



Universitat d'Alacant

Universidad de Alicante

Esta tesis doctoral contiene un índice que enlaza a cada uno de los capítulos de la misma.

Existen asimismo botones de retorno al índice al principio y final de cada uno de los capítulos.

[Ir directamente al índice](#)

Para una correcta visualización del texto es necesaria la versión de [Adobe Acrobat Reader 7.0](#) o posteriores

Aquesta tesi doctoral conté un índex que enllaça a cadascun dels capítols. Existeixen així mateix botons de retorn a l'índex al principi i final de cadascun dels capítols .

[Anar directament a l'índex](#)

Per a una correcta visualització del text és necessària la versió d' [Adobe Acrobat Reader 7.0](#) o posteriors.



Universitat d'Alacant
Universidad de Alicante

Tesis Doctoral

**Plantillas Deformables Extendidas:
Modelización Local Robusta y
Caracterización Basada en Auto-Modelos
para el Reconocimiento y Tracking de
Estructuras Geométricas Activas**

Presentada por

Francisco Escolano Ruiz

Dirigida por

Dr. Ramón Rizo Aldeguer

Programa de Doctorado *Sistemas Informáticos*
Departamento de Tecnología Informática y Computación
Universidad de Alicante

10 de junio de 1997

1º Bº
Ramón Rizo Aldeguer
RAMÓN RIZO ALDEGUER





Universitat d'Alacant
Universidad de Alicante

*A Mamen,
por estar ahí...*



Universitat d'Alacant
Universidad de Alicante

Agradecimientos

El desarrollo de un proyecto de investigación no es un hecho aislado sino que se enmarca en los objetivos y planteamientos de un grupo de investigación y, al mismo tiempo, los resultados obtenidos revierten en éste. Quisiera expresar mi gratitud sincera a todos aquellos que han colaborado directa e indirectamente en la concepción y desarrollo del presente trabajo: a mi director, *Ramón Rizo*, por su apoyo desde el principio, a todos los miembros y amigos del grupo *i3a* (*Juan Puchol*, *Bernardo Ledesma*, *Mar Pujol*, *Rosana Satorre*, *Faraón Llorens*, *Otto Colomina*, *Ginés López*, etc..) y finalmente a *Miguel Cazorla* y *Domingo Gallardo* por su inestimable colaboración y ayuda.



Universitat d'Alacant
Universidad de Alicante

La correspondencia es el proceso que identifica elementos en vistas diferentes como representativos del mismo objeto en distintos instantes de tiempo, manteniendo por ello la identidad perceptual de los objetos sujetos a movimiento o cambio.

Shimon Ullman, *The Interpretation of Visual Motion*, 1979



Universitat d'Alacant
Universidad de Alicante

Índice General

I	Introducción: Objetivos y Estado del Arte	15
1	Presentación, Objetivos y Ámbito de Desarrollo	17
1.1	Presentación	17
1.2	Objetivos: Descripción de la Propuesta	18
1.3	Ámbito de Desarrollo	21
2	Tracking Basado en Modelos Deformables	25
2.1	Introducción	25
2.2	Modelos Deformables y Mecanismos de Tracking	27
2.2.1	Modelos Deformables: Características Generales	27
2.2.2	Mecanismos de Tracking sobre Modelos Deformables	27
2.3	Revisión de Propuestas Previas	28
2.3.1	Modelización Física: Contornos Activos	28
2.3.2	Métodos de Elementos Finitos y Componentes Principales	31
2.3.3	Modelos Algebraicos Basados en Interpolantes	37
2.3.4	Modelos Basados en Superficies Super-Cuádricas	39
2.3.5	Plantillas Deformables Basadas en Modelos Bayesianos	41
2.4	Conclusiones al Análisis de Modelos y Técnicas	45
II	Plantillas Deformables Extendidas	49
3	Caracterización Local: Plantillas Super-Elipsoydales	51
3.1	Campos de Potencial Normalizados	51
3.2	Plantillas Deformables Básicas	53
3.2.1	Plantillas Circulares	53
3.2.2	Ejemplo de Aplicación	55
3.3	Generalización: Plantillas Super-Elipsoydales	56
3.3.1	Directrices de Modelización	56

3.3.2	Definición: Potencial y Función de Energía	56
3.3.3	Dinámicas de Centrado y Rotación	59
3.3.4	Dinámicas Axiales y de Forma	61
3.3.5	Revisión de la Condición de Equilibrio	63
3.4	Experimentación: Análisis Local	64
3.4.1	Objetivos y Condiciones Experimentales	64
3.4.2	Análisis y Discusión	64
3.5	Consideraciones	65
4	Caracterización Estructural: Auto-Restricciones	71
4.1	Modelado Basado en Auto-Restricciones	71
4.1.1	Directrices de Modelización	71
4.1.2	Auto-Restricciones Espaciales Absolutas y Relativas	73
4.1.3	Auto-Restricciones Espacio-Temporales	75
4.2	Aprendizaje: Auto-Restricciones Finales	75
4.2.1	Aprendizaje de Auto-Restricciones	75
4.2.2	Entropía de Restricción	76
4.2.3	Genericidad de Restricción	76
4.2.4	Ejemplo en el Dominio Espacial	77
4.2.5	Ejemplo en el Dominio Espacio-Temporal	78
4.3	Satisfacción de Auto-Restricciones	83
4.3.1	Grado de Satisfacción de Auto-Restricciones	83
4.3.2	Relajación de Restricciones y Nivel de Confianza	84
4.3.3	Ejemplo de Reconocimiento: Dominio Espacial	85
4.3.4	Ejemplo de Reconocimiento: Dominio Espacio-Temporal	86
4.4	Eficiencia de las Auto-Representaciones	88
4.5	Consideraciones	89
5	Tracking Basado en Plantillas Extendidas	91
5.1	Mecanismo de Correspondencia Espacial	91
5.1.1	Directrices de Modelización	91
5.1.2	Minimización de Energía Local	93
5.1.3	Maximización del Grado de Satisfacción	94
5.1.4	Dinámica de Restricciones Absolutas	94
5.1.5	Dinámica de Restricciones Relativas	94
5.1.6	Estructuras Geométricas Óptimas	96
5.1.7	Esquema de Descenso Supervisado	98
5.1.8	Algoritmo de Minimización por Supervisión	98
5.1.9	Ejemplo Básico de Correspondencia	99
5.2	Experimentación en el Dominio Espacial	100

ÍNDICE GENERAL	11
5.2.1	Objetivos y Condiciones Experimentales 100
5.2.2	Análisis y Discusión 101
5.3	Tracking/Correspondencia Espacio-Temporal 107
5.3.1	Planteamiento/Elementos del Tracker 107
5.3.2	Estructuras Espacio-Temporales Óptimas 107
5.3.3	Extensión del Esquema de Descenso Supervisado 107
5.3.4	Algoritmo de Tracking Supervisado con Ventana Unitaria . 108
5.4	Experimentación en el Dominio Espacio-Temporal 110
5.4.1	Objetivos y Condiciones Experimentales 110
5.4.2	Análisis y Discusión 111
5.5	Consideraciones 111
III Dominios de Aplicación	115
Conclusiones	115
6 Dominios de Aplicación	117
6.1	Análisis de Secuencias Intravasculares 117
6.1.1	Introducción y Motivación 117
6.1.2	Modelo de Imagen: Campos de Potencial 120
6.1.3	Extracción de la Pared Arterial: Plantillas Circulares . . . 120
6.1.4	Plantillas Elípticas: Tracking en Angioplastia 121
6.1.5	Análisis Espacio-Temporal: Auto-Restricciones 124
6.2	Reconocimiento Basado en Tracking Gestual 130
6.2.1	Introducción y Motivación 130
6.2.2	Reconocimiento mediante Plantillas Extendidas 131
6.2.3	Modelo de Imagen: Campos de Potencial 132
6.2.4	Tracking Estructural y Reconocimiento 132
6.3	Consideraciones 134
7 Conclusiones y Trabajos Futuros	137
7.1	Aportaciones 137
7.2	Líneas de Continuación 138
7.3	Consideraciones Finales 139
IV Apéndices	141
A Experimentación con Contornos Activos	143
A.1	Reconstrucción Basada en Contornos Activos 143

A.1.1	Reconstrucción Cerebral a partir de Imagen RNM	144
A.1.2	Reconstrucción Ocular a partir de Imagen Microscópica	149
A.2	Análisis del Cansancio Ocular Basado en Snakes	153
A.3	Conclusiones	156
B	Transformada de Karhunen-Loève	157
B.1	Transformada K-L Discreta	157
B.1.1	Transformación Lineal y Ortonormal	157
B.1.2	Selección de Características	158
B.2	Propiedades Relevantes	159
B.2.1	Auto-Espacio y Componentes Principales	159
B.2.2	Efectividad y Otras Propiedades	160
B.3	Ejemplos de Aplicación	162
B.3.1	Modelización Radial del Objeto	162
B.3.2	Experimentación: Obtención de la Transformada K-L	162
V	Referencias	167



Universitat d'Alacant
Universidad de Alicante

Prefacio

¿Cuál es el objetivo de la Visión Artificial?. En los últimos diez años la Visión Activa, enfoque que emana principalmete de los trabajos de Aloimonos (1988), Bajcsy (1988), Ballard (1991) y Krotkow (1987), ha revolucionado el planteamiento clásico, representado por los trabajos de Marr (1982). Este nuevo planteamiento del problema supone un giro conceptual motivado por criterios de eficiencia y caracterizado por los siguientes aspectos:

1. *Percepción Guiada por la Acción*: se trata de que el sistema visual controle la posición del sensor (observador activo), realizando una exploración selectiva (estrategias de observación) de la escena orientada a obtener la información estrictamente necesaria para realizar una tarea concreta. Para ello es preciso diseñar algoritmos que gestionen de forma eficiente los recursos computacionales activando los módulos necesarios.
2. *Tratamiento de Información Dinámica*: el procesamiento de secuencias de imágenes juega un papel central en la resolución del problema perceptual (estructura a partir del movimiento, *tracking*, estéreo, flujo óptico), abandonando dicha fuente de información el carácter auxiliar (secuencias cortas tomadas desde una cámara pasiva) que tenía anteriormente.
3. *Procesamiento Robusto*: en entornos dinámicos, se introduce el control de la incertidumbre asociada a la posición y desplazamiento del sensor, además de la incertidumbre propia de la captación (p.e. calibración automática).
4. *Acercamiento al Planteamiento Biológico*: aunque en los trabajos de Marr ya se reconocía la importancia del estudio del sistema visual humano¹, es en el nuevo enfoque donde esta aproximación puede reportar rendimientos óptimos (Landy et al., 1996):

(a) *Moveo ergo Video*: en la actualidad se encuentra en estudio el impacto del *movimiento ocular* en el rendimiento del sistema visual.

¹Aproximación Neuromórfica o Antropomórfica a la Visión Artificial

- (b) *Cogito ergo Moveo*: así mismo, el análisis del control cognitivo del movimiento ocular pretende examinar las capacidades, estrategias perceptuales humanas y beneficios cualitativos y cuantitativos reportados por el movimiento.
5. *Representaciones Autónomas/Activas*: los nuevos modelos de representación implementan mecanismos interactivos de búsqueda local que permiten la adaptación eficiente a distribuciones concretas de intensidad, aportando soluciones basadas en optimización de funciones de energía (diseño de heurísticas) a la resolución de problemas de segmentación, reconstrucción, estéreo y *tracking*.

Una de las problemáticas centrales de la Visión Activa (Christensen et al., 1993), (Crowley y Christensen, 1994), y en la cual están presentes todos los aspectos que acabamos de enumerar, es el *tracking* o seguimiento de objetos o regiones de interés presentes en la escena (Blake y Yuille, 1992). El objetivo es extraer, de forma selectiva, información paramétrica de utilidad para la tarea a realizar (navegación, reconocimiento, etc.). En este sentido el uso de representaciones activas, y en definitiva, el diseño de funciones de energía para el reconocimiento de patrones está cobrando un auge importante en el ámbito de la Visión Artificial (Pelillo y Hancock, 1997). Estos mecanismos han resultado ser de gran aplicación para simplificar el problema de correspondencia implícito en el *tracking*. La integración de modelos predictivos espacio-temporales con este tipo de representación es un tema de gran interés y constituye el objetivo central de esta tesis.



Universitat d'Alacant
Universidad de Alicante

Parte I

Introducción: Objetivos y Estado del Arte



Universitat d'Alacant
Universidad de Alicante

Capítulo 1

Presentación, Objetivos y Ámbito de Desarrollo

La *motivación científica* del presente trabajo surge del creciente interés propio del ámbito de la Visión Artificial, en general, y de la Visión Activa en particular, en el desarrollo de modelos computacionales (representaciones y técnicas de búsqueda) eficientes para el tratamiento dinámico de la percepción, habida cuenta de la relevancia de esta capacidad en los sistemas inteligentes actuales y futuros.

Esta tesis se orienta fundamentalmente hacia el avance en *la modelización y adquisición de estructuras (conjuntos de primitivas) de geometría variable* como soporte para la resolución del problema computacional de la *correspondencia* y del *tracking no-rígido*¹ formulado como *la predicción y extracción automáticas, usualmente mediante el uso explícito de modelos dinámicos de forma y movimiento, de los parámetros que caracterizan la evolución espacio-temporal de un objeto o de un conjunto de entidades relevantes de la escena.*

1.1 Presentación

En términos genéricos, *las dimensiones de dificultad* de todo problema de análisis (*tracking* y/o reconocimiento) de escenas dinámicas, son las siguientes (O'Rourke y Badler, 1980):

1. *Complejidad de los Objetos*: nivel de abstracción relativo al modelo de objeto o entidad básica en la escena. Cabe considerar los siguientes niveles de complejidad creciente: puntos, aristas, curvas cuadráticas y formas genéricas

¹*Tracking*: suele traducirse por *seguimiento*. No obstante mantendremos la nomenclatura original por su mayor riqueza de significado.

18 CAPÍTULO 1. PRESENTACIÓN, OBJETIVOS Y ÁMBITO DE DESARROLLO

(morfología libre). Asimismo el movimiento puede implicar a objetos 2D o 3D.

2. *Número de Objetos*: cantidad de entidades relevantes presentes en la escena tratables de forma conjunta. Podemos trabajar a nivel de objetos individuales o bien a nivel de estructuras geométricas.
3. *Tipo de Movimiento de los Objetos*: en términos generales cabe considerar movimiento rígido (expresable mediante transformaciones afines, es decir, traslación, rotación y escalado), no-rígido global (implica transformaciones no-afines a nivel global) y no-rígido local (exige transformaciones generales). Estas restricciones de movimiento son aplicables a nivel de cada objeto individual o a nivel de escena.
4. *Nivel de Interpretación*: nivel de abstracción de la información elaborada para el usuario final. Está relacionada con la capacidad del sistema de responder a preguntas sobre la secuencia y con el ámbito de las mismas. En este sentido proponemos considerar los niveles: bajo (movimiento absoluto), medio (nivel sintáctico y consideración de movimiento relativo-estructural) y alto (incluye aproximaciones conceptuales, semánticas y/o cognitivas).

Dichas dimensiones están intrínsecamente relacionadas existiendo un *trade-off* o balance mútuo: un incremento de complejidad en una de ellas repercute en el resto hasta el punto de comprometer la tratabilidad del problema. Por ello deben especificarse niveles razonables. Por ejemplo, el aumento de la complejidad de los objetos viene acompañado de modelos de movimiento también más complejos, lo cual induce la necesidad de limitar el número de entidades, y en consecuencia de interacciones (p.e. oclusión parcial) a tratar, hasta el punto de que los objetos de morfología libre suelen considerarse de manera aislada.

En este sentido nuestra propuesta se orienta al *tratamiento, a nivel de interpretación medio (movimiento relativo-relacional) de escenas compuestas por un conjunto de objetos 2D de complejidad super-elipsoidal, con capacidad de movimiento no-rígido local (traslación, rotación, escalado y cambio de forma parametrizado) a nivel individual, y de movimiento no-rígido global a nivel estructural.*

1.2 Objetivos: Descripción de la Propuesta

El presente trabajo articula (tal como se muestra en la Figura 1.1) el *Modelo de Plantillas Deformables Extendidas* contenido en la parte central de la memoria que, como mostraremos, introduce diversos avances en los problemas objeto de estudio:

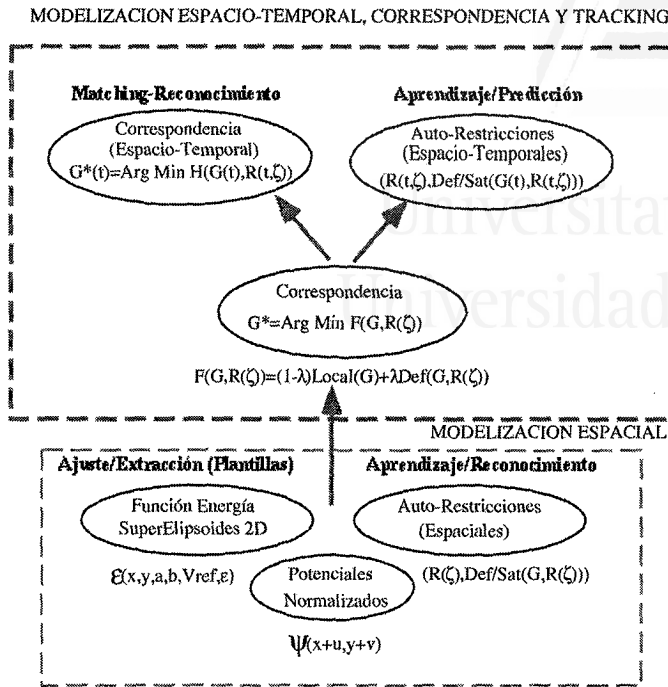


Figura 1.1: Esquema de la propuesta. Elementos principales.

Desde el punto de vista de la *modelización espacial* se pretende aportar:

1. *Plantillas Deformables Super-Elipsoidales*: extensión para el tratamiento automático de formas 2D super-elipsoidales mediante información de intensidad y gradiente sobre potenciales normalizados. Formulación de una función de energía (\mathcal{E}) con un buen comportamiento local (robustez) en lo relativo al control parametrizado del ruido (Capítulo 3).
2. *Auto-Restricciones Geométricas Espaciales*: definición y aprendizaje automático de relaciones geométricas relevantes (conjunto $\mathcal{R}(\zeta)$) entre las plantillas del modelo (conjunto \mathcal{G}) y expresión probabilística del nivel de satisfacibilidad o bien de deficiencia (reconocimiento de una estructura dada). Para ello se toma como soporte formal *la Transformada de Karhunen-Loève* o Análisis de Componentes Principales (ACP) (Capítulo 4).

En cuanto a *modelización espacio-temporal*, *mecanismos de correspondencia y tracking* aportaremos:

20 CAPÍTULO 1. PRESENTACIÓN, OBJETIVOS Y ÁMBITO DE DESARROLLO

1. *Auto-Restricciones Geométricas Espacio-Temporales*: extendiendo las auto-restricciones al caso temporal ($\mathcal{R}(t, \zeta)$) aportamos un mecanismo de caracterización de la secuencia, en el cual se enfatiza la evolución de las relaciones entre objetos. Ello nos permite adiestrar los mecanismos predictivos del *tracking* obteniéndose los parámetros más representativos (Capítulo 4).
2. *Ajuste/Minimización de Energía*: adaptación del modelo deformable integrando de forma ponderada en $F(\mathcal{G}, \mathcal{R}(\zeta))$ el comportamiento local modelado por las distintas \mathcal{E}_i con la satisfacibilidad de las restricciones del modelo estimada mediante proyecciones sobre espacios de dimensión reducida (auto-espacios) (Capítulo 5).
3. *Correspondencia Espacio-Temporal*: trasladando el mecanismo espacial al tratamiento de secuencias integramos la estimación instantánea con el mecanismo predictivo basado en auto-restricciones ($H(\mathcal{G}(t), \mathcal{R}(t, \zeta))$) para estimar la evolución óptima $\mathcal{G}^*(t)$ (Capítulo 5).

El desarrollo del modelo se presenta de forma constructiva. En primer lugar se define el modelo local robusto (Plantillas Super-Elipsoidales) y su mecanismo de ajuste. A continuación se abordan la caracterización y el aprendizaje, mediante ACP, de estructuras (espaciales o espacio-temporales) formadas por conjuntos de plantillas. Seguidamente se integra el modelo local y el estructural para dar solución al problema de correspondencia (espacial y espacio-temporal).

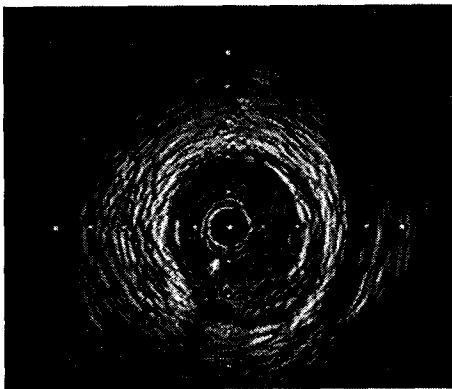


Figura 1.2: Tracking Intravascular (izqda.) e Interpretación Gestual (dcha.)

Desde el punto de vista aplicado mostraremos la adecuación del modelo computacional de Plantillas Deformables Extendidas para avanzar la resolución de dos problemas de interés (ver Figura 1.2) abordados en el Capítulo 6:



Figura 1.3: Reconstrucción Cerebral (izqda.) y Ocular (centro). Análisis de Cansancio (dcha.)

1. *Reconstrucción y Análisis de Órganos:* concretamente en el *tracking* de secciones intravasculares y en el reconocimiento de pautas de deformación arterial en procedimientos médicos como la angioplastia coronaria.
2. *Interpretación Gestual:* abordaremos el tracking de gestos humanos mediante plantillas y desde un punto de vista estructural tomando como referencia el modelo propuesto.

Adicionalmente dicho modelo es aplicable en muchos otros contextos, en combinación con modelos deformables basados en Contornos Activos, (ver Figura 1.3).

El resto de la memoria se estructura como sigue. En el Capítulo 2 abordamos el análisis del *tracking* desde la perspectiva de los modelos deformables así como el estado de las representaciones basadas en ACP y su aplicabilidad a este problema. Una vez presentada la propuesta y sus aplicaciones (Capítulos 3 al 6) y las conclusiones del trabajo (Capítulo 7) finalizamos el documento con dos apéndices que complementan su contenido. En el primero de ellos presentamos diversos resultados experimentales derivados de la aplicación de Contornos Activos a contextos fundamentalmente biomédicos. En el segundo revisamos los fundamentos teóricos del ACP dada la relevancia de este formalismo en nuestra propuesta.

1.3 Ámbito de Desarrollo

Esta tesis está enmarcada en los objetivos del *Grupo de Investigación de Informática Industrial e Inteligencia Artificial (i³a)* del Departamento de Tecnología Informática y Computación de la Universidad de Alicante. En este sentido nuestro trabajo es, *en primer lugar*, fruto de las aportaciones previas realizadas en los últimos cuatro años en el ámbito de diversos proyectos de investigación, orienta-

22 CAPÍTULO 1. PRESENTACIÓN, OBJETIVOS Y ÁMBITO DE DESARROLLO

dos a producción industrial y a biomedicina, que avalan, por otro lado, su interés científico:

1. *Proyecto INTELIPACK*²: cuyo objetivo fue el desarrollo de interfaces inteligentes de posicionamiento. Motivó el estudio de modelos computacionales no-clásicos y de métodos eficientes de Análisis de Imagen (Rizo, 1991), (Rizo et al., 1993), (García, 1993).
2. *Proyecto TDIMAP*³: orientado a la integración de la Morfología Matemática (Guijarro et al., 1993), (Cediel et al., 1993) con Modelos de Computación Altamente Paralelos.
3. *Proyecto MAMOGRAF*⁴: orientado a la detección automática de nódulos malignos y micro-calcificaciones en Mamografías (Ibarra, 1995). Este proyecto tiene una gran influencia en el presente trabajo:
 - (a) En primer lugar *pone de relieve las ventajas de la integración* de diversas metodologías de análisis (Lógica Difusa, Morfología Matemática y Redes Neuronales (Molina et al., 1995), (Florez et al., 1995)).
 - (b) Introducimos el *tratamiento de escenas dinámicas* (evolución de tumores ante tratamiento) (Peñalver et al., 1995) y, por extensión a otras problemáticas biomédicas, el uso de *modelos deformables* como herramienta de segmentación y *tracking* (Cazorla et al., 1995), (Sabuco et al., 1997), (López et al., 1997).
 - (c) Tras una experimentación inicial con Contornos Activos optamos por profundizar en el estudio y extensión de las Plantillas Deformables que, a nuestro juicio, mantienen un mejor compromiso *coste computacional-capacidad expresiva* en el tratamiento de conjuntos de objetos. La aportación al *tracking intravascular* presentada en esta memoria es una consecuencia directa de este planteamiento (Escolano et al., 1997a), (Escolano et al., 1997b).

En segundo lugar este trabajo supone un punto de partida, a nivel de nuestro grupo de investigación, hacia el tratamiento de problemáticas de Visión Activa, en donde el análisis de secuencias basado en modelos deformables juegan un papel central. Actualmente está en desarrollo el *Proyecto SICVA*⁵ orientado hacia el diseño de un vehículo autónomo guiado por Visión. El avance en mecanismos de

² *Empaquetamiento Inteligente de Patrones*: Proyecto DGICYT PB92-0343.

³ *Tratamiento de Imagen: Modelos Altamente Paralelos*: Proyecto Subvencionado por IBM.

⁴ *Detección y Clasificación de Tumores en Mamografías*: Proyecto GV-2215/94.

⁵ *Sistemas de Control Integrado para Vehículos Autónomos*: Proyecto IMPIVA.

1.3. ÁMBITO DE DESARROLLO

23

tracking y estimación del movimiento (Bogas et al.,1995), caracterización y reconocimiento de la secuencia es de gran importancia para la obtención de soluciones eficientes al problema de la navegación autónoma.

Universitat d'Alacant
Universidad de Alicante

24 CAPÍTULO 1. PRESENTACIÓN, OBJETIVOS Y ÁMBITO DE DESARROLLO



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Capítulo 2

Tracking Basado en Modelos Deformables

2.1 Introducción

El problema del *tracking no-rígido* suele abordarse mediante *mecanismos de correspondencia entre tokens o entidades relevantes* (Ullman, 1979). En definitiva este problema se puede formular como la *extracción/estimación de la función espacio-temporal óptima (según algún criterio) que asocia a cada token una trayectoria que codifica su evolución*. Dicho criterio suele venir cuantificado de forma heurística (Cédras y Shah, 1995) y estar condicionado, al menos, por dos restricciones básicas que simplifican y hacen tratable el problema:

1. *Coherencia del Movimiento*: el movimiento obtenido debe ser consistente a *nivel local* (las hipótesis de asignación de instancias consecutivas de objeto deben respetar similitud) y a *nivel global* (el conjunto de asignaciones locales debe generar variabilidad estructural mínima)¹.
2. *Continuidad del Movimiento*: se asume que la variabilidad espacio-temporal instantánea de los objetos está acotada es decir que no se producen cambios bruscos (al menos no superiores a un cierto umbral) en la secuencia. Por lo tanto el problema es predecible con un cierto margen de incertidumbre y es posible obtener trayectorias finales razonablemente suaves.

Teniendo en consideración estos aspectos junto con las dimensiones del problema previamente especificadas, *nos centraremos en el seguimiento de múltiples*

¹Esta doble consistencia es el eje del mecanismo de correspondencia en la medida en que el nivel global regulariza o elimina la ambigüedad local resolviendo el problema de apertura.

objetos globalmente deformables delimitados por contornos cerrados. Por este motivo, y siendo conscientes de la diversidad de enfoques desde los que es abordable el problema del *tracking* (Blake y Yuille, 1992),(Blake et al., 1995),(Mitiche y Bouthemy, 1996) únicamente trataremos aquellas propuestas orientadas al *seguimiento de contornos deformables*. Desde este punto de vista la especificación de un *tracker* o mecanismo de seguimiento va a conjugar dos aspectos básicos:

1. *Modelo Deformable*: definido a su vez por los siguientes elementos:
 - (a) Una *Estructura Geométrica*: compuesta, según el caso, por un conjunto de puntos, primitivas o parámetros y caracterizada por una función objetivo denominada *de potencial o energía*.
 - (b) Un *Modelo de Imagen*: conjunto de distribuciones de intensidad o textura a las que la estructura es sensible. Define el potencial externo ante el que debe reaccionar la estructura.
 - (c) Un *Algoritmo de Ajuste/Matching*: tal que, a partir de una estructura inicial posicionada sobre la imagen, obtiene la *estructura geométrica óptima* es decir aquella que satisface al máximo los condicionamientos estructurales internos al tiempo que se adapta satisfactoriamente al modelo de imagen. Este problema suele abordarse mediante mecanismos numéricos, técnicas de búsqueda o bien conjugando ambas aproximaciones.

2. *Mecanismo de Estimación*: la extracción del movimiento de los contornos debe realizarse de forma robusta (controlando la incertidumbre). El ciclo de control consta de dos fases:
 - (a) *Predicción*: a partir de la posición actual de la estructura, antidiendo al conocimiento aprendido acerca del fenómeno dinámico y de acuerdo con un modelo de incertidumbre, se plantea una extrapolación acerca de su siguiente configuración.
 - (b) *Contraste/Actualización*: comprobación de la hipótesis aplicando el algoritmo de ajuste en las condiciones predichas. Como consecuencia se actualiza el modelo de incertidumbre y con ello el conocimiento actual acerca del fenómeno.

2.2 Modelos Deformables y Mecanismos de Tracking

2.2.1 Modelos Deformables: Características Generales

En términos generales los modelos que abordaremos a continuación se ajustan a un perfil concretado en las características siguientes:

1. *Distribución y Grados de Libertad*: a nivel de capacidad expresiva o de representación el espectro de modelos va desde planteamientos totalmente distribuidos como los *Snakes* (autonomía de punto) hasta aproximaciones basadas en *plantillas deformables* (autonomía de primitiva).
2. *Actividad y Dinámica*: se trata de modelos *activos* ya que reaccionan ante *fuerzas* o condicionantes extrínsecos (modelos de imagen, restricciones controladas, etc.). En consecuencia cada modelo incorpora un mecanismo físico, frecuencial o probabilístico capaz de expresar su dinámica y una heurística o función de potencial que especifica los criterios morfológicos y de imagen que deben guiar el ajuste o *matching*.
3. *Globalidad/Localidad*: usualmente la convergencia de los modelos hacia configuraciones óptimas depende de una inicialización correcta² de los mismos y de la utilización de un método de búsqueda o resolución adecuado. En este sentido estos modelos se diseñan como mecanismos de bajo nivel y su uso correcto debe ser establecido en niveles de abstracción superiores que enfatizen la dimensión estructural del problema.

2.2.2 Mecanismos de Tracking sobre Modelos Deformables

El *tracking* de contornos suele abordarse desde la expresión del *estado actual* en términos del *conjunto de parámetros que caracterizan a la estructura* y la definición de la *conducta temporal* tomando como base ecuaciones de movimiento. Cada mecanismo concreto trata de estimar, con la mejor fidelidad posible, las ecuaciones reales de movimiento. Cabe considerar los siguientes mecanismos típicos de estimación:

1. *Filtrado Nulo*: no se implementa mecanismo predictivo alguno. En cada instante de tiempo se inicia el ajuste a partir del último posicionamiento. Es el tratamiento más simple pero genera trayectorias irregulares (no-suaves) con lo que es necesario aplicar un filtrado de la señal espacio-temporal a posteriori.

²Mediante mecanismos de búsqueda a varias resoluciones o bien mediante variantes de la Transformada de Hough.

2. *Filtrado de Kalman Básico:* (Kalman, 1960), (Gelb, 1974) se basa en la introducción de dos modelos, usualmente lineales ³, uno de sistema y otro de sensor, con sus modelos de error asociados. El error global se ajusta de forma iterativa a medida que incorporamos mediciones o conocimiento acerca de la secuencia. Este esquema permite generar trayectorias suaves (Shah et al., 1993) ya que el filtrado espacio-temporal está implícito. No obstante estos esquemas asumen distribuciones de probabilidad gaussianas. Recientemente se ha propuesto un mecanismo de filtrado bayesiano que permite salvar este problema (Isard y Blake, 1996).
3. *Métodos Frecuenciales/Espectrales:* en determinados casos se mejora la conducta del filtrado de Kalman introduciendo conocimiento adicional (modelo temporal como p.e. ciclicidad) codificado como una expansión frecuencial, modal o armónica o basada en ACP.
4. *Otros Métodos:* otra posibilidad de mejora es modelar la conducta temporal mediante un proceso estocástico, mediante aplicación de teoría de control o mediante integración de ambos aspectos.

A continuación analizaremos los enfoques más relevantes en el ámbito de los modelos deformables, introduciendo, para cada propuesta, los aspectos formales básicos, sus contextos de aplicación y su integración en esquemas de *tracking* de contornos.

2.3 Revisión de Propuestas Previas

2.3.1 Modelización Física: Contornos Activos

Definición

Los *Contornos Activos* o *Snakes* (Kass et al., 1987) introducen el *enfoque activo* en la resolución de problemas de extracción de aristas, *matching* estéreo, y *tracking* de movimiento no-rígido. Estos modelos generalizan, desde un planteamiento variacional e interactivo los mecanismos previos de detección de curvas basados en la optimización del contraste o intensidad y en la uniformidad del contorno, tales como la *resolución directa* (Montanari, 1971), la *búsqueda heurística* (Martelli, 1972), el *etiquetado por relajación* (Zucker, 1977) o el *matching elástico* (Burr, 1986).

Los *snakes* son contornos deformables planos parametrizados $\Gamma(s)$ cuyo comportamiento dinámico está regido por *funciones de energía* (o fuerzas) de manera

³La no-linealidad se aproxima mediante una serie de Taylor (Filtros de Kalman Extendidos).

que su configuración óptima es aquella que minimiza la energía global, la cual se define usualmente de acuerdo con las siguientes ecuaciones:

$$\Gamma : [0, 1] \rightarrow \mathfrak{R}^2, \Gamma(s) = (x(s), y(s)) \quad (2.1)$$

$$\mathcal{E}(\Gamma(s)) = \int_0^1 (\mathcal{E}_{int}(\Gamma(s)) + \mathcal{E}_{ext}(\Gamma(s))) ds \quad (2.2)$$

$$\mathcal{E}_{int}(\Gamma(s)) = \alpha |\Gamma_s|^2 + \beta |\Gamma_{ss}|^2 \quad (2.3)$$

$$\mathcal{E}_{ext}(\Gamma(s)) = \pm \gamma \Psi(\Gamma(s)) \pm \lambda \mathcal{R}(\Gamma(s)) \quad (2.4)$$

donde $\mathcal{E}(\Gamma(s))$ define la energía global compuesta por dos términos principales⁴:

1. \mathcal{E}_{int} caracteriza el *comportamiento interno* del contorno imponiendo restricciones de forma. Se restringe al snake a una expansión uniforme tanto en distancia entre *snaxels*⁵ $|\Gamma_s|^2$ como en diferencias locales de curvatura $|\Gamma_{ss}|^2$. En definitiva α regula la tensión del snake y β su rigidez. En general se puede asignar un peso a cada snaxel. Esto resulta conveniente si se pretende diferenciar el comportamiento local⁶ con el consecuente incremento de coste computacional.
2. \mathcal{E}_{ext} define la incidencia ponderada de *factores externos* tales como el tipo de distribución de intensidad $\pm \gamma \Psi(\Gamma(s))$ o la existencia de condicionantes geométricos $\pm \lambda \mathcal{R}(\Gamma(s))$. En ambos casos el signo indica el sentido de la fuerza simulada (atracción- repulsión). En el primero el contorno puede ser atraído hacia zonas concretas del espacio de imagen⁷. Este potencial puede verse como un valle en el que descansará el snake. En el segundo caso es posible definir puntos específicos que atraen (modelado de *muelles*) o repelen (modelado de *volcanes*) al contorno, p.e. para evitar que éste quede bloqueado en un mínimo local.

Dinámica Lagrangiana

De acuerdo con los principios de la dinámica Lagrangiana el posicionamiento óptimo del snake coincide con un punto de estabilidad o equilibrio. Durante el proceso de ajuste parte de la energía potencial global inicial se transforma en energía cinética y de rozamiento. Las ecuaciones de Euler-Lagrange del movimiento son

⁴Usualmente Γ_s y Γ_{ss} denotan respectivamente $\frac{\partial \Gamma}{\partial s}$ y $\frac{\partial^2 \Gamma}{\partial s^2}$.

⁵Desde el punto de vista discreto esta es la denominación comúnmente aceptada para los puntos del *snake*.

⁶Podemos crear discontinuidades de posición o de curvatura haciendo cero el parámetro correspondiente.

⁷Intensidad Alta, Gradientes, Texturas, etc.

las siguientes:

$$\mu\Gamma_{tt} + \gamma\Gamma_t - \alpha\frac{\partial\Gamma_s}{\partial s} + \beta\frac{\partial^2\Gamma_{ss}}{\partial s^2} = -\nabla(\gamma\Psi(\Gamma(s)) + \lambda\mathcal{R}(\Gamma(s))) \quad (2.5)$$

donde μ y γ representan la densidad de masa y el coeficiente de rozamiento (ambos supuestos constantes para todo el contorno) y asumimos signo positivo en los potenciales de imagen y restricción. En definitiva la configuración final es aquella en que la variación de las fuerzas de inercia, rozamiento, tensión y rigidez se equilibran con la variación de los agentes o factores externos (fuerzas de acoplamiento a las características de imagen).

Técnicas de Minimización

El proceso de ajuste óptimo del *snake* y en definitiva el cumplimiento de la condición de equilibrio expresada en la ecuación 2.5 se implementa desde distintos puntos de vista:

1. *Métodos Numéricos*: particularmente las *Diferencias Finitas* (Kass et al., 1987), los *Elementos Finitos* (Cohen y Cohen, 1993), (Leymarie y Levine, 1993) que consisten en la resolución de sistemas de ecuaciones derivados de las condiciones de equilibrio, y finalmente los *Métodos Probabilísticos* (Szeliski, 1989), (Terzopoulos y Szeliski, 1992) basados en estimación bayesiana.
2. *Métodos de Búsqueda*: tales como la *Búsqueda Voraz* o descenso por gradiente (Williams y Shah, 1992), (Kumar y Goldoff, 1994), en la que se busca el posicionamiento óptimo local para cada *snaxel* (lo cual puede conducir a óptimos relativos), la *Programación Dinámica* (Amini et al., 1990) en la que se recurre a una formulación del problema en términos de un proceso de decisión multietapa discreto.

Implementaciones Alternativas y Aplicabilidad

Partiendo de los principios variacionales básicos se han formulado diversas variantes: en la caracterización geométrica (*Descriptores de Fourier* (Staib, Duncan, 1989), *B-Splines* (Curwen et al., 1991), (Curwen y Blake, 1992), *Representaciones Matriciales* (Lai, 1994)), en su extensión dimensional (*Superficies Activas* (Cohen et al., 1992), *Redes Activas* (Tsumiyama et al., 1989), (Sakaue y Yamamoto, 1991), *Cubos/Tubos Activos* (Takahata et al., 1992), (Bro-Nielsen, 1994)), y en la inclusión de nuevos términos en la función de energía orientados a definir conductas (Balones (Cohen, 1991)). Así mismo se han resuelto problemas como el tratamiento de potenciales complejos (Delagnes et al., 1995) o el de la occlusión parcial de

objetos (Koller et al., 1994). Se ha demostrado la aplicabilidad de estos modelos en diversos contextos, especialmente el biomédico (McIrnerney y Terzopoulos, 1996), y su versatilidad como herramientas de análisis (Terzopoulos et al., 1988) y de síntesis (Witkin y Kass, 1988), (Cohen, 1992), (Lee et al., 1995), (Grzeszczuk y Terzopoulos, 1995)⁸. Como complemento a la descripción realizada puede consultarse el Apéndice A, en el cual se muestran diversos resultados de la aplicación de estos modelos a contextos fundamentalmente biomédicos.

Mecanismos de Tracking para Snakes

Además del filtrado vacío se contempla el filtrado de Kalman (Terzopoulos y Szeliski, 1992), la modelización mediante análisis frecuencial (McEachen et al., 1994) y modal (Curwen y Blake, 1992), el uso de ecuaciones diferenciales estocásticas (Blake et al., 1995) y la relajación del supuesto gaussiano en el filtro de Kalman (Isard y Blake, 1996). El problema clave es la inicialización, estimación y ajuste de los múltiples parámetros que definen la elasticidad del modelo. Recientemente se ha realizado una propuesta (Metaxas y Kakadiaris, 1996), orientada a filtrado no-lineal de Kalman, en la cual se asocia a dichos parámetros una variabilidad dependiente de su contribución local al error.

2.3.2 Métodos de Elementos Finitos y Componentes Principales

Los modelos basados en *Descomposiciones Modales* o *Auto-Representaciones* se fundamentan en los principios de la *Transformada de Karhunen-Loève* (Fukunaga, 1990), existiendo tres tipos de aproximación: el *Análisis de Componentes Principales*, el *Análisis Mecánico Modal (Elementos Finitos)* y el *Análisis de Fourier*. En éste apartado estudiaremos representaciones deformables basados en los dos primeros⁹, siendo sus características definitorias las siguientes:

1. *Espacio de Representación Modal*: dado un conjunto de m formas definidas por n parámetros se trata de realizar un *cambio de dominio* al espacio expandido por un conjunto de autovectores básicos ortogonales (*modos de variación*) $\vec{\Psi}_1, \vec{\Psi}_2, \dots, \vec{\Psi}_t$ con $t \ll n$ centrado en la *forma promedio o básica* \vec{F}_0 , de manera que dada cualquier forma \vec{F} del conjunto y siendo $\alpha_1, \alpha_2, \dots, \alpha_t$ sus *proyecciones* sobre los vectores básicos, podemos representarla en términos de:

$$\vec{F} = \vec{F}_0 + \sum_{i=1}^t \alpha_i \vec{\Psi}_i \quad (2.6)$$

⁸Un aspecto importante en los Contornos Activos en particular y de la mayoría de los Modelos Deformables en general es esta doble versatilidad en análisis y síntesis.

⁹Los modelos deformables basados en Análisis de Fourier se referenciaron en la sección anterior

2. *Reducción de Dimensionalidad*: en base a la condición $t \ll n$ se determina una representación capaz de obtener aproximaciones al grado de precisión (o detalle) deseado. Para ello se toman en consideración los t *autovectores más representativos* (aquellos con mayores autovalores: capturan el porcentaje deseado de variabilidad).
3. *Compacidad y Unicidad*: como resultado pueden obtenerse las componentes menos críticas o con alto nivel de ruido, siendo esta una representación canónica adecuada para la compactación, reconocimiento en tiempo real y reconstrucción robusta. Además, la ortogonalidad de los vectores asegura la unicidad de la representación haciendo tratables los problemas de reconocimiento (Sclaroff y Pentland, 1993).
4. *Matching*: el reconocimiento se define como estimación de la deformación mínima necesaria para incluir el prototipo en la representación. Dependiendo del método se puede utilizar una distancia de Mahalanobis (Duda y Hart, 1973) (aproximación estadística: la representación como distribución) o una métrica basada en la propia deformación.

Modelos Basados en Elementos Finitos

Estos modelos, propuestos originalmente por Pentland y Sclaroff (Pentland y Sclaroff, 1991) toman como base los *Elementos Finitos*, una técnica estándar de ingeniería (Bathe, 1982) que permite simular la conducta dinámica de un objeto, y se basan en:

1. *Representación Física*: se retoma el planteamiento físico de los *snakes* desde un nuevo enfoque: se definen funciones de interpolación que permiten integrar propiedades del material elástico a lo largo de regiones de interés. Con respecto a los métodos de diferencias finitas se aporta una caracterización analítica de la superficie entre nodos (pixels).
2. *Energías y Dinámica*: dada una forma compuesta por n puntos su deformación, se representa por energías definidas a partir de los vectores U de dimensión $3n \times 1$ que representan los desplazamientos $(\Delta x, \Delta y)$ de los puntos nodales 2D con respecto al centro de masas del objeto), de manera que dadas las matrices de orden $2n \times 2n$ que definen respectivamente la masa M , el rozamiento C , la elongación K de los puntos del objeto, y las fuerzas R que actúan sobre ellos, se define la siguiente ecuación de comportamiento:

$$M\ddot{U} + C\dot{U} + KU = R \quad (2.7)$$

3. *Cambio de Dominio*: siendo $KU = R$ la ecuación de equilibrio del sistema de ecuaciones 2.7, se trata de obtener U (los *desplazamientos en el equilibrio*), para lo cual se recurre, con objeto de obtener soluciones eficientes y cerradas, al *Análisis Modal*:

(a) *Transformación Lineal*: para diagonalizar el sistema puede usarse una transformación lineal ortogonal P de manera que $U = P\tilde{U}$ con lo cual realizamos un cambio de base. La transformación óptima se deriva de los *modos de vibración* de la ecuación de equilibrio, resolviéndose el siguiente problema de autovalores:

$$K\vec{\Psi}_i = \omega_i^2 M\vec{\Psi}_i \quad (2.8)$$

$$(\omega_1^2, \vec{\Psi}_1), (\omega_2^2, \vec{\Psi}_2), \dots, (\omega_{2n}^2, \vec{\Psi}_{2n}) \quad (2.9)$$

$$\vec{\Psi}_i^T M \vec{\Psi}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (2.10)$$

$$0 \leq \omega_1^2 \leq \omega_2^2, \dots, \leq \omega_{2n}^2 \quad (2.11)$$

(b) *Modos de Vibración*: cada autovector $\vec{\Psi}_i$ se denomina modo de vibración y representa los desplazamientos nodales de (x, y) , siendo el autovalor ω_i^2 su correspondiente frecuencia, de manera que distinguimos los siguientes *modos cualitativos*:

- i. *Modos de Baja Frecuencia*: (modos rígidos) codifican *traslaciones* en x, y y *rotaciones linealizadas* de manera que nodos opuestos tienen direcciones opuestas.
- ii. *Modos de Frecuencia Media*: representan *deformaciones globales* que dejan el centro de rotación fijo (categorías de objetos compactos): escalados, expansiones y estrechamientos.
- iii. *Modos de Alta Frecuencia*: representan *deformaciones locales* (propios de categorías de objetos muy dispares o con agujeros).

4. *Solución*: siendo $P = [\vec{\Psi}_1, \vec{\Psi}_2, \dots, \vec{\Psi}_{2n}]$, los desplazamientos pueden representarse como:

$$U = P\tilde{U} = \sum_{i=1}^{2n} \tilde{u}_i \vec{\Psi}_i \quad (2.12)$$

es decir a partir de sus coordenadas en la base formada por los autovectores.

Aplicabilidad de los Elementos Finitos

Estas representaciones se han aplicado con éxito en problemas de reconstrucción 3D a partir de medidas de sensor (Pentland y Sclaroff, 1991), y reconocimiento de

formas complejas en biomedicina (Sclaroff y Pentland, 1994) y en contextos más genéricos (bases de datos de categorías de objeto) (Sclaroff, 1995).

Mecanismos de Tracking para Elementos Finitos

En cuanto al *tracking* de movimiento no-rígido (Pentland y Horowitz, 1991), (Pentland y Sclaroff, 1991), la aplicación del análisis modal sobre el *vector velocidad* o *vector de desplazamiento nodal* permite expresar las ecuaciones de movimiento en la nueva base o espacio transformado. De esta manera considerando un número suficiente de modos de variación de baja frecuencia podemos expresar fielmente el movimiento. Al combinar dicho esquema con un filtrado de Kalman se obtiene una estimación suave del movimiento. También se han aplicado mecanismos de elementos finitos para el seguimiento y análisis de deformaciones en 2D y 3D (Nastar y Ayache, 1993).

Modelos de Distribución de Puntos y Formas Activas

Los *Modelos de Distribución de Puntos* (Cootes et al., 1992) se basan, a diferencia de los anteriores, en el ACP. Dada la relevancia de esta técnica en nuestra propuesta dedicamos el Apéndice B de la presente memoria a introducir los fundamentos formales de la misma. Dado un conjunto de ejemplos representativos de una estructura variable (*conjunto de entrenamiento*) el proceso de obtención de sus parámetros característicos es el siguiente:

1. *Representación de Forma*: cada una de las m formas del *conjunto de entrenamiento* se representa como un vector x_i de dimensión $2n$ que codifica las posiciones cartesianas (x_{ik}, y_{ik}) de los n puntos etiquetados que la componen:

$$\vec{x}_i = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{in}, y_{in})^T \quad (2.13)$$

2. *Alineamiento Óptimo*: dado que el modelo se basa en estadísticas de las diferencias entre las coordenadas de puntos equivalentes a lo largo del conjunto, se plantea un mecanismo iterativo de *alineamiento automático* entre una forma inicial \vec{x}_i y cualquier otra \vec{x}_j que obtiene la transformación afín $M(k, \theta)\vec{x}_i + t$ que minimiza la suma ponderada:

$$E = (\vec{x}_j - (M(k, \theta)\vec{x}_i + t))^T W (\vec{x}_j - (M(k, \theta)\vec{x}_i + t)) \quad (2.14)$$

siendo W una matriz diagonal con los pesos w_k proporcionales a la estabilidad de cada punto con respecto a los demás. El ajuste se extiende a todas las formas del conjunto y progresa *renormalizando la forma media recalculada en cada etapa*.

3. Modelado Estadístico y Componentes Principales:

- (a) *Distribución Estadística*: una vez alineado el conjunto de entrenamiento, asumiremos que este define una distribución gaussiana. Dicha distribución puede caracterizarse por la *forma promedio* \bar{x} y por la *matriz de covarianzas* S de dimensión $2n \times 2n$:

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m \vec{x}_i \quad (2.15)$$

$$d\vec{x}_i = \vec{x}_i - \bar{x} \quad (2.16)$$

$$S = \frac{1}{m} \sum_{i=1}^m d\vec{x}_i d\vec{x}_i^T \quad (2.17)$$

- (b) *Componentes Principales/Modos de Variación*: los modos de variación de una forma del conjunto vienen descritos por los $2n$ autovectores unitarios de S , de forma que las variaciones más significativas son descritas por los autovectores con mayores autovalores asociados. En concreto el k -ésimo autovector $\vec{\Psi}_k$ afecta a un punto x_i del modelo moviéndolo a lo largo de un vector paralelo a (dx_{kl}, dy_{kl}) , siendo:

$$\vec{\Psi}_k = (dx_{k1}, dy_{k1}, dx_{k2}, dy_{k2}, \dots, dx_{kn}, dy_{kn})^T \quad (2.18)$$

Siendo $P_t = [\vec{\Psi}_1, \vec{\Psi}_2, \dots, \vec{\Psi}_t]$ la matriz con los k primeros autovectores ($t \ll 2n$) y $b = (b_1, b_2, \dots, b_t)^T$ un vector de pesos o coordenadas en el nuevo dominio, una forma \vec{x}_i del conjunto de entrenamiento puede expresarse de la siguiente manera (ver Sección B.2):

$$\vec{x}_i = \bar{x} + P_t b = \bar{x} + \sum_{k=1}^t b_k \vec{\Psi}_k \quad (2.19)$$

Teniendo en cuenta que $P^T P = I$ las coordenadas b_k son las proyecciones de \vec{x}_i sobre cada eje principal, y pueden obtenerse mediante la ecuación:

$$b = P_t^T (\vec{x}_i - \bar{x}) \quad (2.20)$$

Así mismo se han realizado propuestas para modelizar también las distribuciones de los niveles de gris a los que son sensibles las formas del modelo (Cootes et al., 1993), incorporándose con ello la capacidad de *aprender un modelo de imagen concreto*.

Tomando como base la caracterización descrita, se han desarrollado las *Formas Activas* (Cootes y Taylor, 1992) como modelos deformables que implementan la

interacción dinámica con el espacio de imagen (ajuste a la evidencia de imagen). A partir de su formulación básica se ha desarrollado un algoritmo de ajuste (Cootes et. al, 1994a) que integra la información sobre el modelo estadístico de imagen. Las fases principales del algoritmo son las siguientes

1. *Hipótesis Iniciales*: Partimos de un Modelo de Distribución de Puntos $b = P_t^T(\vec{x} - \bar{x})$ de un objeto dado \vec{x} y de una estimación inicial de su posición, orientación, escala y forma reales en la imagen \vec{X} , dado el origen de coordenadas (x_c, y_c) :

$$\vec{X} = M(k, \theta)\vec{x} + \vec{X}_c \quad (2.21)$$

$$\vec{X}_c = (x_c, y_c, x_c, y_c, \dots, x_c, y_c)^T \quad (2.22)$$

Dicha inicialización puede obtenerse heurísticamente mediante Algoritmos Genéticos (Hill et al., 1992).

2. *Refinamiento Iterativo de Hipótesis*: el cálculo de la nueva forma $(\vec{X} + d\vec{X})$ que mejor se ajusta a las características se realiza de forma iterativa:

- (a) *Cálculo del Desplazamiento*: examinar la *ventana de búsqueda local* centrada en cada punto de acuerdo con el perfil de gris aprendido p y encontrar (dX_i, dY_i) el *mejor desplazamiento normal* a cada uno de los puntos, con lo que se genera el vector:

$$d\vec{X} = (dX_1, dY_1, dX_2, dY_2, \dots, dX_n, dY_n)^T \quad (2.23)$$

- (b) *Cálculo de Transformación Afín*: a partir de los desplazamientos sugeridos se actualiza la traslación (dX_c, dY_c) , escala $(1 + dk)$ y rotación $d\theta$ mediante un procedimiento similar al de alineamiento. A continuación se realiza un *ajuste residual* en el cual se impone la restricción de pertenencia al modelo. Se trata entonces de calcular el ajuste óptimo de los parámetros de forma db de acuerdo con las nuevas posiciones $\vec{x} + d\vec{x}$ derivadas de realizar el cambio de dominio a partir del último $d\vec{X}$ calculado. Este problema puede resolverse mediante ajuste mínimo-cuadrático del error:

$$T = (d\vec{x} - Pdb)^T W (d\vec{x} - Pdb) \quad (2.24)$$

- (c) *Actualización y Matching*: una vez calculado db debemos asegurarnos de que las deformaciones propuestas son consistentes con el modelo. En el caso de que la actualización propuesta exceda el rango de admisibilidad, ésta debe ser reescalada adecuadamente (para que el nuevo

$b' = b + db$ quede dentro del hiper-elipsoide centrado en la media) mediante la ecuación:

$$b'_k \leftarrow b_k \frac{D_{Def}}{D_{Mah}} \quad (2.25)$$

donde D_{Mah} es la distancia de Mahalanobis y D_{Def} es la distancia umbral permitida.

Aplicabilidad del Modelo de Formas Activas

Estos modelos se han utilizado con éxito en diversos contextos y problemáticas entre las que destacan el *tracking* de estructuras biomédicas (Cootes et al., 1994a) y la experimentación en reconocimiento de caras y gestos (Lanitis et al., 1994). Así mismo se han realizado propuestas complementarias de los algoritmos básicos como, por ejemplo, la integración de los Métodos de Elementos Finitos y los Modelos de Distribución de Puntos (Cootes y Taylor, 1995), la optimización de la búsqueda o ajuste de las Formas Activas mediante técnicas de multirresolución (Cootes et al., 1994b) y la integración con métodos probabilísticos bayesianos aplicados en segmentación de secuencias de imágenes (Kervrann y Heitz, 1993).

Mecanismos de Tracking para Formas Activas

En (Baumberg y Hogg, 1994a) se presenta una propuesta orientada al aprendizaje de las configuraciones de los puntos de control de B-Splines y aplicada como base del *tracking* eficiente (Baumberg y Hogg, 1994b) de figuras animadas (p.e. personas en movimiento). El conocimiento acerca del modelo de forma es adquirido mediante ACP, con lo cual se consigue un modelo de baja dimensionalidad que, en combinación con un filtrado de Kalman, permite obtener resultados aceptables y eficientes. En este trabajo se sugiere la incorporación de información adicional en el Modelo de Distribución de Puntos.

2.3.3 Modelos Algebraicos Basados en Interpolantes

Estos modelos (Bookstein, 1989) proporcionan una descripción algebraica de las deformaciones basándose en la construcción de una función de correspondencia entre *frames* consecutivos a partir de un conjunto finito de *matchings*. En términos generales las fases del proceso de extracción y análisis del movimiento son:

1. *Cálculo de Correspondencias Básicas*: dados n puntos característicos (marcadores) $P_i = (x_i, y_i)$ en un *frame* (p.e. puntos de elevada curvatura o intensidad) se trata de encontrar sus homólogos $P'_i = (x'_i, y'_i)$ en el siguiente. Este primer paso dependerá de la aplicación (si la continuidad

espacio-temporal es aceptable suelen utilizarse otros modelos deformables como *Snakes* y mallas).

2. *Interpolación y Correspondencia General*: a continuación se utilizan los puntos de correspondencia conocida para obtener, mediante interpolación de una superficie en los espacios paramétricos $x \times y \times x'$ y $x \times y \times y'$, una función de correspondencia general $f : (x, y) \rightarrow (f_x(x, y), f_y(x, y))$ que asocia a cada punto de un *frame* su correspondiente en el segundo. Siendo $U(r) = r^2 \log(r^2)$, con $r_{ij} = \|P_i - P_j\|$, la ecuación de un interpolante (generalización 2D de un *spline* cúbico) centrado en cada P_i cuyo comportamiento es similar al de un plato de *metal deformable*, la función de correspondencia es la solución de un sistema de ecuaciones:

$$L = \begin{bmatrix} K_{n \times n} & P_{n \times 3} \\ P_{3 \times n}^T & 0_{3 \times 3} \end{bmatrix}_{(n+3) \times (n+3)} \quad (2.26)$$

$$L^{-1}(V \mid 0, 0, 0)^T = (W \mid a_1, a_x, a_y)^T \quad (2.27)$$

$$f(x, y) = a_1 + a_x x + a_y y + \sum_{i=1}^n w_i U(\|P_i - (x, y)\|) \quad (2.28)$$

donde: K es una matriz simétrica que contiene los $U(r_{ij})$, P incluye las posiciones de los P_i y W es un vector de pesos w_i . En consecuencia $f_x(x, y)$ y $f_y(x, y)$ se obtienen cuando sustituimos V por (x'_1, \dots, x'_n) , (y'_1, \dots, y'_n) respectivamente. La ecuación 2.28 se compone de tres *términos de naturaleza afín* (a_1, a_x, a_y) y de la integración de n funciones de interpolación que modelan efectos no-lineales. Este procedimiento es invariante a traslación y rotación del conjunto de marcadores al tiempo que minimiza la *energía de deformación* consiguiendo un error de interpolación aceptable.

3. *Estimación de Parámetros de Movimiento*: el análisis de la deformación puede plantearse desde, al menos, dos ópticas basadas en la extracción de autovectores. En ambos casos se trata de estimar la descomposición del movimiento (deformaciones principales):
- Espectro de la Matriz de Energía*: en el planteamiento original (Bookstein, 1989) la matriz $L_n^{-1} K L_n^{-1}$, donde L_n^{-1} es el bloque $n \times n$ diagonal superior de L^{-1} , se denomina *matriz de energía de deformación*. Los autovectores de esta matriz son descriptores canónicos y ortogonales de los modos de deformación de acuerdo con los desplazamientos de los marcadores.
 - Tensores de Deformación*: otra posibilidad (Kumar y Goldgof, 1994) consiste en estimar mediante mínimos cuadrados la transformación

$P'_i = MP_i + T$ y analizar la descomposición $M = RD$ en los tensores de rotación R y deformación D . Los autovalores de D codifican las deformaciones principales, y los autovectores asociados proporcionan las direcciones ortogonales sobre las que éstas actúan.

Aplicabilidad de Modelos Algebraicos

Esta propuesta está inspirada en investigaciones sobre medición de forma propios del campo de la Biología Cuantitativa¹⁰. En su origen este método se aplicó para analizar el *síndrome de Apert*, una malformación craneofacial congénita caracterizada por el infradesarrollo maxilar. Posteriormente se han aplicado en análisis de movimiento cardíaco (Kumar y Goldgof, 1994), en extracción de diferencias entre mamografías consecutivas (Sallam y Bowyer, 1994) para determinar su grado de ajuste.

2.3.4 Modelos Basados en Superficies Super-Cuádricas

Los Modelos Paramétricos de *Superficies Super-Cuádricas* se introdujeron en primer lugar en el ámbito de los gráficos por computador (Barr, 1981) y posteriormente se aplicaron en Visión Artificial como representaciones adecuadas al análisis y reconocimiento (Pentland, 1986), (Bajcsy y Solina, 1987). Siendo $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$ y $-\pi \leq \phi \leq \pi$ los ángulos correspondientes a latitud y longitud respectivamente, a_x, a_y, a_z los parámetros de escala en cada eje y ϵ_1, ϵ_2 los de cuadratura en el eje Z y en el plano XY respectivamente, se derivan las siguientes definiciones (vectorial e implícita):

$$S(\theta, \phi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_x \cos^{\epsilon_1}(\theta) \cos^{\epsilon_2}(\phi) \\ a_y \cos^{\epsilon_1}(\theta) \sin^{\epsilon_2}(\phi) \\ a_z \sin^{\epsilon_1}(\theta) \end{bmatrix} \quad (2.29)$$

$$\left(\left(\frac{x}{a_x} \right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_y} \right)^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}} + \left(\frac{z}{a_z} \right)^{\frac{2}{\epsilon_1}} = 1 \quad (2.30)$$

Una superficie supercuádrica es una generalización de un elipsoide que permite representar gran variedad de primitivas 3D (cilindros, bloques, troncos, pirámides, conos y elipsoides) y sus *deformaciones globales* variando un reducido número de parámetros ($a_x, a_y, a_z, \epsilon_1, \epsilon_2, \theta, \phi, \psi, t_x, t_y, t_z$), siendo θ, ϕ, ψ los parámetros de orientación en el espacio y t_x, t_y, t_z la traslación del origen local al objeto. Estos

¹⁰En concreto se argumenta que si las correspondencias básicas están en concordancia con criterios de homología biológica la función f modela la comparación de formas mediante deformación como sugirió D'arcy Thompson en 1917.

modelos proporcionan por definición un enlace importante entre el análisis de formas cuantitativo y el cualitativo.

En cuanto al mecanismo de ajuste, teniendo en cuenta que se asume como entrada una imagen de rango (distribución de puntos 3D), el esquema usual consiste en ajustar los parámetros mediante el criterio de maximizar el número de puntos de la imagen que pertenecen a la superficie definida por la curva en el espacio. Este proceso de optimización forma parte de un procedimiento de más alto nivel en el cual se disponen de forma aleatoria un conjunto de curvas que posteriormente se combinan para extraer un modelo de objeto compuesto por partes de complejidad cuadrática (Leonardis et al., 1996).

Modelos Alternativos para Modelización Local

La mejora en capacidad expresiva de los modelos básicos (incorporando modelización de deformaciones locales) se abordó en primer lugar en los *Modelos de Super-Cuádricas Deformables*, propuestos por (Terzopoulos y Metaxas, 1991,1992,1993). Dichos modelos incorporan los principios físico-geométricos propios de los Contornos Activos permitiendo con ello la modelización de deformaciones globales (simétricas y asimétricas) y deformaciones locales. Se trata de combinar el carácter compacto de las Super-Cuádricas (reducido número de parámetros) y por ello su *aptitud para el reconocimiento* con la *sensibilidad al detalle* que aportan los *Snares*. Se han realizado propuestas alternativas al método físico para modelizar deformaciones locales. Se ha argumentado (Chen et al.,1994) que estos modelos no reflejan las interrelaciones entre las componentes del movimiento no-rígido y que su aplicación no es viable dada la *complejidad computacional de la resolución del problema de minimización de energía implícito*. Tal es el caso de la utilización de *Análisis Armónico* (Chen et al.,1994) y de la descomposición jerárquica del movimiento (rígido, deformable global y deformable local) basándonos en el conocimiento apriorístico relativo a la deformación del objeto. Como consecuencia se obtiene una gradación *global-detalle* muy interesante a efectos de capacidad expresiva.

Aplicabilidad y Mecanismos de Tracking para Super-Cuádricas

Los modelos inspirados en este planteamiento paramétrico se han aplicado al *tracking*, mediante filtrado de Kalman, de objetos 3D deformables (Terzopoulos y Metaxas 1992, 1993), al análisis del movimiento cardíaco (Chen et al., 1994) y a la identificación de objetos en escenas 3D (aprovechando las propiedades cualitativas del modelo: descripción de familias de objetos) en general (Dickinson y Metaxas, 1994) y como soporte a estrategias de aprehensión de objetos en Visión

Orientada a Tareas (Ikeuchi y Herbert, 1996).

2.3.5 Plantillas Deformables Basadas en Modelos Bayesianos

Enfoque Transformacional de Grenander

Los Modelos Estadísticos de Plantillas Deformables propuestos en (Grenander et al., 1989,1991) y madurados posteriormente en (Amit et al., 1989,1993,1994), (Christensen, 1994) y otros trabajos, proporcionan una descripción *probabilística bayesiana* de las *transformaciones* implícitas en la deformación a partir de una forma de referencia denominada *plantilla*. Se asume que dichas transformaciones (p.e. homomorfismos) se definen de forma que preservan ciertas propiedades topológicas de la plantilla que caracteriza a una categoría de formas dada. Se asume que la variabilidad de la forma es un *proceso estocástico* y su tratamiento puede abordarse definiendo los siguientes elementos:

1. *Función de Dependencia*: dada la forma de test \vec{f} existe una función de dependencia *a posteriori* $p(\vec{h} | \vec{f})$ que expresa la probabilidad de que \vec{f} haya sido generada por la *transformación* \vec{h} . Según la regla de Bayes:

$$p(\vec{h} | \vec{f}) \propto p(\vec{f} | \vec{h})p(\vec{h}) \quad (2.31)$$

2. *Probabilidad a priori*: $p(\vec{h})$ codifica el *conocimiento estocástico* acerca de las transformaciones (p.e. condicionamientos topológicos). En ocasiones se asume una distribución de Gibbs en función de un *potencial* sobre los desplazamientos inducidos definido de forma variacional (Christensen, 1994). En general este supuesto depende del conocimiento acerca del problema.
3. *Modelo de Verosimilitud*: $p(\vec{f} | \vec{h})$ modela la variabilidad relativa, incluidas las restricciones estructurales, de la forma con respecto a la transformación. También puede definirse un potencial dependiente, en este caso, de una métrica que evalúe la confianza en generar \vec{f} a partir de \vec{h} .
4. *Estimación Bayesiana*: el problema de *matching* entre forma y plantilla se expresa como la obtención de un estadístico sobre $p(\vec{h} | \vec{f})$ tales como en *Máximo a Posteriori* (MAP), el de *Máxima Verosimilitud* (MLE) o el de *Mínima Varianza* (MVE). En el primer caso se trata de encontrar \vec{h}^* tal que:

$$\vec{h}^* = \text{Arg Max } p(\vec{h} | \vec{f}) \quad (2.32)$$

Aplicabilidad del Modelo Transformacional

Estos modelos se han aplicado fundamentalmente en dominios biomédicos, especialmente en reconstrucción (Amit y Maynbeck, 1993) y diagnóstico (Christensen, 1994), ya que son especialmente adecuados para codificar la estructura y variabilidad de la forma y de las distribuciones de intensidad y textura. Ello ha dado lugar a potenciar el desarrollo de representaciones capaces de integrar conocimiento procedente de distintas fuentes y niveles de abstracción. Tal es el caso de los *Diccionarios/Textos Anatómicos Deformables* (Christensen, 1994) los cuales modelizan órganos codificando atributos de: funcionalidad, distribución de intensidad, etiqueta de segmentación, textura, probabilidad de lesión, procedimiento quirúrgico sugerido, etc. Estos modelos constituyen un soporte importante en diversos ámbitos tales como la morfometría aplicada a órganos tales como el cerebro.

Enfoque de Primitivas Deformables de Yuille

Los modelos de *Primitivas/Plantillas Deformables* (Yuille et al., 1989,1990, 1992) están inspirados en el *Modelo Transformacional* y aportan un enlace entre el planteamiento activo de los *Snakes* y el enfoque geométrico pasivo de las técnicas basadas en la *Transformada de Hough* (Hough, 1962). Las características más relevantes de estos modelos son:

1. *Primitivas Deformables y Potenciales de Imagen*: la estructura geométrica viene caracterizada por un conjunto de parámetros relevantes $\Theta = \{\theta_k\}_{k=1,\dots,n}$ y está compuesta por un conjunto de m entidades, *primitivas* o *plantillas*¹¹ $\mathcal{G} = \{g_i(\Theta)\}_{i=1,\dots,m}$ y sensibles a campos de potencial concretos¹² los cuales definen el *modelo de imagen* $\mathcal{I} = \{\Psi_j\}_{j=1,\dots,m}$.
2. *Funciones de Energía*: en términos bayesianos el problema del ajuste de la plantilla consiste en maximizar la probabilidad condicionada:

$$p(\Theta | \mathcal{I}) = \frac{p(\mathcal{I} | \Theta)p(\Theta)}{p(\mathcal{I})} \quad (2.33)$$

siendo $p(\Theta)$ la *probabilidad a priori* o configuración más probable de la estructura y $p(\mathcal{I} | \Theta)$ la verosimilitud asociada al modelo de imagen (conexión entre plantillas y mapas de intensidad concretos). Este problema de maximización se traduce en la minimización de una función de energía global

¹¹Parábolas, círculos, elipses, puntos relevantes etc. En general debe lograrse un equilibrio entre la capacidad expresiva y la eficiencia. Por ello se recomienda seleccionar el menor número de parámetros capaces de expresar adecuadamente la geometría del problema.

¹²Aristas, valles, picos, etc.

$\mathcal{E}(\Theta)$ que rige el comportamiento de las plantillas. Dicha función consta, al menos, de dos términos principales:

- (a) *Energía de Forma*: $\mathcal{E}_{forma}(\Theta)$ determina la satisfacción de preferencias geometrías por la estructura (*métrica de ajuste geométrico*) y permite definir el conocimiento probabilístico a priori $p(\Theta)$.
- (b) *Energía de Imagen*: $\mathcal{E}_{imagen}(\Theta)$ estima el grado de ajuste de la estructura al potencial programado (*métrica de ajuste de imagen*). Este término sirve de base para modelar la verosimilitud $p(\mathcal{I} | \Theta)$.

La definición de estos términos puede enfocarse desde dos puntos de vista:

- (a) *Formulación Aditiva*: se ajusta a las ecuaciones:

$$\mathcal{E}(\Theta) = \mathcal{E}_{forma}(\Theta) + \mathcal{E}_{imagen}(\Theta) \quad (2.34)$$

$$\mathcal{E}_{forma} = \sum_{i=1}^n k_i R_i(\Theta), \mathcal{E}_{imagen} = \sum_{j=1}^m c_j I_j(\Theta) \quad (2.35)$$

$$I_j = \int \int_{D_j \supseteq \Gamma_j} \Psi_j(u, v) dudv \quad (2.36)$$

$$k_i, c_j \in \mathfrak{R} \quad (2.37)$$

siendo R_i relaciones geométricas ponderadas entre los parámetros de la estructura (distancia entre puntos, diferencias de curvatura, escala u orientación) y D_j el dominio correspondiente a la primitiva g_j (la inclusión $D_j \supseteq \Gamma_j$ garantiza robustez).

- (b) *Formulación por Correlación*: presenta una clara similitud con el esquema de votación propio de la Transformada de Hough y se expresa mediante:

$$\mathcal{E}(\Theta) = \mathcal{E}_{forma}(\Theta) * \mathcal{E}_{imagen}(\Theta) \quad (2.38)$$

$$\mathcal{E}_{forma}(\Theta) * \mathcal{E}_{imagen}(\Theta) = \sum_{j=1}^m C_j(\Theta) \quad (2.39)$$

$$C_j(\Theta) = \int \int_{D_j \supseteq \Gamma_j} M(u, v, g_j) \Psi_j(u, v) dudv \quad (2.40)$$

donde $M(u, v, g_j)$ expresa la distancia mínima del punto (u, v) del espacio de imagen a la primitiva g_j (usualmente define el grado de satisfacción de la geometría impuesta por dicha primitiva).

Minimización de Energía

Una vez definida la energía global el ajuste óptimo de la estructura a la imagen suele abordarse fundamentalmente mediante descenso por gradiente. Se plantea un ajuste iterativo que define la dinámica de la estructura. Este comportamiento regido por las ecuaciones:

$$\Theta_{t+1} = \Theta_t + \frac{d\Theta}{dt} \lambda \quad (2.41)$$

$$\frac{d\Theta}{dt} \rightarrow -\frac{\partial \mathcal{E}(\Theta)}{\partial \Theta} \quad (2.42)$$

siendo λ el paso de iteración. La implementación de esta conducta se abordó inicialmente mediante una estrategia de *minimización por etapas*. Cada etapa se caracterizaba por la activación secuencial de un subconjunto de primitivas. Este esquema presenta algunos inconvenientes abordados a posteriori por otros autores (Xie et al., 1994):

1. *Ineficiencia y Mínimos Locales*: por un lado el tiempo de proceso es elevado debido a la implementación secuencial de la minimización, y por otro el ajuste secuencial de los pesos correspondientes a los términos de energía puede generar óptimos locales. Estos problemas se resuelven alterando simultáneamente todos los parámetros e incorporando un método numérico de base con una tasa de convergencia rápida como el de *Levenberg-Marquardt (L-M)*.
2. *Ajuste Experimental de Pesos*: difícil generalización del ajuste de pesos. En este sentido se propone la normalización de todos los términos de energía en el rango $[0, 1]$ (con ello se consigue uniformizar la influencia de cada término), y la asignación de dos posibles pesos globales (energía interna e imagen). Se trata de eliminar la dependencia del conocimiento experto presente en el modelo original.

También se ha abordado el problema mediante mecanismos de *simulated annealing*, que toma como referencia la formulación probabilística. La dependencia de las condiciones iniciales se ha abordado mediante variantes de la transformada de Hough.

Aplicabilidad y Mecanismos de Tracking

Estos modelos se han aplicado con éxito en diversos contextos, en particular el *tracking* facial (Yuille, 1992),(Xie et al., 1994), la extracción de características en imagen biomédica (Lipson et al., 1990) y la identificación de partículas en experimentos de física de alta energía (Yuille et al., 1991). Habitualmente se utiliza el mecanismo de filtrado vacío o *a posteriori*.

Otros Modelos de Plantillas

Recientemente se han propuesto dos nuevos modelos de plantillas deformables bayesianas. En el primero (Dubuisson et al., 1996), aplicado al reconocimiento de objetos en movimiento (segmentación y clasificación de vehículos). Se aporta una plantilla poligonal gobernada por restricciones geométricas predefinidas (no adquiridas) y se integra información de movimiento resolviéndose el problema de minimización mediante *simulated annealing*. El segundo (Jain et al., 1996) es un modelo general de plantillas bayesianas. La novedad con respecto a modelos anteriores es que se define un conjunto de *transformaciones de deformación* de la plantilla prototipo en términos de funciones trigonométricas de desplazamiento expandidas por bases ortogonales. En base a estas transformaciones se definen probabilidades asociadas a la plantilla y se formula una función objetivo a minimizar, aportándose para ello un mecanismo de búsqueda multirresolución.

2.4 Conclusiones al Análisis de Modelos y Técnicas

A partir del análisis de las propuestas previas, tanto a nivel de modelos deformables como de mecanismos de *tracking*, y teniendo en cuenta el tipo de problemas en que estamos interesados, encontramos adecuados los *modelos de primitivas deformables* para modelar los elementos de una estructura. Adicionalmente los *mecanismos modales* (especialmente los basados en PCA) nos permiten aprender restricciones geométricas garantizando economía paramétrica en función de los datos de entrenamiento. Todo ello ha motivado las siguientes líneas de progresión:

1. *Énfasis en lo Estructural*: en la mayoría de los modelos se recurre a un modelo elástico con capacidad expresiva suficiente a nivel global (deformaciones genéricas), donde el riesgo de un mínimo local es mayor, en lugar de *combinar estructuralmente múltiples modelos locales* como ocurre en las Plantillas Deformables de Yuille (aunque en estos modelos las restricciones se especifican de forma manual). Recientemente se ha llamado la atención sobre la necesidad de enfoques estructurales (Amit, 1997) en aplicaciones complejas como el análisis de imágenes cerebrales, lo cual refuerza nuestra hipótesis de trabajo. De esta forma se establece una conexión interesante entre los modelos deformables y los algoritmos de matching de grafos. Para llevar a cabo este enfoque estructural creemos conveniente mejorar los modelos existentes a dos niveles:
 - (a) *A Nivel Local*: dotar a los modelos locales (plantillas) de *propiedades de invarianza*, garantizar un *equilibrio entre coste computacional*

y capacidad expresiva y aportar mecanismos de parametrización de la tolerancia a ruido. En este sentido pensamos que las plantillas de tipo super-elipsoidal dotadas con mecanismos de control de ruido resultan adecuadas a estos criterios.

- (b) *A Nivel Global*: elaborar esquemas de aprendizaje de las configuraciones estructurales más probables así como la definición de mecanismos de minimización de energía que exploten dicha capacidad de forma eficiente, y extender estos planteamientos a nivel espacio-temporal (adiestramiento de *trackers*).
2. *Elementos Finitos frente a Componentes Principales*: En cuanto al mecanismo de aprendizaje se han planteado las siguientes discusiones acerca de la adecuación de Elementos Finitos o ACP como soporte:
- (a) *Individuo frente a Conjunto*: los Modelos de Elementos Finitos toman como base una única forma además de la de referencia, mientras que las Componentes Principales están dirigidas por los datos de un conjunto de entrenamiento. Por ello su adecuación depende del volumen de datos disponibles. En este sentido se ha propuesto integrar ambos métodos en contextos de información pobre bajo el mecanismo de utilizar Elementos Finitos para generar una distribución y a partir de ella realizar el entrenamiento de un Modelo de Distribución de Puntos (Cootes y Taylor, 1995).
 - (b) *Necesidad de Etiquetado y Capacidad Expresiva*: los Modelos de Distribución de Puntos precisan de un etiquetado manual mientras que este problema puede salvarse utilizando Elementos Finitos que además permiten modelar grandes rotaciones (Sclaroff y Pentland, 1993). Se ha argumentado (Pentland y Sclaroff, 1991) que los Elementos Finitos proporcionan una elevada capacidad expresiva y robustez. No obstante las Componentes Principales pueden igualar dicha capacidad expresiva si se dispone de los datos de entrenamiento necesarios (Zhu y Yuille, 1994).
 - (c) *Supuestos de Partida*: los Modelos de Elementos Finitos se basan en el conocimiento de las propiedades del material deformable, lo cual es bastante problemático p.e. para la obtención modelos flexibles realistas (Zhu y Yuille, 1994). En cambio el Análisis de Componentes Principales parte de un supuesto mas asumible, el de distribuciones normales o gaussianas.

En base a estos argumentos se ha comprobado experimentalmente la adecuación de las Componentes Principales al aprendizaje y reconocimiento de primitivas geométricas flexibles descritas en forma paramétrica y capaces de capturar formas de grano-medio correspondientes a seres vivos (Zhu y Yuille, 1994).

3. *Auto-Representaciones*: existen múltiples representaciones, denominadas *auto-representaciones* o *eigenmodels*, basadas en ACP o en Elementos Finitos, aplicadas tanto en el ámbito de los modelos deformables (Formas Activas) y tracking como en otros contextos como: modelización de objetos (Pentland y Williams, 1989); análisis y síntesis del movimiento (Kirby et al., 1993), (Murase y Sakai, 1996); reconocimiento de caras (Kirby y Sirovich, 1990), (Turk y Pentland, 1991), (Pentland et al, 1996); reconocimiento de objetos y *tracking* basados en la apariencia visual (intensidad) (Murase y Nayar, 1993), (Nayar et al., 1995a) y parametrización de tareas de robot en general (Nayar et al., 1995b). No obstante el uso de representaciones de esta naturaleza para el aprendizaje y *matching* de estructuras compuestas por plantillas deformables básicas gobernadas por restricciones, y su extensión espacio-temporal para el adiestramiento de *trackers*, es un aspecto novedoso en este campo.

En los tres capítulos siguientes presentaremos nuestra propuesta para la resolución del problema tomando como referencia el énfasis estructural. En el Capítulo 3 presentaremos un modelo de plantilla super-elipsoidal capaz de parametrizar la sensibilidad al ruido. En el siguiente presentaremos el modelo de Auto-Restricciones (restricciones geométricas aprendibles mediante ACP) y en el Capítulo 5 integraremos las componentes local y estructural para obtener un esquema de correspondencia espacial que seguidamente extenderemos al caso espacio-temporal con lo cual *daremos una solución estructural al problema de tracking de contornos*.



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Parte II

Plantillas Deformables Extendidas



Universitat d'Alacant
Universidad de Alicante

Capítulo 3

Caracterización Local: Plantillas Super-Elipsoidales

En este capítulo presentamos el modelo local de *Plantillas Deformables Super-Elipsoidales* como punto de partida para la caracterización de la geometría de la escena mediante una estructura compuesta por múltiples formas 2D. Una vez establecidos los supuestos acerca del modelo de imagen o campo de potencial a considerar (potenciales normalizados), entraremos en la definición de un modelo de plantilla deformable robusta, basado en intensidad y gradiente de potencial, capaz de capturar formas cerradas de complejidad cuadrática (super-elipsoides¹). Presentaremos la función de energía característica, como extensión de un modelo circular básico, y analizaremos formalmente las condiciones de equilibrio en ausencia de conocimiento probabilístico. Dichas condiciones vienen impuestas por los mínimos de energía. Finalmente completaremos este análisis con una serie de experimentos que nos revelarán tanto la sensibilidad del modelo al ruido y a las condiciones de inicialización como su capacidad de adaptatividad ante cambios de posición, orientación, forma y/o escala.

3.1 Campos de Potencial Normalizados

Definición 3.1 Dados un dominio soporte $\mathcal{S} \subseteq \mathbb{R}^2$ y una función de intensidad de imagen $I(x, y) \in \mathbb{R}$ tal que $I : \mathcal{S} \rightarrow [0, I_{Max}]$ se define como *campo de potencial normalizado* $\Psi_T(x, y) = T(I(x, y))$ la transformación $\Psi_T : \mathcal{S} \rightarrow [0, 1]$ que asocia a

¹En sentido estricto el término super-elipsoide se refiere a objetos 3D. Hemos mantenido esta nomenclatura para el caso 2D atendiendo a criterios de generalidad y teniendo en cuenta que la función de energía que proponemos es fácilmente extensible a plantillas 3D.

52CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

cada píxel $(x, y) \in S$ el *grado de cumplimiento* de la propiedad T por su intensidad asociada $I(x, y)$ ².



Figura 3.1: Imagen y potencial binario

La especificación de T establece la distribución de intensidad a la que es sensible cada primitiva. En este sentido cabe hacer dos consideraciones:

1. Desde un punto de vista genérico T puede verse como *el resultado de un análisis de textura*³, es decir $\Psi_D(x, y) = C_D(I(x, y))$ donde C_D indica la aplicación del clasificador normalizado o *fuzzy* C para obtener el grado de pertenencia de la intensidad $I(x, y)$ a una distribución de probabilidad D .
2. En general asumiremos el tratamiento de potenciales normalizados sobre los cuales aplicaremos un *filtro de suavizado gaussiano*, un *filtrado binario* y una *secuencia adecuada de filtros morfológicos alternados secuenciales* (secuencias de aperturas y/o cierres) (Serra, 1982). Dicho potencial en combinación con el contorno de la región extraído mediante la aplicación de operadores de gradiente nos permitirá delimitar el comportamiento de las plantillas y formular $p(\mathcal{I} \mid \Theta)$.

En la Figura 3.1 presentamos un ejemplo de imagen biomédica (ver Capítulo 6) a partir de la cual se ha derivado un potencial binario. Una vez establecido el punto de partida en lo relativo al modelo de imagen introduciremos, siguiendo el enfoque

²La normalización del potencial puede enfocarse también desde el planteamiento *fuzzy* (Pal y Pal, 1993), lo cual permite incorporar la algoritmia desarrollada en procesamiento y análisis de imagen con este soporte: realzado, filtrado, segmentación y clasificación o *clustering*; con todo lo que ello supone.

³No obstante se han propuesto potenciales más específicos tales como la Transformada de Distancia (Leymarie y Levine, 1992).

Tabla 3.1: Plantilla Circular: Comportamiento

Función de Energía		
$\mathcal{E}(x, y, r)$	$\mathcal{E}_{con}(x, y, r) - \mathcal{E}_{exp}(r)$	
$\mathcal{E}_{exp}(r)$	$F(u, v, r)$	$\mathcal{E}_{con}(x, y, r)$
$\frac{\alpha}{a} r^a$	$(r - \sqrt{u^2 + v^2})$	$\iint_{\mathcal{D}} F(u, v, r) \Psi(x + u, y + v) dudv$
Dinámica por Gradiente		
Dinámica Centro	$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = - \begin{pmatrix} \frac{\partial \mathcal{E}}{\partial x} \\ \frac{\partial \mathcal{E}}{\partial y} \end{pmatrix} = \int_{\Gamma \cap \mathcal{D}} F(u, v, r) (-\vec{\nabla} \Psi) ds$	
Dinámica Radio	$\frac{dr}{dt} = -\frac{\partial \mathcal{E}}{\partial r} = -\int_{\mathcal{D}} \Psi(x + u, y + v) dudv + \alpha r^{a-1}$	

de primitivas deformables de Yuille, el modelo deformable local. En este sentido *prescindiremos inicialmente de introducir conocimiento probabilístico a priori* ya que este aspecto se incorporará, en el capítulo siguiente, desde un punto de vista estructural.

3.2 Plantillas Deformables Básicas

3.2.1 Plantillas Circulares

El *Círculo de Movimiento Libre*⁴, inicialmente propuesto en (Zhu y Yuille, 1994), incorpora una *tolerancia controlada al ruido interior y de frontera*. Su formulación, resumida en la Tabla 3.1, se basa en especificar una función de energía para guiar el comportamiento de una plantilla $g_{\mathcal{D}}(\Theta) = (x, y, r)$ siendo (x, y) la posición del centro, r el radio y \mathcal{D} el dominio circular delimitado por la primitiva. Considerando un potencial binario $\Psi(x + u, y + v)$ ⁵ la función de energía asociada es:

$$\mathcal{E}(x, y, r) = \mathcal{E}_{con}(x, y, r) - \mathcal{E}_{exp}(r) \quad (3.1)$$

donde:

$$\mathcal{E}_{con}(x, y, r) = \iint_{\mathcal{D}} F(u, v, r) \Psi(x + u, y + v) dudv$$

$$F(u, v, r) = (r - \sqrt{u^2 + v^2})$$

⁴Originalmente Denominado FTC: *Free-Travelling Circle*. Es una plantilla deformable orientada a encontrar el eje medio de una forma binaria (cálculo de esqueletos).

⁵Valor 0 en interior del objeto y 1 en el exterior.

54CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

$$\mathcal{E}_{exp}(r) = \frac{\alpha}{a} r^a$$

de manera que el *comportamiento óptimo es aquel que permite compensar la tendencia expansiva de la plantilla con el nivel de potencial no-programado establecido.*

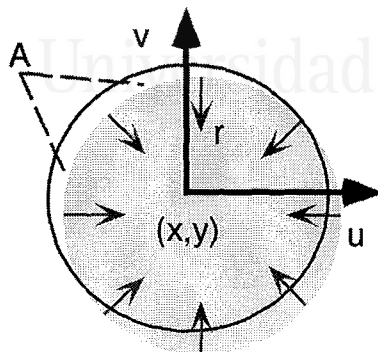


Figura 3.2: Componente de ruido (parámetro A) y fuerzas debidas al gradiente.

Este comportamiento se ilustra en la Figura 3.2 y es consecuencia de las siguientes dinámicas por gradiente:

1. *Dinámica del Centro:* siendo Γ el contorno de la región y $\vec{\nabla}\Psi$ el gradiente⁶:
 - (a) Se considera que $-\vec{\nabla}\Psi$ es una *fuerza de dirección opuesta ponderada por $r - \sqrt{u^2 + v^2}$* inversamente proporcional a la cercanía del pixel Γ al centro de la plantilla.
 - (b) El movimiento del centro es el resultado de la fuerza total integrada a lo largo del contorno de la región interior. El movimiento convergerá cuando todas las fuerzas estén equilibradas.
2. *Dinámica del Radio:* la derivada parcial con respecto al radio consta de dos términos:
 - (a) El primero es el *área de pixels blancos incluidos en el dominio \mathcal{D}* (ver área A en la Figura 3.2) y representa una fuerza que obliga al círculo a disminuir siempre que contenga pixels del exterior de la región.

⁶En el caso binario y en ausencia de ruido de imagen solamente es distinto de cero sobre Γ y es siempre perpendicular al contorno.

- (b) El segundo término representa una *fuerza de expansión*. La plantilla convergerá cuando A sea igual a αr^{a-1} . Este término nos proporciona la cantidad de ruido que se tolerará y va a depender de los valores de α y a :
- i. Si $a = 1$ entonces $A = \alpha$. Así, todos los círculos incluirán la misma cantidad de ruido al final de la convergencia (misma cantidad de pixels blancos). Esto puede no ser deseable pues α puede ser demasiado grande para círculos pequeños y demasiado pequeño para círculos grandes. Si $a = 1$ y $\alpha \rightarrow 0$, el modelo degenera hacia el círculo máximo definido por la transformación de ejes medios.
 - ii. Si $a < 3$ se garantiza convergencia si la inicialización es adecuada. Después de que el círculo alcance el tamaño de la región, el área A crecerá proporcionalmente a r^2 mientras que la fuerza de extensión crece solo $\alpha r^{a-1} < r^2$.

3.2.2 Ejemplo de Aplicación

En las Figura 3.3 se muestra un ejemplo de convergencia de este modelo. En la parte izquierda se muestra la inicialización de dos plantillas circulares. El círculo interior se inicializa en $(x = 123.0, y = 122.0, r = 20.0)$ con $\alpha = 10$ y $a = 2$.

Por otro lado el círculo exterior es concéntrico, siendo $r = 50.0$ $\alpha = 20$, $a = 2$, y reacciona ante un potencial inverso. Además los pixels interiores a la plantilla menor se han ajustado a intensidad máxima para facilitar el ajuste de la plantilla exterior



Figura 3.3: Inicialización y adaptación de círculos.

3.3 Generalización: Plantillas Super-Elipsoidales

3.3.1 Directrices de Modelización

Partiendo del modelo circular básico nos hemos planteado su extensión para el caso de formas cerradas más complejas. En este sentido buscamos siempre un *equilibrio entre la capacidad expresiva de la estructura y la complejidad computacional del procedimiento de ajuste*, así como un *comportamiento robusto ante ruido gaussiano*. Adoptando para el caso de potenciales de intensidad, donde disponemos de información de gradiente espacial, el enfoque de super-cuádricas 3D definidas sobre imágenes de rango (nubes de puntos 3D), implementamos, mediante el *diseño de una función de energía*, la adaptación al potencial y la extracción de los parámetros característicos. En líneas generales el nuevo modelo, resumido en la Tabla 3.2, se ha obtenido a partir de las siguientes directrices:

1. *Generalización de Forma*: la consideración de formas de complejidad cuadrática implica la *sustitución del peso $F(u, v, r)$* tomando como referencia la ecuación implícita asociada al tipo de plantilla. Consecuentemente *la nueva dinámica afecta a parámetros de forma y escala (ejes)*.
2. *Conducta de Rotación*: por otro lado hemos de incorporar una *conducta de giro* que permita rotar la plantilla por inducción del potencial no programado. Explotamos, para ello, *la propiedad de ortogonalidad axial* de las formas consideradas.
3. *Control del Ruido*: el control del ajuste se obtiene especificando *las proporciones de ruido* permitidas tanto para los ejes como para la adaptación de la forma. Esto implica definir condiciones de equilibrio adecuadas y explotarlas para el diseño de la función de energía.

3.3.2 Definición: Potencial y Función de Energía

Definición 3.2 Una *Plantilla Deformable Super-Elipsoidal* es una primitiva 2D variable y expresable por $g_{\mathcal{D}}(\Theta) = (x, y, a, b, \vec{V}_{ref}, \epsilon)$ donde: (x, y) es el centro, a y b son los ejes (mayor y menor respectivamente), $\vec{V}_{ref} = (\cos \theta, \sin \theta)$ es el vector de referencia (siendo θ la orientación en el plano del eje secundario), $\epsilon > 0$ es el parámetro de forma (ver Figuras 3.4, 3.5), y $\mathcal{D} \subseteq \mathbb{R}^2$ el dominio interno acotado por la frontera cerrada Γ .

Tabla 3.2: Plantilla SuperElipsoidal: Comportamiento

Función de Energía	
$\mathcal{E}(x, y, a, b, \vec{V}_{ref}, \epsilon)$	$\mathcal{E}_{con}(x, y, a, b, \epsilon) + \mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon) - \mathcal{E}_{exp}(x, y, a, b, \epsilon)$
$F(u, v, a, b, \epsilon)$	$[1 - (\frac{ u ^\epsilon}{a^\epsilon} + \frac{ v ^\epsilon}{b^\epsilon})]$
$\mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon)$	$(\mathcal{N}(\vec{V}_{ref}) \cdot \mathcal{N}(\vec{\nabla}\Psi(p_1)))^2$
$\mathcal{E}_{con}(x, y, a, b, \epsilon)$	$\iint_{\mathcal{D}} F(u, v, a, b, \epsilon) \Psi(x + u, y + v) dudv$
$\mathcal{E}_{exp}(x, y, a, b, \epsilon)$	$[\alpha(a + b) + \beta\epsilon] \iint_{\mathcal{D}} F(u, v, a, b, \epsilon) dudv$
Dinámica por Gradiente	
Din. del Centro $\frac{\partial(x,y)^T}{\partial t}$	$\int_{\Gamma \cap \mathcal{D}} F(u, v, a, b, \epsilon) (-\vec{\nabla}\Psi) ds - 2\mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon)^{\frac{1}{2}} \mathcal{N}(\vec{V}_{ref}) \mathcal{N}(\vec{\nabla}^2\Psi(p_1$
Din. de Rotación $\frac{\partial\vec{V}_{ref}}{\partial t}$	$-2\mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon)^{\frac{1}{2}} \mathcal{N}(\vec{\nabla}\Psi(p_1))$
Din. Axial $\frac{dr}{dt}$	$-\iint_{\mathcal{D}} \gamma(r, s, \epsilon) \Psi(x + u, y + v) dudv + \alpha \iint_{\mathcal{D}} \gamma(r, s, \epsilon) dudv$
Din. de Forma $\frac{d\epsilon}{dt}$	$-\iint_{\mathcal{D}} \delta(u, v, a, b, \epsilon) \Psi(x + u, y + v) dudv + \beta \iint_{\mathcal{D}} \delta(u, v, a, b, \epsilon) dudv$

58CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

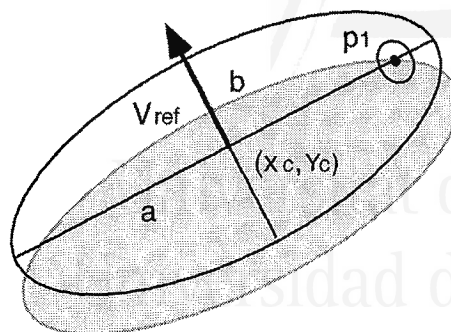
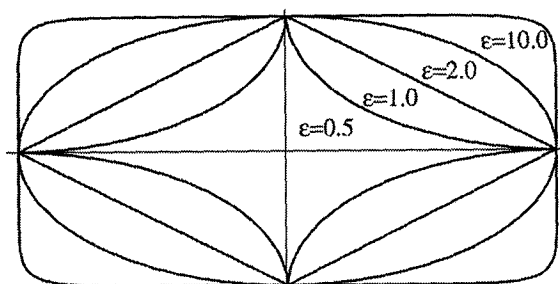


Figura 3.4: Plantilla deformable elíptica

Figura 3.5: Efecto del parámetro ϵ

La variación inducida por el ajuste a un potencial binario $\Psi(x+u, y+v)$ viene regida por la siguiente *función de energía*⁷ (ver Tabla 3.2):

$$\mathcal{E}_{con}(x, y, a, b, \epsilon) + \mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon) - \mathcal{E}_{exp}(x, y, a, b, \epsilon) \quad (3.2)$$

cuyos términos concretamos a continuación:

1. *Energía de Contracción*: la incidencia de potencial no-programado (ruido) en la reducción de los parámetros de posición, escala y forma viene regido por:

$$\mathcal{E}_{con}(x, y, a, b, \epsilon) = \iint_{\mathcal{D}} F(u, v, a, b, \epsilon) \Psi(x+u, y+v) dudv \quad (3.3)$$

donde $F(u, v, a, b, \epsilon) = [1 - (\frac{|u|^\epsilon}{a^\epsilon} + \frac{|v|^\epsilon}{b^\epsilon})]$ es el peso asociado, en función de su distribución espacial, a cada uno de los puntos de ruido pertenecientes en

⁷En este sentido optamos por la definición usual en términos continuos, la cual deberá traducirse en una implementación discreta, tal como veremos en la sección dedicada a experimentación básica.

el dominio de la plantilla. Este peso tiene, como veremos, un fuerte impacto en la dinámica de centrado.

2. *Energía de Rotación*: la conducta de adaptación del vector de referencia se basa en:

$$\mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon) = (\mathcal{N}(\vec{V}_{ref}) \cdot \mathcal{N}(\vec{\nabla}\mathcal{I}(p_1)))^2 \quad (3.4)$$

donde \mathcal{N} indica la normalización del vector que toma como argumento, el operador \cdot representa el producto escalar y p_1 es el primer punto del semieje positivo de a con gradiente no-nulo, es decir:

$$p_1 = (x + i \cos(\theta - \frac{\pi}{2}), y + i \sin(\theta - \frac{\pi}{2})) \mid 0 < i < a, \vec{\nabla}\mathcal{I}(p_1) \neq 0 \quad (3.5)$$

3. *Energía de Expansión*: la contribución de cada uno de los puntos de potencial programado al incremento de escala y curvatura se expresa mediante:

$$\mathcal{E}_{exp}(x_c, y_c, a, b, \epsilon) = [\alpha(a + b) + \beta\epsilon] \int \int_{\mathcal{D}} F(u, v, a, b, \epsilon) du dv \quad (3.6)$$

siendo $\alpha, \beta \in (0, 1]$ las proporciones límite de aceptación de ruido para ambos ejes y para la forma respectivamente. La especificación de este término, similar en forma al término de contracción, permite equilibrar, siempre atendiendo a la ponderación establecida mediante α y β , el efecto del potencial no programado sobre la dinámica de la plantilla. Este término se justifica a partir de las condiciones de equilibrio que seguidamente presentaremos.

En los apartados siguientes detallamos el comportamiento dinámico, ante un potencial dado, de las plantillas cuadráticas.

3.3.3 Dinámicas de Centrado y Rotación

Dinámica de Centrado

Al igual que en el caso circular, el ajuste del centro viene influido por la toma de contacto de la primitiva con el gradiente (frontera):

$$\begin{pmatrix} \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial t} \end{pmatrix} = - \begin{pmatrix} \frac{\partial \mathcal{E}}{\partial x} \\ \frac{\partial \mathcal{E}}{\partial y} \end{pmatrix} = \int_{\Gamma \cap \mathcal{D}} F(u, v, a, b, \epsilon) (-\vec{\nabla}\Psi) ds - 2\mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon)^{\frac{1}{2}} (\mathcal{N}(\vec{V}_{ref})\mathcal{N}(\vec{\nabla}^2\Psi(p_1))) \quad (3.7)$$

60CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

donde $\mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon)^{\frac{1}{2}} = (\mathcal{N}(\vec{V}_{ref}) \cdot \mathcal{N}(\vec{\nabla}\mathcal{I}(p_1)))$. El efecto del gradiente induce un centrado tanto a nivel global de la primitiva (primer término) como a nivel del punto p_1 (segundo término).

A nivel global se tienen en cuenta los *pesos de desplazamiento*, es decir la función $F(u, v, a, b, \epsilon)$, asociados a los puntos del dominio. En la Figura 3.6 se muestra la variación de las funciones de peso para una plantilla con $a = 30$ y $b = 10$ a medida que decrece el valor de ϵ (de arriba a abajo y de izquierda a derecha). La disminución progresiva del parámetro de forma induce concavidad y forma espiculada al tiempo que se reduce el número de puntos del dominio, con lo cual se acentúa la diferencia entre los pesos máximos asociados a los puntos situados en las proximidades de los ejes y el resto de pesos.

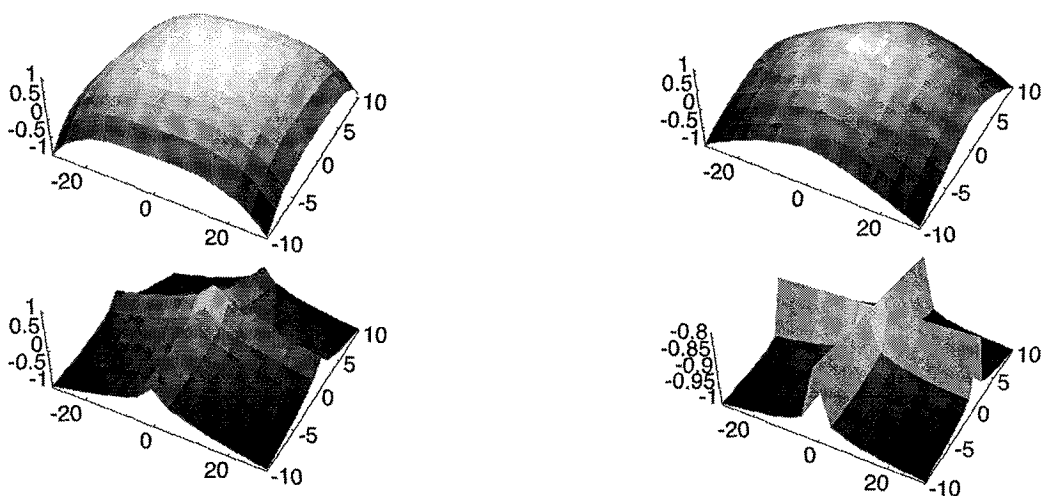


Figura 3.6: Funciones de peso: $\epsilon = 3.5, \epsilon = 1.5, \epsilon = 0.5$ y $\epsilon = 0.03$.

A nivel local el segundo término tiene un efecto secundario (en sentido positivo o negativo) siempre y cuando el vector de referencia no sea ortogonal con respecto al de gradiente en p_1 o a su variación (Laplaciana) en dicho punto, lo cual define la situación de equilibrio.

Dinámicas de Rotación

Dado que la primitiva puede no estar correctamente orientada debe incorporarse un término que induzca un giro para buscar el posicionamiento correcto. Para ello se recurre al cuadrado del producto escalar entre el vector de referencia y el de

gradiente en p_1 con lo cual

$$\begin{pmatrix} \frac{\partial \vec{V}_{ref_x}}{\partial t} \\ \frac{\partial \vec{V}_{ref_y}}{\partial t} \end{pmatrix} = - \begin{pmatrix} \frac{\partial \mathcal{E}}{\partial \vec{V}_{ref_x}} \\ \frac{\partial \mathcal{E}}{\partial \vec{V}_{ref_y}} \end{pmatrix} = -2\mathcal{E}_{rot}(x, y, \vec{V}_{ref}, \epsilon)^{\frac{1}{2}} \mathcal{N}(\vec{\nabla} \Psi(p_1)) \quad (3.8)$$

Esta expresión nos proporciona una estimación de la desviación con respecto a la orientación ideal definida por la *ortogonalidad*. No obstante *para dotar de robustez* al modelo p_1 se promedia en un dominio local definido por una ventana circular (ver Figura 3.4). En caso de que dicho punto no se encuentre (p.e. porque la plantilla no ha capturado el contorno en esa dirección) se propone la búsqueda de p_1 en el semieje negativo. En definitiva esto permite inducir giros hasta que se localiza una orientación localmente óptima.

3.3.4 Dinámicas Axiales y de Forma

Dinámica Axial

La política de variación para cada eje tiene en cuenta el porcentaje límite de ruido admitido. Suponiendo los siguientes cambios de variable $r = a, s = u$ o bien $r = b, s = v$ expresamos de forma compacta dicho comportamiento:

$$\frac{dr}{dt} = - \frac{\partial \mathcal{E}}{\partial r} = - \int \int_{\mathcal{D}} \gamma(r, s, \epsilon) \Psi(x + u, y + v) du dv + \alpha \int \int_{\mathcal{D}} \gamma(r, s, \epsilon) du dv \quad (3.9)$$

siendo $\gamma(r, s, \epsilon) = \frac{\epsilon s \left(\frac{|s|}{r}\right)^{\epsilon-1}}{r^2}$ un factor de ponderación de la componente s de la fuerza que actúa sobre cada punto del dominio (cada componente afecta a un eje). El término de la izquierda integra las fuerzas de los puntos de ruido y el de la derecha integra fuerzas de sentido contrario definidas sobre todos los puntos incluidos en \mathcal{D} . El equilibrio se produce cuando las resultantes sobre el ruido se compensan con las fuerzas de potencial programado, es decir, cuando la proporción global de fuerzas P_r tiende al valor α :

$$P_r = \frac{\int \int_{\mathcal{D}} \gamma(r, s, \epsilon) \Psi(x + u, y + v) du dv}{\int \int_{\mathcal{D}} \gamma(r, s, \epsilon) du dv} \rightarrow \alpha \quad (3.10)$$

En caso de que $P_r > \alpha$ (la resultante del ruido es mayor) se produce un descenso en el eje r para reducir el nivel de ruido. En cambio si $P_r < \alpha$ el radio afectado crecerá. En consecuencia P_r nos proporciona una métrica para estimar la distancia al objetivo (aceptar un determinado nivel de ruido).

62CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

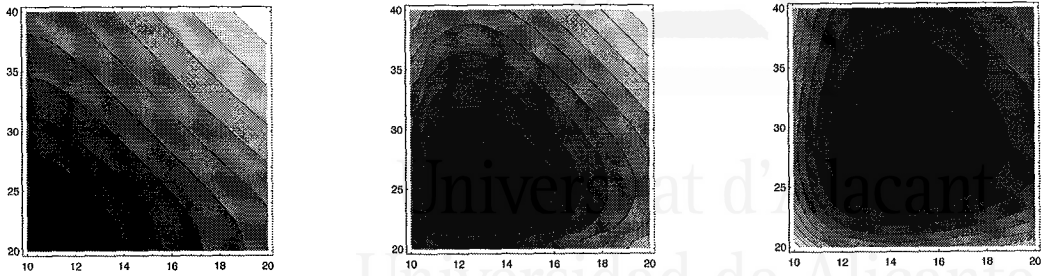


Figura 3.7: Desplazamiento del punto de equilibrio en función de α .

La introducción del parámetro α tiene el efecto de modificar el punto de equilibrio de la plantilla en lo que respecta al posicionamiento final de sus ejes sobre el potencial. Dicho efecto puede explicarse tomando como ejemplo un potencial sin ruido Gaussiano el cual contiene una plantilla elíptica ($\epsilon = 2.0$) con $a = 10$ y $b = 20$. Representando la función de error $(P_a - \alpha)^2 + (P_b - \alpha)^2$ para $a \in [10.0, 20.0]$ y $b \in [20.0, 40.0]$ observamos que: si $\alpha = 0.0$, lo cual corresponde a una situación ideal, (ver Figura 3.7-izquierda) el punto de equilibrio (mínimo de la función de error) para los ejes se sitúa en $(10.0, 20.0)$; cuando $\alpha = 0.2$ (Figura 3.7-centro) el equilibrio se desplaza hacia $(12.5, 25.0)$; y, finalmente, si $\alpha = 0.3$ (Figura 3.7-derecha) el equilibrio se alcanza en $(14.0, 30.0)$. A medida que aumenta α el equilibrio se desplaza y la superficie de error se hace menos restrictiva en el dominio considerado.

Dinámica de Forma

Finalmente, la adaptación interactiva de la plantilla a la forma contenida en la imagen se aborda de forma similar al tratamiento de los ejes, es decir mediante la ecuación:

$$\frac{d\epsilon}{dt} = -\frac{\partial \mathcal{E}}{\partial \epsilon} =$$

$$-\int \int_{\mathcal{D}} \delta(u, v, a, b, \epsilon) \Psi(x + u, y + v) du dv + \beta \int \int_{\mathcal{D}} \delta(u, v, a, b, \epsilon) du dv \quad (3.11)$$

siendo $\delta(u, v, a, b, \epsilon) = -\left[\left(\frac{|u|^\epsilon}{a^\epsilon}\right) \ln\left(\frac{|u|}{a}\right)\right] - \left[\left(\frac{|v|^\epsilon}{b^\epsilon}\right) \ln\left(\frac{|v|}{b}\right)\right]$ la fuerza que pondera a cada punto del dominio. Dicha fuerza tiene sentido positivo ya que $u \leq a$ y $v \leq b$. Al igual que en el caso de los ejes el equilibrio se produce cuando la proporción de fuerzas entre ruido y potencial se hace β . De forma análoga al caso de los ejes definimos el punto de equilibrio de la siguiente manera:

$$P_\epsilon = \frac{\iint_{\mathcal{D}} \delta(u, v, a, b, \epsilon) \Psi(x + u, y + v) du dv}{\iint_{\mathcal{D}} \delta(u, v, a, b, \epsilon) du dv} \rightarrow \beta \quad (3.12)$$

siendo P_ϵ la proporción de ruido admitida por el parámetro de forma. Si $P_\epsilon > \beta$ se induce un descenso de ϵ con lo cual nos aproximamos a figuras cada vez más cóncavas. En caso de que $P_\epsilon < \beta$ se tiende a figuras convexas.

El análisis de la incidencia de β en el punto de equilibrio para el parámetro de forma puede realizarse tomando como referencia la función de error $(P_\epsilon - \beta)^2$ con resultados similares a los obtenidos en el análisis de los radios.

3.3.5 Revisión de la Condición de Equilibrio

En definitiva, el estudio previo de las condiciones de equilibrio para cada uno de los parámetros del modelo nos revela que el comportamiento local de la plantilla está influenciado por tres aspectos básicos:

1. *Condiciones de Inicialización:* en ausencia de conocimiento probabilístico es *condición necesaria* para la convergencia del modelo inicializar la plantilla de manera que su dominio contenga al menos un punto de la figura cuya geometría deseamos extraer⁸.
2. *Incidencia del Gradiente:* las dinámicas de centrado y rotación dependen del contacto de la primitiva deformable con el contorno. Si en los primeros pasos este contacto no se produce la plantilla evolucionará guiada por la dinámica de los ejes y de la forma.
3. *Control del Ruido:* los parámetros α y β establecen la proporción de potencial no programado que estamos dispuestos a aceptar tanto en la adaptación de los ejes como en la de la forma⁹. Este diseño permite *modelar de forma específica* la sensibilidad a cambios de escala por un lado, y de forma por otro, captando las posibles diferencias en orden de magnitud entre los valores de los ejes y los del parámetro ϵ (usualmente $\epsilon \in (0, 5]$). Esto permite reducir el número de parámetros a sintonizar.

En cuanto a *eficiencia de computacional* cada parámetro puede ajustarse de forma independiente con lo cual el coste es divisible (paralelización). Por otro lado el *origen del modelo* que acabamos de describir es la extensión de la plantilla circular

⁸En definitiva el dominio de la plantilla es la ventana local de búsqueda lo cual implica posibles *problemas de apertura* si no se dispone de conocimiento probabilístico local o estructural.

⁹Si permitiésemos $\alpha = 0$ o $\beta = 0$ ello induciría una dinámica de contracción y si ambos son iguales a la unidad la fuerza de expansión progresa hasta admitir el 100% de ruido.

64CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

al caso elíptico. Dicho modelo se presenta en el Capítulo 6 aplicándose al *tracking* en secuencias intravasculares generadas por ultrasonidos.

3.4 Experimentación: Análisis Local

3.4.1 Objetivos y Condiciones Experimentales

La función de energía definida por las Ecuaciones 3.2- 3.6 modela un comportamiento local (asociado a plantillas individuales). En esta sección analizaremos la bondad de la función *en ausencia de conocimiento probabilístico acerca de la geometría del modelo de imagen*. Hemos realizado diversos experimentos orientados a estudiar tres aspectos centrales:

1. *Análisis de Sensibilidad*: evaluar el impacto de las condiciones de inicialización en el resultado final en función de los parámetros no-ajustados. Esto nos permitirá establecer qué parámetros son más sensibles.
2. *Equilibrio y Ruido Relativo*: estudiar el impacto de los coeficientes α y β en el control del potencial no-programado.
3. *Comportamiento Dinámico*: adaptatividad ante variaciones de posición, orientación, forma y escala debidas al carácter dinámico de la imagen.

3.4.2 Análisis y Discusión

Partiendo de un potencial de imagen que contiene un superelipsoide de $x = 120.0$, $y = 120.0$, $a = 70.0$, $b = 40.0$, $\theta = 0.0$ y $\epsilon = 2.0$ estimaremos la respuesta ante distintas condiciones de inicialización y niveles de ruido gaussiano (0%, 10%, 20%). Se han realizado múltiples experimentos. Los mejores resultados para cada caso pueden visualizarse en las Figuras 3.8-3.13 y se resumen en la Tabla 3.3. Cabe distinguir tres bloques de experimentos. El análisis de resultados es el siguiente:

Bloque 3.1 (Traslación y Escala) (*Experimentos 1-3*) En primer lugar, se ha demostrado (ver Figuras 3.8-3.10) la robustez de las dinámicas de centrado y rotación incluso en niveles altos de ruido (con el ajuste adecuado de α y β). La descompensación introducida por el ruido es corregida por la energía de los puntos de potencial incluidos en el dominio de la plantilla. Por otro lado observamos una menor robustez en las dinámicas de rotación y forma (Casos 2.2 y 2.3).

Bloque 3.2 (Rotación y Forma) (*Experimentos 4-6*) En estos experimentos se obtienen, por un lado, buenos resultados (Casos 5.1, 5.2, 6.1, 6.2, 6.3)¹⁰ pero

¹⁰Teniendo en cuenta el problema de apertura.

también se pone de relieve la sensibilidad de las dinámicas de rotación y forma en ausencia de conocimiento previo. Ello se debe, en el caso de la rotación, a que la información de giro se extrae en ventanas locales con lo cual se sufre de un problema de apertura. En el caso de la forma la causa es la configuración formal de los factores de ponderación. Una cantidad elevada de potencial no programado (ruido gaussiano o de frontera) impacta fuertemente a favor de la reducción del parámetro ϵ . El ajuste adecuado de α y β consigue paliar este efecto.

En términos generales, los resultados del modelo local son satisfactorios pero, como suele ser habitual en modelos de plantillas deformables, la función de energía debe ser asistida por un modelo probabilístico. En este sentido pretendemos paliar los problemas de localidad actuando en dos direcciones. Por un lado dotar al modelo del conocimiento previo más relevante. Por otro adquirir este conocimiento mediante el aprendizaje de restricciones geométricas estructurales.

3.5 Consideraciones

A lo largo de este capítulo hemos introducido una nueva función de energía basada en potenciales de intensidad y contorno, capaz de modelar curvas cuadráticas cerradas (super-elipsoides) como primer elemento de nuestro modelo extendido. Hemos estudiado, tanto formal como experimentalmente, el comportamiento local de cada plantilla a nivel individual y *sin conocimiento probabilístico previo*, obteniendo conclusiones parciales acerca de la tolerancia a ruido tanto gaussiano como de frontera. A partir de este punto procede definir un enfoque conceptual que permita adquirir el conocimiento estructural adecuado para integrar eficientemente diversos modelos locales en el dominio espacial y para guiar su evolución a lo largo de una secuencia de imágenes.

66CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

Tabla 3.3: Experimentos 1-6. Evaluación de Función de Energía

Pruebas	Casos	α	β	P_a	P_b	P_c	Iter.
Exp. 1	1.1	0.1	0.1	0.089	0.111	0.058	40
	1.2	0.1	0.1	0.014	0.151	0.101	40
	1.3	0.2	0.2	0.195	0.192	0.166	40
Exp. 2	2.1	0.1	0.1	0.081	0.093	0.058	40
	2.2	0.1	0.1	0.213	0.216	0.166	40
	2.3	0.1	0.1	0.216	0.206	0.143	40
Exp. 3	3.1	0.1	0.1	0.162	0.163	0.099	40
	3.2	0.1	0.1	0.097	0.086	0.073	40
	3.3	0.1	0.1	0.146	0.124	0.110	40
Exp. 4	4.1	0.2	0.0	0.222	0.222	0.124	100
	4.2	0.2	0.2	0.182	0.188	0.112	100
	4.3	0.2	0.2	0.197	0.192	0.166	200
Exp. 5	5.1	0.1	0.1	0.159	0.137	0.100	150
	5.2	0.2	0.1	0.140	0.161	0.094	150
	5.3	0.2	0.2	0.214	0.221	0.189	150
Exp. 6	6.1	0.2	0.1	0.211	0.214	0.147	250
	6.2	0.2	0.1	0.203	0.205	0.140	250
	6.3	0.2	0.05	0.203	0.205	0.125	250
	6.4	0.3	0.2	0.325	0.343	0.218	250

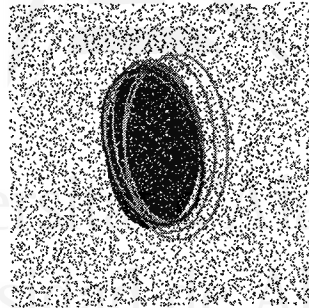
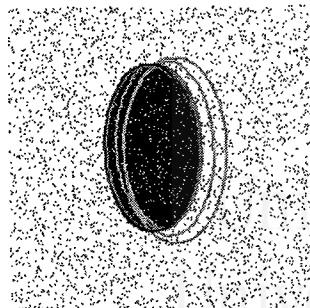
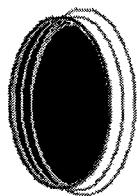


Figura 3.8: Experimento 1

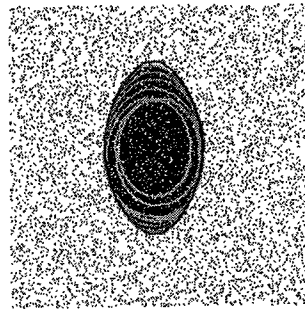
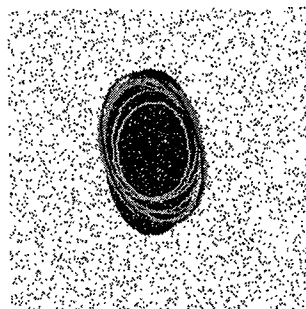


Figura 3.9: Experimento 2

68CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES

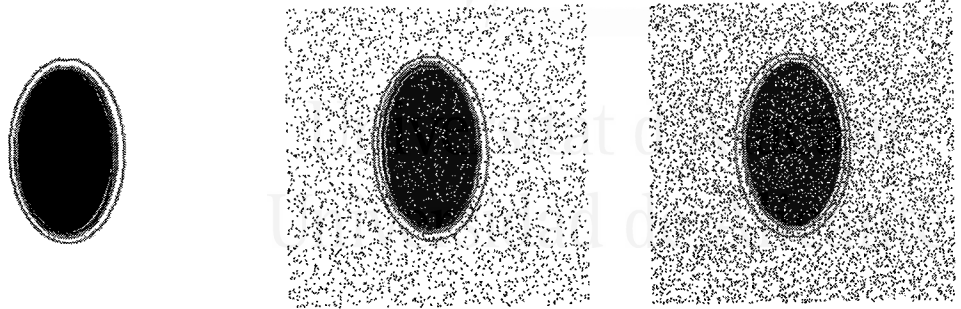


Figura 3.10: Experimento 3

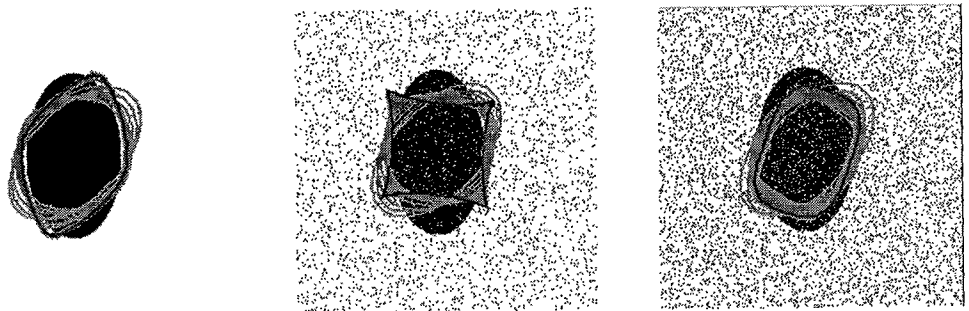


Figura 3.11: Experimento 4

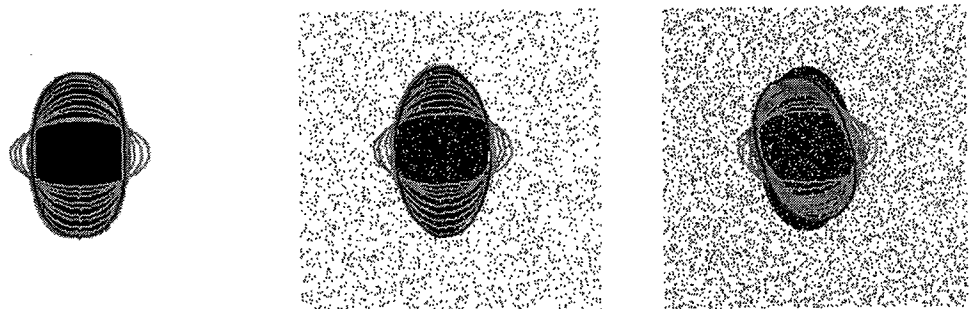


Figura 3.12: Experimento 5



Universitat d'Alacant
Universidad de Alicante

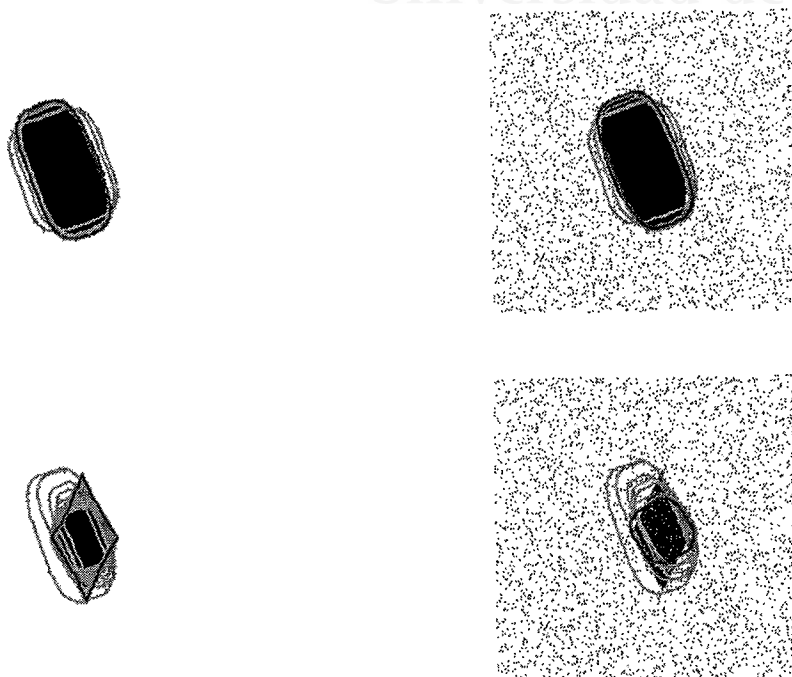


Figura 3.13: Experimento 6

70CAPÍTULO 3. CARACTERIZACIÓN LOCAL: PLANTILLAS SUPER-ELIPSOIDALES



Universitat d'Alacant
Universidad de Alicante

Universitat d'Alacant
Universidad de Alicante

Capítulo 4

Caracterización Estructural: Auto-Restricciones

Una vez establecidas, en el capítulo anterior, las primitivas geométricas soporte y los mecanismos de extracción y ajuste de las mismas a nivel individual, tratamos ahora la definición y adquisición de la información relacional. En este capítulo abordaremos la caracterización estructural mediante el ACP. Como veremos, este enfoque nos permite aprender e incorporar información probabilística compacta al modelo geométrico, lo cual repercutirá, como veremos en el capítulo siguiente, en la mejora del rendimiento de la función de energía. En primer lugar presentaremos el modelo de *Auto-Restricciones* espaciales y espacio-temporales. Especificaremos diversos tipos de restricciones, definidas a su vez sobre los parámetros que caracterizan a conjunto de plantillas super-elipsoidales, y presentaremos el mecanismo de aprendizaje (generación de auto-espacios definidos por distribuciones de probabilidad y caracterizados por los modos principales). A continuación introduciremos la métrica de pertenencia al modelo (satisfacibilidad de auto-restricciones). A lo largo del capítulo aportaremos ejemplos y resultados experimentales de adquisición, caracterización y reconocimiento.

4.1 Modelado Basado en Auto-Restricciones

4.1.1 Directrices de Modelización

En la línea de las Auto-Representaciones¹ el modelo de *Auto-Restricciones* o *Restricciones Propias* se orienta hacia la *expresión de restricciones o condicionantes*

¹*EigenModels*. Como vimos anteriormente esta denominación se aplica a los modelos de representación basados en Componentes Principales o Elementos Finitos.

72 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

geométricos, espaciales y espacio-temporales en términos de un modelo promedio y de sus variaciones o deformaciones más significativas, de manera que el conjunto de tuplas que satisfacen dichas restricciones se representa mediante una distribución de probabilidad acotada. Este enfoque se caracteriza por:

1. *Representación de Baja Dimensionalidad*: adquisición de conocimiento probabilístico sobre la estructura geométrica subyacente de manera que estén representados los parámetros principales (reducción de dimensionalidad). Modelización de la distribución a priori $p(\Theta)$
2. *Integración ACP-Satisfacción-Modelado*: caracterización de restricciones en términos de distribuciones multidimensionales con lo cual se incorpora a las aproximaciones clásicas para resolver el problema de Satisfacción de Restricciones (Freuder y Mackworth, 1994), (Grimson, 1990) técnicas de ACP. Por otro lado se posibilita la síntesis automática de restricciones mediante métodos de minimización de energía (Cohen, 1992).
3. *Ajuste-Minimización de Energía*: los dos aspectos anteriores llevan consigo la propuesta de funciones de energía a nivel global (estructural) que integren (de forma ponderada) el conocimiento adquirido por la interacción de las plantillas individuales con el modelo de imagen, con la satisfacción incremental de restricciones.

Tabla 4.1: Auto-Restricciones Espaciales

Restricciones Absolutas	
Posición	$\vec{r}_{xy} = (x_1, y_1, x_2, y_2, \dots, x_g, y_g)^T$
Escala	$\vec{r}_{ab} = (a_1, b_1, \dots, a_g, b_g)^T$
Orientación	$\vec{r}_{\theta} = (\theta_1, \theta_2, \dots, \theta_g)^T$
Forma	$\vec{r}_{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_g)^T$
Restricciones Relativas	
Distancia	$\vec{r}_{\rho} = (\rho_{12}, \dots, \rho_{1g}, \rho_{23}, \dots, \rho_{2g}, \dots, \rho_{g-2g-1}, \dots, \rho_{g-1g})^T$
Escala	$\vec{r}_{\lambda} = (\lambda_{12}, \dots, \lambda_{1g}, \dots, \lambda_{g-2}, \dots, \lambda_{g-2g-1}, \dots, \lambda_{g-1g})^T$
Rotación	$\vec{r}_{\alpha} = (\alpha_{12}, \dots, \alpha_{1g}, \alpha_{23}, \dots, \alpha_{2g}, \alpha_{g-2}, \dots, \alpha_{g-2g-1}, \dots, \alpha_{g-1g})^T$
Forma	$\vec{r}_{\kappa} = (\kappa_{12}, \dots, \kappa_{1g}, \kappa_{23}, \dots, \kappa_{2g}, \kappa_{g-2}, \dots, \kappa_{g-2g-1}, \dots, \kappa_{g-1g})^T$

4.1.2 Auto-Restricciones Espaciales Absolutas y Relativas

Definición 4.1 Dado un conjunto $\mathcal{G} = \{g_{\mathcal{D}_i}(\Theta)\}_{i=1}^g$ de Plantillas Deformables Super-Elipsoydales, el conjunto característico asociado \mathcal{R} de *Auto-Restricciones o Restricciones Propias Iniciales* es una colección de relaciones geométricas absolutas y/o relativas $r \in \mathcal{R} = \mathcal{R}_{abs} \cup \mathcal{R}_{rel}$ de manera que $r = \mathcal{A}(\vec{r}) \subseteq R^d \mid \vec{r} \in R^d$ y donde cabe distinguir:

1. *Vectores Característicos:* $\vec{r} \in R^d$, cuyas componentes se definen a partir de Θ , indican el tipo de parámetro o aspecto geométrico (escala, orientación, posición o forma) caracterizado.
2. *Subespacios Asociados:* $\mathcal{A}(\vec{r}) \subseteq R^d$, o distribuciones gaussianas o normales d -dimensionales definidas sobre \vec{r} , indican el conjunto de tuplas que satisfacen r .

En base a esta definición establecemos que $\mathcal{R}_{abs} \subseteq \{r_{xy}, r_{ab}, r_{\theta}, r_{\epsilon}\}$ y $\mathcal{R}_{rel} \subseteq \{r_{\rho}, r_{\lambda}, r_{\alpha}, r_{\kappa}\}$ siendo los vectores característicos asociados a dichas restricciones los siguientes (ver Tabla 4.1):

1. *Restricciones Absolutas:* caracterizadas por vectores que incluyen, según el caso, las componentes de escala (\vec{r}_{ab}), orientación (\vec{r}_{θ}), posición (\vec{r}_{xy}), y forma (\vec{r}_{ϵ}) de las plantillas de \mathcal{G} .
2. *Restricciones Relativas:* caracterizadas por vectores que codifican aspectos relativos en términos de restricciones cuadráticas:
 - (a) *Distancia:* (\vec{r}_{ρ}) con $\rho_{ij} = \|(x_i, y_i) - (x_j, y_j)\|^2$.
 - (b) *Orientación:* (\vec{r}_{α}) con $\alpha_{ij} = (\theta_i - \theta_j)^2$.
 - (c) *Escala:* (\vec{r}_{λ}) con $\lambda_{ij} = \|(a_i, b_i) - (a_j, b_j)\|^2$.
 - (d) *Forma:* (\vec{r}_{ϵ}) con $\kappa_{ij} = (\epsilon_i - \epsilon_j)^2$.

A efectos de capacidad expresiva, las restricciones relativas permiten capturar aspectos de invarianza espacial no expresables mediante restricciones absolutas. Tanto las restricciones absolutas como las relativas caracterizan estructuralmente a un conjunto de plantillas, pero *solo a nivel espacial*. La consideración de la dimensión temporal en el modelo previo implica incorporar campos de potencial variables. La extensión a partir de la formulación estática es directa y se presenta a continuación.

74 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

Tabla 4.2: Auto-Restricciones Espacio-Temporales

Restricciones Absolutas	
Posición	$\vec{r}_{xy}(t) = (\vec{x}_1(t), \vec{y}_1(t), \vec{x}_2(t), \vec{y}_2(t), \dots, \vec{x}_g(t), \vec{y}_g(t))^T$
	$\vec{x}_i(t) = (x_i(0), \dots, x_i(T))^T$ $\vec{y}_i(t) = (y_i(0), \dots, y_i(T))^T$
Escala	$\vec{r}_{ab}(t) = (\vec{a}_1(t), \vec{b}_1(t), \dots, \vec{a}_g(t), \vec{b}_g(t))^T$
	$\vec{a}_i(t) = (a_i(0), \dots, a_i(T))^T$ $\vec{b}_i(t) = (b_i(0), \dots, b_i(T))^T$
Rotación	$\vec{r}_\theta(t) = (\vec{\theta}_1(t), \vec{\theta}_2(t), \dots, \vec{\theta}_g(t))^T$
	$\vec{\theta}_i(t) = (\theta_i(0), \dots, \theta_i(T))^T$
Forma	$\vec{r}_\epsilon(t) = (\vec{\epsilon}_1(t), \vec{\epsilon}_2(t), \dots, \vec{\epsilon}_g(t))^T$
	$\vec{\epsilon}_i(t) = (\epsilon_i(0), \dots, \epsilon_i(T))^T$
Restricciones Relativas	
Distancia	$\vec{r}_\rho(t) = (\vec{\rho}_{12}(t), \dots, \vec{\rho}_{1g}(t), \dots, \vec{\rho}_{g-2g}(t), \vec{\rho}_{g-1g}(t))^T$
	$\vec{\rho}_{ij}(t) = (\rho_{ij}(0), \dots, \rho_{ij}(T))^T$
Escala	$\vec{r}_\lambda(t) = (\vec{\lambda}_{12}(t), \dots, \vec{\lambda}_{1g}(t), \dots, \vec{\lambda}_{g-2g}(t), \vec{\lambda}_{g-1g}(t))^T$
	$\vec{\lambda}_{ij}(t) = (\lambda_{ij}(0), \dots, \lambda_{ij}(T))^T$
Rotación	$\vec{r}_\alpha(t) = (\vec{\alpha}_{12}(t), \dots, \vec{\alpha}_{1g}(t), \dots, \vec{\alpha}_{g-2g}(t), \vec{\alpha}_{g-1g}(t))^T$
	$\vec{\alpha}_{ij}(t) = (\alpha_{ij}(0), \dots, \alpha_{ij}(T))^T$
Forma	$\vec{r}_\kappa(t) = (\vec{\kappa}_{12}(t), \dots, \vec{\kappa}_{1g}(t), \dots, \vec{\kappa}_{g-2g}(t), \vec{\kappa}_{g-1g}(t))^T$
	$\vec{\kappa}_{ij}(t) = (\kappa_{ij}(0), \dots, \kappa_{ij}(T))^T$

4.1.3 Auto-Restricciones Espacio-Temporales

Definición 4.2 Dado un conjunto $\mathcal{G}(t) = \{g_{\mathcal{D}_i}(t, \Theta)\}_{i=1}^g$, con $t \in [0, T]$, de Plantillas Deformables Super-Elipsoidales dependientes del tiempo, el conjunto característico asociado $\mathcal{R}(t)$ de *Auto-Restricciones o Restricciones Espacio-Temporales Propias Iniciales*, definidas sobre la ventana temporal² $[0, T]$, es una colección de relaciones geométricas absolutas y/o relativas $r(t) \in \mathcal{R}(t) = \mathcal{R}_{abs}(t) \cup \mathcal{R}_{rel}(t)$ de manera que: $r(t) = \mathcal{A}(\vec{r}(t)) \subseteq R^d \mid \vec{r}(t) \in R^d$ siendo $\vec{r}(t)$ las *Funciones Vectoriales Características* y $\mathcal{A}(\vec{r}(t))$ los subespacios asociados.

La definición de $\mathcal{R}_{abs}(t)$ y $\mathcal{R}_{rel}(t)$ se realiza de forma directa a partir del caso espacial siendo las funciones características, en el caso discreto, las que se especifican en la Tabla 4.2 donde se hace efectiva la discretización en el tiempo. En este caso la invarianza a la escala temporal puede implementarse mediante normalización. Por otro lado el esquema propuesto puede extenderse para incorporar restricciones diferenciales ya sea de primer (velocidad), de segundo orden (aceleraciones) o mixtas.

4.2 Aprendizaje: Auto-Restricciones Finales

4.2.1 Aprendizaje de Auto-Restricciones

Definición 4.3 Dado un *Conjunto de Entrenamiento* $\mathcal{C}(t) = \{\mathcal{G}_i(t)\}_{i=1}^m$ formado por m instancias espaciales o espacio-temporales³, una *Auto-Restricción o Restricción Propia Final* $r^\zeta(t) = \mathcal{A}(\vec{r}(t), \zeta) \subseteq R^e \mid \vec{r}(t) \in R^d$, con *nivel de confianza* $\zeta \in [0, 1]$, es el subespacio de dimensión $0 < e \leq d$ expandido por la aplicación de ACP sobre la distribución $\mathcal{D}(\vec{r}(t), S_r(t))$ de manera que:

1. *Restricciones/Distribución:* a partir del conjunto de entrenamiento considerado extraemos m ejemplos $\vec{s}_{r_i}(t)$ de la restricción $r(t)$. A continuación computamos el vector promedio $\vec{r}(t)$ y la matriz de covarianza $S_r(t)$.
2. *Componentes/Modos de Variación:* los modos de variación de $r^\zeta(t)$ (es decir los modos principales de $r(t)$ para dicho nivel de confianza) vienen definidos por los e *autovectores principales* de $S_r(t)$, es decir $\vec{\Psi}_1, \vec{\Psi}_2, \dots, \vec{\Psi}_e$, siendo e el valor más pequeño para el cual se cumple $\sum_{k=1}^e \lambda_k \geq \zeta \lambda_T \equiv \sum_{k=e+1}^d \lambda_k \leq (1 - \zeta) \lambda_T$, λ_k los autovalores asociados y λ_T su suma o variabilidad total. En base a este análisis se establecen las componentes del subespacio $\mathcal{A}(\vec{r}(t), \zeta)$.

²En términos de implementación discretizaremos dicha ventana para hacer referencia a cada uno de los *frames* de la secuencia de imágenes.

³En adelante y con objeto de simplificar la formulación de definiciones y propiedades relativas a auto-restricciones, recurriremos a una notación dependiente del tiempo como caso más general.

76 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

3. *Representación*: en consecuencia cada una de las tuplas $\vec{s}_r(t)$ que satisfacen $r(t)$ puede expresarse, con un grado de confianza ζ , en el subespacio transformado generado por los autovectores principales, mediante el vector de coordenadas $b = P_e^T(\vec{s}_r(t) - \bar{s}_r(t))$, siendo $P_e = [\vec{\Psi}_1, \vec{\Psi}_2, \dots, \vec{\Psi}_e]$ y $b = (b_1, b_2, \dots, b_e)^T$.

El concepto de restricción aprendida o restricción propia final con su auto-espacio asociado se ilustra en la sección B.2 del Apéndice B donde pueden encontrarse ejemplos de aprendizaje de modelos radiales de objeto.

Definición 4.4 El conjunto de restricciones $\mathcal{R}(\vec{\zeta}, t) = \mathcal{R}_{abs}(\vec{\zeta}_{abs}, t) \cup \mathcal{R}_{rel}(\vec{\zeta}_{rel}, t)$ derivado del aprendizaje a partir de $\mathcal{C}(t)$, siendo $\vec{\zeta} = [\vec{\zeta}_{abs}, \vec{\zeta}_{rel}]^T$ el vector global de niveles de confianza, donde $\vec{\zeta}_{abs} = [\zeta_{xy}, \zeta_{ab}, \zeta_{\theta}, \zeta_{\epsilon}]^T$ y $\vec{\zeta}_{rel} = [\zeta_{\rho}, \zeta_{\lambda}, \zeta_{\alpha}, \zeta_{\kappa}]^T$ son los vectores asociados a restricciones absolutas y relativas (ya sean espaciales o espacio-temporales), se denomina *Conjunto de Restricciones Propias Finales*.

4.2.2 Entropía de Restricción

Definición 4.5 Dada una restricción $r(t) \in \mathcal{R}(t)$ definimos su *Entropía Normalizada* $Ent(r(t)) \in [0, 1]$ como la suma normalizada de los productos de la importancia relativa de cada autovector por su logaritmo:

$$Ent(r(t)) = -\frac{1}{\log_2 d} \sum_{k=1}^d \frac{\lambda_k}{\lambda_T} \log_2 \left(\frac{\lambda_k}{\lambda_T} \right) \quad (4.1)$$

Esta medida, como métrica de incertidumbre que es (Bender, 1996), nos permite estimar la distribución de la variabilidad asociada a los distintos modos, de manera que un valor alto indica que todos los modos son importantes y de igual manera, mientras que un valor bajo indica concentración de variabilidad en unos pocos modos, en cuyo caso se refuerza la utilidad del ACP para caracterizar, desde el punto de vista considerado, la estructura.

4.2.3 Genericidad de Restricción

Definición 4.6 Dada una restricción $r(t) \in \mathcal{R}(t)$ definimos su *Nivel de Genericidad* $Gen(r(t)) \in [0, \infty]$ como la variabilidad promedio de los modos:

$$Gen(r(t)) = \frac{\lambda_T}{e} \in [G_{Min}, G_{Max}] \quad (4.2)$$

El valor de esta medida debe conservarse dentro de unos umbrales máximo y mínimo aceptables ($[G_{Min}, G_{Max}]$) asociados a cada tipo de restricción. Un valor excesivamente bajo indica que el vector promedio es suficiente para describir la distribución, lo cual equivale a auto-espacios de dimensiones reducidas, es decir a restricciones cerradas o muy acotadas). Por contra un valor excesivamente alto indica un carácter muy poco específico y discriminante a efectos de delimitar qué configuraciones la satisfacen (restricciones muy abiertas y poco útiles). Tal como vimos en el capítulo anterior las restricciones de forma y orientación precisan de niveles de genericidad bajos con respecto al resto. De esta forma evitaremos configuraciones erróneas asociadas a mínimos locales.

Tabla 4.3: Dominio Espacial. Entropías y Grados de Genericidad

Propiedad	r_{xy}	r_{ab}	r_{θ}	r_{ϵ}	r_{ρ}	r_{λ}	r_{α}	r_{κ}
$Ent(r)$	0.274	0.255	0.221	0.244	0.155	0.350	0.292	0.211
$Gen(r)$	4.199	4.262	3.944	0.016	9.4×10^5	1.3×10^4	3.8×10^4	2.509

Tabla 4.4: Dominio Espacio-Temporal. Entropías y Grados de Genericidad

Propiedad	$r_{xy}(t)$	$r_{ab}(t)$	$r_{\theta}(t)$	$r_{\epsilon}(t)$	$r_{\rho}(t)$	$r_{\lambda}(t)$	$r_{\alpha}(t)$	$r_{\kappa}(t)$
$Ent(r(t))$	0.153	0.725	0.478	0.807	0.053	0.753	0.48	0.650
$Gen(r(t))$	22.97	0.984	10.40	0.010	4×10^6	15.76	11.6×10^5	0.001

4.2.4 Ejemplo en el Dominio Espacial

Tomando como referencia una escena como la de la parte superior de la Figura 4.1 con $g = 5$ hemos diseñado un conjunto de adiestramiento con $m = 20$ muestras y hemos extraído todas las restricciones posibles. De la aplicación del ACP derivamos un *Mapa de Restricciones* absolutas y otro de restricciones relativas (ver Figura 4.2). En dichos mapas se codifica, para cada restricción la efectividad de

78 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

cada autovalor ($\frac{\lambda}{\lambda_T}$) de manera que a mayor nivel de gris mayor efectividad⁴. El análisis de las entropías normalizadas nos revela que la variabilidad se concentra en unos pocos modos (en todos los casos se obtiene un valor por debajo del 40% (ver Tabla 4.3). Más concretamente se aprecia que en cuanto a *restricciones absolutas*, las principales variaciones se dan en: y_2 (posición), a_3 (escala), θ_1 (orientación) y en ϵ_4 (forma). Este resultado puede corroborarse analizando los auto-espacios tridimensionales representados en la Figura 4.4. En la primera columna se representan las coordenadas por orden de importancia para cada auto-espacio: cada uno de los puntos del conjunto de adiestramiento se proyecta sobre el espacio expandido por los tres modos indicadas entre paréntesis. En la segunda columna se representan las matrices P de modos de variación que expanden los subespacios completos. Por otro lado la mayor variabilidad en las restricciones relativas se da en las relaciones: 2 – 5 en posición, 1 – 3 en escala y orientación, y 4 – 5 en forma. En cuanto a la genericidad en la Tabla 4.3 puede observarse que la restricción de forma relativa es muy cerrada en comparación con el resto de restricciones relativas.

4.2.5 Ejemplo en el Dominio Espacio-Temporal

De manera análoga, y tomando como referencia la parte inferior de la Figura 4.1 hemos diseñado un conjunto de entrenamiento espacio-temporal con $T = 4$ (cinco etapas) y $m = 15$ muestras. En este caso el análisis de entropía (ver Tabla 4.4) derivado de los mapas de restricción que se muestran en la Figura 4.3 nos revela una mayor uniformidad en la importancia de los modos que en el ejemplo espacial, sobre todo en las restricciones absolutas (entropías altas, la mayoría por encima del 50%). El estudio de dichos mapas proporciona, como en el caso espacial, información sobre la importancia relativa de los modos de variación tanto en restricciones absolutas como en restricciones relativas. En las primeras los mejores modos son: $x_1(3)$ (posición), $a_1(1)$ (escala), $\theta_1(2)$ (ángulo) y $\epsilon_2(4)$ (forma). Los mejores modos temporales correspondientes a la relación 1 – 2 son los asociados a los instantes $t = 2$ en posición, $t = 1$ en escala, $t = 2$ en orientación y $t = 4$ en forma. Los auto-espacios de dimensión 3 asociados a las restricciones relativas se representan en la Figura 4.5. Finalmente los índices de genericidad de la Tabla 4.4 muestran que la variabilidad promedio tanto de la forma como de la escala relativas es muy baja a lo largo de la secuencia.

⁴La asociación de modos de variación a parámetros correspondientes a las restricciones es meramente simbólica y solo se aplicará a efectos de ilustración de la variabilidad de dichos parámetros (ya que cada autovalor representa una varianza correspondiente a un parámetro del auto-espacio). Formalmente *los modos caracterizan direcciones principales de variación en el auto-espacio y no en el espacio origen* (ver Apéndice B).

80 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

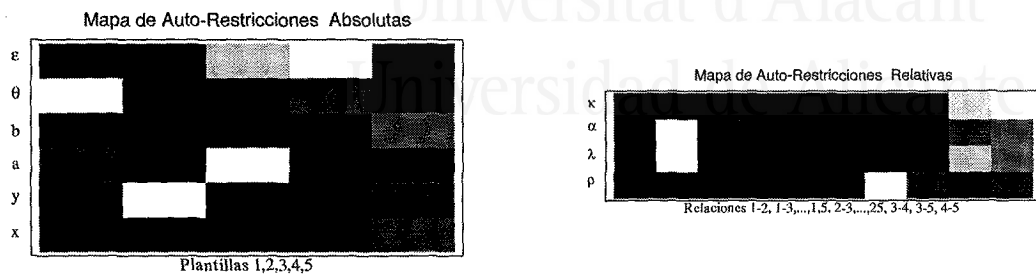


Figura 4.2: Mapas de Auto-Restricciones: efectividad de autovalores

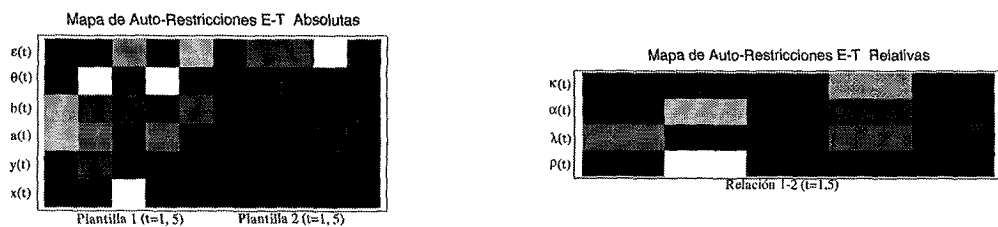


Figura 4.3: Mapas de Auto-Restricciones E-T: efectividad de autovalores

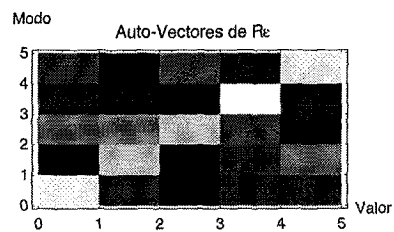
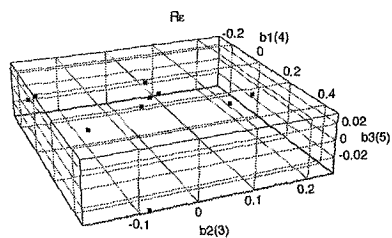
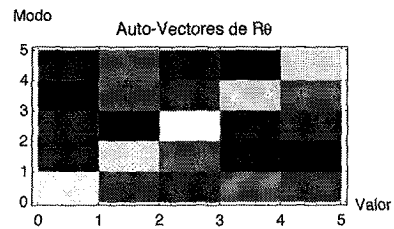
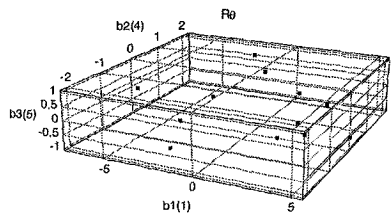
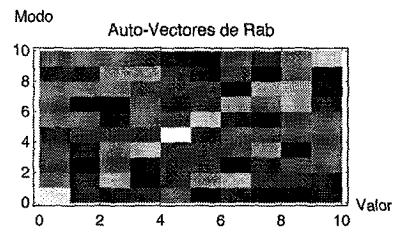
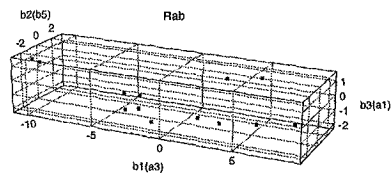
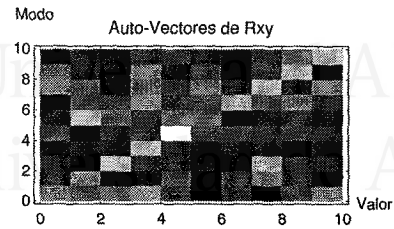
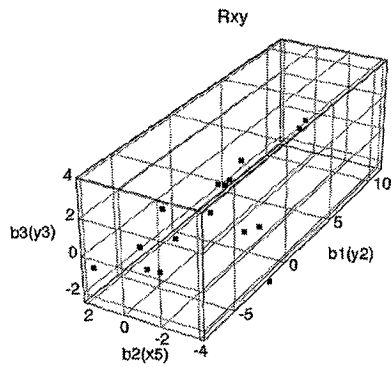


Figura 4.4: Restricciones absolutas espaciales: Auto-Espacios y matrices modales

82 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

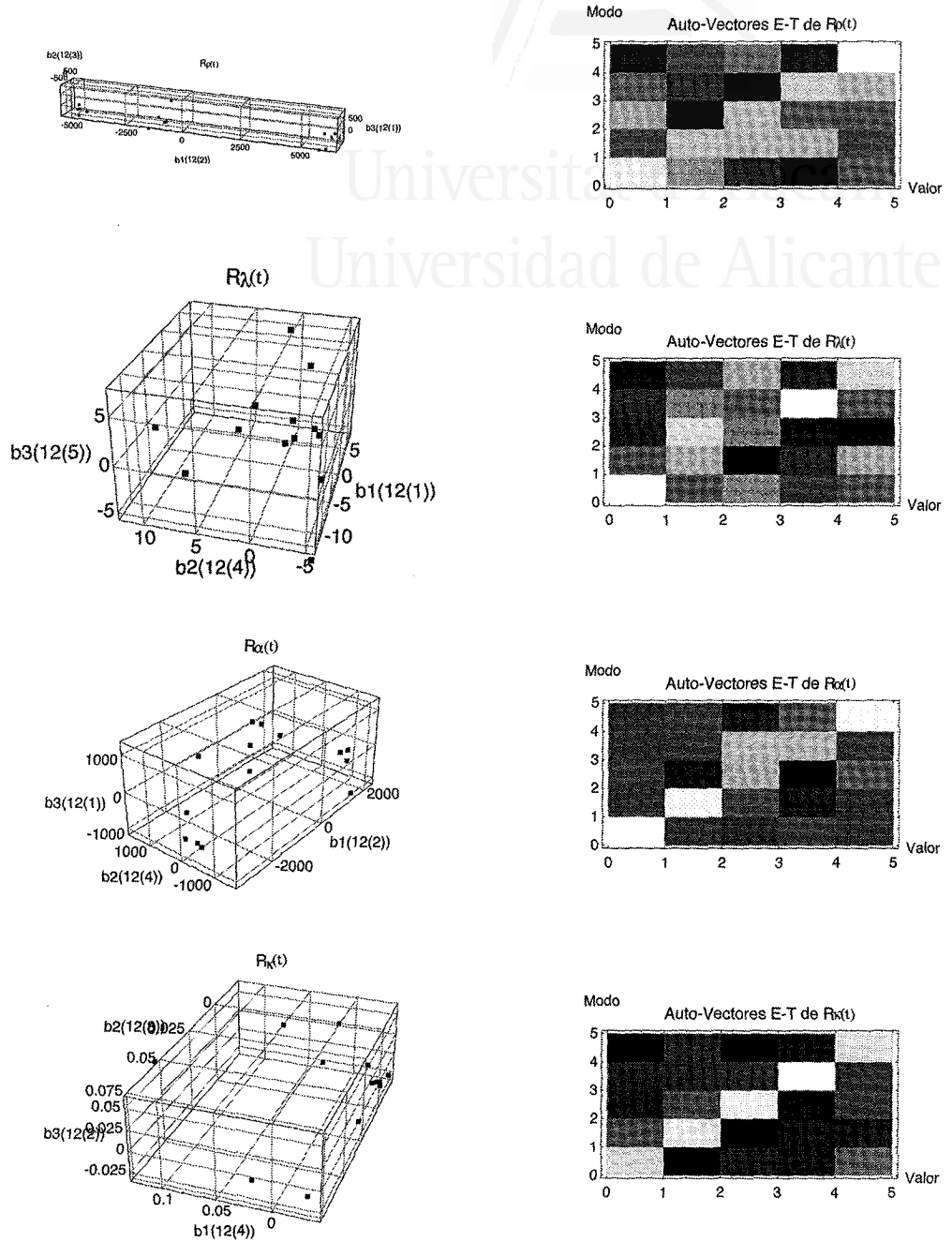


Figura 4.5: Restricciones E-T relativas: Auto-Espacios y matrices modales

4.3 Satisfacción de Auto-Restricciones

4.3.1 Grado de Satisfacción de Auto-Restricciones

Una vez especificadas las restricciones y el mecanismo de aprendizaje entramos en la definición de una métrica de pertenencia al modelo la cual permita guiar el comportamiento de las plantillas. Dado como entrada un conjunto $\mathcal{G}(t)$ de Plantillas Super-Elipsoydales, espaciales o dependientes del tiempo, de manera que no necesariamente pertenezca al conjunto de adiestramiento $\mathcal{C}(t)$:

Definición 4.7 Definimos el *Grado de Satisfacción Local* $Sat(\mathcal{G}(t), r^\zeta(t)) \in [0, 1]$ de una restricción $r(t)$, con nivel de confianza ζ , como:

$$Sat(\mathcal{G}(t), r^\zeta(t)) = 1 - Def(\mathcal{G}(t), r^\zeta(t)) \quad (4.3)$$

$$Def(\mathcal{G}(t), r^\zeta(t)) = \begin{cases} 1 & D_{Mah}(r^\zeta(t), b) > D_{def} \\ \frac{D_{Mah}(r^\zeta(t), b)}{D_{def}} & D_{Mah}(r^\zeta(t), b) \leq D_{def} \end{cases} \quad (4.4)$$

$$D_{Mah}(r^\zeta(t), b) = \sum_{k=1}^e \left(\frac{b_k^2}{\lambda_k} \right) \quad (4.5)$$

donde D_{Mah} es la distancia de Mahalanobis de la proyección b de los parámetros de $r(t)$ extraídos de $\mathcal{G}(t)$, sobre el espacio definido por $r^\zeta(t)$, D_{def} es una constante de normalización, y $Def(\mathcal{G}(t), r^\zeta(t)) \in [0, 1]$ es el *Grado de Deficiencia Local*⁵ de $r^\zeta(t)$ con respecto a \mathcal{G} .

Definición 4.8 Por otro lado, a partir de la definición anterior definimos el *Grado de Satisfacción Global* $Sat(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta})) \in [0, 1]$, con el vector de niveles de confianza $\vec{\zeta}$, como:

$$Sat(\mathcal{G}, \mathcal{R}(t, \vec{\zeta})) = 1 - Def(\mathcal{G}, \mathcal{R}(t, \vec{\zeta})) \quad (4.6)$$

$$Def(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta})) = \begin{cases} 1 & D_{Global}(B, \mathcal{R}(t, \vec{\zeta})) > D_{Def} \\ D_{Global}(B, \mathcal{R}(t, \vec{\zeta})) & D_{Global}(B, \mathcal{R}(t, \vec{\zeta})) \leq D_{Def} \end{cases} \quad (4.7)$$

$$D_{Global}(B, \mathcal{R}(t, \vec{\zeta})) = \frac{\sum_{r \in \mathcal{R}_{abs}(t) \cup \mathcal{R}_{rel}(t)} Def(\mathcal{G}(t), r^\zeta(t))}{Card(\mathcal{R}(t))} \quad (4.8)$$

⁵En definitiva este valor es la distancia de Mahalanobis normalizada por el umbral de rechazo local D_{def} . Este umbral se establece, de acuerdo con el criterio estadístico, a tres desviaciones típicas para cada dimensión (ver Apéndice B).

84 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

siendo B la matriz de coordenadas en los auto-espacios considerados, $Card(\mathcal{R}(t))$ el número de restricciones consideradas y $Def(\mathcal{G}, \mathcal{R}^\zeta) \in [0, 1]$ el *Grado de Deficiencia Global*⁶

4.3.2 Relajación de Restricciones y Nivel de Confianza

Definición 4.9 Dada una restricción espacial o espacio-temporal $r(t)$ de dimensión d considerada a dos niveles de confianza $\zeta_1 > \zeta_2 > 0$, de manera que $d \geq e_1 > e_2$, y siendo $\mathcal{A}_1 = \mathcal{A}(\bar{r}(t), \zeta_1) \subseteq \mathbb{R}^{e_1}$ y $\mathcal{A}_2 = \mathcal{A}(\bar{r}(t), \zeta_2) \subseteq \mathbb{R}^{e_2}$ decimos que $r^{\zeta_2}(t)$ es una *versión relajada* de $r^{\zeta_1}(t)$ ya que:

$$\forall \vec{s}_r(t) \in \mathcal{A}_1 \mid Sat(\vec{s}_r(t), r^{\zeta_1}(t)) > 0$$

se cumple que

$$Sat(Proy(\vec{s}_r(t), \mathcal{A}_2), r^{\zeta_2}(t)) > 0 \quad (4.9)$$

y además

$$Sat(Proy(\vec{s}_r(t), \mathcal{A}_2), r^{\zeta_2}(t)) \geq Sat(\vec{s}_r(t), r^{\zeta_1}(t)) \quad (4.10)$$

siendo $Proy(\vec{s}_r(t), \mathcal{A}_2)$ la proyección de $\vec{s}_r(t)$ sobre el espacio \mathcal{A}_2 .

Esta propiedad puede demostrarse de la siguiente forma. Sea $b = (b_1, b_2, \dots, b_{e_1})^T$ el vector de coordenadas de $\vec{s}_r(t)$ en el auto-espacio expandido por $r^{\zeta_1}(t)$, de manera que los índices de sus componentes *siguen un orden creciente* en relación al modo de variación considerado. Por tanto b puede escribirse de la siguiente forma:

$$b = \underbrace{(b_1, b_2, \dots, b_{e_2}, \dots, b_{e_1})^T}_{b'}$$

siendo b' el vector de coordenadas de $Proy(\vec{s}_r(t), \mathcal{A}_2)$ en el auto-espacio expandido por $r^{\zeta_2}(t)$. Entonces

$$Sat(\vec{s}_r(t), r^{\zeta_1}(t)) > 0 \equiv D_{Mah}(r_1^{\zeta_1}(t), b) \leq D_{def}$$

En consecuencia, ya que los autovalores de la matriz de covarianza son reales y no-negativos (representan varianzas), y teniendo en cuenta que si alguno fuese cero no se habría considerado en ninguno de los auto-espacios, se cumple

$$\sum_{k=1}^{e_1} \left(\frac{b_k^2}{\lambda_k} \right) \equiv \underbrace{\frac{b_1^2}{\lambda_1} + \frac{b_2^2}{\lambda_2} + \dots + \frac{b_{e_2}^2}{\lambda_{e_2}} + \dots + \frac{b_{e_1}^2}{\lambda_{e_1}}}_{D_{Mah}(r_2^{\zeta_2}(t), b')} \leq D_{def}$$

⁶En el caso de la distancia global D_{Global} asumiremos un umbral $D_{Def} = 0.5$. Este índice puede hacerse más restrictivo (considerando p.e. criterios de mínimo que obligen a satisfacer en cierto grado todas las restricciones) o bien pueden establecerse parámetros de ponderación si se estima oportuno a nivel de modelización.

de lo cual se desprende que $D_{Mah}(r_2^\zeta(t), b') \leq D_{def}$ y además se cumple, tal como hemos visto, $D_{Mah}(r_2^\zeta(t), b') \leq D_{Mah}(r_2^\zeta(t), b)$ con lo que queda demostrado el teorema.

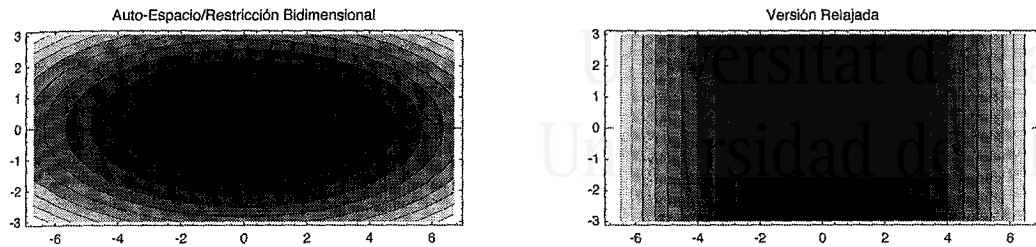


Figura 4.6: Relajación de restricciones. Caso bidimensional

El nivel de confianza asociado a una restricción actúa formalmente como *coeficiente de relajación* en la medida en que si la consideración de niveles de confianza menores implica reducir el número de modos considerados ello se traduce en un incremento del conjunto de vectores que satisfacen la restricción en grado no-nulo (es decir que quedan dentro de los límites estadísticos). En la parte izquierda de la Figura 4.6 se muestra el valor de la distancia de Mahalanobis asociado a cada uno de los puntos de un espacio bidimensional con $\lambda_1 = 5$ y $\lambda_2 = 1$. En la parte derecha se muestra ese valor para una versión relajada en la cual se ha eliminado la segunda componente. Por aplicación del teorema previo todo punto aceptado en el espacio origen se sigue aceptando en el nuevo espacio incluyéndose en este otros que antes no eran aceptados.

4.3.3 Ejemplo de Reconocimiento: Dominio Espacial

Tomando como referencia los resultados obtenidos en la sección 4.2.4. supongamos que consideramos, a efectos de aplicación para el reconocimiento, todas las restricciones absolutas y que $\mathcal{R}_{rel} = \emptyset$. Para estudiar *el efecto del nivel de confianza en el reconocimiento* partimos de una estructura en la cual se presentan desviaciones con respecto a la media en cada una de las restricciones (contrastar los mapas de restricciones espaciales con las gráficas de la Figura 4.7, en donde la línea discontinua indica vector promedio y la continua representa los parámetros de entrada para el reconocimiento). En la Tabla 4.5 se muestran sucesivos grados de aceptación para distintos niveles refinamiento (relajación del nivel de confianza) observándose el aumento del grado de satisfacción global a medida que disminuyen los niveles de confianza (Casos 1-3). En el Caso 4 se muestra el resultado para un vector de entrada cercano al promedio en todas las restricciones. En la Tabla 4.6

86 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

se muestran los niveles de confianza aplicados y los modos de variación utilizados en cada restricción. El descenso en nivel de confianza viene acompañado de una caída monótona del número de modos.

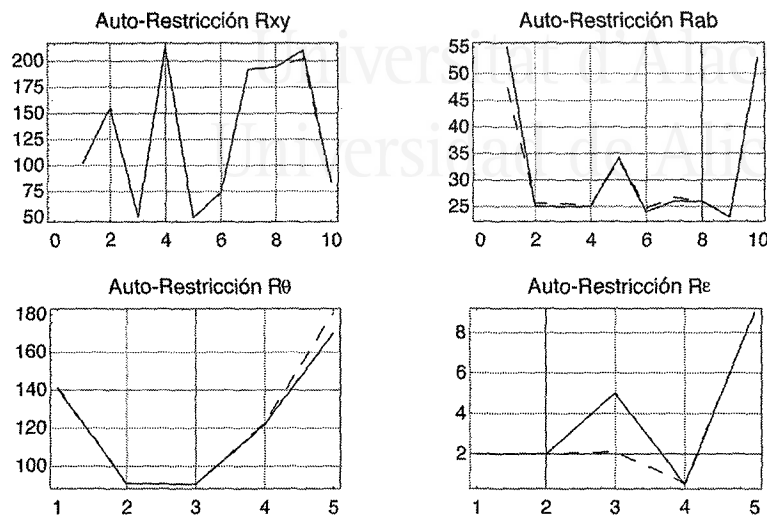


Figura 4.7: Reconocimiento en el dominio espacial

Tabla 4.5: Reconocimiento Espacial con Restricciones Absolutas

Pruebas	$Sat r_{xy}$	$Sat r_{ab}$	$Sat r_{\theta}$	$Sat r_e$	$Sat(\mathcal{R}, \vec{\zeta})$
Caso 1	0.000	0.000	0.056	0.000	0.014
Caso 2	0.203	0.977	0.964	0.000	0.536
Caso 3	0.920	0.991	0.964	0.754	0.907
Caso 4	0.995	0.999	0.984	0.915	0.973

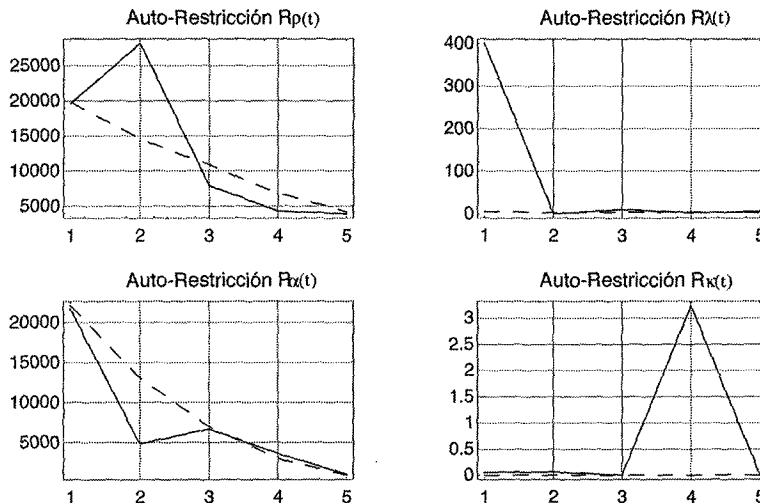
4.3.4 Ejemplo de Reconocimiento: Dominio Espacio-Temporal

En este caso partimos de los resultados obtenidos en la sección 4.2.5 y consideramos todas las restricciones relativas (movimiento relativo) y ninguna absoluta, es

Tabla 4.6: Niveles de Confianza y Modos Espaciales

Pruebas	ζ_{xy}	Modos	ζ_{ab}	Modos	ζ_{θ}	Modos	ζ_{ϵ}	Modos
Caso 1	0.95	3	0.97	3	0.97	2	0.99	2
Caso 2	0.90	2	0.90	2	0.90	1	0.80	2
Caso 3	0.80	1	0.80	1	0.90	1	0.70	1
Caso 4	0.95	3	0.97	3	0.97	2	0.99	2

decir $\mathcal{R}_{abs} = \emptyset$. Si observamos los vectores promedio asociados a cada restricción relativa (Figura 4.8) vemos que la relación entre ambas plantillas se caracteriza por trayectorias decrecientes en el tiempo en cuanto a distancia y ángulo relativo, mientras que se observa la igualdad de forma y escala a lo largo de la secuencia. En este caso las desviaciones de los valores de entrada para cada restricción son mucho mayores, en general, a los límites estadísticos permitidos. Por lo tanto, y tal como se refleja en la Tabla 4.7, la relajación de las restricciones no permitirá reconocer finalmente la entrada incluso cuando se pasa a considerar un único modo en la mayoría de los casos (ver Tabla 4.8).

**Figura 4.8:** Reconocimiento en el dominio Espacio-Temporal

88 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES

Tabla 4.7: Reconocimiento Espacio-Temporal con Restricciones Relativas

Pruebas	$Sat r_\rho(t)$	$Sat r_\lambda(t)$	$Sat r_\alpha(t)$	$Sat r_\kappa(t)$	$Sat \mathcal{R}(t, \vec{\zeta})$
Caso 1	0.815	0.000	0.000	0.000	0.204
Caso 2	0.815	0.000	0.000	0.000	0.204
Caso 3	0.815	0.000	0.000	0.000	0.204
Caso 4	0.999	0.917	0.999	0.917	0.958

Tabla 4.8: Niveles de Confianza y Modos Espacio-Temporales

Pruebas	ζ_ρ	Modos	ζ_λ	Modos	ζ_α	Modos	ζ_κ	Modos
Caso 1	0.90	1	0.90	3	0.90	2	0.90	3
Caso 2	0.80	1	0.70	3	0.70	1	0.70	2
Caso 3	0.70	1	0.50	2	0.50	1	0.50	1
Caso 4	0.90	1	0.90	3	0.90	2	0.90	3

4.4 Eficiencia de las Auto-Representaciones

La caracterización basada en ACP aporta, dependiendo de la naturaleza del conjunto de entrenamiento, una *representación eficiente* de la estructura geométrica (espacial o espacio-temporal) subyacente. La eficiencia se estima tomando como referencia el número g de plantillas del modelo y se estudia a dos niveles:

1. *Complejidad Espacial:* en última instancia el mecanismo de *tracking* basado en correspondencia dispone de recursos de almacenamiento limitado. La codificación mediante ACP permite representar el conjunto de tuplas que satisfacen una restricción $r^\zeta(t)$ mediante un vector característico promedio de dimensión d y e vectores modales de la misma dimensión de manera que:

(a) Si $r^\zeta(t) \in \mathcal{R}_{abs}$ entonces $d = 6g(T + 1)$. En caso contrario $d = \binom{g}{2}(T + 1) = \frac{g!}{2(g-2)!}(T + 1)$.

- (b) El número e de vectores modales depende de la relación $\sum_{k=1}^e \lambda_k \geq \zeta \lambda_T$, y por lo tanto de la entropía de restricción. No obstante el rendimiento a nivel de eficiencia de codificación tiende a ser mayor cuando aumenta T .

2. *Complejidad Temporal*: por otro lado la eficiencia de la representación impacta en el coste temporal asociado al reconocimiento. El coste de aplicación de una restricción es la suma de:

- (a) Coste de extracción de los parámetros de la restricción a partir de $\mathcal{G}(t)$. Dicho coste es cuadrático con relación al número de plantillas para restricciones relativas y lineal para absolutas.
- (b) Coste de proyección de la restricción en el auto-espacio (cálculo de $b = P_e^T(\bar{s}_r(t) - \bar{s}_r(t))$)
- (c) Coste de cálculo de la distancia de Mahalanobis.

Si podemos representar un alto porcentaje de la variabilidad con unos pocos modos de variación (entropía baja) los dos últimos costes descienden. Además, dado que los parámetros relativos a restricciones de distinto tipo son independientes, todo el proceso puede realizarse en paralelo para tipos distintos y en secuencia dentro de cada uno de ellos (absoluto y relativo).

4.5 Consideraciones

El modelo propuesto en este capítulo establece un marco conceptual para la caracterización de estructuras geométricas deformables. Una vez definidos los parámetros característicos disponemos de un mecanismo formal que nos permite diseñar y extraer auto-espacios de forma automática. La métrica de satisfacción de los condicionantes del modelo se construye integrando el grado de pertenencia a cada uno de los espacios definidos por las restricciones. Debe tenerse en cuenta que el nivel de confianza considerado actúa como coeficiente de relajación.

En términos generales, y a nivel del proceso completo de *tracking* y reconocimiento, debemos combinar los posicionamientos debidos a la aplicación de los campos de potencial con los inducidos por la satisfacción de las restricciones definidas. En el capítulo siguiente formularemos el criterio de optimización global y entraremos en estas consideraciones con objeto de dar soluciones eficientes al posicionamiento, en el espacio y en el tiempo, de múltiples plantillas.

90 CAPÍTULO 4. CARACTERIZACIÓN ESTRUCTURAL: AUTO-RESTRICCIONES



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Capítulo 5

Tracking Basado en Plantillas Extendidas

Una vez propuestas tanto la componente local del modelo como su caracterización estructural espacial y espacio-temporal, dedicamos este capítulo a definir el *mecanismo de extracción/reconocimiento espacial* y su extensión para abordar el problema del *tracking basado en modelos espacio-temporales*. Definimos en primer lugar el *Problema de la Correspondencia Espacial* basada en Plantillas Extendidas en términos de la minimización de una función de energía global sujeta a las auto-restricciones del modelo. A continuación aportamos el *Esquema/Algoritmo de Descenso Supervisado* que permite abordar el problema integrando minimización sobre las componentes local y estructural (esta última mediante proyección en auto-espacios). Este esquema se extiende finalmente para el tratamiento del *Tracking Espacio-temporal*. En ambos casos aportamos resultados experimentales de adaptación partiendo de los modelos expuestos en el capítulo anterior.

5.1 Mecanismo de Correspondencia Espacial

5.1.1 Directrices de Modelización

El ajuste de un conjunto de plantillas (estructura geométrica) a un modelo de imagen dado, atendiendo a las auto-restricciones aprendidas, se aborda desde dos consideraciones fundamentales:

1. *Supervisión del Comportamiento Local*: el conocimiento estructural supervisa el comportamiento local de las plantillas de manera que la solución deberá satisfacer, en la medida de lo posible, los condicionantes estructurales establecidos.

2. *Correspondencia Parcial en Auto-Espacios*: parte del proceso de correspondencia se realiza en un espacio de baja dimensión en función de los datos de aprendizaje.

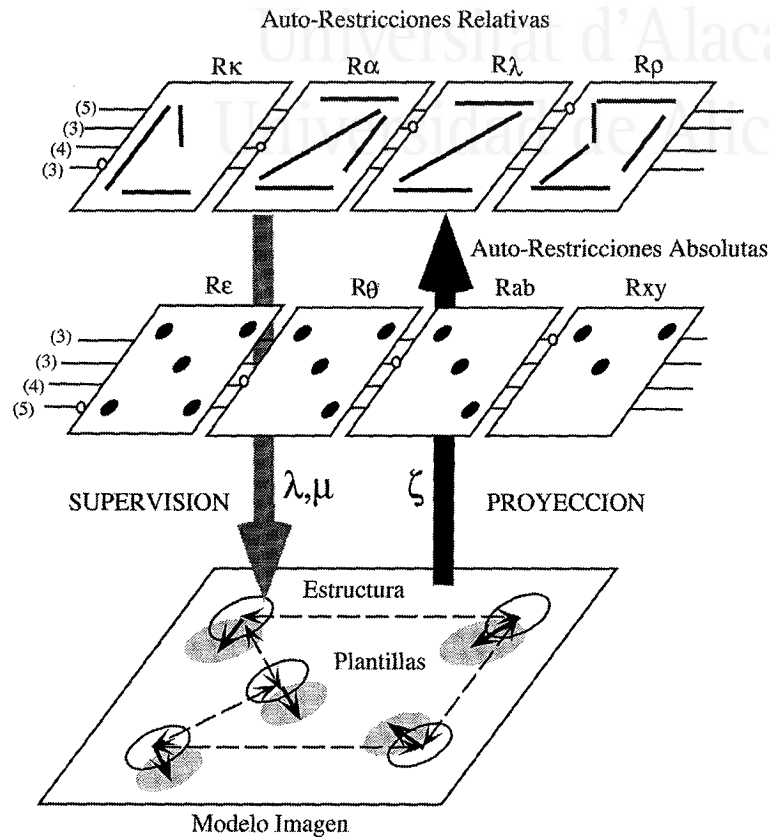


Figura 5.1: Esquema General-Bucle de Retroalimentación

Seguidamente describiremos de forma independiente el comportamiento local de un conjunto de plantillas (*Minimización de Energía Local*) y su regularización estructural (*Maximización del Grado de Satisfacción*). La integración de ambos aspectos da lugar al *Esquema de Descenso Supervisado* la técnica aportada para resolver el problema de la correspondencia espacial.

Dicho esquema implementa un bucle de retro-alimentación entre el espacio de imagen y los auto-espacios paramétricos aprendidos. En términos generales (ver Figura 5.1) una configuración estructural dada se proyecta sobre los auto-espacios

definidos (teniendo en cuenta los niveles de confianza adoptados). En cada auto-espacio representamos de forma simbólica los modos de variación activos (puntos en las restricciones absolutas y por líneas o enlaces en las relativas). En dichos espacios se realiza la parte del proceso de minimización relativa a la eliminación de posibles ambigüedades locales¹. El mecanismo de supervisión sugiere modificaciones que deberán ser ponderadas para obtener el resultado final.

5.1.2 Minimización de Energía Local

Definición 5.1 Dado un conjunto $\mathcal{G} = \{g_{\mathcal{D}_i}(\Theta)\}_{i=1}^g$ de *Plantillas Deformables Extendidas* definimos como *Energía Local Conjunta Local*(\mathcal{G}) la expresión:

$$Local(\mathcal{G}) = \sum_{i=1}^g \mathcal{E}_i(\vec{\Theta}_i) \quad (5.1)$$

siendo $\mathcal{E}_i(\vec{\Theta}_i)$ las energías correspondientes a las plantillas $g_i \in \mathcal{G}$ especificadas respectivamente por los vectores paramétricos $\vec{\Theta}_i$.

Definición 5.2 La derivación en el tiempo de $Local(\mathcal{G})$ para obtener una configuración óptima se realiza, dada la independencia entre los parámetros de cada plantilla, derivando simultáneamente cada una de las energías $\mathcal{E}_i(\vec{\Theta}_i)$. Teniendo en cuenta que:

$$\mathcal{G} \equiv \vec{\Theta} = [\vec{\Theta}_1, \vec{\Theta}_2, \dots, \vec{\Theta}_g]^T$$

definimos la dinámica local $\frac{d\mathcal{G}_{Local}}{dt}$ de las plantillas:

$$\frac{d\mathcal{G}_{Local}}{dt} = \frac{d\vec{\Theta}}{dt} = -W_L \frac{\partial Local(\mathcal{G})}{\partial \vec{\Theta}} = -W_L \left[\frac{\partial \mathcal{E}_1(\vec{\Theta}_1)}{\partial \vec{\Theta}_1}, \frac{\partial \mathcal{E}_2(\vec{\Theta}_2)}{\partial \vec{\Theta}_2}, \dots, \frac{\partial \mathcal{E}_g(\vec{\Theta}_g)}{\partial \vec{\Theta}_g} \right]^T$$

siendo W_L una matriz diagonal de pesos de dimensión $6g$. Este planteamiento inicial no expresa relaciones estructurales entre las plantillas definidas. Implícitamente estamos asumiendo, como punto de partida, que $\mathcal{R} = \emptyset$, con lo cual la solución final está exclusivamente influenciada por la inicialización de las plantillas y por el modelo de imagen. Como veremos a continuación, la consideración de auto-restricciones tiene como función regularizar el resultado local.

¹Dada la independencia de los parámetros de restricciones de tipos distintos dicho proceso es paralelizable.

5.1.3 Maximización del Grado de Satisfacción

Partiendo de la definición de $Sat(\mathcal{G}, \mathcal{R}(\vec{\zeta}))$ su maximización puede abordarse teniendo en cuenta que $Max\ Sat(\mathcal{G}, \mathcal{R}(\vec{\zeta})) \equiv Min\ Def(\mathcal{G}, \mathcal{R}(\vec{\zeta}))$. Es decir, la obtención de la configuración estructural óptima independientemente del modelo de imagen está ligada a *minimizar, para cada restricción, la distancia de Mahalanobis al origen del auto-espacio o configuración promedio*.

$$Min\ D_{Mah}(r^\zeta, \vec{b}) \quad \forall \vec{b} \in B, \forall r^\zeta \in \mathcal{R}(\vec{\zeta})$$

Por lo tanto las dinámicas de adaptación de la estructura en función de este criterio afectarán a los parámetros de \mathcal{G} contemplados en cada restricción, serán inducidas por la satisfacción incremental de dicha restricción y dependerán tanto del tipo de restricción (absoluta o relativa) como del nivel de confianza considerado.

5.1.4 Dinámica de Restricciones Absolutas

Definición 5.3 Dada una restricción $r^\zeta \in \mathcal{R}_{abs}$ y siendo $\vec{\Theta}_r = (\Theta_{r1}, \Theta_{r2}, \dots, \Theta_{r|r|})^T$ el vector de parámetros de \mathcal{G} , con $|r| \in \{g, 2g\}$, correspondientes a la restricción r (en este caso $\vec{\Theta}_r \equiv \vec{r}$), su ajuste $\hat{\Theta}_r$ en cada paso viene dado por la expresión:

$$\hat{\Theta}_r = \vec{\Theta}_r + P_e \widehat{Proy}(\vec{\Theta}_r, r^\zeta)$$

$$\widehat{Proy}(\vec{\Theta}_r, r^\zeta) = \vec{b} - W_{A(r)} \frac{\partial D_{Mah}(r^\zeta, \vec{b})}{\partial \vec{b}}$$

siendo $W_{A(r)}$ una matriz diagonal de pesos cuya dimensión, e , es la inducida por el nivel de confianza ζ , y $\vec{b} = P_e^T (\vec{\Theta}_r - \vec{r}) \equiv Proj(\vec{\Theta}_r, r^\zeta)$ la proyección en el auto-espacio. En consecuencia:

$$\frac{\partial D_{Mah}(r^\zeta, \vec{b})}{\partial \vec{b}} \equiv \left[\frac{2b_1}{\lambda_1}, \frac{2b_2}{\lambda_2}, \dots, \frac{2b_e}{\lambda_e} \right]^T$$

Esta dinámica induce una *aproximación gradual a la configuración promedio de la restricción*, lo cual deberá contrastarse e integrarse con las estimaciones locales.

5.1.5 Dinámica de Restricciones Relativas

Definición 5.4 Dada una restricción $r^\zeta \in \mathcal{R}_{rel}$ el mecanismo de ajuste se basa en el hecho de que dichas restricciones son vectores cuyas componentes se definen, a su vez, a partir de dos o más parámetros de \mathcal{G} incluidos en la restricción absoluta correspondiente. En general podemos expresar el vector de parámetros \vec{r} en términos de:

$$\vec{r} \equiv \varphi(\vec{\Theta}_r) = (\varphi_1(\vec{\Theta}_r), \varphi_2(\vec{\Theta}_r), \dots, \varphi_{|r|}(\vec{\Theta}_r))^T$$

siendo $|r| = \binom{g}{2}$ y $\varphi_1(\vec{\Theta}_r), \varphi_2(\vec{\Theta}_r), \dots, \varphi_{|r|}(\vec{\Theta}_r)$ las funciones² sobre las que se implementa r . Teniendo en cuenta que nuestro objetivo, a este nivel, se concreta en satisfacer:

$$\text{Min} (\varphi_i(\vec{\Theta}_r) - \bar{\varphi}_i(\vec{\Theta}_r))^2 \quad \forall i \in \{1, \dots, |r|\}$$

es decir en obtener aproximaciones aceptables al vector promedio representante de la restricción ($\bar{r} \equiv \bar{\varphi}(\vec{\Theta}_r)$). En consecuencia el ajuste, en cada iteración, de los parámetros de \mathcal{G} afectados por la restricción, puede realizarse en dos pasos:

1. *Proyección en Auto-Espacio/Extrapolación:* en primer lugar *extrapolamos la nueva configuración* de la auto-restricción mediante su proyección en el auto-espacio, obteniendo:

$$\hat{\varphi}(\vec{\Theta}_r) = \bar{\varphi}(\vec{\Theta}_r) + P_e \widehat{Proy}(\varphi(\vec{\Theta}_r), r^\zeta)$$

$$\widehat{Proy}(\varphi(\vec{\Theta}_r), r^\zeta) = \vec{b} - W_{R(r)} \frac{\partial D_{Mah}(r^\zeta, \vec{b})}{\partial \vec{b}}$$

donde aplicamos el mismo razonamiento que el caso de las restricciones absolutas siendo $W_{R(r)}$ la matriz de pesos correspondiente y $\vec{b} = P_e^T(\varphi(\vec{\Theta}_r) - \bar{\varphi}(\vec{\Theta}_r)) \equiv Proj(\varphi(\vec{\Theta}_r), r^\zeta)$ la proyección en el auto-espacio.

2. *Actualización de la Estructura:* una vez realizada la aproximación a la media evaluamos el impacto de este desplazamiento en los parámetros de \mathcal{G} para computar el ajuste de los parámetros de la estructura $\hat{\Theta}_r$

$$\hat{\Theta}_r = \Theta_r - W_{J(r)} J((\varphi_i(\vec{\Theta}_r) - \bar{\varphi}_i(\vec{\Theta}_r))^2, \vec{\Theta})^T (\hat{\varphi}(\vec{\Theta}_r) - \bar{\varphi}(\vec{\Theta}_r))^2$$

siendo $W_{J(r)}$ una matriz diagonal de pesos cuya dimensión es equivalente al número de parámetros de \mathcal{G} afectados por la restricción y $J((\varphi_i(\vec{\Theta}_r) - \bar{\varphi}_i(\vec{\Theta}_r))^2, \vec{\Theta}_r)$ la matriz de derivadas parciales de $(\varphi_i(\vec{\Theta}_r) - \bar{\varphi}_i(\vec{\Theta}_r))^2$ con respecto a $\vec{\Theta}$.

El esquema de actualización de los parámetros de las plantillas de \mathcal{G} con respecto a la satisfacción de una auto-restricción dada, ya sea absoluta o relativa, *se desarrolla parcialmente en un espacio de baja dimensión en función del aprendizaje realizado.*

²Por ejemplo en el caso de la restricción relativa de distancia $\varphi_1(\vec{\Theta}_r) = \rho_{12} = \|(x_1, y_1) - (x_2, y_2)\|^2$.

Con este esquema de actualización, cada restricción afecta a un subconjunto de los parámetros de \mathcal{G} . A continuación ampliamos el contexto de esta definición al tratamiento global de \mathcal{G} en función de todas las restricciones, absolutas y relativas, definidas.

Definición 5.5 Dado un conjunto de restricciones $\mathcal{R}_s(\vec{\zeta}_t)$ con $s, t \in \{abs, rel\}$ definimos la dinámica de \mathcal{G} inducida por la optimización del grado de satisfacción de todas las restricciones definidas en \mathcal{R}_s , con los niveles de confianza especificados en ζ_t , en términos de

$$\hat{\mathcal{G}}_{Def}(\mathcal{R}_s(\vec{\zeta}_t)) = \bigcup_{r \in \mathcal{R}_s} \hat{\Theta}_r$$

En esta definición explotamos la independencia de los parámetros de las restricciones de un tipo con respecto a los asociados con restricciones de tipos diferentes.

5.1.6 Estructuras Geométricas Óptimas

Definición 5.6 Dados un conjunto $\mathcal{G} = \{g_{\mathcal{D}_i}(\Theta)\}_{i=1}^q$ de Plantillas Super-Elipsoydales y su modelo $\mathcal{R}(\vec{\zeta}) = \mathcal{R}_{abs}(\vec{\zeta}_{abs}) \cup \mathcal{R}_{rel}(\vec{\zeta}_{rel})$ de Restricciones Propias Finales y siendo $\lambda \in [0, 1]$ definimos como *Estructura Geométrica Óptima* \mathcal{G}^* el conjunto de Plantillas Super-Elipsoydales que satisface:

$$\mathcal{G}^* = Arg Min F(\mathcal{G}, \mathcal{R}(\vec{\zeta}))$$

$$F(\mathcal{G}, \mathcal{R}(\vec{\zeta})) = (1 - \lambda)Local(\mathcal{G}) + \lambda Def(\mathcal{G}, \mathcal{R}(\vec{\zeta})) \quad (5.2)$$

donde $Local(\mathcal{G})$ se define en la Ecuación 5.1 y $Def(\mathcal{G}, \mathcal{R}(\vec{\zeta}))$ en la Ecuación 4.4. Es decir, de todas las configuraciones estructurales posibles, la óptima es aquella colección de plantillas que compatibiliza el ajuste ponderado al modelo de imagen con la satisfacción al máximo de las auto-restricciones especificadas (minimización de la deficiencia). La extracción automática de la *Estructura Óptima* viene caracterizada por dos aspectos básicos:

1. *Minimización Ponderada*: el parámetro λ permite modular la importancia del ajuste satisfactorio al modelo de imagen y la del ajuste a los condicionantes estructurales previamente adquiridos.
2. *Relajación de Auto-Restricciones*: tal como vimos en el capítulo anterior, los niveles de confianza considerados influyen en la calidad de la solución y en su coste computacional.

Tabla 5.1: Resolución de Correspondencia Espacial

Estructura Geométrica Óptima	
Criterio/Objetivo Heurística	$F(\mathcal{G}, \mathcal{R}(\vec{\zeta})) = (1 - \lambda)Local(\mathcal{G}) + \lambda Def(\mathcal{G}, \mathcal{R}(\vec{\zeta}))$
Criterio Local/Imagen	$Min Local(\mathcal{G}) \rightarrow Min \mathcal{E}_i(\vec{\Theta}_i) \forall i \in \{1, \dots, g\}$
Criterio Estructural	$Min Def(\mathcal{G}, \mathcal{R}(\vec{\zeta})) \rightarrow Min D_{Mah}(r^\zeta, \vec{b}) \forall \vec{b} \in B, \forall r^\zeta \in \mathcal{R}(\vec{\zeta})$
Dinámicas de Adaptación	
Dinámica Local	$\frac{d\mathcal{G}_{Local}}{dt} = \frac{d\vec{\Theta}}{dt}$
Dinámica Absoluta	$\hat{\Theta}_r = \vec{\Theta}_r + P_e \widehat{Proy}(\vec{\Theta}_r, r^\zeta)$
Dinámica Relativa	$\hat{\varphi}(\vec{\Theta}_r) = \vec{\varphi}(\vec{\Theta}_r) + P_e \widehat{Proy}(\varphi(\vec{\Theta}_r), r^\zeta)$
	$\hat{\Theta}_r = \Theta_r - W_{J(r)} J((\varphi_i(\vec{\Theta}_r) - \vec{\varphi}_i(\vec{\Theta}_r)^2, \vec{\Theta}))^T (\hat{\varphi}(\vec{\Theta}_r) - \vec{\varphi}(\vec{\Theta}_r))^2$
Integración	$\hat{\mathcal{G}}_{Def}(\mathcal{R}_s(\vec{\zeta}_t)) = \bigcup_{r \in \mathcal{R}_s} \hat{\Theta}_r$
Esquema de Descenso Supervisado	
Integración	$\mathcal{G}^{i+1} = (1 - \lambda)\hat{\mathcal{G}}_{Local}^{i+1} + \lambda\hat{\mathcal{G}}_{Def}^{i+1}$
Descenso No-Restringido	$\hat{\mathcal{G}}_{Local}^{i+1} = \mathcal{G}^i + \frac{d\mathcal{G}_{Local}^i}{dt}$
Ajuste/Auto- Restricciones	$\hat{\mathcal{G}}_{Def}^{i+1} = (I - W_\mu)\hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{abs}(\vec{\zeta}_{abs})) + W_\mu\hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{rel}(\vec{\zeta}_{rel}))$

5.1.7 Esquema de Descenso Supervisado

Nuestra propuesta de resolución se concreta en un esquema basado en *el descenso por gradiente, supervisado o controlado por la aplicación/verificación de las auto-restricciones del modelo considerado*. Dicho esquema comprende, en general, dos etapas:

1. *Inicialización*: dependiendo de la información disponible acerca de las condiciones iniciales del problema, se trata de proponer una *configuración inicial* \mathcal{G}^0 . En caso de que dicha información se considere insuficiente, se procede a una *búsqueda o muestreo* que permite extraer zonas de la imagen sobre las que posicionar las plantillas.
2. *Refinamiento Iterativo/Verificación*: a partir de la estructura \mathcal{G}^i obtenemos \mathcal{G}^{i+1} combinando los procesos de minimización local de energía correspondientes a las plantillas con la actualización, si procede, de la estructura obtenida, para forzar la satisfacción al máximo de las restricciones del modelo. Este proceso conduce a una solución localmente óptima.

5.1.8 Algoritmo de Minimización por Supervisión

En términos generales este esquema actúa de forma iterativa partiendo de una configuración inicial conocida \mathcal{G}^0 de la cual depende el resultado final. En general, se trata de obtener la configuración \mathcal{G}^{i+1} a partir de \mathcal{G}^i de manera que:

$$\mathcal{G}^{i+1} = (1 - \lambda)\hat{\mathcal{G}}_{Local}^{i+1} + \lambda\hat{\mathcal{G}}_{Def}^{i+1} \quad (5.3)$$

es decir, integrando de forma ponderada el posicionamiento inducido por las variaciones locales de las plantillas con el debido a la mejora incremental en la satisfacción de restricciones geométricas. La nueva estructura \mathcal{G}^{i+1} se computa en dos etapas:

1. *Descenso No-Restringido*: en primer lugar aplicamos las dinámicas locales a cada plantilla para minimizar $Local(\mathcal{G})$, es decir computamos:

$$\hat{\mathcal{G}}_{Local}^{i+1} = \mathcal{G}^i + \frac{d\mathcal{G}_{Local}^i}{dt} \quad (5.4)$$

lo cual tiene solamente en cuenta la respuesta local de las plantillas ante el potencial de imagen. Por tanto este resultado parcial debe ser refinado (siempre y cuando $\lambda \neq 0$, ya que en caso contrario el paso de iteración terminaría en este punto).

2. *Ajuste Basado en Auto-Restricciones*: el refinamiento de la dinámica local supone tener en cuenta el segundo criterio de optimización, es decir $Def(\mathcal{G}, \mathcal{R}(\vec{\zeta}))$. En este paso obtenemos $\hat{\mathcal{G}}_{Def}^{i+1}$ a partir de las variaciones en la satisfacción de las restricciones absolutas y relativas del modelo, es decir, teniendo en cuenta que

$$\hat{\mathcal{G}}_{Def}^{i+1} = (I - W_\mu) \hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{abs}(\vec{\zeta}_{abs})) + W_\mu \hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{rel}(\vec{\zeta}_{rel})) \quad (5.5)$$

siendo W_μ una matriz diagonal de pesos $\mu_i \in [0, 1]$ que modulan la importancia relativa para restricciones de ambos tipos. Para simplificar la notación, cuando todos los pesos sean idénticos sustituiremos la matriz por un solo peso μ . Este proceso inducirá un acercamiento incremental hacia las distribuciones representadas por las restricciones definidas. Parte de esta computación se realiza en un espacio de baja dimensión (auto-espacio).

Este procedimiento se aplica hasta detectar ausencia de cambios. El resultado final, es decir la aproximación al óptimo \mathcal{G}^* depende tanto de las condiciones iniciales como del error asumido una vez seleccionados los niveles de confianza para las restricciones consideradas. En este sentido debe buscarse un equilibrio entre coste computacional (pocos modos reducen el coste) y error cometido.

El esquema propuesto está inspirado en el mecanismo de *matching* de las Formas Activas revisado en el Capítulo 2. El mecanismo de proyección-supervisión se realiza en nuestro trabajo a un nivel de *abstracción estructural*. En la Figura 5.2 se detalla el bucle de retroalimentación referido anteriormente para el caso de una restricción de posición relativa. Solamente se consideran cinco de los diez modos, es decir, los más representativos (baja frecuencia).

5.1.9 Ejemplo Básico de Correspondencia

En la Figura 5.3 se ilustra el comportamiento del algoritmo en función de la ponderación con el parámetro λ . Las restricciones, como modelo estocástico de baja dimensión, juegan un papel de *regularización y filtrado del resultado local*. La confianza en la decisión local se modula mediante λ . El caso de ponderación extremo puede observarse en el siguiente experimento (ver Figura 5.4) en el cual se ha aprendido un modelo básico compuesto por dos plantillas y la restricción posicional absoluta. Partiendo de las mismas condiciones de inicialización el valor de λ tiene un fuerte impacto en el resultado final. Cuando fijamos $\lambda = 0.1$ (izquierda) la plantilla elíptica, situada entre la plantilla aprendida y otra con similares características (elipse central), *no tiene evidencia local suficiente* para decidir un nuevo posicionamiento ya que se induce un mínimo local. En consecuencia *la plantilla*

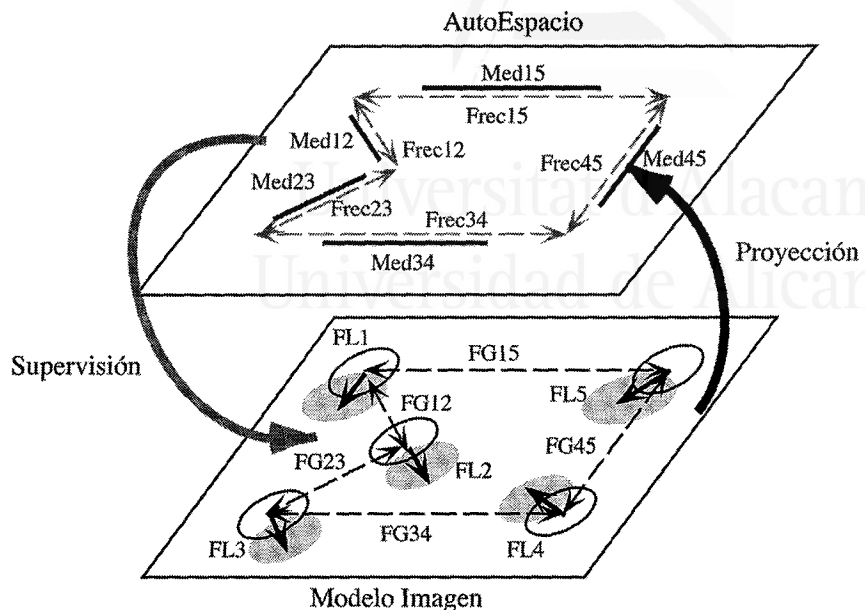


Figura 5.2: Efecto de restricción relativa de posición

tiende a expulsar el potencial no programado reduciendo progresivamente su escala. Por el contrario un peso de $\lambda = 0.5$ (derecha) es suficiente para que la plantilla evite este mínimo y tienda hacia su posición y escala correctas.

En la sección siguiente realizamos un estudio más profundo de la incidencia de este y de otros factores (principalmente el ruido gaussiano). Veremos como la existencia de un modelo geométrico computable de forma eficiente puede salvar situaciones de convergencia hacia configuraciones erróneas.

5.2 Experimentación en el Dominio Espacial

5.2.1 Objetivos y Condiciones Experimentales

En esta sección presentamos un conjunto de experimentos de *correspondencia espacial* dedicados a evaluar el funcionamiento del esquema propuesto bajo dos puntos de vista:

1. *Tests de Robustez*: estudio del impacto de las condiciones de inicialización y de la existencia de potencial no-programado, teniendo en cuenta el cono-

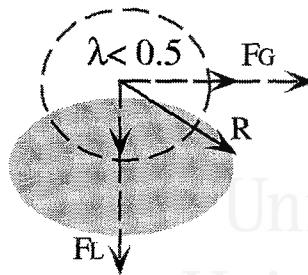


Figura 5.3: Detalle de Integración-Cálculo de Resultante

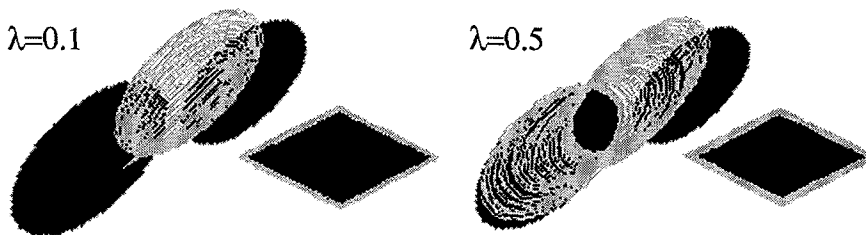


Figura 5.4: Incidencia del parámetro de ponderación

cimiento codificado en las auto-restricciones tanto absolutas como relativas previamente extraídas (*off-line*).

2. *Análisis de Relajación y Ponderación*: estudiar el efecto de la relajación de auto-restricciones, es decir de los parámetros codificados en $\vec{\zeta}$, así como de la ponderación debida al parámetro λ .

5.2.2 Análisis y Discusión

Partiendo de un potencial de imagen que contiene una estructura geométrica promedio (satisface en grado máximo todas las restricciones del modelo espacial aprendido en el capítulo anterior) estimaremos la respuesta ante distintos coeficientes de ponderación y niveles de confianza, todo ello en diversos contextos de ruido gaussiano (20%,40% y 50%). Presentamos dos bloques de experimentos cuyos resultados se resumen en la Tabla 5.2. Los resultados obtenidos, a partir de condiciones de inicialización homogéneas para cada bloque, pueden visualizarse en las Figuras 5.5-5.9:

Tabla 5.2: Evaluación del Esquema de Descenso. Caso Espacial

Pruebas	Casos	α	β	λ	μ	ζ_a	ζ_r	Ruido
Exp. 1	1.1	0.1	0.1	0.0	1.0	0.0	0.9	20%
	1.2	0.3	0.2	0.0	1.0	0.0	0.9	40%
Exp. 2	2.1	0.1	0.1	0.5	1.0	0.0	0.9	20%
	2.2	0.3	0.2	0.7	1.0	0.0	0.9	40%
Exp. 3	3.1	0.1	0.1	0.8	1.0	0.0	0.9	20%
	3.2	0.3	0.2	0.9	1.0	0.0	0.9	40%
Exp. 4	4.1	0.3	0.2	0.5	0.0	0.5	0.0	40%
	4.2	0.3	0.2	0.0	0.0	0.1	0.0	40%
Exp. 5	5.1	0.3	0.2	0.8	0.5	0.1	0.9	40%
	5.2	0.3	0.2	0.8	0.5	0.1	0.9	50%

Bloque 5.1 (Robustez y Ponderación) (*Experimentos 1-3*) Estos experimentos (ver Figuras 5.5-5.7) muestran el comportamiento del algoritmo en contextos ruidosos. El primer caso de cada experimento asume un 20% de ruido y unos coeficientes α y β de 0.1. En el segundo caso el ruido se fija en el 40%. En consecuencia los coeficientes de ponderación deben aumentar para poder obtener buenos resultados (en este caso se considera suficiente fijar α a 0.3 y β a 0.2). Todos los experimentos de este bloque fueron realizados a un mismo nivel de confianza y solamente se tomaron en cuenta restricciones relativas. Además no se aplicaron restricciones angulares o direccionales con lo cual se evidencia todavía más la sensibilidad de las dinámicas locales al ruido (ya que este parámetro junto con el de forma son los más sensibles). Se observa *una mejora incremental* para ambos niveles de ruido a medida que entran en juego las restricciones (aumenta λ). Los niveles elevados de ruido aconsejan, en general, mayor participación de las restricciones en la búsqueda y por tanto menor confianza en el modelo de imagen.

Bloque 5.2 (Niveles de Confianza) (*Experimentos 4-5*) En este caso las condiciones experimentales son idénticas en lo que se refiere a valores de $\alpha = 0.3$ y $\beta = 0.2$ y solamente entran en juego restricciones absolutas ($\mu = 0.0$ y $\zeta_r = 0.0$),

incluyéndose en este caso la restricción angular. En el experimento 4 (nivel de ruido del 40%) se observa la diferencia, en cuanto a calidad de la respuesta, entre incorporar restricciones absolutas a un nivel de confianza del 50% y no hacerlo. Finalmente en el experimento 5 estudiamos el funcionamiento del sistema en situaciones con pocos modos de variación (p.e. cuando existen fuertes requerimientos en el tiempo de respuesta). Los resultados mostrados corresponden a los mejores casos: conseguimos paliar el bajo número de modos, incrementando λ , y reforzando el modelo con restricciones relativas adicionales (salvo la angular) aunque en términos generales se consideren con el mismo peso. Este resultado es equivalente al obtenido trabajando con restricciones muy genéricas y evidencia la necesidad de establecer un compromiso entre calidad y eficiencia.

Estos resultados demuestran que el conocimiento, y no siempre a un nivel de detalle completo, acerca de la estructura permite mejorar el rendimiento de las funciones de energía en condiciones adversas. Por lo tanto disponemos de un esquema de *matching* o ajuste espacial el cual puede extenderse fácilmente al caso espacio-temporal y vertebrar el proceso de *tracking* basado en modelos estructurales.

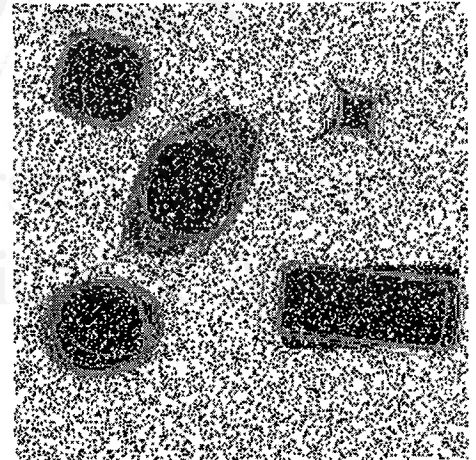
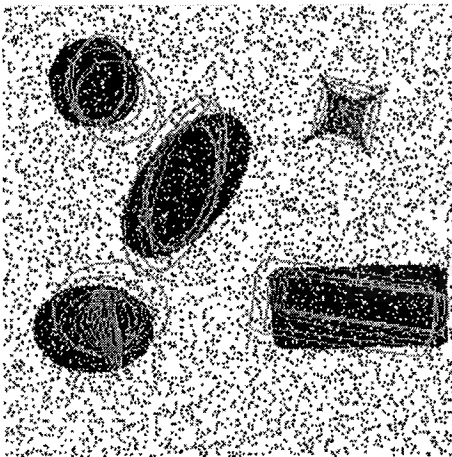


Figura 5.5: Experimento 1

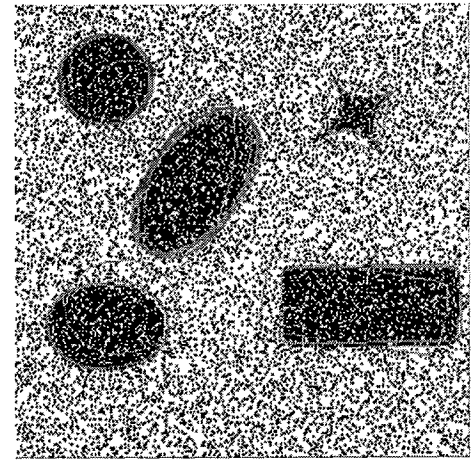
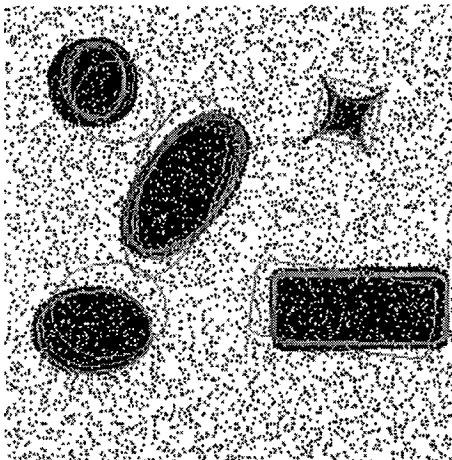


Figura 5.6: Experimento 2



Universitat d'Alacant
Universidad de Alicante

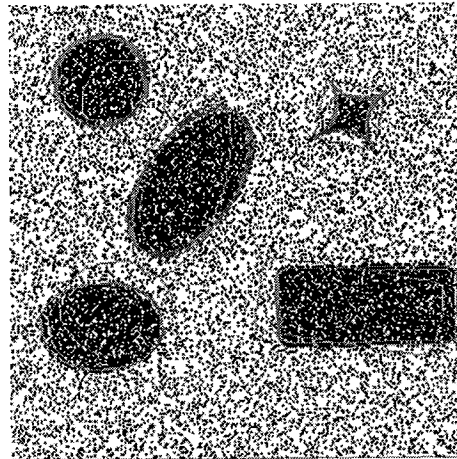
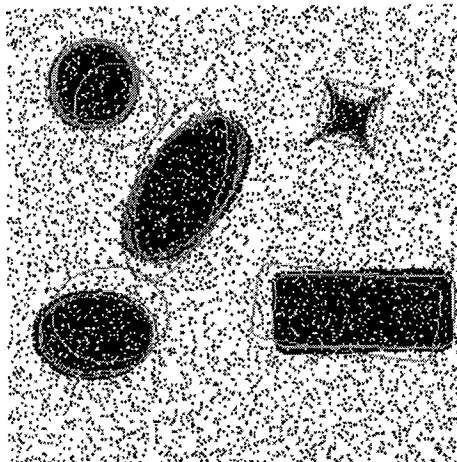


Figura 5.7: Experimento 3

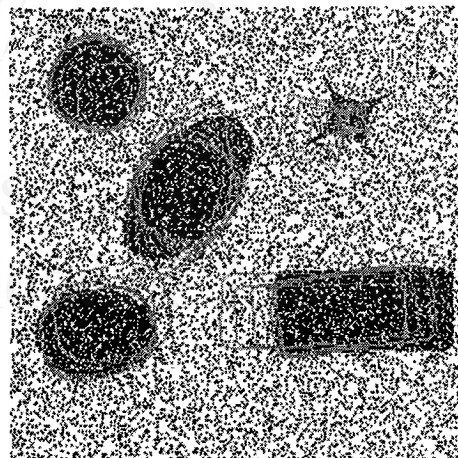
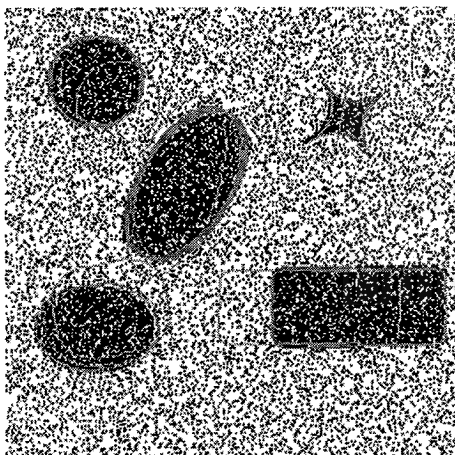


Figura 5.8: Experimento 4

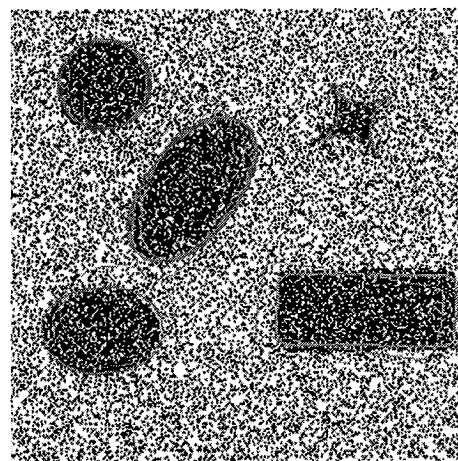
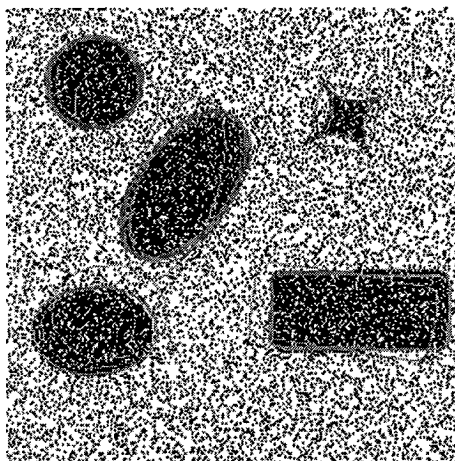


Figura 5.9: Experimento 5

5.3 Tracking/Correspondencia Espacio-Temporal

5.3.1 Planteamiento/Elementos del Tracker

Proponemos un mecanismo de seguimiento estructural de contornos bidimensionales que integra dos elementos fundamentales:

1. *Modelo Espacio-Temporal*: la dinámica de un conjunto de plantillas en una ventana temporal determinada $[0, T]$ viene caracterizada por el conjunto de auto-restricciones espacio-temporales definidas en dicha ventana.
2. *Mecanismo de Predicción-Actualización*: la integración de ambos aspectos (tendencia al promedio) a partir de la *generalización al caso espacio-temporal del mecanismo de correspondencia* permite, dependiendo de las condiciones en el instante t_k , realizar un ajuste poderado en un instante posterior.

5.3.2 Estructuras Espacio-Temporales Óptimas

Definición 5.7 Dados un conjunto $\mathcal{G}(t) = \{g_{\mathcal{D}_i}(t, \Theta)\}_{i=1}^g$ de Plantillas Super-Elipsoidales dependientes del tiempo y su modelo $\mathcal{R}(\vec{\zeta}, t) = \mathcal{R}_{abs}(\vec{\zeta}_{abs}, t) \cup \mathcal{R}_{rei}(\vec{\zeta}_{rei}, t)$ de *Restricciones Espacio-Temporales Propias Finales* y siendo $\lambda \in [0, 1]$ definimos como *Estructura Espacio-Temporal Óptima* $\mathcal{G}^*(t)$ el conjunto de plantillas dependientes del tiempo que satisface:

$$\mathcal{G}^*(t) = \text{Arg Min } H(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta}))$$

$$H(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta})) = (1 - \lambda) \text{Local}(\mathcal{G}(t)) + \lambda \text{Def}(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta})) \quad (5.6)$$

$$\text{Local}(\mathcal{G}(t)) = \int_0^T \text{Local}(\mathcal{G}(t)) dt \quad (5.7)$$

donde $\text{Local}(\mathcal{G}(t)) \approx \sum_{k=0}^T \text{Local}(\mathcal{G}(t_k))$ debido a la discretización. Se trata de encontrar la *trayectoria temporal óptima para todos los parámetros de la estructura*. Este problema es asimilable al problema de encontrar el conjunto de contornos óptimo.

5.3.3 Extensión del Esquema de Descenso Supervisado

Teniendo en cuenta la caracterización espacio-temporal compacta, y frecuentemente de baja-dimensionalidad, obtenida mediante ACP, es asumible almacenar en la memoria del *tracker* y para cada auto-restricción definida, el vector promedio, y la matriz de los modos principales de variación (dependiendo del nivel de confianza asumido). A partir de este punto cabe aplicar dos esquemas *extremos* de resolución:

1. *Ventana Completa*: partiendo de una configuración inicial $\mathcal{G}^0(t)$ de las plantillas en todos los *frames* se trata de aplicar el esquema de descenso descrito anteriormente teniendo en cuenta la incorporación de la dimensión temporal. Este esquema, que requiere elevada capacidad de almacenamiento, sobre todo para ventanas de larga duración, solamente es viable en contextos con bajos requerimientos de respuesta en tiempo real (p.e. reconstrucción de órganos). En este caso no habría mecanismo predictivo.
2. *Ventana Unitaria*: en el otro extremo se sitúa la aproximación *frame-a-frame* en la cual partimos de una configuración inicial espacial $\mathcal{G}^0(t_0)$. Se trata de aplicar el descenso supervisado en cada instante de tiempo recogiendo las recomendaciones del modelo espacio-temporal completo. En este caso el mecanismo predictivo tiene sentido pero la decisión de posicionamiento tomada en un instante es inalterable ya que sólo conservamos en memoria el *frame* actual. El resultado obtenido es un óptimo local de $H(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta}))$.

Entre ambos extremos caben esquemas mixtos (basta con fijar ventana intermedia) cuya aplicabilidad depende de los recursos computacionales disponibles y de los requerimientos en tiempos de respuesta. En este sentido, comprobado ya experimentalmente el esquema de descenso espacial, desarrollaremos a continuación el esquema de *ventana unitaria*.

5.3.4 Algoritmo de Tracking Supervisado con Ventana Unitaria

Partiendo de una configuración espacial conocida $\mathcal{G}^0(t_0)$ se trata de obtener, para cada *frame* $t_k \mid k \in \{0, 1, \dots, T\}$ una configuración espacial $\mathcal{G}^*(t_k)$. El procedimiento consta de dos fases interrelacionadas:

1. *Predicción*: a partir de $\mathcal{G}^*(t_{k-1})$ (o de $\mathcal{G}^0(t_0)$ si $k = 0$) aplicamos p iteraciones para predecir la configuración de partida en el instante t_k . En la iteración i -ésima aplicamos el modelo espacio-temporal adquirido para computar:

$$\hat{\mathcal{G}}_{Pred}^i(t_k) = (I - W_\mu) \hat{\mathcal{G}}_{Def}^i(\mathcal{R}_{abs}(t_k, \vec{\zeta}_{abs})) + W_\mu \hat{\mathcal{G}}_{Def}^i(\mathcal{R}_{rel}(t_k, \vec{\zeta}_{rel})) \quad (5.8)$$

donde $\mathcal{R}_{abs}(t_k, \vec{\zeta}_a)$ y $\mathcal{R}_{rel}(t_k, \vec{\zeta}_b)$ se obtienen proyectando las restricciones espacio-temporales en el instante t_k , es decir suponiendo que para todo instante anterior o posterior a t_k se satisface la configuración promedio de todas las restricciones consideradas. La predicción $\hat{\mathcal{G}}_{Pred}^p(t_k)$ se obtiene a partir de criterios exclusivamente estructurales, es decir sin considerar dinámicas locales, para reducir el riesgo de mínimos locales. El número de iteraciones

necesarias para una buena predicción será tanto mayor cuanto menor sea la continuidad espacio-temporal³.

2. *Actualización/Correspondencia*: tomado como punto de partida $\hat{\mathcal{G}}_{Pred}^p(t_k)$ se procede a aplicar el algoritmo de correspondencia espacial teniendo en cuenta que las restricciones temporales se aplican solamente al instante de tiempo considerado y tomando como base los modos de variación establecidos para cada restricción espacio-temporal. La configuración $\mathcal{G}^{i+1}(t_k)$ se obtiene a partir de $\mathcal{G}^i(t_k)$:

$$\mathcal{G}^{i+1}(t_k) = (1 - \lambda)\hat{\mathcal{G}}_{Local}^{i+1}(t_k) + \lambda\hat{\mathcal{G}}_{Def}^{i+1}(t_k) \quad (5.9)$$

lo cual se computa en dos etapas:

- (a) *Descenso No-Restringido*: en primer lugar se trata de avanzar en la minimización de $Local(\mathcal{G})$, en el instante t_k :

$$\hat{\mathcal{G}}_{Local}^{i+1}(t_k) = \mathcal{G}^i(t_k) + \frac{d\mathcal{G}_{Local}^i(t_k)}{dt} \quad (5.10)$$

- (b) *Ajuste Basado en Auto-Restricciones*: para el refinamiento de la dinámica local en t_k se aplican las restricciones espacio-temporales tal como se hacía en la fase de predicción:

$$\hat{\mathcal{G}}_{Def}^{i+1}(t_k) = (I - W_\mu)\hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{abs}(t_k, \vec{\zeta}_{abs})) + W_\mu\hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{rel}(t_k, \vec{\zeta}_{rel})) \quad (5.11)$$

Este proceso iterativo nos permite obtener $\mathcal{G}^*(t_k)$. Esta fase tiene un doble efecto integrador-corrector: integrador por la mecánica del procedimiento de correspondencia espacial; corrector porque reajusta las desviaciones del modelo predicho con respecto al modelo de imagen siempre atendiendo, y con el efecto de ponderación establecido, al criterio de ajuste estructural.

Las configuraciones espaciales $\mathcal{G}^*(t_0), \mathcal{G}^*(t_1), \dots, \mathcal{G}^*(t_T)$ obtenidas a partir del análisis de la secuencia constituyen, en conjunto, una aproximación local a $\mathcal{G}^*(t)$. Para mejorar la calidad de la respuesta debe ampliarse la ventana, tanto hacia adelante como hacia atrás en el tiempo.

³Un número demasiado bajo de iteraciones posicionará la estructura demasiado cerca de la configuración correspondiente al *frame* anterior.

Tabla 5.3: Tracking Basado en Modelos Espacio-Temporales

Estructura Espacio-Temporal Óptima	
Criterio/Objetivo Heurística	$H(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta})) = (1 - \lambda)Local(\mathcal{G}(t)) + \lambda Def(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta}))$
Criterio Local/Imagen	$Min Local(\mathcal{G})(t) \rightarrow Min Local(\mathcal{G})(t_k) \forall k \in \{0, \dots, T\}$
Criterio Estructural	$Min Def(\mathcal{G}(t), \mathcal{R}(t, \vec{\zeta})) \rightarrow Min Def(\mathcal{G}(t_k), \mathcal{R}(t_k, \vec{\zeta})) \forall k \in \{0, \dots, T\}$
Algoritmo de Ventana Unitaria	
Predicción	$\hat{\mathcal{G}}_{Pred}^i(t_k) = (I - W_\mu)\hat{\mathcal{G}}_{Def}^i(\mathcal{R}_{abs}(t_k, \vec{\zeta}_a)) + W_\mu\hat{\mathcal{G}}_{Def}^i(\mathcal{R}_{rel}(t_k, \vec{\zeta}_b))$
Actualización	$\mathcal{G}^{i+1}(t_k) = (1 - \lambda)\hat{\mathcal{G}}_{Local}^{i+1}(t_k) + \lambda\hat{\mathcal{G}}_{Def}^{i+1}(t_k)$
	$\hat{\mathcal{G}}_{Local}^{i+1}(t_k) = \mathcal{G}^i(t_k) + \frac{d\mathcal{G}_{Local}^i(t_k)}{dt}$
	$\hat{\mathcal{G}}_{Def}^{i+1}(t_k) = (I - W_\mu)\hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{abs}(t_k, \vec{\zeta}_{abs})) + W_\mu\hat{\mathcal{G}}_{Def}^{i+1}(\mathcal{R}_{rel}(t_k, \vec{\zeta}_{rel}))$

5.4 Experimentación en el Dominio Espacio-Temporal

5.4.1 Objetivos y Condiciones Experimentales

En los experimentos realizados para el caso espacial se constató el buen comportamiento del esquema supervisado *en presencia de ruido gaussiano medio-alto y en condiciones de inicialización no-triviales* sobre múltiples plantillas. Por otro lado, el esquema de tracking que acabamos de proponer añade, *como novedades al caso espacial*, la integración entre predicción y actualización y la aplicación de modos espacio-temporales de variación. En consecuencia los experimentos a realizar en esta sección se centrarán en analizar el impacto de tres aspectos:

1. *Discontinuidad Temporal:* necesidad de ajuste adecuado del número de iteraciones de predicción para evitar mínimos locales debidos a la incorporación prematura del modelo de imagen.
2. *Combinación de Restricciones:* dado que las restricciones relativas dotan al modelo de invarianza espacial, el uso exclusivo de restricciones de este tipo a nivel espacio-temporal puede inducir mínimos locales.
3. *Niveles Espacio-Temporales de Confianza:* impacto de la reducción del número de modos espacio-temporales de variación en degradación de la respuesta

del *tracker* o mecanismo de seguimiento.

5.4.2 Análisis y Discusión

Tomaremos como referencia para nuestros experimentos una secuencia de imágenes como la estudiada en el capítulo anterior para ilustrar el reconocimiento espacio-temporal. Dicha escena está formada por dos plantillas elípticas que evolucionan rotacional y traslacionalmente. La evolución posicional de las plantillas de la escena no se corresponde, en la parte central de la secuencia, con la configuración promedio aprendida. La inicialización es ajustada y homogénea para ambos experimentos. En todos los casos $\alpha = \beta = 0.1$. En las Figuras 5.10,5.11 se presentan los resultados. Se han superpuesto los potenciales de imagen correspondientes a todos los instantes de tiempo para una mejor apreciación del resultado final.

Experimento 5.1 (Discontinuidad Temporal) Aplicamos únicamente restricciones absolutas (todas ellas con nivel de confianza 0.9). En primer lugar (Caso 1.1) realizamos una fase de predicción con $p = 10$ iteraciones. En la parte central de la secuencia, en donde hay un leve retraso temporal, el ajuste viene dominado por las restricciones. A continuación (Caso 1.2) reducimos p a 4 iteraciones. Esta reducción implica que el modelo local comienza a aplicarse prematuramente y, en este caso, la búsqueda alcanza un mínimo local de energía.

Experimento 5.2 (Confianza y Combinación de Restricciones) Aplicamos *restricciones absolutas de forma y escala*, cuya variabilidad es muy baja, y *restricciones relativas de posición y orientación*. En el primer caso, el nivel de confianza de las restricciones aplicadas es 0.9. Aún con este nivel de confianza la calidad del resultado es inferior a la obtenida únicamente con restricciones absolutas. En efecto las restricciones relativas soportan invarianza espacial lo cual implica que en el siguiente frame el desplazamiento propuesto puede satisfacer la restricción pero puede ser incoherente con el modelo de imagen. En el segundo caso este problema se agudiza con la reducción a un único modo de variación. El resultado obtenido coincide con el derivado de aplicar, en el Caso 1.2, menos iteraciones de predicción (se cae en el mismo mínimo local).

5.5 Consideraciones

A lo largo de este capítulo hemos integrado el modelo local, definido en el Capítulo 3, con la caracterización estructural presentada en el Capítulo 4, para definir un esquema de correspondencia espacial en términos de la minimización de una función objetivo que combina el modelo de imagen con el conocimiento geométrico.

Dicho esquema es el núcleo del mecanismo de *tracking* basado en conocimiento estructural, y al cual hemos incorporado capacidad predictiva. Así mismo presentamos un conjunto de experimentos representativos que ponen de relieve la tolerancia del modelo al ruido a partir del conocimiento geométrico extraído de forma automática. En el Capítulo 6 demostraremos la aplicabilidad del esquema propuesto en contextos de *reconocimiento automático de gestos*.

Universidad de Alicante

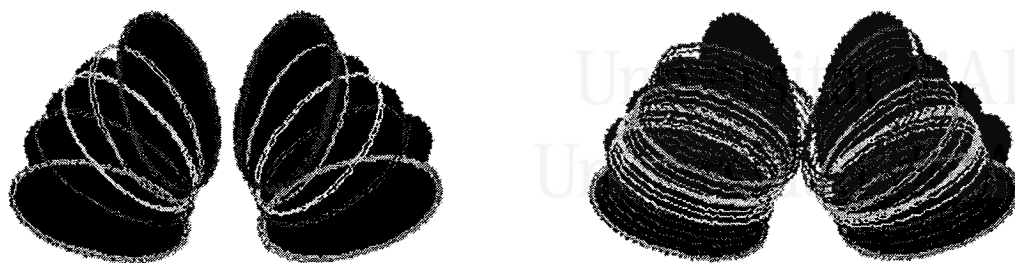


Figura 5.10: Experimento 1

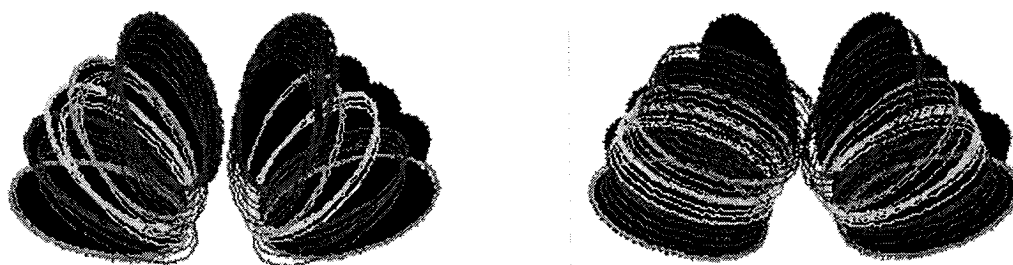


Figura 5.11: Experimento 2



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Parte III

Dominios de Aplicación Conclusiones



Capítulo 6

Dominios de Aplicación

En este capítulo presentamos resultados experimentales vinculados a la aplicación de la técnicas desarrolladas en la parte central de esta disertación a dos contextos en los que el tratamiento de secuencias de imágenes resulta de gran utilidad. El primero es la *reconstrucción, a partir de imágenes de ultrasonidos*, de la pared arterial y *el análisis de la deformación* inducida por procedimientos de angioplastia. Estos experimentos motivaron, por un lado, la extensión del modelo deformable circular al caso elíptico y, a partir de este, su extensión al modelo super-elipsoidal (Capítulo 3). Por otro lado la necesidad de disponer de mecanismos de aprendizaje del conocimiento experto nos llevó a formalizar y desarrollar el modelo de auto-restricciones basadas en ACP (Capítulo 4). La segunda aplicación es de distinta naturaleza y está vinculada al desarrollo de interfaces hombre-máquina capaces de *interpretar gestos*. En este sentido el método de *tracking* y correspondencia desarrollado en el Capítulo 5 se aplica al reconocimiento de gestos formulándose el problema como la obtención del modelo más adecuado a la secuencia de entrada (aquel para el que dicha secuencia satisface más condicionantes geométricos que para ningún otro).

6.1 Análisis de Secuencias Intravasculares

6.1.1 Introducción y Motivación

La extracción y uso de imágenes intravasculares basadas en ecocardiografías constituye una técnica relativamente reciente que permite obtener la estructura interna de las arterias a un detalle suficiente, y mucho mayor que cualquier procedimiento no-invasivo actual (Yock et al.,1989),(Goar et al., 1992), para la planificación de tratamientos coronarios (*bypass*, angioplastia, aterectomía). Esta técnica se basa

en colocar un catéter que incorpora en un extremo un transductor. El catéter es rotado y genera, emitiendo pulsos de ultrasonidos y recibiendo ecos, secciones planas correspondientes a la estructura arterial. En la imagen obtenida (ver Figura 6.1) el centro del catéter se toma como origen del sistema de referencia y pueden apreciarse típicamente tres tipos de eco: *luminosidad del flujo sanguíneo* (intensidades bajas), *placa de calcio o ateroma* (intensidad media) y *pared vascular* (intensidad alta). El análisis del tipo de placa ayuda a los especialistas a seleccionar la modalidad de intervención más aconsejable.

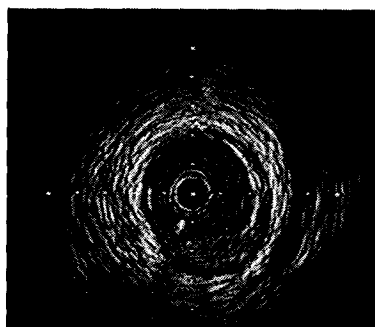


Figura 6.1: Ejemplo de imagen obtenida mediante ecocardiografía.

Resulta atractivo disponer, a efectos prácticos, de una reconstrucción lo más precisa posible de la estructura vascular a partir de secciones, como soporte fundamental del diagnóstico (en lo que se refiere a extracción de información cuantitativa) y de la propia actividad quirúrgica (p.e. en lo relativo a determinar ajustes del mecanismo de activación de *stents* en angioplastias). En este sentido las técnicas previas más relevantes desarrolladas hasta el momento son las siguientes:

1. *Rendering por Pilas de Imágenes*: (Rosenfield et al., 1991), (Krishnaswamy et al., 1992), (Isner et al., 1992), (Roelandt, et. al, 1994)
 - (a) Se asume geometría estática y estructura vascular recta. Las visualizaciones obtenidas se basan en el apilamiento de secciones.
 - (b) La geometría final no se corresponde con la realidad en la mayoría de los casos (distorsión).
 - (c) En aportaciones posteriores (Lengyel et al., 1995) se introducen mejoras en lo relativo a tratamiento de estructuras curvas pero la geometría sigue siendo estática.
2. *Incorporación de Snakes*: (Hyche et al., 1992)

- (a) La identificación espacio-temporal de la estructura arterial mediante modelos deformables se plantea en el ámbito de la angiografía (*tracking* de las paredes externas de la arteria).
- (b) En definitiva esta técnica basada en modelos deformables permite extraer la geometría dinámica real.

3. Integración con Angiografía: (Hyche et al., 1995)

- (a) Con respecto al trabajo anterior se obtiene, en primer lugar, la información estéreo angiográfica (estructura 3D obtenida por *matching* basado en mecanismos variacionales) para lo cual se plantean nuevas funciones de energía (en concreto de definen desplazamientos con respecto al centro de la arteria).
- (b) Dicha información se integra, mediante la sincronización espacio-temporal de transductor y angiografía, con las imágenes correspondientes a las secciones transversales para obtener la visualización de la estructura mediante *rendering* volumétrico basado en interpolación.
- (c) En este sentido el posicionamiento de las secciones está guiado por marcadores (*landmarks*) correspondientes a las intersecciones del árbol vascular.
- (d) Este enfoque se centra en el problema de la visualización, y en definitiva se plantean mejoras en cuatro líneas de trabajo. En primer lugar se trata de tener en cuenta que el catéter no está centrado y que por tanto debe obtenerse el desplazamiento. En segundo lugar urge mejorar la integración ultrasonido-angiografía sobre todo en lo que respecta a la sincronización con el ciclo cardíaco. En tercer lugar se sugiere un mejor ajuste de los coeficientes de los snakes. Finalmente se plantea la mejora de los procesos de *rendering*, fundamentalmente en la obtención de secciones con menor discontinuidad temporal de manera que los supuestos de la interpolación sean aplicables.

4. Temas Abiertos:

- (a) Desde nuestro punto de vista se trata de derivar esfuerzos en la extracción y análisis de la estructura geométrica interna de las secciones, mediante modelos deformables, previo análisis profundo del potencial (Dias y Leitao, 1996).
- (b) Esto supone por un lado automatizar el análisis de la imagen de ultrasonidos (caracterización por textura, definición de potenciales), y

por otro definir y capturar el modelo geométrico definido por las capas circulares o elípticas.

- (c) Disponer de estas técnicas permitirá en primera instancia identificar placas de ateroma (diagnóstico) y en segunda instancia guiar el proceso de angioplastia, con lo cual la información deja de ser meramente gráfica y pasa a ser cuantitativa.

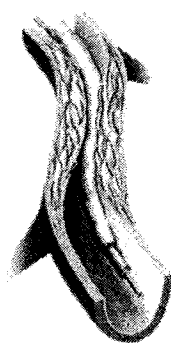


Figura 6.2: Introducción de un catéter con balón.

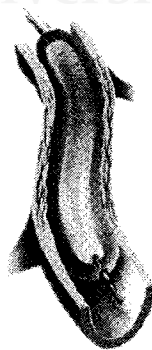


Figura 6.3: Insuflación y compresión de la placa.

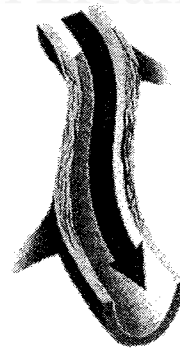


Figura 6.4: Desaparición de la placa y normalización del flujo sanguíneo.

6.1.2 Modelo de Imagen: Campos de Potencial

En primera instancia deseamos un acotamiento de la zona de búsqueda del atero-
ma, por lo que el filtrado inicial de la imagen consta de un umbralizado, aplicándose a continuación un operador morfológico (cierre) que permite la eliminación de estructuras locales que introduzcan distorsiones potenciales y un suavizado del contorno de la región. El resultado de este preproceso se muestra en la Figura 6.5. La Figura 6.6 muestra la magnitud del gradiente de la Figura 6.5 y que será el que proporcionará la energía de centrado de las plantillas.

6.1.3 Extracción de la Pared Arterial: Plantillas Circulares

La extracción de los límites interior y exterior de la pared arterial nos permitirán acotar la zona donde puede estar localizada la placa. De esta manera la aplicación de un reconocedor de textura en dicha zona puede realizarse en tiempo real. Dado el tipo de imagen a tratar se observa la adecuación del modelo circular, presentado



Figura 6.5: Resultado del filtrado aplicado.



Figura 6.6: Gradiente de la imagen filtrada.

anteriormente, para el *tracking* de la estructura geométrica interna de la arteria. Partimos de la inicialización de la primitiva en el interior del vaso, y de las condiciones de experimentación. Los resultados del *tracking* circular se presentan en la Figura 6.8.

En la Figura 6.7 se muestra la interpolación de los distintos círculos¹ encontrados a lo largo de la secuencia.

6.1.4 Plantillas Elípticas: Tracking en Angioplastia

En el contexto de la imagen intravascular obtenida por ultrasonidos, uno de los procedimientos médicos en los que el *tracking* no-rígido puede introducir un cierto grado de automatización es la *Angioplastia Coronaria*. Dicho procedimiento consiste, como puede observarse en las Figuras 6.2, 6.3 y 6.4, en la dilatación, mediante el uso de un catéter con un pequeño balón en su punta, de las lesiones que obstruyen la arteria. Con el hinchado del balón, la presión ejercida comprime y rasga el depósito de materia grasa o placa de ateroma reduciendo o haciendo desaparecer la estrechez o estenosis arterial.

En este sentido el hecho de que la geometría de la forma inducida por el hinchado podía ser elíptica, motivó la extensión del modelo circular. La formulación de la plantilla que veremos a continuación contempla la incorporación de parámetros de ruido y es el *punto de partida para el modelo super-elipsoidal presentado en el Capítulo 3*:

1. *Estructura, Potencial y Energía Global*: la estructura de la plantilla se expresa mediante $g_D(\Theta) = (x, y, a, b, \vec{V}_{ref})$ donde: (x, y) es el centro, a, b son

¹Los círculos interpolados corresponden a los interiores.

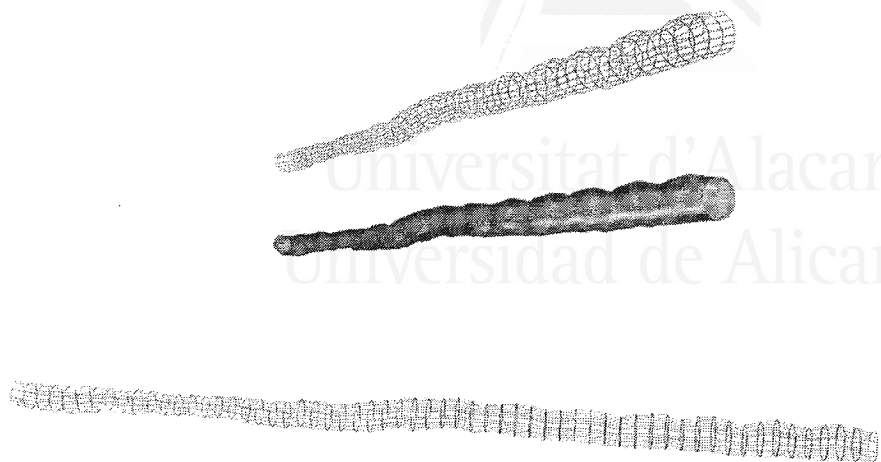


Figura 6.7: Secuencia de círculos deformables.

los ejes (mayor y menor), $\vec{V}_{ref} = (\cos \theta, \sin \theta)$ es el vector de referencia, siendo θ la orientación en el plano del eje secundario y \mathcal{D} el dominio interno acotado por la frontera de la elipse. Basándonos en el potencial binario extraído anteriormente planteamos la siguiente energía global:

$$\mathcal{E}(x, y, a, b, \vec{V}_{ref}) = \mathcal{E}_{con}(x, y, a, b) + \mathcal{E}_{rot}(x, y, \vec{V}_{ref}) - \mathcal{E}_{exp}(a, b) \quad (6.1)$$

cuyos términos y dinámica concretamos a continuación:

- (a) *Energía de Contracción:* el ajuste de los parámetros de posición, curvatura y escala, viene regido por:

$$\mathcal{E}_{con}(x, y, a, b) = \int \int_{\mathcal{D}} \left(1 - \left(\frac{u^2}{a^2} + \frac{v^2}{b^2}\right)\right) \mathcal{I}(x + u, y + v) du dv \quad (6.2)$$

- (b) *Energía de Rotación:* la conducta de adaptación del vector de referencia se basa en:

$$\mathcal{E}_{rot}(x, y, \vec{V}_{ref}) = (\mathcal{N}(\vec{V}_{ref}) \cdot \mathcal{N}(\vec{\nabla} \mathcal{I}(p_1)))^2 \quad (6.3)$$

donde \mathcal{N} indica la normalización del vector que toma como argumento, el operador \cdot representa el producto escalar y p_1 es el primer punto del

semieje positivo de a con gradiente no-nulo, es decir:

$$p_1 = (x + i \cos(\theta - \frac{\pi}{2}), y + i \sin(\theta - \frac{\pi}{2})) \mid 0 < i < a, \vec{\nabla} \mathcal{I}(p_1) \neq 0 \quad (6.4)$$

- (c) *Energía de Expansión*: el término de expansión permite especificar una tolerancia a potencial no-programado distinta para cada eje y viene dado por:

$$\mathcal{E}_{exp}(a, b) = \frac{\alpha_a}{\lambda_a} a^{\lambda_a} + \frac{\alpha_b}{\lambda_b} b^{\lambda_b} \quad (6.5)$$

2. *Dinámica por Gradiente*: teniendo en cuenta que la primitiva se inicializa a partir de un Círculo de Movimiento Libre la derivación en el tiempo de $\mathcal{E}(x, y, a, b, \vec{V}_{ref})$ determina la siguiente dinámica de ajuste:

- (a) *Dinámica del Centro*: al igual que en el caso circular, el ajuste del centro viene influido por la toma de contacto de la primitiva con el gradiente (frontera vascular):

$$\begin{pmatrix} \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial t} \end{pmatrix} = - \begin{pmatrix} \frac{\partial \mathcal{E}}{\partial x} \\ \frac{\partial \mathcal{E}}{\partial y} \end{pmatrix} = \int_{\mathcal{D}} (1 - (\frac{u^2}{a^2} + \frac{v^2}{b^2})) (-\vec{\nabla} \mathcal{I}) ds - 2(\mathcal{N}(\vec{V}_{ref}) \cdot \mathcal{N}(\vec{\nabla} \mathcal{I}(p_1))) (\mathcal{N}(\vec{V}_{ref}) \mathcal{N}(\vec{\nabla}^2 \mathcal{I}(p_1))) \quad (6.6)$$

El efecto del gradiente induce un centrado tanto a nivel global de la primitiva (primer término) como a nivel del punto p_1 .

- (b) *Dinámica de Rotación*: el cuadrado del producto escalar entre el vector de referencia y el de gradiente en p_1 estima la distancia a la orientación correcta (ortogonalidad):

$$\begin{pmatrix} \frac{\partial \vec{V}_{refx}}{\partial t} \\ \frac{\partial \vec{V}_{refy}}{\partial t} \end{pmatrix} = - \begin{pmatrix} \frac{\partial \mathcal{E}}{\partial \vec{V}_{refx}} \\ \frac{\partial \mathcal{E}}{\partial \vec{V}_{refy}} \end{pmatrix} = -2(\mathcal{N}(\vec{V}_{ref}) \cdot \mathcal{N}(\vec{\nabla} \mathcal{I}(p_1))) \mathcal{N}(\vec{\nabla} \mathcal{I}(p_1)) \quad (6.7)$$

- (c) *Dinámica Axial*: la variación de los ejes sigue el mismo esquema que para el modelo circular y en cada caso entran en juego los términos de ruido asociados a cada eje:

$$\frac{da}{dt} = - \frac{\partial \mathcal{E}}{\partial a} = - \iint_{\mathcal{D}} 2 \frac{u^2}{a^3} \mathcal{I}(x + u, y + v) du dv + \alpha_a a^{\lambda_a - 1} \quad (6.8)$$

$$\frac{db}{dt} = -\frac{\partial \mathcal{E}}{\partial b} = -\int \int_{\mathcal{D}} 2 \frac{v^2}{b^3} \mathcal{I}(x+u, y+v) du dv + \alpha_b b^{\lambda_b - 1} \quad (6.9)$$

produciéndose la convergencia cuando ambos niveles de ruido se compensen con el número de pixels admitidos.

El resultado de la aplicación del modelo a una secuencia de hinchado se muestra en la Figura 6.9 donde podemos observar diversas fases del proceso. Los niveles de ruido, seleccionados experimentalmente se han establecido en: $\alpha_a = 15$, $\alpha_b = 15$, $\lambda_a = 1$ y $\lambda_b = 1$. La inicialización previa a partir del modelo circular (con mayor tolerancia al ruido) ha permitido eliminar estructuras internas que induzcan un mal comportamiento de la elipse.

A partir de los buenos resultados obtenidos con este modelo nos planteamos generalizarlo para el caso de plantillas super-elipsoidales. Esto supuso incorporar un nuevo parámetro para modelar la forma del objeto y además modificar la fuerza de expansión. El primer aspecto se resolvió de forma directa a partir del modelo elíptico: incluimos en la dinámica la variación debida al parámetro de forma. La modelización del ruido cambió con respecto al planteamiento aditivo del modelo elíptico debido a la diferencia de escala entre el parámetro asociado al tamaño y el de forma. Todo ello dió lugar al modelo que hemos presentado en el Capítulo 3.

6.1.5 Análisis Espacio-Temporal: Auto-Restricciones

La aplicación de una plantilla deformable a este tipo de secuencia nos proporciona una gran cantidad de información. En la Figura 6.10 se muestra la evolución del radio del círculo deformable utilizado para realizar el seguimiento de la pared interior de la arteria. El eje de ordenadas refleja el número de la secuencia y el de abscisas el tamaño del radio. A partir de la gráfica extraemos las siguientes interpretaciones:

1. *Valle Inicial y Crecimiento*: asociamos el mínimo que comienza en el *frame* número 10 a la presencia del ateroma. A partir de este punto arranca el proceso de hinchado en el que se identifica claramente una tendencia de crecimiento radial hasta un punto de estabilización (*frames* 16 al 40).
2. *Picos Intermitentes*: a partir del *frame* 40 la angioplastia ha terminado y los picos intermitentes observados corresponden al latido cardíaco: teniendo en cuenta un período inter-frame de 0.2 segundos y que el pulso de latido se repite cada 4 frames se estima un período de 0.8 segundos y en consecuencia una frecuencia de latido de 75 pulsaciones por minuto.

Esta circunstancia motiva el desarrollo de esquemas de análisis temporal de la estructura geométrica extraída durante el proceso de *tracking*. En este sentido deben proporcionarse estrategias de aprendizaje y categorización de pautas de evolución que permitan filtrar los picos intermitentes. El conocimiento de experto acerca del tipo de evolución deseable puede instrumentarse mediante ACP. Como ejemplo ilustrativo hemos establecido tres *restricciones espacio-temporales ideales* que debe satisfacer un hinchado correcto: $r_a/r_0(t)$, $r_b/r_0(t)$, $r_a/r_b(t)$ (ver Figura 6.11). Las dos primeras relaciones expresan la evolución promedio de cada eje (el promedio permite filtrar el ruido debido a los pulsos cardíacos) normalizada por el radio interno r_0 . La tercera restricción se refiere a la evolución relativa de los radios (estima la excentricidad) también promediada. Por lo tanto se tiene preferencia por hinchados uniformes a velocidad moderada (pendiente media). Hemos aprendido estas restricciones y las hemos expresado con un número reducido de modos de variación, capturando el 90% de la variabilidad. En la Figura 6.12 presentamos resultados de aceptación (aplicando la distancia de Mahalanobis) para $r_a(t)/r_0$ dado un conjunto de prototipos generados a partir de la variación de la pendiente.

El modelo de restricciones espacio-temporales presentado en el Capítulo 4 viene motivado por la necesidad de disponer de restricciones de este tipo. Su formalización y posterior generalización para el tratamiento de múltiples plantillas arranca de esta aplicación y permite abordar otras como la que presentaremos a continuación.

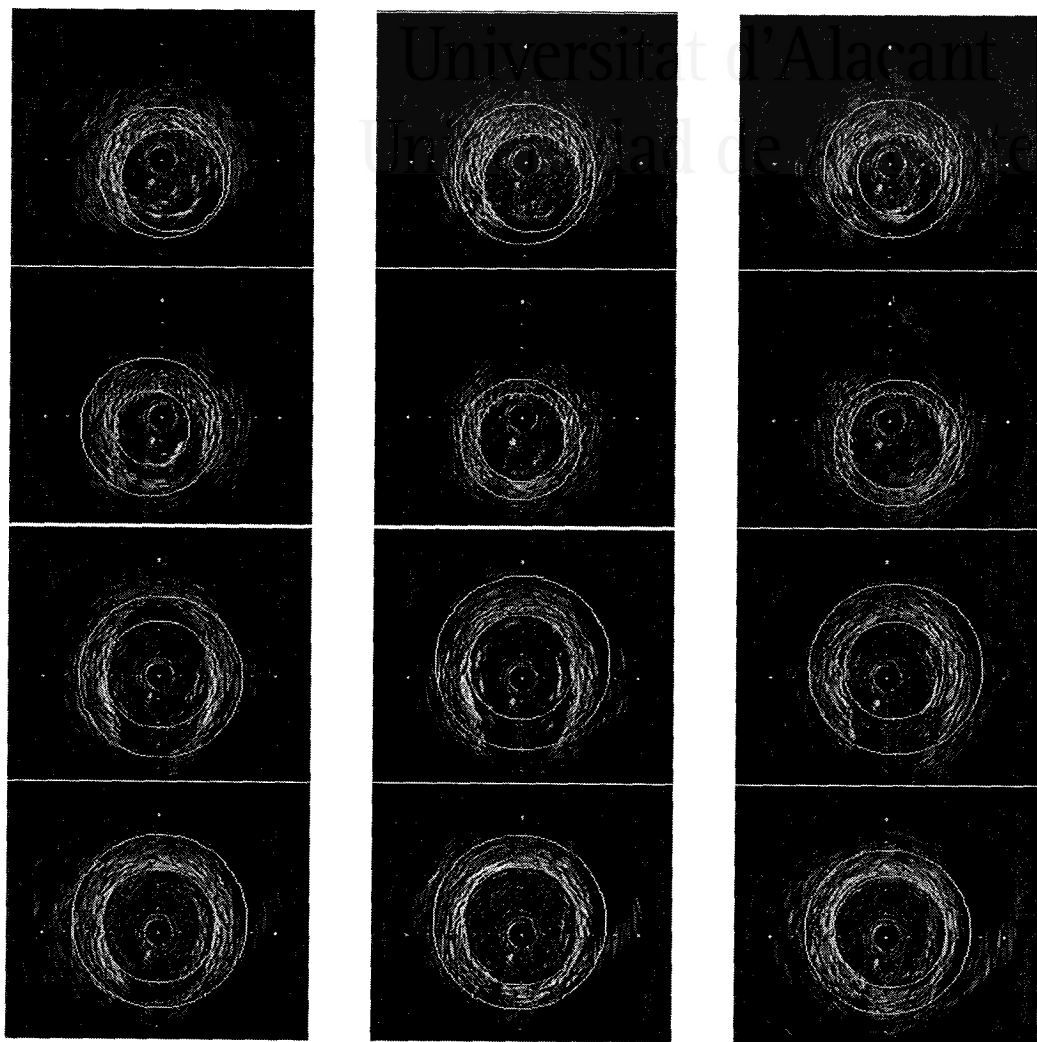


Figura 6.8: Resultados del *tracking* en la secuencia con las dos plantillas circulares (interna y externa).

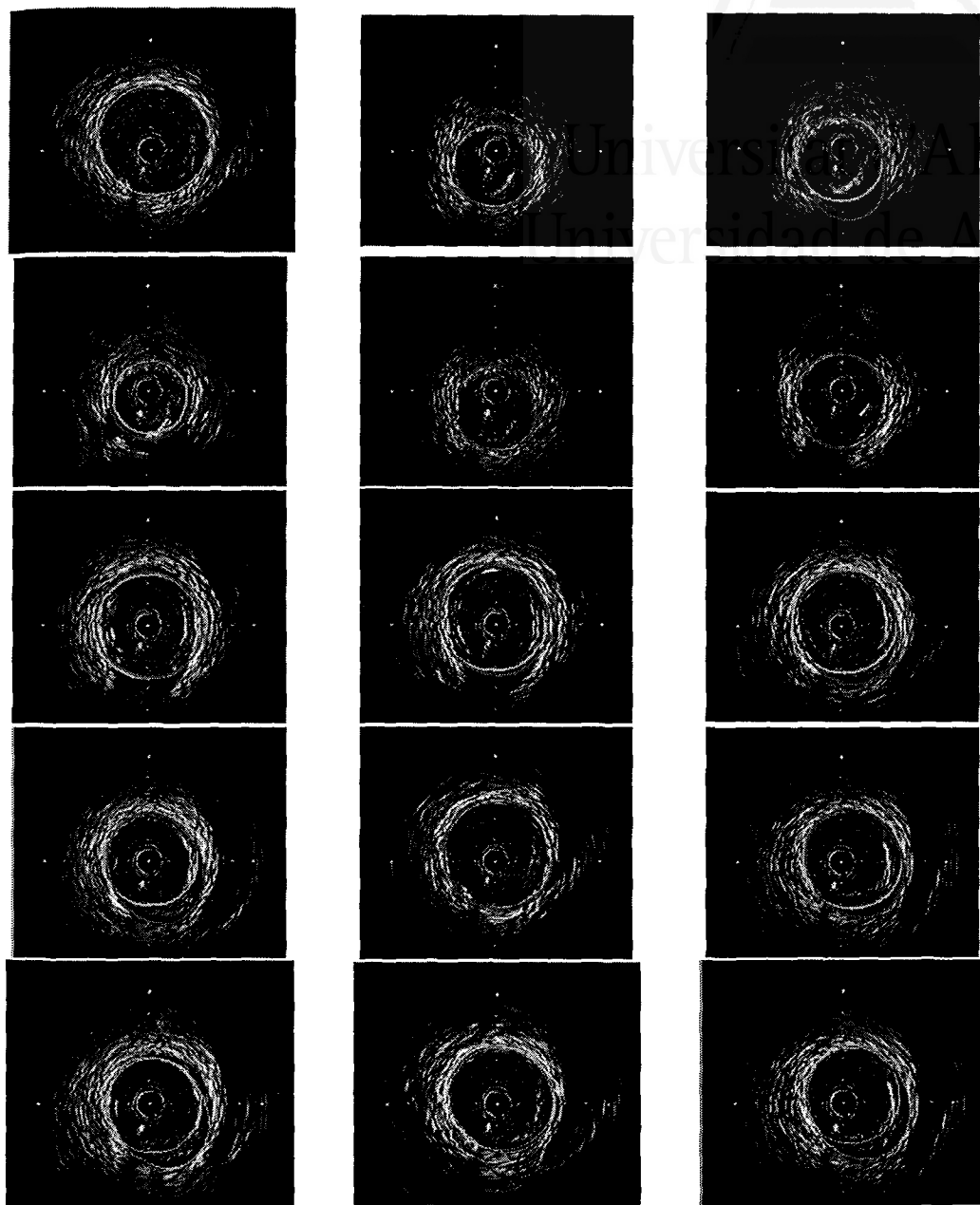


Figura 6.9: Resultados del *tracking* en la secuencia de hinchado

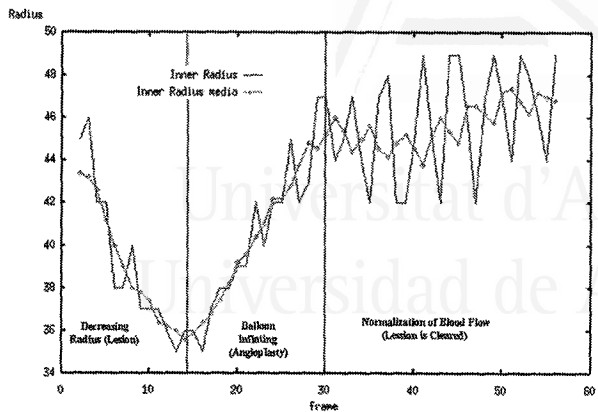


Figura 6.10: Evolución del radio interno.

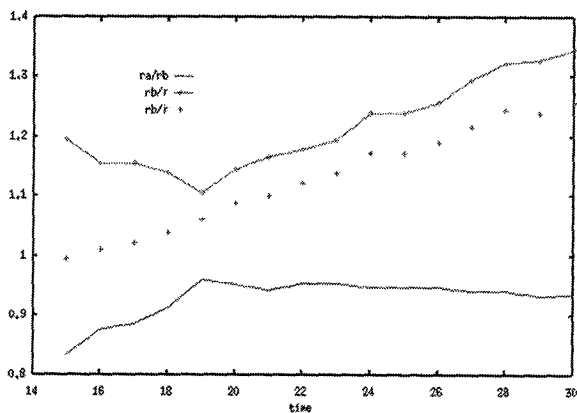


Figura 6.11: Restricciones ideales para el reconocimiento.



Universitat d'Alacant
Universidad de Alicante

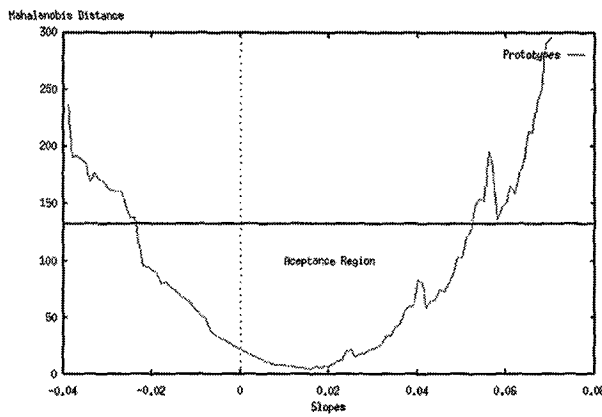


Figura 6.12: Resultados experimentales para $r_a/r_0(t)$.

6.2 Reconocimiento Basado en Tracking Gestual

6.2.1 Introducción y Motivación

El problema del *reconocimiento a partir del movimiento* (Cédras y Shah, 1995), ha cobrado un fuerte interés sobre todo en el ámbito de la interpretación de gestos humanos. Al igual que sucede en otras problemáticas, la Computación ha tomado como referencia experimentos propios de la psicología cognitiva (Johansson, 1973, 1975), (Cutting y Proffitt, 1982). El *tracking* de configuraciones geométricas es, junto con la obtención de *modelos de flujo óptico* (Polana y Nelson, 1994) y de *conjuntos de vistas o de frames clave* (Akita, 1984), (Darrel y Pentland, 1993), (Darrell et al., 1996) uno de los métodos más utilizados en la extracción de información espacio-temporal con vistas a la caracterización y reconocimiento de la secuencia. Estos modelos suelen extraer un modelo de segmentos (*sticks*) (Hogg, 1984) (Rohr, 1994). El reconocimiento a partir de configuraciones geométricas se aborda desde diversos enfoques:

1. *Clustering/Correlación*: en base a características espacio-temporales como la curvatura (Allmen, 1991), la velocidad lineal o angular y las discontinuidades de la trayectoria (Gould y Shah, 1989), o bien mediante cambios de dominio (Campbell y Bobick, 1995).
2. *Arquitecturas Conexionistas*: integración de la variación de forma y movimiento a distintos niveles de abstracción (Goddard, 1992) para el análisis de eventos y la gestión de *escenarios o secuencias interpretadas de cambio*.
3. *Métodos Simbólicos y Cognitivos*: tales como el enfoque de reconocimiento cualitativo (Kuniyoshi et al., 1994), sobre todo en el *aprendizaje visual de tareas*², el uso de *restricciones estructurales* (O'Rourke y Badler, 1980), la gestión de redes semánticas (Tsotsos et al., 1980), la caracterización mediante autómatas (Davis y Shah, 1994) de un conjunto reducido de vocablos del ASL³ o el uso de modelos de Markov ocultos (Yamato et al., 1992).

En general, aunque la información espacio-temporal extraída es más precisa, la aplicabilidad a escenas reales de los esquemas basados en *tracking paramétrico* está limitada ya que la segmentación y adquisición de información geométrica es un proceso costoso que requiere disponer, *a priori*, de conocimiento de alto nivel

² *Teaching by Showing y Learning by Watching*: se trata de que un robot aprenda a generar planes ejecutables y reusables de ensamblaje o manipulación de objetos a partir de una secuencia de imágenes generada por un instructor humano.

³ *American Sign Language*: gestos standard para comunicación con sordomudos.

acerca del fenómeno para simplificar la segmentación y además exige frecuentemente inicialización manual, por lo que disponer de un mecanismo de aprendizaje previo sería de gran utilidad. Dicho mecanismo puede aportarlo el ACP. De hecho esta metodología se ha aplicado habitualmente sobre modelos de vistas (Kirby et al., 1993), (Murase y Sakai, 1996). El desarrollo de los Modelos de Distribución de Puntos ha permitido incorporar ACP a los modelos deformables derivándose esquemas de caracterización de gestos faciales y/o manuales (Lanitis et al., 1994). No obstante dado que el problema se aborda a un nivel de abstracción bajo (puntos de contorno) no se hace uso de información relacional o estructural. En este sentido el uso de Plantillas Extendidas nos va a permitir aprender, extraer y reconocer estructuras geométricas haciendo uso de dicha información.

6.2.2 Reconocimiento mediante Plantillas Extendidas

Caracterizaremos una estructura en movimiento mediante un *modelo gestual* $\mathcal{M}_i = \{\mathcal{G}_i(t), \mathcal{R}_i(\vec{\zeta}, t)\}$ formado por un conjunto de plantillas dependientes del tiempo definidas sobre la ventana $[0, T]$ y sus restricciones asociadas. Dada una serie de modelos gestuales $\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_p$, reconocer una *secuencia de entrada* \mathcal{S} consiste en encontrar el modelo \mathcal{M}_j para el cual se cumple:

$$\forall \mathcal{M}_i \mid i \neq j : \text{Sat}(\mathcal{G}_i^*(t), \mathcal{R}_j(\vec{\zeta}, t)) \geq \text{Sat}(\mathcal{G}_i^*(t), \mathcal{R}_i(\vec{\zeta}, t)) \quad (6.10)$$

Es decir, aquel cuya estructura espacio-temporal óptima, derivada de aplicar el modelo de restricciones mediante *tracking*, satisface en mayor grado los condicionantes geométricos de dicho modelo que los de los demás modelos gestuales considerados. En este sentido la aplicación del mecanismo de seguimiento debe venir regido por las mismas condiciones en cuanto a inicialización, iteraciones predictivas, parámetros de ruido y coeficiente de ponderación entre modelo local y estructural⁴. En la Figura 6.13 se presenta un ejemplo de secuencia con dos modelos gestuales \mathcal{M}_1 y \mathcal{M}_2 . La estructura viene caracterizada por $g = 3$ plantillas elípticas (una sobre la cabeza y dos sobre las manos). La única diferencia entre ambos modelos es la evolución tipo de la mano derecha que en \mathcal{M}_1 es parabólica (trazo claro) y en \mathcal{M}_2 es rectilínea (trazo oscuro). A continuación especificaremos el modelo de imagen y aplicaremos el esquema de descenso por ventana única descrito en el Capítulo 5 para el reconocimiento.

⁴Previamente debe aplicarse, si las ventanas de la entrada y la de cada modelo son distintas, un modelo de desplazamiento y deformación temporal o *Time Warping* para que el resultado sea comparable.

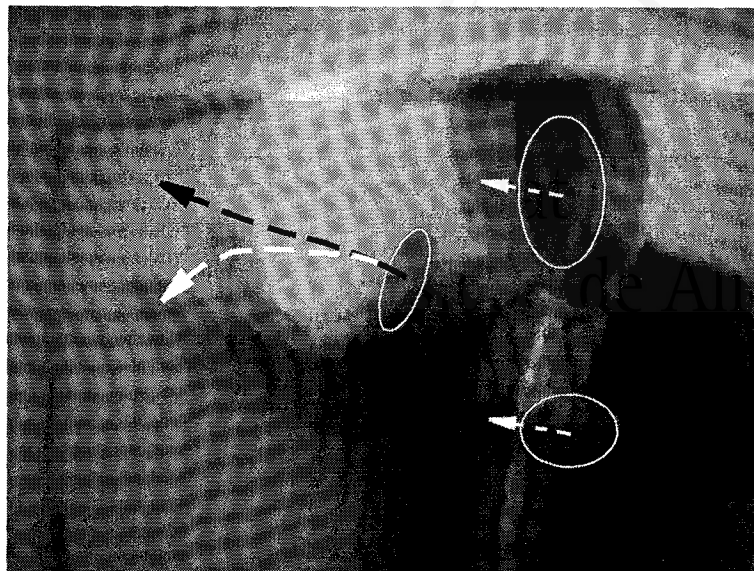


Figura 6.13: Planteamiento: Modelos Geométricos de Movimiento

6.2.3 Modelo de Imagen: Campos de Potencial

Dada la situación de las plantillas sobre la imagen ejemplo, el potencial debe extraer la textura o intensidad asociada a la piel. El proceso seguido (ver Figura 6.14-esquina superior izqda.) ha consistido en aplicar un umbralizado a dos niveles (intensidades superiores e inferiores a la asociada a la piel) seguido de una secuencia de filtros morfológicos (tres erosiones seguidas por dos dilataciones) para eliminar ruido aislado asociado a efectos de iluminación. En la Figura 6.14 ($T = 10$) se muestran los frames de la secuencia tipo utilizada en la experimentación.

6.2.4 Tracking Estructural y Reconocimiento

En el aprendizaje de modelos se ha fijado como constante (salvo errores de discretización) tanto la escala como la forma de las plantillas. Se ha especificado como variable la orientación y posición. Las distribuciones obtenidas son muy específicas para permitir discriminar entre ambos modelos y de baja entropía de restricción. Para ilustrar el comportamiento del método de reconocimiento hemos realizado un experimento de reconocimiento en el que la secuencia de entrada (ver Figura 6.14) se adapta mejor al modelo \mathcal{M}_1 (mano derecha con trayectoria parabólica). Para ello nos basaremos únicamente en restricciones absolutas. Par-

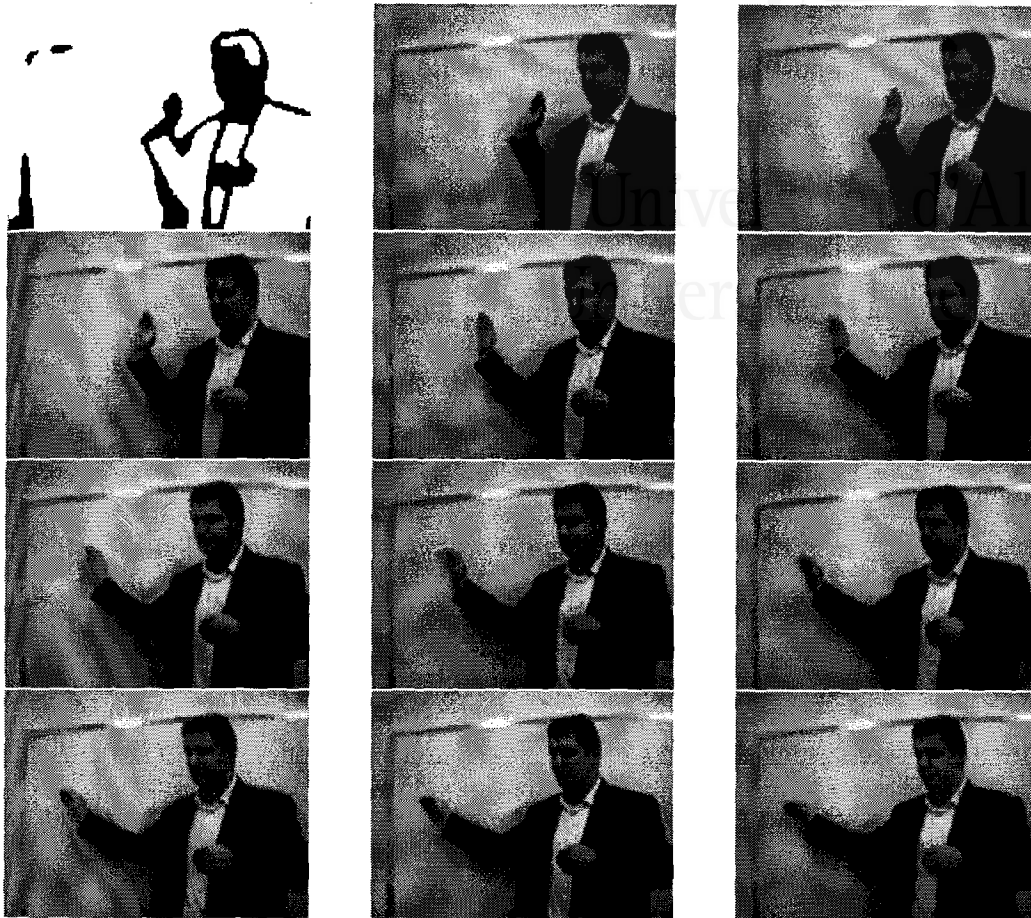


Figura 6.14: Potencial y Frames 0-10

tiendo de condiciones homogéneas de experimentación: $\lambda = 0.2$, $\alpha = \beta = 0.1$, $p = 5$, todas las restricciones a nivel 0.9 y 20 iteraciones de ajuste por frame obtenemos los siguientes resultados:

1. *Ajuste correcto*: la secuencia de entrada se ajusta, en tamaño y orientación espacio-temporales (ver Figura 6.15-arriba), al modelo \mathcal{M}_1 . El descenso por ventana única permite salvar discontinuidades temporales importantes (paso del *frame* 3 al 4) gracias al mecanismo predictivo y al conocimiento adquirido. El valor fijado para λ permite un buen ajuste a dicho modelo teniendo en cuenta el potencial de intensidad.

2. *Ajuste incorrecto*: en cambio la aplicación del segundo modelo (ver Figura 6.15-abajo) aporta un buen comportamiento en las plantillas con una dinámica común. Sin embargo el ajuste de la plantilla asociada a la mano derecha pierde fiabilidad a partir del tercer *frame*, produciéndose una progresiva reducción de la escala debida a la ausencia de potencial y que las restricciones del modelo no pueden salvar como ocurre en el caso de la restricción angular.

Este mismo esquema puede aplicarse para el *control de calidad mediante monitorización y reconocimiento de acciones*. Se trata de aprender el modelo geométrico vinculado a la apariencia geométrica de una acción (movimiento de objetos, fabricación, etc.) y de implementar su seguimiento mediante un mecanismo de *tracking* de plantillas. Una anomalía en el proceso se corresponde con un descenso, por debajo de un umbral preestablecido, del grado de satisfacción de las restricciones del modelo asociado a dicha acción.

6.3 Consideraciones

Los experimentos presentados en este capítulo completan la experimentación que acompaña a la formulación de los modelos y técnicas aportados. En el capítulo siguiente presentaremos nuestras conclusiones y trazaremos las líneas de continuación, tanto en el aspecto formal como en el aplicado, de esta tesis.



Universitat d'Alacant
Universidad de Alicante

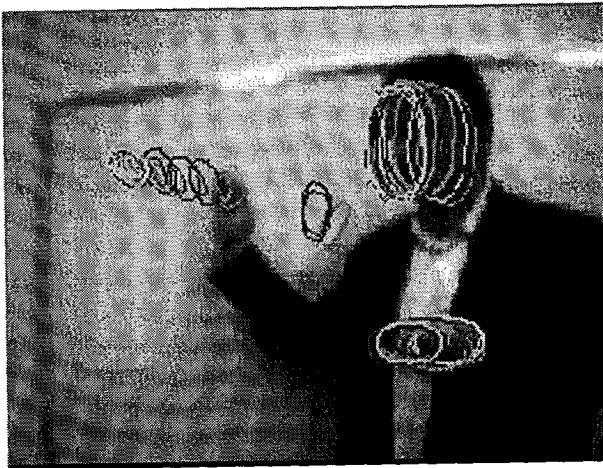
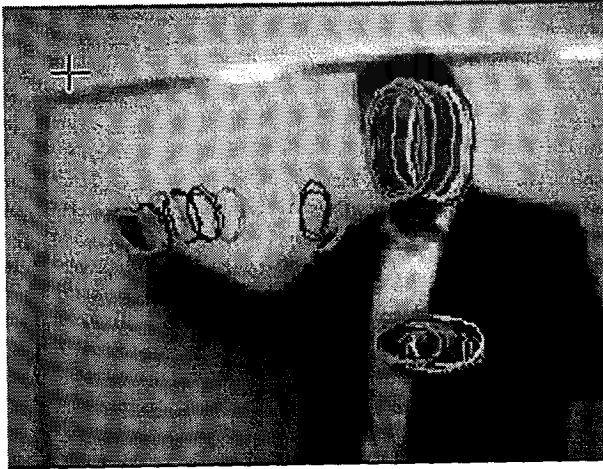
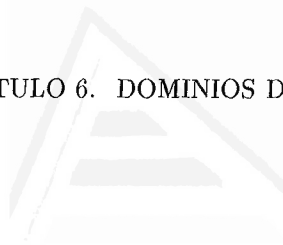


Figura 6.15: Ajuste a los modelos \mathcal{M}_1 (arriba) y \mathcal{M}_2 (abajo).



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Capítulo 7

Conclusiones y Trabajos Futuros

En esta tesis hemos aportado mejoras en el tratamiento del problema de la correspondencia y *tracking* a partir de modelos deformables. Dichas mejoras se enmarcan tanto en el plano teórico (modelos de representación y esquemas computacionales) como en el aplicado (tratamiento de secuencias intravasculares y reconocimiento de gestos). El planteamiento de la disertación se ha realizado de forma constructiva. En primer lugar se ha definido el modelo local de plantilla, a continuación se ha caracterizado, espacial y espacio-temporalmente la estructura y, posteriormente se ha propuesto un esquema de ajuste que integra ambos aspectos. Finalmente hemos extendido dicho esquema y lo hemos completado, dotándolo de un mecanismo predictivo. Todos estos aspectos se han tratado ampliamente a nivel experimental presentándose los resultados más representativos.

7.1 Aportaciones

En cuanto a *modelos de representación y algoritmos de búsqueda para Visión Activa* nuestra propuesta, el Modelo de Plantillas Deformables Extendidas, contiene aportaciones a diversos niveles:

1. *Modelización Local-Control del Ruido*: diseño de una función de energía capaz de parametrizar el nivel de ruido tolerado por el modelo. Modelización de formas bidimensionales con una capacidad expresiva razonable (superelipsoides).
2. *Caracterización Estructural-AutoRestricciones*: incorporación de un mecanismo probabilístico (ACP) capaz de aprender condicionantes geométricos

y de obtener caracterizaciones compactas. Énfasis en el aspecto estructural tanto a nivel espacial como espacio-temporal.

3. *Correspondencia por Minimización Ponderada*: propuesta de un esquema algorítmico de minimización por supervisión que combina el modelo de imagen con la satisfacción de condicionantes geométricos estructurales previamente adquiridos. Integración de la robustez del modelo local con la versatilidad del modelo estructural.
4. *Incorporación de Mecanismos Predictivos*: extensión del mecanismo de correspondencia al caso temporal e incorporación de mecanismos predictivos para la resolución del problema del *tracking* de contornos.

Desde un punto de vista aplicado esta propuesta aporta resultados interesantes en:

1. *Análisis de Secuencias Intravasculares*: a partir de modelos circulares y elípticos es posible realizar reconstrucciones aceptables de la pared intravascular y caracterizar procesos quirúrgicos como la angioplastia coronaria.
2. *Reconocimiento de Gestos*: el *tracking* de plantillas asociadas a partes del cuerpo permite reconocer pautas de evolución (gestos) encontrando aquel modelo cuyos condicionantes estructurales se satisfacen en mayor grado por la secuencia de entrada.
3. Por extensión en problemas similares tales como reconstrucción de órganos complejos o el análisis del cansancio ocular.

7.2 Líneas de Continuación

Este trabajo supone un punto de partida para el desarrollo de nuevos modelos y algoritmos en Visión Activa:

1. *Estrategias de Búsqueda para Plantillas*: desarrollo de técnicas eficientes de localización de objetos (exploración selectiva) para obtener inicializaciones adecuadas de la búsqueda por descenso supervisado.
2. *Algoritmos de Extracción de Profundidad*: tomar como referencia el mecanismo de correspondencia obtenido para abordar el problema de la reconstrucción 3D a partir de pares estéreo o bien de *frames* consecutivos.

3. *Auto-Estructuras*: caracterización de grafos relacionales y estructuras relacionales mediante representaciones modales (especialmente auto-restricciones). Definición de representaciones compactas y de algoritmos de búsqueda heurística que permitan agilizar el tratamiento del problema del reconocimiento.
4. *Paralelización*: instrumentación paralela de algoritmos de *tracking* capaces de explotar la independencia de las restricciones de distinta naturaleza.

En cuanto a las problemáticas a abordar en el futuro las dos más interesantes son:

1. *Navegación Basada en Visión Activa*: la asignación de plantillas deformables a marcadores visuales y la caracterización, mediante ACP, de la variación de su apariencia en función del desplazamiento de una cámara permitirá abordar problemas de localización de robots móviles así como reconocer situaciones tipo.
2. *Análisis de Acciones*: por otro lado la formulación del reconocimiento de gestos mediante plantillas es interesante tanto para el diseño de mecanismos de *control de calidad basado en visión* como en la implementación de *herramientas de realidad virtual* capaces de traducir los gestos interpretados a situaciones de un mundo virtual.

7.3 Consideraciones Finales

El desarrollo de modelos de representación activos es un campo de trabajo relativamente joven (unos diez años) y en actual efervescencia dada la necesidad de disponer de mecanismos eficientes que permitan tratar, mediante técnicas de Visión Activa, problemas de tratamiento de secuencias, que hace sólo algunos años eran inabordables. Por supuesto queda mucho camino por recorrer pero ya se producen los primeros resultados prácticos y, poco a poco, la separación entre modelos locales basados en funciones de energía (heurísticas, al fin y al cabo) y los algoritmos globales basados en *matching* de grafos y otros esquemas algorítmicos se va reduciendo. Sin duda el tratamiento de problemas desde puntos de vista que integren ambos aspectos, bajo el prisma del conocimiento probabilístico previamente adquirido, proporcionará interesantes frutos en los próximos años.



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Parte IV

Apéndices



Universitat d'Alacant
Universidad de Alicante

Apéndice A

Experimentación con Contornos Activos

En este capítulo vamos a mostrar algunos de los dominios donde hemos introducido los Contornos Activos como herramienta para la resolución de problemas, en su mayor parte imágenes biomédicas. En el campo de la biomedicina existen bastantes problemáticas aún por resolver. En concreto nos centramos en facilitar las tareas del personal especializado en el momento de realizar un diagnóstico a partir de imágenes de órganos, obtenidas mediante fotografías microscópicas, resonancias magnéticas, ecocardiografías, por citar algunos ejemplos. Incorporando el conocimiento de experto en modelos geométricos activos e interactuando con el campo de potencial de la imagen se pretende sentar las bases de una línea de investigación orientada al desarrollo de modelos deformables robustos. Los resultados, a partir de Contornos Activos, contenidos en este capítulo sentaron las bases del desarrollo del Modelo de Plantillas Deformables Extendidas.

A.1 Reconstrucción Basada en Contornos Activos

Una problemática que se viene planteando en el contexto de la biomedicina es la necesidad de obtener un modelo o una visión tridimensional de un órgano a partir de secciones/cortes obtenidos mediante Resonancia Nuclear Magnética (RNM), imágenes microscópicas y otras técnicas similares. La obtención de este modelo nos puede ser útil en el reconocimiento y/o detección de deformaciones o tumores en el órgano. También se nos plantea la necesidad de obtener un modelo tridimensional cuando realizamos cálculos volumétricos.

A.1.1 Reconstrucción Cerebral a partir de Imagen RNM

Objetivo En base a una secuencia de imágenes obtenidas mediante RNM, se pretende realizar la reconstrucción 3D de la corteza cerebral y crear así un modelo que pudiera servir de base para la detección automática de malformaciones craneales y/o presencia de tumores.

Problemáticas

- La utilización de imágenes RNM, Figura A.1, implica un alto grado de distorsión, implícito en la propia técnica de obtención de las imágenes. Debido a ello se tuvo que realizar un estudio profundo de la respuesta del tejido cerebral a las resonancias.

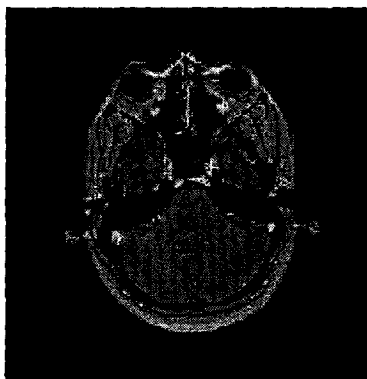


Figura A.1: Sección del cerebro obtenida mediante RNM

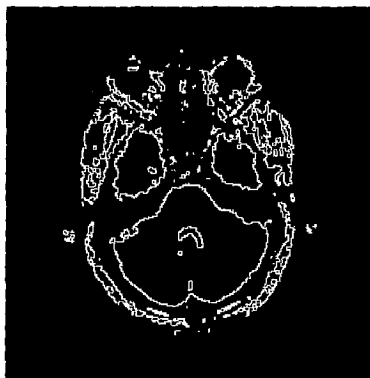


Figura A.2: Campo de potencial obtenido.

El análisis de la imagen tuvo como resultado la aplicación de un umbralizado global, teniendo en cuenta los valores de nivel de gris de respuesta a la resonancia del tejido cerebral. Una vez realizado este umbralizado se aplicaron técnicas de filtrado, máscaras de convolución y operadores morfológicos (apertura, cierre, dilatación, erosión) hasta la obtención de una imagen como la Figura A.2, la cual iba a servir como campo de potencial para la adaptación del *snake*.

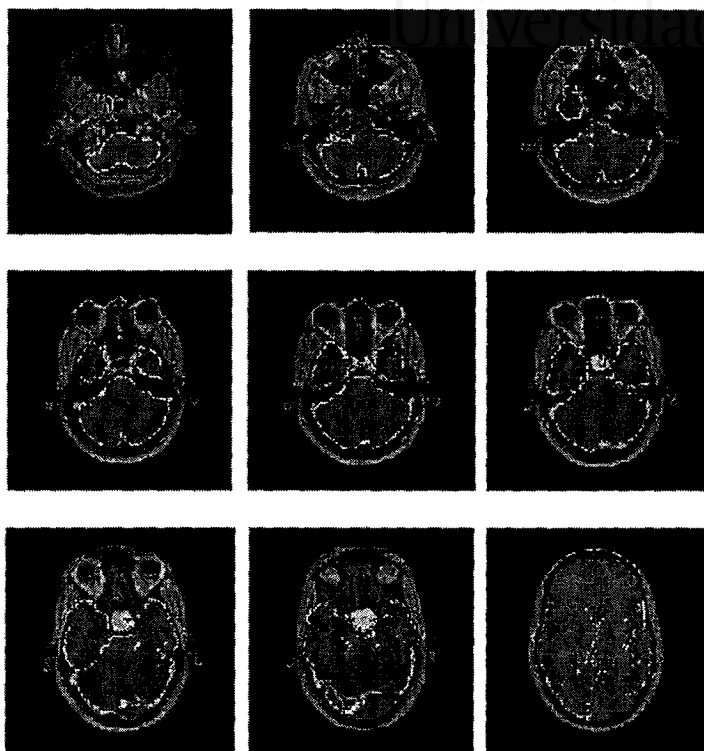


Figura A.3: Seguimiento que se realiza sobre las regiones de interés.

- Uno de los principales problemas que surgen en este tipo de sistemas orientados a reconstrucción es la interacción hombre/máquina. Pretendíamos disminuir lo máximo posible esta interacción. Para ello necesitábamos un modelo que permitiera adaptarse a los cambios en la forma de la región sobre la que pretendíamos realizar el seguimiento. Los *snakes* proporcionaban este seguimiento. El problema que se nos plantea en este modelo es la inicialización. Si el *snake* se inicializa lejos de la superficie de potencial deseada, tendremos problemas en la adaptación. Para ello permitimos que el

usuario seleccione aquellas regiones sobre las que se realizará el seguimiento, indicando además en que instante de tiempo aparecen.

- *Aparición de Nuevas Regiones:* Como se puede observar en la Figura A.3, en el caso del cerebro la región sobre la que queremos realizar el seguimiento en un principio no es única, sino que existe una región inicial y a lo largo de la secuencia van apareciendo el resto. Esto es necesario tenerlo en consideración si nuestro objetivo es un seguimiento semiautomático. En principio permitíamos que fuera el usuario el que indicara cuales eran las regiones de interés y en que sección aparecían. Actualmente estamos trabajando en una modelización geométrica de la secuencia que permita, mediante marcadores morfológicos disponer de dicha información.
- *Fusión de Regiones:* Como antes comentábamos, en principio vamos a realizar el seguimiento de varias regiones. Estas regiones acaban fusionándose en una. Para ello debíamos incorporar al modelo de *snake* una posibilidad de unión. Esto se llevó a cabo detectando el instante en que dos *snakes* se encontraban a una distancia inferior a un cierto umbral (lo cual implicaba su fusión dando lugar a un contorno mayor), tal como se muestra en la Figura A.4.

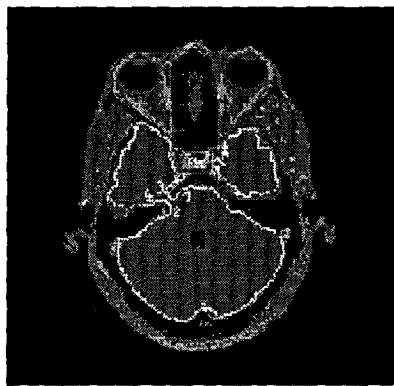


Figura A.4: Fusión de dos *snakes* en uno.

- *Aparición de Mínimos Locales:* Dado que en las secciones aparecen distintos tipos de materia, se origina un excesivo número de contornos que potencialmente pueden atraer a los *snakes* situados sobre las zonas de interés. Además el límite de algunas zonas no está bien definido y el contorno resultante es abrupto. Cuando calculamos el término de continuidad de ese

contorno aparecen mínimos locales entre los puntos a los que se puede expandir los *snaxels* a estudio. La presencia de mínimos locales puede hacer que el *snake* detenga su expansión en dichos puntos de forma incorrecta. Para resolver esta cuestión proponemos realizar un filtrado de los puntos vecinos con mínimos locales.

Experimentación La secuencia de imágenes de la que disponíamos constaba de 74 secciones, tomadas a intervalos regulares. En la Figura A.3 se muestra el resultado de aplicar la técnica de los *snakes* a la secuencia, donde en las últimas imágenes se puede observar los problemas aparecidos. Las Figuras A.5 y A.6 muestran la reconstrucción obtenida a partir de los *snaxels*.

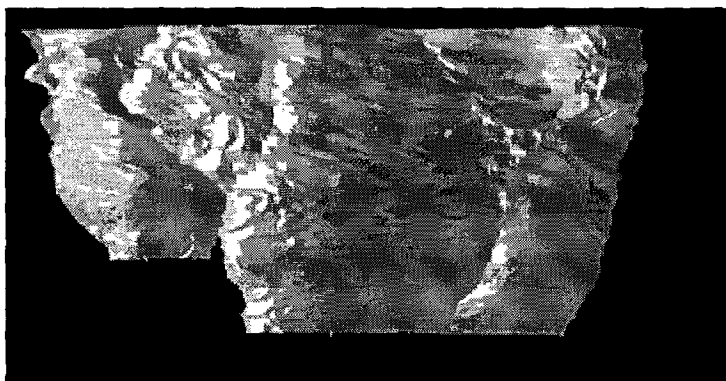


Figura A.5: Vista posterior del cerebro reconstruido.

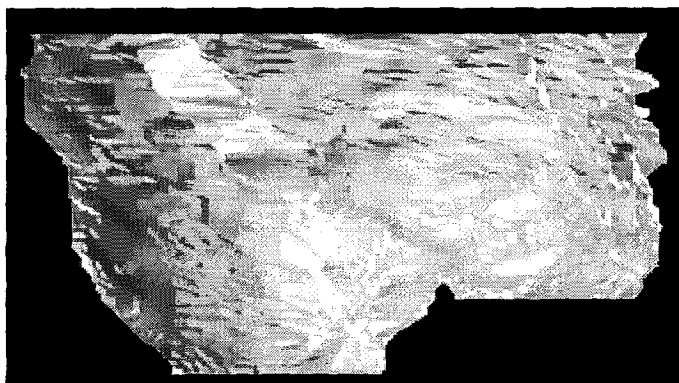


Figura A.6: Vista lateral del cerebro reconstruido.

Conclusiones Debido a la elevada complejidad de las imágenes cerebrales, la determinación y el seguimiento de las características relevantes de las distintas materias es un problema muy complejo. En este sentido hemos abordado el estudio de las diversas problemáticas que surgen en el seguimiento de contornos cerebrales: mínimos locales, discontinuidad espacial, aparición de nuevas regiones y fusión de regiones.

Hemos obtenido, mediante una experimentación exhaustiva, muy buenos resultados en el seguimiento de zonas independientes. Sin embargo queda por resolver el seguimiento coherente y controlado de regiones fusionadas. Dicho problema podría resolverse diseñando un esquema de ponderación adecuado para el *snake* unión e incorporando conocimiento adicional en los términos de la función de energía, incluyendo aspectos como el coloreado de contornos, de forma que cada *snake* sea sensible a un determinado tipo de contorno y no a otro. Con esto solucionaríamos el problema de la atracción, en las etapas previas a la fusión, hacia contornos no deseados.

Por lo tanto nuestras investigaciones futuras en este problema tomarán fundamentalmente dos direcciones:

- Profundizar en la extracción de campos de potencial fiables mediante la incorporación de conocimiento adicional (*snakes/modelos deformables basados en conocimiento probabilístico*) extraído de forma automática: conocimiento sobre la textura (tipo de material de cada región), la morfología (curvatura, concavidades, convexidades) y la organización (inclusión, conexión) de las distintas regiones (estructura 2D). En este sentido se está trabajando en un esquema adecuado de clasificación y en la definición de un modelo estructural de objeto.
- Así mismo profundizaremos en la extracción automática de las condiciones de disparo de los *snakes*. Estamos trabajando en la definición de un modelo de evolución de campos de potencial para detectar aspectos como el disparo de nuevos *snakes*, el crecimiento/decrecimiento del número de *snaxels* y la variación de los factores de ponderación a lo largo de la secuencia. Dicho modelo está basado en la evolución relativa de un conjunto marcadores o puntos característicos que contienen información cualitativa y cuantitativa sobre la textura y morfología de la región a la que representan (estructura 3D).

A.1.2 Reconstrucción Ocular a partir de Imagen Microscópica

Objetivo Pretendemos obtener un modelo 3D del globo ocular de una salamandra, partiendo de una secuencia de imágenes obtenidas mediante una fotografía microscópica de cortes transversales obtenidos de forma manual. En la actualidad los morfólogos realizan, manualmente también, un recuento de los fotorreceptores localizados en el interior de la región que se muestra en la Figura A.7. El objetivo principal en este trabajo era segmentar la zona del globo ocular para, a continuación, obtener datos volumétricos del globo y realizar un cálculo aproximado de la densidad de fotorreceptores contenidos en la retina.

Problemáticas

- Los cortes se realizan de forma manual. No se puede garantizar una continuidad temporal homogénea de la secuencia para que pueda ser tratada con un modelo de *snake* de ventana de búsqueda constante.
- Se realiza un tintado de las láminas, el cual se lleva a cabo también de forma manual, lo que implica una presencia importante de ruido y anisotropía (no-homogeneidad) tal como se puede observar en la Figura A.7.
- Para la segmentación de la región del glóbo ocular se debe realizar un estudio más profundo de las distribuciones de las texturas que aparecen en la región de interés.



Figura A.7: Imagen de un corte del glóbulu ocular de la salamandra.

Experimentación El primer paso consistió en obtener la superficie de potencial sobre la que debería adaptarse el *snake*. Esto no es sencillo debido a la presencia de gran cantidad de ruido debido al tinte incorporado. Los pasos seguidos para la obtención final de la superficie de potencial son los siguientes:

1. Reducir a tres los niveles de gris de la imagen: bajo, medio y alto. Para cada pixel se calcula la media obtenida a partir de los valores de gris de los pixels vecinos dentro de una ventana centrada en él. Este valor de media determina la clasificación del pixel en uno de los tres niveles. Los valores utilizados para cada nivel son:

$$\{0 - 135\} \rightarrow \text{bajo}$$

$$\{136 - 186\} \rightarrow \text{medio}$$

$$\{187 - 255\} \rightarrow \text{alto}$$

El resultado de aplicar esta umbralización se muestra en la Figura A.8.



Figura A.8: Aplicación de umbralización.

2. A continuación eliminamos los pixels de fondo de la imagen para aumentar la rapidez del sistema. La homogeneidad del fondo permite su eliminación, simplificando con ello los cálculos. Esta eliminación la realizamos utilizando un rellenado por semilla desde el borde. También aplicamos un suavizado a la imagen. El resultado se puede apreciar en la Figura A.9.
3. Por último estudiamos la probabilidad de que un pixel pertenezca a uno de los contornos buscados: el interior al globo ocular y el externo. Para ello centramos una ventana de búsqueda en cada pixel y calculamos la probabilidad de que un pixel pertenezca al contorno externo o al interno,



Figura A.9: Eliminación de los pixels del fondo.

obteniendo dos contornos diferenciados cada uno de los cuales atraerá a un *snake* distinto. La diferencia entre los dos *snakes* nos proporciona el área de la región del glóbulo ocular. Un ejemplo del contorno interior se muestra en la Figura A.10.



Figura A.10: Contorno interior del glóbulo ocular.

Conclusiones Los modelos de contornos activos no son válidos para tratar directamente imágenes donde nos encontramos con la presencia de ruido o texturas complejas. El preprocesamiento de la imagen para la utilización correcta de estos modelos debe ser un paso importante y se debe realizar cuidadosamente. La superficie de potencial sobre la que actuará el *snake* debe estar libre, en la medida de lo posible, de atractores que induzcan mínimos locales. En esta aplicación hemos visto que un análisis, en este caso muy básico, de texturas puede ser útil para la obtención de una buena superficie de potencial.

Nuestros trabajos futuros se centrarán en:

- Una vez obtenida el área de la región de interés nos queda realizar una aproximación del número de fotorreceptores contenidos en una pequeña región dentro de la misma. De esta manera obtendríamos información volumétrica del número de fotorreceptores en el globo ocular.
- En combinación con los neurocientíficos se debería encontrar alguna herramienta que mejorara la calidad en la obtención de las imágenes.

A.2 Análisis del Cansancio Ocular Basado en Snakes

Objetivo En este trabajo tratamos la detección del cansancio ocular en la conducción haciendo uso de los modelos de Contornos Activos como herramienta de seguimiento. La detección del cansancio en la conducción se puede basar en la frecuencia con la que se realiza el parpadeo: un parpadeo más lento indica un estado de somnolencia. Calcular la frecuencia de parpadeo consistirá en detectar cuando se cierran y abren los ojos.

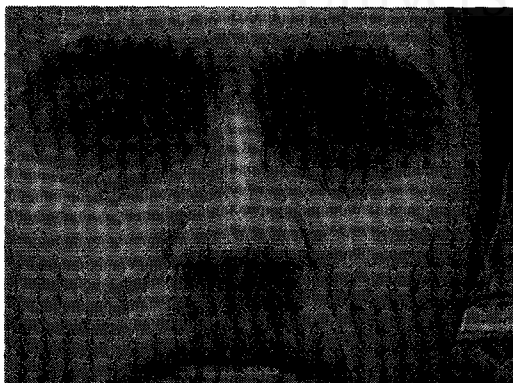


Figura A.11: Tipo de imagen utilizada en la aplicación.

Problemáticas

- En nuestro trabajo contamos con imágenes de un individuo el cual no realiza desplazamientos excesivos. El principal problema al utilizar *snakes* consiste en la adaptación a cambios demasiado bruscos en la superficie de potencial de adaptación. Si trabajáramos con imágenes reales, de un conductor moviéndose constantemente, se debería aplicar una técnica previa de centrado para poder aplicar los contornos correctamente.
- No hemos tenido en cuenta individuos con lentes o pelo cubriendo parte de la cara. También deberíamos tener en cuenta el cambio de iluminación o sombras. Si esto ocurriera se debería aplicar técnicas adicionales para obtener una buena superficie de potencial.
- Uno de los principales problemas del *snake* es la inicialización. Sin embargo podemos disponer de un modelo geométrico de la cara que nos proporcione marcadores morfológicos de las cejas y ojos para poder así inicializar el contorno.

- Cuando realizamos el seguimiento de las trayectorias del ojo y la ceja, existe un intervalo de tiempo el cual los dos *snakes* se sitúan lo suficientemente próximos el uno al otro como para que pueda existir atracción

Experimentación Partimos de imágenes como la mostrada en la Figura A.11 y se debe aplicar un preproceso para obtener una superficie de potencial donde se ajuste el *snake*. Este preproceso tiene como resultado una imagen como la mostrada en la Figura A.12.

Vamos a disponer de dos *snakes*, uno posicionado en la ceja y otro en la parte superior del ojo. La detección de la frecuencia de parpadeo va a venir dada por el tiempo que tardan ambos *snakes* en superar un cierto umbral de proximidad.



Figura A.12: Aplicación del preproceso a la imagen.

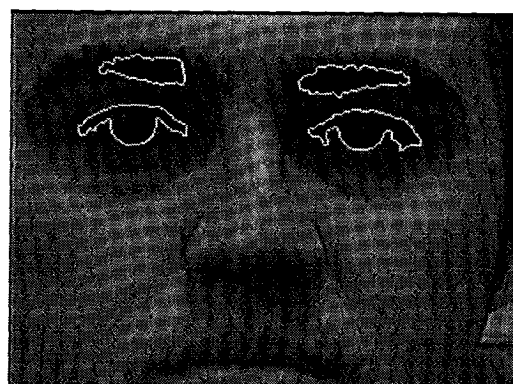


Figura A.13: Snakes inicializados sobre párpados y cejas.

En las gráficas de las Figuras A.14 y A.15 se muestra la evolución en el tiempo

de la distancia promedio entre los *snakes* situados en ceja y ojo. Si dicha distancia es inferior al umbral, los ojos se consideran abiertos¹. Una distancia superior al umbral indica un ojo cerrado. En la Figura A.14 se puede observar como, durante un tiempo, los ojos han permanecido cerrados, por lo que el conductor ha entrado en un estado de somnolencia (mínimos de la función). En la Figura A.15 se muestra un parpadeo normal, donde la alternancia de los valores por encima y debajo del umbral en un intervalo de tiempo determinado indica normalidad.

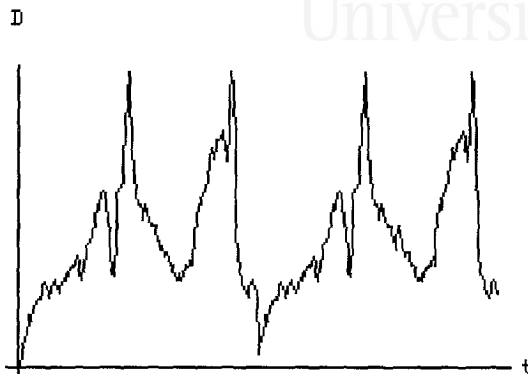


Figura A.14: Gráfica que muestra un parpadeo irregular (dormido).

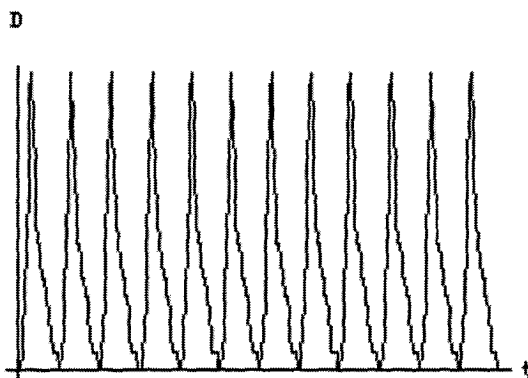


Figura A.15: Gráfica de un parpadeo normal.

Conclusiones Las conclusiones obtenidas con este método se derivan de las problemáticas presentadas. Hemos obtenido un método que detecta el cansancio

¹Notar que la distancia se mide entre ceja y párpado

en la conducción basándonos en la frecuencia de parpadeo del conductor. La experimentación realizada refleja unos buenos resultados en contextos restringidos. La utilización de los modelos de Contornos Activos proporcionan un seguimiento automático de los ojos y las cejas, pero se deben incorporar técnicas adicionales² si pretendemos un seguimiento en condiciones reales. Nuestro línea de trabajo continuará en las dos siguientes:

- Obtención de un modelo geométrico de la cara que permita delimitar las regiones de búsqueda, así como trabajar en la obtención de una segmentación invariante a cambios de iluminación y presencia de objetos en la cara del conductor, tales como pelo o lentes oculares.
- Incorporar un método de aprendizaje (Componentes Principales) que permita reconocer otros patrones que indiquen cansancio.

A.3 Conclusiones

Los experimentos que acabamos de presentar supusieron un punto de partida para el desarrollo de modelos deformables estructurales. Cada uno de estos trabajo nos ha aportado tanto experiencia en el uso de modelos deformables como importantes puntos de vista en aspectos clave. En primer lugar en la reconstrucción a partir de imágenes cerebrales el énfasis se puso en el aspecto estructural (2D y 3D). El análisis de imágenes microscópicas nos aportó experiencia en la extracción de potenciales de imagen relativamente robustos. Finalmente el análisis del cansancio ocular pone de relieve la importancia de disponer de un mecanismo de aprendizaje que permita extraer un modelo temporal para reconocer situaciones tipo. Por otro lado el Modelo de Plantillas Deformables Extendidas sería perfectamente aplicable a cada uno de estos contextos ya que permite introducir robustez local y conocimiento aprendido.

²Marcadores morfológicos, centrado de imágenes, segmentación, etc.



Universitat d'Alacant
Universidad de Alicante

Apéndice B

Transformada de Karhunen-Loève

La *Expansión* de Karhunen-Loève (transformada K-L) (Karhunen, 1946), (Loève, 1955) es una *transformación lineal ortonormal orientada a determinar las características más representativas o efectivas de una distribución o conjunto de muestras de entrada*. En este apéndice completamos la descripción de aspectos básicos del formalismo tomando (Fukunaga, 1990) como referencia. Como ejemplo de aplicación aportamos la modelización radial de objetos y el estudio de su variación mediante esta técnica.

B.1 Transformada K-L Discreta

B.1.1 Transformación Lineal y Ortonormal

Un *vector aleatorio*¹ n -dimensional \vec{x} puede representarse *sin pérdida de información* en función de n vectores *linealmente independientes*:

$$\vec{x} = Pb = \sum_{k=1}^n b_k \vec{\Psi}_k \quad (\text{B.1})$$

siendo $P = [\vec{\Psi}_1, \vec{\Psi}_2, \dots, \vec{\Psi}_n]$ una matriz de *vectores básicos* ($|P| \neq 0$) y $b = (b_1, b_2, \dots, b_n)^T$ las *coordenadas* en la nueva base. Asumimos que dichos vectores

¹En el contexto del reconocimiento de patrones dicho vector puede codificar un conjunto de características de la imagen.

son *ortonormales*:

$$\vec{\Psi}_k^T \vec{\Psi}_l = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \quad (\text{B.2})$$

En consecuencia $P^{-1} = P^T$ y $b = P^T \vec{x}$ ($b_k = \vec{\Psi}_k^T \vec{x}$, $k = 1, 2, \dots, n$). Cada vector $\vec{\Psi}_j$ está asociado a una característica. Hasta el momento hemos realizado un simple cambio de base. A partir de este punto se trata de seleccionar solamente $t < n$ vectores de manera se cometa un *error mínimo* en la aproximación.

B.1.2 Selección de Características

El *criterio de selección* de t características se fundamenta en el razonamiento siguiente. Si reemplazamos los $n - t$ coeficientes no-conocidos por constantes pre-seleccionadas s_k , $k = t + 1, t + 2, \dots, n$ podemos aproximar \vec{x} por:

$$\hat{x}(t) = \sum_{k=1}^t b_k \vec{\Psi}_k + \sum_{k=t+1}^n s_k \vec{\Psi}_k \quad (\text{B.3})$$

El error cometido al considerar solamente un subconjunto de t coordenadas puede expresarse de la siguiente manera:

$$\Delta x(t) = \vec{x} - \hat{x}(t) = \vec{x} - \sum_{k=1}^t b_k \vec{\Psi}_k - \sum_{k=t+1}^n s_k \vec{\Psi}_k = \sum_{k=t+1}^n (b_k - s_k) \vec{\Psi}_k \quad (\text{B.4})$$

El *error cuadrático medio* de $\Delta x(t)$, $\epsilon^2(t) = E(\|\Delta x(t)\|^2)$ es un buen estimador de la efectividad del subconjunto:

$$\begin{aligned} \epsilon^2(t) &= E(\|\Delta x(t)\|^2) = E\left(\sum_{k=t+1}^n \sum_{l=t+1}^n (b_k - s_k)(b_l - s_l) \vec{\Psi}_k^T \vec{\Psi}_l\right) = \\ &= \sum_{k=t+1}^n \sum_{l=t+1}^n E((b_k - s_k)^2) \end{aligned} \quad (\text{B.5})$$

La elección óptima de vectores básicos y términos constantes se obtiene al minimizar la expresión $\epsilon^2(t)$ con respecto a cada s_k :

$$\frac{\partial}{\partial s_k} E((b_k - s_k)^2) = -2[E(b_k) - s_k] = 0 \quad (\text{B.6})$$

obteniéndose $s_k = E(b_k)$, es decir $s_k = \vec{\Psi}_k^T E(\vec{x})$ (reemplazamos los coeficientes no conocidos por sus valores esperados). Una vez aplicado este cambio, podemos reformular $\epsilon^2(t)$ en los términos siguientes:

$$\begin{aligned} \epsilon^2(t) &= \sum_{k=t+1}^n \sum_{k=t+1}^n E((b_k - E(b_k))^2) = \\ &= \sum_{k=t+1}^n \vec{\Psi}_k^T E((\vec{x} - E(\vec{x}))(\vec{x} - E(\vec{x}))^T) \vec{\Psi}_k = \sum_{k=t+1}^n \vec{\Psi}_k^T S \vec{\Psi}_k \end{aligned} \quad (\text{B.7})$$

siendo S la *matriz de covarianza* de \vec{x} . Haciendo uso de la *restricción de ortonormalidad* entre los vectores básicos se demuestra que la selección óptima de los vectores $\vec{\Psi}_k$ es aquella que satisface la siguiente condición:

$$S \vec{\Psi}_k = \lambda_k \vec{\Psi}_k \quad (\text{B.8})$$

es decir, se trata de los autovectores de S . La demostración de esta proposición puede encontrarse en (Fukunaga, 1990). Si sustituimos B.8 en B.7 el error mínimo $\epsilon^2(t)_{opt}$ es:

$$\epsilon^2(t)_{opt} = \sum_{k=t+1}^n \lambda_k \quad (\text{B.9})$$

En definitiva se demuestra que la minimización del error se consigue al *obviar en la representación* las características asociadas a los $n - t$ autovectores con menores autovalores.

B.2 Propiedades Relevantes

B.2.1 Auto-Espacio y Componentes Principales

Como consecuencia del razonamiento anterior obtenemos un nuevo espacio de representación (auto-espacio o espacio de características) de dimensión t definido por:

1. *Origen de Coordenadas:* $E(\vec{x}) = \bar{x}$ es el origen. Normalmente la transformada se aplica sobre $\vec{x} - E(\vec{x})$, es decir una vez computado el desplazamiento con respecto al promedio.
2. *Ejes Principales/Modos de Variación:* vienen dados por los autovectores $\vec{\Psi}_k$. En cuanto a su *interpretación geométrica* el k -ésimo autovector $\vec{\Psi}_k$ establece

la *dirección de variación* de la componente k -ésima en el espacio transformado de manera que afecta a la coordenada x_l provocando un desplazamiento proporcional a dx_{kl} , siendo:

$$\vec{\Psi}_k = (dx_{k1}, dx_{k2}, \dots, dx_{kn})^T \quad (\text{B.10})$$

Combinaciones de distintos autovectores provocan desplazamientos en los elementos del conjunto de entrenamiento. Siendo $P_t = [\vec{\Psi}_1, \vec{\Psi}_2, \dots, \vec{\Psi}_t]$ la matriz con los t primeros autovectores ($t \ll n$) y $b = (b_1, b_2, \dots, b_t)^T$ un vector de pesos o coordenadas en el nuevo dominio (parámetros linealmente independientes) un elemento forma \vec{x}_i del conjunto de entrenamiento puede expresarse como:

$$\vec{x}_i = \bar{x} + P_t b = \bar{x} + \sum_{k=1}^t b_k \vec{\Psi}_k \quad (\text{B.11})$$

Teniendo en cuenta que $P^T P = I$ las coordenadas b_k son las proyecciones de \vec{x}_i sobre cada eje principal, y pueden obtenerse mediante la ecuación:

$$b = P_t^T (\vec{x}_i - \bar{x}) \quad (\text{B.12})$$

3. *Dimensiones de los Ejes*: las variaciones más significativas son descritas por los modos con mayores autovalores asociados. Dichos autovalores λ_k codifican las *varianzas de la distribución*.

En la Figura B.1 puede verse un ejemplo bidimensional en el que se muestra que la eliminación de la segunda dimensión (si su relevancia es baja) en el espacio transformado introduce un error también bajo en la representación. Concretamente, podemos aproximar \vec{x} por $x^* = \bar{x} + \vec{\Psi}_1 b_1 + \vec{\Psi}_2 s_2$.

B.2.2 Efectividad y Otras Propiedades

De la caracterización del nuevo espacio se desprenden los siguientes aspectos relevantes:

1. *Efectividad*: la eliminación de la k -ésima característica $\vec{\Psi}_k$ incrementa el error de la representación en λ_k con lo cual deben retenerse aquellas con mayor autovalor asociado. Esta es una *medida absoluta* de efectividad. La efectividad relativa o normalizada, es decir, la proporción $\frac{\lambda_k}{\lambda_T}$ de variabilidad total λ_T explicada por $\vec{\Psi}_k$ siendo:

$$\lambda_T = \sum_{k=1}^n \lambda_k \quad (\text{B.13})$$

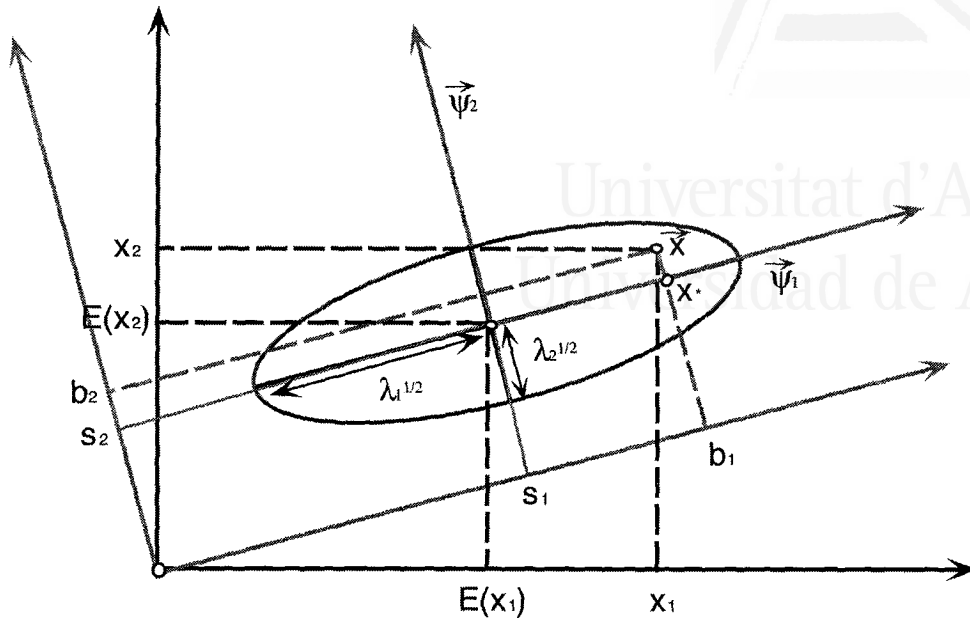


Figura B.1: Expansión K-L bidimensional.

2. *No-Correlación e Independencia*: los valores de las características no están correlacionados entre sí, es decir que $P^T S P$ es diagonal (covarianzas nulas). En el caso especial en que \vec{x} se distribuya normalmente las b_j son independientes. En general y atendiendo a las leyes de la probabilidad haremos uso de dicho supuesto siempre y cuando dispongamos de muestras suficientes en el conjunto de entrenamiento ².
3. *Generación de Elementos de la Distribución*: podemos generar ejemplos variando los parámetros de b dentro de límites aceptables. Bajo el supuesto gaussiano usual, podemos seleccionar aquellos parámetros tales que la distancia de Mahalanobis D_{Mah} entre la forma generada y la media sea inferior a un cierto umbral D_{Def} , siendo éste el criterio de discriminación en reconocimiento:

$$D_{Mah} = b^T S^{-1} b = \sum_{k=1}^t \left(\frac{b_k^2}{\lambda_k} \right) \leq D_{Def} \quad (B.14)$$

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k} \quad (B.15)$$

²Atendiendo al Teorema Central del Límite.

4. *Limitaciones:* la selección de los t autovectores principales introduce el menor error entre todas las selecciones de de vectores básicos ortonormales. Normalmente haremos uso de transformaciones que, como esta, preseven la estructura de la distribución.

B.3 Ejemplos de Aplicación

Como ejemplo de aplicación de esta técnica expondremos a continuación una serie de experimentos orientados a la codificación de procesos de deformación o evoluciones de formas planas. Estos resultados dieron paso a la propuesta final de modelización espacio-temporal aportada.

B.3.1 Modelización Radial del Objeto

Definición B.1 *Un contorno cerrado F sin agujeros se caracteriza mediante un vector n – dimensional, $\vec{F} = (r_0, r_1, \dots, r_{n-1})^T$, es decir, en términos del conjunto de longitudes radiales r_j tomadas a intervalos angulares cada vez más finos (resolución radial incremental) $\frac{\pi}{2}, 0, \frac{3\pi}{2}, \pi, \frac{\pi}{4}, \frac{7\pi}{4}, \frac{5\pi}{4}, \frac{3\pi}{4}, \dots$, que emanan del centro de masas del objeto.*

Definición B.2 *Se define un conjunto radial de entrenamiento C_n^p de grado p y dimensión n como el conjunto de n de contornos cerrados cada uno de los cuales viene caracterizado por un vector \vec{F}_i definido sobre p intervalos angulares.*

B.3.2 Experimentación: Obtención de la Transformada K-L

Definiremos el conjunto de entrenamiento como una serie de formas que codifican la *evolución espacio-temporal* de un objeto y estudiaremos los resultados derivados de la aplicación de la transformada K-L a cada caso. Experimentos similares se han realizado sobre Modelos de Elementos Finitos (Pentland y Horowitz, 1991).

En la Figura B.2 se presenta el modelo radial (esquina superior izquierda) y diversas situaciones sobre imágenes sintéticas dependiendo del *tipo de desplazamiento o evolución del objeto*: deformación horizontal y vertical, expansión/contracción y deformación general. La forma inicial aparece sombreada, mientras que su configuración final se traza de forma discontinua. En todos los casos se ha optado por fijar $n = 8$. Especificaremos el valor de p en cada experimento.

En la Tabla B.1 se presentan las condiciones experimentales así como un resumen numérico de los resultados obtenidos, los cuales comentamos a continuación.

Tabla B.1: Resumen de resultados experimentales.

Exp.	Autoval. λ_k	Efect. $\frac{\lambda_k}{\lambda_T}$	Autovec. $\vec{\Psi}_k$
1	$\lambda_2 = 15.44$	1	$\vec{\Psi}_2 = (0, 1, 0, 0, 0, 0, 0, 0, 0)^T$
2	$\lambda_3 = 2.5$	1	$\vec{\Psi}_3 = (0.7, 0, 0.7, 0, 0, 0, 0, 0, 0)^T$
3	$\lambda_1 = 6.25$	1	$\vec{\Psi}_1 = (0.8, 0, 0.4, 0, 0, 0, 0, 0, 0)^T$
4	$\lambda_2 = 60$	1	$\vec{\Psi}_1 = (0.3, 0.3, 0.3, 0.3, 0.4, 0.4, 0.4, 0.4)^T$
5	$\lambda_1 = 0.89$	0.013	$\vec{\Psi}_1 = (0.4, -0.1, 0.3, -0.4, 0.6, -0.4, 0, -0.2)^T$
	$\lambda_2 = 65.701$	0.932	$\vec{\Psi}_2 = (0.2, 0.4, 0.3, 0.4, 0.4, 0.3, 0.4, 0.3)^T$
	$\lambda_3 = 0$	0	$\vec{\Psi}_3 = (-0.6, 0, -0.1, 0.3, 0.5, 0, 0, -0.5)^T$
	$\lambda_4 = 0.244$	0.003	$\vec{\Psi}_4 = (0.2, -0.8, -0.3, 0.3, 0.1, 0.1, 0.3, 0.2)^T$
	$\lambda_5 = 0.129$	0.002	$\vec{\Psi}_5 = (-0.1, 0.4, -0.7, -0.2, 0.2, -0.3, 0.2, 0.3)^T$
	$\lambda_6 = 0$	0	$\vec{\Psi}_6 = (0, 0, -0.2, -0.6, 0.1, 0.7, 0.1, -0.2)^T$
	$\lambda_7 = 3.508$	0.05	$\vec{\Psi}_7 = (-0.2, 0, 0.2, -0.2, -0.3, -0.2, 0.8, -0.2)^T$
	$\lambda_8 = 0.027$	0	$\vec{\Psi}_8 = (-0.6, -0.2, 0.4, -0.3, 0.2, 0, -0.1, 0.6)^T$

Experimento B.1 (Def. Horizontal) ($p = 5$) Cada forma \vec{F}_i del conjunto puede representarse sin pérdida de información mediante $\vec{F}_i = \vec{F} + \vec{\Psi}_2 b_2$. La única componente significativa de $\vec{\Psi}_2$ es la segunda, es decir la dirección 0, ya que solo en ella se detecta variación. Este resultado es extensible a variaciones en una dirección única (cualquiera de las definidas).

Experimento B.2 (Def. Vert. Isotrópica) ($p = 4$) Idéntico alargamiento del objeto en dos direcciones opuestas. Cada forma \vec{F}_i del conjunto puede representarse sin pérdida de información mediante $\vec{F}_i = \vec{F} + \vec{\Psi}_3 b_3$. Las componentes significativas de $\vec{\Psi}_3$ son las direcciones en que se produce la deformación ($\frac{\pi}{2}$ y $\frac{3\pi}{2}$ en este caso). Dichas componentes mantienen la proporcionalidad de la deformación.

Experimento B.3 (Def Vert. Anisotrópica) ($p = 4$) Alargamiento no-homogéneo en dos direcciones opuestas. Cada forma \vec{F}_i del conjunto puede representarse sin pérdida de información mediante $\vec{F}_i = \vec{F} + \vec{\Psi}_1 b_1$. En este caso el alargamiento en $\frac{\pi}{2}$ es el doble del alargamiento en $\frac{3\pi}{2}$. Las componentes primera y tercera de $\vec{\Psi}_1$ mantienen dicha proporcionalidad.

Experimento B.4 (Expansión/Contracción) ($p = 4$) Crecimiento y Decrecimiento en todas las direcciones. Cada forma \vec{F}_i del conjunto puede representarse sin pérdida de información mediante $\vec{F}_i = \vec{F} + \vec{\Psi}_2 b_2$. Estudiando como en casos anteriores las componentes de $\vec{\Psi}_2$ descubrimos que la variación en las cuatro primeras es 1.414 veces mayor que en las cuatro últimas. Adicionalmente, si estudiamos el comportamiento de dos en dos pasos podemos identificar el signo de la variación analizando el signo de la componente correspondiente.

Experimento B.5 (Deformación General) ($p = 5$) En el caso de deformación general (centrado) el análisis de los autovalores y autovectores nos revela cuales de las variaciones direccionales son significativas a la hora de modelizar la pauta evolutiva del objeto. Atendiendo al criterio de efectividad definido podemos representar cualquier forma \vec{F}_i mediante $\vec{F}_i = \vec{F} + \vec{\Psi}_2 b_2 + \vec{\Psi}_7 b_7$ cometiendo un error medio aproximado del 1.8%.

El caso más general (contemplando traslaciones, rotaciones y escalados) puede estudiarse de forma análoga. En lugar de ello, y de acuerdo con las dimensiones de complejidad especificadas, hemos extendido este planteamiento al tratamiento simultáneo de varios objetos con capacidad de traslación, rotación, escalado y deformación parametrizada. Este nuevo enfoque facilita la especificación de restricciones compactas de movimiento sobre las cuales se basa el reconocimiento.

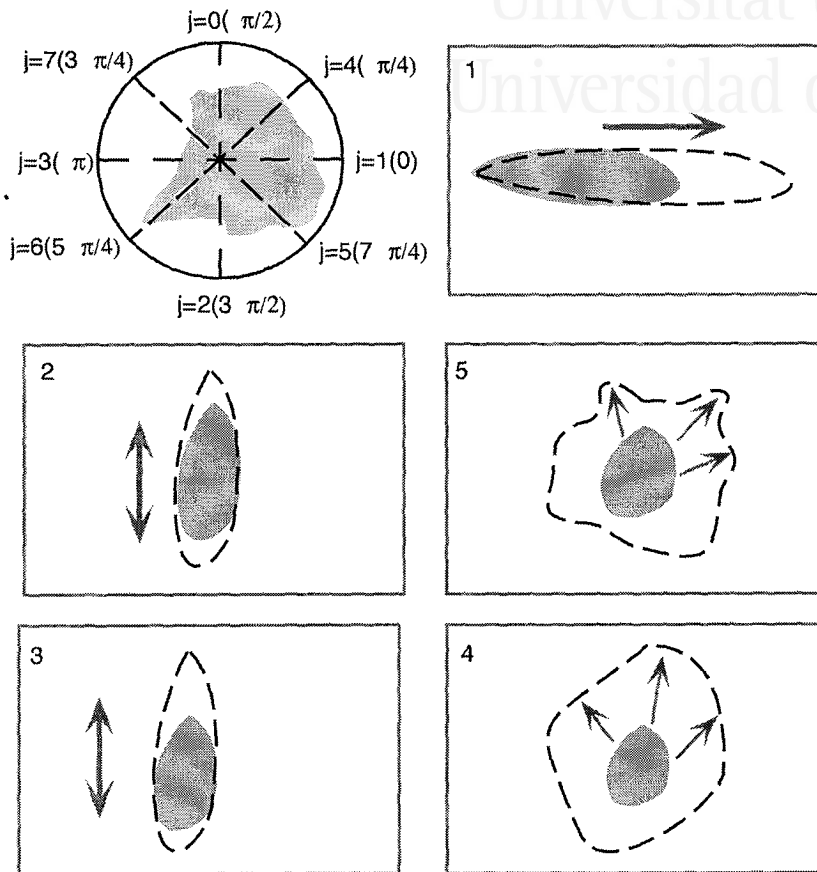


Figura B.2: Modelo radial y casos experimentales.



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Parte V

Referencias



Universitat d'Alacant
Universidad de Alicante

Bibliografía

- (Akita, 1984) K. Akita. Image sequence analysis of real world human motions. *Pattern Recognition*, 17(1):73–83, 1984.
- (Allmen, 1991) M.C. Allmen. *Image Sequence Description Using Spatiotemporal Flow Curves*. PhD thesis, University of Wisconsin-Madison, 1991.
- (Aloimonos *et al.*, 1988) Y. Aloimonos, I. Weiss, y A. Bandopadhyay. Active vision. *International Journal of Computer Vision*, 2:333–356, 1988.
- (Amini *et al.*, 1990) A. A. Amini, T.E. Weymoth, y R.C. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(9):855–867, 1990.
- (Amit y Manbeck, 1993) Y. Amit y K.M. Manbeck. Deformable template models for emission tomography. *IEEE Trans. on Medical Imaging*, 12(2):260–268, 1993.
- (Amit *et al.*, 1989) Y. Amit, U. Grenander, y M. Piccioni. Structural image restoration through deformable templates. *Journal of the American Statistical Association*, 86(414):376–387, 1989.
- (Amit, 1994) Y. Amit. A non-linear variational problem for image matching. *SIAM J. Sci. Computing*, 15(1), 1994.
- (Amit, 1997) Y. Amit. Graphical shape templates for automatic anatomy detection with applications to mri brain scans. *IEEE Trans. on Medical Imaging*, 16(1):28–40, 1997.
- (Bajcsy y Solina, 1987) R. Bajcsy y F. Solina. Three-dimensional object representation revisited. Technical report, Dept. Computer and Information Science, Univ. Pennsylvania, TR No MS-CIS-87-19, 1987.

- (Bajcsy, 1988) R. Bajcsy. Active perception. In *Proc. of the IEEE*, number 76, pages 996–1005, 1988.
- (Ballard, 1991) D.H. Ballard. Animate vision. *Artificial Intelligence*, (48):57–86, 1991.
- (Barr, 1981) A.H. Barr. Superquadrics and angle-preserving transformation. *IEEE Computer Graphics Applications*, 1:11–23, 1981.
- (Bathe, 1982) K.J. Bathe. *Finite Element Procedures in Engineering Analysis*. Prentice-Hall, 1982.
- (Baumberg y Hogg, 1994a) A. Baumberg y D.C. Hogg. An efficient method for contour tracking using active shape models. Technical report, Univ. of Leeds. School of Computer Science. Artificial Intelligence Division. Report 94.11, 1994.
- (Baumberg y Hogg, 1994b) A. Baumberg y D.C. Hogg. Learning flexible models in image sequences. In *European Conference on Computer Vision*, 1994.
- (Bender, 1994) E.A. Bender. *Mathematical Methods in Artificial Intelligence*. IEEE Computer Society Press, 1994.
- (Blake y Yuille, 1992) Eds. A. Blake y A.L. Yuille. *Active Vision*. MIT Press. Cambridge MA, 1992.
- (Blake *et al.*, 1995) A. Blake, M. Isard, y D. Reynard. Learning to track the visual motion of contours. *Artificial Intelligence*, (78):179–212, 1995.
- (Bogas *et al.*, 1997) S. Bogas, V. Penataro, M.J. Segura, F. Escolano, y D. Gallardo. Seguimiento de automóviles en contextos no-restringidos. In *VI Conferencia de la Asociación Española Para la Inteligencia Artificial*, 1997.
- (Bookstein, 1989) F.L. Bookstein. Principal warps: Thin plate splines and the decomposition of deformation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
- (Bro-Nielsen, 1994) M. Bro-Nielsen. Active nets and cubes. Technical report, Institute of Mathematical Modeling. Technical Univ. of Denmark, 1994.
- (Burr, 1986) D.J. Burr. Elastic matching of line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 1986.
- (Campbell y Bobick, 1995) L. Campbell y A. Bobick. Recognition of human body motion using phase space constraints. Technical report, MIT Media Lab. TR. 309, 1995.

- (Cazorla *et al.*, 1995) M.A. Cazorla, P. Caceres, F. Escolano, D. Gallardo, y R. Rizo. Detección automática con snakes y representación 3d sobre imágenes cerebrales. In *VI Conferencia de la Asociación Española Para la Inteligencia Artificial*, 1995.
- (Cediel *et al.*, 1993) M. Cediel, F. Escolano, J.A. Puchol, y R. Rizo. Efecto tamaño y forma del elemento estructurante en compactación morfológica de imágenes. In *V Conferencia de la Asociación Española Para la Inteligencia Artificial*, 1993.
- (Cédras y Shah, 1995) C. Cédras y M. Shah. Motion based recognition: A survey. Technical report, Department of Computer Science. Univ. of Central Florida, 1995.
- (Chen *et al.*, 1994) C.W. Chen, T.S. Huang, y M. Arrott. Modelling, analysis and visualization of left ventricle shape and motion by hierarchical decomposition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16(4):342-356, 1994.
- (Christensen *et al.*, 1993) Eds. H.I. Christensen, K.W. Bowyer, y H. Bunke. *International Journal of Pattern Recognition and Artificial Intelligence*, volume 7. Special Issue: Active Robot Vision: Camera Heads, Model Based Navigation and Reactive Control, 1993.
- (Christensen, 1994) G.E. Christensen. *Deformable Shape Models for Anatomy*. PhD thesis, Sever Institute of Technology. Univ. Washington, 1994.
- (Cohen y Cohen, 1993) L.D. Cohen y I. Cohen. Finite-element methods for active contours models and balloons for 2d and 3d images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1131-1147, 1993.
- (Cohen *et al.*, 1992) I. Cohen, L.D. Cohen, y N. Ayache. Using deformable surface to segment 3d images and infer differential structures. In *Proc. Second European Conference on Computer Vision*, pages 648-652, 1992.
- (Cohen, 1991) L.D. Cohen. On active contour models and balloon. *Computer Vision, Graphics and Image Processing: Image Understanding*, 53(2):211-218, 1991.
- (Cohen, 1992) M.F. Cohen. Interactive space-time control for animation. *Computer Graphics*, 26(2):293-302, 1992.
- (Cootes y Taylor, 1992) T.F. Cootes y C.J. Taylor. Active shape models. smart snakes. In *Proc. British Machine Vision Conference*, pages 266-275, 1992.

- (Cootes y Taylor, 1995) T.F. Cootes y C.J. Taylor. Combining point distribution models with shape models based on finite element analysis. *Image and Vision Computing*, 13(5):403–410, 1995.
- (Cootes *et al.*, 1992) T.F. Cootes, C.J. Taylor, D.H. Cooper, y J. Graham. Trainable method of parametric shape. *Image and Vision Computing*, 10(5):289–294, 1992.
- (Cootes *et al.*, 1993) T.F. Cootes, C.J. Taylor, A. Lanitis, D.H. Cooper, y J. Graham. Building and using flexible models incorporating grey level information. In *Proc. International Conference on Computer Vision*, 1993.
- (Cootes *et al.*, 1994a) T.F. Cootes, A.Hill, C.J. Taylor, y J. Haslam. The use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12(6):339–348, 1994.
- (Cootes *et al.*, 1994b) T.F. Cootes, C.J. Taylor, y A. Lanitis. Active shape models: Evaluating a multi-resolution method for improving image search. In *Proc. British Machine Vision Conference*, 1994.
- (Crowley y Christensen, 1994) Eds. J.L. Crowley y H.I. Christensen. *Vision as Process. Basic Research on Computer Vision Systems*. Springer-Verlag. ES-PRIT Basic Research Series, 1994.
- (Curwen y Blake, 1992) R. Curwen y A. Blake. *Dynamic Contours: Real-Time Active Splines*, chapter 4. Active Vision. Ed. A.L. Yuille, A. Blake. MIT Press, Cambridge MA, 1992.
- (Curwen *et al.*, 1991) R. Curwen, A. Blake, y R. Cipolla. Parallel implementation of lagrangian dynamics for real-time snakes. In *Proc. British Machine Vision Conference*, pages 29–35, 1991.
- (Cutting y Profit, 1982) J.E. Cutting y D.R. Profit. The minimum principle and the perception of absolute and relative motions. *Cognitive Psychology*, 14:211–246, 1982.
- (Darrell y Pentland, 1993) T. Darrell y A.P. Pentland. Space-time gestures. In *Proc. IEEE Conference of Computer Vision and Pattern Recognition*, pages 335–340, 1993.
- (Darrell *et al.*, 1996) T. Darrell, P. Maes, B. Blumberg, y A.P. Pentland. *Exploratory Vision. The Active Eye*, chapter 11. Ed. M. Landy, L.T. Maloney, M. Pavel. Springer-Verlag, 1996.

- (Davis y Shah, 1994) J.W. Davis y M.A. Shah. Gesture recognition. 1994.
- (Delagnes *et al.*, 1995) P. Delagnes, J. Benois, y D. Barba. Active contours approach to object tracking in image sequences with complex background. *Pattern Recognition Letters*, (16):171–178, 1995.
- (Dias y Leitao, 1996) J.M.B. Dias y J.M.N. Leitao. Wall position and thickness estimation from sequences of echocardiographic images. *IEEE Transactions on Medical Imaging*, 15(1):25–37, 1996.
- (Dickinson y Metaxas, 1994) S.J. Dickinson y D. Metaxas. Integrating qualitative and quantitative shape recovery. *International Journal of Computer Vision*, 13(3):1–20, 1994.
- (Dubuisson *et al.*, 1996) M.P. Dubuisson, S. Lakshmanan, y A.K. Jain. Vehicle segmentation and classification using deformable templates. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(3):293–308, 1996.
- (Duda y Hart, 1973) R.O. Duda y P.E. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1973.
- (Escolano *et al.*, 1997a) F. Escolano, M. Cazorla, D. Gallardo, y R. Rizo. Deformable templates for plaque thickness estimation of intravascular ultrasound sequences. In *VII Spanish Symposium on Pattern Recognition and Image Analysis*, 1997.
- (Escolano *et al.*, 1997b) F. Escolano, M. Cazorla, D. Gallardo, y R. Rizo. Deformable templates for tracking and analysis of intravascular ultrasound sequences. In *EMMCVPR-97. International Workshop on Energy Minimization Methods for Computer Vision and Pattern Recognition*, 1997.
- (Florez *et al.*, 1995) F. Florez, F. Ibarra, J.M. García, F. Escolano, y R. Rizo. An optimized neural network system to cursive handwriting recognition. In *VI Spanish Symposium on Pattern Recognition and Image Analysis*, 1995.
- (Freuder y Mackworth, 1994) Eds. E. Freuder y A.K. Mackworth. *Constraint Based Reasoning*. MIT Press/Elsevier, 1994.
- (Fukunaga, 1990) K. Fukunaga. *Introduction to Statistical Pattern Recognition*. New York: Academic, second edition, 1990.
- (García, 1993) J.M. García. *Semicoberturas Heterogéneas de Regiones Bidimensionales Morfológicamente No-restringidas: Modelado Conexionista Aplicado*. PhD thesis, Univ. de Alicante, 1993.

- (Gelb, 1974) Ed. A. Gelb. Applied optimal estimation. *MIT Press. Cambridge MA*, 1974.
- (Goar *et al.*, 1992) F. ST. Goar, F.J. Pinto, E.L. Alderman, P.J. Fitzgerald, M.L. Stadius, y R.L. Popp. Intracoronary ultrasound imaging of angiographically normal coronary arteries: An in vivo comparison with quantitative angiography. *Journal of the American College of Cardiology*, (18):979–987, 1992.
- (Goddard, 1992) N.H. Goddard. *The Perception of Articulated Motion: Recognizing Moving Light Displays*. PhD thesis, University of Rochester, 1992.
- (Gould y Shah, 1989) K. Gould y M.A. Shah. The trajectory primal sketch: a multiscale scheme for representing motion characteristics. 1989.
- (Grenander y Keenan, 1989) U. Grenander y D.M. Keenan. Towards automated image understanding. *Journal of Applied Statistics*, 16(2):207–221, 1989.
- (Grenander *et al.*, 1991) U. Grenander, Y. Chow, y D.M. Keenan. *HANDS: A Pattern Theoretical Stud, of Biological Shapes*. Springer-Verlag New York, 1991.
- (Grimson, 1990) E.W.L. Grimson. *Object Recognition by Computer: The Role of Geometric Constraints*. MIT Press. Cambridge MA, 1990.
- (Grzeszczuk y Terzopoulos, 1995) R. Grzeszczuk y D. Terzopoulos. Automated learning of muscle-actuated locomotion through control abstraction. In *Proc. SIGGRAPH'95*, pages 63–70, 1995.
- (Guijarro *et al.*, 1993) R. Guijarro, F. Escolano, J.A. Puchol, y B. Ledesma. Reconocimiento de patrones basado en cálculo de esqueletos. In *V Conferencia de la Asociación Española Para la Inteligencia Artificial*, 1993.
- (Hill *et al.*, 1992) A. Hill, C.J. Taylor, y T.F. Cootes. Object recognition b,flexible template matching using genetic algorithms. In *Proc. European Conference on Computer Vision*, pages 852–856, 1992.
- (Hogg, 1984) D.C. Hogg. *Interpreting Images of a Known Moving Object*. PhD thesis, University of Sussex, 1984.
- (Hough, 1962) P.V.C. Hough. *A Method and Means for Recognizing Complex Patterns*. US-Patent 3.069.654, 1962.
- (Hyche *et al.*, 1992) M. Hyche, N. Ezquerra, y R. Mullick. Spatiotemporal detection of arterial structure using active contours. In *Proc. Visualization in Biomedical Computing*, pages 52–62, 1992.

- (Hyche *et al.*, 1995) M. Hyche, D.P. Greenberg, y R. Popp. Time-dependent three dimensional intravascular ultrasound. In *Proc. SIGGRAPH'95*, pages 457–464, 1995.
- (Ibarra, 1995) F. Ibarra. *Análisis de Texturas Mediante Coeficiente Morfológico. Modelado Conexionista Aplicado*. PhD thesis, Univ. de Alicante, 1995.
- (Ikeuchi y Herbert, 1996) K. Ikeuchi y M. Herbert. *Exploratory Vision. The Active Eye*, chapter 11. Ed. M. Landy, L.T. Maloney, M. Pavel. Springer-Verlag, 1996.
- (Isard y Blake, 1996) Isard y A. Blake. Contour tracking by stochastic propagation of conditional density. In *Proc. European Conf. Computer Vision*, pages 343–356, 1996.
- (Isner *et al.*, 1992) J.M. Isner, K. Rosenfield, D.W. Losordo, y C. Krishnaswamy. *Clinical Experience with Intravascular Ultrasound as an Adjunct to Percutaneous Revascularization*, chapter 16, pages 186–197. *Intravascular Ultrasound Imaging*. Ed. J.M. Tobis, P.G. Yock. Churchill Livingstone Inc., New York, 1992.
- (Jain, 1996) A.K. Jain. Object matching using deformable templates. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(3):267–278, 1996.
- (Johansson, 1973) G. Johansson. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14(2):210–211, 1973.
- (Johansson, 1975) G. Johansson. Visual motion perception. *Scientific American*, pages 76–88, 1975.
- (Kalman, 1960) R.E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME J. of Basic Engineering*, 1960.
- (Karhunen, 1946) K. Karhunen. *Über Lineare Methoden in der Wahrscheinlichkeitsrechnung*. Ann. Acad. Sci. Fennicae, ser. A1, Math. Phys. (Vol 37), 1946.
- (Kass *et al.*, 1987) M. Kass, A. Witkin, y D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1, 1987.
- (Kervrann y Heitz, 1993) C. Kervrann y F. Heitz. A hierarchical statistical framework for the segmentation of deformable objects in image sequences. Technical report, Institut de Recherche en Informatique et Systèmes Aleatoires, TR No.780, 1993.

- (Kirby y Sirovich, 1990) M. Kirby y L. Sirovich. Application of the karhunen-loève procedure for the characterization of human faces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(1):103-108, 1990.
- (Kirby *et al.*, 1993) M. Kirby, F. Weisser, y G. Dangelmayr. A model problem in the representation of digital image sequences. *Pattern Recognition*, 26(1):63-73, 1993.
- (Koller *et al.*, 1994) D. Koller, J. Weber, y J. Malik. Robust multiple car tracking with occlusion reasoning. Technical report, ECCS Dept. Univ. California at Berkeley, 1994.
- (Krishnaswamy *et al.*, 1992) C. Krishnaswamy, A.J. D'Adamo, y C.M. Sehgal. *Three Dimensional Reconstruction of Intravascular Ultrasound Images*, chapter 13, pages 141-147. *Intravascular Ultrasound Imaging*. Ed. J.M. Tobis, P.G. Yock. Churchill Livingstone Inc., New York, 1992.
- (Krotkow, 1987) E.P. Krotkow. *Exploratory Visual Sensing for Determining Spatial Layout with an Agile Stereo Camera System*. PhD thesis, Univ. of Pennsylvania, 1987.
- (Kumar y Goldgof, 1994) S. Kumar y D. Goldgof. Automatic tracking of spamm grid and the estimation of deformation parameters from cardiac mr images. *IEEE Trans. on Medical Imaging*, 13(1):122-132, 1994.
- (Kuniyoshi *et al.*, 1994) Y. Kuniyoshi, M. Inaba, y H. Inoue. Learning by watching: Extracting reusable task knowledge from visual observation of human performance. *IEEE Trans. on Robotics and Automatation*, 10(6):799-822, 1994.
- (Lai, 1994) K.F. Lai. *Deformable Contours: Modeling, Extraction, Detection and Classification*. PhD thesis, University of Wisconsin-Madison, 1994.
- (Landy *et al.*, 1996) Eds. M. Landy, L.T. Malone, y M. Pavel. *Exploratory Vision: The Active Eye*. Springer-Verlag, 1996.
- (Lanitis *et al.*, 1994) A. Lanitis, C.J. Taylor, T.F. Cootes, y T. Ahmed. Automatic interpretation of human faces and hand gestures using flexible models. Technical report, Department of Medical Biophysics. University of Manchester, 1994.
- (Lee *et al.*, 1995) S.Y. Lee, K-Y. Chua, y S.Y. Shi. Image metamorphosis using snakes and free-form deformations. In *Proc. SIGGRAPH'95*, pages 439-448, 1995.

- (Lengyel *et al.*, 1995) J. Lengyel, D.P. Greenberg, A. Yeung, E. Alderman, y R. Popp. Three dimensional reconstruction and volume rendering of intravascular ultrasound slices imaged on a curved arterial path. In *Proc. CVRMed'95*, 1995.
- (Leonardis *et al.*, 1996) A. Leonardis, F. Solina, y A. Macerl. A direct recovery of spherquadric models in range images using recover-and-select paradigm. Technical report, Computer Vision Lab. Univ. of Ljubljana, 1996.
- (Leymarie y Levine, 1992) F. Leymarie y M.D. Levine. Simulating the grassfire transform using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):56-75, 1992.
- (Leymarie y Levine, 1993) F. Leymarie y M. Levine. Tracking deformable objects in the plane using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6):635-646, 1993.
- (Lipson *et al.*, 1990) P. Lipson, A.L. Yuille, D. O'Keefe, J. Cavanaugh, J. Taaffe, y D. Rosenthal. Deformable templates for feature extraction from medical images. In *Proc. 1st European Conference on Computer Vision*, 1990.
- (Loève, 1955) M. Loève. *Probability Theory*. Von Nonstrand, Princeton, New Jersey, 1955.
- (Lopez *et al.*, 1997) C. Lopez, M. Mi nano, F. Escolano, M. Cazorla, D. Gallardo, y R. Rizo. Snakes based analysis of ocular fatigue. In *VII Spanish Symposium on Pattern Recognition and Image Analysis*, 1997.
- (Marr, 1982) D. Marr. *Vision*. Eds. W.H. Freeman and Co, San Francisco, 1982.
- (Martelli, 1972) A. Martelli. Edge detection using heuristic search methods. *Computer Graphics and Image Processing*, 1:169-182, 1972.
- (McEachen *et al.*, 1994) J.C. McEachen, F.G. Meyer, R.T. Constable, A.Ñehorai, y J.S. Duncan. A recursive filter for phase velocity assisted shape-based tracking of non-rigid motion. Technical report, Yale Univ. School of Medicine, 1994.
- (McIrnerney y Terzopoulos, 1996) T. McIrnerney y D. Terzopoulos. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2), 1996.
- (Metaxas y Kakadiaris, 1996) D. Metaxas y I.A. Kakadiaris. Elastically adaptive deformable models. In *Proc. IV European Conference on Computer Vision*, pages 550-559, 1996.

- (Mitiche y Bouthemy, 1996) A. Mitiche y P. Bouthemy. Computation and analysis of image motion: A synopsis of current problems and methods. *International Journal of Computer Vision*, 19(1):29–55, 1996.
- (Molina *et al.*, 1995) R. Molina, F. Ibarra, J.M. García, y F. Escolano. Automatic detection of microcalcifications in mammograms. neural modeling. In *9th Scandinavian Conference on Image Analysis*, 1995.
- (Montanari, 1971) U. Montanari. On the optimal detection of curves in noisy pictures. *Communications of the ACM*, 14(5):335–345, 1971.
- (Murase y Nayar, 1993) H. Murase y S.K. Nayar. Learning and recognition of 3d objects from appearance. In *Proc. IEEE Workshop on Qualitative Vision*, pages 39–50, 1993.
- (Murase y Sakai, 1996) H. Murase y R. Sakai. Moving object recognition in eigenspace representation: Gait analysis and lip reading. *Pattern Recognition Letters*, 17:155–162, 1996.
- (Nastar y Ayache, 1993) C.Ñastar y N. Ayache. A new physically based model for efficient tracking and analysis of deformations. In *Proc. Geometric Reasoning for Perception and Action*. Springer-Verlag, 1993.
- (Nayar *et al.*, 1995a) S.K. Nayar, H. Murase, y S.A. Nene. Learning, positioning and tracking visual appearance. In *Proc. IEEE International Conference on Robotics and Automation*, 1995.
- (Nayar *et al.*, 1995b) S.K. Nayar, H. Murase, y S.A. Nene. Subspace methods for robot vision. In *Proc. IEEE International Conference on Robotics and Automation*, 1995.
- (ORourke y Badler, 1980) J. ORourke y N.I. Badler. Model based image analysis of human motion using constraint propagation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2:522–536, 1980.
- (Pal y Pal, 1993) N.R. Pal y S.K. Pal. A review of image segmentation techniques. *Pattern Recognition*, 26(9):1277–1294, 1993.
- (Pelillo y Hancock, 1997) Eds. M. Pelillo y E.R. Hancock. *Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer-Verlag. Lecture Notes in Computer Science, 1997.

- (Pentland y Horowitz, 1991) A. Pentland y B. Horowitz. Recovery of non-rigid motion and structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):730–742, 1991.
- (Pentland y Sclaroff, 1991) A. Pentland y S. Sclaroff. Closed-form solutions for physically based shape modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):715–729, 1991.
- (Pentland y Williams, 1989) A. Pentland y J. Williams. Good vibrations: Modal dynamics for graphics and animation. *Computer Graphics*, 1989.
- (Pentland *et al.*, 1996) A.P. Pentland, R.W. Picard, y S. Sclaroff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, 1996.
- (Pentland, 1986) A.P. Pentland. Perceptual organization and the representation of natural form. *Artificial Intelligence*, 28:293–331, 1986.
- (Penyalver *et al.*, 1995) A. Penyalver, F. Escolano, F. Ibarra, J.M. García, y R. Rizo. Neural fuzzy decision systems based detection of shape evolution: Application to breast stellate lesion treatment monitoring. In *VI Spanish Symposium on Pattern Recognition and Image Analysis*, 1995.
- (Polana y Nelson, 1994) R. Polana y R. Nelson. Nonparametric recognition of non-rigid motion. Technical report, Department of Computer Science. Univ. of Rochester, 1994.
- (Rizo *et al.*, 1993) R. Rizo, J.M. García, F. Escolano, J.A. Puchol, F. Ibarra, M. Pujol, J.L. Verdú, y B. Ledesma. Problemas de concurrencia en análisis de imagen y redes neuronales. In *III Jornadas de Concurrencia*, 1993.
- (Rizo, 1991) R. Rizo. *Semicoberturas Homogéneas de Regiones Bidimensionales*. PhD thesis, Univ. Politécnica de Valencia, 1991.
- (Roedlandt *et al.*, 1994) J.R. Roedlandt, C. Di Mario, N.G. Pandian, L. Wenguan, L. Keane, C.J. Slager, P.J. De Feyter, y P.W. Serrius. Three dimensional reconstruction of intracoronary ultrasound images. *Circulation*, (90):1044–1055, 1994.
- (Rohr, 1994) K. Rohr. Towards model-based recognition of human movements in image sequences. *Computer Vision, Graphics and Image Processing: Image Understanding*, 59(1):94–115, 1994.

- (Rosenfield *et al.*, 1991) K. Rosenfield, D.W. Losordo, K. Ramaswamy, J.O. Pastore, E. Langevin, S. Razvi, B.D. Kosowski, y J.M. Isner. Three dimensional reconstruction of human coronary and peripheral arteries from images recorded during two-dimensional intravascular ultrasound examination. *Circulation*, (84):1938–1956, 1991.
- (Sabuco *et al.*, 1997) I. Sabuco, F. Escolano, M. Cazorla, D. Gallardo, y R. Rizo. Snakes based tracking and texture analysis of microscopic images. In *VII Spanish Symposium on Pattern Recognition and Image Analysis*, 1997.
- (Sakaue y Yamamoto, 1991) K. Sakaue y K. Yamamoto. Active net model and its application for region extraction. *JITE of Japan*, 45(10):1155–1163, 1991.
- (Sallam y Bowyer, 1994) M. Sallam y K. Bowyer. Image unwarping for digital subtraction on time sequences of mammograms. Technical report, Dept. of Computer Science and Engineering. Univ. South Florida Tampa, 1994.
- (Sclaroff y Pentland, 1993) S. Sclaroff y A. Pentland. Modal matching for correspondence and recognition. Technical report, MIT Media Lab, Perceptual Computing TR No.201, 1993.
- (Sclaroff y Pentland, 1994) S. Sclaroff y A. Pentland. On modal matching for medical images: Underconstrained shape description and data compression. Technical report, MIT Media Lab, Perceptual Computing TR No.275, 1994.
- (Sclaroff, 1995) S. Sclaroff. Deformable prototypes for encoding shape categories in image databases. Technical report, Boston University CS Dept TR 95-017, 1995.
- (Serra, 1982) J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press. London, 1982.
- (Shah *et al.*, 1993) M. Shah, K. Rangarajan, y P.S. Tsai. Motion trajectories. *IEEE Transactions on Systems, Man and Cybernetics*, 23(4):1138–1150, 1993.
- (Staib y Duncan, 1989) L.H. Staib y J.S. Duncan. Boundary fitting with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1061–1183, 1989.
- (Szeliski, 1989) R. Szeliski. *Bayesian Modeling of Uncertainty in Low-Level Vision*. Kluwer Academics, 1989.

- (Takahata *et al.*, 1992) M. Takahata, M. Imai, y S. Tsuji. Determining motion of non-rigid objects by active tubes. In *Proc. 11th IAPR International Conference on Pattern Recognition*, pages 647–650, 1992.
- (Terzopoulos y Metaxas, 1991) D. Terzopoulos y D. Metaxas. Dynamic 3d models with local and global deformations: Deformable superquadrics. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(7):703–714, 1991.
- (Terzopoulos y Metaxas, 1992) D. Terzopoulos y D. Metaxas. *Active Vision*, chapter 5. Ed. A.L. Yuille, A. Blake. MIT Press, 1992.
- (Terzopoulos y Metaxas, 1993) D. Terzopoulos y D. Metaxas. Shape and non-rigid motion estimation through physics-based synthesis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15(6):580–591, 1993.
- (Terzopoulos y Szeliski, 1992) D. Terzopoulos y R. Szeliski. *Tracking with Kalman Snakes*, chapter 1. *Active Vision*. Ed. A.L. Yuille, A. Blake. MIT Press, Cambridge MA, 1992.
- (Terzopoulos *et al.*, 1988) D. Terzopoulos, A. Witkin, y M. Kass. Constraints on deformable models: Recovering 3d shape and non-rigid motion. *Artificial Intelligence*, 63:91–123, 1988.
- (Tsotsos *et al.*, 1980) J.K. Tsotsos, J. Mylopoulos, H.D. Convey, y S.W. Zucker. A framework for visual motion understanding. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2(6):563–573, 1980.
- (Tsumiyama *et al.*, 1989) Y. Tsumiyama, K. Sakaue, y K. Yamamoto. Active net model for region extraction. *IPSJ SIG Notes, CV 63-2*, 89(96):1–8, 1989.
- (Turk y Pentland, 1991) M.A. Turk y A.P. Pentland. Face recognition using eigenfaces. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–591, 1991.
- (Ullman, 1979) S. Ullman. *The Interpretation of Visual Motion*. MIT Press. Cambridge MA, 1979.
- (Williams y Shah, 1992) D.J. Williams y M. Shah. A fast algorithm for active contours and curvature estimation. *CVGIP: Image Understanding*, 55(1):14–26, 1992.
- (Witkin y Kass, 1988) A. Witkin y M. Kass. Space-time constraints. *Computer Graphics*, 22(4):159–168, 1988.

- (Xie *et al.*, 1994) X. Xie, R. Sudhakar, y H. Zhuang. On improving eye feature extraction using deformable templates. *Pattern Recognition*, 27(6):791–799, 1994.
- (Yamato *et al.*, 1992) J. Yamato, J. Ohya, y K. Ishii. Recognizing human action in time-sequential images using hidden markov models. pages 379–385, 1992.
- (Yock *et al.*, 1989) P. Yock, D. Linker, y A. Angelson. Intravascular ultrasound: Technical development and initial clinical experience. *Journal of the American Society of Echocardiography*, (2):296–304, 1989.
- (Yuille y Halliman, 1992) A.L. Yuille y P.W. Halliman. *Active Vision*, chapter 2. Ed. A.L. Yuille ,A. Blake. MIT Press, 1992.
- (Yuille *et al.*, 1989) A.L. Yuille, D.S. Cohen, y P.W. Halliman. Feature extraction from faces using deformable templates. In *Proc. CVPR*, pages 104–109, 1989.
- (Yuille *et al.*, 1990) A.L. Yuille, K. Honda, y C. Peterson. Particle tracking by deformable templates. In *Proc. International Joint Conference on Neural Networks*, 1990.
- (Yuille, 1990) A.L. Yuille. Generalized deformable models, statistical physics and matching problems. *Neural Computation*, 2:1–24, 1990.
- (Zhu y Yuille, 1994) S.C. Zhu y A.L. Yuille. Forms: A flexible object recognition and modelling system. Technical report, Harvard Robotics Laboratory, TR No.94-1, 1994.
- (Zucker *et al.*, 1977) S. Zucker, R. Hummel, y A. Rosenfeld. An application of relaxation labeling to line and curve enhancement. *IEEE Transactions on Computers*, C-26, 1977.