polinomios 2D de orden 4. En [116] se considera suficiente el conjunto de funciones que resultan de los productos en la mitad izquierda de la matriz, debido a la simetría de la definición.

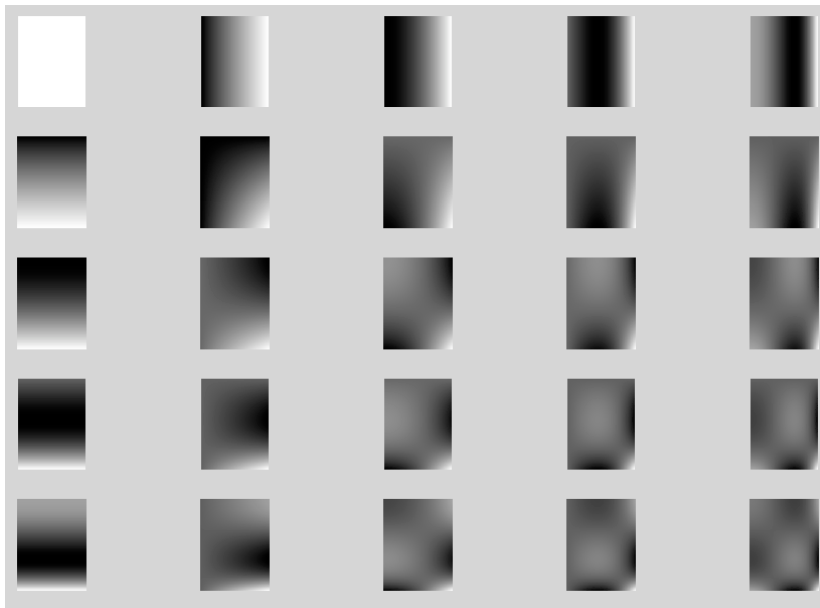


Figura 8.2: Productos de polinomios de Legendre de hasta orden 4.

9. APENDICE 3: ESTRATEGIAS EVOLUTIVAS Y ALGORITMOS MEMÉTICOS

En este apéndice hacemos una introducción breve sobre las estrategias evolutivas y los algoritmos meméticos que sirve para poner en contexto la proposición del algoritmo GradPABIC. La exposición es intuitiva y no incluye proposiciones o pruebas formales.

9.1. Estrategiás evolutivas

En la figura 9.1 se muestra el esquema algorítmico genérico de las estrategias evolutivas [6]. Las estrategias evolutivas son bien conocidas en la literatura como métodos aleatorios de optimización basados en representaciones con números reales, a diferencia de las representaciones binarias de los algoritmos genéticos, y con un énfasis especial en la mutación como operador de cambio. De hecho, en nuestros trabajos hemos prescindido completamente de los operadores de cruce. Las estrategias evolutivas consisten, por tanto, en procesos en los que se itera la selección de unos individuos de la población para que actúen como padres de la siguiente generación. Estos individuos son mutados de acuerdo con perturbaciones que siguen una distribución normal multivariante de media cero y de varianza la que corresponde al individuo. Las varianzas asociadas a los individuos se adaptan en el proceso, lo que puede entenderse como un ajuste automático de la ganancia de un algoritmo de descenso de gradiente [121]. La adaptación de las varianzas se realiza también con un proceso estocástico. Si la población en la iteración t de la estrategia evolutiva se denota como $(x_{1(t)}, \dots, x_{n(t)}, \sigma)$, donde σ es la varianza de la mutación. La mutación se puede formular mediante las siguientes ecuaciones:

$$\sigma' = \sigma \cdot \exp(\tau \cdot N(0, 1)), \quad (9.1)$$

$$x_{i(t+1)} = x_{i(t)} + N_i(0, \sigma'). \quad (9.2)$$

Esto es, se calcula en primer lugar la nueva varianza de la mutación σ' y a continuación se calculan los individuos mutados añadiéndoles una muestra de una variable aleatoria con distribución normal de media cero y varianza σ' . En estas ecuaciones, $N(0, \sigma')$ denota una muestra de una variable aleatoria con distribución normal.

Los algoritmos evolutivos en general pueden ser, o no, elitistas. Un algoritmo elitista siempre escoge para continuar la información más adecuada. En nuestro caso, la siguiente generación estará constituida por los mejores de entre los padres y los hijos, lo que se denomina ES- $(\lambda + \mu)$ en la terminología de estrategias evolutivas. Donde λ es el tamaño de la población (los padres) y μ el número de mutaciones. Las estrategias no elitistas permiten que las nuevas generaciones estén constituidas por individuos no óptimos, con la esperanza de que permitan superar máximos locales. Para ello, la selección sólo se realiza sobre los individuos resultantes de la mutación. Se denomina genéricamente ES- (λ, μ) a estas estrategias de selección.

9.1.1. Estrategias evolutivas para la corrección de la iluminación

En esta sección precisamos los aspectos de la definición de las Estrategias Evolutivas (EE) para el problema de la corrección de la iluminación en IRM, que se usan para comparar con el PABIC en la sección 4.2, para intentar mostrar sus limitaciones como algoritmo evolutivo. Las EE se concretan para el problema de la corrección parametrizada de iluminación si consideramos la extensión de la formulación del PABIC con np individuos. Esto es la población de la EE viene dada por

$$\mathbf{x}_t = \{\mathbf{p}_1, \dots, \mathbf{p}_{np}\}, \quad (9.3)$$

donde \mathbf{p}_i es un vector de parámetros de la combinación lineal de polinomios de Legendre que constituye el estimador del sesgo del campo de iluminación. La función de ajuste para cada individuo \mathbf{p}_i viene dada por la misma ecuación 3.61 dada para el algoritmo PABIC :

$$e(\mathbf{p}_i) = \sum_{\mathbf{x}} \prod_{k=1}^c \left(\frac{f(\mathbf{x})}{\widehat{b}(\mathbf{x}, \mathbf{p}_i)} - \mu_k \right)^2. \quad (9.4)$$

Donde $\{\mu_k; k = 1, \dots, c\}$ son los valores nominales de intensidad de las clases de tejidos en la imagen. Por lo demás la estructura del algoritmo es idéntica a la de las Estrategias Evolutivas descritas anteriormente en la misma sección.

- Random population creation Local optimization of each individual
- Compute fitness
- Repeat
 - Select parents
 - Perform crossover and mutation. Compute fitness
 - Select new population
- Until convergence

Figura 9.1: Algoritmo general de una estrategia evolutiva

9.2. Algoritmos Meméticos

Los algoritmos meméticos han sido propuestos por Moscato [95] quien mantiene información actualizada sobre el estado de desarrollo en la pagina web [96]. En esencia los algoritmos meméticos se inspiran en el modelado de los procesos de información en sistemas sociales. En ellos, cada individuo somete a alteraciones la información antes de propagarla a otros individuos de la sociedad. Un *meme* es un cuanto de información que cada individuo aporta a la cadena de transmisión. En la fusión conceptual de los algoritmos evolutivos y las ideas sobre los sistemas meméticos, lo que nos queda es un proceso evolutivo en el que los individuos seleccionados para constituir la proxima generación son sometidos a un proceso local de optimización, usualmente basados en el descenso del gradiente. La figura 9.2 ilustra el esquema general de un algoritmo memético, tal y como lo entendemos en esta memoria.

9.2.1. Algoritmos meméticos instantaneos

En la figura 9.3 se ilustra la idea de un algoritmo memético instantáneo, que no aplica el resultado de una iteración de descenso de gradiente completa, sino una perturbación guiada por el gradiente de la función de ajuste. Esta clase de algoritmos está inspirado en el algoritmo de descenso de gradiente aleatorio estudiado en [121]. Las estrategias evolutivas y algoritmos genéticos

- Random population creation.
- Local optimization of each individual.
- Compute fitness.
- Repeat
 - Select parents.
 - Perform crossover and mutation. Compute fitness.
 - Select new population.
 - Local optimization of each individual. Compute fitness.
- Until convergence.

Figura 9.2: Algoritmo Memetico general

son sensibles al tamaño de la población, el caso especial de algoritmos cuyas poblaciones están compuestas de un único individuo son casos extremos. El esquema del algoritmo memético de individuo único que se presenta en la figura 9.4 es el que corresponde a nuestra proposición del algoritmo GradPABIC.

Notese que en todo caso el algoritmo memético de individuo único converge hacia un óptimo local, puesto que se trata de un algoritmo elitista. El algoritmo converge al óptimo global en tiempo infinito, de poco valor práctico, puesto que la perturbación aleatorio permite asignar probabilidades no nulas a todos los puntos del espacio de parámetros. Pero el efecto real es una cierta aceleración de la convergencia del algoritmo hacia optimos locales de calidad, debida a la introducción de la mutación guiada por el gradiente de la función objetivo, que incluye información *a priori* en la búsqueda. Incluso podría argüirse que el uso del gradiente introduce información sobre la topología de la función de ajuste en funciones relativamente suaves.

- Random population creation.
- Compute fitness.
- Compute fitness function gradient.
- Repeat
 - Select parents.
 - Perform crossover.
 - Perform mutation based on the fitness gradient.
 - Compute fitness.
 - Select new population.
 - Compute fitness function gradient.
- Until convergence.

Figura 9.3: Algoritmo Memetico instantáneo

- Random population creation.
- Compute fitness.
- Compute fitness function gradient.
- Repeat
 - Select parents.
 - Perform crossover.
 - Perform mutation based on the fitness gradient.
 - Select new population.
 - Compute fitness.
 - Compute fitness function gradient.
- Until convergence.

Figura 9.4: Algoritmo Memetico instantaneo de individuo unico

BIBLIOGRAFÍA

- [1] Aarts E., J. Korst (1989) *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*, John Wiley & Sons
- [2] Adini, Y.; Moses, Y.; Ullman, S. (1997) Face recognition: the problem of compensating for changes in illumination direction, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):721 - 732
- [3] Ahmed M.N., S.N. Yamany, N. Mohamed, A. A. Farag, T. Moriarty, (2002) A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data *IEEE Trans. Med. Imag.*, 21 (3):193-199
- [4] M. S. Atkins, B. T. Mackiewich, (1998) Fully Automatic Segmentation of the Brain in MRI, *IEEE Trans. Med. Imaging* 17(1):98-107
- [5] Bäck T. (1996) *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*, Oxford University Press
- [6] Bäck T., H.-P. Schwefel, (1993) An overview of evolutionary algorithms for parameter optimization, *Evolutionary Computation*, 1(1):1-23
- [7] Beni, G.; X. Liu (1994) A least biased fuzzy clustering method; *IEEE Trans. Patt. Anal. Mach. Int.* 16(9): 954 – 960
- [8] Bensaid, A.M.; Hall, L.O.; Bezdek, J.C.; Clarke, L.P.; Silbiger, M.L.; Arrington, J.A.; Murtagh, R.F. (1996) Validity-guided (re)clustering with applications to image segmentation, *IEEE Trans. Fuzzy Systems* 4(2):112 - 123
- [9] Bezdek J.C. (1980) A convergence theorem for the fuzzy ISODATA clustering algorithm, *IEEE Trans. Patt. Anal. Mach. Int.* 2:1-8

- [10] Bezdek J.C. (1981) Pattern Recognition With Fuzzy Objective Function Algorithms, Plenum Press
- [11] Bezdek J.C., L.O. Hall, L.P. Clarke, (1993) Review of MR image segmentation techniques using pattern recognition, *Med. Phys.* 20(4), pp. 1033-1048.
- [12] Bezdek J.C., S.K. Pal (1991), Fuzzy models for pattern recognition, Piscataway, NJ: IEEE Press
- [13] Blanz, V.; Vetter, T. (2003) Face recognition based on fitting a 3D morphable model, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(9):1063 - 1074
- [14] BrainWeb <http://www.bic.mni.mcgill.ca/brainweb/>
- [15] Castleman K.R. (1995) Digital Image Processing, Prentice-Hall
- [16] Clark, M.C.; Hall, L.O.; Goldgof, D.B.; Clarke, L.P.; Velthuizen, R.P.; Silbiger, M.S. (1994) MRI segmentation using fuzzy clustering techniques, *IEEE Engineering in Medicine and Biology Magazine*, 13(5):730 - 742
- [17] Clark, M.C.; Hall, L.O.; Goldgof, D.B.; Velthuizen, R.; Murtagh, F.R.; Silbiger, M.S. (1998) Automatic tumor segmentation using knowledge-based techniques, *IEEE Trans. Medical Imaging*, 17(2):187-201
- [18] Clarke L.P. , R.P. Velthuizen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, and M.L. Silbinger (1995) MRI Segmentation: Methods and Applications. *Magnetic Resonance Imaging*, 13(3):343-368.
- [19] K.-H. Chuang, M.-J. Chiu, C.-C. Lin, J.-H. Chen (1999) Model-Free Functional MRI Analysis Using Kohonen Clustering Neural Network and Fuzzy -Means, *IEEE Trans. Med. Imaging* 18(12):1117-1118
- [20] Cocosco C.A. , V. Kollokian, R.K.-S. Kwan, and A.C. Evans, (1997) Brainweb: Online interface to a 3d mri simulated brain database, *NeuroImage*, 5(4):425.
- [21] Cohen MS , RM DuBois, MM Zeineh. (2000) Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging. *Hum Brain Mapping* 10(4): 204-211.

- [22] Collins D.L., A.P. Zijdenbos, V. Kollokian, J.G.Sled, N.J. Kabani, C.J.Holmes, A.C. Evans (1998), Design and construction of a realistic digital brain phantom, *IEEE Trans. Med. Imag.*, 17(3) pp. 463 – 468.
- [23] Deans S.R., S. Roderick, (1983) *The Radon Transform and Some of its Applications*. Wiley, New York
- [24] De Carli C., D. G. Murphy, D. Teichberg, G. Campbell, G.S. Sobering (1996) Local histogram correction of MRI spatially dependent image pixel intensity nonuniformity, *J. Mag. Res. Imaging* 6:519-528.
- [25] Dempster A.P., N.M. Laird, D.B. Rubin (1977) Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Stat. Soc.* 39:1-38.
- [26] de Vel O., S. Aeberhard (1999) Line-Based face recognition under varying pose, *IEEE Patt. Anal. MAch. Int.* 21(10):1081-1087
- [27] Dhawan A.P. (2003) *Medical Image Analysis*, IEEE Press, New York.
- [28] Duda R. O., P. E. Hart, (1973) *Pattern Classification and Scene Analysis*, Wiley
- [29] Dunn J.C. (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *J. Cybern.* 3:32-57
- [30] M. Elad, R. Kimmel, D. Shaked, R. Keshetc (2003) Reduced complexity Retinex algorithm via the variational approach, *J. Vis. Commun. Image R.* 14:369–388
- [31] Erdem C.E. , A. T. Erdem (2001) An illumination invariant algorithm for subpixel accuracy image stabilization and its effect on MPEG-2 video compression, *Signal Processing: Image Communication* 16:837-857
- [32] Eschrich, S.; Jingwei Ke; Hall, L.O.; Goldgof, D.B. (2003) Fast accurate fuzzy clustering through data reduction, *IEEE Trans. Fuzzy Systems*,11(2):262 - 270
- [33] Fan A. C.(2003) *A Variational Approach to MR Bias Correction*, Master Thesis, MIT
- [34] Farrar T.C. (1987) *An Introduction To Pulse NMR Spectroscopy*, Farragut Press, Chicago.

- [35] Face database: AR mantenida por Aleix M. Martinez en Purdue , creada en el CVC de la UAB, http://rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html
- [36] Face databases en la CMU: <http://vasc.ri.cmu.edu/idb/html/face/index.html>
- [37] FAcE database Erlangen, <http://www.humanscan.de/support/downloads/facedb.php>, también disponible en <http://home.t-online.de/home/Robert.Frischholz/facedetection/datasets.htm>
- [38] Face database FERET, actualmente distribuida por el NIST http://www.itl.nist.gov/iad/humanid/feret/feret_master.html
- [39] Face database ORL, el sitio original es <http://mambo.ucsc.edu/psl/olivetti.html>, actualmente también se mantiene en <http://www.uk.research.att.com/facedatabase.html>
- [40] Face database Univ. Oulu <http://www.ee.oulu.fi/research/imag/color/pbfd.html>
- [41] Face database: UMIST <http://images.ee.umist.ac.uk/danny/database.html>
- [42] Feng J., B. Rivard, A. Sanchez-Azofeifa (2003) The topographic normalization of hyperspectral data: implications for the selection of spectral end members and lithologic mapping, *Remote Sensing of Environment* 85 221–231
- [43] Fernández E., M. Graña, J. Ruiz Cabello, (2004) On Evolution Strategies for Illumination Correction, *Information Sciences*, in press
- [44] Fernández E., M. Graña, J. Ruiz Cabello, (2004) Simultaneous adaptive segmentation and parametric correction of illumination inhomogeneities in MRI, *MEDSIP2004*.
- [45] Fernández E., M. Graña, J. Ruiz-Cabello, (2004) On a Gradient Based Evolution Strategy for Parametric Illumination Correction, *Electronic Letters*, 2004, 40 (9) , p.531-32.
- [46] Fernández E., M. Graña, J. Ruiz-Cabello, (2004) An Instantaneous Memetic Algorithm for Illumination Correction, *Proc. 2004 Int. Congress of Evolutionary Computation (CEC2004)*.

- [47] Fernández E., M. Graña, J. Ruiz-Cabello, (2004) On a Gradient-based Evolution Strategy for Parametric Illumination Correction, en *Information Processing with Evolutionary Algorithms*, M. Graña, R. Duro, A. d'Anjou, P.P. Wang (eds) Springer Verlag, London, pp.61-72
- [48] Finlayson G. , R. Xu (2003) Illuminant and gamma comprehensive normalisation in log RGB space, *Pattern Recognition Letters* 24 1679–1690
- [49] Geman, D.; Geman, S.; Graffigne, C.; Dong, P. (1990) Boundary detection by constrained optimization, *IEEE Trans. Patt. Anal. Mach. Int.* 12(7):609 - 628
- [50] Geman, D.; Reynolds, G. (1992) Constrained restoration and the recovery of discontinuities, *IEEE Trans. Patt. Anal. Mach. Int.* 14(3) 367 - 383
- [51] Georgiades, A.S.; Belhumeur, P.N.; Kriegman, D.J. (2001), From few to many: illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(6):643 - 660
- [52] Gerig, G.; Kubler, O.; Kikinis, R.; Jolesz, F.A. (1992) Nonlinear anisotropic filtering of MRI data, *IEEE Trans. Medical Imaging* 11(2):221 - 232
- [53] J. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, T. Gevers (2003) Color constancy from physical principles, *Pattern Recognition Letters* 24(11):1653-1662
- [54] Glass J.O., W.E. Reddick, O. Goloubeva, V. Yo, R.G. Steen (2000) Hybrid artificial neural network segmentation of precise and accurate inversion recovery (PAIR) images from normal human brain, *Mag. Res. Imaging* 18:1245-1253.
- [55] Gonzalez R.C., R .E Woods, (1990) *Digital Image Processing*, Academic Press
- [56] Guillemaud M. , Brady R., (1997) Estimating the Bias Field of MR Images, *IEEE Trans. Med. Imaging*, 16(3); 238-251
- [57] Hall, L.O.; Bensaid, A.M.; Clarke, L.P.; Velthuizen, R.P.; Silbiger, M.S.; Bezdek, J.C. (1992), A comparison of neural network and fuzzy clustering

- techniques in segmenting magnetic resonance images of the brain, *IEEE Trans. Neural Networks* 3(5):672 - 682
- [58] Hall, L.O.; Ozyurt, I.B.; Bezdek, J.C. (1999) Clustering with a genetically optimized approach, *IEEE Trans. Evolutionary Computation*, 3(2):103 - 112
- [59] Harris R.K. (1983) *Nuclear Magnetic Resonance Spectroscopy*, Pitman, London
- [60] H.Z. Hel-Or, B.A. Wandell (2002) Object-based illumination classification, *Pattern Recognition* 35:1723–1732
- [61] Held K. , E. Rota Kops, B. J. Krause, W. M. Wells, III, R. Kikinis, H.W. Müller-Gärtner (1997) Markov Random Field Segmentation of Brain MR Images *IEEE Trans. Med. Imaging* 16(6):878-886
- [62] S. Hong, S. Grossberg (2004) A neuromorphic model for achromatic and chromatic surface representation of natural images , *Neural Networks*, 17(5-6):787-808
- [63] Hornak J.S. (1995) *The basics of MRI*, Rochester Institute of Technology, web published doc (<http://www.cis.rit.edu/htbooks/mri/>) (WW324.htm)
- [64] R.L. Hsu, Abdel-Mottaleb M., Jain, A.K. (2002) Face detection in color images *IEEE Trans. Patt. Anal. Mach. Int.* 24(5):696 - 706
- [65] Internet Brain Segmentation Repository
<http://www.cma.mgh.harvard.edu/ibsr/>
- [66] Jurie F., M. Dhome (2002) Real time tracking of 3D objects: an efficient and robust approach, *Pattern Recognition* 35:317–328
- [67] Kamber M., D. Collins, R. Shinghal, G. Francis, A. Evans (1992) Model based 3D segmentation of multiple sclerosis lesions in dual-echo MRI data, *SPIE Visualization Biomed. Comp.* vol. 1808
- [68] J. Kershaw, B. A. Ardekani, I. Kanno (1999) Application of Bayesian Inference to fMRI Data Analysis, *IEEE Trans. Med. Imaging* 18(12):1138-1153

- [69] G. Klancar, M. Kristan, R. Karba (2004) Wide-angle camera distortions and non-uniform illumination in mobile robot tracking, *Robotics and Autonomous Systems* 46(2):125-133
- [70] Kohn M.I., N.K. Tanna, G.T. Herman, S.M. Resnick, P.D. Mozley, R.E. Gur, A. Alavi, R.A. Zimmermann, R.C. Gur (1991) Analysis of brain and cerebrospinal fluid volumes with MR imaging. Part I. Methods, reliability and validation, *Radiology* 18:115-122
- [71] C. Kotropoulos, A. Tefas, I. Pitas (2000) Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions, *Pattern Recognition* 33:1935-1947
- [72] Kwan R.K.-S., A.C. Evans, and G.B. Pike, (1996) An extensible mri simulator for post-processing evaluation, in *Visualization in Biomedical Computing, Proc. VBC '96*,pp.135-140.
- [73] Kwan R.K.-S., A.C. Evans, G.B. Pike (1999) MRI simulation-based evaluation of image-processing and classification methods" *IEEE Trans. Med. Imaging.* 18(11):1085-97.
- [74] Van Laarhoven P. J. M. , E. H. L. Aarts, *Simulated Annealing: Theory and Applications*, Kluwer Academic Publishers, Amsterdam
- [75] La Cascia M., S. Sclaroff, V. Athitsos (2000) Fast, reliable head tracking under varying illumination: an approach based on registration of texture-mapped 3D models, *IEEE trans. Patt. Anal. Mach. Int.* 22(4):322-336
- [76] S.H. Lai (2004) Computation of optical flow under non-uniform brightness variations, *Pattern Recognition Letters*, 25(8):885-892
- [77] Li C.; Goldgof, D.B.; Hall, L.O. (1993) Knowledge-based classification and tissue labeling of MR images of human brain, *IEEE Trans. Medical Imaging*12(4):740 - 750
- [78] Van Leemput K.,F. Maes, D. Vandermeulen, P. Suetens (1999) Automated Model-Based Bias Field Correction of MR Images of the Brain, *IEEE Trans. Med. Imaging* 18(10):885-896

- [79] Van Leemput K., F. Maes, D. Vandermeulen, P. Suetens (1999) Automated Model-Based Tissue Classification of MR Images of the Brain, *IEEE Trans. Med. Imaging* 18(10):897-908
- [80] Liang Z.-P., P.C. Lauterbur (2000), *Principles of Magnetic Resonance Imaging. A signal processing perspective.* IEEE Press, New York
- [81] Likar B., M.A. Viergever, F. Pernus, (2001) Retrospective Correction of MR Intensity Inhomogeneity by Information Minimization. *IEEE Trans Med. Imag.* 20(12):1398-1410.
- [82] Liu C.; Wechsler, H. (2000) Evolutionary pursuit and its application to face recognition, *IEEE Patt. Anal. Mach. Int.* 22(6):570-582
- [83] Liu C.; Wechsler, H. (2002) Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition, *IEEE Trans. Image Processing* 11(4):467 - 476
- [84] Liu C. ; Wechsler, H. (2003) Independent component analysis of Gabor features for face recognition, *IEEE Trans. Neural Networks*, 14(4):919 - 928
- [85] Liu C. (2004) Enhanced independent component analysis and its application to content based face image retrieval, *IEEE trans. Syst. Man Cyb. part B* 34(2):1117-1127
- [86] Liu C. (2004) Gabor-based kernel PCA with fractional power polynomial models for face recognition, *IEEE PAMI* 26(5):572-581.
- [87] G. Lohmann (1998) Extracting Line Representations of Sulcal and Gyral Patterns in MR Images of the Human Brain, *IEEE Trans. Med. Imaging* 17(6):1040-1048.
- [88] Lu J., K.N. Plataniotis, A.N. Venetsanopoulos (2003), Face recognition using kernel direct discriminant analysis algorithms, *IEEE Trans Neural Networks*, 14(1):117-125
- [89] Lundervold, A.; Storvik, G. (1995) Segmentation of brain parenchyma and cerebrospinal fluid in multispectral magnetic resonance images, *IEEE Trans. Med. Imaging* 14(2):339 - 349

- [90] Macovski A., (1996) Noise in MRI, *Magn. Res. Med.* 36(3):494-497.
- [91] Marroquin, J.L.; Vemuri, B.C.; Botello, S.; Calderon, E.; Fernandez-Bouzas, A (2002) An accurate and efficient Bayesian method for automatic segmentation of brain MRI, , *IEEE Trans. Med. Imag.*, 21(8): 934 - 945
- [92] Martinkauppi B. (2002) *Face Colour Under Varying Illumination Analysis and Applications*, Oulu University Press, Oulu, Finland
- [93] Meyer, C.R.; Bland, P.H.; Pipe, J. (1995) Retrospective correction of intensity inhomogeneities in MRI, *IEEE Trans. Med. Imaging*14(1):36 - 41
- [94] Moon, N.; Bullitt, E.; van Leemput, K.; Gerig, G. (2002) Model-based brain and tumor segmentation, , 2002. *Proceedings. 16th International Conference on Pattern Recognition*, Vol. 1 Pages:528 - 531
- [95] Moscato P., (1999) Memetic algorithms:A short introduction, in *New Ideas in Optimization*, D. Corne, F. Glover, and M. Dorigo, Eds. Maidenhead:McGraw-Hill, pp. 219–234.
- [96] Moscato P., pagina web dedicada a algoritmos meméticos: http://www.densis.fee.unicamp.br/~moscato/memetic_home.html
- [97] Nandy D., J. Ben-Arie (2001) Shape from recognition: a novel approach for 3D face shape recovery, *IEEE Trans. Image Proc.* 10(2):206-216
- [98] Pappas, T.N. (1992) An adaptive clustering algorithm for image segmentation, *IEEE Trans. Signal Processing*, 40(4):901 – 914.
- [99] Pham D.L., J. L. Prince (1999) An adaptive fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneities, *Patt. Recog. Letters* 20:57-68
- [100] Pratt W. K. , (2001) *Digital Image Processing: PIKS Inside*, Wiley-Interscience, 3rd Edition
- [101] Parker J.A. (1990), *Image reconstruction in radiology*, CRC Press, Boca Raton, CA
- [102] Press W.H., S.A. Teukolski, W.T. Vetterling, B.P. Flannery (1988) *Numerical Recipes in C*, Cambridge University Press

- [103] Rajapakse J. C., J. N. Giedd, J. L. Rapoport (1997) Statistical Approach to Segmentation of Single-Channel Cerebral MR Images, *IEEE Trans. Med. Imaging* 16(2):176-186
- [104] Rajapakse J. C., F. Kruggel, (1998) Segmentation of MR images with intensity inhomogeneities, *Image and Vision Comp.*, 16(3):165-180.
- [105] W. E. Reddick, J. O. Glass, E. N. Cook, T. D. Elkin, R. J. Deaton, (1997) Automated Segmentation and Classification of Multispectral Magnetic Resonance Images of Brain Using Artificial Neural Networks, *IEEE Trans. Med. Imaging* 16(6):911-918
- [106] H. Rusinek et al. (1991) Alzheimer disease: measuring loss of cerebral gray matter with MR imaging, *Radiology* 178:109-114
- [107] Schempp W.J. (1998), *Magnetic Resonance Imaging*, John Wiley & Sons, New York, NY
- [108] Shashua, A.; Riklin-Raviv, T. (2001) The quotient image: class-based re-rendering and recognition with varying illuminations, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(2):129 - 139
- [109] Shen D., S. Moffat, S.M. Resnick, C. Davatzikos (2002) Measuring size and shape of the hippocampus in MR images using a deformable shape model, *NeuroImage* 15:422-434
- [110] Sim, T.; Baker, S.; Bsat, M. (2003) The CMU pose, illumination, and expression database, *IEEE Trans. Patt. Anal. Mach. Int.* 25(12):1615 - 1618
- [111] Sled J. G., P. Zijdenbos, A. C. Evans, (1997) A Comparison of Retrospective Intensity Non-uniformity Correction Methods for MRI, in *Information Processing in Medical Imaging* 1230:459-464, *Proceedings 15th Int. Conf. IMPI'97*.
- [112] Sled J.G., A.P. Zijdenbos, A.C. Evans, (1998) A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data. *IEEE Trans Med. Imag.*, 17(1):87-97.

- [113] Sled J.G., G.B. Pike, Standing-Wave and RF Penetration Artifacts Caused by Elliptic Geometry: An Electrodynamical Analysis of MRI. *IEEE Trans Med Imag*, 1998, 17(4):653-662.
- [114] Soltanian-Zadeh H., J.P. Windham, D.J. Peck, A.E. Yagle (1992) A comparative analysis of several transformations for enhancement and segmentation of magnetic resonance image scene sequences, *IEEE Trans. Med. Imaging* 11(3):302-318
- [115] H. Spitzer, S. Semo (2002) Color constancy: a biological model and its application for still and video images, *Pattern Recognition* 35:1645–1659
- [116] Styner M., G. Gerig, C. Brechbühler, G. Szekely, (2000) Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Trans. Med. Imag.* 19(3):153-165.
- [117] Sijbers J., A.J. den Dekker, a. van der Linden, M. Verhoye, D. van Dyck (1999), Adaptive anisotropic noise filtering for magnitude MR data, *Magnetic Resonance Imaging* 17(10):1533-1539
- [118] Suri, J.S.; Kecheng Liu; Singh, S.; Laxminarayan, S.N.; Xiaolan Zeng; Reden, L. (2002) Shape recovery algorithms using level sets in 2-D/3-D medical imagery: a state-of-the-art review, *IEEE Trans. Inform. Techn. in Biomedicine*, 6(1):8 - 28
- [119] Weisstein E. W. Legendre Polynomial. From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/LegendrePolynomial.html>
- [120] Wells III W.M., W.E.L. Grimson, R. Kikinis, F.A. Jolez, (1996) Adaptive Segmentation of MRI Data *IEEE Trans. Med. Imag.*, 15:429-442 .
- [121] Widrow B., S.D. Stearns (1985) Adaptive signal processing, Prentice-Hall: Engelwood Cliffs, NJ
- [122] Winkler G. (1995) Image analysis, random fields and dynamic Monte Carlo methods, New York: Springer Verlag
- [123] Wu Y., M.L. Giger, K. Doi, C.J. Vyborny, R.A. Schmidt, C. E. Metz (1993) Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer, *Radiology* 187:81-87

- [124] Yale face database A, <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>
- [125] Yale face database B <http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>
- [126] Zhang, Y.; Brady, M.; Smith, S. (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm, *IEEE Trans. Med. Imag.*, 20(1): 45 – 57
- [127] Zhang M.; Hall, L.O.; Goldgof, D.B. (2002) A generic knowledge-guided image segmentation and labeling system using fuzzy clustering algorithms, *IEEE Trans. Systems, Man and Cybernetics, Part B*, 32(5):571 - 582
- [128] J. Zhao , Y. Su, D. Wang, S. Luo (2003) Illumination ratio image: synthesizing and recognition with varying illuminations, *Pattern Recognition Letters* 24:2703–2710

ÍNDICE DE MATERIAS

- Agrupamiento Borroso, 39, 41, 50
- Algoritmo EM, 59
- Algoritmos de correccion
 - adaptativo de Wells, 56
 - BFCM, 50
 - BMAP, 42, 48
 - GradClassLeg, 71
 - GradPABIC, 70
 - PABIC, 60
- Algoritmos Memeticos, 161
- Analisis de sensibilidad
 - BMAP, 87, 115
 - GradClassLeg, 113
 - Wells, 89, 114
- Artefactos en IRM, 151
- Campo Markoviano, 43
- Coficiente de pertenencia borroso, 50
- Densidad de probabilidad
 - a posteriori, 45, 57
 - a priori, 43, 57
 - Condicional, 43, 56, 57
- Descenso de gradiente, 71, 126
- Estrategias Evolutivas, 159
- Filtrado Homomorfico, 30, 76
- ICM: Iterated Conditional Modes, 38, 41, 45
- Imagenes experimentales
 - Cerebro sintetico, 91
 - IBSR, 108
 - Tablero de ajedrez, 74
 - Yale face database, 127
- IRM funcional, 38, 40
- K-medias, 27, 48, 49
- Maxima probabilidad a posteriori, MAP, 37, 41, 44, 57
- Modelo de iluminacion, 22, 42, 50, 56, 60
- Polinomios de Legendre, 62, 67, 155
- Reconocimiento de caras, 123
- Relacion senal ruido (SNR), 27, 62, 153
- Retinex, 21
- Ruido aditivo, 23, 24, 27, 28, 43, 56, 61, 153
- Secuencia de pulsos
 - Espin-eco, 142
 - Recuperacion de la inversion, 142
- Simulated Annealing, 65
- Sistemas de Coordenadas, 151
- Tiempo de relajacion
 - T1, 139
 - T2, 140
- Vecindario de un pixel, 44