

**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA**

**DEPARTAMENTO DE INFORMÁTICA Y SISTEMAS**



**TESIS DOCTORAL**

**ESTIMACIÓN DEL FLUJO ÓPTICO EN SECUENCIAS DE  
IMÁGENES Y DE LA CARTA DE DISPARIDAD EN  
PARES ESTÉREO: APLICACIÓN A LA  
RECONSTRUCCIÓN TRIDIMENSIONAL**

**JAVIER SÁNCHEZ PÉREZ**

Las Palmas de Gran Canaria, Octubre de 2001

**Título de la tesis:**

***ESTIMACIÓN DEL FLUJO ÓPTICO EN SECUENCIAS DE  
IMÁGENES Y DE LA CARTA DE DISPARIDAD EN PARES  
ESTÉREO: APLICACIÓN A LA RECONSTRUCCIÓN  
TRIDIMENSIONAL***

**Thesis title:**

***DENSE OPTICAL FLOW ESTIMATION IN IMAGE SEQUENCES  
AND DISPARITY MAP COMPUTATION FOR STEREO PAIRS  
APPLIED TO 3D RECONSTRUCTION***

## **Thesis title:**

# ***DENSE OPTICAL FLOW ESTIMATION IN IMAGE SEQUENCES AND DISPARITY MAP COMPUTATION FOR STEREO PAIRS APPLIED TO 3D RECONSTRUCTION***

## **Abstract:**

In this thesis we proposed several variational techniques to deal with the problems of optical flow estimation and stereoscopic reconstruction.

For the first problem we show that a classic optical flow technique by Nagel and Enkelmann can be regarded as an early anisotropic diffusion method with a diffusion tensor. We introduce three improvements into the model formulation that (i) avoid inconsistencies caused by centering the brightness term and the smoothness term in different images, (ii) use a linear scale-space focusing strategy from coarse to fine scales for avoiding convergence to physically irrelevant local minima, and (iii) create an energy functional that is invariant under linear brightness changes. Applying a gradient descent method to the resulting energy functional leads to a system of diffusion-reaction equations. We prove that this system has a unique solution under realistic assumptions on the initial data, and we present an efficient linear implicit numerical scheme in detail. Our method creates flow fields with 100% density over the entire image domain, it is robust under a large range of parameter variations, and it can recover displacement fields that are far beyond the typical one-pixel limits.

For the second problem, we present an energy-based approach to estimate a dense disparity map between two images while preserving its discontinuities resulting from image boundaries. We first derive a simplified expression for the disparity that allows us to easily estimate it from a stereo pair of images using an energy minimization approach. We assume that the epipolar geometry is known and we include this information in the energy model. This method is an extension of the one proposed for the optical flow estimation problem.

We propose methods for the estimation of disparity maps and optical flow fields from grey level images as well as from colour images. Finally, symmetrical methods are introduced to force a congruent solution in both directions.

We show the performance of our methods by comparing with the results given by others for sequences of synthetic and real images.

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA  
DEPARTAMENTO DE INFORMÁTICA Y SISTEMAS  
PERCEPCIÓN ARTIFICIAL Y APLICACIONES

TESIS DOCTORAL

ESTIMACIÓN DEL FLUJO ÓPTICO EN SECUENCIAS DE  
IMÁGENES Y DE LA CARTA DE DISPARIDAD EN PARES  
ESTÉREO: APLICACIÓN A LA RECONSTRUCCIÓN  
TRIDIMENSIONAL

Autor de la tesis

Director de la tesis

Javier Sánchez Pérez

Luis Alvarez León

Las Palmas de Gran Canaria, Octubre de 2001

# TESIS DOCTORAL

Autor: **Javier Sánchez Pérez**  
Director: **Luis Alvarez León**  
Título: **Estimación del flujo óptico en secuencias de imágenes y de la carta de disparidad en pares estéreo: Aplicación a la reconstrucción tridimensional**  
Universidad: **Universidad de Las Palmas de Gran Canaria**  
Departamnt.: **Informática y Sistemas**  
Programa: **Percepción Artificial y Aplicaciones**  
Grado: **Doctor en Informática**  
Fecha: **Octubre 2001**

Tesis doctoral presentada por **D. Javier Sánchez Pérez**, y dirigida por el profesor **D. Luis Alvarez León**, para la obtención del grado de **Doctor en Informática**.

*A mi madre*

# Índice General

Agradecimientos	ix
Introducción	1
<b>I Consideraciones previas y estado del arte</b>	<b>9</b>
<b>1 Difusión anisotrópica</b>	<b>11</b>
1.1 Introducción . . . . .	11
1.2 Perona y Malik . . . . .	13
1.3 Difusión por curvatura media . . . . .	15
1.4 Difusión utilizando tensores de difusión . . . . .	20
1.5 Tensores de difusión lineales . . . . .	25
<b>2 Flujo óptico</b>	<b>31</b>
2.1 Introducción . . . . .	31
2.2 Factores que influyen sobre las imágenes . . . . .	32
2.2.1 Factores fotométricos . . . . .	33
2.2.2 Factores geométricos . . . . .	34
2.3 Clasificación de las distintas técnicas . . . . .	36
2.3.1 Métodos diferenciales . . . . .	36
2.3.2 Métodos basados en la correlación . . . . .	37
2.3.3 Métodos basados en la frecuencia . . . . .	37
2.4 Estudio de métodos diferenciales . . . . .	38
2.5 Imágenes de prueba . . . . .	40
2.5.1 Imágenes sintéticas . . . . .	40
2.5.2 Imágenes reales . . . . .	48
<b>3 Geometría Epipolar</b>	<b>51</b>
3.1 Introducción . . . . .	51
3.2 Geometría de dos cámaras proyectivas . . . . .	52
3.3 Rectificación de imágenes . . . . .	56
3.4 Reconstrucción tridimensional . . . . .	59

<b>4</b>	<b>Visión estereoscópica</b>	<b>63</b>
4.1	Introducción . . . . .	63
4.2	Factores que influyen en los pares estéreo . . . . .	64
4.3	Clasificación de las técnicas . . . . .	66
4.3.1	Basadas en características . . . . .	66
4.3.2	Basadas en áreas . . . . .	66
4.3.3	Basadas en frecuencias . . . . .	67
4.3.4	Basadas en energías . . . . .	67
4.4	Imágenes de prueba . . . . .	67
4.4.1	Imágenes sintéticas . . . . .	67
4.4.2	Imágenes reales . . . . .	71
 <b>II Estimación de correspondencias utilizando una aproximación variacional</b>		 <b>79</b>
<b>5</b>	<b>Introducción</b>	<b>81</b>
5.1	Aproximación variacional . . . . .	83
5.2	Modelo propuesto por Nagel y Enkelmann . . . . .	84
5.3	Recuperando grandes desplazamientos . . . . .	86
<b>6</b>	<b>Cálculo del flujo óptico</b>	<b>93</b>
6.1	Introducción . . . . .	93
6.2	Modelo propuesto . . . . .	94
6.2.1	Centrado consistente . . . . .	94
6.2.2	Relaciones con la difusión anisótropa . . . . .	95
6.2.3	Modelo para el cálculo del flujo óptico con imágenes en color . . . . .	96
6.3	Esquema numérico . . . . .	97
6.4	Parámetros . . . . .	100
6.5	Resultados experimentales . . . . .	102
<b>7</b>	<b>Cálculo de la carta de disparidad</b>	<b>117</b>
7.1	Introducción . . . . .	117
7.2	Modelo propuesto . . . . .	119
7.2.1	El funcional de energía . . . . .	119
7.2.2	Minimizando la energía . . . . .	120
7.2.3	Modelo para el cálculo de la carta de disparidad con imágenes en color . . . . .	122
7.3	Esquema numérico . . . . .	123
7.4	Resultados experimentales . . . . .	124
7.4.1	Imágenes sintéticas . . . . .	124
7.4.2	Imágenes reales . . . . .	125
7.4.3	Resultados experimentales para imágenes en color . . . . .	134

<b>III</b>	<b>Introducción de la simetría en el cálculo de correspondencias</b>	<b>139</b>
<b>8</b>	<b>Introducción</b>	<b>141</b>
8.1	Modelo de energía para las simetrías . . . . .	143
8.2	La elección de la función $\Psi$ . . . . .	144
<b>9</b>	<b>Cálculo del flujo óptico simétrico</b>	<b>147</b>
9.1	Introducción . . . . .	147
9.2	El modelo de energía . . . . .	147
9.3	Minimizando la energía . . . . .	149
9.4	Esquema numérico . . . . .	152
9.5	Experiencias numéricas . . . . .	157
<b>10</b>	<b>Cálculo de la carta de disparidad simétrica</b>	<b>163</b>
10.1	Introducción . . . . .	163
10.2	El modelo de energía . . . . .	163
10.3	Minimizando la energía . . . . .	164
10.4	Esquema numérico . . . . .	167
10.5	Simplificación del modelo . . . . .	170
10.5.1	El modelo de energía . . . . .	171
10.5.2	Minimizando la energía . . . . .	172
10.5.3	Esquema numérico . . . . .	173
10.6	Resultados numéricos . . . . .	175
<b>11</b>	<b>Conclusiones</b>	<b>179</b>
<b>12</b>	<b>Trabajo futuro</b>	<b>183</b>
<b>A</b>	<b>Invariancia a cambios lineales de intensidad</b>	<b>185</b>
<b>B</b>	<b>Existencia y unicidad de soluciones</b>	<b>187</b>
B.1	Marco abstracto . . . . .	187
B.2	Resultado de Existencia y Unicidad . . . . .	188
<b>C</b>	<b>Formas de discretización</b>	<b>193</b>
C.1	Discretización del gradiente . . . . .	193
C.2	Discretización de la divergencia . . . . .	196
	<b>Bibliografía</b>	<b>200</b>

# Agradecimientos

Quisiera agradecer a Luis Alvarez por su dedicación, apoyo y entrega durante todo el transcurso de esta tesis y con quien compartí despacho durante buena parte de ésta.

A mis compañeros del laboratorio de investigación AMI de la Universidad de Las Palmas de Gran Canaria que hicieron del laboratorio algo más que una sala de ordenadores: Alegría, Gonzalo, Mendoza, Octavio y Víctor.

A los miembros del grupo AMI: Agustín, Carmelo, Esther, Fran, Julio, Luis Carrasco, Luis Mazorra y Miguel.

A Karl que vino a nuestro laboratorio para realizar una estancia postdoctoral y con quien tuve la oportunidad de discutir sobre temas que trato en esta tesis.

A Joachim Weickert con quien hemos realizado varios trabajos de investigación y siempre se ha mostrado dispuesto a resolver cualquier tipo de duda.

Al grupo Robotvis, donde hice una pequeña estancia y, en especial, a Rachid Deriche que fue quien me lo propuso y con quien mantenemos una estrecha colaboración. A Gerardo con quien estuve trabajando bastante tiempo y compartí muchos puntos de interés sobre los que aquí se trata. También a Jacques por dejarme vivir en su piso de Antibes.

Al Departamento de Informática y Sistemas de la Universidad de Las Palmas de Gran Canaria por poner los medios suficientes para llevar a cabo este trabajo y por financiar parte de los viajes que realicé durante estos años.

A Elena por su interés e inestimable ayuda en la corrección de este documento.

# Introducción

Los problemas del cálculo del flujo óptico y de la visión estereoscópica son dos problemas clásicos en el campo de la visión por ordenador. El primero consiste en encontrar el desplazamiento aparente de los objetos en una secuencia de imágenes. El segundo, consiste en obtener una reconstrucción tridimensional de una escena a partir de sus proyecciones.

El flujo óptico es el desplazamiento que existe entre los pixels de imágenes bidimensionales. Se puede pensar que éste es una proyección del movimiento de los objetos en escenas tridimensionales, pero en algunos casos este flujo no se corresponde con un desplazamiento real.

En el problema de la visión estereoscópica partimos de dos imágenes que han sido obtenidas por dos cámaras en un instante de tiempo determinado. Estas imágenes contienen la proyección de los puntos de una escena común. Para encontrar esos puntos de la escena, tenemos que saber qué puntos de una imagen se corresponden con qué puntos de la otra. Este proceso es muy similar al del flujo óptico excepto por el hecho de que en la visión estereoscópica existe una relación geométrica entre las imágenes. Como veremos en el capítulo 4, ésta se divide, básicamente, en tres procesos, en los que el de puesta en correspondencia de puntos de las imágenes es fundamental para un método eficiente de reconstrucción. En nuestro trabajo nos centraremos casi en exclusiva en este punto.

Hablamos de *carta o mapa de disparidad* para referirnos al desplazamiento en magnitud que existe entre un par de imágenes estéreo. Muchas veces haremos referencia al problema de la visión estereoscópica como el de la estimación de la carta de disparidad, pero no debemos olvidar que la visión estéreo engloba a éste. Cuando hablamos de *flujo*, nos referimos al desplazamiento que sufren los pixels entre las imágenes. Lo utilizaremos indistintamente tanto para el problema del *flujo óptico* como para el de la visión estéreo, refiriéndonos en este caso al *flujo estéreo*.

En los métodos que proponemos para abordar estos problemas, partimos de una aproximación variacional basada en energías. Partimos de una energía global en la que al minimizar se busca encontrar la solución al flujo como una función densa – sobre

todo el dominio de la imagen – que cumpla con unas restricciones que comentaremos más adelante. Al minimizar esta energía, se genera un sistema de ecuaciones en derivadas parciales (EDP) que posee unas propiedades interesantes de *difusión* – a través de un término basado en la divergencia – y de *reacción* – a través de un término que ejerce una fuerza de atracción entre las imágenes. Para resolver este sistema aplicamos un método de descenso por gradiente que provoca que nuestro método converja rápidamente a un mínimo. Lo implementamos por medio de esquemas numéricos implícitos iterativos que le dan una gran estabilidad frente a un paso temporal grande y aceleran la convergencia a la solución. Para evitar que nuestro método quede atrapado en mínimos locales irrelevantes, utilizamos una estrategia de análisis multiescala lineal y una estrategia multipiramidal. La segunda se suele utilizar con bastante frecuencia en métodos parecidos y se basa en establecer distintas escalas por medio de zooms a las imágenes originales. La primera se basa en definir escalas por medio de convoluciones con núcleos Gaussianos e introduce un marco formal de análisis que tiene un amplio fundamento matemático.

## Alcance del trabajo

Este trabajo se puede dividir en dos partes bien diferenciadas: En la primera introducimos métodos directos para estimar el flujo óptico y la carta de disparidad, y en la segunda incorporamos la simetría a estos métodos con el fin de obtener soluciones *coherentes* en ambas direcciones.

Las energías que proponemos para los métodos directos se dividen en dos términos, uno de *acoplamiento* o *ligadura* y otro de *regularización*. El primero es el encargado de buscar puntos similares entre las imágenes, mientras que el segundo es el que impone que la solución sea regular por zonas. Al igual que muchos otros, nuestro método parte de la *suposición lambertiana* que comentamos en el capítulo 2. Esto es una restricción importante sobre el tipo de imágenes sobre las que podemos trabajar. La complejidad de las imágenes reales hace que en algunos casos nuestro método no sea capaz de captar correctamente el flujo en ciertas zonas, aunque para los métodos que proponemos, las pruebas con imágenes reales que hemos realizado producen soluciones satisfactorias. El segundo término, el de regularización, es el que incorpora las características de difusión que hacen que la solución generada sea suave y densa al 100%. En éste utilizamos el operador de Nagel–Enkelmann que, como justificaremos más adelante, tiene un comportamiento anisótropo en la zona de los contornos e isótropo en zonas homogéneas – en el capítulo 1 tratamos el tema de la difusión anisotrópica y, en particular, veremos el tensor de difusión de Nagel–Enkelmann.

Para los métodos simétricos incorporamos un tercer término que fuerza que las soluciones en ambas direcciones sean inversas entre sí. Se impone esta simetría de

forma débil, es decir, no se impone algebraicamente que las soluciones sean inversas, sino que se busca la solución que minimice el error en las dos direcciones. Debido a las oclusiones y otros problemas que comentamos más adelante, este término debe tener en cuenta que la simetría sólo se puede definir por regiones. La incorporación de las simetrías a nuestros modelos provoca que el esquema numérico que se deriva de las EDPs sea más complejo, siendo difícil encontrar un esquema implícito. En el caso del cálculo de la carta de disparidad simplificamos el modelo por medio de la rectificación de imágenes que nos lleva a una configuración geométrica sencilla y cuyo esquema implícito es inmediato.

Tanto para el problema del cálculo del flujo óptico como para el del mapa de disparidad en una sola dirección, desarrollamos modelos que tienen en cuenta tanto imágenes en niveles de grises como imágenes en color. Suponemos que las imágenes en color vienen definidas por tres canales de color, RGB, cada una correspondiente a un color distinto (rojo, verde y azul).

## Resultados y contribuciones

Durante el desarrollo de los capítulos donde se detallan los métodos que proponemos, explicamos cuáles han sido las contribuciones que hacemos y desarrollamos una serie de pruebas numéricas para evaluar su eficiencia.

En la primera parte, documentamos algunas series de imágenes tanto para el flujo óptico como para la visión estéreo que luego utilizamos para evaluar los métodos que proponemos. Varias de estas secuencias son imágenes generadas de forma sintética que se han utilizado con frecuencia en muchos trabajos relacionados. Para muchas de estas imágenes se conoce el flujo exacto y son bastante útiles de cara a desarrollar estudios comparativos entre los métodos. Nosotros comparamos nuestros métodos basándonos en estudios que se han realizado con antelación en trabajos como el de Barron et al. [BFB94].

Los resultados y contribuciones más importantes que destacamos son las siguientes:

- Como veremos para el caso del flujo óptico, nuestros métodos mejoran los resultados de todos los métodos estudiados en este artículo, considerando los que producen resultados sobre toda la imagen, con una densidad del 100%. Para el caso estéreo, comparamos nuestro método con métodos basados en correlación. Las soluciones que obtenemos son bastante más regulares, preservando los contornos y obteniendo un resultado para todos los puntos del dominio de la imagen.

- También probamos con imágenes reales, obteniendo resultados bastante satisfactorios. Para estas mostramos el flujo o la carta de disparidad y se puede apreciar que son soluciones regulares por zonas (definidas por los contornos de la imagen) con una precisión también elevada. En el caso de la visión estereoscópica realizamos la reconstrucción de los pares estéreo apreciándose, en muchos casos, que se determina correctamente la profundidad de la escena original.
- Los métodos que proponemos aportan ciertas mejoras sobre otros métodos relacionados como puede ser la **invariancia frente a cambios lineales de brillo** de las imágenes. Esto se consigue ajustando los parámetros del método tal y como se explica en el apéndice A.
- Corregimos el método de Nagel–Enkelmann para que se mantenga una **coherencia en la centralización** de los términos de las energías. Sin esta centralización, tal y como demostramos en el capítulo 6 sobre flujo óptico, el resultado que obtenemos no es adecuado cuando los desplazamientos son grandes.
- Incorporamos técnicas de **análisis multiescala y multipiramidal para recuperar grandes desplazamientos**. Utilizamos tanto análisis multiescala lineal, para recuperar grandes desplazamientos a través de convoluciones con Gaussianas, como no lineal, en los que se aplica un esquema de difusión anisótropa. Por otro lado demostramos la existencia y unicidad de las soluciones de nuestro funcional de energía bajo un marco de análisis multiescala.
- Proponemos **métodos simétricos** que mejoran los resultados obtenidos en los no simétricos. Estos dan soluciones coherentes en ambas direcciones y gracias a la forma del término de simetría, se pueden determinar las oclusiones de la imagen.

En las conclusiones (capítulo 11), y en cada capítulo en cuestión, comentamos con mayor detalle estos resultados.

## Publicaciones realizadas

Durante el desarrollo de esta tesis hemos realizado una serie de publicaciones que están relacionadas con:

- La estimación de la geometría epipolar: Realizamos los trabajos [AS99b, AS99a]. En estos trabajos se propone un método no lineal para la estimación de la matriz esencial.

- El flujo óptico: Para el problema de la estimación del flujo óptico en una sola dirección, realizamos los trabajos [AELS99, AWS99, AWS00].
- La visión estéreo: En los trabajos [ADSW00, S00] tratamos el problema del cálculo de la carta de disparidad en una sola dirección.
- Los métodos para imágenes en color: En [AS00a, AS00b] proponemos el método para el cálculo de la carta de disparidad en una dirección utilizando pares estéreo de imágenes en color.
- El cálculo de la carta de disparidad simétrica: Por último hemos realizado un trabajo, [ADPS01], relacionado con el problema de la estimación de la carta de disparidad incorporando la simetría.

## División del documento

El documento se divide básicamente en tres partes.

En la **primera parte**, veremos algunos aspectos importantes que luego utilizaremos a la hora de proponer nuestros métodos. En estos cuatro capítulos hablaremos, en primer lugar, de la difusión anisótropa (capítulo 1) haciendo un pequeño repaso de los métodos de difusión más representativos con algunas variantes y detallando más la difusión por tensores de difusión, en especial el tensor propuesto por Nagel–Enkelmann. Luego trataremos el problema del flujo óptico (capítulo 2), donde explicamos algunos conceptos fundamentales, determinamos los factores que influyen sobre la formación de las imágenes y que repercuten a la hora de realizar una buena estimación, hacemos una clasificación de las técnicas existentes y, por último, comentamos las imágenes sintéticas y reales que utilizamos para realizar las pruebas. En el capítulo 3 hacemos un pequeño apunte sobre la geometría epipolar que nos servirá para entender mejor el primer paso de la visión estereoscópica, realizamos una parametrización de la disparidad que utilizaremos en nuestros métodos de estimación del flujo estéreo, explicamos una forma de rectificación de imágenes a partir de las matrices de proyección de las cámaras que utilizaremos en el método para el cálculo de la carta de disparidad de forma simétrica, y expondremos una forma de realizar la reconstrucción 3D a partir de las matrices de proyección y de las correspondencias entre pixels. Por último, en el capítulo 4, tratamos el problema de la visión estereoscópica. Al igual que en el caso del flujo óptico estudiamos los factores que influyen sobre las imágenes y que repercuten en especial en el caso de la visión estéreo, luego hacemos una clasificación de las técnicas existentes y, por último, comentamos los pares estéreo sintéticos y reales que utilizamos para las pruebas. En los casos sintéticos mostramos el flujo para aquellos que se conocen.

En la segunda y tercera partes se explican nuestros métodos en una sola dirección y simétricos respectivamente. En ambas partes seguimos un esquema muy parecido. Cada una se compone de tres capítulos: El primero es una pequeña introducción en donde se plantea de forma general la energía y donde se comentan características que son comunes a ambos métodos; el segundo es nuestro modelo para el flujo óptico y; el tercero es el correspondiente para el problema de la visión estereoscópica.

En la **segunda parte** proponemos modelos variacionales tanto para el cálculo del flujo óptico (capítulo 6) como para el problema de la visión estereoscópica (capítulo 7) en una sola dirección. En la introducción (capítulo 5) de esta parte tratamos la energía desde un punto de vista general, introduciendo luego el modelo de Nagel–Enkelmann para terminar por comentar las dos estrategias que desarrollamos para considerar grandes desplazamientos entre las imágenes. Estas dos estrategias son las del análisis multiescala lineal, definido por la convolución de las imágenes con una Gaussiana, y la estrategia multipiramidal, que se basa en definir las escalas a través de zooms que se aplican a las imágenes. En el capítulo 6 sobre el flujo óptico proponemos una nueva energía tanto para imágenes en niveles de grises como para imágenes en color. Minimizamos el funcional de energía llegando a una ecuación en derivadas parciales que solucionamos por medio de un método de descenso por gradiente. Desarrollamos un esquema numérico y explicamos los parámetros que están incluidos en el método. Por último, mostramos algunos resultados numéricos de imágenes sintéticas y reales y comparamos numéricamente nuestro método con otros métodos. En el capítulo 7 tratamos el problema del cálculo de la carta de disparidad en una dirección. La organización de este capítulo es similar al anterior. Adicionalmente introducimos una parametrización para incorporar en nuestro modelo la información de la geometría asociada a las cámaras. Al final también mostramos algunos resultados numéricos y hacemos comparaciones con métodos de correlación estándar, mostrando algunas reconstrucciones 3D de pares estéreo reales.

En la **tercera parte** exponemos unos métodos que consideran la simetría para el cálculo del flujo óptico (capítulo 9) y de la carta de disparidad (capítulo 10). Esta parte tiene muchos puntos en común con la anterior, como por ejemplo las estrategias para recuperar grandes desplazamientos. En la introducción proponemos la energía de forma general, haciendo especial énfasis en el término de simetría. La organización de los capítulos son similares a sus homólogos de la parte anterior. En las experiencias numéricas realizamos comparaciones entre los métodos simétricos y los no simétricos, y mostramos en algunos casos los mapas de oclusiones. En el capítulo 10 simplificamos el modelo para imágenes rectificadas, con el fin de poder establecer un esquema numérico implícito de forma sencilla.

En el capítulo 11 se exponen las conclusiones sobre los métodos que proponemos.

En el capítulo 12 comentamos algunos trabajos para el futuro que se podrían continuar a partir de éste. En los apéndices A, B y C tratamos puntualmente algunos temas comunes a los métodos anteriores.

## Aplicaciones

El problema del cálculo del flujo óptico y el de la visión estereoscópica tienen un punto en común que es el de la búsqueda de correspondencias entre pixels de las imágenes. Son muchas las aplicaciones<sup>1</sup> que están relacionadas con este tema como pueden ser: Detección del movimiento en escenas, segmentación de objetos, tracking, medición de la profundidad y movimiento en el entorno, robots móviles, imágenes satélite como las meteorológicas con el fin de medir el movimiento de nubes, aplicaciones militares para el seguimiento de objetivos y navegación de vehículos, imágenes biomédicas tales como el movimiento del corazón, estudio del cerebro, vigilancia de sitios, interfaces y realidad virtual, codificación de señales de vídeo.

---

<sup>1</sup>Información detallada acerca de los experimentos que se realizan en esta tesis se pueden encontrar en el sitio web <http://serdis.dis.ulpgc.es/~jsanchez> en el directorio /research/demos y, en /research/software, se encuentran programas para el cálculo del flujo óptico y de la carta de disparidad, además de programas de tratamiento de imágenes como el XMEgaWave (ver [GT]) y el AMILab (ver [TK01]).

# Parte I

## Consideraciones previas y estado del arte

# Capítulo 1

## Difusión anisotrópica

### 1.1 Introducción

Uno de los factores que más van a influir en los modelos que propondremos en los siguientes capítulos, es el de la difusión anisotrópica. En este capítulo estudiaremos algunos esquemas de difusión típicos en el campo de la visión por ordenador y comentaremos las características de cada uno que nos parecen más interesantes. Explicaremos en detalle el esquema de difusión que utilizamos tanto para los métodos de flujo óptico como para los del cálculo de la carta de disparidad.

La importancia del efecto de la difusión es evidente en muchos campos de la física y en el tratamiento de imágenes está siendo centro de bastante investigación. Muchos métodos de restauración de imágenes, segmentación, realce de contornos, estimación del flujo óptico, etc. utilizan algún proceso de difusión.

En nuestro caso particular, partimos de un problema de minimización de energía en el que se proponen dos términos: uno de ligadura, y otro de regularización. Más adelante, cuando propongamos los métodos objeto de esta tesis, los estudiaremos en profundidad. De forma general, la energía tiene la expresión

$$E(f) = \int_{\Omega} (f - q)^2 dx + \alpha \int_{\Omega} \Phi(\nabla f) dx \quad (1.1.1)$$

El primer término es el de ligadura que ejerce un efecto de atracción entre las dos señales  $f$  y  $q$ , mientras que el segundo es el de regularización, encargado de imponer alguna restricción – generalmente de suavidad – sobre la solución final. Es el segundo término, el de regularización, el que nos da el comportamiento de difusión. En los tipos de energía que veremos, si minimizamos la ecuación (1.1.1) y nos quedamos con la parte que se obtiene del término de regularización, llegamos a la siguiente ecuación

Tabla 1.1: **Esquemas de difusión:** En esta tabla se muestran algunos esquemas de difusión típicos en tratamiento de imágenes.

Esquema de difusión	Ecuación
Perona y Malik [PM90],[CCLM92]	$f_t = \text{div} (g (\ \nabla f\ ) \nabla f)$
Curvatura media [ALM92], [OS88]	$f_t = \ \nabla f\  \text{div} \left( \frac{\nabla f}{\ \nabla f\ } \right) = f_{\xi\xi}$
Difusión tensorial [NE86]	$f_t = \text{div}(D(\nabla f)\nabla f)$

en derivadas parciales:

$$f_t = \text{div} (\Phi(\nabla f) \nabla f) \quad (1.1.2)$$

donde  $\text{div}$  representa la divergencia de una función vectorial. En el caso más simple, cuando  $\Phi(\nabla f) = \|\nabla f\|^2$ , llegamos a la ecuación del calor,

$$f_t = \text{div}(\nabla f) = f_{xx} + f_{yy} \quad (1.1.3)$$

en donde la difusión se ejerce por igual en todas direcciones.

Se ha demostrado, dentro del campo del análisis multiescala (ver [Lin94]), que la solución a esta ecuación es equivalente a aplicar un filtro gaussiano. Este es un esquema de difusión isótropo, es decir, la difusión se realiza por igual en todas direcciones.

Hablamos de difusión anisótropa cuando la dirección de difusión no es paralela al gradiente de la señal [Wei97, Wei98a]. Perona y Malik propusieron uno de los primeros esquemas de difusión en el que no se difuminaba por igual en cada punto de la imagen; ellos lo denominaron difusión anisótropa. Siguiendo los trabajos de Weickert [Wei94, Wei95, Wei96], éste propone esquemas de difusión anisótropos basados en tensores de difusión. Weickert establece en [Wei97] una clasificación de tipos de esquemas de difusión y define el esquema de Perona y Malik como un tipo de difusión no lineal isotrópico. En Alvarez et al. [ALM92] se propone un esquema de difusión anisótropa basada en la curvatura media.

Vamos a describir los esquemas de difusión que creemos más representativos y que nos ayudarán a entender mejor cuál es el funcionamiento de la difusión dentro del problema de la estimación del flujo óptico y del cálculo de la disparidad.

Estudiaremos tres formas distintas de realizar la difusión y sus variantes más representativas.

A lo largo de este capítulo iremos realizando algunas pruebas prácticas con el fin de estudiar el efecto de cada método en las imágenes y realizar algunas comparaciones.

La figura 1.1 es una imagen sintética de una bola y dos rectángulos que utilizaremos en las pruebas de difusión. También utilizaremos una imagen real (figura 1.2) del tenor canario Alfredo Kraus.

En la última sección de este capítulo trataremos un caso especial en el que se utilizan tensores de difusión constantes en el tiempo. En este caso, la imagen destino

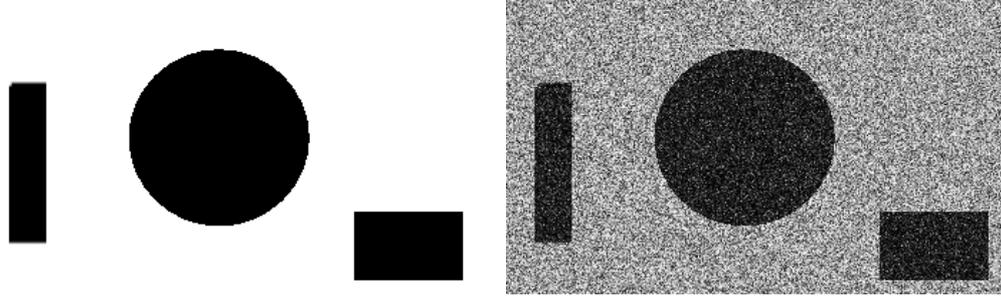


Figura 1.1: Imagen sintética de un círculo y dos rectángulos (izquierda) y la misma imagen con ruido gaussiano (derecha).

de la difusión no coincide con la imagen a partir de la cual se va a calcular el campo de vectores de difusión. Este tipo de esquemas es interesante cuando queremos conservar ciertas características a través del tensor para difuminar otra imagen. Son estos esquemas los que utilizaremos posteriormente para tratar con el problema del flujo óptico y del problema estereoscópico

## 1.2 Perona y Malik

Uno de los métodos que ha sido centro de mayor estudio es la ecuación de Perona y Malik [PM90]:

$$f_t = \text{div} (g (\|\nabla f\|^2) \nabla f) \quad (1.2.1)$$

en donde se elige la función  $g (\|\nabla f\|^2)$  de forma que sea descendiente (p.e  $g(s) =$

$\frac{1}{1+s^2/\lambda^2}$ ,  $g(s) = e^{-\frac{s^2}{\lambda^2}}$ , figura 1.3). Esta función sirve como inhibidora en la zona de los contornos,  $\|\nabla f\| \rightarrow \infty$ , cancelando la difusión cuando y permitiendo que en las zonas homogéneas,  $\|\nabla f\| \rightarrow 0$ , se difumine de forma isotrópica,  $f_t = \text{div}(\nabla f) = \Delta f$ .

Este esquema presenta varios inconvenientes que se han tratado extensivamente en varios trabajos. En Catté et al. [CCLM92], se propone una mejora introduciendo la convolución con una Gaussiana en la estimación de la norma del gradiente para volver el método más robusto bajo la presencia de ruido y consistente desde un punto de vista formal. La ecuación es de la siguiente forma:

$$f_t = \text{div} (g (\|\nabla f_\sigma\|) \nabla f) \quad (1.2.2)$$

donde  $f_\sigma = G_\sigma * f$  representa la convolución con una Gaussiana con desviación típica de  $\sigma$ .

En la figura 1.4 mostramos varios resultados después de aplicar la ecuación de Catté et al. a la figura del círculo (figura 1.1). Para obtener este resultado, hemos implementado un esquema numérico implícito por diferencias finitas.

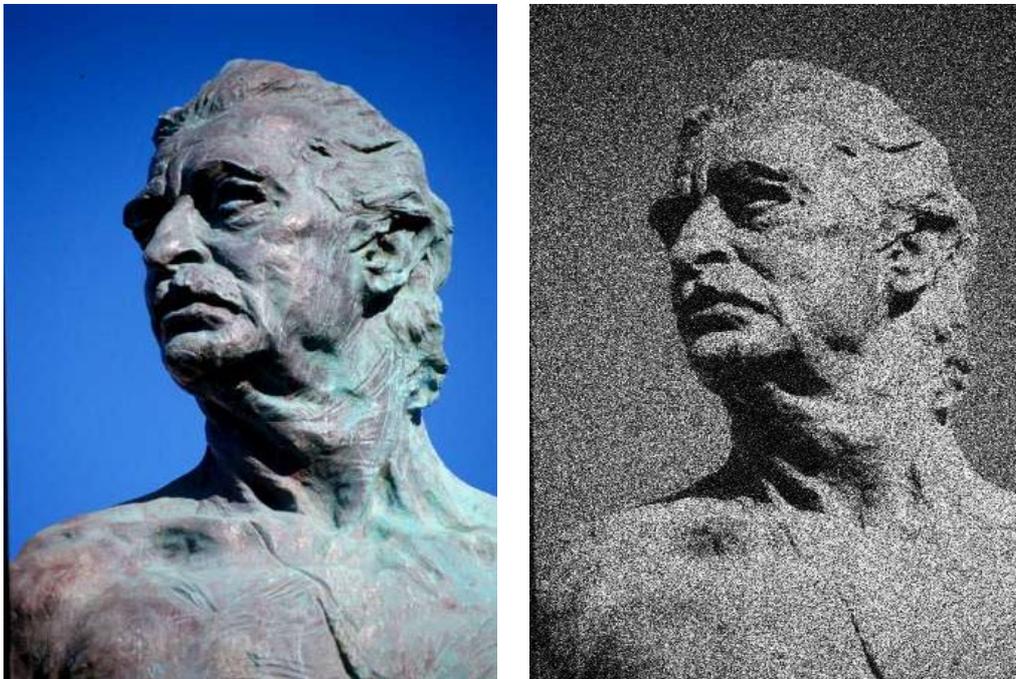


Figura 1.2: Estatua del tenor Alfredo Kraus (izquierda) y la misma imagen con ruido gaussiano (derecha).

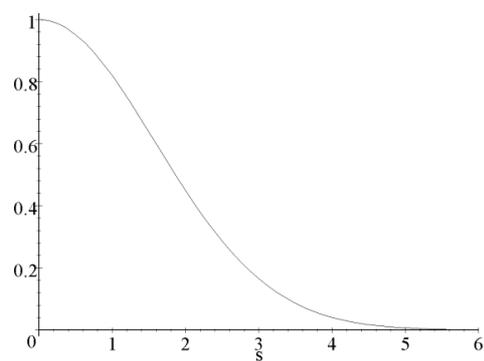


Figura 1.3:  $g(s) = e^{-\frac{s^2}{2}}$ .

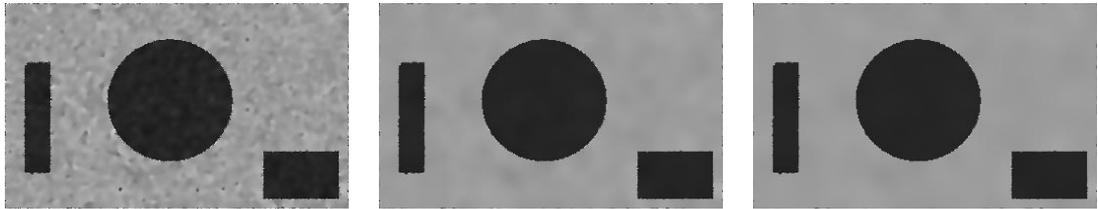


Figura 1.4: Resultado de aplicar el esquema de Catté et al. (1.2.2) a la imagen del círculo y los rectángulos (fig. 1.1). En las tres pruebas hemos utilizado  $\lambda = 2$  y  $\sigma = 1$ : izquierda, solución de la ecuación con 100 iteraciones; centro, con 500 y; derecha, con 1000 iteraciones.

Como podemos ver en las pruebas de la figura 1.4, a medida que aumenta el número de iteraciones, va disminuyendo el nivel de ruido en la imagen, preservando los contornos de los objetos. Se preserva el contorno de los objetos gracias al efecto inhibitor de la función  $g(s)$  cuando el gradiente de la imagen es alto. Sin embargo, también nos damos cuenta que en la zona de los contornos no se ha eliminado completamente el ruido y el borde no se suaviza. Esto se debe a que este método no regulariza en la dirección de los contornos. Con la ecuación de Perona y Malik (1.2.1) estos dos problemas se agravan más y, si el ruido es muy grande, no se consigue quitar completamente el ruido de las zonas homogéneas, formando especies de manchas uniformes.

En las pruebas con una imagen real (figura 1.5) vemos el efecto de las manchas uniformes que mencionábamos antes para la imagen izquierda. Para un valor de  $\lambda$  un poco mayor, imagen de la derecha, vemos que se suaviza demasiado. En la imagen de la izquierda vemos que en la zona de los contornos todavía quedan muchos pixels ruidosos que no se han regularizado y en la de la derecha se consiguen quitar muchos de ellos pero destruyendo muchos segmentos de borde.

El hecho de conservar el ruido en los contornos ya lo habían notado Perona y Malik – este efecto es mayor en la ecuación de Perona y Malik que en la de Catté et al. En realidad, en este esquema no se difumina en otra dirección que no sea en la dirección del gradiente de la imagen, con lo que no se regularizan los contornos con la información de los pixels vecinos.

### 1.3 Difusión por curvatura media

El problema de la regularización de los contornos en la dirección ortogonal al gradiente se consigue con los esquemas basados en la curvatura media. Uno de los primeros en introducir la ecuación de difusión por curvatura media en el campo del tratamiento de imágenes fueron Alvarez et al.[ALM92]. La ecuación de la curvatura media viene



Figura 1.5: Resultado de aplicar el esquema de Catté et al. (1.2.2) a la imagen de Alfredo Kraus (fig. 1.2). En las dos pruebas hemos utilizado  $\lambda = 1$  para la imagen de la izquierda y  $\lambda = 2$  para la de la derecha. El número de iteraciones es de 500 y  $\sigma = 1$ .

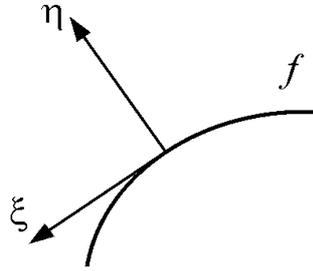


Figura 1.6: Vector  $\eta$ ,  $\xi$ , paralelo y perpendicular al gradiente de la función,  $\nabla f$ , respectivamente.

dada por,

$$f_t = \|\nabla f\| \operatorname{div} \left( \frac{\nabla f}{\|\nabla f\|} \right) = f_{\xi\xi} \quad (1.3.1)$$

En la figura 1.6 vemos que el vector  $\xi$  es perpendicular al gradiente de la señal. Esta ecuación tiene la propiedad de ejercer la difusión de la imagen en la dirección perpendicular al gradiente proporcional al valor de la curvatura en cada punto, es decir, cuanto mayor es la curvatura, mayor será el efecto de la difusión.

Lo que propusieron Alvarez et al. en [ALM92] fue la ecuación de la difusión por curvatura media ponderada con una función  $g(s)$  descendiente, similar a las vistas anteriormente.

$$f_t = g(\|\nabla f_\sigma\|) \|\nabla f\| \operatorname{div} \left( \frac{\nabla f}{\|\nabla f\|} \right) = g(\|\nabla f_\sigma\|) f_{\xi\xi} \quad (1.3.2)$$

Lo que se consigue con la función  $g(s)$  es el efecto inhibitorio que habíamos visto para las ecuaciones (1.2.1) y (1.2.2) a la hora de efectuar la difusión en la dirección de los contornos.

Aplicamos la ecuación de la difusión por curvatura media (1.3.1) a las imágenes que estamos considerando (figuras del círculo 1.1 y de la estatua 1.2) para comprobar su comportamiento. En este caso el método numérico que hemos implementado es de tipo explícito y el incremento de tiempo utilizado es de 0.5. En la figura 1.7 vemos el resultado para el círculo y en la figura 1.8

Vemos que en estas pruebas se regulariza bastante en la dirección de los contornos, desapareciendo el ruido de toda la imagen. Cada curva de nivel de la imagen progresa de forma independiente, reduciendo su curvatura media hasta tomar la forma de un



Figura 1.7: Resultado de aplicar el esquema de *difusión por curvatura media* (1.3.1) a la imagen del círculo y los rectángulos (fig. 1.1). Hemos implementado un esquema numérico explícito por diferencias finitas y hemos variado el número de iteraciones en cada prueba: izquierda, solución de la ecuación con 100 iteraciones; centro, con 250 y; derecha, con 500 iteraciones.

círculo. A medida que pasan las iteraciones, las curvas van disminuyendo de tamaño hasta convertirse en un punto.

Se disminuye el tamaño de los objetos puesto que se difumina de forma proporcional a la curvatura de la imagen: a mayor curvatura, mayor es el efecto de la difusión. Con el modelo de Alvarez et al. (1.3.2) se utiliza la función  $g(s)$  para evitar que las curvas de nivel con módulo de gradiente superior a un umbral disminuyan su tamaño.

El efecto de regularización en las zonas homogéneas es distinto para el método de Perona–Malik que para el de curvatura media. Para el primero, estas zonas se difuminan a través de una difusión isotrópica, mientras que para el segundo, se difuminan haciendo desaparecer las curvas de nivel.

Si desarrollamos la ecuación de Perona–Malik (1.2.1), llegamos a la siguiente expresión:

$$f_t = \frac{g'(\|\nabla f\|)}{\|\nabla f\|} f_{\xi\xi} + g''(\|\nabla f\|) f_{\eta\eta}$$

El término  $f_{\xi\xi}$  representa la difusión en la dirección ortogonal al gradiente y el término  $f_{\eta\eta}$  representa la difusión en la dirección paralela al mismo (ver figura 1.6). Ajustando los coeficientes de estos términos (elección de la función  $\Phi$ , ver [Kor98]), se puede obtener el comportamiento deseado.

En muchos esquemas de difusión, la función  $g(s)$  se elige de forma que en zonas homogéneas tanto  $\frac{g'(\|\nabla f\|)}{\|\nabla f\|}$  como  $g''(\|\nabla f\|)$  tengan el mismo valor y se difumine por igual en todas direcciones, mientras que en las zonas de los contornos se busca que  $g''(\|\nabla f\|)$  se aproxime a cero para evitar la difusión en la dirección del gradiente. Tal y como se comenta en [Kor98] es difícil encontrar una función  $g(s)$  que se comporte como estamos proponiendo. En la práctica se considera muchas veces que existen dos funciones  $g_1(s)$  y  $g_2(s)$  independientes, una por cada término, que se ajustan para



Figura 1.8: Resultado de aplicar el esquema de *difusión por curvatura media* (1.3.1) a la imagen de Alfredo Kraus (fig. 1.2). La imagen de la izquierda es el resultado después de 25 iteraciones del método y la de la derecha 100 iteraciones, con un incremento de tiempo igual a 0.5.

obtener el comportamiento deseado. El inconveniente de esta aproximación está en que no siempre es posible encontrar una expresión para la energía (1.1.1).

## 1.4 Difusión utilizando tensores de difusión

El último tipo de esquema de difusión que vamos a comentar es el de la difusión, tanto lineal como no lineal, utilizando tensores de difusión.

$$f_t = \operatorname{div}(\mathbf{D}(\nabla f_\sigma)\nabla f) \quad (1.4.1)$$

El interés de utilizar un tensor de difusión es el de guiar el proceso de difusión en una dirección u otra dependiendo de alguna característica local de la imagen. Generalmente se suelen utilizar tensores que actúen de forma anisótropa en la zona de los contornos, realizando la difusión en la dirección ortogonal al gradiente, y de forma isotropa en zonas homogéneas.

Uno de los primeros en proponer un tensor de difusión con estas características para guiar el proceso de regularización, fueron Nagel y Enkelman [NE86]. El esquema de difusión que propusieron ellos se basaba en un tensor que no variaba con el tiempo, es decir, que no dependía de la propia señal de difusión. Ellos propusieron el tensor como una matriz de proyección en la dirección ortogonal al gradiente.

$$\mathbf{D}(\nabla h) = \frac{1}{\|\nabla h\|^2 + 2\lambda^2} \begin{pmatrix} \frac{\partial h^2}{\partial y} + \lambda^2 & -\frac{\partial h}{\partial x} \frac{\partial h}{\partial y} \\ -\frac{\partial h}{\partial x} \frac{\partial h}{\partial y} & \frac{\partial h^2}{\partial x} + \lambda^2 \end{pmatrix} \quad (1.4.2)$$

donde  $\lambda$  es un umbral que determina a partir de qué valor de la norma del gradiente se considera que existe un contorno. Si la norma del gradiente es muy superior al de  $\lambda$ ,  $\|\nabla f\| \gg \lambda$ , entonces este tensor se aproxima a una matriz de proyección ortogonal al gradiente

$$\mathbf{D}(\nabla h) = \frac{1}{\|\nabla h\|^2} \begin{pmatrix} \frac{\partial h^2}{\partial y} & -\frac{\partial h}{\partial x} \frac{\partial h}{\partial y} \\ -\frac{\partial h}{\partial x} \frac{\partial h}{\partial y} & \frac{\partial h^2}{\partial x} \end{pmatrix} \quad (1.4.3)$$

de forma que la difusión se comporta de forma anisótropa. Si por el contrario la norma del gradiente es cero, nos encontramos con un tensor de la forma,

$$\mathbf{D}(\nabla h) = \frac{1}{2\lambda^2} \begin{pmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{pmatrix}, \quad (1.4.4)$$

con lo que tenemos un esquema de difusión isotrópico

$$\operatorname{div}(\mathbf{D}(\nabla h)\nabla f) = \operatorname{div}\left(\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\nabla f\right) = \frac{1}{2}\Delta f$$

Si analizamos  $\mathbf{D}(\nabla h)$  se puede deducir fácilmente que esta matriz tiene como autovectores  $v_1 = \nabla h$  y  $v_2 = \nabla h^\perp$ . Los autovalores correspondientes vienen dados por

$$\lambda_1(\|\nabla h\|) = \frac{\lambda^2}{\|\nabla h\|^2 + 2\lambda^2} \quad (1.4.5)$$

$$\lambda_2(\|\nabla h\|) = \frac{\|\nabla h\|^2 + \lambda_2}{\|\nabla h\|^2 + 2\lambda^2} \quad (1.4.6)$$

Observamos que  $\lambda_1 + \lambda_2 = 1$  se cumple independientemente del valor de  $\nabla h$ . En el interior de los objetos tenemos que  $\nabla h \rightarrow 0$  y, por lo tanto,  $\lambda_1 \rightarrow \frac{1}{2}$  y  $\lambda_2 \rightarrow \frac{1}{2}$ . En contornos ideales donde  $\nabla h \rightarrow \infty$ , se obtiene que  $\lambda_1 \rightarrow 0$  mientras que  $\lambda_2 \rightarrow 1$ . Por consiguiente, obtenemos un comportamiento isótropo dentro de las regiones, y en el contorno de los objetos de la imagen el proceso suaviza anisótricamente a lo largo de los bordes. Este comportamiento es muy similar a los filtros de difusión anisótropos para el realzado de contornos [Wei96] y es bastante parecido, en esencia, a la curvatura media considerado en [ALM92]. En este sentido se puede ver el método de Nagel-Enkelmann como un predecesor de las técnicas modernas de EDP para restauración de imágenes.

El modelo que presentaron Nagel-Enkelmann es, por lo tanto, anisótropo y variante en el espacio, pero es lineal en su parte de difusión.

$$f_t = \operatorname{div}(\mathbf{D}(\nabla h)\nabla f) \quad (1.4.7)$$

En los métodos que proponemos se utilizan tensores de difusión temporalmente constantes de este tipo.

Los tensores de difusión no lineales de los filtros de difusión anisótropos son una función de la evolución de la imagen misma como en la ecuación (1.4.1). Iijima propuso filtros de difusión anisótropos lineales parecidos en los sesenta y siete en el contexto de reconocimiento óptico de caracteres (ver [WII99]). Para un tratamiento detallado de los filtros de difusión anisótropos consultar [Wei98a],



Figura 1.9: Resultado de realizar la difusión a través de tensores de difusión (1.4.1) a la imagen de la Bola (fig. 1.1). La imagen de la izquierda es el resultado después de 100 iteraciones del método, la del centro después de 500 iteraciones y la de la derecha con 1000, con un incremento de tiempo igual a 10,  $\lambda = 10$  y  $\sigma = 1$ .

para una clasificación axiomática de la curvatura media y EDP morfológicas relacionadas para el análisis de imágenes consultar [AGLM93], y una reciente colección de artículos acerca de métodos de suavizado de imágenes a través de EDPs [BDH<sup>+</sup>96, CMSE98, tHRFKE97, NJOE99]

En las figuras 1.9 y 1.10 vemos el resultado de aplicar un esquema de difusión utilizando tensores de difusión (1.4.1) variables en el tiempo. Implementamos un método implícito por diferencias finitas centradas<sup>1</sup>

Como podemos ver en ambas figuras, el resultado es similar al visto para el método de Perona–Malik. La ventaja que obtenemos con este es que los contornos se regularizan de forma perpendicular al gradiente, cosa que no ocurría en Perona–Malik. Los contornos son más suaves (se elimina todo el ruido) y no aparecen las "islas" de manchas que aparecían antes. En el caso de los dos rectángulos, se redondean un poco las esquinas, parecido al efecto que se produce en la difusión por curvatura media. Este efecto se puede regular ajustando el valor de  $\sigma$ : entre más pequeño sea, más nos aproximamos al esquema de Perona–Malik y menor es la difusión a lo largo del contorno.

## Análisis de los tensores de difusión

Podemos ver los tensores de difusión como matrices que dirigen el proceso de difusión atendiendo a alguna característica geométrica de la imagen. La forma de selección la dirección se puede realizar a través de matrices de proyección ponderadas por valores que aumentan la magnitud de la difusión. Así pues una forma sencilla e inmediata de proponer una matriz que cumpla estas características es la siguiente

$$\mathbf{D}(\nabla f_\sigma) = \left[ \phi_1(\|\nabla f_\sigma\|) \left( Id - \frac{\nabla f_\sigma \nabla f_\sigma^t}{\|\nabla f_\sigma\|^2} \right) + \phi_2(\|\nabla f_\sigma\|) \frac{\nabla f_\sigma \nabla f_\sigma^t}{\|\nabla f_\sigma\|^2} \right] \quad (1.4.8)$$

<sup>1</sup>En el apéndice C se explica en detalle la forma de discretizar el gradiente de la señal y la divergencia del tipo que estamos viendo en esta sección,  $div(D\nabla f)$



Figura 1.10: Resultado de realizar la difusión a través de tensores de difusión (1.4.1) a la imagen del Alfredo Kraus (fig. 1.2). La imagen de la izquierda es el resultado para  $\lambda = 3$  y la de la derecha para  $\lambda = 5$ , con número de iteraciones igual a 500, incremento de tiempo igual a  $\sigma = 1$ .

donde el primer término es la proyección en la dirección ortogonal al gradiente de la función y el segundo es la proyección en la dirección del gradiente.

En algunos esquemas de difusión se tiene en cuenta la difusión en la dirección del gradiente y se inhibe la difusión en la dirección ortogonal. En otros esquemas se pondera la difusión en ambas direcciones teniendo en cuenta algún criterio geométrico.

Las funciones  $\phi_i$  se utilizan para ponderar entre ambos términos, permitiendo la difusión tanto en la dirección  $\eta$ , como en  $\xi$ , como en ambas a la vez. La elección tradicional para estas funciones es la de realizar la difusión a lo largo de los contornos, en donde la norma del gradiente de la señal es grande y activar la difusión en todas direcciones cuando se trate de una región homogénea.

$$\begin{aligned}\phi_1 &= 1 \\ \phi_2 &= \begin{cases} 0 & \text{si } \|\nabla f_\sigma\| \geq \lambda \\ 1 & \text{si } \|\nabla f_\sigma\| < \lambda \end{cases} \end{aligned} \quad (1.4.9)$$

El comportamiento de este tensor es equivalente al de Nagel–Enkelmann (1.4.2) excepto por el efecto que ejerce  $\lambda$ : Mientras que en Nagel–Enkelmann  $\lambda$  se introduce en la ecuación de forma natural, en este último  $\lambda$  es un umbral perfecto, que discierne entre lo que es un contorno y lo que no. En el de Nagel–Enkelmann el comportamiento es perfectamente anisótropo o isótropo cuando se cumplen las condiciones extremas que vimos en párrafos anteriores. En sitios donde la norma del gradiente no está muy lejos del valor de  $\lambda$ , el comportamiento es distinto. A niveles prácticos, los resultados de los dos tensores de difusión son prácticamente iguales, sin embargo, cuando el rango de valores dentro de la imagen no es muy alto – en imágenes de niveles de grises suele estar entre 0 y 255 –, el efecto de  $\lambda$  sobre el tensor de Nagel–Enkelmann puede ser perjudicial.

Si analizamos el tensor (1.4.8) se puede deducir fácilmente que esta matriz tiene como autovectores  $v_1 = \nabla f^\perp$  y  $v_2 = \nabla f$ . Los autovalores correspondientes vienen dados por

$$\lambda_1(\|\nabla f\|) = \phi_1 \quad (1.4.10)$$

$$\lambda_2(\|\nabla f\|) = \phi_2 \quad (1.4.11)$$

Ajustando los valores de  $\phi_1$  y  $\phi_2$  podemos obtener comportamientos distintos. Por ejemplo, si  $\phi_1 = 1$  y  $\phi_2 = 0$  tenemos un comportamiento anisótropo en toda la imagen, similar en esencia a la difusión por curvatura media; si  $\phi_1 = 0$  y  $\phi_2 = 1$  entonces la difusión se realizará siempre en la dirección del gradiente; y si  $\phi_1 = 1$  y  $\phi_2 = 1$  el comportamiento será isótropo.

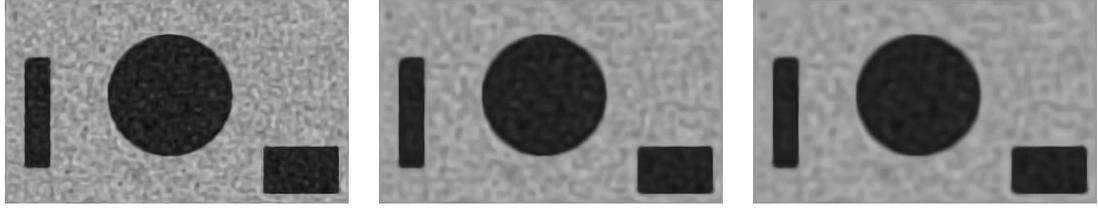


Figura 1.11: Resultado de realizar la difusión a través del tensor de difusión expuesto en (1.4.8) con valores de funciones  $\phi_1 = 1$  y  $\phi_2 = 0$ . La imagen de la izquierda es el resultado después de 100 iteraciones, la del centro después de 500, y la de la derecha con 1000, con un incremento de tiempo igual a 10 y  $\sigma = 1$ .

En las imágenes 1.11 y 1.12 mostramos dos ejemplos utilizando el tensor de difusión que se genera por las funciones  $\phi_1 = 1$  y  $\phi_2 = 0$ . Como podemos ver, la difusión sólo se realiza en la dirección ortogonal al gradiente, haciendo progresar las líneas de nivel independientemente las unas de las otras tal y como ocurriría con el esquema de difusión por curvatura media, pero sin la velocidad de ésta.

Si observamos el esquema (1.4.8) nos damos cuenta que cuando  $\sigma \rightarrow 0$ , se cumple que  $\mathbf{D}(\nabla f_\sigma)$  se convierte en una matriz de proyección alineada con el gradiente de la imagen, luego  $\mathbf{D}(\nabla f)\nabla f = \phi_2(\|\nabla f\|)\nabla f$ , convirtiéndose exactamente en el esquema de Perona–Malik (1.2.1). El papel que juega  $\sigma$  en los tensores de difusión es el de definir el orden del error que se supone posee una imagen. Generalmente se suelen tomar valores pequeños, del orden de 1 píxel (entre 0.5 y 1).

Si comparamos los tres métodos de difusión vistos hasta ahora, llegamos a las siguientes conclusiones: El método de Perona–Malik se puede ver como un caso especial de un tensor de difusión como el visto en (1.4.8) con la característica de que  $\sigma = 0$ . La difusión por curvatura media y por tensores de difusión se asemejan en que ambos difuminan en la dirección de las líneas de nivel, aunque la magnitud de la difusión varía entre uno y otro. Aunque en el caso de los tensores de difusión tenemos un parámetro más,  $\sigma$ , la ventaja que obtenemos con estos es que se puede adaptar el mecanismo de difusión isótropo y anisótropo en un mismo tensor de forma natural.

## 1.5 Tensores de difusión lineales

Casi todos los métodos de difusión que hemos visto en las secciones anteriores dependen exclusivamente de una sola señal. Según la clasificación que se hace en [Wei97] los esquemas que hemos visto anteriormente son todos no lineales, ya que la función de difusividad,  $g(s)$  o la matriz  $\mathbf{D}$ , dependen de la misma señal, excepto el esquema de Nagel–Enkelmann en el que el tensor es constante en el tiempo. En esta sección nos vamos a interesar por estos últimos. Vamos a estudiar la difusión anisótropa lineal.



Figura 1.12: Resultado de realizar la difusión a través del tensor de difusión expuesto en (1.4.8) con valores de funciones  $\phi_1 = 1$  y  $\phi_2 = 0$ , después de 500 iteraciones, con un incremento de tiempo igual a 10 y  $\sigma = 1$ .



Figura 1.13: (Izquierda) Imagen sintética de un círculo y dos rectángulos (señal  $f$  objeto del proceso de difusión) y (derecha) máscara que se va a utilizar para calcular el tensor de difusión (señal  $h$  de la ecuación (1.5.1)).



Figura 1.14: A la izquierda aparece la foto de la Catedral de Arucas y a la derecha la máscara que vamos a utilizar para determinar el tensor de difusión.

En muchas aplicaciones, como ocurre en las técnicas de matching, en el flujo óptico o en la visión estereoscópica, lo que nos interesa es obtener una solución regularizada, pero preservando algunas características que aparecen en otra imagen. En el caso del flujo óptico, por ejemplo, queremos que el flujo final estimado sea regular en zonas homogéneas y se regularize a lo largo de los contornos, preservando las discontinuidades que aparecen en la imagen. En este caso, el flujo es la solución que buscamos y la imagen original (la de partida) suele ser sobre la que se imponen las restricciones para establecer cuáles son los contornos y las regiones homogéneas. Estas restricciones son las que se definen a través del tensor de difusión. Cuando proponemos una ecuación como esta:

$$f_t = \operatorname{div}(\mathbf{D}(\nabla h)\nabla f) \quad (1.5.1)$$

estamos diciendo que queremos realizar una difusión de la señal  $f$  respetando las restricciones definidas por  $\nabla h$ . Es la matriz de proyección  $\mathbf{D}$  la que dirige el proceso de difusión, considerando la información del gradiente de  $h$ .

En las figuras 1.13, 1.14 y 1.15 mostramos tres ejemplos de imágenes y sus máscaras. Las máscaras son iguales a las imágenes originales excepto por algunos trozos que hemos quitado intencionadamente para observar el efecto de la difusión



Figura 1.15: A la izquierda foto de Herve y a la derecha la máscara para el cálculo del tensor de difusión.

lineal a través de tensores de difusión. En el caso del círculo y los rectángulos hemos quitado la mitad del círculo y parte de los rectángulos. En las otras dos imágenes hemos puesto en color blanco el fondo de las imágenes conservando la Catedral y la cara de Herve. Nótese que lo importante al definir las máscaras no es tanto el dar un color a cada zona sino el determinar bien las discontinuidades dentro de las imágenes. Al igual que utilizamos estas máscaras, podríamos haber utilizado otras en donde sólo aparecieran los contornos de la señal, quedándonos con aquellos que nos interesa conservar.

En las figuras 1.16, 1.17 y 1.18 mostramos los resultados después de aplicar el esquema dado en (1.5.1). La señal  $f$  viene dada por la imagen de la izquierda de 1.13, 1.14 y 1.15, y la función que hace de máscara  $h$  está definida por la imagen de la derecha.

En la figura 1.16 se pueden ver cuatro resultados para el ejemplo del círculo y los rectángulos. Como se puede ver, la imagen se difumina de forma isótropa en las zonas donde la máscara es homogénea, y de forma anisótropa en la zona de los contornos. Vemos cómo poco a poco van desapareciendo la mitad del círculo y parte de los rectángulos al mismo tiempo que se respetan la otra mitad del círculo y las otras partes de los rectángulos.

En los ejemplos de las figuras 1.17 y 1.18 vemos cómo se preservan durante más tiempo las zonas donde están definidas las máscaras, mientras que el resto desaparece.

A partir del capítulo 6 aplicaremos este tipo de esquema a los problemas de la estimación del flujo óptico y del cálculo de la carta de disparidad. Estos dos problemas son un claro ejemplo en el que la solución que estamos buscando depende de otra señal, la imagen original, que es la que debe guiar el proceso de difusión.

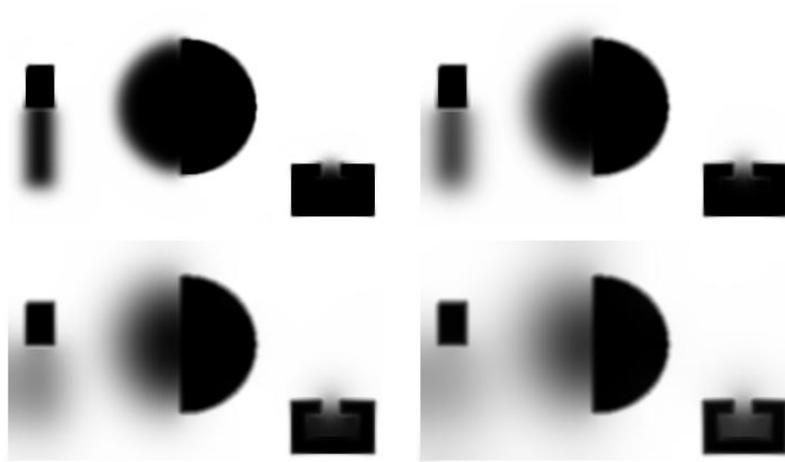


Figura 1.16: Resultado de aplicar el esquema (1.5.1) a la imagen dada en 1.13. La imagen superior izquierda es el resultado después de 500 iteraciones, la superior derecha después de 1500, la inferior izquierda después de 5000 y la inferior derecha después de 10000.



Figura 1.17: Resultado de aplicar el esquema (1.5.1) a la imagen dada en 1.14. La imagen izquierda es el resultado después de 500 iteraciones, la del centro después de 1500 y la derecha después de 5000.



Figura 1.18: Resultado de aplicar el esquema (1.5.1) a la imagen dada en 1.15 después de 500 iteraciones.

# Capítulo 2

## Flujo óptico

### 2.1 Introducción

La definición de flujo óptico fue introducida por primera vez por Gibson [Gib50] en el contexto de la visión humana. El cálculo del flujo óptico consiste en encontrar el desplazamiento aparente de los objetos en una secuencia de imágenes. Recuperar este desplazamiento es uno de los problemas claves en visión por ordenador y se han realizado muchos trabajos de investigación sobre este campo durante los últimos años.

Si representamos una imagen como una aplicación  $I : (x, y, t) \rightarrow I(x, y, t)$  donde  $(x, y)$  representa la coordenada espacial de la imagen y  $t$  el tiempo, se puede ver una secuencia de imágenes como la variación de la intensidad en las coordenadas de la imagen a través del tiempo. Si representamos esto,

$$I(x, y, t) = I(x + dx, y + dy, t + dt)$$

donde  $\vec{v} = (dx, dy)$  es el desplazamiento de la imagen en el punto  $(x, y)$  después de  $dt$  unidades de tiempo – en este documento utilizaremos indistintamente  $\vec{v} = (dx, dy)$  como  $\vec{h} = (u, v)$  para referirnos al vector velocidad en un punto. El desarrollo en series de Taylor de esta función,

$$I(x, y, t) = I(x, y, t) + \nabla I(x, y, t) \begin{pmatrix} dx \\ dy \end{pmatrix} + I_t dt + \mathcal{O}^2$$

que simplificando se nos queda en

$$\nabla I(x, y, t) \begin{pmatrix} dx \\ dy \end{pmatrix} + I_t dt = 0 \tag{2.1.1}$$

La ecuación 2.1.1 se denomina ecuación de *restricción del flujo óptico*. A partir de esta ecuación no es posible determinar  $\vec{v} = (dx, dy)^T$  ya que es un *problema mal condicionado* [BPT88b]. Tan solo se puede estimar  $\vec{v}_\perp$ , la componente en la dirección normal a la curva de nivel, en la dirección del gradiente. Este fenómeno se conoce como el *problema de la apertura* [Ull79].

$$\vec{v}_\perp = \frac{-I_t \nabla I(x, y, t)}{\|\nabla I(x, y, t)\|^2}$$

Para tratar con este problema lo que se suele hacer es imponer alguna restricción sobre el flujo óptico. En muchos métodos se suele imponer de forma implícita un modelo *Lambertiano*, en el que las intensidades de los pixels correspondientes no varían de una imagen a otra. Este es un tema delicado de difícil solución debido a la gran complejidad de las escenas reales. También se suelen imponer restricciones sobre el flujo, denominadas restricciones de suavidad, con el fin de volver el problema bien condicionado, tal y como propusieron Horn y Schuck [HS81].

En este capítulo haremos un análisis de los factores que hacen del cálculo del flujo óptico un problema complicado. Luego comentaremos las técnicas que se han propuesto en los últimos años, refiriéndonos a los trabajos realizados por Mitiche y Boutheimy [MB96], en donde se establece una clasificación de las técnicas más utilizadas, por Barron et al. [BFB94], en donde se evalúa el rendimiento de algunas de estas técnicas, por Jähne y Haussecker [JH98] y por Galvin et al. [GMN98], donde se evalúan ocho algoritmos distintos. Por último comentaremos una serie de imágenes de prueba sintéticas y reales que utilizaremos durante el desarrollo de este trabajo, algunas de las cuales se han utilizado en Barron et al. [BFB94].

## 2.2 Factores que influyen sobre las imágenes

El cálculo del flujo óptico es un problema clásico mal condicionado en visión por ordenador y requiere suposiciones adicionales de regularización. Aparte de este problema, nos encontramos una serie de factores que convierten el problema de la estimación del flujo óptico en un problema de difícil solución. Estos factores están condicionados por el hecho de que las imágenes que tratamos son una representación de una escena del espacio tridimensional. Muchos de los factores derivan de la simple proyección de la escena en una imagen bidimensional; otros, de la riqueza de matices y complejidades que existe en el mundo real.

La definición del flujo óptico es el movimiento aparente de los niveles de intensidad de la imagen. Muchos de los factores que vamos a citar influyen de una u otra manera

en el nivel de intensidad de los puntos de la imagen: para puntos en correspondencia, el nivel de intensidad puede diferir mucho de una imagen a otra.

Podemos considerar que existen dos tipos de factores principales que influyen en las imágenes:

- **Factores fotométricos:** son aquellos que se deben a las fuentes de iluminación, a las características de reflectancia de la superficie de los objetos y otras características ópticas.
- **Factores geométricos:** derivados de la geometría de la imagen y de la escena. Dependen de las características de los objetos y de las relaciones entre estos.

Hablaremos de cada uno de estos tipos en las siguientes secciones, intentando localizar en cada caso los aspectos fundamentales.

### 2.2.1 Factores fotométricos

Entre estos factores podemos considerar las principales fuentes de iluminación, las propiedades de reflectancia de las distintas superficies y la geometría de la escena que afecta a la iluminación mutua y sombras.

Toda superficie posee unas propiedades de reflectancia características, que de forma simplificada se pueden dividir en *superficies lambertianas* – aquellas cuyo brillo es igual, independientemente desde la dirección en que se mire – y *superficies especulares* – aquellas que reflejan toda la luz en una misma dirección, generalmente de forma perpendicular a la superficie. En muchos casos, la función en la superficie de reflectancia se aproxima a una combinación lineal de un término de difusión (comportamiento lambertiano) y un término especular. En la mayoría de los métodos para el cálculo del flujo óptico se suelen suponer que las superficies tienen un tipo de reflectancia lambertiana (*suposición lambertiana*).

Cuando tratamos con imágenes reales, esta suposición no es correcta. El efecto de los brillos hace que los niveles de intensidades de las superficies cambien de una imagen a otra. Esto se debe, por ejemplo, a que los objetos al cambiar de posición no reflejan la luz de igual manera, o si la fuente de luz se desplaza, entonces los brillos cambian en todos los objetos de la imagen. Un ejemplo claro de este efecto es el de una esfera homogénea en intensidad que gira: Si la fuente de luz permanece en la misma posición, parecería que la esfera no se estaría moviendo; sin embargo, si es la fuente de luz la que cambia de posición, entonces se apreciaría un flujo en la esfera que no se correspondería con un movimiento real.

Aparte de estos problemas nos encontramos otros que tienen que ver con las condiciones de luminosidad de la escena. Entre estos podemos destacar:

- **Las sombras:** Depende de la intensidad de las fuentes de luz y de la distribución de los objetos. Cuando un objeto pasa de estar en sombra a ser iluminado, o viceversa, los niveles de grises cambian siendo difícil encontrar su homólogo. De forma general, Es difícil identificar qué zonas están afectadas por las sombras.
- **Los reflejos:** Generalmente, los objetos reflejan la luz, influyendo sobre la irradiación que llega a otros. Dependiendo de las características de reflectancia de los objetos este efecto puede ser más o menos pronunciado.
- **Cambios de posición e intensidad de las fuentes luminosas:** Cuando en una escena hay fuentes de luz que cambian de posición o intensidad, estas influyen sobre los niveles de grises aumentando o disminuyendo su intensidad y creando nuevas zonas de brillos.
- **Transparencias:** En muchas escenas existen objetos transparentes, tales como cristales y objetos de plástico, que aparte de tener un comportamiento altamente especular, son difíciles de detectar y provocan que objetos en movimiento que estén situados detrás de estos sean más difíciles de captar.

Es complicado, por lo tanto, encontrar una solución precisa si no existe una iluminación uniforme en los objetos de la escena, que las superficies reflectantes sean de tipo lambertianas, y que no se encuentren los inconvenientes mencionados en los puntos anteriores. Además, de la ecuación de restricción del flujo óptico (5.0.2) se deduce que para obtener de forma exacta el flujo, los objetos se deberían enfrentar a una traslación pura paralela al plano de la imagen.

## 2.2.2 Factores geométricos

Otro tipo de factores que afectan a la complejidad de las secuencias de imágenes se deben a los estrictamente geométricos, es decir, aquellos que dependen de la forma y posición de los objetos en la escena, de las relaciones que existen entre ellos y de las transformaciones que sufren al pasar de una imagen a otra. En la siguiente lista comentamos los que nosotros creemos más representativos y que suelen aparecer con mayor frecuencia en las secuencias que hemos tratado:

- **Oclusiones:** Cuando un objeto tapa a otro u oculta un trozo de otro al desplazarse. Es difícil estimar el movimiento de la parte que queda oculta. En algunos casos lo que se suele hacer es prever el movimiento dependiendo del desplazamiento que ha sufrido en frames anteriores.
- **Objetos que desaparecen:** El tamaño de las imágenes es limitado, luego puede ocurrir que un objeto que está en una parte de la secuencia, termine

desapareciendo a partir de un frame determinado. También pueden desaparecer los objetos al ser ocultados por otros que están desplazándose en posiciones más próximas a la cámara.

- **Objetos que aparecen:** Es el caso contrario al anterior. A partir de un instante determinado, puede aparecer parcialmente un objeto por uno de los límites de la imagen o puede aparecer detrás de algún otro objeto que lo tapaba completamente.
- **Objetos no rígidos:** Son objetos que pueden cambiar de forma. Es más complicado encontrar primitivas que ayuden a identificar el desplazamiento del objeto, así como resulta difícil determinar de qué manera se deforma.
- **Proyección perspectiva:** El modelo de cámaras que estamos tratando es el proyectivo. La transformación en perspectiva es bastante compleja, haciendo que los objetos de las imágenes puedan cambiar bastante su forma dependiendo de la trayectoria que tengan en la escena y de su posición con respecto a la cámara.
- **Trayectorias arbitrarias:** Como ya comentamos en la sección anterior, el movimiento más adecuado para el cálculo del flujo óptico es el paralelo al plano de la imagen. Cuando tenemos movimientos de rotación, escalado o cualquier otro más complejo, el desplazamiento de los pixels del objeto no es igual, pudiendo incluso desaparecer parte del objeto o deformarse por la proyección en la imagen.

Además de estos factores, nos encontramos con otros problemas que están relacionados con la cámara: 1) El ruido debido a los sensores de obtención de imágenes; 2) Distorsiones de la cámara que dependen de los parámetros intrínsecos de la misma (los veremos en el capítulo 3) y del efecto de la lente.

También es importante tener en cuenta la existencia de múltiples objetos que se mueven en distintas direcciones. En muchas escenas es normal que hayan varios objetos y que sus movimientos no estén alineados. Entonces habrá que tener cuidado de poder discernir correctamente entre los movimientos de cada objeto.

Por último, mencionar que el problema del cálculo del flujo óptico también depende de la magnitud del movimiento de los objetos. En algunos casos, cuando se toman muchas imágenes de una secuencia, es probable que los desplazamientos sean del orden de unos pocos pixels, pero cuando consideramos imágenes en instantes de tiempo grandes, la búsqueda de correspondencias se puede extender a todo el dominio de la imagen, aumentando las posibilidades de error en las estimaciones.

## 2.3 Clasificación de las distintas técnicas

Se ha dedicado mucho esfuerzo al estudio del problema del flujo óptico y se han realizado algunas clasificaciones a través de los años. En esta sección haremos una clasificación de las técnicas de estimación del flujo óptico centrándonos principalmente en los trabajos realizados por Mitiche y Bouthemy [MB96], por Barron et al. [BFB94], Beauchemin y Barron [BB95] y por Fleet [Fle92]. Para profundizar sobre este tema recomendamos leer estos trabajos.

Vamos a comentar las características principales de cada tipo de método y haremos referencia a los métodos más conocidos. Salvando la dificultad que supone clasificar todos los métodos en unos pocos items, comentaremos los subgrupos principales de cada tipo.

### 2.3.1 Métodos diferenciales

En estos métodos lo que se persigue es calcular el flujo óptico a través de las derivadas espaciales y temporales de las intensidades de la imagen. Partimos de la ecuación de restricción del flujo óptico (5.0.2) vista en la introducción de este capítulo, como ecuación fundamental. Como ya mencionamos, la estimación del flujo óptico a través de esta ecuación es un *problema mal condicionado*, además de estar afectado por el *problema de la apertura* que ya comentamos.

Los métodos diferenciales se dividen a su vez en distintos subgrupos. En Beauchemin y Barron [BB95] se identifican hasta cinco subgrupos distintos, aunque nosotros consideraremos los dos más básicos. Estos son los *métodos globales* y los *métodos locales*. Los primeros se caracterizan por utilizar la ecuación de la restricción del flujo óptico conjuntamente con un término de regularización para evitar los problemas anteriores. El primero en introducir uno de estos métodos fueron Horn y Schunk [HS81] al imponer que el flujo óptico variara de forma suave. Ellos propusieron una energía del tipo

$$\int_{\Omega} ((\nabla f \cdot \vec{v} + f_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2)) d\omega$$

con  $\vec{v} = (u, v)$ . A partir de esta idea han surgido muchos métodos con distintos términos de regularización. Otros métodos que se han propuesto son los de Nagel [Nag83, Nag87], Uras et al. [UGVT88]

Los métodos locales no imponen ninguna restricción de regularización, sino que sólo tienen en cuenta la información local de la imagen para estimar el desplazamiento. Para esto, lo que se ha hecho tradicionalmente es utilizar más funciones que las intensidades de las imágenes para evitar que el problema esté mal condicionado, tal

y como propusieron Mitiche et al. [MWA87] o Wohn et al. [WDT83]. Otros autores como Tretiak y Pastor [TP84] o Campani y Verri [CV90] propusieron la utilización de operadores diferenciales. En Lucas y Kanade [LK81, Luc84] se calcula el flujo óptico como una solución de una ecuación de mínimos cuadrados pesados por una función ventana.

### 2.3.2 Métodos basados en la correlación

Estos métodos se basan en la búsqueda de correspondencias utilizando ventanas o patrones alrededor de cada pixel. La ventaja que tienen estos métodos con respecto a los anteriores, es que se utiliza información de los vecinos que en muchos casos hacen que la búsqueda de la correspondencia sea más efectiva.

Lo que se busca con estos métodos es asociar pixels a través de regiones similares en las imágenes que se obtienen por maximización de alguna medida de similaridad. En el caso continuo, la correlación se expresa como

$$C(\vec{v}) = \int_{\Omega} f(\mathbf{x} + \vec{v}) g(\mathbf{x}) d\omega$$

Cuando estamos en el caso discreto entonces una medida de la correlación en un punto que se toma como estándar es la siguiente:

$$C(\vec{v}) = \frac{\sum_{\delta\vec{v}=-\langle a,b \rangle}^{\langle a,b \rangle} \left( f(\mathbf{x} + \delta\vec{v}) - \overline{f(\mathbf{x})} \right) \left( g(\mathbf{x} + \vec{v} + \delta\vec{v}) - \overline{g(\mathbf{x} + \vec{v})} \right)}{(2a+1)(2b+1)\sigma_f(\mathbf{x})\sigma_g(\mathbf{x} + \vec{v})},$$

donde  $(a, b)$  representa las dimensiones de la ventana de correlación,  $\overline{f(\mathbf{x})}$  la media de la imagen  $f$  en esa ventana y  $\sigma_f(\mathbf{x})$  la desviación estándar correspondiente.

Como métodos representativos de este tipo podemos mencionar los de Kories y Zimmerman [KZ86], Sutton et al. [SWP<sup>+</sup>83], Kalivas y Sawchuk [KS91], Little et al. [LBP88].

### 2.3.3 Métodos basados en la frecuencia

Estas técnicas utilizan la transformada de Fourier para calcular el flujo a través del dominio de la frecuencia. Estas suelen ser más adecuadas que las anteriores para extraer el movimiento de objetos que son difíciles de capturar por los anteriores, como puede ser el caso de puntos aleatorios.

Las técnicas basadas en frecuencias se apoyan en el hecho de que un patrón que se mueve de una imagen a otra por una simple traslación, tiene representaciones en Fourier que difieren solo en fase. Castro et al. [DCM87] encontraron una generalización para incluir las rotaciones. El problema de la apertura sigue presente y, por lo tanto, estos métodos deben incorporar términos de regularización igual que ocurría con los métodos diferenciales.

Algunas referencias básicas a técnicas de este tipo son los trabajos de Adelson y Bergen [AB85], Fleet y Jepson [FJ90], Heeger [Hee88], Watson y Ahumada [WJ85].

En el libro de David J. Fleet [Fle92] se propone otra clasificación en la que se añaden otros tipos. Entre éstos están las técnicas basadas en características y las técnicas basadas en energías, aunque éstas se pueden encuadrar dentro de las clasificaciones anteriores.

## 2.4 Estudio de métodos diferenciales

En el presente trabajo vamos a desarrollar un método que está basado en una energía. Esta energía hace uso del gradiente de la imagen aunque no utilizamos la ecuación de restricción del flujo óptico (5.0.2).

Nuestro trabajo está en relación con los métodos diferenciales globales del tipo de Horn y Schunck [HS81] en el que se impone una regularización que refleja la suposición de que el campo de flujo óptico varía de forma suave en el espacio.

Después del trabajo de Horn y Schunck han aparecido muchos trabajos de investigación con el fin de permitir campos de flujos que sean continuos por trozos, es decir, que se imponga la restricción de regularización dentro de los límites de los objetos definidos por sus contornos. Entre estos caben destacar los trabajos de [ADK99, BA91, BA96, BFBG93, Coh93, DKA95, GR96, HB93, KTB96, MP98, NE86, Nes93, PGPO94].

Adelantando un poco el contenido de nuestro trabajo, uno de los objetivos que nos habíamos planteado era el de demostrar que la combinación de ideas sobre análisis multiescala lineales y no lineales nos pueden llevar a métodos diferenciales bien definidos que permiten recuperar desplazamientos grandes.

Consideramos dos imágenes  $I_1(x, y)$  e  $I_2(x, y)$  (definidas  $\mathbb{R}^2$  para simplificar la discusión) que representan dos vistas consecutivas en una secuencia de imágenes. Bajo la suposición de que dos pixels correspondientes tienen igual nivel de gris, la determinación del flujo óptico desde  $I_1(x, y)$  a  $I_2(x, y)$  se reduce a encontrar una función  $\vec{h}(x, y) = (u(x, y), v(x, y))$  tal que

$$I_1(x, y) = I_2(x + u(x, y), y + v(x, y)), \forall (x, y) \in \mathbb{R}^2$$

Tal y como vimos anteriormente, para calcular la ecuación precedente se suele

linealizar generando la ecuación de restricción del flujo óptico

$$I_1(\mathbf{x}) - I_2(\mathbf{x}) \simeq \langle \nabla I_2(\mathbf{x}), h(\mathbf{x}) \rangle$$

donde  $\mathbf{x} = (x, y)$ . La restricción de flujo óptico linealizado se basa en la suposición que los desplazamientos de los objetos,  $\vec{h}(\mathbf{x})$ , son pequeños o que la imagen varía lentamente en el espacio. En otros casos, esta linealización no es válida.

Después de mencionar esto y aunque todavía no hemos expuesto en qué consiste nuestro método, vamos a considerar algunos trabajos que están relacionados: Por ejemplo, en el trabajo de Proesmans et al. [PPG94], [PGPO94] estudiaron una aproximación parecida que también prescindía de la linealización de la restricción del flujo óptico con el fin de estimar grandes desplazamientos. Su método, sin embargo, requiere un sistema de seis ecuaciones en derivadas parciales y su proceso de difusión no lineal es de tipo escalar en vez de utilizar tensores de difusión. Su proceso de difusión se guía por el flujo mientras que el nuestro se guía por la imagen. Otra técnica en EDP que es similar en esencia a la de Proesmans et al. es un método de estero por Shah [Sha93]. Otras técnicas en donde la regularización está guiada por el flujo, y con propiedades de preservación de contornos, incluyen los trabajos de Aubert et al [ADK99], Cohen [Coh93], Deriche [DKA95], Hinterberger [Hin99], Kumar et al. [KTB96], Schnörr [Sch94], Weickert [Wei98b] y Weickert y Schnörr [WS99]. Se han estudiado otras aproximaciones relacionadas con la regularización a través de estimadores estocásticos tales como en Black y Anandan [BA91], [BA96], Blanc-Feraud et al. [BFBG93], Heitz y Bouthemy [HB93] y Mémin y Pérez [MP98]. La aproximación anisotrópica guiada por la imagen de Nagel y Enkelmann ha sido objeto de muchos estudios. Algunos ejemplos incluyen trabajos posteriores de Nagel [Nag87], [Nag90] así como de Schnörr [Sch91a], [Sch91b] y Snyder [Sny91]. Una realización multimalla de este método ha sido descrita por Enkelmann [Enk88] y Anandan [Ana89] ha realizado un trabajo desde el punto de vista piramidal. En Alvarez et al. [AELS99] se estudia un método de flujo óptico isótropo guiado por la imagen. Con respecto a la introducción de un marco de análisis multiescala lineal, nuestro método se puede relacionar también con la aproximación del flujo óptico realizado por Florack et al. [FNN98]. Su método difiere del nuestro en el sentido que es puramente lineal, aplica mecanismos de selección de escalas y no utiliza un término de regularización no lineal que preserve las discontinuidades. Nuestra estrategia de focalización para evitar terminar en mínimos locales irrelevantes también se parece a los algoritmos de *graduated non-convexity (GNC)* propuesto por Blake y Zisseman [BZ87]. Los trabajos de Alvarez et al. [ADSW00, AWS99, AWS00, AS00a] son publicaciones que resumen el trabajo que se expone en esta tesis. Un método de flujo óptico relacionado para la generación de películas entre dos imágenes ha sido propuesto por [Hin99].

## 2.5 Imágenes de prueba

Durante el desarrollo de esta tesis se han utilizado una serie de imágenes para realizar mediciones y comparaciones de los métodos que proponemos.

Algunas de las imágenes que mostramos a continuación son bastante utilizadas en el campo de la visión por ordenador y son referentes para la evaluación de los métodos. Algunas de estas imágenes se han obtenido a partir del trabajo desarrollado por Barron et al. [BFB94], en donde comparan varios métodos utilizando una serie de imágenes sintéticas.

A continuación mostramos algunas imágenes sintéticas para las que mostramos sus resultados exactos. En el siguiente punto mostramos algunas imágenes reales que se han utilizado en varios trabajos de investigación.

### 2.5.1 Imágenes sintéticas

Utilizamos 4 secuencias de imágenes sintéticas. La primera es la secuencia de los cuadrados, en la que hay cuatro cuadrados que se mueven en direcciones distintas y con distinta magnitud. Las siguientes 2 las utilizaron Barron et al. para realizar las pruebas en su trabajo [BFB94]: Una es la secuencia *Yosemite*, que representa un escena compleja de un paisaje natural, con desplazamientos variados y la otra es la secuencia de un cuadrado que se mueve con velocidad constante. La última secuencia es la de las *Torres de Mármol*, copyright de Nagel.

#### Secuencia de los cuadrados

Esta es la primera secuencia sintética que hemos desarrollado nosotros para realizar las pruebas prácticas. Esta secuencia (ver figura 2.1) consiste en cuatro cuadrados que se desplazan de formas diferentes tanto en módulo como en dirección.

En las figuras 2.2 y 2.3 se muestran los desplazamientos exactos para la secuencia de los cuadrados con mapas de niveles de grises. Cuando tratemos el problema del flujo óptico en la siguiente parte, calcularemos los desplazamientos en una sola dirección, mientras que en la parte donde se introduce las simetrías, se calcularán los desplazamientos en las dos direcciones.

Durante esta exposición, utilizaremos indistintamente tanto imágenes de flechas para mostrar los desplazamientos como imágenes en niveles de grises, en donde el nivel de gris indica el desplazamiento en una dirección u otra – si el nivel de gris es oscuro, significa que el desplazamiento es negativo, mientras que si es claro, el desplazamiento es positivo. En estas últimas, utilizaremos dos imágenes, una para el desplazamiento en  $x$  (componente  $u$ ) y otra para el desplazamiento en  $y$  (componente  $v$ ).

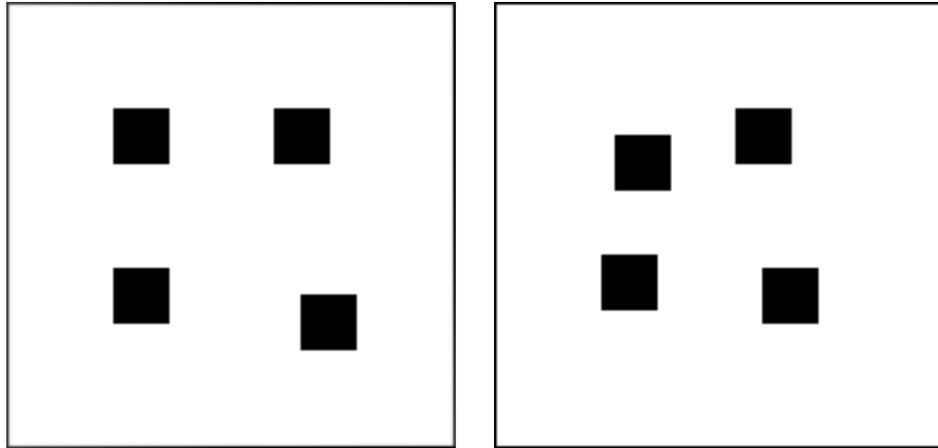


Figura 2.1: **Secuencia de los cuadrados:** En estas dos imágenes hay cuatro cuadrados negros sobre fondo blanco. Los cuadrados se mueven de la imagen izquierda a la derecha de la siguiente manera (de izquierda a derecha y de arriba a abajo): a)  $(10, 5)$ ; b)  $(-5, 0)$ ; c)  $(0, -5)$ ; y d)  $(-10, -10)$ .

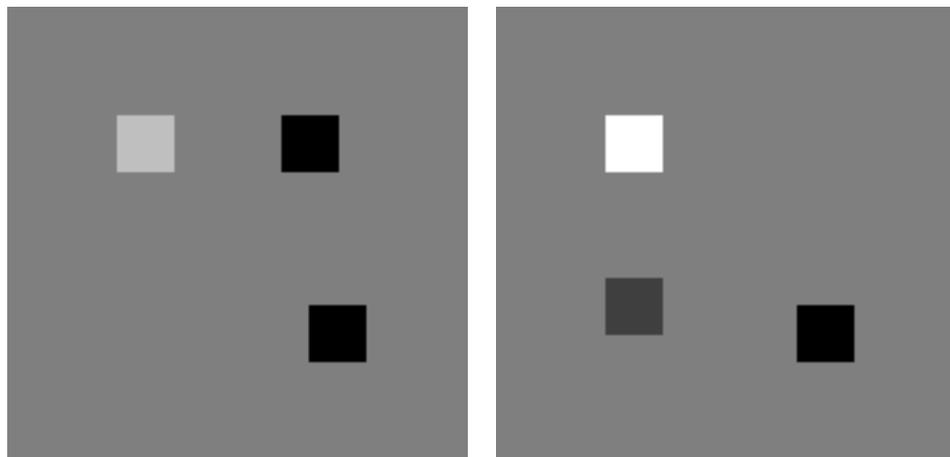


Figura 2.2: **Desplazamiento de izquierda a derecha:** Estas dos imágenes representan el desplazamiento que existe desde la imagen izquierda a la derecha de la secuencia 2.1. Los niveles de grises claros significan desplazamientos positivos y los oscuros, negativos. La imagen de la izquierda representa el desplazamiento en el eje de las  $x$  y la imagen de la derecha en el de las  $y$ .

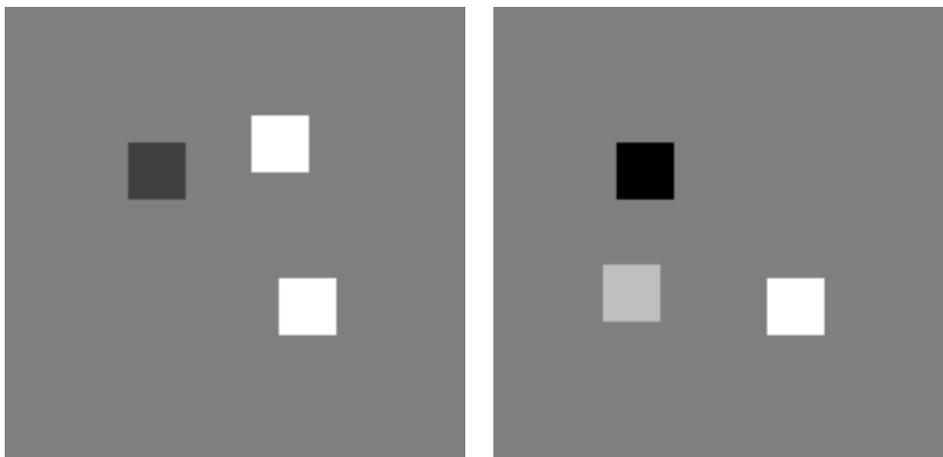


Figura 2.3: **Desplazamiento de derecha a izquierda:** Estas dos imágenes representan el desplazamiento que existe desde la imagen derecha a la izquierda de la secuencia 2.1. La imagen de la izquierda representa el desplazamiento en el eje de las  $x$  y la imagen de la derecha en el de las  $y$ .

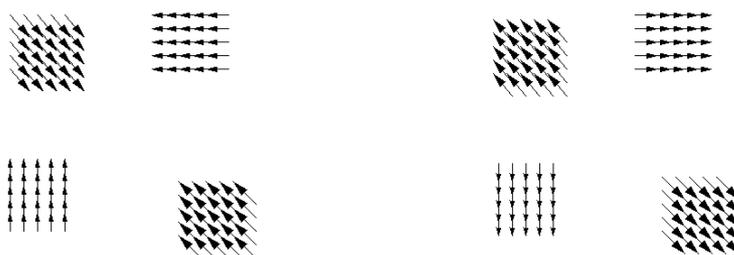


Figura 2.4: **Campo de vectores:** Estas dos imágenes representan los campos de vectores que se producen para la secuencia 2.1. La imagen de la izquierda representa el desplazamiento de izquierda a derecha y la de la derecha, en el sentido contrario.

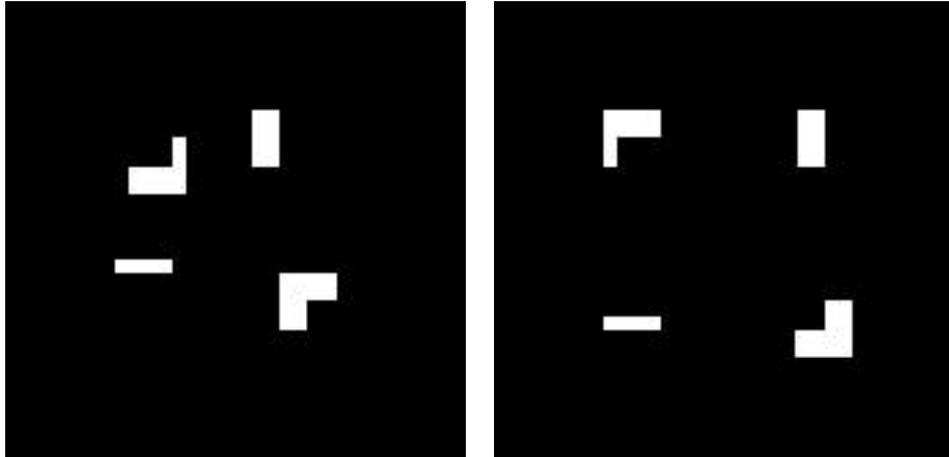


Figura 2.5: **Mapas de oclusiones para la secuencia de los cuadrados:** Estas dos imágenes representan las oclusiones que se producen en la secuencia 2.1. La imagen de la izquierda representa las oclusiones que se producen al calcular el flujo óptico de izquierda a derecha y la de la derecha las oclusiones en el otro sentido.

En la figura 2.4 mostramos los campos de vectores exactos para las dos direcciones. Esta será la otra manera de visualizar los resultados del flujo óptico en las pruebas.

Por último, en la figura 2.5 se muestran las oclusiones para la secuencia 2.1 en las dos direcciones.

### Secuencia Yosemite

En la figura 2.6 se muestra la secuencia Yosemite que representa una foto del parque nacional Yosemite en Sierra Nevada, California . Esta es una secuencia sintética compleja, creada por Lynn Quam y utilizada por Barron et al. [BFB94] para la comparación entre métodos. Los desplazamientos son bastante variados: la parte de las nubes se mueve a 2 pixels/frame, mientras que la parte inferior izquierda se mueve alrededor de 5 pixels/frame en distinta dirección.

En las imágenes 2.7 y 2.8 se muestran los campos de desplazamientos exactos para los frames 9 y 10, en mapas de niveles de grises y en campos de vectores, respectivamente.

En la figura 2.8 se puede apreciar que el desplazamiento es variado: En las nubes el desplazamiento es hacia la derecha con valor de 2 pixels/frame, en la zona de la montaña varía en sentido y módulo, llegando hasta cerca de 5 pixels/frame en la esquina inferior izquierda.

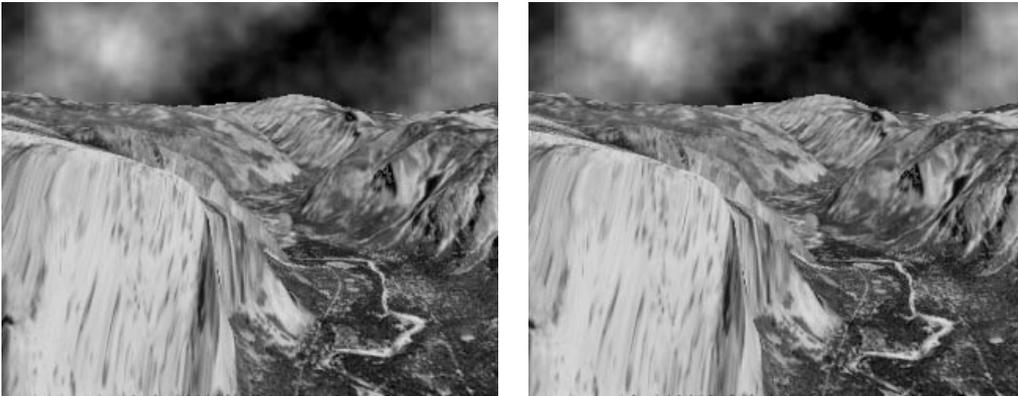


Figura 2.6: **Secuencia de Yosemite:** Frames 9 y 10.

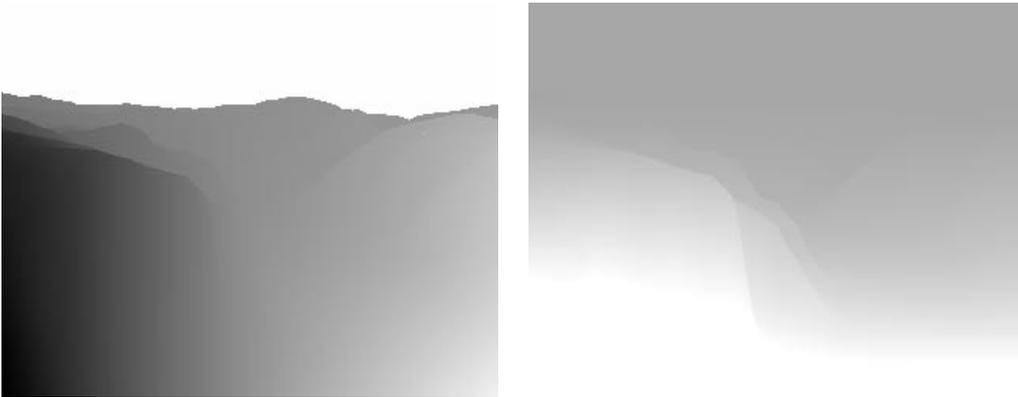


Figura 2.7: **Desplazamiento de izquierda a derecha:** En estas dos imágenes se muestra a la izquierda (resp. derecha) el desplazamiento en  $x$  (resp.  $y$ ) para la secuencia de la figura 2.6.

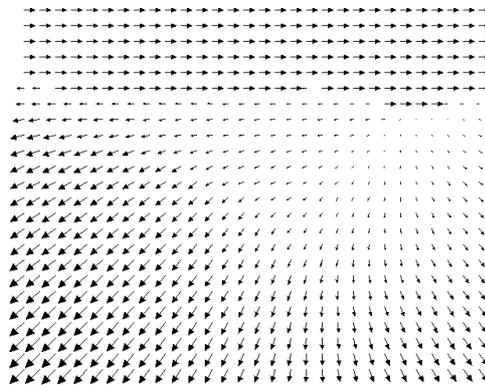


Figura 2.8: **Campos de vectores:** Imagen de flechas para la secuencia 2.6.

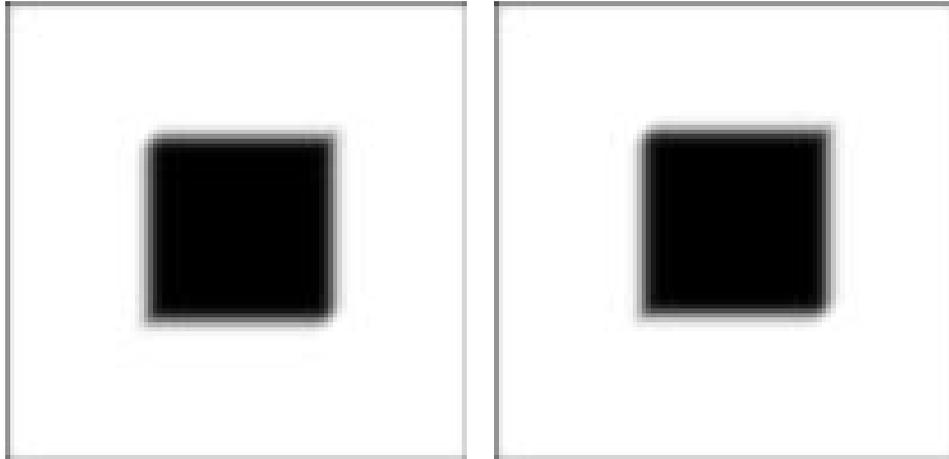


Figura 2.9: **Secuencia del cuadrado:** Frames 11 y 12.

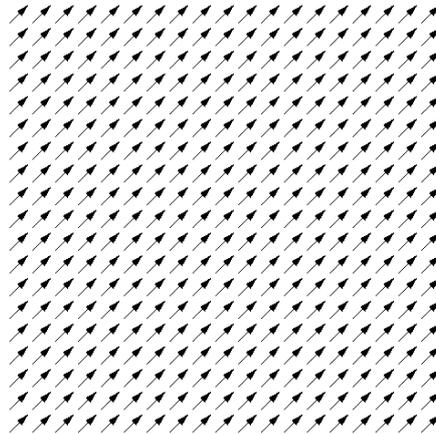


Figura 2.10: **Campo de vectores:** Imagen de flechas para la secuencia del cuadrado 2.9. El desplazamiento es constante de  $(\frac{4}{3}, -\frac{4}{3})$  pixels/frame.

### Secuencia del cuadrado

Esta es una secuencia sencilla (ver figura 2.9) de un cuadrado que se desplaza de forma constante en  $(x, y) = (\frac{4}{3}, -\frac{4}{3})$ . Esta secuencia también fue utilizada por Barron et al. Este desplazamiento está definido para toda la imagen (incluido el fondo). Sin embargo, es difícil estimar el desplazamiento en el fondo, ya que es bastante homogéneo y abarca bastante espacio. En las pruebas que hemos realizado, veremos que la estimación dentro del cuadrado es bastante exacta, pero en el fondo no se obtiene este valor.

En la imagen 2.10 vemos el campo de vectores de desplazamiento para esta secuencia. Los vectores son todos paralelos y de igual magnitud<sup>1</sup>.

<sup>1</sup>No mostramos el desplazamiento para la secuencia del cuadrado con mapas de niveles de gris,

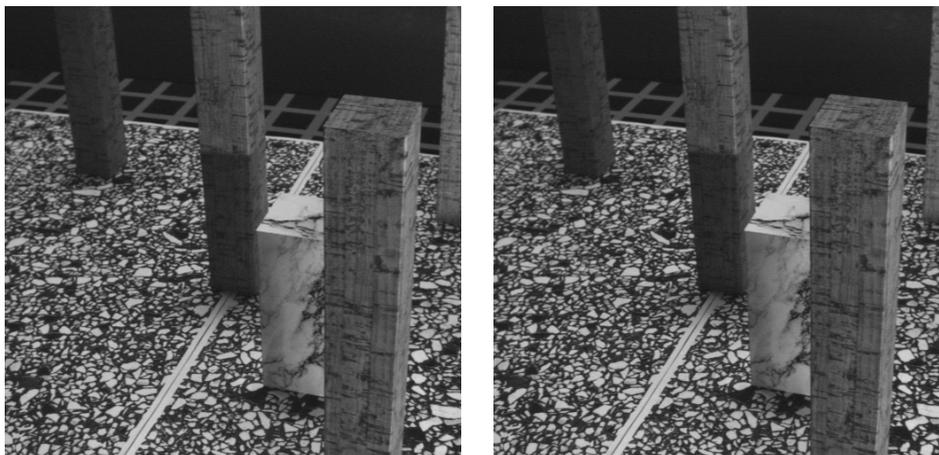


Figura 2.11: **Secuencia de las Torres de Mármol:** Frames 36 y 35.

### Secuencia de las Torres de Mármol

En esta secuencia<sup>2</sup> (figura 2.11) se puede ver una serie de torres y un bloque de mármol que aparece de detrás de una de éstas. La cámara se va desplazando hacia la izquierda al mismo tiempo que el bloque de mármol se desplaza también en esa dirección. El suelo tiene una textura de manchas blancas sobre fondo negro.

Esta secuencia es copyright de H.H.Nagel (KOGS/IAKS, Universidad de Karlsruhe, Alemania). La primera vez que se utilizó esta secuencia fue en el trabajo de Michael Otte y Nagel [ON95] sobre flujo óptico, y está disponible a través del sitio web <http://i21www.ira.uka.de/> en el directorio *image-sequences*.

Si observamos las imágenes 2.12 y 2.13 vemos que las torres y el suelo se desplazan en sentido negativo tanto en  $x$  como en  $y$ . La magnitud del movimiento es mayor en la torre y la parte del suelo que están más pegadas a la parte inferior derecha. El bloque de mármol del centro se mueve en menor magnitud debido a que sigue un poco al movimiento de la cámara.

### Secuencia del cubo

Esta es la secuencia sintética SOFA, cortesía del Grupo de Visión por Ordenador de la Universidad de Heriot-Watt University <http://www.cee.hw.ac.uk/mtc/sofa>. En esta aparecen un cubo y un cilindro que tienen un movimiento de rotación puro de una imagen a la otra (ver figura 2.14). En esta figura se muestran los frames 3 y 5.

---

ya que al ser constante, lo único que veríamos serían dos imágenes en negro

<sup>2</sup>Debemos notar que el flujo óptico está calculado desde el frame 36 hacia el frame 35. Esto se debe a que Nagel propuso un método que estaba centrado, en su término de ligadura, en la imagen  $I_2$ , mientras que nosotros lo centramos en la imagen  $I_1$  por motivos de congruencia en los dos términos, tal y como comentamos en el capítulo 6.

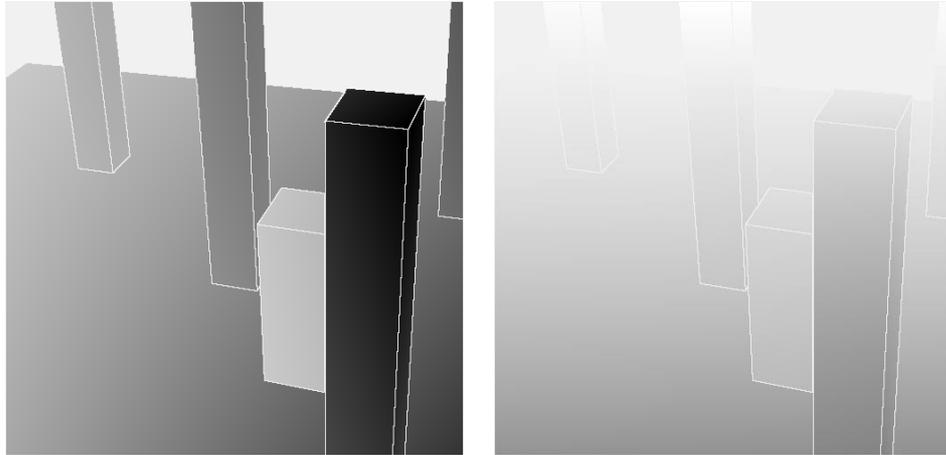


Figura 2.12: **Desplazamiento de derecha a izquierda:** En esta imagen se muestra con mapas de niveles de grises a la izquierda (resp. derecha) el desplazamiento en  $x$  (resp.  $y$ ) para la secuencia de las Torres 2.11.



Figura 2.13: **Campo de vectores:** Imagen de flechas para la secuencia 2.11.

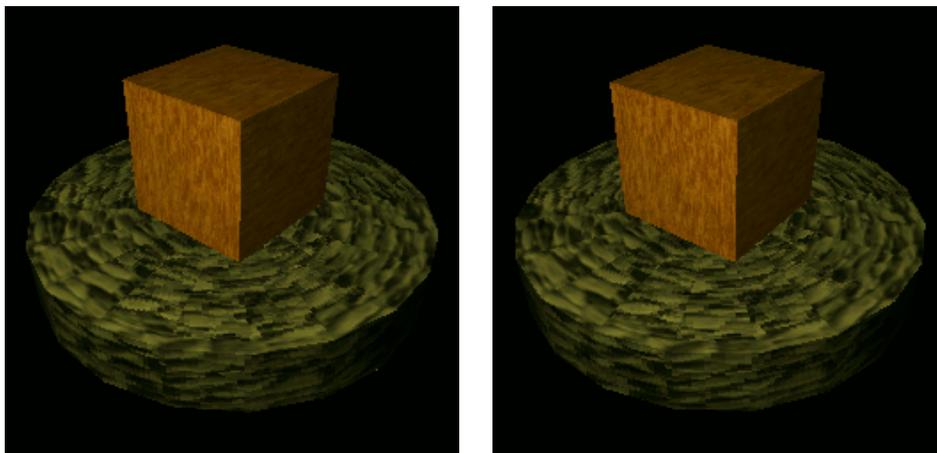


Figura 2.14: **Secuencia del Cubo**

El cubo tiene un color anaranjado con textura en forma de madera, y el cilindro es de color verde, con una textura uniforme.

Esta secuencia la utilizaremos en las pruebas del flujo óptico para imágenes en color.

### 2.5.2 Imágenes reales

Aquí comentamos una sola secuencia, la secuencia del taxi de Hamburgo. Solo mostramos las imágenes de la secuencia ya que el desplazamiento perfecto no lo podemos estimar.

#### Secuencia del Taxi de Hamburgo

Esta es una de las secuencias tradicionales que se suelen utilizar en los trabajos sobre el cálculo del flujo óptico. En ésta se pueden ver un taxi que se mueve en diagonal para cambiar de calle, un coche en la esquina inferior izquierda que se mueve a mayor velocidad hacia la derecha, un camión en la esquina inferior derecha que se desplaza hacia la izquierda y, finalmente, un peatón en la esquina superior izquierda que se desplaza a mucha menor velocidad (casi imperceptible) por la acera. El resto de los objetos no se mueven.

Esta secuencia se puede obtener en el sitio <ftp://csd.uwo.ca> en el directorio `pub/vision`



Figura 2.15: **Secuencia del taxi de Hamburgo:** Frames 15 y 19.

# Capítulo 3

## Geometría Epipolar

### 3.1 Introducción

En este capítulo estudiaremos la geometría que se genera al considerar un par de imágenes estéreo, en el que las imágenes se toman con cámaras proyectivas.

Suponemos que las cámaras son de tipo proyectivo (ver figura 3.1), es decir, los puntos de la escena se proyectan a través del foco de la cámara en el plano proyectivo o imagen. En los capítulos 2 y 3 del libro de Faugeras [Fau93] se introduce de manera formal la geometría proyectiva y se modela la cámara proyectiva. Una cámara se puede modelar como una matriz de proyección  $\tilde{\mathbf{P}}$ , que transforma un punto de dimensión 3 en unas coordenadas de imagen, de dimensión 2. Esta matriz se caracteriza por incluir los *parámetros intrínsecos* y *extrínsecos* de la cámara.

Uno de los pasos previos en el problema de la visión estéreo es conocer la relación que existe entre las cámaras. Para cada punto  $m$  en una de las imágenes existe un correspondiente,  $m'$ , en la otra que se encuentra sobre una línea, denominada línea epipolar. Esto es lo que se denomina la *restricción epipolar*. En [LF96] definen esta restricción a través de una matriz, denominada *matriz fundamental*, mostrando la relación que existe entre las matrices de proyección de las cámaras y la matriz fundamental.

En la matriz fundamental se encuentran representados los parámetros intrínsecos – origen de la imagen, distancia focal, tamaño del pixel y desviación de los ejes – y los parámetros extrínsecos – matriz de rotación y vector de desplazamiento – de las cámaras. Al proceso de obtener esta matriz a través de puntos en correspondencia se denomina calibración. Decimos que un sistema está *calibrado débilmente* cuando conocemos su geometría epipolar o matriz fundamental.

Cuando consideremos el problema del cálculo de la carta de disparidad suponemos que las cámaras están calibradas y que se conoce la geometría epipolar. Aquí no vamos a tratar el problema de la calibración de cámaras. Se han realizado muchos

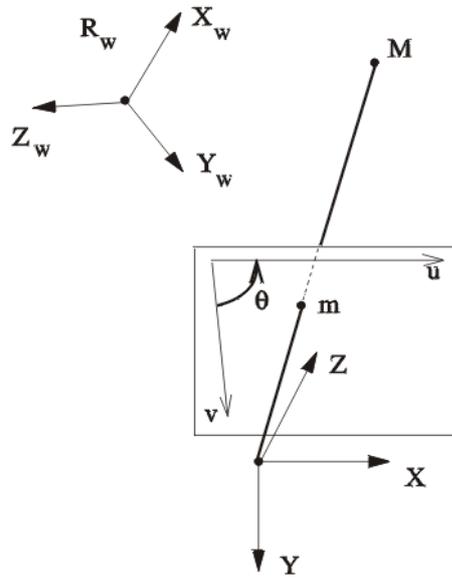


Figura 3.1: El modelo proyectivo y diferentes frames

trabajos sobre el problema de la calibración tales como [Fau93, Har92, Har94a, LH81, LDFP93, LF96, Zha97].

En las siguientes secciones estudiaremos la geometría epipolar, el problema de la rectificación de imágenes y el problema de la reconstrucción tridimensional. Para el primero, haremos un pequeño repaso por algunos aspectos de la geometría para luego exponer una parametrización especial que nos va a permitir determinar la disparidad utilizando sólo una función escalar que representa el desplazamiento sobre la línea epipolar. En la siguiente sección estudiaremos el problema de la rectificación de imágenes con el fin de obtener una geometría más sencilla y que utilizaremos en algún método de los que expondremos en los próximos capítulos. Por último trataremos el problema de la reconstrucción a partir de las matrices de proyección.

## 3.2 Geometría de dos cámaras proyectivas

Se supone que el modelo de la cámara es proyectivo (figura 3.1). Este modelo proyecta un punto 3D  $\mathbf{M} = [X, Y, Z]^t$  a un punto 2D  $\mathbf{m} = [x, y]^t$  a través de una matriz de proyección  $3 \times 4$   $\tilde{\mathbf{P}}$  via  $s\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\mathbf{M}$ , donde  $s$  es un factor de escala distinto de 0 y la notación de  $\tilde{\mathbf{p}}$  es tal que si  $\mathbf{p} = [x, y, \dots]^t$  entonces  $\tilde{\mathbf{p}} = [x, y, \dots, 1]^t$ . En el caso de dos imágenes obtenidas por un sistema estéreo binocular, cada punto físico  $\mathbf{M}$  en el espacio provoca un par de proyecciones 2D  $\mathbf{m}$  y  $\mathbf{m}'$  en las dos imágenes. Las matrices de proyección  $3 \times 4$  se definen por las siguientes relaciones:

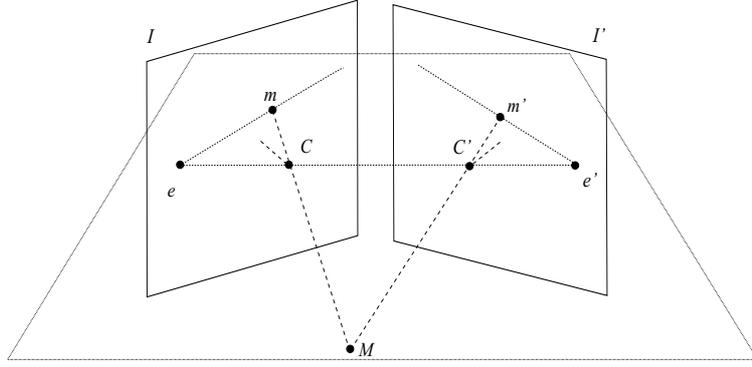


Figura 3.2: **Geometría epipolar:**  $I$  e  $I'$  representan las imágenes.  $C$  y  $C'$  los focos de las cámaras.  $e$  y  $e'$  los epipolos.  $M$  el punto 3D que se proyecta sobre  $m$  y  $m'$  y están situados sobre un mismo plano.

$$s\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} \quad , \quad s'\tilde{\mathbf{m}}' = \tilde{\mathbf{P}}'\tilde{\mathbf{M}}. \quad (3.2.1)$$

Suponiendo que el sistema de coordenadas global está asociado con la primera cámara, las dos matrices de proyección vienen dadas por:

$$\tilde{\mathbf{P}} = [\mathbf{A}|\mathbf{0}] \quad , \quad \tilde{\mathbf{P}}' = [\mathbf{A}'\mathbf{R}|\mathbf{A}'\mathbf{t}], \quad (3.2.2)$$

donde  $\mathbf{R}$  y  $\mathbf{t}$  representan respectivamente la matriz de rotación  $3 \times 3$  y el vector de traslación  $3 \times 1$  que definen el desplazamiento rígido entre las dos cámaras, y  $\mathbf{0}$  denota el vector nulo  $3 \times 1$ . Las matrices  $\mathbf{A}$  y  $\mathbf{A}'$  son las matrices de los parámetros intrínsecos  $3 \times 3$  de las dos vistas, cada una depende de cinco parámetros de la siguiente manera [Fau93]:

$$\mathbf{A} = \begin{bmatrix} \alpha_u & -\alpha_u \cot \theta & u_0 \\ 0 & \alpha_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \mathbf{A}' = \begin{bmatrix} \alpha'_u & -\alpha'_u \cot \theta' & u'_0 \\ 0 & \alpha'_v / \sin \theta' & v'_0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.2.3)$$

Los parámetros  $\alpha_u$  (resp.  $\alpha'_u$ ) y  $\alpha_v$  (resp.  $\alpha'_v$ ) corresponden a las distancias focales en pixels a lo largo de los ejes de la imagen,  $\theta$  (resp.  $\theta'$ ) es el ángulo entre los dos ejes de coordenadas y  $(u_0, v_0)$  (resp.  $(u'_0, v'_0)$ ) son las coordenadas del punto principal de la imagen. El ratio  $\frac{\alpha_v}{\alpha_u}$  (resp.  $\frac{\alpha'_v}{\alpha'_u}$ ) se conoce como el *ratio de aspecto*.

Cuando las dos imágenes se han obtenido con la misma cámara,  $\theta = \theta'$  y  $\frac{\alpha_v}{\alpha_u} = \frac{\alpha'_v}{\alpha'_u}$ ,

incluso aunque las longitudes focales hayan cambiado. En la realidad,  $\theta$  y  $\theta'$  son próximos a  $\frac{\pi}{2}$  para cámaras reales [Fau93]. Adicionalmente, las cámaras modernas tienen pixels casi cuadrados, por lo tanto en este caso, el ratio de aspecto está muy próximo a uno.

Todas estas matrices y parámetros que estas incluyen se pueden calcular con buena precisión a través de un método de calibración clásico [Fau93]. En este caso, el sistema se dice que está *calibrado*.

Eliminando los escalares  $s$  y  $s'$  asociados a las ecuaciones de proyección (3.2.1) así como el punto  $\mathbf{M}$ , se obtiene una ecuación que relaciona el par de proyecciones del mismo punto 3D:

$$\tilde{\mathbf{m}}'^t \mathbf{F} \tilde{\mathbf{m}} = 0.$$

En esta ecuación, la matriz  $\mathbf{F}$  es la matriz fundamental, dada por

$$\mathbf{F} = \mathbf{A}'^* [\mathbf{t}]_{\times} \mathbf{R} \mathbf{A}^{-1}, \quad (3.2.4)$$

donde  $\mathbf{A}'^* = (\mathbf{A}'^{-1})^t$  es la matriz adjunta de  $\mathbf{A}'$ , y  $[\mathbf{x}]_{\times}$  denota la matriz anti-simétrica del vector  $\mathbf{x}$  que está asociado con el producto vectorial. Esta matriz tiene la propiedad  $[\mathbf{x}]_{\times} \mathbf{y} = \mathbf{x} \times \mathbf{y}$  para cada vector  $\mathbf{y}$  y tiene la siguiente forma analítica:

$$[\mathbf{x}]_{\times} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}.$$

Para un punto  $\mathbf{m} = [x, y]^t$  en la primera imagen  $I$ , la matriz fundamental  $\mathbf{F} = (f_{i,j})$  provee la línea epipolar  $\Delta$  de la ecuación  $\tilde{\mathbf{m}}'^t \mathbf{F} \tilde{\mathbf{m}} = 0$  en la imagen  $I'$ . Introduzcamos la notación

$$\begin{aligned} a(x, y) &= f_{11}x + f_{12}y + f_{13}, \\ b(x, y) &= f_{21}x + f_{22}y + f_{23}, \\ c(x, y) &= f_{31}x + f_{32}y + f_{33}. \end{aligned}$$

Utilizando esta notación, la línea epipolar  $\Delta$  se puede escribir como

$$a(x, y)x' + b(x, y)y' + c(x, y) = 0. \quad (3.2.5)$$

Usaremos esta ecuación con el objetivo de introducir una parametrización específica de la disparidad, desarrollada para producir un operador diferencial lineal de segundo orden en la parte de minimización.

### El término de disparidad

Bajo la suposición *Lambertiana* de que pixels correspondientes tienen igual nivel de gris, determinar la disparidad entre un par estéreo se reduce a encontrar una función  $\mathbf{h}(x, y) := (u(x, y), v(x, y))^t$  tal que

$$I(x, y) = I'(x + u(x, y), y + v(x, y)), \quad \forall (x, y) \in \mathbb{R}^2. \quad (3.2.6)$$

donde el punto  $(x', y') = (x + u(x, y), y + v(x, y))$  pertenece a la línea epipolar asociada a  $(x, y)$ .

Sea  $\mathbf{N}$  ( resp.  $\mathbf{T}$  ) el vector normal (resp. tangencial) unitario de la línea epipolar  $\Delta$  dada por la ecuación  $\tilde{\mathbf{m}}'^t \mathbf{F} \tilde{\mathbf{m}} = 0$ , y  $\mathbf{D}$  el vector de disparidad unitario asociado a los puntos  $\mathbf{m}$  y  $\mathbf{m}'$ :

$$\mathbf{N} = \begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} \end{bmatrix} ; \quad \mathbf{T} = \begin{bmatrix} \frac{-b}{\sqrt{a^2+b^2}} \\ \frac{a}{\sqrt{a^2+b^2}} \end{bmatrix} ; \quad \mathbf{D} = \begin{bmatrix} \frac{u}{\sqrt{u^2+v^2}} \\ \frac{v}{\sqrt{u^2+v^2}} \end{bmatrix} \quad (3.2.7)$$

Entonces tenemos que

$$\mathbf{m}' = \mathbf{m} + \delta \mathbf{D} = \mathbf{m} - \gamma \mathbf{N} - \lambda \mathbf{T}$$

donde  $\delta = \sqrt{u^2 + v^2}$  representa la disparidad,

$$\gamma = \frac{\tilde{\mathbf{m}}'^t \mathbf{F} \tilde{\mathbf{m}}}{\sqrt{a^2 + b^2}}$$

representa la distancia (módulo el signo) del punto  $\mathbf{m}$  a su línea epipolar  $\Delta$  en la segunda imagen y  $\lambda$  representa la distancia (módulo el signo) de  $\mathbf{m}_0$ , la proyección del punto  $\mathbf{m}$  en la línea epipolar  $\Delta$ , al punto  $\mathbf{m}'$  que está incluido en la línea epipolar  $\Delta$ .

Utilizando estas ecuaciones se puede derivar fácilmente las siguientes expresiones para los componentes de la disparidad:

$$u(x, y) = \frac{-\lambda(x, y)b(x, y)}{\sqrt{a^2(x, y) + b^2(x, y)}} - \frac{a(x, y)x + b(x, y)y + c(x, y)}{a^2(x, y) + b^2(x, y)} a(x, y), \quad (3.2.8)$$

$$v(x, y) = \frac{\lambda(x, y)a(x, y)}{\sqrt{a^2(x, y) + b^2(x, y)}} - \frac{a(x, y)x + b(x, y)y + c(x, y)}{a^2(x, y) + b^2(x, y)} b(x, y), \quad (3.2.9)$$

donde  $\lambda(x, y)$  representa un parámetro asociado a la profundidad del punto físico considerado. En lo que sigue, utilizaremos la notación  $\mathbf{h}(\lambda(x, y)) = (u(\lambda(x, y), v(\lambda(x, y))))$  para indicar que la disparidad depende de un sólo parámetro  $\lambda(x, y)$ . Nótese que  $\lambda(x, y) = 0$  nos da el punto en la línea epipolar que minimiza la distancia entre la línea epipolar y el punto  $(x, y)$ . Utilizando la restricción epipolar y sabiendo que  $\gamma(x, y)$  se conoce para cada punto  $(x, y)$  gracias al conocimiento de la matriz fundamental, el cálculo del mapa de disparidad  $\delta(x, y)$  se reduce al cálculo de la función  $\lambda(x, y)$ :

$$\delta(x, y)^2 = \lambda(x, y)^2 + \gamma(x, y)^2.$$

En los capítulos 7 y 10 desarrollaremos aproximaciones para estimar con precisión la función  $\lambda(x, y)$  asociada al par estéreo de imágenes. Tal como se ha dicho, esto nos permitirá estimar fácilmente el mapa de disparidad para cada punto de las imágenes y, por lo tanto, recuperar la información de la profundidad en el caso de tratar con sistemas calibrados.

### 3.3 Rectificación de imágenes

El objetivo de la rectificación de imágenes es el de simplificar la geometría epipolar. La geometría más sencilla es aquella en la que las imágenes son fronto-paralelas, es decir, se hayan en el mismo plano y a la misma altura. Esto es equivalente a decir que las líneas epipolares son paralelas y horizontales y que los epipolos se encuentran en el infinito.

Una forma de realizar la rectificación es la de re proyectar los puntos de la imágenes en un plano paralelo situado como se puede ver en la figura 3.3.

Buscar correspondencias entre puntos en cámaras con esta configuración se reduce a desplazarse sólo por la misma ordenada ( $y = y'$ ). La disparidad será, por tanto, el desplazamiento en una sola dirección ( $x = x' + \lambda$ ). En este caso, la disparidad coincide con el desplazamiento dentro de la línea epipolar,  $\delta(x, y)^2 = \lambda(x, y)^2$ . Las ecuaciones (3.2.8) y (3.2.9) nos quedarían de la siguiente manera:

$$u(x, y) = -\lambda(x, y), \tag{3.3.1}$$

$$v(x, y) = 0, \tag{3.3.2}$$

En el capítulo 10, donde se trata la incorporación de las simetrías en el problema de la visión estéreo, propondremos un método especial para este tipo de configuraciones. Suponemos, entonces, que se realiza un proceso de rectificación previo.

Algunos trabajos sobre rectificación de imágenes son [AH88], [Har98], [LZ98], [PD96], [RMC97].

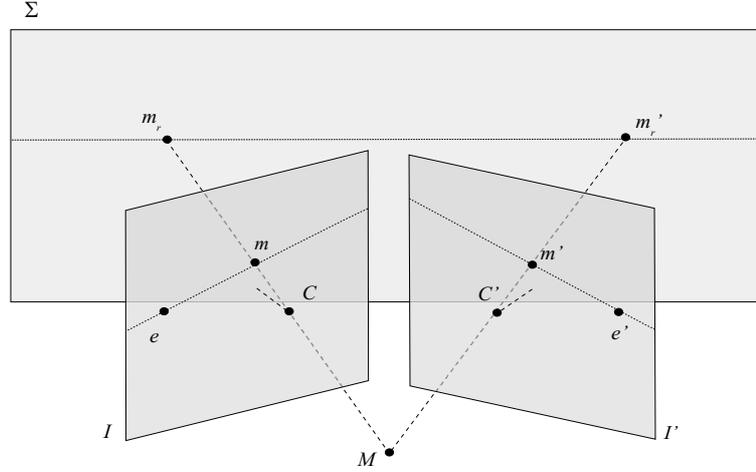


Figura 3.3: **Geometría epipolar rectificada:** Se re proyecta sobre el plano  $\Sigma$  los puntos de las imágenes.

### Rectificación a partir de las matrices de proyección

Se puede rectificar un par estéreo tanto a partir de la información de la geometría epipolar como de las matrices de proyección. Ahora vamos a ver cómo rectificar un par estéreo a partir de sus matrices de proyección.

Partimos de dos matrices de proyección,  $\tilde{\mathbf{P}}$  y  $\tilde{\mathbf{P}}'$ , de  $\mathbb{R}^3$  en  $\mathbb{R}^2$ . Siguiendo la notación y el razonamiento expuesto en [Fau93] vamos a re proyectar los puntos de los planos proyectivos en un plano común de la forma vista en la figura 3.3.

Para esto, lo primero que vamos a hacer es encontrar el plano definido por  $\Sigma$  tal y como se ve en esta figura. Los focos de las cámaras se obtienen como

$$\begin{aligned} \mathbf{C} &= -\mathbf{P}^{-1} p \\ \mathbf{C}' &= -\mathbf{P}'^{-1} p' \end{aligned} \quad (3.3.3)$$

donde  $\mathbf{P}^{-1}$  (resp.  $\mathbf{P}'^{-1}$ ) es la submatriz de  $3 \times 3$  de  $\tilde{\mathbf{P}}$  (resp.  $\tilde{\mathbf{P}}'$ ), y  $p$  (resp.  $p'$ ) es la última columna de  $\tilde{\mathbf{P}}$  (resp.  $\tilde{\mathbf{P}}'$ ). A partir de las matrices de proyección es posible encontrar los parámetros intrínsecos de las cámaras (ver [Fau93]) y, por lo tanto, podemos encontrar el origen de los planos proyectivos, o lo que es lo mismo, la proyección del foco en la imagen. Para la primera matriz,  $\tilde{\mathbf{P}}$ , el centro de la imagen viene dado por  $(u_0, v_0)$  y para  $\tilde{\mathbf{P}}'$  es  $(u'_0, v'_0)$ .

Para que el plano  $\Sigma$  tenga los epipolos en el infinito vamos a imponer que éste sea paralelo a la línea que une  $\mathbf{C}$  y  $\mathbf{C}'$  definida por el vector unitario

$$\vec{e}_1 = \frac{\overrightarrow{\mathbf{C}\mathbf{C}'}}{\|\overrightarrow{\mathbf{C}\mathbf{C}'}\|}$$

Existen infinitos planos que sean paralelos a esta línea. Necesitamos, por tanto, imponer alguna otra condición para definir la orientación del plano. Además, también tenemos que determinar en qué posición se va a encontrar este plano con respecto a los centros de proyección de las cámaras.

Para determinar la orientación, vamos a imponer que el vector normal al plano sea una media de los vectores perpendiculares a los planos de proyección. Estos vectores son:

$$\vec{n} = -\mathbf{P}^{-1} \begin{pmatrix} u_0 \\ v_0 \\ 1 \end{pmatrix}, \quad \vec{n}' = -\mathbf{P}'^{-1} \begin{pmatrix} u'_0 \\ v'_0 \\ 1 \end{pmatrix}$$

El vector unitario normal al plano  $\Sigma$  lo vamos a calcular como la media de los dos anteriores, normalizado:

$$\vec{n}_\Sigma = \frac{\frac{\vec{n} + \vec{n}'}{2}}{\left\| \frac{\vec{n} + \vec{n}'}{2} \right\|}$$

Los vectores  $\vec{e}_1$  y  $\vec{n}_\Sigma$  deben ser ortogonales, con lo que es necesario realizar una rectificación de la manera  $\vec{n}_\Sigma = \vec{n}_\Sigma - \vec{n}_\Sigma \vec{e}_1$ .

Calculamos el otro vector director del plano  $\Sigma$  a través del producto vectorial de los vectores  $\vec{e}_1$  y  $\vec{n}_\Sigma$ ,  $\vec{e}_2 = \vec{e}_1 \times \vec{n}_\Sigma$ . Entonces los vectores  $e_1$  y  $e_2$  son los vectores directores del nuevo plano.

Calculamos un punto del plano desplazándonos una distancia de 1 desde el foco de la primera cámara  $C$  en la dirección dada por el vector perpendicular a  $\Sigma$ ,  $\vec{n}_\Sigma$ :  $\mathbf{p}_\Sigma = \mathbf{C} + \vec{n}_\Sigma$ .

Una vez hemos definido el plano,  $\vec{n}_\Sigma^t (\vec{x} - \mathbf{p}_\Sigma) = 0$ , lo que vamos a hacer es re-proyectar un conjunto de puntos ( $\geq 4$ ) sobre  $\Sigma$  y calcular la homografía que nos lleva de unos a otros.

Cada punto  $(x_i, y_i, 1)^t$  correspondiente al primer plano, define un vector  $\vec{n}_i$  que tiene la dirección de la proyección a través de  $C$  de la forma  $\vec{n}_i = \mathbf{P}^{-1} \cdot (x_i, y_i, 1)^t$ . La línea que pasa por  $C$  y tiene este vector director es  $\mathbf{C} + \lambda_i \vec{n}_i$ .

La intersección de esta línea con el plano viene dado por

$$\vec{n}_\Sigma^t (\mathbf{C} + \lambda_i \vec{n}_i - \mathbf{p}_\Sigma) = 0$$

que despejando obtenemos

$$\lambda_i = \frac{\vec{n}_\Sigma^t \mathbf{C} - \vec{n}_\Sigma^t \mathbf{p}_\Sigma}{\vec{n}_\Sigma^t, \vec{n}_i}$$

Este valor de  $\lambda_i$  nos da la posición dentro de la línea que intersecciona con el plano, pero lo que necesitamos es la posición relativa dentro del plano. Para esto, expresamos el plano en términos de sus dos vectores directores y lo igualamos al punto dado en la intersección,

$$\mathbf{C} + \lambda_i \vec{n}_i = \mathbf{p}_\Sigma + \eta_{i1} \vec{e}_1 + \eta_{i2} \vec{e}_2$$

De aquí deducimos directamente que las coordenadas de los puntos en el nuevo plano vienen dadas por

$$\begin{aligned} \eta_{i1} &= (\mathbf{C} + \lambda_i \vec{n}_i - \mathbf{p}_\Sigma) \vec{e}_1 \\ \eta_{i2} &= (\mathbf{C} + \lambda_i \vec{n}_i - \mathbf{p}_\Sigma) \vec{e}_2 \end{aligned} \quad (3.3.4)$$

Para el segundo plano tenemos que elegir otro conjunto de puntos y repetir todo este proceso, teniendo en cuenta que el foco ahora es  $\mathbf{C}'$  y la matriz de proyección es  $\tilde{\mathbf{P}}'$ .

Una vez obtenidas las nuevas posiciones de los puntos, calculamos las dos homografías – una por cada plano – que nos lleva de cada plano al plano  $\Sigma$ . Aplicamos estas homografías a todos los puntos de los planos para obtener las nuevas posiciones.

De cara a la implementación, no utilizamos las homografías directas, sino que calculamos las inversas, y para cada punto en el plano  $\Sigma$  buscamos su correspondiente en los planos originales, realizando una interpolación de las intensidades de los cuatro vecinos. Además, hay que tener en cuenta que los tamaños de las nuevas imágenes no coincidirán con los antiguos, siendo necesario calcular de antemano el tamaño final.

La forma de rectificación que se propone en [Fau93] es ligeramente distinta a la expuesta aquí. En aquella, lo que se propone es crear un plano  $\Sigma$  a partir de la intersección de los dos planos proyectivos. Esto significa que uno de los vectores directores del plano estará definido por  $\vec{e}_2 = \vec{n} \times \vec{n}' - (\vec{n} \times \vec{n}')^t \vec{e}_1$ . El inconveniente que tiene esto es que si las imágenes están situadas en posiciones próximas a paralelas, entonces el producto vectorial da valores próximos a cero, con lo que el vector director se puede ver afectado negativamente por pequeños errores de redondeo.

## 3.4 Reconstrucción tridimensional

El proceso final dentro de la visión estereó es el de la reconstrucción tridimensional. Como veremos en el siguiente capítulo, una vez calculemos las correspondencias que existen entre los puntos de ambas imágenes, pasaremos a calcular los puntos 3D de la escena.

Existen varias formas de reconstruir una escena a partir de sus proyecciones [Fau92]. Dependiendo del conocimiento que se tenga acerca de la geometría de las cámaras, es posible realizar varios tipos de reconstrucciones. Por ejemplo, si las

cámaras están calibradas de forma débil, entonces se puede realizar una reconstrucción proyectiva, [Har94b]. Si se conoce la información del plano en el infinito, es posible realizar la reconstrucción afín y el caso en el que se conozcan las matrices de proyección de ambas cámaras, entonces podemos realizar una reconstrucción euclídea. Algunos trabajos sobre el problema de la reconstrucción son [MGDP93], [TK92].

En nuestro caso, vamos a trabajar con cámaras para las que se conocen sus matrices de proyección. Vamos a ver entonces cómo a partir de éstas podemos realizar la reconstrucción euclídea.

El proceso consiste básicamente en encontrar la distancia mínima que existe entre las líneas que unen los focos de las cámaras con las proyecciones asociadas a un punto 3D. Si tenemos dos puntos,  $\tilde{\mathbf{m}} = (x, y, 1)^t$  y  $\tilde{\mathbf{m}}' = (x', y', 1)^t$  que son la proyección de un punto  $M$  de la escena, los vectores que marcan la dirección entre estos puntos y su reconstrucción vienen dados por

$$\vec{n} = \mathbf{P}^{-1} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad \vec{n}' = \mathbf{P}'^{-1} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \quad (3.4.1)$$

Estos dos vectores nos indican las direcciones de las rectas que pasando a través de los focos de las cámaras, nos dan una situación de un punto 3D. El punto que estamos buscando,  $\mathbf{Q} = (x, y, z)^t$ , es aquel cuya distancia a las rectas

$$\begin{aligned} \vec{l} &= \mathbf{C} + \lambda \vec{n} \\ \vec{l}' &= \mathbf{C}' + \lambda' \vec{n}' \end{aligned} \quad (3.4.2)$$

sea mínima. La distancia al cuadrado de este punto a cada una de las rectas viene dada por

$$\begin{aligned} dist^2(\mathbf{Q}, \mathbf{l}) &= \|\overrightarrow{\mathbf{C}\mathbf{Q}}\|^2 - \left( \frac{\overrightarrow{\mathbf{C}\mathbf{Q}} \cdot \vec{n}}{\|\vec{n}\|} \right)^2 \\ dist^2(\mathbf{Q}, \mathbf{l}') &= \|\overrightarrow{\mathbf{C}'\mathbf{Q}}\|^2 - \left( \frac{\overrightarrow{\mathbf{C}'\mathbf{Q}} \cdot \vec{n}'}{\|\vec{n}'\|} \right)^2 \end{aligned} \quad (3.4.3)$$

Es posible que por errores que se cometen en la estimación de las matrices de proyección, la precisión de las imágenes y el error en el cálculo de las correspondencias,

estas rectas no intersecten. Buscamos el punto más próximo a las dos, minimizando la suma conjunta de la distancia del punto a las dos rectas,

$$dist^2(\mathbf{Q}, \tilde{\mathbf{I}}, \tilde{\mathbf{I}}') = dist^2(\mathbf{Q}, \tilde{\mathbf{I}}) + dist^2(\mathbf{Q}, \tilde{\mathbf{I}}')$$

Para mayor comodidad suponemos que  $\bar{n} = \frac{\tilde{n}}{\|\tilde{n}\|}$  y  $\bar{n}' = \frac{\tilde{n}'}{\|\tilde{n}'\|}$ . Si desarrollamos la igualdad anterior llegamos a una expresión de la distancia que depende de las tres coordenadas del punto

$$\begin{aligned} dist^2(\mathbf{Q}, \tilde{\mathbf{I}}, \tilde{\mathbf{I}}') = & (x - C_x)^2 + (y - C_y)^2 + (z - C_z)^2 \\ & - ((x - C_x)\bar{n} + (y - C_y)\bar{n} + (z - C_z)\bar{n})^2 \\ & + (x - C'_x)^2 + (y - C'_y)^2 + (z - C'_z)^2 \\ & - ((x - C'_x)\bar{n}' + (y - C'_y)\bar{n}' + (z - C'_z)\bar{n}')^2 \end{aligned} \quad (3.4.4)$$

Para encontrar el mínimo, derivamos con respecto a  $x$ ,  $y$  y  $z$  e igualamos a cero, obteniendo un sistema de ecuaciones de la forma

$$\begin{pmatrix} \left(1 - \frac{2}{\bar{n}\bar{n} + \bar{n}'\bar{n}'}\right) & 1 & 1 \\ 1 & \left(1 - \frac{2}{\bar{n}\bar{n} + \bar{n}'\bar{n}'}\right) & 1 \\ 1 & 1 & \left(1 - \frac{2}{\bar{n}\bar{n} + \bar{n}'\bar{n}'}\right) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{(\mathbf{C}\bar{n}\bar{n} + \mathbf{C}'\bar{n}'\bar{n}') - (\mathbf{C} + \mathbf{C}')}{(\bar{n}\bar{n} + \bar{n}'\bar{n}')}$$

La solución de este sistema de ecuaciones es el punto en el espacio que está más próximo a las dos líneas.

Este proceso se repite para todos los puntos de la imagen. Para cada par de correspondencias tenemos un sistema de ecuaciones distinto que nos da un punto de la escena real.

# Capítulo 4

## Visión estereoscópica

### 4.1 Introducción

El problema de la visión estereoscópica está estrechamente relacionado con el problema del cálculo del flujo óptico en cuanto que el proceso de puesta en correspondencia es común en ambos. Las dificultades y características que nos encontramos en uno y otro son muy parecidas, hasta tal punto que no es raro que muchos autores traten ambos problemas de forma conjunta.

Podemos definir la visión estereoscópica como el proceso por el cual se obtiene una representación tridimensional de una escena a partir de la proyección en un par de imágenes relacionadas entre sí. El objetivo final de la visión estereoscópica es el de reconstruir una escena en el espacio. Este proceso de reconstrucción se puede dividir en tres partes,

- Calibración del par estéreo. Consiste en encontrar la relación geométrica que existe entre las imágenes del par estéreo.
- Estimación de las correspondencias entre las dos imágenes. Dado un punto en una imagen,  $m$ , se busca su correspondiente en la otra,  $m'$ .
- Se calcula la posición 3D del punto a partir de sus proyecciones  $m$  y  $m'$ , haciendo uso de la información dada por la geometría.

El primer proceso está relacionado con el problema de la *calibración de cámaras*. Hablamos de *calibración débil* cuando el objetivo de la calibración es estimar la matriz fundamental,  $\mathbf{F}$ , que relaciona a través de una transformación lineal una imagen con la otra. Nos referimos a *calibración fuerte* al proceso de estimar las matrices de proyección asociadas a las dos cámaras.

El segundo proceso es el de la búsqueda de correspondencias. Nuestro trabajo consiste básicamente en este punto. Para esto, necesitamos conocer la geometría

epipolar asociada al par estéreo y que está definida por la matriz fundamental. Ésta nos permite buscar cada una de las correspondencias a lo largo de una sola línea, denominada línea epipolar.

Una vez determinada todas las correspondencias, podemos proceder a calcular los puntos tridimensionales de la escena. Para esto necesitamos las correspondencias obtenidas en el punto anterior y las matrices de proyección obtenidas por un método de calibración fuerte.

La visión estereoscópica es un proceso delicado que requiere la mayor precisión en cada una de sus partes. Sin una buena estimación de la geometría epipolar es difícil encontrar unas correspondencias precisas. La reconstrucción puede variar considerablemente si se producen errores en las etapas anteriores.

Vamos a definir ahora un concepto que utilizaremos a menudo. Este es el de *mapa de disparidad*. El mapa de disparidad es un mapa – representado generalmente por una imagen de niveles de grises – en el que se representa la norma del vector desplazamiento para cada punto de la imagen,  $\sqrt{u^2 + v^2}$ . Esta carta nos puede dar una idea de la profundidad de la escena, incluso, para ciertas configuraciones geométricas de las cámaras, es posible relacionar esta carta con la profundidad de los puntos. Cuanto más oscuros son los puntos, menos se desplazan de una imagen a otra y cuando más claros, mayor es el desplazamiento. Como ya veremos en las pruebas prácticas, en algunos casos, la claridad de un punto puede estar asociado a la cercanía con la cámara.

Muchos de los temas que tratamos en el capítulo 2 sobre flujo óptico son válidos para este tema. En primer lugar detallaremos algunos factores que afectan en especial a la imágenes estereoscópicas. Luego haremos una clasificación de los distintos métodos que existen en el campo de la visión estéreo y, por último, comentaremos todas las imágenes que utilizamos en las pruebas.

## 4.2 Factores que influyen en los pares estéreo

Se puede decir que estos factores son básicamente los mismos que identificamos para el caso del flujo óptico 2.2, pero existen otros, quizás más sutiles, que pueden afectar más en el caso estereoscópico.

Vamos a identificar estos factores que si bien algunos se pueden dar también en el caso de imágenes para el flujo óptico, influyen con mayor gravedad en el caso estereoscópico.

Cuando estamos intentando estimar el flujo óptico, casi siempre las imágenes se toman con la misma cámara en instantes de tiempo distintos. En este caso, generalmente la mayor parte de la imagen permanece estática y solo se desplaza un número pequeño de objetos. Cuando tratamos con pares estéreo, siempre partimos de dos

cámaras que se sitúan en posiciones distintas del espacio. Las proyecciones de la escena en las imágenes pueden variar mucho dependiendo de la posición relativa que exista entre las dos cámaras. Por lo general, se puede decir que prácticamente todos los puntos se mueven de una imagen a otra y dependiendo de la profundidad de los objetos, estos desplazamientos pueden ser más grandes o más pequeños.

Los factores que hemos identificado son estos:

- **Diferentes tipos de cámaras:** Muchas veces no se dispone de dos cámaras parecidas, con lo que la adquisición de las imágenes puede variar bastante de una a otra. Esto puede provocar que la forma de proyectar en ambas cámaras sea distinto. Con un buen método de calibración este problema se puede solucionar, pero todavía queda el efecto que tiene cada cámara sobre la intensidad de cada pixel. Si los niveles de intensidad varían mucho de una imagen a otra, entonces la suposición lambertiana no se cumple, y los modelos que estamos considerando pueden no ser válidos.
- **Se desplaza un mayor número de puntos:** En los pares estéreo, por lo general, se suelen desplazar prácticamente todos los puntos, ya que si las cámaras están colocadas en posiciones distintas, es difícil que hayan objetos comunes que no se muevan.
- **Mayor variedad de desplazamientos:** Excepto para posiciones relativas específicas – como es el caso de imágenes frontoparalelas – los desplazamientos de una imagen a otra variarán tanto en dirección como en magnitud.
- **Deformación de los objetos:** Quizás como consecuencia del punto anterior, es más fácil encontrar objetos que sufren transformación proyectivas muy fuertes de una imagen a otra dependiendo de la posición relativa de las cámaras.
- **Cambio en las condiciones de luminosidad:** Si la cámara no se mueve y se toman varias imágenes en distintos instantes de tiempo, es probable que las condiciones luminosas no cambien demasiado – excepto el brillo de los objetos en movimiento – sin embargo, en las imágenes estéreo es natural que las condiciones de luminosidad no sean iguales de una cámara a otra, ya que la componente especular hace que las superficies no reflejen la luz por igual en todas direcciones.
- **Geometría epipolar:** Una de las ventajas a la hora de calcular las correspondencias entre pares estéreo es que se puede disponer de la geometría epipolar para facilitar la búsqueda. Esto en principio parece una ventaja, pero tiene sus inconvenientes cuando no se ha calculado con bastante precisión.

- **Reconstrucción:** El objetivo final de la visión estereoscópica es la de obtener una reconstrucción de la escena original. No solo se comete un error en la estimación de las correspondencias, sino que al realizar el proceso inverso al de la proyección se produce otro error en la estimación del punto reconstruido. Esto se debe a que muchas veces los rayos que unen los orígenes de las cámaras con los puntos en correspondencia no intersectan entre sí, teniéndose que aproximar el punto donde se produce la distancia mínima entre las líneas.

### 4.3 Clasificación de las técnicas

La clasificación que proponemos aquí podría ser la misma que la que propusimos para el flujo óptico, pero vamos a agruparlas de la siguiente manera.

#### 4.3.1 Basadas en características

Estos algoritmos establecen correspondencias entre algunas características determinadas que se extraen de las imágenes [Gri85], tales como pixels pertenecientes a contornos [OK85, PMF85], segmentos de línea [AL87, MM88, GR85] o curvas [BB90, Nas92, RF91]. Su principal ventaja es que consiguen información precisa y manipulan pequeñas cantidades de datos, con la consiguiente mejora en tiempo y espacio. Su principal desventaja es la baja densidad de información obtenida en profundidad. Esta clase de métodos se han utilizado ampliamente hace algunos años cuando no era posible recuperar una reconstrucción densa y precisa dentro de un tiempo razonable.

#### 4.3.2 Basadas en áreas

En estas aproximaciones, los mapas de disparidad densos se obtienen correlacionando zonas de las imágenes en las vistas consideradas, suponiendo que estas presentan algún tipo de similaridad [DF94, FHM<sup>+</sup>93, Fua93, Nis84]. Estos métodos proporcionan buenos resultados para áreas relativamente texturadas; sin embargo, estos suelen suponer que la escena observada es localmente fronto-paralela, causando problemas en superficies dobladas, particularmente, cerca de los contornos de oclusión de los objetos. Por último, en el proceso de puesta en correspondencia no tienen en cuenta la información de los contornos que es, normalmente, una información bastante importante y debería ser utilizada con el fin de obtener mapas densos, precisos y viables.

### 4.3.3 Basadas en frecuencias

Otra clase de métodos está basada en la información de las frecuencias de Fourier, que puede ser considerada como una clase de métodos de flujo óptico basadas en el gradiente, con la derivada en el tiempo aproximada por la diferencia entre las imágenes izquierda y derecha de la transformada de Fourier [BFB94, FJ90, FJ93, JJ94, KH75, WWK92]. También se utilizan métodos jerárquicos con el fin de evitar quedar atrapados en mínimos locales.

### 4.3.4 Basadas en energías

Una última aproximación que no sufre de los inconvenientes expuestos anteriormente, consiste en resolver el problema de la puesta en correspondencia a través de una energía a minimizar con algún criterio de regularización [Bar89, BD93, HPB94, PPG94, RD96, Sha93, Yok90]. Después se utiliza un esquema iterativo de la versión discreta de las ecuaciones de Euler-Lagrange para estimar la profundidad. Por ejemplo, en [RD96], los autores proponen un método para calcular directamente el mapa de profundidad  $Z(x, y)$  como el mínimo de la siguiente energía:

$$S(Z) = \iint (I_l(x, y) - I_r(f_1(x, y, Z(x, y))))^2 dx dy + \nu \iint \Phi(|\nabla Z|) dx dy$$

donde  $f_1$  depende de los parámetros intrínsecos y extrínsecos obtenidos de un proceso de calibración previo. El mínimo de la energía anterior se calcula utilizando las ecuaciones de Euler-Lagrange asociada.

## 4.4 Imágenes de prueba

Durante el desarrollo de los métodos propuestos para el problema del cálculo de la carta de disparidad, hemos utilizado una serie de imágenes que comentamos a continuación. Al igual que en el capítulo 2 propondremos algunas imágenes sintéticas y otras reales.

Para cada par estéreo mostraremos la información de la geometría epipolar asociada al mismo, así como las matrices de proyección y, en el caso que se conozca, el mapa de disparidad.

### 4.4.1 Imágenes sintéticas

Aquí proponemos cuatro pares estéreo distintos: el primero es el par estéreo del pasillo creado por el grupo de visión por ordenador de la Universidad de Bonn; el segundo es

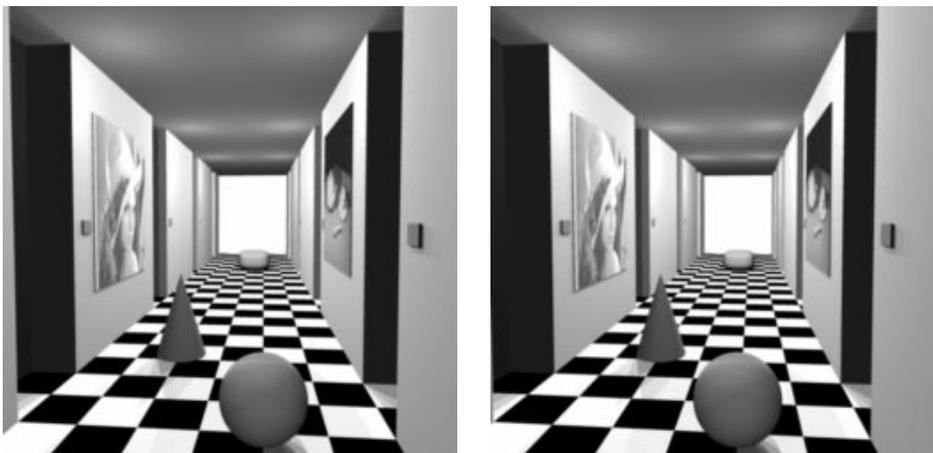


Figura 4.1: Par estéreo del pasillo

una pirámide de escalones con distintos niveles de grises que hemos generado nosotros mismos; el tercero y cuarto son dos tipos de imágenes de puntos aleatorios, en los que se produce una disparidad aunque a simple vista sea difícil detectarla.

### Par estéreo del pasillo

Este par estéreo (figura 4.1) se ha generado de forma sintética a través de un programa de ray tracing. En esta escena nos encontramos con un pasillo en donde hay una esfera, un cono y un objeto al fondo, todos situados a profundidades distintas con respecto a las cámaras. En las paredes hay dos cuadros y la iluminación se reparte a lo largo del pasillo con pares de focos situados a ambos lados del mismo.

Las imágenes se han generado de tal forma que las cámaras están situadas de forma frontoparalela, es decir, las líneas epipolar son horizontales y el desplazamiento solo ocurre en la dirección de las  $x$ . Es fácil deducir que la matriz fundamental es igual a

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

En la figura 4.2 se muestran los mapas de disparidad en los dos sentidos y en la figura 4.3 se muestran las oclusiones también en los dos sentidos.

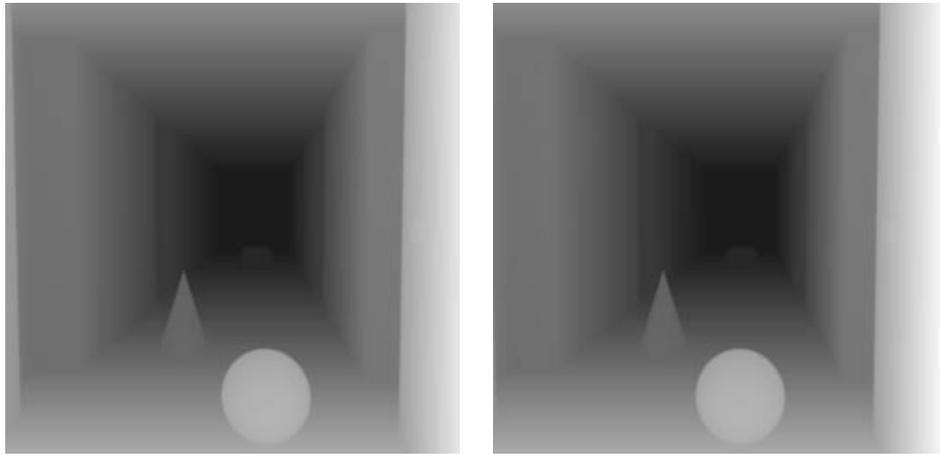


Figura 4.2: **Disparidad del par estéreo del pasillo:** Izquierda, disparidad que se produce al calcular el flujo de izquierda a derecha del par estéreo en 4.1 y, derecha, la disparidad en el sentido contrario.

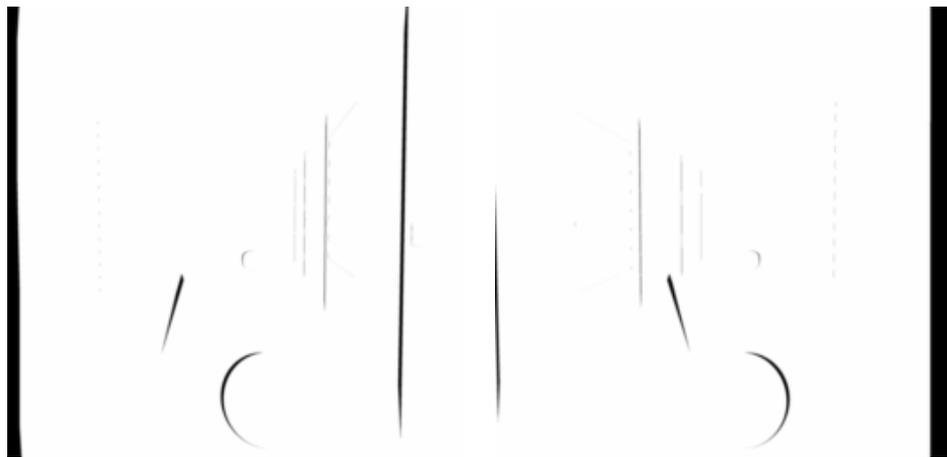


Figura 4.3: **Oclusiones del par estéreo del pasillo:** mapa de oclusiones que se producen al calcular las disparidades de izquierda a derecha (imagen de la izquierda) y en el sentido contrario (imagen de la derecha).

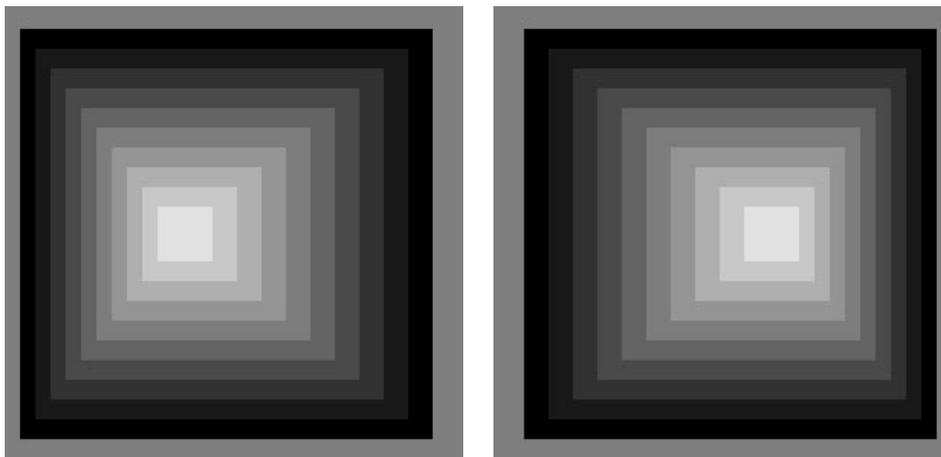


Figura 4.4: Par estéreo de la pirámide.

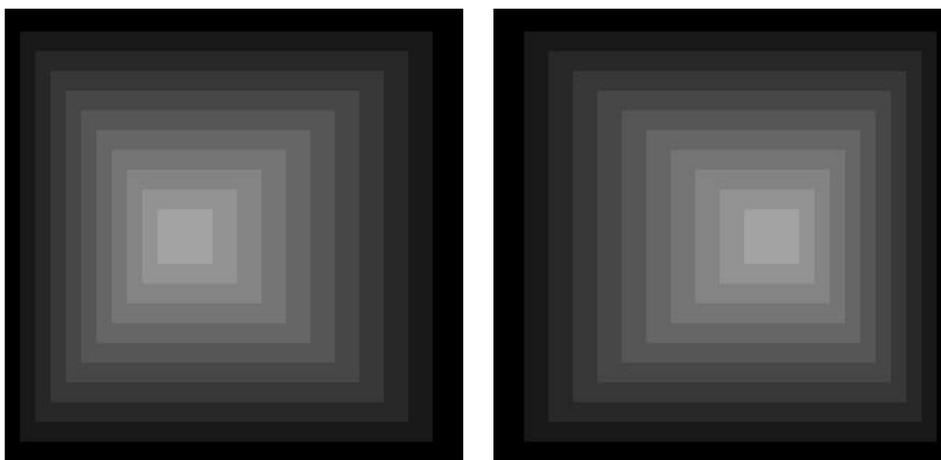


Figura 4.5: Disparidad del par estéreo de la pirámide.

### Par estéreo de la pirámide

En este par estéreo (figura 4.4) se muestra una pirámide escalonada en la que cada escalón tiene un nivel de gris distinto.

Las cámaras también están situadas de forma frontoparalela, luego su matriz fundamental es

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

En la figura 4.5 se muestra la disparidad que se produce al ir en los dos sentidos. Cada escalón se mueve con respecto a su predecesor una cantidad constante.

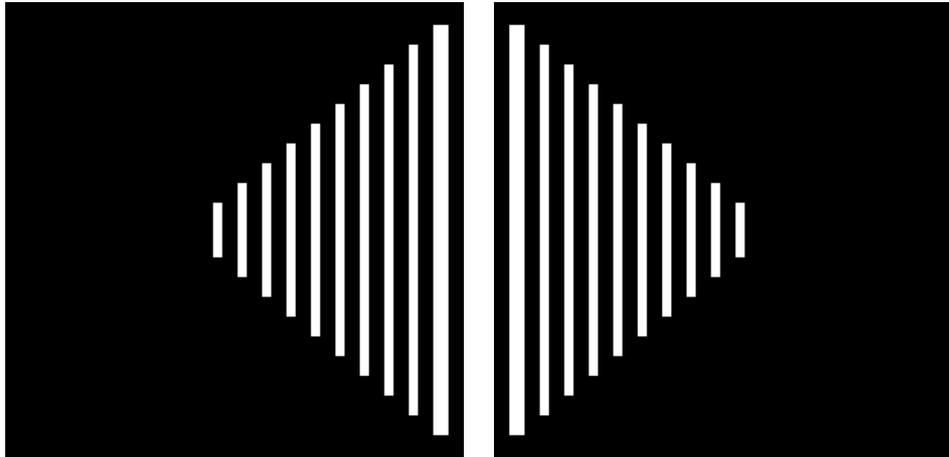


Figura 4.6: Mapa de oclusiones del par estéreo de la pirámide.

Las oclusiones (figura 4.6) que se producen en cada dirección vienen definidas por franjas verticales a la derecha o a la izquierda dependiendo si el desplazamiento es en el sentido de la imagen izquierda a la derecha o viceversa.

#### Pares estéreo aleatorios

En la figura 4.7 se muestra un tipo especial de imágenes estéreo propuestas por Bela Julesz. En estas imágenes se representa una escena tridimensional con puntos aleatorios. Con este par estéreo, Julesz quiso demostrar que el sistema visual humano no necesita formas definidas para detectar una disparidad entre dos imágenes. Para el ser humano resulta más sencillo detectar la profundidad de una escena en donde existen formas reconocibles, pero también es capaz de detectarla en escenas como estas.

Como veremos en las pruebas que expondremos en el capítulo 7, nuestros métodos son capaces de detectar la disparidad en este tipo de imágenes.

En la figura 4.8 se muestra otra imagen parecida desarrollada por nosotros. Es una imagen de puntos aleatorios que presenta una disparidad en forma de pirámide escalonada. El desplazamiento entre los escalones de la pirámide es del orden de tres pixels entre cada escalón.

#### 4.4.2 Imágenes reales

Las imágenes reales que utilizamos son las de los rostros de Herve, Cyrille, Javier y Julio, y la biblioteca del INRIA en Sophi-Antipolis.

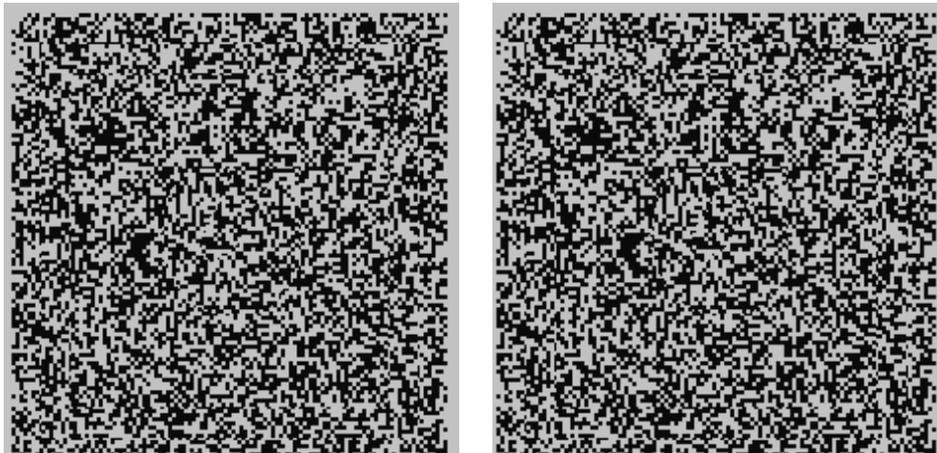


Figura 4.7: Par estéreo de Julesz

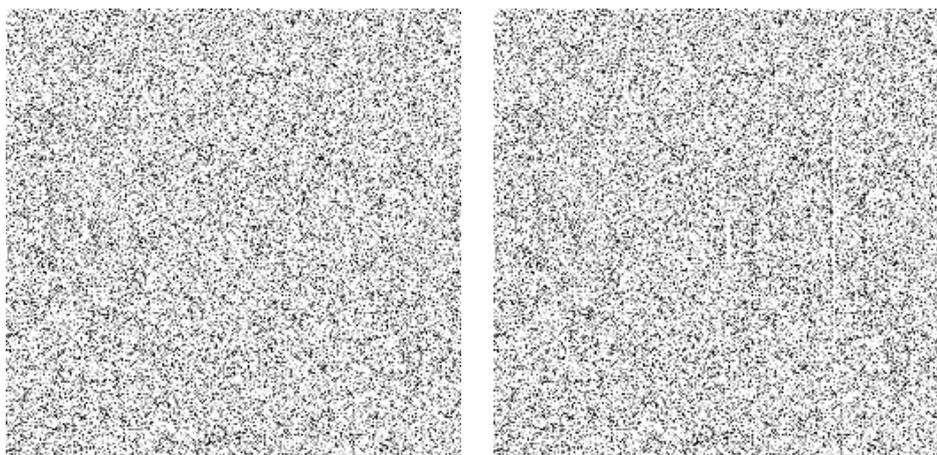


Figura 4.8: Par estéreo de la pirámide de Julesz.

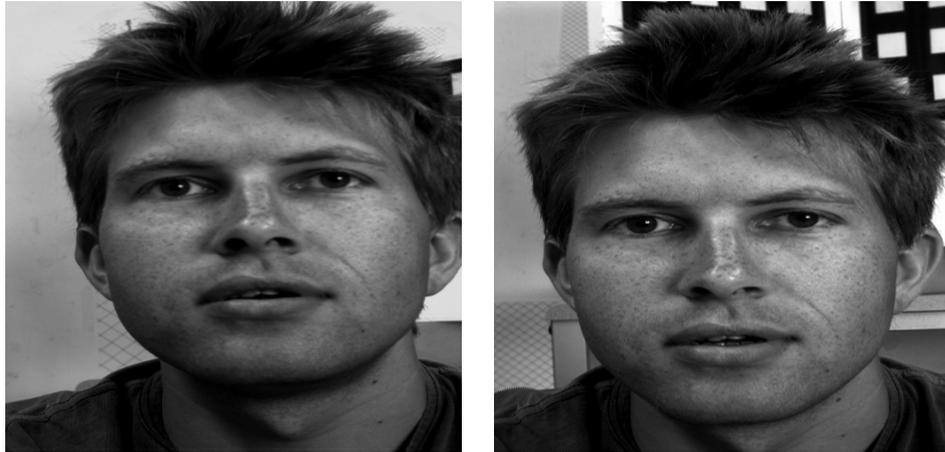


Figura 4.9: Par estéreo de Herve:.

#### Par estéreo de la cara de Herve

Originalmente, esta escena estaba rotada  $90^0$  a la derecha, mostrándose su cara en horizontal. Nosotros aquí la hemos puesto en vertical, solo a efectos de visualización.

La geometría epipolar está definida por

$$\mathbf{F} = \begin{pmatrix} 1.49318e - 07 & 3.39847e - 06 & -0.00255432 \\ -7.18153e - 06 & 1.08563e - 06 & 0.0478208 \\ 0.00307686 & -0.0475878 & 1 \end{pmatrix}$$

Esta geometría define líneas epipolares casi horizontales. Las cámaras están situadas en posiciones casi frontoparalelas, con lo que el mapa de disparidad nos da una buena idea de la profundidad de la escena.

Las matrices de proyección asociadas a las dos cámaras – de la izquierda y de la derecha respectivamente – son

$$\tilde{\mathbf{P}} = \begin{pmatrix} 0.6917 & -0.036930 & -0.531179 & 63.516085 \\ 0.196776 & 1.25842 & 0.158295 & 108.349952 \\ 0.000424 & -0.000075 & 0.000419 & 1.0 \end{pmatrix}$$

$$\tilde{\mathbf{P}}' = \begin{pmatrix} 0.768037 & -0.04413 & -0.417032 & 154.349277 \\ 0.181712 & 1.245176 & 0.160536 & 90.489840 \\ 0.000379 & -0.000044 & 0.000470 & 1.0 \end{pmatrix}$$



Figura 4.10: Par estéreo de la biblioteca del Inria

### Par estéreo de la biblioteca del Inria

En la figura 4.10 mostramos el par estéreo de la biblioteca del INRIA en Sophia-Antipolis, Niza. En esta podemos ver un edificio, unos coches aparcados, una farola y algunos arbustos. Si nos fijamos en las ventanas del edificio a la izquierda, vemos que en una de ellas se produce un cambio importante en el nivel de las intensidades de grises. Esto se debe probablemente a algún efecto de brillo. Veremos en las pruebas que esta ventana dará problemas a la hora de captar su disparidad.

La geometría epipolar está definida por

$$\mathbf{F} = \begin{pmatrix} -1.2283e-08 & 1.51231e-07 & 9.24541e-05 \\ -1.45856e-07 & 1.86386e-08 & -0.00016238 \\ -8.64161e-05 & 0.000150838 & 0.00292565 \end{pmatrix}$$

Las matrices de proyección asociadas a las dos cámaras – de la izquierda y de la derecha respectivamente – son

$$\tilde{\mathbf{P}} = \begin{pmatrix} 0.0970209 & 0.0839468 & 0.01148 & -0.971534 \\ 0.0170976 & -0.0412608 & -0.19218 & -0.0243176 \\ 9.31884e-05 & -0.000244737 & 1.77068e-05 & -0.00333367 \end{pmatrix}$$

$$\tilde{\mathbf{P}}' = \begin{pmatrix} -0.00426314 & -0.0033588 & -0.000573386 & 0.0418263 \\ -0.000753117 & 0.00223661 & 0.00829619 & 0.00271194 \\ -3.77021e-06 & 1.12864e-05 & -1.72595e-06 & 0.000151626 \end{pmatrix}$$

### Par estéreo de Cyrille

En la figura 4.11 mostramos el par estéreo de la cara de Cyrille. Este par estéreo presenta más problemas que el par de Herve, ya que como podemos apreciar en la



Figura 4.11: **Par estéreo de Cyrille**

imagen 4.11 las diferencias de niveles de grises entre las distintas regiones es bastante reducida – en algunas partes se confunde parte de la cara con el fondo – y la rotación que sufre la cabeza es pronunciada.

La geometría epipolar está definida por

$$\mathbf{F} = \begin{pmatrix} -0.0395784 & 0.12248 & -340.293 \\ -0.14318 & 0.0228359 & -334.933 \\ 382.541 & 364.789 & -13102.9 \end{pmatrix}$$

Las matrices de proyección asociadas a las dos cámaras – de la izquierda y de la derecha respectivamente – son

$$\tilde{\mathbf{P}} = \begin{pmatrix} 26.2032 & -17.2914 & 0.17349 & 379.602 \\ 2.33914 & 2.51403 & -34.0252 & 389.735 \\ 0.0112432 & 0.0117628 & -5.39616e-05 & 1 \end{pmatrix}$$

$$\tilde{\mathbf{P}}' = \begin{pmatrix} 24.7377 & -23.2973 & -0.271379 & 399.591 \\ 6.41329 & 6.36073 & -35.7191 & 397.952 \\ 0.0150127 & 0.0104568 & 0.00235911 & 1 \end{pmatrix}$$

#### **Par estéreo de la cara de Julio y Javier**

Para los métodos del cálculo de la disparidad en color utilizamos las fotos de las caras de dos investigadores de nuestro grupo (figuras 4.12 y 4.13). En estas imágenes, hemos cuidado que el fondo fuera bastante homogéneo y hubiese un buen contraste con respecto a las caras. Las cámaras que utilizamos son dos cámaras del mismo tipo controladas por un mando de control remoto.



Figura 4.12: Par estéreo de Julio



Figura 4.13: Par estéreo de Javier

La geometría epipolar está definida por

$$\mathbf{F} = \begin{pmatrix} 73.872423 & 84.146983 & -1205694.051034 \\ -46.314623 & -64.567860 & 858256.364362 \\ 1131211.501774 & -849166.861691 & 59409629.424184 \end{pmatrix}$$

Las matrices de proyección asociadas a las dos cámaras – de la izquierda y de la derecha respectivamente – son

$$\tilde{\mathbf{P}} = \begin{pmatrix} 1396.52875 & -1308.25291 & -15.54100 & 1485116.21910 \\ 201.79457 & 189.23185 & -1872.51273 & 1687977.90930 \\ 0.78465 & 0.61869 & 0.03940 & 4103.09720 \end{pmatrix}$$

$$\tilde{\mathbf{P}}' = \begin{pmatrix} 1501.65593 & -1101.07980 & 73.05620 & 1468058.33906 \\ 385.19944 & 354.19704 & -1764.54265 & 1481735.36465 \\ 0.69466 & 0.69548 & 0.18373 & 3953.11964 \end{pmatrix}$$

Estas matrices las obtuvimos por medio de un patrón de cuadros negros sobre fondo blanco organizados en dos tablas perpendiculares, típico en el campo de la calibración de imágenes.

## Parte II

# Estimación de correspondencias utilizando una aproximación variacional

# Capítulo 5

## Introducción

En esta parte proponemos una aproximación variacional para la estimación del flujo óptico y del mapa de disparidad utilizando el operador de difusión de Nagel y Enkelmann [NE86].

En este capítulo mostramos cómo una técnica clásica para la estimación del flujo óptico por Nagel y Enkelmann [NE86] se puede ver como un método de difusión anisotrópica temprana en la que se emplea un tensor de difusión. En este trabajo introducimos tres mejoras sobre la formulación del modelo que:

1. Evitan inconsistencias debidas a la centralización del término de acoplamiento y del término de regularización en distintas imágenes.
2. Utilizan una técnica de análisis multiescala y multipiramidal desde escalas más gruesas a escalas más finas para evitar la convergencia hacia mínimos locales irrelevantes.
3. Creamos un nuevo funcional de energía que es invariante bajo cambios de brillo lineales<sup>1</sup>.

Aplicando un método de descenso por gradiente al funcional de energía resultante, nos lleva a un sistema de ecuaciones de tipo difusión–reacción. Demostramos que este sistema tiene una única solución<sup>2</sup> bajo suposiciones realistas sobre los datos iniciales, y presentamos en detalle un esquema numérico eficiente.

Nuestros métodos crean campos de flujo 100% densos sobre el dominio completo de la imagen, es robusto bajo una amplia variación de parámetros y puede obtener desplazamientos que están más allá del límite de unos pocos pixels, característico de muchos métodos diferenciales para el cálculo del flujo óptico. Demostramos que se

---

<sup>1</sup>En el apéndice A explicamos los detalles a tener en cuenta para que nuestro método sea invariante a cambios de brillo lineales

<sup>2</sup>Esta demostración se desarrolla en el apéndice B

comporta mejor que los métodos de flujo óptico clásicos 100% densos que se evalúan en Barron et al. [BFB94].

Una clase importante de métodos para la estimación del flujo óptico son aquellos que se refieren a métodos diferenciales. Frecuentemente estos se consideran útiles solo cuando los campos de desplazamiento son pequeños. El objetivo del presente trabajo es demostrar que la combinación de ideas sobre análisis multiescala lineales y no lineales nos pueden llevar a métodos diferenciales bien definidos que permiten recuperar grandes desplazamientos.

Consideramos dos imágenes  $I_1(x, y)$  e  $I_2(x, y)$  (definidas  $\mathbb{R}^2$  para simplificar la discusión) que representan dos vistas consecutivas en una secuencia de imágenes. Bajo la suposición de que dos pixels correspondientes tienen igual nivel de gris, la determinación del flujo óptico desde  $I_1(x, y)$  a  $I_2(x, y)$  se reduce a encontrar una función  $\vec{h}(x, y) = (u(x, y), v(x, y))$  tal que

$$I_1(x, y) = I_2(x + u(x, y), y + v(x, y)), \forall (x, y) \in \mathbb{R}^2 \quad (5.0.1)$$

Para calcular la ecuación precedente se suele linealizar generando lo que se denomina la restricción de flujo óptico linealizado

$$I_1(\mathbf{x}) - I_2(\mathbf{x}) \simeq \langle \nabla I_2(\mathbf{x}), \vec{h}(\mathbf{x}) \rangle \quad (5.0.2)$$

donde  $\mathbf{x} = (x, y)$ . La restricción de flujo óptico linealizado se basa en la suposición de que los desplazamientos de los objetos,  $\vec{h}(\mathbf{x})$ , son pequeños o que la imagen varía lentamente en el espacio. En otros casos, esta linealización no es válida.

Frecuentemente, en vez de la ecuación (5.0.1) se usa la siguiente igualdad

$$I_1(x - u(x, y), y - v(x, y)) = I_2(x, y), \forall (x, y) \in \mathbb{R}^2 \quad (5.0.3)$$

En este caso el desplazamiento está centrado en la imagen  $I_2(x, y)$ .

El cálculo del flujo óptico y de la carta de disparidad son problemas clásicos mal condicionados en visión por ordenador [BPT88a] y requiere suposiciones adicionales de regularización. La regularización de Horn y Schunck [HS81] refleja la suposición de que el campo de flujo óptico varía de forma suave en el espacio. Sin embargo, ya que muchas secuencias de imágenes naturales se describen mejor en términos de trozos de campos de flujo suaves por discontinuidades, se han realizado muchos trabajos de investigación con el fin de modificar la aproximación de Horn y Schunck para permitir estos campos de flujos discontinuos, ver [ADK99, BA91, BA96, BFBG93, Coh93, DKA95, GR96, HB93, KTB96, MP98, NE86, Nes93, PGPO94, Sch94, Wei98b]

## 5.1 Aproximación variacional

En la mayoría de los problemas de estimación del flujo bajo una aproximación variacional, se suele proponer una energía que depende de dos términos

$$E(\vec{h}) = E_l(\vec{h}) + C E_r(\vec{h}) \quad (5.1.1)$$

El primer término es el término de ligadura, en el que se minimiza el error dado por la diferencia de las correspondencias entre ambas imágenes.

$$E_l(\vec{h}) = \int_{\Omega} \Theta(\vec{h}) d\omega \quad (5.1.2)$$

En muchos métodos se supone que se cumple la condición lambertiana y se puede elegir la ecuación de restricción del flujo óptico, o, en nuestro caso,  $\Theta(\vec{h}) = (I_1 - I_2(\mathbf{x} + \vec{h}))^2$ . Cuando en las imágenes que estamos tratando no se cumple esta suposición, entonces se suelen proponer funciones más complejas (ver p.e. [HCF01]) que ayuden a salvar esta dificultad. Este término nos da un esquema de reacción.

El término de regularidad es el que nos va a permitir imponer restricciones sobre la *continuidad* del flujo. Los primeros en proponer un término de este tipo fueron Horn y Schunck [HS81]

$$E_r(\vec{h}) = \int_{\Omega} \Phi(\nabla I_1, \nabla \vec{h}) d\omega \quad (5.1.3)$$

Ejemplos de tipos de funciones  $\Phi(\nabla I_1, \nabla \vec{h})$  son las que hemos estudiado en el capítulo 1 sobre difusión anisótropa. Este término es el que después de minimizar y obtener las ecuaciones de Euler–Lagrange asociadas, nos da un esquema de difusión de los tipos vistos en ese capítulo. Nosotros utilizaremos un esquema de difusión que utiliza un tensor de difusión como el de Nagel–Enkelmann. Tal y como propusimos en el capítulo sobre difusión anisótropa, este tensor de difusión es muy parecido al que propusimos con matrices de proyección en direcciones paralelas y perpendiculares al gradiente de la imagen.

De la energía propuesta en (5.1.1) y tomando

$$\Theta(\vec{h}) = (I_1 - I_2(\mathbf{x} + \vec{h}))^2 \quad (5.1.4)$$

$$\Phi(\nabla I_1, \nabla \vec{h}) = \text{traza}(\nabla \vec{h}^t \mathbf{D}(\nabla I_1) \nabla \vec{h}) \quad (5.1.5)$$

donde  $\mathbf{D}(\nabla I_1)$  es un tensor de difusión de los vistos en el capítulo 1, las ecuaciones de Euler–Lagrange asociadas son de la forma:

$$0 = C \operatorname{div} \left( \mathbf{D}(\nabla I_1) \nabla \vec{h}(\vec{x}) \right) + \left( I_1(\vec{x}) - I_2(\vec{x} + \vec{h}(\vec{x})) \right) \nabla I_2(\vec{x} + \vec{h}(\vec{x})) \quad (5.1.6)$$

Para buscar el mínimo a este sistema, aplicamos un método de descenso por gradiente, haciendo variar el flujo en el tiempo

$$\frac{\partial \vec{h}}{\partial t} = C \operatorname{div} \left( \Phi(\|\nabla \vec{h}\|) \nabla \vec{h}(\mathbf{x}) \right) + \left( I_1(\mathbf{x}) - I_2(\vec{x} + \vec{h}(\mathbf{x})) \right) \nabla I_2(\mathbf{x} + \vec{h}(\mathbf{x})) \quad (5.1.7)$$

Más adelante veremos que en el caso del flujo óptico, este es un sistema de ecuaciones en EDP, mientras que en el caso del cálculo del mapa de disparidad, gracias a la parametrización que vimos en el capítulo 3, ésta es una ecuación simple.

En lo que sigue de este capítulo, hablaremos indistintamente de flujo para referirnos tanto al flujo óptico como el flujo estéreo.

## 5.2 Modelo propuesto por Nagel y Enkelmann

En 1986 Nagel y Enkelmann [NE86] (ver también [Nag83]) propusieron un método para el cálculo del flujo óptico que introducía una importante mejora. Ellos consideraron el siguiente problema de minimización:

$$E_{NE}(\vec{h}) = \int (I_1(x - u(x, y), y - v(x, y)) - I_2(x, y))^2 dx + \\ C \int \operatorname{traza} \left( \left( \nabla \vec{h}^T \right) \mathbf{D}(\nabla I_1) \left( \nabla \vec{h} \right) \right) dx \quad (5.2.1)$$

donde  $C$  es una constante positiva y  $\mathbf{D}(\nabla I_1)$  es una matrix de proyección regularizada en la dirección perpendicular a  $\nabla I_1$  (ver capítulo 1 para más detalles):

$$\mathbf{D}(\nabla I_1) = \frac{1}{|\nabla I_1|^2 + 2\lambda^2} \left\{ \begin{pmatrix} \frac{\partial I_1}{\partial y} \\ -\frac{\partial I_1}{\partial x} \end{pmatrix} \begin{pmatrix} \frac{\partial I_1}{\partial y} \\ -\frac{\partial I_1}{\partial x} \end{pmatrix}^T + \lambda^2 \mathbf{Id} \right\} \quad (5.2.2)$$

En esta fórmula  $\mathbf{Id}$  denota la matriz de identidad. La ventaja de este método es que inhibe la difusión del flujo a través de los contornos de  $I_1$  en posiciones donde  $|\nabla I_1| \gg \lambda$ .

A pesar de este mérito, este método todavía deja lugar para ciertas mejoras:

1. El modelo de Nagel y Enkelmann utiliza un flujo óptico que está centrado en  $I_2$  mientras que la matriz de proyección  $\mathbf{D}$ , en el término de regularización, depende de  $I_1$ . Esta inconsistencia puede crear artefactos para campos de desplazamientos grandes.
2. Evitar la linealización de la restricción del flujo óptico tiene como consecuencia que el funcional de energía (5.2.1) sea no convexo. En este caso, algunos algoritmos populares como los métodos de descenso por gradiente, pueden converger a mínimos locales irrelevantes.
3. Los mínimos del funcional de energía (5.2.1) no son invariantes bajo cambios de intensidad lineales de las imágenes  $I_1$  e  $I_2$ .

En el documento actual abordaremos estos puntos introduciendo tres mejoras en el modelo de Nagel-Enkelmann:

1. Diseñamos un funcional de energía que centra consistentemente tanto la restricción del flujo óptico como el término de regularización en la misma imagen.
2. Conseguimos la convergencia al mínimo global de la energía integrando el método dentro del marco de un análisis multiescala lineal y, también, de una estrategia multipiramidal que permiten ir progresando de escalas más amplias a escalas más finas en pequeños pasos.
3. Introducimos una adaptación de los parámetros  $C$  y  $\lambda$  al rango dinámico de las imágenes tal que el funcional de energía resultante sea invariante bajo cambios de brillo lineales<sup>3</sup>. Esta adaptación es particularmente útil en el contexto de nuestra estrategia de análisis multiescala que altera el rango dinámico de las imágenes.

Aplicando un método de descenso por gradiente a nuestro modelo nos lleva a un sistema de dos ecuaciones ligadas de tipo difusión-reacción, para las que establecemos la existencia y unicidad de soluciones (ver apéndice B). Además, estas ecuaciones se

---

<sup>3</sup>En el apéndice A se explica la manera de que el método sea invariante a los cambios de brillo lineales

pueden relacionar con filtros de difusión anisótropos con tensores de difusión. Presentamos un esquema numérico eficiente que está basado en una discretización implícita por diferencias finitas. Discutimos el papel de los parámetros del modelo y demostramos que nuestro modelo permite una estimación precisa del campo de flujo para un amplio rango de parámetros. Esto se hace considerando tanto secuencias de imágenes sintéticas, para las que se conoce el flujo real, como para secuencias reales.

### 5.3 Recuperando grandes desplazamientos

Uno de los propósitos de este trabajo es el de estimar el flujo óptico para desplazamientos superiores a unos pocos pixels. La suposición que se hace al establecer la ecuación de la restricción del flujo óptico (5.0.2) es la de campos de desplazamientos cortos. Cuando linealizamos la ecuación estamos suponiendo que los desplazamientos de los pixels entre distintas imágenes son pequeños. Si los desplazamientos son grandes, el desarrollo por series de Taylor no es una buena aproximación a la función del flujo.

Es por esto que se hace necesario introducir algún mecanismo que supere este inconveniente. Durante nuestro trabajo, hemos abordado este problema desde dos puntos de vista: a) siguiendo una estrategia multiescala, en donde se utiliza la teoría de los análisis multiescala lineales para conseguir un solape entre las regiones más distantes; b) siguiendo una estrategia multipiramidal, en donde se realizan una serie de zooms a las imágenes con el fin de aproximar los pixels entre sí.

Ambas estrategias son, en esencia, bastante similares, aunque como ya comentaremos, existen diferencias que pueden llevar a soluciones distintas.

A continuación comentamos cada una de las estrategias desarrolladas y hacemos una comparación final entre las dos:

#### **Estrategia multiescala**

El funcional de energía (5.1.1), con la función  $\Theta$  vista anteriormente, puede no ser convexo debido a que el término de acoplamiento no está linealizado. En este caso no podemos esperar que la solución del sistema de ecuaciones elíptico, (5.1.6), sea única. Como consecuencia, el estado asintótico del sistema parabólico (5.1.7), que se utiliza para aproximar el flujo, depende de la aproximación inicial. Podemos esperar que el algoritmo converja a un mínimo local del funcional de energía (5.1.1) que está localizado en la vecindad de la aproximación inicial. Cuando los desplazamientos en la escena son pequeños, la elección normal es tomar  $u \equiv v \equiv 0$  como inicialización del flujo. Para campos de desplazamiento grandes, esto podría no funcionar, necesiándose una mejor aproximación inicial. En este caso introducimos nuestro método en un marco

multiescala lineal [Iij62, WII99]. Considerando el problema a escalas grandes evitamos que el algoritmo converja a mínimos locales irrelevantes. La solución a una escala mayor sirve, luego, como aproximación inicial para resolver el problema a una escala más fina. La estrategia de concentración por escalas tiene una larga tradición en la teoría multiescala lineal [Ber87], y a pesar de que algunas cuestiones teóricas quedan abiertas, no ha perdido su popularidad. Para más detalles sobre la teoría multiescala consultar [Flo97, Iij73, Iij89, Lin94, SNFE97].

Procedemos de la siguiente manera. Primero, introducimos un factor de escala lineal en el sistema de EDP parabólico obteniendo

$$\frac{\partial \vec{h}_\sigma}{\partial t} = C \operatorname{div} \left( \mathbf{D} (\nabla I_1^\sigma) \nabla \vec{h}_\sigma \right) + \left( I_1^\sigma(\mathbf{x}) - I_2^\sigma(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) \right) \nabla I_2^\sigma(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})), \quad (5.3.1)$$

donde  $I_1^\sigma = G_\sigma * I_1$ ,  $I_2^\sigma = G_\sigma * I_2$ ,  $\vec{h}_\sigma(\mathbf{x}) = (u_\sigma(\vec{x}), v_\sigma(\vec{x}))$ , y  $G_\sigma * I_j$  representa la convolución de la imagen  $I_j$  con una Gaussiana de desviación estándar  $\sigma$ .

La convolución con una Gaussiana mezcla la información en las imágenes y nos permite recuperar una conexión entre los objetos de  $I_1$  e  $I_2$ . En nuestra aplicación, esta propiedad global característica de la difusión lineal multiescala es muy importante. Esto lo hace más favorable que los análisis multiescala morfológicos en el sentido de Alvarez et al. [AGLM93] ya que estos últimos no pueden transportar información entre objetos topológicamente desconectados.

Empezamos con una escala grande  $\sigma_0$ . Luego calculamos el flujo óptico  $(u_{\sigma_0}, v_{\sigma_0})$  a escala  $\sigma_0$  como el estado asintótico de la solución del sistema de EDP anterior, utilizando como dato inicial  $u \equiv v \equiv 0$ . Luego, elegimos un número de escalas  $\sigma_n < \sigma_{n-1} < \dots < \sigma_0$ , y para cada escala  $\sigma_i$  calculamos el flujo óptico  $(u_{\sigma_i}, v_{\sigma_i})$  como el estado asintótico del anterior sistema de EDP con aproximación inicial  $(u_{\sigma_{i-1}}, v_{\sigma_{i-1}})$ . El último flujo calculado corresponde a la menor escala  $\sigma_n$ . De acuerdo con la estrategia de muestro logarítmico en la teoría multiescala lineal [Koe84], elegimos  $\sigma_i = \eta^i \sigma_0$  con un ratio de descenso de  $\eta \in (0, 1)$ .

En la figura (5.1) se muestra intuitivamente cuál es el proceso considerando la estrategia de análisis multiescala. Se establecen las escalas para distintos valores de  $\sigma$  y luego se empieza de atrás hacia adelante, calculando el flujo en escalas gruesas y utilizando el resultado en la siguiente escala más fina.

### Estrategia multipiramidal

Una de las aproximaciones más utilizadas a la hora de calcular el flujo con grandes desplazamientos, es la estrategia multipiramidal. Esta estrategia es, en esencia, muy

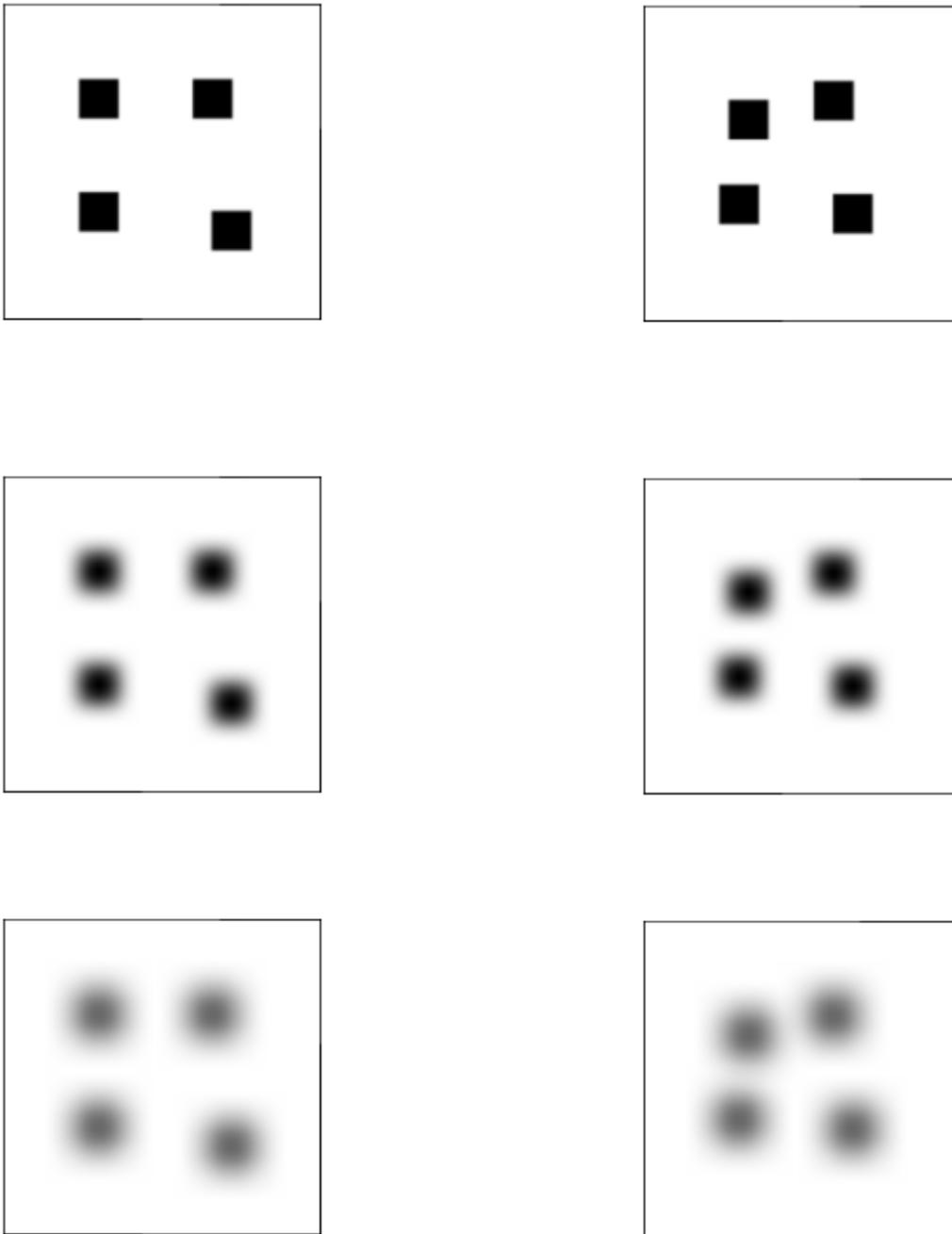


Figura 5.1: **Estrategia multiescala:** A la izquierda se puede ver la secuencia de imágenes a distintas escalas para la imagen de origen (secuencia de los cuadros) y, a la derecha, la secuencia para la imagen destino.

parecida a la propuesta anteriormente. Se establece una secuencia de imágenes a distintas escalas –en este caso las escalas son los zoom que se le aplican a las imágenes originales–, luego se calculan soluciones intermedias que nos sirven de aproximaciones a escalas superiores.

Esta estrategia se ha utilizado frecuentemente en el campo de la estimación de correspondencias. Algunos trabajos relacionados con esta aproximación son los de Anandan [Ana89], Battini et al. [BAK91], Enkelmann [Enk88], Glazer [Gla84], Heitz y Bouthemy [HB93].

El sistema de ecuaciones a resolver es prácticamente igual:

$$\frac{\partial \vec{h}_z}{\partial t} = C \operatorname{div} \left( \mathbf{D} (\nabla I_1^z) \nabla \vec{h}_z \right) + \left( I_1^z(\mathbf{x}) - I_2^z(\mathbf{x} + \vec{h}_z(\mathbf{x})) \right) \nabla I_2^z(\mathbf{x} + \vec{h}_z(\mathbf{x})), \quad (5.3.2)$$

donde  $z$  representa la escala o zoom para el que se calcula la solución del sistema,  $I_{1z} = I_1(z\mathbf{x})$ ,  $I_{2z} = I_2(z\mathbf{x})$  y  $\vec{h}_z(\mathbf{x}) = (u(z\mathbf{x}), v(z\mathbf{x}))$ .

Al igual que para la estrategia anterior, se establece un proceso por el que dada una serie de escalas para las imágenes,  $z_i$ , se calculan soluciones intermedias que nos sirven de aproximación inicial a las siguientes escalas. Generalmente se suele utilizar la escala  $z_i = 2z_{i-1}$ , de forma que una escala superior sea equivalente a realizar un *zoom* a la imagen original por una potencia entera de 2. Entonces se eligen valores de la forma  $z_n > z_{n-1} > \dots > z_1 > z_0$ . Se puede considerar que  $z_0$  es la escala para el tamaño de las imágenes originales. A medida que aumenta la escala disminuye su tamaño. El efecto que conseguimos con esto es que pixels que están distantes en imágenes a escalas pequeñas se aproximen en escalas superiores.

El proceso empezaría por la escala superior  $z_n$ . Empezar por esta escala es equivalente a aplicar a las imágenes un zoom de  $2^n$  de la forma  $I_{iz_n}(\mathbf{x}) = I_i(2^n \mathbf{x})$ . Resolvemos el sistema para esta escala y utilizamos la solución obtenida como aproximación inicial para la siguiente escala. A diferencia de la estrategia anterior, es necesario realizar un ajuste entre iteración e iteración de forma que se actualicen los tamaños de escalas consecutivas. Para esto hay que modificar la solución de la forma  $h_{z_{n-1}}(2\mathbf{x}) = h_{z_n}(\mathbf{x})$

En la figura (5.2) se muestra intuitivamente cuál es el proceso considerando la estrategia multipiramidal. Se establecen los distintos zoom y luego, empezando desde las imágenes más pequeñas, se va calculando el flujo que se utiliza en la siguiente escala aplicando la transformación explicada anteriormente.

Una de las ventajas fundamentales del proceso multipiramidal frente al de multiescala es su coste computacional. En la aproximación multiescala el proceso se desarrolla para imágenes cuyas dimensiones son iguales a las originales y el paso de un escala a otra se realiza de forma suave –el valor de  $\eta$  suele estar próximo a 1 ( $\simeq 0.9$ ). El resultado de esto es que se tarda más tiempo en calcular el sistema (5.1.6), y se

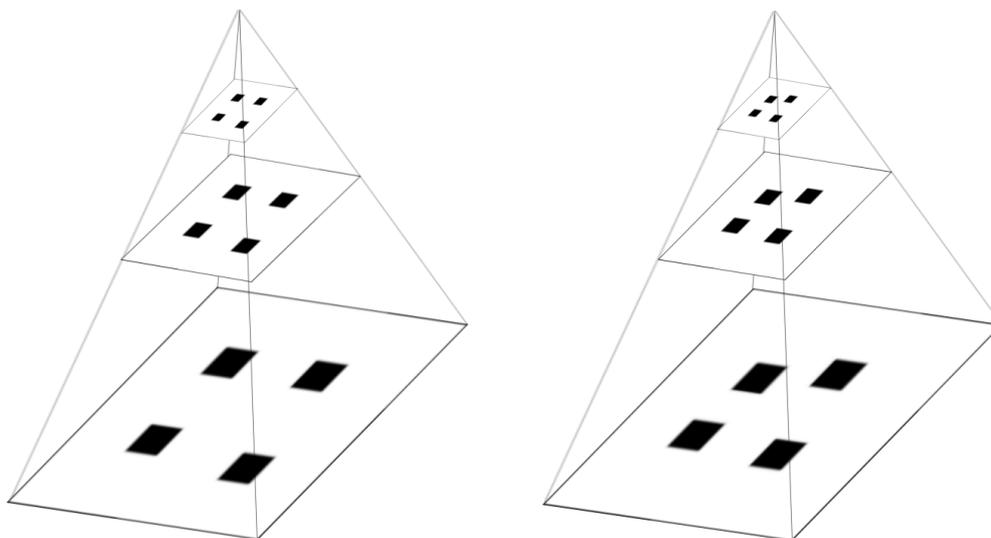


Figura 5.2: **Estrategia multipiramidal:** A la izquierda se muestra la secuencia de zooms para la imagen de origen (secuencia de los cuadros) y a la derecha se muestra la secuencia para la imagen de destino.

necesitan un mayor número de escalas. En la estrategia multipiramidal, a medida que aumenta  $z_i$ , disminuye el tamaño de las imágenes y, por tanto, el tiempo necesario para calcular el sistema. El número de escalas que se establece generalmente para esta última suele ser bastante reducido –del orden de 2 a 6.

La estrategia multipiramidal es, sin embargo, más sensible a la elección del número de escalas. Con el análisis multiescala basta seleccionar un  $\sigma$  que alcance el máximo desplazamiento esperado para poder encontrar con precisión todos los desplazamientos; si, además, el ratio de descenso,  $\eta$ , es lo suficientemente suave, entonces la solución final será buena.

Con la aproximación piramidal podemos pensar que si realizamos muchos zoom sobre las imágenes, los pixels estarán lo suficientemente cerca y encontraremos una buena solución. Esto no siempre es cierto. Aquí nos encontramos con una mayor dificultad a la hora de elegir la escala: 1) si el número de zoom es bajo, podemos no acercarnos a los pixels más alejados y, por lo tanto, no encontrar sus correspondencias, ó 2) si el número es alto, nos podemos encontrar con un problema un poco más sutil. En este caso, el problema viene dado por el paso de escalas superiores a inferiores. Este paso, consiste en la transformación  $h_{z_{n-1}}(2\mathbf{x}) = h_{z_n}(\mathbf{x})$ , o lo que es igual a realizar un zoom inverso. Esto puede dar lugar a que pequeños errores en escalas grandes se conviertan en grandes errores en escalas más finas y se puedan ir ampliando a medida que llegamos a tamaños más grandes de imágenes. Este problema es especialmente importante en zonas homogéneas o zonas con mucha textura, en donde es difícil

encontrar la correspondencia más adecuada.

Para evitar esto, no se realizan muchos zoom y en la escala más gruesa se utiliza un método de correlación para estimar la primera aproximación al flujo.

# Capítulo 6

## Cálculo del flujo óptico

### 6.1 Introducción

En este capítulo vamos a proponer un nuevo método para la estimación del flujo óptico basado en el operador de Nagel–Enkelmann.

Vamos a particularizar el modelo visto en el capítulo anterior para el caso especial del flujo óptico. Partimos de una energía a minimizar con un término de ligadura y otro de regularización del tipo visto en la sección 5.1. En el término de regularización usamos el operador de Nagel–Enkelmann para estimar un flujo continuo por regiones. Deducimos las ecuaciones de Euler–Lagrange que nos llevan a un sistema elíptico de ecuaciones en derivadas parciales. En este sistema tenemos un esquema de difusión anisótropo que define el flujo preservando las discontinuidades de la imagen y regularizando en las zonas homogéneas.

Las imágenes con las que trabajamos son tanto de niveles de grises como en color. Proponemos un método para estimar el flujo para ambos tipos de imágenes. Las imágenes en color están definidas por tres canales (RGB) cada una para un color distinto. El método para el color tiene en cuenta esta información a la hora de estimar el flujo.

Insertamos nuestro método en una estrategia multiescala y multipiramidal de la forma vista en la sección 5.3. Esto nos permite estimar desplazamientos que están más allá de unos pocos pixels. En las pruebas prácticas veremos ejemplos que sobrepasan los 10 pixels de distancia.

En la sección 6.4 explicamos en detalle los parámetros del método tanto para la estrategia multiescala como la multipiramidal.

Una vez propuesto el método, y deducidas las ecuaciones en derivadas parciales, procedemos a resolver el sistema a través de un esquema numérico implícito.

Por último, mostramos algunos resultados numéricos. Comparamos nuestro método con otros estudiados en [BFB94]. Como veremos, nuestro método supera a todos los

métodos cuyas densidades son del 100%, superando, incluso, a otros cuyas densidades son muy inferiores.

## 6.2 Modelo propuesto

### 6.2.1 Centrado consistente

Hemos visto que en el funcional de energía (5.2.1) se utiliza un término de acoplamiento y uno de regularización que están centrados en imágenes distintas. Nuestros experimentos muestran que esta inconsistencia puede llevar a artefactos cuando los campos de desplazamiento son grandes. Como remedio, consideramos un funcional de energía modificado donde tanto el término de acoplamiento como el de regularización están relacionados con  $I_1$  :

$$E(\vec{h}) = \int (I_1(x, y) - I_2(x + u(x, y), y + v(x, y)))^2 dx + \\ C \int \text{traza} \left( \left( \nabla \vec{h}^t \right) \mathbf{D} (\nabla I_1) \left( \nabla \vec{h} \right) \right) dx \quad (6.2.1)$$

Las ecuaciones de Euler-Lagrange asociadas vienen dadas por el sistema de EDP

$$C \operatorname{div} (\mathbf{D} (\nabla I_1^\sigma) \nabla u_\sigma) + \left( I_1^\sigma(\mathbf{x}) - I_2^\sigma(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) \right) \frac{\partial I_2^\sigma}{\partial x}(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) = 0, \quad (6.2.2)$$

$$C \operatorname{div} (\mathbf{D} (\nabla I_1^\sigma) \nabla v_\sigma) + \left( I_1^\sigma(\mathbf{x}) - I_2^\sigma(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) \right) \frac{\partial I_2^\sigma}{\partial y}(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) = 0. \quad (6.2.3)$$

En este trabajo estamos interesados en las soluciones de las ecuaciones (6.2.2) y (6.2.3) en el caso de campos de desplazamientos grandes que no tienen por qué variar suavemente en el espacio. Por lo tanto, no utilizamos una linealización de la restricción del flujo óptico (5.0.2) en el sistema anterior.

### 6.2.2 Relaciones con la difusión anisótropa

Obtenemos la solución de las ecuaciones de Euler-Lagrange (6.2.2) y (6.2.3) calculando el estado asintótico ( $t \rightarrow \infty$ ) del sistema parabólico

$$\frac{\partial u_\sigma}{\partial t} = C \operatorname{div} (\mathbf{D} (\nabla I_1^\sigma) \nabla u_\sigma) + \left( I_1^\sigma(\mathbf{x}) - I_2^\sigma(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) \right) \frac{\partial I_2^\sigma}{\partial x}(\mathbf{x} + \vec{h}_\sigma(\vec{x})), \quad (6.2.4)$$

$$\frac{\partial v_\sigma}{\partial t} = C \operatorname{div} (\mathbf{D} (\nabla I_1^\sigma) \nabla v_\sigma) + \left( I_1^\sigma(\mathbf{x}) - I_2^\sigma(\mathbf{x} + \vec{h}(\mathbf{x})) \right) \frac{\partial I_2^\sigma}{\partial y}(\mathbf{x} + \vec{h}_\sigma(\vec{x})). \quad (6.2.5)$$

Estas ecuaciones también aparecen cuando se aplica el método de descenso por pasos con el fin de minimizar la energía (6.2.1).

Curiosamente, este sistema de ecuaciones de difusión-reacción revela un tensor de difusión que se parece al utilizado en los filtros de difusión anisótropa para el realzado de contornos. Si analizamos  $\mathbf{D} (\nabla I_1^\sigma)$  nos damos cuenta que esta matriz tiene como autovectores  $v_1 = \nabla I_1^\sigma$  y  $v_2 = \nabla I_1^{\sigma,\perp}$ . Los autovalores correspondientes vienen dados por

$$\lambda_1 (|\nabla I_1^\sigma|) = \frac{\lambda^2}{|\nabla I_1^\sigma|^2 + 2\lambda^2} \quad (6.2.6)$$

$$\lambda_2 (|\nabla I_1^\sigma|) = \frac{|\nabla I_1^\sigma|^2 + \lambda_2}{|\nabla I_1^\sigma|^2 + 2\lambda^2} \quad (6.2.7)$$

Observamos que  $\lambda_1 + \lambda_2 = 1$  se cumple independientemente del valor de  $\nabla I_1$ . En el interior de los objetos tenemos que  $\nabla I_1 \rightarrow 0$  y, por lo tanto,  $\lambda_1 \rightarrow \frac{1}{2}$  y  $\lambda_2 \rightarrow \frac{1}{2}$ . En contornos ideales donde  $\nabla I_1 \rightarrow \infty$ , se obtiene que  $\lambda_1 \rightarrow 0$  mientras que  $\lambda_2 \rightarrow 1$ . Por consiguiente, obtenemos un comportamiento isótropo dentro de las regiones, y en el contorno de los objetos de la imagen el proceso suaviza anisótricamente a lo largo de los bordes. Este comportamiento es muy similar a los filtros de difusión anisótropos para el realzado de contornos [Wei96] y es bastante parecido, en esencia, a la curvatura media considerado en [ALM92]. En este sentido se puede ver el método de Nagel-Enkelmann como un predecesor de las técnicas modernas de EDP para restauración de imágenes.

Debemos observar, sin embargo, una diferencia estructural: Las ecuaciones del flujo óptico (6.2.4) y (6.2.5) utilizan un tensor de difusión constante temporalmente, mientras que los tensores de difusión no lineales de los filtros de difusión anisótropos

son una función de la evolución de la imagen misma. Por tanto, el modelo de Nagel-Enkelmann es anisótropo y variante en el espacio, pero es lineal en su parte de difusión. Iijima propuso filtros de difusión anisótropos lineales relacionados en los sesenta y setenta en el contexto de reconocimiento óptico de caracteres; ver [WII99]. Para un tratamiento detallado de los filtros de difusión anisótropos consultar [Wei98a], para una clasificación axiomática de la curvatura media y EDP morfológicas relacionadas para el análisis de imágenes consultar [AGLM93], y una reciente colección de artículos acerca de métodos de suavizado de imágenes a través de EDPs [BDH<sup>+</sup>96, CMSE98, tHRFKE97, NJOE99]

### 6.2.3 Modelo para el cálculo del flujo óptico con imágenes en color

Para el cálculo del flujo óptico con imágenes en color tenemos que tener en cuenta que las imágenes están formadas por tres canales de intensidades. En nuestro caso, utilizaremos canales RGB – canal para el color rojo, otro para el verde y otro para el azul –. Vamos a incorporar esta información al modelo de energía anterior. La energía que proponemos es como sigue:

$$E(\vec{h}) = \sum_{i=0}^3 \int (I_{1,i}(x, y) - I_{2,i}(x + u(x, y), y + v(x, y)))^2 dx + \\ C \int \text{traza} \left( \left( \nabla \vec{h}^t \right) \mathbf{D}(\nabla I_{1,max}) \left( \nabla \vec{h} \right) \right) dx \quad (6.2.8)$$

En el término de ligadura incorporamos un sumatorio para tener en cuenta la información de los tres canales, obteniendo la diferencia entre las intensidades de los correspondientes canales en las dos imágenes. En el término de regularización obtenemos una estimación del gradiente de la imagen como el máximo en norma de los gradientes de cada canal.

$$\nabla I_{1,max} = \max\{\nabla I_{1,R}, \nabla I_{1,G}, \nabla I_{1,B}\}$$

Minimizando obtenemos las ecuaciones de Euler–Lagrange y aplicando un método

de descenso por gradiente, llegamos al sistema de ecuaciones

$$\begin{aligned} \frac{\partial u_\sigma}{\partial t} &= C \operatorname{div} (\mathbf{D} (\nabla I_{1,\max}^\sigma) \nabla u_\sigma) \\ &+ \sum_{i=0}^3 \left( I_{1,i}^\sigma(\mathbf{x}) - I_{2,i}^\sigma(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) \right) \frac{\partial I_{2,i}^\sigma}{\partial x}(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})), \end{aligned} \quad (6.2.9)$$

$$\begin{aligned} \frac{\partial v_\sigma}{\partial t} &= C \operatorname{div} (\mathbf{D} (\nabla I_{1,\max}^\sigma) \nabla v_\sigma) \\ &+ \sum_{i=0}^3 \left( I_{1,i}^\sigma(\mathbf{x}) - I_{2,i}^\sigma(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})) \right) \frac{\partial I_{2,i}^\sigma}{\partial y}(\mathbf{x} + \vec{h}_\sigma(\mathbf{x})). \end{aligned} \quad (6.2.10)$$

Existen formas más complejas de tratar el problema del color, tal y como se ha descrito en algunos trabajos sobre restauración y difusión de imágenes [BC98, SR96, DR01, Wei99]. La forma más sencilla que se podría considerar sería la de calcular una nueva imagen como la media entre los tres canales y utilizar ésta en todo el proceso. El inconveniente de ésta es que se pierde información al promediar pudiendo desaparecer o minimizar el efecto de algunos contornos. La idea de usar el máximo del gradiente para determinar los contornos es que si el gradiente para un canal es alto, es probable que este sea el que defina el contorno del objeto, al mismo tiempo que no se pierde en velocidad de cómputo.

### 6.3 Esquema numérico

Ahora vamos a describir un algoritmo eficiente para nuestro modelo de flujo óptico. Discretizamos el sistema parabólico (6.2.4) y (6.2.5) por diferencias finitas [MM94]. Todas las derivadas espaciales se aproximan por diferencias centradas, y para la discretización en la dirección de  $t$  utilizamos un esquema lineal implícito. Supongamos

que  $\mathbf{D} (\nabla I_1^\sigma) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ . Entonces nuestro esquema lineal implícito tiene la siguiente

estructura

$$\frac{u_{i,j}^{k+1} - u_{i,j}^k}{\tau} = T_l + T_r$$

El primer término es el que corresponde a la parte de ligadura y el segundo al de regularización. Veamos cada uno por separado.

$$\begin{aligned}
T_l &= I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \left( I_1(\mathbf{x}_{i,j}) - I_2(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \right) \\
&\quad + u_{i,j}^k I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) + v_{i,j}^k I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \\
&\quad - u_{i,j}^{k+1} I_{2,x}^2(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) - v_{i,j}^{k+1} I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k), \tag{6.3.1}
\end{aligned}$$

Para el término de regularización tenemos en cuenta la discretización de la divergencia que se explica en el apéndice C.

$$\begin{aligned}
T_r &= C \left( \frac{a_{i+1,j} + a_{i,j}}{2} \frac{u_{i+1,j}^{k+1} - u_{i,j}^{k+1}}{h_1^2} + \frac{a_{i-1,j} + a_{i,j}}{2} \frac{u_{i-1,j}^{k+1} - u_{i,j}^{k+1}}{h_1^2} \right. \\
&\quad + \frac{c_{i,j+1} + c_{i,j}}{2} \frac{u_{i,j+1}^{k+1} - u_{i,j}^{k+1}}{h_2^2} + \frac{c_{i,j-1} + c_{i,j}}{2} \frac{u_{i,j-1}^{k+1} - u_{i,j}^{k+1}}{h_2^2} \\
&\quad + \frac{b_{i+1,j+1} + b_{i,j}}{2} \frac{u_{i+1,j+1}^{k+1} - u_{i,j}^{k+1}}{2h_1h_2} + \frac{b_{i-1,j-1} + b_{i,j}}{2} \frac{u_{i-1,j-1}^{k+1} - u_{i,j}^{k+1}}{2h_1h_2} \\
&\quad \left. - \frac{b_{i+1,j-1} + b_{i,j}}{2} \frac{u_{i+1,j-1}^{k+1} - u_{i,j}^{k+1}}{2h_1h_2} - \frac{b_{i-1,j+1} + b_{i,j}}{2} \frac{u_{i-1,j+1}^{k+1} - u_{i,j}^{k+1}}{2h_1h_2} \right) \tag{6.3.2}
\end{aligned}$$

De igual manera desarrollamos el esquema numérico para la componente  $v$

$$\frac{v_{i,j}^{k+1} - v_{i,j}^k}{\tau} = T_l + T_r$$

El término de ligadura,

$$\begin{aligned}
T_l = & I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \left( I_1(\mathbf{x}_{i,j}) - I_2(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \right. \\
& + u_{i,j}^k I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) + v_{i,j}^k I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \left. \right) \\
& - u_{i,j}^{k+1} I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) - v_{i,j}^{k+1} I_{2,y}^2(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k). \tag{6.3.3}
\end{aligned}$$

El término de regularización,

$$\begin{aligned}
T_r = & C \left( \frac{a_{i+1,j} + a_{i,j}}{2} \frac{v_{i+1,j}^{k+1} - v_{i,j}^{k+1}}{h_1^2} + \frac{a_{i-1,j} + a_{i,j}}{2} \frac{v_{i-1,j}^{k+1} - v_{i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{c_{i,j+1} + c_{i,j}}{2} \frac{v_{i,j+1}^{k+1} - v_{i,j}^{k+1}}{h_2^2} + \frac{c_{i,j-1} + c_{i,j}}{2} \frac{v_{i,j-1}^{k+1} - v_{i,j}^{k+1}}{h_2^2} \\
& + \frac{b_{i+1,j+1} + b_{i,j}}{2} \frac{v_{i+1,j+1}^{k+1} - v_{i,j}^{k+1}}{2h_1h_2} + \frac{b_{i-1,j-1} + b_{i,j}}{2} \frac{v_{i-1,j-1}^{k+1} - v_{i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{b_{i+1,j-1} + b_{i,j}}{2} \frac{v_{i+1,j-1}^{k+1} - v_{i,j}^{k+1}}{2h_1h_2} - \frac{b_{i-1,j+1} + b_{i,j}}{2} \frac{v_{i-1,j+1}^{k+1} - v_{i,j}^{k+1}}{2h_1h_2} \right) \tag{6.3.4}
\end{aligned}$$

Aunque este esquema puede parecer a primera vista un tanto complicado, realmente es fácil de implementar. La notación es autoexplicativa: por ejemplo,  $\tau$  es el paso temporal,  $h_1$  y  $h_2$  denotan el tamaño del pixel en la dirección de  $x$  y  $y$ , respectivamente,  $u_{i,j}^k$  aproxima  $u_\sigma$  en algún punto de la malla  $\mathbf{x}_{i,j}$  en el instante  $k\tau$ , e  $I_{1,x}$  es

una aproximación de  $G_\sigma * \frac{\partial I_1}{\partial x}$ . Calculamos valores del tipo  $I_2(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k)$  a través de una interpolación lineal.

La idea que reside detrás de estos esquemas implícitos es la de utilizar discretizaciones implícitas con el fin de mejorar las propiedades de estabilidad, así como crear sistemas de ecuaciones lineales. Implícito significa que las variables del instante de tiempo nuevo aparecen en ambas partes de la ecuación. Para ecuaciones no lineales, métodos completamente implícitos requerirán una solución computacionalmente cara de sistemas de ecuaciones no lineales. Los métodos lineales implícitos son computacionalmente más rápidos y se consiguen a través de los desarrollos de Taylor, de

forma que tales ecuaciones se convierten en lineales. En nuestro caso, conseguimos esto usando los términos de Taylor de primer orden.

$$\begin{aligned} I_2(\mathbf{x}_{i,j} + \vec{h}_{i,j}^{k+1}) &\approx I_2(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \\ &+ (u_{i,j}^{k+1} - u_{i,j}^k) I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \\ &+ (v_{i,j}^{k+1} - v_{i,j}^k) I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{i,j}^k) \end{aligned}$$

en una discretización implícita completa, y discretizando  $G_\sigma * \frac{\partial I_2}{\partial x}$  y  $G_\sigma * \frac{\partial I_2}{\partial y}$  de forma explícita. Un análisis consistente demuestra que el esquema precedente es de segundo orden en el espacio y de primer orden en el tiempo.

Resolvemos el sistema de ecuaciones lineal resultante de forma iterativa con el algoritmo de Gauß–Seidel simétrico. Para entender mejor su estructura, supongamos que queremos resolver un sistema lineal de la forma  $A\vec{w} = \vec{b}$  donde  $A = D - L - U$  y  $D$  es una matriz diagonal,  $L$  es una matriz triangular inferior, y  $U$  una matriz triangular superior. Entonces las iteraciones del algoritmo de Gauß–Seidel simétrico vienen dadas por

$$\begin{aligned} (\mathbf{D} - \mathbf{L}) \vec{w}^{(n+1/2)} &= \mathbf{b} + \mathbf{U} \mathbf{w}^{(n)}, \\ (\mathbf{D} - \mathbf{U}) \vec{w}^{(n+1)} &= \mathbf{b} + \mathbf{L} \mathbf{w}^{(n+1/2)} \end{aligned}$$

donde el índice superior denota la iteración. Los sistemas se resuelven directamente utilizando eliminaciones directas e inversas, respectivamente.

En una versión anterior de nuestro trabajo [AWS99] estudiamos un esquema explícito que no requería resolver sistemas de ecuaciones lineales. La aproximación lineal implícita que empleamos esta vez, sin embargo, nos ha llevado a acelerar de una a dos veces la velocidad del algoritmo, ya que permite intervalos de tiempo significativamente más grandes sin crear problemas de estabilidad.

## 6.4 Parámetros

Nuestro algoritmo para el cálculo del flujo óptico depende de un número de parámetros que tienen un significado intuitivo:

- El parámetro de regularización  $C$  especifica el balance entre el término de suavizado y la restricción del flujo óptico<sup>1</sup>. Valores grandes llevan a campos

---

<sup>1</sup>En el apéndice A se explica cómo se calcula el parámetro  $C$

de desplazamiento más suaves, rellenando la información a partir de los contornos de la imagen donde hay una mayor precisión.

- El factor de isotropía  $s$  determina el parámetro de contraste  $\lambda$  a través del histograma acumulativo de las magnitudes de los gradientes de las imágenes<sup>2</sup>. Si elegimos  $s = 0.7$  significa que el término de suavizado difumina isotrópicamente en el 70% de las posiciones en la imagen, mientras que el 30% de las posiciones se asume que pertenecen a los contornos de la imagen, donde el suavizado se realiza de forma anisótropa a lo largo de los bordes.
- El intervalo de tiempo  $\tau$  y el tiempo de parada  $T$  para resolver el sistema (5.3.1) o el sistema (5.3.2) en cada escala  $\sigma_m$  son parámetros puramente numéricos. En nuestras experiencias, fijamos  $\tau = 10$  y  $T = 500$  y los resultados obtenidos están suficientemente próximos al estado asintótico. Utilizando valores más pequeños de  $\tau$  o mayores para  $T$  hacen más lento el algoritmo sin mejorar la calidad de los campos de flujos.

En estos esquemas  $\sigma$  representa el factor de escala tal y como vimos en la introducción de esta parte (capítulo 5). Este factor representa, por tanto, la desviación típica si consideramos una estrategia multiescala, o el número de zooms si se trata de una estrategia multipiramidal. Para la primera tenemos los siguientes:

- La escala  $\sigma_0$  denota la desviación estándar de la mayor Gaussiana. En general,  $\sigma_0$  se elige de acuerdo al máximo desplazamiento esperado.
- El ratio de descenso  $\eta \in (0, 1)$  para el cálculo de las escalas  $\sigma_m = \eta^m \sigma_0$ . Podemos esperar una buena concentración si  $\eta$  está cerca de 1.
- La escala más pequeña viene dada por  $\sigma_n$ . Debe estar cerca de la escala original de la imagen con el fin de lograr una localización del flujo óptima.

Para la estrategia multipiramidal tenemos únicamente el número de zooms,  $z_i$ , que es el que determina cuántos zoom se le van a aplicar a las imágenes originales. Normalmente este valor suele estar entre 3 y 6 dependiendo del tamaño original. Por ejemplo, para imágenes del orden de  $512 \times 512$  solemos realizar 4 zooms, mientras que para imágenes inferiores, 3 zooms suele ser suficiente.

En la próxima sección veremos que los resultados de nuestro método se ven afectados a penas por una gran variación de los parámetros. Como consecuencia se pueden utilizar valores por defecto para la mayoría de los parámetros.

---

<sup>2</sup>En el apéndice A se explica cómo se calcula el parámetro  $s$

## 6.5 Resultados experimentales

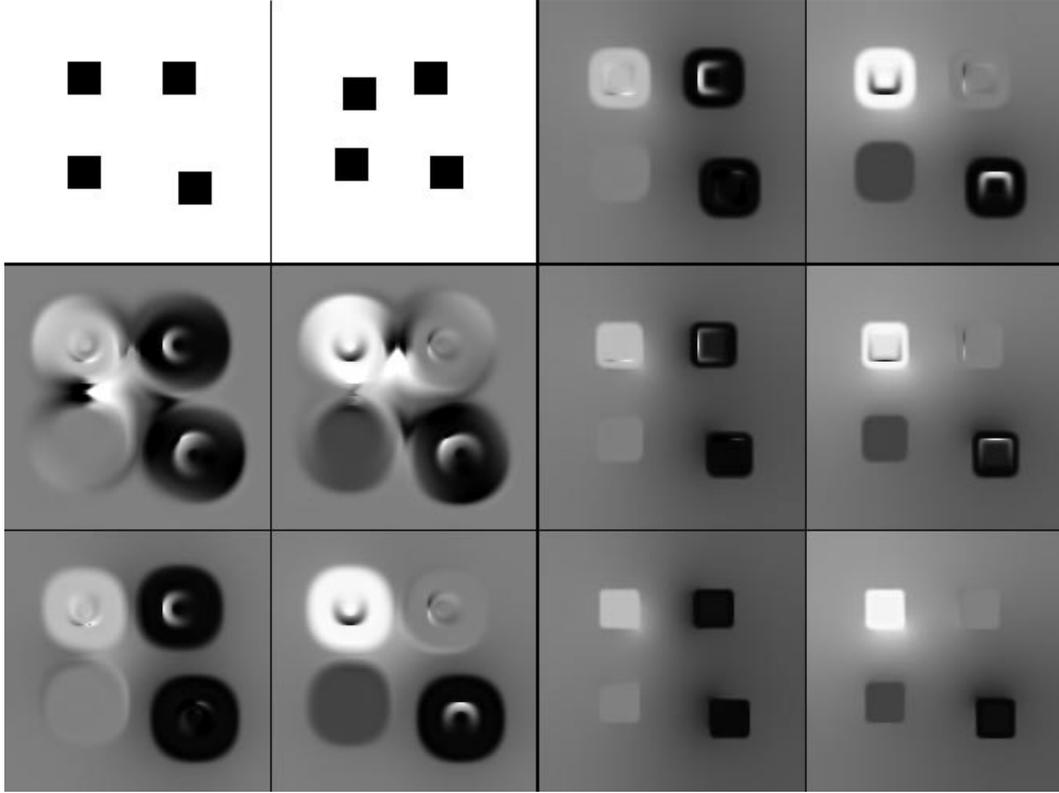


Figura 6.1: Cálculo del flujo óptico para las imágenes de los cuadrados con  $\alpha = 0.6$ ,  $s = 0.1$ , y  $\eta = 0.95$ . (a) **Primera fila, par izquierdo:** Imágenes originales. (b) **Segunda fila, par izquierdo:** Componentes del flujo óptico  $(u, v)$  para  $\sigma_0 = 10$ . (c) **Tercera fila, par izquierdo:** Resultado del flujo óptico después de pasar a la escala con  $\sigma_{12} = 5.7$ . (d) **Primera fila, par derecho:**  $\sigma_{25} = 2.9$ . (e) **Segunda fila, par derecho:**  $\sigma_{37} = 1.4$ , (f) **Tercera fila, par derecho:**  $\sigma_{50} = 0.8$ .

La imagen 6.1 muestra nuestro primer experimento. Utilizamos una imagen sintética compuesta de cuatro cuadrados negros sobre fondo blanco. Cada cuadrado se mueve en una dirección distinta y con una magnitud diferente: Sobre la suposición de que el eje de las  $x$  está orientado desde la izquierda a la derecha y que el eje de las  $y$  está orientado desde arriba hacia abajo, el cuadrado superior izquierdo se mueve con  $(u, v) = (5, 10)$ , el cuadrado superior derecho se desplaza con  $(u, v) = (-10, 0)$ , el cuadrado inferior izquierdo se mueve con  $(u, v) = (0, -5)$ , y el cuadrado inferior derecho sufre una traslación de  $(-10, -10)$ . Con el fin de visualizar el campo de desplazamiento  $(u, v)$  utilizamos dos imágenes de niveles de grises  $(u_{gl}, v_{gl})$  definidas por  $u_{gl} = 128 + 12u$  y  $v_{gl} = 128 + 12v$ . En la imagen 6.1 nos damos cuenta que el flujo estimado mejora significativamente gracias a la técnica multiescala desde  $\sigma_0 = 10$  a

$\sigma_{50} = 0.8$ : las discontinuidades del flujo evolucionan y el campo de desplazamiento aproxima el campo de movimiento verdadero.

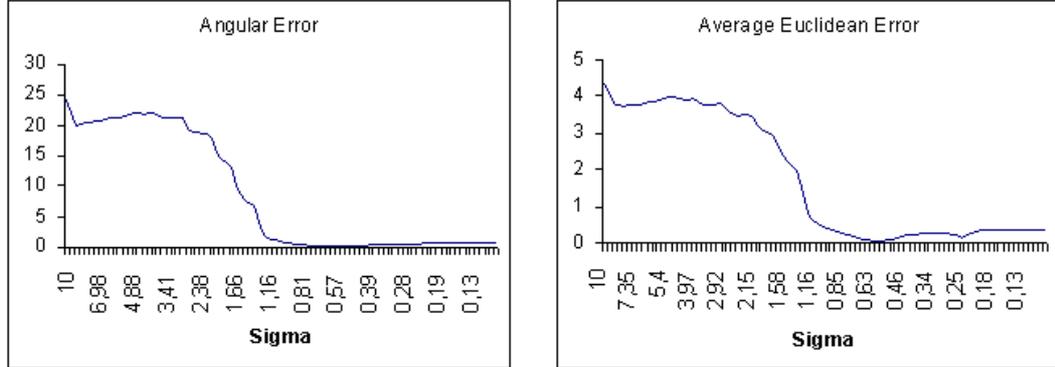


Figura 6.2: **Izquierda:** Error angular medio del cálculo del flujo óptico para los cuadrados del primer frame de la figura 6.1. **Derecha:** Error euclídeo medio correspondiente.

Esta observación cualitativa se confirma en las evaluaciones cuantitativas que se visualizan en la imagen 6.2. La figura de la izquierda muestra el error angular medio dentro de los cuatro cuadrados del marco. El error angular  $\Psi_e$  se ha calculado de la misma manera que en Barron et al. [BFB94] usando

$$\Psi_e = \arccos \left( \frac{u_c u_e + v_c v_e + 1}{\sqrt{(u_c^2 + v_c^2 + 1)(u_e^2 + v_e^2 + 1)}} \right) \quad (6.5.1)$$

donde  $(u_c, v_c)$  denota el flujo óptico correcto, y  $(u_e, v_e)$  el flujo estimado. La figura de la derecha muestra el error euclídeo  $\sqrt{(u_e - u_c)^2 + (v_e - v_c)^2}$  promediado sobre los pixels de los cuatro cuadrados del primer frame.

En ambos casos se puede observar que el error se reduce drásticamente gracias a la estrategia de análisis multiescala, en donde se van reduciendo las escalas hasta alcanzar un valor muy pequeño – valor de la  $\sigma$  de la Gaussiana próxima a la escala real de la imagen. Si se reduce aún más  $\sigma$  nos lleva a un incremento ligero de los errores. Esto se puede deber a los efectos causados por la discretización y la cuantización. Sólo evaluamos los errores en el interior de los cuadrados debido a la homogeneidad del fondo. El flujo no está bien definido en esta zona ya que cualquier desplazamiento del fondo es compatible con la secuencia de imágenes.

La imagen 6.3 demuestra la importancia de un centrado consistente de los términos de regularización y ligadura. En este experimento, el término de regularización sigue al gradiente de  $I_2$ , mientras que el término de ligadura está centrado en  $I_1$ , p.e.

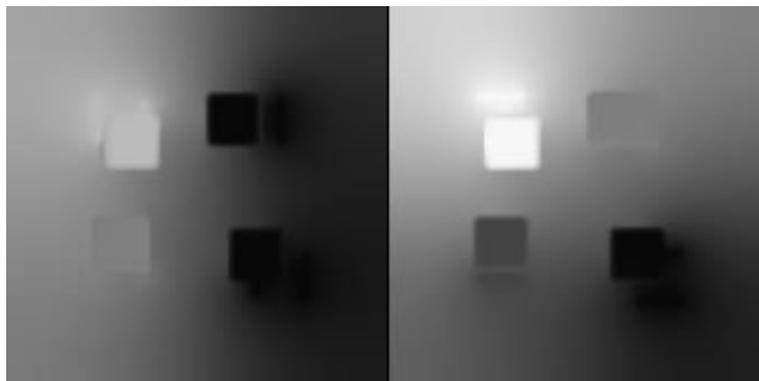


Figura 6.3: Resultado cuando la restricción del flujo óptico y del término de regularización están centrados en imágenes distintas. Las imágenes y los parámetros son idénticos a los de la figura 6.1.

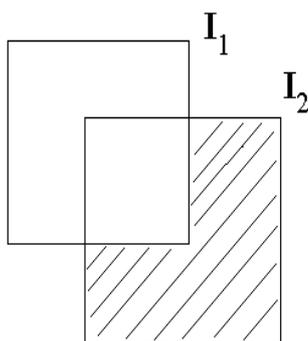


Figura 6.4: Ilustración del problema de las oclusiones. Un cuadro se mueve desde  $I_1$  a  $I_2$ . La región sombreada en la imagen  $I_1$  no tiene correspondencia en  $I_2$ .

estamos buscando desplazamientos desde  $I_1$  a  $I_2$ . Debido a que la localización de los bordes de  $I_1$  e  $I_2$  difieren para desplazamientos largos, no es sorprendente que el campo de desplazamiento que respeta las discontinuidades de  $I_2$  falle en las posiciones de los bordes de  $I_1$ .

Generalmente, cuando un objeto se mueve a través de la secuencia de imágenes, el fondo se oculta parcialmente. Este problema de oclusiones se ilustra en la figura 6.4. En la dirección del movimiento del objeto una región del fondo queda oculto, de manera que los puntos de esta región (la zona sombreada de la imagen 6.4) no tienen correspondencias en  $I_2$ , y la restricción del flujo óptico ya no es válido. En estas regiones del fondo aparecen algunas pequeñas inhomogeneidades como se pueden ver en la figura 6.1. Sin embargo, observamos que el término de regularización de la



Figura 6.5: Cálculo del flujo óptico para la secuencia del taxi (frames 15 y 19) con  $\alpha = 0.6$ ,  $s = 0.1$ ,  $\sigma_0 = 10$ ,  $\sigma_n = 0.8$ , y  $\eta = 0.95$ .

energía ayuda a reducir tal efecto.

Para el siguiente experimento utilizamos la secuencia clásica del taxi de Hamburg. Estos datos están disponibles en el sitio <ftp://csd.uwo.ca> en el directorio `pub/vision`. En lugar de tomar dos frames consecutivos – como generalmente se suele hacer – consideramos los frames 15 y 19. El coche oscuro de la izquierda crea un desplazamiento grande de aproximadamente 12 pixels. En las figuras 6.5 y 6.6 presentamos el flujo calculado. El máximo flujo calculado es de 11.68, que es una buena aproximación del desplazamiento real del coche oscuro. Es interesante notar que aunque el movimiento del peatón en la esquina superior derecha de la escena es difícil de reconocer en las imágenes 6.5, la imagen vectorial 6.6 muestra que este movimiento no pasa desapercibido.

En las figuras 6.7 y 6.8 mostramos los resultados de nuestro método para la secuencia de las torres de mármol. Esta secuencia es copyright de H.-H. Nagel (KOGS/IAKS, Universidad de Karlsruhe, Germany). Ha sido grabada y evaluada por primera vez por Michael Otte [ON95] y está disponible en [http://i21www.ira.uka.de/image\\_sequences](http://i21www.ira.uka.de/image_sequences). La secuencia de las torres de mármol describen una escena sintética donde la cámara se está moviendo. En nuestro experimento tomamos los frames 20 y 25, y utilizamos los mismos parámetros que para la escena del taxi.

En el siguiente ejemplo realizamos una comparación cuantitativa con las técnicas de flujo óptico clásicas descritas en el artículo de Barron *et al.* [BFB94]. Este lo



Figura 6.6: Vectores del flujo óptico para la figura 6.5.

realizamos utilizando los datos de los desplazamientos correctos así como sus utilidades de evaluación que están disponibles en el sitio <ftp://csd.uwo.ca> en el directorio `pub/vision`. Debería notarse que los resultados de [BFB94] han sido alcanzados con secuencias de test donde los desplazamientos son pequeños, mientras que nuestro método está diseñado para campos de desplazamientos grandes. Más aún, sus métodos también utilizan un pre-suavizado en el tiempo lo cual envuelve más de dos frames en donde nosotros sólo utilizamos dos. A pesar de estas limitaciones vamos a mostrar que podemos obtener resultados competitivos con nuestro método.

En la comparación, nos centramos en aquellos métodos en [BFB94] que crean campos de flujo 100% densos. Para muchas tareas tales como la inferencia del movimiento y de la estructura de superficies esto es una propiedad muy deseable. Métodos locales que provocan menores densidades pueden tener que ser complementados con estrategias adicionales para rellenar los lugares donde no hay resultados disponibles. Su eficacia práctica puede depender considerablemente de este tipo de post-proceso. Aproximaciones variacionales con términos de regularización no requieren tal post-proceso ya que automáticamente generan campos de flujo 100% densos.

En las imágenes 6.9 y 6.10 mostramos el flujo óptico calculado para la secuencia del *Cuadrado* que muestra un cuadro moviéndose con velocidad  $(4/3, -4/3)$ . La tabla 6.1 lo compara con los resultados dados en Barron et al. para algunas técnicas clásicas de flujo óptico que crean campos de flujo 100% densos. Se puede ver que nuestro método revela errores menores que estos métodos. En particular, esto también demuestra que nuestras tres modificaciones mejoran sustancialmente el método de Nagel. Mientras que la implementación del método de Nagel en [BFB94] da un error angular de  $34.57^\circ$ ,

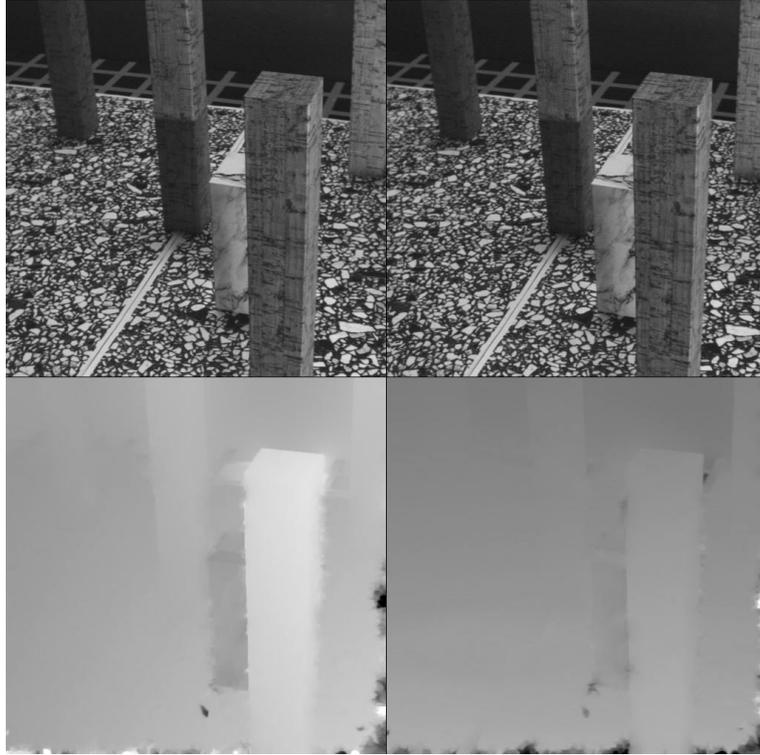


Figura 6.7: Flujo óptico resultado para la secuencia de las torres de mármol (frames 20 y 25) con  $\alpha = 0.6$ ,  $s = 0.1$ ,  $\sigma_0 = 10$ ,  $\sigma_n = 0.8$ , y  $\eta = 0.95$ .

nuestro método genera un error de  $10.97^\circ$ . En este ejemplo Barron et al. asumen que el fondo se mueve en la misma dirección que el cuadrado. Si embargo, como el fondo es constante el desplazamiento no está bien definido en esta zona. Si centramos nuestra atención en el error del flujo óptico calculado dentro del cuadrado, obtenemos un error angular medio de  $0.85$ . Este muestra que el flujo calculado es muy preciso dentro del cuadrado.

Vamos a dirigir nuestra atención a la secuencia de pruebas sintética más compleja realizada por [BFB94], la secuencia de Yosemite con el cielo nublado. Contiene desplazamientos de hasta cinco pixels. Los resultados de nuestro flujo óptico se muestran en las imágenes 6.11 y 6.12, y una comparación con otros métodos se puede encontrar en la tabla 6.2. De nuevo, nuestra técnica mejora todos los métodos propuestos en [BFB94] en los cuales los campos de flujo son del 100% densos. Con un error angular de  $5.53^\circ$  incluso alcanza la calidad de estimación de métodos con densidades entorno al 30%, y la desviación estándar de  $7.40^\circ$  es más pequeña que la desviación estándar de todos los métodos que han sido evaluados en [BFB94]: el mejor método (Lucas y Kanade con  $\lambda_2 \geq 5.0$ ) tenía un error angular medio de  $3.22^\circ$  con una desviación estándar de  $8.92^\circ$  y una densidad de solo el 8.7%.

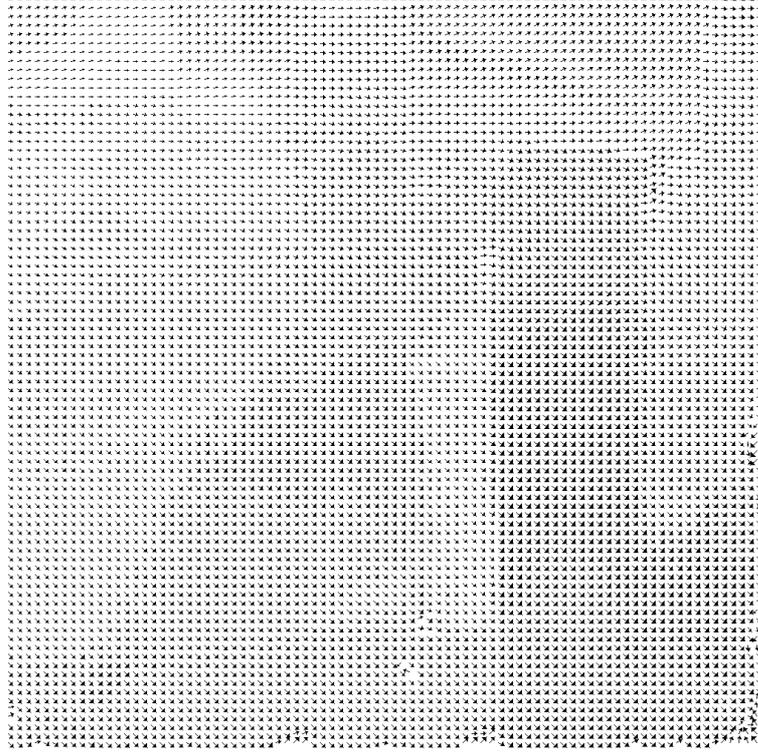


Figura 6.8: Vectores del flujo óptico para la figura 6.7.

Con el fin de evaluar la robustez de nuestro algoritmo con respecto a la elección de parámetros, presentamos en la tabla 6.3 los errores para la secuencia de Yosemite tomando diferente valores para los parámetros. Para simplificar la presentación, fijamos la escala más fina a  $\sigma_n = 1$ , y como parámetros numéricos utilizamos  $\tau = 10$  y  $T = 500$ . Estos parámetros son casi independientes de la imagen y pueden, por lo tanto, ser colocados a valores por defecto. Por esto, variamos solo los parámetros  $\alpha$ ,  $s$ ,  $\eta$  y  $\sigma_0$  en la tabla 6.3.

Antes que nada, se puede apreciar que nuestro método supera a todos los métodos en [BFB94] con densidad del 100% no solo en el caso de parámetros óptimos sino también para un amplio rango de valores. Estudiemos ahora la influencia de los parámetros con más detenimiento.

Una observación importante a partir de la tabla 6.3 es que el ratio de descenso  $\eta$  tiene una influencia importante en el resultado: Valores cercanos a 0.5, tales como se suelen utilizar de forma implícita en los algoritmos típicos piramidales, no son óptimos. Un enfoque lento con  $\eta = 0.95$  da resultados significativamente mejores. Nuestra experiencia con otras imágenes nos sugiere que  $\eta$  se puede fijar a este valor para todas las aplicaciones.

Si se elige un valor demasiado pequeño para el factor de isotropía  $s$  no se empeoran

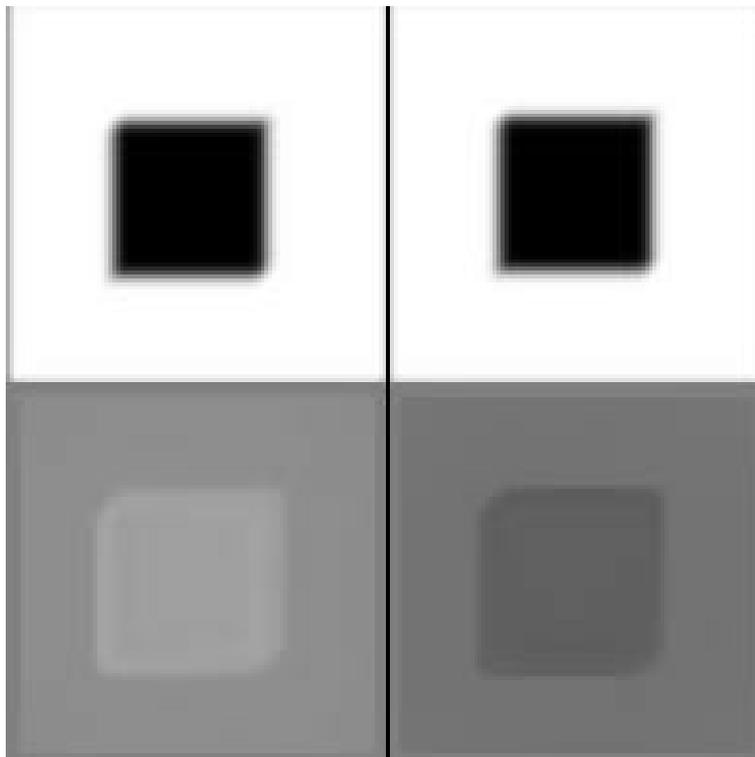


Figura 6.9: Cálculo del flujo óptico para la secuencia del *Cuadrado* con  $\alpha = 0.6$ ,  $s = 0.1$ ,  $\sigma_0 = 10$ ,  $\sigma_n = 1$ , y  $\eta = 0.95$ .

mucho los resultados, mientras que para valores más grandes el término de suavizado se convierte en isotrópico en casi todos sitios y aproxima el esquema de Horn and Schunck [HS81]. Con el fin de evitar resultados deteriorados, proponemos fijar  $s = 0.1$ , que significa que el método suaviza anisótricamente en un 90% de las localizaciones.

Si observamos el parámetro de suavizado  $\alpha$ , nuestro método parece ser bastante robusto frente a sobre y subestimaciones. Por lo tanto hemos utilizado un valor fijo de 0.6 para todos los experimentos.

Tal y como se mencionó anteriormente, la escala inicial  $\sigma_0$  se debería elegir de forma que se cubra el desplazamiento más largo esperado. Encontramos que las sobrestimaciones son menos críticas que las subestimaciones. Esto también confirma la utilización de la estrategia multiescala. Valores muy pequeños incrementan el peligro de acabar en un mínimo local irrelevante. Realmente,  $\sigma_0$  era básicamente el único parámetro que teníamos que adaptar con el fin de analizar diferentes secuencias de imágenes. Debido a que su interpretación física es clara, esta adaptación era simple.

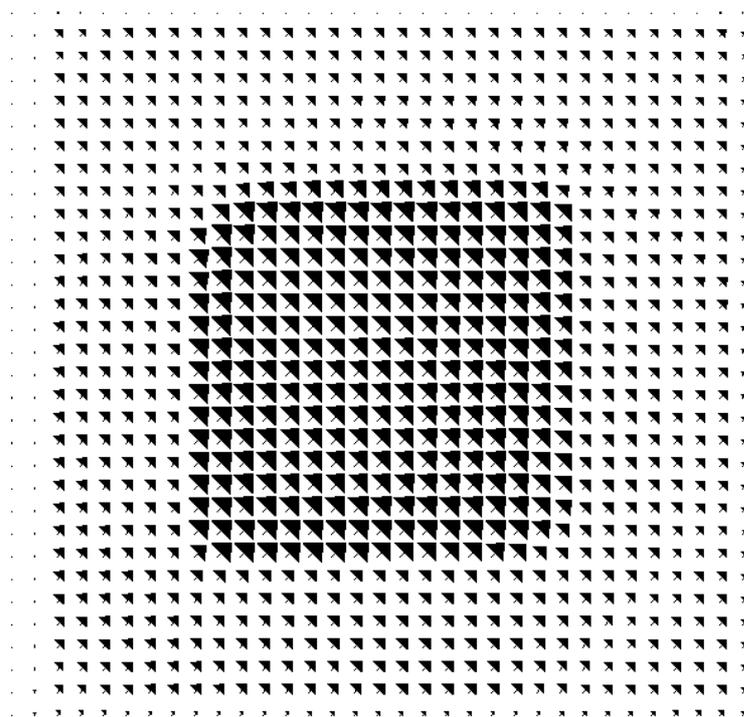


Figura 6.10: Vectores del flujo óptico para la figura 6.9.

## Resultados experimentales para imágenes en color

En esta prueba hemos utilizado la secuencia del cubo (ver figura 2.14). En la figura 6.13 se muestra el resultado del flujo en niveles de grises. Como se puede ver, el desplazamiento en  $x$  (imagen inferior izquierda) es positivo en una mitad del cilindro y negativo en la otra, al igual que pasa en la otra dirección.

En esta prueba hemos utilizado la estrategia multipiramidal. Los parámetros utilizados son  $\alpha = 1.0$ ,  $s = 0.5$ ,  $T = 500$ ,  $dt = 10$  y el número de zooms 3. El tamaño original de estas imágenes es de  $256 \times 256$ , con lo que la escala inicial implica un tamaño de  $16 \times 16$ .

Como vemos en la figura 6.14 se ha detectado bastante bien el efecto de la rotación. Efectivamente, la rotación es en el sentido contrario a las agujas del reloj y centrado en la base del cilindro.

Tabla 6.1: Comparación con los resultados de los métodos en [BFB94] 100% densos y nuestro método para la secuencia del *Cuadrado*.

Técnica	Error medio	Desviación Estándar	Densidad
Horn and Schunck (original)	47.21°	14.60°	100%
Horn and Schunck (modificado)	32.81°	13.67°	100%
Nagel	34.57°	14.38°	100%
Anandan (sin umbral)	31.46°	18.31°	100%
Singh (paso 1)	49.03°	21.38°	100%
Singh (paso 2)	46.12°	18.64°	100%
nuestro método	10.97°	9.60°	100%

Tabla 6.2: Comparación con los resultados de los métodos en [BFB94] 100% densos y nuestro método para la secuencia de Yosemite.

Técnica	Error medio	Desviación Estándar	Densidad
Horn and Schunck (original)	31.69°	31.18°	100%
Horn and Schunck (modificado)	9.78°	16.19°	100%
Nagel	10.22°	16.51°	100%
Anandan (sin umbral)	13.36°	15.64°	100%
Uras et al. (sin umbral)	8.94°	15.61°	100%
Singh (paso 2)	10.03°	13.13°	100%
nuestro método	5.53°	7.40°	100%

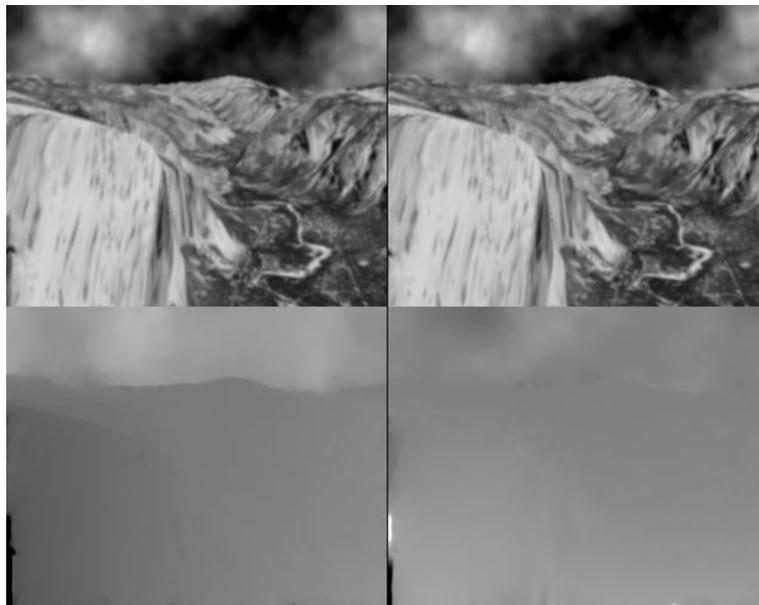


Figura 6.11: Cálculo del flujo óptico para la secuencia de Yosemite con  $\alpha = 0.6$ ,  $s = 0.1$ ,  $\sigma_0 = 5$ ,  $\sigma_n = 1$ , y  $\eta = 0.95$ .

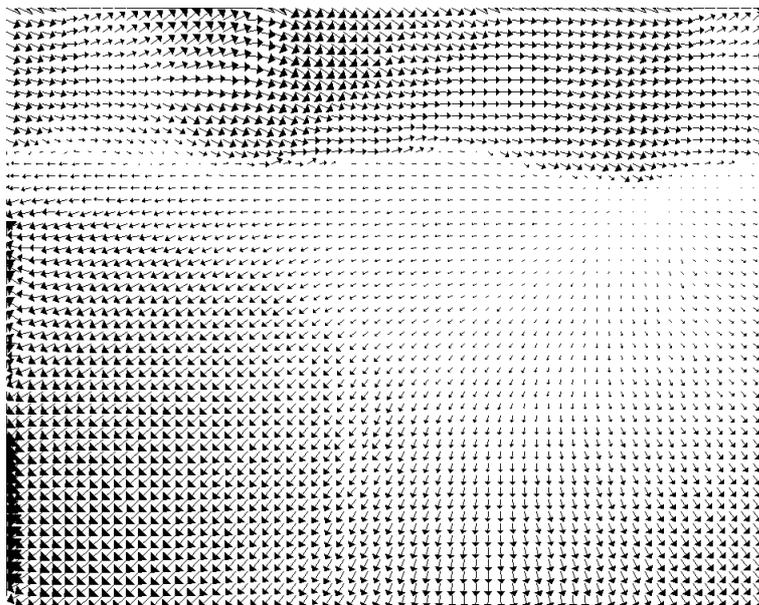


Figura 6.12: Vectores del flujo óptico para la figura 6.11.

Tabla 6.3: Errores para la secuencia de Yosemite utilizando diferentes parámetros del algoritmo

regulariz. $\alpha$	escal.inic. $\sigma_0$	fact.isotr. $s$	ratio descenso $\eta$	error angul.	desv. estándar.
0.4	5	0.1	0.90	5.61°	7.46°
0.5	"	"	"	5.57°	7.41°
0.6	"	"	"	5.55°	7.37°
0.7	"	"	"	5.56°	7.33°
1.0	"	"	"	5.69°	7.24°
0.6	1	0.1	0.90	16.83°	15.23°
"	2.5	"	"	5.92°	7.31°
"	5	"	"	5.55°	7.37°
"	10	"	"	5.54°	7.37°
"	15	"	"	5.81°	8.45°
0.6	5	0.01	0.90	5.70°	7.92
"	"	0.1	"	5.55°	7.37°
"	"	0.2	"	5.70°	7.31°
"	"	0.5	"	6.38°	8.14°
"	"	0.8	"	7.31°	9.76°
"	"	0.9	"	7.64°	10.37°
"	"	0.99	"	8.04°	11.21°
0.6	5	0.1	0.50	7.25°	7.58°
"	"	"	0.70	6.14°	7.36°
"	"	"	0.80	5.75°	7.33°
"	"	"	0.95	5.53°	7.40°
"	"	"	0.99	5.56°	7.45°

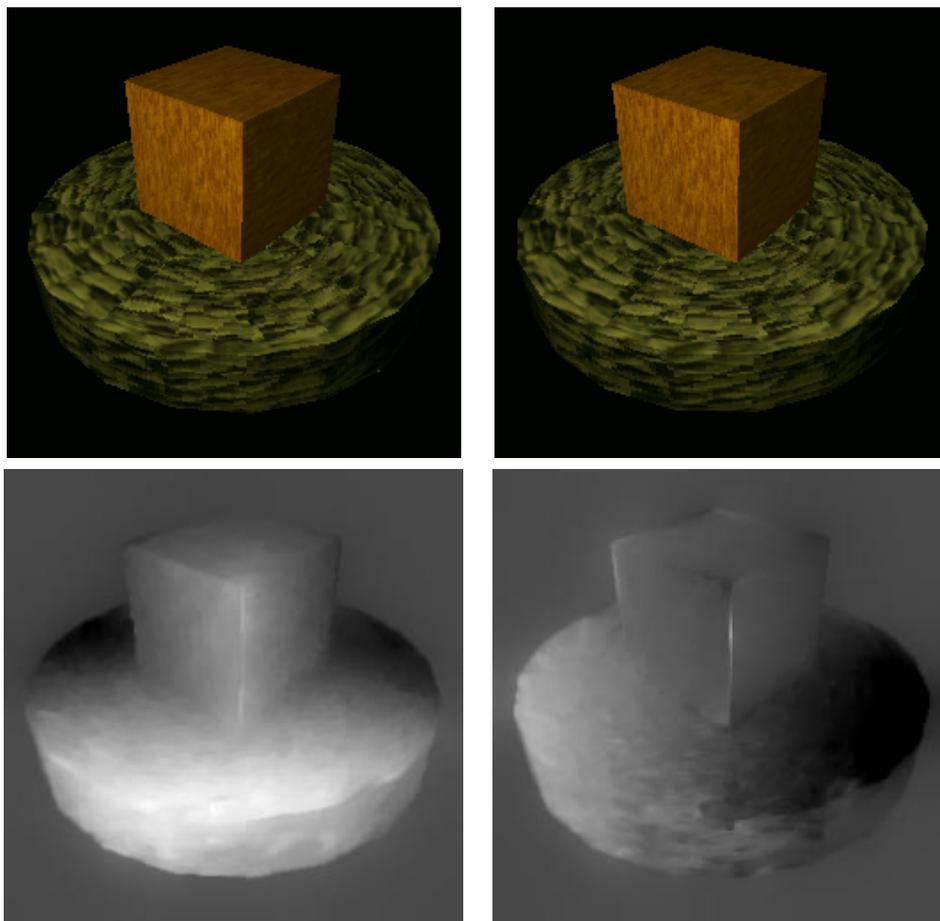


Figura 6.13: Cálculo del flujo óptico para la secuencia del cubo (frames 3 y 5) con  $\alpha = 1.0$ ,  $s = 0.5$  y  $zoom = 3$ .

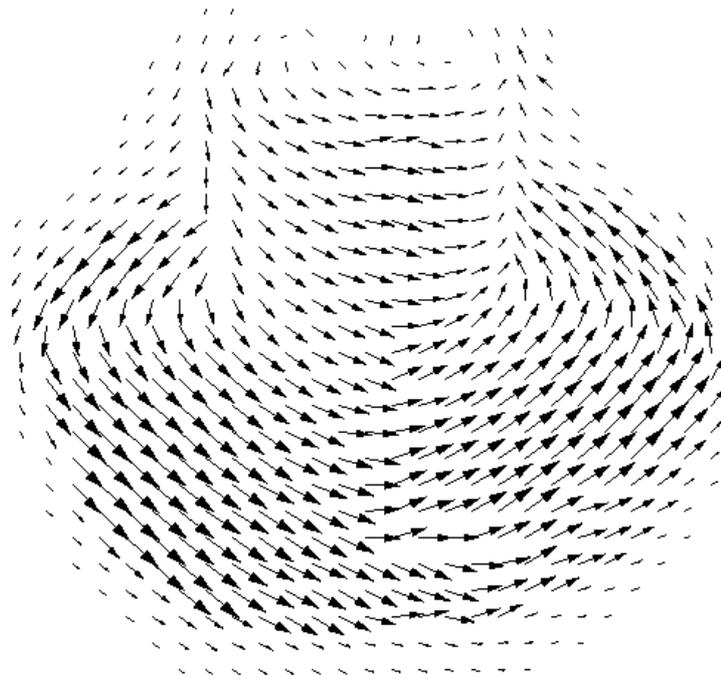


Figura 6.14: Vectores del flujo óptico para la figura 6.13.

# Capítulo 7

## Cálculo de la carta de disparidad

### 7.1 Introducción

En este trabajo presentamos un método para la estimación de mapas de disparidad densos basado en una aproximación por energía. En este método trabajamos con dos imágenes estéreo para las cuales proponemos un método que preserva las discontinuidades producidas por los contornos de la imagen. Primero derivamos una expresión simplificada de la disparidad que nos permite estimarla fácilmente a partir de un par estéreo de imágenes utilizando un método de minimización por energía. Suponemos que la geometría epipolar es conocida e introducimos esta información en el modelo de energía. Se preservan las discontinuidades a través de un término de regularización basado en el operador de Nagel-Enkelmann. Deducimos las ecuaciones de Euler-Lagrange para el funcional de energía y aproximamos la solución de la EDP resultante utilizando un método de descenso por gradiente. Con el fin de evitar que el proceso quede atrapado en mínimos locales irrelevantes a través de las iteraciones, utilizamos una estrategia de enfoque basado en la técnica de multiescala lineal y una estrategia piramidal. Demostramos la existencia y unicidad del sistema de EDP parabólica en el apéndice B. Presentamos algunos resultados sobre imágenes reales y sintéticas que muestran la capacidad de este método de EDP y multiescala.

Presentamos una aproximación variacional para el cálculo de mapas de disparidad densos a partir de dos imágenes estéreo *calibradas de forma débil*. Para resolver este problema, primero hacemos uso de la *matriz fundamental* (ver Faugeras 1993 [Fau93]) para derivar las ecuaciones que relacionan dos pixels correspondientes en las dos vistas, y luego combinamos herramientas de regularización y análisis multiescala para estimar iterativa y jerárquicamente el mapa de disparidad. La solución que se obtiene a una escala grande se utiliza para restringir la búsqueda en escalas más finas. Minimizamos un término de energía que tiene en cuenta las restricción de las líneas epipolares así como la información de los contornos a través de un término de

regularización apropiado.

El método que presentamos aquí, es un método basado en energía, y se puede presentar como una generalización del método presentado en [RD96] para el caso de calibración débil, con las siguientes originalidades:

- Como ya se ha dicho, el método presentado en este artículo considera un sistema estéreo calibrado de forma débil, (el sistema no está calibrado como en [RD96]), y solo se conoce la matriz fundamental. Los parámetros intrínsecos y extrínsecos no se conocen.
- Consecuentemente, se calcula una clase de profundidad proyectiva relacionada con la bien conocida medida de disparidad. En [RD96] los autores estiman directamente la información de profundidad. No tenemos que tratar con ningún proceso intermedio como el de rectificación [AH88, PKG99, RMC97, ZF94]. La información de la profundidad proyectiva se expresa directamente como una función  $m \rightarrow D(m)$  del punto de la imagen.
- Este método aborda el problema de determinar con precisión el mapa de disparidad denso mientras que suaviza y regulariza este mapa de disparidad a lo largo de los contornos de la imagen de niveles de grises e inhibe el suavizado a través de las discontinuidades de la imagen. En [RD96] el proceso de regularización se lleva a cabo con el fin de inhibir el suavizado del mapa de profundidad a través de las discontinuidades de la profundidad. Por otro lado, en  $S(Z)$  la conservación de las discontinuidades de la profundidad  $Z$  se obtiene por un operador de difusión no lineal asociado a la función  $\Phi$ , mientras que en nuestro método, la conservación de las discontinuidades en el mapa de disparidad se obtiene utilizando un operador lineal anisótropo que nos permite desarrollar discontinuidades en el mapa de disparidad a través de los bordes de la imagen  $I_1$ . Esto se logra considerando el término de regularización de Nagel-Enkelmann que ya ha demostrado ser muy útil en la estimación del flujo óptico [AWS99, NE86].
- Otro punto importante está relacionado con la forma en que se procede para evitar converger a mínimos locales irrelevantes. Se utiliza una estrategia de enfoque multiescala lineal, en el que se introduce nuestro método, al igual que se ha aplicado de forma exitosa en la estimación del flujo óptico [AWS99, AWS00]. Esto también relaciona nuestro trabajo con métodos estéreo que utilizan análisis multiescala lineales tales como [JM92, Lin98, ND95, NMN<sup>+</sup>99].
- Finalmente, con el objetivo de mostrar que nuestro método lleva a un concepto matemáticamente correcto, demostramos la existencia y unicidad de la solución

de la ecuación parabólica que gobierna nuestro método.

El trabajo está organizado de la siguiente manera: En la sección 7.2, presentamos el formalismo del proceso de puesta en correspondencia y comentamos todas las características relevantes de nuestro modelo. En la sección 7.3 comentamos los aspectos numéricos del algoritmo, y en la sección 7.4 presentamos algunos resultados experimentales.

## 7.2 Modelo propuesto

### 7.2.1 El funcional de energía

Con el fin de estimar  $\lambda(x, y)$  se puede proceder de manera clásica e intentar recuperar esta información importante utilizando un esquema simple de correlación. Desafortunadamente, esta solución no nos dará una solución correcta y precisa, en particular en las regiones donde el mapa de disparidad puede presentar algunas discontinuidades, como suele ser el caso próximo a los contornos de los objetos de la imagen. Es bien conocido que el mapa de disparidad que se obtiene utilizando este método clásico, tiende a suavizar mucho a través de los contornos de las imágenes. La idea que nos gustaría formalizar y desarrollar aquí es la de estimar la función  $\lambda(x, y)$  tal que sea suave solo a lo largo de los contornos de la imagen y no a través de ellos. Esto nos lleva a considerar la minimización del siguiente funcional de energía:

$$\begin{aligned}
 E(\lambda) &= \int_{\Omega} (I_1(x, y) - I_2(x + u(\lambda(x, y)), y + v(\lambda(x, y))))^2 dx dy \\
 &+ C \int_{\Omega} \Phi(\nabla I_1, \nabla \lambda) dx dy
 \end{aligned}
 \tag{7.2.1}$$

donde  $\Omega$  es el dominio de la imagen,  $C$  es una constante positiva, y  $\Phi(\nabla I_1, \nabla \lambda)$  determina el término de regularización. Se han estudiado distintos tipos de términos de regularización tal y como se ilustra en la tabla 7.1.

En todos estos modelos excepto en la regularización de Tikhonov, la idea subyacente es la de preservar las discontinuidades de la función en la regularización. Sin embargo, solo el modelo de Nagel y Enkelmann, propuesto en el contexto de la teoría de la estimación del flujo óptico [Nag83, NE86], procede a regularizar la solución a través de la información de los niveles de grises. Utilizan una matriz de proyección regularizada en la dirección perpendicular al  $\nabla I_1$ :

$$\mathbf{D}(\nabla I_1) = \frac{1}{|\nabla I_1|^2 + 2\nu^2} \left\{ \begin{bmatrix} \frac{\partial I_1}{\partial y} \\ -\frac{\partial I_1}{\partial x} \end{bmatrix} \begin{bmatrix} \frac{\partial I_1}{\partial y} \\ -\frac{\partial I_1}{\partial x} \end{bmatrix}^t + \nu^2 \mathbf{Id} \right\}. \quad (7.2.2)$$

En esta formulación,  $\mathbf{Id}$  denota la matriz identidad. Ya que estamos interesados en estimar el mapa de disparidad discontinuo de acuerdo a la información de los niveles de grises, centraremos nuestra atención en el método de Nagel y Enkelmann, primero, debido a su simplicidad, ya que el operador diferencial de segundo orden es lineal, y segundo, ya que este método ha demostrado su eficacia numerosas veces en el contexto de la estimación del flujo óptico [AWS99, AWS00, BFB94, Enk88, Nag83, Nag87, NE86, Sch91a, Sny91].

Mencionamos aquí la relación que existe entre la aproximación que proponemos en este artículo y el propuesto en [RD96] dentro del marco estéreo. Una aplicación sencilla de la idea en [RD96] resulta de seleccionar un término de regularización que tiene en cuenta las discontinuidades dentro del mapa de disparidad. Selecciona un término de regularización como  $\Phi(|\nabla\lambda|)$  donde  $\Phi$  es una función especial diseñada para preservar las discontinuidades (como la propuesta por Charbonnier et al. [CBFAB94]; ver también [DF95]). Tal y como ya habíamos mencionado, esta elección no tiene en cuenta la información de los contornos de la imagen que puede ser relevante en nuestra aplicación. Otra solución interesante podría ser la de considerar un término de regularización  $\Phi((\nabla\lambda)^T \mathbf{D}(\nabla I_1) \nabla\lambda)$  con la misma función  $\Phi$ .

## 7.2.2 Minimizando la energía

Teniendo en cuenta la pararmetrización del flujo visto en el capítulo 3 sobre geometría epipolar, y con el fin de minimizar el funcional de energía, resolvemos su ecuación de Euler-Lagrange asociada,

Tabla 7.1: Algunas regularizaciones utilizadas en proceso de imágenes y visión por ordenador.

Authors	$\Phi(\nabla I_1, \nabla\lambda)$
Tikhonov [TA77]	$ \nabla\lambda ^2$
Perona y Malik [PM90]	$-\frac{k^2}{2} e^{-( \nabla\lambda /k)^2}$
Geman y Reynolds [GR92]	$\frac{( \nabla\lambda /k)^2}{1+( \nabla\lambda /k)^2}$
Rudin et al. [ROF92]	$ \nabla\lambda $
Charbonnier et al. [CBFAB94]	$k^2 \sqrt{1 + ( \nabla\lambda /k)^2}$
Nagel y Enkelmann [NE86]	$(\nabla\lambda)^T D(\nabla I_1) \nabla\lambda$

$$C \operatorname{div} (\mathbf{D} (\nabla I_1^\sigma) \nabla \lambda_\sigma) + (I_1(\mathbf{x}) - I_2^{\lambda_\sigma}(\mathbf{x})) \frac{a \left( \frac{\partial I_2^\sigma}{\partial y} \right)^{\lambda_\sigma}(\mathbf{x}) - b \left( \frac{\partial I_2^\sigma}{\partial x} \right)^{\lambda_\sigma}(\mathbf{x})}{\sqrt{a^2 + b^2}} = 0, \quad (7.2.3)$$

donde

$$\mathbf{x} = (x, y)^t, \quad (7.2.4)$$

$$I_2^{\lambda_\sigma}(\mathbf{x}) = I_2^\sigma(\mathbf{x} + \vec{h}(\lambda_\sigma(\mathbf{x}))), \quad (7.2.5)$$

$$\left( \frac{\partial I_2^\sigma}{\partial x} \right)^{\lambda_\sigma}(\mathbf{x}) = \frac{\partial I_2^\sigma}{\partial x}(\mathbf{x} + \vec{h}(\lambda_\sigma(\mathbf{x}))), \quad (7.2.6)$$

$$\left( \frac{\partial I_2^\sigma}{\partial y} \right)^{\lambda_\sigma}(\mathbf{x}) = \frac{\partial I_2^\sigma}{\partial y}(\mathbf{x} + \vec{h}(\lambda_\sigma(\mathbf{x}))). \quad (7.2.7)$$

Obtenemos la solución a la ecuación anterior calculando el estado asintótico ( $t \rightarrow \infty$ ) de la ecuación parabólica

$$\frac{\partial \lambda_\sigma}{\partial t} = C \operatorname{div} (\mathbf{D} (\nabla I_1^\sigma) \nabla \lambda_\sigma) + (I_1^\sigma(\mathbf{x}) - I_2^{\lambda_\sigma}(\mathbf{x})) \frac{a \left( \frac{\partial I_2^\sigma}{\partial y} \right)^{\lambda_\sigma}(\mathbf{x}) - b \left( \frac{\partial I_2^\sigma}{\partial x} \right)^{\lambda_\sigma}(\mathbf{x})}{\sqrt{a^2 + b^2}}. \quad (7.2.8)$$

Observamos que en este método de difusión-reacción la matriz  $\mathbf{D}(\nabla I_1^\sigma)$  juega un papel de tensor de difusión. Sus autovectores son  $v_1 = \nabla I_1^\sigma$  y  $v_2 = \nabla I_1^{\sigma \perp}$  y los autovalores correspondientes vienen dados por

$$\lambda_1(|\nabla I_1^\sigma|) = \frac{\nu^2}{|\nabla I_1^\sigma|^2 + 2\nu^2}, \quad (7.2.9)$$

$$\lambda_2(|\nabla I_1^\sigma|) = \frac{|\nabla I_1^\sigma|^2 + \nu^2}{|\nabla I_1^\sigma|^2 + 2\nu^2}. \quad (7.2.10)$$

En el interior de los objetos tenemos que  $|\nabla I_1^\sigma| \rightarrow 0$  y por lo tanto  $\lambda_1 \rightarrow 1/2$  y  $\lambda_2 \rightarrow 1/2$ . En bordes ideales donde  $|\nabla I_1^\sigma| \rightarrow \infty$  obtenemos que  $\lambda_1 \rightarrow 0$  y  $\lambda_2 \rightarrow 1$ . Por

lo tanto, tenemos un comportamiento isótropo dentro de las regiones, y en el contorno de las imágenes el proceso suaviza de forma anisótropa a lo largo de los bordes. Este comportamiento es similar al filtrado por difusión anisótropa en el realce de contornos [Wei96, Wei98a], y también está muy próximo en esencia a la evolución por curvatura media considerada en [ALM92].

### 7.2.3 Modelo para el cálculo de la carta de disparidad con imágenes en color

Para el cálculo de la carta de disparidad con imágenes en color tenemos que tener en cuenta que las imágenes están formadas por tres canales de intensidades. En nuestro caso, utilizaremos canales RGB – canal para el color rojo, otro para el verde y otro para el azul –. Vamos a incorporar esta información al modelo de energía anterior. La energía que proponemos es como sigue:

$$\begin{aligned}
E(\lambda) &= \sum_{i=0}^3 \int_{\Omega} \left( I_{1,i}(x, y) - I_{2,i}(\mathbf{x} + \vec{h}(\lambda(\mathbf{x}))) \right)^2 d\mathbf{x} \\
&+ C \int_{\Omega} \Phi(\nabla I_{1,max}, \nabla \lambda) d\mathbf{x}
\end{aligned} \tag{7.2.11}$$

En el término de ligadura incorporamos un sumatorio para tener en cuenta la información de los tres canales, obteniendo la diferencia entre las intensidades de los correspondientes canales en las dos imágenes. En el término de regularización obtenemos una estimación del gradiente de la imagen como el máximo en norma de los gradientes de cada canal.

$$\nabla I_{1,max} = \max\{\nabla I_{1,R}, \nabla I_{1,G}, \nabla I_{1,B}\}$$

Minimizando obtenemos las ecuaciones de Euler–Lagrange y aplicando un método de descenso por gradiente, llegamos a la ecuación en EDP

$$\begin{aligned}
\frac{\partial \lambda}{\partial t} &= C \operatorname{div} (\mathbf{D} (\nabla I_{1,max}^\sigma) \nabla \lambda) \\
&+ \sum_{i=0}^3 \left( I_{1,i}^\sigma - I_{2,i}^\sigma(\mathbf{x} + \vec{h}(\lambda)) \right) \frac{a \left( \frac{\partial I_{2,i}^\sigma}{\partial y} \right)^\lambda (\mathbf{x}) - b \left( \frac{\partial I_{2,i}^\sigma}{\partial x} \right)^\lambda (\mathbf{x})}{\sqrt{a^2 + b^2}}
\end{aligned} \tag{7.2.12}$$

### 7.3 Esquema numérico

Ahora vamos a describir un algoritmo eficiente para nuestro método. Discretizamos la ecuación parabólica (7.2.8) por diferencias finitas<sup>1</sup> (ver [MM94] para una introducción a este tema). Todas las derivadas espaciales se aproximan por diferencias centradas, y para la discretización en la dirección de  $t$  utilizamos un esquema lineal implícito.

Sea  $\mathbf{D}(\nabla I_1^\sigma) = \begin{pmatrix} d & f \\ f & e \end{pmatrix}$ . Entonces nuestro esquema lineal implícito tiene la estructura

$$\begin{aligned}
\frac{\lambda_{i,j}^{k+1} - \lambda_{i,j}^k}{\tau} = C & \left( \frac{d_{i+1,j} + d_{i,j}}{2} \frac{\lambda_{i+1,j}^{k+1} - \lambda_{i,j}^{k+1}}{h_1^2} + \frac{d_{i-1,j} + d_{i,j}}{2} \frac{\lambda_{i-1,j}^{k+1} - \lambda_{i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{f_{i,j+1} + f_{i,j}}{2} \frac{\lambda_{i,j+1}^{k+1} - \lambda_{i,j}^{k+1}}{h_2^2} + \frac{f_{i,j-1} + f_{i,j}}{2} \frac{\lambda_{i,j-1}^{k+1} - \lambda_{i,j}^{k+1}}{h_2^2} \\
& + \frac{e_{i+1,j+1} + e_{i,j}}{2} \frac{\lambda_{i+1,j+1}^{k+1} - \lambda_{i,j}^{k+1}}{2h_1h_2} + \frac{e_{i-1,j-1} + e_{i,j}}{2} \frac{\lambda_{i-1,j-1}^{k+1} - \lambda_{i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{e_{i+1,j-1} + e_{i,j}}{2} \frac{\lambda_{i+1,j-1}^{k+1} - \lambda_{i,j}^{k+1}}{2h_1h_2} - \frac{e_{i-1,j+1} + e_{i,j}}{2} \frac{\lambda_{i-1,j+1}^{k+1} - \lambda_{i,j}^{k+1}}{2h_1h_2} \right) \\
& + \left( I_1(\mathbf{x}_{i,j}) - I_2^{\lambda_{i,j}^k}(\mathbf{x}_{i,j}) - \frac{\lambda_{i,j}^{k+1} - \lambda_{i,j}^k}{\sqrt{a_{i,j}^2 + b_{i,j}^2}} \left( a_{i,j} I_{2,y}^{\lambda_{i,j}^k}(\mathbf{x}_{i,j}) - b_{i,j} I_{2,x}^{\lambda_{i,j}^k}(\mathbf{x}_{i,j}) \right) \right) \\
& \cdot \frac{a_{i,j} I_{2,y}^{\lambda_{i,j}^k}(\mathbf{x}_{i,j}) - b_{i,j} I_{2,x}^{\lambda_{i,j}^k}(\mathbf{x}_{i,j})}{\sqrt{a_{i,j}^2 + b_{i,j}^2}} \tag{7.3.1}
\end{aligned}$$

donde  $\tau$  es el paso en el tiempo,  $h_1$  y  $h_2$  denotan el tamaño del pixel en la dirección de  $x$  e  $y$ , respectivamente,  $\lambda_{i,j}^k$  aproxima  $\lambda_\sigma$  en algún punto de la malla  $\mathbf{x}_{i,j}$  en el instante

$k\tau$ , e  $I_{2,x}^{\lambda_{i,j}^k}(\mathbf{x}_{i,j})$  es una aproximación de  $\frac{\partial I_2^{\sigma,\lambda}}{\partial x}(\mathbf{x}_{i,j}, k\tau)$ . Calculamos valores fuera de los puntos por interpolación lineal, y resolvemos el sistema de ecuaciones lineal resultante iterativamente a través del algoritmo iterativo de Gauß–Seidel simétrico [AWS00].

<sup>1</sup>Ver apéndice C sobre discretización del gradiente y de la divergencia

El esquema lineal implícito ofrece la ventaja de poder utilizar intervalos de tiempo más grandes con el fin de acelerar la convergencia a un estado estable. Utilizamos como paso  $\tau = 10$  y detenemos el proceso después de 50 iteraciones.

## 7.4 Resultados experimentales

### 7.4.1 Imágenes sintéticas

#### Par estéreo del pasillo

En nuestro primer experimento utilizamos un par estéreo sintético clásico donde se conoce la disparidad exacta. En este caso las líneas epipolares son paralelas al eje horizontal y la matriz fundamental  $F$  está dada por

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

Calculamos el error como el promedio de los valores absolutos de la diferencia entre la disparidad calculada y la disparidad exacta. No tenemos en cuenta una franja de 15 pixels del contorno de la imagen para evitar su influencia en los cálculos de los errores.

Comparamos nuestro método con una versión mejorada de la técnica de correlación clásica, donde usamos algunas técnicas de interpolación para conseguir precisión sub-píxel así como una carta de disparidad del 100%. Con esta técnica de correlación encontramos que el mejor tamaño para la ventana es de  $23 \times 23$ . Para este tamaño de ventana el error es de 0.4978.

La tabla 7.2 muestra los resultados obtenidos utilizando nuestro método para diferentes configuraciones de parámetros, y en la figura 7.1 presentamos el mejor resultado que obtuvimos con nuestro método. Tiene un error medio de 0.2639. La tabla 7.2 muestra también que nuestro método es más robusto con respecto a la elección de los parámetros. En realidad, en la mayoría de los casos los resultados son mejores que para el método de correlación con el tamaño de ventana óptimo.

#### Par estéreo de imágenes con puntos aleatorios

En las próximas pruebas utilizamos imágenes de puntos aleatorios para probar el método. La disparidad que existe para el ejemplo de Julesz (figura 7.2) es, como se puede ver, una especie de  $\mathbf{e}$  que se desliza con un flujo constante de 2 con respecto al fondo de una imagen a la otra.

Tabla 7.2: Resultado de nuestro método aplicado al par estéreo de la figura 7.1.

$\sigma_0$	$\sigma_n$	$\alpha$	$\lambda_0$	$s$	$\eta$	Error medio
7	0.8	0.5	-3	0.15	0.95	0.26
10	0.8	0.5	-3	0.15	0.95	0.36
4	"	"	"	"	"	3.23
7	1.0	0.5	-3	0.15	0.95	0.27
"	0.6	"	"	"	"	0.28
7	0.8	0.6	-3	0.15	0.95	0.27
"	"	0.4	"	"	"	4.49
7	0.8	0.5	-1	0.15	0.95	0.28
"	"	"	-5	"	"	0.58
7	0.8	0.5	-3	0.10	0.95	0.30
"	"	"	"	0.20	"	0.30
7	0.8	0.5	-3	0.15	0.99	0.30
"	"	"	"	"	0.90	0.33

El resultado de la correlación es bastante exacto en gran parte de la imagen, aunque en la zona de los bordes se ven muchos grupos de manchas erróneas. Con nuestro método desaparecen prácticamente todas estas manchas, aunque se suaviza un poco más el resultado.

En la figura 7.3 se muestra otro ejemplo de imágenes de puntos aleatorios en las que la disparidad viene representada por una pirámide escalonada. Se repite el mismo efecto que para el ejemplo anterior en el que la correlación no es capaz de encontrar correspondencias en muchas zonas (más que en el anterior) y nuestro método encuentra una solución densa aunque un poco más regularizada.

## 7.4.2 Imágenes reales

### Par estéreo de Hervé

En el siguiente experimento utilizamos un par estéreo real obtenido a partir del web site del INRIA <http://www-sop.inria.fr/robotvis/demo>. En este caso las líneas epipolares no son paralelas al eje horizontal, y la matriz fundamental viene dada por

$$\mathbf{F} = \begin{bmatrix} 0.000000149 & 0.00000340 & -0.0026 \\ -0.00000718 & 0.00000109 & 0.048 \\ 0.0031 & -0.048 & 1 \end{bmatrix}$$

En la imagen 7.4 presentamos la disparidad calculada ( $\sqrt{u^2 + v^2}$ ) usando el método

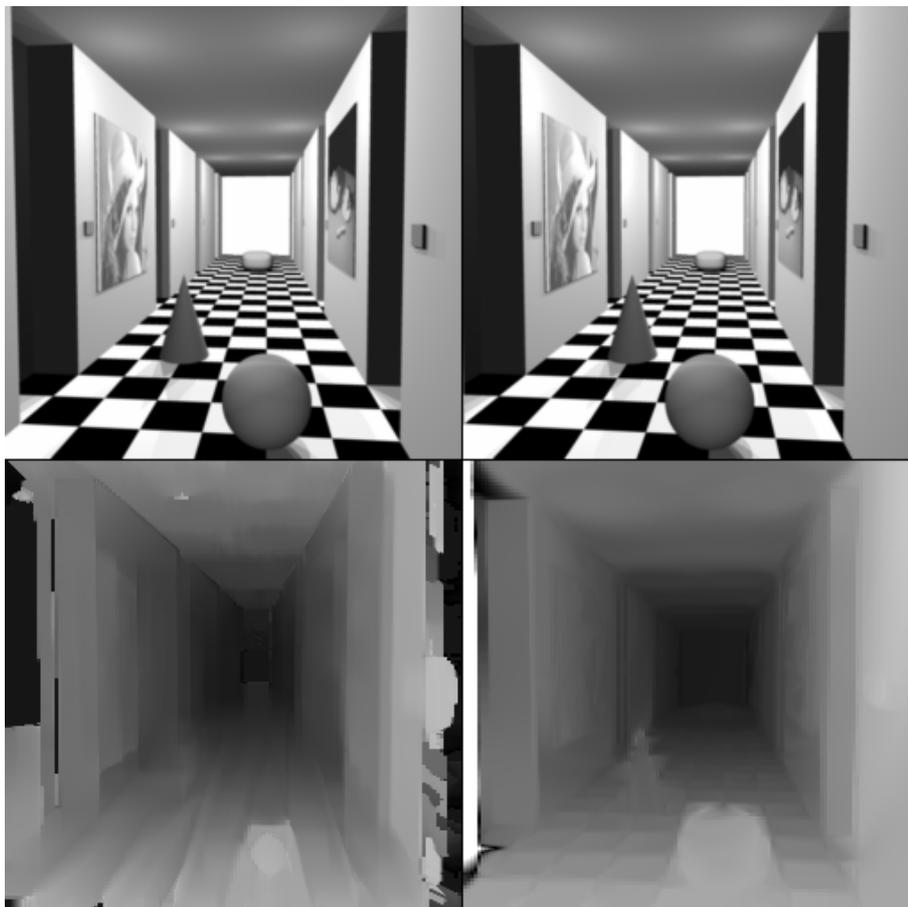


Figura 7.1: De izquierda a derecha y de arriba abajo: el par estéreo original, la carta de disparidad calculada utilizando una ventana de correlación de  $23 \times 23$ , y la carta de disparidad obtenido con nuestro método.

de correlación así como nuestro método con la técnica de correlación como inicialización. La figura 7.5 muestra algunas vistas de la reconstrucción 3-D de la cara en el par estéreo, haciendo uso de la carta de disparidad obtenido por nuestro método. La reconstrucción parece muy realista.

### Par estéreo de la Biblioteca del Inria

En la figura 7.6 mostramos un experimento con la biblioteca del INRIA en Sophia-Antipolis. En ésta se muestra en la parte superior el par estéreo y en la parte inferior el resultado que se obtiene de la correlación a la izquierda y aplicando nuestro método a la derecha. A diferencia con el ejemplo visto anteriormente, la geometría epipolar hace que sea más difícil apreciar a través de la carta de disparidad la profundidad real de la escena.

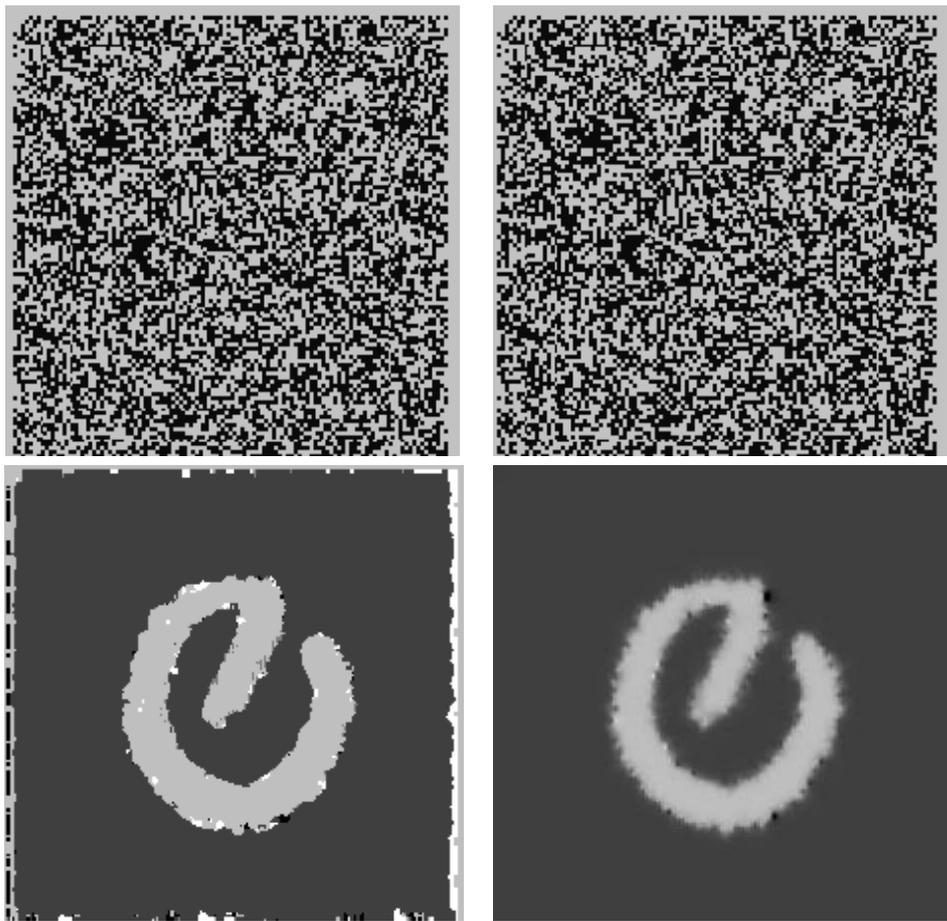


Figura 7.2: Arriba: El par estéreo original. Abajo izquierda: La carta de disparidad calculada utilizando una ventana de correlación de  $3 \times 3$ . Abajo derecha: Nuestro método ( $\alpha = 1$ ,  $s = 0.5$ ,  $zoom = 3$ ) empezando desde  $u \equiv v \equiv 0$ .

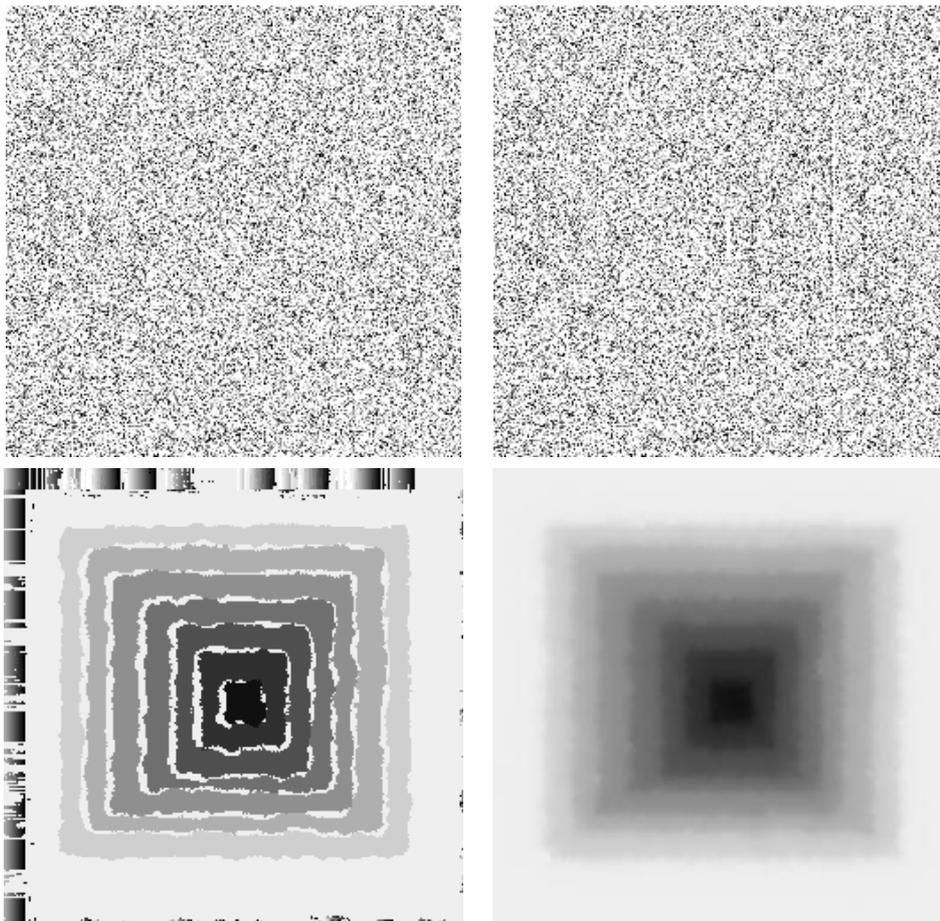


Figura 7.3: Arriba: El par estéreo original. Abajo izquierda: La carta de disparidad calculada utilizando una ventana de correlación de  $4 \times 4$ . Abajo derecha: Nuestro método ( $\alpha = 0.3$ ,  $s = 0.5$ ,  $zoom = 3$ ) empezando desde  $u \equiv v \equiv 0$ .

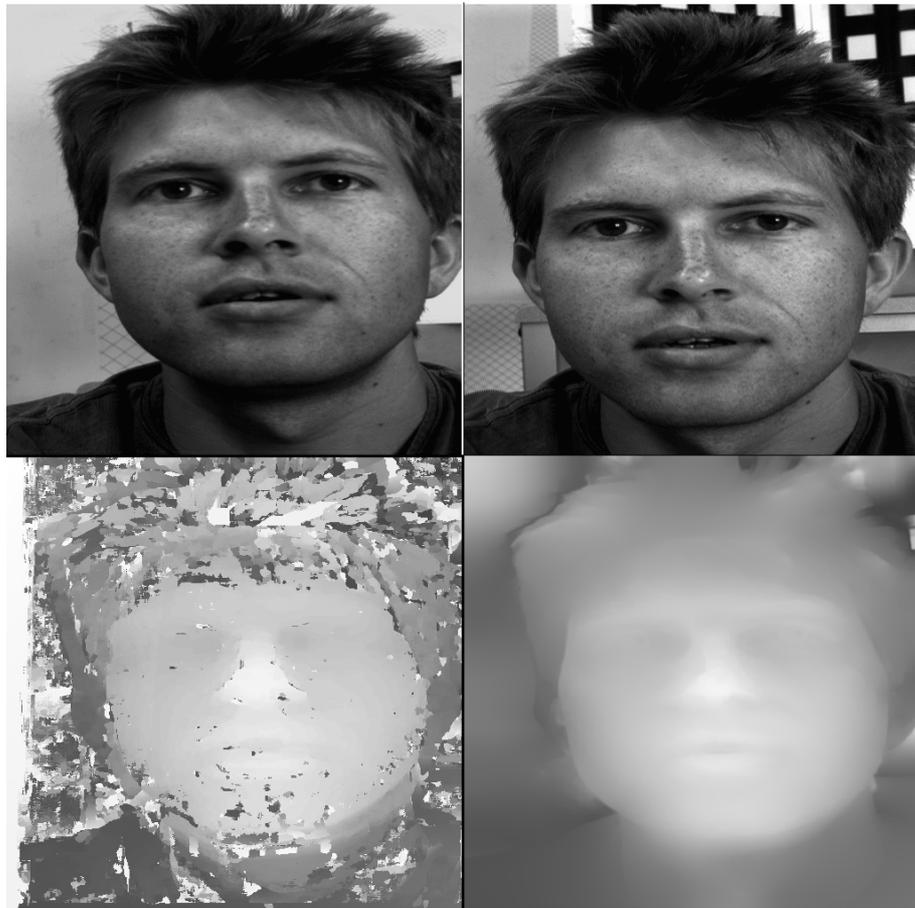


Figura 7.4: Arriba: El par estéreo original. Abajo izquierda: La carta de disparidad calculada utilizando una ventana de correlación de  $13 \times 13$ . Abajo derecha: Nuestro método ( $\sigma_0 = 7$ ,  $\sigma_n = 0.8$ ,  $\alpha = 1$ ,  $s = 0.5$ ,  $\eta = 0.95$ ), con el resultado de la correlación como inicialización.

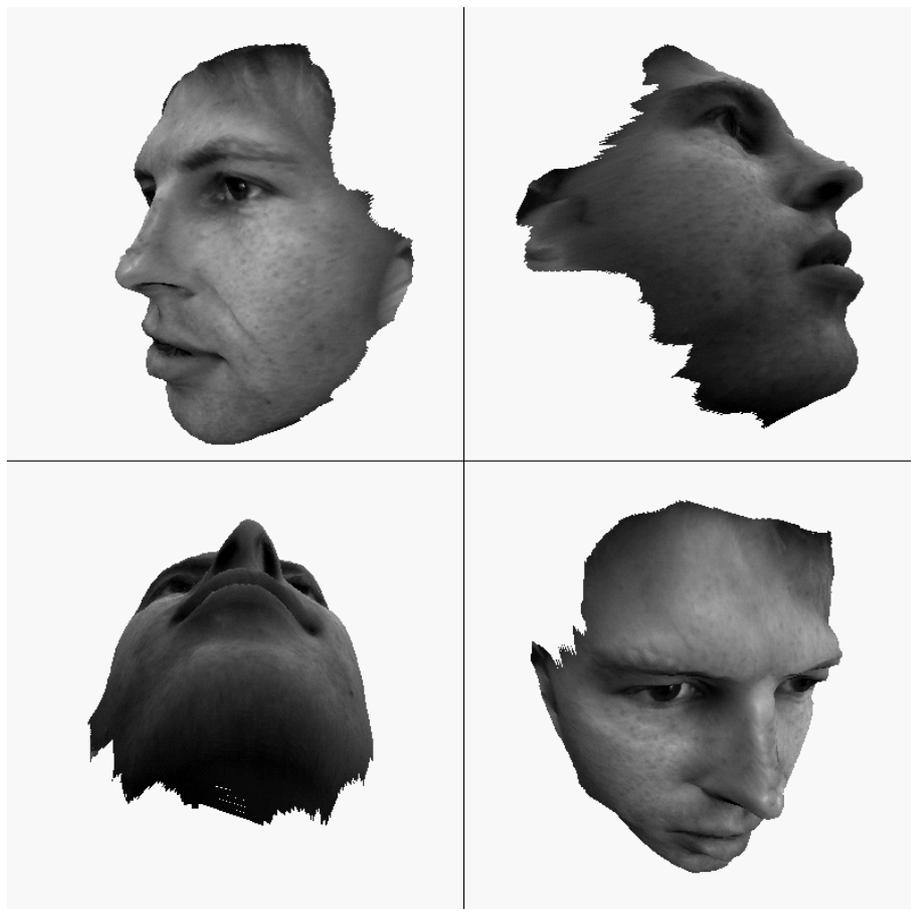


Figura 7.5: Cuatro vistas de la reconstrucción 3-D del par estéreo de la figura 7.4, utilizando la carta de disparidad obtenida por nuestro método.

Si nos fijamos en la ventana del edificio que está más a la izquierda, vemos que no se obtiene una buena disparidad. Esto se debe a que hay un cambio de brillo importante, haciendo que las intensidades en las dos imágenes no sean iguales, no cumpliéndose la suposición lambertiana.

En la figura 7.7 mostramos cuatro vistas de la reconstrucción de la biblioteca. Como vemos, se capta perfectamente la profundidad de los distintos objetos (coches, edificio, farola), además de conseguir que los ángulos que forman las paredes del edificio sean prácticamente perpendiculares.

### Par estéreo de Cyrille

En la imagen 7.8 presentamos un experimento con otro par estéreo del INRIA. También representa la cara de un ser humano, pero la geometría epipolar difiere bastante del ejemplo de Hervé. Sin embargo, como se puede ver en la figura 7.9, también en



Figura 7.6: Arriba: El par estéreo original. Abajo izquierda: La carta de disparidad calculada utilizando una ventana de correlación de  $13 \times 13$ . Abajo derecha: Nuestro método ( $\alpha = 0.3$ ,  $s = 0.5$ ,  $zoom = 3$ ) empezando desde  $u \equiv v \equiv 0$ .

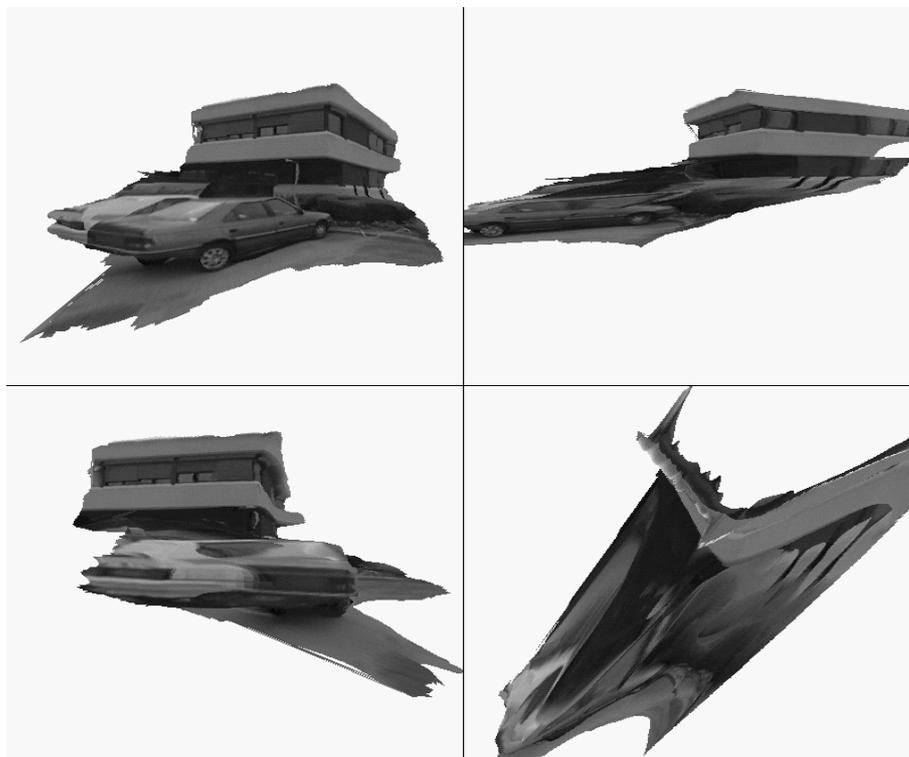


Figura 7.7: Cuatro vistas de la reconstrucción 3-D del par estéreo de la figura 7.6, utilizando la carta de disparidad de nuestro método.

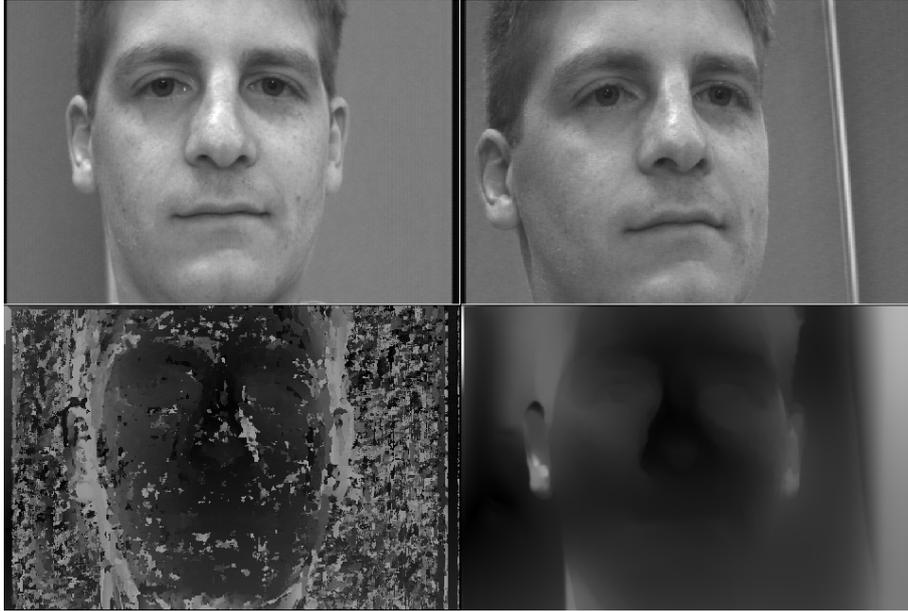


Figura 7.8: Arriba: El par estéreo original. Abajo izquierda: La carta de disparidad calculada utilizando una ventana de correlación de  $13 \times 13$ . Abajo derecha: Nuestro método ( $\sigma_0 = 7$ ,  $\sigma_n = 0.8$ ,  $\alpha = 1$ ,  $s = 0.5$ ,  $\eta = 0.95$ ), con el resultado de la correlación como inicialización.

este caso la reconstrucción 3-D es buena con nuestro método. Un pequeño artefacto se hace visible: debido a problemas de oclusiones no es posible reconstruir perfectamente la región entre el ojo izquierdo y la nariz. Aquí podemos ver como el término de regularización de Nagel–Enkelmann intenta propagar la información desde la vecindad con el fin de guardar las inconsistencias lo más pequeñas posibles. Como consecuencia, el ojo se extiende de alguna manera dentro de la zona de oclusión.

Nos gustaría señalar que hemos utilizado exactamente los mismos valores de parámetros en ambos pares estéreos reales. En el caso del par sintético – donde no hay ruido presente – tan solo hemos reducido los valores del parámetro de isotropía  $s$  y del parámetro de regularización  $\alpha$ . Esto permite mayores discontinuidades en la carta de disparidad. En pares estéreos reales se requiere mayor regularidad (p.e. mayor valor para  $s$  y  $\alpha$ ) debido a dos razones: Primero, hay presencia de ruido, y segundo, en situaciones reales la suposición lambertiana que establece que puntos correspondientes en ambas imágenes tienen igual nivel de brillo, no se cumple estrictamente. Lo que es realmente importante es que, incluso si el número de parámetros del algoritmo es alto, se pueden obtener buenos resultados en diferentes situaciones utilizando un conjunto de parámetros por defecto. Nuestros experimentos ilustran que esto se mantiene incluso para geometrías epipolares muy diferentes. Esto vuelve a confirmar la robustez del algoritmo que ya se había señalado en la tabla 7.2.

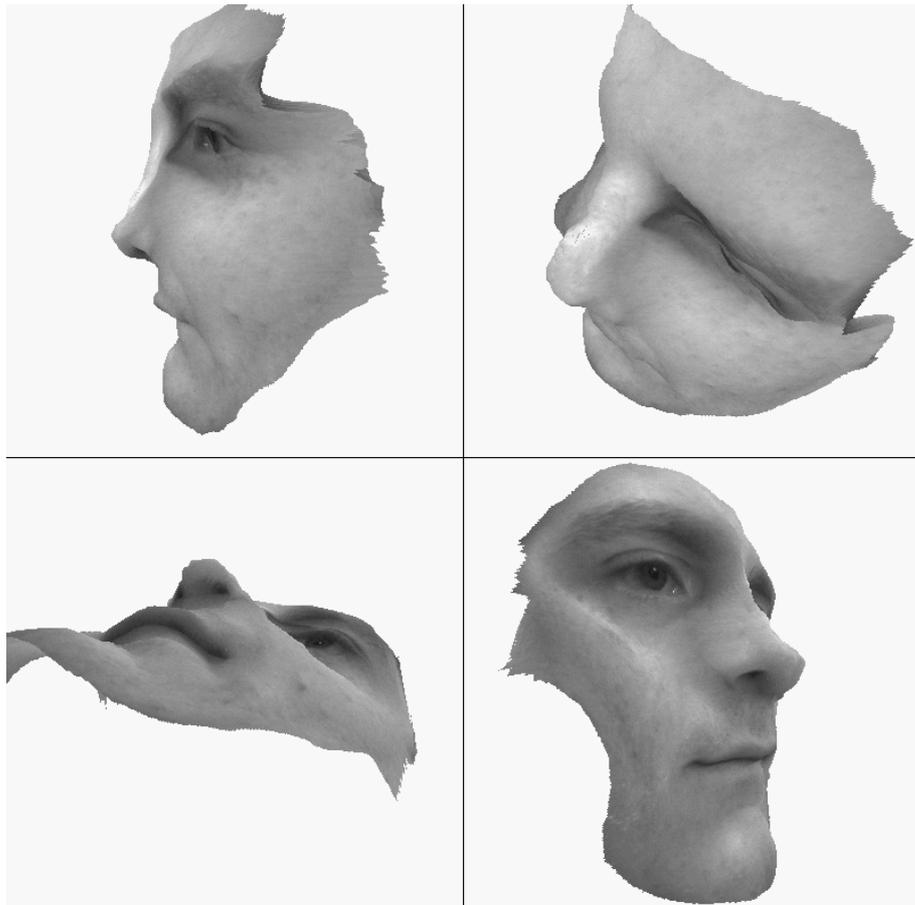


Figura 7.9: Cuatro vistas de la reconstrucción 3-D del par estéreo de la figura 7.8, utilizando la carta de disparidad de nuestro método.



Figura 7.10: Par estéreo de Julio (ver capítulo 4 para más detalles acerca del mismo).

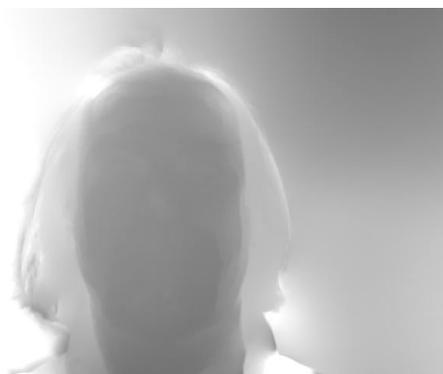


Figura 7.11: Disparidad para el par estéreo de Julio. Se ha elegido la estrategia multipiramidal con parámetros  $z = 4$ ,  $\alpha = 1$ ,  $s = 0.5$ ,  $T = 500$  y  $dt = 10$ .

### 7.4.3 Resultados experimentales para imágenes en color

En esta sección utilizamos pares estéreos en color para probar el método dado por (7.2.12). Utilizamos dos imágenes reales creadas por nosotros mismos tal y como explicamos en el capítulo 4.

#### Par estéreo de Julio

En la figura 7.10 y 7.11 mostramos el par estéreo de Julio y el mapa de disparidad que obtenemos al aplicar nuestro método.

Donde se obtiene mayor precisión es en la zona interior a la cara, ya que hay mayor contraste y más variación de niveles de grises. En la zona del fondo y del pelo la solución varía bastante de un lado a otro debido a que al ser zonas tan homogéneas, se propagan más fácilmente los errores.

La reconstrucción 3D (figura 7.12) muestra cuatro vistas de la cara de Julio desde distintos puntos de vista. Se capta bastante bien la profundidad de las distintas partes

de la cara.

### **Par estéreo de Javier**

El último ejemplo para imágenes en color es el par estéreo de Javier. Esta foto se tomó junto con las del par anterior, por lo tanto, la geometría epipolar es la misma.

La disparidad que se obtiene (figura 7.14) tiene las mismas características que las comentadas para el par de Julio.

La reconstrucción (figura 7.15) parece más fina que en el caso anterior. Se aprecia un pequeño error en los ojos debido probablemente al efecto del cambio de brillo de una imagen a la otra.



Figura 7.12: Distintas vistas de la reconstrucción tridimensional para el par estéreo de Julio.



Figura 7.13: Par estéreo de Javier (ver capítulo 4 para más detalles acerca del mismo).



Figura 7.14: Disparidad para el par estéreo de Javier. Se ha elegido la estrategia multipiramidal con parámetros  $z = 4$ ,  $\alpha = 1$ ,  $s = 0.5$ ,  $T = 500$  y  $dt = 10$ .

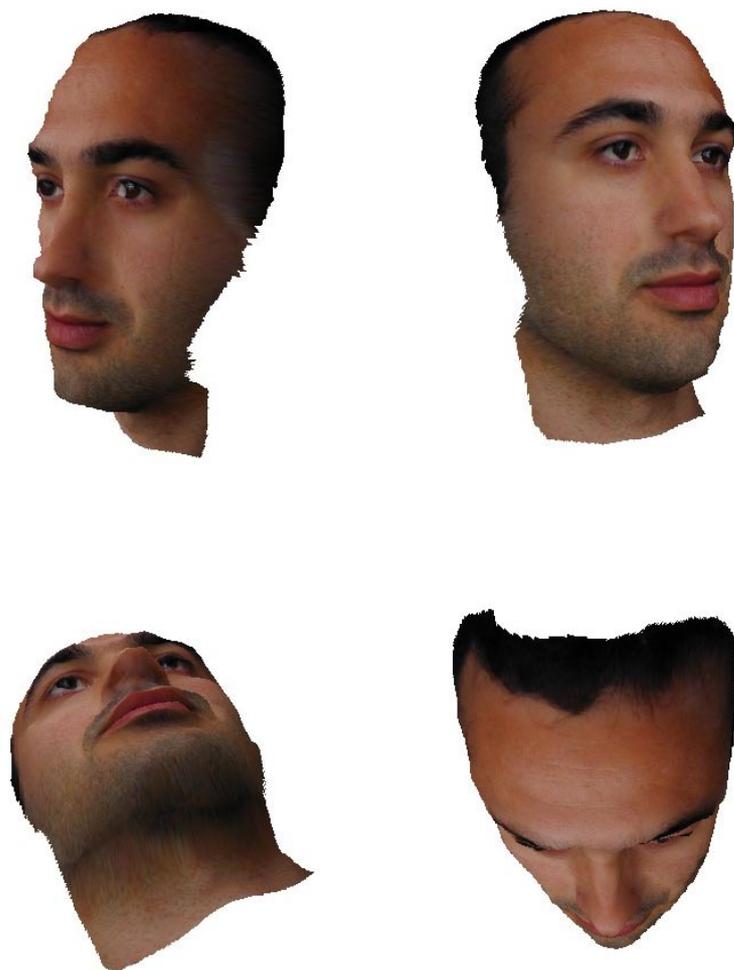


Figura 7.15: Distintas vistas de la reconstrucción tridimensional para el par estéreo de Javier.

## Parte III

# Introducción de la simetría en el cálculo de correspondencias

# Capítulo 8

## Introducción

Los métodos variacionales que hemos estudiado en los capítulos anteriores, no generan una solución simétrica, a pesar que el problema de la estimación del flujo óptico y del mapa de disparidad deberían ser independientes del orden en que se tomen las imágenes para el cálculo. Generalmente, esto no es cierto para la mayoría de los métodos variacionales que existen actualmente.

El problema de la simetría se ha tratado recientemente en varios trabajos relacionados con imágenes médicas como pueden ser [Chr99, MY99, CR00]. En el primero se propone un método en el que se introduce un término de simetría, en el que se minimiza la diferencia entre las funciones inversas de correspondencias, y una restricción difeomórfica, en el que se hace uso de modelos de la mecánica continua para forzar que las soluciones sean inversas una de la otra. En el último se exponen algunas causas por las que las energías del tipo que nosotros manejamos son asimétricas. Las razones básicas por las que no se cumple la simetría son:

1. **Puede no existir  $h_1^{-1}$ :** No se puede asegurar que  $h_1$  sea biyectiva y, por lo general, debido al tipo de imágenes que estamos considerando, para ciertas regiones como pueden ser las oclusiones no existe  $h_1^{-1}$ .
2. **No se preserva el orden de las energías:** Se puede tener que  $E_1(h_1) < E_2(h_2)$  pero puede que no se cumpla que  $E_1(h_2^{-1}) < E_2(h_1^{-1})$ .
3. **El espacio de las transformaciones no es estable:** El espacio de las transformaciones permitidas  $\mathcal{T}$  puede no ser un grupo y, por lo tanto, no estable por inversión, por ejemplo, podemos tener  $T \in \mathcal{T}$  y  $T^{-1} \notin \mathcal{T}$ .
4. **Mínimos locales:** El proceso puede caer en mínimos locales distintos dependiendo del sentido en que se tomen las imágenes.

En la formulación de la energía no se impone ninguna restricción para que el flujo estimado sea simétrico.

Los métodos que estamos considerando están pensados para imágenes tomadas del mundo real, en donde la complejidad de las escenas puede ser bastante alta. Los problemas que se derivan de estas son varios – tal y como se discutió en la sección 2.2 – pero, por lo general, los que más repercuten sobre la estimación simétrica son las oclusiones y las apariciones y desapariciones de objetos. Debido a estos problemas, no podemos hablar de una simetría perfecta entre imágenes, sino de simetrías por regiones. En otro tipo de imágenes, p.e. en algún tipo de imágenes médicas, esto no es un problema, ya que el efecto de las oclusiones es prácticamente despreciable. Sin embargo, en la mayoría de las imágenes con que tratamos, el problema de las oclusiones suele ser importante.

Por esto es necesario introducir algún mecanismo que sea capaz de discernir entre regiones existentes en ambas imágenes de regiones que aparecen en una imagen y desaparecen en la otra.

En esta parte vamos a modificar los modelos descritos anteriormente para incorporar la simetría en el cálculo de las correspondencias. Empezaremos por dar una idea general de cómo vamos a modificar las energías propuestas anteriormente, para luego detallar para el problema del flujo óptico y del cálculo de mapas de disparidad cómo se expresarían sus correspondientes energías y cuáles serían las ecuaciones en EDP para las que se proponen soluciones numéricas.

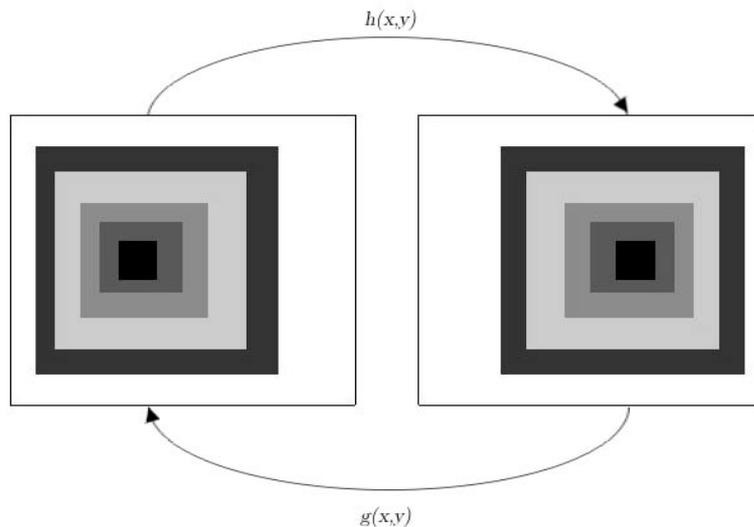


Figura 8.1: Las funciones de correspondencia  $\vec{h}(x, y)$  y  $\vec{g}(x, y)$  deben ser inversas la una de la otra. Debe existir una correspondencia uno a uno entre las dos funciones excepto en aquellas zonas donde se producen oclusiones.

En la figura 8.1 se muestra de forma intuitiva el problema de la simetría. En nuestras pruebas prácticas hemos podido comprobar que difícilmente  $\vec{h}(x, y) = \vec{g}^{-1}(x, y)$ .

En nuestros modelos incorporamos un nuevo término que minimiza la diferencia entre las funciones de correspondencia directa e inversa, así como tiene en cuenta los puntos para los que no existe una buena correspondencia – aquellos que pertenecen a regiones de oclusión.

Como veremos en las pruebas prácticas, el efecto de introducir un término de simetría mejora considerablemente aquellas regiones donde los métodos no simétricos producían grandes diferencias entre  $\vec{h}(x, y)$  y  $\vec{g}(x, y)$ . Generalmente, estas son las regiones homogéneas, donde es difícil estimar las correspondencias debido a que el término de ligadura no es capaz de discernir entre puntos de la región, o las regiones donde existe textura muy homogénea, donde es fácil que los métodos confundan una correspondencia con otra. En estas zonas, se producían soluciones más suaves que las obtenidas con los métodos anteriores. En las regiones donde se obtenían valores precisos para los métodos no simétricos (p.e. zonas de los contornos) se siguen obteniendo en los métodos simétricos.

## 8.1 Modelo de energía para las simetrías

En esta sección vamos a proponer, de forma general, una energía que introduce las simetrías. Básicamente, las energías propuestas hasta ahora se pueden descomponer en dos términos: 1) De acoplamiento, que se encarga de aproximar los puntos de las dos imágenes, y 2) de regularización, que es el encargado de que la solución sea suave<sup>1</sup>.

$$E(\vec{h}_1) = \int_{\Omega} \left( I_1(\mathbf{x}) - I_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right)^2 + \alpha \int_{\Omega} \Phi \left( \mathbf{D}(\nabla I_1), \vec{h}_1 \right)$$

$$E(\vec{h}_2) = \int_{\Omega} \left( I_2(\mathbf{x}) - I_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right)^2 + \alpha \int_{\Omega} \Phi \left( \mathbf{D}(\nabla I_2), \vec{h}_2 \right)$$

La primera ecuación calcula las correspondencias desde la imagen  $I_1$  hasta  $I_2$ , mientras que la segunda las calcula en sentido inverso. No existe ninguna restricción para que ambas estimaciones sean inversas la una de la otra. Lo que se suele hacer

---

<sup>1</sup>En esta ecuación la función  $\Phi$  depende de un tensor de difusión. Esta es la forma en que hemos utilizado las energías propuestas.

muchas veces es estimarlas en paralelo y descartar aquellas correspondencias que no tengan un mínimo de precisión determinado en los dos sentidos. Otra cosa que se suele hacer para que las soluciones se aproximen a las inversas, es ir calculando iterativamente una solución, luego calcular su inversa y utilizar esta para calcular la solución en el sentido inverso.

Una forma natural de incorporar la simetría en estas energías es a través de un nuevo término en el que se considere la diferencia entre las soluciones en las dos direcciones.

$$\begin{aligned}
 E_1(\vec{h}_1, \vec{h}_2) &= \int_{\Omega} \left( I_1(\mathbf{x}) - I_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right)^2 + \alpha \int_{\Omega} \Phi \left( \mathbf{D}(\|\nabla I_1\|), \nabla \vec{h}_1 \right) \\
 &\quad + \beta \int_{\Omega} \Psi \left( \left\| \vec{h}_2(x) + \vec{h}_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right\| \right) \quad (8.1.1)
 \end{aligned}$$

$$\begin{aligned}
 E_2(\vec{h}_1, \vec{h}_2) &= \int_{\Omega} \left( I_2(\mathbf{x}) - I_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right)^2 + \alpha \int_{\Omega} \Phi \left( \mathbf{D}(\|\nabla I_2\|), \nabla \vec{h}_2 \right) \\
 &\quad + \beta \int_{\Omega} \Psi \left( \left\| \vec{h}_1(\mathbf{x}) + \vec{h}_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right\| \right) \quad (8.1.2)
 \end{aligned}$$

Se fuerza la consistencia en la transformación inversa minimizando las diferencias cuadradas entre  $h_1$  y  $h_2$ . Esta aproximación es similar a la propuesta por Christensen en [Chr99], aunque en éste se calcula la inversa de  $h_1$  y trabaja con cuatro incógnitas en vez de dos.

## 8.2 La elección de la función $\Psi$

La función  $\Psi$  afecta a la norma de la suma de  $h_1$  y  $h_2$  en las ecuaciones (8.1.1) y (8.1.2). Gracias a la función  $\Psi$  podemos diferenciar entre puntos que tienen buenas

correspondencias,  $\Psi \left( \left\| \vec{h}_2(\mathbf{x}) + \vec{h}_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right\|^2 \right) \cong 0$ , de los puntos donde no hay

una buena correspondencia,  $\Psi \left( \left\| \vec{h}_2(\mathbf{x}) + \vec{h}_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right\|^2 \right) \gg 0$ .

En las pruebas que hemos realizado, hemos considerado dos tipos de funciones  $\Psi(s)$ :

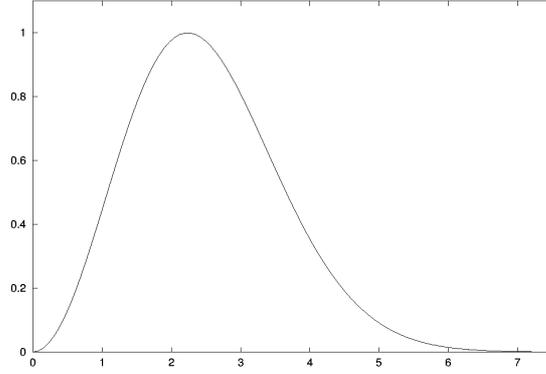


Figura 8.2:  $\Psi(s^2) = \frac{s^2}{\gamma} e^{1-\frac{s^2}{\gamma}}$

$$\Psi(s) = s, \quad (8.2.1)$$

$$\Psi(s) = \frac{s}{\gamma} e^{1-\frac{s}{\gamma}} \quad (8.2.2)$$

En la función (8.2.2)  $\gamma$  es un umbral para el cual se considera que hay una buena correspondencia. La primera elección para la función  $\Psi$  (8.2.1) se corresponde con mínimos cuadrados estándar y la segunda, se puede utilizar para tratar con outliers. Si utilizamos este tipo de funciones, podemos extender el dominio de integración a la imagen completa teniendo en cuenta aquellos subdominios donde el flujo es suficientemente simétrico.

Con la primera función, intentaremos encontrar las simetrías incluso aún si el punto en cuestión no tiene correspondencia. Se le dará mayor peso a aquellos puntos cuyas soluciones directa e inversa estén muy alejados y se intentarán aproximar entre sí.

Con la segunda función incorporamos de forma natural un mecanismo para excluir los puntos que probablemente no tengan correspondencia en ambas imágenes. Cuando la diferencia entre la función directa y la inversa sobrepasa el umbral  $\gamma$ , la función 8.2.2 decae, con lo que esos puntos tienden a tener un menor efecto en la minimización, es decir, no se tiende a aproximar las dos funciones en ese punto.

Como veremos en los ejemplos, los puntos donde falla la simetría se puede asociar con las oclusiones de las dos imágenes.

# Capítulo 9

## Cálculo del flujo óptico simétrico

### 9.1 Introducción

En este capítulo vamos a proponer una energía simétrica para el problema de la estimación del flujo óptico. Modificaremos los modelos que habíamos propuesto en el capítulo 6 para considerar la simetría.

En las distintas pruebas que hemos realizado, utilizamos las dos versiones de funciones  $\Psi$  (8.2.1) y (8.2.2) vistas en el capítulo anterior. Hacemos una comparación con el método no simétrico, mostrando resultados numéricos para el ejemplo de los Cuadrados y de las Torres de Mármol. Para estos dos ejemplos se consigue una mejora considerable y para el ejemplo de los Cuadrados podremos ver los mapas de oclusiones.

### 9.2 El modelo de energía

La energía que vamos a proponer está compuesta por dos – una para cada dirección de cálculo – que están ligadas en una sola para la estimación conjunta de las funciones de correspondencia. Cada una de las energías constituye la estimación del flujo en cada una de las direcciones. Ambas energías están relacionadas por su término de simetría, mientras que el resto de sus términos son similares a los ya vistos en la sección 6.2.

Cada ecuación de energía depende de las dos funciones de correspondencia,  $\vec{h}_1$  y  $\vec{h}_2$ , pero en cada una se intenta encontrar la solución a una sola de las funciones. La primera energía viene dada por:

$$\begin{aligned}
E_1(\vec{h}_1, \vec{h}_2) &= \frac{1}{2} \int_{\Omega} \left( I_1(\mathbf{x}) - I_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right)^2 \\
&\quad + \frac{C}{2} \int_{\Omega} \text{traza} \left( \nabla \vec{h}_1^t \mathbf{D}(\nabla I_1) \nabla \vec{h}_1 \right) \\
&\quad + \frac{\beta}{2} \int_{\Omega} \Psi \left( \left\| \vec{h}_1(\mathbf{x}) + \vec{h}_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right\|^2 \right) \quad (9.2.1)
\end{aligned}$$

Como podemos ver, esta energía es equivalente a la propuesta en la sección 6.2, en el sentido  $I_1$  hacia  $I_2$ . Se añade un tercer término, término de simetría, que es el que liga las dos incógnitas.

Análogamente para la segunda función de correspondencia tenemos que,

$$\begin{aligned}
E_2(\vec{h}_1, \vec{h}_2) &= \frac{1}{2} \int_{\Omega} \left( I_2(\mathbf{x}) - I_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right)^2 \\
&\quad + \frac{C}{2} \int_{\Omega} \text{traza} \left( \nabla \vec{h}_2^t \mathbf{D}(\nabla I_2) \nabla \vec{h}_2 \right) \\
&\quad + \frac{\beta}{2} \int_{\Omega} \Psi \left( \left\| \vec{h}_2(\mathbf{x}) + \vec{h}_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right\|^2 \right) \quad (9.2.2)
\end{aligned}$$

o lo que es igual a la energía de estimación del flujo en el sentido de  $I_2$  hacia  $I_1$  incorporando el término de simetría.

El interés de plantear estas ecuaciones es el de poder estimar conjuntamente tanto  $\vec{h}_1$  como  $\vec{h}_2$  al mismo tiempo. Podríamos estimar primero una de ellas con el método de flujo óptico directo y utilizarla después a la hora de estimar la función de correspondencia en el sentido inverso. Esto simplificaría bastante la forma de estimar el flujo pero la solución que se obtendría se vería distorsionada por la primera estimación. La manera más adecuada es la de estimarlas de forma conjunta combinando ambas energías en una sola,

$$E(\vec{h}_1, \vec{h}_2) = E_1(\vec{h}_1, \vec{h}_2) + E_2(\vec{h}_1, \vec{h}_2)$$

### 9.3 Minimizando la energía

La función de energía anterior se minimiza como la solución asintótica de las ecuaciones de Euler-Lagrange asociadas. Con el fin de encontrar un mínimo para esta energía, procedemos de forma iterativa como sigue: Primero empezamos con una aproximación inicial  $(\vec{h}_1^0, \vec{h}_2^0)$  y luego obtenemos  $(\vec{h}_1^{n+1}, \vec{h}_2^{n+1})$  a partir de  $(\vec{h}_1^n, \vec{h}_2^n)$  como una iteración de un método de descenso por gradiente. El inconveniente de esta energía es que al derivar las ecuaciones de Euler-Lagrange, nos queda un sistema de ecuaciones en derivadas parciales muy complejas, de difícil resolución.

Lo que hacemos es

$$E(\vec{h}_1^{n+1}, \vec{h}_2^{n+1}) = E_1(\vec{h}_1^{n+1}, \vec{h}_2^n) + E_2(\vec{h}_1^n, \vec{h}_2^{n+1})$$

de forma que la energía,  $E_i$ , que calcula cada función de correspondencia depende sólo de  $\vec{h}_i^{n+1}$ , en el instante  $n + 1$ . En cada iteración suponemos que  $\vec{h}_1^n$  y  $\vec{h}_2^n$  son constantes y  $E_1$  y  $E_2$  dependen solo de una variable.

La aproximación inicial,  $(\vec{h}_1^0, \vec{h}_2^0)$ , se obtiene por un método clásico de correlación en la escala de resolución más gruesa. Derivando las ecuaciones de Euler-Lagrange obtenemos,

$$\begin{aligned} \frac{\partial E}{\partial \vec{h}_1^{n+1}} (\vec{h}_1^{n+1}, \vec{h}_2^{n+1}) &= - \left( I_1 - I_2^{\vec{h}_1^{n+1}} \right) \nabla I_2^{\vec{h}_1^{n+1}} - C \operatorname{div} \left( \mathbf{D}(\nabla I_1) \nabla \vec{h}_1^{n+1} \right) \\ &+ \beta \Psi' \left( \left\| \vec{h}_1^{n+1} + (\vec{h}_2^n) \vec{h}_1^{n+1} \right\|^2 \right) \left( \mathbf{Id} + \nabla (\vec{h}_2^n) \vec{h}_1^{n+1} \right) \\ &\cdot \left( \vec{h}_1^{n+1} + (\vec{h}_2^n) \vec{h}_1^{n+1} \right) \end{aligned} \quad (9.3.1)$$

$$\begin{aligned} \frac{\partial E}{\partial \vec{h}_2^{n+1}} (\vec{h}_1^{n+1}, \vec{h}_2^{n+1}) &= - \left( I_2 - I_1^{\vec{h}_2^{n+1}} \right) \nabla I_1^{\vec{h}_2^{n+1}} - C \operatorname{div} \left( \mathbf{D}(\nabla I_2) \nabla \vec{h}_2^{n+1} \right) \\ &+ \beta \Psi' \left( \left\| \vec{h}_2^{n+1} + (\vec{h}_1^n) \vec{h}_2^{n+1} \right\|^2 \right) \left( \mathbf{Id} + \nabla (\vec{h}_1^n) \vec{h}_2^{n+1} \right) \\ &\cdot \left( \vec{h}_2^{n+1} + (\vec{h}_1^n) \vec{h}_2^{n+1} \right) \end{aligned} \quad (9.3.2)$$

Aplicando un método de descenso por gradiente para buscar el mínimo y con el fin de recuperar largos desplazamientos – tal y como vimos en el cálculo del flujo

óptico en una sola dirección (ver sección 5.3)– introducimos el método en un marco de análisis multiescala lineal o multipiramidal, obteniendo un sistema de cuatro ecuaciones diferenciales:

$$\begin{aligned}
\frac{\partial u_{1,\sigma}^{n+1}}{\partial t} &= \left( I_1^\sigma - I_2^{\sigma, \vec{h}_{1,\sigma}^{n+1}} \right) \left( \frac{\partial I_2^\sigma}{\partial x} \right)^{\vec{h}_{1,\sigma}^{n+1}} + C \operatorname{div} \left( \mathbf{D}(\nabla I_1^\sigma) \nabla u_{1,\sigma}^{n+1} \right) \\
&\quad - \beta \Psi' \left( \left\| \vec{h}_{1,\sigma}^{n+1} + (\vec{h}_{2,\sigma}^n)^{\vec{h}_{1,\sigma}^{n+1}} \right\|^2 \right) \left[ \left( 1 + \left( \frac{\partial u_{2,\sigma}^n}{\partial x} \right)^{\vec{h}_{1,\sigma}^{n+1}} \right) \left( u_{1,\sigma}^{n+1} + (u_{2,\sigma}^n)^{\vec{h}_{1,\sigma}^{n+1}} \right) \right. \\
&\quad \left. + \left( \frac{\partial u_{2,\sigma}^n}{\partial y} \right)^{\vec{h}_{1,\sigma}^{n+1}} \left( v_{1,\sigma}^{n+1} + (v_{2,\sigma}^n)^{\vec{h}_{1,\sigma}^{n+1}} \right) \right] \tag{9.3.3}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial v_{1,\sigma}^{n+1}}{\partial t} &= \left( I_1^\sigma - I_2^{\sigma, \vec{h}_{1,\sigma}^{n+1}} \right) \left( \frac{\partial I_2^\sigma}{\partial y} \right)^{\vec{h}_{1,\sigma}^{n+1}} + C \operatorname{div} \left( \mathbf{D}(\nabla I_1^\sigma) \nabla v_{1,\sigma}^{n+1} \right) \\
&\quad - \beta \Psi' \left( \left\| \vec{h}_{1,\sigma}^{n+1} + (\vec{h}_{2,\sigma}^n)^{\vec{h}_{1,\sigma}^{n+1}} \right\|^2 \right) \left[ \left( \frac{\partial v_{2,\sigma}^n}{\partial x} \right)^{\vec{h}_{1,\sigma}^{n+1}} \left( u_{1,\sigma}^{n+1} + (u_{2,\sigma}^n)^{\vec{h}_{1,\sigma}^{n+1}} \right) \right. \\
&\quad \left. + \left( 1 + \left( \frac{\partial v_{2,\sigma}^n}{\partial y} \right)^{\vec{h}_{1,\sigma}^{n+1}} \right) \left( v_{1,\sigma}^{n+1} + (v_{2,\sigma}^n)^{\vec{h}_{1,\sigma}^{n+1}} \right) \right] \tag{9.3.4}
\end{aligned}$$

Estas dos ecuaciones son para el cálculo del flujo en la dirección de la imagen  $I_1$  a  $I_2$  y

$$\begin{aligned}
\frac{\partial u_{2,\sigma}^{n+1}}{\partial t} &= \left( I_2^\sigma - I_1^{\sigma, \vec{h}_{2,\sigma}^{n+1}} \right) \left( \frac{\partial I_1^\sigma}{\partial x} \right)^{\vec{h}_{2,\sigma}^{n+1}} + C \operatorname{div} \left( \mathbf{D}(\nabla I_2^\sigma) \nabla u_{2,\sigma}^{n+1} \right) \\
&\quad - \beta \Psi' \left( \left\| \vec{h}_{2,\sigma}^{n+1} + (\vec{h}_{1,\sigma}^n)^{\vec{h}_{2,\sigma}^{n+1}} \right\|^2 \right) \left[ \left( 1 + \left( \frac{\partial u_{1,\sigma}^n}{\partial x} \right)^{\vec{h}_{2,\sigma}^{n+1}} \right) \left( u_{2,\sigma}^{n+1} + (u_{1,\sigma}^n)^{\vec{h}_{2,\sigma}^{n+1}} \right) \right. \\
&\quad \left. + \left( \frac{\partial u_{1,\sigma}^n}{\partial y} \right)^{\vec{h}_{2,\sigma}^{n+1}} \left( v_{2,\sigma}^{n+1} + (v_{1,\sigma}^n)^{\vec{h}_{2,\sigma}^{n+1}} \right) \right] \tag{9.3.5}
\end{aligned}$$

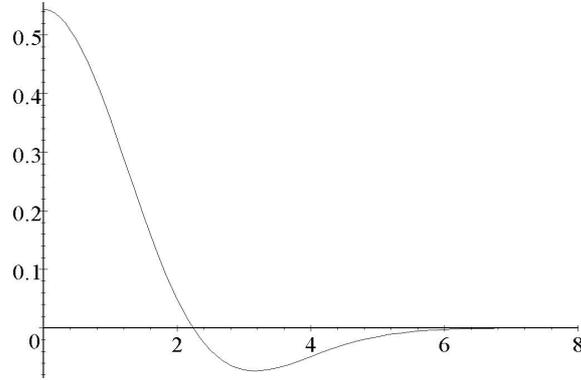


Figura 9.1:  $\Psi'(s^2) = \frac{s^2}{\gamma} e^{1-\frac{s^2}{\gamma}}$  con  $\gamma = 5$ .

$$\begin{aligned}
\frac{\partial v_{2,\sigma}^{n+1}}{\partial t} &= \left( I_2^\sigma - I_1^{\sigma, \vec{h}_{2,\sigma}^{n+1}} \right) \left( \frac{\partial I_1^\sigma}{\partial y} \right)^{\vec{h}_{2,\sigma}^{n+1}} + C \operatorname{div} (\mathbf{D}(\nabla I_2^\sigma) \nabla v_{2,\sigma}^{n+1}) \\
&- \beta \Psi' \left( \left\| \vec{h}_{2,\sigma}^{n+1} + (\vec{h}_{1,\sigma}^n)^{\vec{h}_{2,\sigma}^{n+1}} \right\|^2 \right) \left[ \left( \frac{\partial v_{1,\sigma}^n}{\partial x} \right)^{\vec{h}_{2,\sigma}^{n+1}} \left( u_{2,\sigma}^{n+1} + (u_{1,\sigma}^n)^{\vec{h}_{2,\sigma}^{n+1}} \right) \right. \\
&\left. + \left( 1 + \left( \frac{\partial v_{1,\sigma}^n}{\partial y} \right)^{\vec{h}_{2,\sigma}^{n+1}} \right) \left( v_{2,\sigma}^{n+1} + (v_{1,\sigma}^n)^{\vec{h}_{2,\sigma}^{n+1}} \right) \right] \quad (9.3.6)
\end{aligned}$$

para el cálculo en la dirección inversa.

Estas ecuaciones en EDP son muy similares a las vistas en la sección 5.3 con la excepción del tercer término de la parte derecha que representa la parte de la simetría.

Si analizamos con mayor detenimiento este término, vemos que depende exclusivamente de las funciones de correspondencia en ambas direcciones, además de mezclar la información de las dos en cada una de las ecuaciones. La función  $\Psi'$  funciona como un factor que inhibe o habilita el mecanismo de simetría.

Como podemos ver en la figura 9.1, eligiendo la función  $\Psi$  exponencial con  $\gamma = 5$ ,  $\Psi'$  cruza el eje de las  $x$  en  $\sqrt{5}$ , para luego acercarse asintóticamente a 0 a medida que nos alejamos en el eje de coordenadas. Efectivamente, nuestro propósito al elegir la función  $\Psi$  era el de hacer simétricas las funciones de correspondencia en los puntos cuyos valores fueran inferior al umbral  $\gamma$ . Esto se puede ver en las ecuaciones en EDP anteriores, ya que cuando se anula  $\Psi'$  obtenemos exactamente el modelo tradicional, es decir, la estimación directa sin considerar la simetría, mientras que para valores

pequeños conseguimos que se *active* la simetría. Si observamos con mayor detenimiento la gráfica, vemos que a partir de  $x^2 = \sqrt{5}$  la curva está próxima a cero, lo que significa que a partir de este valor, la simetría no se tiene en cuenta y las funciones progresan libremente.

## 9.4 Esquema numérico

Vamos a describir un algoritmo eficiente para nuestro modelo de flujo óptico simétrico. Discretizamos el sistema parabólico formado por las ecuaciones (9.3.3), (9.3.4), (9.3.5) y (9.3.6) por diferencias finitas. Todas las derivadas espaciales se aproximan por diferencias centradas, y para la discretización en la dirección de  $t$  utilizamos un esquema lineal implícito. Discretizamos las derivadas de primer orden,  $\frac{\partial u}{\partial x}$  y  $\frac{\partial u}{\partial x}$ , utilizando máscaras invariantes por rotaciones de  $45^\circ$ <sup>1</sup>

Al igual que en la sección 6.3 suponemos que

$$\mathbf{D}(\nabla I_1^\sigma) = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \text{ y } \mathbf{D}(\nabla I_2^\sigma) = \begin{pmatrix} a' & b' \\ b' & c' \end{pmatrix}$$

Empezamos estudiando cada ecuación por separado. Para el componente de  $\vec{h}_1$  en la dirección de las  $x$  tenemos

$$\frac{u_{1,i,j}^{k+1} - u_{1,i,j}^k}{\tau} = T_l + T_r + T_s, \quad (9.4.1)$$

donde  $T_l$  representa el término de ligadura,  $T_r$  el término de regularización y  $T_s$  el de simetría. Vamos a estudiar cómo discretizar cada uno por separado.

Los dos primeros términos, de ligadura y regularización, son parecidos a los ya expuestos en la sección 6.3. Aplicamos los desarrollos en serie de Taylor centrados en  $(u_{1,i,j}^k, v_{1,i,j}^k)$  para linealizar el término de ligadura,

$$\begin{aligned} T_l = & I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) \left( I_1(\mathbf{x}_{i,j}) - I_2(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) \right. \\ & + u_{1,i,j}^k I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) + v_{1,i,j}^k I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) \left. \right) \\ & - u_{1,i,j}^{k+1} I_{2,x}^2(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) - v_{1,i,j}^{k+1} I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) \end{aligned} \quad (9.4.2)$$

<sup>1</sup>En el apéndice C se explica detalladamente la forma de discretización para el gradiente y para la divergencia.

Para discretizar el término de regularización tenemos en cuenta la discretización propuesta en el apéndice C para la divergencia, obteniendo el siguiente esquema,

$$\begin{aligned}
T_r = C & \left( \frac{a_{i+1,j} + a_{i,j}}{2} \frac{u_{1,i+1,j}^{k+1} - u_{1,i,j}^{k+1}}{h_1^2} + \frac{a_{i-1,j} + a_{i,j}}{2} \frac{u_{1,i-1,j}^{k+1} - u_{1,i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{c_{i,j+1} + c_{i,j}}{2} \frac{u_{1,i,j+1}^{k+1} - u_{1,i,j}^{k+1}}{h_2^2} + \frac{c_{i,j-1} + c_{i,j}}{2} \frac{u_{1,i,j-1}^{k+1} - u_{1,i,j}^{k+1}}{h_2^2} \\
& + \frac{b_{i+1,j+1} + b_{i,j}}{2} \frac{u_{1,i+1,j+1}^{k+1} - u_{1,i,j}^{k+1}}{2h_1h_2} + \frac{b_{i-1,j-1} + b_{i,j}}{2} \frac{u_{1,i-1,j-1}^{k+1} - u_{1,i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{b_{i+1,j-1} + b_{i,j}}{2} \frac{u_{1,i+1,j-1}^{k+1} - u_{1,i,j}^{k+1}}{2h_1h_2} - \frac{b_{i-1,j+1} + b_{i,j}}{2} \frac{u_{1,i-1,j+1}^{k+1} - u_{1,i,j}^{k+1}}{2h_1h_2} \right) \quad (9.4.3)
\end{aligned}$$

Por último, discretizamos el término de simetría,

$$\begin{aligned}
T_s = -\beta \Psi' & \left( \left\| \vec{h}_{1,i,j}^k + (\vec{h}_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right\|^2 \right) \left[ \left( 1 + (u_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right) \left( u_{1,i,j}^{k+1} + (u_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right) \right. \\
& \left. + (u_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \left( v_{1,i,j}^k + (v_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right) \right] \quad (9.4.4)
\end{aligned}$$

De la misma forma discretizamos la componente  $v_1$  del vector  $\vec{h}_1$  como sigue,

$$\frac{v_{1,i,j}^{k+1} - v_{1,i,j}^k}{\tau} = T_l + T_r + T_s, \quad (9.4.5)$$

donde  $T_l$ ,  $T_r$  y  $T_s$  tienen el mismo significado que en la ecuación (9.4.1).

Analizamos cada término por separado: Término de ligadura,

$$\begin{aligned}
T_l = I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) & \left( I_1(\mathbf{x}_{i,j}) - I_2(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) \right. \\
& + u_{1,i,j}^k I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) + v_{1,i,j}^k I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) \\
& \left. - u_{1,i,j}^{k+1} I_{2,x}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) I_{2,y}(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) - v_{1,i,j}^{k+1} I_{2,y}^2(\mathbf{x}_{i,j} + \vec{h}_{1,i,j}^k) \right) \quad (9.4.6)
\end{aligned}$$

Término de regularización,

$$\begin{aligned}
T_r = C & \left( \frac{a_{i+1,j} + a_{i,j}}{2} \frac{v_{1,i+1,j}^{k+1} - v_{1,i,j}^{k+1}}{h_1^2} + \frac{a_{i-1,j} + a_{i,j}}{2} \frac{v_{1,i-1,j}^{k+1} - v_{1,i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{c_{i,j+1} + c_{i,j}}{2} \frac{v_{1,i,j+1}^{k+1} - v_{1,i,j}^{k+1}}{h_2^2} + \frac{c_{i,j-1} + c_{i,j}}{2} \frac{v_{1,i,j-1}^{k+1} - v_{1,i,j}^{k+1}}{h_2^2} \\
& + \frac{b_{i+1,j+1} + b_{i,j}}{2} \frac{v_{1,i+1,j+1}^{k+1} - v_{1,i,j}^{k+1}}{2h_1h_2} + \frac{b_{i-1,j-1} + b_{i,j}}{2} \frac{v_{1,i-1,j-1}^{k+1} - v_{1,i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{b_{i+1,j-1} + b_{i,j}}{2} \frac{v_{1,i+1,j-1}^{k+1} - v_{1,i,j}^{k+1}}{2h_1h_2} - \frac{b_{i-1,j+1} + b_{i,j}}{2} \frac{v_{1,i-1,j+1}^{k+1} - v_{1,i,j}^{k+1}}{2h_1h_2} \right) \quad (9.4.7)
\end{aligned}$$

Término de simetría,

$$\begin{aligned}
T_s = -\beta\Psi' & \left( \left\| \vec{h}_{1,i,j}^k + (\vec{h}_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right\|^2 \right) \left[ (v_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \left( u_{1,i,j}^{k+1} + (u_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right) \right. \\
& \left. + \left( 1 + (v_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right) \left( v_{1,i,j}^{k+1} + (v_{2,i,j}^k)^{\vec{h}_{1,i,j}^k} \right) \right] \quad (9.4.8)
\end{aligned}$$

El esquema numérico para el vector  $\vec{h}_2$  en la dirección de las  $x$  viene dado por

$$\frac{u_{2,i,j}^{k+1} - u_{2,i,j}^k}{\tau} = T_l + T_r + T_s, \quad (9.4.9)$$

Los tres términos son similares a los correspondientes para  $u_1$  dados en (9.4.2), (9.4.3) y (9.4.4)

Término de ligadura,

$$\begin{aligned}
T_l = I_{1,x}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) & \left( I_2(\mathbf{x}_{i,j}) - I_1(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) \right) \\
& + u_{2,i,j}^k I_{1,x}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) + v_{2,i,j}^k I_{1,y}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) \\
& - u_{2,i,j}^{k+1} I_{1,x}^2(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) - v_{2,i,j}^{k+1} I_{1,y}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) I_{1,x}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) \quad (9.4.10)
\end{aligned}$$

Término de regularización,

$$\begin{aligned}
T_r = C & \left( \frac{a'_{i+1,j} + a'_{i,j}}{2} \frac{u_{2,i+1,j}^{k+1} - u_{2,i,j}^{k+1}}{h_1^2} + \frac{a'_{i-1,j} + a'_{i,j}}{2} \frac{u_{2,i-1,j}^{k+1} - u_{2,i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{c'_{i,j+1} + c'_{i,j}}{2} \frac{u_{2,i,j+1}^{k+1} - u_{2,i,j}^{k+1}}{h_2^2} + \frac{c'_{i,j-1} + c'_{i,j}}{2} \frac{u_{2,i,j-1}^{k+1} - u_{2,i,j}^{k+1}}{h_2^2} \\
& + \frac{b'_{i+1,j+1} + b'_{i,j}}{2} \frac{u_{2,i+1,j+1}^{k+1} - u_{2,i,j}^{k+1}}{2h_1h_2} + \frac{b'_{i-1,j-1} + b'_{i,j}}{2} \frac{u_{2,i-1,j-1}^{k+1} - u_{2,i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{b'_{i+1,j-1} + b'_{i,j}}{2} \frac{u_{2,i+1,j-1}^{k+1} - u_{2,i,j}^{k+1}}{2h_1h_2} - \frac{b'_{i-1,j+1} + b'_{i,j}}{2} \frac{u_{2,i-1,j+1}^{k+1} - u_{2,i,j}^{k+1}}{2h_1h_2} \right) \quad (9.4.11)
\end{aligned}$$

Término de simetría,

$$\begin{aligned}
T_s = -\beta \Psi' & \left( \left\| \vec{h}_{2,i,j}^k + (\vec{h}_{1,i,j}^k)^{\vec{h}_{2,i,j}^k} \right\|^2 \right) \left[ \left( 1 + (u_{1,i,j}^k)_x^{\vec{h}_{2,i,j}^k} \right) \left( u_{2,i,j}^{k+1} + (u_{1,i,j}^k)^{\vec{h}_{2,i,j}^k} \right) \right. \\
& \left. + (u_{1,i,j}^k)_y^{\vec{h}_{2,i,j}^k} \left( v_{2,i,j}^k + (v_{1,i,j}^k)^{\vec{h}_{2,i,j}^k} \right) \right] \quad (9.4.12)
\end{aligned}$$

De la misma forma discretizamos la componente  $v_2$  del vector  $\vec{h}_2$  como sigue,

$$\frac{v_{2,i,j}^{k+1} - v_{2,i,j}^k}{\tau} = T_l + T_r + T_s, \quad (9.4.13)$$

donde  $T_l$ ,  $T_r$  y  $T_s$  tienen el mismo significado que en la ecuación (9.4.9).

Analizamos cada término por separado: Término de ligadura,

$$\begin{aligned}
T_l = I_{1,y}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) & \left( I_2(\mathbf{x}_{i,j}) - I_1(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) \right) \\
& + u_{2,i,j}^k I_{1,x}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) + v_{2,i,j}^k I_{1,y}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) \\
& - u_{1,i,j}^{k+1} I_{1,x}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) I_{1,y}(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) - v_{2,i,j}^{k+1} I_{1,y}^2(\mathbf{x}_{i,j} + \vec{h}_{2,i,j}^k) \quad (9.4.14)
\end{aligned}$$

Término de regularización,

$$\begin{aligned}
T_r = C & \left( \frac{a'_{i+1,j} + a'_{i,j}}{2} \frac{v_{2,i+1,j}^{k+1} - v_{2,i,j}^{k+1}}{h_1^2} + \frac{a'_{i-1,j} + a'_{i,j}}{2} \frac{v_{2,i-1,j}^{k+1} - v_{2,i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{c'_{i,j+1} + c'_{i,j}}{2} \frac{v_{2,i,j+1}^{k+1} - v_{2,i,j}^{k+1}}{h_2^2} + \frac{c'_{i,j-1} + c'_{i,j}}{2} \frac{v_{2,i,j-1}^{k+1} - v_{2,i,j}^{k+1}}{h_2^2} \\
& + \frac{b'_{i+1,j+1} + b'_{i,j}}{2} \frac{v_{2,i+1,j+1}^{k+1} - v_{2,i,j}^{k+1}}{2h_1h_2} + \frac{b'_{i-1,j-1} + b'_{i,j}}{2} \frac{v_{2,i-1,j-1}^{k+1} - v_{2,i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{b'_{i+1,j-1} + b'_{i,j}}{2} \frac{v_{2,i+1,j-1}^{k+1} - v_{2,i,j}^{k+1}}{2h_1h_2} - \frac{b'_{i-1,j+1} + b'_{i,j}}{2} \frac{v_{2,i-1,j+1}^{k+1} - v_{2,i,j}^{k+1}}{2h_1h_2} \right) \quad (9.4.15)
\end{aligned}$$

Término de simetría,

$$\begin{aligned}
T_s = -\beta\Psi' & \left( \left\| \vec{h}_{2,i,j}^k + (\vec{h}_{1,i,j}^k)^{\vec{h}_{2,i,j}^k} \right\|^2 \right) \left[ (v_{1,i,j}^k)_{x}^{\vec{h}_{2,i,j}^k} \left( u_{2,i,j}^{k+1} + (u_{1,i,j}^k)^{\vec{h}_{2,i,j}^k} \right) \right. \\
& \left. + \left( 1 + (v_{1,i,j}^k)_y^{\vec{h}_{2,i,j}^k} \right) \left( v_{2,i,j}^{k+1} + (v_{1,i,j}^k)^{\vec{h}_{2,i,j}^k} \right) \right] \quad (9.4.16)
\end{aligned}$$

Hubiese sido más correcto introducir dos índices,  $k$  y  $n$ , en la discretización, ya que formalmente tenemos dos esquemas iterativos, uno para el nivel de la energía que es el que se corresponde con el índice  $n$ , y otro para el método de descenso por gradiente, que se corresponde con el índice  $k$ . El primero es un mecanismo que introducimos para simplificar las ecuaciones, procediendo en varios pasos a la hora de estimar la solución final, mientras que el segundo es el índice temporal correspondiente al método de descenso por gradiente.

A la hora de implementar el algoritmo, sólo tuvimos en cuenta un esquema iterativo, o lo que es igual a suponer que el índice  $k$  sólo varía en un instante de tiempo.

No olvidemos que además de estos esquemas iterativos, hemos introducido estas ecuaciones en un proceso multiescala que va refinando las soluciones de escalas más gruesas a escalas más finas.

Los parámetros son los mismos que los utilizados en el método de flujo óptico no simétrico, además del parámetro  $\beta$ , que es el que multiplica al término de simetría, y del parámetro  $\gamma$  si elegimos la función  $\Psi$  exponencial. En las pruebas numéricas el

valor de  $\beta$  se suele tomar más pequeño que el valor de  $\alpha$ . Más adelante veremos la influencia de este parámetro en el resultado del flujo.

## 9.5 Experiencias numéricas

En esta sección mostramos algunos resultados para imágenes sintéticas y realizamos comparaciones entre el método estándar y el simétrico.

Hemos probado nuestro algoritmo con el par de imágenes mostrado en la figura 9.2. Las oclusiones en estas imágenes son importantes, así que utilizamos la versión robusta de la función  $\Psi$ . En la figura 9.3 mostramos el desplazamiento real que existe de la imagen de la izquierda a la imagen de la derecha – fila superior – y viceversa, – en la primera columna representamos el desplazamiento en  $x$  y en la segunda en  $y$ .

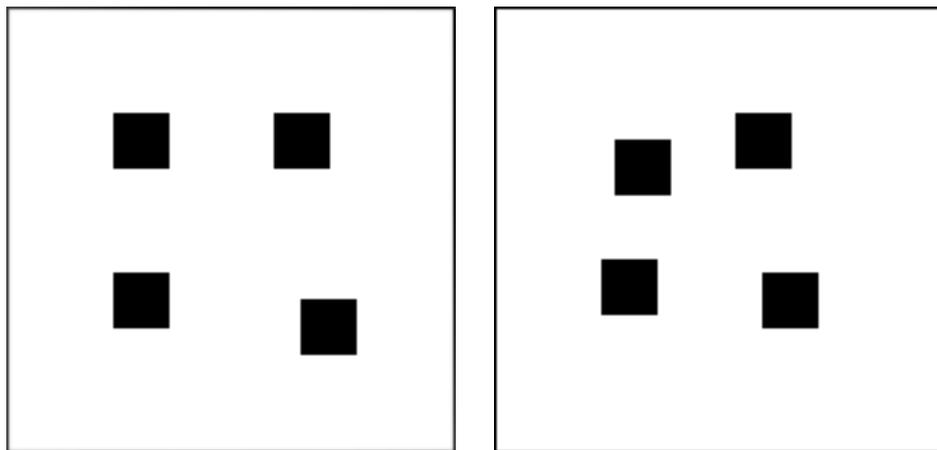


Figura 9.2: Secuencia de los cuadrados

En la figura 9.4 representamos el resultado obtenido con nuestro método simétrico. Se han llevado a cabo tres niveles de resolución con  $\alpha = 1, \beta = 0.5, \gamma = 5, T = 200, dt = 10$  y  $\nu = 0.1$ .

En la figura 9.5 se muestran las oclusiones tanto exactas – fila superior – como estimadas. Los puntos de las oclusiones son aquellos para los que las condiciones de simetría no se cumplen – sobrepasan el umbral fijado de  $\gamma = 5$ . Como se puede ver, estas localizaciones se corresponden bastante bien con las oclusiones reales. Sin embargo, estas regiones se han visto erosionadas debido al término de regularización

En la tabla 9.1 se muestran los errores que se producen tanto para el método simétrico como para el no simétrico. Como se puede ver, los errores que se producen en el método simétrico son bastante inferiores al no simétrico.

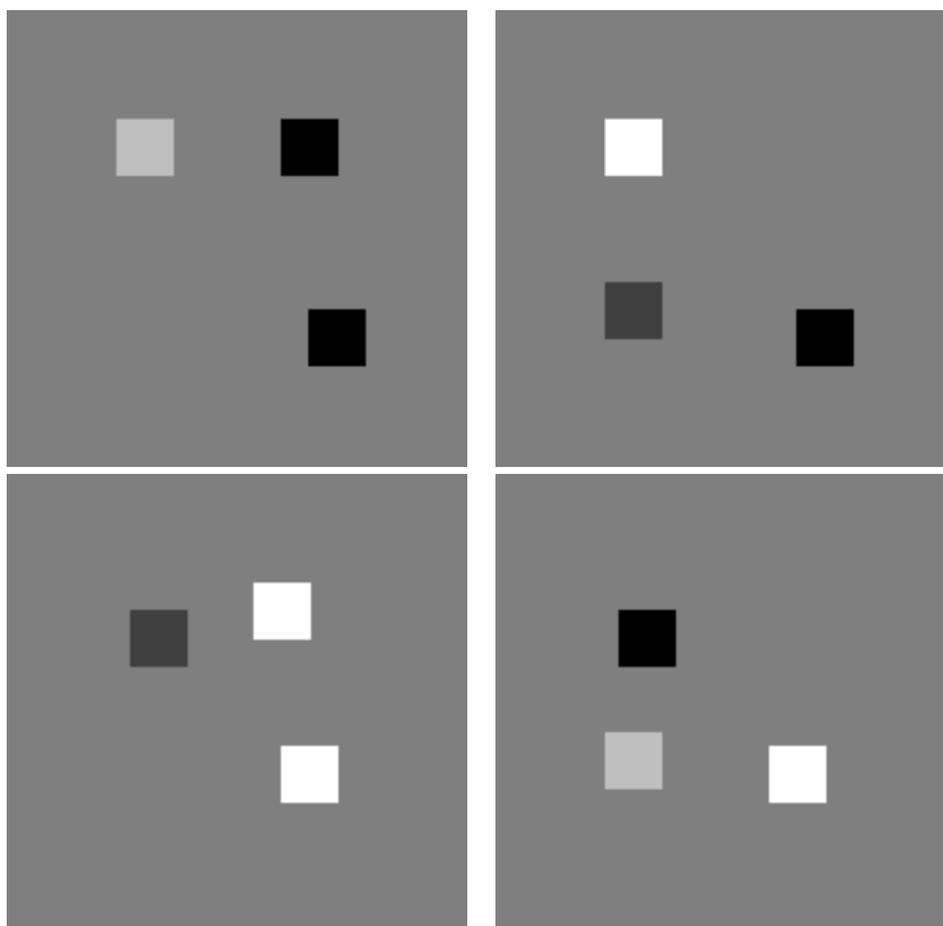


Figura 9.3: Desplazamiento real para la secuencia de los cuadrados: en la fila superior representamos el desplazamiento de izquierda a derecha – con la primera columna para la componente  $x$  y la segunda para  $y$ . En la segunda fila, representamos el caso inverso.

En la figura 9.6 mostramos la secuencia de las torres de mármol suministradas por el grupo KOGS/IAKS de la Universidad de Karlsruhe.

En la figura 9.7 se muestra la disparidad perfecta para las Torres de mármol.

En la figura 9.8 representamos el resultado obtenido con nuestro método simétrico. Se han llevado a cabo tres niveles de resolución con  $\alpha = 1, \beta = 0.5, \gamma = 5, T = 200, dt = 10$  y  $\nu = 0.1$ . Se compara el resultado con el método no simétrico. Se utiliza la función  $\Psi$  lineal,  $\Psi(s) = s$ .

Si comparamos los resultados sobre las imágenes de las torres de mármol, vemos que el método simétrico genera una solución más suave en la región del suelo, la columna del medio tiene un flujo más homogéneo, desaparece el gran error que se produce para el caso no simétrico en la esquina inferior izquierda y la componente vertical del flujo es más consistente con la solución exacta en la zona del fondo.

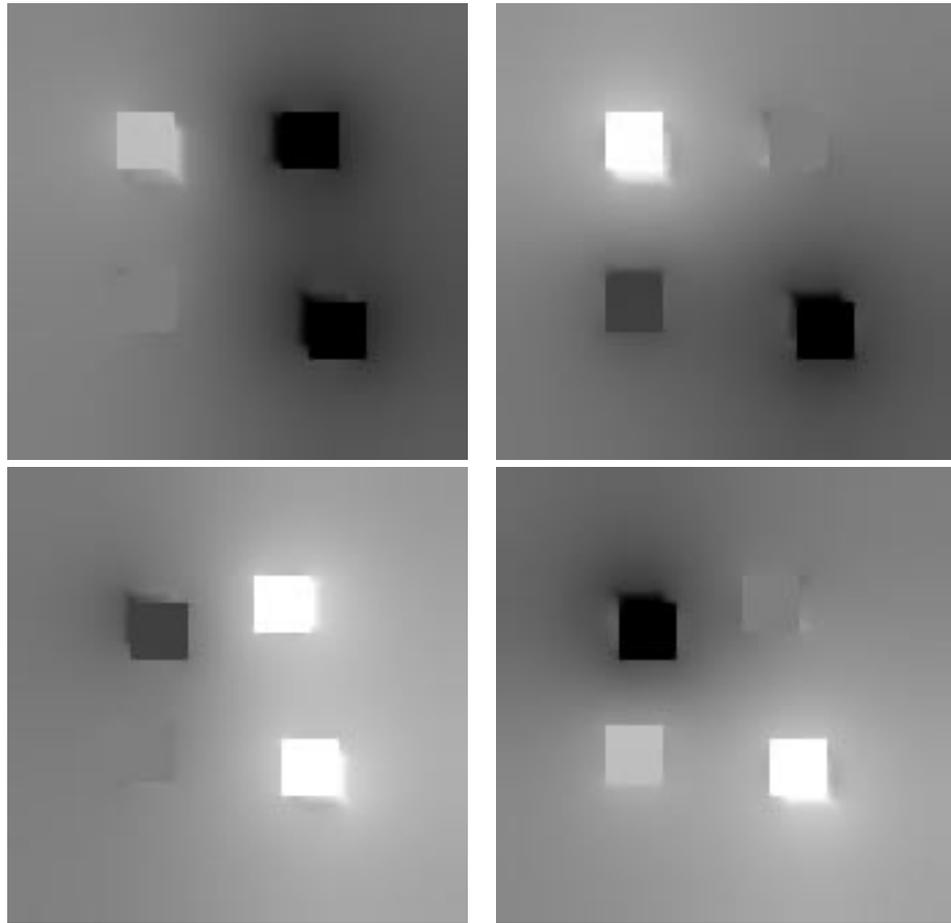


Figura 9.4: Flujo óptico que resulta de aplicar nuestro método simétrico a la secuencia de los cuadrados. En la izquierda (resp. derecha) la componente  $u$  (resp.  $v$ ) del flujo. Arriba (resp. abajo) se muestra  $\vec{h}_1$  (resp.  $\vec{h}_2$ ).

Tenemos que tener en cuenta que el número de iteraciones,  $T$ , lo hemos fijado a 200, muy por debajo de las 500 iteraciones que realizábamos en las pruebas anteriores. Esto puede implicar que el método simétrico converge más rápidamente hacia la solución.

El efecto del término de simetría parece haber mejorado considerablemente el resultado en las zonas donde existe mucha textura y en las zonas homogéneas. Solo con el término de regularización no es posible realizar una buena estimación del flujo en las zonas homogéneas, debido a la incertidumbre que se produce en el término de ligadura.

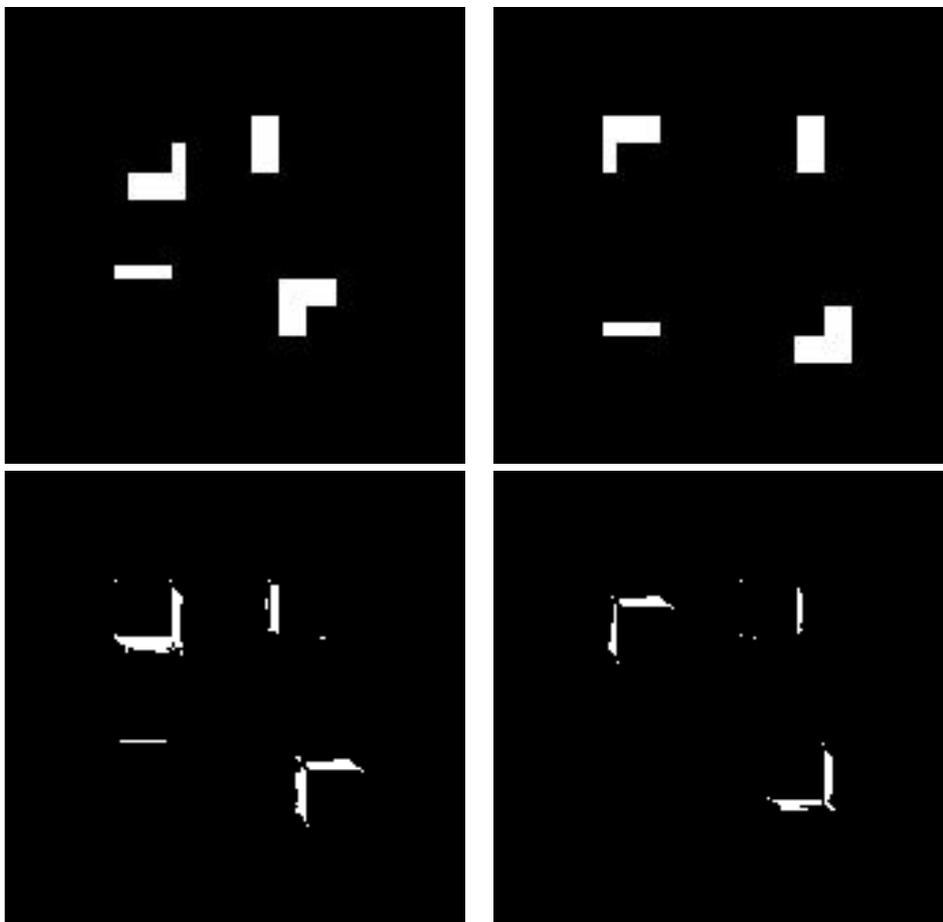


Figura 9.5: Oclusiones reales (fila superior) y estimadas (inferior) para la secuencia de los cuadrados.

Tabla 9.1: Errores euclídeo y angular medios que se producen para el ejemplo de los cuadrados 9.4 para el método simétrico y para el no simétrico. "ID" significa de izquierda a derecha y "DI" lo contrario.

		Error Euclídeo	Error Angular
No simétrico	ID	0.16	0.50
	DI	0.18	0.58
Simétrico	ID	0.081	0.18
	DI	0.084	0.19

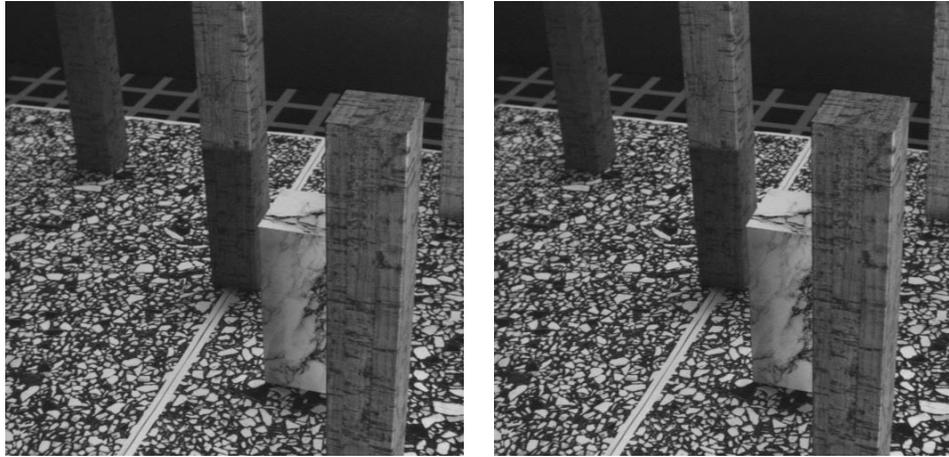


Figura 9.6: Secuencia de las Torres de mármol

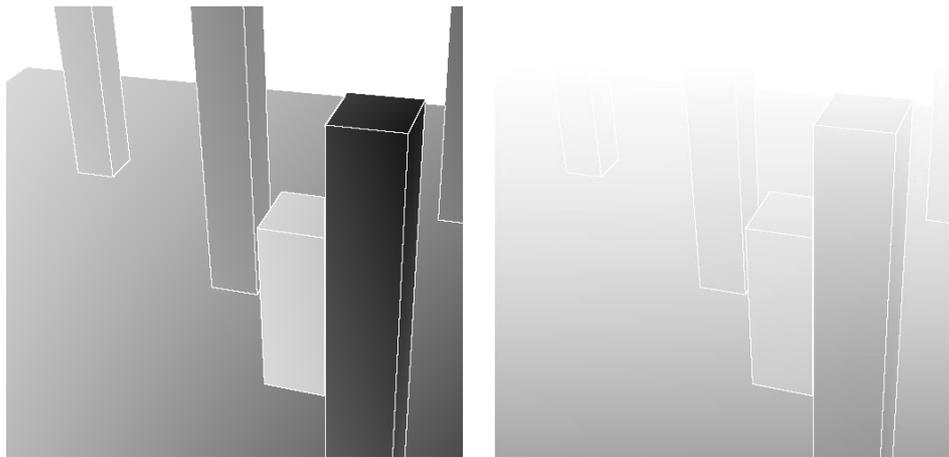


Figura 9.7: Desplazamiento real para la secuencia de las Torres: en la fila superior representamos el desplazamiento de izquierda a derecha – con la primera columna para la componente  $u$  y la segunda para  $v$ . En la segunda fila, representamos el caso inverso.

Tabla 9.2: Errores euclídeo y angular medios que se producen para el ejemplo de las torres 9.8 para el método simétrico y para el no simétrico.

	Error Euclídeo	Error Angular
No simétrico	0.37	9.4
Simétrico	0.17	5.2

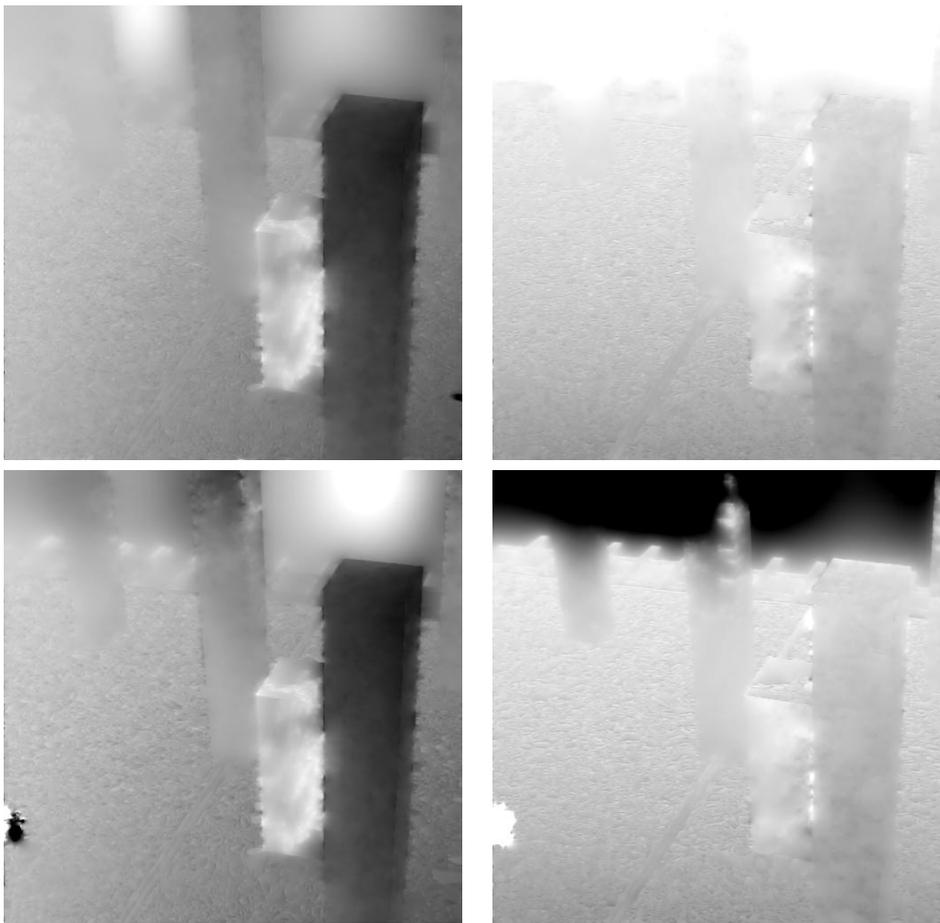


Figura 9.8: Flujo óptico que resulta de aplicar los métodos simétrico (arriba) y no simétrico (abajo) a la secuencia de las Torres. En la izquierda (resp. derecha) la componente  $u$  (resp.  $v$ ) del flujo.

# Capítulo 10

## Cálculo de la carta de disparidad simétrica

### 10.1 Introducción

En este capítulo vamos a estudiar cómo modificar nuestro modelo para introducir la simetría en el cálculo de la carta de disparidad. Al igual que vimos para el cálculo del flujo óptico, el modelo propuesto para la carta de disparidad está basado en una aproximación variacional. En este capítulo propondremos dos modelos distintos: Uno general, considerando cualquier configuración para la posición de las cámaras – geometría epipolar arbitraria – y otra para imágenes fronto-paralelas – aquellas en las que las líneas epipolares son horizontales.

El primero implica un modelo más complejo para el que hemos implementado un esquema numérico explícito. El segundo es más sencillo y se puede encontrar un esquema implícito para implementarlo. Además, gracias a la rectificación de imágenes, el proceso es mucho más rápido. Como ya hemos mencionado en el capítulo 3 podemos transformar cualquier par estéreo en un par fronto-paralelo. Con esta transformación se puede perder resolución, pero se obtiene un método más sencillo y rápido.

### 10.2 El modelo de energía

Tal y como vimos en el capítulo 3 sobre geometría epipolar, vamos a utilizar la parametrización del vector de desplazamiento teniendo en cuenta la geometría epipolar.

$$\vec{h}_i(x, y) = \begin{pmatrix} u_i(x, y) \\ v_i(x, y) \end{pmatrix} = \begin{pmatrix} \frac{-\lambda_i(x, y)b_i(x, y)}{\sqrt{a_i^2(x, y)+b_i^2(x, y)}} - \frac{a_i(x, y)x+b_i(x, y)y+c_i(x, y)}{a_i^2(x, y)+b_i^2(x, y)} a_i(x, y) \\ \frac{\lambda_i(x, y)a_i(x, y)}{\sqrt{a_i^2(x, y)+b_i^2(x, y)}} - \frac{a_i(x, y)x+b_i(x, y)y+c_i(x, y)}{a_i^2(x, y)+b_i^2(x, y)} b_i(x, y) \end{pmatrix}$$

La incógnita de nuestro problema es entonces  $\lambda_i$ , que según vimos en la sección 3.2, representa el desplazamiento dentro de la línea epipolar.

Al igual que hicimos con el flujo óptico, añadimos un nuevo término en la ecuación variacional con el fin de minimizar la diferencia entre la disparidad de izquierda a derecha y la disparidad de derecha a izquierda. Nuestra energía variacional viene dada por:

$$\begin{aligned} E_1(\lambda_1, \lambda_2) &= \frac{1}{2} \int_{\Omega} \left( I_1(\mathbf{x}) - I_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right)^2 + \frac{C}{2} \int_{\Omega} \nabla \lambda_1^t \mathbf{D}(\nabla I_1) \nabla \lambda_1 \\ &+ \frac{\beta}{2} \int_{\Omega} \Psi \left( \left\| \vec{h}_1(\mathbf{x}) + \vec{h}_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right\|^2 \right) \end{aligned} \quad (10.2.1)$$

Esta energía minimiza la disparidad desde la imagen  $I_1$  a  $I_2$ . Análogamente, tenemos que la energía en el sentido inverso viene dada por:

$$\begin{aligned} E_2(\lambda_1, \lambda_2) &= \frac{1}{2} \int_{\Omega} \left( I_2(\mathbf{x}) - I_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right)^2 + \frac{C}{2} \int_{\Omega} \nabla \lambda_2^t \mathbf{D}(\nabla I_2) \nabla \lambda_2 \\ &+ \frac{\beta}{2} \int_{\Omega} \Psi \left( \left\| \vec{h}_2(\mathbf{x}) + \vec{h}_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right\|^2 \right) \end{aligned} \quad (10.2.2)$$

donde la función  $\Psi(s)$  es de la forma que ya se ha explicado en la introducción de esta parte. La energía simétrica está dada por la suma de las energías anteriores

$$E(\lambda_1, \lambda_2) = E_1(\lambda_1, \lambda_2) + E_2(\lambda_1, \lambda_2)$$

### 10.3 Minimizando la energía

Con el fin de converger a un mínimo de esta energía, procedemos de forma iterativa como sigue: Primero, empezamos con una aproximación inicial  $(\lambda_1^0, \lambda_2^0)$  y luego

obtenemos cada  $(\lambda_1^{n+1}, \lambda_2^{n+1})$  a partir de la anterior  $(\lambda_1^n, \lambda_2^n)$  como una iteración de un método de descenso por gradiente

$$E(\lambda_1^{n+1}, \lambda_2^{n+1}) = E_1(\lambda_1^{n+1}, \lambda_2^n) + E_2(\lambda_1^n, \lambda_2^{n+1})$$

Aplicando Euler-Lagrange obtenemos:

$$\begin{aligned} \frac{\partial E}{\partial \lambda_1^{n+1}}(\lambda_1^{n+1}, \lambda_2^{n+1}) &= - \left( I_1 - I_2^{\lambda_1^{n+1}} \right) \nabla I_2^{\lambda_1^{n+1}} \begin{pmatrix} \frac{-b_1}{\sqrt{a_1^2 + b_1^2}} \\ \frac{a_1}{\sqrt{a_1^2 + b_1^2}} \end{pmatrix} - C \operatorname{div}(\mathbf{D}(\nabla I_1) \nabla \lambda_1) \\ &+ \beta \Psi' \left( \left\| \vec{h}_1^{n+1} + (\vec{h}_2^n)^{\lambda_1^{n+1}} \right\|^2 \right) \left( \mathbf{Id} + \nabla (\vec{h}_2^n)^{\lambda_1^{n+1}} \right) \\ &\cdot \begin{pmatrix} \frac{-b_2}{\sqrt{a_2^2 + b_2^2}} \\ \frac{a_2}{\sqrt{a_2^2 + b_2^2}} \end{pmatrix} \left( \vec{h}_1^{n+1} + (\vec{h}_2^n)^{\lambda_1^{n+1}} \right) \end{aligned} \quad (10.3.1)$$

Para la primera energía y

$$\begin{aligned} \frac{\partial E}{\partial \lambda_2^{n+1}}(\lambda_1^{n+1}, \lambda_2^{n+1}) &= - \left( I_2 - I_1^{\lambda_2^{n+1}} \right) \nabla I_1^{\lambda_2^{n+1}} \begin{pmatrix} \frac{-b_2}{\sqrt{a_2^2 + b_2^2}} \\ \frac{a_2}{\sqrt{a_2^2 + b_2^2}} \end{pmatrix} - C \operatorname{div}(\mathbf{D}(\nabla I_2) \nabla \lambda_2) \\ &+ \beta \Psi' \left( \left\| \vec{h}_2^{n+1} + (\vec{h}_1^n)^{\lambda_2^{n+1}} \right\|^2 \right) \left( \mathbf{Id} + \nabla (\vec{h}_1^n)^{\lambda_2^{n+1}} \right) \\ &\cdot \begin{pmatrix} \frac{-b_2}{\sqrt{a_2^2 + b_2^2}} \\ \frac{a_2}{\sqrt{a_2^2 + b_2^2}} \end{pmatrix} \left( \vec{h}_2^{n+1} + (\vec{h}_1^n)^{\lambda_2^{n+1}} \right) \end{aligned} \quad (10.3.2)$$

Obtenemos una solución para estas ecuaciones a través de un método de descenso por gradiente obteniendo el siguiente sistema de ecuaciones diferenciales en derivadas parciales:

$$\begin{aligned}
 \frac{\partial \lambda_1^{n+1}}{\partial t} &= C \operatorname{div} (\mathbf{D}(\nabla I_1) \nabla \lambda_1^{n+1}) + (I_1 - I_2^{\lambda_1^{n+1}}) \nabla I_2^{\lambda_2^{n+1}} \begin{pmatrix} \frac{-b_1}{\sqrt{a_1^2 + b_1^2}} \\ \frac{a_1}{\sqrt{a_1^2 + b_1^2}} \end{pmatrix} \\
 &\quad - \beta \Psi' \left( \left\| \vec{h}_1^{n+1} + (\vec{h}_2^n) \vec{h}_1^{n+1} \right\|^2 \right) \\
 &\quad \cdot \left[ \begin{pmatrix} 1 + \left( \frac{\partial u_2^n}{\partial x} \right)^{\lambda_1^{n+1}} \\ \left( \frac{\partial u_2^n}{\partial y} \right)^{\lambda_1^{n+1}} \end{pmatrix}^T \begin{pmatrix} \frac{-b_1}{\sqrt{a_1^2 + b_1^2}} \\ \frac{a_1}{\sqrt{a_1^2 + b_1^2}} \end{pmatrix} (u_1^{n+1} + (u_2^n)^{\lambda_1^{n+1}}) \right. \\
 &\quad \left. + \begin{pmatrix} \left( \frac{\partial v_2^n}{\partial x} \right)^{\lambda_1^{n+1}} \\ 1 + \left( \frac{\partial v_2^n}{\partial y} \right)^{\lambda_1^{n+1}} \end{pmatrix}^T \begin{pmatrix} \frac{-b_1}{\sqrt{a_1^2 + b_1^2}} \\ \frac{a_1}{\sqrt{a_1^2 + b_1^2}} \end{pmatrix} (v_1^{n+1} + (v_2^n)^{\lambda_1^{n+1}}) \right] \quad (10.3.3)
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial \lambda_2^{n+1}}{\partial t} &= C \operatorname{div} (\mathbf{D}(\nabla I_2) \nabla \lambda_2^{n+1}) + (I_2 - I_1^{\lambda_2^{n+1}}) \nabla I_1^{\lambda_1^{n+1}} \begin{pmatrix} \frac{-b_2}{\sqrt{a_2^2 + b_2^2}} \\ \frac{a_2}{\sqrt{a_2^2 + b_2^2}} \end{pmatrix} \\
 &\quad - \beta \Psi' \left( \left\| \vec{h}_2^{n+1} + (\vec{h}_1^n) \vec{h}_2^{n+1} \right\|^2 \right) \\
 &\quad \cdot \left[ \begin{pmatrix} 1 + \left( \frac{\partial u_1^n}{\partial x} \right)^{\lambda_2^{n+1}} \\ \left( \frac{\partial u_1^n}{\partial y} \right)^{\lambda_2^{n+1}} \end{pmatrix}^T \begin{pmatrix} \frac{-b_2}{\sqrt{a_2^2 + b_2^2}} \\ \frac{a_2}{\sqrt{a_2^2 + b_2^2}} \end{pmatrix} (u_2^{n+1} + (u_1^n)^{\lambda_2^{n+1}}) \right. \\
 &\quad \left. + \begin{pmatrix} \left( \frac{\partial v_1^n}{\partial x} \right)^{\lambda_2^{n+1}} \\ 1 + \left( \frac{\partial v_1^n}{\partial y} \right)^{\lambda_2^{n+1}} \end{pmatrix}^T \begin{pmatrix} \frac{-b_2}{\sqrt{a_2^2 + b_2^2}} \\ \frac{a_2}{\sqrt{a_2^2 + b_2^2}} \end{pmatrix} (v_2^{n+1} + (v_1^n)^{\lambda_2^{n+1}}) \right] \quad (10.3.4)
 \end{aligned}$$

## 10.4 Esquema numérico

Vamos a describir un algoritmo eficiente para nuestro modelo simétrico. Discretizamos el sistema parabólico formado por las ecuaciones (10.3.3) y (10.3.4) por diferencias finitas. Todas las derivadas espaciales se aproximan por diferencias centradas, y para la discretización en la dirección de  $t$  utilizamos un esquema lineal implícito. Discretizamos las derivadas de primer orden,  $\frac{\partial u}{\partial x}$  y  $\frac{\partial u}{\partial y}$ , utilizando máscaras invariantes por rotaciones<sup>1</sup> de 45°.

Definimos las posiciones de los tensores de difusión para cada dirección de la siguiente manera

$$\mathbf{D}(\nabla I_1^\sigma) = \begin{pmatrix} d & e \\ e & f \end{pmatrix} \text{ y } \mathbf{D}(\nabla I_2^\sigma) = \begin{pmatrix} d' & e' \\ e' & f' \end{pmatrix}$$

A continuación vamos a ver un esquema numérico para este modelo. Empezamos con la discretización de  $\lambda_1$ :

$$\frac{\lambda_{1,i,j}^{k+1} - \lambda_{1,i,j}^k}{\tau} = T_l + T_r + T_s \quad (10.4.1)$$

Estudiamos cada término por separado: Linealizamos el término de ligadura utilizando los desarrollos en serie de Taylor de primer orden,

$$T_l = \left( I_1(\mathbf{x}_{i,j}) - I_2^{\lambda_{1,i,j}^k}(\mathbf{x}_{i,j}) - \frac{\lambda_{1,i,j}^{k+1} - \lambda_{1,i,j}^k}{\sqrt{a_{1,i,j}^2 + b_{1,i,j}^2}} \left( a_{1,i,j} I_{2,y}^{\lambda_{1,i,j}^k}(\mathbf{x}_{i,j}) - b_{1,i,j} I_{2,x}^{\lambda_{1,i,j}^k}(\mathbf{x}_{i,j}) \right) \right) \\ + \frac{a_{1,i,j} I_{2,y}^{\lambda_{1,i,j}^k}(\mathbf{x}_{1,i,j}) - b_{1,i,j} I_{2,x}^{\lambda_{1,i,j}^k}(\mathbf{x}_{1,i,j})}{\sqrt{a_{1,i,j}^2 + b_{1,i,j}^2}} \quad (10.4.2)$$

---

<sup>1</sup>En el apéndice C se explica detalladamente la forma de discretización para el gradiente y para la divergencia.

Término de regularización,

$$\begin{aligned}
 T_r = C & \left( \frac{d_{i+1,j} + d_{i,j}}{2} \frac{\lambda_{1,i+1,j}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_1^2} + \frac{d_{i-1,j} + d_{i,j}}{2} \frac{\lambda_{1,i-1,j}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_1^2} \right. \\
 & + \frac{f_{i,j+1} + f_{i,j}}{2} \frac{\lambda_{1,i,j+1}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_2^2} + \frac{f_{i,j-1} + f_{i,j}}{2} \frac{\lambda_{1,i,j-1}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_2^2} \\
 & + \frac{e_{i+1,j+1} + e_{i,j}}{2} \frac{\lambda_{1,i+1,j+1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1h_2} + \frac{e_{i-1,j-1} + e_{i,j}}{2} \frac{\lambda_{1,i-1,j-1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1h_2} \\
 & \left. - \frac{e_{i+1,j-1} + e_{i,j}}{2} \frac{\lambda_{1,i+1,j-1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1h_2} - \frac{e_{i-1,j+1} + e_{i,j}}{2} \frac{\lambda_{1,i-1,j+1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1h_2} \right) \quad (10.4.3)
 \end{aligned}$$

Término de simetría,

$$\begin{aligned}
 T_s = -\beta \Psi' & \left( \left\| \vec{h}_1^k + (\vec{h}_2^k) \vec{h}_1^k \right\|^2 \right) \\
 & \cdot \left[ \left( \frac{a_{1,i,j} (u_{2,i,j}^k)_y^{\lambda_{1,i,j}^k} - b_{1,i,j} \left( 1 + (u_{2,i,j}^k)_x^{\lambda_{1,i,j}^k} \right)}{\sqrt{a_{1,i,j}^2 + b_{1,i,j}^2}} \right) \left( u_{1,i,j}^{k+1} + (u_{2,i,j}^k)^{\lambda_{1,i,j}^k} \right) \right. \\
 & \left. + \frac{a_{1,i,j} \left( 1 + (v_{2,i,j}^k)_y^{\lambda_{1,i,j}^k} \right) - b_{1,i,j} (v_{2,i,j}^k)_x^{\lambda_{1,i,j}^k}}{\sqrt{a_{1,i,j}^2 + b_{1,i,j}^2}} + \left( v_{1,i,j}^{k+1} + (v_{2,i,j}^k)^{\lambda_{1,i,j}^k} \right) \right] \quad (10.4.4)
 \end{aligned}$$

Tenemos que hacer notar que  $\lambda_1(x, y) \neq -\lambda_2(x + \lambda_1(x, y), y)$  para el caso general. Sólo se cumple esta igualdad cuando las líneas epipolares son paralelas. En la siguiente sección estudiaremos el caso especial en el que la líneas epipolares son horizontales.

Ahora exponemos el esquema numérico para  $\lambda_2$ : Término de ligadura,

$$\begin{aligned}
T_l = & \left( I_2(\mathbf{x}_{i,j}) - I_1^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) - \frac{\lambda_{2,i,j}^{k+1} - \lambda_{2,i,j}^k}{\sqrt{a_{2,i,j}^2 + b_{2,i,j}^2}} \left( a_{2,i,j} I_{1,y}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) - b_{2,i,j} I_{1,x}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) \right) \right) \\
& \cdot \frac{a_{2,i,j} I_{1,y}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) - b_{2,i,j} I_{1,x}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j})}{\sqrt{a_{2,i,j}^2 + b_{2,i,j}^2}} \quad (10.4.5)
\end{aligned}$$

Término de regularización,

$$\begin{aligned}
T_r = & C \left( \frac{d'_{i+1,j} + d'_{i,j}}{2} \frac{\lambda_{2,i+1,j}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_1^2} + \frac{d'_{i-1,j} + d'_{i,j}}{2} \frac{\lambda_{2,i-1,j}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{f'_{i,j+1} + f'_{i,j}}{2} \frac{\lambda_{2,i,j+1}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_2^2} + \frac{f'_{i,j-1} + f'_{i,j}}{2} \frac{\lambda_{2,i,j-1}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_2^2} \\
& + \frac{e'_{i+1,j+1} + e'_{i,j}}{2} \frac{\lambda_{2,i+1,j+1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} + \frac{e'_{i-1,j-1} + e'_{i,j}}{2} \frac{\lambda_{2,i-1,j-1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{e'_{i+1,j-1} + e'_{i,j}}{2} \frac{\lambda_{2,i+1,j-1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} - \frac{e'_{i-1,j+1} + e'_{i,j}}{2} \frac{\lambda_{2,i-1,j+1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} \right) \quad (10.4.6)
\end{aligned}$$

Término de simetría,

$$\begin{aligned}
T_s = & -\beta \Psi' \left( \left\| \vec{h}_2^k + (\vec{h}_1^k) \vec{h}_2^k \right\|^2 \right) \\
& \cdot \left[ \left( \frac{a_{2,i,j} (u_{1,i,j}^k)_y^{\lambda_{2,i,j}^k} - b_{2,i,j} \left( 1 + (u_{1,i,j}^k)_x^{\lambda_{2,i,j}^k} \right)}{\sqrt{a_{2,i,j}^2 + b_{2,i,j}^2}} \right) \left( u_{2,i,j}^{k+1} + (u_{1,i,j}^k)^{\lambda_{2,i,j}^k} \right) \right. \\
& \left. + \frac{a_{2,i,j} \left( 1 + (v_{1,i,j}^k)_y^{\lambda_{2,i,j}^k} \right) - b_{2,i,j} (v_{1,i,j}^k)_x^{\lambda_{2,i,j}^k}}{\sqrt{a_{2,i,j}^2 + b_{2,i,j}^2}} + \left( v_{2,i,j}^{k+1} + (v_{1,i,j}^k)^{\lambda_{2,i,j}^k} \right) \right] \quad (10.4.7)
\end{aligned}$$

Si observamos con detenimiento el término de simetría, vemos que éste no está expresado de forma implícita. Esto se debe a que la expresión no es lineal en las variables  $\lambda_1$  y  $\lambda_2$ , siendo además compleja su linealización a través de los desarrollos en serie de Taylor.

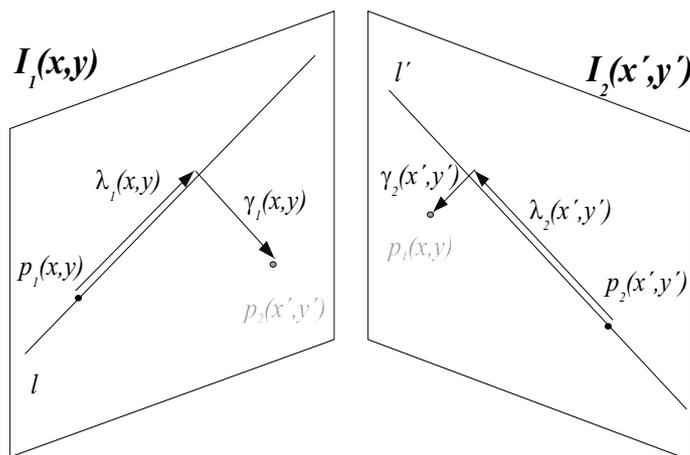


Figura 10.1: **Geometría epipolar arbitraria:** Generalmente, cuando se adquiere un par estéreo, se genera una geometría epipolar arbitraria. Esto lleva a que normalmente no se satisfaga la igualdad  $\lambda_1(x, y) = -\lambda_2(x + \lambda_1(x, y), y)$ . En este caso se satisface  $\sqrt{\lambda_1^2(\mathbf{x}) + \gamma_1^2(\mathbf{x})} = \sqrt{\lambda_2^2(\mathbf{x} + \vec{h}_1) + \gamma_2^2(\mathbf{x} + \vec{h}_1)}$ .

Este modelo tiene en cuenta cualquier tipo de geometría epipolar (ver figura 10.1). Si la geometría epipolar generase líneas epipolares paralelas (figura 10.2), el modelo se simplificaría bastante ya que  $\gamma_i(x, y)$  se anularía en todos los casos.

El objetivo de la siguiente sección será el de proponer un nuevo modelo en el que las líneas epipolares sí sean paralelas (figura 10.2), minimizando la diferencia entre  $\lambda_1$  y  $\lambda_2$ .

## 10.5 Simplificación del modelo

Se puede simplificar el modelo anterior realizando una rectificación del par estéreo. Gracias a la información de las matrices de proyección, podemos proyectar las dos imágenes en un plano paralelo de forma que las líneas epipolares sean fronto-paralelas.

En la sección 3.3 se detalla el proceso de rectificación de imágenes que hemos utilizado en nuestras pruebas.

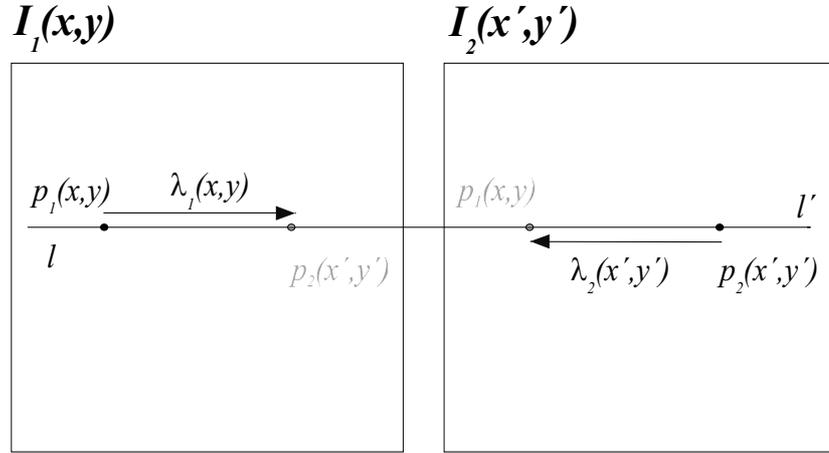


Figura 10.2: **Imágenes frontoparalelas:** Este tipo de imágenes cumple que sus líneas epipolares son paralelas (en este caso horizontales). Este tipo de imágenes se suele utilizar mucho ya que satisface  $\lambda_1(x, y) = -\lambda_2(x + \lambda_1(x, y), y)$ , lo que simplifica mucho los modelos a tratar.

### 10.5.1 El modelo de energía

Esta rectificación nos va a permitir minimizar la energía teniendo en cuenta un desplazamiento en una sola dimensión. Esta simplificación afecta al término de simetría de la siguiente manera:

$$\begin{aligned}
 E_1(\lambda_1, \lambda_2) &= \frac{1}{2} \int_{\Omega} \left( I_1(\mathbf{x}) - I_2(\mathbf{x} + \vec{h}_1(\mathbf{x})) \right)^2 + \frac{C}{2} \int_{\Omega} \nabla \lambda_1 \mathbf{D}(\nabla I_1) \nabla \lambda_1 \\
 &\quad + \frac{\beta}{2} \int_{\Omega} \Psi \left( \|\lambda_1(\mathbf{x}) + \lambda_2(\mathbf{x} + \lambda_1(\mathbf{x}))\|^2 \right)
 \end{aligned} \tag{10.5.1}$$

Esta energía minimiza la disparidad desde la imagen  $I_1$  a  $I_2$ . Análogamente, tenemos que la energía en el sentido inverso viene dada por:

$$\begin{aligned}
 E_2(\lambda_1, \lambda_2) &= \frac{1}{2} \int_{\Omega} \left( I_2(\mathbf{x}) - I_1(\mathbf{x} + \vec{h}_2(\mathbf{x})) \right)^2 + \frac{C}{2} \int_{\Omega} \nabla \lambda_2 \mathbf{D}(\nabla I_2) \nabla \lambda_2 \\
 &\quad + \frac{\beta}{2} \int_{\Omega} \Psi \left( \|\lambda_2(\mathbf{x}) + \lambda_1(\mathbf{x} + \lambda_2(\mathbf{x}))\|^2 \right)
 \end{aligned} \tag{10.5.2}$$

La energía simétrica está dada por la suma de las energías anteriores

$$E(\lambda_1, \lambda_2) = E_1(\lambda_1, \lambda_2) + E_2(\lambda_1, \lambda_2)$$

### 10.5.2 Minimizando la energía

Con el fin de converger a un mínimo de esta energía, procedemos de forma iterativa como sigue: primero, empezamos con una aproximación inicial  $(\lambda_1^0, \lambda_2^0)$  y luego obtenemos cada  $(\lambda_1^{n+1}, \lambda_2^{n+1})$  a partir de la anterior  $(\lambda_1^n, \lambda_2^n)$  como una iteración de un método de descenso por gradiente

$$E(\lambda_1^{n+1}, \lambda_2^{n+1}) = E_1(\lambda_1^{n+1}, \lambda_2^n) + E_2(\lambda_1^n, \lambda_2^{n+1})$$

Minimizando la energía compuesta llegamos al siguiente sistema de ecuaciones en derivadas parciales:

$$\begin{aligned} \frac{\partial E}{\partial \lambda_1^{n+1}}(\lambda_1^{n+1}, \lambda_2^{n+1}) &= - \left( I_1 - I_2^{\lambda_1^{n+1}} \right) \frac{\partial I_2^{\lambda_1^{n+1}}}{\partial \lambda_1^{n+1}} - C \operatorname{div}(\mathbf{D}(\nabla I_1) \nabla \lambda_1) \\ &+ \beta \Psi' \left( \left\| \lambda_1^{n+1} + (\lambda_2^{n+1}) \lambda_1^{n+1} \right\|^2 \right) \left( 1 + \frac{\partial (\lambda_2^n) \lambda_1^{n+1}}{\partial x} \right) \\ &\cdot \left( \lambda_1^{n+1} + (\lambda_2^n) \lambda_1^{n+1} \right) \end{aligned} \quad (10.5.3)$$

Para la primera energía y

$$\begin{aligned} \frac{\partial E}{\partial \lambda_2^{n+1}}(\lambda_1^{n+1}, \lambda_2^{n+1}) &= - \left( I_2 - I_1^{\lambda_2^{n+1}} \right) \frac{\partial I_1^{\lambda_2^{n+1}}}{\partial \lambda_2^{n+1}} - C \operatorname{div}(\mathbf{D}(\nabla I_2) \nabla \lambda_2) \\ &+ \beta \Psi' \left( \left\| \lambda_2^{n+1} + (\lambda_1^n) \lambda_2^{n+1} \right\|^2 \right) \left( 1 + \frac{\partial (\lambda_1^n) \lambda_2^{n+1}}{\partial x} \right) \\ &\cdot \left( \lambda_2^{n+1} + (\lambda_1^n) \lambda_2^{n+1} \right) \end{aligned} \quad (10.5.4)$$

Aplicando un método de descenso por gradiente para el cálculo del mínimo,

$$\begin{aligned}
\frac{\partial \lambda_1^{n+1}}{\partial t} &= C \operatorname{div} (\mathbf{D}(\nabla I_1) \nabla \lambda_1^{n+1}) + (I_1 - I_2^{\lambda_1^{n+1}}) \nabla I_2^{\lambda_2^{n+1}} \begin{pmatrix} \frac{-b_1}{\sqrt{a_1^2 + b_1^2}} \\ \frac{a_1}{\sqrt{a_1^2 + b_1^2}} \end{pmatrix} \\
&\quad - \beta \Psi' \left( \left\| \lambda_1^{n+1} + (\lambda_2^{n+1})^{\lambda_1^{n+1}} \right\|^2 \right) \left( 1 + \frac{\partial (\lambda_2^n)^{\lambda_1^{n+1}}}{\partial x} \right) \left( \lambda_1^{n+1} + (\lambda_2^n)^{\lambda_1^{n+1}} \right)
\end{aligned} \tag{10.5.5}$$

$$\begin{aligned}
\frac{\partial \lambda_2^{n+1}}{\partial t} &= C \operatorname{div} (\mathbf{D}(\nabla I_2) \nabla \lambda_2^{n+1}) + (I_2 - I_1^{\lambda_2^{n+1}}) \nabla I_1^{\lambda_1^{n+1}} \begin{pmatrix} \frac{-b_2}{\sqrt{a_2^2 + b_2^2}} \\ \frac{a_2}{\sqrt{a_2^2 + b_2^2}} \end{pmatrix} \\
&\quad - \beta \Psi' \left( \left\| \lambda_2^{n+1} + (\lambda_1^n)^{\lambda_2^{n+1}} \right\|^2 \right) \left( 1 + \frac{\partial (\lambda_1^n)^{\lambda_2^{n+1}}}{\partial x} \right) \left( \lambda_2^{n+1} + (\lambda_1^n)^{\lambda_2^{n+1}} \right)
\end{aligned} \tag{10.5.6}$$

### 10.5.3 Esquema numérico

Vamos a describir un algoritmo eficiente para nuestro modelo de simétrico. Discretizamos el sistema parabólico formado por las ecuaciones (10.5.5) y (10.5.6) por diferencias finitas. Todas las derivadas espaciales se aproximan por diferencias centradas, y para la discretización en la dirección de  $t$  utilizamos un esquema lineal implícito.

A continuación vamos a ver un esquema numérico para este modelo. Empezamos con la discretización de  $\lambda_1$ :

$$\frac{\lambda_{1,i,j}^{k+1} - \lambda_{1,i,j}^k}{\tau} = T_l + T_r + T_s \tag{10.5.7}$$

Estudiamos cada término por separado: Linealizamos el término de ligadura utilizando los desarrollos en serie de Taylor de primer orden,

$$\begin{aligned}
 T_l = & \left( I_1(\mathbf{x}_{i,j}) - I_2^{\lambda_{1,i,j}^k}(\mathbf{x}_{i,j}) - \frac{\lambda_{1,i,j}^{k+1} - \lambda_{1,i,j}^k}{\sqrt{a_{1,i,j}^2 + b_{1,i,j}^2}} \left( a_{1,i,j} I_{2,y}^{\lambda_{1,i,j}^k}(\mathbf{x}_{i,j}) - b_{1,i,j} I_{2,x}^{\lambda_{1,i,j}^k}(\mathbf{x}_{i,j}) \right) \right) \\
 & \cdot \frac{a_{1,i,j} I_{2,y}^{\lambda_{1,i,j}^k}(\mathbf{x}_{1,i,j}) - b_{1,i,j} I_{2,x}^{\lambda_{1,i,j}^k}(\mathbf{x}_{1,i,j})}{\sqrt{a_{1,i,j}^2 + b_{1,i,j}^2}} \quad (10.5.8)
 \end{aligned}$$

Término de regularización,

$$\begin{aligned}
 T_r = & C \left( \frac{d_{i+1,j} + d_{i,j}}{2} \frac{\lambda_{1,i+1,j}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_1^2} + \frac{d_{i-1,j} + d_{i,j}}{2} \frac{\lambda_{1,i-1,j}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_1^2} \right. \\
 & + \frac{f_{i,j+1} + f_{i,j}}{2} \frac{\lambda_{1,i,j+1}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_2^2} + \frac{f_{i,j-1} + f_{i,j}}{2} \frac{\lambda_{1,i,j-1}^{k+1} - \lambda_{1,i,j}^{k+1}}{h_2^2} \\
 & + \frac{e_{i+1,j+1} + e_{i,j}}{2} \frac{\lambda_{1,i+1,j+1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1 h_2} + \frac{e_{i-1,j-1} + e_{i,j}}{2} \frac{\lambda_{1,i-1,j-1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1 h_2} \\
 & \left. - \frac{e_{i+1,j-1} + e_{i,j}}{2} \frac{\lambda_{1,i+1,j-1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1 h_2} - \frac{e_{i-1,j+1} + e_{i,j}}{2} \frac{\lambda_{1,i-1,j+1}^{k+1} - \lambda_{1,i,j}^{k+1}}{2h_1 h_2} \right) \quad (10.5.9)
 \end{aligned}$$

Término de simetría,

$$T_s = -\beta \Psi' \left( \left\| \lambda_{1,i,j}^k + (\lambda_{2,i,j}^k)^{\lambda_{1,i,j}^k} \right\|^2 \right) \left( 1 + (\lambda_{2,i,j}^k)^{\lambda_{1,i,j}^k} \right) \left( \lambda_{1,i,j}^{k+1} + (\lambda_{2,i,j}^k)^{\lambda_{1,i,j}^k} \right) \quad (10.5.10)$$

Tenemos que hacer notar que en este caso sí se cumple que  $\lambda_1(x, y) = -\lambda_2(x + \lambda_1(x, y), y)$  ya que las líneas epipolares son paralelas.

Ahora exponemos el esquema numérico para  $\lambda_2$ : Término de ligadura,

$$\begin{aligned}
T_l = & \left( I_2(\mathbf{x}_{i,j}) - I_1^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) - \frac{\lambda_{2,i,j}^{k+1} - \lambda_{2,i,j}^k}{\sqrt{a_{2,i,j}^2 + b_{2,i,j}^2}} \left( a_{2,i,j} I_{1,y}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) - b_{2,i,j} I_{1,x}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) \right) \right) \\
& \cdot \frac{a_{2,i,j} I_{1,y}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j}) - b_{1,i,j} I_{1,x}^{\lambda_{2,i,j}^k}(\mathbf{x}_{i,j})}{\sqrt{a_{2,i,j}^2 + b_{2,i,j}^2}} \quad (10.5.11)
\end{aligned}$$

Término de regularización,

$$\begin{aligned}
T_r = & C \left( \frac{d'_{i+1,j} + d'_{i,j}}{2} \frac{\lambda_{2,i+1,j}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_1^2} + \frac{d'_{i-1,j} + d'_{i,j}}{2} \frac{\lambda_{2,i-1,j}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_1^2} \right. \\
& + \frac{f'_{i,j+1} + f'_{i,j}}{2} \frac{\lambda_{2,i,j+1}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_2^2} + \frac{f'_{i,j-1} + f'_{i,j}}{2} \frac{\lambda_{2,i,j-1}^{k+1} - \lambda_{2,i,j}^{k+1}}{h_2^2} \\
& + \frac{e'_{i+1,j+1} + e'_{i,j}}{2} \frac{\lambda_{2,i+1,j+1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} + \frac{e'_{i-1,j-1} + e'_{i,j}}{2} \frac{\lambda_{2,i-1,j-1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} \\
& \left. - \frac{e'_{i+1,j-1} + e'_{i,j}}{2} \frac{\lambda_{2,i+1,j-1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} - \frac{e'_{i-1,j+1} + e'_{i,j}}{2} \frac{\lambda_{2,i-1,j+1}^{k+1} - \lambda_{2,i,j}^{k+1}}{2h_1h_2} \right) \quad (10.5.12)
\end{aligned}$$

Término de simetría,

$$T_s = -\beta \Psi' \left( \left\| \lambda_{2,i,j}^k + (\lambda_{1,i,j}^k)^{\lambda_{2,i,j}^k} \right\|^2 \right) \left( 1 + (\lambda_{1,i,j}^k)^{\lambda_{2,i,j}^k} \right) \left( \lambda_{2,i,j}^{k+1} + (\lambda_{1,i,j}^k)^{\lambda_{2,i,j}^k} \right) \quad (10.5.13)$$

## 10.6 Resultados numéricos

La primera prueba que presentamos a continuación es para el par estéreo del pasillo. En la figura 10.3 se muestra el resultado del método simétrico para cámaras frontoparalelas. Como se puede apreciar, el resultado es bastante parecido en ambos sentidos.

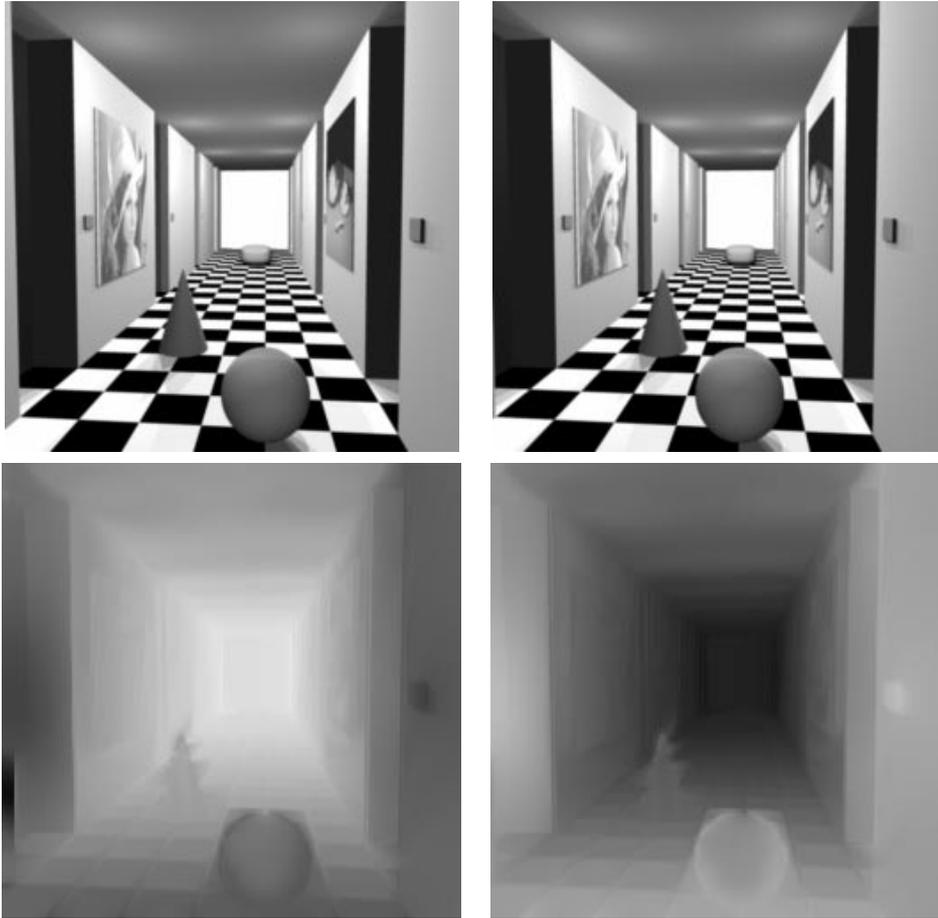


Figura 10.3: Resultado de aplicar el método estéreo simétrico dado por el sistema 10.5.5 y 10.5.6. En la imagen inferior izquierda se muestra el resultado para  $\lambda_1$  y en la derecha para  $\lambda_2$ . Parámetros:  $zoom = 4$ ,  $\alpha = 0.3$ ,  $\beta = 0.1$  y  $s = 0.15$ . Se utilizó la función  $\Psi$  lineal.

El segundo ejemplo que proponemos es el de un par real rectificado. En la figura 10.4 se muestra el resultado para el par estéreo de Javier.

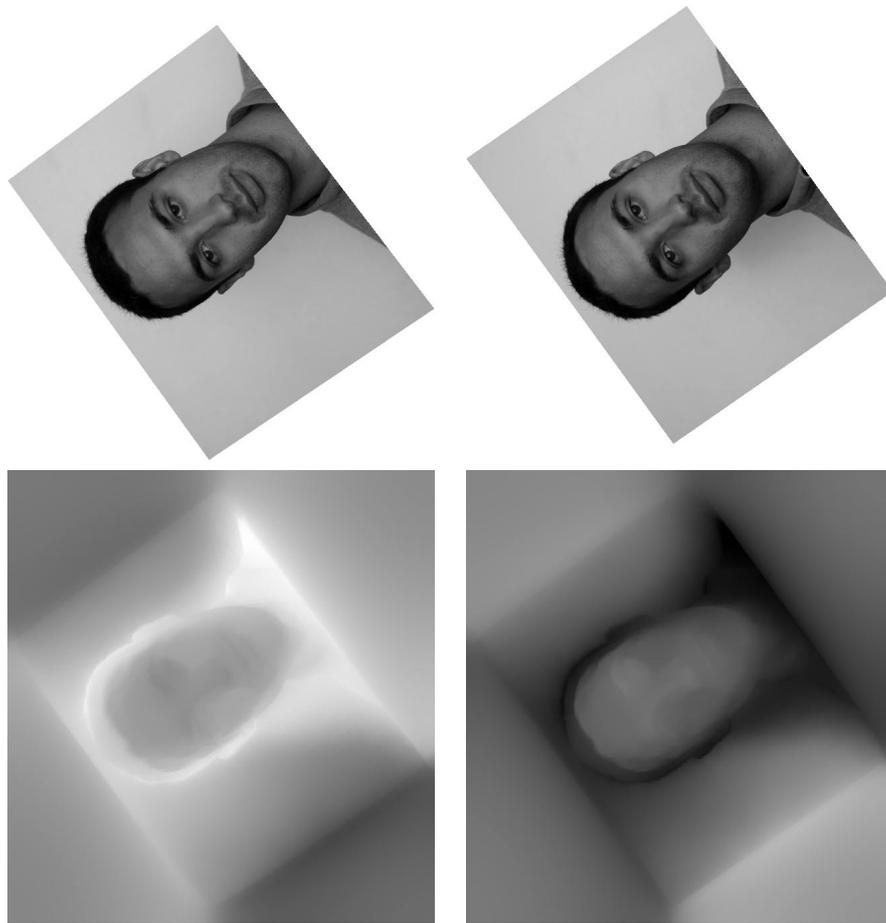


Figura 10.4: Resultado de aplicar el método estéreo simétrico dado por el sistema 10.5.5 y 10.5.6. En la imagen inferior izquierda se muestra el resultado para  $\lambda_1$  y en la derecha para  $\lambda_2$ .

# Capítulo 11

## Conclusiones

Normalmente, cuando los investigadores en visión por ordenador tratan con métodos variacionales para el cálculo del flujo óptico, suelen linealizar la restricción del flujo óptico. Excepto para aquellos casos en los que las imágenes varían suficientemente despacio en el espacio, la linealización solo funciona para desplazamientos pequeños. En este trabajo hemos introducido tres mejoras dentro de un método clásico por Nagel y Enkelmann donde no se ha utilizado linealización. Identificamos este método con filtros de difusión anisótropa ligados con un término de reacción no lineal. Demostramos que este sistema parabólico está bien definido desde un punto de vista matemático, y presentamos un esquema lineal implícito en diferencias finitas para encontrar una solución numérica eficiente. Con el fin de evitar que el algoritmo converja a mínimos locales irrelevantes, introducimos nuestro método en una estrategia de análisis multiescala y multipiramidal para buscar la solución desde una escala más gruesa a otra más fina. Un análisis cuantitativo detallado utilizando secuencias de prueba con información exacta acerca de su disparidad mostró los siguientes resultados.

- El método puede recuperar desplazamientos de más de 10 pixels con una buena precisión.
- Es significativamente mejor que el método original de Nagel y todos los demás métodos 100% densos que se evalúan en Barron et al. [BFB94].
- El método es robusto. La resolución permanece estable a pesar del gran rango de valores a que se sometieron los parámetros. Esto permite utilizar valores por defecto para los parámetros para muchas aplicaciones.

A pesar de estas propiedades favorables, todavía queda sitio para futuras mejoras y extensiones de este algoritmo. Por ejemplo, esperamos que nuestro método se pueda beneficiar de los resultados obtenidos de las imágenes anteriores. Se puede interpolar

el flujo de los frames precedentes con el fin de tener una primera aproximación del flujo en el frame actual [BA91]. Si esta primera estimación es razonablemente buena para nuestro algoritmo, no haría falta introducir grandes escalas de corrección utilizando grandes valores para  $\sigma_0$  o realizando muchos zoom. Un valor inicial más pequeño para  $\sigma_0$  aceleraría el proceso de enfoque y haría que el algoritmo fuese más rápido.

Originalmente empezamos con el problema del cálculo del flujo óptico y lo extendimos al problema de la reconstrucción 3D a partir de pares estéreos en niveles de grises y en color [ADSW00, AS00a]. Esperamos que nuestro método que combina ecuaciones de difusión-reacción anisotrópicas con técnicas de análisis multiescala lineal, pueda servir como motivación para el estudio de otras combinaciones de aproximaciones multiescala lineales y no lineales para resolver problemas de visión por ordenador.

Con respecto al problema de la estimación de la carta de disparidad densa entre dos imágenes, partimos del modelo que propusimos para el flujo óptico. Gracias a que la geometría epipolar es conocida, nos permite simplificar el modelo de energía así como la ecuación de Euler-Lagrange asociada. Utilizamos un esquema lineal implícito por diferencias finitas para realizar una aproximación numérica eficiente. De igual manera, con el fin de evitar que el algoritmo converja a mínimos locales irrelevantes, lo introducimos dentro de una técnica de análisis multiescala o multipiramidal para encontrar la solución desde una escala más gruesa a otra más fina.

Nuestro método combina algunas técnicas desarrolladas en el contexto de la estimación del flujo óptico [AWS99, NE86] con otras técnicas desarrolladas en el contexto de la estimación de la carta de disparidad [RD96] que tiene en cuenta las restricciones geométricas asociadas a un par estéreo. Creemos que la combinación de estas ideas es fructífera en el sentido que produce nuevas herramientas para estimar campos de disparidad densos que se benefician de los esfuerzos de la investigación tanto en visión estéreo como en la estimación del flujo óptico.

Los resultados experimentales obtenidos con nuestro método son muy prometedores. En el caso de un par estéreo sintético donde se conoce de forma exacta la carta de disparidad, demostramos que los resultados obtenidos mejoran los que se obtienen con una técnica de correlación clásica. También hemos evaluado nuestro método en pares estéreos de imágenes reales. En estos casos hemos reconstruido la geometría 3D a partir de la disparidad calculada a través de nuestro método. Los resultados que se obtuvieron con éstas son también buenos. Incluso en las regiones de oclusiones los artefactos son moderados gracias al efecto del término de regularización.

La incorporación de la simetría en el cálculo del flujo nos permite obtener una solución congruente en ambas direcciones. Los puntos que en los métodos directos se encontraban con alta precisión, en los métodos simétricos se mantienen. En las zonas homogéneas se pueden observar mejoras considerables con respecto a los métodos

estándar. En las zonas con mucha textura, también, como es el caso del suelo para la prueba de las Torres de Mármol, se obtiene una solución más regular que con respecto a los otros métodos. Una última conclusión que se deriva de las pruebas numéricas, es que la solución en el método simétrico varía menos de una escala a otra, consiguiendo que la velocidad de convergencia a la solución final sea más rápida que con respecto al método en una sola dirección.

Para recuperar grandes desplazamientos hemos introducido nuestros métodos dentro de una estrategia de análisis multiescala y también dentro de una estrategia multipiramidal. Aunque no hemos realizado comparaciones numéricas entre ambas aproximaciones podemos concluir de forma general que el análisis multiescala da unos resultados más precisos que el multipiramidal, en contraposición a la velocidad en la que la estrategia multipiramidal produce tiempos muy inferiores. Podemos ver la estrategia multipiramidal como un caso especial de la aproximación multiescala en la que se escoge un ratio de descenso entorno a  $\frac{1}{2}$  – descontando el efecto de la reducción de las imágenes. Por último mencionar que la estrategia de análisis multiescala es menos sensible a la elección de los parámetros. El único parámetro que diferencia ambas aproximaciones es la escala:  $\sigma$ , la desviación estándar en el caso multiescala y el número de zooms en el caso multipiramidal. Eligiendo  $\sigma$  como el mayor desplazamiento que existe en la imagen es suficiente para poder estimar todos los desplazamientos, sin embargo, elegir el número de zooms no es tan trivial. Si elegimos un valor muy grande (muchos zooms sobre las originales), podemos provocar que al ir subiendo de escala, por el propio mecanismo para pasar de una escala a otra, se vaya magnificando el efecto del error, y si elegimos un valor pequeño, es probable que no se encuentren los desplazamientos más grandes. Si utilizamos alguna técnica de correlación para obtener una primera estimación del flujo, también es más probable que los errores que ésta produzca se propaguen más en el método multipiramidal.

# Capítulo 12

## Trabajo futuro

Durante esta tesis se nos han planteado otros trabajos que se podrían desarrollar para mejorar los métodos propuestos, pero que por cuestiones de tiempo y esfuerzo no hemos podido presentar. Entre estos podemos destacar:

**Acelerar el método.** Nuestros métodos están basados en técnicas de optimización que utilizan esquemas numéricos iterativos. Durante el desarrollo de los programas no nos hemos centrado mucho en la optimización con respecto al tiempo. Para el método estéreo se puede mejorar el tiempo rectificando las imágenes, como ya hemos hecho en el caso simétrico. Suponemos que por la naturaleza de nuestros algoritmos es un problema complejo. Un primer paso sería el de optimizar el código que hemos desarrollado. En el caso de los métodos simétricos se podría estudiar algún mecanismo para paralelizar el cálculo del flujo en cada sentido. Otra posibilidad sería la de introducir restricciones en las búsquedas dependiendo de las características de las imágenes.

Modelo que considere una **suposición no Lambertiana**. Hemos supuesto que las superficies que vamos a tratar reflejan la luz de forma difusa y uniforme. Esto no suele ser cierto en la mayoría de las imágenes reales. El inconveniente que implica este efecto es que el término de ligadura que hemos propuesto en nuestros modelos no es un buen estimador para buscar correspondencias en muchos casos. Se debería proponer un término más complejo que tenga en cuenta mayor información acerca de las propiedades de reflectancia de las superficies de la escena y de la posición de las fuentes de luz que inciden sobre los objetos.

Desde el principio hemos trabajado con imágenes que representan escenas reales que puede percibir cualquier persona. El fundamento de nuestro modelo es fácilmente extrapolable a cualquier tipo de imágenes, tales como imágenes médicas. Para esto tendríamos que proponer **otras formas para el término de ligadura**, que sean capaz de definir la transformación de los niveles de intensidad de una imagen a otra.

**Desplazamientos grandes en imágenes complejas.** Nuestro método está clasificado dentro de los métodos diferenciales en los que para buscar las correspondencias se tiene en cuenta la información de los vecinos a través del cálculo del gradiente. Esto implica que el desplazamiento que se va a encontrar será el que le lleve al pixel más próximo que se le parezca, también influenciado por los desplazamientos de los vecinos. Esto suele ser suficiente para muchos tipos de imágenes, sin embargo, cuando nos encontramos con casos más complejos, por ejemplo, objetos cuya textura se pueda confundir con la de otros objetos cercanos o con el fondo, es fácil que nuestro método estime que los pixels se muevan en otro sentido. Esto se debe a que no se tiene en cuenta la forma de los objetos.

Uno de los temas que también nos hemos planteado es el de utilizar modelos más complejos para trabajar con **imágenes en color** de forma que sea más robusta la estimación de los contornos de los objetos de la imagen a través del gradiente.

Una extensión directa del método estéreo es el de **utilizar una serie de imágenes** (más de dos). Por ahora sólo hemos utilizado dos imágenes para reconstruir una escena tridimensional. Utilizar más imágenes ayudará a tener mayor información acerca de los puntos 3D, a parte de reducir el efecto de las oclusiones.

Uno de los trabajos que nos gustaría realizar, y que sería una continuación de éste, es el de la **estimación del mapa de disparidad al mismo tiempo que se estima la geometría epipolar** asociada a las cámaras. Hasta ahora, la determinación de la matriz fundamental era un paso previo al cálculo de la disparidad, pero sería interesante que todo el proceso se desarrollara de forma automática. Además, es probable que gracias al mecanismo de puesta en correspondencia, se consigan correspondencias más precisas que mejoren el cálculo de la matriz fundamental.

El **problema de las oclusiones** influye negativamente sobre el flujo final, ya que no solo se produce un error en éstas, sino que lo propagan a otras partes de la imagen. Un método de estimación de correspondencias debería incorporar algún mecanismo para tener en cuenta la información de las oclusiones y evitar que éstas influyan en otras partes. En los métodos simétricos hemos creado un mecanismo para detectar las oclusiones, pero no utilizamos esta información para mejorar el resultado.

Para los métodos simétricos, igual que para los métodos unidireccionales, sería también interesante tratar imágenes en color.

En este trabajo hemos introducido nuestros métodos dentro de un marco de **análisis multiescala y multipiramidal** para obtener grandes desplazamientos. Se podrían realizar comparaciones numéricas entre las soluciones que dan ambas estrategias con el fin de establecer cuantitativamente la diferencia que existe entre ellas.

# Apéndice A

## Invariancia a cambios lineales de intensidad

En este apéndice vamos a estudiar cómo modificar los parámetros del funcional de energía para que la solución sea invariante bajo transformaciones lineales de los niveles de grises de la forma  $(I_1, I_2) \rightarrow (kI_1, kI_2)$ , donde  $k$  es una constante. La elección de los parámetros tiene una fuerte dependencia del contraste de la imagen. Esto es especialmente problemático cuando el método es insertado en una estrategia de análisis multiescala ya que la cantidad de suavizado influye en el rango de contraste de las imágenes regularizadas  $I_l^\sigma$  e  $I_r^\sigma$ .

Nuestra reciente investigación sobre flujo óptico [AWS99] mostró que este problema se puede resolver como sigue. Calculamos  $C$  y  $\lambda$  especificando dos parámetros  $\alpha, s \in (0, 1)$  tal que

$$C = \frac{\alpha}{\max_{\mathbf{x}}(|(\nabla I_1^\sigma)(\mathbf{x})|^2)},$$
$$s = \int_0^\lambda \mathcal{H}_{|\nabla I_1^\sigma|}(z) dz$$

donde  $\mathcal{H}_{|\nabla I_1^\sigma|}(z)$  representa el histograma normalizado de  $|\nabla I_1^\sigma|$ . Llamamos  $s$  el *factor de isotropía*. Cuando  $s \rightarrow 0$ , el operador de difusión se transforma en anisotrópico en todas las posiciones, y cuando  $s \rightarrow 1$ , nos lleva a la difusión isotrópica en todos sitios. En aplicaciones prácticas de nuestro método es suficiente especificar los parámetros  $\alpha$  y  $s$  en vez de  $C$  y  $\lambda$ . Los parámetros  $C$  y  $\lambda$  se ajustan automáticamente al rango dinámico de la imagen en cada paso del proceso multiescala.

# Apéndice B

## Existencia y unicidad de soluciones

En este apéndice demostramos la existencia y unicidad de soluciones de los sistemas parabólicos (5.3.1) y (5.3.2) donde  $\mathbf{D}(\nabla I_1^\sigma)$  es el tensor propuesto por Nagel–Enkelmann dado por (1.4.2). Los parámetros  $C$  y  $\lambda$  pueden ser números reales positivos arbitrarios. En particular, estos se pueden estimar como se describió en el apéndice anterior. Primero, introducimos un marco abstracto donde estudiamos el sistema anterior. Este marco se utiliza luego para demostrar la existencia y unicidad de soluciones.

### B.1 Marco abstracto

Por simplicidad, asumimos que las imágenes están definidas completamente en el espacio  $\mathbb{R}^2$ . Suponemos que las imágenes de entrada  $I_1$  e  $I_2$  pertenecen al espacio funcional  $L^2(\mathbb{R}^2)$ . Sea  $H = L^2(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$ , y, denotamos por  $A : \mathcal{D}(A) \subset H \rightarrow H$  el operador diferencial definido por

$$A(\vec{h}) = -C \begin{pmatrix} \operatorname{div}(\mathbf{D}(\nabla I_1^\sigma) \nabla u_\sigma) \\ \operatorname{div}(\mathbf{D}(\nabla I_1^\sigma) \nabla v_\sigma) \end{pmatrix},$$

donde  $\mathcal{D}(A)$  es el dominio de  $A$ . Si  $I_1 \in L^2(\mathbb{R}^2)$  entonces  $I_1^\sigma \in W^{1,\infty}(\mathbb{R}^2)$ , así que  $\nabla I_1^\sigma$  está acotado y los autovalores del tensor de difusión  $\mathbf{D}(\nabla I_1^\sigma)$  son estrictamente positivos. Por tanto, como  $C > 0$ , el operador  $A(\vec{h})$  es un operador monótono maximal. Para más detalles acerca de los operadores monótonos maximales nos dirigimos a Brezis [Bre73]. Introduzcamos la función  $F : H \rightarrow H$  definida por

$$F(\vec{h}) = \left( I_1^\sigma - I_2^\sigma(\mathbf{Id} + \vec{h}) \right) \nabla I_2^\sigma(\mathbf{Id} + \vec{h}).$$

Entonces el problema de evolución abstracta se puede escribir como

$$\begin{cases} \frac{d\vec{h}_\sigma}{dt} + A\vec{h}_\sigma = F(\vec{h}_\sigma) \text{ in } H, \forall t \in [0, T] \\ \vec{h}_\sigma(0) = \vec{h}^0 \text{ in } H. \end{cases} \quad (\text{B-1})$$

Cualquier solución clásica  $\vec{h}_\sigma \in C^1([0, T]; H) \cap C([0, T]; \mathcal{D}(A))$  de (B-1) está dada por

$$\vec{h}_\sigma(t) = S(t)\vec{h}^0 + \int_0^t S(t-s)F(\vec{h}_\sigma(s))ds, \quad (\text{B-2})$$

donde  $\{S(t)\}_{t>0}$  es la contracción del semigrupo asociado al problema homogéneo.

**Definición.** Decimos que  $h \in C([0, T]; H)$  es una solución generalizada de (B-1) si satisface (B-2).

## B.2 Resultado de Existencia y Unicidad

Con el fin de demostrar la existencia y unicidad, tenemos que establecer primero un lema.

**Lema 1.** Supongamos que  $I_1, I_2 \in L^2(\mathbb{R}^2)$ , entonces  $F$  es Lipschitz-continuo, y la constante Lipschitz  $L$  depende de las funciones  $I_1$  e  $I_2$  y en  $\sigma$ .

### Demostración:

Primero remarcamos que si  $I_1, I_2 \in L^2(\mathbb{R}^2)$ , entonces tenemos en particular que  $I_2^\sigma \in W^{1,\infty}(\mathbb{R}^2)$  y  $I_1^\sigma \in L^\infty(\mathbb{R}^2)$ . Sea  $\vec{h}_1, \vec{h}_2 \in H$ . Para el componente  $i$ -ésimo de

$F(\vec{h}_1) - F(\vec{h}_2)$ ,  $i = 1, 2$ , tenemos la siguiente estimación puntual.

$$\begin{aligned}
|F_i(\vec{h}_1) - F_i(\vec{h}_2)| &= |(I_1^\sigma - I_2^\sigma(\mathbf{Id} + \vec{h}_1))\partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_1) \\
&\quad - (I_1^\sigma - I_2^\sigma(\mathbf{Id} + \vec{h}_2))\partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_2)|, \\
&\leq |I_2^\sigma(\mathbf{Id} + \vec{h}_1)\partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_1) - I_2^\sigma(\mathbf{Id} + \vec{h}_2)\partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_2)| \\
&\quad + |I_1^\sigma| \cdot |\partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_1) - \partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_2)|, \\
&\leq \frac{1}{2} |\partial_i(|I_2^\sigma|^2)(\mathbf{Id} + \vec{h}_1) - \partial_i(|I_2^\sigma|^2)(\mathbf{Id} + \vec{h}_2)| \\
&\quad + \|I_1^\sigma\|_\infty \cdot |\partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_1) - \partial_i I_2^\sigma(\mathbf{Id} + \vec{h}_2)|, \\
&\leq \frac{1}{2} C_{Lip}(\partial_i(|I_2^\sigma|^2)) \cdot |\vec{h}_1 - \vec{h}_2| + \|I_1^\sigma\|_\infty \cdot C_{Lip}(\partial_i I_2^\sigma) \cdot |\vec{h}_1 - \vec{h}_2| \\
&\leq \left( \frac{1}{2} C_{Lip}(\partial_i(|I_2^\sigma|^2)) + \|I_1^\sigma\|_\infty \cdot C_{Lip}(\partial_i I_2^\sigma) \right) \cdot |\vec{h}_1 - \vec{h}_2|,
\end{aligned}$$

donde  $C_{Lip}(f)$  denota la constante Lipschitz de la función  $f$ . Finalmente deducimos que

$$\begin{aligned}
\|F(\vec{h}_1) - F(\vec{h}_2)\|_H &= \|F_1(\vec{h}_1) - F_1(\vec{h}_2)\|_{L^2} + \|F_2(\vec{h}_1) - F_2(\vec{h}_2)\|_{L^2} \\
&\leq \sum_{i=1}^2 \left( \frac{1}{2} C_{Lip}(\partial_i(|I_2^\sigma|^2)) + \|I_1^\sigma\|_\infty \cdot C_{Lip}(\partial_i I_2^\sigma) \right) \cdot \|\vec{h}_1 - \vec{h}_2\|_H.
\end{aligned}$$

Concluimos la demostración del lema imponiendo

$$L = \sum_{i=1}^2 \left( \frac{1}{2} C_{Lip}(\partial_i(|I_2^\sigma|^2)) + \|I_1^\sigma\|_\infty \cdot C_{Lip}(\partial_i I_2^\sigma) \right).$$

Esto demuestra la afirmación.

Ahora podemos establecer la existencia y unicidad del resultado para el problema (6.2.2)-(6.2.3).

**Teorema 2.** *Supongamos que  $I_1, I_2 \in L^2(\mathbb{R}^2)$  entonces, para todo  $\vec{h}^0 \in H$ , existe una solución única generalizada  $\vec{h}_\sigma(t) \in C([0, \infty); H)$  de (6.2.2)-(6.2.3).*

**Demostración:**

Las suposiciones sobre  $I_1$  e  $I_2$  nos permiten aplicar el lema 1. Supongamos que  $\vec{h}_1(t)$  y  $\vec{h}_2(t)$  son soluciones de (B-2) para condiciones iniciales  $\vec{h}_1(0)$  y  $\vec{h}_2(0)$ , entonces tenemos, haciendo uso del hecho que  $A$  es monótono (lo que implica que  $\|S(t)f\|_H \leq \|f\|_H$ ), y de la continuidad Lipschitz de  $F$  el siguiente estimador.

$$\|\vec{h}_1(t) - \vec{h}_2(t)\|_H \leq \|\vec{h}_1(0) - \vec{h}_2(0)\|_H + L \int_0^t \|\vec{h}_1(s) - \vec{h}_2(s)\|_H ds.$$

Aplicando el lema Gronwall–Bellman [Bre73] nos da

$$\|\vec{h}_1(t) - \vec{h}_2(t)\|_H \leq e^{Lt} \cdot \|\vec{h}_1(0) - \vec{h}_2(0)\|_H,$$

lo que implica la unicidad de la solución. Ahora consideremos el espacio de Banach definido por

$$E = \{\vec{h} \in C([0, \infty); H) : \sup_{t \geq 0} \|\vec{h}(t)\|_H e^{-Kt} < \infty\}$$

inducida por la norma  $\|\vec{h}\|_E = \sup_{t \geq 0} \|\vec{h}(t)\|_H e^{-Kt}$ . Sea  $\phi : E \rightarrow C([0, \infty); H)$  definida por

$$\phi(\vec{h})(t) = S(t)\vec{h}^0 + \int_0^t S(t-s)F(\vec{h}(s))ds.$$

Si  $K > L$ , entonces  $\phi(E) \subset E$ , y  $\phi$  es  $\frac{L}{K}$ -Lipschitz ya que

$$\begin{aligned} \|\phi(\vec{h}_1) - \phi(\vec{h}_2)\|_E &= \sup_{t \geq 0} \|\phi(\vec{h}_1)(t) - \phi(\vec{h}_2)(t)\|_H e^{-Kt}, \\ &\leq \sup_{t \geq 0} \int_0^t L \|\vec{h}_1(s) - \vec{h}_2(s)\|_H ds e^{-Kt} \\ &\leq \sup_{t \geq 0} L \|\vec{h}_1 - \vec{h}_2\|_E \cdot e^{-Kt} \int_0^t e^{Ks} ds \\ &\leq \sup_{t \geq 0} \frac{L}{K} \|\vec{h}_1 - \vec{h}_2\|_E \cdot e^{-Kt} (e^{Kt} - 1) \\ &\leq \frac{L}{K} \|\vec{h}_1 - \vec{h}_2\|_E. \end{aligned}$$

Deducimos que  $\phi$  es una contracción, y por el teorema de punto fijo de Banach existe una única  $\vec{h}_\sigma$  tal que  $\phi(\vec{h}_\sigma) = \vec{h}_\sigma$ . Esta es la solución generalizada de (B-1), y la demostración está concluida<sup>1</sup>.

---

<sup>1</sup>La demostración de la existencia y unicidad está basada en suposiciones más bien débiles sobre las imágenes iniciales  $I_1$  e  $I_2$ . Solo suponemos integrabilidad cuadrática. No tienen que ser continuos y pueden estar corrompidos por ruido o por efectos de cuantización, como suele ser común para imágenes del mundo real. El comportamiento de la solución cuando  $\sigma$  tiende hacia 0 es un problema matemático aún por definir. Si  $I_1$  e  $I_2$  no son suficientemente suaves, no podemos esperar un buen comportamiento asintótico cuando  $\sigma$  tiende a 0. Esto sugiere que las imágenes originales  $I_1$  e  $I_2$  deberían ser procesadas siempre por una pequeña cantidad de suavizado Gaussiano. En nuestros experimentos se observa que puede ser una ventaja parar el proceso multiescala cuando  $\sigma$  alcanza el orden del tamaño de un pixel.

# Apéndice C

## Formas de discretización

### C.1 Discretización del gradiente

Para discretizar el gradiente vamos a utilizar la máscara propuesta en [AM94], en donde se propone una discretización invariante por rotaciones de  $45^\circ$ .

Nos fijamos primero en los desarrollos en serie de Taylor de la señal  $u(t, x, y)$ .

Desarrollos para la variable temporal

$$u(t + k, x, y) = u + ku_t + O(k^2) \quad (\text{C-1})$$

$$u(t - k, x, y) = u - ku_t + O(k^2) \quad (\text{C-2})$$

Desarrollos para las variables espaciales –  $h_1$  es el incremento en  $x$  y  $h_2$  en  $y$ :

$$u(t, x + h_1, y) = u + h_1u_x + \frac{h_1^2}{2}u_{xx} + O(h_1^3) \quad (\text{C-3})$$

$$u(t, x - h_1, y) = u - h_1u_x + \frac{h_1^2}{2}u_{xx} + O(h_1^3) \quad (\text{C-4})$$

$$u(t, x, y + h_2) = u + h_2u_y + \frac{h_2^2}{2}u_{yy} + O(h_2^3) \quad (\text{C-5})$$

$$u(t, x, y - h_2) = u - h_2u_y + \frac{h_2^2}{2}u_{yy} + O(h_2^3) \quad (\text{C-6})$$

$$\begin{aligned}
u(t, x + h_1, y + h_2) &= u + h_1 u_x + h_2 u_y + \frac{1}{2} (h_1^2 u_{xx} + 2h_1 h_2 u_{xy} + h_2^2 u_{yy}) \\
&+ \mathcal{O} \left( (h_1^2 + h_2^2)^{\frac{3}{2}} \right)
\end{aligned} \tag{C-7}$$

$$\begin{aligned}
u(t, x - h_1, y - h_2) &= u - h_1 u_x + h_2 u_y + \frac{1}{2} (h_1^2 u_{xx} + 2h_1 h_2 u_{xy} + h_2^2 u_{yy}) \\
&+ \mathcal{O} \left( (h_1^2 + h_2^2)^{\frac{3}{2}} \right)
\end{aligned} \tag{C-8}$$

$$\begin{aligned}
u(t, x + h_1, y - h_2) &= u + h_1 u_x - h_2 u_y + \frac{1}{2} (h_1^2 u_{xx} - 2h_1 h_2 u_{xy} + h_2^2 u_{yy}) \\
&+ \mathcal{O} \left( (h_1^2 + h_2^2)^{\frac{3}{2}} \right)
\end{aligned} \tag{C-9}$$

$$\begin{aligned}
u(t, x - h_1, y + h_2) &= u - h_1 u_x - h_2 u_y + \frac{1}{2} (h_1^2 u_{xx} - 2h_1 h_2 u_{xy} + h_2^2 u_{yy}) \\
&+ \mathcal{O} \left( (h_1^2 + h_2^2)^{\frac{3}{2}} \right)
\end{aligned} \tag{C-10}$$

Introducimos la notación  $u_{i,j}^n \cong u(kn, hi, hj)$  para realizar la discretización del gradiente, teniendo en cuenta un entorno de  $3 \times 3$  puntos. Para simplificar suponemos que  $h_1 = h_2 = h$ . Obtenemos los siguientes esquemas discretos:

Restando las ecuaciones (C-3), (C-4) y (C-5), (C-6) obtenemos una expresión para el gradiente de la forma

$$(u_{i,j})_x = \frac{u_{i+1,j} - u_{i-1,j}}{2h} \tag{C-11}$$

$$(u_{i,j})_y = \frac{u_{i,j+1} - u_{i,j-1}}{2h} \tag{C-12}$$

en donde sólo se tienen en cuenta los vecinos anterior y posterior al punto.

Para diseñar nuestro esquema para el cálculo del gradiente, tendremos en cuenta que la norma del gradiente es invariante por rotaciones. Para esto, consideramos una

señal con un borde y la señal rotada  $45^\circ$ .

1	1	1
0	0	0
0	0	0

1	1	0
1	0	0
0	0	0

Si calculamos la norma del gradiente para esta señal utilizando el esquema anterior, vemos que para la primera esta vale  $\|\nabla u\| = 1$  mientras que para la segunda vale  $\|\nabla u\| = \sqrt{2}$

Operando con las ecuaciones (C-7), (C-8), (C-9) y (C-10), obtenemos la siguiente expresión para el gradiente

$$\begin{aligned}(u_{i,j})_x &= \frac{u_{i+1,j+1} - u_{i-1,j+1} + u_{i+1,j-1} - u_{i-1,j-1}}{4h} \\ (u_{i,j})_y &= \frac{u_{i+1,j+1} - u_{i+1,j-1} + u_{i-1,j+1} - u_{i-1,j-1}}{4h}\end{aligned}\quad (\text{C-13})$$

en la que sólo se tienen en cuenta los vecinos de las esquinas y no los inmediatos. Si calculamos la norma del gradiente para las dos señales, vemos que tampoco son iguales.

Podemos generalizar el cálculo para el gradiente ponderando los dos esquemas propuestos anteriormente de la siguiente manera:

$$\begin{aligned}(u_{i,j})_x &= (1 - \gamma) \frac{u_{i+1,j} - u_{i-1,j}}{2h} + \gamma \frac{u_{i+1,j+1} - u_{i-1,j+1} + u_{i+1,j-1} - u_{i-1,j-1}}{4h} \\ (u_{i,j})_y &= (1 - \gamma) \frac{u_{i,j+1} - u_{i,j-1}}{2h} + \gamma \frac{u_{i+1,j+1} - u_{i+1,j-1} + u_{i-1,j+1} - u_{i-1,j-1}}{4h}\end{aligned}\quad (\text{C-14})$$

donde  $\gamma$  es un parámetro a elegir. Imponiendo que la norma del gradiente de la señal y la norma de la señal rotada  $45^\circ$  sean iguales, se obtiene que  $\gamma = 2 - \sqrt{2}$ .

Esto es equivalente a utilizar máscaras de la forma

$$\frac{1}{4h} \begin{array}{|c|c|c|} \hline -(2 - \sqrt{2}) & 0 & (2 - \sqrt{2}) \\ \hline -2(\sqrt{2} - 1) & 0 & 2(\sqrt{2} - 1) \\ \hline -(2 - \sqrt{2}) & 0 & (2 - \sqrt{2}) \\ \hline \end{array}$$

para la derivada en la dirección de las  $x$  y

$$\frac{1}{4h} \begin{array}{|c|c|c|} \hline -(2 - \sqrt{2}) & -2(\sqrt{2} - 1) & -(2 - \sqrt{2}) \\ \hline 0 & 0 & 0 \\ \hline (2 - \sqrt{2}) & 2(\sqrt{2} - 1) & (2 - \sqrt{2}) \\ \hline \end{array}$$

en la otra dirección.

## C.2 Discretización de la divergencia

En los métodos que utilizamos para el cálculo del flujo óptico y de la carta de disparidad, nos encontramos siempre con un término de la forma  $div(\mathbf{D}(\|\nabla v\|)\nabla u)$ . Vamos a ver ahora como discretizar este término.

El tensor de difusión,  $\mathbf{D}$ , es una matriz de dimensión  $2 \times 2$  simétrica

$$\mathbf{D}(\nabla v) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

donde cada elemento depende de la norma del gradiente de  $u$ .

Desarrollando la divergencia tenemos:

$$\begin{aligned} div(\mathbf{D}(\nabla v)\nabla u) &= div \begin{pmatrix} au_x + bu_y \\ bu_x + cu_y \end{pmatrix} = \frac{\partial (au_x + bu_y)}{\partial x} + \frac{\partial (bu_x + cu_y)}{\partial y} \\ &= \frac{\partial (au_x)}{\partial x} + \frac{\partial (bu_y)}{\partial x} + \frac{\partial (bu_x)}{\partial y} + \frac{\partial (cu_y)}{\partial y} \end{aligned} \quad (\text{C-1})$$

Si discretizamos cada término por separado, aplicando primero diferencias hacia

adelante y luego hacia atrás, llegamos a lo siguiente:

$$\begin{aligned}
\frac{\partial (a u_x)}{\partial x} &\cong \left( a_{i,j} \frac{u_{i+1,j} - u_{i,j}}{h_1} \right)_x \\
&\cong \frac{\left( a_{i,j} \frac{u_{i+1,j} - u_{i,j}}{h_1} \right)_{i,j} - \left( a_{i,j} \frac{u_{i+1,j} - u_{i,j}}{h_1} \right)_{i-1,j}}{h_1} \\
&\cong \frac{a_{i+1,j} + a_{i,j}}{2} \frac{u_{i+1,j} - u_{i,j}}{h_1^2} + \frac{a_{i-1,j} + a_{i,j}}{2} \frac{u_{i-1,j} - u_{i,j}}{h_1^2} \tag{C-2}
\end{aligned}$$

para el primer término. El último término es similar a este pero derivando con respecto a  $y$ :

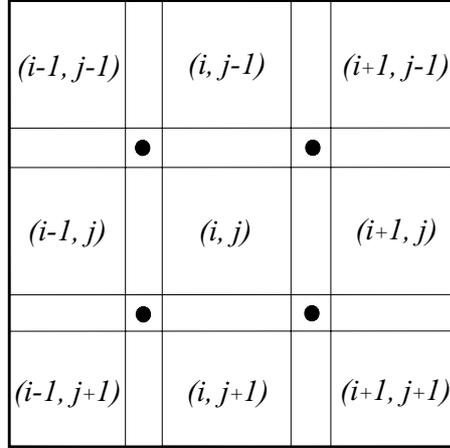
$$\begin{aligned}
\frac{\partial (c u_y)}{\partial y} &\cong \left( c_{i,j} \frac{u_{i,j+1} - u_{i,j}}{h_2} \right)_y \\
&\cong \frac{\left( c_{i,j} \frac{u_{i,j+1} - u_{i,j}}{h_2} \right)_{i,j} - \left( c_{i,j} \frac{u_{i,j+1} - u_{i,j}}{h_2} \right)_{i,j-1}}{h_2} \\
&\cong \frac{c_{i,j+1} + c_{i,j}}{2} \frac{u_{i,j+1} - u_{i,j}}{h_2^2} + \frac{c_{i,j-1} + c_{i,j}}{2} \frac{u_{i,j-1} - u_{i,j}}{h_2^2} \tag{C-3}
\end{aligned}$$

Las variables  $a$ ,  $b$ , y  $c$  son funciones que dependen de la norma del gradiente de  $v$  y están definidas en todo el dominio de la señal. En esta discretización aproximamos  $a_{i+1,j}$  por una media entre  $\frac{a_{i+1,j} + a_{i,j}}{2}$ .

Ahora vamos a discretizar los otros dos términos. Para deducir su forma de discretización, vamos a basarnos en la derivada direccional. Supongamos que tenemos

un vector de la forma  $\vec{n} = \left( \frac{h_1}{\sqrt{2h_1h_2}}, \frac{h_2}{\sqrt{2h_1h_2}} \right)^t$  en la dirección de la diagonal al pixel

(ver figura C.1 de  $(i-1, j-1)$  hacia  $(i+1, j+1)$ ). La derivada en esta dirección viene dada por

Figura C.1: Ventana de  $3 \times 3$  sobre el pixel  $(i, j)$ 

$$\begin{aligned}
\frac{\partial (b u_{\vec{n}})}{\partial \vec{n}} &= \frac{\partial (b \nabla u^t \cdot \vec{n})}{\partial \vec{n}} = \nabla (b \nabla u^t \cdot \vec{n}) \cdot \vec{n} \\
&\cong \frac{b_{i+1, j+1} + b_{i, j}}{2} \frac{u_{i+1, j+1} - u_{i, j}}{2h_1 h_2} + \frac{b_{i-1, j-1} + b_{i, j}}{2} \frac{u_{i-1, j-1} - u_{i, j}}{2h_1 h_2} \quad (C-4)
\end{aligned}$$

en donde, al igual que antes, se aproxima cada componente del tensor de difusión por una media de cada uno de los elementos de la diagonal con el asociado al del pixel en cuestión.

Si realizamos el mismo proceso para el vector  $\vec{n}' = \left( -\frac{h_1}{\sqrt{2h_1 h_2}}, \frac{h_2}{\sqrt{2h_1 h_2}} \right)^t$  en la dirección de la otra diagonal (figura C.1 de  $(i+1, j-1)$  hacia  $(i-1, j+1)$ ) llegamos a

$$\begin{aligned}
\frac{\partial (b u_{\vec{n}'})}{\partial \vec{n}'} &= \frac{\partial (b \nabla u^t \cdot \vec{n}')}{\partial \vec{n}'} = \nabla (b \nabla u^t \cdot \vec{n}') \cdot \vec{n}' \\
&\cong \frac{b_{i+1, j-1} + b_{i, j}}{2} \frac{u_{i+1, j-1} - u_{i, j}}{2h_1 h_2} + \frac{b_{i-1, j+1} + b_{i, j}}{2} \frac{u_{i-1, j+1} - u_{i, j}}{2h_1 h_2} \quad (C-5)
\end{aligned}$$

Si restamos las dos ecuaciones anteriores, (C-4)-(C-5), obtenemos los dos términos que nos quedan por definir

$$\begin{aligned}
\frac{\partial (b u_y)}{\partial x} + \frac{\partial (b u_x)}{\partial y} &= \frac{\partial (b u_{\vec{n}})}{\partial \vec{n}} - \frac{\partial (b u_{\vec{n}'})}{\partial \vec{n}'} \\
&\cong \frac{b_{i+1,j+1} + b_{i,j}}{2} \frac{u_{i+1,j+1} - u_{i,j}}{2h_1h_2} + \frac{b_{i-1,j-1} + b_{i,j}}{2} \frac{u_{i-1,j-1} - u_{i,j}}{2h_1h_2} \\
&\quad - \frac{b_{i+1,j-1} + b_{i,j}}{2} \frac{u_{i+1,j-1} - u_{i,j}}{2h_1h_2} - \frac{b_{i-1,j+1} + b_{i,j}}{2} \frac{u_{i-1,j+1} - u_{i,j}}{2h_1h_2}
\end{aligned} \tag{C-6}$$

Agrupando todos los términos anteriores, (C-2), (C-3) y (C-6), la discretización de la divergencia se aproxima por

$$\begin{aligned}
\operatorname{div}(\mathbf{D}(\nabla v) \nabla u) &\cong \left( \frac{a_{i+1,j} + a_{i,j}}{2} \frac{u_{i+1,j} - u_{i,j}}{h_1^2} + \frac{a_{i-1,j} + a_{i,j}}{2} \frac{u_{i-1,j} - u_{i,j}}{h_1^2} \right. \\
&\quad + \frac{c_{i,j+1} + c_{i,j}}{2} \frac{u_{i,j+1} - u_{i,j}}{h_2^2} + \frac{c_{i,j-1} + c_{i,j}}{2} \frac{u_{i,j-1} - u_{i,j}}{h_2^2} \\
&\quad + \frac{b_{i+1,j+1} + b_{i,j}}{2} \frac{u_{i+1,j+1} - u_{i,j}}{2h_1h_2} + \frac{b_{i-1,j-1} + b_{i,j}}{2} \frac{u_{i-1,j-1} - u_{i,j}}{2h_1h_2} \\
&\quad \left. - \frac{b_{i+1,j-1} + b_{i,j}}{2} \frac{u_{i+1,j-1} - u_{i,j}}{2h_1h_2} - \frac{b_{i-1,j+1} + b_{i,j}}{2} \frac{u_{i-1,j+1} - u_{i,j}}{2h_1h_2} \right)
\end{aligned} \tag{C-7}$$

# Bibliografía

- [AB85] Edward H. Adelson and James R. Bergen. Spatiotemporal energy models for the perception of motion. *J. of the Optical Society of America A*, 2(2):284–299, 1985.
- [ADK99] G. Aubert, R. Deriche, and P. Kornprobst. Computing optical flow via variational techniques. *SIAM Journal on Applied Mathematics*, 60(1):156–182, 1999.
- [ADPS01] L. Alvarez, R. Deriche, T. Papadopoulo, and J. Sánchez. Symmetrical dense optical flow estimation with occlusions detection. *Preprint*, 2001.
- [ADSW00] L. Alvarez, R. Deriche, J. Sánchez, and J. Weickert. Dense disparity map estimation respecting image derivatives: a pde and scale-space based approach. Report 3874, ROBOTVIS, INRIA, 2004 route des Lucioles, BP 93, 06902 Sophia–Antipolis Cedex, France, 2000. Accepted for publication in *Journal of Visual Communication and Image Representation*.
- [AELS99] L. Alvarez, J. Esclarín, M. Lefébure, and J. Sánchez. A pde model for computing the optical flow. *Proc. XVI Congreso de Ecuaciones Diferenciales y Aplicaciones (C.E.D.Y.A. XVI, Las Palmas de Gran Canaria)*, pages 1349–1356, Sept. 1999.
- [AGLM93] L. Alvarez, F. Guichard, P.L. Lions, and J.M. Morel. Axioms and fundamental equations of image processing. *Arch. for Rat. Mechanics*, 123(3):199–257, 1993.
- [AH88] N. Ayache and C. Hansen. Rectification of images for binocular and trinocular stereovision. *In Ninth International Conference on Pattern Recognition*, pages 11–16, October 1988.
- [AL87] N. Ayache and F. Lustman. Fast and reliable passive trinocular stereovision. *In Proceedings of the 1st International Conference on Computer Vision*, pages 422–427, June 1987.

- [ALM92] L. Alvarez, P.L. Lions, and J.M. Morel. Image selective smoothing and edge detection by nonlinear diffusion (ii). *SIAM J. on Numerical Analysis*, 29(3):845–866, 1992.
- [AM94] L. Alvarez and L. Mazorra. Signal and image restoration using shock filter and anisotropic diffusion. *SIAM J. on Numerical Analysis*, 31(2):590–605, 1994.
- [Ana89] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2(3):283–310, 1989.
- [AS99a] L. Alvarez and J. Sánchez. Sobre algunos problemas reales en visión por ordenador que conducen a sistemas de ecuaciones algebraicos. In *Proceedings of EACA '99*, 1:123–132, September 1999.
- [AS99b] L. Alvarez and J. Sánchez. Sobre un problema de optimización no-lineal en visión estereoscópica. aspectos computacionales. In *Proceedings of CEDYA XVI*, 1:1357–1364, September 1999.
- [AS00a] L. Alvarez and J. Sánchez. 3-d geometry reconstruction using a color image stereo pair and partial differential equations. Report 6, Instituto Universitario de Ciencias y Tecnologías Cibernéticas, University of Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas, Spain, 2000.
- [AS00b] L. Alvarez and J. Sánchez. Reconstrucción 3d de objetos a partir de un par estéreo de imágenes en color utilizando ecuaciones en derivadas parciales. *CEIG 2000. X Congreso Español de Informática Gráfica*, 3:249–262, Junio 2000.
- [AWS99] L. Alvarez, J. Weickert, and J. Sánchez. A scale-space approach to non-local optical flow calculations. *Scale-space theories in computer vision, Lecture Notes in Computer Science*, 1682:235–246, 1999.
- [AWS00] L. Alvarez, J. Weickert, and J. Sánchez. Reliable estimation of dense optical flow fields with large displacements. *International Journal of Computer Vision*, 39(1):41–56, 2000. Una versión extendida de este artículo se encuentra en el Technical Report n°2 del Instituto Universitario de Ciencias y Tecnologías Cibernéticas de la Universidad de Las Palmas de Gran Canaria, Spain.

- [BA91] M.J. Black and P. Anandan. Robust dynamic motion estimation over time. *Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR '91, Maui)*, IEEE Computer Society Press, pages 292–302, Junio 3–6 1991.
- [BA96] Michael J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, January 1996.
- [BAK91] Roberto Battiti, Edoardo Amaldi, and Christof Koch. Computing optical flow across multiple scales: an adaptive coarse-to-fine strategy. *International Journal of Computer Vision*, 6(2):133–145, 1991.
- [Bar89] S.T. Barnard. Stochastic Stereo Matching Over Scale. *International Journal of Computer Vision*, 3:17–32, 1989.
- [BB90] A.T. Brint and M. Brady. Stereo Matching of Curves. *Image and Vision Computing*, 8(1):50–56, 1990.
- [BB95] S. S. Beauchemin and J. L. Barron. The computation of optical flow. *ACM Computing Surveys*, 27(3):433–467, September 1995.
- [BC98] P. Blomgren and T.F. Chan. Color tv: Total variations methods for restoration of vector-valued images. *IEEE Transactions on Image Processing*, 7(3):304–309, 1998.
- [BD93] B. Bascle and R. Deriche. Energy based methods for 2D curve tracking, reconstruction, and refinement of 3D curves and applications. *In Proceedings Geometric Methods in Computer Vision II*, 2031:282–293, Julio 12-13 1993.
- [BDH<sup>+</sup>96] M.-O. Berger, R. Deriche, I. Herlin, J. Jaffré, and J.-M. Morel (Eds.). Icaos '96: Images, wavelets and pdes. *Lecture Notes in Control and Information Sciences*, 219, 1996.
- [Ber87] F. Bergholm. Edge focusing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9:726–741, 1987.
- [BFB94] J. L. Barron, David J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

- [BFBG93] L. Blanc-Féraud, M. Barlaud, and T. Gaidon. Motion estimation involving discontinuities in a multiresolution scheme. *Optical Engineering*, 32(7):1475–1482, 1993.
- [BPT88a] M. Bertero, T.A. Poggio, and V. Torre. Ill-posed problems in early vision. *Proc. IEEE*, 76:869–889, 1988.
- [BPT88b] Mario Bertero, Tomaso Poggio, and Vincent Torre. Ill-posed problems in early vision. *In Proc. of the IEEE*, 76:869–890, 1988.
- [Bre73] H. Brezis. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de hilbert, 1973.
- [BZ87] A. Blake and A. Zisserman. *Visual reconstruction*. MIT Press, Cambridge, 1987.
- [CBFAB94] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. *Two deterministic half-quadratic regularization algorithms for computed imaging*, volume 2 of *Proc. IEEE Int. Conf. Image Processing (ICIP-94)*, pages 168–172. IEEE Computer Society Press, Los Alamitos, Nov. 13–16 1994.
- [CCLM92] F. Catté, T. Coll, P.L. Lions, and J.M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, February 1992.
- [Chr99] Gary E. Christensen. Consistent linear-elastic transformations for image matching. In *IPMI*, pages 224–237, 1999.
- [CMSE98] V. Caselles, J.M. Morel, G. Sapiro, and A. Tannenbaum (Eds.). special issue on partial differential equations and geometry-driven diffusion in image processing and analysis. *IEEE Trans. Image Proc*, 7(3), March 1998.
- [Coh93] I. Cohen. Nonlinear variational method for optical flow computation. *Proc. Eighth Scandinavian Conf. on Image Analysis (SCIA '93)*, 1:523–530, May 25–28 1993.
- [CR00] Pascal Cachier and David Rey. Symmetrization of the non-rigid registration problem using inversion-invariant energies: Application to multiple sclerosis. *In Proc. of MICCAI 2000*, pages 472–481, 2000.

- [CV90] Marco Campani and Alessandro Verri. Computing optical flow from an overconstrained system of linear algebraic equations. *In Proc. 3rd International Conference on Computer Vision*, pages 22–26, 1990.
- [DCM87] E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:700–703, 1987.
- [DF94] Frédéric Devernay and Olivier D. Faugeras. Computing differential properties of 3-D shapes from stereoscopic images without 3-D models. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 208–213, June 21–23 1994.
- [DF95] R. Deriche and O. Faugeras. Les edp en traitement des images et vision par ordinateur. Technical Report 2697, INRIA Sophia–Antipolis, 1995.
- [DKA95] R. Deriche, P. Kornprobst, and G. Aubert. Optical-flow estimation while preserving its discontinuities: A variational approach. *Proc. Second Asian Conf. Computer Vision (ACCV '95)*, 2:290–295, December 5–8 1995.
- [DR01] D.Tschumperlé and R.Deriche. Constrained and unconstrained pdes for vector image restoration. *In Proc of 12th Scandinavian Conference on Image Analysis*, June 2001.
- [Enk88] W. Enkelmann. Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences. *Computer Vision, Graphics and Image Processing*, 43(2):150–177, 1988.
- [Fau92] Olivier D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig. *In Proc. of the 2nd European Conference on Computer Vision*, 588:563–578, May 1992.
- [Fau93] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, 1993.
- [FHM<sup>+</sup>93] Olivier Faugeras, Bernard Hotz, Hervé Mathieu, Thierry Viéville, Zhengyou Zhang, Pascal Fua, Eric Théron, Laurent Moll, Gérard Berry, Jean Vuillemin, Patrice Bertin, , and Catherine Proy. Real time correlation based stereo: algorithm implementations and applications. *International Journal of Computer Vision*, 1993.

- [FJ90] David J. Fleet and Allan D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1):77–104, 1990.
- [FJ93] David J. Fleet and Allan D. Jepson. Stability of phase information. *IEEE Trans. Patt. Anal. Mach. Intel.*, 15:1253–1268, 1993.
- [Fle92] David J. Fleet. *Measurement of Image Velocity*. Engineering and Computer Science. Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061 USA, 1992.
- [Flo97] L. Florack. *Image structure*. Kluwer, Dordrecht, 1997.
- [FNN98] L.M.J. Florack, W.J. Niessen, and M. Nielsen. The intrinsic structure of the optic flow incorporating measurement duality. *Int. J. Comput. Vision*, 27:263–286, 1998.
- [Fua93] Pascal Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. *Machine Vision and Applications*, 6(1):35–49, 1993. Available as INRIA research report 1369.
- [Gib50] J. J. Gibson. *The Perception of the Visual World*. Houghton Mifflin, Boston, Massachusetts, 1950.
- [Gla84] F. Glazer. Multilevel relaxation in low-level computer vision. *Multiresolution Image Processing and Analysis*, pages 312–320, 1984.
- [GMN98] B. Galvin, B. McCane, and K. Novins. Recovering motion fields: An analysis of eight optical flow algorithms. *Proc. 1998 British Machine Vision Conference (BMVC '98, Southampton)*, September 14–17 1998.
- [GR85] G. Medioni and R. Nevatia. Segment based stereo matching. *Computer Vision, Graphics and Image Processing*, 31(1):2–18, 1985.
- [GR92] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14:367–383, 1992.
- [GR96] F. Guichard and L. Rudin. Accurate estimation of discontinuous optical flow by minimizing divergence related functionals. *Proc. Third Int. Conf. on Image Processing (ICIP '96, Lausanne)*, 1:497–500, Sept. 1996.

- [Gri85] W. E. L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(1):17–34, 1985.
- [GT] E. González and A. Trujillo. Xmegawave. a window oriented image processing software.
- [Har92] R. Hartley. Estimation of relative camera positions for uncalibrated cameras. *In Proc. European Conference on Computer Vision*, pages 579–587, 1992.
- [Har94a] R. Hartley. An algorithm for self calibration from several views. *In Proc. Computer Vision and Pattern Recognition*, pages 908–912, 1994.
- [Har94b] R. Hartley. Projective reconstructions and invariants from multiple images. *PAMI*, 16(10):1036–1040, 1994.
- [Har98] R. Hartley. Theory and practice of projective rectification. *International Journal of Computer Vision. To appear*, 1998.
- [HB93] Fabrice Heitz and Patrick Bouthemy. Multimodal motion estimation of discontinuous optical flow using markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(12):1217–1232, December 1993.
- [HCF01] G. Hermosillo, C. Ched’Hotel, and O. Faugeras. A variational approach to multi-modal image matching. Technical Report 4117, INRIA Sophia-Antipolis, 2004, Route des Lucioles, B.P.93, 06902 Sophia Antipolis Cedex (France), February 2001.
- [Hee88] David J. Heeger. Optical flow using spatio-temporal filters. *International Journal of Computer Vision*, 1(4):279–302, 1988.
- [Hin99] W. Hinterberger. *Generierung eines Films zwischen zwei Bildern mit Hilfe des optischen Flusses*. M.sc. thesis, Industrial Mathematics Institute, University of Linz, Austria, 1999.
- [HPB94] F. Heitz, P. Pérez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. *CVGIP : Image Understanding*, 59(1):125–134, January 1994.
- [HS81] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203, August 1981.

- [Iij62] T. Iijima. Basic theory on normalization of pattern (in case of typical one-dimensional pattern). *Bulletin of the Electrotechnical Laboratory*, 26:368–388, 1962. (in Japanese).
- [Iij73] T. Iijima. *Pattern recognition*. Corona-sha, 1973. (in Japanese).
- [Iij89] T. Iijima. *Theory of pattern recognition*, volume 6 of *Series of Basic Information Technology*. Morishita Publishing, 1989. (in Japanese).
- [JH98] B. Jähne and H. Haussecker. Performance characteristics of low-level motion estimators in spatiotemporal images. *Proc. Workshop on Evaluation and Validation of Computer Vision Algorithms*, March 16–20 1998.
- [JJ94] M.R.M. Jenkin and A.D. Jepson. Recovering local surface structure through local phase difference methods. *CVGIP: Image Understanding*, 59:72–93, 1994.
- [JM92] D.G. Jones and J. Malik. *A computational framework for determining stereo correspondence from a set of linear spatial filters*, volume 588 of *Lecture Notes in Computer Science*, pages 395–410. Springer, Berlin, computer vision – eccv ’92 edition, 1992.
- [KH75] C. D. Kuglin and D. C. Hines. The phase correlation image alignment method. *In Conference on Cybernetics and Society*, pages 163–165, Septiembre 1975.
- [Koe84] J.J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.
- [Kor98] P. Kornprobst. *Contribution a la restauration d’images et a l’analyse de séquences: approches variationnelles et solutions de viscosité*. PhD thesis, Universidad de Niza, Noviembre 1998.
- [KS91] D.S. Kalivas and A.A. Sawchuk. A region matching motion estimation algorithm. *CVGIP*, 54(2):275–288, 1991.
- [KTB96] A. Kumar, A. R. Tannenbaum, and G. J. Balas. Optical flow: a curve evolution approach. *IEEE Transactions on Image Processing*, 5(4):598–610, 1996.
- [KZ86] R. Kories and G. Zimmerman. A versatile method for the estimation of displacement vector fields from image sequences. *In IEEE Proc.*

- of Workshop on Motion: Representation and Analysis*, pages 101–106, 1986.
- [LBP88] J.J. Little, H.H. Butthoff, and T. Poggio. Parallel optical flow using local voting. *In Proc. of ICCV*, pages 454–459, 1988.
- [LDFP93] Q.-T. Luong, R. Dericé, O. D. Faugeras, and T. Papadopoulos. On determining the fundamental matrix: analysis of different methods and experimental results. Technical Report 1894, INRIA, 1993.
- [LF96] Q.-T. Luong and O. D. Faugeras. Fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17(1):43–75, 1996.
- [LH81] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
- [Lin94] T. Lindeberg. *Scale Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994.
- [Lin98] T. Lindeberg. A scale selection principle for estimating image deformations. *Image and Vision Computing*, 16(14):961–977, 1998.
- [LK81] B.D. Lucas and T. Kanade. An iterative image-registration technique with an application to stereo vision. *In DARPA Proc. of Image Understanding Workshop*, pages 121–130, 1981.
- [Luc84] B.D. Lucas. *Generalized image matching by the method of differences*. PhD thesis, Carnegie Mellon University, 1984.
- [LZ98] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. *In Proceedings of the Conference on Computer Vision and Pattern Recognition*, 1998.
- [MB96] A. Mitiche and Patrick Bouthemy. Computation and analysis of image motion: A synopsis of current problems and methods. *International Journal of Computer Vision*, 19(1):29–55, 1996.
- [MGDP93] T. Moons, L. Van Gool, M. Van Diest, and E. Pauwels. Affine reconstruction from perspective image pairs. *In Proceeding of the DARPA-ESPRIT workshop on Applications of Invariants in Computer Vision*, pages 249–266, October 1993.

- [MM88] James H. McIntosh and Kathleen M. Mutch. Matching straight lines. *Computer Vision, Graphics and Image Processing*, 43(3):386–408, September 1988.
- [MM94] K.W. Morton and L.M. Mayers. *Numerical Solution of partial differential equations*. Cambridge University Press, Cambridge, 1994.
- [MP98] E. Mémin and P. Pérez. Dense estimation and object-based segmentation of the optical flow with robust techniques. *IEEE Trans. Image Proc*, 7:703–719, 1998.
- [MWA87] A. Mitiche, Y. F. Wang, and J. K. Aggarwal. Experiments in computing optical flow with the gradient-based, multiconstraint method. *Pattern Recognition*, 20(2):173–179, 1987.
- [MY99] Miller and Younes. Group actions, homeomorphisms, and matching: A general framework, 1999.
- [Nag83] Hans-Hellmut Nagel. Constraints for the estimation of displacement vector fields from image sequences. *Proc. Eighth Int. Joint Conf. on Artificial Intelligence (IJCAI '83, Karlsruhe)*, pages 945–951, August 8–12 1983.
- [Nag87] Hans-Hellmut Nagel. On the estimation of optical flow(): Relations between different approaches and some new results. *Artificial Intelligence*, 33(3):299–324, November 1987.
- [Nag90] Hans-Hellmut Nagel. Extending the 'oriented smoothness constraint' into the temporal domain and the estimation of derivatives of optical flow. *Computer vision – ECCV '90, Lecture Notes in Computer Science*, 427, Springer, Berlin:139–148, 1990.
- [Nas92] Nasser M. Nasrabadi. A stereo vision technique using curve-segments and relaxation matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(5):566–572, May 1992.
- [ND95] M. Nielsen and R. Deriche. Binocular dense depth reconstruction using isotropy constraint. *Theory and applications of image processing II – selected articles from the 9th Scandinavian Conference on Image Analysis*, pages 27–140, 1995.
- [NE86] Hans-Hellmut Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from

- images sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(5):565–593, 1986.
- [Nes93] Paolo Nesi. Variational approach to optical flow estimation managing discontinuities. *Image and Vision Computing*, 11(7):419–439, 1993.
- [Nis84] Keith Nishihara. Practical real-time imaging stereo matcher. *Optical Engineering*, 23(5):536–545, Agosto 1984.
- [NJOE99] M. Nielsen, P. Johansen, O.F. Olsen, and J. Weickert (Eds.). Scale-space theories in computer vision. *Lecture Notes in Computer Science*, Springer, Berlin, 1682, 1999.
- [NMN<sup>+</sup>99] M. Nielsen, R. Maas, W.J. Niessen, L.M.J. Florack, and B.M. ter Haar Romeny. Binocular stereo from grey-scale images. *J. Math. Imag. Vision*, 10:103–122, 1999.
- [OK85] Y. Ohta and Takeo Kanade. Stereo by intra- and inter-scanline search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2):139–154, 1985.
- [ON95] Michael Otte and Hans-Hellmut Nagel. Estimation of optical flow based on higher-order spatiotemporal derivatives in interlaced and non-interlaced image sequences. *Artificial Intelligence*, 78(1/2):5–43, November 1995.
- [OS88] S. Osher and S. Sethian. Fronts propagating with curvature dependent speed: algorithms based on the hamilton-jacobi formalism. *J. Comp. Physics*, 79:12–49, 1988.
- [PD96] D. Papadimitriou and T. Dennis. Epipolar line estimation and rectification for stereo images pairs. *IEEE Transactions on Image Processing*, 3(4):672–676, April 1996.
- [PGPO94] M. Proesmans, L. Van Gool, E. Pauwels, and A. Oosterlinck. Determination of optical flow and its discontinuities using non-linear diffusion. *J.-O. Eklundh (Ed.), Computer vision – ECCV ’94, Lecture Notes in Computer Science*, 801, Springer:295–304, 1994.
- [PKG99] M. Pollefeys, R. Koch, and L. Van Gool. *A simple and efficient rectification method for general motion*, pages 496–501. In Proceedings of the 7th International Conference on Computer Vision. IEEE Computer Society Press, Corfu, Greece, Sept. 21–24 1999.

- [PM90] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:429–439, 1990.
- [PMF85] S. B. Pollard, John E. W. Mayhew, and John P. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient constraint. *Perception*, 14:449–470, 1985.
- [PPG94] M. Proesmans, E. Pauwels, and L. Van Gool. Coupled geometry-driven diffusion equations for low-level vision. *Geometry-driven diffusion in computer vision*, pages 191–228, 1994.
- [RD96] L. Robert and R. Deriche. Dense depth map reconstruction: A minimization and regularization approach which preserves discontinuities. *Computer vision – ECCV '96, Lecture Notes in Computer Science*, 1064:439–451, 1996.
- [RF91] Luc Robert and Olivier Faugeras. *Curve-based stereo: Figural continuity and curvature*. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society Press, Lahaina, Hawaii, June 3–6 1991.
- [RMC97] S. Roy, J. Meunier, and I. Cox. *Cylindrical rectification to minimize epipolar distortion*, pages 393–399. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society Press, San Juan, Puerto Rico, Junio 17–19 1997.
- [ROF92] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [S00] J. Sánchez. Variational multi-scale-space approach to dense disparity map estimation. *RecPad2000*, In *Proc of 11<sup>th</sup> Portuguese Conference on Pattern Recognition*:435–438, May 2000.
- [Sch91a] Christoph Schnörr. Determining optical flow for irregular domains by minimizing quadratic functionals of a certain class. *International Journal of Computer Vision*, 6(1):25–38, April 1991.
- [Sch91b] Christoph Schnörr. On functionals with greyvaluecontrolled smoothness terms for determining optical flow. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15:1074–1079, April 1991.

- [Sch94] Christoph Schnörr. Segmentation of visual motion by minimizing convex non-quadratic functionals. *Proc. 12th Int. Conf. Pattern Recognition (ICPR 12, Jerusalem)*, IEEE Computer Society Press, Los Alamitos, A:661–663, Oct. 9–13 1994.
- [Sha93] J. Shah. A nonlinear diffusion model for discontinuous disparity and half-occlusions in stereo. *Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '93, New York)*, IEEE Computer Society Press, Los Alamitos, pages 34–40, June 15–17 1993.
- [SNFE97] J. Sporring, M. Nielsen, L. Florack, and P. Johansen (Eds.). *Gaussian scale-space theory*. Kluwer, Dordrecht, 1997.
- [Sny91] M. A. Snyder. On the mathematical foundations of smoothness constraints for the determination of optical flow and for surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(11):1105–1114, November 1991.
- [SR96] G. Sapiro and D.L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing*, 5(11):1582–1585, 1996.
- [SWP<sup>+</sup>83] M.A. Sutton, W.J. Walters, W.H. Peters, W.F. Ranson, and S.R. McNeil. Determination of displacement using an improved digital correlation method. *Image and Vision Computing*, 1(3):133–139, 1983.
- [TA77] A.N. Tikhonov and V.Y. Arsenin. Solutions of ill-posed problems. *John Wiley & Sons*, 1977.
- [tHRFKE97] B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever (Eds.). *Scale-space theory in computer vision*, volume 1252 of *Lecture Notes in Computer Science*. Springer, Berlin, 1997.
- [TK92] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. *International Journal of Computer Vision*, 9(2):137–154, 1992.
- [TK01] A. Trujillo and K. Krissian. Amilab: A 2d/3d image processing software. *In Proc. IEEE Eight International Conference on Computer Vision*, 2:744–744, July 2001.
- [TP84] O. Tretiak and L. Pastor. Velocity estimation from image sequences with second order differential operators. *Proc. 7th International Conference on Pattern Recognition*, pages 20–22, 1984.

- [UGVT88] S. Uras, F. Girosi, Alessandro Verri, and V. Torre. A computational approach to motion perception. *Biological Cybernetics*, 60:79–97, 1988.
- [Ull79] S. Ullman. *The interpretation of visual motion*. MIT Press, Cambridge, London, 1979.
- [WDT83] KwangYun Wohn, Larry S. Davis, and P. Thrift. Motion estimation based on multiple local constraints and nonlinear smoothing. *Pattern Recognition*, 16(6):563–570, 1983.
- [Wei94] J. Weickert. Scale-space properties of nonlinear diffusion filtering with a diffusion tensor. Report 110, Laboratory of Technomathematics, University of Kaiserslautern, P.O.Box 3049, 67653 Kaiserslautern, Germany, October 1994.
- [Wei95] J. Weickert. *Multiscale Texture Enhancement*, volume 970 of *Computer analysis of images and patterns; Lecture Notes*, pages 230–237. Springer, Berlin, 1995.
- [Wei96] J. Weickert. Theoretical foundations of anisotropic diffusion in image processing. *Computing*, 11:221–236, 1996.
- [Wei97] J. Weickert. *A Review of Nonlinear Diffusion Filtering*, volume 1252 of *Scale-Space Theory in Computer Vision; Lecture Notes in Computer Science*, pages 3–28. Springer, Berlin, 1997.
- [Wei98a] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart, 1998.
- [Wei98b] J. Weickert. On discontinuity-preserving optic flow. *Proc. Computer Vision and Mobile Robotics Workshop (CVMR '98)*, pages 115–122, Sept. 17–18 1998. S. Orphanoudakis, P. Trahanias, J. Crowley, N. Katevas (Eds.).
- [Wei99] J. Weickert. Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, 17:201–212, 1999.
- [WII99] J. Weickert, S. Ishikawa, and A. Imiya. Linear scale-space has first been proposed in japan. *J. Math. Imag. Vision*, 10:237–252, 1999.
- [WJ85] A.B. Watson and A.J. Ahumada Jr. Model of human visual motion sensing. *J. Opt. Soc. Am.*, 2(2):322–341, 1985.

- [WS99] J. Weickert and C. Schnörr. *Räumlich-zeitliche Berechnung des optischen Flusses mit nichtlinearen flußabhängigen Glattheitstermen*, pages 317–324. Springer, Berlin, w. förstner, j.m. buhmann, a. faber, p. faber (eds.) edition, 1999.
- [WWK92] J. Wiklund, C.J. Westelius, and H. Knutsson. Hierarchical phase based disparity estimation. *Proc. Second Int. Singapore Conf. on Image Proc.*, pages 128–131, Sept. 7–11 1992.
- [Yok90] N. Yokoya. Stereo Surface Reconstruction by Multiscale-Multistage Regularization. Technical Report TR-90-45, ETL Electrotechnical Laboratory, Tsukuba, Japan, November 1990.
- [ZF94] C. Zeller and O. Faugeras. *Applications of non-metric vision to some visually guided tasks*, pages 132–136. In Proceedings of the International Conference on Pattern Recognition. Computer Society Press, Jerusalem, Israel, Octubre 1994. Longer version available as INRIA Research Report RR2308.
- [Zha97] Zhengyou Zhang. Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, 27(2):161, 1997. Also as Technical Report 2927 at the Inria Sophia–Antipolis.