

**UNIVERSIDAD DE LA LAGUNA**

**Contribuciones al reconocimiento  
de objetos desde primitivas  
de elementos de contorno**

**Autor: Pérez Nava, Fernando**

**Director: Antonio Falcón Martel**

**Departamento de Estadística, Investigación Operativa  
y Ciencias de la Computación**

*A la grata memoria de mi padre.  
A Gloria por su comprensión y estímulo,  
a Raquel por su sonrisa.  
A mi familia por su apoyo durante todos estos años.*

## ***AGRADECIMIENTOS***

Quiero agradecer a todos aquellos que han hecho posible la realización de este trabajo. En primer lugar a D. Antonio Falcón Martel por su dirección y al grupo de Inteligencia Artificial y Sistemas en la Universidad de Las Palmas de Gran Canaria. A toda la gente del DEIOC en la Universidad de La Laguna y en particular al área de Inteligencia Artificial por su compañerismo durante estos años. Mi gratitud para Gloria y nuestra Raquel por su paciencia y su apoyo constante. Mi agradecimiento más profundo a mi familia y en particular a mis padres que han hecho todo lo que estaba en su mano para que esto fuera posible.

# Índice

## **1 INTRODUCCIÓN** **1**

---

## **2 RECONOCIMIENTO Y SEGUIMIENTO DE OBJETOS BASADO EN MODELOS** **5**

<b>2.1</b>	<b>EL PROBLEMA DEL MODELADO DE OBJETOS</b>	<b>8</b>
2.1.1	CUESTIONES BÁSICAS DEL MODELADO DE OBJETOS	8
2.1.2	REPRESENTACIÓN DE MODELOS BASADOS EN CONTORNOS	13
2.1.3	MODELADO DE OBJETOS DEFORMABLES	26
<b>2.2</b>	<b>EMPAREJAMIENTO DE OBJETOS BASADO EN MODELOS</b>	<b>36</b>
2.2.1	EMPAREJAMIENTO DE MODELOS RÍGIDOS	37
2.2.2	EMPAREJAMIENTO DE OBJETOS DEFORMABLES	48
<b>2.3</b>	<b>CLASIFICACIÓN DE OBJETOS</b>	<b>54</b>
2.3.1	FUNCIONES ESPECÍFICAS	54
2.3.2	FUNCIONES DE AJUSTE COMBINADAS	54
2.3.3	PROBABILIDAD A POSTERIORI DE LOS MODELOS	54
2.3.4	MEDIDAS BASADAS EN CLASIFICADORES	55
<b>2.4</b>	<b>SEGUIMIENTO DE OBJETOS BASADO EN MODELOS</b>	<b>56</b>
2.4.1	CUESTIONES BÁSICAS DEL SEGUIMIENTO BASADO EN MODELOS	57
2.4.2	SEGUIMIENTO DE MODELOS RÍGIDOS	60
2.4.3	SEGUIMIENTO DE MODELOS DEFORMABLES	62

## **3 MODELADO DE CONTORNOS DEFORMABLES MEDIANTE WAVELETS** **67**

<b>3.1</b>	<b>ANÁLISIS DE FUNCIONES MEDIANTE WAVELETS</b>	<b>67</b>
3.1.1	INTRODUCCIÓN	67
3.1.2	LA TRANSFORMADA CONTINUA MEDIANTE WAVELETS	68
3.1.3	SERIES DE WAVELETS	69
3.1.4	ANÁLISIS A MÚLTIPLES RESOLUCIONES	70
3.1.5	WAVELETS ORTOGONALES	71
3.1.6	WAVELETS BIORTOGONALES	72
3.1.7	APROXIMACIÓN MEDIANTE WAVELETS	73
3.1.8	LA TRANSFORMADA RÁPIDA MEDIANTE WAVELETS	73
3.1.9	WAVELETS EN UN INTERVALO	74
3.1.10	WAVELETS Y SUAVIDAD DE FUNCIONES	75
3.1.11	MODELOS ESTADÍSTICOS BASADOS EN WAVELETS	76
3.1.12	DERIVADAS BASADAS EN WAVELETS	78
3.1.13	MULTIWAVELETS	79
3.1.14	FAMILIAS DE WAVELETS	80
<b>3.2</b>	<b>ANÁLISIS DE CONTORNOS MEDIANTE WAVELETS</b>	<b>82</b>
3.2.1	DESCOMPOSICIÓN DE CONTORNOS MEDIANTE WAVELETS	82
3.2.2	DESCOMPOSICIÓN DE CONTORNOS MEDIANTE MULTIWAVELETS	83
3.2.3	UNA COMPARATIVA ENTRE DISTINTAS REPRESENTACIONES	85

3.2.4	NORMAS PARA CURVAS	87
3.2.5	NORMAS PARA CURVAS EN ESPACIOS DE SUAVIDAD	88
3.2.6	MODELOS DE ESPACIO DE FORMAS	90
<b>3.3</b>	<b>MODELOS DE DEFORMACIÓN DE CONTORNOS MEDIANTE WAVELETS</b>	<b>92</b>
3.3.1	DEFORMACIÓN DE CONTORNOS MEDIANTE WAVELETS	92
3.3.2	AJUSTE DE CONTORNOS MEDIANTE WAVELETS	94
3.3.3	MODELOS PROBABILÍSTICOS DE DEFORMACIÓN BASADOS EN WAVELETS	97
3.3.4	EJEMPLOS DE REGULARIZACIÓN	112
<b>3.4</b>	<b>MODELOS DINÁMICOS BASADOS EN WAVELETS</b>	<b>115</b>
3.4.1	INTRODUCCIÓN	115
3.4.2	MODELOS AUTOREGRESIVOS DE PRIMER ORDEN (AR1)	115
3.4.3	MODELOS AUTOREGRESIVOS DE SEGUNDO ORDEN (AR2)	120

## **4 APLICACIONES** **127**

<b>4.1</b>	<b>SEGUIMIENTO DE CONTORNOS</b>	<b>127</b>
4.1.1	SEGUIMIENTO DE CONTORNOS MEDIANTE MODELOS EN EL ESPACIO DE SOBOLEV	127
4.1.2	SEGUIMIENTO DE CONTORNOS CON MODELOS NO GAUSIANOS	143
<b>4.2</b>	<b>EMPAREJAMIENTO DE CONTORNOS</b>	<b>154</b>
4.2.1	INTRODUCCIÓN	154
4.2.2	REPRESENTACIÓN INVARIANTE DE CONTORNOS MEDIANTE MULTIWAVELETS	155
4.2.3	ALGORITMO DE EMPAREJAMIENTO	156
4.2.4	EJEMPLOS	159

## **5 CONCLUSIONES, PRINCIPALES APORTACIONES Y POSIBLES EXTENSIONES** **161**

<b>5.1</b>	<b>CONCLUSIONES Y PRINCIPALES APORTACIONES</b>	<b>161</b>
<b>5.2</b>	<b>POSIBLES EXTENSIONES Y LÍNEAS DE DESARROLLO</b>	<b>163</b>

## **REFERENCIAS** **165**

## **APÉNDICE A** **173**

## **APÉNDICE B** **180**

# 1 Introducción

La visión por ordenador basada en modelos [Pope 1994] es una de las principales tendencias en el ámbito de la visión artificial y está basada en resolver los problemas computacionales que se presentan mediante modelos matemáticos. En este trabajo se ha utilizado esta aproximación para estudiar el problema del reconocimiento y seguimiento de objetos. Éste puede describirse como la identificación y localización de qué objetos están presentes en una escena dada y su seguimiento en las siguientes imágenes de una secuencia. Para ello se parte de algún tipo de información acerca de un conjunto de objetos que podrían aparecer en una imagen, junto con una secuencia de escenas que posiblemente los contenga. Este problema aparece generalmente en un amplio espectro de aplicaciones y afecta a campos del conocimiento donde la información visual es significativa como dato (Medicina, Biología, Comunicaciones, Robótica,...). La forma de resolver este problema difiere de acuerdo con la tipología de conocimiento que se emplee. En este trabajo la información acerca de un objeto vendrá dada mediante primitivas que proporcionan un modelo geométrico de su forma, un modelo de deformación y un modelo de su evolución temporal.

Fundamentalmente, el reconocimiento y seguimiento de los objetos presentes en la escena, se lleva a cabo mediante el ajuste entre ciertas características del objeto presentes en la imagen y ciertas características comparables en el modelo del objeto, contando con una posible predicción acerca de su localización y apariencia basada en el análisis temporal de su comportamiento.

Para dar una solución a este problema deben resolverse cuatro cuestiones fundamentales:

1. **El problema del modelado**, es decir, cual es la forma de abstraer los elementos más significativos de la apariencia del objeto.
2. **El problema del emparejamiento**, que proporciona el método de emparejamiento entre el modelo del objeto y la imagen que lo contiene.
3. **El problema de la clasificación**, que indica en base a los resultados del proceso de emparejamiento si el modelo está presente en la imagen.
4. **El problema del seguimiento**, que utiliza la información temporal para realizar un emparejamiento en imágenes sucesivas

Estos problemas no son independientes y la elección de un tipo de modelado condiciona frecuentemente el tipo de emparejamiento debido a que éste se lleva a cabo entre las representaciones del objeto y de la imagen. Es evidente que la elección del tipo de emparejamiento condiciona también el método de seguimiento elegido.

En esta tesis se han abordado los problemas anteriores desde modelos geométricos basados en contornos. Se utiliza la descripción de curvas por medio de wavelets

[Chuang 1996] para proporcionar un conjunto de primitivas y soluciones basadas en éstas para los problemas anteriores. En particular en esta tesis se aporta:

- Un modelo probabilístico para la descripción de la deformación de contornos basado en wavelets.
- Un modelo dinámico de la evolución de contornos en el tiempo basado en wavelets.
- La integración de estos dos modelos en métodos de seguimiento basados en el filtro de Kalman y propagación de la densidad condicional “Condensation”.

En el marco de la representación con wavelets se ha propuesto una extensión a los mismos: los multiwavelets [Chui 1996]. En este trabajo se ha propuesto también una nueva representación de curvas basada en multiwavelets y se ha abordado el problema de emparejamiento rígido con esta representación.

Esquemáticamente estas aportaciones pueden describirse como:

	Modelado	Emparejamiento	Clasificación	Seguimiento
Modelos rígidos	Descripción de curvas basada en multiwavelets	Emparejamiento de curvas basada en multiwavelets	Clasificación de curvas basada en multiwavelets	
Modelos deformables	Modelo probabilístico de deformación para descripciones basadas en wavelets	Ajuste de curvas para descripciones basadas en wavelets		Modelo dinámico de deformación para descripciones basadas en wavelets  Integración con varios métodos de seguimiento

**Tabla 1.1 Principales aportaciones de esta tesis.**

Esta tesis consta de esta introducción y cuatro capítulos más cuyos contenidos se detallan a continuación:

- En el segundo capítulo se expone una revisión bibliográfica de las aportaciones propuestas por diversos autores a los problemas de modelado, emparejamiento, clasificación y seguimiento tanto para modelos rígidos como para modelos deformables. Para ello se ha dedicado una sección a cada uno de los problemas que se presentan.

De esta forma:

- En la Sección 2.1 se aborda el problema de modelado haciendo énfasis en los problemas comunes para modelos rígidos y modelos deformables (secciones 2.1.1, 2.1.2) y dedicando la Sección 2.1.3 a los problemas específicos de modelado de objetos deformables.

- En la Sección 2.2 se aborda el problema del emparejamiento tanto para modelos rígidos (Sección 2.2.1) como para modelos deformables (Sección 2.2.2).
  - En la Sección 2.3 se aborda el problema de la clasificación de objetos
  - En la Sección 2.4 se aborda el problema de seguimiento dedicando la Sección 2.4.2 a los modelos rígidos y la Sección 2.4.3 a los deformables.
- 
- En el tercer capítulo se exponen las aportaciones propuestas en este trabajo en distintos aspectos de la representación de curvas basada en wavelets y multiwavelets y se proporciona un completo análisis teórico de aspectos fundamentales de esta representación como los modelos de deformación o modelos dinámicos.

Así:

- En la Sección 3.1 se aborda el problema de la representación de funciones mediante wavelets partiendo de la transformada continua mediante wavelets y llegando al análisis a múltiples resoluciones. Se detalla también la conexión entre la descomposición en series de wavelets y los espacios de suavidad.
  - En la Sección 3.2 se aborda la representación de curvas mediante wavelets y se introduce su representación mediante multiwavelets haciéndose una comparativa entre las distintas representaciones.
  - En la Sección 3.3 se proponen distintos modelos de deformación para representaciones de curvas basadas en wavelets en espacios de suavidad de Besov y se resuelve el problema de ajuste de la curva a la imagen.
  - En la Sección 3.4 se proponen modelos dinámicos para la representación anterior y en particular se introducen los modelos autoregresivos de primer y segundo orden.
- 
- En el cuarto capítulo se muestran las aplicaciones de los modelos de dinámicos y de deformación estudiados en el capítulo anterior a problemas de seguimiento de formas tanto en ambientes controlados como en ambientes complejos. Se propone además un algoritmo de emparejamiento basado en multiwavelets que es robusto en condiciones de oclusión parcial debido a la utilización de información local.

De esta forma:

- En la Sección 4.1 se aborda el problema de seguimiento mediante los modelos teóricos propuestos en el Capítulo 3. Para ello se expone la forma de utilizar los modelos dinámicos a priori en espacios de Sobolev mediante el filtro de Kalman abordándose problemas de seguimiento en ambientes controlados. Posteriormente se introduce el problema de seguimiento en ambientes generales mediante la utilización de los modelos dinámicos a priori señalados anteriormente y el filtro Condensation.
- En la Sección 4.2 se utiliza la representación de curvas mediante multiwavelets presentada en el Capítulo 3 y se introduce un algoritmo de emparejamiento robusto frente a oclusiones.

- En el Capítulo 5 se muestran las conclusiones de esta tesis y se hace un repaso a las principales aportaciones de este trabajo. Además se señalan posibles líneas de investigación que, a partir de los resultados presentados en esta tesis, completan o extienden los procedimientos de representación, reconocimiento y seguimiento presentados.

## 2 Reconocimiento y Seguimiento de objetos basado en modelos

A continuación se mostrará una visión global de los problemas de la visión por ordenador basada en modelos y la aproximación utilizada en este trabajo para su resolución. Además se introducirá una notación común correspondiente a la formalización de estos problemas.

Se asumirá que se dispone de un conjunto de modelos  $M_1, M_2, \dots, M_n$ . Cada modelo  $M_i$  está definido por un conjunto de primitivas que tomará generalmente la forma de un vector de parámetros  $w$  sobre un espacio paramétrico  $Q$ . Así cada valor particular del vector  $w$  proporciona un valor específico del modelo que se denota como  $M_i(w)$ .

Por otra parte disponemos de una imagen en la que pretendemos determinar si el modelo se encuentra presente en base a unos datos  $D$  que se extraen de ésta.

Así, los problemas descritos en la Introducción pueden enunciarse como:

### 1. El problema del modelado, que necesita:

- Determinar una representación común para el conjunto de modelos mediante un espacio paramétrico  $Q$ .
- Para cada modelo, determinar el conjunto de posibles parámetros de éste  $W_i \subset Q$ . En el caso rígido  $W_i$  debe de permitir únicamente movimientos rígidos del modelo. En el caso deformable generalmente  $W_i = \mathbb{R}^n$  donde  $n$  es la dimensión del espacio de deformaciones.
- Si se modela un objeto deformable, determinar que deformaciones son más preferidas o probables. Esto suele hacerse comparando el modelo  $M_i(w)$  con un modelo de referencia  $M_i(\bar{w}_i)$  obteniéndose una función de costo o energía de deformación:

$$E_{def} = E_{def}(w, M_i, \bar{w}_i) \quad \text{Ec. 2.1}$$

### 2. El problema del emparejamiento pretende determinar el valor de $w$ de forma que el modelo específico $M_i(w)$ se ajuste al conjunto de datos $D$ extraídos de la imagen

- Determinar un criterio de ajuste implícito o explícito del modelo  $M_i(w)$  a la imagen:

$$E_{aju} = E_{aju}(w, M_i, D) \quad \text{Ec. 2.2}$$

- Determinar una función de ajuste combinado que, si es necesario, utilice ambos criterios :

Caso deformable:

$$E_{tot}(w, \mathbf{t}, M_i, D, \overline{w}_i) = Y (E_{aju}(w, M_i, D), E_{def}(w, M_i, \overline{w}_i), \mathbf{t}) \quad \text{Ec. 2.3}$$

donde  $\mathbf{t}$  es un parámetro de regularización que define la importancia relativa del término de ajuste y el de deformación

- Caso rígido:

$$E_{tot}(w, \mathbf{t}, M_i, D, \overline{w}_i) = E_{aju}(w, M_i, D) \quad \text{Ec. 2.4}$$

( $E_{aju}$  no depende en este caso de  $\mathbf{t}$  y  $\overline{w}_i$ , pero se mantiene la notación)

• Resolver el problema:

$$\text{Hallar } \hat{w} = \arg \min_{w \in W_i} E_{tot}(w, \mathbf{t}, M_i, D, \overline{w}_i) \quad \text{Ec. 2.5}$$

Es interesante señalar que el caso deformable se puede utilizar como una técnica para resolver el caso rígido si se diseña una función  $E_{def}(w, M_i, \overline{w}_i)$  que actúe como función de penalización para buscar la factibilidad ( $w \in W_i$ ) del problema.

En general el problema de optimización global es muy complicado debido a complejidad del término  $E_{aju}(w, M_i, D)$  que provoca la presencia de un gran número de óptimos locales. En determinadas circunstancias (por ejemplo en problemas de seguimiento) es posible obtener un valor  $\hat{w}^-$  cercano a la solución óptima y pueden utilizarse métodos de optimización local.

3. **El problema de la clasificación**, que puede determinarse a través de alguna función  $E_{clas}$  que mida el ajuste del modelo a la imagen a través de una función discriminante  $Z$ :

$$E_{clas} = Z(w, \mathbf{t}, M_i, D, \overline{w}_i) \quad \text{Ec. 2.6}$$

4. **El problema del seguimiento**, que utilizará distintos conjuntos de datos  $D_{1..t} = \{D_1, D_2, \dots, D_{t-1}, D_t\}$ . A continuación se presenta un esquema general de seguimiento:

• Determinar una predicción  $\hat{w}^-$  de la posición del modelo en función de la estimación en imágenes anteriores del objeto mediante una función  $F_{(t)}$ :

$$\hat{w}_t^- = F_{(t)}(\hat{w}_{t-1}) \quad \text{Ec. 2.7}$$

- Determinar una función de la deformación de la predicción  $\hat{w}_t^-$  en la imagen  $t$  construida a partir de la evolución temporal del objeto y el conjunto de datos  $D_{1..t}$ :

$$E_{def(t)} \equiv E_{def(t)}(w_t, M_i, \hat{w}_t^-) \quad \text{Ec. 2.8}$$

- Construir una función de ajuste combinado:

$$E_{tot(t)}(w_t, \mathbf{t}_t, M_i, D_{1..t}, \hat{w}_t^-) = Y(E_{aju(t)}(w_t, M_i, D_t), E_{def(t)}(w_t, M_i, \hat{w}_t^-), \mathbf{t}_t) \quad \text{Ec. 2.9}$$

- Resolver

$$\hat{w}_t = \arg \min_{w_t \in W_i} E_{tot(t)}(w_t, \mathbf{t}_t, M_i, D_{1..t}, \hat{w}_t^-) \quad \text{Ec. 2.10}$$

en este caso es posible utilizar métodos locales de optimización, pues en general  $\hat{w}_t$  estará en un entorno de  $\hat{w}_t^-$ .

- Actualizar  $E_{def}$ ,  $E_{aju}$ ,  $E_{tot}$ ,  $\mathbf{t}$

Nótese que si hay una única imagen con datos  $D_1$  y la predicción es  $\hat{w}_1^- = \bar{w}_i$  esta formulación es equivalente al problema de deformación.

A continuación se pasarán a detallar los problemas anteriores y las soluciones más relevantes que se han aportado en este área.

## 2.1 El Problema del Modelado de Objetos

En este apartado se describirán las distintas técnicas de modelado de objetos, tanto para objetos rígidos como para objetos deformables. De acuerdo con la aproximación anterior, podemos considerar que la diferencia entre un modelo rígido y otro deformable es únicamente la presencia de una función de deformación  $E_{def}(W, M_i, \overline{W}_i)$ , por tanto ambos presentan problemas comunes de modelado. De esta forma en la Sección 2.1.1 se planteará el problema de la representación de modelos y en la Sección 2.1.3 se introducirán diversas formas de la función de deformación  $E_{def}$ .

### 2.1.1 Cuestiones básicas del Modelado de Objetos

Para proporcionar un modelo del objeto cuya presencia y seguimiento se desea establecer en una secuencia de imágenes es necesario resolver dos cuestiones fundamentales:

1. **El problema de las características**, es decir, qué características elegir, y cómo calcularlas de forma eficiente.
2. **El problema de la representación**, cuya solución proporciona un modelo del objeto basado en dichas características.

#### 2.1.1.1 El Problema de la elección de características

El problema de la elección de las características en el esquema de la representación basada en modelos es un paso fundamental, pues en ellas están basadas las primitivas que son las componentes del modelo.

Características de la imagen como: bordes, esquinas, curvas, agujeros y curvatura de los contornos se suelen utilizar para describir los componentes de un objeto. Estas características y sus relaciones espaciales se combinan para generar las representaciones de los objetos. Debido a que estas características representan elementos de alto nivel correspondientes a propiedades físicas de la escena, son menos sensibles al ruido que los datos en bruto a partir de la imagen; proporcionando además una elevada compresión de los datos de entrada.

Generalmente la decisión de qué características usar, es subjetiva y dependiente de la aplicación. De hecho, prácticamente cualquier función de la apariencia de un objeto podría usarse como primitiva. Sin embargo, en la práctica se buscan primitivas que tengan esencialmente las siguientes propiedades:

- **Las primitivas deberían hacer explícita cualquier información necesaria para la tarea a realizar.** Para una tarea particular de reconocimiento de objetos esto significa que las primitivas deben capturar todas las distinciones necesarias para diferenciar los mismos.

- **Las primitivas deberían reflejar las regularidades y estructuras del mundo externo.** Las primitivas deberían, por ejemplo, aprovechar las propiedades que son invariantes con respecto a algunas transformaciones, ya sean de tipo geométrico o debidas a cambios en la iluminación de la escena.

El proceso de extracción de características generalmente empieza mediante la segmentación de la imagen en regiones de propiedades similares, de forma que estas regiones pertenezcan a partes de un mismo objeto. Posteriormente, a partir de las regiones, se extraen las características que determinarán las primitivas utilizadas para representar el modelo del objeto.

A partir de la forma de obtención de las primitivas éstas pueden clasificarse en dos grandes apartados:

- **Primitivas basada en características globales.**  
En este tipo, las características se calculan sobre todo el objeto. Las principales ventajas de esta representación son:
  - Características definidas de forma simple y fáciles de calcular.
  - Generación de modelos de forma sencilla.
  - Utilización de herramientas estadísticas ampliamente utilizadas.
  - Proceso de emparejamiento realizado de forma sencilla y eficiente.

Por la propia naturaleza de las funciones de imagen implicadas en el cómputo, su principal desventaja reside en que los objetos no deben tocarse u ocultarse unos a otros, ni contener defectos.

- **Primitivas basadas en características locales.**  
La especificidad de este método se establece por el hecho de que las características se calculan sobre determinadas zonas del objeto. Las principales ventajas de este tipo de métodos son fundamentalmente:
  - Las características se definen de forma local y por tanto son computables de forma eficiente.
  - La capacidad de trabajar con objetos que tienen puntos de contacto, presentan ocultamientos o bien contienen defectos.

Las principales desventajas derivan del hecho de:

- La utilización de métodos de emparejamiento generalmente heurísticos.
- Proceso de emparejamiento o “*matching*” complejo.
- Generación compleja de modelos.

#### 2.1.1.2 El Problema de la representación

En el reconocimiento de objetos basado en modelos son necesarios dos esquemas de representación: uno para representar el modelo del objeto, y otro para

representar el contenido de la imagen. Ambas representaciones tienen habitualmente significados distintos (generalmente el modelo describe la forma del objeto y la imagen las manifestaciones visibles de dicha forma) sin embargo, para facilitar el problema del emparejamiento entre modelo e imagen, las dos representaciones deberían estar íntimamente relacionadas.

#### 2.1.1.2.1 Características deseables de una representación

Las cualidades que debería tener una correcta representación de un objeto se sintetizan en:

- **Generalidad y sensibilidad**, esto es, la representación debe ser capaz de describir todos los objetos relevantes y preservar las diferencias entre éstos.
- **Unicidad**. Debe haber una única descripción para cada objeto. De esta manera, elementos idénticos tendrán descripciones idénticas, con lo que el problema de comparación se simplifica considerablemente.
- **Estabilidad**. Los cambios pequeños en el objeto deberían generar cambios pequeños en su descripción.
- **Eficiencia**. Debe ser posible el calcular de forma rápida y sencilla la descripción a partir de los datos de entrada. También debe ser posible comparar distintas descripciones de formas eficientemente.

#### 2.1.1.2.2 La elección de un sistema de coordenadas.

Para describir las posiciones relativas de las distintas primitivas, una representación debe emplear algún sistema de coordenadas. Hay principalmente dos formas de definir este sistema de coordenadas para una representación de un objeto tridimensional:

- **Representación basada en el objeto.**

Este tipo de representación describe al objeto 3D mediante sus coordenadas en el espacio con respecto a un sistema de coordenadas [Dickinson 1992]. Así se define al objeto de forma exacta e independientemente del punto de vista, por lo que permite el trabajar con uno arbitrario y con un nivel de precisión también arbitrario. La principal ventaja de este tipo de representación es una descripción del objeto más concisa y precisa. Mientras que su principal desventaja es la dificultad de comparar las coordenadas de la proyección del objeto 3D con las del objeto en la escena.

- **Representación basada en varias vistas del objeto.**

Este tipo de representación describe el objeto 3D mediante un conjunto de vistas 2D del objeto denominadas vistas características. Cada vista característica describe

como aparece el objeto desde un punto de vista o desde varios puntos de vista que generan vistas similares. Esta representación es adecuada cuando el punto de vista está fijo y únicamente son posibles un pequeño número de posiciones estables del objeto. Este tipo de representación trata cada una de las vistas de forma independiente, reduciendo el problema del reconocimiento a un problema bidimensional mediante el uso de primitivas basadas en características 2D y sus relaciones. Para cada punto de vista, un conjunto suficiente de características y sus relaciones se utilizan para modelar el objeto.

La principal ventaja de las representaciones basadas en vistas reside en que es posible realizar el proceso de emparejamiento 2D/2D sin necesidad del cálculo de la proyección del objeto 3D. Éste se lleva a cabo mediante la comparación de descripciones bidimensionales (Ver Sección 2.2) y por tanto no es necesario el proyectar el modelo en el proceso de emparejamiento.

Entre sus principales desventajas están:

- La representación es más compleja debido a que deben de almacenarse varias vistas del mismo objeto para su representación.
- La representación del objeto es menos precisa. Si se quiere limitar la complejidad de la representación, es necesario limitar el número de vistas, esto hace que se pierda capacidad en la representación.
- Mayor complejidad en el reconocimiento. Esto es debido a que existen varios modelos que aparecen adicionalmente asociados a las vistas que deben ser considerados. Este hecho puede compensarse debido a que el emparejamiento de cada modelo requiere una comparación 2D/2D en vez de una 3D/2D o 3D/3D. Además existen características [Mundy 1992], que permanecen invariantes en distintas vistas lo que facilita también el proceso de emparejamiento.

### 2.1.1.2.3 Invarianza

Debido a que es deseable identificar un objeto bajo diferentes condiciones de posición e iluminación, es preciso contar con primitivas que sean invariantes con respecto a cambios de dichas condiciones.

Por ejemplo, si asumimos que un objeto está representado por su silueta y se mueve en un plano perpendicular al eje óptico de la cámara o en la dirección del eje, el modelo del objeto sufrirá una similaridad euclídea, determinada por una rotación, traslación y escalado. Si relajamos la hipótesis de perpendicularidad y permitimos que el plano se incline, el efecto proyectivo se aproxima mediante una transformación afín determinada además por distintos valores de deformaciones sobre dos ejes.

La invarianza frente a transformaciones puede llevarse a cabo de forma implícita mediante la definición de un modelo invariante frente a dicha transformación o de forma directa mediante la estimación de los parámetros de la transformación. Los métodos implícitos requieren estimar menos parámetros y por tanto son generalmente más eficientes que los métodos explícitos puesto que éstos proporcionan la información acerca de la posición del objeto después de resolver el problema del emparejamiento.

Se ha realizado un considerable esfuerzo para intentar identificar y emplear primitivas que sean completamente invariantes con respecto al punto de vista. Estas primitivas están basadas en estructuras geométricas, llamadas invariantes o invariantes geométricos, que permanecen constantes frente a una clase de transformaciones (ver Tabla 2. 2). Estos invariantes pueden calcularse a partir de un conjunto de puntos como la razón doble, mediante los coeficientes de curvas algebraicas, utilizando un conjunto derivadas de orden superior calculadas en un punto de la curva, o a través de ciertas combinaciones de las coordenadas de las primitivas. Sin embargo, su utilización es compleja: el uso de un conjunto de puntos tiene el problema de su detección y agrupamiento, mientras que el uso de los coeficientes de curvas o de derivadas genera el problema de estimar de forma precisa que dichas cantidades.

Transformación	Grados de libertad	Invariantes
Proyección entre planos	8	Concurrencia y colinearidad Orden de contacto: intersección (orden 1); tangencia (orden 2); inflexión (orden 3). Discontinuidades en la tangente Razón doble
Transformación afín	6	Paralelismo Razón de longitudes de segmentos colineales o paralelos Razón de áreas Combinaciones lineales de vectores
Similaridad euclídea	4	Razón de longitudes de segmentos
Movimiento rígido	3	Longitudes, ángulos, áreas

Tabla 2. 2 Propiedades geométricas invariantes en transformaciones entre planos [Mundy 1992].

#### 2.1.1.2.4 Organización de las primitivas

En el caso general de trabajar con primitivas locales y una vez determinado el sistema de coordenadas las primitivas deben de organizarse en algún tipo de estructura.

- **Modelos basados en características estructurales.**

Este tipo de modelos se construyen utilizando primitivas basadas en descripciones abstractas y precisas, como por ejemplo arcos, líneas y esquinas. Estas características son de naturaleza local y cada una describe una porción del objeto. La organización de estas características está altamente estructurada, como por ejemplo en una lista ordenada. Este tipo de métodos permite el uso de técnicas de reconocimiento sintáctico. El emparejamiento se lleva a cabo entonces mediante análisis sintáctico.

- **Modelos basados en grafos.**

En este tipo de modelos, los objetos se representan estructuralmente mediante grafos. Las relaciones geométricas entre las distintas primitivas definen la estructura

relacional, representada mediante un grafo en el cual cada nodo representa una característica local, y está etiquetado con una lista de propiedades o con la pertenencia a una clase. Los arcos representan características relacionales que unen pares de nodos y están etiquetadas con una lista de valores de relación, como por ejemplo distancia y orientación relativa. El reconocimiento del objeto se convierte entonces en un problema de emparejamiento de grafos.

### **2.1.2 Representación de modelos basados en contornos**

En este trabajo, el conjunto de primitivas estará formado a partir de los elementos de contorno extraídos utilizando los bordes de la imagen. Las razones por las cuales el empleo de los bordes de una imagen es frecuente en diversas áreas de la visión artificial es debido al hecho de que los límites de los objetos tienden a mostrarse como discontinuidades en la función de intensidad de la imagen. Los experimentos realizados en personas muestran que la detección de contornos es un elemento fundamental del sistema de visión humano [Yen 1996]. La razón por la cual los seres humanos están dotados con un sistema de visión capaz de trabajar a partir de la utilización de la información de los bordes puede ser debido a que la habilidad de detectar discontinuidades en superficies bidimensionales da una ventaja competitiva. Las discontinuidades en las superficies bidimensionales son en general unidimensionales, esto es, curvas. Tales curvas pueden representar la presencia de otro objeto que se diferencia del fondo debido a la discontinuidad en sus texturas, cambios en la geometría local de los objetos, o en el cambio de iluminación debido a la distancia. Además una descripción de una imagen en términos de bordes es mas concisa y compacta que los valores en bruto de los elementos de la imagen. La elevada cantidad de información que proporcionan los bordes de una imagen hace que una señal adecuadamente filtrada pueda ser reconstruida a partir de información extraída de sus bordes [Hildreth 1987], [Mallat 1992].

Sin embargo, la detección precisa por un ordenador de los contornos de un objeto es una tarea difícil debido a razones de naturaleza semántica, pues la búsqueda de los contornos de un objeto hace necesaria la definición para el ordenador de un modelo del objeto, y a razones de tipo físico, debido a la presencia de ruido, y a las diferentes fuentes que pueden haberlos generado. Esto ha hecho se utilicen diferentes tipos de detectores de bordes, desde los más sencillos: operador de Roberts [Roberts 1965], Sobel [Gonzalez 1987] y Prewitt [Prewitt 1970], hasta los más sofisticados: Laplaciano del Gausiano (LoG) [Marr 1980], operador de Canny [Canny 1986], (ver [Ziou 1998] para una revisión de las diferentes técnicas). Dependiendo de la complejidad de la escena estos bordes pueden transformarse en contornos mediante análisis local [Gonzalez 1987], la transformada de Hough [Duda 1972], [Ballard 1982], o programación dinámica [Martelli 1972]

#### **2.1.2.1 Técnicas de representación de contornos.**

Entre los distintos métodos de descripción de contornos, es una práctica habitual distinguir entre dos familias de técnicas diferentes, aquellas que describen el interior de una región y aquellas que describen su frontera. Sin embargo, para ambos

tipos de descriptores un requerimiento esencial es el de su invarianza frente a un cierto tipo de transformaciones geométricas. En [Gonzalez 1987], [Sonka 1993], [Loncaric 1998] es posible encontrar un estudio sobre los distintos tipos de descriptores de contornos. A continuación se presentan los métodos más significativos.

### 2.1.2.2 Descriptores de Contornos.

Los descriptores de contorno son útiles cuando se pretenden caracterizar los bordes extraídos de una escena. Algunas técnicas basadas exclusivamente en el contorno de los objetos, se describen a continuación.

#### 2.1.2.2.1 Códigos de cadena.

Una forma de comprimir la descripción exhaustiva de un contorno es utilizar los códigos de cadena. [Freeman 1961]. Dado un punto del contorno, el siguiente punto se busca en un conjunto de direcciones basado en el modelo de entornos de la imagen, conectividad a 4 o conectividad a 8 (Ver Figura 2. 1). En el caso de que los códigos representen contornos cerrados éstos puedan transformarse en invariantes frente a traslación mediante la elección de un punto de comienzo de forma que la sucesión resultante de códigos de cadena forme un entero de magnitud mínima [Tsai 1985]. La “derivada”, (diferencia a módulo cuatro u ocho), del código de cadena tiene la propiedad de ser invariante frente a rotación. Los códigos de cadena se utilizan habitualmente para codificar la frontera de objetos binarios en sistema de procesamiento de imágenes. Su principal ventaja es la obtención de una representación compacta de la frontera de forma eficiente, mientras que su principal desventaja es la elevada sensibilidad al ruido de la codificación junto con el bajo contenido semántico de los símbolos empleados en la representación.

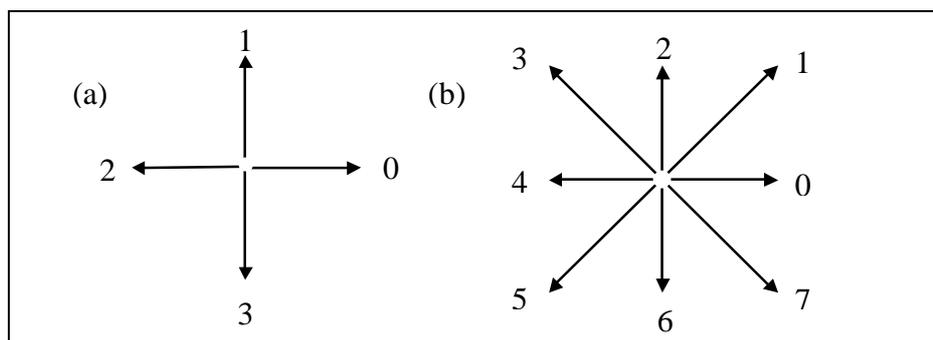


Figura 2. 1 Direcciones para conectividad a 4 (a) y conectividad a 8 (b)

#### 2.1.2.2.2 Aproximaciones poligonales.

La representación de un contorno por medio de segmentos con un determinado tipo de propiedades es otra opción para la descripción de contornos. Si cada segmento

tiene asociado un tipo o clase, el contorno puede ser descrito como una cadena de tipos de segmento, siendo entonces el código de cadena un caso particular de esta representación. Una aproximación poligonal representa un contorno por medio de un polígono. Ésta técnica se basa en el hecho de que cualquier contorno puede ser aproximado, con precisión arbitraria por un sucesión de líneas rectas. (Ver Figura 2. 2) Existen varias técnicas para la aproximación de curvas, como puede encontrarse en [Pavlidis 1977] y [Gonzalez 1987]. (Ver [Rosin 1997] para una comparativa). En la práctica la aproximación busca optimizar algún tipo de disimilaridad y el tiempo de cálculo es del orden  $O(n^2 \log n)$  con  $n$  el número de puntos del contorno, para varios de estos tipos [Imai 1988]. La principal ventaja de éstas aproximaciones es la elevada compresión que se obtiene en la descripción para contornos regulares, mientras que sus principales desventajas son la sensibilidad frente a variaciones del contorno.

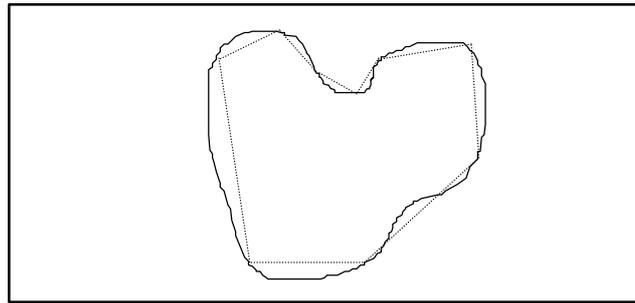


Figura 2. 2 Contorno de un objeto y aproximación poligonal (en línea discontinua).

### 2.1.2.2.3 Firmas (Signatures)

Una firma es una representación funcional unidimensional del contorno. Ejemplos de firmas son: el valor del arco de curva a la izquierda o sobre la tangente en un punto [O'Rourke 1985] o la distancia del centroide de un contorno a los puntos de éste [Gonzalez 1987] (Ver Figura 2. 3). En este caso, la principal ventaja de este tipo de representación es su capacidad de representar de forma concisa el contorno, mientras que su principal desventaja radica en la dificultad en la normalización frente a transformaciones del contorno. Los momentos unidimensionales y los descriptores de Fourier [Peerson 1977] y [Gonzalez 1987] utilizan las firmas como representación intermedia, antes de calcular los descriptores definitivos.

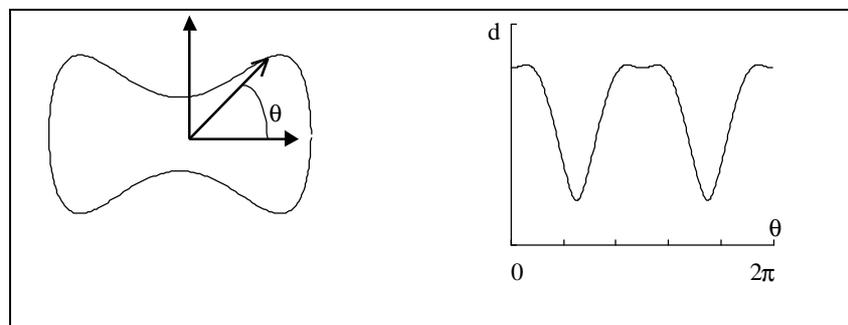


Figura 2. 3. Distancia  $d$  del contorno a su centroide en función del ángulo  $\theta$

2.1.2.2.4 Transformadas de Fourier del contorno.

Los descriptores de Fourier permiten la codificación, reconstrucción y clasificación de curvas bidimensionales. En [Ballard 1982] y [Sonka 1993] puede encontrarse diversos descriptores basados en la transformada de Fourier sobre el contorno o sobre alguna de sus firmas. En [Zahn 1972] puede encontrarse la primera definición elaborada de los descriptores de Fourier, siendo ésta refinada más tarde por [Persoon 1977]. La forma de calcular estos descriptores es la siguiente:

Sea un contorno cerrado  $\mathbf{C}$  con representación paramétrica  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))$ . Se reparametriza la curva de forma que ésta se recorra con velocidad unidad mediante un nuevo parámetro  $s$

$$s = s(u) = \int_a^u |\dot{\mathbf{r}}(v)| dv \quad \text{con } \dot{\mathbf{r}} = \frac{d\mathbf{r}}{du} \tag{Ec. 2.11}$$

y se hace que el parámetro  $s$  varíe entre 0 y  $2\pi$  mediante el cambio de variable  $v = 2\pi s / L$ , donde  $L$  es la longitud de la curva. De esta forma se puede calcular la expansión de Fourier de las funciones  $x$  e  $y$  obteniéndose:

$$\begin{pmatrix} x(s) \\ y(s) \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \sum_{k=1}^{\infty} \begin{pmatrix} a_k & b_k \\ c_k & d_k \end{pmatrix} \begin{pmatrix} \cos ks \\ \sin ks \end{pmatrix} \quad \text{donde:} \tag{Ec. 2.12}$$

$$a_k = \frac{1}{2p} \int_0^{2p} x(s) \cos ks ds \quad b_k = \frac{1}{2p} \int_0^{2p} x(s) \sin ks ds$$

$$c_k = \frac{1}{2p} \int_0^{2p} y(s) \cos ks ds \quad d_k = \frac{1}{2p} \int_0^{2p} y(s) \sin ks ds$$

Puesto que al trabajar con imágenes digitales las coordenadas de los puntos del contorno son discretas los descriptores pueden calcularse mediante la transformada discreta de Fourier. Los descriptores de Fourier pueden hacerse invariantes frente a traslación rotación y reflexión [Jain 1989].

Otro tipo de descriptores de Fourier puede obtenerse a partir de una función unidimensional  $\mathbf{x}^*$  (ver Ec. 2.14) obtenida del contorno de la siguiente forma:

Se reparametriza la curva de forma que ésta se recorra con velocidad unidad mediante (Ec. 2.11). De esta forma se obtiene otra representación de la curva que puede expresarse en función del parámetro  $s$ . Se denotará por  $\mathbf{q}(s)$  a la dirección angular de la curva  $\mathbf{C}$  para el valor de la longitud de arco  $s$  y se definirá la función de acumulación angular  $\mathbf{x}(s) = \mathbf{q}(s) - \mathbf{q}(0)$  (Figura 2. 4):

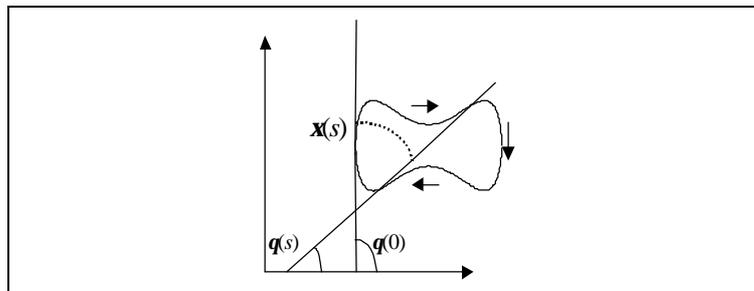


Figura 2. 4. Definición de las funciones  $q$  y  $x$

Esta función así definida es periódica con período  $L$ , con  $L$  la longitud de la curva. Puesto que el período ideal para realizar la transformada de Fourier es  $2\pi$ ,  $\mathbf{x}(s)$  se normaliza usando un nuevo parámetro  $u$ , que varía entre 0 y  $2\pi$ . Para compensar el hecho de que cuando se recorre la curva se acumula un ángulo de  $-2\pi$ , se añade un término aditivo para obtener una función periódica y se define:

$$\mathbf{x}^*(u) = \mathbf{x}\left(\frac{Lu}{2p}\right) + u \quad \text{Ec. 2.13}$$

Finalmente se realiza la expansión de Fourier de  $\xi^*(s)$ , lo cual nos lleva a los descriptores  $(A_k, \mathbf{a}_k)$ :

$$\mathbf{x}^*(t) = \mathbf{m}_0 + \sum_{k=1}^{\infty} A_k \cos(ku - a_k) \quad \text{Ec. 2.14}$$

Los coeficientes  $A_k$  son las amplitudes armónicas de la expansión de Fourier y los coeficientes  $\mathbf{a}_k$  son los ángulos de fase. Puede demostrarse que los  $A_k$  son invariantes frente a rotaciones y punto de inicio en el contorno. La normalización frente a cambio de escala se obtiene dividiendo cada uno de los  $A_k$  por  $A_1$  que suele ser además el coeficiente de mayor magnitud. El ruido debido a la discretización del contorno hace que la capacidad de discriminación de los  $\mathbf{a}_k$  sea pequeña [Gonzalez 1987]. Las principales ventajas de este método residen en la facilidad de normalización de los descriptores y en el hecho de que sea posible reconstruir la curva a partir de los coeficientes. La principal desventaja reside en que los contornos a procesar han de ser cerrados.

#### 2.1.2.2.5 Descripción mediante b-splines

La representación de curvas mediante splines [Boor 1978] para obtener curvas se ha utilizado frecuentemente en los gráficos por ordenador debido a su capacidad para representar formas perceptualmente agradables debido a su suavidad. Un spline de orden  $d$  es una función polinomial a trozos, consistente en la concatenación de segmentos polinomiales, donde cada polinomio tiene orden (número de coeficientes) igual a  $d$  y se unen en un conjunto de puntos denominados nodos. Las curvas sencillas pueden representarse con un número reducido de segmentos. Si son más complejas puede aumentarse el grado del polinomio igual a  $d-1$ , pero generalmente es preferible aumentar el número de segmentos. Generalmente el grado del polinomio utilizado es cuadrático ( $d=3$ ) o cúbico ( $d=4$ ).

Los b-splines son una forma conveniente computacionalmente de representar los splines. Mediante esta representación una función spline  $x(u)$  puede construirse mediante la combinación lineal de  $N_b$  funciones básicas  $\mathbf{j}_k(u)$ . En el caso general, cada función básica consiste en  $d$  polinomios definidos sobre un segmento en el eje  $u$ . Generalmente se toma cada segmento de longitud unidad y éstos se unen en los nodos como se muestra en la Figura 2. 5

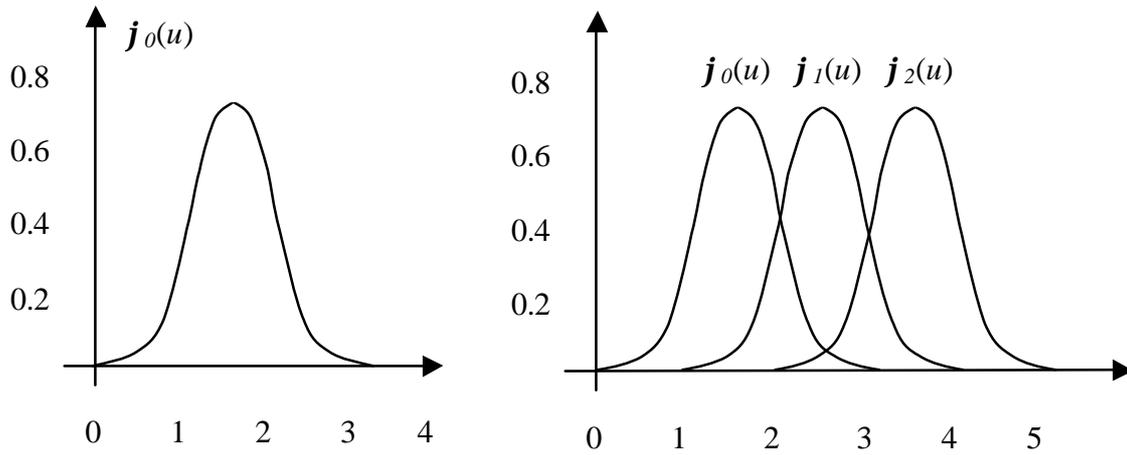


Figura 2. 5 (Izquierda) Función spline cuadrática  $j_0(u)$ . Los Nodos en  $s=0,1,2,3,4$  marcan la transición entre los segmentos polinomiales de la función. (Derecha) en el caso regular con segmentos espaciados en los valores enteros de  $u$ , cada función básica es una copia trasladada de la anterior

De esta forma la función puede expresarse como:

$$x(u) = \sum_{k=0}^{N_b-1} a_k^x \mathbf{j}_k(u) \quad \text{Ec. 2.15}$$

Donde  $a_k^x$  son los pesos aplicados a las funciones básicas respectivas  $\mathbf{j}_k(u)$ . Esto puede expresarse con notación matricial como:

$$x(u) = (\mathbf{a}^x)^T \mathbf{j}(u), \quad \mathbf{j}(u) = (\mathbf{j}_0(u), \mathbf{j}_1(u), \dots, \mathbf{j}_{N_b-1}(u))^T \quad \text{Ec. 2.16}$$

$$\mathbf{a}^x = (a_0^x, a_1^x, \dots, a_{N_b-1}^x)^T$$

Por convenio, las funciones básicas se construyen de forma que sumen la unidad en todos los puntos:

$$\forall u, \quad \sum_{k=0}^{N_b-1} \mathbf{j}_k(u) = 1 \quad \text{Ec. 2.17}$$

En el caso más simple de un b-spline cuadrático con nodos regularmente espaciados en intervalos de longitud unidad, la primera función básica tiene la forma:

$$\mathbf{j}_0(u) = \begin{cases} u^2/2 & 0 \leq u \leq 1 \\ 3/4 - (u - 3/2)^2 & 1 \leq u \leq 2 \\ (u - 3)^2/2 & 2 \leq u \leq 3 \\ 0 & \text{otro caso} \end{cases} \quad \text{Ec. 2.18}$$

Mientras que las siguientes son copias trasladadas:

$$\mathbf{j}_k(u) = \mathbf{j}_0(u - k) \quad \text{Ec. 2.19}$$

En los casos prácticos, se trabaja siempre con un número finito de funciones básicas por lo que es necesario generar bases finitas sobre un intervalo  $0 \leq u \leq L$ . Estas pueden ser periódicas o aperiódicas. En el caso periódico las funciones básicas son  $\mathbf{j}_0, \mathbf{j}_1, \dots, \mathbf{j}_{L-1}$  definidas de forma periódica sobre el intervalo  $[0, L]$ . Una base no periódica es más difícil de construir requiriendo nodos “múltiples” en los extremos del intervalo mediante la disminución del grado de continuidad del spline a través de la repetición de su valor. Para obtener una representación operativa es necesario obtener los coeficientes del modelo de b-splines para una función discretizada  $x(k)$ . Para splines de grado 0 (constantes a trozos) y splines de grado 1 (lineales a trozos), es sencillo puesto que  $a_k^x = x(k)$ . Para splines de orden superior, la situación es más compleja. Tradicionalmente, el problema de interpolación mediante splines ha sido resuelto mediante técnicas de álgebra lineal resolviendo un sistema con bandas diagonales [Boor 1978]. Recientemente [Unser 1993] se han desarrollado técnicas más eficientes basadas en filtrado digital y de complejidad lineal.

Para construir una representación paramétrica de una curva en el plano  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))$ , basta aplicar la representación en b-splines a ambas funciones paramétricas obteniendo:

$$\mathbf{r}(u) = \sum_{k=0}^{N_b-1} \mathbf{a}_k \mathbf{j}_k(u), \quad \mathbf{a}_k = (a_k^x, a_k^y)^T \quad \text{Ec. 2.20}$$

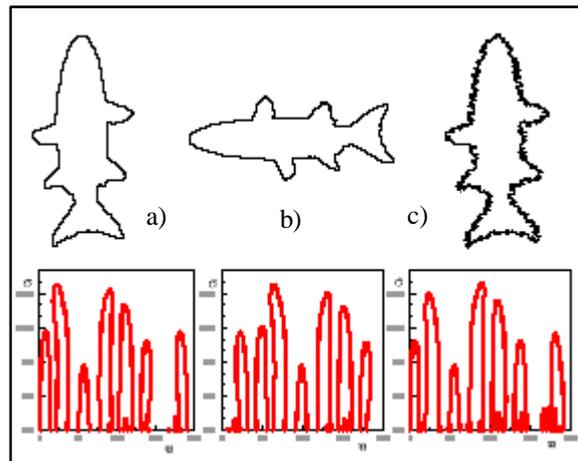
La principal ventaja de esta representación es la obtención de curvas suaves de forma eficiente mientras que su principal inconveniente es la dificultad de representar curvas con cambios bruscos en la orientación.

La representación de contornos basada en b-splines ha sido utilizada en parametrización de contornos activos “snakes” [Brigger 2000] o problemas de seguimiento [Blake 1998].

#### 2.1.2.2.6 Técnicas en el espacio de escalas

Este grupo de técnicas se basan en la representación del contorno del objeto mediante el espacio de escalas. En [Witkin 1983] se propone un filtrado en el espacio de escalas que proporciona una representación útil para la descripción de las características significativas del objeto. La representación se crea mediante el seguimiento de la posición de los puntos de inflexión en señales filtradas mediante filtros de paso bajo gaussianos con amplitudes variables. Los puntos de inflexión que se mantienen presentes en la representación se definen como características significativas del objeto.

La técnica del espacio escalas puede utilizarse para la descripción de formas planas utilizando la imagen del espacio de escalas [Mokhtarian 1992]. Para ello se calcula la curvatura a lo largo del contorno y se suaviza mediante un filtro gaussiano de diferente amplitud. Los cruces con cero de la función de curvatura constituyen la imagen de la función de curvatura en el espacio de escalas que se construye colocando en un eje el valor del parámetro del contorno y en el otro el valor  $\mathbf{s}$  que define el filtro gaussiano. Se obtiene entonces una descripción jerárquica de formas que invariante a traslación, escalado, y rotación (ver Figura 2. 6).



**Figura 2. 6** Espacio de escalas de curvatura para un contorno (a), una rotación de éste (b) y una versión ruidosa (c). (de [Mokhtarian 1996]).

La principal ventaja de esta representación es su invarianza frente a similitudes euclídeas mientras que su principal inconveniente es su tiempo de cálculo. Esta representación multiescala ha sido extendida a b-splines [Wang 1999] y wavelets [Yoon 1998].

#### 2.1.2.2.7 Transformada wavelet del contorno.

En este apartado se situará la representación de contornos basada en wavelets en el contexto de la representación de contornos. Un estudio detallado de sus propiedades se realizará en la Sección 3.1 del Capítulo 3.

La transformada mediante wavelets aparece como una solución a la falta de localización de la transformada de Fourier. Mediante la transformada de Fourier, una función definida en el dominio temporal o espacial se transforma en una función en el dominio de frecuencias. La transformada de Fourier asume implícitamente que la función original es de naturaleza periódica. Es por esta razón que la transformada de Fourier tiene problemas con las componentes de la señal localizadas en el espacio. En los últimos años, han sido descubiertas nuevas familias de funciones ortogonales [Mallat 1989] que definen transformaciones de forma que este problema queda resuelto. Estas funciones básicas son las llamadas “wavelets”, y al contrario que los senos y cosenos que forman las funciones básicas de la transformada de Fourier, no necesitan tener soporte infinito. Este soporte compacto permite a la transformada wavelet el trasladar una función en el dominio del espacio a otra función que está localizada no sólo en frecuencia sino también en el dominio espacial.

Los wavelets que se utilizarán en este trabajo son funciones  $\mathbf{y}(u)$  cuyas traslaciones y dilataciones:

$$\frac{1}{\sqrt{|s|}} \mathbf{y}\left(\frac{u-t}{s}\right)_{(s,t) \in \mathbb{R} - \{0\} \times \mathbb{R}}, \quad u \in \mathbb{R} \quad \text{Ec. 2.21}$$

pueden ser usadas para la expansión de las funciones de  $L_2(\mathbb{R})$ . Asociado a una función wavelet puede encontrarse una función de escala  $\mathbf{f}$  que proporciona una aproximación

de la función  $f$  a una escala determinada. La expansión de una función  $f$  en este conjunto de funciones forman su transformada wavelet.

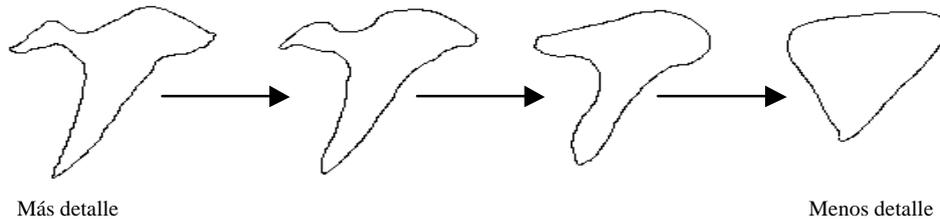
Puede demostrarse que existen wavelets  $\mathbf{y}(u)$  tales que:

$$(\sqrt{2^j} \mathbf{y}(2^j u - k))_{(j,k) \in \mathbb{Z} \times \mathbb{Z}} \quad \text{Ec. 2.22}$$

forman un conjunto ortogonal. Esto permite utilizar dicho conjunto de funciones para la descomposición ortogonal de funciones (Ec. 2.23) obteniéndose entonces una descripción ortogonal de una función  $f$  mediante la transformada wavelet diádica. Esta transformada puede calcularse de forma eficiente mediante la transformada rápida de wavelets cuya complejidad es lineal [Mallat 1989].

$$f(u) = \sum_{k \in \mathbb{Z}} c_{J_0, k} 2^{\frac{J_0}{2}} f(2^{J_0} u - k) + \sum_{j=J_0+1}^{\infty} \sum_{k \in \mathbb{Z}} d_{j, k} 2^{\frac{j}{2}} \mathbf{y}(2^j u - k) \quad \text{Ec. 2.23}$$

De esta forma, la transformada wavelet permite estudiar la señal original a distintos niveles de resolución, dependiendo del valor del parámetro de escala. Su aplicación a un contorno va mostrando los detalles del contorno a las diferentes escalas en una representación a varias resoluciones.



**Figura 2. 7 Descomposición del contorno a varias escalas mediante wavelets**

La utilización de wavelets para la descripción de contornos ha sido propuesta por diferentes autores. Como ejemplo de esta metodología en [Pfeiffer 1995] se utilizaron los descriptores basados en wavelets para compararlos con la Transformada de Fourier elíptica mostrándose una mejor compresión de la información. En [Chuang 1996] puede encontrarse una teoría matemática de los descriptores basados en wavelets así como algunas aplicaciones. En su trabajo demostraron que los descriptores basados en wavelets tienen propiedades deseables tales como una representación multiescala, invarianza, unicidad, estabilidad y localización espacial.

La principal ventaja de esta representación es la elevada compresión de la información y la velocidad de cálculo de la representación (orden lineal) que proporciona, mientras que su principal inconveniente es la falta de invarianza frente al punto de comienzo del contorno.

Otro tipo de aplicaciones incluyen: reconocimiento de caracteres [Wunsch 1995], contornos activos “snakes” basados en wavelets [Knoll 1999], [Yoshida 1997], reconocimiento de objetos [Tieng 1997], [Cheick 2000], selección de escalas naturales [Quddus 2000] y aplicaciones en sistemas multimedia [Gatica-Pérez 1999], [Muller 1999].

### 2.1.2.3 Descriptores de regiones

A continuación se mostrarán brevemente para ilustrar las diferencias con los métodos de descripción de contornos, una serie de métodos que representan un objeto por su extensión interior en lugar de por su contorno.

#### 2.1.2.3.1 Momentos

El concepto de momentos como características invariantes de la imagen fue introducido en [Hu 1962]. La invarianza se consigue frente a cambios de escala, traslación, rotación y reflexión. Estos invariantes fueron utilizados para el reconocimiento de imágenes de diferentes tipos de aviones a partir de sus siluetas por [Gonzalez 1987]. A continuación se muestra la presentación de los momentos invariantes dada por [Khotanzad 1990].

Sea  $f(x,y) \geq 0$  una función real en una región finita  $S$ . Se define el momento de orden  $p+q$  como :

$$m_{pq} = \int_S f(x,y)x^p y^q dx dy \quad p, q = 0,1,2,\dots \quad \text{Ec. 2.24}$$

o, en el caso de que se tenga una imagen bidimensional de tamaño  $M \times M$   $\{ f(x,y); x, y = 0, \dots, M-1 \}$  se define como:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} x^p y^q f(x,y) \quad \text{Ec. 2.25}$$

Para hacer estos momentos invariantes a la traslación se define el momento central como:

$$\mathbf{m}_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} (x - \bar{x})^p (y - \bar{y})^q f(x,y) \quad \text{con: } \bar{x} = \frac{m_{01}}{m_{00}} \text{ e } \bar{y} = \frac{m_{10}}{m_{00}} \quad \text{Ec. 2.26}$$

Los momentos centrales se pueden normalizar para que sean invariantes frente al cambio de escala definiendo:

$$\mathbf{h}_{pq} = \frac{\mathbf{m}_{pq}}{\mathbf{m}_{pq}^g} \quad \mathbf{g} = \frac{p+q}{2} + 1 \quad \text{Ec. 2.27}$$

A partir de estos momentos se han desarrollado un conjunto de funciones no lineales definidas sobre  $\mathbf{h}_{pq}$  que son invariantes a la rotación, traslación y cambio de escala. Son las siguientes:

$$G_1 = \mathbf{h}_{20} + \mathbf{h}_{02}$$

$$G_2 = (\mathbf{h}_{20} - \mathbf{h}_{02})^2 + 4\mathbf{h}_{11}^2$$

Ec. 2.28

$$G_3 = (\mathbf{h}_{30} - 3\mathbf{h}_{12})^2 + (3\mathbf{h}_{21} - \mathbf{h}_{03})^2$$

$$G_4 = (\mathbf{h}_{30} + \mathbf{h}_{12})^2 + (\mathbf{h}_{21} + \mathbf{h}_{03})^2$$

$$G_5 = (\mathbf{h}_{30} - 3\mathbf{h}_{12}) (\mathbf{h}_{30} + \mathbf{h}_{12}) \left[ (\mathbf{h}_{30} + \mathbf{h}_{12})^2 - 3(\mathbf{h}_{21} + \mathbf{h}_{03})^2 \right] \\ + (3\mathbf{h}_{21} - \mathbf{h}_{03}) (\mathbf{h}_{21} + \mathbf{h}_{03}) \\ \cdot \left[ 3(\mathbf{h}_{30} + \mathbf{h}_{12})^2 - (\mathbf{h}_{21} + \mathbf{h}_{03})^2 \right]$$

$$G_6 = (\mathbf{h}_{20} - \mathbf{h}_{02}) \left[ (\mathbf{h}_{30} + \mathbf{h}_{12})^2 - (\mathbf{h}_{21} + \mathbf{h}_{03})^2 \right] + 4\mathbf{h}_{11} (\mathbf{h}_{30} + \mathbf{h}_{12})(\mathbf{h}_{03} + \mathbf{h}_{21})$$

Los valores numéricos de  $G_1$  hasta  $G_6$  son muy pequeños. Para solventar los problemas de precisión, generalmente se utiliza el logaritmo del valor absoluto de estas seis funciones como características para representar la imagen. La principal ventaja de este método es su adecuación para discriminar entre regiones sencillas, y su principal desventaja es la existencia de información redundante entre los coeficientes lo que no permite discriminar entre regiones complejas.

### 2.1.2.3.2 Transformada de Fourier 2-D

Mientras que los métodos anteriores analizan el objeto en el dominio del espacio, la transformada de Fourier analiza la señal en el dominio de frecuencias. La transformada de Fourier ha sido ampliamente utilizada como herramienta en el procesamiento de imágenes [Gonzalez 1978], [Bow 1992] y [Russ 1992].

Para ello se modela la imagen como una función  $f(x,y)$  definida en  $\mathbb{R}^2$  y se calcula su transformada de Fourier como:

$$F\{f(x, y)\} = F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(xu+yv)} dx dy \quad \text{Ec. 2.29}$$

En general, la transformada de Fourier es una función compleja  $F(u,v) = |F(u,v)| e^{jq(u,v)}$ . A  $|F(u,v)|$  se le llama espectro de Fourier y a  $q(u,v)$  se le llama fase de la transformada de Fourier. Si la función continua está discretizada en una matriz de tamaño  $M \times M$ , la transformada discreta de Fourier se describe como:

$$F\{f(x, y)\} = F(u, v) = \frac{1}{M} \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} f(x, y) e^{\frac{-j2\pi(ux+vy)}{M}} \quad \text{Ec. 2.30}$$

A la hora de trabajar con representaciones invariantes es habitual el trabajar con el módulo  $|F(u, v)|$  debido a su invarianza frente a rotaciones y traslaciones, es decir, se elimina la información de fase. Las principales ventajas de este método residen en la facilidad de normalización de los descriptores y en el hecho de que sea posible reconstruir la imagen a partir de los coeficientes. La principal desventaja reside en que el número de características a utilizar es muy grande y en que el valor de los valores de  $F(u, v)$  para valores próximos de  $u$  y de  $v$  contienen prácticamente la misma información para propósitos de clasificación.

### 2.1.2.3.3 Desarrollos ortogonales

La transformada discreta de Fourier realiza la expansión de  $f(x, y)$  en un conjunto de funciones ortogonales. Utilizando otros conjuntos de funciones ortogonales  $\{f_{nm}\}$  se pueden obtener otros tipos de expansiones. Así por ejemplo, en [Jain 1989] se utilizan los polinomios ortogonales de Legendre, en [Teague 1980] se utilizan los polinomios ortogonales de Zernike, en [Méndez 1984] se utilizan las funciones de Fourier-Bessel. La invarianza frente a rotaciones para estos métodos suele obtenerse a partir de una descomposición radial-angular de las funciones ortogonales en la forma:

$$f_{nm}(x, y) = f_{nm}(\mathbf{r}, \mathbf{q}) = R_{nm}(\mathbf{r}) e^{jm\mathbf{q}} \quad \text{Ec. 2.31}$$

Lo que permite que el módulo de los descriptores sea invariante frente a rotaciones. La función radial podrá expresarse de diferentes formas. Así para la transformada de Zernike se tiene:

$$R_{nm}(\mathbf{r}) = \sum_{k=0}^{n-|m|/2} \frac{(-1)^s [(n-k)!]^{n-2k} \mathbf{r}^{n-2k}}{s! (\frac{n+|m|}{2} + k)! (\frac{n+|m|}{2} - k)!} \quad \text{Ec. 2.32}$$

En el caso de utilizar la transformada de Fourier-Bessel la función radial tiene la siguiente forma:

$$R_{nm}(\mathbf{r}) = \frac{j_{nm} J_m(\mathbf{r})}{j_{nm}^2 - \mathbf{r}^2} \quad \text{Ec. 2.33}$$

donde:

$J_m$  es la función de Bessel de orden  $m$ , y

$j_{nm}$  es el  $n$ -ésimo cero de la función de Bessel de orden  $m$ .

Otro tipo de funciones radiales que generan desarrollos ortogonales se caracterizan a través de la transformada de Fourier-Laguerre. Esta transformada tiene una función radial:

$$R_{nm}(\mathbf{r}) = e^{-\frac{r^2}{2}} \mathbf{r}^n L_n^m(\mathbf{r}^2) \quad \text{Ec. 2.34}$$

con  $L_n^m$  polinomio de Laguerre de ordenes  $n, m$

Otro tipo de representaciones ortogonales utilizadas incluyen la transformada wavelet bidimensional [Mallat 1989] que es una generalización de la transformada wavelet unidimensional y se genera a partir de una función de escala unidimensional  $f$  y una función wavelet unidimensional  $y$  mediante:

$$\begin{aligned} F(x, y) &= f(x)f(y) && \text{Ec. 2.35} \\ Y^1(x, y) &= f(x)y(y) \\ Y^2(x, y) &= y(x)f(y) \\ Y^3(x, y) &= y(x)y(y) \end{aligned}$$

Los coeficientes asociados a la función de escala  $F$  proporcionan una aproximación de la función bidimensional, mientras que los asociados a  $Y^1$ ,  $Y^2$  y  $Y^3$  proporcionan detalles asociados con la dirección horizontal, vertical y diagonal respectivamente como se ve en la Figura 2. 8

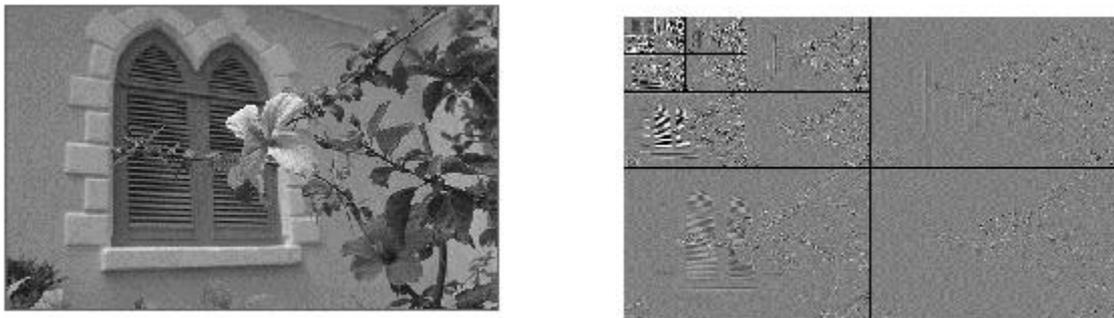


Figura 2. 8 (Izquierda) imagen original. (Derecha) transformada wavelet de la imagen

La transformada wavelet bidimensional ha sido utilizada para el reconocimiento de caracteres [Tang 1996], análisis de imágenes médicas [Carmona 1995], compresión [Lewis 1992], restauración [Banham 1996] o emparejamiento [Pan 1996].

La principal ventaja de este tipo de descripción es que la condición de ortogonalidad permite una descripción no redundante de las regiones mientras que su principal inconveniente es su costo computacional.

### 2.1.3 Modelado de Objetos Deformables

Las distintas representaciones descritas hasta ahora representan la forma de un modelo rígido cuya descripción consiste en un conjunto de primitivas junto con sus relaciones. Para extender el ámbito de los métodos basados en modelos a reconocer también objetos no rígidos, o bien se permite que varios modelos rígidos representen cada objeto para ser capaces de detectar todas las posibles deformaciones, lo cual en general no es factible debido a que sería necesario trabajar con una enorme base de modelos, o bien permitimos que el modelo se deforme.

De esta forma se introducen los modelos deformables (ver [McInerney 1996], [Cheung 1998] para una revisión) que poseen la capacidad de representar objetos no rígidos mediante la deformación de su modelo asociado. Para ello se utiliza generalmente una función de deformación  $E_{def}$  (Ec. 2.1). También es posible utilizar un conjunto de restricciones en los parámetros libres que determinan la deformación, en este caso  $W_i \neq \emptyset$ . Las ventajas que utilizar modelos deformables son entre otros:

- Capacidad de detectar objetos no rígidos de la imagen.
- Mayor robustez frente a ruido y oclusión.
- Capacidad de segmentar la imagen basada en el conocimiento; es decir la segmentación se hace con el objetivo de extraer un objeto en vez de regiones homogéneas.

#### 2.1.3.1 Técnicas de modelado de objetos deformables

Las distintas aproximaciones a los modelos deformables pueden dividirse en tres grandes categorías [Cheung 1998]: modelos paramétricos, semi-paramétricos y modelos no paramétricos.

En la aproximación paramétrica se asume que el modelo y su deformación está caracterizada por un pequeño conjunto de parámetros. Por el contrario, en la aproximación no paramétrica generalmente el número de parámetros que definen la deformación es muy *grande* y difícilmente interpretable.

Por ejemplo, la representación de un círculo puede ser hecha paramétricamente especificando simplemente la localización del centro y el radio o bien utilizando un conjunto de puntos ordenados que serán interpolados para generar una aproximación de éste. La última representación requieren un número relativamente grande de puntos como parámetros comparado con la primera, la cual requiere únicamente tres. En el caso extremo donde el número de puntos sea igual al número de "píxeles" que forman el círculo, la representación se convierte en puramente no paramétrica. Esta perspectiva sugiere que la distinción entre paramétrico y no paramétrico forma un continuo de dos casos extremos en vez de dos clases claramente delimitadas. Sin embargo para facilitar la descripción se llamarán a los modelos con *muchos* parámetros como no paramétricos y aquellos caracterizados con un conjunto compacto de parámetros como paramétricos. Los modelos con que combinan ambos tipos de representación a diferentes niveles se llamarán semiparamétricos.

En esta sección se estudiarán las características de deformación de los modelos, mientras que la función de ajuste a la imagen y los métodos de optimización se estudiarán en la Sección 2.2.2 .

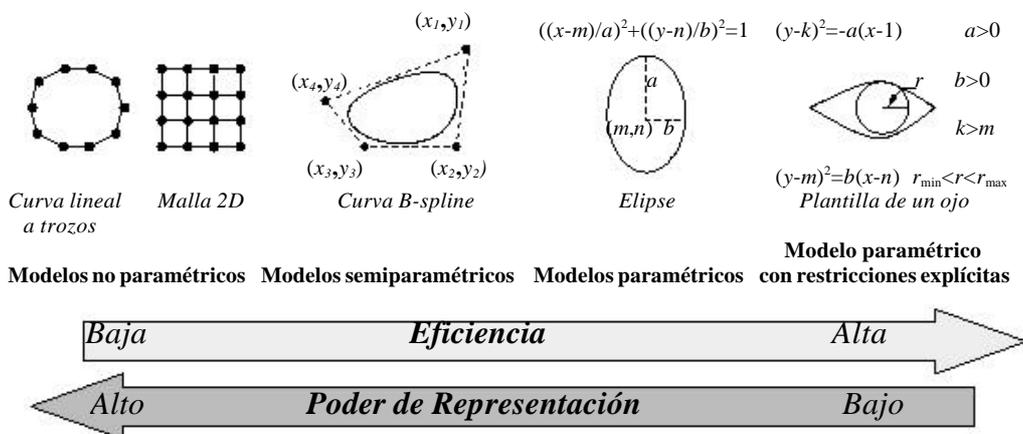


Figura 2. 9 Ejemplos de modelos con diferentes niveles de parametrización [Cheung 1998].

### 2.1.3.1.1 Modelos no paramétricos

Los modelos no paramétricos utilizan generalmente una función de deformación  $E_{def}$  que penaliza la falta de suavidad del modelo mediante la teoría física de la elasticidad. Esta función de deformación crece cuando el objeto se deforma de una determinada posición de reposo. Dos ejemplos clásicos son: los snakes [Kass 1987] y la arquitectura de enlaces dinámicos [Lades 1993].

#### 2.1.3.1.1.1 Snakes

Un buen ejemplo de los modelos deformables no paramétricos es el modelo de los contornos activos “snakes”. Consiste en un conjunto de puntos ordenados llamados “snaxels” y cada uno de ellos puede moverse (o deformarse localmente) libremente. Con un número suficiente de snaxels, se puede representar cualquier forma de manera precisa lo cual implica un gran poder de representación. Sin embargo, ésta representación no es eficiente para representar formas suaves como círculos donde el número de snaxels para una descripción precisa puede ser considerable.

Aunque los snakes admiten una representación continua en la práctica están representados por un conjunto de puntos ordenados. La deformación puede ser medida mediante la distancia entre los puntos que lo componen y mediante la suavidad del modelo, que se corresponden a la magnitud de la primera y segunda derivada espacial con respecto a los puntos que representan el modelo.

Matemáticamente estas deformaciones pueden ser formuladas como sigue: Sea  $\mathbf{p}_k = (x_k, y_k)^T$  las coordenadas del k-ésimo punto y  $\mathbf{w} = (\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2 \dots)$ . Se definen entonces las siguientes funciones:

- Distancia entre puntos

$$E_{def1}(w; M_i, \bar{w}_i) = \sum_k |p_k - p_{k+1}|^2 \quad \text{Ec. 2.36}$$

- Suavidad

$$E_{def2}(w; M_i, \bar{w}_i) = \sum_k |p_{k-1} - 2p_k + p_{k+1}|^2 \quad \text{Ec. 2.37}$$

Las derivadas espaciales presentes en la función de deformación puede ser interpretadas como una fuerza interna del modelo, que tiende a unir los puntos y hacer la variación en la curvatura del contorno tan pequeña como sea posible. Estas dos funciones se combinan en:

$$E_{def}(w; M_i, \bar{w}_i) = \sum_k h_1(p_k) |p_k - p_{k+1}|^2 + \sum_k h_2(p_k) |p_{k-1} - 2p_k + p_{k+1}|^2 \quad \text{Ec. 2.38}$$



Figura 2. 10 Snakes con baja (izquierda) y alta (derecha) energía de deformación

El principal inconveniente de este modelo es debido a que el modelo de deformación es demasiado general ya que la forma de la curva en reposo es un punto.

### 2.1.3.1.1.2 Arquitectura de enlaces dinámicos

Otra aproximación que no utiliza contornos [Lades 1993] está basada en el emparejamiento de grafos etiquetados mediante el uso de una malla regular bidimensional en la cual los nodos se pueden mover libremente. A diferencia de los snakes, cada nodo de la malla tiene un conjunto de atributos que se utilizan para representar un objeto. La deformación del modelo en este caso se mide con respecto a una malla de referencia y se describe a partir de la distancia entre los nodos que lo componen. Matemáticamente estas deformaciones pueden ser formuladas como se describe a continuación:

Sea  $p_{ij} = (x_{ij}, y_{ij})^T$  las coordenadas del punto en la fila  $i$  y columna  $j$  de la malla regular y sea  $w = (p_{00}^T, p_{01}^T, p_{02}^T \dots)$  y un modelo de referencia  $\bar{w}_i = (q_{00}^T, q_{01}^T, q_{02}^T, \dots)$  entonces:

$$E_{def}(w; M_i, \bar{w}_i) = \sum_{k,l,m,n} |(p_{kl} - p_{mn}) - (q_{kl} - q_{mn})|^2 \quad \text{Ec. 2.39}$$



Figura 2.11 Modelo de enlaces dinámicos

Este modelo se ha aplicado al reconocimiento de caras con porcentajes de reconocimiento superiores al 80%.

### 2.1.3.1.2 Modelos paramétricos

Para superar las limitaciones de eficiencia de los modelos no paramétricos, se han propuesto modelos deformables con una representación paramétrica. Los modelos paramétricos pueden representar una gran variedad de formas con una escala arbitraria. La deformación de los modelos paramétricos generalmente se mide a través de una distancia de los parámetros del modelo a partir de su valores medios, obtenidos o bien a priori o a través de entrenamiento. Para ello se suele contar con una matriz de covarianzas que relaciona la variación de sus distintas componentes y que puede ser simplemente la matriz identidad si la deformación de modelo es independiente parámetro a parámetro o puede ser una matriz de covarianzas más compleja si se dispone de un conjunto de datos. En este caso, el criterio se convierte en la distancia de Mahalanobis:

$$E_{def}(w; M_i, \bar{w}_i) = (w - \bar{w}_i)^T S^{-1} (w - \bar{w}_i) \quad \text{Ec. 2.40}$$

A continuación se muestran una serie de ejemplos basados en esta metodología.

#### 2.1.3.1.2.1 Curva elíptica deformable

La curva elíptica deformable [Staib 1992] está basada en la transformada elíptica de Fourier del contorno (Ec. 2.12). A partir de los datos en bruto de la representación  $w_{bruto} = (a_0, c_0, a_1, b_1, c_1, d_1, \dots)$  se definen los semi-invariantes:

$$\begin{aligned} A_i^2 &= \frac{a_i + \sqrt{a_i^2 - 4b_i^2}}{2} & B_i^2 &= \frac{2b_i^2}{a_i + \sqrt{a_i^2 - 4b_i^2}} \\ \mathbf{q}_i &= \arctg \frac{A_i c_i + B_i b_i}{A_i a_i - B_i d_i} & \mathbf{g}_i &= \arctg \frac{B_i a_i - A_i d_i}{A_i c_i + B_i d_i} \\ \mathbf{a}_i &= a_i^2 + b_i^2 + c_i^2 + d_i^2 & \mathbf{b}_i &= a_i d_i - b_i c_i \end{aligned} \quad \text{Ec. 2.41}$$

con  $A_i$  positivo y  $B_i$  del mismo signo que  $b_i$  para eliminar la ambigüedad del signo. Estos parámetros se denominan refinados  $w_{\text{ref}} = (a_0, c_0, A_1, B_1, \mathbf{q}_1, \mathbf{g}_1, \dots)$ . Por último para obtener parámetros invariantes frente a similitudes se utiliza:

$$\forall k > 1 \quad \mathbf{q}'_k = \mathbf{q}_k - \mathbf{q}_{k-1}, \quad \mathbf{g}'_k = \mathbf{g}_k - \mathbf{g}_{k-1} - \mathbf{g}_1, \quad A'_k = A_k / A_1, \quad B'_k = B_k / A_1 \quad \text{Ec. 2.42}$$

$$\mathbf{q}'_1 = \mathbf{q}_1, \quad \mathbf{g}'_1 = 0, \quad A'_1 = A_1, \quad B'_1 = B_1$$

lo que proporciona el vector de parámetros:  $w = (a_0, c_0, A'_1, B'_1, \mathbf{q}'_1, \mathbf{g}'_1, \dots)$  y la función de deformación con respecto a contorno de referencia  $\bar{w}_i$  es la distancia de Mahalanobis (Ec. 2.40) que puede interpretarse como una distribución a priori (ver Figura 2. 12) del vector de parámetros  $w$ :

$$P(w | M_i, \bar{w}_i) \propto \exp((w - \bar{w}_i)^T S^{-1} (w - \bar{w}_i)) \quad \text{Ec. 2.43}$$

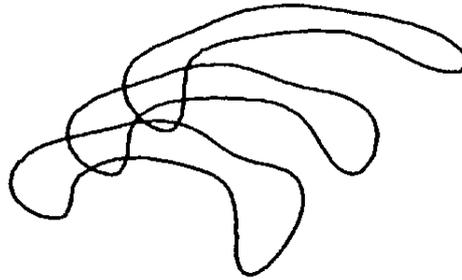


Figura 2. 12 Realizaciones del modelo probabilístico (de [Staib 1992])

Aunque esta representación es completa (es decir el número de parámetros coincide con el número de puntos) en la práctica se utilizan los primeros armónicos por lo cual se la considera una representación paramétrica.

### 2.1.3.1.2.2 Contornos activos basados en B-splines

Una aproximación similar ha sido desarrollada por [Blake 1998]. En este caso, la descomposición de la curva está basada en la descomposición de la función paramétrica en b-splines (ver Sección 2.1.2.2.5). Una de las aportaciones de este trabajo es la introducción de los espacios de formas en los cuales el problema de la invariancia puede expresarse de manera precisa y general. Para obtener la invariancia de la función de deformación respecto a un cierto subespacio lineal (por ejemplo el subespacio de traslaciones  $S$  definido por una matriz de proyección  $R$ ) se obliga a la función de deformación a operar fuera del subespacio, calculándose por tanto la deformación frente a un contorno de referencia  $\bar{w}_i$  como:

$$E_{\text{def}}(w; M_i, \bar{w}_i) = (w - \bar{w}_i)^T R' (w - \bar{w}_i) \quad R' = (I - R)^T (I - R) \quad \text{Ec. 2.44}$$

que puede interpretarse como una distribución a priori degenerada del vector de parámetros  $w$ :

$$P(w | M_i, \bar{w}_i) \propto \exp((w - \bar{w}_i)^T R' (w - \bar{w}_i)) \quad \text{Ec. 2.45}$$

Aunque esta representación es completa en la práctica se utiliza una aproximación a la figura de interés y en base a ésta se determinan los parámetros.

### 2.1.3.1.2.3 Snakes basados en la descomposición radial mediante wavelets

Un ejemplo de la utilización de wavelets para la representación de contornos deformables puede encontrarse en [Knoll 1999]. En este caso, se parte de una representación radial del contorno  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))$  utilizando la parametrización natural de la curva (Ec. 2.11) y su centroide  $(x_c, y_c)$ :

$$r(s) = \sqrt{(x(s) - x_c)^2 + (y(s) - y_c)^2}, \quad \mathbf{q}(s) = \arctg\left(\frac{(y(s) - y_c)}{(x(s) - x_c)}\right) \quad \text{Ec. 2.46}$$

y se descompone la parte radial mediante wavelets obteniendo:

$$r(s) = \sum_{k \in \mathbb{Z}} c_{J_0, k} 2^{\frac{J_0}{2}} \mathbf{f}(2^{J_0} s - k) + \sum_{j=J_0+1}^{\infty} \sum_{k \in \mathbb{Z}} d_{j, k} 2^{\frac{j}{2}} \mathbf{j}(2^j s - k) \quad \text{Ec. 2.47}$$

y por tanto el vector de parámetros  $w = (c_{J_0, 0}, \dots, d_{J_0, 0}, \dots)$ .

La información de la fase  $\mathbf{q}(s)$  solo se utiliza para la reconstrucción del contorno. En este modelo, la deformación frente a un modelo de referencia se lleva a cabo mediante el añadido de un nuevo término a la función de deformación clásica de los snakes:

$$F_S = \mathbf{I}_S (\mathbf{r} - \mathbf{r}_{rec}) \quad \text{Ec. 2.48}$$

donde  $\mathbf{r}_{rec}$  es el contorno obtenido a partir de un emparejamiento inicial entre la imagen y un objeto desde :

$$|d_{rec, j, k}| = \min(|d_{modelo, j, k} + U|, |d_{objeto, j, k}|), \quad j > j_U \quad \text{Ec. 2.49}$$

donde  $U$  es un valor de umbral. Es decir, se combina la utilización de funciones de deformación y restricciones sobre los parámetros.

En [Yoshida 1997] puede encontrarse otra aproximación basada en la descomposición de contornos en wavelets. En este caso la suavidad del modelo se obtiene mediante:

$$d_{j, k} = 0 \quad j > j_U \quad \text{Ec. 2.50}$$

limitándose por tanto la información a escalas finas en beneficio de una representación a baja escala. Sobre los coeficientes no nulos propone un modelo probabilístico similar al de (Ec. 2.43) pero al proponer una densidad a priori constante el modelo de deformación queda en realidad definido por (Ec. 2.50).

### 2.1.3.1.3 Modelos semiparamétricos

A pesar de la ventaja de utilizar modelos paramétricos, éstos se usan generalmente para representar formas suaves lo cual limita significativamente su poder de representación. Para aprovechar las ventajas de los modelos paramétricos y no paramétricos puede adoptarse una aproximación que utilice ambos modelos.

#### 2.1.3.1.3.1 *b*-snakes

La utilización de snakes parametrizados mediante b-splines permite utilizar un pequeño número de puntos de control (parte no paramétrica) y segmentos de curva generadas mediante splines cúbicos (parte paramétrica). De esta forma la curva se describe mediante (ver Ec. 2.15- Ec. 2.20)

$$\mathbf{r}(u) = \sum_{k=0}^{N_b-1} \mathbf{a}_k \mathbf{j}_k(u), \quad \mathbf{a}_k = (a_k^x, a_k^y)^T \quad \text{Ec. 2.51}$$

utilizando la discretización mediante b-splines puede demostrarse que el término de suavidad (Ec. 2.37) se define mediante :

$$\sum_{k \in Z} (b_1^3 * d^{(2)} * \mathbf{r})(k) (d^{(2)} * \mathbf{r})(k) \quad \text{Ec. 2.52}$$

donde \* denota la convolución discreta y los núcleos  $b_1^3$ , la discretización del b-spline y  $d^{(2)}$  la segunda diferencia y puede por tanto calcularse de forma eficiente [Brigger 2000]. En dicho trabajo se propone también controlar la suavidad de la representación de forma implícita mediante el escalado de los b-splines:

$$\mathbf{r}(u) = \sum_{k=0}^{N_b-1} \mathbf{a}_k \mathbf{j}_k^h(u), \quad \mathbf{a}_k = (a_k^x, a_k^y)^T, \quad \mathbf{j}_k^h(u) = \mathbf{j}_0(u/h - k) \quad \text{Ec. 2.53}$$

que aporta resultados equivalentes a la formulación de los snakes con mayor eficiencia.

#### 2.1.3.1.3.2 Modelos de elementos finitos

Este modelo semiparamétrico utiliza el método de elementos finitos (M.E.F.) y consiste en una malla que se trata como si estuviera compuesta de material elástico. La técnica

del Análisis Modal proporciona un conjunto de deformaciones lineales del objeto equivalentes a los modos resonantes de vibración del objeto original. Más específicamente se define una función de forma polinomial de desplazamiento con respecto a un conjunto nodos de referencia  $\bar{w}_i$  como

$$\mathbf{u}(w) = \mathbf{H}(w)\mathbf{U} \quad \text{Ec. 2.54}$$

que permite calcular los desplazamientos de un punto  $w$  con respecto al conjunto de nodos de referencia en  $\bar{w}_i$  por medio de una matriz de interpolación  $\mathbf{H}(w)$  y los desplazamientos entre el conjunto de nodos de referencia presentes en la matriz  $\mathbf{U}$ . Para la mayor parte de las aplicaciones es necesario calcular la tensión  $e(w)$  debida a la deformación. Esta se calcula mediante:

$$\mathbf{e}(w) = \mathbf{B}(w)\mathbf{U} \quad \text{Ec. 2.55}$$

donde  $\mathbf{B}(w)$  es la matriz de desplazamientos debido a la tensión. Las filas de  $\mathbf{B}$  se obtienen mediante una combinación adecuada de las filas de la matriz de interpolación. A partir de éstas se construye la matriz de masa y rigidez:

$$\mathbf{M} = \int_V \mathbf{r}\mathbf{H}^T \mathbf{H} dV, \quad \mathbf{K} = \int_V \mathbf{B}^T \mathbf{C}\mathbf{B} dV \quad \text{Ec. 2.56}$$

donde  $\mathbf{r}$  es la densidad de la masa y  $\mathbf{C}$  es la matriz de material que expresa la ley de esfuerzo-tensión particular del material. De esta forma se pueden calcular los modos naturales de vibración del objeto mediante la resolución del problema de autovectores generalizado:

$$\mathbf{K}\mathbf{T} = \mathbf{M}\mathbf{T}\mathbf{H}^2, \quad \mathbf{Q}, \mathbf{H} \text{ matrices de autovectores y autovalores} \quad \text{Ec. 2.57}$$

y la energía de deformación puede definirse como:

$$E_{def}(w; \mathbf{M}_i, \bar{w}_i) = \tilde{\mathbf{U}}^T \mathbf{H}^2 \tilde{\mathbf{U}}, \quad \tilde{\mathbf{U}} = \mathbf{T}^T \mathbf{M}\mathbf{U} \quad \text{Ec. 2.58}$$

donde  $\mathbf{U}$  está formada por los desplazamientos entre el conjunto de nodos determinados por  $w$ .

La malla proporciona un buen poder de representación (debido a su naturaleza no paramétrica) mientras que la matriz de interpolación, que interpola un punto arbitrario basado en los nodo de su entorno, actúa como un modelo paramétrico local que ayuda a mantener el número de nodos en la malla pequeño (debido a su naturaleza paramétrica) y mejorar la eficiencia del modelo.

### 2.1.3.2 Incorporación de información a priori

La mayor parte de las representaciones mencionadas en la última sección están generadas a partir de modelos de deformación local y permiten que el modelo se

deforme con gran libertad. Sin embargo, para el reconocimiento de objetos es deseable modelar cierto conjunto de formas y restringir su deformación utilizando alguna información del modelo. Por ejemplo, es deseable que un ojo deformado siga siendo un ojo y no llegue a convertirse en otro objeto. Para ello es necesario añadir restricciones de forma explícita o implícita al modelo. Por ejemplo en [Yuille 1992] se define una plantilla paramétrica y específica para algunas formas. Por ejemplo, una plantilla del ojo consiste en dos curvas parabólicas como la parte superior e inferior de este y un círculo para el iris, y está parametrizado por el centro y el radio del círculo junto con los coeficientes de la parábola superior e inferior (Figura 2. 13). Bajo restricciones generales, el modelo sigue siendo una plantilla del ojo aunque algunas realizaciones de éste sean menos probables de ocurrir en la realidad. Utilizando la misma idea, en [Dubuisson 1996] puede encontrarse la utilización de un polígono deformable junto con algunas restricciones para representar coches de distintos tipos.

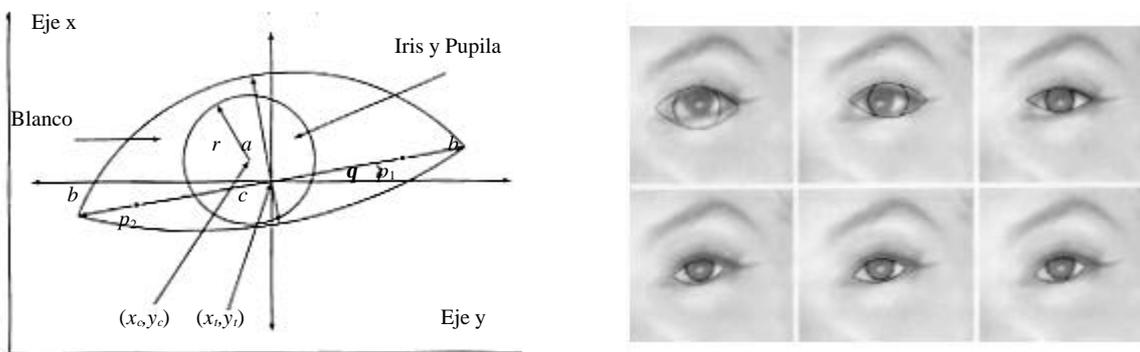


Figura 2. 13 Plantilla deformable de un ojo [Yuille 1992]

En lugar de utilizar algún conocimiento a priori para introducir restricciones en el modelo, se puede utilizar el aprendizaje para automatizar este proceso. En [Cootes 1992] se encuentra el modelo de distribución de puntos (MDP), que es básicamente un vector formado mediante la concatenación de las coordenadas  $x$  e  $y$  de un conjunto ordenado de puntos. Basado en el MDP se genera una forma media y una matriz de covarianzas del conjunto de modelos que aproximan de forma óptima al conjunto de muestras. Las componentes principales de la matriz de covarianzas con mayor magnitud de los autovalores se extraen como los modos de deformación posible y la desviación de la forma media del modelo se expresa entonces con una combinación lineal de dichas componentes principales. Adicionalmente, se pueden extraer perfiles de intensidad del gradiente perpendiculares al modelo del contorno en los puntos ordenados que forman el conjunto entrenamiento y obtener de esta forma el perfil medio del gradiente junto con su correspondiente matriz de covarianzas. Esto proporciona atributos al MDP y lo hace un modelo basado en características. Este modelo asume una distribución unimodal de los parámetros. En general, para un mejor modelado, se puede utilizar una distribución multimodal, como por ejemplo una mezcla de gaussianas.

La mayor limitación para construir modelos con aprendizaje es el requerimiento de una gran cantidad de datos de entrenamiento. En el caso de tener un conjunto limitado de datos, la media y la matriz de covarianzas puede ser inicializadas utilizando algún conocimiento a priori (sin entrenamiento). El entrenamiento se utiliza para ajustar de forma final el modelo y por tanto se necesita mucho menos datos. Así, se suele usar un modelo MEF para la inicialización y combinarlo con un modelo MDP durante el

entrenamiento. De esta forma se obtiene un modelo híbrido que posee un mejor rendimiento para conjuntos limitados de datos.

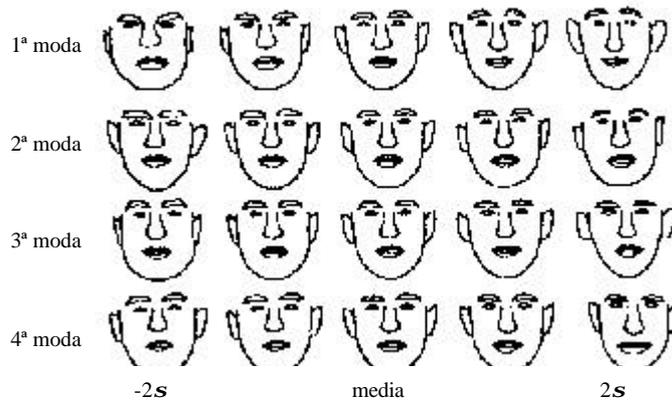


Figura 2. 14 Diferentes componentes principales para una cara [Cootes 1992]

### 2.1.3.3 Invarianza en modelos deformables

La invarianza de modelos deformables es un problema de considerable importancia. La forma de lograr esta invarianza se lleva a cabo mediante la definición de un modelo implícito invariante o bien de forma explícita mediante la estimación directa de los parámetros de la transformación.

El modelo de contornos activos generalizados [Lai 1994] es un buen ejemplo de la utilización de modelos implícitos para conseguir la invarianza frente a transformaciones afines. Utiliza una matriz de forma, definida de manera que sea invariante frente a transformaciones afines, para representar los distintos objetos. Una aproximación similar [Blake 1998] utiliza el concepto general de espacio de formas para obtener invarianzas sobre determinados subespacios lineales. En los métodos explícitos la transformación afín se estima directamente, aislando de esta forma la deformación global debida a la proyección para obtener la invarianza frente al punto de vista. Ejemplos de este tipo de métodos pueden encontrarse en [Yuille 1992].

## 2.2 Emparejamiento de Objetos basado en modelos

El proceso del emparejamiento de objetos intenta buscar una correspondencia entre el modelo y la imagen. Para ello se dispone de una función de costo  $E_{aju}$  que proporciona la bondad del ajuste del modelo específico a la imagen. En el caso de tener un modelo deformable la función de costo tiene un término de deformación  $E_{def}$  que se une con el anterior para proporcionar una función de ajuste combinada. En general se realizan algunos de los siguientes pasos:

- **Detección de características.**

Se realiza un procesamiento de la imagen para detectar características que serán representadas.

- **Agrupamiento.**

El objetivo de este proceso es el de producir grupos de características que sean más informativos que las características individuales, y por tanto produzcan un mejor guiado en la selección y emparejamiento de modelos.

- **Indexado.**

El problema de indexado consiste en seleccionar modelos para los que es probable que se produzca un emparejamiento con las características extraídas de la imagen. El objetivo es evitar el tener que realizar el proceso de emparejamiento para cada uno de los modelos.

- **Emparejamiento o “*matching*”.**

Dada una imagen y un modelo del objeto, ambas representados por sus características, el objetivo es encontrar un emparejamiento entre los dos y estimar como el modelo del objeto está posicionado en la imagen. De todas las soluciones factibles al problema del emparejamiento, se desea aquella que maximiza la medida de ajuste a la imagen.

Aunque tanto el emparejamiento de modelos rígidos como el de modelos deformables intentan optimizar una función de costo  $E_{tot}$ . Los algoritmos de emparejamiento utilizan diferentes técnicas. Esto es debido al hecho de que en el emparejamiento de objetos rígidos, el espacio  $W_i$  de valores del parámetro  $w$  es un espacio de pequeña dimensión coincidente con los grados de libertad de la transformación (ver Tabla 2. 2). Este espacio  $W_i$  generalmente se define a través del movimiento rígido definido en un espacio de formas  $S=L(\mathbf{H}, \overline{\mathbf{w}})$  (ver Sección 3.2.6) definido por una función  $\mathbf{H}$  y un modelo de referencia  $\overline{\mathbf{w}}$ . Así puede escribirse de forma general:

$$W = \{w / w = \mathbf{H}(x, \overline{\mathbf{w}})\} \quad \text{Ec. 2.59}$$

Puesto que el espacio de parámetros de la transformación determinado por el vector  $\mathbf{x}$  es de baja dimensionalidad el emparejamiento rígido suele operar sobre dicho espacio. Por tanto se separará el estudio del emparejamiento de modelos rígidos y deformables.

### 2.2.1 Emparejamiento de modelos rígidos

En este caso, el objetivo del proceso de emparejamiento es el intentar encontrar en la imagen una transformación rígida del modelo. Para ello se busca optimizar una determinada función de ajuste  $E_{tot}$  que en el caso rígido es igual a  $E_{aju}(W, M_i, D)$ .

#### 2.2.1.1 Criterios de ajuste a la imagen

En el caso rígido es posible encontrar dos clases de criterios para la definición del ajuste del modelo a la imagen, el criterio de consistencia y los basados en distancias.

- **Basados en consistencia**

Este tipo de métodos están basados en transformar el conjunto de primitivas del modelo de referencia  $\bar{W}_i = (P_1, P_2, \dots, P_n)$  mediante una transformación rígida  $H(x, \bar{W}_i) = (H(x, P_1), H(x, P_2), \dots, H(x, P_n))$ . Cada primitiva transformada se asigna con una primitiva cercana de la imagen y la calidad del emparejamiento se evalúa mediante un escalar  $r(H(x, P_i))$ , generalmente  $r$  vale 0 ó 1. En este caso, si se puede emparejar la primitiva  $r$  toma el valor 1 y si no  $r$  toma el valor 0. La función de ajuste se define como:

$$E_{aju}(W, M_i, D) = \max_x \sum_{i=1}^n r(H(x, P_i)) \quad \text{Ec. 2.60}$$

- **Basados en distancia**

Otro tipo de métodos de emparejamiento están basados en el establecimiento de métricas para conjuntos de puntos y por ello se suelen utilizar con primitivas puntuales. Un ejemplo es la distancia dirigida de Hausdorff donde se parte de dos conjuntos de puntos del modelo  $W = (p_1, p_2, \dots, p_n)$  e imagen  $V = (q_0, q_1, \dots, q_m)$  y se calcula:

$$E_{aju}(W, M_i, D) = \bar{d}(W, V) = \max_{p \in W} \min_{q \in V} d(p, q) \quad \text{Ec. 2.61}$$

con  $d$  una distancia entre puntos de  $\mathbb{R}^n$ . Otros tipos de distancias son posibles y se revisarán en la Sección 2.2.1.2.2.

#### 2.2.1.2 Métodos de emparejamiento

Existe una gran cantidad de métodos en el caso rígido para el reconocimiento basado en modelos. En esta sección, se describirán los métodos más importantes.

2.2.1.2.1 Métodos basados en consistencia

Los métodos basados en consistencia se caracterizan por su énfasis en emparejar de forma consistente las características del modelo y la imagen. Este conjunto de emparejamientos entonces determina la posición del modelo. A continuación se discutirán algunos métodos de emparejamiento junto con sus correspondientes complejidades. Para ello denotaremos por  $n$  al número de primitivas del modelo y a  $m$  el número de primitivas de la imagen.

2.2.1.2.1.1 Árbol de interpretación

El método del árbol de interpretación para el reconocimiento basado modelos fue desarrollado por [Grimson 1987]. Éste método usa modelos bidimensionales e imágenes, cuyas primitivas son los segmentos de líneas, y busca rotaciones bidimensionales y traslaciones en el espacio de transformaciones. El método busca en un árbol para encontrar un conjunto emparejamientos ente las primitivas del modelo y las primitivas de la imagen. El árbol de interpretación es un árbol que tiene  $m+1$  niveles, donde cada nivel se corresponde con una de las primitivas de la imagen mas una primitiva nula, y cada nodo en el árbol (excepto las hojas) tiene un factor de ramificación de  $n+1$ , correspondientes a las primitivas del modelo y a la rama llamada "primitiva nula". En cada nodo interno del árbol, el algoritmo de búsqueda intenta encontrar un emparejamiento para la primitiva de la imagen correspondiente al nivel de ese nodo. El camino desde ese nodo hasta la raíz corresponde a todos emparejamientos ya se realizados. La búsqueda intenta, en el caso de que sea posible, emparejar la primitiva de la imagen con cada una de las primitivas del modelo. Para ello se verifica cada primitiva del modelo para ver si es compatible con una primitiva de la imagen; realizándose un emparejamiento salvo en el caso de que no sean compatibles (por ejemplo, si sus longitudes son muy distintas). Si son compatibles, entonces se comprueba ese emparejamiento, uno a uno, contra todos los otro emparejamientos de la solución actual (es decir del camino de vuelta hacia la raíz). La forma de verificar que dos emparejamiento son consistentes está basado en la búsqueda de una transformación que lleve a las primitivas del modelo cerca de las correspondientes primitivas de la imagen.

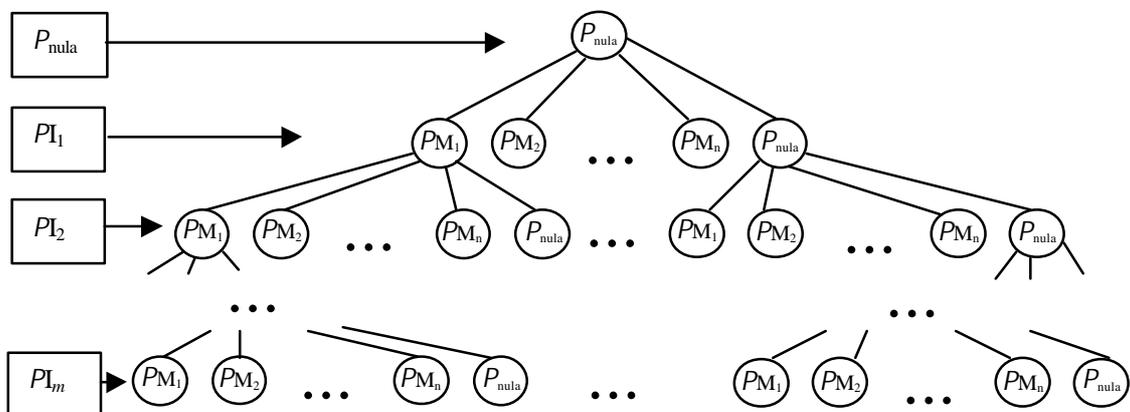


Figura 2. 15 Árbol de interpretación.  $PI$  y  $PM$  denotan primitivas de la imagen y modelo respectivamente.

Si el nuevo emparejamiento es consistente (uno a uno) con todos emparejamientos previos, se continua explorando el subárbol a partir de ese nodo. Si no existe ninguna primitiva del modelo que pueda ser emparejada con la primitiva de la imagen, debido a la inconsistencia con los emparejamientos previos, entonces la "primitiva nula" se empareja con la primitiva de la imagen. Siempre que se alcance una hoja del árbol, se verifica el conjunto actual de emparejamientos para ver si existe una consistencia global. La búsqueda hasta ese momento solamente ha comprobado la consistencia uno a uno, aunque esto no garantiza que todos los emparejamientos sean consistentes bajo una determinada transformación, ya que puede que no haya una transformación que lleve cada primitiva del modelo cerca de su correspondiente primitiva en la imagen. Si no existe tal transformación, entonces la hoja se rechaza y la búsqueda vuela hacia atrás. La verificación de la consistencia global no se realiza en cada nodo debido a su poca potencia en la poda y su mayor costo computacional. La principal desventaja de este método es que el peor caso su tiempo de ejecución es proporcional al tamaño de árbol que es exponencial.

#### 2.2.1.2.1.2 Alineamiento

El conjunto de técnicas de alineamiento se caracterizan por su estructura de "hipótesis y verificación". Esencialmente, estos métodos eligen un subconjunto (de tamaño determinado) de las primitivas de la imagen y un subconjunto (del mismo tamaño) de las primitivas del modelo, y forman un emparejamiento entre dichos conjuntos. Este emparejamiento se utiliza entonces para calcular un conjunto de posibles transformaciones del modelo que lleva las primitivas elegidas sobre las primitivas de la imagen. Los tamaños de los conjuntos de características se eligen de forma que el número de posibles transformaciones sea una pequeña constante. Por ejemplo, si tanto el modelo y la imagen son conjuntos de puntos bidimensionales y el espacio de transformaciones es el espacio de rotaciones, traslaciones, y cambios de escala, entonces los subconjuntos serían de tamaño dos. En este caso solamente existe una transformación que puede alinear el subconjunto de primitivas del modelo con el subconjunto de las primitivas de la imagen. Una vez que se han elegido los subconjuntos y se ha calculado el alineamiento, se verifican las transformaciones mediante la proyección de las primitivas adicionales del modelo, para ver si están suficientemente cerca de sus correspondientes primitivas en la imagen. Si un número determinado de primitivas del modelo está suficientemente cerca de algunas primitivas de la imagen, entonces la transformación se acepta; sino, se prueba con otro subconjunto.

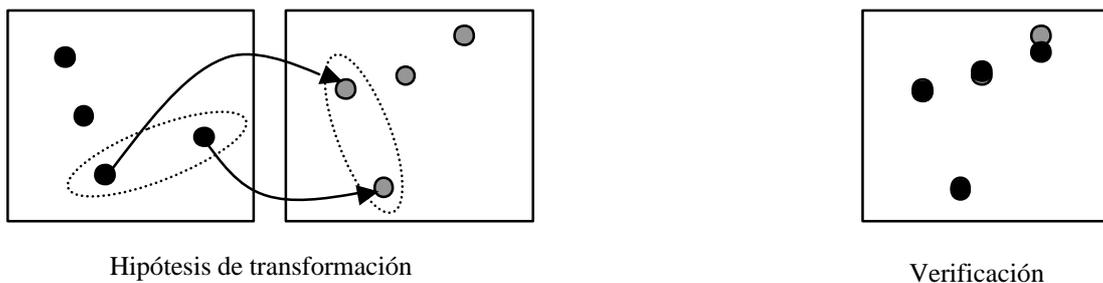
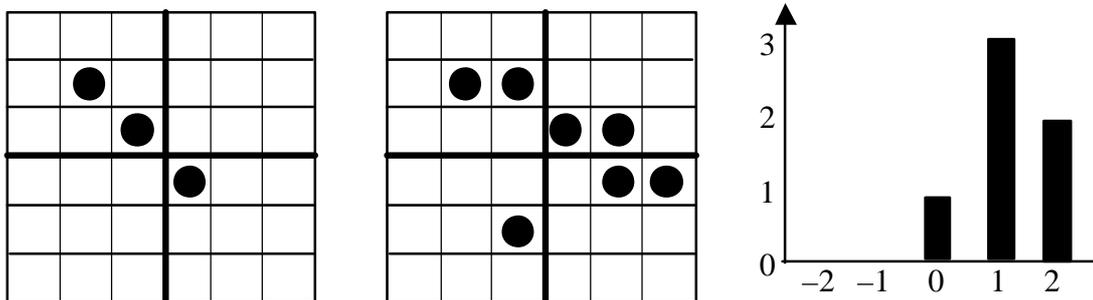


Figura 2. 16 El método de alineamiento

El principal problema de este método es el del ajuste: pequeñas variaciones la posición de los tres puntos emparejados puede llevar a grandes errores en la posición de los siguientes puntos del modelo. Su complejidad tiempo es también bastante elevada  $O(n^4m^3)$ .

### 2.2.1.2.1.3 Transformada de Hough

Los métodos basados en la transformada de Hough, también llamados métodos de agrupamiento de parámetros, intentan calcular los parámetros de la transformación del modelo mediante la generación de un gran número de emparejamientos entre subconjuntos del modelo y de la imagen, la determinación del conjunto de transformaciones que pueden haber generado tal emparejamiento (y a diferencia de los métodos de alineamiento, este conjunto no tiene por qué contener un pequeño conjunto de transformaciones sino que puede ser un subespacio del espacio de transformaciones), y la construcción de un histograma en el espacio de parámetros de transformación mediante la discretización de dicho espacio en celdas. Cada celda del subespacio correspondiente a alguna transformación de este conjunto se incrementa (se dice que recibe un voto). Para buscar la transformación óptima se busca en los máximos del histograma (buscando las transformaciones más votadas).



**Figura 2. 17** Determinación de la traslación horizontal del modelo (izquierda) en la imagen (centro) mediante el histograma (derecha) de la transformada de Hough.

Por ejemplo, en [Thompson 1987] puede encontrarse una aplicación de esta técnica se utiliza para resolver el problema de localizar un modelo tridimensional en un imagen bidimensional, bajo movimiento rígido en el modelo de perspectiva débil. Como primitiva se utilizan pares de vértices: en este caso un par de vértice está compuesto por un vértice y dos de su aristas incidentes, junto con otro vértice (posiblemente distante y no conectado). Un emparejamiento entre un par de vértices del modelo y de la imagen determina una única transformación que alinea esas dos primitivas. La imagen se procesa para extraer aristas, y todos los posibles pares de vértices que se forman a partir de esas aristas. Cada par de vértices del modelo se emparejan uno a uno con cada par de vértices de la imagen; y se agrupan las transformaciones que se generan. Puesto que el espacio transformaciones es seis-dimensional, agruparlas en un histograma directamente no es factible; en vez de ello, se agrupan en histogramas independientes: un histograma bidimensional indexado por dos de los parámetros rotacionales, un histograma

unidimensional indexado por el tercer parámetro rotacional y un histograma tridimensional indexado por los tres parámetros traslación. Los histogramas se colocan en cascada. Cada máximo por encima de un determinado umbral en el histograma inicial bidimensional hará que se realice un histograma unidimensional sobre las transformaciones definidas por dicho máximo. Los máximos en ese histograma generan finalmente, un histograma tridimensional cuyos máximos se utilizan para determinar la transformación del modelo.

La principal desventaja de este tipo de métodos consiste en que no toda celda que tenga un gran número de votos representa una solución correcta. En algunos casos, debido a la forma en la que se realiza la discretización del espacio de transformaciones, una celda puede conseguir varios votos que no sean consistentes con una única transformación. Además, cuando la imagen contienen una gran cantidad de ruido, se pueden generar agrupaciones aleatorias de votos, de forma que puede tenerse que examinar una gran cantidad de máximos antes de que se encuentre una solución correcta. Por otra parte la complejidad computacional es también alta  $O(n^3m^3)$  para transformaciones afines.

#### 2.2.1.2.1.4 *Dispersión (Hashing) Geométrica*

La idea en la que se basa la dispersión geométrica parte de disponer de un modelo bidimensional compuesto de puntos bajo la influencia de una transformación afín, entonces si se conoce la posición de tres puntos del modelo en la imagen se puede determinar dicha transformación y por tanto la transformación de los restantes puntos del modelo. De hecho, si esos puntos se utilizan para definir el origen y una base para localización de los restantes puntos del modelo y la transformación de esos tres puntos se utiliza para definir el origen y una base en la imagen de para los restantes puntos del modelo transformado, entonces las coordenadas (en la base original) de un punto del modelo son las mismas que las coordenadas (en la base transformada) de su posición transformada (Figura 2. 18). Estás coordenadas son por tanto invariantes bajo la aplicación de cualquier transformación afín. Nótese que un punto cualquiera del modelo puede tener varias coordenadas afines diferentes generadas a partir de bases distintas.

Esta invarianza de las coordenadas afines fue utilizada en [Lamdan 1988] para localizar modelos bidimensionales en imágenes bidimensionales bajo la influencia de una transformación afín. También es posible utilizar este método para otro tipo de transformaciones, por ejemplo para una similitud sólo se necesitan dos puntos para definir una base.



Figura 2. 18 Invarianza de coordenadas frente a transformación de la base

Utilizando la invarianza en las coordenadas, el modelo es preprocesado para producir una tabla de dispersión. Para ello se examina cada base y para cada una, se determinan las coordenadas de todos los demás puntos del modelo. Estas coordenadas se introducen (convenientemente discretizadas) en una tabla junto con la información de que base los produjo. A continuación se extraen primitivas basadas en puntos de la imagen. Se eligen tres puntos de la imagen, y se calculan las coordenadas afines de los restantes puntos de la imagen. Las coordenadas de cada punto se utilizan para buscar en la tabla de dispersión y se determinan cuales son las bases correspondientes incrementándose una tabla de votación (ver Figura 2. 19). Si los 3 puntos de la imagen fueron realmente las primitivas del modelo transformado, entonces la base que forman será una de las utilizadas para construir la tabla; y tal base debería de ser determinada siempre que uno de los restantes puntos del objeto se utilice para buscar en la tabla y por tanto debería ser más votada que cualquiera de las otras.

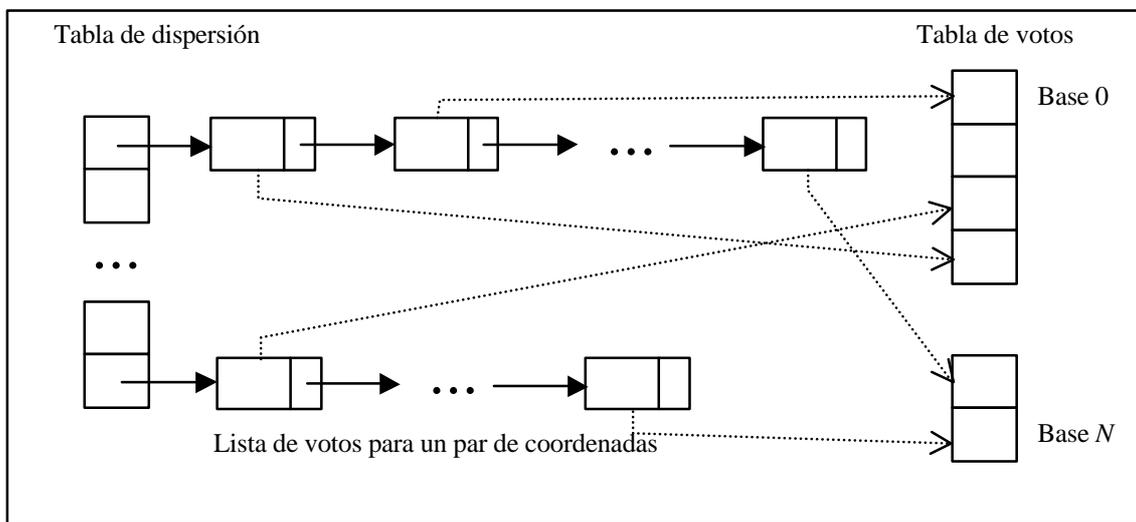


Figura 2. 19 Esquema de votación para una base

Si no todos los grupos de tres puntos de la imagen fueron parte del objeto, entonces ninguna base se debe generar preferentemente; en este caso, se elige otra tripleta y se procede hasta que se localiza la base correcta. Los principales inconvenientes de este método consisten en que los errores en las coordenadas afines puede ser grandes incluso con pequeños errores en las coordenadas de las posiciones de los tres puntos que forman la base y que el número de celdas en la tabla para buscar puede ser grande. Por otra parte es posible tal y como ocurre en la transformada de Hough que se generen falso positivos debidos a alineamientos aleatorios. La complejidad de este método es de  $O(n^4+m^3)$  (asumiendo un costo unitario de examinar la tabla de dispersión).

### 2.2.1.2.2 Métodos basados en distancias

Otra clase de algoritmos intentan optimizar una función de distancia (ver [Alt 1999], [Veltkamp 1999] para una revisión y estudio de la complejidad computacional). En esta sección se asume que los elementos de  $W=(p_1, p_2, \dots, p_n)$  son un conjunto de puntos del modelo que se desean emparejar con un conjunto de puntos

$v = (q_0, q_1, \dots, q_m)$  extraídos de la imagen  $D$ . El emparejar los conjuntos de puntos significa encontrar una correspondencia entre los puntos de  $w$  y los puntos de  $v$ .

#### 2.2.1.2.2.1 Emparejamiento de cuello de botella

Supóngase que tanto  $w$  como  $v$  constan de  $n$  puntos y  $d(p, q)$  es la distancia entre dos puntos. La distancia de cuello de botella es el mínimo de la distancia máxima  $d(p, f(p))$  entre todas las correspondencias  $f$  uno a uno entre  $w$  y  $v$ . La distancia puede ser calculada en tiempo  $O(n^{3/2} \log n)$  si  $d$  es la distancia euclídea.

Es de interés el calculo de la distancia mínima para un grupo de transformaciones para ello encontraremos la transformación óptima  $f_{opt}$  tal que  $d(f(w), v)$  sea mínimo. Para movimientos rígidos (traslación y rotación), se puede encontrar una solución en tiempo  $O(n^6 \log n)$ , mientras que para traslaciones únicamente, se puede calcular el tiempo  $O(n^5 \log^2 n)$ .

En el caso de que sea posible la presencia de ruido, se considera el problema de encontrar la distancia de cuello de botella menor entre todos emparejamientos uno a uno entre  $k$  puntos de  $w$  y  $k$  puntos de  $v$ . Esta distancia puede ser calculada en  $O(m \log m + n^{3/2} \log m)$ .

#### 2.2.1.2.2.2 Emparejamiento de mínimo peso.

La distancia mínimo total es el mínimo de la suma de las distancias  $d(p, f(p))$  entre todas las correspondencias  $f$  uno a uno entre  $w$  y  $v$ . Puede calcularse en tiempo  $O(n^{2+\epsilon})$ , donde  $\epsilon$  es la constante positiva que puede ser elegida arbitrariamente pequeña mediante una elección apropiada de otras constantes del algoritmo.

#### 2.2.1.2.2.3 Emparejamiento uniforme.

La distancia más uniforme es el mínimo entre todas las correspondencias  $f$  uno a uno entre  $w$  y  $v$  de la diferencia entre el máximo y el mínimo de las distancias  $d(p, f(p))$ . Este emparejamiento se llama también emparejamiento balanceado. La distancia puede ser calculada en tiempo  $O(n^{10/3} \log n)$ . El problema de encontrar la distancia uniforme más pequeñas entre todos emparejamientos uno a uno utilizando  $k$  puntos de  $w$  y  $k$  puntos de  $v$  tiene la misma complejidad

#### 2.2.1.2.2.4 Emparejamiento de mínima desviación.

El emparejamiento de mínima desviación es el mínimo sobre todas las correspondencias  $f$  uno a uno entre  $w$  y  $v$  de la diferencia entre el máximo y la distancia media  $d(p, f(p))$ . Este emparejamiento puede ser calculado en tiempo  $O(n^{10/3+\epsilon})$ .

2.2.1.2.2.5 Distancia de Hausdorff

En muchas aplicaciones, no todos los puntos del conjunto  $W$  tienen que tener correspondencia en  $V$ , debido a la oclusión y al ruido. Típicamente, los dos conjuntos tienen tamaño distinto, de forma que no existe una correspondencia uno a uno entre todos los puntos. En este caso, una medida de disimilitud que se utiliza es la distancia de Hausdorff. La distancia de Hausdorff para conjuntos de puntos finitos se describe como sigue: la distancia dirigida de Hausdorff  $\vec{d}(W, V)$  se define como el máximo sobre todos los puntos de  $W$  de las distancias a un punto de  $V$ . La distancia de Hausdorff  $d(W, V)$  es el máximo de  $\vec{d}(W, V)$  y  $\vec{d}(V, W)$ :

$$d(W, V) = \max(\vec{d}(W, V), \vec{d}(V, W)), \quad \vec{d}(W, V) = \max_{p \in W} \min_{q \in V} d(p, q) \quad \text{Ec. 2.62}$$

esta distancia puede ser calculada utilizando diagramas de Voronoi en tiempo  $O((n+m)\log(n+m))$ .

La traslación que minimiza la distancia de Hausdorff puede ser determinada en tiempo  $O(mn(\log(mn))^2)$  si la métrica es  $L_1$  o  $L_\infty$ . Para el uso de la métrica  $L_p$   $p=2,3...$  puede ser calculada en tiempo  $O(mn(m+n)\alpha(nm)\log(n+m))$  ( $\alpha(n)$  es la inversa de la función de Ackerman).

Dado un valor real  $\epsilon$ , decidir si hay un movimiento rígido (traslación y rotación) de forma que la distancia de Hausdorff sea menor que  $\epsilon$ , puede ser calcularse en tiempo  $O((m+n)m^2n^2\log(mn))$ . El cálculo del valor óptimo se hace en  $O(m+n)^6\log(m+n)$ . El cálculo la distancia óptima aproximada bajo movimiento rígido puede ser el hecho en tiempo:  $O(m+n)\log(m+n)$

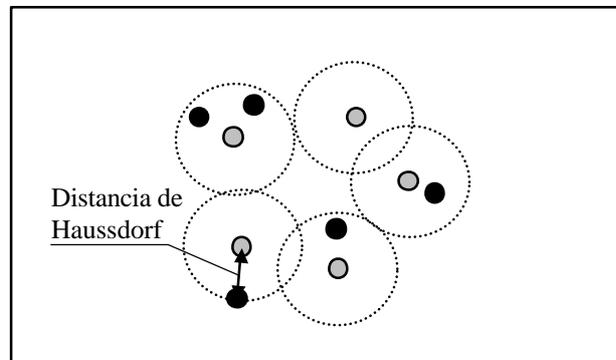


Figura 2. 20 Distancia de Hausdorff entre dos conjuntos

2.2.1.2.2.6 Subdivisión del espacio de transformaciones

El emparejamiento de conjuntos de puntos finitos, de imágenes, bajo traslaciones y escalado se hace [Huttenlocher 1993] dividiendo el espacio de transformaciones. En vez de utilizar la distancia de Hausdorff, se utiliza la distancia parcial de Hausdorff, que es el máximo de las dos distancias parciales dirigidas de Hausdorff:

$$d_k(W, V) = \max(\vec{d}_k(W, V), \vec{d}_k(V, W)), \quad \vec{d}_k(W, V) = k - \underset{p \in W}{\text{ésimo}} \min_{q \in V} d(p, q) \quad \text{Ec. 2.63}$$

la distancia parcial de Hausdorff no es una métrica puesto que no cumple la desigualdad triangular. El tiempo de ejecución depende de la profundidad en la división del espacio de transformaciones.

### 2.2.1.2.3 Métodos de emparejamiento de contornos

El hecho de que un contorno esté compuesto de puntos hace que sea posible aplicar para su reconocimiento cualquiera de los métodos anteriores. No obstante, los conceptos de similitud entre puntos no son necesariamente buenas medidas de similitud entre contornos. Por otra parte, el alto costo computacional del emparejamiento de puntos hace aconsejable el desarrollo de técnicas específicamente dedicadas a contornos. En esta sección se hará un breve resumen de las principales técnicas, y su dominio de aplicación.

#### 2.2.1.2.3.1 Códigos de cadena

Debido a la alta sensibilidad al ruido del código de cadena y al hecho de que cambios arbitrarios en escala y rotación cambian su descripción, esta representación no se suele utilizar en problemas de reconocimiento sino como un paso intermedio para la utilización de otras técnicas. Sin embargo, es posible su utilización cuando los contornos no tienen una excesiva complejidad. Por ejemplo en [Parizeau 1998] se emplean los códigos de cadena para el reconocimiento de dígitos. El proceso de reconocimiento está basado en la comparación de las cadenas correspondientes a los diferentes dígitos por medio de la distancia de edición. Por medio de esta distancia se comparan dos cadenas mediante la asignación de un coste a las operaciones elementales de inserción, borrado y cambio que permiten transformar una cadena en otra. Utilizando algoritmos genéticos se calcula el costo óptimo de dichas operaciones obteniéndose errores de reconocimiento del orden del 10%.

#### 2.2.1.2.3.2 Aproximaciones poligonales

Una técnica para el emparejamiento de aproximaciones poligonales es la basada en la distancia de Hausdorff. Dados dos polígonos con  $n$  y  $m$  vértices respectivamente, la distancia mínima de Hausdorff bajo traslaciones se calcula en tiempo  $O((mn)^2 (\log(m+n))^3)$ . Debido a la alta complejidad, tiene sentido implementar algoritmos de aproximación para encontrar una transformación que proporcione una distancia de Hausdorff a un número constante de veces la distancia óptima. Para emparejamientos bajo traslaciones esto puede hacerse en tiempo  $O((m+n) \log(m+n))$  para un factor  $1 + \sqrt{2}$ . La distancia mínima de Hausdorff incluyendo rotaciones puede calcularse en tiempo  $O((mn)^4 (m+n) (\log(m+n)))$ , por tanto es de interés obtener también un algoritmo aproximado. Este puede obtenerse en tiempo  $O((mn) \log(mn) \log^*(mn))^1$ .

---

<sup>1</sup> La notación  $O(\log^*(n))$  es el número de veces que debe aplicarse el logaritmo a  $n$  para obtener un número menor que la unidad y es una función de crecimiento muy lento.

2.2.1.2.3.3 Firmas

Existen diferentes tipos de emparejamiento de firmas puesto que éstas pueden ser definidas de formas distintas. Sin embargo todas ellas están basadas en la minimización de alguna medida de disimilaridad entre funciones unidimensionales. Así por ejemplo para la signatura de O'Rourke (Ver Sección 2.2.1.2.3.3), el algoritmo de optimización está basado en la deformación de las firmas para ajustar una a la otra [Kruskal 1983]. Utilizando programación dinámica puede llevarse a cabo en tiempo  $O(nm)$ .

2.2.1.2.3.4 Transformadas de Fourier del contorno

De las distintas formas de calcular la transformada de Fourier del contorno, las más utilizadas son los coeficientes de la transformada elíptica (Ec. 2.12). El emparejamiento se lleva a cabo mediante la normalización previa frente a transformaciones ( punto de inicio, rotación, traslación, escalado) de los coeficientes para posteriormente calcular la distancia  $L_2$  de las descripciones del contorno con los del resto de los modelos mediante cualquier clasificador estándar. En [Park 1999] puede encontrarse una aplicación de esta técnica con un porcentaje de éxito del 80 al 90%. El principal inconveniente consiste en que los coeficientes son globales y por tanto todos los coeficientes son sensibles frente a oclusiones. No obstante se han realizado intentos de emparejamiento con oclusiones [Gorman 1988], donde los contornos se segmentan y cada segmento es descrito mediante la transformada elíptica. Utilizando técnicas de programación dinámica se busca el emparejamiento óptimo con oclusiones.

2.2.1.2.3.5 Métodos basados en b-splines

Existen diversos métodos de emparejamiento basados en b-splines. El más sencillo consiste en el cálculo de invariantes (Ec. 2.24) a partir de la representación de la curva mediante b-splines. En [Jacob 2000] puede encontrarse como realizar éste cálculo de forma eficiente utilizando el teorema de Green.

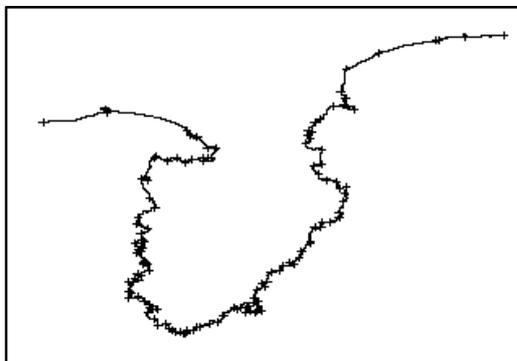


Figura 2. 21 Puntos dominantes extraídos del contorno. Factor de compresión 15.17. Error medio 0.0057 (de [Wang 1997])

Otra aproximación al reconocimiento está basada en la detección de puntos dominantes en el contorno [Wang 1999] en el cual se calcula la imagen del espacio de escalas de la curvatura (ver Sección 2.1.2.2.6) utilizando una descripción de la curva mediante b-splines (Figura 2. 21). Posteriormente puede aplicarse cualquier técnica de emparejamiento de puntos (Sección 2.2.1.2) o emparejamiento en el espacio de escalas.

2.2.1.2.3.6 Métodos en el espacio de escalas

Este tipo de métodos de emparejamiento están basados en la búsqueda de una correspondencia entre la información a varias escalas del contorno. En [Ueda 1993] se propone un algoritmo basado en los segmentos determinados por puntos de inflexión a múltiples escalas extraídos de un contorno. Utilizando esta información se realiza la búsqueda de la correspondencia entre los distintos segmentos de ambos contornos desde la escala más fina a la más gruesa mediante el reemplazo de un número impar de segmentos consecutivos en una escala más fina por otro en una escala más gruesa. El emparejamiento óptimo es equivalente a la búsqueda de la correspondencia óptima entre segmentos de forma que la suma de las diferencias a múltiples escalas sea mínimo y se resuelve mediante programación dinámica.

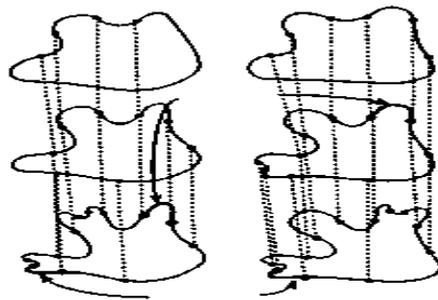


Figura 2. 22 Emparejamiento a múltiples escalas (de [Ueda 1993])

Otra aproximación multiescala utiliza el espacio de escalas de curvatura [Moktharian 1996]. Para ello se extraen los máximos de los contornos determinados por los cruces con cero de la curvatura y se determina la traslación óptima entre los dos espacios de escala de forma que el mayor número de máximos del modelo estén en un entorno de los máximos de la imagen

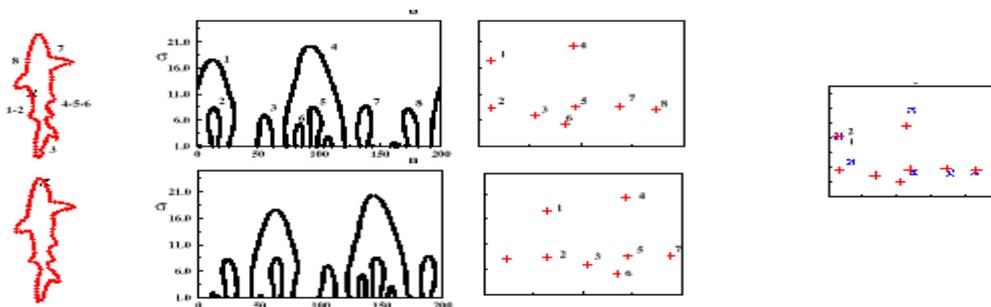


Figura 2. 23 (Izquierda) Contornos a emparejar, (Derecha) Emparejamiento en el espacio de escalas. (de [Moktharian 1996]).

2.2.1.2.3.7 Transformadas wavelet del contorno

La transformada wavelet de contorno también puede utilizarse en problemas de reconocimiento. En [Cheick 2000] se propone utilizar el módulo máximo de la transformada en el espacio de escalas de la descomposición mediante wavelets para el reconocimiento. Para ello, en cada nivel de la transformada, se calcula un valor de similitud en cada nivel de descomposición basado en comparar si dos máximos están suficientemente cerca y la diferencia de sus magnitudes se sitúan por debajo de un umbral. Para hacer robusto el proceso de emparejamiento no se utilizan todos los niveles de la transformación sino únicamente los asociados a escalas más burdas de las funciones wavelets. Los autores obtienen resultados comparables a la descomposición del espacio de escalas (sección anterior) en un tiempo menor.

Otras técnicas incluyen un algoritmo de emparejamiento basado en el cálculo de la transformada continua mediante wavelets [Yoon 1998] que permite abordar problemas con cambio de escala y oclusión. Finalmente en [Tieng 1997] puede encontrarse un algoritmo de emparejamiento invariante frente a transformaciones afines mediante wavelets utilizando una variación de la técnica del módulo máximo.

**2.2.2 Emparejamiento de objetos deformables**

A continuación se exponen el conjunto de técnicas para el emparejamiento de objetos deformables haciendo énfasis en las distintas formulaciones de la función de ajuste y en los diferentes procedimientos de optimización.

2.2.2.1 Criterios de ajuste a la imagen

El criterio de ajuste a la imagen, que mide la discrepancia entre los datos y el valor actual del modelo, proporciona una guía para que la deformación del modelo se ajuste a los contenidos de la imagen. Existe muchos criterios diferentes para objetos deformables en la literatura que puede ser simplificado en cuatro clases: basados en imagen, basados en bordes, basados en regiones y basados en movimiento.

• **Basados en la imagen.**

Este criterio busca el ajuste del modelo a “píxeles” de alta intensidad. Para ello se define una función de ajuste a los datos como:

$$E_{ajud}(w; M_i, D) = -\sum_k |I(p_k(w; D))| \tag{Ec. 2.64}$$

donde  $I(p; D)$  denota la intensidad de la imagen D en  $p$  y  $\{p_k(w), \forall k\}$  denota la posición de las coordenadas del modelo definidas por  $w$ .

• **Basados en bordes.**

Se utiliza cuando los bordes de la imagen son las características de interés, y se definen utilizando el gradiente de la función de intensidad:

$$E_{aju2}(w; M_i, D) = -\sum_k |G(\mathbf{p}_k(w; D))| \quad \text{Ec. 2.65}$$

donde  $G(\mathbf{p}; D)$  denota el gradiente de la imagen  $D$  en  $\mathbf{p}$ .

Si se desea utilizar también la dirección del gradiente puede definirse como:

$$E_{aju3}(w; M_i, D) = -\sum_k |1 - \mathbf{h}_v G(\mathbf{p}_k(w; D))| \quad \text{Ec. 2.66}$$

donde  $\mathbf{h}_v$  es el vector normal en  $\mathbf{p}(w)$ . Para hacer que el modelo favorezca regiones de gradiente constante, puede definirse:

$$E_{aju4}(w; M_i, D) = -\sum_k |G(\mathbf{p}_k(w; D)) - \bar{G}| \quad \text{Ec. 2.67}$$

con  $\bar{G}$  el valor medio del gradiente sobre los puntos del modelo.

- **Basados en regiones.**

En este caso se busca llevar el modelo hacia regiones con niveles de gris homogéneos. Este criterio puede definirse como:

$$E_{aju5}(w; M_i, D) = -\sum_k |H(\mathbf{p}_k(w; D))| \quad \text{Ec. 2.68}$$

donde  $H$  es la función indicador de los “pixels” de la región que pueden ser calculados mediante algún algoritmo de segmentación.

- **Basados en primitivas**

Se pretende llevar a las primitivas hacia sus correspondientes posiciones en la imagen. El criterio puede definirse como:

$$E_{aju6}(w; M_i, D) = -\sum_{k,l} |\mathbf{r}_{kl} P_l(\mathbf{p}_k(w; D))| \quad \text{Ec. 2.69}$$

donde  $P_l$  representa el mapa de distancias de las primitivas del tipo  $l$  y  $\mathbf{r}_{kl}$  es el coeficiente de pertenencia de la primitiva localizada en  $\mathbf{p}_k(w)$  a la clase  $l$ .

- **Basados en movimiento**

Se utilizan para el seguimiento de objetos deformables que se mueven en una secuencia. Puede definirse como:

$$E_{aju7}(w; M_i, D) = -\sum_k M_v(\mathbf{p}_k(w; D)) \quad \text{Ec. 2.70}$$

donde  $M_v$  es una función indicador de los “píxeles” que se mueven.

- **Términos adicionales**

En algunos problemas se han utilizado algunos términos adicionales para mejorar el rendimiento. Por ejemplo, optimizar  $E_{defl}$  sin ninguna fuerza externa hace que el modelo se colapse a un punto. Para evitarlo, puede utilizarse una fuerza de inflado que lo evita como:

$$E_{ad\tilde{n}}(w; M_i, D) = -A(w) \quad \text{Ec. 2.71}$$

donde  $A(w)$  es el área del contorno del modelo definido por  $w$ .

### 2.2.2.2 Regularización

La deformación del modelo  $E_{def}$  y el ajuste a los datos  $E_{aju}$  son generalmente dos criterios en conflicto y es necesario establecer un compromiso entre ambos. La forma más usual es definir una función criterio que sea una combinación de ambas. Tal técnica, se llama regularización, y se ha utilizado durante mucho tiempo para tareas en Visión por Computador [Poggio 1985]. Para ello se suele recurrir por simplicidad a una suma pesada de los términos de deformación y ajuste mediante un parámetro de regularización  $t$ .

$$E_{tot}(w, t; M_i, D, \bar{w}_i) = t E_{def}(w; M_i, \bar{w}_i) + E_{aju}(w, M_i, D) \quad \text{Ec. 2.72}$$

Desde el punto de vista de la regularización, la extracción de contornos u objetos de una imagen, es un problema mal definido, es decir, tiene insuficientes restricciones y no tiene una única solución.  $E_{def}$  actúa como un regularizador para hacer que el proceso de la extracción o el emparejamiento se convierta en un problema bien definido.

La determinación del valor del parámetro  $t$  es importante para un buen emparejamiento. Si  $t$  es muy pequeño,  $E_{def}$  domina, lo que implica que el modelo no se deformará mucho y el ajuste a los datos será pobre. Por el contrario si  $t$  es muy grande  $E_{aju}$  dominará y el resultado será muy sensible al ruido. En la mayor parte de los casos, el valor de  $t$  se determina de forma empírica. Para hacer el proceso más robusto, se varía de un valor grande a pequeño [Yuille 1992], se utiliza validación cruzada [Shahraray 1989], o técnicas bayesianas [Cheung 1995].

### 2.2.2.3 Métodos de emparejamiento

El emparejamiento de modelos deformables se lleva a cabo mediante la optimización de la función criterio  $E_{tot}$ . Puesto que en la mayoría de los casos, la función criterio tiene una gran cantidad de mínimos locales, se necesita un algoritmo de optimización local con una buena inicialización o en otro caso un buen algoritmo de optimización global.

#### 2.2.2.3.1 Inicialización

Tal y como se ha mencionado, una buena inicialización es muy importante para optimizar un mundo función criterio de gran complejidad y no lineal, que normalmente contiene una gran cantidad de óptimos locales sin importancia. La mayor parte de las aproximaciones asume que es la interacción con el usuario la que permite que éste de forma manual coloque el modelo cerca de las zonas de interés. Por ejemplo en

[Neuenschwander 1994] se comienza con una “snake” completa cuyos puntos finales se localizan de forma precisa en algunas posiciones deseadas. Entonces, la fuerza de la imagen entra en efecto progresivamente, comenzando a partir de los dos puntos finales.

Aunque estas aproximaciones evitan una gran cantidad de extremos locales, fallan, sin embargo, en casos con una gran cantidad de ruido. Otras aproximaciones incluyen la transformada generalizada de Gauss [Lai 1994] o los momentos del contorno [Blake 1998].

### 2.2.2.3.2 Optimización

La parte principal en el proceso de emparejamiento de modelos deformables es la de la optimización. Se han desarrollado diferentes métodos de optimización con éxito dispar. En general, si el énfasis se hace en el mejor rendimiento del emparejamiento, deben elegirse métodos de optimización global. En otro caso, para reducir la complejidad computacional con propósitos prácticos, es necesario utilizar métodos de optimización local. Por otra parte, la optimización a múltiples escalas se muestra como un buen compromiso entre los métodos globales y locales.

#### 2.2.2.3.2.1 Optimización local

Los métodos de optimización local intentan determinar una mejora de la función de ajuste a partir de la información proporcionada por una estimación inicial. Pueden dividirse en continuos y discretos.

- **Métodos de optimización continuos.**

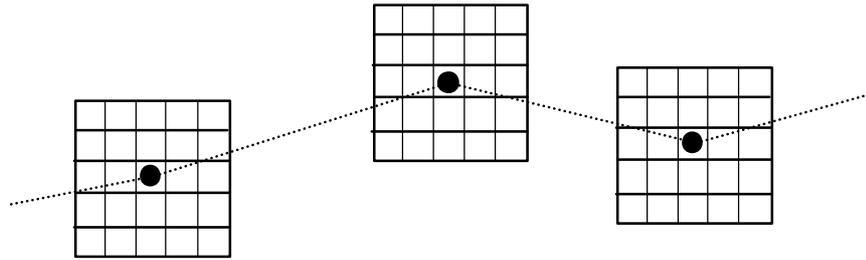
Para minimizar la función criterio para la emparejamiento del modelo, el método más simple con el que se puede empezar es el método del gradiente. Si se utiliza la formulación de los snakes, el proceso es equivalente a resolver la ecuación de Euler-Lagrange:

$$\underbrace{\frac{\partial(w_1 r)}{\partial s} - \frac{\partial^2(w_2 r)}{\partial s^2}}_{\text{fuerzas internas}} + \underbrace{\nabla F}_{\text{fuerza externa}} = 0 \quad \text{Ec. 2.73}$$

Para mejorar la convergencia, pueden utilizarse el método del gradiente conjugado o el método Levenberg-Marquardt [Xie 1994].

- **Métodos de optimización discretos**

Los métodos mencionados anteriormente utiliza coordenadas bidimensionales  $x$  e  $y$  como el dominio del problema. De forma alternativa, algunos métodos trabajan en una malla discreta sobre la imagen discretizada. Por ejemplo, en [Williams 1992] se utiliza un modelo inicial del contorno, definido por una lista ordenada puntos y se colocan unas ventana de búsqueda sobre cada punto. Utilizando un algoritmo “greedy” o ávido, se realiza una búsqueda secuencial en cada una de las ventanas se para optimizar la función criterio.



**Figura 2. 24 Optimización discreta mediante ventanas de búsqueda.**

#### 2.2.2.3.2.2 Optimización global

Los métodos de optimización local sólo pueden dar soluciones subóptimas. Para ser capaces de resolver el problema de forma óptima, se necesitan métodos de optimización global, como por ejemplo el “simulated annealing” [Kervrann 1994] o algoritmos genéticos [Lanitis 1993]. De forma alternativa, utilizando una malla discreta sobre la imagen se han utilizado también técnicas de programación dinámica [Geiger 1995] o el algoritmo A\* [Cooper 1979]. La optimalidad de estos algoritmos depende del tamaño de la ventana de búsqueda y de la resolución de la malla.

#### 2.2.2.3.2.3 Optimización multiescala

La optimización global es un problema computacionalmente costoso. La optimización multiescala [Lai 1994], [Brigger 2000] parece una buena aproximación puesto que, por una parte, puede evitar una gran cantidad de mínimo locales y al mismo tiempo el costo computacional que se requiere es mucho menor que el de los métodos globales.

#### 2.2.2.3.2.4 Soluciones analíticas

Es de resaltar que algunas implementaciones son capaces de ofrecer soluciones analíticas. Por ejemplo los MEF utilizados en [Pentland 1991] son capaces de obtener la solución siempre que los datos de entrada a la imagen estén distribuidos en los nodos a la hora de la formulación del modelo. En [Blake 1998] puede encontrarse otro ejemplo de solución analítica mediante la definición cuadrática de las funciones de deformación y ajuste.

### 2.2.2.4 Interpretación Probabilística

Es interesante señalar que tal y como han señalado muchos investigadores [Staib 1992], [Dubuisson 1996], [Blake 1998], el emparejamiento de modelos deformables

puede ser formulado como la estimación máximo a posteriori (MAP) con una adecuada formulación probabilística.

Desde un punto de vista probabilístico, la deformación de un modelo  $M_i(w)$ , cuantificado por  $E_{def}(w, M_i, \bar{w}_i)$ , puede ser interpretado como la incertidumbre de que se deforme de un modelo de referencia  $M_i(\bar{w}_i)$ . El desajuste a los datos  $E_{aju}(w, M_i, D)$  puede ser interpretado como la incertidumbre de que los datos de la imagen  $D$  provengan realmente del modelo  $M_i(w)$ . Mediante la distribución de Gibbs, estas incertidumbres pueden ser representadas como probabilidades, siendo  $E_{def}$  una distribución a priori de los parámetros del modelo y  $E_{aju}$  una función de verosimilitud de los datos de la imagen. Mediante la regla de Bayes,  $E_{tot}$  se corresponde a la distribución a posteriori. Es decir:

Probabilidad a priori :

$$P(w | \mathbf{t}; M_i, \bar{w}_i) \propto \exp(-\mathbf{t} E_{def}(w; M_i, \bar{w}_i))$$

Función de verosimilitud

**Ec. 2.74**

$$P(D | w, M_i) \propto \exp(-E_{aju}(w; M_i, D))$$

Distribución a posteriori

$$P(w | D, \mathbf{t}; M_i, \bar{w}_i) \propto \exp(-E_{tot}(w, \mathbf{t}; M_i, D, \bar{w}_i))$$

Por tanto hallar el mínimo de  $E_{tot}$  es equivalente a hallar el estimador máximo a posteriori (MAP) de  $P(w | D, \mathbf{t}; M_i, \bar{w}_i)$  con respecto a  $w$ .

## 2.3 Clasificación de Objetos

La forma más sencilla de realizar la clasificación basada en modelos consiste en realizar un emparejamiento de todos los modelos con la imagen, calcular alguna función discriminante  $Z$  para cada uno de ellos y elegir como candidato a aquel modelo que proporcione un mejor valor.. En general las funciones discriminantes están basadas en cuatro tipos [Cheung 1998], funciones específicas, funciones de ajuste, probabilidad a posteriori de los modelos y medidas basadas en clasificadores discriminantes.

### 2.3.1 Funciones específicas

Mediante el conocimiento de alguna de las características del problema es posible determinar algunas funciones de ajuste específicas. Por ejemplo en [Burr 1981] se utiliza una distancia basada en las incompatibilidades posicionales y direccionales entre los modelos y los datos para la clasificación. Otra aproximación consiste en analizar la distribución espacial de las partes de los modelos no emparejadas para inferir una oclusión por otro objeto [Breuel 1993].

### 2.3.2 Funciones de ajuste combinadas

La diferencia principal entre la utilización de la función de ajuste combinada  $E_{tot}$  y una función específica es que ésta además de para la clasificación sirve también para el emparejamiento. Esto proporciona el medio de integrar el emparejamiento y la clasificación en un único módulo. Este es el método más utilizado en la práctica [Lades 1993], [Dubuisson 1996], aunque no está justificado teóricamente.

### 2.3.3 Probabilidad a posteriori de los modelos

Una forma de clasificación justificada teóricamente es comparar las probabilidades a posteriori para los distintos modelos  $P(M_i | D)$ . Si todas las clases son igualmente probables el maximizar  $P(M_i | D)$  es equivalente a maximizar  $P(D | M_i)$ . Utilizando la regla de Bayes este término puede expresarse como:

$$P(D | M_i) = \int \frac{P(D | w, M_i)P(w | \mathbf{t}; M_i)}{P(w | D, \mathbf{t}; M_i)} P(\mathbf{t} | M_i) dt \quad \text{Ec. 2.75}$$

De la ecuación puede inferirse que si se pudiera realizar la integral de forma analítica no haría falta resolver el problema de emparejamiento y hallar el valor óptimo de los

parámetros presentes en  $\hat{w}$ . En la práctica sin embargo o bien la integral no se puede resolver analíticamente o bien es muy costoso resolverla computacionalmente. En general se suele utilizar el valor óptimo  $\hat{w}$  para calcular una aproximación [MacKay 1992].

#### **2.3.4 Medidas basadas en clasificadores**

Una forma de evitar que la falta de precisión en el modelado y la formulación de la función discriminante afecte al rendimiento del clasificador es la utilización de los valores de  $E_{def}$  y  $E_{aju}$  para distintos modelos y utilizarlos en una red neuronal [Williams 1994] o un clasificador estadístico. El rendimiento del clasificador se optimiza durante la fase de entrenamiento de forma que pueden reducirse las imprecisiones en la descripción del modelo. Las limitaciones de esta metodología consisten en la necesidad de un gran conjunto de datos para entrenamiento y la necesidad de un reentrenamiento de la red cada vez que se añade un nuevo modelo.

## 2.4 Seguimiento de Objetos basado en modelos

El movimiento es una fuente de gran importancia en la variación temporal que se encuentra en la secuencias de imágenes. Esta variación es inducida por el movimiento de los distintos objetos en una escena tridimensional y por el movimiento de la cámara.

La variación temporal de la posición de los objetos en una secuencia hace que en muchas circunstancias sea necesario el seguimiento de un objeto o número de objetos a lo largo de dicha secuencia. Un ejemplo puede ser el seguimiento un objeto en tareas de manipulación, vigilancia, conducción automática, etc.

Existen dos aproximaciones fundamentales al seguimiento de objetos. La primera no utiliza modelos y está basada en el campo de flujo óptico ( ver [Stiller 1999] para una revisión) es decir, un campo sobre la imagen que proporciona la velocidad de cada punto de ésta. A partir de esa información se puede inferir el movimiento, la estructura, o ambas para todos los objetos de la imagen. La segunda aproximación está basada modelos y puede definirse como la actualización de la configuración, en el plano de la imagen o en el espacio tridimensional, de valor específico del modelo del objeto  $M_i(w_t)$  a lo largo de una secuencia de imágenes. Dado que los modelos de los objetos están compuestos de primitivas y de las relaciones entre éstas, el seguimiento de primitivas, es un elemento fundamental en el problema del seguimiento de objetos basado en modelos. Las relaciones entre las primitivas existentes en el modelo, como por ejemplo, que las primitivas pertenecen a un objeto rígido, o que las primitivas determinan la forma de un contorno bidimensional, permiten una mayor consistencia en la estimación dado que las relaciones entre éstas restringen los posibles resultados del problema.

De acuerdo al tipo de movimiento que realizan los objetos, su evolución temporal puede clasificarse de acuerdo al siguiente esquema [Kambhamettu 1994]:

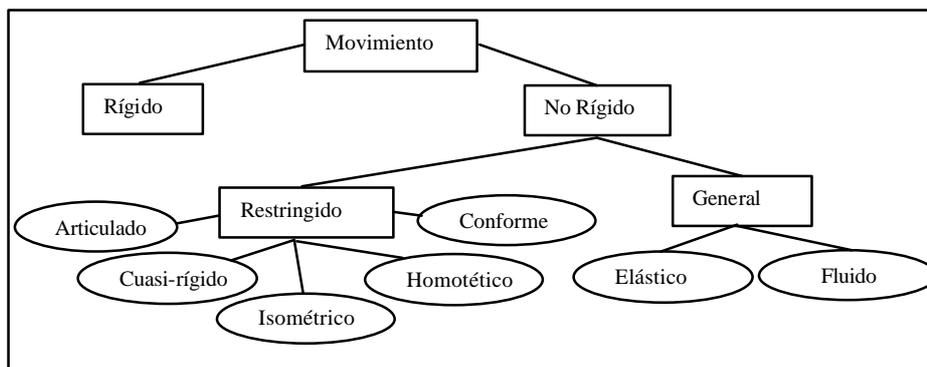


Figura 2. 25 Tipos de movimiento (de [Kambhamettu 1994])

En esta sección se abordarán los movimientos asociados a los tipos de modelos que se han presentado en las secciones anteriores, el movimiento rígido (asociado al modelo rígido) y el movimiento elástico (asociado al modelo deformable).

### 2.4.1 Cuestiones básicas del seguimiento basado en modelos

En el proceso del seguimiento de un objeto basado modelos se pueden identificar una serie de pasos que en todo o en parte presentan la mayoría de las aplicaciones. Estos son, tal como puede verse en la siguiente figura, la predicción, detección y emparejamiento, y actualización de los parámetros del modelo.

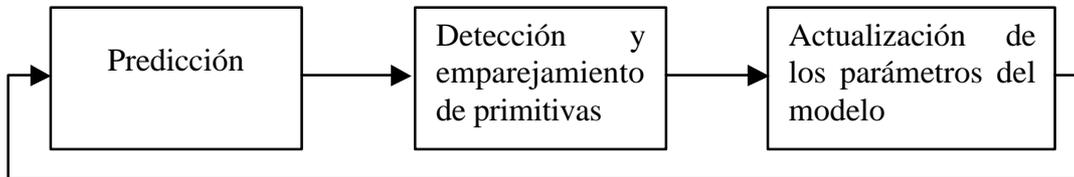


Figura 2. 26 El bucle de seguimiento

o bien siguiendo la notación de la Sección 2:

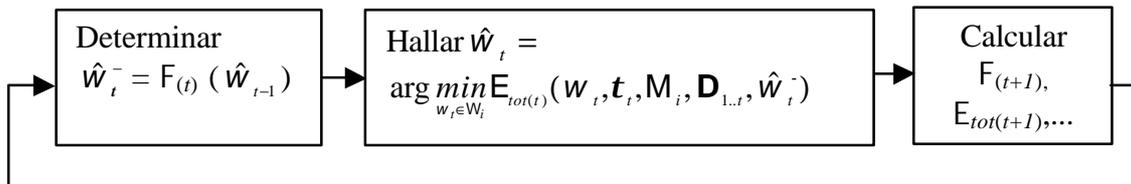


Figura 2. 27 Notación del bucle de seguimiento

Cada uno de estos pasos realiza:

- **Predicción de la localización**

Puesto que las primitivas son generalmente densas espacio-temporalmente, mantener el seguimiento en la posición y velocidad de una primitiva a lo largo del tiempo lleva generalmente a una mejor predicción de la localización de la primitiva en la siguiente imagen de la secuencia. Existen muchos modelos para el comportamiento dinámico de las primitivas, con el balance habitual entre calidad de la predicción y complejidad computacional.

- **Detección y seguimiento de primitivas**

Independientemente del modelo que se utilice para la predicción de la posición de las primitivas en la imagen, una vez que se calcule la posición de las primitivas en la imagen, debe haber algún proceso para buscar las primitivas correspondientes en la imagen. El problema de la correspondencia se ve simplificado debido a la presencia de la predicción permitiendo técnicas locales de búsqueda. En efecto, la predicción limita el conjunto de posibles primitivas en la imagen con las que es posible emparejar una primitiva del modelo a una determinada región alrededor de la posición predicha simplificando por tanto el proceso de emparejamiento.

- **Actualización de los parámetros del modelo**

El último paso del bucle lo constituye la utilización de la nueva posición del objeto para realizar una actualización de los parámetros que determinan la

predicción espacial de la primitiva, la función de ajuste total en la siguiente imagen teniendo en cuenta la calidad de la solución conseguida por el proceso de emparejamiento y los parámetros de regularización. En general para esta actualización se utilizan técnicas probabilísticas (ver Sección 2.4.1.1) como en filtro de Kalman o similares.

### 2.4.1.1 Utilización de técnicas probabilísticas

La utilización de técnicas probabilísticas es común en los métodos de seguimiento porque proporcionan un marco teórico para la determinación de las incertidumbres inherentes al proceso de seguimiento y define una metodología para integrar el proceso de toma de medidas en la imagen con el modelo dinámico del objeto.

#### 2.4.1.1.1 Esquema general

La aproximación probabilística del problema de seguimiento, al igual que en el caso del estudio de deformaciones, asume que cualquier cantidad de interés debe estimarse de la probabilidad a posteriori de  $w_t$  que define un valor específico del modelo dada la secuencia de imágenes  $D_{1..t}$  hasta tiempo  $t$ . Debe por tanto calcularse la distribución  $P(w_t | D_{1..t})$ . El esquema que se muestra a continuación permite una estimación recursiva de dicha distribución de probabilidad. Para ello se asume:

1. Se dispone de una distribución de probabilidad que describe la variación del vector de parámetros en el tiempo. Además el conjunto de estados  $w_{1..t}$  forman una cadena de Markov, es decir

$$P(w_t | w_{1..t-1}) = P(w_t | w_{t-1}) \quad \text{Ec. 2.76}$$

2. Se dispone de una distribución de probabilidad que describe los datos observados de la imagen  $D_{1..t}$  en función del vector  $w_{1..t}$ . Además se asume que los datos de la imagen son condicionalmente independientes de la evolución de los parámetros, es decir:

$$P(D_{1..t} | w_{1..t}) = \prod_{r=1..t} P(D_r | w_r) \quad \text{Ec. 2.77}$$

Entonces puede demostrarse que [Doucet 2000]:

$$P(w_t | D_{1..t}) \propto P(D_t | w_t) P(w_t | D_{1..t-1}) = \quad \text{Ec. 2.78}$$

$$P(D_t | w_t) \int_{w_{t-1}} P(w_t | w_{t-1}) P(w_{t-1} | D_{1..t-1}) dw_{t-1}$$

lo que permite una formulación recursiva del proceso de estimación.

### 2.4.1.1.2 El filtro lineal de Kalman

En el caso de que las distribuciones anteriores sean lineales y gaussianas aparecen las ecuaciones del filtro de Kalman. Si el modelo dinámico y de medidas cumple:

$$w_t | w_{t-1} \sim N(\mathbf{G}_t w_{t-1}, \mathbf{U}_t), \quad D_t | w_t \sim N(\mathbf{F}_t w_t, \mathbf{V}_t) \quad \text{Ec. 2.79}$$

entonces:

$$w_{t-1} | D_{1..t-1} \sim N(\hat{\mathbf{w}}_{t-1}^-, \mathbf{S}_{t-1}), \quad w_t | D_{1..t-1} \sim N(\hat{\mathbf{w}}_t^-, \mathbf{G}_t \mathbf{S}_{t-1} \mathbf{G}_t^T), \quad \hat{\mathbf{w}}_t^- = \mathbf{G}_t \hat{\mathbf{w}}_{t-1}^- \quad \text{Ec. 2.80}$$

y por tanto:

$$w_t | D_{1..t} \sim N(\hat{\mathbf{w}}_t^- + \mathbf{K}_t \mathbf{e}_t, \mathbf{R}_t - \mathbf{K}_t \mathbf{F}_t \mathbf{R}_t) \quad \text{Ec. 2.81}$$

lo que permite una estimación recursiva de los parámetros:

$$\begin{aligned} \hat{\mathbf{w}}_t &= \mathbf{G}_t \hat{\mathbf{w}}_{t-1}^- + \mathbf{K}_t \mathbf{e}_t \\ \mathbf{S}_t &= \mathbf{R}_t - \mathbf{K}_t \mathbf{F}_t \mathbf{R}_t \end{aligned} \quad \text{Ec. 2.82}$$

donde  $\mathbf{K}_t$  es la “ganancia de Kalman”,  $\mathbf{e}_t$  es el error de predicción y  $\mathbf{R}_t$  es la covarianza del vector de parámetros antes de la observación.

Estas cantidades se calculan también de forma iterativa como:

$$\begin{aligned} \mathbf{K}_t &= \mathbf{R}_t \mathbf{F}_t^T (\mathbf{V}_t + \mathbf{F}_t \mathbf{R}_t \mathbf{F}_t^T)^{-1} \\ \mathbf{e}_t &= D_t - \mathbf{F}_t \mathbf{G}_t \hat{\mathbf{w}}_{t-1}^- \\ \mathbf{R}_t &= \mathbf{G}_t \mathbf{S}_{t-1} \mathbf{G}_t^T + \mathbf{U}_t \end{aligned} \quad \text{Ec. 2.83}$$

Para aplicar estas ecuaciones es necesario conocer los parámetros  $\mathbf{G}_t$ ,  $\mathbf{U}_t$ ,  $\mathbf{F}_t$ , y  $\mathbf{V}_t$ , y realizar una estimación inicial de la media del vector de parámetros y su matriz de covarianzas;  $\hat{\mathbf{w}}_t^-$  y  $\mathbf{S}_0$ . Después las ecuaciones proceden de forma recursiva.

En este caso, de acuerdo con la notación del inicio del Capítulo 2

$$\begin{aligned} \hat{\mathbf{w}}_t^- &= \mathbf{F}_t (\hat{\mathbf{w}}_{t-1}^-) = \mathbf{G}_t \hat{\mathbf{w}}_{t-1} \\ E_{def(t)}(w_t, M_i, \hat{\mathbf{w}}_t^-) &= (w_t - \hat{\mathbf{w}}_t^-)^T (\mathbf{G}_t \mathbf{S}_{t-1} \mathbf{G}_t^T)^{-1} (w_t - \hat{\mathbf{w}}_t^-)^T \\ E_{aju(t)}(w, M_i, D_t) &= (D_t - w_t)^T \mathbf{V}_t^{-1} (D_t - w_t)^T \end{aligned} \quad \text{Ec. 2.84}$$

y no es necesaria definir ninguna constante de regularización puesto que está implícita en el modelo dinámico. Otra ventaja de esta formulación es la posibilidad de calcular la verosimilitud de los datos antes de realizar los cambios en el modelo:

$$P(D_t | D_{1..t-1}) = N(\mathbf{F}_t w_{t-1}, \mathbf{V}_t + \mathbf{F}_t (\mathbf{G}_t \mathbf{S}_{t-1} \mathbf{G}_t^T) \mathbf{F}_t^T) \quad \text{Ec. 2.85}$$

En terminología bayesiana esta verosimilitud es la evidencia del conjunto de datos  $D_t$ . Un conjunto de datos con poca evidencia corresponde con períodos en los cuales la estructura estadística del proceso cambia (no estacionalidad) o a períodos con un alto nivel de ruido. Este hecho suele utilizarse para descartar correspondencias en el seguimiento de primitivas.

### 2.4.1.1.3 Otras aproximaciones.

El filtro de Kalman es el filtro lineal óptimo para la estimación del vector de parámetros cuando las funciones de distribución son gaussianas. Existen otras extensiones estándar a este filtro cuando aparecen distribuciones no gaussianas como el filtro extendido de Kalman (FEK) [Gelb 1974] para el caso unimodal no lineal, el filtro probabilístico de asociación de datos (FPAS) [Bar-Shalom 1988] para el caso multimodal mezcla de gaussianas o el filtro “Condensation” [Isard 1998a] para funciones de distribución generales.

## 2.4.2 Seguimiento de modelos rígidos

Los métodos de seguimiento de objetos rígidos están basados en la actualización a través de toda la secuencia de imágenes de la posición tridimensional del objeto. Para ello se basan en la resolución del problema del emparejamiento entre las primitivas del modelo e imagen aprovechando la predicción proporcionada por la imagen anterior. De esta forma se restringe el conjunto de asignaciones entre primitivas del modelo y primitivas de la imagen. A continuación se muestran algunos ejemplos

### 2.4.2.1 Método integrado de Lowe

Este método [Lowe 1992] está diseñado para el seguimiento de modelos 3D basados en segmentos. Para ello, en cada imagen se calcula la transformación  $H(x, \bar{w})$  que proyecta de forma óptima el modelo en la imagen. Dada una predicción  $\hat{x}_t^-$  de los parámetros de la transformación, se definen las funciones:

$$E_{def(t)}(w_t, M, \hat{w}_t^-) = \| \mathbf{W} (x_t - \hat{x}_t^-) - \mathbf{d}_t \|^2 \quad \text{Ec. 2.86}$$

con  $w_t = H(x_t, \bar{w})$ ,  $\hat{w}_t^- = H(\hat{x}_t^-, \bar{w})$ ,  $\mathbf{W}$  una matriz de regularización y  $\mathbf{d}_t$  un desplazamiento esperado en los parámetros de la transformación.

El término de ajuste  $E_{aju(t)}(w_t, M, D_t)$  se define como una medida del error entre el modelo y la imagen. Esta medida se lineariza en la predicción  $\hat{x}_t^-$  obteniéndose:

$$E_{aju(t)}(w_t, M, D_t) = \| \mathbf{J} (x_t - \hat{x}_t^-) - \mathbf{e}_t \|^2 \quad \text{Ec. 2.87}$$

con  $J$  el jacobiano de la transformación entre modelo e imagen evaluado en  $\hat{x}_t^-$  y  $e_t$  el error de emparejamiento y se define la función de ajuste combinado:

$$E_{aju(t)}(w_t, M, D_t) + t E_{def(t)}(w_t, M, \hat{w}_t^-) \quad \text{Ec. 2.88}$$

Para su optimización global se utiliza una variante del método de alineamiento (ver Sección 2.2.1.2.1.2). Esto es, se parte del número de segmentos mínimo en modelo e imagen para definir de forma unívoca una transformación  $H(x, \bar{w})$  del modelo en la imagen y se proyectan el resto de los segmentos determinándose la compatibilidad entre todos los segmentos proyectados y los correspondientes en la imagen obteniéndose un vector de error  $e_t$ .

En el método de Lowe se introducen dos variantes: En primer lugar, puesto que se tiene una predicción de la posición del modelo a partir de la imagen anterior, en la determinación del número de segmentos mínimo se utilizan en primer lugar aquellos segmentos de la imagen más compatibles con la predicción del modelo mediante una búsqueda primero en profundidad. Esta búsqueda lo hace más robusto pero también menos eficiente computacionalmente. En segundo lugar, una vez asignados los segmentos del modelo se refina el parámetro de la transformación mediante la optimización de Ec. 2.88. fijado el valor de  $e_t$ . En este método no se asume la existencia de un modelo dinámico aunque puede llevarse a cabo utilizando un valor  $d_t \neq 0$ .

#### 2.4.2.2 Sistema RAPID

El método RAPID [Harris 1992] tiene características similares al de Lowe sin embargo el proceso de búsqueda es más sencillo y por lo tanto es más rápido a costa de cierta pérdida de robustez. El proceso de búsqueda está basado en los llamados “puntos de control” sobre el contorno del objeto y la búsqueda de correspondencias se limita a las direcciones perpendiculares a los bordes proyectados del modelo buscando el borde más importante en la imagen (Ver Figura 2. 28)

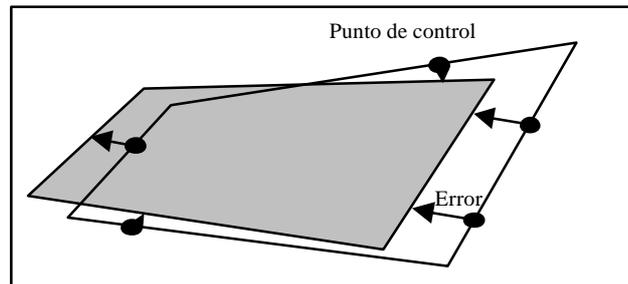


Figura 2. 28 Medidas efectuadas en el método de Harris

Posteriormente, en base a estos errores se calcula una corrección en el vector de transformación y se calcula la nueva posición del objeto mediante un filtro de Kalman. Este método ha sido mejorado para hacerlo más robusto utilizando técnicas de aleatorización para la eliminación de emparejamientos incorrectos y para la utilización de diversos modelos de cámaras en [Armstrong 1995].

### 2.4.2.3 Método robusto de Marchand

Este método [Marchand 1999] utiliza también un conjunto de puntos de control  $w_i = (p_0, p_1, p_2 \dots)$  (se omite el subíndice temporal en los  $p_i$  para simplificar la notación) en los sobre el contorno para realizar la búsqueda en direcciones normales a éste. En vez de buscar el borde más importante, para cada  $p_i$  se utiliza un núcleo  $M_v$ , donde  $v$  es la dirección normal a la proyección del punto y se calcula su nueva posición  $\hat{q}_i$  como:

$$\hat{q}_i = \arg \max_{q_i \in S_i} | l_t * M_v(p_i) + l_{t+1} * M_v(q_i) | \quad \text{Ec. 2.89}$$

donde  $S_i$  es un segmento centrado en  $p_i$  y  $l * M_v$  denota la convolución de la imagen y el núcleo. Esta asignación se realiza siempre que el valor del máximo sea mayor que un umbral  $I$ . A partir de este conjunto en bruto de transformaciones se evalúa la transformación afín que transforma modelo a imagen por medio de un estimador robusto y se obtiene una predicción refinada del modelo en la imagen. Para obtener la posición definitiva del modelo en la imagen se utiliza como función de ajuste:

$$E_{aju2(t)}(w_t; M, D_t) = - \sum_k |G(p_k(w_t; D_t))| \quad \text{con } w_t = H(X_t, \bar{W}), \quad \text{Ver Ec. 2.65}$$

o bien:

$$E_{aju3(t)}(w_t; M, D_t) = - \sum_k |1 - h_v G(p_k(w_t; D_t))| \quad \text{con } w_t = H(X_t, \bar{W}), \quad \text{Ver Ec. 2.66}$$

sobre los puntos visibles del modelo.

A continuación se realiza un proceso de optimización sobre la función de ajuste que proporciona la posición definitiva del objeto.

### 2.4.3 Seguimiento de modelos deformables

A continuación se muestran algunas técnicas de seguimiento de modelos deformables centrándonos en los modelos deformables descritos en la Sección 2.1.3. Dentro de éstos se describirán con más detalle los basados en modelos probabilísticos ya que dichas técnicas serán utilizadas en el capítulo 3.

#### 2.4.3.1 Kalman snakes

Las “snakes” ya fueron definidas en la Sección 2.1.3.1.1 dentro de los modelos deformables basados en contornos, la definición estática de su comportamiento se hace en base a la minimización de una función de energía representando las fuerzas elásticas y de tensión para asegurar que forma una curva suave, y un término dependiente de la imagen que mueve al contorno hacia los puntos de interés. Matemáticamente puede expresarse como:

$$\underbrace{\frac{\partial(w_1\mathbf{r})}{\partial s} - \frac{\partial^2(w_2\mathbf{r})}{\partial s^2}}_{\text{fuerzas internas}} + \underbrace{\nabla F}_{\text{fuerza externa}} = 0 \quad \text{Ver Ec. 2.73}$$

Donde los coeficientes  $w_1$  y  $w_2$  representan constantes que pesan la elasticidad y tensión respectivamente y  $F$  representa una imagen que asigna mayor peso a las características de interés. En el caso dinámico, el comportamiento de la “snake” cambia a:

$$\underbrace{\mathbf{r} \frac{\partial^2 \mathbf{r}}{\partial t^2}}_{\text{fuerza inercial}} = - \underbrace{\left( \mathbf{g} \frac{\partial \mathbf{r}}{\partial t} + \frac{\partial(w_1\mathbf{r})}{\partial s} + \frac{\partial^2(w_2\mathbf{r})}{\partial s^2} \right)}_{\text{fuerzas internas}} + \underbrace{\nabla F}_{\text{fuerza externa}} \quad \text{Ec. 2.90}$$

que es la ley de newton del movimiento para una “snake” con masa y bajo la influencia de fuerzas externas e internas. Los nuevos coeficientes son  $\mathbf{r}$ , la densidad de masa y  $\mathbf{g}$  la resistencia viscosa del medio en el que se mueve la “snake”. Además en este caso la fuerza externa  $F$  depende del tiempo.

La resolución del sistema en Ec. 2.90 se hace mediante discretización utilizando diferencias finitas o métodos numéricos de integración en el tiempo aplicados a simular el sistema resultante de ecuaciones diferenciales ordinarias. Cuando las ecuaciones de la evolución temporal del snake se interpretan de forma probabilística aparece el “Kalman snake” [Terzopoulos 1992].

### 2.4.3.2 Splines activos

En [Blake 1998] puede encontrarse un modelo probabilístico para realizar el seguimiento de objetos mediante contornos basados en b-splines utilizando el filtro de Kalman. Para ello se parte de una curva parametrizada como un b-spline y se define un modelo de deformación (ver Sección 2.1.3.1.2.2) como una probabilidad a priori sobre los parámetros:

$$P(w | M_i, \bar{w}_i) \propto \exp((w - \bar{w}_i)^T \mathbf{R} (w - \bar{w}_i)) \quad \text{Ver Ec. 2.45}$$

A partir de este modelo se define una probabilidad de ajuste del modelo a la imagen como:

$$P(D | w, M_i) \propto \exp\left( k \sum_{i=1}^n |D_i - \mathbf{r}(w)_i|^2 \right) \quad \text{Ec. 2.91}$$

con  $\mathbf{r}(w)_i$  el  $i$ -ésimo “punto de control” sobre la curva definida por los parámetros del b-spline,  $D_i$  el punto correspondiente al borde más importante sobre un segmento definido sobre la normal a la curva en dicho punto. En caso de que dicho segmento no se encuentre debido a ruido u oclusión se elimina dicho término de la suma en Ec. 2.91 (Ver Figura 2. 29).

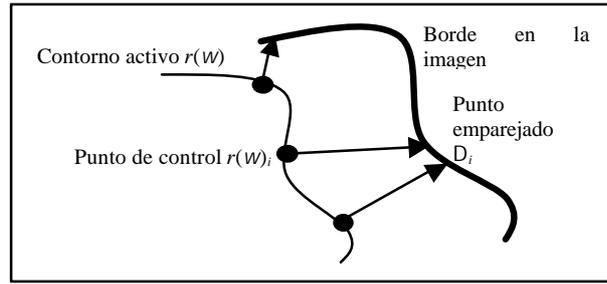


Figura 2. 29 Emparejamiento en contornos activos

Además se define el modelo dinámico autoregresivo de segundo orden:

$$w_t - \bar{w}_i = A_2(w_{t-2} - \bar{w}_i) + A_1(w_{t-1} - \bar{w}_i) + B_0 n_k, \quad n_k \sim N(0, I) \quad \text{Ec. 2.92}$$

Las distribuciones gaussianas de las ecuaciones anteriores permiten hacer el seguimiento mediante un filtro de Kalman. Este modelo tiene múltiples ventajas, por una parte el proceso de regularización entre las medidas tomadas de la imagen y el modelo dinámico es realizado automáticamente por el filtro, por otra parte el filtro de Kalman ofrece una distribución a priori de las medidas antes de realizar la búsqueda sobre la imagen (Ec. 2.85) lo que permite determinar de forma teórica la amplitud del segmento de búsqueda alrededor de cada punto. Además se ofrecen métodos de aprendizaje para todos los parámetros del algoritmo.

### 2.4.3.3 Algoritmo “Condensation” (Conditional Density Propagation)”

El algoritmo desarrollado en la sección anterior es capaz de realizar el seguimiento siempre que la cantidad de ruido en la imagen sea moderada. En el caso de que haya una cantidad sustancial de ruido la distribución de probabilidad en Ec. 2.91 se vuelve multimodal y por tanto se pierde la distribución gaussiana en la que se basa el filtro de Kalman. Para este caso general la distribución de probabilidad sigue la regla:

$$P(w_t | D_{1..t}) \propto P(D_t | w_t) \int_{w_{t-1}} P(w_t | w_{t-1}) P(w_{t-1} | D_{1..t-1}) dw_{t-1} \quad \text{Ver Ec. 2.78}$$

para aproximar la distribución \$P(w\_t | D\_{1..t})\$ en el algoritmo “Condensation” [Isard 1998a] se utilizan los llamados “filtros de partículas” [Doucet 2000] que son un proceso de estimación recursiva con la siguiente formulación:

Dado un conjunto de muestras \$S\_{t-1} = \{(w\_{t-1}^1, p\_{t-1}^1), (w\_{t-1}^2, p\_{t-1}^2), \dots, (w\_{t-1}^n, p\_{t-1}^n)\}\$

con \$p\_{t-1}^i\$ la probabilidad de que \$w\_{t-1}^i\$ sea elegido y que aproxima el muestreo de la distribución \$P(w\_{t-1} | D\_{1..t-1})\$

- Paso 1: Elegir un elemento de \$\tilde{w}\_{t-1}^i\$ de \$S\_{t-1}\$ al azar utilizando los \$p\_{t-1}^i\$ y generar una muestra \$w\_t^i\$ de la distribución \$P(w\_t | w\_{t-1} = \tilde{w}\_{t-1}^i)\$ (Utilizando por ejemplo el modelo

dinámico de Ec. 2.92). Esto genera un conjunto de muestras que aproxima la distribución de  $P(w_t | D_{1..t-1})$ .

- Paso 2: Para cada muestra  $w_t^i$  evaluar:  $p_t^i = P(D_t | w_t = w_t^i)$  y normalizarlos de forma que sumen uno:  $\sum_{i=1}^n p_t^i = 1$ . Construir  $S_t = \{(w_t^1, p_t^1), (w_t^2, p_t^2), \dots, (w_t^n, p_t^n)\}$ , incrementar  $t$  e ir al paso 1.

Para estimar (ver Figura 2. 30) la esperanza a posteriori  $E[g(w_t) | D_{1..t}]$  puede utilizarse:

$$E[g(w_t) | D_{1..t}] \approx \frac{\sum_{i=1}^n g(w_t^i) p_t^i}{\sum_{i=1}^n p_t^i} \quad \text{Ec. 2.93}$$

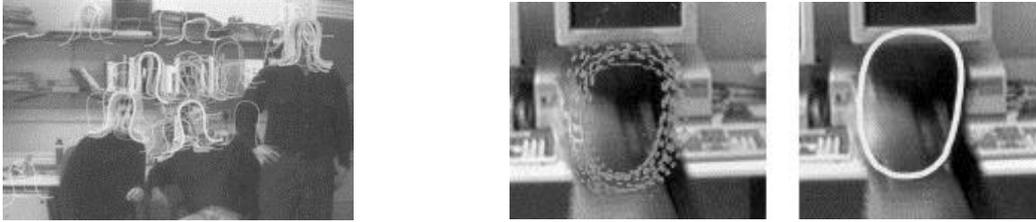


Figura 2. 30 (Izquierda) Muestras de la distribución a posteriori de las curvas. El nivel de gris representa la probabilidad. (Centro y derecha) Distribución de curvas y estimación de la media. (de [Blake 1998]).

La principal ventaja del método Condensation es su generalidad pues permite utilizar cualquier tipo de modelo dinámico o función de ajuste a los datos siempre que esté formulada de forma probabilística. Proporciona además una forma elegante de unir la incertidumbre de las mediciones sobre la imagen con la incertidumbre de la dinámica del objeto lo que lo hace prácticamente libre de parámetros. Este método se considera el “estado del arte” en métodos de seguimiento. No obstante también presenta el inconveniente de los métodos basados en muestreo: su costo computacional, aunque modificaciones de éste método que incluyen muestreo por importancia (algoritmo Icondensation) [Isard 1998b] son capaces de efectuar el seguimiento en tiempo real.

#### 2.4.3.4 Otros métodos

Otros métodos de interés en el seguimiento de objetos deformables incluyen:

- **Seguimiento basado en elementos finitos**  
Este método [Nastar 1994] está basado en la extensión de la metodología de deformación basada en elementos finitos (ver Sección 2.1.3.1.3.2) a deformaciones temporales, para ello se define la evolución de un sistema en el tiempo bajo un conjunto de fuerzas exteriores mediante:

$$M\ddot{U} + C\dot{U} + KU = R_t \quad \text{Ec. 2.94}$$

con  $M$  la matriz de masa,  $K$  la matriz de rigidez y  $C$  la matriz de amortiguamiento. Mediante análisis modal (ver Sección 2.1.3.1.3.2) las amplitudes y velocidades modales pueden ser estimadas dinámicamente mediante integración temporal o utilizando un filtro de Kalman

- **Métodos basados en la distancia de Hausdorff**

Están basados en que el método de comparación entre modelo e imagen se realiza mediante la distancia parcial de Hausdorff (ver Sección 2.2.1.2.2.5). El movimiento del objeto se descompone en dos partes: un movimiento bidimensional, correspondiente a un cambio de posición del modelo y una deformación bidimensional [Huttenlocher 1992]. Su principal ventaja es la robustez y facilidad de descripción del modelo.

- **Método del subconjunto**

Está basado en la discretización del espacio de deformaciones de la curva [Freedman 1999]. Es un método aproximado de optimización global y su complejidad aproximada es  $O(nmDd^{-s})$  donde  $n$  es el número de puntos de la curva,  $m$  es el número de posibles correspondencias para cada punto de la curva,  $Dd$  es la distancia a la solución óptima y  $s$  es la dimensión del espacio de deformación. Su principal ventaja es la robustez (al ser un algoritmo de optimización global). No tiene modelo dinámico por lo que requiere que los desplazamientos entre imágenes sucesivas sean pequeños

## 3 Modelado de contornos deformables mediante wavelets

En este capítulo se proponen nuevas técnicas para el modelado y emparejamiento de curvas mediante representaciones basadas en wavelets. En las siguientes secciones se presenta un estudio teórico de las propiedades de dicha representación para modelar las posibles deformaciones de la curva en imágenes estáticas así como su evolución temporal en secuencias de imágenes. Estas deformaciones están basadas en la extensión de la representación de funciones mediante wavelets en espacios de Besov (sección 3.1.10) a la representación de curvas en dichos espacios. Estos modelos a priori de deformación se contrastan con las medidas tomadas de la imagen mediante un proceso de regularización (sección 2.2.2.2). El proceso de optimización resultante se aborda ofreciendo una solución aproximada en el caso general y una solución analítica en espacios de Sobolev. Finalmente en este capítulo se presenta la definición de modelos dinámicos basados en wavelets y se analizan sus propiedades para los modelos autoregresivos de primer y segundo orden. Estos modelos dinámicos serán utilizados en el capítulo 4 en problemas de seguimiento de objetos en imágenes.

### 3.1 Análisis de funciones mediante wavelets

En esta sección se realiza una breve introducción al análisis de funciones unidimensionales mediante wavelets y multiwavelets. Para ello se parte del concepto general de transformada continua mediante wavelets y se analiza la forma de reducir su redundancia mediante el concepto de series de wavelets. Seguidamente se introduce el concepto de análisis a múltiple resoluciones que permite llegar a una de las características más importantes de las series de wavelets como es el cálculo de los coeficientes de la representación mediante un algoritmo de complejidad lineal. Posteriormente se introducen los modelos estadísticos basados en wavelets y su relación con los espacios de Besov y se finaliza con un estudio de las propiedades de las distintas familias de wavelets.

#### 3.1.1 Introducción

La teoría de los wavelets (ver [Jawerth 1994], [Wojtaszczyk 1997], [Resnikoff 1998] para una introducción) está basada en la representación de funciones generales en términos de otras funciones más simples a diferentes escalas y en distintas posiciones. La idea de la representación de funciones en base a otras más simples no es nueva: es bien sabido a partir de la teoría de Fourier que una señal puede ser expresada como una suma, posiblemente infinita, de senos y cosenos. La gran desventaja de una expansión de Fourier es que posee resolución en frecuencia pero carece de resolución espacial.

El análisis mediante wavelets es la solución más reciente a los inconvenientes de la transformada de Fourier. En esta aproximación, el procedimiento empleado consiste en la adopción de una función wavelet prototipo, llamada *wavelet de análisis* o *wavelet*

*madre*. De esta forma puede realizarse el análisis espacial mediante una versión contraída de alta frecuencia del wavelet prototipo. Esto permite un análisis más preciso de funciones que tienen cambios bruscos localizados espacialmente.

La versatilidad de este análisis se demuestra en el gran número de campos en el que ha encontrado aplicación, como por ejemplo: astronomía, acústica, codificación de señales, procesamiento de señales e imágenes, óptica, física aplicada o bien en matemática teórica mediante su utilización en la resolución de ecuaciones diferenciales.

### 3.1.2 La transformada continua mediante wavelets

El punto de partida para la descomposición de funciones mediante wavelets lo constituye la transformada continua mediante wavelets (TCW). Para ello se parte de una función wavelet prototipo  $\mathbf{y}(u) \in L_2(\mathbb{I})$  que se traslada y dilata de forma continua mediante los parámetros de traslación y dilatación  $t$  y  $s$  respectivamente:

$$\mathbf{y}_{s,t}(u) = \frac{1}{\sqrt{|s|}} \mathbf{y}\left(\frac{u-t}{s}\right) \quad u \in \mathbb{I} \quad \text{Ec. 3.1}$$

donde el factor de proporcionalidad asegura que sus normas en  $L_2(\mathbb{I})$  sean independientes de  $s$ . La TCW de una función  $f \in L_2(\mathbb{I})$  se define como:

$$W(s,t) = \langle f, \mathbf{y}_{s,t} \rangle = \int_{-\infty}^{\infty} f(u) \mathbf{y}_{s,t}(u) du \quad \text{Ec. 3.2}$$

Se asumirá que el wavelet prototipo  $\mathbf{y}$  y su transformada de Fourier  $\hat{\mathbf{y}}$  son funciones con centros  $\bar{u}$  y  $\bar{w}$  y radios  $\Delta_u$  y  $\Delta_w$  finitos, definidos como:

$$\bar{u} = \frac{1}{\|\mathbf{y}\|_2^2} \int_{-\infty}^{\infty} u |\mathbf{y}(u)|^2 du, \quad \Delta_u = \frac{1}{\|\mathbf{y}\|_2^2} \int_{-\infty}^{\infty} (u - \bar{u})^2 |\mathbf{y}(u)|^2 du \quad \text{con:} \quad \text{Ec. 3.3}$$

$$\|\mathbf{y}\|_2 = \|\mathbf{y}\|_{L_2(\mathbb{R})} = \langle \mathbf{y}, \mathbf{y} \rangle$$

y de forma similar para  $\bar{w}$  y  $\Delta_w$ . Entonces de Ec. 3.2 y Ec. 3.3 puede verse que la TCW en  $(s,t)$  recoge información fundamentalmente sobre los intervalos  $[t+s\bar{u}-s\Delta_u, t+s\bar{u}+s\Delta_u]$  y del intervalo de frecuencias  $[(\bar{w}-\Delta_w)/s, (\bar{w}+\Delta_w)/s]$  que determinan la ventana espacial y de frecuencias para  $(s,t)$ .

Si el wavelet prototipo  $\mathbf{y}$  satisface la condición de admisibilidad [Jaweth 1994]:

$$C_y = \int_{-\infty}^{\infty} \frac{|\hat{\mathbf{y}}(w)|^2}{w} dw < \infty \quad \text{Ec. 3.4}$$

entonces la TCW tiene inversa:

$$f(u) = \frac{1}{C_y} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(s,t) \mathbf{y}_{s,t}(u) \frac{dsdt}{s^2} \quad \text{Ec. 3.5}$$

la condición de admisibilidad implica que  $\mathbf{y}$  tiene que oscilar por lo que se comporta como una onda.

La propiedad de la localización de los wavelets permite que la TCW se pueda utilizar en la detección de singularidades. Puede probarse [Jawerth 1994] que si una función  $f$  es continua de Hölder (o Lipschitz) de orden  $0 < \alpha < 1$ <sup>1</sup>, si y solo si la TCW tiene un comportamiento asintótico determinado por:

$$W(s,t) = o(s^{\alpha+1/2}) \quad \text{para } s \rightarrow 0 \quad \text{Ec. 3.6}$$

Un resultado similar puede obtenerse para  $\alpha \geq 1$ ; en este caso el wavelet prototipo debe tener un número de momentos nulos mayor que  $\alpha$ :

$$\int_{-\infty}^{\infty} u^p \mathbf{y}(u) du = 0 \quad \text{para } 0 \leq p \leq \alpha \quad \text{Ec. 3.7}$$

esta propiedad de decaimiento rápido junto a las propiedades de onda hacen que  $\mathbf{y}$  recibe el nombre de wavelet u “onda pequeña”.

### 3.1.3 Series de Wavelets

La definición de la TCW permite el estudio teórico de la transformación pero posee una gran redundancia. En general es deseable una representación con el menor número posible de componentes. Para ello se introduce las series de wavelets (SW) en la que el wavelet prototipo solo puede ser trasladado y escalado en pasos discretos. Esto se hace mediante la modificación de la representación en Ec. 3.1 obteniéndose:

$$\mathbf{y}_{j,k}(u) = \frac{1}{\sqrt{s_0^j}} \mathbf{y}\left(\frac{u - s_0^j \mathbf{t}_0 k}{s_0^j}\right), \quad j, k \in \mathbb{Z} \quad \text{Ec. 3.8}$$

donde  $s_0 > 1$  es un factor de dilatación y  $\mathbf{t}_0 > 0$  es el paso espacial. Generalmente se utiliza  $s_0=2$  y  $\mathbf{t}_0=1$  obteniéndose de esta forma un muestreo diádico. La SW queda entonces definida mediante:

$$WS(j,k) = W(s_0^j, k \mathbf{t}_0 s_0^j) = \int_{-\infty}^{\infty} f(u) \mathbf{y}_{j,k}(u) du \quad \text{Ec. 3.9}$$

El resultado de la transformación de una función  $f$  será un conjunto de coeficientes que se denominan descomposición en la serie de wavelets. La utilidad de esta descomposición radica en que es posible reconstruir la función a partir de dichos

<sup>1</sup>  $f$  es continua de Hölder (o Lipschitz) de orden  $0 < \alpha < 1$ , si y solo si  $|f(u+h)-f(u)| = o(h^\alpha)$

coeficientes (existe la inversa de la transformación). Para ello debe cumplirse que la energía de los coeficientes esté acotada entre dos constantes positivas:

$$A\|f\|^2 \leq \sum_{j,k} |\langle f, \mathbf{y}_{j,k} \rangle|^2 \leq B\|f\|^2 \quad \text{Ec. 3. 10}$$

Cuando esto ocurre, la familia de funciones básicas  $\mathbf{y}_{j,k}$  se denomina marco o *frame*. En el caso de que  $A=B$  la familia se comporta como una base ortogonal. Cuando  $A \neq B$  la reconstrucción también es posible pero en este caso el wavelet de descomposición es distinto del de reconstrucción.

En el caso ortogonal una función arbitraria puede reconstruirse mediante la suma de las funciones básicas pesadas por los coeficientes de la transformación:

$$f(u) = \sum_{j,k} WS(j,k) \mathbf{y}_{j,k}(u) \quad \text{Ec. 3. 11}$$

### 3.1.4 Análisis a múltiples resoluciones

La descomposición de una función  $f$  en términos de series de wavelets puede ser reinterpretado dentro del marco del análisis a múltiples resoluciones [Mallat 1989].

Un análisis multiresolución, (AMR) de  $L_2(\mathbb{R})$  se define como una sucesión de subespacios cerrados  $V_j$  de  $L_2(\mathbb{R})$ ,  $j \in \mathbb{Z}$ , con las propiedades:

1.  $V_j \subset V_{j+1}$  Ec. 3. 12
2.  $f(u) \in V_j \Leftrightarrow f(2u) \in V_{j+1}$
3.  $f(u) \in V_0 \Leftrightarrow f(u+1) \in V_0$
4. clausura  $\bigcup_{j=-\infty}^{\infty} V_j = L_2(\mathbb{R})$  y  $\bigcap_{j=-\infty}^{\infty} V_j = \{\mathbf{0}\}$
5. Existe una función de escala  $\mathbf{j} \in V_0$ , con integral no nula de forma que el conjunto  $\{\mathbf{j}(u-l), l \in \mathbb{Z}\}$  es una base de Riesz de  $V_0$ <sup>2</sup>.

Puesto que  $V_0 \subset V_1$ ,  $\mathbf{j}(u)$  puede escribirse como una combinación lineal de  $\{\mathbf{j}(2u-l)\}$  que son versiones escaladas y trasladadas de  $\mathbf{j}(u)$ . De forma más precisa:

$$\mathbf{j}(u) = 2 \sum_k h_k \mathbf{j}(2u-k) \quad \text{Ec. 3. 13}$$

Es entonces inmediato comprobar que el conjunto de funciones  $\{\mathbf{j}_{j,l}, l \in \mathbb{Z}\}$ , con:  $\mathbf{j}_{j,l}(u) = \mathbf{j}(2^j u - l)$ ,  $j, l \in \mathbb{Z}$  es una base de Riesz de  $V_j$ .

Se dice que los wavelets tienen soporte compacto si únicamente un número finito de los  $h_k$  son no nulos en la relación anterior.

---

<sup>2</sup> El conjunto  $\{f_l, l \in \mathbb{Z}\}$  es una base de Riesz si es un marco o *frame* de vectores linealmente independientes

Dado un conjunto de subespacios anidados  $V_j$ , la propiedad 1 del AMR implica que existen subespacios  $W_k$ , que son el complemento de  $V_j$  en  $V_{j+1}$ , es decir:

$$V_{j+1} = V_j \oplus W_j, \quad j \in Z \quad \text{Ec. 3. 14}$$

Puesto que los subespacios  $V_j$  están anidados entonces:

$$V_j = V_j \oplus \bigoplus_{k=0}^{j-1} W_{j+k}, \quad j < J \quad \text{y} \quad L_2(\mathbb{R}) = \bigoplus_{j \in Z} W_j \quad \text{Ec. 3. 15}$$

Además los subespacios  $W_j$  cumplen:  $f(u) \in W_j \Leftrightarrow f(2u) \in W_{j+1}$

Dada una función de escala  $\mathbf{j}$  en  $V_0$ , el principio básico del AMR es la existencia de otra función  $\mathbf{y}$  en  $W_0$  llamada wavelet de forma que el conjunto  $\{\mathbf{y}(u-l), l \in Z\}$  es una base de Riesz de  $W_0$ . El conjunto de funciones wavelets  $\mathbf{y}_{j,l}(u) = \mathbf{y}(2^j u - l)$ ,  $j, l \in Z$  es entonces una base de Riesz de  $L_2(\mathbb{R})$ .

De esta forma, si  $P_j$  es el operador de proyección en  $V_j$  y  $Q_j$  es el operador de proyección en  $W_j$  una función  $f$  puede escribirse como:

$$f(u) = \sum_j Q_j f(u) = \sum_{j,l} d_{j,l} \mathbf{y}_{j,l}(u) \quad \text{Ec. 3. 16}$$

que es la descomposición de la función  $f$  en series de wavelets.

### 3.1.5 Wavelets ortogonales

La clase de wavelets ortogonales es particularmente interesante e introduce el concepto de análisis multiresolución ortogonal. Este es AMR en el que los subespacios wavelet  $W_j$  se definen como el complemento ortogonal de  $V_j$  en  $V_{j+1}$ . En consecuencia, los espacios  $W_j$  con  $j \in Z$  son todos mutuamente ortogonales.

Una condición suficiente para un AMR ortogonal es que los subespacios  $V_0$  y  $W_0$  sean ortogonales. Una función de escala ortogonal es una función  $\mathbf{j}$  definida de forma que el conjunto determinado por  $\{\mathbf{j}(u-l), l \in Z\}$  sea una base ortogonal de  $V_0$ , y una función wavelet ortogonal es una función  $\mathbf{y}$  tal que el conjunto  $\{\mathbf{y}(u-l), l \in Z\}$  es una base ortogonal de  $W_0$ , puesto que los subespacios  $W_j$  son ortogonales, el conjunto de funciones:  $\{\mathbf{y}_{j,l}, j, l \in Z\}$  es entonces una base ortonormal de  $f \in L_2(\mathbb{R})$ .

Esto permite que pueda escribirse para una función  $f \in L_2(\mathbb{R})$ :

$$P_j f(u) = \sum_l \langle f, \mathbf{j}_{j,l} \rangle \mathbf{j}_{j,l}(u), \quad Q_j f(u) = \sum_l \langle f, \mathbf{y}_{j,l} \rangle \mathbf{y}_{j,l}(u) \quad \text{Ec. 3. 17}$$

Por tanto, para una función  $f \in L_2(\mathbb{R})$  se tiene:

$$f(u) = \sum_{j,l} d_{j,l} \mathbf{y}_{j,l}(u), \quad d_{j,l} = \langle f, \mathbf{y}_{j,l} \rangle \quad \text{Ec. 3. 18}$$

que es la descomposición de  $f$  como serie de wavelets ortogonales.

### 3.1.6 Wavelets biortogonales

La condición de ortogonalidad establece una fuerte limitación en la construcción de wavelets. Por ejemplo, únicamente existe un wavelet (wavelet de Haar) que sea real, de soporte compacto simétrico y ortogonal. Para ello se ha introducido la noción de wavelet biortogonal. En este análisis se busca una función de escala dual  $\tilde{\mathbf{j}}$  y una función wavelet dual  $\tilde{\mathbf{y}}$  que general un AMR dual con subespacios  $\tilde{V}_j$  y  $\tilde{W}_j$  que verifican:

$$\tilde{V}_j \perp W_j \text{ y } V_j \perp \tilde{W}_j \quad \text{Ec. 3. 19}$$

por lo que:

$$W_j \perp \tilde{W}_{j'}, \quad j \neq j' \quad \text{Ec. 3. 20}$$

Puesto que las funciones duales definen un AMR deben cumplir:

$$\tilde{\mathbf{j}}(u) = 2 \sum_k \tilde{h}_k \tilde{\mathbf{j}}(2u-k) \quad \text{y} \quad \tilde{\mathbf{y}}(u) = 2 \sum_k \tilde{g}_k \tilde{\mathbf{y}}(2u-k) \quad \text{Ec. 3. 21}$$

por lo que en este caso el operador de proyección toma la forma:

$$P_j f(u) = \sum_l \langle f, \tilde{\mathbf{j}}_{j,l} \rangle \mathbf{j}_{j,l}(u), \quad Q_j f(u) = \sum_l \langle f, \tilde{\mathbf{y}}_{j,l} \rangle \mathbf{y}_{j,l}(u) \quad \text{Ec. 3. 22}$$

y

$$f(u) = \sum_{j,l} \langle f, \tilde{\mathbf{y}}_{j,l} \rangle \mathbf{y}_{j,l}(u) \quad \text{Ec. 3. 23}$$

nótese como la expresión anterior puede ser vista como una serie de wavelets donde las condiciones impuestas sobre  $\mathbf{y}$  son menos restrictivas que en el caso anterior.

En particular escribiendo  $\mathbf{j}(2u-k) \in V_1$  en las bases de  $V_0$  y  $W_0$  se tiene:

$$\mathbf{j}(2u-k) = \sum_l \tilde{h}_{k-2l} \mathbf{j}(u-l) + \sum_l \tilde{g}_{k-2l} \mathbf{y}(u-l) \quad \text{Ec. 3. 24}$$

Aunque la función de escala y wavelet no sean ortogonales, el AMR puede ser ortogonal. Cuando esto ocurre, la función de escala y wavelet biortogonal se llaman semiortogonales.

### 3.1.7 Aproximación mediante wavelets

La descomposición de una función  $f$  en series de wavelets permitieron caracterizar las condiciones de regularidad en la sección 3.1.2 hasta el orden  $a < M$  donde  $M$  es el número de momentos nulos del wavelet. Otra consecuencia es que puede derivarse la razón de convergencia de las aproximaciones mediante wavelets para funciones suaves: Si  $f$  es derivable de orden  $n$  ( $f \in C^n$ ) entonces [Jawerth 1994]:

$$\|P_j f - f\|_2 = o(h^n), \quad h = 2^{-j} \quad \text{Ec. 3. 25}$$

El número de momentos nulos del wavelet prototipo está también relacionado con su regularidad [Wojtaszczyk 1997].

### 3.1.8 La transformada rápida mediante wavelets

Una vez presentadas las características principales de los desarrollos de funciones en series de wavelets es necesario mostrar como calcular los coeficientes de la descomposición. El resultado fundamental es que existe un algoritmo de orden lineal que es capaz de realizar la transformación llamado la transformada rápida mediante wavelets (TRW).

El algoritmo está basado en la descomposición de subespacios siguiente: puesto que  $V_j = V_{j-1} \oplus W_{j-1}$ , una función  $v_j \in V_j$  puede escribirse de forma unívoca como la suma de una función  $v_{j-1} \in V_{j-1}$  y otra función  $w_{j-1} \in W_{j-1}$ :

$$\begin{aligned} v_j(u) &= \sum_k c_{j,k} \mathbf{j}_{j,k}(u) = v_{j-1}(u) + w_{j-1}(u) = \\ &= \sum_l c_{j-1,l} \mathbf{j}_{j-1,l}(u) + \sum_l d_{j-1,l} \mathbf{y}_{j-1,l}(u) \end{aligned} \quad \text{Ec. 3. 26}$$

Utilizando Ec. 3. 21 se puede pasar entre las dos representaciones:

$$c_{j-1,l} = \langle v_j, \tilde{\mathbf{j}}_{j-1,l} \rangle = \sqrt{2} \left\langle v_j, \sum_k \tilde{h}_{k-2l} \tilde{\mathbf{j}}_{j,k} \right\rangle = \sqrt{2} \sum_k \tilde{h}_{k-2l} c_{j,k} \quad \text{Ec. 3. 27}$$

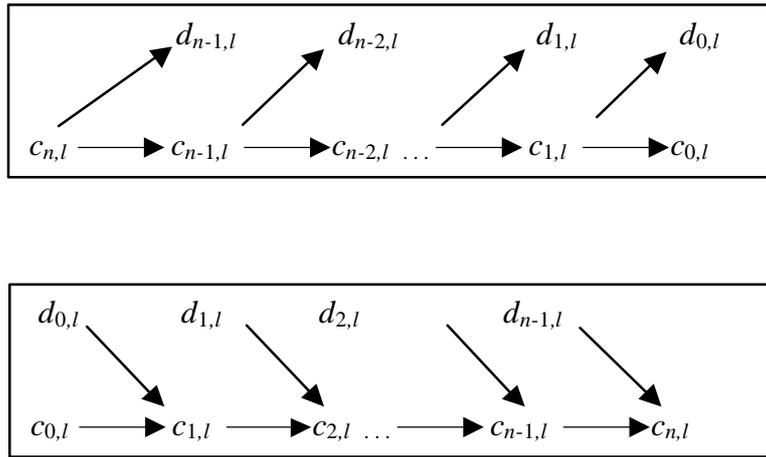
y de forma similar:

$$d_{j-1,l} = \sqrt{2} \sum_k \tilde{g}_{k-2l} c_{j,k} \quad \text{Ec. 3. 28}$$

El paso en la otra dirección de análoga complejidad. Utilizando Ec. 3. 24 se tiene

$$c_{j,k} = \sqrt{2} \sum_l h_{k-2l} c_{j-1,l} + \sqrt{2} \sum_l g_{k-2l} d_{j-1,l} \quad \text{Ec. 3. 29}$$

La aplicación recursiva de éstas fórmulas proporcionan la TRW. Las ecuaciones Ec. 3. 27 y Ec. 3. 28 definen la transformación hacia delante y la Ec. 3. 29 la transformación hacia atrás (ver Figura 3. 1)



**Figura 3. 1** Esquema de descomposición (arriba) y esquema de reconstrucción (abajo).

No obstante, se plantea el siguiente problema para una completa determinación del algoritmo: dada una función  $f$ , calcular con determinada precisión y costo computacional, los coeficientes  $c_{n,l}$  de una función en el subespacio  $V_n$  necesarios para calcular la TRW. Una solución trivial (que será la adoptada en la tesis) es:

$$c_{n,l} = f(l/2^n) \tag{Ec. 3. 30}$$

Aunque pueden utilizarse otros procedimientos basados en interpolación o cuadraturas [Walter 1992], [Sweldens 1994].

### 3.1.9 Wavelets en un intervalo

Hasta ahora se han discutido los wavelets sobre la recta real. Para muchas aplicaciones es necesario restringirse a un conjunto compacto sobre el que está definida la función. Para ser más específico el problema que se plantea sobre la recta real es la definición de los wavelets sobre un intervalo que sin pérdida de generalidad se asumirá el  $[0,1]$ .

Dada una función  $f$  en el intervalo  $[0,1]$ , la solución mas obvia es extenderla a la recta real definiendo la función  $f$  como  $f(u)=0$  fuera del intervalo, y utilizar la teoría anterior. Sin embargo en general, esta solución provoca discontinuidades fuera en los límites del intervalo.

Otra aproximación (que será la utilizada en esta tesis) está basada en considerar que la función es periódica con periodo 1,  $f(u+1)=f(u)$ . En este caso, los coeficientes a una escala dada satisfacen:

$$\langle f, \mathbf{y}_{j,k} \rangle = \langle f, \mathbf{y}_{j,k+2^j} \rangle, \quad k \in \mathbb{Z} \tag{Ec. 3. 31}$$

Esto permite reescribir las expansiones de funciones en la recta real en el intervalo  $[0,1]$  mediante la periodización de las funciones básicas como:

$$\mathbf{j}_{j,l}^*(u) = \mathbf{c}_{[0,1]}(u) \sum_m \mathbf{j}_{j,l}(u+m), \quad 0 \leq l \leq 2^j, j \geq 0 \quad \text{Ec. 3.32}$$

De forma análoga se define  $\mathbf{y}_{j,l}^*$ . Estas funciones periodizadas tienen las siguientes propiedades:

- La función de escala en  $V_0$  verifica  $\mathbf{j}_{0,0}^* = 1$
- Si el soporte de  $\mathbf{y}_{j,l}^*$  está contenido en  $[0,1]$  entonces  $\mathbf{y}_{j,l}^* = \mathbf{y}_{j,l}$

La primera propiedad será utilizada al estudiar el comportamiento de la representación frente transformaciones lineales de la función. La segunda nos indica que para resoluciones grandes los wavelets son similares para el caso general y el del intervalo lo que nos permite trasladar las propiedades entre ambos. En la práctica, hay un límite en la resolución de la señal determinada por un cierto subespacio  $V_n$  por tanto se tiene  $j \leq n$ .

### 3.1.10 Wavelets y suavidad de funciones

La teoría de espacios de suavidad de funciones tiene cada vez mayor importancia en el ámbito del procesamiento de señales e imágenes. En esta tesis se considerará la familia de espacios de Besov  $B_q^a(L_p(I))$  [Wojtaszczyk 1997], donde  $I$  será un dominio finito (generalmente el intervalo  $[0,1]$ ), y donde  $0 < a < \infty$ ,  $0 < p \leq \infty$ ,  $0 < q \leq \infty$ . Puede decirse que estos espacios tienen “ $a$  derivadas” en  $L_p(I)$ ; el parámetro  $q$  permite hacer mayores distinciones en la suavidad. Estos espacios se definen de la siguiente forma:

Dada una función  $f$  de variable real se considera para  $r > 0$  y  $h > 0$  la  $r$ -ésima diferencia de  $f$  como:

$$\Delta_h^{(r)} f(u) = \sum_{k=0}^r \binom{r}{k} (-1)^k f(u+kh), \quad u \in I_h^r = \{u \in I / u+rh \in I\} \quad \text{Ec. 3.33}$$

el módulo de suavidad de  $L_p(I)$  para  $0 < p \leq \infty$  se define como:

$$\mathbf{w}_r(f, u)_p = \sup_{|h| \leq u} \|\Delta_h^{(r)} f\|_{L_p(I_h^r)} \quad \text{Ec. 3.34}$$

La seminorma<sup>3</sup> de Besov de índice  $(a,p,q)$  se define para  $r > a$ , donde  $1 \leq p, q \leq \infty$  por:

<sup>3</sup> Una seminorma es una norma excepto por el hecho de que un elemento de norma cero no es necesariamente nulo

$$|f|_{B_q^a(L_p(I))} = \left( \int_0^\infty \left( \frac{\mathbf{w}_r(f, u)_p}{t^a} \right)^q \frac{dt}{t} \right)^{1/q}, \quad 0 \leq q \leq \infty \quad \text{Ec. 3. 35}$$

$$|f|_{B_q^a(L_p(I))} = \sup_{0 < t < \infty} \left( \frac{\mathbf{w}_r(f, u)_p}{t^a} \right), \quad q = \infty$$

La norma para el espacio de Besov se define entonces como:

$$\|f\|_{B_q^a(L_p(I))} = \|f\|_{L_p(I)} + |f|_{B_q^a(L_p(I))} \quad \text{Ec. 3. 36}$$

El espacio de Besov  $B_q^a(L_p(I))$  es entonces el conjunto de funciones que pertenecen a  $L_p(I)$  y verifican  $|f|_{B_q^a(L_p(I))} < \infty$ .

Una propiedad interesante de los wavelets es que proporcionan una caracterización simple de los espacios de Besov  $B_q^a(L_p(I))$ ,  $0 < a < \infty$ ,  $0 < p \leq \infty$ ,  $0 < q \leq \infty$ :

**Teorema 3. 1 [Wojtaszczyk 1997]**

Dadas una función de escala  $\mathbf{j}$  y wavelet  $\mathbf{y}$  con  $M > a$  momentos nulos, entonces si  $f$  es una función de periodo 1 y admite la descomposición en wavelets de la sección 3.1.8 la norma de Besov  $\|f\|_{B_q^a(L_p(I))}$  es equivalente a:

$$\|f\|_{B_q^a(L_p(I))} \equiv |c_{0,0}| + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p/2-1)} |d_{j,k}|^p \right)^{q/p} \right)^{1/q} \quad \text{Ec. 3. 37}$$

La interpretación de los parámetros es la siguiente: una  $p$ -norma de los coeficientes de la expansión se realiza en cada escala  $j$ , una  $q$ -norma se toma a través de las escalas, y el parámetro de suavidad  $a$  controla el decaimiento de los  $d_{jk}$  a través de las escalas. En esta tesis se llevará a cabo una generalización de esta norma al caso de los contornos.

**3.1.11 Modelos estadísticos basados en wavelets**

Los modelos estadísticos permiten el estudio de fenómenos en los cuales está presente la incertidumbre y ésta se puede expresar de forma probabilística. Los modelos estadísticos más simples para wavelets se obtienen asumiendo que cada coeficiente es independiente de los demás. Utilizando la independencia de los coeficientes, el modelo estadístico se reduce a especificar la distribución marginal de cada uno de los coeficientes. La distribución marginal de los coeficientes suele ser aproximada por la distribución gaussiana generalizada (DGG) que contiene entre otras a la distribución gaussiana. Otra distribución que suele utilizarse también es una mezcla de varias distribuciones gaussianas. En esta tesis se propondrá una generalización de la DGG para modelar los coeficientes del contorno.

La distribución gaussiana generalizada  $DGG_n(0, \mathbf{s}^2)$  de media 0, con parámetro de forma  $\mathbf{n}$  y varianza  $\mathbf{s}^2$  se define a partir de la función de densidad:

$$f(x) = \frac{\mathbf{n}\mathbf{h}(\mathbf{n})}{2\Gamma(1/\mathbf{n})} \frac{1}{\mathbf{s}} \exp\left(-(\mathbf{h}(\mathbf{n})|x|/\mathbf{s})^{\mathbf{n}}\right), \quad \mathbf{h}(\mathbf{n}) = \sqrt{\frac{\Gamma(3/\mathbf{n})}{\Gamma(1/\mathbf{n})}} \quad \text{Ec. 3. 38}$$

donde  $\Gamma$  hace referencia a la función gamma.

El modelo DGG contiene las distribuciones gaussiana y laplaciana como casos especiales, usando  $\mathbf{n} = 2$  y  $\mathbf{n} = 1$  respectivamente.

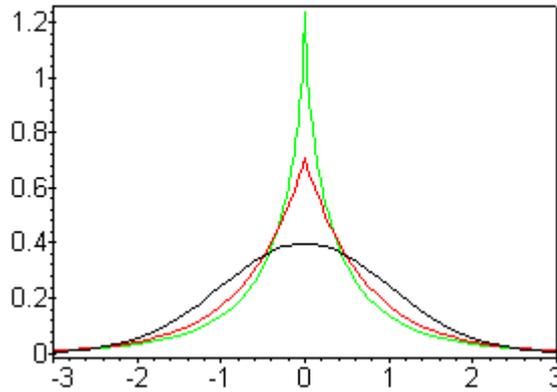


Figura 3. 2 Función de densidad GGD(0,1) para  $n = 2, n = 1, n = 0.7$  de mayor a menor achatamiento.

Una propiedad de la distribución que será utilizada más adelante es la siguiente:

**Proposición 3. 1 [Box 1973]**

Dada una variable aleatoria  $X$  distribuida como  $DGG_n(0, \mathbf{s}^2)$  se verifica que la esperanza de los momentos en valor absoluto verifican:

$$E(|X|^l) = D(\mathbf{n}, l) \mathbf{s}^l, \quad D(\mathbf{n}, l) = \left( \Gamma\left(\frac{l+1}{\mathbf{n}}\right) / \Gamma\left(\frac{1}{\mathbf{n}}\right) \right) \mathbf{h}(\mathbf{n})^{-l} \quad \text{Ec. 3. 39}$$

En un modelo DGG independiente, cada coeficiente se genera de forma independiente de acuerdo a la distribución anterior. Para que el modelo sea tratable, todos los coeficientes en cada escala se asumen independientes e idénticamente distribuidos (iid). Este hecho se conoce como modelo iid en escala. Por tanto bajo el modelo anterior se tiene:  $d_{j,k} \underset{iid}{\sim} DGG_{\mathbf{n}_j}(0, \mathbf{s}_j^2)$ .

El parámetro de forma  $\mathbf{n}_j$  en la escala  $j$  se determina para caracterizar la forma en pico y las amplias colas de la distribución (Figura 3. 2). En aplicaciones prácticas, se simplifica el modelo asumiendo que el parámetro de escala  $\mathbf{n}_j$  es el mismo para todas las

escalas. La varianza  $\mathbf{s}_j^2$  representa la energía de la señal en la escala  $j$ . Ésta puede estimarse de forma empírica a través de los datos o bien puede especificarse de forma que decaiga exponencialmente. Con este modelo de decaimiento se tiene:

$$d_{j,k} \underset{iid}{\sim} \text{DGG}_n(0, \mathbf{s}_j^2), \quad \mathbf{s}_j = 2^{-jb} \mathbf{s}_0 \quad \text{Ec. 3. 40}$$

Debido a que el decaimiento de los wavelets a través de las escalas determinan la suavidad de la señal, puede demostrarse que las realizaciones del modelo anterior pertenecen a los espacios de Besov. De forma más precisa:

**Teorema 3. 2 [Choi 1999]**

Dada una función unidimensional  $f$ , supongamos que cada coeficiente de su expansión se distribuye según el modelo de Ec. 3. 40 con  $\mathbf{b} > 0$  y  $\mathbf{s}_0 > 0$ . Entonces para  $0 < p, q < \infty$ , las realizaciones de dicho modelo están casi seguro en  $B_q^a(L_p(I))$  si y solo si  $\mathbf{b} > \mathbf{a} + 1/2$

Este teorema será generalizado al caso de contornos en la sección 3.3.1

**3.1.12 Derivadas basadas en wavelets**

A lo largo de la tesis será necesario el evaluar la derivada de los contornos para calcular su orientación local. En esta sección se muestra como derivar una función dada su expansión en  $V_n$ . Por simplicidad se abordará únicamente el caso ortogonal y periódico que será el utilizado en la tesis. De esta forma, si se tiene la descomposición de  $f$  en  $V_n$ :

$$P_n f(u) = \sum_k c_{n,k} \mathbf{j}_{n,k}(u) \quad \text{Ec. 3. 41}$$

su derivada será:

$$\frac{dP_n f(u)}{du} = \sum_k c_{n,k} \frac{d\mathbf{j}_{n,k}(u)}{du} \quad \text{Ec. 3. 42}$$

esta derivada debe ser reprojectada en  $V_n$ :

$$P_n \frac{dP_n f(u)}{du} = \sum_{k,l} c_{n,k} r_{k,l} \mathbf{j}_{n,l}(u), \quad r_{k,l} = \left\langle \frac{d\mathbf{j}_{n,k}(u)}{du}, \mathbf{j}_{n,l}(u) \right\rangle \quad \text{Ec. 3. 43}$$

puede comprobarse que  $r_{k,l} = r_{k-l,n}$ . Por tanto para el cálculo de la derivada únicamente es necesario conocer el conjunto de coeficientes  $r_{k,l}$ . Los valores de los coeficientes  $r_{k,l}$

para la familia de wavelets de Daubechies utilizadas en esta tesis pueden encontrarse en [Restrepo 1997]. El cálculo de la derivada en otros subespacios  $V_j$  es similar.

### 3.1.13 Multiwavelets

Los multiwavelets constituyen una extensión de los wavelets basada en la utilización de varias funciones de escala. Es sabido que mediante la utilización de un único wavelet (caso escalar) no es posible obtener simultáneamente las propiedades de soporte compacto, ortogonalidad, simetría y momentos nulos. Los multiwavelets no poseen dichas limitaciones y pueden por tanto ofrecer un rendimiento superior que los wavelets escalares.

De la misma forma que en el caso escalar, los multiwavelets están basados en la idea del análisis a múltiples resoluciones (AMR) (ver Sección 3.1.4). Para los multiwavelets la noción del AMR es la misma salvo que en este caso  $V_0$  está generado por traslaciones de  $r$  funciones básicas que de forma vectorial se escriben como:  $\mathbf{F}(u) = (\mathbf{j}^1(u), \dots, \mathbf{j}^r(u))^T$ . Dicho vector satisface una ecuación (matricial) de dilatación análoga al caso escalar:

$$\mathbf{F}(u) = \sum_k \mathbf{H}_k \mathbf{F}(2u - k) \quad \text{Ec. 3. 44}$$

donde ahora los coeficientes  $\mathbf{H}_k$  son matrices de orden  $r$  por  $r$  y no escalares.

Asociados con estas funciones de escala hay  $r$  wavelets básicas que de forma vectorial se escriben como:  $\mathbf{?}(u) = (\mathbf{y}^1(u), \dots, \mathbf{y}^r(u))^T$  y que satisfacen la ecuación:

$$\mathbf{?}(u) = \sum_k \mathbf{G}_k \mathbf{F}(2u - k) \quad \text{Ec. 3. 45}$$

donde de nuevo los coeficientes  $\mathbf{G}_k$  son matrices de orden  $r$  por  $r$  y no escalares.

De esta forma, una función  $f$  puede escribirse en el caso ortogonal como:

$$f(u) = \sum_{j,l} \mathbf{D}_{j,l}^T \mathbf{?}_{j,l}(u), \quad \mathbf{D}_{j,l} = (\langle f, \mathbf{y}_{j,l}^1 \rangle, \dots, \langle f, \mathbf{y}_{j,l}^r \rangle)^T \quad \text{Ec. 3. 46}$$

Análogamente al caso escalar, existe un algoritmo para el cálculo de los coeficientes de complejidad lineal.

### 3.1.14 Familias de wavelets

La elección del wavelet prototipo para realizar un análisis AMR no es única. Dependiendo de las propiedades deseadas pueden encontrarse distintas familias:

### 3.1.14.1 Wavelets ortogonales

Algunas familias de wavelets ortogonales [Jawerth 1994] son:

- **Wavelets de Haar**  
 Su función de escala es la función indicador  $\chi_{[0,1]}$ . Su correspondiente función wavelet es el *wavelet de Haar*  $\psi_{Haar}(u) = \chi_{[0,1/2]}(u) - \chi_{[1/2,1]}(u)$ . Su interés es básicamente teórico debido a la no continuidad de la función de escala.
- **Wavelet de Shannon**  
 Su función de escala es la función de muestreo de Shannon  $\text{sinc}(u)$  y su wavelet asociado es  $\psi_{Shannon}(u) = (\text{sen}(2\pi u) - \text{sen}(\pi u)) / (\pi u)$ . Tiene el inconveniente de que su soporte no es compacto.
- **Wavelet de Meyer**  
 Pertenecen a  $C^\infty$  y tienen decrecimiento mayor que cualquier polinomio. Su transformada de Fourier tiene soporte compacto. La función de escala y wavelet son simétricos alrededor de 0 y  $1/2$  respectivamente y el wavelet tiene un número infinito de momentos nulos.
- **Wavelets de Battle-Lemarié**  
 Se construyen mediante la ortogonalización de un b-spline y tienen decrecimiento exponencial. El wavelet con  $M$  momentos nulos es un polinomio a trozos de grado  $M-1$  que pertenece a  $C^{M-2}$ .
- **Wavelets de Daubechies**  
 Es la familia de wavelets ortogonales más ampliamente utilizada. Son una familia de wavelets indizada por  $M$ , donde  $M$  es el número de momentos nulos y se suele denotar por  $D_M$ . Su soporte lo constuye el intervalo  $[0, 2M-1]$ . Su principal desventaja es que salvo para el caso  $M=1$  (wavelet de Haar) no puede ser simétrica. Su regularidad se incrementa con  $N$  y es aproximadamente igual a  $0.2075M$  para  $M$  suficientemente grande.  
 Las condiciones anteriores (soporte compacto y número de momentos nulos) no determinan completamente la familia por lo que suelen añadirse más condiciones para definirla unívocamente. En esta tesis se utilizará la condición de mínima asimetría cuando se utilicen los wavelets de Daubechies

### 3.1.14.2 Wavelets biortogonales

Algunas familias de wavelets biortogonales [Jawerth 1994] son las siguientes:

- **Wavelets biortogonales basados en splines**  
 Fueron construidos por Cohen, Daubechies y Feaveau. Las funciones de escala son las funciones b-splines cardinales y los wavelets son también

splines. Todas las funciones tienen soporte compacto y son simétricas. Su principal desventaja es que para soportes pequeños tienen poca regularidad.

- Wavelets semiortogonales basados en splines  
Fueron construidos por Chui y Wang. Las funciones de escala son las funciones b-splines cardinales de orden  $m$  y los wavelets son splines con soporte  $[0, 2m-1]$ . Todas las funciones tienen una expresión analítica. Su principal desventaja es que las funciones duales no tienen soporte compacto.

### 3.1.14.3 Multiwavelets

Algunas familias de multiwavelets son [Strela 1996], [Tham 2000]:

- Multiwavelet GMH  
Fue construida por Geronimo, Hardin y Massopust. Las dos funciones de escala son simétricas y las funciones wavelet forman un par simétrico-antisimétrico. Son ortonormales y tienen orden 2 de aproximación.
- Multiwavelet CL  
Fue construida por Chui y Lian. Las dos funciones de escala forman un par simétrico (una función de escala es la reflexión de la otra).  
Las funciones wavelet forman un par simétrico-antisimétricas y ortonormales y tienen también orden dos de aproximación.
- Multiwavelet SA  
Fue construida por Shen, Tan y Tham, las funciones wavelet son simétricas o antisimétricas y tienen diferentes subfamilias con distintos órdenes de aproximación (desde orden 1 a orden 3) y soporte.

### 3.2 Análisis de contornos mediante wavelets

En esta sección se introduce la representación de contornos mediante wavelets y multiwavelets y se realiza una comparativa con las representaciones de contornos más importantes presentadas en la sección 2.1.2. Además se introduce la definición de normas para curvas (lo que permite compararlas) y se extienden las normas en espacios de Besov definidas en la sección 3.1.10 a los contornos. Finalmente se introducen los espacios de formas que permitirán estudiar la invarianza de estas normas frente a transformaciones geométricas.

#### 3.2.1 Descomposición de contornos mediante wavelets

De la misma forma que en el caso de la descomposición de funciones, la descomposición de contornos mediante wavelets está basado en la idea del análisis a múltiples resoluciones (AMR). Para el caso de los contornos, en vez de tener una función se parte de una representación paramétrica del contorno  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))^T$ . La noción del AMR (Figura 3. 3) es la misma salvo que en este caso el espacio sobre el que se realiza es  $L_2(\mathbb{I}) \times L_2(\mathbb{I})$ . El subespacio  $V_0$  está generado por traslaciones de un vector de funciones básicas que de forma vectorial se escribe como:  $\mathbf{F}(u) = (\mathbf{j}(u), \mathbf{j}(u))^T$ . Asociados con estas funciones de escala hay un vector de wavelets que se escribe como:  $\mathbf{?}(u) = (\mathbf{y}(u), \mathbf{y}(u))^T$

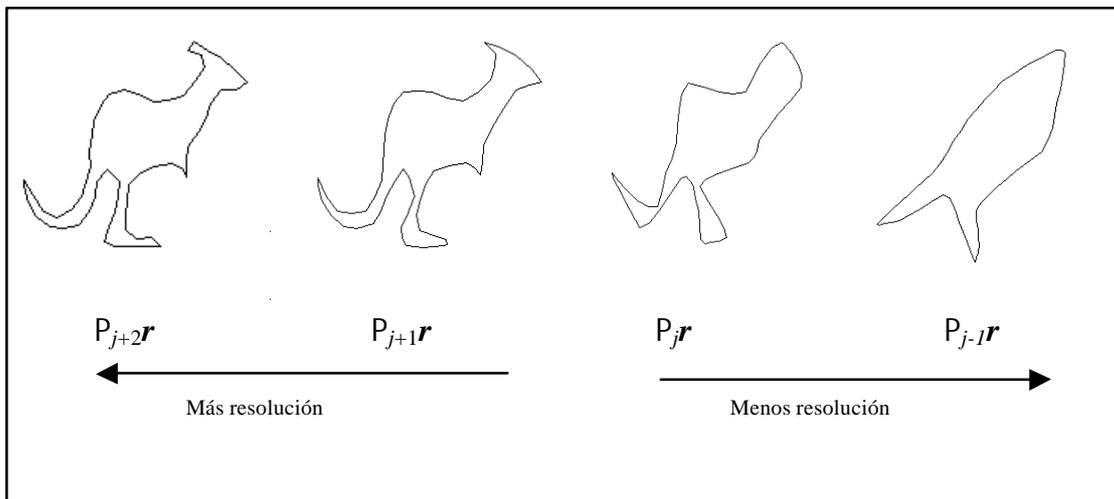


Figura 3. 3 Análisis a múltiples resoluciones y proyección sobre los distintos subespacios  $V_j$ .

De esta forma, las funciones paramétricas  $x$  e  $y$  del contorno  $\mathbf{C}$  pueden escribirse como:

$$x(u) = \sum_{j,l} d_{j,l,x} \mathbf{y}_{j,l}(u), \quad y(u) = \sum_{j,l} d_{j,l,y} \mathbf{y}_{j,l}(u) \quad \text{Ec. 3. 47}$$

y el contorno  $\mathbf{C}$  admite una descomposición:

$$\mathbf{r}(u) = \sum_{j,l} \mathbf{d}_{j,l} \mathbf{y}_{j,l}(u), \quad \mathbf{d}_{j,l} = (d_{j,l;x}, d_{j,l;y})^T \quad \text{Ec. 3. 48}$$

La forma más general de la descomposición utilizando también funciones de escala es:

$$\mathbf{r}(u) = \sum_l \mathbf{c}_{j_0,l} \mathbf{j}_{j_0,l}(u) + \sum_{j \geq j_0,l} \mathbf{d}_{j,l} \mathbf{y}_{j,l}(u), \quad \mathbf{c}_{j_0,l} = (c_{j_0,l;x}, c_{j_0,l;y})^T, \quad \mathbf{d}_{j,l} = (d_{j,l;x}, d_{j,l;y})^T \quad \text{Ec. 3. 49}$$

En el caso periódico (contorno cerrado) se tiene la siguiente descomposición:

$$\mathbf{r}(u) = \mathbf{c}_{0,0} \mathbf{j}_{0,0}(u) + \sum_{\substack{j \geq 0 \\ 0 \leq l < 2^j}} \mathbf{d}_{j,l} \mathbf{y}_{j,l}(u) \quad \text{Ec. 3. 50}$$

que en el caso de tener un contorno con  $N=2^J$  puntos muestreados toma la forma (Ec. 3. 30):

$$\mathbf{r}(u) = \mathbf{c}_{0,0} \mathbf{j}_{0,0}(u) + \sum_{\substack{j=0 \\ 0 \leq l < 2^j}}^{J-1} \mathbf{d}_{j,l} \mathbf{y}_{j,l}(u) \quad \text{Ec. 3. 51}$$

Las anteriores relaciones (Ec. 3. 51) se pueden expresar de forma concisa como:

$$x(u) = \mathbf{G}_w(u)^T \mathbf{w}^x, \quad y(u) = \mathbf{G}_w(u)^T \mathbf{w}^y, \quad \mathbf{r}(u) = \mathbf{F}_w(u) \mathbf{w}, \quad \text{con:} \quad \text{Ec. 3. 52}$$

$$\mathbf{G}_w(u) = \begin{pmatrix} \mathbf{j}_{0,0}(u) \\ \mathbf{y}_{0,0}(u) \\ \vdots \\ \mathbf{y}_{j,l}(u) \\ \vdots \\ \mathbf{y}_{J-1,2^{J-1}-1}(u) \end{pmatrix}, \quad \mathbf{w}^x = \begin{pmatrix} c_{0,0;x} \\ d_{0,0;x} \\ \vdots \\ d_{j,l;x} \\ \vdots \\ d_{J-1,2^{J-1}-1;x} \end{pmatrix}, \quad \mathbf{w}^y = \begin{pmatrix} c_{0,0;y} \\ d_{0,0;y} \\ \vdots \\ d_{j,l;y} \\ \vdots \\ d_{J-1,2^{J-1}-1;y} \end{pmatrix}$$

$$\mathbf{F}_w(u) = \mathbf{I}_2 \otimes \mathbf{G}_w(u)^T = \begin{pmatrix} \mathbf{G}_w(u)^T & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w(u)^T \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} \mathbf{w}^x \\ \mathbf{w}^y \end{pmatrix}$$

De forma análoga al caso de una única función, el cálculo de los coeficientes del vector  $\mathbf{w}$  tiene complejidad lineal.

### 3.2.2 Descomposición de contornos mediante multiwavelets

En [Pérez 1999a] se introduce la descomposición de contornos mediante multiwavelets siendo su desarrollo similar a la descomposición de contornos mediante wavelets. Se parte por tanto de un análisis multiresolución de  $L_2(\mathbb{I}) \times L_2(\mathbb{I})$ . En este caso  $V_0$  está

generado por traslaciones de  $2r$  funciones básicas que de forma vectorial se escriben como:  $F(u) = (\mathbf{j}^1(u), \dots, \mathbf{j}^r(u), \mathbf{j}^1(u), \dots, \mathbf{j}^r(u))^T$ . Asociadas con estas funciones de escala hay  $2r$  wavelets básicas que de forma vectorial se escriben como:  $\mathbf{?}(u) = (\mathbf{y}^1(u), \dots, \mathbf{y}^r(u), \mathbf{y}^1(u), \dots, \mathbf{y}^r(u))^T$

De esta forma, las funciones paramétricas  $x$  e  $y$  pueden escribirse como:

$$x(u) = \sum_{j,l} \mathbf{D}_{j,l;x}^T \mathbf{?}_{j,l}(u), \quad y(u) = \sum_{j,l} \mathbf{D}_{j,l;y}^T \mathbf{?}_{j,l}(u) \quad \text{Ec. 3. 53}$$

y el contorno  $\mathbf{C}$  admite una descomposición:

$$\mathbf{r}(u) = \sum_{j,l} \mathbf{D}_{j,l} \mathbf{?}_{j,l}(u), \quad \mathbf{D}_{j,l} = \begin{pmatrix} \mathbf{D}_{j,l;x}^T \\ \mathbf{D}_{j,l;y}^T \end{pmatrix} \quad \text{Ec. 3. 54}$$

La forma más general de la descomposición utilizando también funciones de escala es:

$$\mathbf{r}(u) = \sum_l \mathbf{C}_{j_0,l} \mathbf{F}_{j_0,l}(u) + \sum_{j \geq j_0,l} \mathbf{D}_{j,l} \mathbf{?}_{j,l}(u) \quad \text{Ec. 3. 55}$$

En el caso periódico (contorno cerrado) se utilizará la siguiente descomposición análoga a la de (Ec. 3. 51) para el caso de los wavelets escalares:

$$\mathbf{r}(u) = \mathbf{C}_{0,0} \mathbf{F}_{0,0}(u) + \sum_{\substack{j=0 \\ 0 \leq l < 2^j}}^{J-1} \mathbf{D}_{j,l} \mathbf{?}_{j,l}(u) \quad \text{Ec. 3. 56}$$

De forma análoga al caso escalar, la anterior relación (Ec. 3. 53) se pueden expresar de forma concisa como:

$$x(u) = \mathbf{G}_w(u)^T \mathbf{W}^x, \quad y(u) = \mathbf{G}_w(u)^T \mathbf{W}^y, \quad \mathbf{r}(u) = \mathbf{F}_w(u) \mathbf{W}, \quad \text{con:} \quad \text{Ec. 3. 57}$$

$$\mathbf{G}_w(u) = \begin{pmatrix} \mathbf{F}_{0,0}(u) \\ \mathbf{?}_{1,0}(u) \\ \vdots \\ \mathbf{?}_{j,l}(u) \\ \vdots \\ \mathbf{?}_{J-1,2^{J-1}-1}(u) \end{pmatrix}, \quad \mathbf{W}^x = \begin{pmatrix} \mathbf{C}_{0,0;x} \\ \mathbf{D}_{1,0;x} \\ \vdots \\ \mathbf{D}_{j,l;x} \\ \vdots \\ \mathbf{D}_{J-1,2^{J-1}-1;x} \end{pmatrix}, \quad \mathbf{W}^y = \begin{pmatrix} \mathbf{C}_{0,0;y} \\ \mathbf{D}_{1,0;y} \\ \vdots \\ \mathbf{D}_{j,l;y} \\ \vdots \\ \mathbf{D}_{J-1,2^{J-1}-1;y} \end{pmatrix}$$

$$\mathbf{F}_w(u) = \mathbf{I}_2 \otimes \mathbf{G}_w(u)^T = \begin{pmatrix} \mathbf{G}_w(u)^T & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w(u)^T \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} \mathbf{W}^x \\ \mathbf{W}^y \end{pmatrix}$$

De forma análoga al caso de una única función, el cálculo de los coeficientes de la matriz  $\mathbf{W}$  tiene complejidad lineal.

### 3.2.3 Una comparativa entre distintas representaciones

Las propiedades de los wavelets las hacen un candidato atractivo para la representación de curvas. A la hora de establecer una comparativa entre las distintas representaciones pueden establecerse varios criterios: costo computacional, error de representación o propiedades de los coeficientes. En cuanto al costo computacional, la representación basada en wavelets es óptima puesto que su costo es lineal. Por tanto pasaremos a continuación a realizar una comparativa entre los distintos errores de aproximación. Un inconveniente a la hora de realizar este estudio es el hecho de que no hay ninguna representación que sea siempre mejor que otra. Esto es debido a que para cada representación puede construirse un contorno a medida de forma que el error de aproximación sea nulo para dicha representación y no nulo para el resto. Otra aproximación consiste en determinar un conjunto de contornos (por ejemplo el conjunto de contornos provenientes de imágenes) y evaluar el error medio sobre dicho conjunto. Esta aproximación tiene el inconveniente de que no existe ningún modelo para dicho conjunto de contornos por lo que un estudio analítico no puede realizarse. La opción que se ha tomado en esta tesis consiste en la selección de un conjunto de contornos provenientes de imágenes *naturales* y llevar a cabo un estudio del error de aproximación sobre dicho conjunto [Pérez 2000b].

Como conjunto de imágenes se ha tomado un conjunto de siluetas que determinan un conjunto de contornos en el plano (Figura 3. 4). Para cada silueta se ha tomado únicamente el contorno exterior de la figura. Todos los contornos se han remuestreado para obtener un total de 256 puntos por contorno. Puesto que en las aplicaciones en las que se usará esta representación es de gran importancia el costo computacional, no se ha realizado ningún procesamiento sobre el conjunto de puntos del contorno.

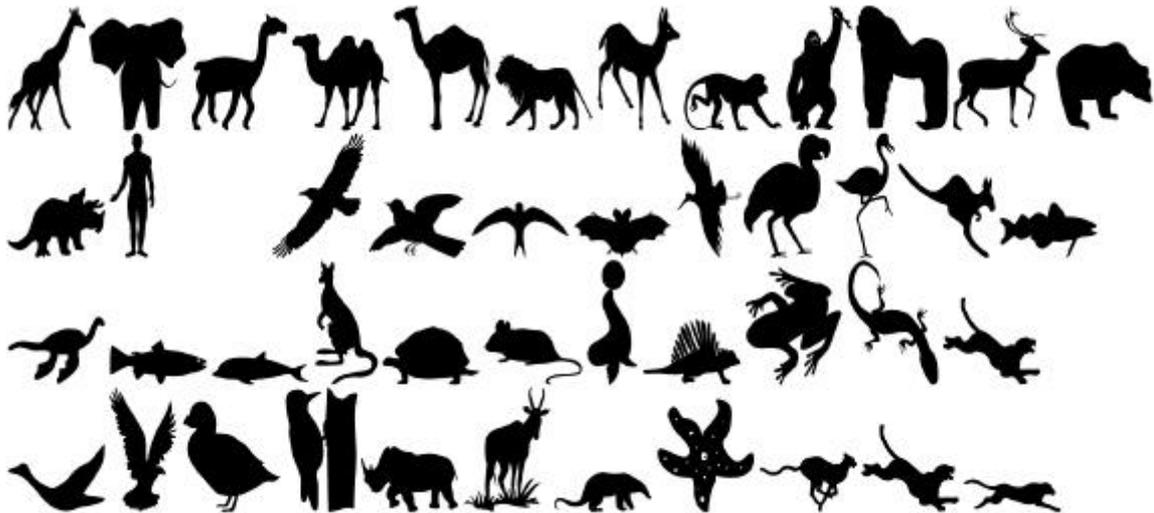


Figura 3. 4 Conjunto de imágenes de prueba

Dado este conjunto de imágenes se calculó el error medio entre la curva reconstruida a partir del conjunto de coeficientes y la curva original para diferentes familias de wavelets ortogonales: Haar,  $D_4$  y  $D_8$  de Daubechies (mínima asimetría), distintas

familias de multiwavelets: GMH de Gerónimo-Hardin y Massopoust, CL de Chu y Lian y SA de Shen-Tan y Tham (Ver Sección 3.1.14). Para realizar una comparación con las técnicas tradicionales se incluyen también los resultados para la transformada elíptica de Fourier y los b-splines (Ver Sección 2.1.2). Puesto que todas las transformaciones (salvo los b-splines) son ortogonales se han tomado los  $m$  coeficientes más importantes (coeficientes de máxima norma) con  $m=32, 64, 128$ . Para los b-splines el contorno se ha remuestreado para los distintos valores de  $m$  y después se han calculado los descriptores. Los resultados obtenidos del error cuadrático medio (en pixels) para el conjunto de imágenes han sido:

	Fourier	b-splines	Haar	D4	La8	GMH	CL	SA(4)
n= 32	4.0	3.9	6.1	2.8	2.4	2.5	2.7	2.3
n= 64	1.8	2.1	3.3	1.2	1.0	1.1	1.1	1.0
n=128	0.8	0.9	1.7	0.6	0.4	0.5	0.5	0.5

Tabla 3. 1 Errores cuadráticos medios para el conjunto de imágenes

Los resultados numéricos muestran que en general la mejor compresión se obtiene con el multiwavelet SA(4) y el wavelet  $D_8$ . Después de estos, los mejores resultados se obtiene con los multiwavelets GMH y CL, el wavelet  $D_4$ , el descriptor elíptico de Fourier y el wavelet de Haar. En la Figura 3. 5 se muestra el error cuadrático medio ( $m=32$ ) para los descriptores elípticos de Fourier (línea gruesa), el wavelet  $D_8$  y el multiwavelet SA(4). En esta figura podemos observar como el multiwavelet SA(4) obtiene resultados ligeramente superiores que el wavelet  $D_8$ . El peor contorno tanto para el  $D_8$  y SA(4) se muestra en la Figura 3. 5 donde el contorno original (línea gruesa) se muestra con la aproximación mediante el wavelet  $D_8$  (línea discontinua) y el multiwavelet SA(4) (línea fina)

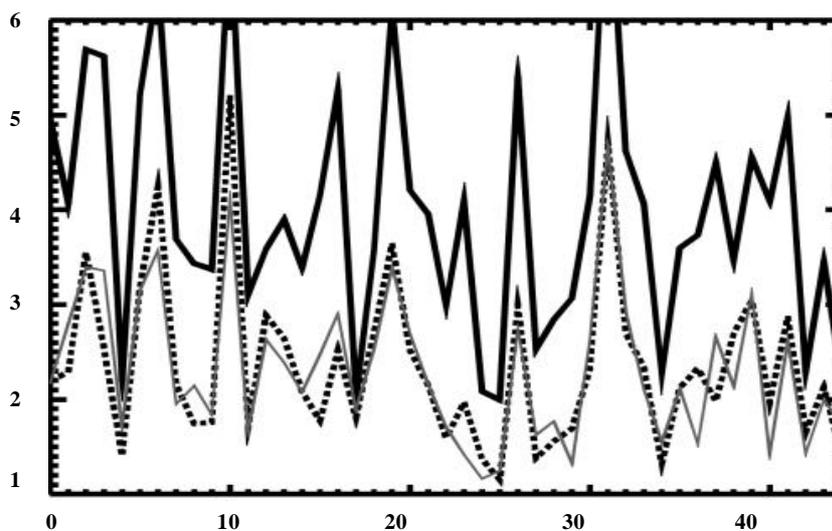


Figura 3. 5 Comparación del error cuadrático medio para todas las imágenes

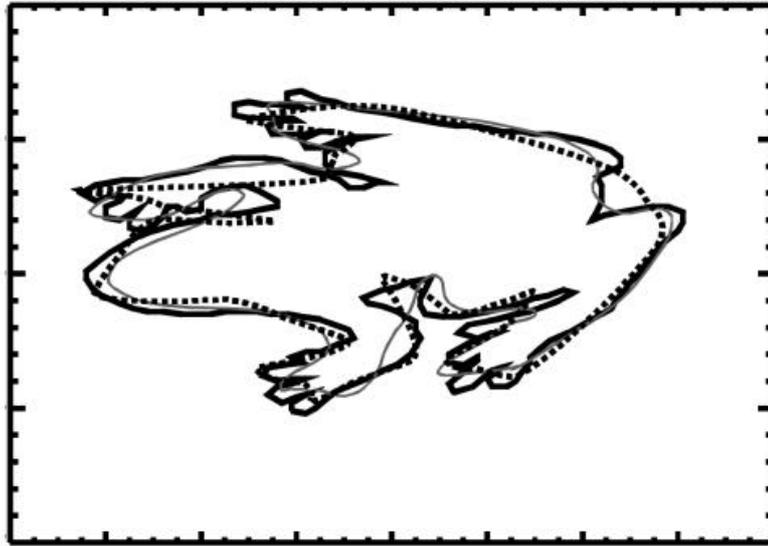


Figura 3. 6 Contorno de error máximo

En el procesamiento de imágenes no se busca únicamente una descripción con un error cuadrático medio pequeño. Generalmente es útil obtener una pequeña distancia máxima entre el contorno original y su aproximación. Para comprobar la capacidad de la descripción basada en wavelets para obtener un bajo error de este tipo, se tomó el percentil 95 de la distancia máxima para el wavelet  $D_8$  y multiwavelet SA(4) para todas las imágenes de la base de datos obteniéndose los siguientes resultados (ver Tabla 3. 2)

	n= 32	n =64	n =128
LA8 wavelet	4.14	1.72	0.76
SA(4) multiwavelet	4.09	1.80	0.79

Tabla 3. 2 El percentil 95 del error máximo

Los resultados muestran que el error en ambas descripciones es similar. En este trabajo se utilizarán ambos tipos de wavelets. El de Daubechies en las secciones 4.1 y el SA(4) en la 4.2

### 3.2.4 Normas para curvas

Una vez expuesta la representación de curvas utilizando wavelets (Sección 3.2.1), el siguiente paso es determinar la forma de evaluar normas y productos escalares para curvas. La norma para curvas es especialmente interesante para compararlas, pues para ello se suele utilizar la norma de su diferencia. La forma usual de calcular la norma de una curva en  $L_2(\mathbb{R}) \times L_2(\mathbb{R})$  es la siguiente:

$$\|\mathbf{r}\|_2^2 = \int |\mathbf{r}(u)|^2 du = \int x(u)^2 + y(u)^2 du \quad \text{Ec. 3. 58}$$

o de forma equivalente:

$$\|\mathbf{r}\|_2^2 = \mathbf{w}^T \mathbf{F}_N \mathbf{w} \quad , \quad \mathbf{F}_N = \int \mathbf{F}_w(u)^T \mathbf{F}_w(u) du \quad \text{Ec. 3. 59}$$

donde  $w$  y  $F$  se definen como en (Ec. 3. 52) esta norma proporciona el producto escalar:

$$\langle \mathbf{w}_1, \mathbf{w}_2 \rangle = \mathbf{w}_1^T \mathbf{F}_N \mathbf{w}_2 \quad , \quad \text{Ec. 3. 60}$$

la utilización de funciones ortogonales, resulta especialmente interesante pues la norma se simplifica a:

$$\|\mathbf{r}\|_2^2 = \mathbf{w}^T \mathbf{w} = \|\mathbf{w}\|_2^2 \quad \text{Ec. 3. 61}$$

La norma descrita anteriormente está definida sobre  $L_2(\cdot) \times L_2(\cdot)$  y es la norma para curvas generalmente utilizada para comparar curvas.

### 3.2.5 Normas para curvas en espacios de suavidad

Aprovechando la relación existente en el caso unidimensional entre la descomposición en wavelets y los espacios de suavidad se proponen en esta tesis nuevas normas sobre curvas en espacios de suavidad distintos de  $L_2(\cdot) \times L_2(\cdot)$ . Al igual que en la sección 3.1.10 se considerará la familia de espacios de Besov  $B_q^a(L_p(I))$  donde  $I$  será en general un dominio finito (el intervalo  $[0,1]$ ), y donde  $0 < a < \infty$ ,  $0 < p \leq \infty$ ,  $0 < q \leq \infty$ . Estos espacios se generalizarán para curvas de la siguiente forma:

Dada una curva  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))^T$  se considera para  $m > 0$  y  $h > 0$  la  $m$ -ésima diferencia de  $f$  como:

$$\Delta_h^{(m)} \mathbf{r}(u) = \sum_{k=0}^m \binom{m}{k} (-1)^k \mathbf{r}(u + kh), \quad u \in I_h^r = \{u \in I / u + rh \in I\} \quad \text{Ec. 3. 62}$$

se define entonces:

$$\|\mathbf{r}\|_{L_p(I)} = \begin{cases} \left( \int_I \|\mathbf{r}\|^p du \right)^{1/p} = \left( \int_I \sqrt{x(u)^2 + y(u)^2}^p du \right)^{1/p} & p < \infty \\ \sup_{u \in I} \|\mathbf{r}\| = \sup_{u \in I} \sqrt{x(u)^2 + y(u)^2} & p = \infty \end{cases} \quad \text{Ec. 3. 63}$$

el módulo de suavidad de  $L_p(I)$  para  $0 < p \leq \infty$  se define como:

$$\mathbf{w}_m(\mathbf{r}, u)_p = \sup_{|h| \leq u} \|\Delta_h^{(m)} \mathbf{r}\|_{L_p(I'_h)}, \quad \text{Ec. 3. 64}$$

La seminorma de Besov de índice  $(\mathbf{a}, p, q)$  se define para  $m > \mathbf{a}$ , donde  $1 \leq p, q \leq \infty$  por:

$$\begin{aligned} |\mathbf{r}|_{B_q^{\mathbf{a}}(L_p(I))} &= \left( \int_0^\infty \left( \frac{\mathbf{w}_m(\mathbf{r}, u)_p}{t^{\mathbf{a}}} \right)^q \frac{dt}{t} \right)^{1/q}, \quad 0 \leq q \leq \infty \\ |\mathbf{r}|_{B_q^{\mathbf{a}}(L_p(I))} &= \sup_{0 < t < \infty} \left( \frac{\mathbf{w}_m(\mathbf{r}, u)_p}{t^{\mathbf{a}}} \right), \quad q = \infty \end{aligned} \quad \text{Ec. 3. 65}$$

La norma para el espacio de Besov se define entonces como:

$$\|\mathbf{r}\|_{B_q^{\mathbf{a}}(L_p(I))} = \|\mathbf{r}\|_{L_p(I)} + |\mathbf{r}|_{B_q^{\mathbf{a}}(L_p(I))} \quad \text{Ec. 3. 66}$$

El espacio de Besov de curvas  $B_q^{\mathbf{a}}(L_p(I))$  es entonces el conjunto de curvas que pertenecen a  $L_p(I) \times L_p(I)$  y  $|\mathbf{r}|_{B_q^{\mathbf{a}}(L_p(I))} < \infty$ .

A continuación se presentan algunas propiedades de esta norma:

### Proposición 3.2

Dada una curva  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))^T$ , las siguientes normas son equivalentes:

$$\|\mathbf{r}\|_{B_q^{\mathbf{a}}(L_p(I))} \equiv \|x\|_{B_q^{\mathbf{a}}(L_p(I))} + \|y\|_{B_q^{\mathbf{a}}(L_p(I))} \quad \text{Ec. 3. 67}$$

### Demostración

La demostración puede encontrarse en el Apéndice A

### Corolario 3. 1

Dada una curva  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))^T$  entonces  $\mathbf{r} \in B_q^{\mathbf{a}}(L_p(I))$  si y solo si sus funciones coordenadas  $x$  e  $y$  están en dicho espacio.

### Demostración

Evidente aplicando el teorema anterior

### Teorema 3. 3

Dada una curva  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))^T$  y su expansión en series de wavelets (Ec. 3. 49) entonces:

$$\|\mathbf{r}\|_{B_q^a(L_p(I))} \equiv \sqrt{|c_{0,0;x}|^2 + |c_{0,0;y}|^2} + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p/2-1)} \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right)^p \right)^{q/p} \right)^{1/q}$$

**Demostración**

La demostración puede encontrarse en el Apéndice A

La importancia del teorema anterior es que nos permite relacionar la norma de la curva  $\mathbf{r}(u)$  con su descomposición mediante wavelets.

**3.2.6 Modelos de espacio de formas**

En la práctica es deseable distinguir entre el vector de coeficientes  $\mathbf{w} \in \mathbf{Q}$  que describe la forma de un objeto y el vector de forma  $\mathbf{x} \in \mathbf{S}$ . Donde  $\mathbf{S}$  es un espacio de formas. Mientras que  $\mathbf{Q}$  tiene dimensión  $n_Q = 2N$  donde  $N$  es el número de funciones en la descomposición mediante wavelets,  $\mathbf{S}$  se construye a partir de un espacio vectorial de dimensión  $n_S$  que suele ser considerablemente menor que  $n_Q$ . La razón de la introducción de este espacio de formas es la reducción de grados de libertad en las posibles deformaciones del objeto, siendo de esta forma más fácil centrarse en aquellas de interés como por ejemplo, traslaciones, similaridades euclídeas o transformaciones afines.

De forma precisa se define un espacio de formas, a través de un espacio de transformaciones entre  $\mathbf{S}$  y  $\mathbf{Q}$  de manera que  $\mathbf{w} = \mathbf{H}(\mathbf{x}, \bar{\mathbf{w}})$ ,  $\mathbf{x} \in \mathbf{S}$ . Cuando el espacio de transformaciones es lineal puede escribirse como  $L(\mathbf{H}, \bar{\mathbf{w}})$  que es el conjunto de transformaciones lineales entre el vector del espacio de formas  $\mathbf{x} \in \square^{n_S}$  y el vector de parámetros  $\mathbf{w} \in \square^{n_Q}$  definidas mediante una “matriz de forma”  $\mathbf{H}$  es orden  $n_Q \times n_S$ :

$$\mathbf{w} = \mathbf{H}\mathbf{x} + \bar{\mathbf{w}} \tag{Ec. 3. 68}$$

El vector  $\bar{\mathbf{w}}$  representa una curva base que es transformada de forma lineal. A continuación se discuten los tres espacios de formas mas simples: el espacio de traslaciones, el espacio de similaridades euclídeas y el espacio de transformaciones afines.

**3.2.6.1 El espacio de traslaciones**

El espacio de traslaciones de una curva base  $\bar{\mathbf{r}}$  con vector de parámetros  $\bar{\mathbf{w}}$  constituye un espacio de formas 2-dimensional  $\mathbf{S}$  con matriz de forma:

$$H = \begin{pmatrix} T & \mathbf{0} \\ \mathbf{0} & T \end{pmatrix} \quad \text{Ec. 3. 69}$$

$$T = (1,0,\dots,0), \quad \mathbf{0} = (0,0,\dots,0)$$

$T$  y  $\mathbf{0}$  son vectores de tamaño  $N$  igual al número de puntos de la curva

Sin pérdida de generalidad se asume que  $\bar{w}$  tiene su centroide en el origen. De esta forma el vector  $\mathbf{x}=(a,b)$ , desplaza la curva determinada por  $\bar{w}$   $a$  unidades en el eje  $x$  y  $b$  unidades en el eje  $y$ .

### 3.2.6.2 El espacio de similaridades euclídeas

El espacio de similaridades euclídeas de una curva base  $\bar{r}$  con vector de parámetros  $\bar{w}$  constituye un espacio de formas 4-dimensional  $\mathbf{S}$  con matriz de forma:

$$H = \begin{pmatrix} T & \mathbf{0} & \bar{w}^x & -\bar{w}^y \\ \mathbf{0} & T & \bar{w}^y & \bar{w}^x \end{pmatrix} \quad \text{Ec. 3. 70}$$

$$T = (1,0,\dots,0), \quad \mathbf{0} = (0,0,\dots,0)$$

Sin pérdida de generalidad se asume que  $\bar{w}$  tiene su centroide en el origen de forma que las dos primeras columnas están asociadas con la traslación de la curva y las dos últimas con rotación y escalado.

De esta forma el vector  $\mathbf{x}=(a,b,k(\cos(q)-1), k(\sin(q)))$ , desplaza la curva determinada por  $\bar{w}$   $a$  unidades en el eje  $x$  y  $b$  unidades en el eje  $y$ , rotándose también la curva con ángulo  $q$  y escalándose con factor de magnificación  $k$ .

### 3.2.6.3 El espacio de transformaciones afines

El espacio de transformaciones euclídeas de una curva base  $\bar{r}$  con vector de parámetros  $\bar{w}$  constituye un espacio de formas 6-dimensional y puede interpretarse como el conjunto de transformaciones lineales que pueden aplicarse a  $\bar{r}$  :

$$r = t + M\bar{r} \quad \text{Ec. 3. 71}$$

donde  $t$  es un vector de traslación bidimensional y  $M=(m_{ij})$  es una matriz cuadrada de orden 2.

El espacio generado  $\mathbf{S}$  tiene como matriz de forma:

$$H = \begin{pmatrix} T & \mathbf{0} & \bar{w}^x & \mathbf{0} & \mathbf{0} & \bar{w}^y \\ \mathbf{0} & T & \mathbf{0} & \bar{w}^y & \bar{w}^x & \mathbf{0} \end{pmatrix} \quad \text{Ec. 3. 72}$$

$$T = (1,0,\dots,0), \quad \mathbf{0} = (0,0,\dots,0)$$

Sin pérdida de generalidad se asume de nuevo que  $\overline{\mathbf{W}}$  tiene su centroide en el origen de forma que las dos primeras columnas están asociadas con la traslación de la curva y las cuatro últimas con la transformación afín.

De esta forma el vector  $\mathbf{x}=(a,b,m_{11}-1,m_{22}-1,m_{21},m_{12})$ , desplaza la curva determinada por  $\overline{\mathbf{W}}$   $a$  unidades en el eje  $x$  y  $b$  unidades en el eje  $y$ , transformándose también la curva con la matriz  $\mathbf{M}=(m_{ij})$ .

### 3.2.6.4 Normas en el espacio de formas

Dado que siempre que sea posible se trabaja en el espacio de formas  $\mathbf{S}$  es necesario que las normas definidas sean consistente con el espacio paramétrico  $\mathbf{Q}$ . Esto se consigue definiendo:

$$\|\mathbf{x}\|_{\mathbf{S}} = \|\mathbf{H}\mathbf{x}\|_{\mathbf{Q}} \quad \text{Ec. 3. 73}$$

esto es, la norma sobre  $\mathbf{S}$  se induce a partir de la norma sobre  $\mathbf{Q}$ .

En el caso de trabajar con la norma sobre  $L_2(\cdot)$  se simplifica a:

$$\|\mathbf{x}\|_{\mathbf{S}}^2 = \mathbf{x}^T \mathbf{S}_N \mathbf{x}, \quad \mathbf{S}_N = \mathbf{H}^T \mathbf{F}_N \mathbf{H} \quad \text{Ec. 3. 74}$$

De nuevo en el caso ortogonal la norma se simplifica al ser  $\mathbf{F}_N$  (Ec. 3. 59) la identidad.

## 3.3 Modelos de deformación de contornos mediante wavelets

En esta sección se estudia el proceso de regularización mediante wavelets utilizando los modelos de deformación definidos anteriormente. Además se extienden los modelos probabilísticos de los wavelets para generar un modelo probabilístico de deformación a priori de los contornos y se introducen técnicas para el emparejamiento de dichos modelos. Estas técnicas consisten en un algoritmo aproximado de emparejamiento en espacios de Besov y un algoritmo exacto para espacios de Sobolev.

### 3.3.1 Deformación de contornos mediante wavelets

Los modelos de deformación permiten determinar la curva  $\mathbf{C} \equiv \mathbf{r}(u)$  que mejor se ajusta a un conjunto de datos  $\mathbf{D}$  de la imagen que en este caso tiene la forma de una curva  $\mathbf{r}_D$ . Se definirá un criterio de ajuste de la curva  $\mathbf{r}$  a la imagen:

$$E_{aju} = E_{aju}(\mathbf{r}, \mathbf{r}_D) \quad \text{Ec. 3. 75}$$

que se utiliza junto con la función de deformación a una curva de referencia  $\overline{\mathbf{r}}$  :

$$E_{def} = E_{def}(\mathbf{r}, \bar{\mathbf{r}}) \quad \text{Ec. 3. 76}$$

para determinar una función de ajuste combinado que, utilice ambos criterios y que tomará la forma:

$$E_{tot}(\mathbf{r}, \mathbf{t}, \mathbf{r}_D, \bar{\mathbf{r}}) = \mathbf{t} E_{def}(\mathbf{r}, \bar{\mathbf{r}}) + E_{aju}(\mathbf{r}, \mathbf{r}_D) \quad \text{Ec. 3. 77}$$

donde  $\mathbf{t}$  es un parámetro de regularización.

Para definir la función de deformación se utilizará la extensión de la norma de Besov a curvas propuesta en la sección 3.2.5. De esta forma además de penalizar la diferencia con respecto a la curva de referencia  $\bar{\mathbf{r}}$  se hace especialmente costoso las deformaciones no suaves con respecto de la curva de referencia. Así se define una función de energía  $E_{def}(\mathbf{r}, \bar{\mathbf{r}})$  que de forma explícita puede escribirse en función de los coeficientes de la expansión de los contornos en wavelets como se muestra a continuación.

Dada una curva para la que se calculan los coeficientes de la expansión en wavelets como se muestra en Ec. 3. 52:

$$x(u) = \mathbf{G}_w(u)^T \mathbf{w}^x, \quad y(u) = \mathbf{G}_w(u)^T \mathbf{w}^y, \quad \mathbf{r}(u) = \mathbf{F}_w(u) \mathbf{w}, \quad \text{con:}$$

$$\mathbf{G}_w(u) = \begin{pmatrix} \mathbf{j}_{0,0}(u) \\ \mathbf{y}_{01,0}(u) \\ \vdots \\ \mathbf{y}_{j,j}(u) \\ \vdots \\ \mathbf{y}_{J-1,2^{J-1}-1}(u) \end{pmatrix}, \quad \mathbf{w}^x = \begin{pmatrix} c_{0,0;x} \\ d_{0,0;x} \\ \vdots \\ d_{j,j;x} \\ \vdots \\ d_{J-1,2^{J-1}-1;x} \end{pmatrix}, \quad \mathbf{w}^y = \begin{pmatrix} c_{0,0;y} \\ d_{0,0;y} \\ \vdots \\ d_{j,j;y} \\ \vdots \\ d_{J-1,2^{J-1}-1;y} \end{pmatrix}$$

$$\mathbf{F}_w(u) = \mathbf{I}_2 \otimes \mathbf{G}_w(u)^T = \begin{pmatrix} \mathbf{G}_w(u)^T & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w(u)^T \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} \mathbf{w}^x \\ \mathbf{w}^y \end{pmatrix}$$

entonces si  $\mathbf{r}(u) = \mathbf{F}_w(u) \mathbf{w}$ ,  $\bar{\mathbf{r}}(u) = \mathbf{F}_w(u) \bar{\mathbf{w}}$ , y

$$\Delta \mathbf{w} = \mathbf{w} - \bar{\mathbf{w}} = \begin{pmatrix} \Delta c_{0,0;x} \\ \Delta d_{0,0;x} \\ \vdots \\ \Delta d_{j,j;x} \\ \vdots \\ \Delta d_{J-1,2^{J-1}-1;x} \\ \Delta c_{0,0;y} \\ \Delta d_{0,0;y} \\ \vdots \\ \Delta d_{j,j;y} \\ \vdots \\ \Delta d_{J-1,2^{J-1}-1;y} \end{pmatrix} = \begin{pmatrix} c_{0,0;x} \\ d_{0,0;x} \\ \vdots \\ d_{j,j;x} \\ \vdots \\ d_{J-1,2^{J-1}-1;x} \\ c_{0,0;y} \\ d_{0,0;y} \\ \vdots \\ d_{j,j;y} \\ \vdots \\ d_{J-1,2^{J-1}-1;y} \end{pmatrix} - \begin{pmatrix} \bar{c}_{0,0;x} \\ \bar{d}_{0,0;x} \\ \vdots \\ \bar{d}_{j,j;x} \\ \vdots \\ \bar{d}_{J-1,2^{J-1}-1;x} \\ \bar{c}_{0,0;y} \\ \bar{d}_{0,0;y} \\ \vdots \\ \bar{d}_{j,j;y} \\ \vdots \\ \bar{d}_{J-1,2^{J-1}-1;y} \end{pmatrix}$$

se define la función de deformación como:

$$E_{def}(\mathbf{r}, \bar{\mathbf{r}}) \equiv E_{def}(W, \bar{W}) = \tag{Ec. 3. 78}$$

$$|\Delta c_{0,0;x}|^2 + |\Delta c_{0,0;y}|^2 + \sum_{j=0}^{J-1} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p/2-1)} \left( \sqrt{|\Delta d_{j,k;x}|^2 + |\Delta d_{j,k;y}|^2} \right)^p \right)^{q/p}$$

de esta forma, el término de deformación hace que la curva  $\mathbf{r}$  tienda hacia la curva de referencia  $\bar{\mathbf{r}}$ .

La definición de la función de deformación en Ec. 3. 78 no es satisfactoria pues no es invariante frente a transformaciones de la curva de referencia  $\bar{\mathbf{r}}$ . En general, se desea que  $\bar{\mathbf{r}}$  influya solamente en la forma de  $\mathbf{r}$ , pero no en su posición u orientación que deben extraerse de la imagen. Una forma de conseguir esta invarianza frente a un conjunto T de transformaciones es definir:

$$E_{def}^T(\mathbf{r}, \bar{\mathbf{r}}) = \min_{T \in T} E_{def}(\mathbf{r}, T(\bar{\mathbf{r}})) \tag{Ec. 3. 79}$$

de esta forma si por ejemplo T es el conjunto de traslaciones se tiene:

$$E_{def}^T(\mathbf{r}, \bar{\mathbf{r}}) = \sum_{j=0}^{J-1} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p/2-1)} \left( \sqrt{|\Delta d_{j,k;x}|^2 + |\Delta d_{j,k;y}|^2} \right)^p \right)^{q/p} \tag{Ec. 3. 80}$$

### 3.3.2 Ajuste de contornos mediante wavelets

Para establecer el modelo completo de ajuste mediante wavelets es necesario determinar la función de ajuste  $E_{aju}$  ésta se define mediante:

$$E_{aju}(\mathbf{r}, \bar{\mathbf{r}}) = \|\mathbf{r} - \mathbf{r}_D\|_2^2 \tag{Ec. 3. 81}$$

con lo cual el problema de ajuste queda definido como:

$$\min_{\mathbf{r}} t E_{def}^T(\mathbf{r}, \bar{\mathbf{r}}) + E_{aju}(\mathbf{r}, \mathbf{r}_D) \tag{Ec. 3. 82}$$

la solución al problema en Ec. 3. 82 para un espacio de Besov general es muy compleja (aunque se presentará una solución general aproximada en la sección 3.3.3.3), sin embargo existen espacios de suavidad de interés para los que es posible encontrar una solución analítica. Más concretamente en esta tesis se trabajará fundamentalmente en el espacio de Sobolev  $W^a(L_2(I))$  es decir, el espacio de Besov  $B_2^a(L_2(I))$ . En este espacio se tiene:

$$E_{def}(\mathbf{r}, \bar{\mathbf{r}}) \equiv E_{def}(W, \bar{W}) = \tag{Ec. 3. 83}$$

$$\Delta c_{0,0;x}^2 + \Delta c_{0,0;y}^2 + \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} 2^{2aj} (\Delta d_{j,k;x}^2 + \Delta d_{j,k;y}^2)$$

que de forma matricial puede expresarse como:

$E_{def}(W, \bar{W}) = (W - \bar{W})^T S^{-1} (W - \bar{W})^T$ , donde  $S$  se define mediante la siguiente matriz diagonal:

$$S = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 2^{-2a_0} & & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & & \vdots \\ 0 & 0 & \dots & 2^{-2a(J-1)} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & & 0 & 0 & 2^{-2a_0} & & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 2^{-2aJ(J-1)} \end{pmatrix} \quad \text{Ec. 3. 84}$$

siendo  $J = \log_2 N$  donde  $N$  es el número de puntos de la curva discretizada.

De nuevo, esta definición de la función de deformación no es invariante frente a transformaciones de la curva de referencia determinada por el vector de parámetros  $\bar{W}$ . Es por tanto necesario calcular la función de deformación  $E_{def}^T(\mathbf{r}, \bar{\mathbf{r}})$  invariante para un determinado conjunto de transformaciones. Para las transformaciones de la sección 3.2.6.3 (transformaciones afines) el conjunto de transformaciones puede escribirse como:  $T(\bar{\mathbf{r}}) = \mathbf{M}\bar{\mathbf{r}} + \mathbf{t}$  donde  $\mathbf{M}$  es una matriz de orden  $2 \times 2$  y  $\mathbf{t}$  es un vector de orden  $2 \times 1$ . Desarrollaremos a continuación este caso (el resto es similar). La función de deformación se expresa como:

$$E_{def}^T(\mathbf{r}, \bar{\mathbf{r}}) = \min_{T \in \mathcal{T}} E_{def}(\mathbf{r}, T(\bar{\mathbf{r}})) = \min_{\mathbf{M}, \mathbf{t}} E_{def}(\mathbf{r}, \mathbf{M}\bar{\mathbf{r}} + \mathbf{t}) \quad \text{Ec. 3. 85}$$

para calcular los valores de  $\mathbf{M}$  y  $\mathbf{t}$  óptimos para la transformación podemos resolver el problema de forma equivalente (utilizando los modelos de espacios de formas de la sección 3.2.6 como:

$$\begin{aligned} \min_{\mathbf{M}, \mathbf{t}} E_{def}(\mathbf{r}, \mathbf{M}\bar{\mathbf{r}} + \mathbf{t}) &\equiv \min_{\mathbf{X}} E_{def}(W, \mathbf{H}\mathbf{X} + \bar{W}) = \\ &\min_{\mathbf{X}} (W - \bar{W} - \mathbf{H}\mathbf{X})^T S^{-1} (W - \bar{W} - \mathbf{H}\mathbf{X})^T \end{aligned} \quad \text{Ec. 3. 86}$$

y por tanto:

### Proposición 3. 3

Dado el problema  $\min_{M,t} E_{def}(\mathbf{r}, \mathbf{M}\bar{\mathbf{r}}+t)$ , que puede expresarse de forma equivalente como:  $\min_{\mathbf{x}} (w-\bar{w}-\mathbf{H}\mathbf{x})^T \mathbf{S}^{-1} (w-\bar{w}-\mathbf{H}\mathbf{x})$  entonces:

- La solución óptima al problema es  $\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{S}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{S}^{-1} (w - \bar{w})$
- Se verifica  $\mathbf{E}_{def}^T(\mathbf{r}, \bar{\mathbf{r}}) = (w-\bar{w})^T \mathbf{S}^{-1/2} \mathbf{P} \mathbf{S}^{-1/2} (w-\bar{w})$  donde  $\mathbf{P}$  es la matriz de proyección  $\mathbf{P} = \mathbf{I} - \mathbf{S}^{-1/2} \mathbf{H} (\mathbf{H}^T \mathbf{S}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{S}^{-1/2}$

### Demostración

- Basta derivar la forma cuadrática y resolver :

$$\frac{\partial}{\partial \mathbf{x}} (w-\bar{w}-\mathbf{H}\mathbf{x})^T \mathbf{S}^{-1} (w-\bar{w}-\mathbf{H}\mathbf{x}) = -2 \mathbf{H}^T \mathbf{S}^{-1} (w-\bar{w}) + 2 \mathbf{H}^T \mathbf{S}^{-1} \mathbf{H} (w-\bar{w}) = \mathbf{0}.$$

La solución de la ecuación es:  $\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{S}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{S}^{-1} (w - \bar{w})$

- Basta sustituir  $\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{S}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{S}^{-1} (w - \bar{w})$  en la forma cuadrática.

### Proposición 3. 4

Dado el problema  $\min_{\mathbf{r}} t \mathbf{E}_{def}^T(\mathbf{r}, \bar{\mathbf{r}}) + E_{aju}(\mathbf{r}, \bar{\mathbf{r}})$

- Es equivalente a  $\min_w t (w-\bar{w})^T \mathbf{S}^{-1/2} \mathbf{P} \mathbf{S}^{-1/2} (w-\bar{w}) + (w-w_D)^T \mathbf{I} (w-w_D)$
- La solución óptima del problema es:  $\hat{w} = (\mathbf{I} + t \mathbf{S}^{-1/2} \mathbf{P} \mathbf{S}^{-1/2})^{-1} (w_D + t \mathbf{S}^{-1/2} \mathbf{P} \mathbf{S}^{-1/2} \bar{w})$

### Demostración

- Evidente pues  $E_{aju}(\mathbf{r}, \mathbf{r}_D) = (w-w_D)^T \mathbf{I} (w-w_D)$

- De nuevo basta con calcular la parcial con respecto a  $w$  e igualar a  $\mathbf{0}$

### Proposición 3. 5

En las condiciones anteriores si  $\mathbf{S}=\mathbf{I}$  donde  $\mathbf{I}$  es la matriz identidad y que corresponde al caso  $\mathbf{H}=\mathbf{0}$ . La solución óptima es

$$\hat{w} = \mathbf{P} \left( \frac{t}{t+1} \bar{w} + \frac{1}{t+1} w_D \right) + (\mathbf{I} - \mathbf{P}) w_D$$

### Demostración

$$\begin{aligned} \hat{w} &= (tP + I)^{-1}(tP\bar{w} + w_D) = ((t+1)P + I - P)^{-1}(tP\bar{w} + w_D) = \\ &= \left( \left( \frac{1}{t+1} \right)P + (I - P) \right) (tP\bar{w} + w_D) = \\ &= \left( \left( \frac{1}{t+1} \right)P(t\bar{w} + w_D) + (I - P)w_D \right) = P \left( \frac{t}{t+1}\bar{w} + \frac{1}{t+1}w_D \right) + (I - P)w_D \end{aligned}$$

La importancia de la proposición anterior es que afirma que en el caso estándar de que las curvas tengan su norma sobre  $L_2(I)$  la solución óptima puede encontrarse gracias a la ortogonalidad de los wavelets sin tener que realizar ninguna inversión matricial. Además la expresión anterior permite una interpretación sencilla: en el subespacio determinado por el conjunto de transformaciones toda la información se obtiene de la imagen a partir de  $w_D$  mientras que en el subespacio de deformación se toma una media ponderada entre la información de la imagen y la del contorno de referencia  $\bar{w}$ , siendo esta ponderación función del parámetro de regularización  $t$ .

A continuación se muestra un experimento mostrando el proceso de ajuste. En la Figura 3. 7 (de [Pérez 2000a]) se puede ver la posición inicial del contorno que se toma como contorno de referencia en la Figura 3. 8 puede verse el resultado del proceso de ajuste.



Figura 3. 7 Posición Inicial



Figura 3. 8 Ajuste a la imagen

Una vez definido el procedimiento de ajuste basado en la función de deformación resulta de utilidad su interpretación probabilística pues proporciona la posibilidad de interpretar el parámetro de regularización del problema de ajuste para un espacio de Besov general y permitirá dar a la función de ajuste su forma definitiva.

### 3.3.3 Modelos probabilísticos de deformación basados en wavelets

Tal y como se señaló en la sección 2.2.2.4, el emparejamiento de modelos deformables puede ser formulado como la estimación máximo a posteriori (MAP) con una adecuada formulación probabilística.

Así se tiene:

Probabilidad a priori :

$$P(w | t; \bar{w}) \propto \exp(-t E_{def}(w, \bar{w}))$$

Función de verosimilitud

Ec. 3. 87

$$P(w_D | w) \propto \exp(-E_{aju}(w, w_D))$$

Distribución a posteriori

$$P(w | w_D; \bar{w}) \propto \exp(-t E_{def}(w, \bar{w}) + E_{aju}(w, w_D))$$

### 3.3.3.1 Formulación de la distribución a priori

Utilizando la forma de la función de deformación para espacios de Besov (Ec. 3. 78 ) se tiene:

$$\begin{aligned} P(w | t; \bar{w}) \propto \exp(-t E_{def}(w, \bar{w})) = \\ \exp\left(-t \left( |\Delta c_{0,0;x}|^2 + |\Delta c_{0,0;y}|^2 + \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p/2-1)} \left( \sqrt{|\Delta d_{j,k;x}|^2 + |\Delta d_{j,k;y}|^2} \right)^p \right)^{q/p} \right)\right) = \\ \exp\left(-t \left( |\Delta c_{0,0;x}|^2 + |\Delta c_{0,0;y}|^2 \right)\right) \exp\left(-t \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p/2-1)} \left( \sqrt{|\Delta d_{j,k;x}|^2 + |\Delta d_{j,k;y}|^2} \right)^p \right)^{q/p} \right) \end{aligned} \quad \text{Ec. 3. 88}$$

Por tanto la función de probabilidad se divide en dos términos, el primero está asociado con la traslación de la curva y tiene distribución normal:

$$\begin{pmatrix} \Delta c_{0,0;x} \\ \Delta c_{0,0;y} \end{pmatrix} \sim N_2(\mathbf{0}, \mathbf{D}), \quad \mathbf{D} = \begin{pmatrix} \mathbf{s}_{TR}^2 & 0 \\ 0 & \mathbf{s}_{TR}^2 \end{pmatrix}, \quad \mathbf{s}_{TR}^2 = \frac{1}{2t} \quad \text{Ec. 3. 89}$$

donde  $\mathbf{s}_{TR}^2$  representa la incertidumbre asociada con una posible traslación de la posición de la curva.

Para el estudio del segundo término se asumirá  $q=p$  con lo que se éste queda definido como sigue:

$$\begin{aligned} \exp\left(-t \sum_{j \geq 0} \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p/2-1)} \left( \sqrt{|\Delta d_{j,k;x}|^2 + |\Delta d_{j,k;y}|^2} \right)^p \right) = \\ \prod_{j \geq 0} \prod_{k=0}^{2^j-1} \exp\left(-t 2^{ajp} 2^{j(p/2-1)} \left( \sqrt{|\Delta d_{j,k;x}|^2 + |\Delta d_{j,k;y}|^2} \right)^p \right) \end{aligned} \quad \text{Ec. 3. 90}$$

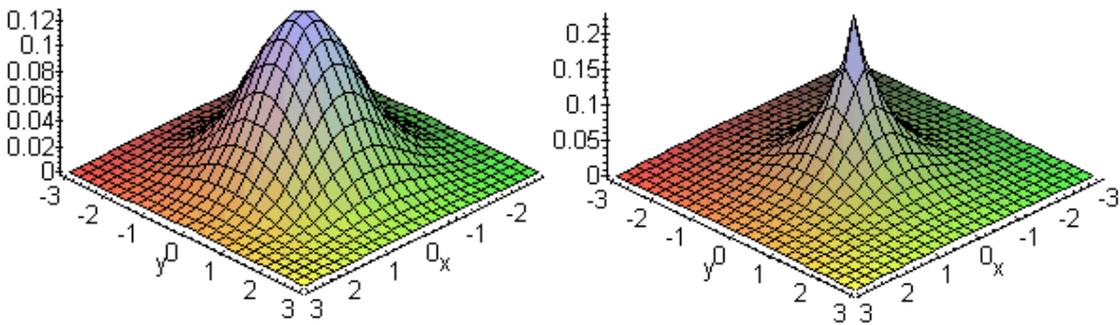
para el estudio de este segundo término se propone en esta tesis una generalización de la distribución generalizada gaussiana ( Ver 3.1.11), así se tiene:

**Definición**

La distribución gaussiana generalizada bidimensional simétrica de media  $\mathbf{0}$ ,  $DGGS_n(\mathbf{0}, \mathbf{s}^2)$  con parámetro de forma  $n$  y varianza  $\mathbf{s}^2$  se define a partir de la función de densidad:

$$f(x, y) = \frac{nh(n)^2}{2\Gamma(2/n)} \frac{1}{ps^2} \exp\left(-\left(h(n)\sqrt{x^2 + y^2}/s\right)^n\right), \quad h(n) = \sqrt{\frac{\Gamma(3/n)}{\Gamma(1/n)}} \quad \text{Ec. 3. 91}$$

donde  $\Gamma$  representa la función gamma.



**Figura 3. 9** Función de densidad  $GGD(\mathbf{0},1)$  para  $n = 2$  (izquierda),  $n = 1$  (derecha).

De esta forma, el segundo término expresa la distribución conjunta de variables DGGS independientes donde cada par de variables sigue la distribución:

$$\begin{pmatrix} \Delta d_{j,k;x} \\ \Delta d_{j,k;y} \end{pmatrix} \sim DGGS_p(\mathbf{0}, \mathbf{s}_{j,k}^2), \quad \mathbf{s}_{j,k} = \frac{h(p)}{t^{1/p}} 2^{-j(a+1/2-1/p)} \quad \text{Ec. 3. 92}$$

Por tanto el modelo probabilístico asociado a Ec. 3. 90 es la distribución conjunta de variables independientes con distribución DGGS y varianzas decrecientes exponencialmente. Este hecho puede formalizarse de forma rigurosa en el siguiente teorema:

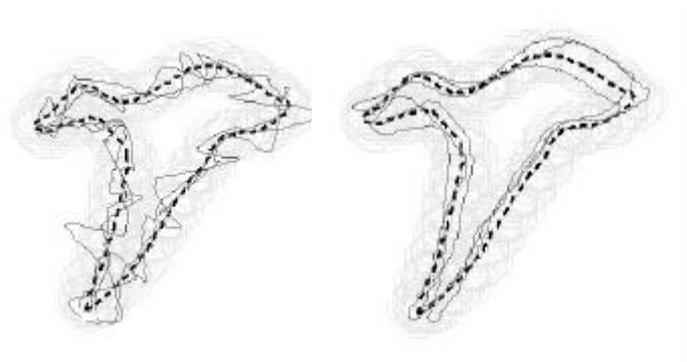
**Teorema 3.4**

Supongamos que  $\mathbf{d}_{j,k} = (d_{j,k;x}, d_{j,k;y})^T \stackrel{iid}{\sim} \text{DGGSP}(\mathbf{0}, \mathbf{s}_j^2)$  con  $\mathbf{s}_j = 2^{-jb} \mathbf{s}_0$  con  $b > 0$  y  $\mathbf{s}_0 > 0$  entonces para  $0 < p, q < \infty$  las curvas realizaciones del modelo están casi seguro en  $B_q^a(L_p(I))$  si y solo si  $b > a + 1/2$ .

**Demostración**

La demostración es sigue las líneas del caso unidimensional [Abramovich 1998] y puede encontrarse en el Apéndice A.

A continuación se presentan algunas realizaciones del modelo probabilístico en un contorno de prueba que se muestra en línea discontinua. Tal y como se predice por el modelo, cuando se incrementa  $b$  aparecen deformaciones más suaves. Alrededor de la figura se muestra también en gris claro un intervalo de confianza del 99% para los puntos de la curva.



**Figura 3.10** Realizaciones del modelo probabilístico. Los valores del parámetro son  $b=0$  para la imagen de la izquierda y  $b=1.6$  para la de la derecha. El valor de  $p$  es igual a 2.

Es importante señalar que con el modelo actual las varianzas debida a traslación y al modelo de deformación  $\mathbf{s}_{TR}^2$  y  $\mathbf{s}_{j,k}^2$  respectivamente están ligadas a través del parámetro de regularización  $t$ . La interpretación probabilística permite que ambas cantidades sean determinadas independientemente proporcionando así una mayor flexibilidad al modelo.

**3.3.3.1.1 Modelos probabilísticos en el espacio de Sobolev**

Un caso especialmente interesante se produce cuando  $p=2$ . En este caso las realizaciones del modelo probabilísticos están casi seguro en el espacio de Sobolev  $W^a(L_2(I))$ . En este caso todas las distribuciones son gaussianas e independientes lo que permitirá el desarrollo analítico del estimador MAP. Por tanto:

$$\mathbf{d}_{j,k} \sim N_2(\bar{\mathbf{d}}_{j,k}, 2^{-2jb} \mathbf{s}_{DEF}^2 \mathbf{I}) \quad \text{Ec. 3. 93}$$

lo cual, junto con el coeficiente de traslación independiente de los anteriores :

$$\mathbf{c}_{0,0} \sim N_2(\bar{\mathbf{c}}_{0,0}, \mathbf{s}_{TR}^2 \mathbf{I}) \quad \text{Ec. 3. 94}$$

determina completamente el modelo.

Todo lo expuesto anteriormente puede establecerse de forma precisa como:

$$P(\mathbf{w} | \bar{\mathbf{w}}) \propto \exp\left(-\frac{1}{2}(\mathbf{w} - \bar{\mathbf{w}})^T \mathbf{S}^{-1} (\mathbf{w} - \bar{\mathbf{w}})\right) \quad \text{Ec. 3. 95}$$

$$\mathbf{w} = (\mathbf{c}_{0,0,x}, \mathbf{d}_{0,0,x}, \dots, \mathbf{d}_{j,k,x}, \dots, \mathbf{d}_{J-1,2^{J-1}-1,x}, \\ , \mathbf{c}_{0,0,y}, \mathbf{d}_{0,0,y}, \dots, \mathbf{d}_{j,k,y}, \dots, \mathbf{d}_{J-1,2^{J-1}-1,y}) \\ j = 0 \dots J-1, k = 0 \dots 2^j - 1$$

$$\mathbf{S} = \begin{pmatrix} \mathbf{s}_{TR}^2 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \mathbf{s}_{DEF}^2 2^{-2b0} & & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & & \vdots \\ 0 & 0 & \dots & \mathbf{s}_{DEF}^2 2^{-2b(J-1)} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \mathbf{s}_{TR}^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \mathbf{s}_{DEF}^2 2^{-2b0} & & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & \mathbf{s}_{DEF}^2 2^{-2b(J-1)} \end{pmatrix}$$

donde  $N=2^J$  es el número de puntos de la curva discretizada.

### 3.3.3.1.2 Aprendizaje de los parámetros

Se mostrará a continuación dos métodos para la estimación de los parámetros del modelo [Pérez 2001]. El primero está basado en las ecuaciones de máxima verosimilitud (EMV) que se aplicará en el caso gaussiano mientras que el segundo está basado en el desplazamiento cuadrático medio y que puede aplicarse en el caso de la distribución gaussiana generalizada.

- **Estimación por máxima verosimilitud (EMV) en el caso gaussiano**

#### Proposición 3. 6

Dado un conjunto de muestras de descriptores  $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ , las ecuaciones de máxima verosimilitud son:

$$\widehat{\bar{w}} = \bar{w}$$

$$\hat{\mathbf{s}}_{TR}^2 = \frac{s_{00} + s_{NN}}{2}$$

$$\frac{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} (i-1)r_i}{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} r_i} = \frac{N(\log_2 N - 2) + 2}{(N-1)} \quad \hat{\mathbf{s}}_{DEF}^2 = \frac{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} r_i}{2(N-1)}$$

Ec. 3. 96

donde:

$$\bar{w} = \frac{1}{m} \sum_{i=1}^m w_i, \mathbf{S} = (s_{ij}) = \frac{1}{m} \sum_{i=1}^m (w_i - \bar{w})(w_i - \bar{w})^T, r_i = \sum_{j=2^{i-1}}^{2^i-1} s_{jj} + s_{j+n, j+n}$$

y  $N$  es el número de puntos de la curva discretizada.

### Demostración

Basta derivar con respecto a los parámetros e igualar a  $\mathbf{0}$ .

Todos los parámetros pueden calcularse fácilmente excepto  $\mathbf{b}$ . Su ecuación debe ser resuelta numéricamente.

- **Estimación por desplazamiento cuadrático medio**

En caso de tener una estimación de  $\mathbf{b}$  y el desplazamiento cuadrático medio desde la curva de referencia se puede utilizar el siguiente resultado:

### Proposición 3. 7

Sea una curva descrita como en Ec. 3. 88 con  $\mathbf{S}$  una matriz de varianzas-covarianzas arbitraria entonces el desplazamiento cuadrático medio  $\bar{\mathbf{r}}^2$  a lo largo de la curva viene dado por:

$$\bar{\mathbf{r}}^2 = \frac{\text{Traza}(\mathbf{S})}{N} \quad \text{Ec. 3. 97}$$

donde  $N$  es el número de puntos de la curva discretizada.

### Demostración

Basta adaptar la demostración de [Blake 1998]

Por tanto  $\mathbf{s}_{TR}^2$  y  $\mathbf{s}_{DEF}^2$  pueden ser estimados a partir de  $\mathbf{b}$  y el desplazamiento cuadrático medio debido a translación y deformación  $\bar{\mathbf{r}}_{TR}^2$  y  $\bar{\mathbf{p}}_{DEF}^2$  respectivamente utilizando la siguiente proposición.



$$\exp\left(-\frac{1}{2\mathbf{s}_D^2}\left(|c_{0,0;x} - c_{D,0,0;x}|^2 + |c_{j,k;y} - c_{D,0,0;x}|^2 + \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} |d_{j,k;x} - d_{D,j,k;x}|^2 + |d_{j,k;y} - d_{D,j,k;x}|^2\right)\right)$$

$$\min_{M,t} \left\{ \exp\left(-\frac{1}{2\mathbf{s}_{TR}^2} |c_{0,0;x} - (\bar{c}_{0,0;x} + t_1)|^2 + |c_{0,0;y} - (\bar{c}_{0,0;y} + t_2)|^2\right) \right. \quad \text{Ec. 3. 100}$$

$$\left. \exp\left(-\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \left(\frac{\mathbf{h}(p)}{\mathbf{g}_{jk}} \sqrt{|d_{j,k;x} - m_{11}\bar{d}_{j,k;x} + m_{12}\bar{d}_{j,k;y}|^2 + |d_{j,k;y} - m_{21}\bar{d}_{j,k;x} + m_{22}\bar{d}_{j,k;y}|^2}\right)^p\right)\right\}$$

$$\mathbf{g}_{jk} = \mathbf{s}_{DEF} 2^{-bj}$$

nótese como el problema se descompone en:

### Problema 1

Hallar es estimador MAP para  $(c_{0,0;x}, c_{0,0;y})$  a partir de

$$\exp\left(-\frac{1}{2\mathbf{s}_D^2}\left(|c_{0,0;x} - c_{D,0,0;x}|^2 + |c_{j,k;y} - c_{D,0,0;x}|^2\right)\right)$$

$$\min_t \left\{ \exp\left(-\frac{1}{2\mathbf{s}_{TR}^2} |c_{0,0;x} - (\bar{c}_{0,0;x} + t_1)|^2 + |c_{0,0;y} - (\bar{c}_{0,0;y} + t_2)|^2\right) \right\}$$

### Problema 2

Hallar es estimador MAP para  $(d_{j,k;x}, d_{j,k;y})$  a partir de

$$\exp\left(-\frac{1}{2\mathbf{s}_D^2}\left(\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} |d_{j,k;x} - d_{D,j,k;x}|^2 + |d_{j,k;y} - d_{D,j,k;x}|^2\right)\right)$$

$$\min_M \left\{ \exp\left(-\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \left(\frac{\mathbf{h}(p)}{\mathbf{g}_{jk}} \sqrt{|d_{j,k;x} - m_{11}\bar{d}_{j,k;x} + m_{12}\bar{d}_{j,k;y}|^2 + |d_{j,k;y} - m_{21}\bar{d}_{j,k;x} + m_{22}\bar{d}_{j,k;y}|^2}\right)^p\right)\right\}$$

$$\mathbf{g}_{jk} = \mathbf{s}_{DEF} 2^{-bj}$$

Entonces se tiene:

### Proposición 3. 9

Dado el Problema 1 definido anteriormente se tiene:

- a) El valor óptimo de  $t=(t_1, t_2)$  es:  $\hat{t}=(c_{0,0;x} - \bar{c}_{0,0;x}, c_{0,0;y} - \bar{c}_{0,0;y})$

b) El estimador MAP para  $\mathbf{c}_{0,0}=(c_{0,0;x},c_{0,0;y})$  es:  $\hat{\mathbf{c}}_{0,0}=(c_{D,0,0;x},c_{D,0,0;y})$

**Demostración**

a) El problema de la determinación de  $\mathbf{t}$  es equivalente a:

$$\min_b |c_{0,0;x} - (\bar{c}_{0,0;x} + t_1)|^2 + |c_{0,0;y} - (\bar{c}_{0,0;y} + t_2)|^2$$

y por tanto su valor óptimo es:  $\hat{\mathbf{t}}=(c_{0,0;x} - \bar{c}_{0,0;x}, c_{0,0;y} - \bar{c}_{0,0;y})$

b) Dado el valor de  $\mathbf{t}$  obtenido anteriormente el problema de la determinación de  $\mathbf{c}_{00}$  es equivalente a:

$$\min_{\mathbf{c}_{0,0}} |c_{0,0;x} - c_{D,0,0;x}|^2 + |c_{j,k;y} - c_{D,0,0;x}|^2$$

y por tanto su valor óptimo es:  $\hat{\mathbf{c}}_{0,0}=(c_{D,0,0;x},c_{D,0,0;y})$

El Problema 2 puede entonces formularse como:

$$\min_w \left\{ h(w - w_D) \min_x g(w - (\bar{H}x + \bar{w})) \right\}$$

$$h(w - w_D) = \exp \left( -\frac{1}{2\mathbf{s}_o^2} \left( \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} |d_{j,k;x} - d_{D,j,k;x}|^2 + |d_{j,k;y} - d_{D,j,k;x}|^2 \right) \right)$$

$$g(w - (\bar{H}x + \bar{w})) = \exp \left( -\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \left( \frac{\mathbf{h}(p)}{\mathbf{g}_{jk}} \sqrt{|d_{j,k;x} - m_{11}\bar{d}_{j,k;x} + m_{12}\bar{d}_{j,k;y}|^2 + |d_{j,k;y} - m_{21}\bar{d}_{j,k;x} + m_{22}\bar{d}_{j,k;y}|^2} \right)^p \right)$$

con:

$$w = \begin{pmatrix} d_{0,0;x} \\ d_{1,0;x} \\ \vdots \\ d_{j,l;x} \\ \vdots \\ d_{J-1,2^{J-1}-1;x} \\ d_{0,0;y} \\ d_{1,0;y} \\ \vdots \\ d_{j,l;y} \\ \vdots \\ d_{J-1,2^{J-1}-1;y} \end{pmatrix} \quad \bar{w} = \begin{pmatrix} \bar{d}_{0,0;x} \\ \bar{d}_{1,0;x} \\ \vdots \\ \bar{d}_{j,l;x} \\ \vdots \\ \bar{d}_{J-1,2^{J-1}-1;x} \\ \bar{d}_{0,0;y} \\ \bar{d}_{1,0;y} \\ \vdots \\ \bar{d}_{j,l;y} \\ \vdots \\ \bar{d}_{J-1,2^{J-1}-1;y} \end{pmatrix} \quad w_D = \begin{pmatrix} d_{D,0,0;x} \\ d_{D,1,0;x} \\ \vdots \\ d_{D,j,l;x} \\ \vdots \\ d_{D,J-1,2^{J-1}-1;x} \\ d_{D,0,0;y} \\ d_{D,1,0;y} \\ \vdots \\ d_{D,j,l;y} \\ \vdots \\ d_{D,J-1,2^{J-1}-1;y} \end{pmatrix}$$

$$\bar{\mathbf{H}} = \begin{pmatrix} \bar{d}_{0,0;x} & \bar{d}_{0,0;y} & 0 & 0 \\ \bar{d}_{1,0;x} & \bar{d}_{1,0;y} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \bar{d}_{j,l;x} & \bar{d}_{j,l;y} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \bar{d}_{J-1,2^{J-1}-1;x} & \bar{d}_{J-1,2^{J-1}-1;y} & 0 & 0 \\ 0 & 0 & \bar{d}_{0,0;y} & \bar{d}_{0,0;y} \\ 0 & 0 & \bar{d}_{1,0;y} & \bar{d}_{1,0;y} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \bar{d}_{j,l;y} & \bar{d}_{j,l;y} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \bar{d}_{J-1,2^{J-1}-1;y} & \bar{d}_{J-1,2^{J-1}-1;y} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} m_{11} \\ m_{12} \\ m_{21} \\ m_{22} \end{pmatrix}$$

nótese como para simplificar la notación se mantiene el uso de  $w$  aunque éste no tiene ahora la componente asociada a traslación. La resolución del Problema 2 es compleja debido a la doble optimización en  $w$  y  $x$ . Por ello se formula de forma equivalente como:

### Teorema 3.5

Sea el problema  $\min_w \{h(w - w_D) \min_x g(w - (\bar{\mathbf{H}}x + \bar{w}))\}$ , entonces la solución óptima es:  $w_D + \mathbf{P}(\hat{\mathbf{x}} + \bar{w} - w_D)$  con  $\mathbf{P} = \mathbf{I} - \bar{\mathbf{H}}(\bar{\mathbf{H}}^T \bar{\mathbf{H}})^{-1} \bar{\mathbf{H}}^T$  y  $\hat{\mathbf{x}}$  la solución óptima de  $\min_x h(\mathbf{P}(\mathbf{x} + \bar{w} - w_D))g(\mathbf{x})$  (Problema 3).

### Demostración

El problema a optimizar es equivalente a:

$$\min_{w, x} f(w, x), \quad f(w, x) = h(w - w_D)g(w - (\bar{\mathbf{H}}x + \bar{w}))$$

descomponiendo  $h$

$$h(w - w_D) = h(\mathbf{P}(w - w_D) + (\mathbf{I} - \mathbf{P})(w - w_D)) = h(\mathbf{P}(w - w_D))h(\mathbf{I} - \mathbf{P})(w - w_D)$$

definiendo entonces  $\mathbf{x} = w - (\bar{\mathbf{H}}x + \bar{w})$  se tiene:

$$\begin{aligned} f(w, x) &= h(w - w_D)g(w - (\bar{\mathbf{H}}x + \bar{w})) = h(\mathbf{P}(\mathbf{x} + \bar{w} - w_D))h(\mathbf{I} - \mathbf{P})(w - w_D)g(\mathbf{x}) \\ &\geq h(\mathbf{P}(\mathbf{x} + \bar{w} - w_D))g(\mathbf{x}) \end{aligned}$$

por tanto :

$$\begin{aligned} \min_{w, x} f(w, x) &\geq \min_x h(\mathbf{P}(\mathbf{x} + \bar{w} - w_D))g(\mathbf{x}) = h(\mathbf{P}(\hat{\mathbf{x}} + \bar{w} - w_D))g(\hat{\mathbf{x}}) = \\ &= f(w_D + \mathbf{P}(\hat{\mathbf{x}} + \bar{w} - w_D), -(\bar{\mathbf{H}}^T \bar{\mathbf{H}})^{-1} \bar{\mathbf{H}}^T (\hat{\mathbf{x}} + \bar{w} - w_D)). \end{aligned}$$

Desafortunadamente el problema:  $\min_{\mathbf{x}} h(\mathbf{P}(\mathbf{x} + \bar{\mathbf{w}} - \mathbf{w}_D))g(\mathbf{x})$  es complejo de resolver debido a la presencia de la matriz de proyección  $\mathbf{P}$  y a la no linealidad de la función  $g$ . En esta tesis se dará un método de resolución aproximada en la siguiente sección. No obstante existe un caso en los cuales la resolución es posible de forma analítica: cuando el proceso de optimización se hace sobre el espacio de Sobolev y que será presentado en la Sección 3.3.3.5.

### 3.3.3.4 El caso general

Como se ha comentado anteriormente la solución del problema de emparejamiento en el caso general es muy compleja debido a la complejidad de la función objetivo. A continuación se muestra un procedimiento aproximado para la optimización de la función objetivo.

La base de dicho procedimiento es el último teorema de la sección anterior junto con una aproximación suave de la norma en  $L_p(\cdot)$  elevada a la p-ésima potencia [Belge 1999]:

$$\|\mathbf{x}\|_p^p \approx \sum_i \left( |x_i|^2 + \mathbf{e} \right)^{p/2} - \mathbf{e}^{p/2} \quad \text{Ec. 3. 101}$$

utilizando esta aproximación el Problema 3 queda definido como:

$$\begin{aligned} \min_{\mathbf{x}} h(\mathbf{P}(\mathbf{x} - d\mathbf{w}))g(\mathbf{x}) = \\ \min_{\mathbf{x}} \exp\left(-\frac{1}{2\mathbf{S}_D^2}(\mathbf{x} - d\mathbf{w})^T \mathbf{P}(\mathbf{x} - d\mathbf{w})\right) \cdot \exp\left(-\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \left(\frac{h(p)}{\mathbf{g}_{jk}} \sqrt{\mathbf{x}_{j,k;x}^2 + \mathbf{x}_{j,k;y}^2}\right)^p\right) \approx \\ \min_{\mathbf{x}} \exp\left(-\frac{1}{2\mathbf{S}_D^2}(\mathbf{x} - d\mathbf{w})^T \mathbf{P}(\mathbf{x} - d\mathbf{w})\right) \cdot \exp\left(-\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \left(\frac{h(p)}{\mathbf{g}_{jk}}\right)^p \left(\mathbf{x}_{j,k;x}^2 + \mathbf{x}_{j,k;y}^2 + \mathbf{e}\right)^{p/2} - \mathbf{e}^{p/2}\right) \\ d\mathbf{w} = \mathbf{w}_D - \bar{\mathbf{w}} \end{aligned}$$

teniendo en cuenta que  $\mathbf{P}$  es una matriz de proyección, derivando con respecto a  $\mathbf{x}$  e igualando a  $\mathbf{0}$  se tiene que la solución óptima  $\hat{\mathbf{x}}$  verifica:

$$\begin{aligned} \left(\mathbf{P} + \frac{p}{2} \mathbf{D}_{\hat{\mathbf{x}}}\right) \hat{\mathbf{x}} = \mathbf{P} d\mathbf{w}, \quad \text{con:} \\ \mathbf{D}_{\hat{\mathbf{x}}} = \begin{pmatrix} \mathbf{E}_{\hat{\mathbf{x}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{\hat{\mathbf{x}}} \end{pmatrix} \end{aligned} \quad \text{Ec. 3. 102}$$

$$E_{\hat{\mathbf{x}}} = \frac{\mathbf{h}(p)^p}{\mathbf{h}(2)^2} \begin{pmatrix} \frac{(\mathbf{s}_o)^2}{(\mathbf{g}_{00})^p \left( \|\hat{\mathbf{x}}_{0,0}\|_2^2 + \mathbf{e} \right)^{1-p/2}} & 0 & \dots & 0 \\ 0 & & & \vdots \\ \vdots & & & 0 \\ 0 & 0 & \dots & \frac{(\mathbf{s}_o)^2}{(\mathbf{g}_{00})^p \left( \|\hat{\mathbf{x}}_{J-1,2^{J-1}-1}\|_2^2 + \mathbf{e} \right)^{1-p/2}} \end{pmatrix}$$

$$\|\hat{\mathbf{x}}_{j,k}\|_2^2 = (\hat{\mathbf{x}}_{j,k;x})^2 + (\hat{\mathbf{x}}_{j,k;y})^2$$

El sistema en Ec. 3. 102 se puede resolver de forma iterativa como [Belge 1999]:

$$\left( \mathbf{P} + \frac{p}{2} \mathbf{D}_{\hat{\mathbf{x}}_k} \right) \hat{\mathbf{x}}_{k+1} = \mathbf{P} d w , \quad \text{Ec. 3. 103}$$

donde  $\mathbf{D}_{\hat{\mathbf{x}}_k}$  se obtiene sustituyendo  $\hat{\mathbf{x}}$  por la solución actual  $\hat{\mathbf{x}}_k$  en Ec. 3. 102

Por tanto se tiene:

$$\hat{\mathbf{x}}_{k+1} = \left( \mathbf{P} + \frac{p}{2} \mathbf{D}_{\hat{\mathbf{x}}_k} \right)^{-1} \mathbf{P} d w , \quad \text{Ec. 3. 104}$$

Todos los resultados anteriores pueden expresarse por medio del siguiente teorema:

### Teorema 3. 6

Dado el problema de cálculo del estimador MAP:

$$\begin{aligned} & \min_{\substack{c_{00} \\ d_{j,k}}} \left\{ \exp \left( -\frac{1}{2\mathbf{s}_D^2} \left( |c_{0,0;x} - c_{D,0,0;x}|^2 + |c_{j,k;y} - c_{D,0,0;x}|^2 \right) \right) \right. \\ & \exp \left( -\frac{1}{2\mathbf{s}_D^2} \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} |d_{j,k;x} - d_{D,j,k;x}|^2 + |d_{j,k;y} - d_{D,j,k;x}|^2 \right) \cdot \\ & \left. \min_{M,t} \left\{ \exp \left( -\frac{1}{2\mathbf{s}_{TR}^2} |c_{0,0;x} - (\bar{c}_{0,0;x} + t_1)|^2 + |c_{0,0;y} - (\bar{c}_{0,0;y} + t_2)|^2 \right) \right. \right. \\ & \left. \left. \exp \left( -\sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \left( \frac{\mathbf{h}(p)}{\mathbf{g}_{jk}} \sqrt{|d_{j,k;x} - m_{11} \bar{d}_{j,k;x} + m_{12} \bar{d}_{j,k;y}|^2 + |d_{j,k;y} - m_{21} \bar{d}_{j,k;x} + m_{22} \bar{d}_{j,k;y}|^2} \right)^p \right) \right) \right\} \right\} \end{aligned}$$

$$\text{con } \mathbf{g}_{jk} = \mathbf{s}_{DEF} 2^{-bj}$$

entonces:

a) La solución exacta para  $\mathbf{c}_{00}$  es:  $\hat{\mathbf{c}}_{0,0} = (c_{D,0,0;x}, c_{D,0,0;y})$

b) El problema admite una solución aproximada para

$$\mathbf{w} = \left( d_{0,0;x} \quad \cdots \quad d_{j,j;x} \quad \cdots \quad d_{J-1,2^{J-1}-1;x} \quad d_{0,0;y} \quad \cdots \quad d_{j,j;y} \quad \cdots \quad d_{J-1,2^{J-1}-1;y} \right)^T \text{ como:}$$

$$\mathbf{w}_D + \left( \frac{2}{p} \mathbf{P}(\mathbf{D}_x)^{-1} + \mathbf{I} \right)^{-1} \mathbf{P}(\bar{\mathbf{w}} - \mathbf{w}_D) \text{ con } \mathbf{P} = \mathbf{I} - \bar{\mathbf{H}}(\bar{\mathbf{H}}^T \bar{\mathbf{H}})^{-1} \bar{\mathbf{H}}^T, \text{ y } \hat{\mathbf{x}} \text{ verificando}$$

$$\left( \mathbf{P} + \frac{p}{2} \mathbf{D}_x \right) \hat{\mathbf{x}} = \mathbf{P}(\mathbf{w}_D - \bar{\mathbf{w}}) \text{ con } \mathbf{D}_x \text{ definido en Ec. 3. 102.}$$

### Demostración

Es consecuencia de los resultados anteriores

Un subcaso interesante aparece cuando el conjunto de transformaciones para llevar a cabo la invarianza coincide con el conjunto de traslaciones.

### Corolario 3. 2

En las condiciones del problema definido en el problema anterior, si el conjunto de transformaciones coincide con el conjunto de traslaciones se tiene:

a) La solución exacta para  $\mathbf{c}_{00}$  es:  $\hat{\mathbf{c}}_{0,0} = (c_{D,0,0;x}, c_{D,0,0;y})$

b) El problema admite una solución aproximada para

$$\mathbf{w} = \left( d_{0,0;x} \quad \cdots \quad d_{j,j;x} \quad \cdots \quad d_{J-1,2^{J-1}-1;x} \quad d_{0,0;y} \quad \cdots \quad d_{j,j;y} \quad \cdots \quad d_{J-1,2^{J-1}-1;y} \right)^T \text{ como:}$$

$$\hat{d}_{j,k;x} = \hat{\mathbf{I}}_{j,k} d_{D,j,k;x} + (1 - \hat{\mathbf{I}}_{j,k}) \bar{d}_{j,k;x}$$

$$\hat{d}_{j,k;y} = \hat{\mathbf{I}}_{j,k} d_{D,j,k;y} + (1 - \hat{\mathbf{I}}_{j,k}) \bar{d}_{j,k;y}$$

$$\hat{\mathbf{I}}_{j,k} = \frac{(2/p) \mathbf{h}(2)^2 \mathbf{g}_{jk}^p \left( \left\| \hat{\mathbf{x}}_{j,k} \right\|_2^2 + \mathbf{e} \right)^{1-p/2}}{(2/p) \mathbf{h}(2)^2 \mathbf{g}_{jk}^p \left( \left\| \hat{\mathbf{x}}_{j,k} \right\|_2^2 + \mathbf{e} \right)^{1-p/2} + \mathbf{h}(p)^p \mathbf{s}_0^2}$$

con :

$$\hat{\mathbf{x}}_{j,k;x} = \hat{\mathbf{I}}_{j,k} (d_{D,j,k;x} - \bar{d}_{j,k;x})$$

$$\hat{\mathbf{x}}_{j,k;y} = \hat{\mathbf{I}}_{j,k} (d_{D,j,k;y} - \bar{d}_{j,k;y})$$

### Demostración

Puesto que en este caso  $\mathbf{P}=\mathbf{I}$  se tiene:

$$w_D + \left( \frac{2}{p} (D_{\hat{x}})^{-1} + I \right)^{-1} (\bar{w} - w_D) \text{ con } D_{\hat{x}} \text{ definido en Ec. 3. 102., y } \hat{x} \text{ verificando}$$

$$\left( I + \frac{p}{2} D_{\hat{x}} \right) \hat{x} = (w_D - \bar{w})$$

Como  $D_{\hat{x}}$  es diagonal se tiene:

$$(\hat{w})_i = (w_D)_i + \frac{1}{(2/p)(D_{\hat{x}}^{-1})_{ii} + 1} (\bar{w} - w_D)_i =$$

$$(w_D)_i \frac{(2/p)(D_{\hat{x}}^{-1})_{ii}}{(2/p)(D_{\hat{x}}^{-1})_{ii} + 1} + \frac{1}{(2/p)(D_{\hat{x}}^{-1})_{ii} + 1} (\bar{w})_i$$

Es interesante señalar que el problema anterior puede resolverse de forma exacta cuando  $p=1$  de la siguiente forma:

**Proposición 3. 10**

Dada la función:  $p(s) = (s-t)^2 + m|s|$ . Entonces la solución óptima es:

$$s = \begin{cases} 0 & |t| \leq m/2 \\ t - \text{sgn}(t)m/2 & |t| \geq m/2 \end{cases}$$

**Demostración**

Evidente mediante optimización de funciones de una variable.

Entonces se tiene:

**Proposición 3. 11**

En las condiciones del problema anterior, si el conjunto de transformaciones coincide con el conjunto de traslaciones y  $p=1$  se tiene:

a) La solución exacta para  $c_{00}$  es:  $\hat{c}_{0,0} = (c_{D,0,0;x}, c_{D,0,0;y})$

b) El problema admite una solución exacta para

$$w = (d_{0,0;x} \quad \dots \quad d_{j,j;x} \quad \dots \quad d_{J-1,2^{J-1}-1;x} \quad d_{0,0;y} \quad \dots \quad d_{j,j;y} \quad \dots \quad d_{J-1,2^{J-1}-1;y})^T$$

como:

$w_D + (\hat{x} + \bar{w} - w_D)$  con  $\hat{x}$  definido de forma que dados  $\hat{x}_{j,k} = (\hat{x}_{j,k;x}, \hat{x}_{j,k;y})$  y

$d_{w_{j,k}} = (d_{D,j,k;x} - \bar{d}_{j,k;x}, d_{D,j,k;y} - \bar{d}_{j,k;y})$  ambos vectores tienen la misma dirección y:

$$\|\hat{\mathbf{x}}_{j,k}\|_2 = \begin{cases} \left\| d\mathbf{w}_{j,k} \right\|_2 - \sqrt{2} \frac{\mathbf{s}_D^2}{\mathbf{s}_{TR}} & \text{si } \left\| d\mathbf{w}_{j,k} \right\|_2 \geq \sqrt{2} \frac{\mathbf{s}_0^2}{\mathbf{s}_{DEF} 2^{-bj}} \\ 0 & \text{en otro caso} \end{cases}$$

### Demostración

Basta aplicar la proposición anterior.

A continuación pasamos a desarrollar un caso de especial interés como es el emparejamiento en espacios de Sobolev y que tiene la particularidad de admitir una solución analítica.

#### 3.3.3.5 El caso gaussiano

El problema de emparejamiento en espacios de Sobolev se corresponde con el caso en que  $p$  es igual a 2. Entonces se tiene:

### Teorema 3.7

Dado el problema del cálculo del estimador MAP para  $p=2$  de Teorema 3.6 entonces:

a) La solución exacta para  $c_{00}$  es:  $\hat{c}_{0,0} = (c_{D,0,0;x}, c_{D,0,0;y})$

b) El problema admite una solución exacta para

$$\mathbf{w} = \left( d_{0,0;x} \quad \cdots \quad d_{j,j;x} \quad \cdots \quad d_{J-1,2^{J-1}-1;x} \quad d_{0,0;y} \quad \cdots \quad d_{j,j;y} \quad \cdots \quad d_{J-1,2^{J-1}-1;y} \right)^T \text{ como:}$$

$$\mathbf{w}_D + (\mathbf{P}\mathbf{D}^{-1} + \mathbf{I})^{-1} \mathbf{P}(\bar{\mathbf{w}} - \mathbf{w}_D) \text{ con } \mathbf{P} = \mathbf{I} - \bar{\mathbf{H}}(\bar{\mathbf{H}}^T \bar{\mathbf{H}})^{-1} \bar{\mathbf{H}}^T \text{ y } \mathbf{D} \text{ definida como matriz}$$

$$\mathbf{D} = \mathbf{s}_D^2 \mathbf{S}^{-1}$$

### Demostración

$$\min_{\mathbf{x}} (\mathbf{P}(\mathbf{x} + \bar{\mathbf{w}} - \mathbf{w}_D)) g(\mathbf{x}) = \min_{\mathbf{x}} \exp\left(-\frac{1}{2\mathbf{s}_D^2} (\mathbf{x} + \bar{\mathbf{w}} - \mathbf{w}_D)^T \mathbf{P} (\mathbf{x} + \bar{\mathbf{w}} - \mathbf{w}_D)\right) \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{S}^{-1} \mathbf{x}\right)$$

la solución óptima del problema anterior es:

$$\hat{\mathbf{x}} = (\mathbf{P} + \mathbf{s}_D^2 \mathbf{S}^{-1})^{-1} \mathbf{P}(\mathbf{w}_D - \bar{\mathbf{w}}) = (\mathbf{P} + \mathbf{D})^{-1} \mathbf{P}(\mathbf{w}_D - \bar{\mathbf{w}})$$

y por tanto la solución óptima del problema de ajuste es:

$$\begin{aligned}
 w_D + P((P + D)^{-1}P(w_D - \bar{w}) + (\bar{w} - w_D)) &= \\
 w_D + (I - P(P + D)^{-1})P(\bar{w} - w_D) &= \\
 w_D + (PD^{-1} + I)^{-1}P(\bar{w} - w_D) &
 \end{aligned}$$

Nótese como se obtiene la misma solución que en el caso general salvo que en este caso  $D_{\hat{x}}=D$

Análogamente al caso general si no hay invarianza frente a transformaciones  $P=I$  por lo que la solución óptima es:

$$w_D + (D^{-1} + I)^{-1}(\bar{w} - w_D) \quad \text{Ec. 3. 105}$$

esto es:

$$\begin{aligned}
 \hat{d}_{j,k;x} &= d_{D,j,k;x} \frac{2^{-2jb} \mathbf{s}_{DEF}^2}{2^{-2jb} \mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} + \frac{\mathbf{s}_D^2}{2^{-2jb} \mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} \bar{d}_{j,k;x} \\
 \hat{d}_{j,k;y} &= d_{D,j,k;y} \frac{2^{-2jb} \mathbf{s}_{DEF}^2}{2^{-2jb} \mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} + \frac{\mathbf{s}_D^2}{2^{-2jb} \mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} \bar{d}_{j,k;y}
 \end{aligned} \quad \text{Ec. 3. 106}$$

lo que indica que cuanto mayor es la escala menor es la influencia de los datos y por tanto el efecto de regularización es mayor. En el caso de hacer la regularización en  $L_2(I)$  se tiene:

$$\begin{aligned}
 \hat{d}_{j,k;y} &= d_{D,j,k;x} \frac{\mathbf{s}_{DEF}^2}{\mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} + \frac{\mathbf{s}_D^2}{\mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} \bar{d}_{j,k;x} \\
 \hat{d}_{j,k;y} &= d_{D,j,k;y} \frac{\mathbf{s}_{DEF}^2}{\mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} + \frac{\mathbf{s}_D^2}{\mathbf{s}_{DEF}^2 + \mathbf{s}_D^2} \bar{d}_{j,k;y}
 \end{aligned} \quad \text{Ec. 3. 107}$$

de forma que el contorno solución es una media ponderada entre el contorno extraído de los datos y el contorno medio. Las ponderaciones se hacen en función de la incertidumbre asociada tanto al modelo como a los datos.

### 3.3.4 Ejemplos de regularización

A continuación se muestran algunos ejemplos de emparejamiento utilizando las técnicas anteriores:

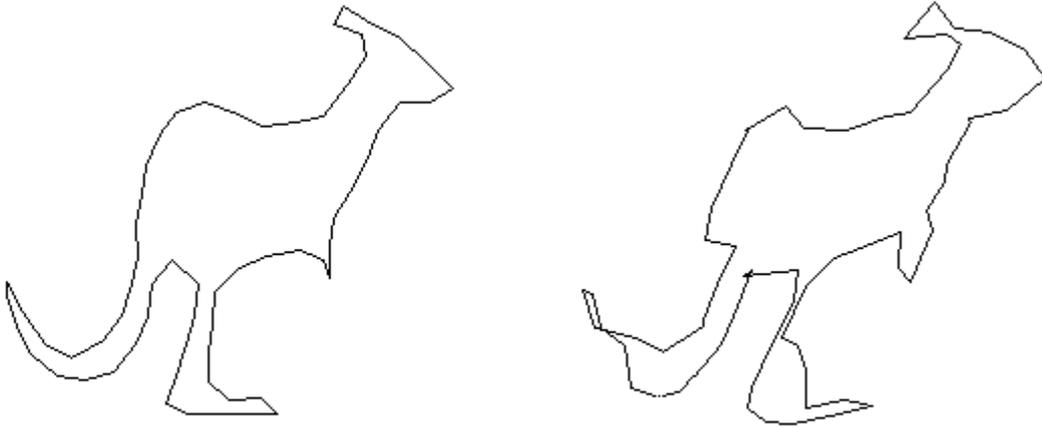


Figura 3. 11 Contorno de prueba (izquierda) y contorno con ruido uniforme añadido (derecha)

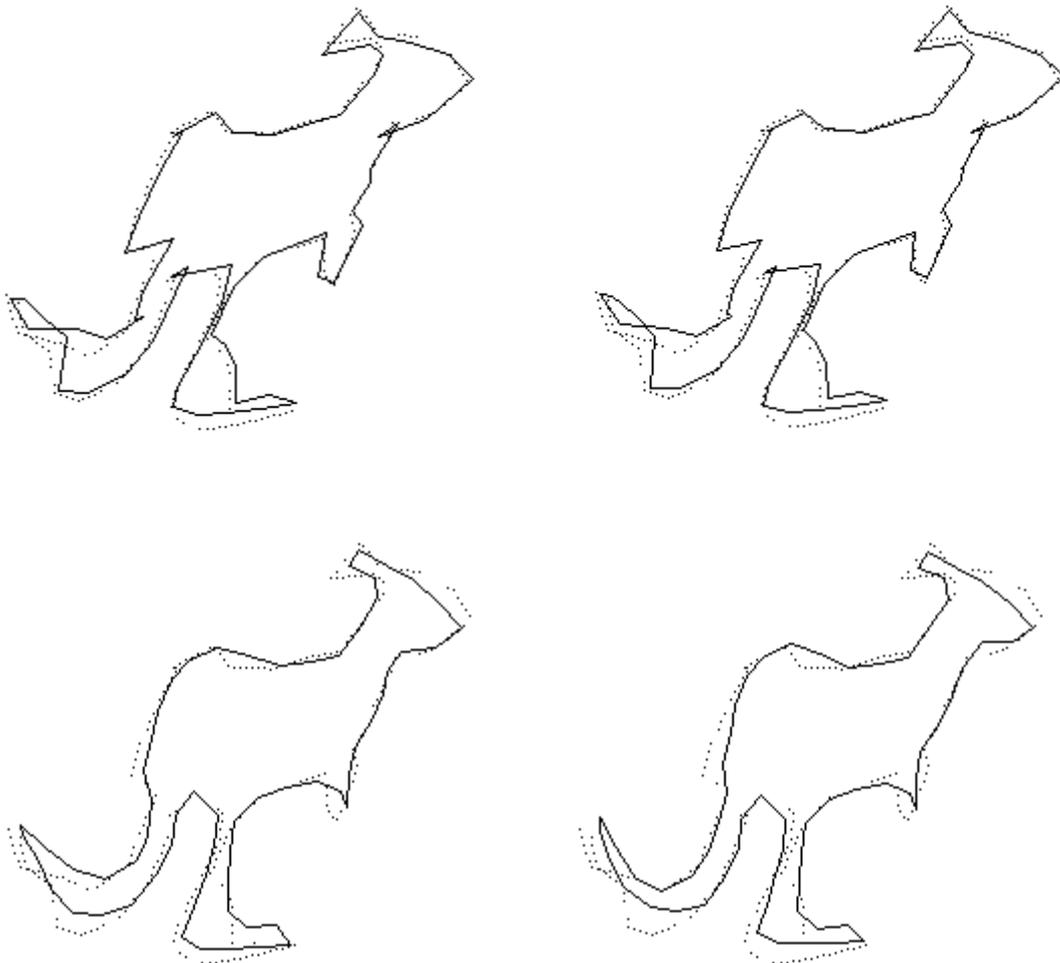
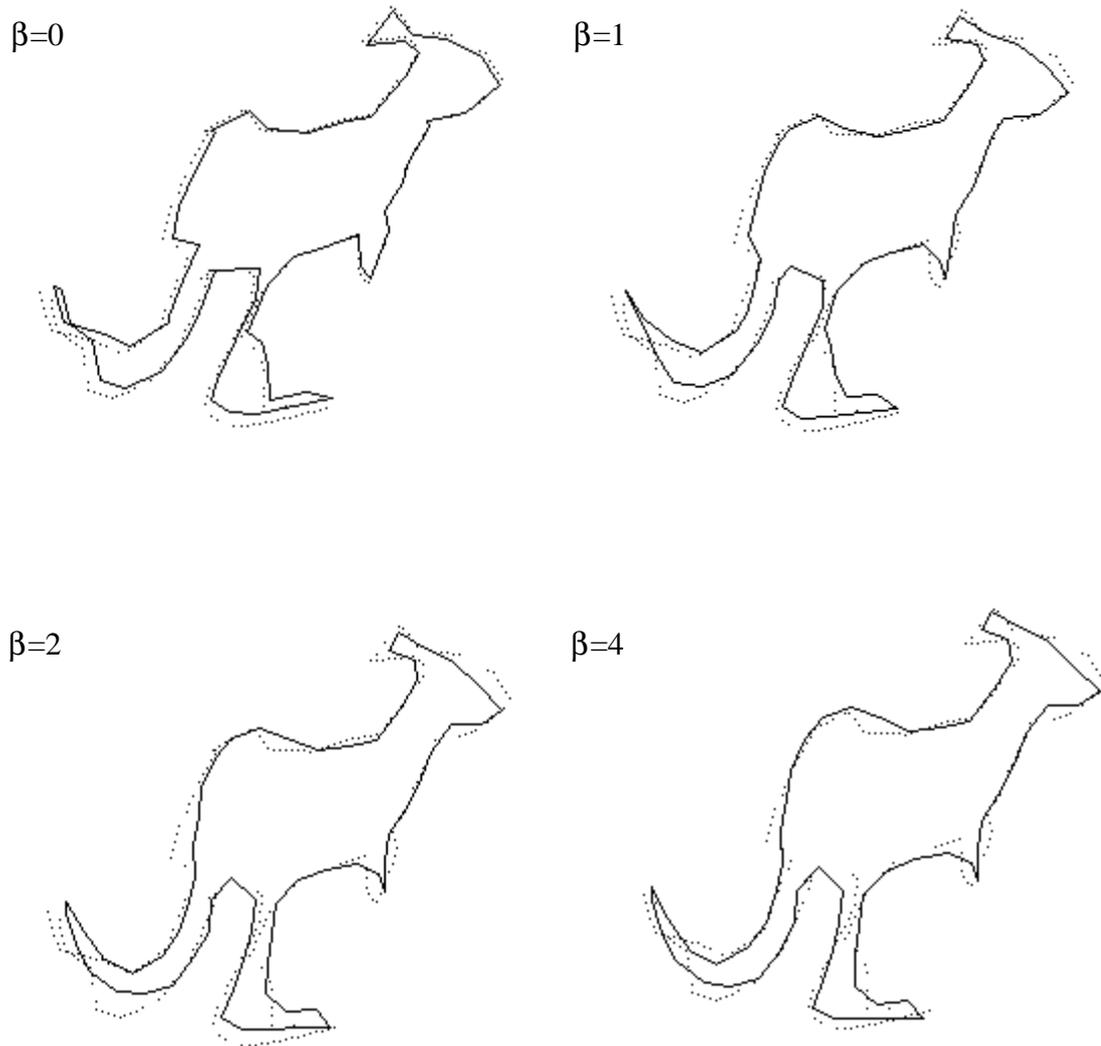


Figura 3. 12 Proceso de regularización para  $p=1$  para distintos valores de  $b$ . En línea discontinua se muestra el contorno ruidoso y el línea continua el resultado de la regularización



**Figura 3. 13** Proceso de regularización para  $p=2$  con distintos valores de  $b$ . En línea discontinua se muestra el contorno ruidoso y el línea continua el resultado de la regularización

El ruido uniforme ha sido generado de forma que el desplazamiento cuadrático medio sea igual a 1 unidad mientras que se asume que el desplazamiento cuadrático medio debido a la regularización es de 2 unidades para poder realizar una comparación independientemente del valor de  $b$ . Nótese como al incrementar el valor de  $b$  (y por tanto aportar la información de que el contorno subyacente es suave se va produciendo una eliminación sucesiva del ruido uniforme)

### 3.4 Modelos dinámicos basados en wavelets

A continuación se presentan varios modelos dinámicos basados en wavelets para la representación de la evolución temporal de la posición espacial de contornos así como de su apariencia [Pérez 2001]. En particular se introducen los modelos dinámicos autoregresivos de primer y segundo orden que serán utilizados con la distribución gaussiana generalizada como modelo de ruido estocástico y se discuten varias de sus propiedades.

#### 3.4.1 Introducción

Las herramientas presentadas en la sección 3.3 permiten el emparejamiento de una curva en una imagen estática. En esta sección se estudiarán procedimientos para el seguimiento de curvas en secuencias de imágenes (ver sección 2.4). Tal y como ocurre con imágenes individuales es de gran importancia la presencia de modelos a priori para el contorno a emparejar, sin embargo dichos modelos deben de extenderse para utilizar la *coherencia* de los movimientos típicos de contornos a lo largo del tiempo.

Un modelo estadístico adecuado debe por tanto proporcionar no sólo un modelo a priori para la forma del contorno en la primera imagen de la secuencia, similar al caso estático de la sección anterior, sino que además, debe proporcionar un modelo a priori para los posibles movimientos del contorno. Este modelo dinámico a priori se aplicará entre pares de imágenes sucesivas. Debe de tener una parte determinística, que genera el desplazamiento esperado entre imágenes sucesivas junto con un componente estocástico que modele la incertidumbre asociada con dicho desplazamiento.

Se dispondrá por tanto (ver sección 2.4) de una distribución de probabilidad que describe la variación del vector de parámetros en el tiempo. Además, suele asumirse que el conjunto de estados  $w_{1..t}$  forman una cadena de Markov, es decir

$$P(w_t | w_{1..t-1}) = P(w_t | w_{t-1}) \quad \text{Ec. 3. 108}$$

puesto que de esta forma el modelo probabilístico queda simplificado. Lo que queda por tanto es la expresión concreta de la función de probabilidad  $P(w_t | w_{t-1})$  que será estudiada en esta tesis para curvas representadas mediante wavelets.

#### 3.4.2 Modelos autoregresivos de primer orden (AR1)

La forma más simple de expresar la función de probabilidad  $P(w_t | w_{t-1})$  es mediante un modelo autoregresivo de primer orden (AR1) este modelo se pondrá de forma:

$$w_t = A w_{t-1} + D + B n_t \quad \text{Ec. 3. 109}$$

donde  $\mathbf{n}_t$  es un conjunto de vectores independientes e igualmente distribuidos de media  $\mathbf{0}$  y matriz de covarianzas igual a la identidad,  $\mathbf{A}$  y  $\mathbf{B}$  son matrices reales y  $\mathbf{D}$  un vector de desplazamiento constante.

Una forma no tan general pero de fácil interpretación es:

$$w_t - \bar{w} = \mathbf{A} (w_{t-1} - \bar{w}) + \mathbf{B}n_t \quad \text{Ec. 3. 110}$$

ambas formas se relacionan a través de :

$$(\mathbf{I}-\mathbf{A}) \bar{w} = \mathbf{D} \quad \text{Ec. 3. 111}$$

Estas definiciones nos llevan a la siguiente proposición

**Proposición 3. 12**

Dado el modelo de Ec. 3. 110 entonces:

- a)  $E(w_t) = \bar{w} + \mathbf{A} E(w_{t-1} - \bar{w})$
- b)  $\text{Var}(w_t) = \mathbf{A} \text{Var}(w_{t-1}) \mathbf{A}^T + \mathbf{B} \mathbf{B}^T$

**Demostración**

Es trivial utilizando las propiedades de medias y varianzas.

Este resultado permite calcular la media y matriz de varianzas de la distribución asintótica del proceso a través del siguiente corolario

**Corolario 3. 3**

Dado el modelo de Ec. 3. 110. entonces si existen los límites

$$\mathbf{E}_\infty = \lim_{t \rightarrow \infty} E(w_t), \quad \mathbf{V}_\infty = \lim_{t \rightarrow \infty} \text{Var}(w_t) \text{ se tiene } \mathbf{E}_\infty = \bar{w}, \quad \mathbf{V}_\infty = \mathbf{A} \mathbf{V}_\infty \mathbf{A}^T + \mathbf{B} \mathbf{B}^T$$

**Demostración**

Evidente utilizando las proposición anterior.

3.4.2.1 Modelos con innovación DGGS

Dado el empleo de la distribución gaussiana generalizada simétrica (3.3.1) en el estudio de problemas estáticos su empleo en problemas dinámicos vendrá dada a través de la modelización de la incertidumbre aleatoria en Ec. 3. 110. Se define entonces el vector  $\mathbf{n}_t$  como:

$$\mathbf{n}_t = (n'_{0,0;x}, n_{0,0;x}, n_{1,0;x}, \dots, n_{j-1,2^{j-1}-1;x}, n'_{0,0;y}, n_{0,0;y}, n_{1,0;y}, \dots, n_{j-1,2^{j-1}-1;y}) \text{ con:} \quad \text{Ec. 3. 112}$$

$$\mathbf{n}'_{0,0} = (n'_{0,0;x}, n'_{0,0;y}) \sim DGGS_2(\mathbf{0},1)$$

$$\mathbf{n}_{j,k} = (n_{j,k;x}, n_{j,k;y}) \sim DGGS_p(\mathbf{0},1)$$

y donde las variables  $\mathbf{n}_{j,k}$  son independientes.

Para este modelo se tiene:

**Proposición 3. 13**

Dado el vector definido en Ec. 3. 112 se tiene para todo  $t$   $E(\mathbf{n}_t)=\mathbf{0}$ ,  $\text{Var}(\mathbf{n}_t)=\mathbf{I}$

**Demostración**

Evidente a partir de Proposición 3. 1

Un subcaso especialmente interesante lo constituye cuando el vector de innovación  $\mathbf{n}_t$  está en el espacio de Sobolev. Si además se asume que la matriz  $\mathbf{B}$  sigue el modelo de varianzas decrecientes se tiene:

$$\mathbf{n}_t \sim N(\mathbf{0}, \mathbf{I}) \tag{Ec. 3. 113}$$

En este caso, tanto la distribución de  $w_t$  como la distribución asintótica son normales.

3.4.2.2 Aprendizaje en modelos de primer orden

A continuación se presentan dos métodos para la determinación de los parámetros en los modelos de primer orden. En primer lugar se presenta el método general para el aprendizaje en modelos separados por subespacios. Posteriormente se hace el análisis para el caso de los espacios de Sobolev.

3.4.2.2.1 Aprendizaje en modelos separados en subespacios

Se presenta en esta sección una solución parcial para el problema del aprendizaje en espacios de Besov para modelos de primer orden. Se asumirá que el parámetro  $p$  es conocido (sus valores típicos están en el intervalo [0.5, 2]) o bien se ha estimado por separado (por ejemplo mediante el método de los momentos [Mallat 1989]) y se pasará a estimar el resto de los parámetros.

Para esta estimación se utilizará el concepto de separación en subespacios. Se dispone por tanto de una serie de subespacios lineales ortogonales  $S_i$  determinados por una matriz de proyección  $\mathbf{P}_i$ . Se asume entonces:

$$\mathbf{A} = \sum_i a_i \mathbf{P}_i, \quad \mathbf{B} = \sum_i b_i \mathbf{P}_i \mathbf{S}^{1/2}, \quad a_i, b_i \in \mathbb{R} \tag{Ec. 3. 114}$$

donde se asumirá que  $\mathbf{S}$  es la matriz definida en Ec. 3. 84 con lo que el modelo dinámico se descompone en los distintos subespacios:

$$P_i W_t - P_i \bar{W} = a_i (P_i W_{t-1} - P_i \bar{W}) + b_i P_i S^{1/2} n_t$$

Ec. 3. 115

entonces se verifica:

**Proposición 3. 14**

Dado el modelo de Ec. 3. 115 entonces si  $|a_i| < 1$  se tiene

$$E_{\infty i} = P_i \bar{W}, \quad V_{\infty i} = \frac{b_i^2}{1 - a_i^2} P_i S P_i$$

donde  $E_{\infty i}$  y  $V_{\infty i}$  son la media y varianzas de la distribución límite en el subespacio.

**Demostración**

Basta comprobar que  $V_{\infty i}$  verifica la ecuación b) del Corolario 3. 3 dentro del subespacio.

A continuación se presenta un resultado que permite determinar los parámetros asociados con la parte estocástica si se posee una estimación del movimiento determinístico y el desplazamiento medio del contorno en el subespacio.

**Corolario 3. 4**

Dado el modelo dinámico de Ec. 3. 115 supongamos que existe una distribución límite dentro del subespacio  $S_i$ . Entonces para obtener un desplazamiento cuadrático medio  $\bar{r}_i^2$  dentro del subespacio  $S_i$ , el valor correspondiente de  $b_i$  debe ser:

$$b_i = \bar{r}_i \sqrt{(1 - a_i^2) \frac{N}{\text{Traza}(P_i S)}}$$

**Demostración**

Basta aplicar la Proposición 3. 14 y la Proposición 3. 7.

Uniendo los resultados para todos los subespacios se tiene:

**Teorema 3. 8**

Dado el modelo de Ec. 3. 110 entonces si  $|a_i| < 1$  para todo  $i$  se tiene:

$$\text{Traza}(V_{\infty}) = \sum_i \text{Traza}(V_{\infty i})$$

**Demostración**

Basta comprobar que la ecuación b) del Corolario 3. 3 admite una solución de la forma

$$V_{\infty} = \sum_{i,j} I_{i,j} P_i S P_j, \quad I_{i,i} = \frac{b_i^2}{1-a_i^2} \quad \text{y aplicar la propiedad } \text{Traza}(\mathbf{AB}) = \text{Traza}(\mathbf{BA})$$

La importancia de este teorema es que permite descomponer el desplazamiento cuadrático medio total como suma de los desplazamientos cuadráticos dentro de los intervalos como se indica en el siguiente corolario.

**Corolario 3. 5**

Dado el modelo de Ec. 3. 110 entonces si  $|a_i| < 1$  para todo  $i$  y además  $\bar{\mathbf{r}}_i^2$  es el desplazamiento cuadrático medio dentro del subespacio  $S_i$  y  $\bar{\mathbf{r}}^2$  es el desplazamiento cuadrático medio total se tiene:  $\bar{\mathbf{r}}^2 = \sum_i \bar{\mathbf{r}}_i^2$

**Demostración**

Es consecuencia evidente de la demostración anterior.

Un caso de interés para el cual no existe distribución límite es el caso de velocidad constante  $a_i=1$ . En esta situación puede utilizarse el siguiente teorema:

**Teorema 3. 9**

Dado el modelo dinámico señalado en Ec. 3. 110 . Si  $a_i=1$  no existe distribución límite en el subespacio  $S_i$  . Entonces dado un valor exacto para  $w_0$  el desplazamiento medio  $\bar{\mathbf{r}}_i^2(k)$  en el subespacio  $S_i$  en tiempo  $k$  verifica:

$$\bar{\mathbf{r}}_i^2(k) = b_i \sqrt{\frac{\text{Traza}(\mathbf{P}_i \mathbf{S})}{N}} \sqrt{k}$$

**Demostración**

Basta comprobar que  $\text{Var}(w_t) = k \text{Var}(w_{t-1})$

**3.4.2.2.2 Aprendizaje en espacios de Sobolev**

Los espacios de Sobolev tienen la ventaja de que las distribuciones son gaussianas. En este caso se pasa a una estimación por máxima verosimilitud cuyas ecuaciones dado un conjunto de muestras  $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ , son:

$$\begin{aligned}
 \hat{\mathbf{A}} &= \mathbf{R}'_{01} (\mathbf{R}'_{11})^{-1} & \hat{S}_{TR}^2 &= \frac{\hat{c}_{00} + \hat{c}_{NN}}{2} & \text{Ec. 3. 116} \\
 \hat{\mathbf{W}} &= (\mathbf{I} - \hat{\mathbf{A}}) \hat{\mathbf{D}} \\
 \frac{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} (i-1) p_i}{\sum_{i=1}^{\log_2 n} 2^{2(i-1)\hat{b}} p_i} &= \frac{N(\log_2 N - 2) + 2}{(N-1)} & \hat{S}_{DEF}^2 &= \frac{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} p_i}{2(N-1)}
 \end{aligned}$$

donde:

$$\begin{aligned}
 \mathbf{R}_i &= \sum_{k=2}^m w_{k-i}, \quad \mathbf{R}_{ij} = \sum_{k=2}^m w_{k-i} w_{k-j}^T, \quad \mathbf{R}'_{ij} = \mathbf{R}_{ij} - \frac{1}{m-1} \mathbf{R}_i \mathbf{R}_j^T & \text{Ec. 3. 117} \\
 \hat{\mathbf{D}} &= \frac{1}{m-1} (\mathbf{R}_0 - \hat{\mathbf{A}}_1 \mathbf{R}_1) \\
 \hat{\mathbf{C}} &= \frac{1}{m-1} (\mathbf{R}_{00} - \hat{\mathbf{A}}_1 \mathbf{R}_{10} - \hat{\mathbf{D}} \mathbf{R}_0^T) \quad p_i = \sum_{j=2^{i-1}}^{2^i-1} \hat{c}_{ij} + \hat{c}_{j+n, j+n}
 \end{aligned}$$

Todas las ecuaciones se resuelven fácilmente excepto la ecuación en  $\mathbf{b}$  que tiene que ser resuelta numéricamente. Este resultado es una adaptación de [Blake 1998].

### 3.4.2.3 Limitaciones de los modelos de primer orden

Los modelos de primer orden, aunque poseen algunos de los requerimientos de los modelos dinámicos, tienen importantes limitaciones.

En primer lugar, el modelo de Ec. 3. 110 aunque representa un movimiento direccional con ruido, tiene una dirección media que es fija y debe ser conocida a priori. Un buen modelo, sin embargo debería utilizar una distribución a priori que favoreciera un movimiento suave de velocidad aproximadamente constante pero con dirección arbitraria y que permitiera que la velocidad cambiara lentamente en magnitud y dirección. Un segundo requerimiento de un modelo dinámico es que sea capaz de modelar oscilaciones que ocurren en muchos modelos dinámicos de interés. Sin embargo los modelos de primer orden no tienen componentes oscilatorios por lo que no pueden representar los movimientos oscilatorios de forma adecuada. Por tanto es necesario utilizar modelos de mayor complejidad y ésta es la razón por la cual se introducen los modelos dinámicos autoregresivos de segundo orden.

### 3.4.3 Modelos autoregresivos de segundo orden (AR2)

Un modelo autoregresivo de segundo orden (AR2) es una extensión natural del proceso de primer orden donde ahora:  $P(w_t | w_{1..t-1}) = P(w_t | w_{t-1}, w_{t-2})$  y

$$w_t - \bar{w} = A_1 (w_{t-1} - \bar{w}) + A_2 (w_{t-2} - \bar{w}) + Bn_t \quad \text{Ec. 3. 118}$$

donde  $n_t$  es un conjunto de vectores independientes e igualmente distribuidos de media  $\mathbf{0}$  y matriz de covarianzas igual a la identidad,  $A_1$ ,  $A_2$  y  $B$  son matrices reales

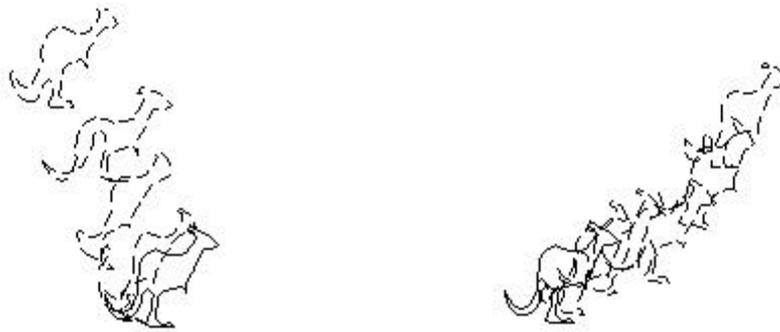


Figura 3. 14 Dos simulaciones (contornos en línea discontinua) del modelo AR2 sobre un contorno de prueba (en línea continua). Nótese como el modelo predice además de la posición espacial del contorno su apariencia debida a la deformación.

Este modelo admite la representación:

$$\tilde{w}_t - \tilde{w} = \tilde{A}(\tilde{w}_{t-1} - \tilde{w}) + \tilde{B}\tilde{n}_t \quad \text{con} \quad \text{Ec. 3. 119}$$

$$\tilde{w}_t = \begin{pmatrix} w_{t-1} \\ w_t \end{pmatrix}, \quad \tilde{w} = \begin{pmatrix} \bar{w} \\ \bar{w} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ A_2 & A_1 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} \mathbf{0} \\ B \end{pmatrix}$$

### Proposición 3. 15

Dado el modelo de Ec. 3. 118 entonces:

- $E(\tilde{w}_t) = \tilde{w} + \tilde{A}E(\tilde{w}_{t-1} - \tilde{w})$
- $\text{Var}(\tilde{w}_t) = \begin{pmatrix} V_{t-1,t-1} & V_{t,t-1}^T \\ V_{t,t-1} & V_{t,t} \end{pmatrix}, \quad V_{t,t} = \text{Var}(w_t), \quad V_{t,t-1} = E((w_t - \bar{w}_t)(w_{t-1} - \bar{w}_{t-1})^T)$
- $\text{Var}(\tilde{w}_t) = \tilde{A} \text{Var}(\tilde{w}_{t-1}) \tilde{A}^T + \tilde{B}\tilde{B}^T$

### Demostración

Es similar a la demostración de Proposición 3. 12

**Proposición 3. 16**

Dado el modelo de Ec. 3. 118 entonces si existen los límites

$$\mathbf{E}_\infty = \lim_{t \rightarrow \infty} \mathbf{E}(w_t), \quad \mathbf{V}_\infty = \lim_{t \rightarrow \infty} \text{Var}(w_t) \text{ se tiene:}$$

- a)  $\mathbf{E}_\infty = \bar{w}$
- b)  $\mathbf{V}_\infty = \mathbf{A}_2 \mathbf{V}_\infty \mathbf{A}_2^T + \mathbf{A}_1 \mathbf{V}_\infty \mathbf{A}_1^T + \mathbf{A}_1 \mathbf{V}_\infty^{21} \mathbf{A}_2^T + (\mathbf{A}_1 \mathbf{V}_\infty^{21} \mathbf{A}_2^T)^T + \mathbf{B} \mathbf{B}^T$  con  
 $\mathbf{V}_\infty^{21} = \mathbf{A}_2 (\mathbf{V}_\infty^{21})^T + \mathbf{A}_1 \mathbf{V}_\infty$

**Demostración**

Evidente utilizando la proposición anterior.

3.4.3.1 Modelos con innovación DGGS

De forma análoga al caso del modelo AR1 se utilizarán en el modelo AR2 variables aleatorias con distribución DGGS. Se define entonces el vector  $\tilde{\mathbf{n}}_t$  como:

$$\tilde{\mathbf{n}}_t = (\mathbf{m}_t, \mathbf{n}_t) \tag{Ec. 3. 120}$$

donde  $\mathbf{m}_t$  y  $\mathbf{n}_t$  se definen como en Ec. 3. 112 siendo sus componentes  $m_{j,k}$  y  $n_{j,k}$  variables independientes.

De nuevo un subcaso especialmente interesante lo constituye cuando el vector de innovación  $\mathbf{n}_t$  está en el espacio de Sobolev. Si además se asume el modelo de varianzas decrecientes se tiene:

$$\mathbf{n}_t \sim N(\mathbf{0}, \mathbf{I}) \tag{Ec. 3. 121}$$

En este caso, tanto la distribución de  $w_t$  como la distribución asintótica son normales.

3.4.3.2 Aprendizaje en modelos de segundo orden

A continuación se presentan dos métodos para la determinación de los parámetros en los modelos de segundo orden. En primer lugar se presenta el método general para el aprendizaje en modelos separados por subespacios. Posteriormente se hace el análisis para el caso de los espacios de Sobolev.

3.4.3.2.1 Aprendizaje en modelos separados en subespacios

En esta sección se presenta una solución parcial para el problema del aprendizaje en espacios de Besov para modelos de segundo orden. De nuevo se asumirá que el parámetro  $p$  es conocido (sus valores típicos están en el intervalo [0.5, 2]) o bien se han

estimado por separado (por ejemplo mediante el método de los momentos [Mallat 1989] y se pasará a estimar el resto de los parámetros. Para esta estimación se utilizará el concepto de separación en subespacios (Sección 3.4.2.2). Se dispone por tanto de una serie de subespacios lineales ortogonales  $S_i$  determinados por una matriz de proyección  $P_i$ . Se asume entonces:

$$A_1 = \sum_i a_{1,i} P_i, \quad A_2 = \sum_i a_{2,i} P_i, \quad B = \sum_i b_i P_i S^{1/2}, \quad a_{1,i}, a_{2,i}, b_i \in \mathbb{R} \quad \text{Ec. 3. 122}$$

donde se asumirá que  $S$  es la matriz definida en Ec. 3. 84 con lo que el modelo dinámico se descompone en los distintos subespacios determinados por las matrices de proyección  $P_i$ :

$$P_i W_t - P_i \bar{W} = a_{1,i} (P_i W_{t-1} - P_i \bar{W}) + a_{2,i} (P_i W_{t-2} - P_i \bar{W}) + b_i P_i S^{1/2} n_t \quad \text{Ec. 3. 123}$$

entonces se verifica:

### Proposición 3.17

Dado el modelo de Ec. 3. 123 entonces si  $\frac{(1 - a_{2,i})}{(1 + a_{2,i})(1 - a_{1,i} - a_{2,i})(1 + a_{1,i} - a_{2,i})} \geq 0$  se tiene

$$E_{\infty i} = P_i \bar{W}, \quad V_{\infty i} = \frac{b_i^2 (1 - a_{2,i})}{(1 + a_{2,i})(1 - a_{1,i} - a_{2,i})(1 + a_{1,i} - a_{2,i})} P_i S P_i$$

donde  $E_{\infty i}$  y  $V_{\infty i}$  son la media y varianzas de la distribución límite en el subespacio determinado por  $P_i$ .

### Demostración

Basta comprobar que  $V_{\infty i}$  verifica la ecuación b) en Proposición 3. 16 dentro del subespacio.

### Teorema 3. 10

Dado el modelo dinámico de Ec. 3. 123 supongamos que existe una distribución límite dentro del subespacio  $S_i$ . Entonces para obtener un desplazamiento cuadrático medio  $\bar{r}_i$ , el valor correspondiente de  $b_i$  debe ser:

$$b_i = \bar{r}_i \sqrt{\frac{(1 + a_{2,i})(1 - a_{1,i} - a_{2,i})(1 + a_{1,i} - a_{2,i})}{(1 - a_{2,i})} \frac{N}{\text{Traza}(P_i S)}}$$

**Demostración**

Basta aplicar Proposición 3.17 y Proposición 3. 10.

Este resultado nos permite determinar los parámetros asociados con la parte estocástica si se posee una estimación del movimiento determinístico y el desplazamiento medio del contorno en el subespacio.

**Teorema 3. 11**

Dado el modelo de Ec. 3. 118 y Ec. 3. 123 entonces si se cumple para todo  $i$  :

$$\frac{(1 - a_{2,i})}{(1 + a_{2,i})(1 - a_{1,i} - a_{2,i})(1 + a_{1,i} - a_{2,i})} \geq 0 \text{ se tiene:}$$

$$\text{Traza}(\mathbf{V}_\infty) = \sum_i \text{Traza}(\mathbf{V}_{\infty i})$$

**Demostración**

Basta comprobar que la ecuación b) en Proposición 3. 16 admite una solución de la

forma  $\mathbf{V}_\infty = \sum_{i,j} \mathbf{I}_{i,j} \mathbf{P}_i \mathbf{S} \mathbf{P}_j$ ,  $\mathbf{I}_{i,i} = \frac{b_i^2 (1 - a_{2,i})}{(1 + a_{2,i})(1 - a_{1,i} - a_{2,i})(1 + a_{1,i} - a_{2,i})}$  y aplicar la propiedad  $\text{Traza}(\mathbf{AB}) = \text{Traza}(\mathbf{BA})$

La importancia de este teorema radica en que tal y como ocurre en el caso de los modelos de primer orden permite descomponer el desplazamiento cuadrático medio total como suma de los desplazamientos cuadráticos dentro de los subespacios como se indica en el siguiente corolario.

**Corolario 3. 6**

Dado el modelo de Ec. 3. 118 y Ec. 3. 123 entonces si para todo  $i$  se verifica :

$$\frac{(1 - a_{2,i})}{(1 + a_{2,i})(1 - a_{1,i} - a_{2,i})(1 + a_{1,i} - a_{2,i})} \geq 0$$

y además  $\bar{\mathbf{r}}_i^2$  es el desplazamiento cuadrático medio dentro del subespacio  $S_i$  y  $\bar{\mathbf{r}}^2$  es el desplazamiento cuadrático medio total se tiene:

$$\bar{\mathbf{r}}^2 = \sum_i \bar{\mathbf{r}}_i^2$$

**Demostración**

Es consecuencia evidente de la demostración anterior.

Un caso de interés para el cual no existe distribución límite es el caso de velocidad constante  $a_{2,i}=-1$  y  $a_{1,i}=2$ . En esta situación puede utilizarse el siguiente teorema:

### Teorema 3. 12

Dado el modelo dinámico señalado en Ec. 3. 123. Si  $a_{2,i}=-1$  y  $a_{1,i}=2$  no existe distribución límite en el subespacio  $S_i$ . Entonces dado un valor exacto para  $W_0$  el desplazamiento medio  $\bar{r}_i^2(k)$  en el subespacio  $S_i$  en tiempo  $k$  verifica:

$$\bar{r}_i(k) \approx b_i \sqrt{\frac{\text{Traza}(\mathbf{P}_i \mathbf{S})}{3N}} \sqrt{k^3}$$

### Demostración

La demostración es similar a la del Teorema 3. 9

#### 3.4.3.2.2 Aprendizaje en espacios de Sobolev

Los espacios de Sobolev tienen la ventaja de que las distribuciones son gaussianas. En este caso se pasa a una estimación por máxima verosimilitud cuyas ecuaciones son:

$$\begin{aligned} \hat{\mathbf{A}}_2 &= (\mathbf{R}'_{02} - \mathbf{R}'_{01} (\mathbf{R}'_{11})^{-1} \mathbf{R}'_{12}) (\mathbf{R}'_{22} - \mathbf{R}'_{21} (\mathbf{R}'_{11})^{-1} \mathbf{R}'_{12})^{-1} & \text{Ec. 3. 124} \\ \hat{\mathbf{A}}_1 &= (\mathbf{R}'_{01} - \hat{\mathbf{A}}_2 \mathbf{R}'_{21}) (\mathbf{R}'_{11})^{-1} \\ \hat{\mathbf{W}} &= (\mathbf{I} - \hat{\mathbf{A}}_2 - \hat{\mathbf{A}}_1)^{-1} \hat{\mathbf{D}} \end{aligned} \quad \hat{\mathbf{S}}_{TR}^2 = \frac{\hat{c}_{00} + \hat{c}_{NN}}{2}$$

$$\frac{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} (i-1) p_i}{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} p_i} = \frac{N(\log_2 N - 2) + 2}{(N-1)} \quad \hat{\mathbf{S}}_{DEF}^2 = \frac{\sum_{i=1}^{\log_2 N} 2^{2(i-1)\hat{b}} p_i}{2(N-1)}$$

donde:

$$\begin{aligned} \mathbf{R}_i &= \sum_{k=3}^m w_{k-i}, \quad \mathbf{R}_{ij} = \sum_{k=3}^m w_{k-i} w_{k-j}^T, \quad \mathbf{R}'_{ij} = \mathbf{R}_{ij} - \frac{1}{m-2} \mathbf{R}_i \mathbf{R}_j^T & \text{Ec. 3. 125} \\ \hat{\mathbf{D}} &= \frac{1}{m-2} (\mathbf{R}_0 - \hat{\mathbf{A}}_2 \mathbf{R}_2 - \hat{\mathbf{A}}_1 \mathbf{R}_1) \\ \hat{\mathbf{C}} &= \frac{1}{m-2} (\mathbf{R}_{00} - \hat{\mathbf{A}}_2 \mathbf{R}_{20} - \hat{\mathbf{A}}_1 \mathbf{R}_{10} - \hat{\mathbf{D}} \mathbf{R}_0^T) \quad p_i = \sum_{j=2^{i-1}}^{2^i-1} \hat{c}_{jj} + \hat{c}_{j+n \ j+n} \end{aligned}$$

Todas las ecuaciones se resuelven fácilmente excepto la ecuación en  **$b$**  que tiene que ser resuelta numéricamente . Este resultado es una adaptación de [Blake 1998].

A continuación en el siguiente capítulo se presentan algunas aplicaciones obtenidas con esta representación y que son aplicación directa de estos resultados teóricos.

## 4 Aplicaciones

En este capítulo se mostrarán algunas aplicaciones de las representaciones de contornos basadas en wavelets y multiwavelets a problemas de reconocimiento y seguimiento. En primer lugar, se mostrarán las aplicaciones de los resultados teóricos en espacios de deformación desarrollados en la sección 3.3 y los resultados teóricos obtenidos para modelos dinámicos en la sección 3.4 como técnica para resolver el problema del seguimiento de objetos. Para ello se mostrarán los resultados de seguimiento en espacios de Sobolev utilizando el filtro de Kalman y en espacios generales utilizando el filtro Condensation. Finalmente se muestra una aplicación de las representaciones de curvas basadas en multiwavelets para problemas de reconocimiento bajo transformaciones afines y se muestra un algoritmo de emparejamiento basado en esta representación, que es capaz de realizar el reconocimiento de utilizando información local y permitiendo por tanto soportar oclusiones parciales de los contornos.

### 4.1 Seguimiento de contornos

En la sección 3.4.3 los modelos dinámicos se caracterizaron mediante un modelo autoregresivo de segundo orden que proporcionan una distribución de probabilidad  $P(w_t)$ , que evoluciona temporalmente y que representa una distribución a priori para el vector  $w$  en cada instante de tiempo  $t$ . En esta sección esta información a priori generada por el modelo dinámico se unirá con la proveniente de la imagen tal y como se presentó en la sección 2.4. El resultado es la fusión de la información a priori del objeto con la información obtenida a partir de las observaciones generadas a partir de la imagen tal y como se realizó en la sección 3.2 para imágenes estáticas, pero extendido ahora a secuencias de imágenes.

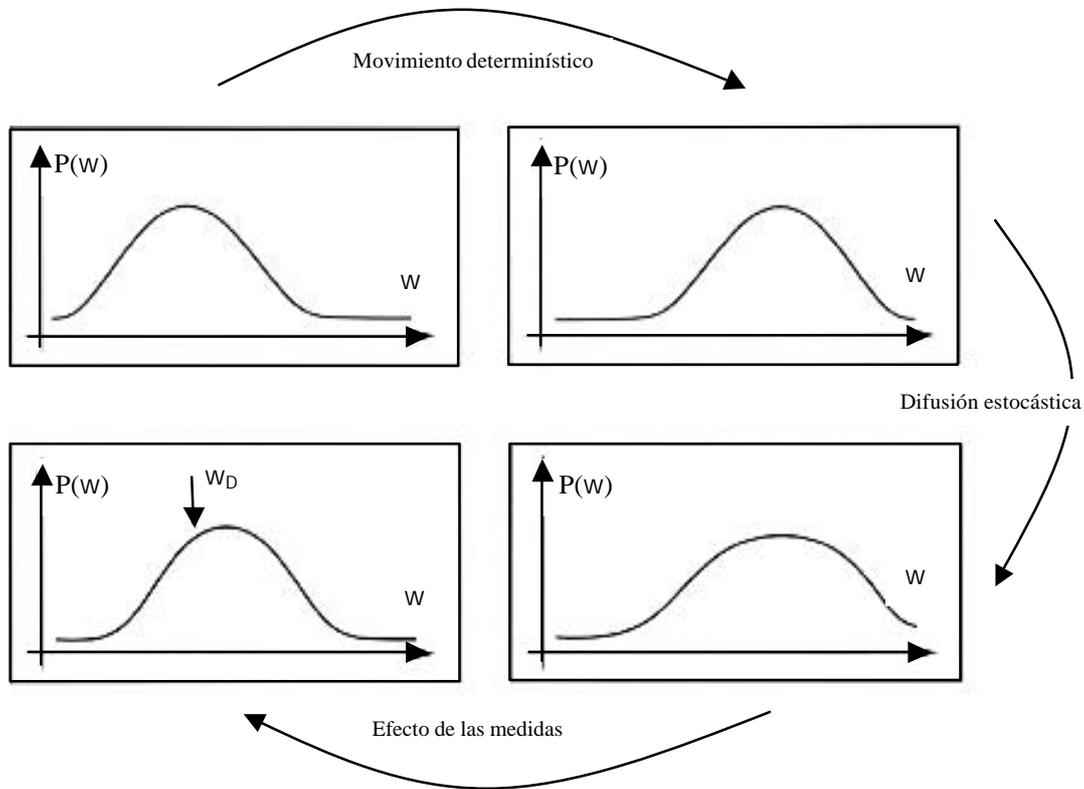
#### 4.1.1 Seguimiento de contornos mediante modelos en el espacio de Sobolev

Se presenta a continuación una aproximación al seguimiento de contornos mediante modelos a priori en espacios de Sobolev lo que lleva aparejado la utilización de la distribución probabilística gaussiana y el filtro de Kalman. En esta sección se presenta el algoritmo de seguimiento utilizando los modelos dinámicos basados en wavelets en espacios de Sobolev y se discuten algunos ejemplos de su aplicación.

##### 4.1.1.1 Introducción

El mecanismo natural para esta fusión de información generada por el modelo dinámico cuando las distribuciones son gaussianas tal y como ocurre en el espacio de Sobolev (sección 3.4.3.1) es el filtro de Kalman. Este filtro, presentado en la sección

2.4.1 es una herramienta de amplia utilización en problemas de seguimiento (sección 2.4.3). El filtro de Kalman permite determinar la evolución de la función de densidad gaussiana para el vector  $w$  que representa el contorno tal y como se ilustra en la Figura 4.1.



**Figura 4.1 Filtro de Kalman como proceso de propagación de densidad [Blake 1998]. En el caso de una distribución a priori gaussiana junto con un proceso estocástico y densidad de observación también gaussiano y asumiendo una dinámica lineal, la densidad del estado tiene una distribución gaussiana.**

#### 4.1.1.2 El filtro de Kalman

El filtro de Kalman es la forma de combinar el modelo de predicción dinámica con las observaciones obtenidas de la imagen en el caso gaussiano y cuando el modelo dinámico es lineal. Para ello necesita que el modelo dinámico y de medidas cumpla:

$$w_t | w_{t-1} \sim N(w_{t-1}, U_t), \quad D_t | w_t \sim N(w_t, V_t) \quad \text{Ec. 4.1}$$

a continuación se proporciona la forma exacta de ambas distribuciones para el caso de modelos dinámicos en espacios de Sobolev.

#### 4.1.1.2.1 El modelo dinámico

En la sección 3.4.2.3 ya se discutieron las limitaciones de los modelos autoregresivos de primer orden y por esa razón se introdujeron los modelos autoregresivos de segundo orden en la sección 3.4.3. En esta sección se asumirá por tanto que el modelo dinámico se puede poner como un modelo autoregresivo de segundo orden. Para poner el modelo dinámico en la forma de Ec 3.118, únicamente debe observarse la ecuación 3.119 y trabajar con el vector de estado extendido  $\tilde{w}_t = (w_{t-1}, w_t)$  de forma que el modelo autoregresivo de segundo orden quede escrito como:

$$\tilde{w}_t - \bar{w} = \tilde{A}(\tilde{w}_{t-1} - \bar{w}) + \tilde{B}\tilde{n}_t \quad \text{Ver Ec. 3.119}$$

$$\tilde{w}_t = \begin{pmatrix} w_{t-1} \\ w_t \end{pmatrix}, \quad \bar{w} = \begin{pmatrix} \bar{w} \\ \bar{w} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ A_2 & A_1 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} \mathbf{0} \\ B \end{pmatrix}$$

donde  $\tilde{B}$  está definido en función de la matriz  $S$  de la sección 3.3.3.1.1 de forma que la innovación estocástica pertenezca al espacio de Sobolev y

$$\tilde{w}_t - \bar{w} \mid \tilde{w}_{t-1} - \bar{w} \sim N(\tilde{A}(\tilde{w}_{t-1} - \bar{w}), \tilde{B}\tilde{B}^T) \quad \text{y} \quad G_t = \tilde{A}, \quad U_t = \tilde{B}\tilde{B}^T \quad \text{Ec. 4.2}$$

con lo que los modelos dinámicos propuestos en la sección 3.4.3 son por tanto empleados en el filtro de Kalman.

Para finalizar la descripción del filtro entonces solo queda determinar la forma del modelo de observaciones

#### 4.1.1.2.2 El modelo de observaciones

En la sección 3.3.2 se definió el ajuste entre modelo y contorno como:

$$E_{aju}(\mathbf{r}, \bar{\mathbf{r}}) = \|\mathbf{r} - \mathbf{r}_D\|_2^2 \quad \text{Ver Ec. 3.81}$$

donde  $\mathbf{r}_D$  es una curva extraída de la imagen y que constituye el conjunto de datos  $D$ . En la sección 3.3.3 se vio que este modelo llevaba a la distribución:

$$P(\mathbf{r}_D \mid \mathbf{r}) \propto \exp\left(\frac{-1}{2\mathbf{S}_D^2} \|\mathbf{r} - \mathbf{r}_D\|_2^2\right) \quad \text{Ver Ec. 3.87}$$

o de forma equivalente:

$$P(\mathbf{r}_D - \bar{\mathbf{r}} \mid \mathbf{w} - \bar{\mathbf{w}}) \propto \exp\left(\frac{-1}{2\mathbf{S}_D^2} \|\mathbf{r}_D - \bar{\mathbf{r}} - \mathbf{F}_w(\mathbf{w} - \bar{\mathbf{w}})\|_2^2\right) \approx \text{Ec. 4.3}$$

$$\exp\left(\frac{-1}{2\mathbf{S}_D^2} \sum_{i=1}^M \|(r_{D_i} - \bar{r})(u_i) - F_w(u_i)(\mathbf{w} - \bar{\mathbf{w}})\|_2^2\right)$$

donde  $\mathbf{F}_w$  es la matriz definida en la sección 3.2.6 .

Entonces si  $\mathbf{r}_D(u) = (x_D(u), y_D(u))^T$ ,  $\bar{\mathbf{r}}(u) = (\bar{x}(u), \bar{y}(u))^T$  se tomará un conjunto de  $M$  observaciones de la imagen que se pondrán en el vector de datos  $D_i$  de la siguiente forma:

$$D_i = ((x_{D_i} - \bar{x})(u_1), \dots, (x_{D_i} - \bar{x})(u_M), (y_{D_i} - \bar{y})(u_1), \dots, (y_{D_i} - \bar{y})(u_M))^T \quad \text{Ec. 4.4}$$

y se define la matriz  $\mathbf{C}$  como:

$$\mathbf{C} = \mathbf{D}\mathbf{Q}\mathbf{H} \text{ con} \quad \text{Ec. 4.5}$$

$$\mathbf{Q} = \mathbf{I}_2 \otimes (\mathbf{G}_w(u_1), \mathbf{G}_w(u_2), \dots, \mathbf{G}_w(u_M))^T, \mathbf{H} = (\mathbf{0}_{2N \times 2N}, \mathbf{I}_{2N})$$

$$\mathbf{D} = \mathbf{I}_2 \otimes (\mathbf{v}_1, \dots, \mathbf{v}_M)^T \text{ y } \mathbf{v}_i = (0, \dots, 0, 1, 0, \dots, 0)^T$$

entonces el modelo de observación queda como:

$$D_i \mid \tilde{\mathbf{w}}_i - \bar{\mathbf{w}} \sim N(\mathbf{C}\mathbf{w}_i, \mathbf{L}), \mathbf{L} = \frac{1}{\mathbf{S}_D^2} \mathbf{I}_M \equiv D_i \mid \mathbf{w}_i - \tilde{\mathbf{w}} \quad \text{Ec. 4.6}$$

con lo que las matrices de filtro de Kalman toman la forma:  $\mathbf{F}_i = \mathbf{C}$ ,  $\mathbf{V}_i = \mathbf{L}$  y el problema de seguimiento queda definido en el marco del filtro de Kalman.

De esta forma, con este modelo definido anteriormente, el filtro de Kalman de la sección 2.4.1.1.2 se transforma en:

<b>Predicción</b>	<b>Predicción</b>
$\hat{\mathbf{w}}_i^- = \mathbf{G}_i \hat{\mathbf{w}}_{i-1}$	Calcular $\tilde{\mathbf{w}}_i - \bar{\mathbf{w}} = \tilde{\mathbf{A}}(\tilde{\mathbf{w}}_{i-1} - \bar{\mathbf{w}})$ , es decir: $\hat{\mathbf{w}}_i^- = \bar{\mathbf{w}} + \mathbf{A}_1(\hat{\mathbf{w}}_{i-1} - \bar{\mathbf{w}}) + \mathbf{A}_2(\hat{\mathbf{w}}_{i-2} - \bar{\mathbf{w}})$ $\hat{\mathbf{w}}_{i-1}^- = \hat{\mathbf{w}}_{i-1}$
$\mathbf{R}_i = \mathbf{G}_i \mathbf{S}_{i-1} \mathbf{G}_i^T + \mathbf{U}_i$	$\mathbf{R}_i'' = \mathbf{S}_{i-1}$ $\mathbf{R}_i' = \mathbf{A}_2 \mathbf{S}_{i-1}^T + \mathbf{A}_1 \mathbf{S}_{i-1}$ $\mathbf{R}_i = \mathbf{A}_2 \mathbf{S}_{i-1}'' \mathbf{A}_2^T + \mathbf{A}_1 \mathbf{S}_{i-1}' \mathbf{A}_2^T + \mathbf{A}_2 \mathbf{S}_{i-1}' \mathbf{A}_1^T + \mathbf{A}_1 \mathbf{S}_{i-1} \mathbf{A}_1^T + \mathbf{B}_0 \mathbf{B}_0^T$
<b>Obtención de <math>D_i</math></b>	<b>Obtención de <math>D_i</math></b>
<b>Asimilación</b>	<b>Asimilación</b>

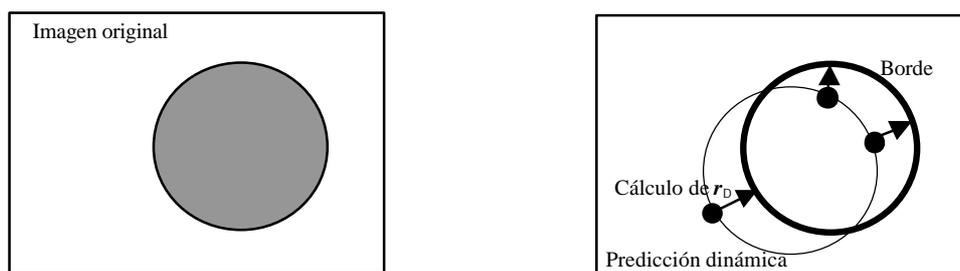
$K_t = R_t F_t^T (V_t + F_t R_t F_t^T)^{-1}$	$K'_t = R'_t (DQ)^T (L + DQR_t (DQ)^T)^{-1}$ $K_t = R_t (DQ)^T (L + DQR_t (DQ)^T)^{-1}$
$S_t = R_t - K_t F_t R_t$	$S''_t = R''_t - K'_t DQR'_t$ $S'_t = R'_t - K_t DQR'_t$ $S_t = R_t - K_t DQR_t$
$e_t = D_t - F_t G_t \hat{w}_{t-1}$	$e_t = D_t - DQ \hat{w}_t^-$
$\hat{w}_t = G_t \hat{w}_{t-1} + K_t e_t$	$\hat{w}_{t-1} = \hat{w}_t^- + K'_t e_t$ $\hat{w}_t = \hat{w}_t^- + K_t e_t$

**Tabla 4.1 Filtro de Kalman para el problema dinámico**

A continuación se pasa a determinar el vector  $D_t$  a partir de la imagen.

#### 4.1.1.2.2.1 El cálculo de $D_t$ : El caso ideal

Para acabar de desarrollar el filtro de Kalman es necesario definir la forma de calcular  $r_D$  a partir de los datos de la imagen. Para ello, se calcula la predicción del modelo dinámico  $\hat{w}_t^- = G_t \hat{w}_{t-1}$  (Ver Tabla 4.1), se calcula la curva que define como  $\hat{r}_t^- = F_t \hat{w}_t^-$  y, en el caso ideal, se asume que en un entorno de los puntos  $\hat{r}_t^-(u_1), \hat{r}_t^-(u_2), \dots, \hat{r}_t^-(u_M)$  es posible encontrar el punto correspondiente a cada elemento del contorno mediante alguna regla como por ejemplo el punto más cercano en la imagen de bordes (ver Figura 4.2)



**Figura 4.2 (Izquierda) Imagen de ejemplo. (Derecha) Borde de la imagen en línea gruesa, predicción del contorno por el filtro de Kalman en línea fina y cálculo de  $r_D$  a partir de la predicción señalado mediante flechas.**

Para acelerar el proceso de búsqueda de  $r_D$  puede utilizarse la verosimilitud de los datos a priori (Ec 1.85)

$$P(D_t | D_{1:t-1}) = N(F_t \hat{w}_{t-1}, V_t + F_t (G_t S_{t-1} G_t^T) F_t^T) = N(\hat{w}_t^-, L + DQR_t (DQ)^T) \quad \text{Ec. 4.7}$$

que permite calcular una región de confianza para la curva de datos en el plano a partir de la distribución de los datos ya encontrados en la secuencia y buscar únicamente en dicha región.

4.1.1.2.2 Observación con ruido

Cuando hay ruido en la imagen o bien ocultación por otros objetos es inevitable que el proceso de búsqueda definido en la sección anterior falle para algunos puntos del contorno predicho  $\hat{r}_t^-(u_1), \hat{r}_t^-(u_2), \dots, \hat{r}_t^-(u_M)$  al no ser posible determinar el elemento correspondiente para cada elemento del contorno predicho (ver Figura 4.3)

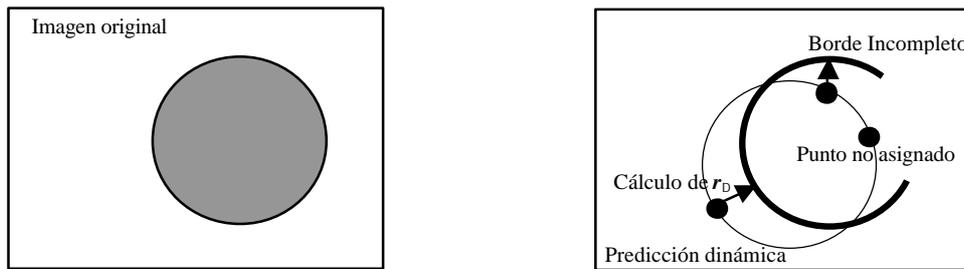


Figura 4.3 (Izquierda) Imagen de ejemplo. (Derecha) Borde incompleto de la imagen en línea gruesa, predicción del contorno por el filtro de Kalman en línea fina y cálculo de  $r_D$  a partir de la predicción señalado mediante flechas con un elemento del contorno sin asignación.

En ese caso, el algoritmo se modifica de la siguiente forma: en cada iteración  $t$  habrá un subconjunto de índices de forma que para los puntos del conjunto  $\{\hat{r}_t^-(u_{i1}), \hat{r}_t^-(u_{i2}), \dots, \hat{r}_t^-(u_{iM(t)})\} \subset \{\hat{r}_t^-(u_1), \hat{r}_t^-(u_2), \dots, \hat{r}_t^-(u_M)\}$  es posible encontrar correspondencia. En ese caso basta redefinir:

$$D_t = (v_{t1}, \dots, v_{tM(t)})^T \text{ y } L = \frac{1}{S_D^2} I_{M(t)} \quad \text{Ec. 4.8}$$

en el caso de que no sea posible encontrar ninguna correspondencia el proceso de observación falla por completo en el instante  $t$ . En ese caso, basta con definir  $K_t = \mathbf{0}$  y  $e_t = \mathbf{0}$  y continuar con el filtro en la siguiente iteración.

## 4.1.1.2.2.3 Búsquedas a través de la normal

Uno de los problemas de medir la diferencia de curvas utilizando normas es la falta de invarianza frente a cambio de parametrización. Dos curvas similares pueden tener una diferencia sustancial salvo que las parametrizaciones de las curvas sean similares. Esto es un problema especialmente cuando la curva de datos  $r_D$  está obtenida de la imagen utilizando las técnicas de la sección anterior, puesto que su parametrización se hereda de la curva predicha por el modelo dinámico. Una solución a este problema es redefinir el proceso de ajuste de forma que se tenga en cuenta el efecto del cambio de parametrización. No obstante, el problema del ajuste óptimo teniendo en cuenta la parametrización es muy complejo. Una solución aproximada consiste en definir [Blake 1998] :

$$E_{aju}(r, \bar{r}) = \|r - r_D\|_{n(r)} = \int ((r(u) - r_D(u)) \cdot n(u))^2 du \quad \text{Ec. 4.9}$$

donde  $n(u)$  es el vector normal a la curva  $r$ . Por tanto del vector diferencia entre ambas curvas se calcula la norma de la proyección sobre el vector normal. (Ver Figura 4.4)

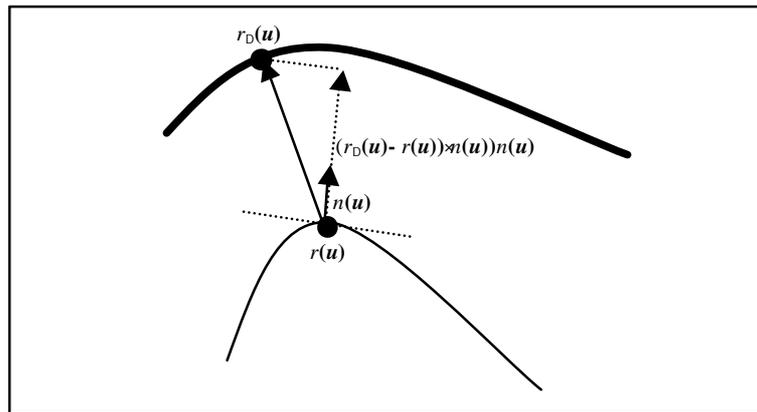


Figura 4.4 Desplazamiento normal

Esta invarianza aproximada es la base para que la búsqueda de la sección anterior se lleve a cabo siguiendo las normales de la curva predicha por el modelo dinámico. Estas normales son calculadas mediante las técnicas de derivación de curvas basadas en wavelets de la sección 3.1.12.

Con la introducción del desplazamiento normal la distribución del vector de observaciones queda definido como:

$$D_t = N_t d_t \quad \text{Ec. 4.10}$$

$$d_t = ((x_D - \bar{x})(u_{t1}), \dots, (x_D - \bar{x})(u_{tM(t)}), (y_D - \bar{y})(u_{t1}), \dots, (y_D - \bar{y})(u_{tM(t)}))$$

donde si se  $n(u) = (n_x(u), n_y(u))^T$  es el vector normal a la curva predicha por el modelo dinámico, se define la matriz  $N_t$  como:

$$\begin{aligned}
 N_t &= \left( \text{Diag}(\mathbf{n}_x(u_{t1}), \mathbf{n}_x(u_{t2}), \dots, \mathbf{n}_x(u_{tM(t)})), \text{Diag}(\mathbf{n}_y(u_{t1}), \mathbf{n}_y(u_{t2}), \dots, \mathbf{n}_y(u_{tM(t)})) \right) \quad \text{Ec. 4.11} \\
 \text{Diag}(\mathbf{n}_x(u_{t1}), \mathbf{n}_x(u_{t2}), \dots, \mathbf{n}_x(u_{tM})) &= \begin{pmatrix} \mathbf{n}_x(u_{t1}) & 0 & \dots & 0 \\ 0 & \mathbf{n}_x(u_{t2}) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \mathbf{n}_x(u_{tM(t)}) \end{pmatrix} \\
 \text{Diag}(\mathbf{n}_y(u_{t1}), \mathbf{n}_y(u_{t2}), \dots, \mathbf{n}_y(u_{tM})) &= \begin{pmatrix} \mathbf{n}_y(u_{t1}) & 0 & \dots & 0 \\ 0 & \mathbf{n}_y(u_{t2}) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \mathbf{n}_y(u_{tM(t)}) \end{pmatrix}
 \end{aligned}$$

y el nuevo modelo de observación se define a través de la matriz  $C_t$  definida como:

$$\begin{aligned}
 C_t &= N_t I_2 \otimes D_t Q H \quad \text{con} \quad \text{Ec. 4.12} \\
 Q &= I_2 \otimes (G_W(u_1), G_W(u_2), \dots, G_W(u_N))^T, H = (\mathbf{0}_{2N \times 2N}, I_{2N})
 \end{aligned}$$

de la siguiente forma:

$$D_t | \tilde{w}_t - \bar{w} \sim N(C_t w_t, L_t), L_t = \frac{1}{S_D^2} I_{M(t)} \equiv D_t | w_t - \tilde{w} \quad \text{Ec. 4.13}$$

y por tanto las matrices del modelo de observación son ahora:  $F_t = C_t$ ,  $V_t = L_t$  lo que permite de nuevo la aplicación directa del algoritmo de Tabla 4.1

### 4.1.1.3 Inicialización

El algoritmo de Tabla 4.1 genera un estimación para el tiempo  $t$  a partir de la estimación para  $t-1$ . Por tanto, para el uso del algoritmo hace falta determinar las condiciones iniciales para  $t=0$ . Es decir, los valores de  $\tilde{w}_0 = (w_0 - \bar{w}, w_{-1} - \bar{w})$  y  $S_0$ .

En [Blake 1998] se proporcionan tres métodos para inicializar el filtro de Kalman en procesos de seguimiento:

#### 4.1.1.3.1 Posición inicial conocida y estática

En este caso, se posee una estimación inicial del contorno del objeto  $r_{D_0}$  que define un vector de parámetros  $w_0$  del objeto de interés utilizándose entonces  $\tilde{w}_0 = (w_0 - \bar{w}, w_0 - \bar{w})$  y  $S_0$  generada a partir de una distribución a priori del contorno construida en la sección 3.3.3. A partir de esta estimación inicial se deja iterar el filtro de Kalman utilizando siempre como conjunto de medidas el contorno  $r_{D_0}$  hasta

que se obtiene la convergencia. En ese momento se empiezan a tratar las nuevas imágenes con nuevas medidas  $r_{D_t}$

#### 4.1.1.3.2 Posición inicial desconocida

Si la posición inicial es desconocida se parte del vector  $\tilde{w}_0 = (\overline{w} - \overline{w}, \overline{w} - \overline{w}) = (\mathbf{0}, \mathbf{0})$  y  $S_0$  generada a partir de una distribución a priori similar a la de la sección anterior. A partir de este punto inicial se deja iterar el filtro de Kalman en ausencia de toda medida hasta que se obtiene la convergencia (por tanto este método requiere la no utilización de los modelos de velocidad constante de la sección 3.4.3.2.1). En ese momento, se empiezan a tratar las nuevas imágenes con nuevas medidas  $r_{D_t}$

#### 4.1.1.3.3 Velocidad inicial conocida

En muchas aplicaciones es posible determinar una estimación a priori de la velocidad inicial del objeto  $v_0$ . En este caso se utiliza como valor inicial  $\tilde{w}_0 = (w_0 - \overline{w}, w_0 - v_0 - \overline{w})$  y la matriz inicial de varianzas se obtiene como en la sección 4.1.1.3.1. Una vez obtenida dicha matriz empiezan a tratar las nuevas imágenes con nuevas medidas  $r_{D_t}$

#### 4.1.1.4 Rendimiento

El costo computacional del algoritmo depende de la complejidad de la curva y del modelo dinámico. Cuando los espacios de deformación son sencillos el filtro de Kalman es capaz de realizar el seguimiento en tiempo real, sin embargo, en el peor caso (deformación libre) la complejidad computacional es de orden  $o(N^3)$  donde  $N$  es el número de puntos de la curva. Por tanto el incremento de la dimensión del vector de parámetros  $w$  penaliza de forma importante el tiempo de computación. Este costo viene dado por la multiplicación e inversión de varias matrices de orden  $N$ . La capacidad de representación de los wavelets (ver sección 3.2.3) ofrece por tanto reducciones sustanciales del tiempo de ejecución en comparación con otras representaciones.

#### 4.1.1.5 Ejemplos

A continuación se presentan dos ejemplos de seguimiento de contornos utilizando la representación basada en wavelets y el algoritmo presentado en la Tabla 4.1.

Es necesario señalar que la utilización de distribuciones gaussianas limita el tipo de secuencias que pueden ser tratadas por el algoritmo. En efecto, la del unimodalidad

de esta distribución no la hace apropiada para problemas de excesiva complejidad y requiere ambientes controlados y con poco ruido.

El filtro Condensation que se presentará en la sección 4.1.2 permitirá abordar problemas más generales y complejos.

#### 4.1.1.5.1 Ejemplo 1: Seguimiento de una mano

A continuación se muestra un primer ejemplo de seguimiento utilizando el filtro de Kalman. El problema que se pretende resolver es el seguimiento de una mano manejando un ratón sobre una alfombrilla situada en una mesa. Para ello se dispone de una secuencia de imágenes en color y se pretende detectar en cada instante la posición de la mano. El problema se ve complicado por la similitud entre el color de la mesa y el color de la mano lo que provoca que las regiones de cambio de contraste sean la región mano-ropa y la mano-alfombrilla por lo que las medidas del filtro de Kalman provendrán en exclusiva de esas dos regiones

- *Preprocesamiento de imagen*

Puesto que el proceso de medidas del filtro de Kalman trabajará sobre la información de bordes de la imagen es necesario preprocesar la imagen para realzar la información relevante para el proceso de medidas y eliminar en la medida de lo posible cualquier otra información que pudiera “distraer” al filtro de Kalman al realizar el seguimiento. Para ello, tras leer la imagen en color y con el fin de detectar la mano, se calcula la tonalidad a partir de las componentes de color y se aplica un filtro rápido y sencillo como el filtro de Sobel [Ziou 1998] para detectar los bordes (ver Figura 4.5)



**Figura 4.5** Preprocesamiento de la imagen. Imagen original en color (impresa en blanco y negro) a la izquierda. Imagen de tonalidad (centro). Detalle de la imagen de bordes con el contorno y región de búsqueda en color gris claro(derecha).

Con respecto a la imagen anterior es necesario hacer algunos comentarios. En primer lugar, los cambios bruscos de negro a blanco que aparecen en la imagen central son debidos a la circularidad de la tonalidad (está definida como un ángulo) y no son bordes reales. Por esta razón no aparecen en la imagen de la derecha. Nótese como la igualdad en tonalidad entre mesa y mano hacen difícil el problema de seguimiento pues como se ve en la imagen de la derecha no existen bordes (representados en negro) entre

la mano y la mesa. En esto consiste la potencia de la técnica de regularización bayesiana que es capaz de representar el contorno completo (representado en gris en la imagen de la derecha) a partir de la información parcial obtenida de la imagen y la representación a priori del modelo.

La capacidad de la representación probabilística de proporcionar una región de confianza para el proceso de búsqueda (Ec. 4. 7) se representa en la figura anterior mediante las flecha en gris que salen del contorno en su dirección normal. Es en esta región en la que el contorno busca información para actualizar la predicción generada por el modelo dinámico (contorno en gris) con la obtenida a partir de la imagen. Esta región de confianza es el resultado de dos requerimientos contradictorios: una región de confianza muy pequeña puede hacer que el contorno no obtenga información de la imagen y una región muy grande introduce información de otros objetos que “distraen” el filtro de Kalman. La región de búsqueda obtenida a partir del filtro ( que en este caso un 95% de confianza) la genera el filtro de Kalman a partir del modelo e imagen.

- *Parámetros de seguimiento*

La componente determinista del cambio en posición y la deformación de la mano en este ejemplo tiene dos componentes fundamentales, la primera componente es un movimiento de traslación mientras que la componente de deformación se puede modelar a través de una deformación debida a una transformación afín. Por otra parte, dado que la mano está confinada en la alfombrilla y debido a la naturaleza del movimiento del ratón se espera un movimiento lento tanto en su componente traslacional como afín y con poca indeterminación espacial en el subespacio afín, mientras que la indeterminación espacial debida a la traslación de ciñe a movimientos en el ámbito de la alfombrilla.

Para la descripción de los parámetros deterministas se utiliza el modelo de oscilador armónico cuyas características de presentan en el Apéndice B y así se tendrán los valores :  $a_2 = -\exp(-2d\mathbf{i})$ ,  $a_1 = 2\exp(-2d\mathbf{i})\cos(2\pi f)$  con:

Traslación	$d=1$	$f=0$	$r=30$
Deformación	$d=0.1$	$f=0$	$r=3$

**Tabla 4.2 Parámetros dinámicos para el ejemplo del seguimiento de la mano**

La elección de los parámetros refleja que no se espera movimiento oscilatorio ( $f = 0$ ) y se espera mayor velocidad de movimiento en el subespacio de traslación que en el de deformación. Además hay un desplazamiento medio de 30 pixeles en el subespacio de traslación mientras que el subespacio de deformación tiene un desplazamiento medio menor correspondiente al rango más reducido de transformaciones. La incertidumbre debida a las medidas es de  $s_D=2$  pixeles y el parámetro de deformación en el espacio de Sobolev a  $b=2.25$ . Como función wavelet se ha utilizado la  $D_8$  de Daubechies.

- *Resultados del seguimiento*

En la Figura 4.6 se muestran los resultados del proceso de seguimiento para este ejemplo

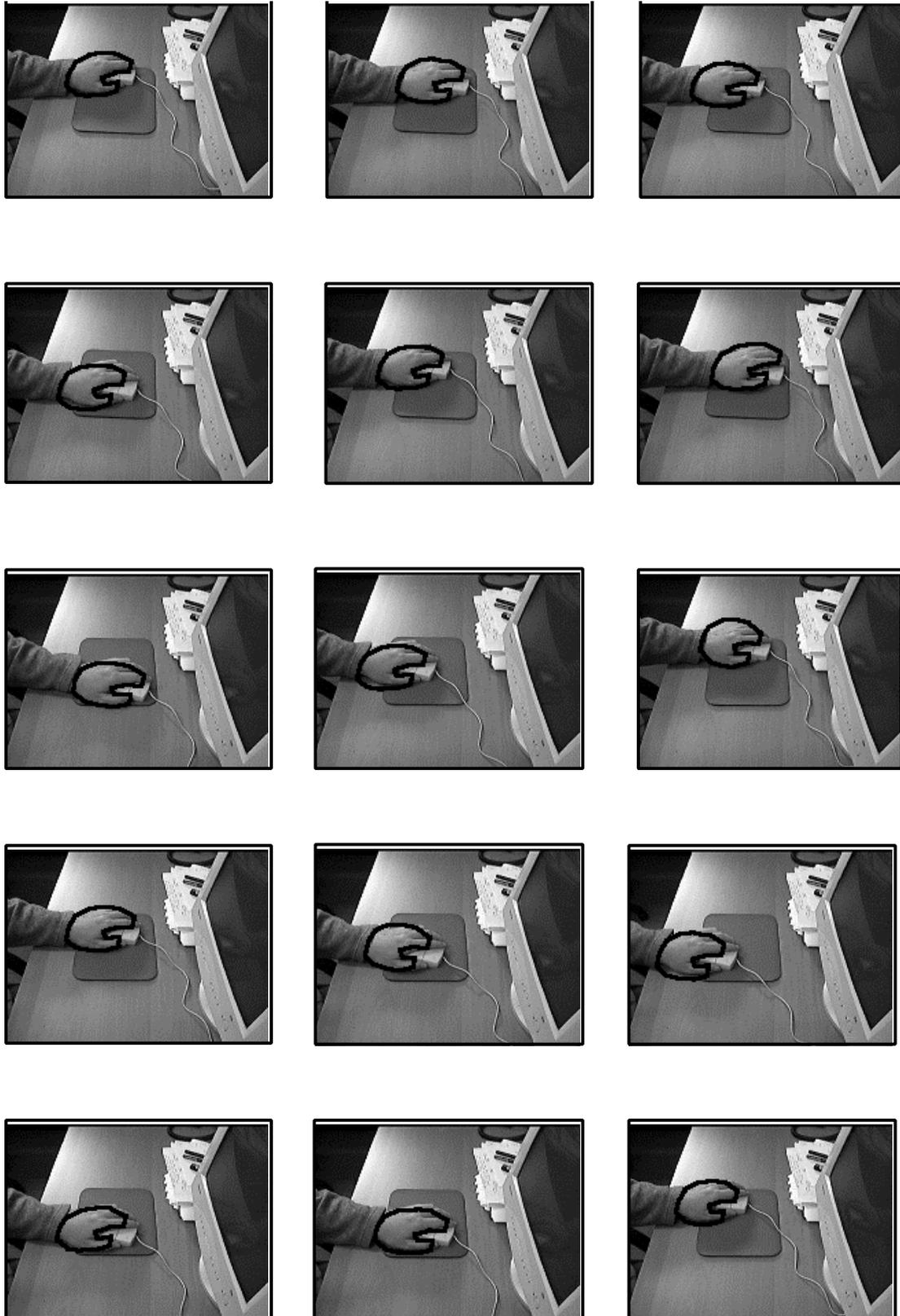


Figura 4.6. Imágenes 1-15 de la secuencia de la mano. Resultados del proceso de seguimiento. En negro se muestra la media de la distribución gaussiana.

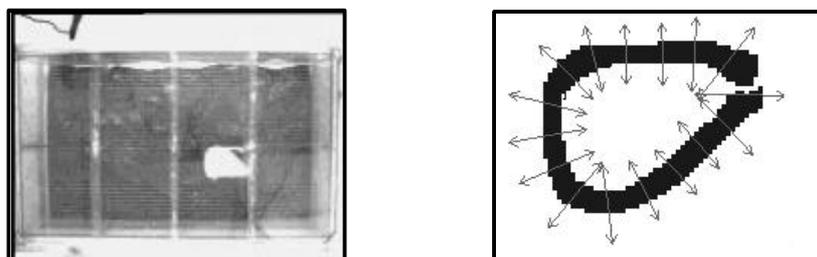
Puede verse que el algoritmo de seguimiento es capaz de seguir la mano en su movimiento sobre la alfombrilla. Nótese que tal y como puede verse comparando las imágenes 6 y 7 de la secuencia, el movimiento de la mano no tiene únicamente una componente traslacional sino que también tiene una componente afín que es capaz de determinarse a través del modelo. Ocasionalmente el algoritmo de seguimiento se “distrae” con otros bordes. Por ejemplo en la imagen 4 la parte inferior del contorno duda entre el borde de la mano y el borde proveniente del borde entre la alfombrilla y la mesa. Sin embargo, utilizando la información aportada por el resto de los bordes de la imagen, el algoritmo es capaz de seguir la mano correctamente (imágenes 5 y 6).

#### 4.1.1.5.2 Ejemplo 2: Seguimiento de un ratón

A continuación se muestra el segundo ejemplo de seguimiento utilizando el filtro de Kalman. El problema que se pretende resolver es el seguimiento de un ratón situado sobre una celda en un entorno de laboratorio. Para ello se dispone de una secuencia de imágenes en blanco y negro y se pretende detectar en cada instante el eje cabeza-principio de la cola del ratón. En este caso, con respecto al modelado, la dificultad del problema viene dada por la gran variabilidad de la forma del ratón. Con respecto al proceso de medidas, al estar la secuencia tomada en un entorno controlado de laboratorio se facilita la tarea de tomar medidas, siendo las principales fuentes de distracción del algoritmo la cola del ratón y el reflejo de éste sobre las paredes de su celda.

- *Preprocesamiento de imagen*

Al igual que en el ejemplo anterior, puesto que el proceso de medidas del filtro de Kalman trabajará sobre la información de bordes de la imagen es necesario preprocesar la imagen para realzar la información relevante para el proceso de medidas y eliminar en la medida de lo posible cualquier otra información que pudiera “distraer” al filtro de Kalman al realizar el seguimiento. Para ello, tras leer la imagen en blanco y negro y con el fin de detectar el ratón, se calcula la diferencia entre la imagen leída y una imagen de las celdas sin el ratón. Esta diferencia se umbraliza y al resultado se le aplica el filtro de Sobel para detectar los bordes (Figura 4.7).



**Figura 4.7** Preprocesamiento de la imagen. Imagen original en blanco y negro (a la izquierda). Detalle de la imagen de bordes con el contorno y región de búsqueda en color gris claro(derecha).

La capacidad de la representación probabilística de proporcionar una región de confianza para el proceso de búsqueda (Ec. 4. 7) se representa en la figura anterior mediante las flecha en gris que salen del contorno en su dirección normal. Es en esta región en la que el contorno busca información para actualizar la predicción generada por el modelo dinámico (contorno en gris) con la obtenida a partir de la imagen.

- *Parámetros de seguimiento*

La componente determinista del cambio en posición y la deformación del ratón en este ejemplo tiene dos componentes fundamentales, la primera componente es un movimiento de traslación mientras que la componente de deformación se puede modelar a través de una rotación y una contracción y expansión del eje cabeza-cola del ratón que será aproximada por medio de una transformación afín. Por otra parte dado que el ratón está confinada en la celda y a la naturaleza del movimiento del ratón se espera un movimiento lento tanto en su componente traslacional mientras que se prevé un movimiento más ágil y con mayor indeterminación espacial en el subespacio afín

Para la descripción de los parámetros deterministas se utiliza el modelo de oscilador armónico del Apéndice B y así se tendrá:  $a_2 = -\exp(-2d\mathbf{i})$ ,  $a_1 = 2\exp(-2d\mathbf{i})\cos(2\pi f)$  con:

Traslación	$d=2$	$f=0$	$r=20$
Deformación	$d=5$	$f=0$	$r=6$

**Tabla 4.3** Parámetros dinámicos para el ejemplo del seguimiento del ratón

La elección de los parámetros refleja que no se espera movimiento oscilatorio ( $f = 0$ ) y se espera mayor velocidad de movimiento en el subespacio afín que en el de traslación. Además hay un desplazamiento medio de 20 pixeles en el subespacio de traslación mientras que el subespacio de deformación tiene un desplazamiento medio menor correspondiente al rango de transformaciones afines presentes en la imagen.

La incertidumbre debida a las medidas es de  $s_D=2$  pixeles y el parámetro de deformación en el espacio de Sobolev a  $b=2.25$ . Como función wavelet se ha utilizado la  $D_8$  de Daubechies.

- *Resultados del seguimiento*

En la figura 4.8 se muestran los resultados del proceso de seguimiento para este ejemplo. Puede verse que, el algoritmo de seguimiento sigue al ratón en su movimiento en la celda. Nótese que tal y como puede verse comparando las imágenes 6 y 15 de la secuencia, el movimiento de la mano no tiene únicamente una componente traslacional sino que también tiene una componente afín que es capaz de determinarse a través del modelo.

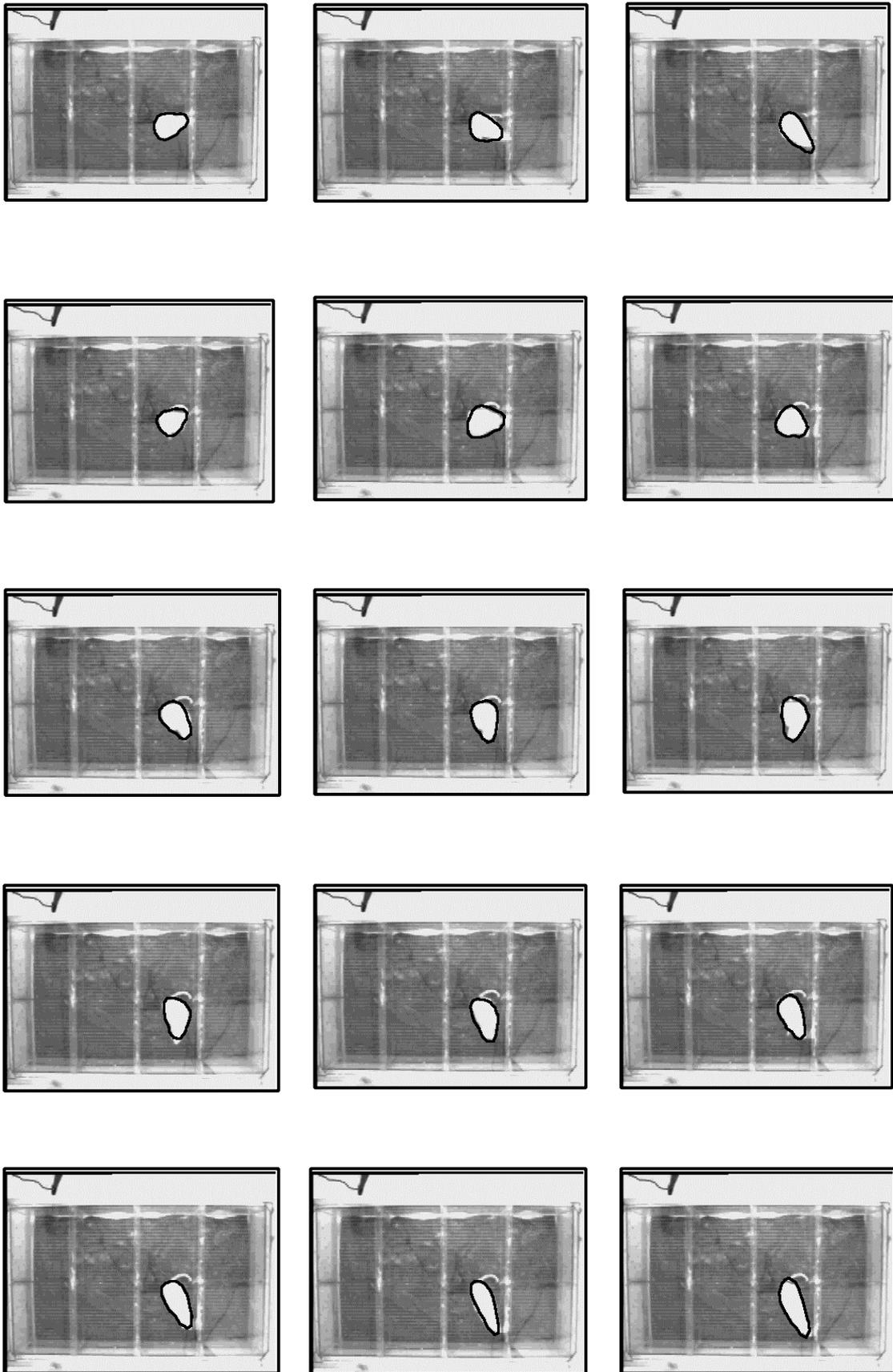
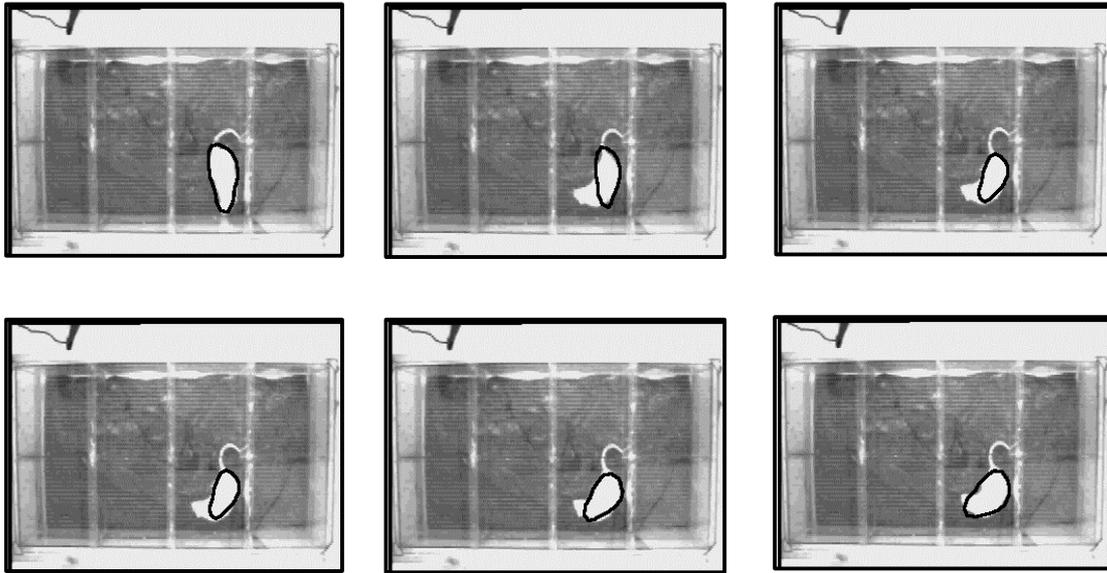
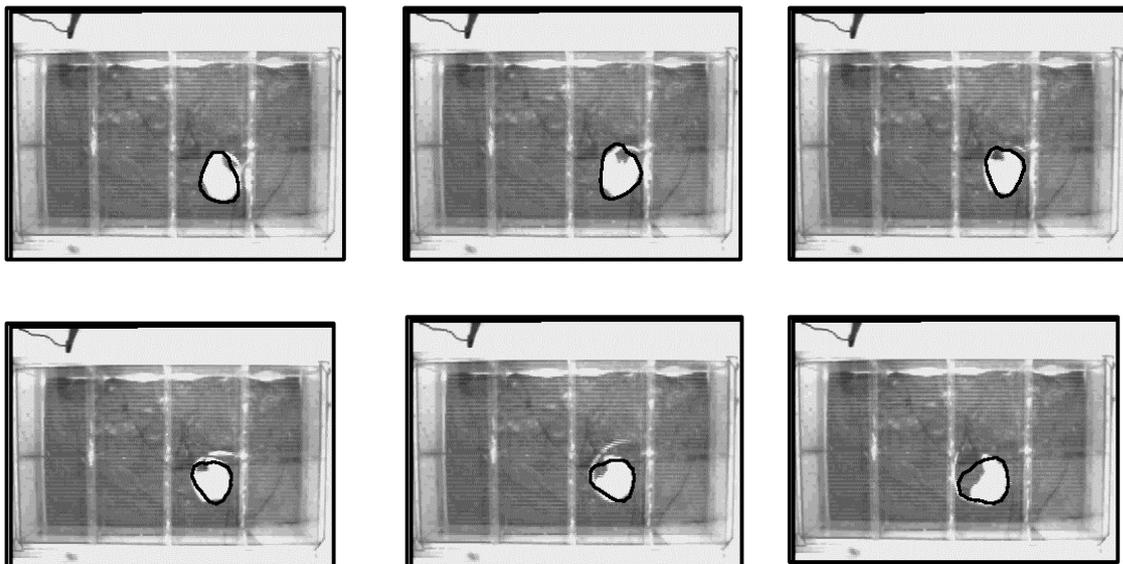


Figura 4. 8 Imágenes 1-15 de la secuencia del ratón. Resultados del proceso de seguimiento. En negro se muestra la media de la distribución gaussiana.

Este problema de seguimiento, sin embargo, es mucho más complejo que el anterior. En efecto, la aproximación de las transformaciones del ratón por medio de una transformación afín solamente es válida para ciertos momentos de la secuencia y no reproduce en su totalidad todos los posibles movimientos del ratón. Por otra parte, el modelo dinámico es sólo una aproximación a la realidad, pues el ratón se para y mueve en ocasiones de forma brusca por lo que sería necesario modelos más complejo tanto en cuanto a forma como a dinámica para realizar un seguimiento exento de fallos.



**Figura 4.9** Problemas en la secuencia del ratón. Nótese el cambio en forma entre las imágenes 1 y 2. El cambio entre dichas imágenes ocurrió en 1/15 de segundo (velocidad de muestreo). En este caso el algoritmo fue capaz de recuperarse.



**Figura 4.10** Problemas en la secuencia del ratón. Nótese el cambio en forma entre las imágenes 1 y 2. El cambio entre dichas imágenes ocurrió en 2/15 de segundo (dos veces la velocidad de muestreo). En este caso el algoritmo no fue capaz de recuperarse.

Debemos señalar, no obstante, que estrictamente hablando el contorno continua siguiendo al ratón. Sin embargo, la información de la posición cabeza-cola no es la correcta y es en ese sentido en el que se dice que el algoritmo se pierde.

#### 4.1.2 Seguimiento de contornos con modelos no gaussianos

Se presenta a continuación una aproximación al seguimiento de contornos en espacios generales donde ya no es válida la utilización de la distribución probabilística gaussiana. En esta sección, no obstante se utilizará la distribución en el espacio de Sobolev para el modelo a priori dinámico. La utilización de funciones de distribución generales será realizada mediante el filtro Condensation. En esta sección se presenta el algoritmo de seguimiento utilizando los modelos dinámicos basados en wavelets y se discuten algunos ejemplos de su aplicación.

##### 4.1.2.1 Introducción

El filtro de Kalman como estimador lineal recursivo es un caso especial, aplicado a distribuciones gaussianas, de un proceso más general de propagación de probabilidades. En el caso de que exista una gran cantidad de ruido en la imagen no es razonable el asumir que la distribución sea gaussiana y por tanto unimodal a lo largo del tiempo.

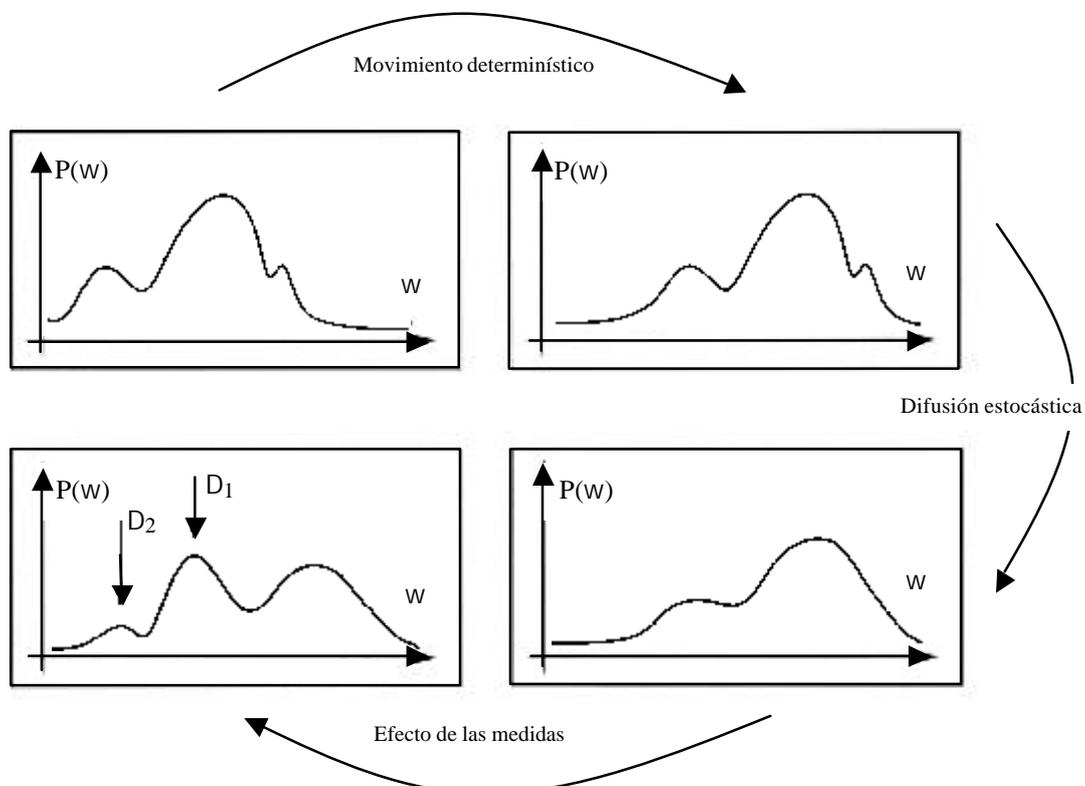


Figura 4.11 Difusión estocástica para modelos no gaussianos (De [Blake 1998])

En este caso, aunque el modelo dinámico pueda mantener una distribución gaussiana el efecto de las observaciones obtenidas a partir de la imagen será el de obligar a la función de densidad a aumentar la probabilidad en la vecindad de éstas. En ambientes complejos o ruidosos hay varias de estas observaciones que compiten entre sí (ver Figura 4.11) provocando funciones de densidad multimodales lo que promueve el uso de modelos no gaussianos.

#### 4.1.2.2 El filtro Condensation

El filtro Condensation (ver sección 2.4.3.3) se utiliza para aproximar una distribución arbitraria y en general no gaussiana  $P(w | D)$  para ello genera una variable aleatoria  $\tilde{w}$  con distribución  $P(\tilde{w})$  que aproxima a  $P(w | D)$ . Primero se genera un conjunto de muestras  $S = \{s^{(i)}, i = 1 \dots N_s\}$  con probabilidad a priori  $P(w)$  y se elige un índice  $n \in \{1 \dots N_s\}$  con probabilidad  $p^{(n)}$  donde

$$p^{(n)} = \frac{P_D(s^{(n)})}{\sum_{j=1}^N P_D(s^{(j)})}, \quad P_D(s^{(j)}) = P(D|w = s^{(j)}) \quad \text{Ec. 4. 14}$$

los valores elegidos de esta forma tienen una distribución que se aproxima a  $P(w | D)$  y que converge a ésta cuando el número de elementos de  $S$  tiende a infinito.

##### 4.1.2.2.1 El modelo dinámico

El algoritmo Condensation es capaz de manejar cualquier modelo dinámico para describir el desplazamiento del contorno entre imágenes. En esta sección se emplearán los modelos dinámicos autoregresivos de segundo orden basados en wavelets de la sección 3.4.3 y que fueron empleados anteriormente en el desarrollo del filtro de Kalman. De esta forma se tendrá:

$$\tilde{w}_t - \overline{\tilde{w}} = \tilde{A}(\tilde{w}_{t-1} - \overline{\tilde{w}}) + \tilde{B}\tilde{n}_t \quad \text{con} \quad \text{Ver Ec. 3.119}$$

$$\tilde{w}_t = \begin{pmatrix} w_{t-1} \\ w_t \end{pmatrix}, \quad \overline{\tilde{w}} = \begin{pmatrix} \overline{w} \\ \overline{w} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ A_2 & A_1 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} \mathbf{0} \\ \mathbf{B} \end{pmatrix}, \quad \tilde{n}_t = \begin{pmatrix} m_t \\ n_t \end{pmatrix}$$

donde:

$$\mathbf{n}_t = (n'_{0,0;x}, n_{0,0;x}, \dots, n_{j-1,2^{j-1}-1;x}, n'_{0,0;y}, n_{0,0;y}, \dots, n_{j-1,2^{j-1}-1;y}) \quad \text{con:} \quad \text{Ec. 4. 15}$$

$$n'_{0,0} = (n'_{0,0;x}, n'_{0,0;y}) \sim DGGS_2(\mathbf{0}, 1)$$

$$n_{j,k} = (n_{j,k;x}, n_{j,k;y}) \sim DGGS_p(\mathbf{0}, 1)$$

siendo las variables  $n_{j,k}$  y  $m_{j,k}$  independientes.

De esta forma se puede aprovechar la versatilidad del algoritmo Condensation para trabajar no sólo con distribuciones gaussianas sino con distribuciones *DGGS*.

#### 4.1.2.2.2 El modelo de observaciones

El efecto de las observaciones externas  $D_t$  es el de proporcionar un efecto reactivo al proceso de difusión que promueve que la función de densidad tenga máximos en la vecindad de éstas. Se procederá por tanto a modelar la densidad de observación  $P(D | w)$ . Para ello se procede a definirla en una dimensión y se generaliza después a dos dimensiones.

En una dimensión las observaciones se reducen a un conjunto de posiciones escalares  $D = (d_1, d_2, \dots, d_m)$  y la densidad de observación tiene la forma  $P(D | x)$  donde  $x$  es una posición escalar. Se utilizará entonces como función de densidad [Isard 1998a]

$$p(D | x) \propto 1 + \frac{1}{\sqrt{2psa}} \sum_m e^{-\frac{(d_m-x)^2}{2s^2}} \quad \text{Ec. 4. 16}$$

donde  $s$  representa la incertidumbre en la posición de los  $d_i$  y  $a$  equilibra la distribución en el caso de que ninguno de los  $d_i$  del vector  $D$  sea la característica correspondiente a  $x$ .

Para dos dimensiones  $D$  es en principio todo el conjunto de características detectadas en la imagen. No obstante para reducir el tiempo computacional la densidad se evalúa como el producto de densidades unidimensionales a lo largo de la curva. Una mayor velocidad se obtiene eligiendo únicamente aquella característica más cercana a la predicha por el modelo dinámico.

Se pasa por tanto a continuación a explicar la forma del filtro Condensation [Isard 1998a] en el marco de los modelos dinámicos autoregresivos de segundo orden basados en wavelets.

#### Filtro “Condensation”

##### Paso 0

Generar un conjunto de muestras  $S = \{s_0^{(i)}, p_0^{(i)}, c_0^{(i)}, i = 1 \dots N_s\}$  para el tiempo  $t=0$  donde:

$s_0^{(i)}$  es una muestra de la distribución a priori de la curva

$p_0^{(i)}$  es la probabilidad de que esta curva se elija de  $S$  (inicialmente  $1/N_s$ )

$c_0^{(i)}$  es la distribución de probabilidad acumulativa para  $S$

Hacer  $t=1$

**Paso 1**

A partir de  $S$  (el conjunto previo de muestras) construir un nuevo conjunto de muestras  $S_{\text{nuevo}}$  como sigue:

Hacer  $i=1$

Mientras  $i \leq N_S$  hacer

**Paso 1.1 (Selección)**

Seleccionar una muestra  $s_t^{(i)}$  como sigue:

- Generar un número aleatorio  $r$  uniformemente distribuido en  $[0,1]$
- Encontrar por búsqueda binaria el menor  $j$  para el que  $c_{t-1}^{(j)} \geq r$
- Hacer  $s_t^{(i)} = s_{t-1}^{(j)}$

**Paso 1.2 (Predicción)**

Generar una predicción dinámica muestreando de:

$$s_t^{(i)} - \tilde{w} = \tilde{A}(s_{t-1}^{(i)} - \tilde{w}) + \tilde{B}\tilde{n}_t$$

**Paso 1.3 (Medición)**

Medir y obtener los pesos para esta nueva posición en términos de las medidas observadas  $D_t$ :

$$p_t^{(i)} = p(D_t | \tilde{w}_t = s_t^{(i)})$$

Guardar esos valores in  $S_{\text{nuevo}}$ . Incrementar  $i$

fin del mientras

Normalizar los pesos de forma que  $\sum_i p_k^{(i)} = 1$ , y evaluar las probabilidades acumulativas como:  $c_k^{(0)} = 0$ ,  $c_k^{(i)} = c_k^{(i-1)} + p_k^{(i)}$ ,  $i = 1 \dots N_S$

Hacer  $S = S_{\text{nuevo}}$ . Incrementar  $t$ . Ir al Paso 1

**Tabla 4.4 Algoritmo Condensation**

4.1.2.3 Inicialización

La inicialización del filtro Condensation es análoga a la inicialización definida para el filtro de Kalman en la sección 4.1.1.3, es decir se selecciona entre cada una de las tres posibilidades: posición inicial conocida, posición inicial desconocida o velocidad inicial conocida y se realizan iteraciones del filtro Condensation hasta obtener

la convergencia. La única diferencia es que en este caso la convergencia se estudia sobre el conjunto  $S$  en lugar de sobre los parámetros de la distribución normal.

#### 4.1.2.4 Rendimiento

El estudio del rendimiento del filtro Condensation es complicado debido a que su complejidad es función de la dificultad en aproximar la función de probabilidad  $P(W | D)$  por medio del conjunto de muestras en  $S$ . No obstante, la descripción de curvas basadas en wavelets ofrece varias características que las hacen apropiadas para este filtro. En primer lugar, la capacidad para representar la curva que evoluciona a lo largo del tiempo con un pequeño número de coeficientes y la capacidad de cambio entre modelo (vector de características) e imagen (curva) con costo lineal provocan que esta sea una descripción óptima desde el punto de vista del tiempo computacional. Por otra parte, la generalidad de la representación que permite combinar la capacidad de representar cambios bruscos en la curva junto con la definición a medida de la suavidad intrínseca de ésta permite una mayor amplitud de representación que la que pueden ofrecer otros modelos basados en b-splines o descripciones de Fourier. Esto permite obtener seguimientos cercanos al tiempo real para modelos sencillos de deformación y entornos no demasiado complejos.

#### 4.1.2.5 Ejemplos

A continuación se presentan dos ejemplos de seguimiento utilizando la representación basada en wavelets y el algoritmo presentado en la Tabla 4.4. El filtro Condensation que se utiliza en esta sección permite abordar problemas más generales y complejos. En primer lugar se presentará una secuencia de seguimiento de una mano donde a diferencia de la sección 4.1.1.5.2 el fondo de la imagen, que tiene gran complejidad, no es estático y por lo tanto no puede eliminarse mediante el cálculo de la diferencia con una imagen de referencia. En el segundo caso, se muestra un problema de seguimiento en exteriores donde de nuevo el fondo es complejo y el objeto que se sigue puede confundirse con otros objetos de la imagen.

##### 4.1.2.5.1 Ejemplo 1: Seguimiento de una mano

A continuación se muestra un primer ejemplo de seguimiento utilizando el filtro Condensation. El problema que se pretende resolver es el seguimiento de una mano con una lata que se mueve sobre un fondo complejo. Para ello se dispone de una secuencia de imágenes en blanco y negro y se pretende detectar en cada instante la posición del conjunto mano-lata. En este caso el preprocesamiento está orientado a intentar eliminar en lo posible la mayor cantidad de bordes no significativos de la imagen para lo cual ésta se suavizará con un filtro gaussiano tal y como se muestra en la siguiente sección

- *Preprocesamiento de imagen*

Como en los anteriores ejemplos, el filtro Condensation trabajará sobre la información de bordes de la imagen que es necesario preprocesar para eliminar en la cualquier otra información que pudiera “distraer” al filtro al realizar el seguimiento. Para ello, tras leer la imagen en blanco y negro y con el fin de eliminar la mayor cantidad posible de bordes espurios se realiza un suavizado de la imagen mediante un filtro gaussiano, y se aplica el filtro de Sobel para detectar los bordes (Figura 4.12)



Figura 4.12 Preprocesamiento de la imagen. Imagen original en blanco y negro (a la izquierda). Imagen de bordes con el contorno en línea fina(derecha).

Nótese como el suavizado gaussiano permite obtener los bordes más significativos. Aún así, existe una gran cantidad de bordes tanto dentro de la lata como en la pared .

- *Parámetros de seguimiento*

La componente determinista del cambio en posición y la deformación de la mano en este ejemplo tiene dos componentes fundamentales, la primera componente es un movimiento de traslación mientras que la componente de deformación se puede modelar a través de una deformación debida a una transformación afín. Por otra parte dado que es posible que la mano se mueva por toda la imagen se utilizará el modelo de velocidad constante (sección 3.4.3) para la componente traslacional, mientras que por el contrario se espera un movimiento lento y con poca indeterminación espacial en el subespacio afín.

Para la descripción de los parámetros deterministas se utiliza el modelo de oscilador armónico del Apéndice B y así se tendrá:  $a_2 = -\exp(-2d\mathbf{i})$ ,  $a_1 = 2\exp(-2d\mathbf{i})\cos(2\pi f)$  con:

Traslación	$d=0$	$f=0$	$g=35$
Deformación	$d=5$	$f=0$	$r=2$

Tabla 4.5 Parámetros dinámicos para el ejemplo del seguimiento de la mano

La elección de los parámetros refleja que no se espera movimiento oscilatorio ( $f = 0$ ) y se espera mayor velocidad de movimiento en el subespacio de traslación que en el de deformación. Además hay un desplazamiento medio de 2 píxeles en el subespacio de

deformación correspondiente al rango reducido de transformaciones afines presentes en la imagen.

La incertidumbre debida a las medidas es de  $s_D=2$  píxeles y el parámetro de deformación en el espacio de Sobolev a  $b=2.25$ . La familia de wavelets utilizada ha sido la  $D_8$  de Daubechies.

- *Resultados del seguimiento*

En la figura 4.14 se muestran los resultados del proceso de seguimiento para este ejemplo. En este caso se ha utilizado una población de  $N_S=100$  elementos. En la imagen se muestra una estimación del máximo a posteriori de la distribución a partir de la población. Nótese que en este caso no puede utilizarse la media pues la distribución puede ser multimodal y en el mejor de los casos la media coincidiría únicamente con una de las modas de la distribución. Para calcular el estimador MAP se ha examinado la población y se ha elegido el elemento de máxima probabilidad. Este es un procedimiento rápido proporciona una estimación burda del estimador (no utiliza por ejemplo la coherencia temporal entre una estimación y la siguiente). Aunque existen métodos para mejorar dicha estimación [Doucet 2000] no se han utilizado por su costo computacional.

La multimodalidad de la distribución puede verse en el siguiente ejemplo [Pérez 2000c] donde se representan los diez elementos de mayor probabilidad del conjunto  $S$  en color blanco y se ve en la imagen central la aparición de dos modas. La primera aparece alrededor de la mano, mientras que la segunda es consecuencia de una alineación accidental de bordes: la parte inferior de la mano la parte inferior del monitor y la parte inferior del papel sobre la mesa. No obstante, a medida que el proceso de seguimiento continúa esta segunda moda desaparece y el conjunto  $S$  vuelve a concentrarse alrededor de la primera moda.



**Figura 4.13** Ejemplo de multimodalidad en la distribución del contorno. En color blanco se muestran los 10 elementos de mayor probabilidad del conjunto  $S$ .

En la figura 4.14 puede verse como se realiza un correcto seguimiento de la mano con ocasionales distracciones (por ejemplo en la imagen 2) y que son debidos fundamentalmente a la elección de un conjunto  $S$  con un número pequeño de muestras. Nótese que este seguimiento se realiza a pesar de los cambios en iluminación presentes en la secuencia.



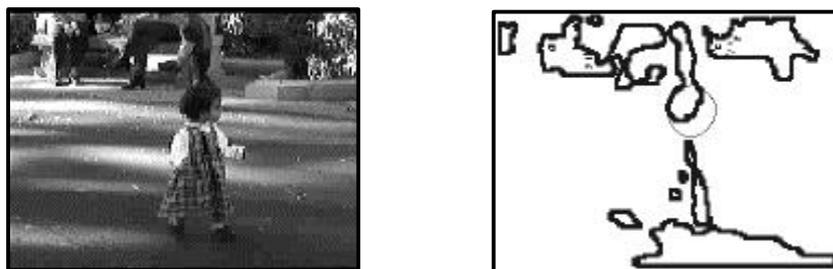
Figura 4.14 Imágenes 1-15 de la secuencia de la lata. Resultados del proceso de seguimiento. En negro se muestra la estimación MAP del contorno.

#### 4.1.2.5.2 Ejemplo 2: Seguimiento en exteriores

A continuación se muestra el segundo ejemplo de seguimiento utilizando el filtro Condensation. El problema que se pretende resolver es el seguimiento de la cabeza de una niña que se mueve sobre un fondo complejo. Para ello se dispone de una secuencia de imágenes en blanco y negro y se pretende detectar en cada instante cual es la posición de la cabeza en la imagen. En este caso el preprocesamiento está orientado a intentar eliminar en lo posible la mayor cantidad de bordes no significativos de la imagen para lo cual se realizará una segmentación de la imagen mediante intervalos de niveles de gris tal y como se muestra en la siguiente sección

- *Preprocesamiento de imagen*

En este ejemplo el seguimiento de la cabeza se realizará mediante la localización en cada imagen de una región circular oscura correspondiente al cabello de la niña. Para ello, tras leer la imagen en blanco y negro y con el fin de eliminar la mayor cantidad posible de bordes no esenciales se segmenta la imagen utilizando niveles de gris y posteriormente se realiza un suavizado de la imagen mediante un filtro gaussiano, y se aplica el filtro de Sobel para detectar los bordes (Figura 4.15)



**Figura 4.15** Preprocesamiento de la imagen. Imagen original en blanco y negro (a la izquierda). Imagen de bordes con el contorno en línea fina(derecha).

Nótese como el suavizado gaussiano permite obtener los bordes más significativos. Aún así, existe una gran cantidad de bordes que pueden despistar al algoritmo.

- *Parámetros de seguimiento*

La componente determinista del cambio en posición y la deformación en este ejemplo tiene dos componentes fundamentales, la primera componente es un movimiento de traslación mientras que la componente de deformación se puede modelar a través de una deformación debida a una transformación afín. Por otra parte dado que es posible que la cabeza se mueva por toda la imagen se utilizará el modelo de velocidad constante (sección 3.4.3) para la componente traslacional, mientras que por el contrario se espera un movimiento muy lento y con poca indeterminación espacial en el subespacio afín.

Para la descripción de los parámetros deterministas se utiliza el modelo de oscilador armónico del Apéndice B y así se tendrá:  $a_2 = -\exp(-2d\mathbf{i})$ ,  $a_1 = 2\exp(-2d\mathbf{i})\cos(2\pi f)$  con:

Traslación	$d=0$	$f=0$	$g=70$
Deformación	$d=5$	$f=0$	$r=1$

**Tabla 4.6** Parámetros dinámicos para el ejemplo del seguimiento de la niña

La elección de los parámetros refleja que no se espera movimiento oscilatorio ( $f = 0$ ) y se espera mayor velocidad de movimiento en el subespacio de traslación que en el de deformación. Además hay un desplazamiento medio de 1 pixel en el subespacio de deformación correspondiente al rango reducido de transformaciones afines presentes en la imagen.

La incertidumbre debida a las medidas es de  $s_D=2$  pixeles y el parámetro de deformación en el espacio de Sobolev  $b=2.25$ . La familia de wavelets utilizada ha sido la  $D_8$  de Daubechies.

- *Resultados del seguimiento*

En la figura 4.17 se muestran los resultados del proceso de seguimiento para este ejemplo. En este caso, se ha utilizado una población de  $N_S=100$  elementos. En la imagen se muestra una estimación del máximo a posteriori de la distribución a partir de la población. En este caso, una de las dificultades del algoritmo es la presencia de movimientos bruscos en la imagen, paradas y movimientos repentinos que provocan una mayor incertidumbre en el modelo. Esta mayor incertidumbre se ve reflejada en la mayor dispersión de los elementos del conjunto  $S$  tal y como se ve en la Figura 4.16.



**Figura 4.16** Incertidumbre en la distribución del contorno. En color blanco se muestran los 10 elementos de mayor probabilidad del conjunto  $S$ .

En la figura 4.17 puede verse como se realiza un correcto seguimiento de la cabeza con ocasionales distracciones (por ejemplo en la imagen 6). Nótese que este seguimiento se realiza a pesar de la presencia de otros objetos en movimiento en la escena y los grandes cambios de iluminación en dentro de las imágenes.

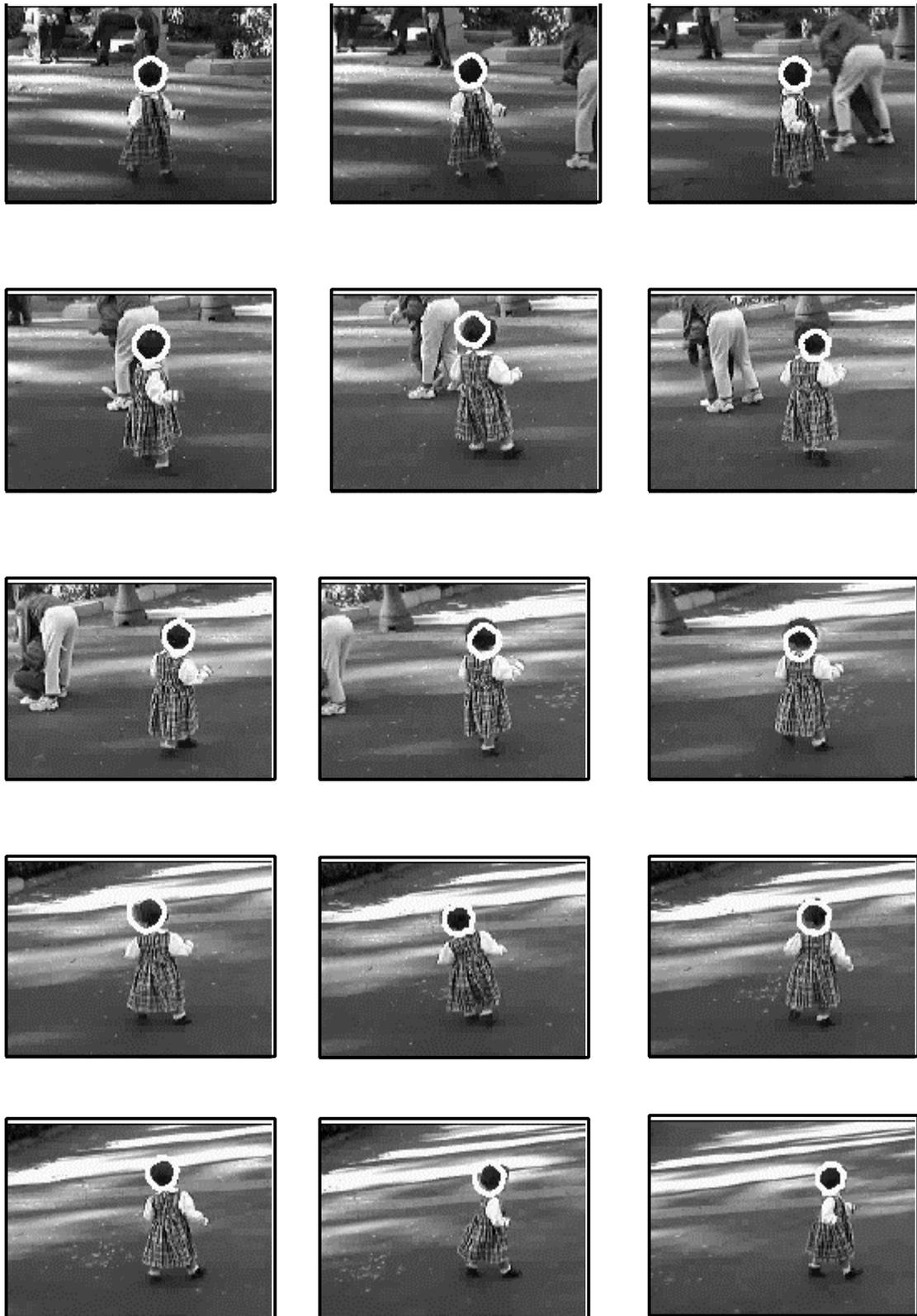


Figura 4.17 Imágenes 1-15 de la secuencia de exteriores. Resultados del proceso de seguimiento. En blanco se muestra la estimación MAP del contorno.

## 4.2 Emparejamiento de contornos

Una vez mostradas las aplicaciones de los modelos basados en wavelets para problemas de seguimiento, se pasan a abordar aplicaciones de los modelos basados en multiwavelets en problemas de reconocimiento. El problema del emparejamiento de contornos y las técnicas para su resolución ya ha sido abordado en la sección 2.2. En esta sección se vuelve a examinar el problema en el contexto de la inicialización del proceso de seguimiento en uno de los casos no examinados anteriormente, esto es, en el caso de que se conozca cual es la forma del contorno pero se desconozca su posición y en general cualquier otra transformación geométrica que haya podido tener. Este problema puede verse además complicado por la presencia de ruido y oclusión por otros objetos. En la sección 2.2.1.2.3.7 ya se han presentado técnicas para resolver este problema utilizando representaciones del contorno basadas en wavelets para el caso de las similitudes euclídeas. A continuación se presenta una extensión de esta metodología basada en multiwavelets para el caso afín mediante un algoritmo que es capaz de trabajar también bajo la presencia de oclusiones parciales del objeto de interés.

### 4.2.1 Introducción

Una de las principales dificultades para a la hora de reconocer objetos a partir de imágenes es el hecho de que la forma del objeto depende del punto de vista. Es por tanto necesario desarrollar descripciones geométricas que no se vean afectadas por un cambio en la posición relativa del objeto y la cámara. Una representación invariante frente a tales transformaciones geométricas constituye entonces el problema básico que debe resolverse.

La distorsión que sufre la forma de un objeto al ser observado mediante una cámara bajo una orientación arbitraria puede ser modelada como una transformación proyectiva. (ver sección 2.1.1.2.3). Sin embargo, cuando las dimensiones del objeto son pequeñas comparadas con la distancia de la cámara al objeto, puede asumirse el modelo de perspectiva *débil* y se asume entonces que puede aproximarse la transformación proyectiva por medio de una transformación afín.

A la hora de construir una representación invariante del objeto frente a transformaciones afines se han utilizado dos tipos de técnicas fundamentales: técnicas en el dominio espacial y técnicas en el dominio de alguna transformación. Dada la capacidad de las representaciones multiescala (secciones 2.1.1.2.6 y 2.1.1.2.7) para representar de forma precisa los cambios ocurridos en el dominio espacial han aparecido recientemente un nuevo tipo de técnicas basadas en estas representaciones [Tieng 1997]. Éstas, sin embargo, necesitan alguna medida global del contorno (por ejemplo el centroide) para desarrollar un algoritmo de emparejamiento lo cual las hace de poca utilidad si no está presente el contorno completo debido a oclusiones, cambios en la condiciones de iluminación o ruido.

En las siguientes secciones se muestra como estas limitaciones aparecen en el caso afín debido a la utilización de una única función wavelet en la descomposición del contorno y se propondrá el uso de un par de wavelets independientes, (es decir de multiwavelets) para resolver esta limitación. De esta forma se presentará también un algoritmo de emparejamiento local basado en esta nueva formulación y se demostrará su

capacidad para emparejar contornos aun en la presencia de ocultamientos parciales por otros objetos.

#### 4.2.2 Representación invariante de contornos mediante multiwavelets

En esta sección se presentan los elementos básicos para construir una representación afín invariante del contorno del objeto mediante multiwavelets. Como se ha señalado anteriormente (sección 2.1.1.2.3) las transformaciones afines en el plano constituyen un espacio de seis grados de libertad y pueden representarse mediante una matriz cuadrada de orden 2 denotada por  $\mathbf{M}=(m_{ij})$  y un vector de traslación bidimensional  $\mathbf{t}$ .

Dada entonces una curva:  $\mathbf{r}(u) = (x(u), y(u))^T$  la parametrización natural afín se define como [Buchin 1983]:

$$s = \int \left| \frac{d\mathbf{r}}{du}, \frac{d^2\mathbf{r}}{d^2u} \right|^{1/3} du \quad \text{Ec. 4. 17}$$

donde  $|\cdot|$  denota el determinante de ambos vectores.

Esta parametrización es invariante frente a transformaciones equiafines ( $|\mathbf{M}|=1$ ) y permite definir la longitud afín de forma análoga a la longitud usual de una curva en el caso euclídeo.

Entonces se elige una familia de mutiwavelets ortogonal de orden dos obteniéndose una representación del contorno (sección 3.2.2):

$$\mathbf{r}(s) = \mathbf{C}_{0,0} \mathbf{F}_{0,0}(s) + \sum_{\substack{j=0 \\ 0 \leq l < 2^j}}^{J-1} \mathbf{D}_{j,l} \mathbf{F}_{j,l}(s) \quad \text{Ver Ec. 3. 56}$$

y se definen los coeficientes:

$$N_{j,l} = |\mathbf{D}_{j,l}|, \quad \text{donde } \mathbf{D}_{j,l} \text{ está definido :} \quad \text{Ec. 4. 18}$$

$$\mathbf{D}_{j,l} = \begin{pmatrix} \langle x(s), \mathbf{y}_{0;j,l}(s) \rangle & \langle y(s), \mathbf{y}_{0;j,l}(s) \rangle \\ \langle x(s), \mathbf{y}_{1;j,l}(s) \rangle & \langle y(s), \mathbf{y}_{1;j,l}(s) \rangle \end{pmatrix}$$

Así se tiene:

#### Proposición 4.1

Los coeficientes  $N_{j,l}$  definidos anteriormente son invariantes para la transformación equiafín  $\mathbf{r} = \mathbf{M}\mathbf{r}_0 + \mathbf{t}$  con  $|\mathbf{M}|=1$ .

#### Demostración

Basta comprobar que si  $|\mathbf{M}|=1$

$$\mathbf{M}\mathbf{r}(s) + \mathbf{t} = \mathbf{C}'_{0,0} \mathbf{F}_{0,0}(s) + \sum_{\substack{j=0 \\ 0 \leq l < 2^j}}^{J-1} \mathbf{D}'_{j,l} \mathbf{F}_{j,l}(s), \quad \mathbf{D}'_{j,l} = \mathbf{M}\mathbf{D}_{j,l} \text{ entonces basta aplicar } |\mathbf{AB}|=|\mathbf{A}||\mathbf{B}|$$

### 4.2.3 Algoritmo de emparejamiento

El algoritmo de emparejamiento [Pérez 1999b] está basado en la representación definida en la sección anterior. Hace uso del hecho de que los coeficientes de la representación invariante del contorno se calculan de forma local para obtener un algoritmo de emparejamiento local. Esto lo hace apropiado para emparejamientos parciales de formas u oclusiones. Este algoritmo hace uso de una generalización al caso afín de los mapas de coeficientes [Yoon 1998] provenientes de la transformada continua mediante wavelets.

#### Algoritmo de emparejamiento

Dados dos contornos  $r_1$  y  $r_2$  el algoritmo de emparejamiento procede como sigue:

##### Paso 1 (Parametrización afín)

Parametrizar  $r_1$  y  $r_2$  en su forma natural afín

##### Paso 2 (Cálculo de los coeficientes)

Evaluar los coeficientes  $N_{j,l}^i$  para los contornos  $r_i$   $i=1,2$ . Estos coeficientes se almacenan en los mapas de coeficientes  $C_i(j,l) = N_{j,l}^i$ ,  $i=1,2$ .

##### Paso 3 (Búsqueda de puntos de interés)

Realizar una búsqueda de puntos de interés en ambos mapas. (Un punto  $P$  se define como punto de interés para una medida  $\Lambda$  si  $\Lambda(P)$  es un máximo local en el mapa  $C$ )

##### Paso 3 (Emparejamiento basado en puntos de interés)

Dado el conjunto de puntos de interés  $P=\{P_1, P_2, \dots, P_n\}$  y  $Q=\{Q_1, Q_2, \dots, Q_m\}$  en los mapas  $C_1$  y  $C_2$ .

Definir  $M=\{M_{ij}, i=1, \dots, n \quad j=1, \dots, m\}$  como el conjunto de emparejamientos donde  $M_{ij}$  denota un emparejamiento entre un punto  $P_i$  en  $P$  y  $Q_j$  en  $Q$ .

Para cada emparejamiento  $M_{ij}$

Estimar los coeficientes de la matriz de transformación afín  $M$  y  $t$  a partir de las funciones wavelets y de escala.

Buscar el conjunto de emparejamientos en  $M$  que soporten dicha hipótesis  $Sop(M_{ij})$  (es decir que definan la misma transformación afín)

Si el número de emparejamientos es suficientemente grande

Segmentar  $r_1$  y  $r_2$  basado en los emparejamientos (Ver Figura 4.18)

Esto proporciona segmentaciones emparejadas:

$\text{Segm}_1 = \{S_{1;i}\}$  y  $\text{Segm}_2 = \{S_{2;i}\}$

Para cada uno de los segmentos emparejados  $S_{1;i}$  y  $S_{2;i}$

Evaluar el error de emparejamiento

Si el error de emparejamiento es mayor que un umbral

Retirar  $S_{1;i}$  y  $S_{2;i}$  de  $\text{Segm}_1$  y  $\text{Segm}_2$

Si  $(\text{longitud afín de } r_1 \text{ segmentado}) / (\text{longitud afín de } r_1) > \text{umbral del modelo}$  y  $(\text{longitud afín de } r_2 \text{ segmentado}) / (\text{longitud afín de } r_2) > \text{umbral de la imagen}$  entonces se ha encontrado un emparejamiento. Parar.

Tabla 4.7 Algoritmo de emparejamiento

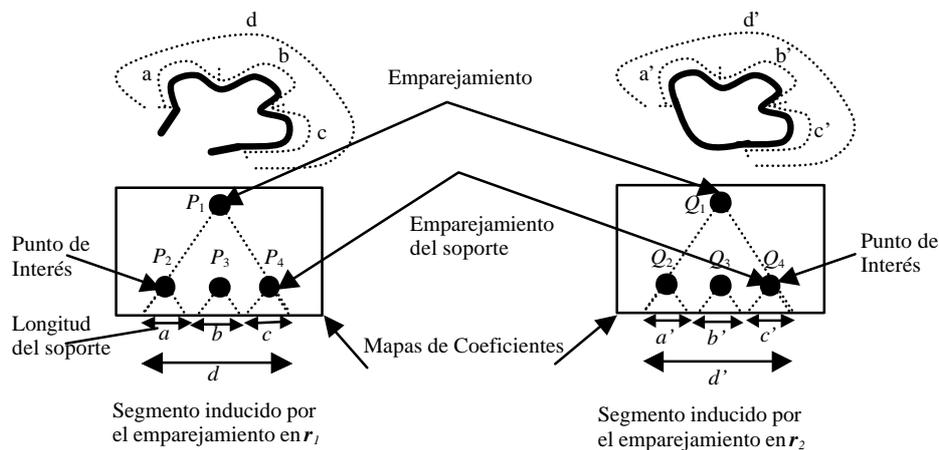


Figura 4.18 Segmentación inducida por el emparejamiento

Para determinar de forma completa el algoritmo hay varios elementos que deben ser definidos de forma explícita:

- La medida de interés  $\Lambda$

Esta medida se determina mediante una normalización en escala de los coeficientes  $N_{j,l}^i$  mediante la fórmula:

$$\Lambda(N_{j,l}^i) = \frac{|N_{j,l}^i|}{s_j^4}$$

Ec. 4.19

donde  $||$  denota el valor absoluto del coeficiente y  $s_j$  la longitud del soporte del multiwavelet ortogonal en la escala  $j$ .

- Estimación de la transformación afín y su comparación

Dado un emparejamiento  $M_{ij}$  es necesario estimar la transformación afín entre  $P_i$  y  $Q_j$ . Puesto que la transformación queda definida mediante una matriz cuadrada de orden 2  $M=(m_{ij})$  y un vector de traslación bidimensional  $t$  es necesaria la estimación de 6 parámetros. La estimación se realiza de la siguiente forma: asumiendo que una de las funciones de escala tiene integral nula, la matriz  $M$  se obtiene a partir de dicha función de escala y una de las funciones wavelet. el vector  $t$  se estima a partir de los coeficientes obtenidos de la función de escala con integral no nula.

Para obtener los emparejamientos soporte es necesario comparar transformaciones afines. Esto se hace como sigue:

La matriz  $M$  se descompone como:

$$M = S(I)R(q)D(a)R(m), \text{ donde:} \tag{Ec. 4. 20}$$

$$S(I) = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \text{ es una matriz de escala}$$

$$R(q) = \begin{pmatrix} \cos(q) & -\sin(q) \\ \sin(q) & \cos(q) \end{pmatrix} \text{ es una matriz de rotación}$$

$$D(a) = \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \text{ es una matriz de deformación}$$

La igualdad entre matrices se comprueba mediante la comparación de los ángulos  $q$  y  $m$  así como a partir de las diferencias entre los logaritmos de los parámetros  $I$  y  $a$  (en ambos casos la comparación se hace contra un umbral de error).

La igualdad entre los vectores de traslación se comprueba mediante la comparación de sus componentes (también contra un umbral de error).

- Evaluación del error de emparejamiento.

Una vez que se define la transformación entre modelo e imagen los coeficientes de la representación del contorno en el modelo se transforman mediante la matriz estimada  $M$  y el vector de traslación estimado  $t$ . Puesto que se utilizan multiwavelets ortogonales es posible calcular el error entre ambas representaciones de forma sencilla mediante la norma de su diferencia.

En la práctica para acelerar los cálculos no se utiliza el conjunto completo de coeficientes en el cálculo del error. En vez de esto se utiliza el coeficiente

multiwavelet no utilizado en el cálculo de la transformación afín para evaluar el error. Puesto que esta comparación se efectúa entonces a baja resolución, para evitar problemas debido a una evaluación por defecto del error se introduce el concepto de emparejamientos soporte en el algoritmo.

#### 4.2.4 Ejemplos

A continuación se presentan un conjunto de experimentos para evaluar el algoritmo. En primer lugar se presenta un ejemplo libre de oclusiones que puede resolverse mediante las formulaciones de emparejamiento afín previas. Posteriormente se presenta un ejemplo con oclusión que no puede resolverse mediante los algoritmos de emparejamiento previos basados en wavelets. La familia de multiwavelets utilizada ha sido la SA(4) (sección 3.1.14) y el número de puntos por contorno ha sido 256.

##### 4.2.4.1 Ejemplo sin oclusiones

En este primer ejemplo un conjunto de seis contornos de pájaros de le han presentado al algoritmo (ver Figura 4.19)

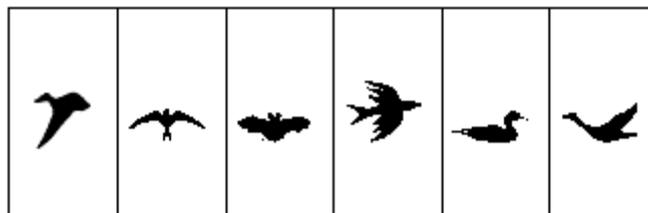


Figura 4.19 Contornos de las aves

Dichos contornos han sido transformados de forma aleatoria con los parámetros elegidos de forma uniforme en:  $I \in [5/4, 4/5]$ ,  $q \in [0, 2\pi]$ ,  $a \in [4/3, 3/4]$ ,  $m \in [0, 2\pi]$  de forma que se han generado un conjunto de 100 muestras. Los resultados del algoritmo de clasificación fueron:

Contorno	1	2	3	4	5	6
Clasificación correcta	97	92	97	100	99	95

Tabla 4.8 Resultados de clasificación

Los resultados muestran que prácticamente cada contorno pudo ser reconocido por el algoritmo. En la Figura 4.20 se muestran los mapas de coeficientes tanto para el modelo como para el contorno de la imagen en uno de los emparejamientos del experimento. Los puntos de interés se muestran en blanco y los valores de la medida de interés  $\Lambda$  se muestran en niveles de gris.

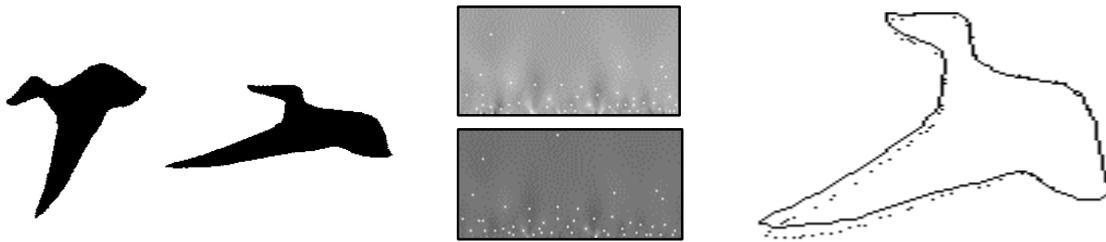


Figura 4.20 Contornos a emparejar (izquierda). Valores de la medida de interés y puntos de interés (centro). Resultado del emparejamiento con la posición predicha por el algoritmo en línea discontinua (derecha).

#### 4.2.4.2 Ejemplo con oclusiones

Para comprobar la capacidad del algoritmo para el emparejamiento frente a oclusiones parciales se han realizado un conjunto de experimentos. En la Figura 4.21 se muestra el resultado de uno de dichos experimentos.

En la dicha figura pueden verse dos imágenes obtenidas a partir de hojas de árbol. A continuación se muestran los contornos deformados por una transformación afín con parámetro de deformación  $a=4/3$  y superpuestos el uno sobre el otro. Por último en la derecha puede verse la localización de los dos modelos (mostrado en línea discontinua) en la imagen compuesta siendo el algoritmo capaz de recuperar la transformación afín entre modelo e imagen.

Es interesante señalar que el hecho de que la recuperación de una transformación afín sea más compleja que la de una similaridad euclídea produce que sea necesario obtener más segmentos de contorno no ocultos de cada figura que en el caso euclídeo para recuperar la transformación.



Figura 4.21 Contornos a emparejar (izquierda). Contornos transformados y superpuestos (centro). Resultado del emparejamiento con la posición predicha por el algoritmo en línea discontinua (izquierda).

Los resultados muestran que este algoritmo puede utilizarse para el proceso de emparejamiento del contorno en una imagen inicial. La utilización de multiwavelets permite obtener una representación invariante frente a transformaciones afines y una descripción local que la hace invariante frente a oclusiones.

## 5 Conclusiones, principales aportaciones y posibles extensiones

Las contribuciones más significativas de este trabajo se sitúan en dos aspectos diferenciados. De una parte, el soporte teórico construido en el Capítulo 3 posibilita la extensión de las representaciones de curvas basadas en wavelets a modelos deformables. Además la introducción de las representaciones de curvas basadas en multiwavelets proporciona la capacidad de obtener una nueva aproximación al problema de emparejamiento y reconocimiento de curvas basados en representaciones multiescala. De otra parte, los procedimientos construidos en base a dicha representación son utilizados en problemas de seguimiento y reconocimiento lo que permite disponer de elementos de utilidad práctica inmediata. A continuación se presentarán las conclusiones y principales aportaciones de este trabajo. Posteriormente se propondrán algunas extensiones y mejoras futuras de las representaciones y procedimientos desarrollados.

### 5.1 Conclusiones y principales aportaciones

Las conclusiones y principales aportaciones de este trabajo han sido las siguientes:

1.- Con el objetivo de orientar el desarrollo de este trabajo, se ha realizado previamente una revisión estructurada e integradora de los problemas de modelado, emparejamiento, clasificación y seguimiento de objetos. Para ello se han revisado las cuestiones básicas del modelado de objetos y examinado con detalle las representaciones propuestas en la literatura para el modelado de objetos basado en contornos haciendo hincapié en los modelos deformables que forman una de las bases de este trabajo. A continuación se ha abordado el problema del emparejamiento de objetos tanto para modelos rígidos como deformables haciendo una revisión de las técnicas y métodos más importantes. Posteriormente se ha abordado el problema de la clasificación de objetos y por último se han mostrado las técnicas más significativas de seguimiento de objetos basados en modelos con énfasis en los modelos deformables y modelos dinámicos.

2.- Para construir las representaciones de contornos basadas en wavelets y multiwavelets se ha realizado un examen de los elementos más importantes del análisis de funciones mediante wavelets. Para ello se ha tomado como punto de partida la transformada continua mediante wavelets obteniéndose las series de wavelets con su interpretación por medio de los análisis a múltiples resoluciones lo que permite obtener la transformada rápida mediante wavelets. Se han discutido las propiedades de aproximación mediante wavelets y se ha ligado la descomposición de funciones mediante esta representación con los espacios de suavidad de Besov. Posteriormente se ha examinado el tópico de derivación mediante wavelets y se ha pasado a la extensión del concepto de wavelets a multiwavelets.

3.- Una vez introducida la representación funcional mediante wavelets se introduce la representación de curvas mediante wavelets y se aporta la representación de curvas mediante multiwavelets realizándose una comparativa con otras representaciones previas. Se han obtenido resultados que indican un poder de compresión similar al de los wavelets con una mejor capacidad perceptual en la representación debido a la simetría de las funciones básicas.

4.- En esta memoria se aportan también los modelos de deformación de contornos en espacios de Besov mediante una generalización de los espacios de deformación de funciones. Se han discutido sus propiedades y resuelto el problema de ajuste de contornos en los espacios de Sobolev mediante técnicas de regularización dándose una interpretación de la solución del problema de ajuste en los distintos subespacios.

5.- Para inscribir el problema de regularización dentro del marco bayesiano se proponen en esta tesis modelos probabilísticos de deformación basados en wavelets mediante una extensión de la distribución gaussiana generalizada. En este marco se proporciona un método de resolución aproximada al problema de ajuste en espacios generales de Besov mediante la aproximación de Belge y se dan métodos exactos en espacios de Sobolev que permiten interpretar de forma probabilística los parámetros de regularización del apartado anterior. Este modelo teórico de ajuste se verifica experimentalmente mediante una serie de ejemplos señalándose como el conocimiento del grado de suavidad de la curva es una información de especial relevancia a la hora de resolver el problema de ajuste.

6.- La generalización de los modelos anteriores a problemas dinámicos hace que se propongan modelos autoregresivos de primer y segundo orden con innovación generada por variables de distribución gaussiana generalizada para modelar el conocimiento a priori de la dinámica del objeto. Además se proponen diversos métodos para el aprendizaje de los valores de los parámetros tanto el método general de separación en subespacios, como el método de la máxima verosimilitud en espacios de Sobolev.

7.- La aplicación de los modelos dinámicos anteriores en espacios de Sobolev proporcionan la utilización del filtro de Kalman en problemas de seguimiento. Para la resolución de dichos problemas se adapta el filtro de Kalman para la utilización de los modelos autoregresivos de segundo orden y se muestran algunos resultados computacionales. Éstos muestran como en ambientes controlados y con modelos precisos del objeto la utilización del filtro de Kalman permite seguir el objeto a lo largo de la imagen. Se muestran también cuales son las limitaciones del método: ambientes complejos o descripciones aproximadas del modelo de deformación del objeto.

8.- La aplicación de los modelos dinámicos en espacios de medidas no gaussianas que aparecen cuando hay un ruido sustancial en el fondo ha sido también objeto de estudio. Se ha adaptado el filtro Condensation para la utilización de los modelos autoregresivos de segundo orden basados en wavelets y se muestran algunos resultados

computacionales. Éstos muestran como en ambientes complejos y con modelos precisos del objeto la utilización del filtro Condensation con la representación basada en wavelets es capaz de seguir a los objetos a lo largo de la imagen.

9.- Por último, se ha abordado el uso de la representación basada en multiwavelets para problemas de reconocimiento. Se ha presentado un algoritmo que utiliza información local para realizar el problema de emparejamiento, capaz por tanto de ser robusto frente a oclusiones, y se han discutido los resultados computacionales obtenidos sobre una serie de imágenes de ejemplo.

## **5.2 Posibles extensiones y líneas de desarrollo**

A continuación se presentan las principales extensiones y líneas de desarrollo de los elementos presentados en esta memoria:

### **1.- Generalizar los modelos de deformación a multiwavelets**

Los modelos de deformación que se han presentado están basados en descomposiciones del contorno basados en wavelets. Dada la mejora en la representación perceptual que introducen los multiwavelets una posible línea de desarrollo de este trabajo sería la extensión de dichos modelos de deformación a los multiwavelets.

### **2.- Introducir los subespacios de deformación en los modelos de espacios de formas.**

Los modelos de deformación en espacios de formas proporcionan una forma potente de precisar las deformaciones permitidas en el modelo. Una vía interesante de trabajo es la unión de los subespacios de traslación, similaridades o afines con los subespacios que provienen de la proyección de la curva de referencia mediante un análisis a múltiples resoluciones. De esta forma se dotaría a los espacios de forma de información de deformación además de la puramente geométrica.

### **3.- Estudiar modelos dinámicos no lineales**

Los modelos lineales autoregresivos presentados tienen dificultades a la hora de modelar objetos que presentan cambios bruscos en su velocidad o dirección de movimiento. Es por ello interesante el estudio de otro tipo de modelos dinámicos que sean capaces de abordar estas situaciones. Además la utilización de modelos dinámicos con distribución asintótica no gaussiana puede ser una alternativa a explorar fuera de los espacios de Sobolev.

### **4.- Utilizar técnicas de muestreo basadas en importancia en el filtro Condensation**

El filtro Condensation ha sido enriquecido recientemente (filtro Icondensation [Isard 1998b]) con una variante cuya técnica de muestreo está basada en muestreo por importancia. Esta variante permite que el filtro supere uno de sus principales problemas:

la pérdida de diversidad en el conjunto poblacional  $S$ . La aplicación de las técnicas desarrolladas con este nuevo filtro deben de incrementar la robustez del modelo.

5.- Utilizar métodos de seguimiento multiescala

Las técnicas multiescala, de amplia aplicación en la Visión por Ordenador, permiten incrementar la robustez y disminuir la velocidad de cómputo de los métodos que las utilizan. La aplicación de técnicas multiescala ligando la descomposición a múltiples resoluciones del modelo con la descripción a múltiples resoluciones de la imagen deben permitir incrementar el rendimiento de las técnicas presentadas.

---

## Referencias

- [Abramovich 1998] F. Abramovich, T. Sapatinas, B.W. Silverman. Wavelet thresholding via a Bayesian approach. *J. R. Statist. Soc. B*, 60, 725-749, 1998.
- [Alt 1999] H. Alt and L. Guibas. Discrete geometric shapes: Matching, interpolation, and approximation. J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 121-153. Elsevier Science Publishers B.V. North-Holland, 1999.
- [Armstrong 1995] M. Armstrong, A. Zisserman. Robust object tracking. *Proc. Asian Conference on Computer Vision*, 1, 58-61, 1995.
- [Ballard 1982] D.H. Ballard, C.M. Brown. *Computer Vision*. Prentice-Hall, 1982.
- [Banham 1996] M. R. Banham, A. K. Katsaggelos. Spatially adaptive wavelet-based multiscale image restoration. *IEEE Trans. Image Processing*, 5(4), 619-633, 1996.
- [Bar-Shalom 1988] Y. Bar-Shalom, T.E. Fortmann. *Tracking and Data Association*. Academic Press Limited, San Diego, 1988.
- [Belge 2000] M. Belge, E. L. Miller, M. Kilmer, Wavelet domain image restoration with adaptive edge-preserving regularization, *IEEE Trans. on Image Processing*, 9(4), 598-608, 2000.
- [Blake 1998] A. Blake, M. Isard. *Active Contours*, Springer-Verlag, 1998.
- [Boor 1978] C. De Boor. *A practical Guide to Splines*. Springer Verlag, 1978.
- [Bow 1992] S.T. Bow. *Pattern recognition and image preprocessing*. Marcel Dekker Inc., 1992.
- [Box 1973] G. Box, G. Tiao. *Bayesian inference in statistical analysis*. Addison-Wesley, 1973.
- [Breuel 1992] T. M. Breuel. Fast recognition using adaptive subdivisions of transformation space. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 445-451, 1992.
- [Breuel 1993] T. M. Breuel. Higher-order statistics in object recognition. *Technical Report 93-02, IDIAP*, 1993
- [Brigger 2000] P. Brigger, J. Hoeg, M. Unser. B-spline snakes: a flexible tool for parametric contour detection. *IEEE Trans. Image Processing*, 9(9), 1484-1496, 2000.
- [Buchin 1983] S. Buchin. *Affine Differential Geometry*. Science Press, 1983.
- [Burr 1981] D.J. Burr. Elastic matching of line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(6), 708-713, 1981.
- [Canny 1986] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679-698, 1986.
- [Carmona 1995] R.A. Carmona, W.L. Hwang, R.D. Frostig. Wavelet analysis for brain- function imaging. *IEEE Transactions on Medical Imaging*, 14(3), 556-564, 1995.

- [Cheick 2000] F.A. Cheick , A. Quddus, M. Gabbouj. Contour-based object recognition using wavelet transform. Proc. of the European Signal processing conference, 2000.
- [Cheung 1995]. K.-W. Cheung, D.-Y. Yeung, R. T. Chin. A unified framework for handwritten character recognition using deformable models. Proc. Second Asian Conference on Computer Vision, (1), 344-348, 1995.
- [Cheung 1998] K.W. Cheung, D.Y. Yeung, R.T. Chin. On deformable models for visual pattern recognition. Technical Report HKUST-CS98-8, Department of Computer Science, Hong Kong University of Science and Technology, 1998.
- [Choi 1999] H. Choi, R. Baraniuk. Wavelet-domain statistical models and Besov spaces. Proc. of SPIE Technical conference on Wavelet Applications in Signal Processing VII, Denver, 1999.
- [Chuang 1996] G.C.H. Chuang, C.C.J. Kuo. Wavelet descriptor of planar curves: Theory and Applications. IEEE Trans. on Image Processing, 5(1), 56-70, 1996.
- [Chui 1996] C.K. Chui, J. Lian. A study on orthonormal multiwavelets. Appl. numer. Math. (20), 273-298, 1996.
- [Cooper 1979] D.Cooper. Maximum likelihood estimation of Markov process blob boundaries in noisy Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1, 372-384, 1979.
- [Cootes 1992] T.F.Cootes, C.J.Taylor. Active Shape Models - 'Smart Snakes'. Proc. British Machine Vision Conference, 266-275, Springer-Verlag, 1992.
- [Dickinson 1992] S.J. Dickinson, A.P. Pentland, A. Rosenfeld. 3-D Shape Recovery Using Distributed Aspect Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence. 14(2). 1992.
- [Doucet 2000] A. Doucet, S.J. Godsill, C. Andrieu. On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. Statistics and Computing, 10(3), 197-208, 2000.
- [Dubuisson 1996] M. P. Dubuisson, A. Jain, S. Lakshmanan. Vehicle segmentation using deformable templates. IEEE Trans. Pattern Anal. and Machine Intelligence, 18(3), 293-308, 1996.
- [Duda 1972] R. O. Duda, P.E. Hart. Pattern Classification & Scene Analysis. John-Wiley & Sons, 1972.
- [Freedman 1999] D. Freedman, M. Brandstein. A subset approach to contour tracking in clutter. Proceedings ICCV, 242-247, 1999.
- [Freeman 1961] H. Freeman. On the encoding of arbitrary geometric configurations. IRE Trans. Elec. Computers, 10, 260-268, 1961.
- [Gatica-Pérez 1999] D. Gatica-Pérez, F. García-Ugalde. Compact shape representation of planar curves based on a wavelet shape descriptor for multimedia applications Vision Interface, 1999.
- [Geiger 1995] D. Geiger, A. Gupta, L.A. Costa, J. Vlontzos. Dynamic programming for detecting, tracking, and matching deformable contours. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(3), 294-302, 1995.

- 
- [Gelb 1974] A. Gelb, Applied Optimal Estimation, M.I.T. Press, 1974.
- [Gonzalez 1987] R. C. Gonzalez, P. Wintz. Digital Image Processing. AddisonWesley, 1987.
- [Gorman 1988] J.W. Gorman, O.R. Mitchell, F.P. Kuhl, F.P., Partial Shape Recognition Using Dynamic Programming. IEEE Transactions on Pattern Analysis and Machine Intelligence, 10(2), 257-266, 1988.
- [Grimson 1987] W. Grimson, T. Lozano-Pérez. Localizing Overlapping Parts by Searching the Interpretation Tree. IEEE Transactions on Pattern Analysis and Machine Intelligence, 9(4), 1987.
- [Harris 1992] C. G. Harris. Tracking with rigid models. A. Blake and A. Yuille, editors, Active Vision. MIT Press, Cambridge, MA, 1992.
- [Hildreth 1987] E. Hildreth. Edge Detection. S.C. Shapiro, editor. Encyclopedia of Artificial Intelligence, John Wiley, 1, 257-267, 1987.
- [Hu 1962] M.K. Hu. Visual Pattern Recognition by moment invariants. IRE transactions on Information theory, 8(2), 179-187, 1962.
- [Huttenlocher 1990] D. Huttenlocher, S. Ullman. Recognizing 3D solid objects by alignment with an image. Int. J. of Comp. Vision, 5(2),195-212, 1990.
- [Huttenlocher 1992] Daniel P. Huttenlocher, Jae J. Noh, and William J. Rucklidge. Tracking non-rigid objects in complex scenes. Technical Report CUCS TR 92-1320, Department of Computer Science, Cornell University, 1992.
- [Huttenlocher 1993] D. P. Huttenlocher, G. A. Klauderman, W. J. Rucklidge, Comparing images using the Hausdorff-distance, IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(9), 850-863, 1993.
- [Imai 1988] H. Imai and M. Iri. Polygonal approximations of a curve-formulations and algorithms. G. T. Toussaint, editor, Computational Morphology, 71-86. North-Holland, Amsterdam, 1988.
- [Isard 1998] M. Isard and A. Blake. ICONDENSATION: unified lowlevel and high-level tracking in a stochastic framework. Proc. European Conf. on Computer Vision, 893-908, 1998
- [Isard 1998] M. Isard. Visual motion analysis by probabilistic propagation of conditional density. PhD thesis, Department of Engineering Science, University of Oxford.
- [Jacob 2000] M. Jacob, T. Blu, M. Unser. Exact Computation of Area Moments for Spline and Wavelet Curves. Proceedings of the Fifteenth International Conference on Pattern, 2000.
- [Jain 1989] A.K. Jain. Fundamentals of digital image processing. Prentice Hall, 1989.
- [Jawerth 1994] B. Jawerth, W. Sweldens. An overview of wavelet based multiresolution analyses. SIAM Rev. 36(3), 377-412, 1994.
- [Kambhamettu 1994] C. Kambhamettu, D. B. Goldgof, D. Terzopoulos, Thomas S. Huang. Nonrigid motion analysis. T. Young, editor, Handbook of PRIP: Computer vision, 2 , 405-430, Academic Press, San Diego, 1994.
- [Kass 1987] M. Kass, A. Witkin, D. Terzopoulos. Snakes: active contour models. First International Conference on Computer Vision, 259-268, 1987.

- [Kervrann 1994] C. Kervrann, F. Heitz. A hierarchical statistical framework for the segmentation of deformable objects in image sequences. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 724-727, Seattle, 1994.
- [Khotanzad 1990] A. Khotanzad, J. Lu. Classification of invariant image representations using a neural network. IEEE Trans. on signal processing, 6, 1028-1038, 1990
- [Knoll 1999] C. Knoll, M. Alcañiz, V. Grau, C. Monserrat, M.C. Juan .Outlining of the prostate using snakes with shape restrictions based on the wavelet transform, Pattern Recognition, 32, 1767-1781, 1999.
- [Kruskal 1983] J. B. Kruskal. An overview of sequence comparison. D. Sankoff and J. B. Kruskal, editors, Time warps, string edits, and macromolecules, 1-44, Addison-Wesley, Amsterdam, Netherlands, 1983.
- [Lades 1993] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Wurtz, W. Konen. Distortion Invariant Object Recognition in the Dynamic Link Architecture. IEEE Transactions on Computers, 42(3), 300-311, 1993.
- [Lai 1994] K.F. Lai, Deformable contours: Modelling, extraction, detection and clasification, PhD thesis, Electrical and Computing Engineering Department, University of Wisconsin-Madison, 1994.
- [Lamdan 1988] Y. Lamdan, H. J. Wolfson. Geometric hashing: A general and efficient model-based recognition scheme. Proc. International Conference on Computer Vision, 238- 249, 1988.
- [Lanitis 1993] A. Lanitis. Automatic Location of Facial Characteristics using Genetic Algorithms. Wolfson Image Analysis Unit Internal Report, June 1993.
- [Lewis 1992] A. Lewis, G. Knowles. Image compression using 2-D wavelet transform. IEEE Transactions on Image Processing, 1(2), 244--250, 1992.
- [Leymarie 1989] F. Leymarie, M.D. Levine, Shape features using curvature morphology. Visual Communications and Image Processing IV, 390-401,1989.
- [Loncaric 1998] S. Loncaric. A survey of shape analysis techniques. Pattern Recognition, 31(8), 983-1001, 1998.
- [Lowe 1992] D. G. Lowe, Robust Model-based Motion Tracking Through the Integration of Search and Estimation. International Journal of Computer Vision, 8(2), 113-122, 1992.
- [MacKay 1992] D.J.C. MacKay. Bayesian interpolation. Neural Computation, 4(3):415--447, May 1992.
- [Mallat 1989] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), 674-693, 1989.
- [Mallat 1992] S G Mallat, S. Zhong Characterization of signals from multiscale edges. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(7), 710-732, 1992.
- [Marchand 1999] E. Marchand, P. Bouthemy, F. Chaumette, and V. Moreau. Robust real-time Visual Tracking Using a 2D-3D Model-Based Approach. International Conference on Computer Vision, 262-268, September 1999.
- [Marr 1980] D. Marr, E. Hildreth. Theory of Edge Detection. Proceedings of the Royal Society of London, B, 207, 187-217, 1980.

- [Martelli 1972] A. Martelli. Edge detection using heuristic search methods. *Computer Graphics and Image Processing*, 1,169-182, 1972.
- [McInerney 1996] T. McInerney, D. Terzopoulos. Deformable models in medical images analysis: a survey. *Medical Image Analysis*, 1(2),91-108, 1996.
- [Méndez 1984] J. A. Méndez, A. Falcón. On the orthogonal expansions of images for low level recognition. *Proc. IEEE first Conf. on A.I. Applications*, 1984.
- [Mokhtarian 1992] F. Mokhtarian, A.K. Mackworth. A theory of multiscale, curvature based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8), 789-805, 1992.
- [Mokhtarian 1996] F. Mokhtarian, S. Abbasi, Josef Kittler. Efficient and robust retrieval by shape content through curvature scale space. *Proc. International Workshop on Image Databases and MultiMedia Search*, 35-42, 1996.
- [Müller 1999] K. Müller, J.-R. Ohm. Contour description using wavelets. *Workshop on Image Analysis for Multimedia Interactive Services*, 1999.
- [Mundy 1992] J. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press, 1992.
- [Nastar 1994] C. Nastar. Vibration modes for nonrigid motion analysis in 3D images. *Proc. of the Third European Conference on Computer Vision*, 1994.
- [Neuenschwander 1994] W. Neuenschwander, P. Fua, G. Szekely, O. Kubler. Initializing snakes. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 658-663, 1994.
- [O'Rourke 1985] J. O'Rourke. *Curve similarity with signatures*, 295-317. North-Holland 1985.
- [Pan 1996] H. Pan. *General Stereo Matching Using Symmetric Complex Wavelets Wavelet Applications in Signal and Image Processing IV* 1996.
- [Parizeau 1998] M. Parizeau, N. Ghazzali, J. F. Hebert, Optimizing the Cost Matrix for Approximate String Matching Using Genetic Algorithms, *Pattern Recognition* 31 (4), 431-440, 1998.
- [Park 1999] J.S. Park, D. H. Chang. 2-D Invariant Descriptor for shape based image retrieval. 26(2), 554-556,1999.
- [Pavlidis 1977] T. Pavlidis, *Structural Pattern Recognition*. Springer Verlag, 1977.
- [Peerson 1977] E. Peerson , K. Fu. Shape discrimination using Fourier descriptors. *IEEE Trans Sys, Man, and Cybernetics* 7(3) ,170-178, 1977.
- [Pérez 1999a] F. Pérez, A. Falcón. Detección Invariante frente a transformaciones afines de puntos dominantes mediante wavelets. *Actas del IV Simposio Iberoamericano de Reconocimiento de Patrones*, 345-351, La Habana, Cuba, 1999.
- [Pérez 1999b] F. Pérez, A. Falcón. Affine contour matching based on Multiwavelets. N.M. Namazi editor. *Signal and Image Processing*. ISBN: 0-88986-271-0 (cd-rom), 1999.
- [Pérez 2000a] F. Pérez, A. Falcón. A wavelet contour representation for fitting and tracking applications. *Proc. of Engineering of Intelligent Systems*. Paisley. Escocia. ICSC Academic Press ISBN: 3-906454-21-5 (cd-rom), 2000.

- [Pérez 2000b] F. Pérez, A. Falcón. Planar shape representation based on multiwavelets. Proc. of the X European Signal Processing Conference. Tampere, Finlandia. ISBN: 952-15-0443-9 (cd-rom), 2000.
- [Pérez 2000c] F. Pérez, A. Falcón. Contour tracking in complex environments using wavelet basis. Proc of the International Conference on Modelling and Simulation, 671-678 Las Palmas de Gran Canaria, 2000.
- [Pérez 2001] F. Pérez, A. Falcón. Dynamic models for wavelet representations of Shape. C. Arcelli, L.P. Cordella, G. Sanniti di Baja editors, Visual Form 2001, Lecture Notes in Computer Science LNCS-2059, 431-439, Springer-Verlag, 2001.
- [Petland 1991] A. Pentland, 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.
- [Pfeiffer 1995] M. Pfeiffer, M. Pandit. A parametrical description of plane curves using wavelet descriptors, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 4, 2531-2534, 1995.
- [Poggio 1985] T. Poggio, V. Torre, C. Koch. Computational vision and regularization theory, Nature, 317, 600-610, 1985.
- [Pope 1994] R. Pope. Model-Based Object Recognition: A Survey of Recent Research. Technical Report 94-04, The University of British Columbia - Department of Computer Science, 1994.
- [Prewitt 1970] J.M.S. Prewitt. Object enhancement and extraction. B.S. Likin, A. Rosenfeld editors Picture Processing and Psychopictorics, 75-149., Academic Press., 1970.
- [Quddus 2000] A. Quddus, M. Gabbouj. Selection of natural scale in discrete wavelet domain using eigenvalues. Proceedings of the European Signal processing conference 2000.
- [Resnikoff 1998] H.L. Resnikoff, R.O. Wells Jr. Wavelet Analysis: The Scalable Structure of Information. Springer-Verlag, New York, 1998.
- [Restrepo 1997] Juan Mario Restrepo, Gary K. Leaf. Inner Product Computations Using Periodized Daubechies Wavelets. International Journal of Numerical Methods in Engineering, 40, 3557-3578, 1997.
- [Roberts 1965] L. G. Roberts. Machine Perception of Three-Dimensional Solids. Tippett, et al. editors, Optical and Electro-optical Information Processing, MIT Press, 159-197, 1965.
- [Rosin 1997] P. L. Rosin. Techniques for assessing polygonal approximations of curves. IEEE Transactions On Pattern Analysis and Machine Intelligence, 19(6), 659-666, 1997.
- [Russ 1992] J. C. Russ. The image processing handbook. CRC Press Inc, 1992
- [Saund 1992] E. Saund. Putting knowledge in visual shape representation. Artificial Intelligence, 54, 71-119, 1992.
- [Sclaroff 1995] S. Sclaroff, A. Pentland. Modal matching for correspondence and recognition. IEEE Transactions On Pattern Analysis and Machine Intelligence, 17(6), 545-561, 1995.
- [Shahraray 1989] B. Shahraray, D. J. Anderson, Optimal Estimation of Contour Properties by Cross-Validated Regularization. IEEE Transactions On Pattern Analysis and Machine Intelligence. 11, 600-610, 1989.

- 
- [Sonka 1993] M. Sonka, V. Hlavac, R. Boyle. Image processing, analysis and machine vision. Chapman and Hall, London, 1993
- [Staib 1992] L.H. Staib, J.S. Duncan. Boundary finding with parametrically deformable models. IEEE Transactions On Pattern Analysis and Machine Intelligence, 14(11), 1061-1075, 1992.
- [Stiller 1999] C. Stiller, J. Konrad, Estimating motion in image sequences: A tutorial on modeling and computation of 2D motion, IEEE Signal Process. Mag., 16, 70-91, 1999.
- [Strela 1996] V.Strela, P.Heller, G.Strang, P.Topiwala, C.Heil. The application of multiwavelet filter banks to signal and image processing. IEEE Trans. on Image Processing ,1996.
- [Sweldens 1994] W. Sweldens and R. Piessens. Quadrature Formulae and Asymptotic Error Expansions for wavelet approximations of smooth functions. SIAM J. Numer. Anal., 31, 1240-1264, 1994.
- [Szeliski 1997] R. Szeliski, J. Coughlan. Spline-Based Image Registration. International Journal of Computer Vision, 22(3), 199-218, 1997.
- [Tang 1996] Y. Tang, H. Ma, B. Li, J. Liu. Two-dimensional wavelet transform in document analysis. Proceedings of the 1st International Conference on Multimodal Interface, 274-279, 1996.
- [Teague 1980] M. Teague. Image analysis via the general theory of moments. J. Opt. Soc. Am. 70, 920-930, 1980
- [Terzopoulos 1992] D. Terzopoulos, R. Szeliski. Tracking with Kalman Snakes. A. Blake, A. Yuille editors, Active Vision. MIT Press, 1992.
- [Tham 2000] J. Y. Tham, L. Shen, S. L. Lee and H. H. Tan. General Approach for Analysis and Application of Multiwavelet Transforms. IEEE Trans. Signal Processing, 48(2), 457-464, 2000.
- [Thompson 1987] D. W. Thompson, J. L. Mundy. Three-dimensional model matching from an unconstrained viewpoint. Proceedings of IEEE Conference on Robotics and Automation, 208-220, 1987.
- [Tieng 1997] Q. M. Tieng and W. Bowles. Wavelet Based Affine Invariant Representation: A Tool for Recognizing Planar Objects in 3D Space. IEEE Trans. Pattern Analysis and Machine Intelligence, 19(8), 846-857, 1997.
- [Tsai 1985] W. Tsai, S. Yu. Attributed string matching with merging for shape recognition. IEEE Transactions On Pattern Analysis and Machine Intelligence 7, 453-462, 1985.
- [Ueda 1993] N. Ueda, S. Suzuki. Learning visual models from shape contours using multiscale convex/concave structure matching. IEEE Transactions On Pattern Analysis and Machine Intelligence 15(4), 337-352, 1993.
- [Unser 1993] M. Unser, A. Aldroubi, M. Eden, B-spline signal processing: Part II-efficient design and applications, IEEE Trans. Signal Processing. 41(2), 834-847, 1993.
- [Veltkamp 1999] R. C. Veltkamp, M. Hagedoorn. State-of-the-art in shape matching. Tech. Report ( UU-CS-1999-27). University of Utrecht. 1999.
- [Vernon 1987] D. Vernon. Two dimensional object recognition using partial contours. Image an Vision computing, 5(1), 21-27, 1987

- [Walter 1992] G. G. Walter. A sampling theorem for wavelet subspaces, *IEEE Trans. Inform. Theory*, 38, 881-884, 1992.
- [Wang 1999] Y. P. Wang, S. L. Lee, K. Torachi, Multiscale curvature based shape representation using B-spline wavelets, *IEEE Trans. Image Processing*, 8(11), 1586-1592, 1999.
- [Williams 1992] D.J. Williams and M. Shah. A fast algorithm for active contours and curvature estimation. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 55(1), 14-26, 1992.
- [Williams 1994] C. K. I. Williams, Combining deformable models and neural networks for handprinted digit recognition. PhD thesis, Dept. of Computer Science, University of Toronto, 1994.
- [Witkin 1983] A. P. Witkin, Scale Space Filtering. *Proceedings of the International Joint Conference on Artificial Intelligence*, 1019-1022, 1983.
- [Wojtaszczyk 1997] P. Wojtaszczyk. *A Mathematical Introduction to Wavelets*. London Mathematical Society Student Texts, 37, Cambridge University Press, Cambridge, 1997.
- [Wunsch 1995] P. Wunsch, A. Laine. Wavelet descriptors for multiresolution recognition of handprinted characters. *Pattern Recognition* 28(8), 1237-1249, 1995.
- [Xie 1994] X. Xie, R. Sudhakar, H. Zhuang. On improving eye feature extraction using deformable templates. *Pattern Recognition*, 27(6), 791-799, 1994.
- [Yen 1996] S. C. Yen, L. H. Finkel. Salient Contour Extraction by Temporal Binding in a Cortically-Based Network. M. Jordan, M. Mozer, T. Petsche (Editors). *Advances in Neural Information Processing Systems 9*. MIT Press. 1996.
- [Yoon 1998] S. H. Yoon, J.H. Kim, W.E. Alexander, S.M. Park, K.H. Sohn. An optimum solution for scale-invariant object recognition based on the multiresolution approximation. *Pattern Recognition*, 31(7), 889-908, 1998.
- [Yoshida 1997] H. Yoshida, S. Katsuragawa, Y. Amit, K. Doi. Wavelet Snake for Classification of Nodules and False Positives in Digital Chest Radiographs. *Proc IEEE Engineering in Medicine and Biology Society*, 509-512, 1997.
- [Yuille 1992] A. Yuille, P. Hallinan, D. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision* 8 (2), 99-111, 1992.
- [Zahn 1972] T. Zahn, R. Z. Roskies. Fourier descriptors for plane closed curves. *IEEE-T on computers*, C-21(3), 269-281, 1972.
- [Ziou 1998] D. Ziou, S. Tabbone, Edge Detection Techniques - An Overview , *Int. Journal of Pattern Recognition and Image Analysis*, 8(4), 537-559, 1998.

## Apéndice A

A continuación se demuestran los teoremas 3.2, 3.3 y 3.4

### Lema A.1

Dada la función:

$$g(m) = \frac{(1+m)^p}{1+m^p}, m \geq 0$$

Si  $p \geq 1$  entonces  $1 \leq g(m) \leq 2^{p-1}$

Si  $p < 1$  entonces  $2^{p-1} \leq g(m) \leq 1$

### Demostración

Si  $p=1$  la demostración es trivial, supongamos que  $p \neq 1$

La función  $g$  está bien definida y es derivable en  $[0, \infty)$  entonces puede comprobarse que

$g'(m) = 0 \Leftrightarrow m=1$  que es un máximo para  $p > 1$  y un mínimo para  $p < 1$

Por tanto: si  $p > 1$   $g(m) \leq g(1)=2^{p-1}$ , si  $p < 1$   $g(m) \geq g(1)=2^{p-1}$

Además:

$$\lim_{m \rightarrow \infty} g(m) = \lim_{m \rightarrow \infty} \frac{(1+m)^p}{1+m^p} = 1 \text{ y } g(0)=1 \text{ con lo que se completa la demostración}$$

### Lema A. 2

Dado  $p > 0$  existen  $C_p, D_p$  constantes positivas (que dependen de  $p$ ) de forma que:

$$C_p (|x|^p + |y|^p) \leq (|x| + |y|)^p \leq D_p (|x|^p + |y|^p)$$

### Demostración

Si  $p=1$  basta tomar  $C_p=D_p=1$ , supongamos entonces  $p \neq 1$ .

Si  $|x|=0$  y  $|y|=0$  cualquier valor de  $C_p, D_p$  es válido.

Si  $|x|=0$  entonces cualquier valor  $C_p \leq 1, D_p \geq 1$  es válido.

Supongamos entonces  $|x| \neq 0$

Definamos:

$$f(x, y) = \frac{(|x| + |y|)^p}{|x|^p + |y|^p}, \quad \text{entonces } f(x, y) = \frac{(1 + |y/x|)^p}{1 + |y/x|^p} = g(|y/x|)$$

Por tanto (Lema A.1)

Si  $p > 1$  entonces  $1 \leq f(x, y) \leq 2^{p-1}$  si  $|x| \neq 0$  y la cota también es válida para  $|x|=0$  de acuerdo con el caso anterior.

Si  $p < 1$  entonces  $2^{p-1} \leq f(x, y) \leq 1$  si  $|x| \neq 0$  y la cota también es válida para  $|x|=0$  de acuerdo con el caso anterior.

### Lema A.3

Se verifica:

- a)  $|x|^p \leq (|x| + |y|)^p, \quad |y|^p \leq (|x| + |y|)^p$
- b)  $C_0(|x|^p + |y|^p) \leq (|x| + |y|)^p, \quad C_0 = 1/2$

### Demostración

- a) Evidente
- b) Utilizando a) o bien el lema anterior.

### Lema A.4

Se verifica:

$$C_{1/2}(|x| + |y|) \leq \sqrt{(|x|^2 + |y|^2)} \leq D_{1/2}(|x| + |y|)$$

### Demostración

Basta tomar  $p=1/2$  y sustituir  $|x|^2, |y|^2$  en el lema A.2

### Lema A.5

Existen  $A_p, B_p$  constantes positivas de forma que:

$$A_p(|x|^p + |y|^p) \leq \left(\sqrt{(|x|^2 + |y|^2)}\right)^p \leq B_p(|x|^p + |y|^p)$$

### Demostración

Por el lema anterior:  $C_{1/2}(|x| + |y|) \leq \sqrt{(|x|^2 + |y|^2)} \leq D_{1/2}(|x| + |y|)$  entonces:

$$\begin{aligned} \left(\sqrt{|x|^2 + |y|^2}\right)^p &\leq (D_{1/2})^p (|x| + |y|)^p \leq (D_{1/2})^p D_p (|x|^p + |y|^p) \\ C_p (C_{1/2})^p (|x|^p + |y|^p) &\leq (C_{1/2})^p (|x| + |y|)^p \leq \left(\sqrt{|x|^2 + |y|^2}\right)^p \end{aligned}$$

**Teorema A.1**

Las siguientes normas son equivalentes:

$$\|r\|_{L_p(I)} \equiv \|x\|_{L_p(I)} + \|y\|_{L_p(I)}$$

**Demostración**

El caso  $p=\infty$  sale de forma trivial utilizando el Lema A.3 y A.4

En el caso  $p<\infty$ :

$$\|r\|_{L_p(I)} = \left( \int_I \left| \sqrt{x(u)^2 + y(u)^2} \right|^p du \right)^{1/p}$$

entonces utilizando el lema anterior:

$$A_p \int_I (|x(u)|^p + |y(u)|^p) du \leq \int_I \left( \sqrt{|x(u)|^2 + |y(u)|^2} \right)^p du \leq B_p \int_I (|x(u)|^p + |y(u)|^p) du$$

entonces:

$$\begin{aligned} \left( \int_I \left( \sqrt{|x(u)|^2 + |y(u)|^2} \right)^p du \right)^{1/p} &\leq (B_p)^{1/p} \left( \int_I |x(u)|^p du + \int_I |y(u)|^p du \right)^{1/p} \\ &\leq (B_p)^{1/p} D_p \left( \left( \int_I |x(u)|^p du \right)^{1/p} + \left( \int_I |y(u)|^p du \right)^{1/p} \right) \end{aligned}$$

la otra desigualdad se resuelve de manera análoga.

**Proposición A.1**

Existen  $E_p, F_p$  constantes positivas de forma que:

$$E_p \left( \|x\|_{B_q^s(L_p(I))} + \|y\|_{B_q^s(L_p(I))} \right) \leq \|r\|_{B_q^s(L_p(I))} \leq F_p \left( \|x\|_{B_q^s(L_p(I))} + \|y\|_{B_q^s(L_p(I))} \right)$$

**Demostración:** Basta utilizar la demostración anterior sobre  $w_m(r,u)_p$  y aplicar las acotaciones de los Lemas A.2 y A.3

**Teorema 3.2**

Las siguientes normas son equivalentes:  $\|r\|_{B_q^a(L_p(I))} \equiv \|x\|_{B_q^a(L_p(I))} + \|y\|_{B_q^a(L_p(I))}$

**Demostración:**

$$\|r\|_{B_q^a(L_p(I))} = \|r\|_{L_p(I)} + |r|_{B_q^a(L_p(I))} \leq H_p (\|x\|_{L_p(I)} + \|y\|_{L_p(I)}) + F_p (|x|_{B_q^a(L_p(I))} + |y|_{B_q^a(L_p(I))}) \leq \max(H_p, F_p) (\|x\|_{B_q^a(L_p(I))} + \|y\|_{B_q^a(L_p(I))})$$

la otra desigualdad es similar.

**Lema A.6**

Existen dos constantes  $K, L$  tal que:

$$\begin{aligned} & K \left( |c_{0,0;x}| + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;x}|^p \right)^{q/p} \right)^{1/q} + |c_{0,0;y}| + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;y}|^p \right)^{q/p} \right)^{1/q} \right) \leq \\ & \sqrt{|c_{0,0;x}|^2 + |c_{0,0;y}|^2} + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right)^p \right)^{q/p} \right)^{1/q} \leq \\ & L \left( |c_{0,0;x}| + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;x}|^p \right)^{q/p} \right)^{1/q} + |c_{0,0;y}| + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;y}|^p \right)^{q/p} \right)^{1/q} \right) \end{aligned}$$

**Demostración:**

$$\begin{aligned} & \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right) \leq D_{1/2} (|d_{j,k;x}| + |d_{j,k;y}|) \Rightarrow \\ & \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right)^p \right)^{q/p} \leq \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} D_{1/2}^p (|d_{j,k;x}| + |d_{j,k;y}|)^p \right)^{q/p} \leq \\ & \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} D_{1/2}^p D_p (|d_{j,k;x}|^p + |d_{j,k;y}|^p) \right)^{q/p} \\ & (D_p)^{q/p} (D_{1/2})^q D_{q/p} \left( \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;x}|^p \right)^{q/p} + \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;y}|^p \right)^{q/p} \right) \Rightarrow \end{aligned}$$

$$\begin{aligned} & \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right) \right)^{q/p} \right)^{1/q} \\ & (D_p)^{1/p} (D_{1/2}) (D_{q/p})^{1/q} \left( \sum_{j \geq 0} \left( \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;x}| \right)^{q/p} + \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;y}| \right)^{q/p} \right) \right)^{1/q} \\ & (D_p)^{1/p} (D_{1/2}) (D_{q/p})^{1/q} D_{1/q} \left( \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;x}| \right)^{q/p} \right)^{1/q} + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} |d_{j,k;y}| \right)^{q/p} \right)^{1/q} \right) \end{aligned}$$

Puesto que  $\sqrt{|c_{0,0;x}|^2 + |c_{0,0;y}|^2} \leq D_{1/2} (|c_{0,0;x}| + |c_{0,0;y}|)$  se verifica la desigualdad “ $\leq$ ”. La otra desigualdad se realiza de forma similar

### Teorema 3.3

Dada una curva  $\mathbf{C} \equiv \mathbf{r}(u) = (x(u), y(u))^T$  entonces:

$$\|\mathbf{r}\|_{B_q^a(L_p(I))} \equiv \sqrt{|c_{0,0;x}|^2 + |c_{0,0;y}|^2} + \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{ajp} 2^{j(p-2)} \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right)^p \right)^{q/p} \right)^{1/q}$$

### Demostración

Este teorema se demuestra utilizando el lema anterior y el teorema 3.2 que indica que las normas  $\|\mathbf{r}\|_{B_q^a(L_p(I))} \equiv \|x\|_{B_q^a(L_p(I))} + \|y\|_{B_q^a(L_p(I))}$  son equivalentes.

A continuación se demuestra el Teorema 3.4

### Proposición A.2

Dadas dos variables aleatorias  $X$  e  $Y$  que se distribuyen como  $DGGS_n(\mathbf{0}, \mathbf{s}^2)$  entonces:

- $E\left(\left(\sqrt{X^2 + Y^2}\right)^p\right) = D_2(v, p+1) \mathbf{s}_j^p, \quad D_2(v, p+1) = \frac{\mathbf{h}(\mathbf{n})\Gamma(1/\mathbf{n})}{\Gamma(2/\mathbf{n})} D(\mathbf{n}, p+1)$
- $\text{Var}\left(\left(\sqrt{X^2 + Y^2}\right)^p\right) \leq D_2(v, p+1) \mathbf{s}_j^{2p}$

**Demostración**

a)

$$\begin{aligned} \mathbb{E}\left(\left(\sqrt{X^2 + Y^2}\right)^p\right) &= \int_0^\infty \int_0^{2p} r^p \frac{nh^2(\mathbf{n})}{2\Gamma(2/\mathbf{n})ps^2} \exp(-(\mathbf{h}(\mathbf{n})r/\mathbf{s})^p) r \, dr d\mathbf{q} = \\ &= 2p \int_0^\infty r^p \frac{nh^2(\mathbf{n})}{2\Gamma(2/\mathbf{n})ps^2} \exp(-(\mathbf{h}(\mathbf{n})r/\mathbf{s})^p) r \, dr = \frac{1}{\mathbf{s}} \frac{h(\mathbf{n})\Gamma(1/\mathbf{n})}{\Gamma(2/\mathbf{n})} \int_0^\infty |r|^{p+1} \frac{nh(\mathbf{n})}{2\Gamma(1/\mathbf{n})\mathbf{s}} \exp(-(\mathbf{h}(\mathbf{n})|r|/\mathbf{s})^p) dr = \\ &= \frac{h(\mathbf{n})\Gamma(1/\mathbf{n})}{\Gamma(2/\mathbf{n})} D_2(\mathbf{n}, p+1)\mathbf{s}^p = D_2(\mathbf{n}, p+1)\mathbf{s}^p \end{aligned}$$

b)

$$\begin{aligned} \text{Var}\left(\left(\sqrt{X^2 + Y^2}\right)^p\right) &= \mathbb{E}\left(\left(\left(\sqrt{X^2 + Y^2}\right)^p\right)^2\right) - \mathbb{E}\left(\left(\sqrt{X^2 + Y^2}\right)^p\right)^2 \leq \mathbb{E}\left(\left(\left(\sqrt{X^2 + Y^2}\right)^p\right)^2\right) = \\ &= \mathbb{E}\left(\left(\sqrt{X^2 + Y^2}\right)^{2p}\right) = D_2(\mathbf{n}, 2p+1)\mathbf{s}^{2p} \end{aligned}$$

**Teorema 3.4**

Supongamos que  $\mathbf{d}_{j,k} = (d_{j,k;x}, d_{j,k;y})^T \stackrel{iid}{\sim} \text{DGGS}_n(\mathbf{0}, \mathbf{s}_j^2)$  con  $\mathbf{s}_j = 2^{-j\mathbf{b}}\mathbf{s}_0$  con  $\mathbf{b} > 0$  y  $\mathbf{s}_0 > 0$  entonces para  $0 < p, q < \infty$  las curvas realizaciones del modelo están casi seguro en  $B_q^a(L_p(I))$  si y solo si  $\mathbf{b} > \mathbf{a} + 1$ .

**Demostración**

Sea  $\mathbf{Z}_j = (z_{j,0}, z_{j,1}, \dots, z_{j,2^j-1})$  con  $z_{j,k} = \sqrt{d_{j,k;x}^2 + d_{j,k;y}^2} / \mathbf{s}_j$  entonces:

$$\begin{aligned} \mathbb{E}\left(\left\|\mathbf{Z}_j\right\|_p^p\right) &= \sum_{k=0}^{2^j-1} \mathbb{E}(z_{j,k}^p) = \frac{1}{\mathbf{s}_j^p} \sum_{k=0}^{2^j-1} \mathbb{E}\left(\left(\sqrt{d_{j,k;x}^2 + d_{j,k;y}^2}\right)^p\right) = 2^j D_2(\mathbf{n}, p+1) \\ \text{Var}\left(\left\|\mathbf{Z}_j\right\|_p^p\right) &= \sum_{k=0}^{2^j-1} \text{Var}(z_{j,k}^p) = \frac{1}{\mathbf{s}_j^{2p}} \sum_{k=0}^{2^j-1} \text{Var}\left(\left(\sqrt{d_{j,k;x}^2 + d_{j,k;y}^2}\right)^p\right) \leq 2^j D_2(\mathbf{n}, 2p+1) \end{aligned}$$

Dado un  $\epsilon > 0$  la desigualdad de Chebyshev implica que:

$$\sum_{j=0}^\infty P\left(\left|2^{-j}\left\|\mathbf{Z}_j\right\|_p^p - D_2(\mathbf{n}, p+1)\right| > \epsilon\right) \leq D_2(\mathbf{n}, 2p+1)\epsilon^{-2} \sum_{j=0}^\infty 2^{-j}$$

por tanto aplicando el primer lema de Borel-Cantelli:

$$2^{-j}\left\|\mathbf{Z}_j\right\|_p^p \xrightarrow{\text{casi seguro}} D_2(\mathbf{n}, p+1), \text{ entonces:}$$

$$\begin{aligned}
 & \left( \sum_{j \geq 0} \left( \sum_{k=0}^{2^j-1} 2^{\mathbf{a}j p} 2^{j(p/2-1)} \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right)^p \right)^{q/p} \right)^{1/q} \\
 &= \left( \sum_{j \geq 0} \left( 2^{\mathbf{a}j p} 2^{j(p/2-1)} \sum_{k=0}^{2^j-1} \left( \sqrt{|d_{j,k;x}|^2 + |d_{j,k;y}|^2} \right)^p \right)^{q/p} \right)^{1/q} = \\
 & \left( \sum_{j \geq 0} \left( 2^{\mathbf{a}j p} 2^{j(p/2-1)} 2^j 2^{-j} \mathbf{s}_j^p \|\mathbf{z}_j\|_p^p \right)^{q/p} \right)^{1/q} = \left( \sum_{j \geq 0} \left( 2^{\mathbf{a}j p} 2^{j(p/2-1)} 2^j 2^{-j b p} \mathbf{s}_0^p 2^{-j} \|\mathbf{z}_j\|_p^p \right)^{q/p} \right)^{1/q} = \\
 & \left( \sum_{j \geq 0} \left( 2^{j(\mathbf{a}p + p/2 - bp)} \mathbf{s}_0^p 2^{-j} \|\mathbf{z}_j\|_p^p \right)^{q/p} \right)^{1/q}
 \end{aligned}$$

entonces la serie converge si y solo si  $p(\mathbf{a}+1/2-\mathbf{b}) < 0$  es decir si  $\mathbf{b} > \mathbf{a}+1/2$

## Apéndice B

Los modelos autoregresivos más simple que permite modelar tanto un movimiento traslacional como un movimiento oscilatorio son los de segundo orden. Éstos tienen una estructura típicamente resonante con una evolución temporal típica en la forma de una oscilación atenuada de la forma  $\exp(-\mathbf{b}t)\cos(2\pi ft)$ .

Para describir un oscilador armónico con constante de atenuación  $\mathbf{b}$  y frecuencia de oscilación  $f$  [Blake 1998] es necesario expresar el modelo autoregresivo de segundo orden con matrices  $A$  y  $B$  definidas como:

$$A = \begin{pmatrix} 0 & 1 \\ a_1 & a_2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ b_0 \end{pmatrix}$$

con  $a_2 = -\exp(-2\mathbf{b}t)$  y  $a_1 = 2\exp(-\mathbf{b}t)\cos(2\pi ft)$  donde  $t$  representa el intervalo temporal de muestreo.

Para obtener una varianza asintótica  $\mathbf{r}^2$  se determina  $b_0$  como:

$$b_0 = \mathbf{r} \sqrt{\frac{(1+a_2)(1-a_1-a_2)(1+a_1-a_2)}{(1-a_2)}}$$

En el caso del modelo de velocidad constante en el que  $\mathbf{b} = f = 0$  o de forma equivalente  $a_2 = -1$  y  $a_1 = 2$  la varianza asintótica crece de forma semicúbica como:

$$\mathbf{r}(t) = \gamma_0 t^{3/2} \text{ donde } b_0 = \gamma_0 t^{3/2}$$

Para determinar los modelos armónicos en los modelos separados por subespacios de la sección 3.4.3.2.1 se determinan los valores de  $a_1$  y  $a_2$  de la forma anterior y se determina el valor de  $b_0$  o  $\gamma_0$  a partir de los resultados de dicha sección:

### Teorema 3. 10

Dado el modelo dinámico de Ec 3.123 supongamos que existe una distribución límite dentro del subespacio  $S_i$ . Entonces para obtener un desplazamiento cuadrático medio  $\bar{\mathbf{r}}_i$ , el valor correspondiente de  $b_i$  debe ser:

$$b_i = \bar{\mathbf{r}}_i \sqrt{\frac{(1+a_{2,i})(1-a_{1,i}-a_{2,i})(1+a_{1,i}-a_{2,i})}{(1-a_{2,i})}} \frac{N}{\text{Traza}(\mathbf{P}_i \mathbf{S})}$$

**Teorema 3. 12**

Dado el modelo dinámico señalado en Ec. 3.123. Si  $a_{2,i}=-1$  y  $a_{1,i}=2$  no existe distribución límite en el subespacio  $S_i$ . Entonces dado un valor exacto para  $w_0$  el desplazamiento medio  $\bar{\mathbf{r}}_i^2(k)$  en el subespacio  $S_i$  en tiempo  $k$  verifica:

$$\bar{\mathbf{r}}_i(k) \approx b_i \sqrt{\frac{\text{Traza}(\mathbf{P}_i \mathbf{S})}{3N}} \sqrt{k^3}$$

o de forma equivalente:

$$\bar{\mathbf{r}}_i(t) \approx \mathbf{g}_i t^{3/2}, \quad b_i = \frac{\mathbf{g}_i}{\sqrt{\frac{\text{Traza}(\mathbf{P}_i \mathbf{S})}{3N}}} t^{3/2}$$