

Algunos problemas de interpolación en procesamiento de imágenes

V. Caselles *

December 31, 2004

Abstract

Problemas de interpolación aparecen en diversos contextos del procesamiento de imágenes: cartografía digital, compresión de imágenes basados en regiones, restauración y manipulación de imágenes, restauración de *agujeros* en superficies digitalizadas, fotografía digital, o imágenes satélite. Presentamos algunos de ellos y discutimos una posible solución basada en la morfología matemática. Adaptando algunas de sus ideas básicas sobre el contenido geométrico de las imágenes a cada caso concreto pueden construirse modelos variacionales o basados en ecuaciones en derivadas parciales cuya solución numérica proporciona la interpolación buscada.

1 Introducción

Problemas de interpolación aparecen en diversos contextos en procesamiento de imágenes. Vamos a presentar algunos de ellos en esta introducción sin pretender ser exhaustivo. La selección de estos problemas está motivada tanto por su interés práctico como por razones de exposición: vamos a situarlos en un contexto común de interpolación de estructuras topográficas, idea que desarrollaremos en detalle en el resto de este trabajo. Quedarán de lado muchos otros problemas de interpolación (entre ellos los relacionados con la conversión analógico-digital) que se abordan mejor desde el punto de vista de los splines o del análisis de Fourier. El lector interesado en una buena introducción a los mismos puede consultar [48].

En el contexto de la cartografía digital, se plantea el problema de la reconstrucción de un modelo digital de elevación de terreno a partir de una familia de sus curvas de nivel (ver Figura 1). Las curvas de nivel seleccionadas son generalmente representativas de la elevación real y es poco probable que oscilaciones importantes de los datos no hayan sido reflejadas por las mismas. Ello implica que los métodos de interpolación razonables no deben crear oscilaciones artificiales entre las curvas de nivel dadas. Uno de los métodos más simples de interpolación entre curvas de nivel para las que conocemos su altura es la solución de la ecuación de Laplace

$$\Delta u := u_{xx} + u_{yy} = 0 \tag{1}$$

en el dominio D comprendido entre dos curvas de nivel siendo la condición de frontera la dada por la altura a la que se encuentran dichas curvas. Los resultados que se obtienen con este modelo no se consideran aceptables ya que produce algunos artefactos descritos como creación de terrazas. Para corregir este defecto ([24]) los investigadores han propuesto las estrategias siguiente:

- 1) Calcular las líneas de gradiente de la solución u de (1) o de cualquier otra interpolación suave entre curvas de nivel.
- 2) Interpolar a lo largo de las líneas de gradiente a partir de los valores de altura donde la línea de gradiente cruza las curvas de nivel.

*Dept. Tecnología, Universitat Pompeu Fabra, Passeig de Circumvalació, 8, 08003 Barcelona, Spain, vicent.caselles@upf.edu

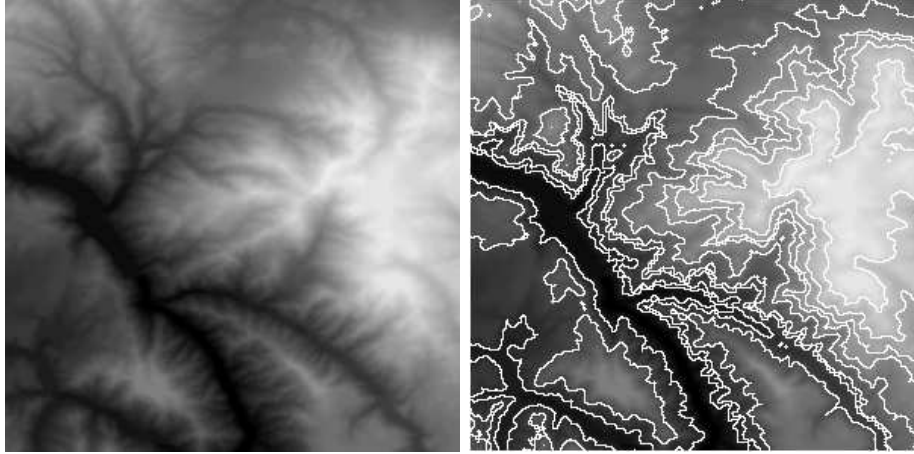


Figure 1: a. Visualización como imagen de un modelo digital de elevación de terreno. b) Algunas de sus curvas de nivel.

Si en el paso 2) adoptamos una interpolación lineal y llamamos v a la solución obtenida, podemos resumir el algoritmo anterior con la ecuación

$$\langle D^2v(\nabla u), \nabla u \rangle = 0 \quad \text{en } D \quad (2)$$

con condiciones de frontera $v|_{\partial D}$ dadas por la altura de las curvas de nivel que constituyen ∂D , donde D^2v representa la matriz Hessiana de v y ∇u el gradiente de u .

Obviamente, en estos métodos de interpolación la información no se difunde a través de las curvas de nivel dadas y puede que éstas sean demasiado visibles en la superficie reconstruida ya que las pendientes de la superficie pueden no coincidir a un lado y otro de la curva de nivel. El modelo más simple que permite incorporar información sobre las pendientes de la superficie en las curvas de nivel está basado en la ecuación

$$\Delta\Delta u = 0 \quad \text{en } D \quad (3)$$

donde el operador $\Delta\Delta u = u_{xxxx} + 2u_{xxyy} + u_{yyyy}$ recibe el nombre de bi-laplaciano. Podemos resolver esta ecuación si especificamos los valores de u y de $\nabla u \cdot \nu^D$ en la frontera de D (siendo ν la normal unitaria exterior a ∂D). El defecto de este modelo es que puede provocar la aparición de oscilaciones espúreas y ello impide su utilización. Esto no es nada sorprendente ya que la ecuación (3) no satisface el principio del máximo. Como veremos en la Sección 3, si pretendemos interpolar datos en un dominio de forma que se satisfaga el principio del máximo, nos veremos obligados a adoptar modelos basados en ecuaciones elípticas de segundo orden. El principio del máximo es el requisito que garantiza que nuestro interpolador va a respetar las estructuras geométricas presentes en los datos ya que no se van a crear nuevos extremos; en particular, no van a aparecer oscilaciones espúreas.

Los métodos de compresión de imágenes basados en regiones – llamados métodos de segunda generación–, que intentan alcanzar altas tasas de compresión, proponen una representación de la imagen en contornos y texturas. Los contornos conforman una partición de la imagen y la información debe ser reconstruida por el receptor en cada región a partir de la información dada por contornos y texturas [45]. Los contornos pueden obtenerse de diferentes maneras, por ejemplo, pueden estar formados por una selección de curvas de nivel de la imagen atendiendo a criterios de coste de codificación. El resto de los píxeles de la imagen ha de recuperarse a partir de la información disponible por el receptor. Para ello, J.R. Casas ([45]) propuso un algoritmo de interpolación basado en la propagación de los valores de la imagen en los contornos. Los frentes de propagación de los valores confluyen en un frente intermedio al que asignamos el valor promedio. Esto determina un nuevo sistema de frentes que de nuevo vuelven a confluir proporcionando nuevos valores intermedios. Iterando este esquema puede llegarse a una reconstrucción de la imagen en la region considerada D . Una variante de este algoritmo puede ser interpretada en términos de la solución de la ecuación en derivadas parciales [18]



Figure 2: a. Imagen dada. b) Curvas de nivel correspondientes a los múltiplos de 20.

$$\langle D^2u(\nabla u), \nabla u \rangle = 0 \quad \text{en } D \quad (4)$$

con condiciones de frontera dadas por los valores de la imagen en ∂D .

Los dos casos planteados hasta ahora conducen al mismo problema de interpolación. Las soluciones propuestas, bien de forma directa o indirecta, pueden interpretarse en términos de la solución de una ecuación en derivadas parciales de tipo elíptico.

La necesidad de restaurar o retocar imágenes abarca desde la clásica restauración de pinturas hasta la actual de fotografías y películas. Los objetivos pueden ir desde evitar el deterioro (caso de rayas o motas de polvo en películas) hasta añadir o suprimir detalles, objetos o personas (desde la supresión de un texto en una imagen hasta la infame desaparición de enemigos políticos [32]). La Figura 3 nos muestra un ejemplo de aplicación de las técnicas de 'inpainting' para suprimir un texto impreso sobre una fotografía. Por comodidad, nos referiremos a todo este tipo de aplicaciones como 'inpainting', que es la palabra inglesa para referirse a la restauración de pinturas. Con el desarrollo de la tecnología digital, se usan cada vez más las técnicas digitales en 'inpainting' de imágenes, aunque éstas se aplican aún de forma manual, siendo necesario el desarrollo de métodos robustos que permitan su automatización.



Figure 3: a. Fotografía de New-Orleans con texto. B) Imagen restaurada.

Podemos clasificar los trabajos sobre inpainting de imágenes digitales en tres grandes grupos. El primero de ellos contendría los relativos a la restauración de películas [33, 34, 35], el segundo abarcaría los trabajos relativos a la síntesis de texturas, y el tercero los relativos a inpainting de estructuras geométricas.

Kokaram et al. [33, 34, 35] utilizan una estimación de movimiento y modelos autoregresivos para corregir rayas u otros defectos en películas a partir de la información de los cuadros adyacentes (básicamente, copiándola de forma conveniente). Esta técnica puede aplicarse tanto a imágenes estáticas como a películas en las que las zonas a corregir se extienden a varios cuadros.

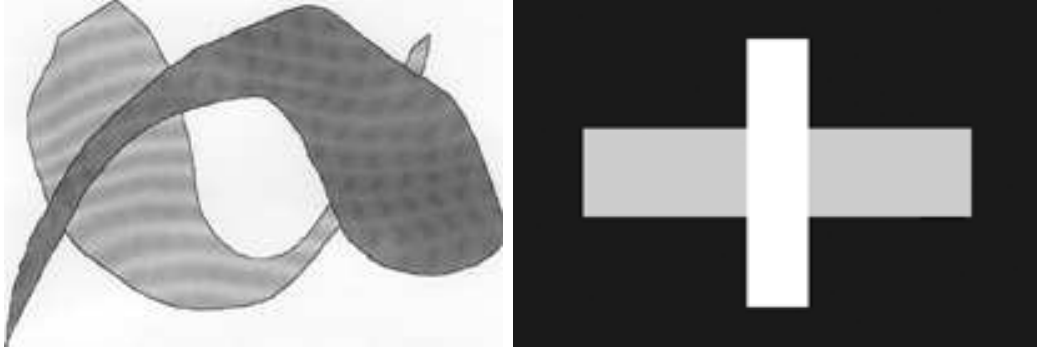


Figure 4: a) Nuestra interpretación de esta imagen sigue el principio de *buena continuación*. b) Ejemplo esquemático de junción en T .

Entre los trabajos dedicados a la síntesis de texturas uno de los más notables es el de Efros y Leung basado en una interpretación de la imagen como un campo de Markov [22]. En este trabajo, a partir de una textura de muestra, los autores proponen sintetizar una realización de la misma de forma incremental considerando a cada paso entornos similares en la muestra y copiándolos, llegando a resultados espectaculares. Entre las mejoras propuestas a las técnicas de Efros-Leung hay que destacar el trabajo de Criminisi-Perez-Toyama [21] en el cual se orienta el proceso de síntesis priorizando los píxeles de los contornos de la imagen (puntos donde el gradiente es alto, que intuitivamente corresponden a contornos de objetos). Otros autores como Hirani y Totsuka [26] combinan información espacial y del dominio de frecuencias para rellenar una región con una textura seleccionada por el usuario.

Mencionemos finalmente los métodos geométricos para interpolar información en regiones de la imagen donde, por una razón u otra, no disponemos de ella. Llamaremos *agujero* de la imagen a cada una de estas zonas, véanse las letras en la Figura 3.a. El trabajo de D. Mumford, M. Nitzberg and T. Shiota [42] es una de las primeras contribuciones a la recuperación de la geometría de una imagen 2D. Dicho trabajo no estaba directamente dedicado a la restauración de *agujeros*, sino a la segmentación de imágenes en objetos ordenados de acuerdo con su profundidad en la escena. El algoritmo de segmentación debería ser capaz de calcular los objetos, ordenarlos según su profundidad, y determinar las fronteras de los mismos ocultas por otros objetos. Para ello, se apoyaron en un principio básico de la teoría de la Gestalt: nuestro sistema visual es capaz de completar las fronteras parcialmente ocultas y su completación tiende a respetar el *principio de buena continuación* [31] (ver Figura 4.a). Cuando un objeto oculta otro, las fronteras forman una configuración particular llamada junción en T en la que la parte visible de la frontera del objeto ocultado llega a la frontera del objeto que le oculta formando una T (véase la Figura 4.b). En estas condiciones, nuestro sistema visual continúa de forma suave la frontera ocultada entre dos junciones en T . Nitzberg-Mumford-Shiota [42] propusieron un funcional de energía para segmentar una escena que toma en cuenta la profundidad de los objetos y la energía de las fronteras ocultas entre T -junciones. Esta energía intenta calcular curvas lo más cortas posibles que respeten el principio de buena continuación. Obviamente, es necesario dar una interpretación matemática de estos dos principios para que sean operativos. Dados dos T -junciones p y q y las tangentes τ_p y τ_q a los contornos que en ellas terminan, los autores propusieron la elástica como curva de buena continuación entre p y q , es decir, la curva que minimiza la energía

$$\int_C (\alpha + \beta \kappa^2) ds, \quad \alpha, \beta > 0, \quad (5)$$

donde el mínimo se calcula entre todas las curvas C que unen p y q y tienen por tangentes τ_p y τ_q en esos puntos, siendo κ la curvatura de C y ds su longitud de arco. La elástica ha sido usada reiteradamente en procesamiento de imágenes ([47, 27, 51, 46]) en relación con el cálculo de contornos subjetivos (véase la Figura 10.b) y el artículo de Mumford [41] contiene una discusión muy interesante sobre la misma.

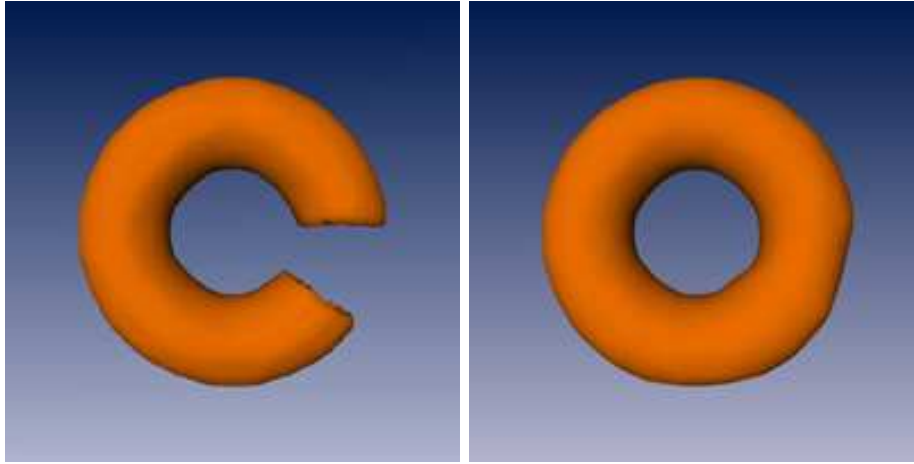


Figure 5: Un toro roto y su reconstrucción.

Inspirados por Mumford-Nitzberg-Shiota, Masnou y Morel [37, 39, 38] propusieron un modelo variacional para interpolar las curvas de nivel de una imagen (uniendo los pares de junctiones en T por rectas) y llamaron a este proceso des-ocultación. La elástica ha inspirado también los modelos variacionales introducidos en [6, 8, 7] que interpolan conjuntamente las curvas de nivel y su campo de direcciones. Los desarrollos algorítmicos de estos trabajos permiten la interpolación de imágenes 3D y la reconstrucción de agujeros en superficies. Todo ello será presentado con más detalle en la Sección 4.3. El papel jugado por la elástica en el caso bi-dimensional lo juega ahora el funcional de Willmore (la integral del cuadrado de la curvatura media).

Mientras que las imágenes digitales en niveles de gris pueden ser representadas como funciones $u : \{1, \dots, M\} \times \{1, \dots, N\} \rightarrow \mathbb{R}$ (tomando de hecho un conjunto discreto de valores), las imágenes en color lo son como funciones $\vec{u} : \{1, \dots, M\} \times \{1, \dots, N\} \rightarrow \mathbb{R}^3$, es decir, como tres imágenes representando la energía de la luz en tres bandas espectrales que podemos llamar rojo, verde y azul. La imagen de intensidad correspondiente, a la que llamaremos a partir de ahora imagen pancromática, se calcula sumando las contribuciones de los tres canales de \vec{u} con pesos convenientes. En el contexto de las imágenes satélite y para economizar el espacio de almacenamiento, se adquiere la imagen pancromática a la máxima resolución de muestreo espacial, y los tres canales de color (más eventualmente un canal infrarrojo) a una resolución s veces más pequeña donde $s = 2$ ó 4 . En la Figura 6 mostramos una imagen pancromática y los tres canales de color (R, G, B) a una resolución 4 veces más pequeña. En este caso un píxel de la imagen pancromática representa 0.7 metros mientras que un píxel de la imagen en color representa 2.8 metros. El objetivo consiste en recuperar los canales de color a la resolución de la imagen pancromática [9].

Veremos cómo los problemas planteados en esta introducción pueden tratarse desde una perspectiva común: como problemas de interpolación del mapa topográfico de una imagen. El mapa topográfico es la estructura básica que permite describir la información geométrica de una imagen y a él está dedicada la Sección 2 que nos permitirá abordar el resto de problemas en las secciones que siguen.

2 La geometría de las imágenes en niveles de gris

2.1 Imágenes digitales

Simplificando las cosas, podemos decir que una imagen digital es como una fotografía, diferenciándose tan sólo en su codificación. Consideremos primero, para simplificar, imágenes de intensidad o niveles de gris (fotografías en blanco y negro) y no en color. Para digitalizar una fotografía y guardarla en la memoria de un ordenador, la dividimos en pequeños cuadrados, a los que llamaremos píxeles, y colocamos en cada uno de ellos un número que representa la luminancia o nivel de gris del píxel. El valor numérico 0 representa el negro, el valor numérico 255 representa el blanco, y los enteros



Figure 6: a) Imagen pancromática a la resolución de 0.7 m/píxel, denotada por u en la Sección 6. b) Canales de color $\vec{u} = (R, G, B)$ a la resolución de 2.8 m/píxel.

1, 2, ..., 254 representan los niveles de gris intermedios entre el negro y el blanco. Después de la división de la imagen en píxeles y de la codificación de los niveles de gris en números enteros de 0 a 255, tenemos la imagen digitalizada, es decir, la hemos convertido en una matriz de enteros que podemos almacenar en la memoria del ordenador. Visualizarla requiere, de alguna manera, reinvertir este proceso, transformar la imagen digital en cantidades de luz que representen la sensación asociada al nivel de gris correspondiente. Este proceso tiene lugar cuando miramos la imagen en la pantalla del ordenador. La Figura 7.a muestra una imagen de la que hemos extraído un fragmento representado en la Figura 7.b, siendo los valores numéricos del nivel de gris los de la Figura 7.c.

Diremos pues que una imagen digital es una función $u : \{0, 1, \dots, M - 1\} \times \{0, 1, \dots, N - 1\} \rightarrow [0, 255] \cap \mathbb{Z}$ donde, para simplificar, hemos escogido como rango de valores de u los valores enteros de 0 a 255 y $M, N \in \mathbb{Z}$. Una imagen analógica sería una función u definida en un rectángulo de \mathbb{R}^2 , que para simplificar podemos tomar como el cuadrado unidad $[0, 1] \times [0, 1]$, con valores reales. Digitalizar una imagen analógica supone discretizar tanto su dominio de definición (operación llamada muestreo) como su rango de valores (operación llamada cuantización).



Figure 7: a. Lena, b. Extracto de Lena, c. Niveles de gris

2.2 La geometría de las imágenes de intensidad



Figure 8: El despacho del INRIA.

Ante una imagen de cualquier escena natural podemos hacernos la siguiente pregunta básica, ¿cuáles son las formas geométricas presentes en la imagen ?. ¿De otra forma, cuáles son los objetos presentes en la imagen ?. ¿Cómo reconstruirlos ? La mayoría de estas preguntas sencillas son harto complejas y no podemos decir por ahora que exista una respuesta satisfactoria. Como ocurre muchas veces en la ciencia, conviene modificar ligeramente la pregunta de manera que seamos capaces de darle una respuesta que, aunque no nos satisfaga completamente, nos permita llegar a una conclusión operativa. Transformamos la pregunta anterior en la pregunta siguiente: ¿Cuáles son las formas **computables** presentes en una imagen ?. No vamos a abordar esta cuestión desde un punto de vista filosófico, más bien buscaremos una definición operativa de la idea de forma u objeto en una imagen. Ilustremos, antes de nada, la complejidad de esta tarea con la imagen de la Figura 8 que representa un despacho con un par de mesas y sillas. Esta observación, que es una tarea muy sencilla para la visión humana, supone de hecho una construcción muy elaborada: una observación más atenta de la imagen nos permite ver que lo que llamamos 'silla' no es más que un conjunto de manchas de niveles de gris muy diferentes que agrupamos perceptualmente y a lo que llamamos silla. Una definición de silla debería incluir todos los tipos diferentes de sillas, cuesta creer que la tarea de reconocerlas sea factible para un ordenador ...

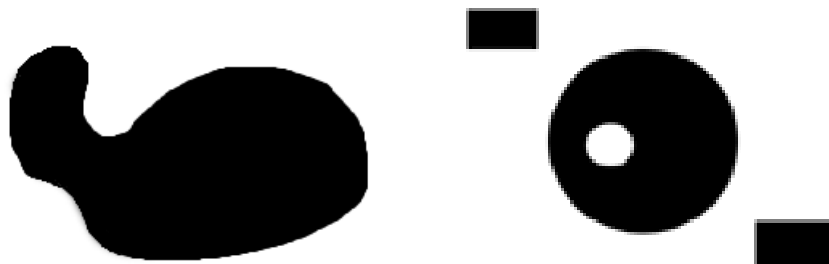


Figure 9: Imágenes binarias

Antes de dar nuestra respuesta a la pregunta anterior, vamos a considerar un caso más sencillo aunque no exento de dificultades: el caso de imágenes con dos niveles de gris, digamos blanco y negro, o imágenes binarias. La Figura 9 muestra ejemplos de imágenes binarias. Vemos unos 'objetos negros sobre fondo blanco', (¿'objetos blancos sobre un fondo negro' ?). La ambigüedad figura-fondo está ilustrada en la Figura 10.a. Otro tipo de dificultades que presentan las imágenes binarias está ilustrada por el triángulo de Kanizsa, Figura 10.b, que nos presenta un contorno ilusorio, ya que no existe una discontinuidad del nivel de gris en el contorno del triángulo, pero que a pesar de ello vemos con mucha fuerza perceptiva. No vamos a entrar en más detalles relativos a los problemas presentes en

la percepción de las imágenes binarias y mandamos para ello a la obra de los psicólogos de la Gestalt, entre ellos a la obra de G. Kanizsa, accesible en castellano [31].

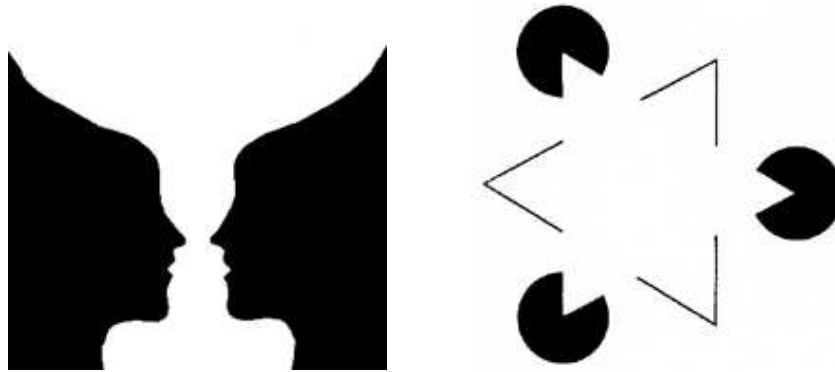


Figure 10: a) ¿Vaso o caras ? b) El triángulo de Kanizsa

A todos nos ha pasado alguna vez que hemos llegado al ordenador, mirado la pantalla y nos hemos dado cuenta de que estaba algo oscura por estar desajustado el contraste. En un caso así y para ver mejor las imágenes de la pantalla modificamos el contraste hasta que nuestra visión es satisfactoria. Algunos aumentaremos más que otros la luminosidad de la pantalla pero todos estaremos de acuerdo en que no por ello no estamos viendo la misma imagen. Admitimos que tanto antes como después del ajuste de contraste ambas imágenes nos proporcionan la misma información geométrica. Tomaremos este hecho básico como axioma. Para precisarlo, vamos a dar algunas definiciones.

Definición 1 Llamaremos cambio de contraste a cualquier función $g : \mathbb{R} \rightarrow \mathbb{R}$ continua y estrictamente creciente.

La noción de cambio de contraste pretende formalizar el ajuste de contraste del que hablamos en nuestra experiencia compartida. No se trata más que de un reordenamiento (biyectivo) de los niveles de gris. Esta noción contiene, entre otras cosas, un modelo matemático para los cambios de iluminación, ellos serían casos particulares de cambios de contraste. Podemos ahora enunciar nuestro axioma básico.

Axioma básico: Dada una imagen u y un cambio de contraste g , las imágenes u y $g(u)$ poseen la misma información geométrica.

En otras palabras, el contenido geométrico de una imagen u está asociado a su clase de equivalencia en el conjunto de las imágenes módulo los cambios de contraste. Ejemplos de cambios de contraste se producen cuando regulamos el contraste de la pantalla de un ordenador o cuando hacemos un cambio de iluminación global. De hecho, cualquier sensor o cámara traduce el continuum de intensidades luminosas siguiendo una curva de contraste específica, que en la práctica nos es desconocida. Todo ello llevó al físico y psicólogo de la Gestalt M. Wertheimer [49] a formular como principio básico que el nivel de gris no es un observable. Las imágenes son percibidas módulo un cambio de contraste arbitrario y desconocido. Una doctrina del análisis de imágenes, la Morfología Matemática, reconoce la invarianza por cambios de contraste como un axioma fundamental y propone que las operaciones que se realicen sobre las imágenes lo sean [44]. Si aceptamos este axioma, debemos aceptar sus consecuencias. Este axioma nos permite responder la pregunta que enunciamos al principio, nos permite decir cuáles son las formas geométricas de una imagen.

Antes de dar respuesta a la pregunta formulada más arriba, recordemos un hecho básico de las funciones con valores reales. Las funciones reales están dadas por sus conjuntos de nivel.

Definición 2 Consideremos la imagen $u : [0, 1]^2 \rightarrow \mathbb{R}$. Llamaremos conjuntos de nivel superiores (resp. inferiores) de u los conjuntos de la forma

$$X^\lambda u = \{x \in [0, 1]^2 : u(x) \geq \lambda\} \quad (\text{resp. } X_\lambda u = \{x \in [0, 1]^2 : u(x) \leq \lambda\}), \quad \lambda \in \mathbb{R}.$$

Por comodidad, utilizaremos la notación $[u \geq \lambda]$ o $[u \leq \lambda]$, $\lambda \in \mathbb{R}$, para denotar el conjunto de nivel superior, o inferior, de nivel λ .

Una función con valores reales puede describirse en términos de sus conjuntos de nivel. Enunciemos este resultado en términos precisos.

Proposición 1 Si $u : [0, 1]^2 \rightarrow [a, b]$ es una función, los conjuntos de nivel superiores de u ,

$$X^\lambda := X^\lambda u = \{x \in [0, 1]^2 : u(x) \geq \lambda\}, \quad \lambda \in \mathbb{R},$$

cumplen las propiedades

(i) $X^\lambda \subseteq X^\mu$ si $\lambda > \mu$, $X^a = [0, 1]^2$, $X^\lambda u = \emptyset$ para todo $\lambda > b$.

(ii) $X^\lambda = \bigcap_{\mu < \lambda} X^\mu$ para todo $\lambda > -\infty$.

Recíprocamente, si $(X^\lambda)_\lambda$ es una familia de conjuntos que cumple las propiedades (i) y (ii), entonces la función $u : [0, 1]^2 \rightarrow [a, b]$ definida por la fórmula

$$u(x) = \sup\{\lambda : x \in X^\lambda\} \tag{6}$$

tiene por conjuntos de nivel los conjuntos $X^\lambda u = X^\lambda$.

No vamos a dar en detalle la demostración de este hecho, por otro lado simple. No es más que una traducción de las propiedades elementales de los números reales.

Llamaremos descripción topográfica de una función a la descripción de la misma en términos de sus conjuntos de nivel tal como ha sido dada en la Proposición 1. Efectivamente, ésta es es prácticamente similar a la descripción de un mapa de alturas por curvas de nivel. Todos estamos acostumbrados a leer la altura del mapa mirando las curvas de nivel, pudiendo visualizar mentalmente la topografía, siempre y cuando conozcamos la altura de cada curva. Vamos a aclarar, pues, por qué hemos utilizado entonces conjuntos en vez de curvas de nivel. La descripción de la función altura en términos de curvas de nivel es perfecta si por todo punto pasa una y sólo una curva como ocurre con las funciones suaves. Sin embargo si la función es discontinua a lo largo de una curva rectificable (una curva de clase C^1 , por ejemplo), entonces no podemos afirmar que por los puntos de esta curva pasa una curva de nivel. De hecho, esto ocurriría en un acantilado, por citar un ejemplo. Los acantilados de una imagen representan los contornos de los objetos que aparecen, en muchos casos, como discontinuidades del nivel de gris. Para evitar esta complicación técnica hemos utilizado los conjuntos de nivel como descriptores de la topografía de una función con valores reales. Las curvas de nivel pueden recuperarse tomando las fronteras de los conjuntos de nivel, siempre que estas fronteras sean curvas, lo cual depende ya del modelo funcional que utilizemos para describir las imágenes (es decir, a qué tipo de espacio funcional pertenecen). Por ejemplo, si la función es suave podemos definir sus curvas de nivel por las fronteras (relativas al dominio $[0, 1]^2$)

$$\partial X^\lambda u, \quad \lambda \in \mathbb{R}.$$

Observemos que siendo complementarios $\partial X^\lambda u = \partial X_\lambda u$ para cada $\lambda \in \mathbb{R}$, de manera que podemos restringirnos a considerar solamente las fronteras de los conjuntos de nivel superiores. Genéricamente, es decir, para casi todos los niveles $\lambda \in \mathbb{R}$, se cumplirá que

$$\partial X^\lambda u = [u = \lambda]$$

con la excepción eventual de los rellanos.

Definición 3 Llamaremos curvas de nivel de u a las fronteras de sus conjuntos de nivel, sean éstos superiores o inferiores.

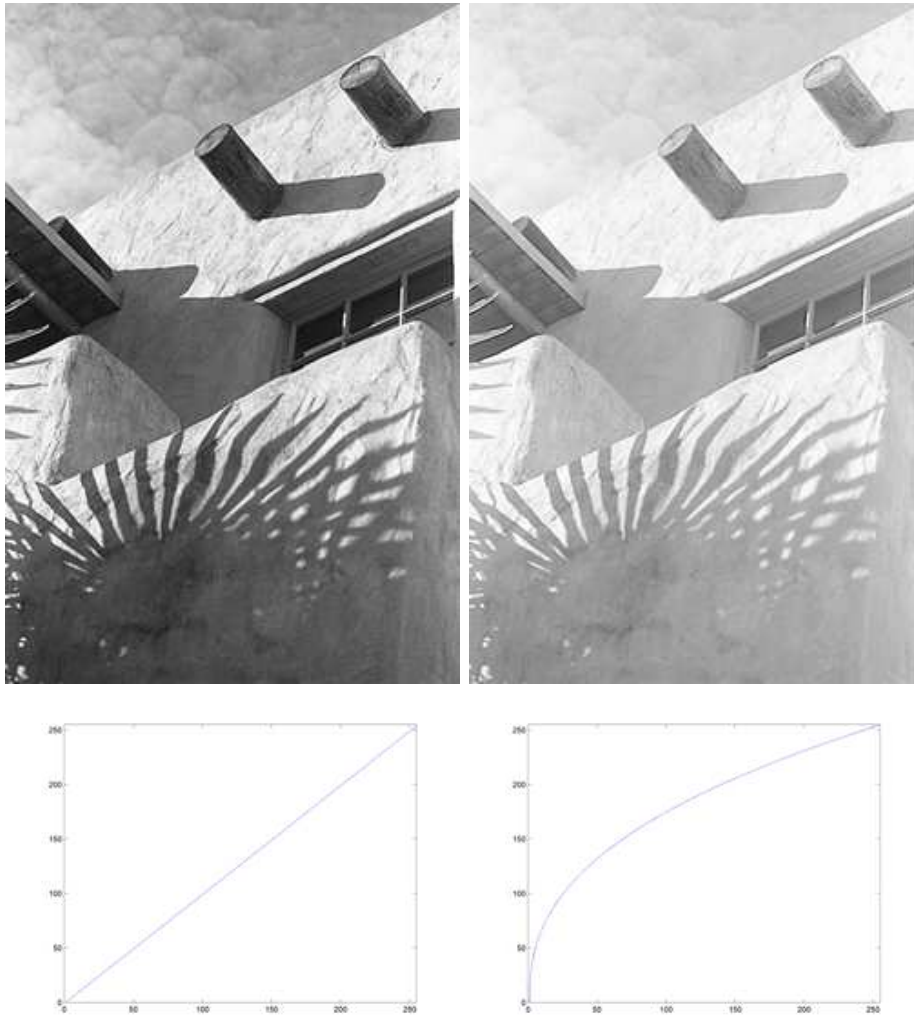


Figure 11: Cambios de contraste



Figure 12: Curvas de nivel de Lena correspondientes a los múltiplos de 20.

Más adelante mencionaremos otros modelos funcionales más generales y más adaptados a las imágenes que las funciones suaves y para los cuales las curvas de nivel són efectivamente uniones a lo sumo numerables de curvas de Jordan.

La Figura 12 representa las curvas de nivel – correspondientes a niveles múltiples de 20 – de la Figura 7.a. Podemos ver de manera compacta la geometría de la imagen 7.a.

Los conjuntos de nivel constituyen los invariantes básicos de una imagen módulo cambios de contraste, en otras palabras, se cumple:

Proposición 2 *Dos imágenes relacionadas por un cambio de contraste comparten sus familias de conjuntos de nivel.*

En efecto, consideremos una imagen $u : [0, 1]^2 \rightarrow \mathbb{R}$ y un cambio de contraste g . Observemos que, siendo continua y estrictamente creciente, g es una biyección de \mathbb{R} sobre su imagen. Entonces, para cada $\lambda \in g(\mathbb{R})$, se cumple

$$X^\lambda g(u) := \{x \in [0, 1]^2 : g(u(x)) \geq \lambda\} = \{x \in [0, 1]^2 : u(x) \geq g^{-1}(\lambda)\} = X^{g^{-1}(\lambda)}u.$$

En otras palabras las familias de conjuntos de nivel coinciden aunque tengamos que modificar las alturas correspondientes.

En términos de la interpretación topográfica, si miramos las curvas de nivel de un mapa sin ver la altura correspondiente estamos mirando de hecho muchos mapas, todos ellos relacionados por cambios de contraste. Para explicar de manera más precisa este último hecho necesitaríamos cambios de contraste más generales, que incluirían cambios de contraste locales, e incluso, inversiones de contraste. Dejamos esto a la reflexión del lector.

Para completar este círculo de ideas tan sólo queda por enunciar el siguiente resultado.

Proposición 3 *Si u y v son dos imágenes cuyas familias de conjuntos de nivel superiores coinciden, existe un cambio de contraste generalizado g tal que $v = g(u)$.*

Observemos que el enunciado habla de un cambio de contraste generalizado, de hecho de una función no decreciente $g : \mathbb{R} \rightarrow \mathbb{R}$. Todo esto forma parte de los detalles técnicos de las demostraciones y no es nuestro objetivo aquí insistir en ello.

Esta excursión nos permite responder a lo que podemos considerar como formas de una imagen. Admitiendo el axioma básico, todas las imágenes relacionadas por un cambio de contraste comparten la misma información geométrica y, según las Proposiciones 2 y 3, esta información consiste en los conjuntos de nivel comunes a dichas imágenes.

2.3 Cambios de contraste locales

Los conjuntos de nivel son pues objetos básicos en tratamiento y análisis de imágenes. Para tener una descripción más local de los objetos de una imagen algunos autores ([44], [15]) han propuesto considerar las componentes conexas de los conjuntos de nivel como objetos básicos. El argumento principal es que los cambios de contraste no son globales sino locales y dependen de las propiedades de reflectancia de los objetos. Por tanto, no solo los cambios de contraste globales, sino también los locales son irrelevantes a la hora de analizar una imagen. En [15] definimos una noción de cambio de contraste local y vimos como las componentes conexas de los conjuntos de nivel son invariantes bajo tales cambios de contraste. Esto conduce al concepto de **mapa topográfico**, es decir, la familia de componentes conexas de los conjuntos de nivel superiores $[u \geq \lambda]$ o inferiores $[u \leq \lambda]$. Buscando el nivel más grande de generalidad podríamos definir como *objetos básicos de la imagen u la σ -álgebra $\Sigma(u)$ engendrada por las componentes conexas de los conjuntos de bi-nivel $[\lambda \leq u < \mu]$* . Para funciones suaves, los átomos de esta σ -álgebra serían las componentes conexas de los isoniveles $[u = \lambda]$.

Este planteamiento posee aplicaciones interesantes a la comparación de imágenes. Para mencionar alguna de ellas mencionemos [10], donde los autores proponen comparar imágenes satélite de la misma

región geográfica tomadas en tiempos distintos (o correspondientes a diferentes canales espectrales) comparando los mapas topográficos. Mencionemos también el algoritmo de alineación de imágenes desarrollado en [40], donde el autor se propone calcular la transformación entre dos imágenes de la misma escena tomadas por la misma o por cámaras distintas de manera que la relación entre ambas está dada de manera aproximada por una transformación afín. La alineación de imágenes basada en componentes conexas de conjuntos de nivel funciona de manera eficaz cuando las técnicas clásicas de correlación fallan: cuando ambas imágenes no se corresponden con fotografías tomadas casi simultáneamente. Situaciones más generales requieren aún objetos básicos más generales como fragmentos de curvas de nivel [29].

2.4 Algunos modelos funcionales de imágenes

No hemos hablado hasta ahora de qué tipo de funciones pueden representar imágenes, éste es, de hecho, un problema muy complejo. Aunque no vamos a resolver este problema en esta sección, vamos a presentar un modelo que parece suficiente para muchas aplicaciones, entre ellas: segmentación de imágenes, restauración en el sentido de la deconvolución y supresión de ruido, aumento de resolución, inpainting, estudio de filtros de la morfología matemática, y un largo etcétera.

El modelo de que vamos a hablar fue introducido por L. Rudin, S. Osher y E. Fatemi en el contexto de la deconvolución y eliminación de ruido (restauración) en imágenes [43]. También fue utilizado en segmentación de imágenes por la escuela de De Giorgi para estudiar el modelo de segmentación de Mumford-Shah. Finalmente, fue utilizado también en [1] para fundamentar los desarrollos esbozados en las secciones anteriores, dando un contenido riguroso a los modelos de la morfología matemática para funciones con discontinuidades de salto, y poder estudiar tanto algunas familias de filtros del análisis de imágenes como los modelos variacionales de inpainting, de los que daremos cuenta en la Sección 4.

El objeto del trabajo [1] fue proponer un modelo funcional en el que se pudiera describir el mapa topográfico en términos de curvas de Jordan (orientadas). Este modelo funcional, llamado WBV (que se lee en inglés Weak BV, funciones débilmente de variación acotada), es una variante del espacio de las funciones de variación acotada BV.

Definición 4 Diremos que $u : [0, 1]^2 \rightarrow \mathbb{R}$ es una función de WBV si para casi todos los niveles $\lambda \in \mathbb{R}$, el conjunto de nivel $[u \geq \lambda]$ tiene perímetro finito.

La única parte técnica de esta definición es la noción de conjunto de perímetro finito. Ésta es una noción técnica de la que daremos algún detalle más abajo. Para seguir los desarrollos que siguen basta tener una comprensión intuitiva de lo que estamos hablando. Un conjunto tiene perímetro finito si su frontera (rigurosamente hablando, su frontera esencial) tiene una longitud finita. Por ejemplo, el conjunto acotado por una curva de Jordan rectificable tiene perímetro finito. Más generalmente, un conjunto como el anterior, acotado exteriormente por una curva de Jordan de longitud finita y por una cantidad a lo sumo numerable de agujeros delimitados por curvas de Jordan de longitud finita y cuya suma de longitudes sea finita, es también un conjunto de perímetro finito en \mathbb{R}^2 . Llamemos a estos conjuntos 'simples'. Finalmente, una reunión finita o numerable de conjuntos 'simples' cuya suma de longitudes de sus fronteras sea finita, es un conjunto de perímetro finito. Los conjuntos 'simples' pueden ser disjuntos o pueden tocarse en sus fronteras, siempre que al tocarse se toquen en un conjunto de longitud nula. Mientras la situación sea la dibujada aquí la descripción de estas uniones numerable de conjuntos 'simples' puede hacerse sin ambigüedad. Hemos demostrado en [1] que esta situación es la más general posible para los conjuntos de perímetro finito en \mathbb{R}^2 . En otras palabras, todo conjunto de perímetro finito de \mathbb{R}^2 es una unión numerable de conjuntos 'simples' que, si se tocan, lo hacen en un conjunto de longitud nula de su frontera. Además podemos considerar cada conjunto 'simple' como una componente conexas del conjunto. Hablaremos, pues, de componente conexas en vez de conjunto 'simple'.

Si u es una imagen de WBV , casi todos sus conjuntos de nivel tienen perímetro finito y están formados por uniones, a lo sumo numerables, de componentes conexas, conjuntos delimitados por curvas de Jordan rectificables. Nuestra conclusión puede resumirse:

Si $u \in WBV$, entonces las formas geométricas de u están dadas por las curvas de Jordan que forman las fronteras exteriores e interiores de las componentes conexas de los conjuntos de nivel de u .

Definición de función de variación acotada. Sea Ω un abierto de \mathbb{R}^N . Diremos que u es una función de variación acotada en Ω si $u \in L^1(\Omega)$ y su gradiente en el sentido de las distribuciones Du es una medida de Radon vectorial cuya variación total $|Du|(\Omega) := \sup\{\int_{\Omega} u \operatorname{div} \varphi \, dx : \varphi \in C_0^\infty(\Omega)\}$ es finita. Denotaremos por $BV(\Omega)$ al espacio de funciones de variación acotada en Ω . Recordemos que $C_0^\infty(\Omega)$ representa el espacio de las funciones infinitamente diferenciables y de soporte compacto en Ω .

Diremos que un conjunto medible $E \subseteq \mathbb{R}^N$ tiene perímetro finito en Ω si $\chi_E \in BV(\Omega)$. El perímetro de E en Ω se define por $P(E, \Omega) := |D\chi_E|(\Omega)$. Un resultado fundamental debido a de Giorgi [23] establece que $D\chi_E = \nu \mathcal{H}^{N-1}|_{\partial^* E}$ donde $\partial^* E$ es un subconjunto de la frontera topológica de E para el cual tiene sentido definir la normal exterior ν en el sentido de teoría geométrica de la medida, y \mathcal{H}^{N-1} representa la medida de Hausdorff $N - 1$ -dimensional. La frontera $\partial^* E$ recibe el nombre de frontera esencial de E .

Observación: Las funciones de WBV son funciones de BV módulo un cambio de contraste, es decir, para cada función $u \in WBV$ existe una función $g : \mathbb{R} \rightarrow \mathbb{R}$ continua, acotada y estrictamente creciente tal que $g(u)$ es una función de variación acotada, i.e., $g(u) \in BV$.

Las funciones continuas como modelos de imágenes. A pesar de que las funciones BV o WBV son modelos mejores para describir imágenes, debido a que permiten discontinuidades de salto a lo largo de curvas rectificables, en algunos contextos, como los dos primeros problemas de interpolación mencionados en la Sección 1 las funciones continuas son un buen modelo para abordarlos. En ambos casos, pretendemos interpolar una función en un dominio en cuya frontera conocemos su valor. Para ello resolvemos una ecuación en derivadas parciales como (1) ó (4) y la solución obtenida es una función continua.

3 Una aproximación axiomática al problema de interpolación

Los dos primeros ejemplos considerados en la introducción pueden formularse como problemas de interpolación de datos escalares dispersos dados sobre un conjunto de curvas y/o puntos del plano. Para abordarlo, seguiremos el planteamiento axiomático propuesto en [18].

Los principales operadores de interpolación aparecen inmediatamente considerando qué tipo de media (lineal o no lineal) debe satisfacer la función a interpolar $u(x)$. Supongamos que conocemos $u(x)$ en \mathbb{R}^2 (que suponemos discretizado) salvo un punto $x_0 \in \mathbb{R}^2$. ¿Qué valor deberíamos escoger en $x_0 \in \mathbb{R}^2$? Planteemos algunas posibilidades:

- 1) $u(x_0)$ es el valor medio de los píxeles vecinos.
- 2) $u(x_0)$ es el valor mediano de los píxeles vecinos.
- 3) $u(x_0)$ se obtiene propagando el valor de los píxeles vecinos (según precisaremos luego).

Si suponemos ahora que la función u es un interpolante de sí misma, es decir, que para todo punto x_0 , $u(x_0) = (\text{valor medio de } (u(x)))$ en un entorno de x_0 independientemente de lo que signifique 'valor medio', entonces suponiendo que u sea de clase C^2 , bajo las posibilidades anteriores obtenemos:

- 1) $u(x_0) = \frac{1}{4}(u(x_0 + (h, 0)) + u(x_0 - (h, 0)) + u(x_0 + (0, h)) + u(x_0 - (0, h)))$. Calculando la diferencia y haciendo $h \rightarrow 0$, el desarrollo de Taylor permite escribir

$$\Delta u(x_0) = \frac{\partial^2 u}{\partial x_1^2}(x_0) + \frac{\partial^2 u}{\partial x_2^2}(x_0) = 0.$$

Ésta es la que podríamos llamar interpolación estándar. El cálculo anterior puede extenderse a muchas otras medias lineales.

2) $u(x_0) = \text{valor mediano de } \{u(y), y \in D(x_0, h)\}$, donde $D(x_0, h)$ es un disco de radio h . Después de algunos cálculos, puede demostrarse en este caso ([25]) que haciendo $h \rightarrow 0$ se obtiene

$$\text{curv}(u)(x_0) = \frac{1}{|Du|^3} D^2u(Du^\perp, Du^\perp) = \frac{u_{x_2}^2 u_{x_1 x_1} - 2u_{x_1} u_{x_2} u_{x_1 x_2} + u_{x_1}^2 u_{x_2 x_2}}{(u_{x_1}^2 + u_{x_2}^2)^{3/2}} = 0$$

donde $\text{curv}(u)(x_0)$ es la curvatura de la curva de nivel que pasa por x_0 y Du^\perp es perpendicular a Du , $|Du^\perp| = |Du|$, siendo $Du = (u_{x_1}, u_{x_2})$ el gradiente de u y D^2u la matriz Hessiana de u . En lo sucesivo utilizaremos la notación $A(x, y) = \sum_{i,j=1}^2 a_{ij} x_i y_j$, siendo $A = (a_{ij})_{i,j}$ una matriz 2×2 y $x, y \in \mathbb{R}^2$.

3) En el caso de propagación de valores discutido por J.R. Casas, si u es de clase C^2 en x y si $u(x)$ se obtiene promediando los valores propagados de píxeles vecinos siguiendo la dirección del gradiente, podemos escribir

$$u(x) = \frac{1}{2} (u(x + hDu) + u(x - hDu)) + o(h^2).$$

Haciendo $h \rightarrow 0$ y calculando el desarrollo de Taylor, obtenemos

$$D^2u(Du, Du) = 0.$$

En conclusión, después de analizar los algoritmos anteriores, hemos visto que la función interpolante satisface una de las ecuaciones en derivadas parciales de tipo elíptico siguientes

$$\Delta u = 0. \tag{7}$$

$$D^2u(Du^\perp, Du^\perp) = 0, \tag{8}$$

$$D^2u(Du, Du) = 0. \tag{9}$$

Un análisis axiomático detallado de los operadores de interpolación permite obtener los operadores anteriores e identificar todos los posibles operadores de interpolación que satisfacen una serie razonable de requisitos. De hecho, veremos cómo los tres operadores analizados (7), (8), (9) describen esencialmente todas las posibilidades de que disponemos para generar tales operadores. El segundo (8) no admite, en general, una solución. El primero de ellos (7) es un operador de interpolación excelente y estándar pero no permite interpolar puntos aislados. Como es bien sabido, $\Delta u = 0$, $u = 0$ sobre $\partial D(0, r)$, $u(0, 0) = 1$ no tiene solución. La ecuación (9), en cambio, produce el cono $u(x) = |x| - 1$, como resultado de la interpolación. Dicha ecuación (9) es, digamos, nueva en el contexto de procesamiento de imágenes. Para ella, usando los resultados de Aronsson [5] y Jensen [28], se demuestra que existe un interpolante Lipschitz continuo si los datos están dados por funciones Lipschitzianas en un conjunto de curvas y/o puntos. Aunque nos inspiramos en los trabajos de Casas [14] para llegar a dicho operador, el método de este autor no produce necesariamente una interpolación continua de los datos, en contraste con el operador de Aronsson [5] (véase también [28]).

Sea C un conjunto de curvas de Jordan en \mathbb{R}^2 . Para cada $\Gamma \in C$, sea $F(\Gamma)$ el conjunto de funciones continuas sobre Γ . Consideramos un operador de interpolación como una aplicación E que asocia a cada curva $\Gamma \in C$ y cada $\varphi \in F(\Gamma)$ una única función $E(\varphi, \Gamma)$ definida en la región $D(\Gamma)$ interior a Γ , satisfaciendo los axiomas:

(A1) Principio de comparación:

$$E(\varphi, \Gamma) \leq E(\psi, \Gamma)$$

para cualquier $\Gamma \in C$ y cualesquiera $\varphi, \psi \in F(\Gamma)$ tales que $\varphi \leq \psi$.

(A2) Principio de estabilidad:

$$E(E(\varphi, \Gamma) |_{\Gamma'}, \Gamma') = E(\varphi, \Gamma) |_{D(\Gamma')}$$

para cualquier $\Gamma \in C$, cualquier $\varphi \in F(\Gamma)$ y $\Gamma' \in C$ tales que $D(\Gamma') \subseteq D(\Gamma)$.

Denotemos por $SM(2)$ el conjunto de las matrices simétricas 2×2 .

(A3) Principio de regularidad: Sean $A \in SM(2)$, $p \in \mathbb{R}^2 - \{0\}$, $c \in \mathbb{R}$ y

$$Q(y) = \frac{A(y-x, y-x)}{2} + \langle p, y-x \rangle + c.$$

(donde $\langle x, y \rangle = \sum_{i=1}^2 x_i y_i$). Entonces

$$\frac{E(Q |_{\partial D(x,r)}, \partial D(x,r))(x) - Q(x)}{r^2/2} \rightarrow F(A, p, c, x) \quad \text{cuando } r \rightarrow 0+ \quad (10)$$

donde $F : SM(2) \times \mathbb{R}^2 - \{0\} \times \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función continua.

Junto con estos axiomas básicos, consideremos los siguientes axiomas que expresan un conjunto de invarianzas geométricas de los operadores de interpolación.

(A4) Invarianza por traslación:

$$E(\tau_h \varphi, \Gamma - h) = \tau_h E(\varphi, \Gamma)$$

siendo $\tau_h \varphi(x) = \varphi(x+h)$, $h \in \mathbb{R}^2$, $\varphi \in F(\Gamma)$, $\Gamma \in \mathcal{C}$.

(A5) Invarianza por rotación:

$$E(R\varphi, R\Gamma) = RE(\varphi, \Gamma)$$

siendo $R\varphi(x) = \varphi(R^t x)$, y R una matriz de rotación en \mathbb{R}^2 , $\varphi \in F(\Gamma)$, $\Gamma \in \mathcal{C}$.

(A6) Invarianza por traslación del nivel de gris:

$$E(\varphi + c, \Gamma) = E(\varphi, \Gamma) + c$$

para cualesquiera $\Gamma \in \mathcal{C}$, $\varphi \in F(\Gamma)$ y $c \in \mathbb{R}$.

(A7) Invarianza por cambio de escala lineal del nivel de gris:

$$E(\lambda\varphi, \Gamma) = \lambda E(\varphi, \Gamma) \quad \text{para cualquier } \lambda \in \mathbb{R}$$

(A8) Invarianza por zoom:

$$E(\delta_\lambda \varphi, \lambda^{-1} \Gamma) = \delta_\lambda E(\varphi, \Gamma)$$

donde $\delta_\lambda \varphi(x) = \varphi(\lambda x)$, $\lambda > 0$.

Mencionemos tan sólo un resultado para ilustrar la clasificación de los operadores de interpolación que satisfacen los axiomas anteriores (véase [18]). Como referencia para el concepto de solución de viscosidad mencionemos [20]. Dado $p \in \mathbb{R}^2$, $p \neq 0$, sea R_p la matriz de rotación tal que $R_p^t p = |p|e_1$ donde $e_1 = (1, 0)$.

Teorema 1 *Supongamos que el operador E satisface los axiomas (A1) – (A8). Sean $\varphi \in \mathcal{C}(\Gamma)$, $u = E(\varphi, \Gamma)$. Entonces u es la solución de viscosidad de*

$$\begin{aligned} G(R_{\nabla u}^t D^2 u R_{\nabla u}) &= 0 \quad \text{in } D(\Gamma) \\ u|_{\Gamma} &= \varphi \end{aligned} \quad (11)$$

donde $G(A) = F(A, e_1)$, $A \in SM(2)$. Por consiguiente, G es una función continua y no decreciente de A tal que $G(\lambda A) = \lambda G(A)$ para todo $\lambda \in \mathbb{R}$ y toda $A \in SM(2)$.

Supongamos que E satisface (A1) – (A8). Sea

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

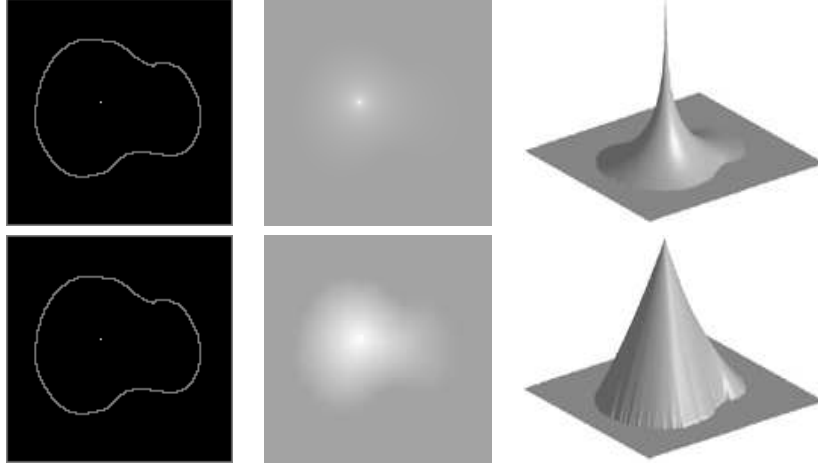


Figure 13: Arriba y de derecha a izquierda: el dato inicial, resultado de la interpolación usando el Laplaciano y su representación 3D. Abajo y de derecha a izquierda: el dato inicial, resultado de la interpolación usando el AMLE y su representación 3D.

y escribamos, por simplicidad $G(a, b, c)$, en vez de $G(A)$. Sea $\nu = \frac{p}{|p|}$, $p \in \mathbb{R}^2$, $p \neq 0$. Entonces $R_p(x) = e_1 \otimes \nu(x) + e_2 \otimes \nu^\perp(x)$ y podemos escribir

$$R_p^t A R_p = \begin{pmatrix} D^2 u(\nu, \nu) & D^2 u(\nu, \nu^\perp) \\ D^2 u(\nu^\perp, \nu) & D^2 u(\nu^\perp, \nu^\perp) \end{pmatrix}.$$

Podemos escribir la ecuación (11) como

$$G \left(D^2 u \left(\frac{Du}{|Du|}, \frac{Du}{|Du|} \right), D^2 u \left(\frac{Du}{|Du|}, \frac{Du^\perp}{|Du|} \right), D^2 u \left(\frac{Du^\perp}{|Du|}, \frac{Du^\perp}{|Du|} \right) \right) = 0.$$

Como discutimos en [18], suponiendo que la función G es diferenciable en $(0, 0, 0)$, el único operador de la familia anterior capaz de interpolar datos dados sobre curvas y/o puntos y mantener la solución Lipschitz cuando los datos lo sean es el operador asociado a la ecuación

$$D^2 u \left(\frac{Du}{|Du|}, \frac{Du}{|Du|} \right) = 0. \quad (12)$$

Si Ω es un dominio en \mathbb{R}^N , podemos preguntar si existe una función u Lipschitz en Ω tal que

$$\|Du\|_{L^\infty(\tilde{\Omega}; \mathbb{R}^N)} \leq \|Dw\|_{L^\infty(\tilde{\Omega}; \mathbb{R}^N)}$$

para todo $\tilde{\Omega} \subseteq \Omega$ y toda función w tal que $u - w$ es Lipschitz en $\tilde{\Omega}$ y $u = w$ en $\partial\tilde{\Omega}$. Si existe, tal función se llama un interpolante AMLE (Absolutely Minimizing Lipschitz Extension) de $w|_{\partial\Omega}$ en Ω . Aronsson demostró [5] que si u es un AMLE y es de clase C^2 en Ω , entonces u es una solución clásica de

$$D^2 u(Du, Du) = 0 \quad \text{en } \Omega. \quad (13)$$

Luego, Jensen demostró [28] que si u es un AMLE, entonces u es una solución de (13) en el sentido de viscosidad, siendo la solución única en este sentido. Las ecuaciones (12) y (13) son la misma desde el punto de vista de las soluciones de viscosidad. En un trabajo posterior, Frédéric Cao [13] demostró que los resultados de interpolación AMLE pueden extenderse a ciertos datos discontinuos.

Digamos finalmente que los resultados anteriores pueden extenderse a la interpolación de datos sobre superficies (véase la Figura 15).

En la Figura 13 podemos ver el resultado de la interpolación de un cono con el Laplaciano y el AMLE. Aunque estrictamente hablando no podemos imponer el valor de una función armónica (solución de (7)) en un punto, la Figura 13 nos muestra la solución numérica. La Figura 14.b nos muestra el resultado de la interpolación de las curvas de nivel de la Figura 2. Este resultado puede compararse con la Figura 14.a donde se muestra la imagen constante a trozos siendo su nivel de gris en cada región de 2 la constante correspondiente al nivel de gris de una de sus fronteras.



Figure 14: Izquierda: Imagen cuantizada de 20 en 20. Derecha: Interpolación AMLE de la imagen de la izquierda.

4 Modelos variacionales en inpainting

4.1 Reconstrucción del mapa topográfico basada en la elástica

Como propuso Masnou en [37, 39], el problema de inpainting puede verse como un problema de completación del mapa topográfico. Si u es una función de variación acotada en $D = [0, 1]^2$, de acuerdo con los resultados expuestos en la Sección 2.4, casi todos los conjuntos de nivel u tienen perímetro finito y pueden descomponerse en componentes conexas [1] de forma que la frontera esencial de cada componente conexa está formada por una curva de Jordan exterior y, a lo sumo, un conjunto numerable de curvas de Jordan interiores, todas ellas rectificables. Además, dos de tales curvas pueden tocarse, si lo hacen, en un conjunto de longitud nula [1]. Ya que los conjuntos de nivel superiores de u forman una familia monótona de conjuntos, la fórmula de reconstrucción

$$u(x) = \sup\{\lambda : x \in X^\lambda u\}, \quad (14)$$

es válida para casi todo $x \in D$ [1].



Figure 15: Interpolación de un modelo de elevación digital a partir de algunas curvas de nivel sobre la esfera. a) datos sobre la esfera, b) la interpolación AMLE, c) el mismo resultado representado como un grafo sobre la esfera.

Supongamos que $\tilde{\Omega}$ es un subconjunto abierto de D cuya frontera es Lipschitziana. Supongamos que conocemos la imagen $u_0 : D \setminus \tilde{\Omega} \rightarrow [a, b]$, $0 \leq a < b$, y queremos reconstruir la imagen u_0 en $\tilde{\Omega}$ usando la información de u_0 en $D \setminus \tilde{\Omega}$. Llamaremos a $\tilde{\Omega}$ un *agujero* de la imagen. Supondremos que la función u_0 es una función de variación acotada en $D \setminus \tilde{\Omega}$. Sabemos que el mapa topográfico de u_0 en $D \setminus \tilde{\Omega}$ está formado por curvas de Jordan. Además, podemos suponer, ampliando para ello el *agujero* si fuese necesario, que para casi todo nivel λ las curvas de nivel de $X^\lambda u_0$ intersectan transversalmente a la frontera del *agujero* en un número finito de puntos [37]. Denotemos por $\Lambda \subseteq \mathbb{R}$ a la familia de dichos niveles. S. Masnou formuló el problema de inpainting [37, 39, 38] como el problema de reconstrucción del mapa topográfico de u_0 dentro de $\tilde{\Omega}$. Dado $\lambda \in \Lambda$ y dos puntos $p, q \in X_\lambda u_0 \cap \partial \tilde{\Omega}$ denotamos por τ_p y τ_q a los vectores tangentes a la curva de nivel $\partial^* X_\lambda u_0$ en p y q , respectivamente. Siguiendo a Mumford-Nitzberg-Shiota, la curva de unión de p y q propuesta por Masnou [37, 39] es la curva Γ contenida en $\tilde{\Omega}$ que minimiza la energía

$$\int_{\Gamma} (\alpha + \beta |\kappa|^p) d\mathcal{H}^1 + (\tau_p, \tau_{\Gamma}(p)) + (\tau_q, \tau_{\Gamma}(q)) \quad (15)$$

donde κ representa la curvatura de Γ , y $\tau_{\Gamma}(p)$ y $\tau_{\Gamma}(q)$ representan las tangentes a Γ en los puntos p y q , respectivamente; siendo $(\tau_p, \tau_{\Gamma}(p))$ el ángulo que forman los vectores τ_p y $\tau_{\Gamma}(p)$, y $(\tau_q, \tau_{\Gamma}(q))$ el ángulo respectivo para q . Las constantes α, β son positivas y $p \geq 1$. La solución del problema de inpainting se obtiene minimizando el funcional

$$\int_{-\infty}^{+\infty} \sum_{\Gamma \in F_\lambda} \left(\int_{\Gamma} (\alpha + \beta |\kappa|^p) d\mathcal{H}^1 + (\tau_p, \tau_{\Gamma}(p)) + (\tau_q, \tau_{\Gamma}(q)) \right) d\lambda \quad (16)$$

donde F_λ denota la familia of curvas de unión asociadas al conjunto de nivel $X^\lambda u_0$. Como hemos indicado más arriba, para casi todo nivel $\lambda \in \mathbb{R}$, la familia F_λ es finita, de forma que la suma in (16) is finita de modo genérico. En su trabajo [37, 39] los autores demostraron que existe un mínimo para el problema de inpainting en $\tilde{\Omega}$ para todo $p \geq 1$ y, en el caso del exponente $p = 1$, propusieron un algoritmo de programación dinámica para calcular las uniones óptimas entre puntos de $\partial X_\lambda u_0 \cap \partial \tilde{\Omega}$. Observemos que para $p = 1$ las curvas de unión son líneas rectas. Con esta estrategia, calcularon el mínimo del funcional (16) [37, 38]. Una variación del funcional (16) fué propuesta en [2]. Los autores observaron que calculando la integral $\int_{\Gamma} (\alpha + \beta |\kappa|^p) d\mathcal{H}^1$ no sólo a lo largo de la curva de unión sino también sobre un trozo de la curva de nivel exterior a $\tilde{\Omega}$ el criterio (16) podía escribirse como

$$\int_{-\infty}^{+\infty} \sum_{\Gamma \in F_\lambda} \left(\int_{\Gamma} (\alpha + \beta |\kappa|^p) d\mathcal{H}^1 \right) d\lambda \quad (17)$$

donde ahora las curvas de F_λ contienen la curva de unión y un trozo de curva de nivel de u_0 en $\Omega \setminus \tilde{\Omega}$ para un dominio $\Omega \supseteq \tilde{\Omega}$. Para proceder con esta estrategia los autores suponen que las curvas de nivel de u_0 están en $W^{2,p}$ en $\Omega \setminus \tilde{\Omega}$. En este caso, para funciones u de clase C^2 , el funcional (17) puede escribirse de la forma

$$\int_{\Omega} |\nabla u| \left(\alpha + \beta \left| \operatorname{div} \frac{\nabla u}{|\nabla u|} \right|^p \right) dx \quad (18)$$

usando la convención de que el integrando es 0 cuando $|\nabla u| = 0$. Este funcional fué estudiado en [2] cuando el dominio D y el *agujero* $\tilde{\Omega}$ son subconjuntos de \mathbb{R}^N con $N \geq 2$, así como su correspondiente relajación que como demostraron los autores coincide con el funcional

$$\int_R \int_{\partial[u \geq t]} (\alpha + \beta |H_{[u \geq t]}|^p) d\mathcal{H}^{N-1} dt \quad (19)$$

para funciones $u \in C^2(\Omega)$, $N \geq 2$, $p > N - 1$, siendo $H_{[u \geq t]}$ la curvatura media de $[u \geq t]$.

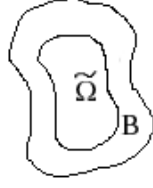


Figure 16: El agujero y la banda

4.2 Interpolación conjunta de campos de vectores y niveles de gris

Supondremos en esta sección que D es un hiperrectángulo de \mathbb{R}^N , $N \geq 2$, que representa el dominio de la imagen. Sean $\Omega, \tilde{\Omega}$ dos dominios abiertos y acotados de \mathbb{R}^N cuya frontera es Lipschitziana y supongamos que $\tilde{\Omega} \subset\subset \Omega \subset\subset D$. Para simplificar la presentación supondremos que Ω no toca a la frontera de D . Definimos $B := \Omega \setminus \tilde{\Omega}$, conjunto al que llamaremos la banda alrededor de $\tilde{\Omega}$ (véase la Figura 16). Supondremos que la imagen u_0 está dada en $D \setminus \tilde{\Omega}$ y es de variación acotada en este dominio. En una serie de trabajos [6, 8, 7] los autores propusieron restaurar el agujero $\tilde{\Omega}$ de la imagen utilizando tanto el nivel de gris como las tangentes (o las normales) a las curvas de nivel conocidas en B intentando prolongar las curvas de nivel de u_0 hacia el interior de $\tilde{\Omega}$ de acuerdo con el principio de la *buena continuación*.

Denotemos por θ_0 al campo de direcciones del gradiente de u_0 en $D \setminus \tilde{\Omega}$, es decir, θ_0 es un campo de vectores con valores en \mathbb{R}^N que satisface $\theta_0(x) \cdot Du_0(x) = |Du_0(x)|$ como medidas en B , y $|\theta_0(x)| \leq 1$ casi por todas partes en B . Supondremos que $\theta_0(x)$ has tiene una traza en $\partial\Omega$.

Planteamos el problema de inpainting de siguiente forma: ¿Podemos extender (de forma razonable) el par de funciones (u_0, θ_0) definidas en B a un par de funciones (u, θ) dentro de $\tilde{\Omega}$? Por supuesto, hemos de precisar qué entendemos por razonable.

Nos proponemos dar una formulación variacional de este problema y discutir un funcional de energía diseñado a tal efecto. Los datos del problema están dados sobre B y debemos procurar que la solución (u, θ) esté cerca de los datos en B . El campo de vectores debería satisfacer $\theta \cdot \nu^\Omega = \theta_0 \cdot \nu^\Omega$, $|\theta| \leq 1$ en Ω , y debería estar relacionado con u por la relación $\theta \cdot Du = |Du|$, es decir, deberíamos imponer que θ es el campo de direcciones del gradiente de u . La condición $|\theta(x)| \leq 1$ debería interpretarse como una relajación de este hecho. De hecho puede ocurrir que $\theta(x) = 0$ y, en este caso, no podemos normalizar el campo. Hemos de tener en mente que lo ideal sería que $\theta = \frac{Du}{|Du|}$, siendo u una función suave tal que $Du(x) \neq 0$ para todo $x \in \Omega$. Finalmente, deberíamos imponer que θ_0 se prolongue de manera suave en el campo θ en Ω . Impondremos esto observando que, si θ representa las direcciones de las normales a las hipersuperficies $u(x) = \lambda$, $\lambda \in \mathbb{R}$, entonces un término como $\text{div}(\theta)$ representa su curvatura media. Para imponer la continuación suave de las hipersuferficies de nivel de u_0 dentro de Ω exigiremos que $\text{div}(\theta) \in L^p(\Omega)$, $p > 1$. Por consistencia, supondremos que $\text{div} \theta_0 \in L^p(B)$.

Interpretando la energía de la elástica en este contexto, proponemos minimizar el funcional

$$\begin{aligned} & \text{Minimize } \int_{\Omega} |\text{div}(\theta)|^p (\gamma + \beta |\nabla K * u|) dx \\ & |\theta| \leq 1, |Du| - \theta \cdot Du = 0 \text{ in } \Omega \\ & |u| \leq M \\ & u = u_0 \text{ in } B, \theta \cdot \nu^\Omega|_{\partial\Omega} = \theta_0 \cdot \nu^\Omega|_{\partial\Omega}, \end{aligned} \tag{20}$$

donde $p > 1$, $\gamma > 0$, $\beta \geq 0$, K es un núcleo regularizador de clase C^1 tal que $K(x) > 0$ casi por todas partes, $M = \sup_{x \in B} |u_0(x)|$, y ν^Ω es la normal unitaria exterior a Ω . La convolución de Du con K en (20) es útil para demostrar la existencia de un mínimo de (20).

Minimizamos el funcional en el espacio $\mathcal{E}_p(\Omega, B, \theta_0)$ formado por los pares de funciones (u, θ) donde $u \in BV(\Omega)$, y θ es un campo vectorial en $L^\infty(\Omega, \mathbb{R}^N)$, $|\theta| \leq 1$, $\text{div}(\theta) \in L^p(\Omega)$, $\theta \cdot Du = |Du|$,

$$\theta \cdot \nu^\Omega = \theta_0 \cdot \nu^\Omega \text{ en } \partial\Omega.$$

Comentarios sobre el modelo (20).

A) ¿Cuál es el papel de la banda B ? Para discutir esta cuestión consideremos la Figura 17.a que representa una cinta gris en un fondo negro y que ha sido parcialmente tapada por un cuadrado blanco $\tilde{\Omega}$. Supongamos que los lados del cuadrado son perpendiculares a las curvas de nivel de la cinta. En estas condiciones, la componente normal de θ_0 en $\partial\tilde{\Omega}$ es nula. Por consiguiente, siendo las condiciones de frontera naturales del problema las dadas por $\theta_0 \cdot \nu^{\tilde{\Omega}}|_{\partial\tilde{\Omega}}$, se cumple que $\theta_0 \cdot \nu^{\tilde{\Omega}}|_{\partial\tilde{\Omega}} = 0$. En particular, el campo $\theta = 0$ satisface esta condición. Si no somos capaces de propagar θ dentro de $\tilde{\Omega}$ estamos ante una situación desagradable, ya que ello implica que no podremos propagar los valores de u desde la frontera hacia el interior de Ω . Si escribimos el funcional (20) con $\theta = 0$, $\alpha = 1$, el funcional resultante coincide con la variación total [43]. La decisión de extender la cinta gris o el fondo negro hacia el interior del agujero Ω se tomará en función del perímetro de las discontinuidades de la función resultante en el agujero. Para superar esta dificultad, introducimos la banda B alrededor de Ω como una manera de garantizar que el campo θ esté presente en el funcional. En la Figura 17.b mostramos el resultado de la interpolación con $\theta = 0$ en $\tilde{\Omega}$. En la Figura 17.c mostramos el resultado de la interpolación obtenido con el funcional (20) usando la banda B y el campo θ en ella. Aunque la banda B podría extenderse a todo el dominio $D \setminus \tilde{\Omega}$, en la práctica usamos tan sólo la información en una banda de unos pocos píxeles alrededor del agujero.

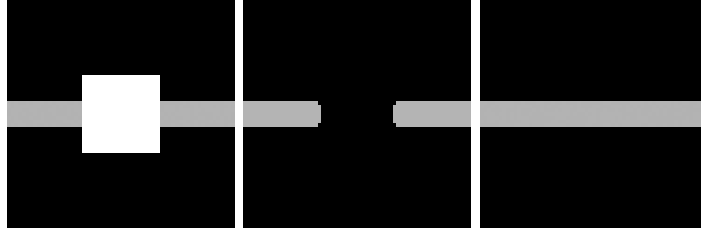


Figure 17: a) Una cinta tapada por un cuadrado. b) Resultado obtenido usando (20) con $\theta = 0$. c) Resultado obtenido usando (20).

B) Si $N = 2$ y u es la función característica de la región interior a una curva regular (de clase C^2) C entonces podemos escribir la integral

$$\beta \int_{\Omega} |\operatorname{div}(\theta)|^p |Du| + \alpha \int_{\Omega} |Du| \quad (21)$$

como $\int_C (\alpha + \beta |\kappa|^p) ds$, donde κ es la curvatura de C . Si $p = 2$, esta energía es la de la elástica de Euler (5). La elástica (5) fué propuesta en [42] cómo técnica para recuperar los contornos de un objeto ocultos por otros con el objetivo de segmentar una imagen teniendo en cuenta la relación de profundidad entre objetos de la escena, ya que proporciona curvas suaves y cortas. En términos de funciones características el funcional (5) toma la forma

$$\int |\nabla u| \left(\alpha + \beta \left| \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) \right|^2 \right). \quad (22)$$

El inconveniente es que este funcional no es semicontinuo inferiormente [12]. Como se demuestra en [2], el funcional (22) coincide con el propuesto por Masnou-Morel [37, 38, 39] para funciones suaves.

El funcional puede ser considerado también como una formulación relajada de la elástica. En nuestro caso, hemos introducido el campo θ como una variable independiente relacionada con u por $\theta \cdot Du = |Du|$. La convolución de Du con K permite evitar las dificultades planteadas por un estudio directo de (22).

C) Los coeficientes γ y β son > 0 . La positividad de γ implica que $\operatorname{div}(\theta) \in L^p$ lo que a su vez garantiza la regularidad de las curvas de nivel de u ([7]). Si $\beta > 0$ podemos tomar en consideración la contribución de la curvatura sobre la curva de nivel correspondiente a la frontera del objeto.

D) En la práctica usamos el funcional (20) para interpolar los conjuntos de nivel $[u_0 \geq \lambda]$, y obtener los conjuntos de nivel $X^\lambda u$ de una función u , que se reconstruye en Ω usando la fórmula (6). Para garantizar que los conjuntos de nivel reconstruidos corresponden a una función u hemos de imponer que $X^{\lambda+1}u \subseteq X^\lambda u$. De hecho calculamos una solución verificando esta propiedad.

Aunque el funcional (20) puede ser directamente usado para interpolar funciones, la energía pondera las discontinuidades de la imagen por el tamaño del salto lo cual tiene un efecto difusivo sobre el resultado. Si trabajamos con los conjuntos de nivel todos ellos son tratados en pie de igualdad y la energía (20) sólo contiene cantidades geométricas (longitud y curvatura).

E) La descomposición de la imagen u_0 en conjuntos de nivel superiores y la reconstrucción de la solución introduce un sesgo a favor de dichos conjuntos respecto a los conjuntos de nivel inferiores. El efecto puede verse en la Figura 18. La Figura 18.a nos muestra la imagen que queremos reconstruir. La descomposición en conjuntos de nivel superiores proporciona la solución de la Figura 18.b, ya que favorece que el objeto cuyo nivel es 210 pase por encima del objeto cuyo nivel es 0. Su hubiésemos usado los conjuntos de nivel inferiores, las cintas negras habrían pasado por encima de las blancas. En cualquier caso, ello demuestra que no existe una solución única de este problema y nuestra elección proporciona una de ellas.

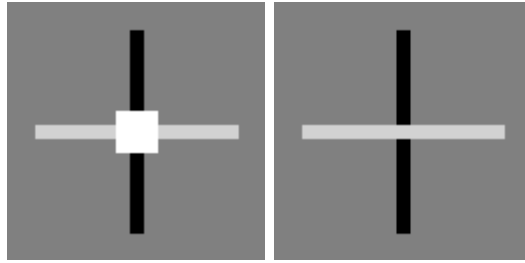


Figure 18: a) Imagen con *agujero*. b) Imagen reconstruida usando (20). La descomposición de la imagen en conjuntos de nivel superiores provoca que la banda blanca pase por encima de la negra.

Un análisis numérico directo del modelo (22) fué abordado en [19] aproximando el campo $\frac{Du}{|Du|}$ por $\frac{Du}{\sqrt{\epsilon^2 + |Du|^2}}$, $\epsilon > 0$. Las soluciones de este modelo aproximado convergen, después de extraer una subsucesión, hacia una solución de (20) [7] (el estudio de los operadores diferenciales involucrados es el objeto de [4]). Para este análisis es necesario incorporar las condiciones de frontera en forma variacional.

4.3 Algunos experimentos

Ejemplos para imágenes 2D. Los experimentos que siguen muestran algunos resultados obtenidos usando el funcional (20) para el valor $p = 2$.

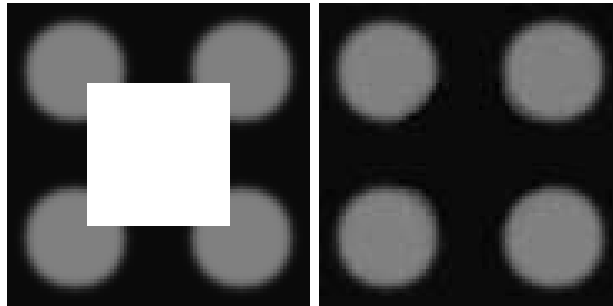


Figure 19: Cuatro discos y su reconstrucción.

La Figura 19 nos muestra el resultado de la interpolación de cuatro discos parcialmente cubiertos por un cuadrado. La Figura 20.a es un detalle de la boca de Lena con un *agujero* y la Figura 20.b nos

muestra el resultado obtenido usando (20). La figura 21.a nos muestra una imagen con un texto que queremos eliminar, la Figura 21.b nos muestra el correspondiente resultado.

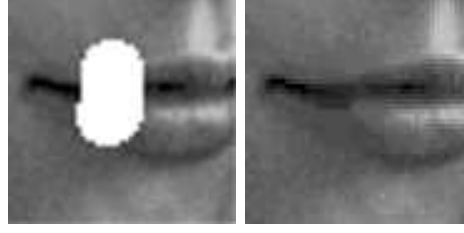


Figure 20: a) Detalle de la boca de Lena. b) Reconstrucción obtenida usando (20).



Figure 21: Imagen con texto y su reconstrucción.

Reconstrucción de agujeros en superficies. Podemos usar el funcional (20) para reconstruir un *agujero* en una superficie \mathcal{S} de \mathbb{R}^3 . Supongamos, para fijar ideas, que \mathcal{S} es una superficie compacta y conexa y \mathcal{M} es una parte desconocida de \mathcal{S} que no pudo ser obtenida al escanearla. De manera que en lo sucesivo identificaremos \mathcal{S} con su parte conocida. Consideremos un paralelepípedo Q en \mathbb{R}^3 conteniendo en su interior al *agujero* \mathcal{M} y a una parte de \mathcal{S} (see Figs. 23.a, 23.b). Denotemos por $\partial\mathcal{M}$ a la frontera del *agujero* (una curva o un conjunto de curvas en \mathbb{R}^3). Aunque \mathcal{M} es desconocida, sí conocemos su frontera relativa en \mathcal{S} . Consideremos en entorno \mathcal{F} de $\mathcal{S} \cap Q$ definido por

$$\mathcal{F} = \{x \in Q : d(x, \mathcal{S} \cap Q) < \alpha d(x, \partial\mathcal{M})\}, \quad \alpha > 0,$$

donde $d(x, \mathcal{S} \cap Q)$ denota la distancia de x a $\mathcal{S} \cap Q$. Supondremos que $\mathcal{F} \setminus (\mathcal{S} \cap Q)$ esta formado por dos componentes conexas que podemos identificar como vecinas a las dos caras de \mathcal{S} (ver Figura 22). Denotemos por $d^{\mathcal{F}}(x)$ a la distancia de un punto $x \in \mathcal{F}$ a $\mathcal{S} \cap Q$. Cambiando el signo de $d^{\mathcal{F}}$ en una de las caras de la superficie podemos definir la función distancia signada a $\mathcal{S} \cap Q$ en \mathcal{F} , que denotaremos por $d_s^{\mathcal{F}}(x)$. El campo de vectores dado por

$$N(x) = \nabla d_s(x) \quad \text{en } \mathcal{F}$$

es una extensión del campo de vectores formado por las normales unitarias a $\mathcal{S} \cap Q$ a su entorno \mathcal{F} . Consideraremos esta información como fiable y la usaremos como restricción.

Para adaptar el funcional (20) a la reconstrucción de *agujeros* de superficies debemos explicitar el *agujero* $\tilde{\Omega}$ y las funciones (u_0, θ_0) en una banda alrededor de $\tilde{\Omega}$. Para ello consideramos una bola \mathcal{B} (u otro abierto homeomorfo a una bola) tal que $\overline{\mathcal{B}} \subset\subset Q$ conteniendo a $\partial\mathcal{M}$ en su interior. Definimos el *agujero* $\tilde{\Omega}$ eliminando de \mathcal{B} los puntos de \mathcal{F} . Definimos la banda $B = Q \setminus \overline{\tilde{\Omega}}$.

Por otra parte, definimos $u_0 : Q \setminus \overline{\tilde{\Omega}} \rightarrow R$ como una función característica. Los valores $u_0(x) = 1$ y $u_0(x) = 0$ representan el interior y el exterior de \mathcal{S} , respectivamente. Inicialmente, indexamos las dos componentes conexas de $\mathcal{F} \setminus (\mathcal{S} \cap Q)$ por los valores $u_0 = 1$ (interior) y $u_0 = 0$ (exterior). Propagando estos valores al resto de $Q \setminus \overline{\tilde{\Omega}}$ obtenemos u_0 en B . Podemos proceder a esta indexación de forma consistente: no podemos conectar dos puntos con valores diferentes sin cruzar \mathcal{S} (véase la Figura 22).

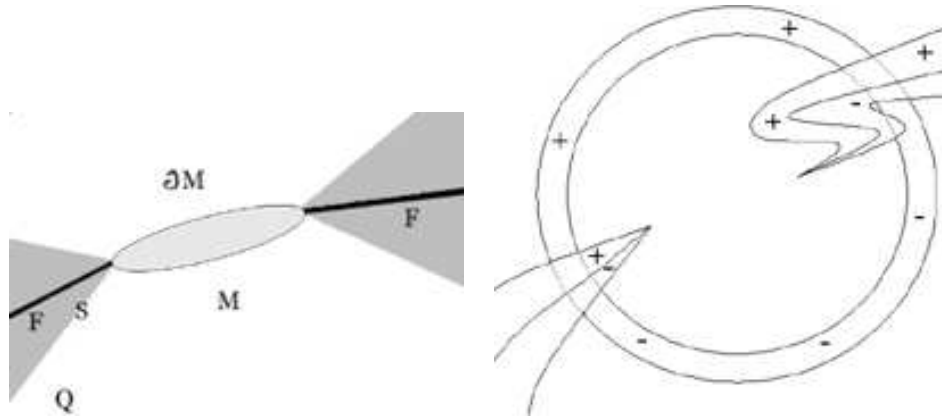


Figure 22: a) Una sección de la superficie S con el agujero M y el entorno F . b) Asignación del signo en las dos caras de S .

Llamamos A al conjunto de puntos x de $Q \setminus \overline{\Omega}$ tales que $u_0(x) = 1$, de manera que $u_0(x) = \chi_A(x)$. Minimizando (20) podemos reconstruir el conjunto A dentro de $\overline{\Omega}$. Definimos el campo $\theta_0 = \nabla d_s$ en F , y lo extendemos al resto de $Q \setminus \Omega$ extendiendo primero d_s a $Q \setminus \overline{\Omega}$ y tomando $\theta_0 = \nabla d_s$. Observemos que $Du_0 = \nu \mathcal{H}^2|_{S \cap Q}$ donde ν es la normal unitaria exterior a S y $\mathcal{H}^2|_{S \cap Q}$ representa la medida de Hausdorff en $S \cap Q$. Por consiguiente, se cumple que $\theta_0 \cdot Du_0 = |Du_0|$. Minimizamos (20) usando $u = u_0$ y $\theta = \theta_0$ en B como restricción. La aproximación numérica usada es la descrita en [6, 7].

Las Figuras 23.a, 23.b nos muestran dos agujeros en una versión digitalizada del David de Miguel Ángel (imágenes cortesía de Marc Levoy y el Michelangelo Project [36])). Hemos representado en estas imágenes el paralelepípedo Q conteniendo el agujero. Las Figuras 23.c y 23.d representan la superficie (triangulada) alrededor del agujero. Las Figuras 23.e y 23.f representan la reconstrucción obtenida minimizando el funcional (20). Para visualizar estas imágenes nos hemos servido del software AMIRA Visualization and Modeling System [3].

5 Comentarios sobre el método de Efros-Leung

El algoritmo anterior puede usarse para imágenes en secuencias de video como si fuesen datos 3D. Experimentos de este tipo pueden verse en [7]. Sin embargo, los datos de vídeo no pueden considerarse datos 3D geométricos y los experimentos realizados en [7] para secuencias de imágenes no son convincentes. En ello influye la baja tasa de muestreo temporal para garantizar la suavidad de las superficies de nivel. Por otro lado, los cambios de iluminación y el movimiento de los objetos provoca distorsiones en los mismos o la aparición o desaparición de objetos de la escena. Sin entrar en mucho detalle, podemos aprovechar el algoritmo propuesto por Efros-Leung para restaurar algunos agujeros en secuencias de vídeo (por ejemplo rayas en películas). Este algoritmo aprovecha la redundancia de la imagen: busca en la imagen un entorno similar a la parte conocida del entorno del punto p que se quiere reconstruir y copia dicha información en p . Si hay varios entornos igualmente semejantes, determina cuál de ellos copiar de forma aleatoria. Una mejora notable de este algoritmo ha sido propuesta en [21]: se calcula una prioridad para determinar el orden en que se van a considerar los píxeles del agujero. Esta prioridad depende del modulo y orientación del gradiente en cada píxel.

La Figura 24 nos muestra un ejemplo de reconstrucción de un agujero en una imagen de una secuencia de vídeo calculada usando el algoritmo de Efros-Leung habiendo compensado el movimiento de la secuencia. Las imágenes a) y b) nos muestran las imágenes 74 y 76 de una secuencia estándar llamada Foreman. La imagen en c) nos muestra la imagen 75 de la secuencia con un agujero tapando los ojos. La imagen d) nos muestra la reconstrucción. La información recuperada en la imagen 75 se ha buscado en las dos imágenes consecutivas 74 y 76.

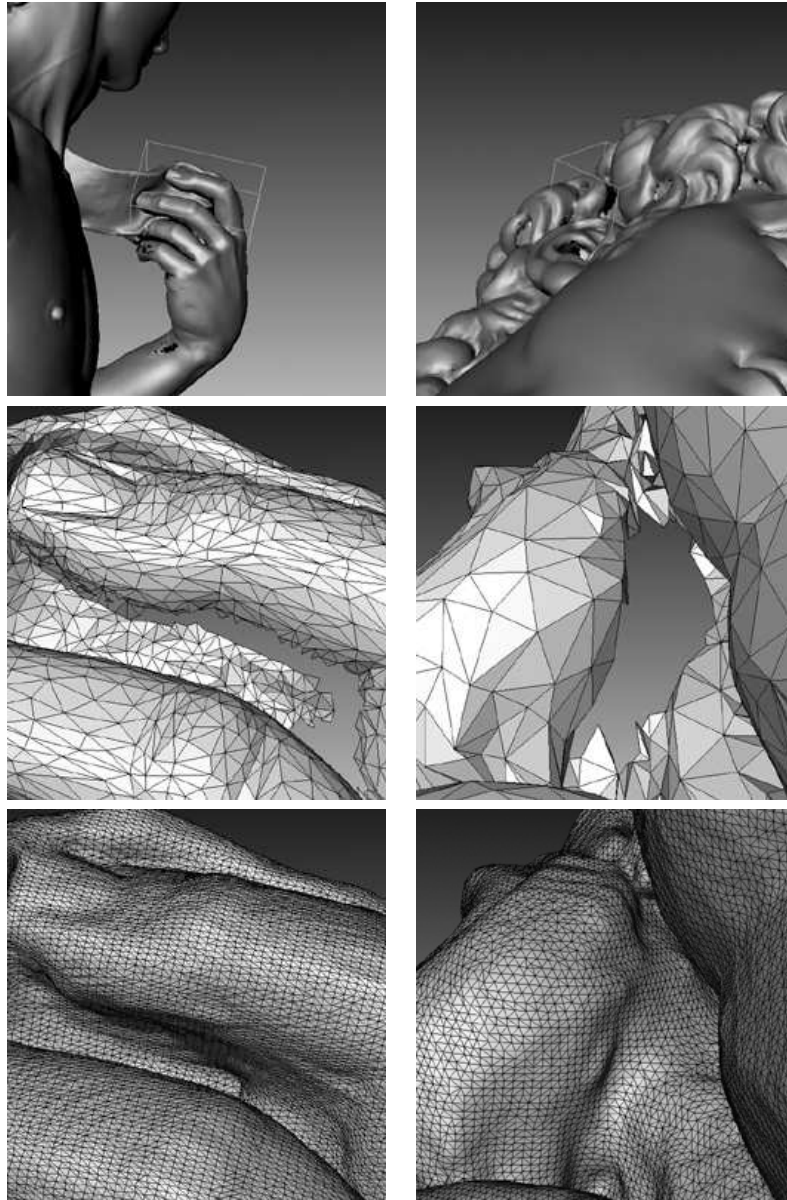


Figure 23: De arriba a abajo y de izquierda a derecha: a) La mano izquierda del David de Miguel Ángel con la caja Q . b) Parte de la cabeza y del pelo mostrando la caja Q . c) Un zoom de un detalle de a) mostrando un *agujero* sobre su superficie. d) Un zoom de un detalle de b) mostrando un *agujero* en el pelo. e) Reconstrucción del *agujero* de c). f) Reconstrucción del *agujero* de d).

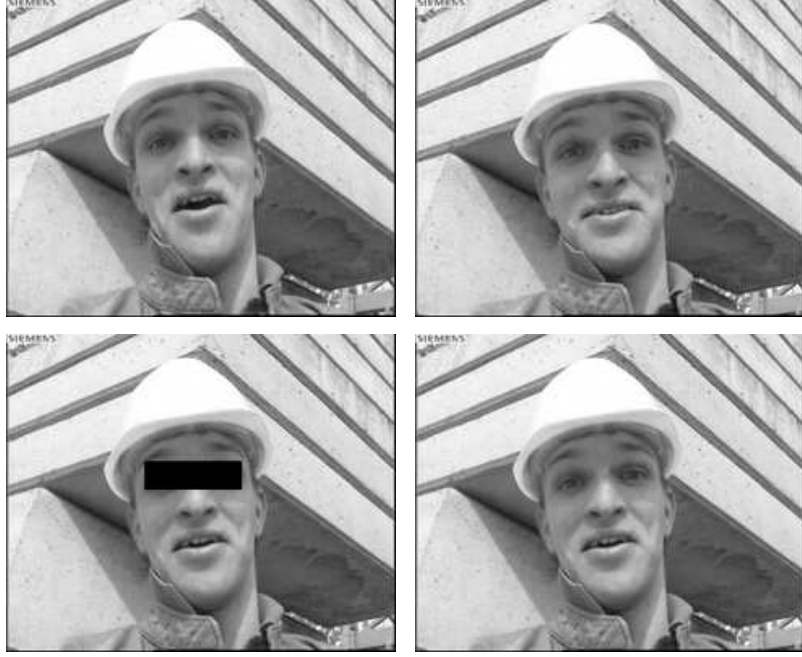


Figure 24: Arriba: a) y b) Las imágenes 74 y 76 de la secuencia Foreman. Abajo: c) La imagen 75 de la secuencia Foreman con una *agujero* en los ojos. d) Reconstrucción obtenida usando el algoritmo de Efron-Leung con compensación de movimiento.

6 Fusión de imágenes multi-espectrales

Como en el resto de este trabajo, $D = [0, 1]^2$. Mientras que una imagen de intensidad puede representarse por una función $u : D \rightarrow \mathbb{R}$, una imagen multispectral se representa por una función vectorial $\vec{u} : D \rightarrow \mathbb{R}^m$ donde m representa el número de canales espectrales. En el caso de imágenes en color $m = 3$ y $\vec{u} = (R, G, B)$ donde R, G, B representan los canales rojo, verde y azul, respectivamente. Si a ellos añadimos el canal infrarrojo, como ocurre en algunos satélites, entonces $m = 4$. Mientras cada canal representa la energía de la luz en una banda del espectro, la suma ponderada de todas ellas produce la imagen de intensidad o imagen pancromática. Para fijar ideas, supondremos que $m = 3$ y denotaremos $\vec{u} = (X_1, X_2, X_3)$ para representar a los canales (R, G, B) . En este caso,

$$u(x) = \alpha_1 X_1(x) + \alpha_2 X_2 + \alpha_3 X_3, \quad (23)$$

siendo $\alpha_1, \alpha_2, \alpha_3 \geq 0$, $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

En las imágenes digitales, la única información disponible es una versión muestreada y cuantizada de u , $u(i, j)$, donde (i, j) recorre un conjunto discreto de puntos (habitualmente sobre una retícula) y $u(i, j)$ recorre un conjunto discreto de valores que en muchos casos se corresponden con los valores $0, 1, \dots, 255$. Suponiendo que las muestras han sido tomadas respetando el teorema de muestreo de Shannon, entonces podemos recuperar la función $u(x)$, $x \in \mathbb{R}^2$, a partir de las muestras $u(i, j)$. En el caso de imágenes satélite, esto es razonable como aproximación y podemos suponer que $u(x)$ es interpolable en D a partir de las muestras.

Típicamente (como ocurre en el caso de los satélites SPOT) la imagen pancromática es muestreada a más alta resolución que los canales espectrales y se plantea el problema de recuperar dichos canales a la resolución de la imagen pancromática. Supondremos que la imagen $u(i, j)$ tiene un tamaño de $N \times N$ píxeles mientras los canales espectrales tienen una resolución s veces más baja y, por tanto, un tamaño de $\frac{N}{s} \times \frac{N}{s}$ donde $s = 2$ o $s = 4$. Denotaremos a los canales espectrales a baja resolución por el vector $\vec{u}^s = (X_1^s, X_2^s, X_3^s)$. El problema consiste en recuperar \vec{u} a partir de u y de \vec{u}^s teniendo en cuenta la restricción (23) y las restricciones dadas por el modelo de adquisición de imágenes. En nuestro caso, para cada canal, el valor X_i^s en un píxel a baja resolución está relacionado con el valor

de X_i en los píxeles a alta resolución por un filtro pasa-bajo y un submuestreo. Si k_1, k_2, k_3 denotan las respuestas impulsionales de los filtros correspondientes a los canales X_1, X_2, X_3 , supondremos que k_i son operadores lineales de $L^2(D)$ en $C(D)$ y se cumple que

$$X_n^s(i, j) = k_n * X_n(i, j), \quad \forall (i, j) \in S, \quad n = 1, 2, 3. \quad (24)$$

Obviamente, la recuperación de X_n , $n = 1, 2, 3$, a partir de u y de X_n^s , $n = 1, 2, 3$, satisfaciendo las relaciones (23) y (24) es un problema mal puesto. Para resolverlo seguiremos la estrategia habitual: regularización. En nuestro caso, propondremos una regularización basada en la información geométrica contenida en la imagen pancromática.

Hemos definido la geometría de una imagen de intensidades en términos de su mapa topográfico. ¿Cuál es el contenido geométrico de una imagen en color? Obviamente, la respuesta a esta pregunta es compleja y, estrictamente hablando, la geometría de una imagen en color no se puede reducir a la geometría de su luminancia. De hecho es fácil construir un contraejemplo en el que podemos ver objetos en color en una imagen de luminancia constante. Pero, ¿qué ocurre en las imágenes naturales? ¿Se crean contornos en imágenes en color con luminancia constante? En el trabajo [16] adoptamos la hipótesis (**GC**) de que, esencialmente, *el contenido geométrico de una imagen en color está contenido en el mapa topográfico de su correspondiente imagen de intensidad* y la sometimos a verificación experimental. Para ello, modificamos los canales de color de forma que sus curvas de nivel estuviesen contenidas en las curvas de nivel de su imagen de intensidad y procedimos a su comparación con la imagen original. Los resultados parecen confirmar nuestra hipótesis y podemos resumir brevemente la conclusión de [16]: *la sustitución de los colores de una imagen por sus esperanzas condicionales respecto a los niveles de gris no altera nuestra percepción de la imagen en color*. Nada hay de sorprendente en esto, como es experiencia común, las fotografías o el cine en blanco y negro permiten convencernos de que las formas contenidas en las imágenes pueden codificarse en la escala de niveles de gris, experiencia corroborada por una gran parte de la investigación en procesamiento de imágenes.

Admitimos en lo que sigue que la hipótesis (*GC*) es razonable para los canales *RGB* en imágenes satélite. De acuerdo con la Sección 2.2, la geometría de u está dada por sus curvas de nivel. Pero necesitamos una herramienta más adaptada al diseño de un funcional de energía y para ello vamos a describir la geometría de u a partir del campo de normales unitarias a sus curvas de nivel. Suponiendo que $u \in BV(D)$, definimos el campo θ como la derivada de Radon-Nykodym de la medida Du respecto a su variación $|Du|$, de manera que, formalmente, se satisface la relación $\theta \cdot Du = |Du|$, y $|\theta| \leq 1$ casi por todas partes. Llamaremos a θ el campo de normales unitarias a los conjuntos de nivel de u . En la práctica, a nivel computacional, definimos θ por la relación $\theta(x) = \frac{Du(x)}{|Du(x)|}$ si $Du(x) \neq 0$, y $\theta(x) = 0$ si $Du(x) = 0$, $x \in D$.

Suponiendo que $X_1, X_2, X_3 \in BV(D)$, para imponer que la geometría de los canales de color está determinada por θ impondremos las relaciones

$$|DX_n| = \theta \cdot DX_n, \quad n = 1, 2, 3, \quad (25)$$

o bien, de otra forma, impondremos

$$\theta^\perp \cdot DX_n = 0, \quad n = 1, 2, 3, \quad (26)$$

donde θ^\perp representa la rotación en el sentido anti-horario de $\pi/2$ de θ , es decir, el campo de tangentes a las curvas de nivel de u . Para fijar ideas, impondremos (26).

Obviamente, tanto las relaciones (26) como las restricciones (23) y (24) pueden no ser satisfechas de forma exacta y es necesario imponerlas en un contexto variacional. Teniendo en cuenta todas estas restricciones, proponemos minimizar el siguiente funcional:

$$\sum_{n=1}^3 \gamma_n \int_{\Omega} |\theta^\perp \cdot DX_n|^2 dx + \lambda \int_{\Omega} \left(\sum_{n=1}^3 \alpha_n X_n - u \right)^2 dx + \mu \sum_{n=1}^3 \int_{\Omega} \Pi_S \left((k_n * X_n(x) - X_n^S(x))^2 \right) dx \quad (27)$$

con las restricciones $0 \leq X_n \leq M_n := \max_{(i,j) \in S} \max \left(\frac{u(i,j)}{\alpha_n}, X_n^S(i,j) \right)$, siendo $\gamma_n, \lambda, \mu > 0$, $n = 1, 2, 3$. Observemos que Π_S representa el peine de Dirac sobre la retícula de puntos S , es decir, $\Pi_S = \sum_{(i,j) \in S} \delta_{(i,j)}$. Para más precisiones sobre el espacio funcional en que minimizamos (27) nos referimos a [9]. Digamos para finalizar que el efecto del primer término del funcional es la difusión de los canales de color a lo largo de las curvas de nivel de la intensidad. Un descenso de gradiente aplicado a este término lleva a la ecuación de difusión

$$\frac{\partial X}{\partial t} = \operatorname{div} \left(\langle \theta^\perp, DX \rangle \theta^\perp \right) \quad (28)$$

para cada canal $X = X_1, X_2, X_3$ ([9]).

La Figura 25.a nos muestra el resultado obtenido a partir de la imagen pancromática u y los canales de color \vec{u} de la Figura 6 cuya resolución es de 0.7 y 2.8 m/píxel, respectivamente. La Figura 25.b nos muestra una imagen de referencia con la que podemos comparar visualmente. Para una comparación cuantitativa referimos al lector a [9].



Figure 25: a) Resultado obtenido a partir de las imágenes u y \vec{u} de la Figura 6. b) Imagen de referencia para comparar el resultado. $(R_{ref}, G_{ref}, B_{ref})$.

Agradecimientos: Quisiera agradecer a C. Ballester, M. Bertalmío, F. Andreu y J.M. Mazón sus comentarios en la preparación de este trabajo. Este trabajo ha sido parcialmente financiado por el Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya, por el proyecto PNP GC, referencia BFM2003-02125 y por el CNES.

References

- [1] L. Ambrosio, V. Caselles, S. Masnou and J.M. Morel, *Connected Components of Sets of Finite Perimeter and Applications to Image Processing*, Journal of the Euro. Math. Soc., 3, 39-92 (2001).
- [2] L. Ambrosio and S. Masnou. A Direct Variational Approach to a Problem Arising in Image Reconstruction. *Interfaces and Free Boundaries*, 5:63–81, 2003.
- [3] Amira. Amira Visualization and Modeling System. <http://www.AmiraVis.com>, 2004.
- [4] F. Andreu, V. Caselles, and J.M. Mazón, *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*. Progress in Mathematics, vol. 223, 2004. Birkhauser.

- [5] G. Aronsson, *Extension of functions satisfying Lipschitz conditions*, Ark. for Math. **6**, (1967) 551–561.
- [6] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by Joint Interpolation of Vector Fields and Grey Levels. *IEEE Transactions on Image Processing*, 10:1200–1211, 2001.
- [7] C. Ballester, V. Caselles, and J. Verdera. Disocclusion by Joint Interpolation of Vector Fields and Gray Levels. *Multiscale Modeling and Simulation*, 2:80–123, 2003.
- [8] C. Ballester, V. Caselles, J. Verdera, M. Bertalmio, and G. Sapiro. A Variational Model for Filling-in Gray level and Color Images. In *ICCV 2001*, vol. 1, pp. 10–16, Vancouver, Canada.
- [9] C. Ballester, V. Caselles, B. Rougé, and J. Verdera: Une méthode géométrique de fusion des images P+XS. Bulletin de la Société Française de Photogrammetrie et Teledetection, Num. 169, pp. 53-63, 2003. Una versión ulterior aparecerá en IJCV.
- [10] C. Ballester, E. Cubero-Castan, M. Gonzalez and J.M. Morel, *Contrast Invariant Image Intersection*, Cahiers de l'ENS Cachan, 9817, 1998.
- [11] C. Ballester, V. Caselles and P. Monasse, *The tree of shapes of an image* ESAIM: Control, Optimisation and Calculus of Variations, 9, pp. 1-18, 2003.
- [12] G. Bellettini, G. Dal Maso, and M. Paolini. Semicontinuity and Relaxation Properties of a Curvature Depending Functional in 2d. *Ann. Scuola Normale Sup. di Pisa, Cl. Sci.*, 20:247–297, 1993.
- [13] F. Cao, *Absolutely minimizing Lipschitz extension with discontinuous boundary data*, C.R. Acad. Sci. Paris **327**, (1998) 563-568.
- [14] J.R. Casas, *Image compression based on perceptual coding techniques*, PhD thesis, Dept. of Signal Theory and Communications, UPC, Barcelona, Spain, March 1996.
- [15] V. Caselles, B. Coll and J.M. Morel, *Topographic Maps and Local Contrast Changes in Natural Images*, International Journal of Computer Vision 33(1), 5-27 1999.
- [16] V. Caselles, B. Coll and J.M. Morel, *Geometry and Color in Natural Images*, Journal of Mathematical Imaging and Vision, vol. 16, pp. 89-107, 2002.
- [17] V. Caselles, L. Igual and O. Sander, *An Axiomatic Approach to Scalar Data Interpolation on Surfaces*, Preprint 2004.
- [18] V. Caselles, J.M. Morel and C. Sbert, *An axiomatic approach to image interpolation*, IEEE Transactions on Image Processing **7**, (1998) 376–386.
- [19] T. Chan, S. Kang, and J. Shen. Euler’s Elastica and Curvature Based Inpaintings. *SIAM J. Appl. Math.*, 63:564–592, 2002.
- [20] M. G. Crandall, H. Ishii and P. L. Lions, User’s guide to viscosity solutions of second order partial differential equations, *Bull. Am. Math. Soc.* 27 (1992) pp. 1-67.
- [21] A. Criminisi, P. Perez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE-TIP*, 13(9):1200–1212, 2004.
- [22] A. Efros and T. Leung. Texture synthesis by non-parametric sampling. In *IEEE International Conference on Computer Vision*, pages 1033–1038, 1999.
- [23] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Math., CRC Press, 1992.

- [24] W.R. Franklin and A. Said, *Lossy compression of elevation data*, 7th Int. Symposium on Spatial Data Handling, 1996.
- [25] F. Guichard and J.M. Morel, *Introduction to Partial Differential Equations on image processing*, Tutorial, ICIP-95, Washington.
- [26] A. Hirani and T. Totsuka. Combining Frequency and Spatial Domain Information for Fast Interactive Image Noise Removal. In *SIGGRAPH 96*, pages 269–276, 1996.
- [27] B. Horn. The Curve of Least Energy. *ACM Trans. on Mathematical Software*, 9:441–460, 1982.
- [28] R. Jensen, *Uniqueness of Lipschitz extensions: Minimizing the sup norm of the gradient*, Arch. Rat. Mech. Anal. **123**, (1993) 51–74.
- [29] J.L. Lisani, *Comparaison automatique d’images par leurs formes*, Ph.D Thesis, Université de Paris-Dauphine, July, 2001.
- [30] J.L. Lisani, P. Monasse, L. Moisan and J.M. Morel, *Affine Invariant Mathematical Morphology Applied to Generic Shape Recognition Algorithm*, Proceedings of the International Symposium on Mathematical Morphology 2000, Palo Alto, June 2000.
- [31] G. Kanizsa, *Gramática de la Visión*, Ed. Paidós, 1986.
- [32] D. King. *The Commissar Vanishes*. Henry Holt and Company, 1997.
- [33] A. Kokaram. On missing data treatment for degraded video and film archives: a survey and a new bayesian approach. *IEEE Transactions on Image Processing*, 13:397–415, 2004.
- [34] A. Kokaram, R. Morris, W. Fitzgerald, and P. Rayner. Detection of Missing Data in Image Sequences. *IEEE Transactions on Image Processing*, 11:1496–1508, 1995.
- [35] A. Kokaram, R. Morris, W. Fitzgerald, and P. Rayner. Interpolation of Missing Data in Image Sequences. *IEEE Transactions on Image Processing*, 11:1509–1519, 1995.
- [36] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The Digital Michelangelo Project: 3D Scanning of Large Statues. In *Computer Graphics (SIGGRAPH) 2000*, pages 269–276, 1996.
- [37] S. Masnou. *Filtrage et Desocclusion d’Images par Méthodes d’Ensembles de Niveau*. PhD thesis, Université Paris-Dauphine., 1998.
- [38] S. Masnou. Disocclusion: a Variational Approach using Level Lines. *IEEE-TIP*, 11:68–76, 2002.
- [39] S. Masnou and J. Morel. Level Lines Based Disocclusion. In *Proc. 5th IEEE International Conference on Image Processing*, pages 259–263, Chicago, Illinois, 1998.
- [40] P. Monasse, *Image Registration*, Cahiers de l’ENS Cachan, 1999.
- [41] D. Mumford. Elastica and Computer Vision. In C. Bajaj, editor, *Algebraic geometry and its applications*, pages 491–506. Springer-Verlag, 1994.
- [42] M. Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation, and Depth*. Springer-Verlag, Berlin, 1993.
- [43] L.I. Rudin, S. Osher and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*. Physica D, 60:259-269, 1992.
- [44] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, 1982.

- [45] P. Salembier, P. Brigger, J.R. Casas, M. Pardàs, *Morphological Operators for Image and Video Compression*, IEEE Trans. on Image Processing, 5 (1996), pp. 881-897.
- [46] K. Thornber and L. Williams. Characterizing the Distribution of Completion Shapes with Corners Using a Mixture of Random Processes. *Pattern Recognition*, 33:543–553, 2000.
- [47] S. Ullman. Filling-in the Gaps: the Shape of Subjective Contours and a Model for Their Generation. *Biological Cybernetics*, 75:1–6, 1976.
- [48] M. Unser, Splines, a perfect fir for signal and image processing. IEEE Signal Processing Magazine, pp. 22-38, 1999.
- [49] M. Wertheimer, *Untersuchungen zur Lehre der Gestalt, II* Psychol. Forschung, 4:301-350, 1923.
- [50] D. Williams, *Probability with Martingales*, Cambridge Mathematical Textbooks 1991.
- [51] L. Williams and D. Jacobs. Local Parallel Computation of Stochastic Completion Fields. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition CVPR '96*, pages 161–168, 1996.