Irregularly sampled image restoration and interpolation

Gabriele Facciolo Furlan

Tesi Doctoral UPF / 2011

Supervised by
Prof. Vicent Caselles Costa
Departament de Tecnologies de la Informació i les Comunicacions

Para el Pato

Agradecimientos

Quisiera, antes que nada, agradecer a mi director de tesis Vicent Caselles, por compartir sus conocimientos conmigo, por su apoyo, por sus enseñanzas, por su disponibilidad y comprensión. Por ello le estaré siempre agradecido.
Muchas gracias a todos los que han contribuido a la realización de este trabajo: Andrés Almansa, Pablo Arias, Fean-François Aujol, Enric Meinhardt, Álvaro Pardo, Guillermo Sapiro y Federico Lecumberry. Este documento es en gran medida resultado de su colaboración, y les estoy muy agradecido por eso.
Gracias a todos los compañeros del grupo: Coloma Ballester, Marcelo Bertalmío, Juan Cardelino, Sira Ferradans, Lluís Garrido, Gloria Haro, Laura Igual, Vanel Lazcano, Adrian Marques, Edoardo Provenzi y Rida Sadek, por tantas y enriquecedoras discusiones.
Agradezco el apoyo del programa ALFA (Computer Vision Foundations and Applications), y de la Agencia Espacial Francesa (CNES) a través del proyecto MISS (Mathematics for Space Stereo Images). Especialmente a Andrés Almansa, Jean-Michel Morel, Gregory Randall y Bernard Rougé, por brindarme la oportunidad de participar de estas actividades.
Y por supuesto, muchas gracias a los compañeros y amigos de la UPF y de Cachan, con ellos he compartido muy buenos momentos en estos años: Ralph Andreijack, Antoni Buades, Pablo Cancela, Oscar Civit, Mariella Dimiccoli, Vicenç Gonez, Rafael Grompone, Jeremie Jakubowicz, Andreas Kaltenbrunner, Dani Martí, Héctor Palacios, Neus Sabater, Catalina Tobón, Rodrigo Verschae y a todas aquellas personas que me han ayudado de alguna manera y me olvido de nombrar. Todos han contribuido a hacer de esta una experiencia muy valiosa para mi. Tampoco puedo omitir a mis amigos de Uruguay, a quienes dejé "casi abandonados" los ultimos 5 años, pero que siempre me han expresado su apoyo constante e incondicional.
Finalmente, pero no por eso menos importanes, gracias a mis padres Maria y Rafaello y mis hermanos Alessandro y Andrea, por apoyarme durante todos estos años. Y muy especialmente, gracias a mi esposa Laura Podestá, por su cariño, comprensión (y compresión), por estar siempre en los momentos difíciles y por acompañarme al otro lado del mundo.

Abstract

The generation of urban digital elevation models from satellite images using stereo reconstruction techniques poses several challenges due to its precision requirements. In this thesis we study three problems related to the reconstruction of urban models using stereo images in a low baseline disposition. They were motivated by the MISS project, launched by the CNES (*Centre National d'Etudes Spatiales*), in order to develop a low baseline acquisition model. The first problem is the restoration of irregularly sampled images and image fusion using a band limited interpolation model. A novel restoration algorithm is proposed, which incorporates the image formation model as a set of local constraints, and uses of a family of regularizers that allow to control the spectral behavior of the solution. Secondly, the problem of interpolating sparsely sampled images is addressed using a self-similarity prior. The related problem of image inpainting is also considered, and a novel framework for exemplar-based image inpainting is proposed. This framework is then extended to consider the interpolation of sparsely sampled images. The third problem is the regularization and interpolation of digital elevation models imposing geometric restrictions. The geometric restrictions come from a reference image. For this problem three different regularization models are studied: an anisotropic minimal surface regularizer, the anisotropic total variation and a new piecewise affine interpolation algorithm.

Resumen

La generación de modelos urbanos de elevación a partir de imágenes de satélite mediante técnicas de reconstrucción estereoscópica presenta varios retos debido a sus requisitos de precisión. En esta tesis se estudian tres problemas vinculados a la generación de estos modelos partiendo de pares estereoscópicos adquiridos por satélites en una configuración con baseline pequeño. Estos problemas fueron motivados por el proyecto MISS, lanzado por el CNES (*Centre National d'Etudes Spatiales*) con el objetivo de desarrollar las técnicas de reconstrucción para imágenes adquiridas con baseline pequeños. El primer problema es la restauración de imágenes muestreadas irregularmente y la fusión de imágenes usando un modelo de interpolación de banda limitada. Se propone un nuevo método de restauración, el cual usa una familia de regularizadores que permite controlar el decaimiento espectral de la solución e incorpora el modelo de formación de imagen como un conjunto de restricciones locales. El segundo problema es la interpolación de imágenes muestreadas en forma dispersa usando un prior de auto similitud, se considera también el problema relacionado de inpainting de imágenes. Se propone un nuevo framework para inpainting basado en ejemplares, el cual luego es extendido a la interpolación de imágenes muestreadas en forma dispersa. El tercer problema es la regularización e interpolación de modelos digitales de elevación imponiendo restricciones geométricas las cuales se extraen de una imagen de referencia. Para este problema se estudian tres modelos de regularización: un regularizador anisótropo de superficie mínima, la variación total anisótropa y un nuevo algoritmo de interpolación afín a trozos.

Contents

Chapter 1

Overview

The context. The demand for high resolution remote sensing products on urban areas has experimented a rapid growth in the last years. In the case of urban Digital Elevation Models (DEM) and high resolution imagery for civil and defense applications, not only has grown the precision exigency but also the amount of data, making impractical any non-extensive and non-automatic technique.

The main sources of information for obtaining digital elevation models are: cartography and cadastral registers, LIDAR scanning (light detection and ranging or laser imaging detection and ranging), or obtained from aerial images (airborne or satellite) by applying stereo reconstruction techniques.

In the case of stereoscopic images, the depth of the objects in a three dimensional scene is computed by comparing two projections taken from two slightly different positions. The distance between the two projection centers is called *baseline*, and the ratio between the baseline and the distance to the scene is the *stereoscopic coefficient*. The depth of a point is inversely proportional to its disparity (the distance between its projections in the two images). A higher disparity means that the object is near to the cameras. Occlusions are a common artifact of the stereoscopic reconstruction, they correspond to portions of the images whose depth cannot be determined because they are only visible from one of the projections.

Since self-occlusions are common in urban imagery, the best acquisition direction to avoid them is exactly from the zenith (this is also true for the LIDAR scanning). On the other hand, the accuracy of the measured disparity (and the depth of the object) depends on the resolution of the images. Thus, to obtain the maximal precision it is necessary to increase the stereoscopic coefficient. However a big stereoscopic coefficient means more occlusions.

Interestingly, as it is shown in [Del04, SMA10] the precision of a stereo correlation algorithm does not depend on the stereoscopic coefficient, thus, allowing to compute sub-pixel disparities with small baseline stereo images. The small baseline assures that there are almost no occlusions, and, being the distance between the projections is very small, the images obtained are also quasi-simultaneous. The drawback of stereo correlation is that not all image points can

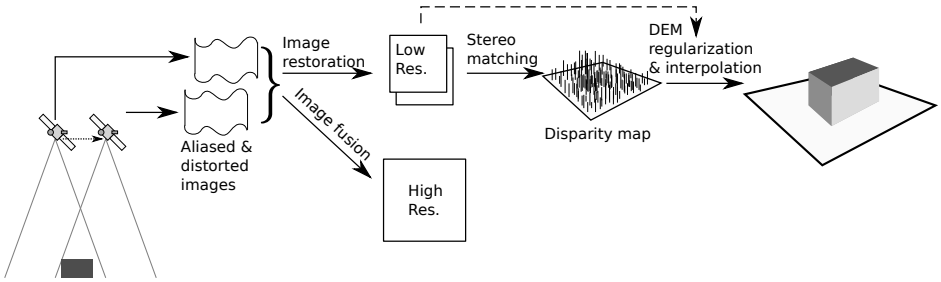


Figure 1.1: Digital elevation model acquisition using stereo reconstruction. The schema represents the processes from the stereo pair to the interpolated and regularized elevation model, through the restoration and stereo matching. The initial stereo pair can also be fused into a higher resolution image.

be faithfully matched, and its depth accurately computed, the remaining points must be interpolated. In that context the challenge is to interpolate sparse data so that the model fits the geometry of the scene.

On the other hand the stereo pair must be restored before applying any correlation algorithm. Most of the airborne imaging sensors are push-broom scanners. Instead of using a rectangular sensor array, a push-broom sensor acquires the image progressively by a line of sensors arranged perpendicular to the direction of motion. Because of satellite vibrations the resulting images are non regularly sampled. So, a regularly sampled pair must be restored from an irregularly sampled one.

A second issue that must be solved before applying a stereo algorithm is the alias. To enable super-resolution applications, by design the sensors produce aliased images. However, stereo correlation algorithms are very sensitive to the presence of alias in the stereo pair, so the alias must be removed before computing the disparity [Van06].

Finally, let us mention that since the stereo pair is quasi-simultaneous and the images are aliased, both images can be merged into a higher resolution one. In this case the samples of both images must be fused into a single image.

The generation of urban digital elevation models and fused images from satellite stereo images is schematically summarized in Figure 1.1.

Organization of the thesis. This thesis is organized in three independent parts, each part has its own introduction where we describe and motivate the succeeding chapters.

- The first part is devoted to the restoration of irregularly sampled images. We propose a variational method based on a band limited interpolation model, and we study the problem of image fusion (super-resolution) as a restoration from irregular samples.

- In the second part we explore the use of self-similarity as a prior for interpolating sparsely sampled images (we restrict us to the case where the samples are at integer positions). We start by introducing a new framework for exemplar based inpainting which we subsequently adapt to the interpolation of sparsely sampled images.

- The main subject of the third part is the use of anisotropic regularization applied to the interpolation and regularization of digital elevation models.

Although the three parts cover different topics, they can be understood within the remote sensing context. The first part is motivated by the problems of satellite image restoration and fusion of several images into a higher resolution one. For both of them we use the band limited interpolation model as a prior. In the second part we use the self-similarity prior for interpolating sparsely sampled images. This permits to increase the resolution of the images by adding features that are not in the data but are copied from other portions of the same image (or eventually also from other images). The third part is devoted to the interpolation of digital elevation models obtained from aerial images using a geometric regularization prior. Indeed, we consider three anisotropic regularization algorithms for interpolating elevation models which enforce different smoothness priors: minimal surface, bounded variation and piecewise affine. The anisotropy embeds the geometric structure extracted from a reference image.

Let us give a detailed description of the contents of each chapter.

Part I. Image restoration for irregularly sampled images

Chapter 2: We introduce the image acquisition model for irregularly sampled images, and justify the regularization approach taken in Chapter 3 to solve the restoration problem. Then, we extend the discussion to image fusion and we point out the limitations of the total variation regularizer when used in an underdetermined restoration problem. We introduce the "spectral projector" which will be used in Chapters [FAAC09] and 4.

Chapter 3: We propose a restoration algorithm for band limited images that considers irregular sampling, denoising, and deconvolution. We explore the application of a family of regularizers that allow us to control the spectral behavior of the solution and we incorporate the image acquisition model as set of local constraints.

Chapter 4: We extend the band limited restoration model to the the problem of image fusion. Considering some particular types of sampling we propose two methods for estimating the "spectral projector" for image fusion. Then we show that the restoration of aliased images with a spectral projector induces a bias in the noise estimation. This bias is difficult to estimate since it depends on the statistics of each image, and is still an open problem.

Part II. Interpolation from sparse samples

Chapter 5: In this chapter we introduce the problem of reconstructing an image from a set of sparse samples taken on a regular grid. As opposed to the interpolation methods studied in the previous part, we adopt here a non-local prior for imposing the regularity on the solution. The proposed exemplar-based framework is applied to inpainting (Chapter 6) and interpolation problems (Chapter 7).

Chapter 6: In this chapter we propose and analyze a variational framework for exemplar based inpainting. From this framework we derive and test three inpainting algorithms.

Chapter 7: In this chapter we extend the framework proposed in Chapter 6 to consider the interpolation of sparsely sampled images.

Part III. Digital elevation model interpolation

Chapter 8: We motivate the problem of digital elevation model interpolation from sparse or incomplete data in the context where a reference image of the scene is available. In this setting the information provided by the reference image can be incorporated into the regularization by means of different smoothness priors.

Chapter 9: In this chapter we consider an anisotropic minimal surface regularizer. The anisotropy, derived from the reference image, constrains the regularization of the model so that diffusion only occurs along the isophotes of the image. In this context we considered two data terms based on L^2- and L^1-norms.

Chapter 10: In this chapter we use an anisotropic total variation as regularizer for the interpolation of elevation data, and we explore the connections of the underlying functional with the geodesic active contour segmentation model.

Chapter 11: Finally, we propose a piecewise interpolation algorithm for sparse data that incorporates geometric information from a reference image by means of a geodesic distance defined on the image domain. This geodesic distance allows to compute a Voronoi diagram that respect the edges of the reference image, and permits to determine a set of neighboring samples for each point of the image. The neighborhoods are later used to robustly fit a piecewise affine model for each cell of the diagram. The resulting piecewise interpolation of the data adapts to the geometric information provided by the reference image.

Chapter 12: This chapter summarizes the conclusions of this thesis and a brief discussion about the future directions.

Part I

Image restoration for irregularly sampled images

Chapter 2

Image restoration and irregular sampling

In this part we address the problem of image restoration from irregular samples, in particular we will focus on the restoration of satellite images, where the micro-vibrations are responsible of the type of distortions we are considering here. In this chapter we provide some background about the formulation of a restoration problem as an energy minimization problem and its relation with the Bayesian methods.

2.1 Introduction

The general model of an image acquisition system may be written as:

$$z(\xi_i) = (h * u)(\xi_i) + n_{\xi_i}, \qquad \xi_i \in \Xi, \tag{2.1}$$

where $\Xi = \{\xi_i\}_{i=1}^{N^2} \subseteq \mathbb{R}^2$ is a known and finite set of regular or irregular samples, u is the ideal undistorted image, h is a blurring kernel and n is the usual Gaussian noise.

We consider the problem of reconstructing the signal u from the set of observations $z(\xi_k)$ knowing the sampling geometry Ξ, the blurring kernel h, and the statistics of the noise n. The problem of recovering u from z arise in several real-world image acquisition systems. Let us overview three different problems of interest that arise in practice:

- image restoration under regular or irregular sampling,
- image denoising,
- image fusion posed as a restoration problem from irregular samples.

2.2 The problem of image restoration

For simplicity, let us consider first the restoration problem in the case of regular sampling. The case of irregular sampling will be discussed in Section 2.2.1.

The problem of recovering u from z is ill-posed. Notice that even assuming zero noise and an invertible kernel h, the sampling rate of Ξ may be insufficient to recover the desired details of u. Several methods have been proposed to recover a plausible solution u, most of them impose directly or indirectly some kind of regularity prior on the solution.

The Bayesian approach to image restoration Let $\Omega = \{1, ..., N\}^2$ be a discrete domain. We denote by X the Euclidean space $\mathbb{R}^{N\times N}$. The Euclidean norm in X will be denoted by $\|\cdot\|_X$. An image is represented by a vector $(u(x))_{x\in\Omega}$ (we may also use the notation $u(i,j)$ to denote a pixel of u).

The ingredients of the Bayesian approach for solving the restoration problems are:

- The image formation model:

$$z(x) = h * u(x) + n(x) \qquad x \in \Omega, \tag{2.2}$$

  where h is a discrete convolution kernel and $n(x)$ is a white Gaussian noise with zero mean and standard deviation σ.

- An a priori probability density for "perfect" original signals:

$$p_\beta(u) \propto \exp\left(-\beta J_d(u)\right), \tag{2.3}$$

  where $\beta > 0$, and $J_d(u)$ is an energy density.

Using Bayes rule, the posterior density of u given z is

$$p_\beta(u|z) = \frac{1}{C(z)} p(z|u) p_\beta(u), \tag{2.4}$$

where $C(z)$ is the normalization constant making the mass of $p_\beta(u|z)$ to be 1. The conditional distribution of z given u would be a delta at the function $h * u$ if there were no noise. On the other hand, since the noise statistic is known and $n = z - h * u$, then the probability density of z given u is Gaussian:

$$p(z|u) \propto \exp\left(-\frac{\|z - h * u\|_X^2}{2\sigma^2}\right),$$

and the posterior probability becomes

$$p_\beta(u|z) = \frac{1}{C} \exp\left(-\left(\frac{\|z - h * u\|_X^2}{2\sigma^2} + \beta J_d(u)\right)\right). \tag{2.5}$$

The maximization of the a posteriori density (2.5) or (MAP) is equivalent to the minimization problem (2.6) provided that $\beta\sigma^2 = \lambda$,

$$\min_{u\in X} \quad J_d(u) + \frac{1}{2\lambda}\|h * u - z\|_X^2. \tag{2.6}$$

The MAP estimator is not the only, nor necessarily the best, choice for computing u. Indeed the image with the highest probability, given by the resolution of MAP, might be a very rare event. The Conditional Means for instance, provides an estimation of u that consists in computing the expected value of u given z:

$$E(u|f) = \frac{1}{C}\int_{\mathbb{R}^{N^2}} u p_\beta(u|z)\, du = \frac{1}{C}\int_{\mathbb{R}^{N^2}} u \exp\left(-\frac{\|z - h*u\|_X^2}{2\sigma^2} - \beta J_d(u)\right) du. \tag{2.7}$$

The Conditional Means estimate requires to compute an integral in a high dimensional space. Lassas-Siltanen [LS04] and Louchet-Moisan [LM08] propose to approximate this integral with a Markov Chain Monte-Carlo algorithm (MCMC). However due to the high computational cost of these techniques, it seems that these approaches are yet not efficient enough for complex restoration problems.

We notice that the MAP strategy generalizes to the continuous setting [Cha04], while the Conditional Means does not [LS04].

Regularization approaches In what follows we adopt the MAP strategy. The MAP formulation provides a justification of the usual regularization approach to solve ill-conditioned problems [SGG+08]. In this setting the selection of the regularizer amounts to impose different priors over the reconstructed signal.

Assuming a linear regularization, the solution of (2.2) is estimated by minimizing a functional of the form

$$J_\lambda(u) = \| Hu - z \|_X^2 + \lambda \| Qu \|_X^2, \tag{2.8}$$

which yields the estimate

$$u_\lambda = (H^t H + \lambda Q^t Q)^{-1} H^t z, \tag{2.9}$$

where $Hu = h * u$, and Q is a regularization operator. Observe that to obtain u_λ we have to solve a system of linear equations. The role of Q is, on one hand, to move the small eigenvalues of H away from zero while leaving the large eigenvalues unchanged, and, on the other hand, to incorporate the a priori (smoothness) knowledge that we have on u.

If we treat u and n as random vectors and we select $\lambda = 1$ and $Q = R_s^{-1/2} R_n^{1/2}$ with R_s and R_n the image and noise covariance matrices, then (2.9) corresponds to the Wiener filter that minimizes the mean square error between the original and restored images.

One of the first regularization methods consisted in choosing between all possible solutions of (2.2) the one which minimized the discrete Sobolev (semi) norm of u $\sum_{x\in\Omega} |Du(x)|^2$ which corresponds to $Qu = Du$. In the continuous case it would be the usual Sobolev seminorm $\int_{\mathbb{R}^2} |Du|^2 dx$. Notice that for this choice of Q the solution of (2.8) given by (2.9) can be expressed as

$$\hat{u}(\omega) = \frac{\overline{\hat{h}(\omega)}}{|\hat{h}(\omega)|^2 + 4\lambda|\omega|^2\pi^2/N^2}\hat{z}(\omega).$$

From the above formula we see that high frequencies of z (which are mostly attributed to the noise) are attenuated by the smoothness constraint.

This formulation was an important step, but the results were not satisfactory, mainly due to the inability of the previous functional to resolve discontinuities (edges) and oscillatory textured patterns. The smoothness required by the finiteness of the Dirichlet integral constraint is too restrictive. Indeed, functions in $W^{1,2}(\mathbb{R}^2)$ (i.e., functions $u \in L^2(\mathbb{R}^2)$ such that $Du \in L^2(\mathbb{R}^2)$) cannot have discontinuities along rectifiable curves. These observations motivated the introduction of Total Variation in image restoration problems by L. Rudin, S. Osher and E. Fatemi in their work [ROF92]. The a priori hypothesis is that functions of bounded variation (the BV model) [AFP00] are a reasonable functional model for many problems in image processing, in particular, for restoration problems [ROF92]. Typically, functions of bounded variation have discontinuities along rectifiable curves, being continuous in some sense (in the measure theoretic sense) away from discontinuities [AFP00]. The discontinuities could be identified with edges. The ability of total variation regularization to recover edges is one of the main features which advocates for the use of this model but its ability to describe textures is less clear, even if some textures can be recovered, up to a certain scale of oscillation. An interesting experimental discussion of the adequacy of the BV-model to describe real images can be found in [GM01].

On the basis of the BV model, Rudin-Osher-Fatemi [ROF92] proposed to use the Total Variation as regularization. And solve the following minimization problem

$$\min_{u \in X} J(u) + \frac{1}{2\lambda} \sum_{x \in \Omega} |h * u(x) - z(x)|^2, \tag{2.10}$$

where $J(u) = \sum_{x \in \Omega} |Du(x)|$ is the discrete Total Variation regularizer. The typical constrained formulation of a regularization method [TA77] consists in choosing between all possible solutions of (2.1) the one that minimizes the functional $J(u)$. We may interpret λ as a penalization parameter which controls the tradeoff between the goodness of fit of the constraint and the smoothness term given by the Total Variation. In this formulation, a methodology is required for a correct choice of λ.

Let us mention that the regularization parameter λ in classical variational or bayesian formulations can be eliminated if we introduce the image model (2.1) using a constrained minimization approach. Since our information on the noise is reduced to its mean and variance and we can easily guarantee that the mean value is preserved, we can write a variational model where the constraint is expressed as an upper estimate of the noise variance σ^2. The constrained restoration problem becomes

$$\begin{aligned} &\min_{u} \quad J(u), \\ &\text{subject to } \sum_{x \in \Omega} |(h * u)(x) - z(x)|^2 \leq N^2 \sigma^2. \end{aligned} \tag{2.11}$$

Again, the regularizer $J(u)$ embodies our a-priori knowledge of the image, specifying its smoothness properties.

Notice that the image acquisition model (2.2) is only incorporated through a global constraint. Assuming that $h * 1 = 1$ (energy preservation), the additional constraint that $\sum_{x\in\Omega} h * u\,(x) = \sum_{x\in\Omega} z(x)$ is automatically satisfied by its minima [CL97].

The problem (2.11) is solved using the method of Lagrange multipliers, which leads again to the unconstrained formulation (2.10), however the parameter λ is adjusted to satisfy the constraint.

The connections between (2.11) and (2.10) were studied by Chambolle and Lions in [CL97] where they proved that both problems are equivalent for some positive value of the Lagrange multiplier λ.

2.2.1 Restoration of irregularly sampled images

Several applications in satellite imaging require to restore (or interpolate) an image from an irregularly distributed grid of samples. Some examples where irregular sampling grids arise naturally are the following:

Microvibrations. The vibrations of airborne Push-Broom imaging systems, produce an irregular sampling, which must be taken into account during the restoration.

Superresolution. Restore a single image by fusing information from several views of the same scene.

Geophysical potential fields such as the earth's gravitational field which can only be measured at sparse locations [Rau98].

In this context a general image acquisition system may be modeled by the following image formation model

$$z(\xi_k) = (h * u)(\xi_k) + n_{\xi_k}, \qquad \xi_k \in \Xi, \tag{2.12}$$

where $\Xi = \{\xi_k\}_{k=1}^{N^2} \subseteq \mathbb{R}^2$ is a finite set of regular or irregular samples, $u : \mathbb{R}^2 \to \mathbb{R}$ is the ideal undistorted image, $h : \mathbb{R}^2 \to \mathbb{R}$ is a blurring kernel whose Fourier spectrum $\hat{h}$ has most of its energy concentrated in the spectral support of u, z is the observed sampled image which is represented as a function $z : \Xi \to \mathbb{R}$, and n_{ξ_k} is, as usual, a white Gaussian noise with zero mean and standard deviation σ.

Even knowing the exact sampling geometry Ξ, the blurring kernel h, and the statistics of the noise n, the problem of recovering u from z is very ill-posed in a variety of situations that arise in real-world image acquisition systems. The main difficulty comes from the fact that perfect reconstruction from irregular samples can be ensured only if the sampling rate is either much larger than in the regular case (relative to the bandwidth of h), or is like in the regular case

but under very constrained sampling geometries that rarely appear in practice. (See [Alm02, Chapter 2] for a discussion on this problem). In any case, even if perfect reconstruction is sometimes possible in theory with a relatively small and realistic sampling rate, it is still very ill posed: This limits its practical interest whenever noisy measurements are involved.

Again the unavoidable strategy to solve this ill conditioning is regularization. The typical constrained formulation of a regularization method [TA77] consists in choosing between all possible solutions of (2.12) the one that minimizes the functional $J(u)$

$$\begin{aligned} &\min_u \quad J(u), \\ &\text{subject to } \sum_{\xi_k \in \Xi} |(h * u)(\xi_k) - z(\xi_k)|^2 \leq N^2 \sigma^2, \end{aligned} \tag{2.13}$$

where the regularizer $J(u)$ embodies our a-priori knowledge of the image, specifying its smoothness properties. In this formulation the image acquisition model has been incorporated as a single constraint. Indeed, we only know the statistics of the noise, its mean and variance. As in (2.11) the variance knowledge is formulated as a global constraint.

In [ABCH08] (see also [ACHR06, BCRS03, RO94]), the authors proposed a restoration model using a local estimate of the noise variance. The local formulation contributes to reducing the unnaturally-looking aspect of images obtained from global Total Variation based minimization, thus improving the recovery of textures. Following this proposal we replace the constraint

$$\sum_{\xi_k \in \Xi} |(h * u)(\xi_k) - z(\xi_k)|^2 \leq N^2 \sigma^2,$$

by a set of local constraints

$$G * |\Delta_\Xi (h * u) - z|^2 (\xi_k) \leq \sigma^2, \qquad \forall \xi_k \in \Xi,$$

where the sampling operator $\Delta_\Xi : C(\mathbb{R}^2) \to \ell^2(\Xi)$ is given by $\Delta_\Xi(v) = \{v(\xi_k)\}_{k=1}^{N^2}$ and G is a discrete convolution kernel such that $G(\xi) > 0$ for all $\xi \in \Xi$ and $\sum_k G(\xi_k) = 1$.

2.3 Denoising

A particular and important case contained in the above formulation is the denoising problem which corresponds to the case where $h = \delta$, so that the model (2.2) becomes

$$z(x) = u(x) + n(x), \qquad x \in \Omega, \tag{2.14}$$

where n is an additive Gaussian white noise of zero mean and variance σ^2.

In this case the MAP model is:

$$\min_{u \in X} J(u) + \frac{1}{2\lambda} \|u - z\|_X^2.$$

As above, the case where $J(u)$ is the discrete Total Variation has been studied in depth [AVCM04] and has had a strong influence in the development of recent methods for image denoising.

Recently, nonlocal regularizers have been proposed for image denoising [PMD$^+$10, PBC09]. They were motivated by a variational formulation of the nonlocal means filter [BCM05]. The underlying idea of the non local means filter is to give a probability model for the value of each pixel conditioned by its neighborhood. Then is pixel is restored by averaging the values of pixels with similar neighborhoods.

2.4 Image fusion

The fusion problem also fits the irregular sampling model (2.12). The objective of image fusion (super-resolution) is to recover a super-resolved version of the image using samples from several low resolution images. For the super-resolution to be effective the samples of each low resolution image must contain alias (i.e., must be sampled at a rate below the Nyquist rate corresponding to the highest frequency contained in $h * u$). Otherwise, according to the Shannon's theorem, a single image would be sufficient to recover the high resolution image.

Although the fusion problem fits our model, the application of (2.13) poses certain challenges.

For simplicity let us explain these problems in the discrete case, with a regular sampling omitting h from the model. So, strictly speaking we are not in the setting of (2.12) because h is not there to limit the bandwidth of the signal u. But, as a consequence, there is plenty of alias.

In this setting, let us consider a zooming problem where the data z is modeled by the subsampling (of factor k) of the discrete image $u \in X$

$$z(i,j) = u(k\,i, k\,j) + n(i,j) \quad (i,j) \in \{1, ..., N/k\}^2.$$

The objective is to recover an image resembling u. As noted in [MG02], the direct resolution of (2.13) (with $h = \delta$) using the Total Variation regularizer produces solutions with "point-like" artifacts at the positions of the samples (see figure 2.1) .

These artifacts are due to the unability of Total variation to impose the values of the solution on a set of isolated points. The corresponding solutions are "point-like". Let us explain intuitively why. By the coarea formula we can interpret the Total Variation as the sum of the perimeters of all the level sets of the image u. In a delta-like solution these perimeters are very small, while u still satisfies the restriction. Other interpolations will have level sets with larger perimeters, and hence higher Total Variation.

The solution proposed in [MG02] to avoid these artifacts is to introduce a spectral projector in the data term. This amounts to an increase the support of the restrictions in the spatial domain, which is equivalent to introduce a kernel P into the restoration model that cuts the frequencies that cannot be represented

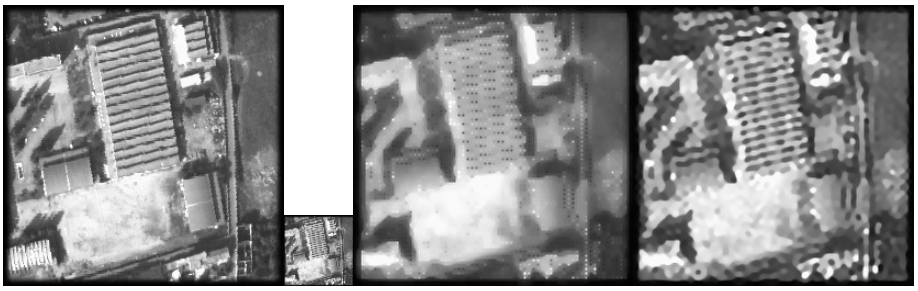


Figure 2.1: Restoration with zoom. From left to right: The original image u, the subsampled image z (4x) with noise of variance $\sigma = 1$. The restoration according to $\min_u \int |\nabla u|$ s.t. $\|S_4 u - z\|^2 \leq \sigma^2$ where S_4 is the subsampling operator, and the restoration using the projector $\min_u \int |\nabla u|$ s.t. $\|S_4 P u - z\|^2 \leq \sigma^2$.

by the sampling (i.e., it removes the alias frequencies from z)

$$z(i,j) = (Pu)(k\,i, k\,j) + n(i,j).$$

For regular grids, not even necessarily square, it is possible to the estimate this antialias projector if we assume certain statistics of the signal u (see [Alm02]).

When we considered in the previous section the restoration from irregular samples, we did not mentioned any spectral projector. The reason being that we implicitly assumed that the number of samples is almost the same as the number of restored pixels, and in that case we do not expect alias.

However, in the case of image fusion (super-resolution), our objective is to recover a zoomed version of the image by overlying samples from several images. The geometry of the combined grid may not be regular, and therefore there may be small and large gaps in the grid. The issue is to deal with the large gaps so that we avoid the "point-like" artifacts mentioned in the zoom problem.

As before, to cope with the problem of alias, we can use a spectral projector to remove the ambiguous frequencies from the solution. However since the sampling is irregular, we cannot define the projector in terms of non representable frequencies. Some heuristics for the estimation of the spectral projector in the case of irregular sampling are proposed in Chapter 4

On the other hand, even knowing (or after estimating) a spectral projector for an irregular sampling set, there is still a problem in the definition of the data term. The inclusion of the spectral projector in the restoration formulation is an "artifice" that allows us to avoid the artifacts mentioned before, however the acquisition model is still $z(i,j) = u(k\,i, k\,j) + n(i,j)$, so the data term

$$\sum_{(i,j)\in\{1,\ldots,N/k\}} ((Pu)(k\,i, k\,j) - z(i,j))^2 \leq \frac{N^2}{4}\sigma^2$$

will have a systematic bias in the estimation of the noise's variance. The bias corresponds to the noise frequencies that are being eliminated by the spectral

projector. With the current tools this bias cannot be estimated in general. This makes the constrained problem less accurate and prone to the apparition of artifacts due to noise over-fitting. These issues will be discussed in Chapter 4.

Chapter 3

Irregular to regular sampling, denoising and deconvolution

In this chapter we preset a restoration algorithm for band limited images that considers irregular (perturbed) sampling, denoising, and deconvolution. We explore the application of a family of regularizers that allow us to control the spectral behavior of the solution. And we incorporate the image acquisition model as set of local constraints. The experiments we present are focused on the restoration of satellite images, where the micro-vibrations are responsible of the type of distortions we are considering here.

3.1 Introduction

A general image acquisition system may be modeled by the following image formation model

$$z(\xi_k) = (h * u)(\xi_k) + n_{\xi_k}, \qquad \xi_k \in \Xi, \tag{3.1}$$

where $\Xi = \{\xi_k\}_{k=1}^{N^2} \subseteq \mathbb{R}^2$ is a finite set of regular or irregular samples, $u : \mathbb{R}^2 \to \mathbb{R}$ is the ideal undistorted image, $h : \mathbb{R}^2 \to \mathbb{R}$ is a blurring kernel whose Fourier spectrum $\hat{h}$ has most of its energy concentrated in the spectral support of u, z is the observed sampled image which is represented as a function $z : \Xi \to \mathbb{R}$, and n_{ξ_k} is, as usual, a white Gaussian noise with zero mean and standard deviation σ.

Reconstructing a signal $u : \mathbb{R}^2 \to \mathbb{R}$ over an infinite support from a finite set of samples $z(\xi_k)$ is not possible without imposing restrictions. As in most works, in order to simplify this problem, we shall assume that the functions h and u are periodic of period N in each direction. Let us denote by Ω_N the interval $[0, N[^2$ and assume that h, u are functions defined in Ω_N. To fix ideas, we assume that $h, u \in L^2(\Omega_N)$, so that $h * u$ is a continuous function in Ω_N [GW90] (which may be extended to a continuous periodic function in $\mathbb{R}^2$). Then the samples $(h * u)(\xi_k)$, $\xi_k \in \Xi$, are well defined.

We shall concentrate on the particular case of perturbed sampling and we shall assume that Ξ is a set of N^2 samples, which take the particular form

$$\Xi = \mathbb{Z}^2 \cap \Omega_N + \varepsilon(\mathbb{Z}^2 \cap \Omega_N), \tag{3.2}$$

where $\varepsilon : \mathbb{R}^2 \to \mathbb{R}^2$ is a "smooth and small" perturbation function in the sense that $\operatorname{supp} \hat{\varepsilon} \subseteq [-\frac{N}{T_\varepsilon}, \frac{N}{T_\varepsilon}]^2$ for some period $T_\varepsilon > 2$ corresponding to the maximum vibration frequency and the mean amplitude of the perturbation $([\int_{\Omega_N} |\varepsilon(x)|^2 \, dx / \int_{\Omega_N} dx]^{\frac{1}{2}})$ is small with respect to 1 pixel (we refer to Section 3.2 for a model (3.7) of this perturbation and also for a general overview of irregular sampling aspects).

Even knowing the exact sampling geometry Ξ, the blurring kernel h, and the statistics of the noise n, the problem of recovering u from z is very ill-posed in a variety of situations that arises in real-world image acquisition systems. The main difficulty comes from the fact that perfect reconstruction from irregular samples can be ensured only with either a sampling rate much larger than in the regular case (relative to the bandwidth of h), or similar to the regular critical sampling rate but under very constrained sampling geometries that rarely appear in practice. (See [Alm02, Chapter 2] for a discussion on this problem). In any case, even if perfect reconstruction is sometimes possible in theory, with a relatively small and realistic sampling rate, it is still very ill posed, which limits its practical interest whenever noisy measurements are involved.

A common and unavoidable strategy to solve this ill conditioning is regularization. In its constrained formulation, we choose between all possible solutions of (3.1) the one that minimizes a regularization functional $J(u)$. Thus, the image acquisition model (3.1) is introduced as a constraint. In this way, we eliminate the regularization parameter appearing in classical variational or bayesian formulations.

Since our knowledge of the noise is limited to its mean and variance, the constraint is expressed as an upper estimate of the noise variance σ^2. The constrained restoration problem becomes

$$\begin{aligned} &\min_u \quad J(u), \\ &\text{subject to } \sum_{\xi_k \in \Xi} |(h * u)(\xi_k) - z(\xi_k)|^2 \leq N^2 \sigma^2, \end{aligned}$$

where the regularizer $J(u)$ embodies our a-priori knowledge of the image, specifying its smoothness properties.

One of the first regularization methods consisted in choosing between all possible solutions of (3.1) the the one which minimized the Dirichlet integral $\int_{\Omega_N} |Du|^2 \, dx$. This formulation was an important step, but the results were not satisfactory, mainly due to the unability of the previous functional to resolve discontinuities (edges) and oscillatory textured patterns, the information corresponding to high frequencies of z being attenuated by it. Indeed, functions in

$W^{1,2}(\Omega)$ (i.e., functions $u \in L^2(\Omega)$ such that $Du \in L^2(\Omega)$) cannot have discontinuities along rectifiable curves. These observations motivated the introduction of Total Variation ($TV(u) = \int_{\Omega_N} |Du|$) in image restoration problems by L. Rudin, S. Osher and E. Fatemi in their work [ROF92]. The a priori hypothesis is that functions of bounded variation (the BV model) ([AFP00, EG92]) are a reasonable functional model for many problems in image processing, in particular, for restoration problems (see [ROF92]). Typically, functions of bounded variation have discontinuities along rectifiable curves, being continuous in the measure theoretic sense away from discontinuities. The discontinuities could be identified with edges. The ability of total variation regularization to recover edges is one of the main features which advocates for the use of this model (its ability to describe textures is less clear, some textures can be recovered, but up to a certain scale of oscillation).

Due to its strong ability to preserve discontinuities, the total variation may not be the best choice in the case of satellite images which, in order to look natural, have to keep some minimal smoothness due to the frequency cut of the optical system. This characteristic can be modeled by the spectral decay of the Fourier coefficients in the class of images we are looking for. In the case of satellite images, this spectral behavior can be estimated by statistical measures of the decay of Fourier coefficients.

We explore in this chapter a family of regularizers that takes into account the spectral decay of the Fourier coefficients in the class of images we are looking for. The general class of regularizers we consider is

$$J_A(u) = \int_{\Omega_N} |A(D)u|, \tag{3.3}$$

where $A(D)u$ is defined in terms of the coefficients of its Fourier series

$$\mathcal{F}(A(D)u)(\omega) = A(i\omega)\hat{u}(\omega) \quad \omega \in \mathbb{Z}^2.$$

Note that $J_A(u) < \infty$ imposes a frequency penalization according to the profile $A(i\omega)$. In practice we choose $A(i\omega)$ so that $|A(i\omega)| \sim |\frac{2\pi}{N}\omega|^p$ for large $|\omega|$, $1 \leq p \leq 2$. This is in consonance with the approach of Gröchenig and Strohmer [GS00] that proposes to incorporate an a priori decay in the restoration process (see Section 3.2).

We also incorporate the image acquisition model as a set of constraints. As opposed to the use of a single global constraint, in [ABCH08] (see also [ACHR06, BCRS03, RO94]), the authors proposed a restoration model using a local estimate of the noise variance. The local formulation contributes to reducing the unnaturally-looking aspect of images obtained from global Total Variation based minimization, thus improving the recovery of textures. Following the mentioned proposal we replace the constraint

$$\sum_{\xi_k \in \Xi} |(h * u)(\xi_k) - z(\xi_k)|^2 \leq N^2\sigma^2,$$

by

$$G * |\Delta_\Xi(h * u) - z|^2(\xi_k) \leq \sigma^2, \qquad \forall \xi_k \in \Xi, \tag{3.4}$$

where the sampling operator $\Delta_\Xi : C(\mathbb{R}^2) \to \ell^2(\Xi)$ is given by $\Delta_\Xi(v) = \{v(\xi_k)\}_{k=1}^{N^2}$ and G is a discrete convolution kernel such that $G(\xi) > 0$ for all $\xi \in \Xi$ and $\sum_k G(\xi_k) = 1$.

Combining the two ideas described above, the use of a regularizer that takes into account the spectral decay of images in a certain class (3.3), and the incorporation of the image acquisition model as a set of local constraints (3.4), we propose the following constrained variational model for restoring u

$$\begin{aligned} &\min_u \quad \int_\Omega |A(D)u|, \\ &\text{subject to } \left[G * |\Delta_\Xi (h * u) - z|^2\right](\xi_k) \leq \sigma^2 \quad \forall \xi_k \in \Xi. \end{aligned} \tag{3.5}$$

The constrained formulation (3.5) can be solved using the unconstrained formulation

$$\min_u \max_{(\lambda_k)\geq 0} \int_\Omega |A(D)u| + \frac{1}{2} \sum_{\xi_k \in \Xi} \lambda_k \left\{ \left[G * |\Delta_\Xi (h * u) - z|^2\right](\xi_k) - \sigma^2 \right\}, \tag{3.6}$$

where $\lambda_k \geq 0$ are Lagrange multiplier that have to be chosen so that the constraints (3.4) are satisfied. Let us say explicitly that both the blurring kernel h and the sampling grid Ξ (alternatively the grid perturbation function ε) are assumed to be known exactly and that the only thing known about the noise n_{ξ_k} is that it is a white Gaussian noise with zero mean and known variance σ^2. Several methods exist to estimate all these parameters [JBFZ02] for a given acquisition device, and we shall not address this question here.

The case of recovering an irregularly sampled image on a regular sampling grid was considered in [Alm02], but the blurring kernel h was assumed to be an ideal window (with Nyquist frequency cutoff), i.e., $\hat{h} = \chi_{[-1/2,1/2]^2}$. Different numerical algorithms were tested in [Alm02] in the case where the sampling set is perturbed according to (3.2) and they worked relatively well only within a low frequency spectral region $R \subseteq [-\alpha, \alpha]^2$, where $\alpha \approx \frac{1}{2} - 1/T_\varepsilon$. When attempting to recover $\hat{u}$ in the high frequency band $[-1/2, 1/2]^2 \backslash R$ serious theoretical and numerical problems appeared and, actually, restoration errors were most important there. Subsequently, the restoration problem (3.5) was studied in [ACHR06] when $J(u)$ is the total variation and the image acquisition model was incorporated as a set of local constraints on a partition of the image obtained as a result of a segmentation. The use of local constraints (3.4) was advocated in [ABCH08] and we also adopt this technique here.

Let us finally mention that many numerical algorithms have been proposed to minimize total variation (or similar models) subject to a global constraint as in (3.5) [ROF92, GR92, VO96, Cha04, CL97, CZC95, CM99, DMR00, BFCAB95,

CBFAB97, SGG$^+$08]. Imposing local constraints in a partition of the image was proposed in [RO94] and further developed in [BCRS03, ACHR06, ABCH08]. In [LM04] the authors combined total variation minimization with a set of constraints of type $|\langle h * u - z, \psi \rangle| \leq \tau$ where ψ varies along an orthonormal basis of wavelets (or a family of them) and $\tau > 0$. The aim was also to construct an algorithm which preserves textures and has good denoising properties. As we will do here, these constraints were incorporated using Uzawa's algorithm. In [GSZ03], the authors proposed to minimize total variation subject to a family of local constraints which control the local variance of the oscillatory part of the signal. The constraints are introduced via Lagrange multipliers with an approach similar to the one used in [ROF92]. This amounts to adding a spatially varying fidelity term that locally controls the extent of denoising over image regions depending on their content. Besides the fact that we use Uzawa's algorithm and we try to address the problem of deconvolution and denoising of irregularly sampled images, the work [GSZ03] is quite similar to our approach. Finally, in [SGG$^+$08], the authors proposed a non-convex data attachment term with a larger weight (depending on the inverse of the modulus of the gradient) in flat areas than in textured ones. This goes in the same sense as our model, that is, to be able to denoise flat regions of the image while keeping the oscillations in textured areas.

Summary of the contributions of this chapter. Our goal is to propose an algorithm for image denoising and deconvolution. The specificity of the considered problem comes from the fact that the images we deal with are irregularly sampled. On the one hand, we carry out a thorough and accurate modelization of the problem based on local constraints (3.4), and propose a family of frequency adaptive regularization functionals (3.3) that incorporates a priori image smoothness model. On the other hand, we prove the convergence of our algorithm. Finally, the study of the local constraint model combined with standard error propagation allows to derive useful stopping criteria for the algorithm.

Organization of this chapter. In Section 3.2 we introduce the problem of irregular to regular sampling and we discuss the ACT algorithm of Gröchenig and Strohmer [GS00]. In Section 3.3 we introduce our frequency adaptive variational restoration model with local constraints in the continuous setting. The corresponding discrete model is discussed in Section 3.4 along with the adaptation of a computational improvement introduced by L. Moisan in [Moi07]. In Section 3.5 we study the existence, uniqueness and numerical approximation to the model introduced in the previous section. This study is completed in Section 3.6 where we describe a Quasi-Newton algorithm for the solution of the Euler Lagrange equation corresponding to the energy in (3.6) for fixed values of the Lagrange multipliers (λ_k). In Section 3.7 we propose a practical stopping condition for the restoration algorithm for the local constraint model. In Section 3.8 we display experiments concerning restoration and zooming of irregularly sampled images. Section 3.9 summarizes the main conclusions of this

work.

3.1.1 Preliminaries and notations

Let us introduce some notation that will be used throughout this chapter.

For any function $u \in L^2(\Omega_N)$ (assuming periodicity of period N in each direction) we denote its Fourier coefficients as

$$\hat{u}(p,q) = \frac{1}{N^2} \int_{\Omega_N} u(x,y) e^{-2\pi i \frac{(px+qy)}{N}} \, dx \, dy \qquad \text{for } (p,q) \in \mathbb{Z}^2.$$

As in [MG02], our plan is to compute a band limited approximation to the solution of the restoration problem. To do so, assume for simplicity that M is an even number and define

$$\mathcal{B}_M := \{u \in L^2(\Omega_N) : \hat{u} \text{ is supported in } I_M\}, \text{ where } I_M := \{-\frac{M}{2}+1, \dots, \frac{M}{2}\}^2.$$

Notice that since $\mathcal{B}_M$ is a finite dimensional vector space of dimension M^2, then it can be identified with $\mathbb{R}^{M^2}$ by mapping $u \in \mathcal{B}_M$ to the vector $\vec{u} = \left(u(r\frac{N}{M}, l\frac{N}{M})\right)_{r,l=0}^{M-1}$. Moreover, if $u \in \mathcal{B}_M$ we may write

$$u(x,y) = \sum_{-\frac{M}{2} < p,q \leq \frac{M}{2}} \hat{u}(p,q) e^{2\pi i \frac{(px+qy)}{N}}.$$

where

$$\hat{u}(p,q) = \frac{1}{M^2} \sum_{0 \leq r,l < M} u\left(r\frac{N}{M}, l\frac{N}{M}\right) e^{-2\pi i \frac{(pr+ql)}{M}}, \qquad -\frac{M}{2} < p,q \leq \frac{M}{2}.$$

Then the values $u\left(r\frac{N}{M}, l\frac{N}{M}\right)$, $0 \leq r,l < M$, can be recovered as the discrete inverse Fourier transform of $\hat{u}(p,q)$. Hence $u \in \mathcal{B}_M$ can also be identified with the vector of Fourier coefficients $\hat{u} \in \mathbb{C}^{M^2}$. Note that we shall mainly study here the critical sampling case $M = N$, however we will keep two different symbols M for the bandwidth and N for the domain size in order to simplify the exposition of certain parts of the algorithm.

We intend to solve the restoration problem in the class of band-limited functions $\mathcal{B}_M$. Later on we will comment on this choice. We will also use the operator notation for the Fourier transform that applied to the function u returns the vector of its Fourier coefficients: $\hat{u} = \mathcal{F}u$. Conversely $\mathcal{F}^*\hat{u} = u$ denotes the inverse transform (we may also denote it as $\hat{u}^\vee$) then $\mathcal{F}^*\mathcal{F} = Id$.

3.2 Irregular to regular sampling

Opposed to digital photographs, satellite images are generally not acquired by a squared array of sensors but by a sweeping bar of sensors known as TDI (Time

Delay Integrator) [Sch07]. This acquisition geometry called *push-broom* is widely applied in aerospace imaging applications and, nowadays, it provides the highest resolution in earth imaging applications. As a consequence of this progressive acquisition mode, the micro-vibrations of the satellite together with irregularities in sensors position result in perturbed sampling sets. In most cases, the knowledge of certain vibration modes and the analysis of acquired images help to estimate, very accurately, the perturbations in the sampling grid, which can be modeled [Alm02] by

$$\varepsilon(x) = \sum_{k=1}^{q} a_k(x) cos(2\pi \langle \frac{\omega_k}{N}, x \rangle + \phi_k), \qquad x \in \mathbb{R}^2, \tag{3.7}$$

for some $q \geq 1$, where $a_k(x)$ are smooth modulation functions and the vibration frequencies ω_k are an order of magnitude (or even more) below the Nyquist frequency of the sampling rate. The bound on the modulation functions is inversely proportional to ω_k and the number of vibration modes is small. This results in smooth and small perturbations with $|\varepsilon(x)|$ no larger than a few pixels, and perturbation slope $|\nabla\varepsilon(x)|$ no larger than about one tenth of a pixel per pixel. As a consequence these perturbations are hardly noticeable and we should talk of *perturbed sampling* rather than irregular sampling in those cases. Even if the image distortion is not evident from a geometrical point of view it is very important to correct the perturbations in image registration applications where a sub-pixel accuracy is necessary.

In order to be less dependent on a particular physical instrument, in our experiments we used a simplified version of this model which still captures its main characteristics, namely the perturbation function $\varepsilon = (\varepsilon_1, \varepsilon_2)$ is simulated as a discrete colored noise, i.e. for $\omega \in \mathbb{Z}^2$ we define

$$\begin{array}{ll} \widehat{\varepsilon_i}(\omega) \sim N(0, \tilde{\sigma}^2) & \text{if } |\omega| \leq N/T_\epsilon, \\ = 0 & \text{otherwise,} \end{array} \tag{3.8}$$

where $\tilde{\sigma}$ is chosen in such a way that the standard deviation of $\varepsilon_i(x)$ is A for $i \in \{1, 2\}$. This gives $\tilde{\sigma} = \frac{AT_\epsilon}{2}$ (we have taken the Fourier transform as an isometry). Thus the behavior of the perturbation is characterized by the two parameters "amplitude" A and maximal vibration frequency N/T_ε (or "minimal vibration period" T_ε). The precise values of A and T_ε used in our experiments will be specified in the experiments Section.

There are many works in the literature dealing with the irregular to regular sampling problem. Let us mention that, according to Kadec's theorem [Kad64], we have a perfect recovery of the signal if we consider a perturbed sampling with small perturbations $|\varepsilon(x)| \leq 0.11$. Recall also that Beurling-Landau's theorem [Lan67], ensures perfect reconstruction of a function from its samples for arbitrary *stable sampling sets* [Lan67], but it requires the (lower) sampling density to be larger than 1. These conditions are very restrictive and do not hold true for most of the image restoration applications. For a comparison between several iterative methods we refer the reader to [AM01, CFH91, Alm02].

3.2.1 The ACT algorithm

One of the best performing reconstruction methods available for irregular to regular sampling is the ACT algorithm (for Adaptive-weights Conjugate-gradient on Toeplitz-matrix) introduced by Gröchenig *et. al.* in [GS00]. This method represents a discrete image u as a trigonometric polynomial of order $M/2$ in each variable (for simplicity of notation we shall assume that M is an even number) so that the interpolation at the sampling points $\Xi = \{\xi_k\}_{k=1}^{N^2} \subseteq \mathbb{R}^2$ becomes

$$u(\xi_k) = \sum_{t \in \{-\frac{M}{2}+1,\ldots,\frac{M}{2}\}^2} a_t e^{\frac{2\pi i}{M}\langle t,\xi_k\rangle}, \quad k \in \{1,\ldots,N^2\}. \tag{3.9}$$

Thus, if z represents the irregularly sampled data we may write [GS00]

$$z = Sa, \quad \text{where} \quad S = ((s_{kt})),\ s_{kt} = e^{\frac{2\pi i}{M}\langle t,\xi_k\rangle}, \tag{3.10}$$

i.e. S is the Vandermonde matrix associated to the trigonometric polynomial in (3.9). Note that S maps $a \in \ell^2(\{-\frac{M}{2}+1,\ldots,\frac{M}{2}\}^2)$ to $z = \{z(\xi_k)\} \in \ell^2(\Xi)$ as given in (3.9). The bandwidth M of the trigonometric polynomial is chosen to be $M \leq N$, so the system (3.10) is expected to be determined or over-determined.

Following [GS00], the ACT algorithm recovers the coefficients a of the trigonometric polynomial by solving the least squares problem

$$\arg\min_a \|\sqrt{W}(Sa - z)\|^2, \tag{3.11}$$

where the matrix $W = \mathrm{diag}(\{w_k\}_{k=1..N^2})$ assigns weights that are inversely proportional to the sampling density at ξ_k:

$$w_k = \mathrm{area}(V_k) \qquad \text{where} \qquad V_k := \left(\{x : |x-\xi_k| < |x-\xi_j|,\ \forall j \neq k\}\right). \tag{3.12}$$

If we interpret the discrete samples $z(\xi_k)$ as a piecewise constant function $\sum_k z(\xi_k)\chi_{V_k}$, then the weights w_k guarantee the isometry between the irregular sampling on the grid Ξ and its function representation, thus compensating the local variations in the sampling density. Moreover, by using the weights W, Gröchenig and Strohmer provide an explicit estimate for the rate of convergence of the ACT algorithm [GS00].

The system of normal equations associated to (3.11) is

$$S^*WSa = S^*Wz\,, \tag{3.13}$$

where S^* denotes the adjoint matrix of S (a notation that will be used throughout this chapter). Observe that the $M^2 \times M^2$ matrix S^*WS has a Toeplitz structure [GS00], and thus, S^*WSa is efficiently computed in $\mathcal{O}(M^2\log_2(M^2))$ time using Fourier methods. Moreover the entries of $T := S^*WS$ and $b := S^*Wz$ can be approximated using the NFFT [PST01] in $\mathcal{O}(M^2\log_2(M^2))$ time each [GS00]. Finally, (3.13) is solved using a *conjugate gradient* (CG) method. The following algorithm summarizes the method.

Algorithm I: ACT algorithm for a fixed bandwidth M
Requires: N^2 irregular samples in vector z.
Ensures: M^2 regular samples in vector u.

1. Compute $T = S^*WS$ and $b = S^*Wz$ using the NFFT.
2. Solve $Ta = b$ using conjugate gradients.
3. Compute the regular samples $u(i\frac{N}{M}, j\frac{N}{M})$ for $(i,j) \in \{0,\dots,M-1\}^2$ by applying the inverse FFT to a.

Let us note that in the more realistic cases where T is not invertible or ill-conditioned, the CG solver acts as a regularizer and chooses the minimum norm solution a among those that satisfy (3.13). This is a constrained variational formulation that can be written using a Lagrange multiplier $\lambda > 0$ as the unconstrained minimization problem

$$\min_a \|a\|^2 + \lambda \|Ta - b\|^2.$$

This formulation also applies to the following two variants of the ACT algorithm [GS00] that incorporate an a-priori spectral decay rate $|\hat{u}(\omega)| \leq L\phi(\omega)$, for some $L > 0$, of the image class to be restored (when available). For satellite images this estimation has been performed by Almansa in [Alm02] and it is given by $\phi(\omega) = (1 + |\frac{2\pi}{N}\omega|)^{-p}$ for some value of p near 1.6.

- The first proposed variant solves $CTa = Cb$ (instead of (3.13)) using the CG algorithm, where $C = diag(\{\phi(\omega)\}_{\omega \in \{-\frac{M}{2}+1,\dots,\frac{M}{2}\}^2})$. Notice that we can write the problem as

$$(\mathrm{ACT}_D) \qquad \min_a \|a\|^2 + \lambda \|C(Ta - b)\|^2, \tag{3.14}$$

  for some $\lambda > 0$. Since the weighting matrix C is applied to the residuals $Ta - b$, then it affects the relative cost of the errors by shifting the cost from high frequencies towards the lower ones. This shift is reflected in CG search directions, altering the intermediate solutions so that they will fit the low frequency before the higher ones.

- The second ACT variant solves $Td = b$ where $a = Cd$, using the CG algorithm. Re-writing it as an optimization problem, observe that the weights appear now in the regularity term

$$(\mathrm{ACT}_R) \qquad \min_a \|C^{-1}a\|^2 + \lambda \|Ta - b\|^2. \tag{3.15}$$

  The spectral weights C^{-1} are now penalizing the apparition of higher frequencies in the solution a and not in the residual.

In either case, if T is invertible and the CG algorithm converged, then the solution of both variants coincides with the solution of (3.13). But the CG iteration is truncated before its convergence mainly due to the ill-conditioning of the operator T. So, the solutions obtained by the above methods differ because of the different search directions. As it can be observed experimentally, incorporating the spectral decay indeed reduces the restoration errors, specially when applied to the regularity term in (3.15) (see Table 3.2 in Section 3.8). In that case, it amounts to finding a solution in a class of functions with a particular spectral decay.

Remark 1. *(Global constraint & stopping condition)* In [GS00] the authors also proposed to extend the ACT algorithm in order to consider the presence of Gaussian noise n (with standard deviation σ) in the image formation model: $Sa + n = z$. This extension is implemented as a stopping condition for the CG algorithm (Step 2 of Algorithm I)

$$\text{stop CG if} \quad \|Sa - z\|^2 \leq \tau N^2 \sigma^2 \quad \text{with} \quad \tau \simeq 1$$

that is designed to avoid the over-fitting of the solution and hence the amplification of the noise. Thus, the ACT with residual-based stop conditions can be seen as a numerical approximation to the minimization of $\|Ca\|^2$ subject to the constraints above.

Remark 2. *(Fourier vs. spline models)* Notice that the ACT Algorithm is based on the underlying assumption that the data can be represented by a trigonometric polynomial. Other interpolation models like the B-Spline have been used in the literature [ASHU05b]. In this work we will restrict ourselves to trigonometric polynomials mainly because convolutions are more easily modeled in this setting, but we intend to explore the use of B-splines in the future.

3.3 A frequency adaptive restoration model with local constraints

In this section we sketch main elements of the restoration model adopted in this chapter. To simplify the presentation the model is first defined in a continuous setting, and the discretization details are deferred to Section 3.4. The proposed model is inspired from the ideas presented in the previous section on the ACT, as well as on a recently proposed ACT+TV extension (next subsection), and the use of local constraints (Subsection 3.3.3) as in [ABCH08].

3.3.1 The ACT + TV extension

In [ACHR06] the authors proposed to combine the ACT algorithm written as (3.11) with total variation regularization, i.e.,

$$\begin{aligned} &\min_u \int_{\Omega_N} |Du|, \\ &\text{subject to } \|\sqrt{W}(S\hat{u} - z)\|^2 \leq N^2\sigma^2. \end{aligned} \tag{3.16}$$

For convenience, we refer to this model as ACT+TV. As reported in [ACHR06] (see also Table 3.2 in Section 3.8) we observe an improvement of ACT+TV with respect to the original ACT algorithm in terms of MSE error, this improvement is mainly attributed to the edge preserving ability of the TV regularizer.

3.3.2 Frequency Adaptive Regularization (FAR)

Inspired by (3.15) we propose to integrate the spectral weight priors given by the matrix C into the ACT+TV formulation.

From the ACT+TV formulation we keep: *(i)* the use of the spatial L^1-norm of a (pseudo)differential operator as a regularizer (instead of the L^2-norm $\|C^{-1}a\|$ in (3.15)), because of the better edge-preserving capabilities of the L^1-norm; and *(ii)* the formulation as an optimization problem under constraints, which enables us to choose automatically the regularization parameter λ in (3.15).

From the ACT_R we keep the idea of using a frequency adapted regularizer. For that we define the function $\omega \in \mathbb{Z}^2 \to A(i\omega) \in \mathbb{C} \times \mathbb{C}$. We assume that

$$\begin{aligned} &A(0) = 0, \\ &A(i\omega) \neq 0 \quad \forall \omega \neq 0, \\ &\text{and} \quad |A(i\omega)| \leq L\left(1 + \left|\frac{2\pi}{N}\omega\right|^{\kappa}\right), \ \forall \omega \in \mathbb{Z}^2, \text{ for some } L > 0, \kappa \geq 0. \end{aligned} \tag{3.17}$$

If $u \in C^\infty(\Omega_N)$ can be extended as a smooth and periodic function to $\mathbb{R}^2$, we define $A(D)u$ by its Fourier coefficients

$$\widehat{A(D)u}(\omega) = A(i\omega)\hat{u}(\omega) \qquad \omega \in \mathbb{Z}^2.$$

We define the regularizer functional

$$J_A(u) = \int_{\Omega_N} |A(D)u| = \int_{\Omega_N} |(A(i\omega)\hat{u}(\omega))^\vee|.$$

Notice that $J_A(u)$ can be defined for any $u \in L^2(\Omega_N)$ such that $A(D)u$ is a Radon measure. In practice this will not be a problem since we only consider band limited functions.

The total variation $J(u) = \int_{\Omega_N} |Du|$ corresponds to the choice $A(i\omega) = \frac{2\pi}{N} i\omega$. Recall that functions with finite total variation are a good model for image

restoration since they permit to recover the discontinuities of the image. But, in practice, digital images may exhibit a stronger decay in its Fourier coefficients than $|\frac{2\pi}{N}\omega|^{-1}$ and other functional models can be better suited. The operator $A(D)$ permits us to penalize the frequencies according to the profile $A(i\omega)$. Notice that

$$|A(i\omega)\hat{u}(\omega)| = |\widehat{A(D)u}(\omega)| \leq \int_{\Omega_N} |A(D)u| =: L$$

for all $\omega \in \mathbb{Z}^2$, hence

$$|\hat{u}(\omega)| \leq \frac{L}{|A(i\omega)|}.$$

If $|A(i\omega)| \sim |\frac{2\pi}{N}\omega|^{\kappa}$ for large $|\omega|$, then $|\hat{u}(\omega)|$ decreases as $|\frac{2\pi}{N}\omega|^{-\kappa}$. In this way we can favor a specific decay rate of the Fourier coefficients of u as a prior.

Remark 3. In the ACT_R algorithm the frequency adaptive regularizer has the form $C^{-1}\hat{u} = \phi^{-1}(\omega)\hat{u}(\omega)$ where $\phi(\omega) = (1 + |2\pi\omega/N|)^{-k}$, $k \geq 0$. In practice we consider functions $A(i\omega) = \frac{2\pi i\omega}{N}\psi(\omega)$ where $\psi(\omega) = (1 + |2\pi\omega/N|)^{p}$, $p \geq 0$ ($p = k - 1$ gives the same decay). Then we can write $A(D)u = D\Psi u$, where $\widehat{\Psi u} = \psi\hat{u}$, and $A(D)u$ can be seen as the derivative of Ψu a filtered version of u. In addition the effect of the new regularizer can be compared, in terms of frequency decay, to the ACT_R preconditioning.

3.3.3 Local constraints

Following [ABCH08] we will interpret the acquisition model as a set of local constraints (3.4) and we will search for a solution of the restoration problem that minimizes the frequency adaptive regularizer $J_A(u)$ with a data-fitting term derived from the acquisition model and consisting in a set of local constraints (as in (3.5)). For convenience we will describe these constraints directly in the discrete model (see the next section).

3.4 A discrete regularization model

Let us stress here the fact that the regularization functional J_A defined in the previous section is adapted to the restoration of functions with infinite resolution while its numerical approximation restricts the solution to be in a finite dimensional space. We are going to adopt here the following practical point of view. Since our data consists of a finite set of samples, we are going to reconstruct a sampled version of the image and therefore we work in a finite dimensional space. This reflects the fact that digital images have finite resolution. Usually, the restored image is modeled as a piecewise constant function (the values given on the set of pixels), but we consider here images as bandlimited functions with a finite number frequencies, since this is a reasonable model for restoring digital images. Moreover this model is adapted to compute convolutions and permits to include an a priori decay of the Fourier coefficients.

Based on the above considerations, we propose the following discrete regularization functional, If $u \in \mathcal{B}_M$, then we define

$$J_A^d(u) = \frac{1}{M^2} \sum_{0 \le r,l < M} \left| A(D)u\left(\frac{rN}{M}, \frac{lN}{M}\right)\right|. \tag{3.18}$$

If $A(i\omega)$ satisfies (3.17), then $J_A(u)$ and J_A^d are seminorms in $\mathcal{B}_M$ and the only function $u \in \mathcal{B}_M$ such that $J_A(u) = 0$ (resp. such that $J_A^d(u) = 0$) is $u =$ constant. Thus $J_A(u)$ and J_A^d are norms in the finite dimensional quotient space $\mathcal{B}_M/\mathbb{R}$, hence they are equivalent. Notice that if $u \in \mathcal{B}_M$ and we define $J_A^{d,k}$ by replacing M by k in (3.18), then $J_A^{d,k}(u) \to J_A(u)$ as $k \to \infty$. Unless we intend to zoom and restore the images, we take $M = N$, where N^2 is the number of data.

Assume that the input data (the measurements) consist of N^2 samples $\{z(\xi_k)\}_k$, and let $\bar{\sigma}, \beta > 0$, and $G \in \ell^\infty(\mathbb{Z}^2)$ be a discrete, positive, normalized convolution kernel such that $G(r,l) \ge 0$ and $\sum_{(r,l)\in\mathbb{Z}^2} G(r,l) = 1$. Then we propose to minimize the functional

$$\begin{aligned} &\min_{u\in\mathcal{B}_N} J_{\beta,A}(u) := \sum_{0\le r,l<N} \sqrt{\beta^2 + |A(D)u\,(r,l)|^2}\quad , \\ &\text{subject to } \left[G * |\Delta_\Xi\,(h*u) - z|^2\right](\xi_k) \le \bar{\sigma}^2 \quad \forall \xi_k \in \Xi, \\ &\text{and } \sum_{0\le r,l<N} u\,(r,l) = \sum_{\xi_k\in\Xi} w_k z(\xi_k) =: \bar{z}^w. \end{aligned} \tag{3.19}$$

To avoid the non differentiability of J_A^d at 0, it is a common approximation to use the regularizer $J_{\beta,A}$ instead.

Therefore we will minimize $J_{\beta,A}(u)$ on $\mathcal{B}_N$ subject to the family of constraints in (3.19). Notice that we have incorporated the image acquisition model (3.1) as a set of local constraints. The convolution of G and $v \in \ell^\infty(\Xi)$ is defined in the usual way by imposing an arbitrary regular grid structure on Ξ, *i.e.* $(G * v)(\xi_k) = \sum_{l\in\mathbb{Z}^2} G(l-k)v(\xi_l)$. In the case of perturbed sampling this regular grid structure may be determined by the original unperturbed grid, otherwise it may be based on a nearest neighbor computation.

Notice that we have used the value $\bar{\sigma} > 0$ as an estimate of the standard deviation of the noise. We will make tests with $\bar{\sigma} = \sigma$ and also with values of $\bar{\sigma}$ different from σ. Also, the effective support of G must permit the statistical estimation of the variance of the noise. In Section 3.7 we will come back to the noise estimation issue and the choice of $\bar{\sigma}$.

In the rest of the chapter, we assume that the blurring kernel h satisfies

$$h \in L^2(\Omega_N), \qquad \operatorname{supp}\hat{h} \subseteq [-\frac{M}{2}, \frac{M}{2}]^2, \qquad \text{and} \qquad \hat{h}(0,0) = 1. \tag{3.20}$$

If $u \in \mathcal{B}_M$, then we can compute the convolution $h * u$ using the Fourier transform: $\widehat{h*u}\,(p,q) = \hat{h}\,(p,q)\,\hat{u}\,(p,q)$.

The last equality constraint in (3.19) fixes the global mean of u to be the mean of the samples z, weighted by the areas w_k (see (3.12)). This constraint is necessary to assure the uniqueness of the solution, since the data fitting is provided only by inequality constraints, and the solution may be undetermined up to a constant (in the kernel of the regularizer) in areas where the variance of z is smaller than σ^2 [BCRS03]. The details of the uniqueness proof are given in Section 3.5.

Now, our purpose is to prove that the constrained formulation of (3.19) can be solved using Uzawa's method once we guarantee that the assumptions of Uzawa's method [Fau88] hold. But before that, we comment on a improved discretization for approximating $J_A(u)$.

3.4.1 An improved discretization formula

In this section we follow the proposal made by Moisan in [Moi07] to improve the discretization of the total variation formula. The basic observation is that, the computation of the Dirichlet integral in $[0,N)^2$ ($\int_{[0,N)^2} |\nabla u|^2$) cannot be done accurately unless we previously zoom the image u by a factor of two. The same argument applies to the case of $\int_{\Omega_N} |A(D)u|^2$, but not for $\int_{\Omega_N} \sqrt{\beta^2 + |A(D)u|^2}$ with $\beta \geq 0$. Indeed, in the last case, an exact computation would involve an infinite number of samples. However, as Moisan has shown for the TV case [Moi07], doubling the number of variables leads to a good approximation of the above integral, being a good compromise between precision and algorithmic efficiency.

Let us analyze the implications of Moisan's discretization for our restoration model. To do so, we need to introduce some notation. For each $M \in \mathbb{N}$, we denote by X_M the Euclidean space $\mathbb{R}^{M\times M}$. The Euclidean scalar product and the norm in X_M will be denoted by $\langle\cdot,\cdot\rangle_{X_M}$ and $\|\cdot\|_{X_M}$, respectively, but in absence of ambiguities we will omit the subindex. X_M represents the space of images $\mathcal{B}_M$ sampled in the regular grid $\{0,\dots,M-1\}\times\{0,\dots,M-1\}$ (or given by its Fourier coefficients $\hat{u} \in \ell^2(I_M)$. Let us introduce the operator

$$P : X_N \to X_{2N}\,, \quad P\{u(k,l)\} = \left\{\bar{u}\left(\tfrac{r}{2},\tfrac{l}{2}\right)\right\}_{r,l\in\{0,\dots,2N-1\}}\,,$$

where $\{u(r,l)\}_{r,l\in\{0,\dots,N-1\}} \in X_N$ and $\bar{u}$ is the function of $\mathcal{B}_N$ defined by the samples $\{u(r,l)\}_{r,l\in\{0,\dots,N-1\}}$. Observe that $\widehat{Pu}(p,q) = \mathbf{1}_{I_N}(p,q)\hat{u}(p,q)$. So we may consider the operator $A(D)$ as acting on X_{2N} or as acting on X_N. Notice that if $u \in X_N$ we may write $A(D)Pu = PA(D)u$. From now on we will avoid (except in ambiguous cases) the use of subindexes to specify the function spaces of norms and scalar products, and the function space should be clear from the context.

Thus, our final restoration model is

$$\begin{aligned} &\min_{u\in X_N} K_{\beta,A}(u) := \sum_{0\le r,l<2N} \sqrt{\beta^2 + |A(D)Pu\,(r,l)|^2}\quad , \\ &\text{subject to } \left[G * |\Delta_\Xi\,(h*u) - z|^2\right](\xi_k) \le \bar\sigma^2 \quad \forall \xi_k \in \Xi\ , \\ &\text{and } \sum_{0\le r,l<N} u\,(r,l) = \sum_{\xi_k\in\Xi} w_k z(\xi_k) =: \bar z^w. \end{aligned} \tag{3.21}$$

3.5 The well-posedness of the model and its numerical solution

Proposition 1. *Assume that* (3.20) *holds. Then there exists a unique minimum* $u \in X_N$ *of (3.21).*

Proof. Let u_m be a minimizing sequence of (3.21). Since $A(D)Pu_m$ is bounded in X_{2N}, and $\omega = 0$ is the only vanishing frequency for $A(i\omega)$ we deduce that $v_m := u_m - \widehat{u_m}(0,0)$ is bounded in X_N. Now, since $\widehat{u_m}(0,0)$ is constrained to be $\bar z^\omega$, we have that u_m is bounded in X_N. By extracting a subsequence, if necessary, we may assume that $u_n \to u$. It is immediate to see that u satisfies the constraints. Since $K_{\beta,A}$ is lower semicontinuous, we have that u is a minimum of (3.21).

Now, let u_1, u_2 be two minima of (3.21). If $A(D)Pu_1 \neq A(D)Pu_2$, letting $\overline{u} = (u_1+u_2)/2$, then the strict convexity of $K_{\beta,A}$ proves that $K_{\beta,A}(\overline{u}) < \inf_{u\in X_N} K_{\beta,A}$, a contradiction. Thus $A(D)Pu_1 = A(D)Pu_2$ and we have uniqueness modulo constants, i.e., $u_1 - u_2 = c$ for some $c \in \mathbb{R}$. Since $\hat u_1(0,0) = \hat u_2(0,0)$ we deduce that $c = 0$, and therefore, $u_1 = u_2$. □

Remark 4. Proposition 1 is also true if instead of assuming the average constraint in (3.21) we assume that $\inf_{c\in\mathbb{R}} G * (z-c)^2 > \bar\sigma^2$. This can be proved as in [BCRS03, CL97]. But in that case, we should also use a different (gradient descent based) algorithm to minimize (3.21) as described in [BCRS03].

From now on, we assume that the constraints are qualified, that is there is $u \in X_N$ such that

$$\hat u(0,0) = \bar z^w \qquad \text{and} \qquad \left[G * |\Delta_\Xi(h*u) - z|^2\right](\xi_k) < \bar\sigma^2, \qquad \forall \xi_k \in \Xi, \tag{3.22}$$

which implies that the set of functions satisfying the constraints is non-empty.

We prove that the solution of (3.21) can be computed by adapting Uzawa's algorithm. Let $\mu > 0$ and $\lambda = (\lambda_k)_{k=1}^{N^2} \geq 0$. Define the Lagrangian function

$$\mathcal{L}^{\mu}(u, \{\lambda\}) = K_{\beta,A}(u) + \mu \left(\hat{u}(0,0) - \bar{z}^{w}\right)^2 + \\ + \sum_{\xi_k \in \Xi} \frac{\lambda_k}{2} ([G * |\Delta_{\Xi}(h * u) - z|^2](\xi_k) - \bar{\sigma}^2).$$

In order to adapt Uzawa's algorithm we need the following result which can be proved as in the proof of Proposition 1.

Theorem 2. *For each $\lambda = (\lambda_k)_{k=1}^{N^2} \geq 0$, there is a unique solution u of*

$$\min_{u \in X_N} \mathcal{L}^{\mu}(u, \{\lambda\}).$$

Proof. Since $\mathcal{L}^{\mu}(u, \{\lambda\})$ is lower semicontinuous in u, it suffices to prove that any minimizing sequence u_n is bounded. Since $\mathcal{L}^{\mu}(u_m, \{\lambda\})$ is bounded, we know that $A(D)(u_m)$ and $(\widehat{u_m}(0,0) - \bar{z}^{w})^2$ are bounded. The boundedness of $A(D)(u_m)$ implies that $u_m - \widehat{u_m}(0,0)$ is bounded. Since $(\widehat{u_m}(0,0) - \bar{z}^{w})^2$ is bounded, then $\widehat{u_m}(0,0)$ is also bounded. □

We solve (3.21) with Uzawa's algorithm.

Algorithm II: Restoration with local constraints

1. Choose any set of values $\lambda_k^0 \geq 0$, $k = 1, \ldots, N^2$, and $\mu^0 > 0$.
   Iterate from $p = 0$ until convergence of λ^p the following steps:

2. With the values of λ^p, μ^p solve:

$$u_p = \arg\min_{u} \mathcal{L}^{\mu^p}(u, \{\lambda^p\}). \tag{3.23}$$

3. Update μ and λ in the following way:

$$\mu^{p+1} = \mu^p + 1,$$

$$\lambda_k^{p+1} = \max(\lambda_k^p + \rho_p([G * |\Delta_{\Xi}(h * u_p) - z|^2](\xi_k) - \bar{\sigma}^2), 0) \quad \forall \xi_k, \tag{3.24}$$

   where $0 < \rho_* \leq \rho_p \leq \rho^*$.

Proposition 3. *Assume that there exists $u \in X_N$ such that $\hat{u}(0,0) = \bar{z}^{w}$ and $z(\xi_k) = h * u(\xi_k)$ $\forall \xi_k \in \Xi$. Then Uzawa's algorithm converges to the solution of (3.21).*

To prove Proposition 3 we need to reformulate problem (3.21) as

$$\min_{u \in X_N} \max_{\lambda \geq 0, \alpha_+, \alpha_- \geq 0} \mathcal{L}(u, \{\lambda\}, \alpha_+, \alpha_-), \tag{3.25}$$

where $\lambda = (\lambda_k)_{k=1}^{N^2}, \alpha_+, \alpha_- \geq 0$,

$$\mathcal{L}(u, \{\lambda\}, \alpha_+, \alpha_-) = K_{\beta,A}(u) + \textstyle\sum_{\xi_k \in \Xi} \lambda_k([G * |\Delta_\Xi(h * u) - z|^2](\xi_k) - \bar{\sigma}^2) + \\ + \alpha_+ \varphi_+(u) + \alpha_- \varphi_-(u),$$

and

$$\varphi_+(u) := \hat{u}(0,0) - \bar{z}^w \qquad \text{and} \qquad \varphi_-(u) = -\hat{u}(0,0) + \bar{z}^w,$$

so that the equality constraint is written as the two inequalities $\varphi_+(u) \leq 0$, $\varphi_-(u) \leq 0$.

Since, by Proposition 1, problem (3.21) has a solution, the classical existence result of saddle points (see [Fau88]) proves the existence of a solution of (3.25). Indeed the following result is classical and can be found, for instance, in [Fau88] (Theorems 4 and 6, pp. 59-61) or [Cia98] (Theorem 9.3.2).

Theorem 4. *Assume that* (3.22) *holds. Let u be the solution of (3.21). Then there are $(\{\lambda\}, \alpha_+, \alpha_-) \geq 0$ such that $(u, \{\lambda\}, \alpha_+, \alpha_-)$ is a solution of (3.25), i. e., a saddle point of $\mathcal{L}(\cdot,\cdot,\cdot,\cdot)$. If $(u, \{\lambda\}, \alpha_+, \alpha_-)$ is a solution of (3.25), then u is a solution of (3.21).*

Since we will need it below, let us compute the gradient of $K_{\beta,A}(u)$. For any $v \in X_N$ we have

$$\begin{aligned}\langle \nabla K_{\beta,A}(u), v\rangle_{X_N} &= \left\langle \frac{A(D)Pu}{\sqrt{\beta^2 + |A(D)Pu|^2}}, A(D)Pv \right\rangle_{X_{2N}} \\ &= \left\langle P^* A(D)^* \left(\frac{A(D)Pu}{\sqrt{\beta^2 + |A(D)Pu|^2}} \right), v \right\rangle_{X_N}\end{aligned}$$

for each $v \in X_N$ vanishing on the boundary of $\{0, \ldots, N-1\}^2$. Thus, we may write

$$\nabla K_{\beta,A}(u) = A(D)^* P^* \left(\frac{A(D)Pu}{\sqrt{\beta^2 + |A(D)Pu|^2}} \right) \in X_N.$$

Now, we notice that $P^* f$ is just the restriction operator (subsampling operator) that considers only the samples of $f \in X_{2N}$ in the grid $\{0, \ldots, N-1\}^2$.

Finally, using this and the last two formulas, we deduce that

$$\frac{A(D)Pu}{\sqrt{\beta^2 + |A(D)Pu|^2}} \cdot \nu^{\{0,\ldots,N-1\}^2} = 0,$$

where $\nu^{\{0,\ldots,N-1\}^2}$ is the discrete normal.

As usual, we denote by $\|v\|_q = \left(\sum_{i,j=1}^{N^2} |v(i,j)|^q\right)^{1/q}$ for any $v \in X_N$, $1 \leq q < \infty$. We denote $\|v\|_\infty = \max_{(i,j)\in\{1,\ldots,N\}} |v(i,j)|$. And for simplicity in the cases

where there is no ambiguity, we shall omit the subindexes for the L_2-norm, then $\|u\|_2$=$\|u\|$.

Proof of Proposition 3. Let us write $Q(u) = (\widehat{u}(0,0) - \overline{z}^w)^2$, $R(u) = G * |\Delta_\Xi(h * u) - z|^2$. To adapt the convergence proof of Uzawa's method to our case, we need to prove that
(a) If U is a bounded subset of X_N then there is a constant $\alpha > 0$ such that

$$\langle \nabla K_{\beta,A}(u) - \nabla K_{\beta,A}(v), u - v\rangle + \mu\langle \nabla Q(u) - \nabla Q(v), u - v\rangle \geq \alpha\|u - v\|^2$$

for all $u, v \in U$.
(b) $R(u)$ is Lipschitz on bounded sets of X_N and
(c) the sequence u_p constructed in Step 2 of the above algorithm is bounded in X_N.
To prove (a) we use the inequality [Tru09]

$$\left\langle \frac{\xi}{\sqrt{\beta^2 + |\xi|^2}} - \frac{\xi'}{\sqrt{\beta^2 + |\xi'|^2}}, \xi - \xi' \right\rangle \geq \beta^2 \frac{|\xi - \xi'|^2}{(\beta^2 + |\xi|^2 + |\xi'|^2)^{3/2}} \quad \forall \xi, \xi' \in \mathbb{R}^k$$

with $k = 2$ and we compute

$$\begin{aligned}\langle \nabla K_{\beta,A}(u) - \nabla K_{\beta,A}(v), u - v\rangle \geq \\ &\geq \sum_{(r,l)} \frac{|A(D)Pu(r,l)) - A(D)Pv(r,l)|^2}{\left(\beta^2 + |A(D)Pu(r,l)|^2 + |A(D)Pv(r,l)|^2\right)^{3/2}} \\ &\geq \alpha \sum_{(r,l)} |A(D)Pu(r,l) - A(D)Pv(r,l)|^2,\end{aligned}$$

where $(r,l) \in \{0,1,\ldots,2N-1\}$, the constant $\alpha > 0$ depends on the bound for U, and

$$\langle \nabla Q(u) - \nabla Q(v), u - v\rangle = 2(\hat{u}(0,0) - \hat{v}(0,0))^2.$$

Then (a) follows as a consequence of the two previous inequalities.
(b) Assume that $U \subseteq X_N$ is a bounded set. Let $u, \overline{u} \in U$. Since $\|G\|_1 \leq 1$, we have

$$\begin{aligned}\|R(u) - R(v)\| &\leq \|G\|_1 \|(h * u - z)^2 - (h * v - z)^2\| \\ &\leq 2\|z\|_\infty \|h * (u - v)\| + \|h * (u + v)\|_\infty \|h * (u - v)\| \\ &\leq C\|u - v\|,\end{aligned}$$

where C is a constant depending on the norms of h and z and on the bound for U.
(c) To prove that $\{u_p\}_p$ is bounded we observe that

$$\mathcal{L}^\mu(u_p, \{\lambda^p\}) \leq \mathcal{L}^\mu(u, \{\lambda^p\}), \qquad \forall u \in X_N, \tag{3.26}$$

for all p. Choosing $u \in X_N$ such that $\hat{u}(0,0) = \bar{z}^w$ and $z = \Delta_\Xi(h * u)$, we obtain that

$$K_{\beta,A}(u_p) + \mu^p Q(u_p) \leq K_{\beta,A}(u),$$

hence $\{u_p\}_p$ is bounded in X_N.

Now, we can adapt the proof of Uzawa's method to our case (see Theorem 5 in [Fau88], Sect. 3.1). Since u_p satisfies (3.26) then $\forall u \in X_N$ we have

$$\langle \nabla K_{\beta,A}(u_p), u - u_p \rangle + \mu^p \langle \nabla Q(u_p), u - u_p \rangle + \langle \lambda^p, R(u) - R(u_p) \rangle \geq 0. \quad (3.27)$$

Let u^* be the solution of problem (3.21). Since, by Theorem 4, we have

$$\mathcal{L}(u^*, \{\lambda\}, \alpha_+, \alpha_-) \leq \mathcal{L}(u, \{\lambda\}, \alpha_+, \alpha_-) \quad \forall u \in X_N,$$

we also have

$$\begin{aligned} \langle \nabla K_{\beta,A}(u^*), u - u^* \rangle + \langle \lambda, R(u) - R(u^*) \rangle + \\ + \alpha_+(\varphi_+(u) - \varphi_+(u^*)) + \alpha_-(\varphi_-(u) - \varphi_-(u^*)) \geq 0 \end{aligned} \quad (3.28)$$

for all $u \in X_N$. Since u^* is a solution of (3.21), we have that

$$\nabla Q(u^*) = 0,$$

and we can add $\mu^p \langle \nabla Q(u^*), u - u^* \rangle$ to the inequality (3.28). Taking $u = u^p$ in this form of the second inequalities, and $u = u^*$ in (3.27) and adding both of them we obtain

$$\begin{aligned} &\langle \nabla K_{\beta,A}(u_p) - \nabla K_{\beta,A}(u^*), u_p - u^* \rangle + \\ &+ \mu^p \langle \nabla Q(u_p) - Q(u^*), u_p - u^* \rangle - \\ &- \alpha_+(\varphi_+(u_p) - \varphi_+(u^*)) - \alpha_-(\varphi_-(u_p) - \varphi_-(u^*)) + \\ &+ \langle \lambda^p - \lambda, R(u_p) - R(u^*) \rangle \leq 0. \end{aligned}$$

Since

$$\langle \nabla K_{\beta,A}(u_p) - \nabla K_{\beta,A}(u^*), u_p - u^* \rangle \geq \alpha \| A(D)Pu_p - A(D)Pu^* \|^2$$

and

$$\begin{aligned} \mu^p \langle \nabla Q(u_p) - Q(u^*), u_p - u^* \rangle - \\ - \alpha_+(\varphi_+(u_p) - \varphi_+(u^*)) - \alpha_-(\varphi_-(u_p) - \varphi_-(u^*)) \\ = 2\mu_p(\widehat{u_p}(0,0) - \widehat{u^*}(0,0))^2 - (\alpha_+ - \alpha_-)(\widehat{u_p}(0,0) - \widehat{u^*}(0,0)) \\ \geq \mu^p(\widehat{u_p}(0,0) - \widehat{u^*}(0,0))^2 \end{aligned}$$

for p large enough, we have

$$\begin{aligned} \langle \lambda^p - \lambda, R(u_p) - R(u^*) \rangle \leq \\ \leq -\alpha \| A(D)Pu_p - A(D)Pu^* \|^2 - \mu^p(\widehat{u_p}(0,0) - \widehat{u^*}(0,0))^2 \\ \leq -\alpha_0 \| u_p - u^* \|. \end{aligned} \quad (3.29)$$

Now, the proof follows in a standard way. Let us give the details for the sake of completeness. Using (3.24), we have

$$\|\lambda^{p+1} - \lambda\|_{\leq} \|\lambda^p - \lambda + \rho_p(R(u_p) - R(u^*))\|.$$

Taking squares, we have

$$\|\lambda^{p+1} - \lambda\|^2 \leq \|\lambda^p - \lambda\|^2 + 2\rho_p\langle\lambda^p - \lambda, R(u_p) - R(u^*)\rangle + \rho_p^2\|R(u_p) - R(u)\|^2.$$

Using (3.29) and (b), we have

$$\|\lambda^{p+1} - \lambda\|^2 \leq \|\lambda^p - \lambda\|^2 - 2\alpha_0\rho_p\|u_p - u^*\|^2 + \rho_p^2 L^2\|u_p - u^*\|^2.$$

for some $L > 0$. If we choose ρ_p such that

$$2\alpha_0\rho_p - L^2\rho_p^2 \geq \gamma > 0,$$

that is, $0 < \rho_* \leq \rho_p \leq \rho^*$, we have

$$\|\lambda^{p+1} - \lambda\|^2 \leq \|\lambda^p - \lambda\|^2 - \gamma\|u_p - u^*\|^2.$$

Then we deduce that $\|\lambda^p - \lambda\|$ is decreasing and, thus, has a limit $\ell \geq 0$. Then letting $p \to \infty$ we have that $\|u_p - u^*\| \to 0$.

3.6 A Quasi-Newton algorithm for the solution of (3.23)

The purpose of this section is to explain the algorithm used to solve problem (3.23) in Algorithm II. For convenience, let us denote the convolution and irregular sampling operators, as $\mathcal{S}u = \Delta_\Xi(h * u)$ for any $u \in X_N$.

Observe that the Euler-Lagrange equation corresponding to (3.23) is

$$A(D)^T\left(P^*\frac{A(D)Pu}{\sqrt{\beta^2 + |A(D)Pu|^2}}\right) + 2\mu(\hat{u}(0,0) - \overline{z}^w) + \mathcal{S}^*(G * \lambda)(\mathcal{S}u - z) = 0. \tag{3.30}$$

To shorten our expressions, let us define the following operators:

$$\mathcal{M}u = (\sqrt{G * \lambda})\mathcal{S}u, \qquad \text{so that} \qquad \mathcal{M}^*u = \mathcal{S}^*(\sqrt{G * \lambda}\,u),$$

$$\mathcal{N}u = \hat{u}(0,0), \qquad b = \mathcal{S}^*((G * \lambda)z) + 2\mu\overline{z}^w,$$

$$\mathcal{A}[v](u) = A(D)^T P^*\left(\frac{A(D)Pu}{\sqrt{\beta^2 + |A(D)Pv|^2}}\right)$$

and

$$\mathcal{T}[v](u) = \mathcal{A}[v](u) + 2\mu\mathcal{N}u + \mathcal{M}^*\mathcal{M}u.$$

where $u, v \in X_N$.

We want to solve (3.30) with a fixed point iteration:

$$\mathcal{A}[u^t](u^{t+1}) + 2\mu\mathcal{N}u^{t+1} + \mathcal{M}^*\mathcal{M}u^{t+1} = b. \tag{3.31}$$

The rest of this section is devoted to show that such a fixed point algorithm converges to the minimizer of (3.23). Notice that a proof based on convex analysis (the half-quadratic regularization approach) can be found in [CL97] in the continuous case, or in [CBFAB97, AK06] for the discrete case. Further analysis can be found in [AV94, DV97, VO96]. Here, we will extend the proof proposed by Chan and Mulet in [CM99]. The advantage of such an approach is that only basic algebra is needed. Moreover, the linear convergence rate of this algorithm can be shown explicitly. The difference with the approach in [CM99] relies on the fact that the operator $\mathcal{M}$ is not assumed to be invertible. In our case the presence of the mean constraint $\mu\mathcal{N}$ permits to prove the same convergence result as in [CM99], without an invertibility hypothesis on $\mathcal{M}$.

Remark 5. Note that computing $\mathcal{M}^*\mathcal{M}$ in (3.31) entails the computation of an operator with a Toeplitz structure. As in the ACT algorithm (Section 3.2.1) this is efficiently computed in $\mathcal{O}(N^2 \log_2 N^2)$ steps.

Remark 6. As mentioned in Remark 3, we may write $A(D)u = D\Psi u$. This permits to use the change of variables $v = \Psi u$ and write the constrained restoration problem (3.21) in terms of v. Then the regularizer coincides with the total variation applied to v. The solution of (3.21) can be recovered as $u = \Psi^{-1}v$. This has no effect on the developments of this chapter so we keep the notation $A(D)u$.

3.6.1 Existence of u^t and its boundedness

The sequence u^t will be defined iteratively using (3.31).

Proposition 5. *The equation (3.31) has a unique solution $u^{t+1} \in X_N$ which is the minimizer of*

$$\mathcal{E}(u) = \left\| \frac{A(D)Pu}{(\beta^2 + |A(D)Pu^t|^2)^{1/4}} \right\|^2 + \frac{1}{2}\|\mathcal{M}u - z'\|^2 + \mu\|\mathcal{N}u - \overline{z}^w\|^2, \tag{3.32}$$

where $\left\| \frac{A(D)Pu}{(\beta^2+|A(D)Pu^t|^2)^{1/4}} \right\|^2 = \sum_{(r,l)} \left| \frac{A(D)Pu(r,l)}{(\beta^2+|A(D)Pu^t(r,l)|^2)^{1/4}} \right|$ *and* $z' = \sqrt{G * \lambda}\, z$.

Proof. It is standard that (3.32) admits a unique solution $u^{t+1} \in X_N$. Moreover, (3.31) is the Euler-Lagrange equation associated to (3.32) and solutions of (3.31) are minimizers of (3.32). □

Proposition 6. (i) *There exists $K_0 > 0$ such that*

$$\|A(D)Pu^t\| \leq K_0.$$

(ii) $\mathcal{T}[u^t]$ is a bounded coercive operator. Indeed we have

$$\langle \mathcal{T}[u^t]u, u\rangle \geq \alpha \|u\|^2 \tag{3.33}$$

for some $\alpha > 0$ independent of t.
(iii) The sequence u^t is uniformly bounded.

Proof. Since u^{t+1} is a minimizer of $\mathcal{E}(u)$, we have $\mathcal{E}(u^{t+1}) \leq \mathcal{E}(0) = \frac{1}{2}\|z'\|^2 + \mu\|\bar{z}^w\|^2$, and thus:

$$\left\| \frac{1}{(\beta^2 + |A(D)Pu^t|^2)^{1/4}} A(D)Pu^{t+1} \right\|^2 \leq \mathcal{E}(0) = \frac{1}{2}\|z'\|^2 + \mu\|\bar{z}^w\|^2.$$

We have

$$\left| \frac{A(D)Pu^{t+1}}{(\beta^2 + |A(D)Pu^t|^2)^{1/4}} \right|^2 \geq \frac{\left|A(D)Pu^{t+1}\right|^2}{\|\sqrt{\beta^2 + |A(D)Pu^t|^2}\|_\infty}$$

Thus

$$\begin{aligned} \left\|A(D)Pu^{t+1}\right\|^2 &\leq \mathcal{E}(0) \left\| \sqrt{\beta^2 + |A(D)Pu^t|^2} \right\|_\infty \\ &\leq \mathcal{E}(0)\sqrt{\beta^2 + \|A(D)Pu^t\|_\infty^2}. \end{aligned}$$

But since we deal with finite dimensional spaces, there exists $L > 0$ which does not depend on u^t such that $\|A(D)Pu^t\|_\infty \leq L\|A(D)Pu^t\|$. Hence we deduce that

$$\left\|A(D)Pu^{t+1}\right\|^2 \leq \mathcal{E}(0)\sqrt{\beta^2 + L^2\|A(D)Pu^t\|^2}.$$

Assume that $\left\|A(D)Pu^t\right\| \leq K$. Using (3.6.1), to get that $\left\|A(D)Pu^{t+1}\right\| \leq K$, it is sufficient to choose $K > 0$ large enough so that

$$\mathcal{E}(0)\sqrt{\beta^2 + L^2K^2} \leq K^2.$$

(ii) The boundedness of $\mathcal{T}[u^t]$ is immediate and we omit its proof. Let us prove that $\mathcal{T}_t$ is a coercive operator. Using the bounds in Step *(i)*, we have

$$\left\langle \frac{1}{\sqrt{\beta^2 + |A(D)Pu^t|^2}} A(D)Pu, A(D)Pu \right\rangle \geq \frac{1}{\sqrt{\beta^2 + L^2K_0^2}} \|A(D)Pu\|^2.$$

Hence

$$\langle \mathcal{A}[u^t]u, u\rangle \geq \frac{1}{\sqrt{\beta^2 + L^2K_0^2}} \|A(D)Pu\|^2 \geq \alpha_0\|u_0\|^2,$$

where we wrote $u = u_0 + c$, with $c = \mathcal{N}u$ and $\mathcal{N}u_0 = 0$, and we used the fact that $A(D)P$ is a linear operator whose kernel are the constants. Using $\mathcal{A}[u^t]c = 0$ and $\langle u_0, c\rangle = 0$, we get

$$\begin{aligned}\langle \mathcal{T}[u^t](u_0+c), u_0+c\rangle &\geq \langle \mathcal{A}[u^t](u_0+c), u_0+c\rangle + 2\mu\langle \mathcal{N}(u_0+c), u_0+c\rangle \\ &\geq \langle \mathcal{A}[u^t]u_0, u_0\rangle + 2\mu c^2 \\ &\geq \alpha_0\|u_0\|^2 + 2\mu c^2.\end{aligned}$$

Thus, we deduce (3.33).
(iii) From (3.33), we know that $\langle \mathcal{T}[u^t]u^{t+1}, u^{t+1}\rangle \geq \alpha\|u^{t+1}\|^2$. But from (3.30), we know that $\langle \mathcal{T}[u^t]u^{t+1}, u^{t+1}\rangle = \langle b, u^{t+1}\rangle \leq \|b\|\|u^{t+1}\|$. We deduce that $\|u^{t+1}\| \leq \frac{\|b\|}{\alpha}$. □

3.6.2 Convergence of the fixed point algorithm

For simplicity, given $\lambda = (\lambda_k)_{k=1}^{N^2}$, we write $\mathcal{L}(u) = \mathcal{L}^\mu(u, \{\lambda\})$. Recall that

$$\nabla_u \mathcal{L}(u) = \mathcal{T}[u](u) - 2\mu\overline{z}^w - \mathcal{M}^* z'.$$

Let us finally define

$$\mathcal{G}(v,u) = \mathcal{L}(u) + \langle v-u, \nabla_u\mathcal{L}(u)\rangle + \frac{1}{2}\langle v-u, \mathcal{T}[u](v-u)\rangle.$$

Proposition 7. *The following inequality holds for any $u, v \in X_N$:*

$$\mathcal{L}(v) \leq \mathcal{G}(v,u). \tag{3.34}$$

Proof. We follow the proof in [CM99]. Since

$$\mathcal{G}(v,u) - \mathcal{L}(v) = \mathcal{L}(u) - \mathcal{L}(v) + \langle v-u, \nabla_u\mathcal{L}(u)\rangle + \frac{1}{2}\langle v-u, \mathcal{T}[u](v-u)\rangle,$$

standard computations lead to

$$\mathcal{G}(v,u) - \mathcal{L}(v) = \sum_{(r,l)} \left(a - \overline{a} + \frac{1}{2a}(\overline{a}^2 - a^2)\right)$$

with $a = \sqrt{\beta^2 + |A(D)Pu(r,l)|^2}$ and $\overline{a} = \sqrt{\beta^2 + |A(D)Pv(r,l)|^2}$, where $(r,l) \in \{0,1,\ldots,2N-1\}$. Since $a, \overline{a} > 0$, and $a - \overline{a} + \frac{1}{2a}(\overline{a}^2 - a^2) = \frac{(a-\overline{a})^2}{2a} \geq 0$, we have that $\mathcal{G}(v,u) - \mathcal{L}(v) \geq 0$. □

Proposition 8. (i) *The function u^{t+1} defined by (3.31) is such that:*

$$u^{t+1} = \mathit{argmin}_v \mathcal{G}(v, u^t) \tag{3.35}$$

i.e.: $0 = \nabla_u\mathcal{L}(u^t) + \mathcal{T}[u^t](u^{t+1} - u^t)$.
(ii) *We have* $\lim_{t\to+\infty}\|u^{t+1} - u^t\| = 0$.

Proof. (i) Let us denote by $\bar{u} = \text{argmin}_v \mathcal{G}(v, u^t)$. We thus have $0 = \nabla_u \mathcal{L}(u^t) + \mathcal{T}[u^t](\bar{u} - u^t)$. And this last equation is precisely equation (3.31), which implies that $\bar{u} = u^{t+1}$.
(ii) From (3.34) and (3.35), we have $\mathcal{L}(u^{t+1}) \leq \mathcal{G}(u^{t+1}, u^t) \leq \mathcal{G}(u^t, u^t) \leq \mathcal{L}(u^t)$, i.e. $(\mathcal{L}(u^t))$ is decreasing. Now, from (3.34) and (3.35), we have:

$$\begin{aligned}\mathcal{L}(u^{t+1}) &\leq \mathcal{G}(u^{t+1}, u^t) \\ &= \mathcal{L}(u^t) + \langle u^{t+1} - u^t, \nabla_u \mathcal{L}(u^t)\rangle + \frac{1}{2}\langle u^{t+1} - u^t, \mathcal{T}[u^t](u^{t+1} - u^t)\rangle \\ &= \mathcal{L}(u^t) - \frac{1}{2}\langle u^{t+1} - u^t, \mathcal{T}[u^t](u^{t+1} - u^t)\rangle\end{aligned}$$

Using (3.33), we deduce: $\frac{1}{2}\alpha\|u^{t+1} - u^t\|^2 \leq \frac{1}{2}\langle u^{t+1} - u^t, \mathcal{T}[u^t](u^{t+1} - u^t)\rangle \leq \mathcal{L}(u^t) - \mathcal{L}(u^{t+1})$ and (ii) follows. □

We are now in position to state a convergence result.

Theorem 9. *The sequence u^t defined by (3.31) converges to the solution of (3.23).*

Proof. From Proposition 6.(iii), we know that u^t is uniformly bounded. There exists u such that we can extract a convergent subsequence, which we still denote by u^t, with $u^t \to u$ as $t \to +\infty$. From Proposition 8.(ii) , we know that u^{t+1} is also convergent and $u^{t+1} \to u$ as $t \to +\infty$. Letting $t \to +\infty$ in equation (3.31), we deduce that u is the solution of (3.30) (and thus of (3.23)). By uniqueness of the solution of (3.23), we conclude that the whole sequence u^t converges to u, solution of (3.23). □

We end this section by stating a result about the convergence rate. We denote by $\tilde{u}$ the solution of Problem (3.23). We use the following notations:

$$\gamma_t := \frac{\mathcal{G}(\tilde{u}, u^t) - \mathcal{L}(\tilde{u})}{\frac{1}{2}\langle \tilde{u} - u^t, \mathcal{T}[u^t](\tilde{u} - u^t)\rangle}$$

and

$$\zeta := 1 - \lambda_{\min}(\mathcal{T}[\tilde{u}]^{-1}\nabla_u^2 \mathcal{L}(\tilde{u}))$$

where $\lambda_{\min}(M)$ denotes the smallest eigenvalue of the matrix M; in particular if M is positive definite then $\lambda_{\min}(M) > 0$

Proposition 10.
1. $\mathcal{L}(u^{t+1}) - \mathcal{L}(\tilde{u}) \leq \gamma_t(\mathcal{L}(u^t) - \mathcal{L}(\tilde{u}))$.
2. *$\zeta < 1$ and $0 \leq \gamma_t \leq \zeta$, for t sufficiently large. In particular, $\mathcal{L}(u^t)$ has a linear convergence rate of at most ζ.*
3. *u^t is r-linearly convergent with a convergence rate of at most $\sqrt{\zeta}$.*

Proof. We refer the interested reader to the proof of Theorem 6.1 in [CM99] which can easily be extended to our case. □

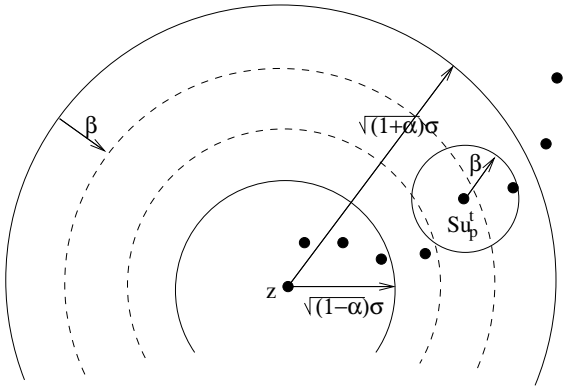


Figure 3.1: The relaxed optimization problem allows solutions within a band. From the depicted sequence of solutions u_p^t (black dots) the ones that are truly satisfying the constraints are those inside the dotted ring, since it accounts for the approximation error η. $\mathcal{S}$ stands for the irregular sampling and convolution operators, Su_p^t is also a sequence defined over the irregular samples and not over the solution space.

3.7 Band constraints and stopping conditions

Coming back to the optimization problem (3.21), since both the functional $K_{\beta,A}(u)$ and the constraints are convex, and the constraint's feasible set V does not contain the absolute minimun of $K_{\beta,A}(u)$, then the solution lies in the boundary of V. In practice, computing a solution in ∂V is not only computationally too expensive (due to the size of the problems), but also unnecessary because of the noise. Since we rely on noise estimates that have a certain accuracy, exceeding this accuracy in the data fitting is useless (as we will see later in this section). Moreover, as it has been observed in all numerical experiments [BFCAB95, Cha04, CL97, CZC95, CM99, CBFAB97, DMR00, GR92, AG06, GSZ06, ROF92, VO96], using total variation as regularizer in denoising or restoration generally carries some loss of texture and it is not desirable to compute the solution that (absolutely) minimizes the TV but to keep a solution with a slightly higher TV value in order to avoid the loss of textures.

As a consequence, to avoid this degradation, the rule of thumb has been ever since to remove less noise than noise is actually present in the image. Gilboa studied this in [AG06] and concluded that in terms of SNR the optimal selection of $\bar{\sigma}$ is between 0.7 and 0.8 times the value of σ. In what follows we will modify the constraints to account for this change

$$\left[G * \left|\Delta_\Xi (h * u) - z\right|^2\right](\xi_k) \leq \bar{\sigma}^2 \quad \forall \xi_k \in \Xi\,,$$

where $\bar{\sigma} < \sigma$. That is, in order to keep more texture in u we do not remove all noise and we could rewrite the acquisition model as $z = \mathcal{S}(u_0 + \mathcal{S}^{-1}(n - \bar{n})) + \bar{n}$

where $\bar{n}$ is a noise with variance $\bar{\sigma}^2$ and we identify $u_0 + \mathcal{S}^{-1}(n - \bar{n})$ with u.

Motivated by this observation and by the fact that, due to noise, there is always some uncertainty in the surroundings of ∂V, we will avoid the computational overhead of getting exactly to ∂V by stopping the algorithm as soon as the solution is close to it. This is the notion behind the *band constraint*

$$(1-\alpha)\bar{\sigma}^2 \leq \left[G * |\Delta_\Xi (h * u) - z|^2\right](\xi_k) \leq (1+\alpha)\bar{\sigma}^2 \quad \forall \xi_k \in \Xi, \tag{3.36}$$

with $\alpha > 0$. The constraint described by equation (3.36) is clearly non-convex, and therefore it cannot be integrated in the method presented here. But since Uzawa's algorithm always pushes the solution towards the boundary of the feasible set, then (at least in practice) it can be used to stop Uzawa's loop by testing if (3.36) is fulfilled.

In what follows we will see that even considering relaxed constraints like (3.36), imposing all of the local constraints simultaneously is not statistically correct, since each constraint's estimator behaves as a random variable. Then we will see how this relaxation of constraints is used to early stop Uzawa's iterations and how this helps to improve the efficiency of our implementation.

In our experiments, we have chosen $\bar{\sigma} = 0.8\sigma$ and α such that $0.8(1+\alpha) < 1$.

3.7.1 Expected number of satisfied *band* constraints

Let us summarize the arguments of [ABCH08] and adapt them to the case of band constraints. Each local constraint relies on a local estimate of the residual variance of the form

$$S_G(\xi_k) = \left[G * |\Delta_\Xi (h * u) - z|^2\right](\xi_k) = \left[G * |\bar{n}|^2\right](\xi_k), \tag{3.37}$$

where G is a Gaussian or uniform window centered at the interest point and n_k denotes a zero mean Gaussian noise and variance $\bar{\sigma}^2$ (recall that we are going to remove only a noise of variance $\bar{\sigma}^2 < \sigma^2$). Since S_G is a random variable itself, the number of satisfied constraints is estimated by the probability $P[S_G \leq (1+\alpha)\bar{\sigma}^2]$ in the case of the ball shaped constraint, or $P[(1-\alpha)\bar{\sigma}^2 \leq S_G \leq (1+\alpha)\bar{\sigma}^2]$ for the band constraint. Observe that a constraint of the type $S_G = \bar{\sigma}^2$ (that is in practice imposed when solving exactly (3.21)) was already doomed to failure since it has zero probability to occur $P[S_G = \bar{\sigma}^2] = 0$.

Using the Central Limit Theorem gives only a loose estimate of the probability distribution of S_G. To improve the estimation of the expected number of satisfied constraints let us simplify S_G. By approximating the discrete convolution with G (of standard deviation $\tilde{r}$) by a mean over a disk I of radius $r = 2\tilde{r}$ we can define a simpler estimator $S_I = \frac{1}{|I|}\sum_{k\in I} \bar{n}_k^2$.

Then, for the case of the ball constraint, the expected number of satisfied constraints is the number of pixels times the following probability

$$P[S_I \leq (1+\alpha)\bar{\sigma}^2] = P\left[\frac{1}{|I|}\sum_{k\in I}\bar{n}_k^2 \leq (1+\alpha)\bar{\sigma}^2\right] = P\left[\sum_{k\in I}\left(\frac{\bar{n}_k}{\bar{\sigma}}\right)^2 \leq (1+\alpha)|I|\right].$$

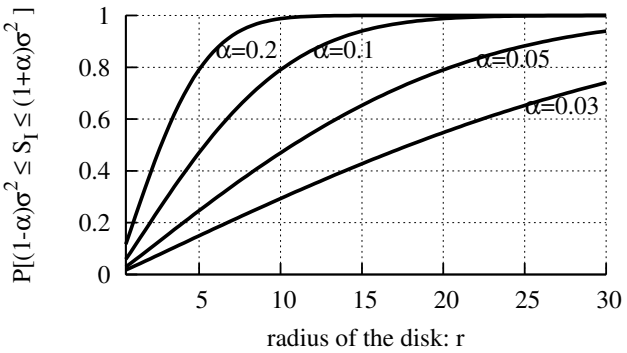


Figure 3.2: Relation between the expected number of satisfied constraints and the radius r of the disk used for the local noise estimation. The graph shows the curves computed for different widths of the band α.

Notice that in the rightmost equation $\sum_{k\in I}\left(\frac{\bar{n}_k}{\bar{\sigma}}\right)^2$ is a sum of $|I|$ squared normalized Gaussian random variables, so it follows a chi-square distribution with $|I|$ degrees of freedom ($\chi^2(|I|)$). And the probability can be computed using the incomplete gamma function $\Gamma(a,x)=\int_x^\infty t^{a-1}e^{-t}dt$

$$P[S_I \leq (1+\alpha)\bar{\sigma}^2] = P\left[\chi^2(|I|) \leq (1+\alpha)|I|\right] = \Gamma\left(\frac{(1+\alpha)|I|}{2}, \frac{|I|}{2}\right).$$

In the case of the band constraint the expected number of satisfied constraints $N(\alpha, r)$ is the number of pixels times the following probability

$$P\left[(1-\alpha)\bar{\sigma}^2 \leq S_I \leq (1+\alpha)\bar{\sigma}^2\right] = \Gamma\left(\frac{(1+\alpha)|I|}{2}, \frac{|I|}{2}\right) - \Gamma\left(\frac{(1-\alpha)|I|}{2}, \frac{|I|}{2}\right). \tag{3.38}$$

Equation (3.38) expresses the expected proportion of satisfied constraints as a function of the radius of the disk r ($|I| = \pi r^2$) and the width of the band α. We plot in Figure 3.2 this function, for different values of α to give an intuition of its behavior. Notice that the expected number of satisfied constraints decreases as we reduce the band width α or the radius r. This permits to determine one parameter as a function of the other two, i.e. using a disk of radius $r = 13$ (or a Gaussian with standard deviation 7.0) and defining a band of width $0.2\bar{\sigma}^2$ ($\alpha = 0.1$) gives 89% of satisfied constraints. In practice, either we specify α, r and then the expected number of satisfied constraints is $N(\alpha, r)$, or we give α and N_α and we compute r so that $N(\alpha, r) = N_\alpha$. We have taken the second option in the experiments displayed in Section 3.8.

Remark 7. Observe that in (3.37) the estimation of the noise variance corresponds to the case when the mean of the random variable is zero. Indeed, the global mean is enforced to be zero in (3.21). We should also impose that the local means are zero with a new set of constraints, otherwise

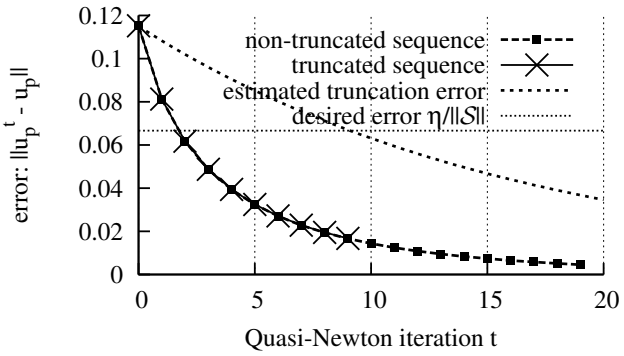


Figure 3.3: Truncation of the Quasi-Newton method. The first sequence shows the error evolution along the Quasi-Newton iterations, without applying the stopping condition for the CG algorithm. The second sequence was obtained stopping the CG with the empirical bound, observe that the two sequences are indistinguishable. The third sequence depicts the estimated error used to effectively stop the Quasi-Newton iterations, as soon as the desired error is achieved.

Table 3.1: Relation between the width of the constraint band and the restoration time. Reducing the width of the band increases the computational cost of the restoration algorithm. In the table the value of η was selected according to the rule $\eta = \frac{1}{3}(\sqrt{1+\alpha}\,\bar{\sigma} - \sqrt{1-\alpha}\,\bar{\sigma})$, and $\bar{\sigma} = 1$. All the reported times correspond to experiments ran on a 1.6Ghz CPU restoring a 256×256 pixels image.

Band Width parameter: α	**Effective Band** $\sqrt{(1-\alpha)}\,\bar{\sigma} + \eta$	$\sqrt{(1+\alpha)}\,\bar{\sigma} - \eta$	**Total running time**
0.60	0.84	1.05	3 min 3 sec
0.34	0.93	1.04	3 min 37 sec
0.22	0.96	1.03	4 min 58 sec
0.10	0.98	1.02	18 min

S_G will be an overestimation of the noise variance. The local mean constraint $\sum_{\xi_k \in \Xi} |G * (\Delta_\Xi(h * u) - z)|^2(\xi_k) = 0$ can be added in (3.21), and the formalism developed in this chapter fits the new formulation. But to avoid the computational overhead of its implementation, and since the overestimation plays in favor of the relaxation arguments presented earlier in this section, we will not include it in the present formulation.

3.7.2 Practical stopping conditions for an efficient implementation

Using (3.38) we may derive a practical rule to stop Uzawa's loop. Indeed, the user specifies α and the proportion of constraints N_α / N that must lie within the band of width α, and the algorithm deduces the radius r of the kernel G such

that $N(\alpha, r) = N_\alpha$. Then we iterate the Uzawa's loop in Algorithm II until the number of pixels that satisfy the constraint (3.38) is at least N_α.

The truncation error of the Quasi-Newton has three sources: *(i)* truncation of the Quasi-Newton iterations themselves, *(ii)* truncation of the nested CG loop, and *(iii)* propagation of the CG error along QN iterations. Here we summarize how to estimate and control the combination of the three errors for a given (global) target error bound on the QN result $\|u_p - u_p^t\|^2 \leq \eta / \|\mathcal{S}\|$. Using standard error propagation analysis [DB03] and the knowledge that QN is at least linearly convergent we can estimate the global error determining *(ii)* and bound the inverses of the operators $\mathcal{T}[u^t]$ and their dependence on u^t. In our case these bounds are estimated empirically, and the CG error is approximated by its residual. Figure 3.3 shows that this procedure is quite effective in practice. First our CG stopping condition makes the truncated QN sequence indistinguishable from the non-truncated QN sequence (i.e. the one with CG iterated until exact convergence is reached, and thus not affected by CG truncation errors *(ii)* and their propagation *(iii)*). This is because the actual QN truncation error (with either CG truncated or not) is considerably over-estimated by our error propagation analysis as shown in Figure 3.3. This also means that the desired error is achieved much faster than predicted by the our error bounds.

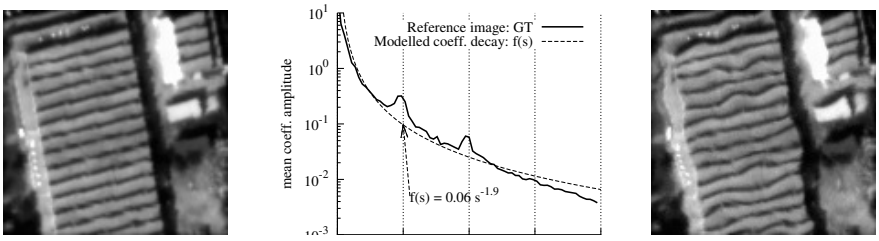
Let us now take into account the truncation error of the Quasi-Newton method in the determination of the band constraints. Assume that we are computing the Quasi-Newton solution u_p and that we have controlled the errors $\|\mathcal{S}u_p - \mathcal{S}u_p^t\| \leq \|\mathcal{S}\|\|u_p - u_p^t\| \leq \eta$, where u_p^t denotes the solution obtained at the t-th iteration of the method. Then we may erode the band by η as seen in Figure 3.1. In this way we ensure that if we stop the Quasi-Newton solution with the criterion $\|\mathcal{S}\|\|u_p - u_p^t\| \leq \eta$, then the truncated solution u_p satisfies (3.38). Clearly η must satisfy the inequality $\eta < \frac{1}{2}(\sqrt{1+\alpha}\sigma - \sqrt{1-\alpha}\sigma)$ and we have taken $\eta := \frac{1}{3}(\sqrt{1+\alpha}\sigma - \sqrt{1-\alpha}\sigma)$, which equalizes the widths of the three bands in Figure 3.1. Figure 3.1 illustrates the band and its reduced version for a single constraint.

Finally, we notice that the computational complexity of the algorithm increases as we reduce the width of the band. Indeed, taking $\alpha \to 0$ makes it harder to satisfy the constraints. And to illustrate this we display in Table 3.1 the computation times corresponding to different values of α.

3.8 Experiments

We will test the proposed algorithm in three contexts, first in the irregular to regular sampling and denoising task, then in Subsection 3.8.2 we add the deconvolution, and finally in Subsection 3.8.3 we consider the full restoration problem, with deconvolution, denoising, and zooming.

We compare the results obtained by the ACT Algorithm [GS00] (when applicable), the algorithm described in [ACHR06] (ACT+TV), and the proposed algorithm discussing different choices of $A(D)$ in each context. In addition we will also consider a re-formulation of our local constraint algorithm based on



(a) Reference Image. (b) Mean Fourier coefficient amplitude for the reference image. (c) Noisy and irregularly sampled image.

Figure 3.4: Image used in the denoising experiment. At left: the reference image is non-aliased, and it has 149×149 pixels. It was multiplied by a smooth window vanishing on the borders in order to avoid periodization artifacts (not shown). In this image the mean Fourier coefficient amplitudes decay like $(2\pi\omega/N)^{-1.9}$ (center). At right, is shown the perturbed sampled image. The perturbations have an amplitude of $A = 0.88$ pixels (standard deviation of $\varepsilon(x)$), and were simulated according to (3.8) as colored noise with the spectral content inside $[-0.5/T_\varepsilon, 0.5/T_\varepsilon]^2$ for $T_\varepsilon = 10$.

the minimization of the L^2-norm of the pseudo-differential operator $A(D)$. The L^2-norm formulation is faster and leads to a linear system (which is solved directly with Conjugate Gradient, avoiding the need of a fixed point loop), and it is derived by replacing the regularizer $K_{A,\beta}$ in (3.21) by

$$\mathrm{L}_A(u) := \sum_{0 \le r,l \le 2N} |A(D)Pu(r,l)|^2. \tag{3.39}$$

All the experiments were performed with simulated images. The perturbations $\varepsilon(x)$ were generated according to the model (3.8), with an amplitude $A = 0.88$ (standard deviations of $\varepsilon(x)$), and where supp $\hat{\varepsilon} \subseteq [-\frac{0.5}{T_\varepsilon}, \frac{0.5}{T_\varepsilon}]^2$ for $T_\varepsilon = 10$. The perturbed samples z were computed very accurately (usually 10^{-8}) by approximating the irregular sampling formula (3.9) with the transposed NFFT [PST01]. Finally, we added a white noise of standard deviation σ to the irregular samples. In the experiment displayed here, we have taken $\sigma = 1$ gray levels (i.e.the noise power is 890 times smaller than the image power, $SNR = 29.5dB$). The perturbed image shown in Figure 3.4, which corresponds to the denoising experiment, was simulated according to this procedure.

To quantify the errors we adopt the classical root mean squared error measure $\mathrm{RMSE}(u, u_0) = (\frac{1}{N}\|u - u_0\|^2)^{1/2}$ against the ground truth image (denoted as u_0), and the *method noise* for a qualitative analysis. The method noise was originally aimed at comparing denoising algorithms. It consists in subtracting the restored image to the noisy one, and studying the remaining noise. In our context assuming an image formation model like (3.1) and denoting u the image

Table 3.2: Comparison of the algorithms in the irregular to regular sampling and denosing task. These results correspond to the restoration of the image shown in Figure 3.4, corrupted by a white Gaussian noise with standard deviation $\sigma = 1$. The error column is obtained by comparing the restored image u with the ground truth u_0, where $\text{RMSE}(u, u_0) = (\frac{1}{N}\|u - u_0\|^2)^{1/2}$. Its values evidence that all the algorithms achieve errors similar to the noise variance. In all the experiments the power of the removed noise was set to be $\frac{1}{N}\|\mathcal{S}u - z\|^2 \sim 0.908$, a little below the noise level $\sigma^2 = 1$, but this allows to improve both the result's RMSE and the visual quality of the restored images.

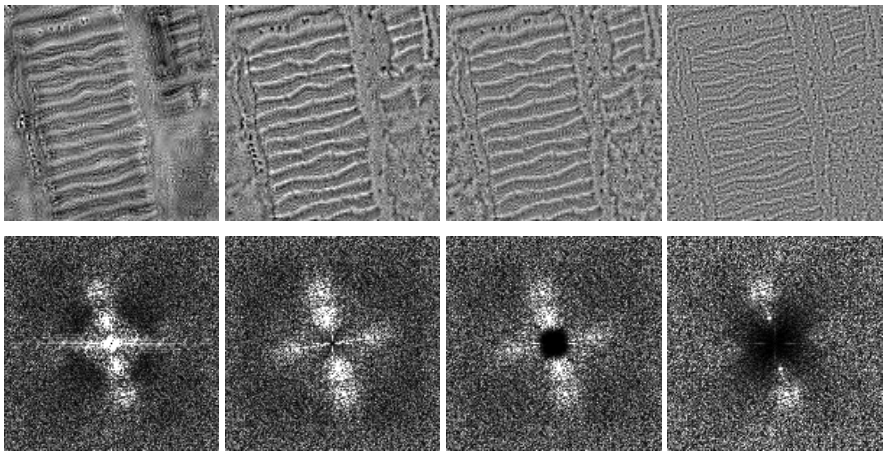
Algorithm	**Regularizer**	**RMSE**
ACT [GS00] or (3.11)		1.354
ACT_D (3.14) residual preconditioning		1.121
ACT_R (3.15) regularity preconditioning		1.049
ACT+TV [ACHR06] or (3.16)		0.961
QN+TV $\lvert A(i\omega)\rvert = \lvert\frac{2\pi}{N}\omega\rvert$	$K_{\beta,A}(u)$ eq. (3.21)	0.874
QN+FAR $\lvert A(i\omega)\rvert = \lvert\frac{2\pi}{N}\omega\rvert^{1.6}$		0.776
QN+FAR $\lvert A(i\omega)\rvert = \lvert\frac{2\pi}{N}\omega\rvert^{1.9}$		0.757
L^2-norm of the gradient	$\text{L}_A(u)$ eq. (3.39)	0.880
L^2-norm of $A(D)u$ with $\lvert A(i\omega)\rvert = \lvert\frac{2\pi}{N}\omega\rvert^{1.9}$		0.773

obtained by a restoration algorithm, the method noise becomes $(z - \mathcal{S}u)$, where z are the noisy samples and $\mathcal{S}$ stands for the irregular sampling and convolution operators. Since the restoration is expected to recover the original image $u \simeq u_0$, the method noise should be as similar to a white noise as possible. In addition, since we would like the original image u_0 not to be altered by the restoration method, the method noise should not exceed the actual noise variance, which justifies our local constraints approach.

3.8.1 Denoising

Observe in Table 3.2 that the proposed algorithm outperforms (in terms of RMSE) the ACT and ACT+TV [ABCH08] algorithms, in the denoising experiment. Also notice in Figure 3.5 how the method noise of the proposed algorithm retains less structure, meaning that the method removes just the noise with less alteration of the texture.

The variants of the algorithm based on the minimization of an L^2-norm (3.39), give results that are comparable with the ones obtained with the FAR regularizer, and in both reported cases the L^2 results are indistinguishable from the ones obtained by FAR. This observation advocates for the local formulation of the constraints and it can be justified by the density of the samples in this denosing experiment, and by the lack of both deconvolution and zoom. We can conclude that, for the denoising (without either deconvolution nor zoom) the advantage of a nonlinear regularizer over the L^2 one is marginal, mostly because no spectral extrapolation is needed. However minimizing the L^2-norm



(a) ACT_R eq. (3.15) regularizer precondition RMSE= 1.049

(b) ACT+TV eq. (3.16) RMSE= 0.961

(c) QN+TV eq. (3.21) with $|A(i\omega)| \sim |\frac{2\pi\omega}{N}|^1$ RMSE= 0.864

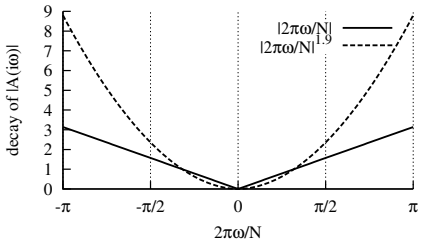
(d) QN+FAR eq. (3.21) with $|A(i\omega)| \sim |\frac{2\pi\omega}{N}|^{1.9}$ RMSE= 0.757

Figure 3.5: Method noise of the different algorithms in the experiment of irregular to regular sampling plus denoising. The images in the first row display the method noise $(z - \mathcal{S}u)$ for different methods, less visible structure indicates a better reconstruction. The images display the grayscale range $[-3, 3]$ scaled to $[0, 255]$ (the full grayscale range of the image z is $[0, 155]$). In the second row we show the corresponding Fourier transforms, the spectrum highlights the structures that are barely visible in the spatial domain.

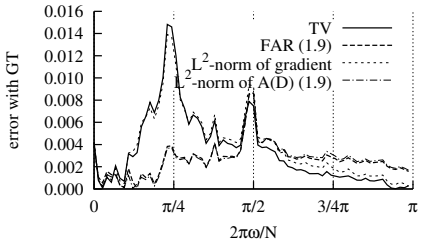
is not expected to perform well in tasks that entail a spectral extrapolation like deconvolving or zooming.

The imposition of the spectral profile produces a consistent improvement in all the cases ($|\frac{2\pi}{N}\omega|$ vs. $|\frac{2\pi}{N}\omega|^{1.6}$ vs. $|\frac{2\pi}{N}\omega|^{1.9}$), and imposing the profile corresponding to the coefficient decay of the reference image (Figure 3.4) produces the best results. We can study the performance of the spectral profiles by analyzing the frequency distribution of the errors of the restored images with respect to the reference image (GT), shown in Figure 3.6. There we can see that using the linear profile (total variation), the low frequencies are heavily penalized and most of the errors come from them, but imposing the profile corresponding to this image ($|\frac{2\pi}{N}\omega|^{1.9}$) we reduce the errors in the low frequency range.

Finally, let us spend a word to note the impact of the practical stopping conditions proposed in Section 3.7, since they allow a significant speed up of the algorithm reducing the execution time from 200 sec (stopping after 50 Uzawa's iterations) to 30 sec for images of size 149×149 pixels.



(a) Spectral profiles.

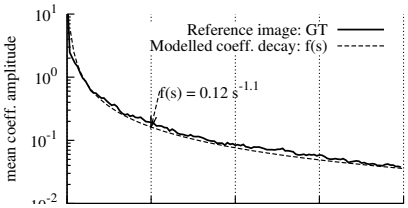


(b) Errors in the restored images.

Figure 3.6: Denoising experiments with different selections of spectral profiles $A(i\omega)$. Using the spectral profile that fits the model for the ground truth image, reduces the error, especially in the mid-low frequencies.

(a) Reference Image.

(b) Noisy, blurred and irregularly sampled image.



(c) Fourier coefficient decay of the reference image.

Figure 3.7: Image used in the deconvolution experiment. The Fourier coefficients' amplitudes in the reference image (257×257 pixels) decay as $(2\pi\omega/N)^{-1.1}$.

Table 3.3: Comparison of the algorithms in the irregular to regular sampling and deconvolution task.

Algorithm	Regularizer	RMSE
ACT+TV [ACHR06]		9.035
QN+FAR $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1.6}$	$K_{\beta,A}(u)$ eq. (3.21)	8.751
QN+TV $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1}$		8.542
QN+FAR $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{0.4}$		8.779
QN+FAR $\|A(i\omega)\| \sim \|\frac{2\pi}{N}\omega\|^{0.4}$ for $\|\omega\| \geq \frac{N}{4}$ and $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1}$ for $\|\omega\| <= \frac{N}{4}$		8.337
regular QN+FAR $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1.1}$	$K_{\beta,A}(u)$ eq. (3.21)	8.393
L^2-norm of the gradient	$\mathrm{L}_A(u)$ eq. (3.39)	9.112
L^2-norm of $A(D)u$ with $\|A(i\omega)\| \sim \|\frac{2\pi}{N}\omega\|^{0.4}$ for $\|\omega\| \geq \frac{N}{4}$ and $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1}$ for $\|\omega\| <= \frac{N}{4}$		9.044

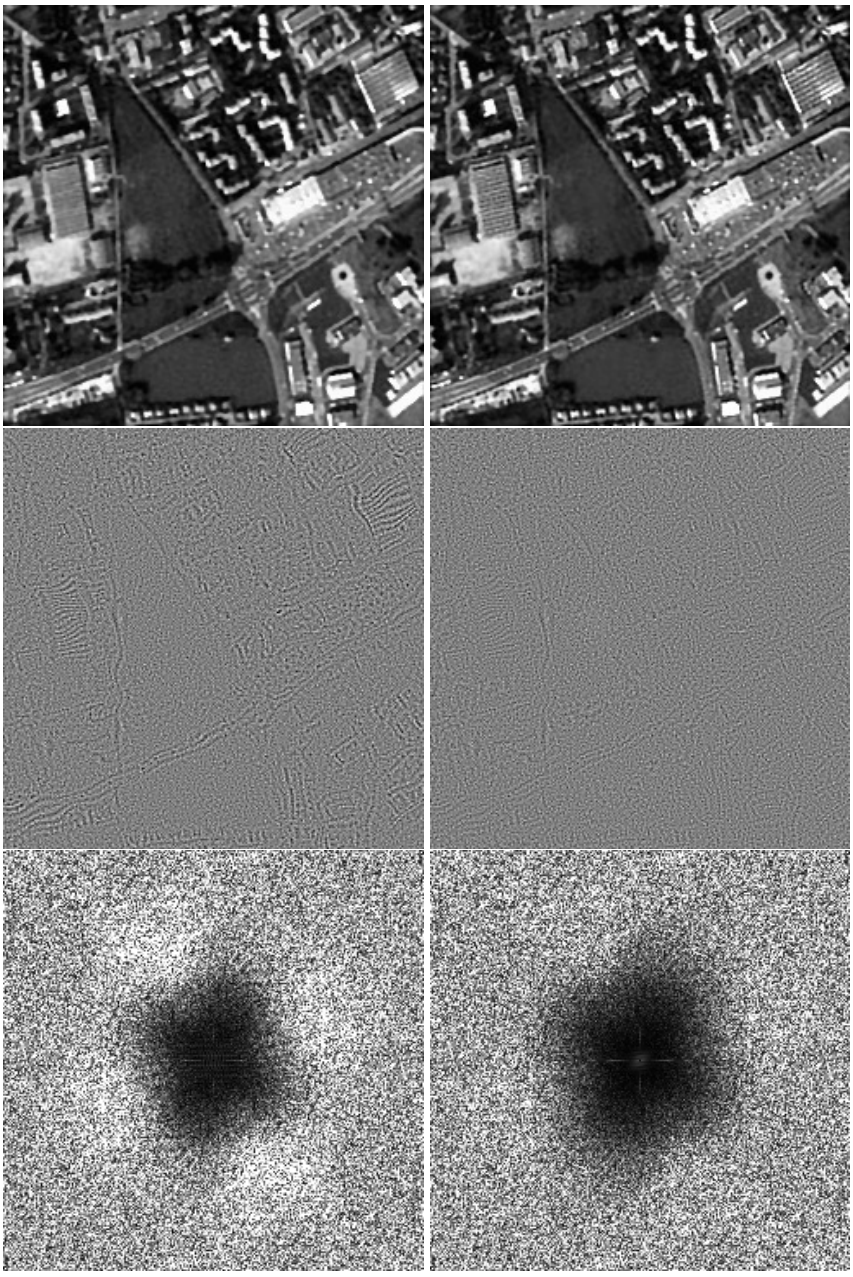
3.8.2 Denoising and deconvolution

We consider in this section the denoising and deconvolution of irregularly sampled images. For that we include in our image formation model the MTF corresponding to SPOT 5 HRG (High Resolution Geometric) satellite with *Hipermode* sampling [HML94]. Shortly, Hipermode is a push-broom acquisition mode that uses two shifted bars of sensors to sample on a double-density grid. The MTF associated to this system is modeled by

$$\hat{h}(p,q) = \underbrace{\operatorname{sinc}_\pi\left(\frac{2p}{N}\right)\operatorname{sinc}_\pi\left(\frac{2q}{N}\right)e^{-\beta_1|\frac{p}{N}|}}_{\text{sensor integration blur}}\underbrace{e^{-\frac{\alpha_1}{N}\sqrt{p^2+q^2}}}_{\text{optics \& aperture diffraction}}\underbrace{\operatorname{sinc}_\pi\left(\frac{p}{N}\right)}_{\text{motion blur}},$$
$$-\frac{N}{2} < p,q \leq \frac{N}{2},$$

where $\alpha_1 = 3.73$, $\beta_1 = 1.75$, and where $\operatorname{sinc}_\pi(x) = \frac{\sin(\pi x)}{\pi x}$ if $x \neq 0$ and $\operatorname{sinc}_\pi(x) = 1$ if $x = 0$. This function has its first zero crossing at frequency $N/2 = 1/2Hz$, while at frequency $1/4Hz$ the power of the MTF is only 1%, meaning that outside the spectral support $[-N/2, N/2]^2$ (or $[-1/2, 1/2]^2$ Hz) of the filter there is almost no information. To simulate a blurred, noisy and irregularly sampled image with $N \times N$ samples we start by applying the MTF $\hat{h}$ to the image $u_0 \in X_N$. The samples z (Figure 3.7) are then obtained by sampling $h * u_0$ on the irregular grid and adding the noise.

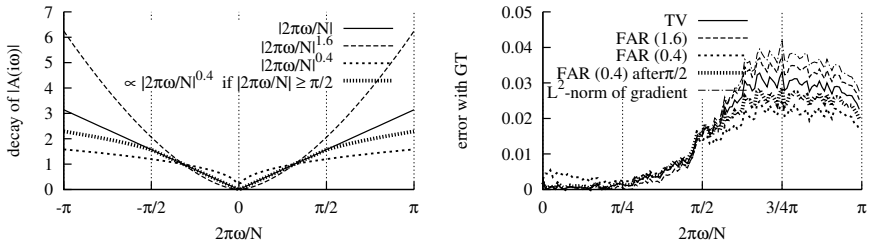
In Figure 3.8 and Table 3.3 we report the results of the restoration experiments corresponding to deconvolution with the Hipermode MTF. The small variability in the results is an indicator that the MTF $\hat{h}$ makes the restoration harder per-se, this is also confirmed by restoring a regularly sampled image and observing that in that case the error is similar to the irregular one (see Table 3.3). Anyway, the proposed algorithm outperforms ACT+TV [ACHR06] mainly thanks to the local formulation of the constraints (see Figure 3.9).



Global Constraint (RMSE=9.035) Local Constraints (RMSE=8.542)

Figure 3.8: Restoration with deconvolution computed with ACT+TV [ACHR06] (left), and with the proposed algorithm (right) that solves (3.21) . In the first row are shown the restored images. In the second row, the method errors, that are re-scaled from [-5,5] (the range of the image is [0,255]). The third row is shown the Fourier transform of the method noise (it should resemble the white noise).

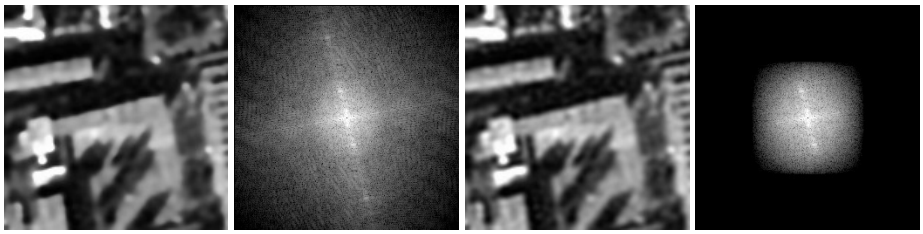
Figure 3.9: Two details of images deconvolved using total variation and L^2-norm of the gradient as regularizers, both with local and global constraints. From left to right is shown the result of the proposed algorithm (3.21), ACT+TV [ACHR06] (global constraint), the L^2-norm with local constraints and with a global constraint. Note that the last two results are more blurred that the total variation ones, and as seen in Table 3.3 have a higher RMSE.



(a) Spectral profiles.

(b) Errors in the restored images.

Figure 3.10: Deconvolution experiments with different selections of spectral profiles $A(i\omega)$. Using the spectral profile that fits the model for the ground truth image, reduces the error.



TV result (RMSE=6.079) L^2-norm result (RMSE=6.270)

Figure 3.11: Deconvolution and zoom using TV and the L^2-norm of the gradient as regularizers. The images show a detail of the zoomed images and its spectrum. Note that the L^2 regularizer recovers the low frequencies but is unable to extrapolate to higher ones.

We observed that manipulating the spectral profile $A(D)$ does not produce improvements consistent with the denoising case (see Figure 3.10). This is due, on one hand, to the fact that the reference image has a spectral decay different from the target image class ($|\frac{2\pi}{N}\omega|^{1.6}$). Indeed the Fourier coefficients of the reference image decay as ($|\frac{2\pi}{N}\omega|^{1.1}$), which explains the good performance of the total variation (see Table 3.3). On the other hand, in Figure 3.10(b) we see that the total variation controls the error in the low frequency range, but in the high frequency range it is too conservative and does not promote the spectral extrapolation.

The previous observation motivates the following experiment, building a profile that combines, the decay of the total variation in the low frequency range, with a decay similar to $|\frac{2\pi}{N}\omega|^{0.4}$ in the high frequency range. The result of this experiment is shown in Table 3.3 and its profile is depicted in Figure 3.10(b), there we can confirm the desired effect.

Remark 8. The results exhibited in Tables 3.3 and 3.4 confirm that the regularizers based on the minimization of L^2-norm, are not adequate to restore images with deconvolution and/or zoom, due to the inability of this type of regularizers to extrapolate the spectrum. This difference is also illustrated in Figure 3.9 and in Figure 3.11 for the case of deconvolution and zoom.

3.8.3 Extension to zooming

Zooming requires to interpolate and restore the image while preserving and enhancing the shapes, this can be seen as a spectrum extrapolation problem. The basic idea is to fit in as much as this is possible, the low frequency components of the restored and zoomed image to the original data, and to extrapolate the spectrum to the rest of the frequency domain by means of the regularization functional. The regularization allows to recover some high frequencies, which is indeed much more convenient than just filling them with zeros, a technique which is known to produce ringing.

Since the FAR regularizer allows to control the spectral behavior of the solution, in particular the extrapolated part, it will allow to improve the zoom results. Let us first extend the formulation (3.21) to consider the restoration of irregularly sampled images with a zoom of factor n

$$\begin{aligned} &\min_{u \in X_{nN}} K^n_{\beta,A}(u) := \sum_{0 \le r,l < 2nN} \sqrt{\beta^2 + |A(D)Pu\,(r,l)|^2} \quad , \\ &\text{subject to } \left[G * |\Delta_\Xi\,(h * p * u) - z|^2 \right](\xi_k) \le \sigma^2 \quad \forall \xi_k \in \Xi \qquad (3.40) \\ &\text{and } \sum_{0 \le r,l < nN} u\,(r,l) = \sum_{\xi_k \in \Xi} w_k z(\xi_k). \end{aligned}$$

The zoomed and restored image u is a vector of size $nN \times nN$ (we recall that the size of z is $N \times N$), and p is a spectral projector (e.g. $\hat{p} = \chi_R$ or a prolate function) on a low-band region R which depends both on the MTF and the sampling set. In the context we are considering here, $\hat{p}$ will be different from zero in the frequency band corresponding to the resolution of the data $[-1/2, 1/2]^2$ and the constraint is saying that the data is explained by the lower frequency part $h * p * u$ of $h * u$. The regularization functional $K^n_{\beta,A}(u)$ penalizes the oscillations that may appear when we extrapolate the high frequencies in the spectral region $[-n/2, n/2]^2 \setminus [-1/2, 1/2]^2$. Let us mention that, as discussed in [Alm02] in the context of regular sampling, the right choice of the spectral region R permits to reduce the aliasing effects, but we shall not consider this problem here. For us, if we want to restore and zoom the image u by a factor n, $\hat{p}$ will be different from zero on the region $R = [-1/2, 1/2]^2$ and zero on $[-n/2, n/2]^2 \setminus R$. This is a way to impose that the restored image fits the data z at low frequencies and the high ones are extrapolated via the minimization of $K^n_{\beta,A}(u)$. This minimization problem (3.40) with $\hat{p}(\omega) = \mathbf{1}_{I_N}(\omega)$ is a direct extension of the oversampling and denoising method introduced by Malgouyres and Guichard [MG02] to the more general case of irregular to regular sampling, deconvolution, denoising and oversampling.

The experiments shown here correspond to a 2X zoom, the images were simulated using the procedure described for the deconvolution case, with a filter $\hat{h}$ extended with zeros up to a double size, and where the irregular sampling is performed at a double spacing. The resulting image is shown in Figure 3.12. Since the restored image and the reference image u_0 have the same size, they can be directly compared. In Figure 3.12 are shown the distorted and the reference image as well as two restorations.

Let us first comment on the stair-casing effect that is noticeable in the bottom left image of Figure 3.12. It is a common observation that the total variation *introduces* a staircasing effect in the restored images, but let us point out that in our case where the derivatives are computed analytically this effect may not appear. It appears in the bottom left image of Figure 3.12 due to the poor discretization of the total variation. Notice that the same image processed with a finer approx-

Table 3.4: Deconvolution denoising and 2X zoom experiments.

Algorithm	Regularizer	RMSE
ACT+TV [ACHR06]		6.236
QN+FAR $\|A(i\omega)\| \sim \|\frac{2\pi}{N}\omega\|^{1.6}$ for $\|\omega\| > \frac{N}{4}$	$K_{\beta,A}(u)$ eq (3.21)	5.957
QN+TV $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1}$		5.965
QN+FAR $\|A(i\omega)\| \sim \|\frac{2\pi}{N}\omega\|^{0}$ for $\|\omega\| > \frac{N}{4}$ and $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1}$ for $\|\omega\| <= \frac{N}{4}$		5.996
L^2-norm of the gradient	$\mathrm{L}_A(u)$ eq. (3.39)	6.270
L^2-norm of $A(D)u$ with $\|A(i\omega)\| \sim \|\frac{2\pi}{N}\omega\|^{1.6}$ for $\|\omega\| \geq \frac{N}{4}$ and $\|A(i\omega)\| = \|\frac{2\pi}{N}\omega\|^{1}$ for $\|\omega\| <= \frac{N}{4}$		6.270

imation as proposed in section 3.4.1 does not exhibit this artifact; see bottom right image in Figure 3.12. As we mentioned above, this effect is a product of the coarse discretization of the total variation and was negligible in the cases of restoration without zoom. But when zooming it is important to avoid this effect since it produces unnatural looking images in spite of the fact that the RMSE errors of both images are similar.

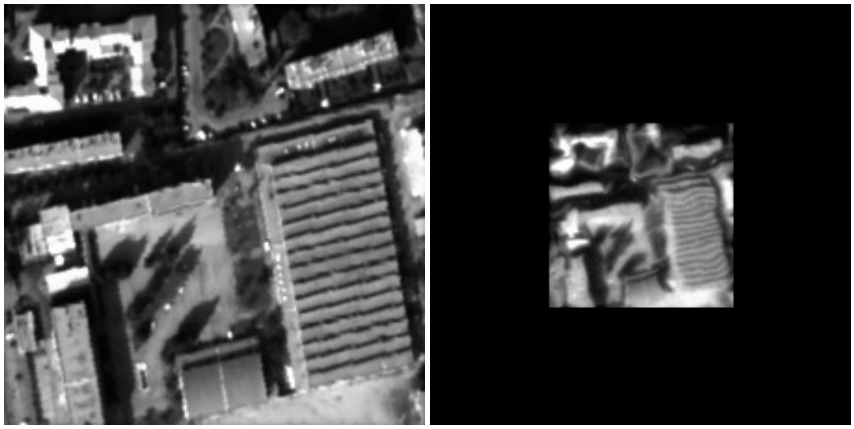
Since the spectrum is extrapolated, the quality of the zoomed image is not affected by the penalty in the frequencies imposed by the regularization term. In Table 3.4 we display the results obtained with different penalization profiles. Notice that the results are better in the case of L^1-norm than L^2-norm.

In contrast with the previous applications, in the present case, not removing all the noise leads to some artifacts. The noise is defined over the original grid (be either regular or irregular), but any residual of the original noise becomes a low frequency colored noise in the zoomed image that is visible as artifacts.

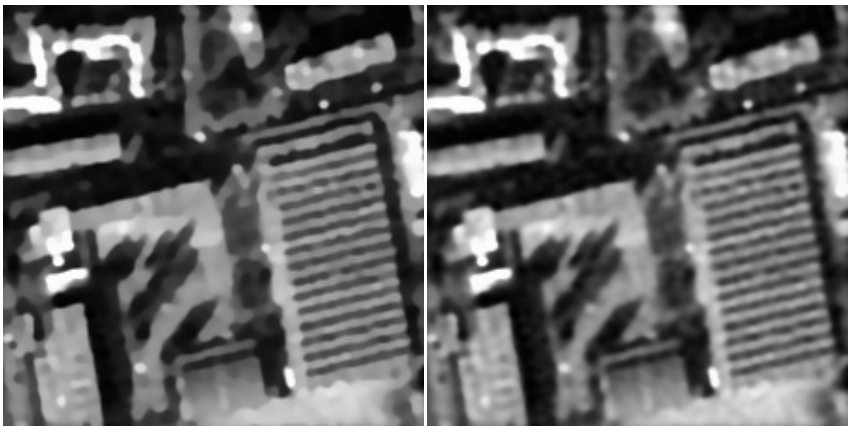
3.9 Conclusions

We have proposed a model for the restoration of band limited images that considers irregular (perturbed) sampling, denoising, and deconvolution. In addition, a theoretically sound discretization of this model and an algorithm that solves the resulting discrete optimization problem efficiently using automatically optimized thresholds and stopping conditions for a prescribed output precision level are presented.

An experimental evaluation (which actually requires all three modeling, discretization, and algorithmical elements) shows that our restoration model, and its accurate discrete approximation, actually improves the performance of previous methods in terms of both signal to noise ratio (Tables 3.2, 3.3 and 3.4) and adequacy of the statistical properties of the computed solution with respect to the assumed image acquisition model. As we can see from Figures 3.5 and 3.8 our approach permits to recover part of the structure lost with a global constraint leading to a method noise that resembles more the postulated gaussian white noise. As observed by Nikolova in [Nik07], most denoising methods tend



Reference image | Irregular samples



Zoomed with QN+FAR using (3.19) | Zoomed with QN+FAR using (3.21)

Figure 3.12: Deconvolution denoising and zoom. We display the QN+FAR results using a spectral decay like $(\frac{2\pi\omega}{N})^{1.6}$, which coincides with the decay of the reference image. The staircasing effect in *(c)* (RMSE: 6.003) is due to the poor discretization of the regularizer in (3.19). In *(d)* (RMSE: 5.957) we observe the result obtained with a finer discretization of the regularizer (3.21).

to produce a method noise with a distribution which largely differs from the postulated Gaussian white noise hypothesis.

The key ingredients that make our restoration model (Section 3.3) produce good results are the combination of

1. the modeling of the image acquisition system through local constraints as in [ABCH08, ACHR06],

2. the use of a special kind of *frequency adapted regularization* in the spirit of the ACT algorithm proposed by H.G. Feichtinger, K. Gröchenig, M. Rauth and T. Strohmer [GS00], and

3. the minimization of an L^1-norm (instead of L^2-norm) of a frequency-adapted pseudo-differential operator. However the L^2-norm produces competitive results in the case of denoising without deconvolution nor zoom (see Table 3.2).

All three aspects together collaborate to provide a better model of both, image acquisition and a priori image regularity, that leads to better accordance of the statistical properties of our solution with respect to the postulated models and the reference image. An accurate discrete approximation to this model (Section 3.4), both allows the use of convenient and efficient numerical algorithms (Section 3.6), and leads to well-behaved solutions (Section 3.5). Finally, the study of the local constraints (Section 3.7) leads to a trade-off between the precision of the noise estimation and the number of constraints that can be ensured, that is used to derive the stopping criterion for the different stages of the algorithm.

Experiments (Section 3.8) highlighting the importance of the modeling and discretization aspects described earlier are presented, namely: the quality boost provided by the local formulation of the constraints (Figure 3.5), the importance of the discretization of the regularizer (Figure 3.12) and the usefulness of the frequency adapted regularizer in favoring certain spectral behaviors of the solutions (Figure 3.5 and 3.8).

Our prospects for future work include adapting our algorithm for using splines as the underlying interpolation model, and the introduction of anti-aliasing filters to handle the case of aliased data. The exploitation of self-similarity proporties of most natural images would also be useful to increase resolution while reducing noise and aliasing effects.

Chapter 4

Open issues related to image fusion

In this chapter we discuss some issues related to the application of the restoration tools developed in Chapter 3 to the the problem of image fusion. We will explain the need of a "spectral projector" for restoring aliased images, and we will propose methods for computing spectral projectors in a couple of cases. With the spectral projector we can formulate the restoration of aliased images, where we show that the noise estimation is biased by the presence of the spectral projector. This bias is difficult to estimate since it depends on the statistics of each image, and is still an open problem.

4.1 From irregular sampling to image fusion

In the previous chapter we have studied the problem of restoring an irregularly sampled image in an alias free case. It remains the problem of restoring an irregularly sampled image in presence of alias. This situation may appear in the context of image fusion, where several images are registered so that samples can be combined into a single grid. The distortions of each grid can be an affine, projective or depth dependent, in any case, the resulting combined grid will not regular in general. Also, for the fusion to be effective, the images must contain alias, so we face exactly the type of problem we mentioned: irregular sampling restoration in the presence of alias.

As we will see next, the restoration techniques developed for irregular sampling may not be directly applicable to the case of grid fusion. In the case of microvibrations, the sample density (and therefore the resolution) of the acquired images is known and it is the same as in the restored image. But in the case of image fusion the objective is to increase the resolution of the restored image, therefore the final resolution must be estimated.

We will distinguish three types of sampling geometries, resulting from image fusion: 1) Two shifted grids; 2) Locally constant shift: If one of the grids suffers a smooth and low frequency distortion, then the effect of the distortions can be approximated locally by a constant shift; and 3) Completely irregular. This can be the result of superimposing grids distorted by an arbitrary of projective transformation

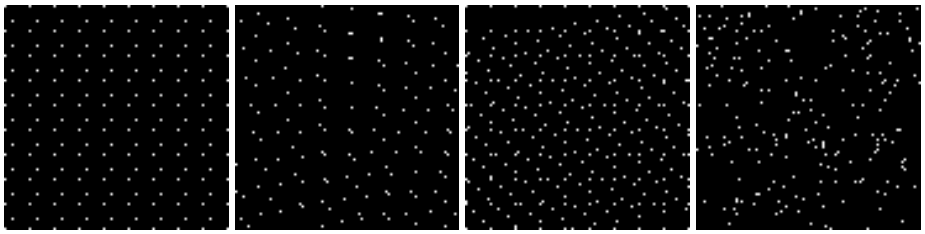


Figure 4.1: Grid Superposition (from left to right): two shifted grids, two grids smoothly distorted, two affine transformed grids, and two random grids. Notice how the sample spacing becomes more unpredictable as we move to right.

The Figure 4.1 illustrates the effect of combining different types of grids according to the previous cases.

Reducing the problem to restoration from irregular samples

Let us consider the model

$$z(\xi_i) = (h * u)(\xi_i) + n_{\xi_i}, \qquad \xi_i \in \Xi, \tag{4.1}$$

where the sampling Ξ is not sufficient to recover $h * u$ from z (we say that z is aliased). The difference with the model (3.1) is that the sampling Ξ does not satisfy the Nyquist criterion. This means that, the sampling may be not enough for capturing the high frequency components of $h * u$ and these may appear as artifacts, which are called "aliasing" in the regular sampling case.

The direct application of the restoration functional (3.5) to this problem leads to the apparition of artifacts in the solution due to the aliased data. As seen in Chapter 2, if the support of the solution u is overestimated, then the restoration may present artifacts at places where the sampling presents large gaps. This is the case of image fusion, since after combining the samples of the aliased (low resolution) images the resulting sampling grid may contain small and large gaps. The difficulty is to deal with the large gaps so that we avoid the "point-like" artifacts mentioned in Chapter 2 for the zoom problem.

Given the sampling Ξ, we call *spectral projector* the low-pass filter p such that $\hat{p}$ keeps only the un-aliased portion of the spectrum and these frequencies can be recovered without ambiguity from z. The notion of spectral projector is useful in two related problems: a) Zooming in case of regular sampling; b) Image fusion. Both require a spectral projector adapted to the sampling grid and/or to the expected alias levels.

When we considered the restoration from irregular samples in Chapter 2, we did not mention the need of a spectral projector for the alias. The reason was our assumption that the irregular sampling was sufficient to recover $h * u$ with no alias.

In [Mal00] (and also in section 3.8.3) the authors consider the zoom case in absence of alias, by solving:

$$\min J(u) \quad s.t. \quad \sum_{\xi \in \Xi} |(p * h * u)(\xi) - z(\xi)|^2 \leq |\Xi| \sigma^2,$$

where Ξ is a regular sampling grid that verify the Nyquist criterion for the image z, u is the zoomed image (thus, of higher resolution) and $\hat{p}$ is an spectral projector onto the part of the spectrum that is representable with the samples (note that since there is no alias then $\hat{z} = \hat{p}\hat{z}$). The interest of this method lies in its ability to extrapolate the spectrum in region complementary to $\hat{p}$ ($\hat{q} := (\hat{p})^c$), thanks to the total variation regularizer.

In [ADR04] the authors considered the restoration problem with regular sampling but in presence of alias. They proposed the use of the spectral projector $\hat{p}$ for removing the aliased frequencies whose expected power is smaller than its alias power, and solved:

$$\min J(u) \quad s.t. \quad \sum_{\xi \in \Xi} |(p * h * u)(\xi) - (p * z)(\xi)|^2 \leq |\Xi| b^2,$$

where the value of b^2 is manually chosen to obtain a reasonable restoration.

Generalizing the previous formulation to the case of an irregular sampling, we can incorporate the spectral projector p into an image restoration functional by means of global constraints:

$$\min_u \quad J(u), \quad \text{s.t.} \quad \sum_{\xi \in \Xi} \left[(p * h * u)\,(\xi) - z(\xi)\right]^2 \leq |\Xi|(\sigma^2 + \texttt{Bias}). \tag{4.2}$$

Observe that since z is a set of irregular samples, we cannot use the spectral projector p as in the regular case, so we must modify the constraint to consider the energy contribution of the alias which is integrated as a bias in the noise estimation. The bias is formulated as:

$$\texttt{Bias} := \sum_{\xi \in \Xi} |(q * h * u)(\xi)|^2 - 2((h * u - z)(\xi)(q * h * u)(\xi),$$

where $\hat{q} := (\hat{p})^c$. This bias accounts for to the error in the modeling of the noise in (4.2) due to the introduction of the spectral projector. The value of the bias depends on the image being restored, and will in general lead to an underestimation of the noise to be removed. Indeed, replacing $(h * u - z) = n$ in the definition of `Bias` we observe that when u is the ideal undistorted image then $\texttt{Bias} \geq 0$ (because the noise is non correlated with $(q * h * u)$). However in practice u is obtained from z which is in turn related to the noise, so this condition may not be satisfied and should be enforced. But imposing the non correlation of the noise residual with $(q * h * u)$ as a constraint leads to a non convex problem.

As a workaround to avoid the non convexity associated to the previous model we propose to solve it for each possible value of $\texttt{Bias} > 0$ without imposing the non correlation constraint between the noise residual with $(h * q * u)$. Let us denote the solution of the previous problem by u_b. If Bias is too small then u_b may not exist (i.e., the solution u_b cannot attain the desired noise statistics). On the other hand a big value leads to an over-relaxation of the problem and, therefore, to an over-smoothed result. The idea is to look for the smallest value of $\texttt{Bias} > 0$ for which the problem has a solution that satisfies the genuine noise residual constraint $\sum_{\xi \in \Xi} [(h * u)(\xi) - z(\xi)]^2 \leq |\Xi|\sigma^2$.

Still remains the problem of determining the alias-free part of the spectrum for a given sampling and a given class of images. In the following sections we will address this problem. For that we consider two types of grids: the union of two translated grids, and an irregular grid. The first case corresponds to a generalization of the Supermode acquisition geometry, in this case the underlying sampling geometry is regular and we can accurately estimate the spectral support of the solution. In the second case the geometry of the grid is unknown, so we resort to an heuristic argument to determine the spectral support of the solution. We verify that in the case of a regular grid, our heuristic produces reasonable results. In Section 4.4 we will display some restoration experiments obtained using the spectral projectors computed with the proposed methods.

4.2 Computing the spectral projection in Generalized Supermode: fusion of translated grids

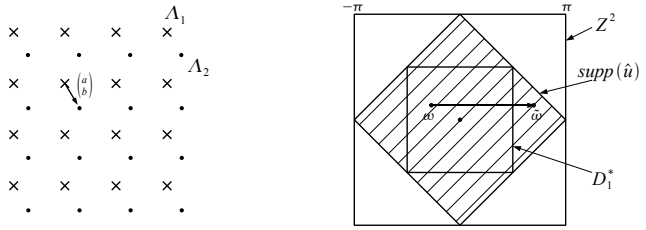
Consider two aliased observations z_1 and z_2 of the same scene taken with a small fronto-parallel camera translation. Our objective is to restore the scene u with the highest possible resolution. We assume that the two images are acquired with two sampling grids which are translated by a vector $s = (a, b)^t$ (see Figure 4.2(a)), that is:

$$\begin{aligned} \Lambda_1 &= 2\mathbb{Z}^2 \\ \Lambda_2 &= \Lambda_1 + (a, b)^t. \end{aligned}$$

The signal u is assumed to be band limited with its spectrum entirely contained in $R = supp(\hat{u})$ as depicted in Figure 4.2(a) (or at least with negligible higher frequency contributions). If the reciprocal cells D_i^* of the sampling grids Λ_i are contained in R (as shown in Figure 4.2(b)), then the sampling process can be modeled by:

$$\begin{aligned} z_1 &= \Pi_{\Lambda_1} u + n_1 \\ z_2 &= \Pi_{\Lambda_2} u + n_2, \end{aligned} \tag{4.3}$$

where n_i is the usual zero mean additive gaussian white noise. By abusing the 1D notation of the Fourier transform, the same process can be re-written in terms of the Fourier coefficients $\hat{u}$ (we denote $\omega = 2\pi \frac{k}{N}$):



(a) Sampling Grids configuration.

(b) Spectral support and aliases. $\tilde{\omega}$ is the alias of the frequency ω.

Figure 4.2: Fusion in the case of two shifted grids.

$$z_1(x) = \frac{1}{N}\sum_{k=0}^{N-1} \hat{u}(\omega)e^{i\omega x} + n_1 \qquad \forall x \in 2\mathbb{Z}^2, \tag{4.4}$$

$$z_2(x) = \frac{1}{N}\sum_{k=0}^{N-1} \hat{u}(\omega)e^{i\omega s}e^{i\omega x} + n_2 \quad \forall x \in 2\mathbb{Z}^2.$$

Since $supp(\hat{u}) \supset D_1^*$, most of the frequencies ω inside the cell D_1^* will be aliased (see Figure 4.2). But there will be only one alias contribution $\tilde{\omega}$ per frequency since the spectrum of u is contained in the diamond shaped region region R in Figure 4.2(b). And the alias corresponding to each frequency will be a function of the quadrant. Let us define

$$\gamma \in \left\{ \begin{pmatrix} 0 \\ \pi \end{pmatrix}, \begin{pmatrix} 0 \\ -\pi \end{pmatrix}, \begin{pmatrix} \pi \\ 0 \end{pmatrix}, \begin{pmatrix} -\pi \\ 0 \end{pmatrix} \right\}$$

as the vector such that, for a given frequency ω, it determines the location of its first alias $\tilde{\omega} = \omega + \gamma$. Its value will depend on the quadrant containing ω (see Figure 4.2(b)).

The alias can be expressed in terms of the Fourier coefficients of z_i (recall that z is undersampled). This relation is given, for each pair ω and $\omega + \gamma$, by the following linear system:

$$\begin{bmatrix} \hat{z}_1(\omega) \\ \hat{z}_2(\omega) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ e^{i\omega s} & e^{i(\omega+\gamma)s} \end{bmatrix}}_{A_\omega} \begin{bmatrix} \hat{u}(\omega) \\ \hat{u}(\omega+\gamma) \end{bmatrix} + \begin{bmatrix} \hat{n}_1(\omega) \\ \hat{n}_2(\omega) \end{bmatrix}. \tag{4.5}$$

Our ability to recover the frequencies $\hat{u}(\omega)$ and $\hat{u}(\omega+\gamma)$ will depend on the values taken by the vector s and the power of the noise. For instance, if $s = (0,0)^t$

the matrix A_ω will not be invertible, because $e^{i\omega s} = e^{i(\omega+\gamma)s} = 1$. Observe that this is the case where the two sampling grids coincide ($s = (0,0)^t$), so it is impossible to recover both frequencies. This situation repeats for $s \in \{(2,2)^t, (0,2)^t, (2,0)^t\}$. The best case corresponds to $s = (1,1)^t$, then the matrix becomes $A_\omega = \begin{pmatrix} 1 & 1 \\ e^{i\omega} & -e^{i\omega} \end{pmatrix}$ and the system can be solved with a condition number $\kappa(A_\omega) = 1 \quad \forall\omega$.

4.2.1 Computing the spectral projector

We have seen that we cannot solve the system for all displacements s. Our objective in this section is to determine when it is worth to solve (4.5) to recover $\hat{u}(\omega)$ and $\hat{u}(\omega+\gamma)$; sometimes the best option is to discard one frequency (zero padding) or maybe both (ω and $\omega+\gamma$).

For each pair of frequencies ω and $\omega+\gamma$, let us define the random vectors $\mathrm{n} := (\hat{n}_1, \hat{n}_2)^t$ and $\mathrm{u} := (\hat{u}(\omega), \hat{u}(\omega+\gamma))^t$. Then the samples are given by $\mathrm{z} := (z_1(\omega), z_2(\omega))^t = A_\omega \mathrm{u} + \mathrm{n}$. For a matrix Q we define the restoration as $\tilde{\mathrm{u}} = Q\mathrm{z}$, and the residual with respect to the original vector as $\mathrm{u} - \tilde{\mathrm{u}}$. Ideally, in absence of noise Q should be the inverse of A_ω. To study the quality of the restoration, we introduce the expected approximation error for the frequencies ω and $\omega+\gamma$ as:

$$H(\omega) = tr(cov(\mathrm{u} - \tilde{\mathrm{u}}, \mathrm{u} - \tilde{\mathrm{u}})).$$

Replacing the definitions of $\tilde{\mathrm{u}}$ and z, and using independence of the vectors u (the ideal solution) and n (the noise), we get:

$$H(\omega) = tr(Q(A_\omega E\{\mathrm{u}\mathrm{u}^*\}A_\omega^* + E\{\mathrm{n}\mathrm{n}^*\})Q^* - 2E\{\mathrm{u}\mathrm{u}^*\}A_\omega^* Q^*) + E\{\mathrm{u}\mathrm{u}^*\}, \quad (4.6)$$

where X^* denotes the adjoint matrix (conjugate transpose) of X. The computation of $H(\omega)$ depends on the a-priori information about the statistics of n and u. For the noise $E\{\mathrm{n}\mathrm{n}^*\} = diag(\sigma^2, \sigma^2)$ is the variance of the gaussian noise. As for the statistics of the undistorted image u, the spectral decay of $\hat{u}(\omega)$ is known for the class of images we are dealing with (see [Alm02, pg. 29]), that is $E\{|\hat{u}(\omega)|^2\} = \left(\frac{1}{\epsilon+|\omega|^{1.6}}\right)^2 =: F(\omega)$ and therefore $E\{\mathrm{u}\mathrm{u}^*\} = diag(F(\omega), F(\omega+\gamma))$.

We will consider three possible ways to recover $\tilde{\mathrm{u}}$:

1. **inverting the matrix A_ω:** in this case $Q_S = A_\omega^{-1}$ and we denote $H_S(\omega)$ the corresponding error as computed with (4.6),

2. **discarding the alias frequency $\omega+\gamma$:** then $Q_Z = \begin{pmatrix} 0.5 & 0.5 \\ 0 & 0 \end{pmatrix}$ and the error is $H_Z(\omega)$,

3. **or discarding both frequencies:** then $Q_D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and the error is

$$H_D(\omega) = E\{|\Delta\hat{u}(\omega+\gamma)|^2\} + E\{|\Delta\hat{u}(\omega)|^2\} = F(\omega+\gamma) + F(\omega).$$

For each pair of frequencies ω and $\omega + \gamma$, we will opt for the operation that produces the least expected approximation error (i.e., that minimizes $H(\omega)$). So our definition of the alias-free cell will be:

$$RC = \{\omega, \omega + \gamma : H_S(\omega) < min(H_Z(\omega), H_D(\omega))\} \cup \{\omega : H_Z(\omega) < H_D(\omega)\}.$$

In section 4.4 we show an example of a projector computed according to this criterion.

4.3 Pessimistic Reciprocal Cell: arbitrary grid case

When the sampling is not regular then the estimation of the spectral support is not easy since the concept of Fourier transform is not well defined. In this section we propose a heuristic for estimating a spectral projector that seeks to guarantee the invertibility of the irregular sampling operator. For simplicity we perform the following development in one dimension.

Suppose that $u(x)$ is given by the trigonometric polynomial with coefficients $\hat{u}_k$:

$$u(x) = \sum_{k=0}^{N-1} \hat{u}_k e^{\frac{2\pi i}{N} kx} \quad x \in [0, 1].$$

Let $\Xi = \{\xi_n\}$ be a set of sample locations taken in the interval $[0, 1]$. The evaluation of the trigonometric polynomial at arbitrary locations permits to sample the band limited function u. We denote the sampling operator as a matrix S:

$$S = \begin{bmatrix} e^{\frac{2\pi i}{N} 0\xi_1} & \dots & e^{\frac{2\pi i}{N}(N-1)\xi_1} \\ \vdots & \dots & \vdots \\ e^{\frac{2\pi i}{N} 0\xi_N} & \dots & e^{\frac{2\pi i}{N}(N-1)\xi_N} \end{bmatrix} \quad s.t. \quad S\hat{u} = z.$$

If we suppose that the function u is band limited, then it is possible to determine the coefficients $\hat{u}_k$ from a set of N regularly spaced samples of u by computing its Fourier transform

$$\hat{u}_k = \int_0^1 u(x) e^{-\frac{2\pi i}{N} kx} dx, \qquad k \in \{0...N-1\}.$$

But when the sampling is irregular one could only approximate the previous integral by the Riemann sum:

$$\tilde{u}_k = \sum_{n=1}^{|\Xi|} u(\xi_n) e^{-\frac{2\pi i}{N} k\xi_n} \Delta_n, \qquad k \in \{0...N-1\},$$

where Δ_n is the area of the n-th cell in the Voronoi diagram induced by the samples (but restricted to the domain $[0, 1]$). The matrix associated to this operation

involves the adjoint of the matrix S:

$$S^*W = \begin{bmatrix} e^{\frac{-2\pi i}{N}0\xi_1} & \dots & e^{-\frac{2\pi i}{N}0\xi_N} \\ \vdots & \dots & \vdots \\ e^{\frac{-2\pi i}{N}(N-1)\xi_1} & \dots & e^{-\frac{2\pi i}{N}(N-1)\xi_N} \end{bmatrix} \begin{bmatrix} \Delta_1 & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & \Delta_N \end{bmatrix}.$$

Let us call the operator S^*W *pseudo Fourier transform*. The quality of the approximation of this transform depends on the sampling grid and on each particular frequency; being better at low frequencies. In particular, if the samples are taken at regular intervals, this approximation coincides with the discrete Fourier transform.

The properties of the discrete Fourier transform are distorted in this approximation. For instance, the orthogonality is lost, so the response of a sinusoidal signal is no longer a delta, instead it is spread across many frequencies. Still we can ask questions like: how well it approximates the real DFT ?, is it possible to reconstruct a signal from it ? and up to which frequency? Figures 4.3 and 4.4 show images where a image with a single frequency is analyzed. Observe that in both figures the peak of the pseudo Fourier transform corresponds to the main frequency of the original image, but the transform also contains many nonzero responses ("aliases") at other frequencies.

4.3.1 Grid Potential

A natural question to ask is whether it is possible to undo this alias. Let us substitute $z = S\hat{u}$ into the pseudo transform of S^*Wz to obtain:

$$D\hat{u} := S^*WS\hat{u} = \tilde{u}.$$

The matrix $D \in \mathbb{C}^{N\times N}$:

$$D(l,k) = \sum_{n=1}^{|\Xi|} e^{\frac{2\pi i}{N}l\xi_n} e^{-\frac{2\pi i}{N}k\xi_n}\Delta_n = \sum_{n=1}^{|\Xi|} e^{\frac{2\pi i}{N}(l-k)\xi_n}\Delta_n \qquad l,k \in \{0...N-1\}$$

characterizes the response of the sampling Ξ to the different frequencies l. In other words, it tells us the way each frequency of $\hat{u}$ is represented by the pseudo Fourier transform. We call this matrix *grid potential*. Inverting D permits to recover $\hat{u}$ from $\tilde{u} = S^*Wz$. But unfortunately the inversion of D is generally ill-posed.

4.3.2 The pessimistic reciprocal cell

We propose to analyze the matrix D to determine the frequencies of $\hat{u}$ that can be recovered (without alias), thus stabilizing the inversion of D. To perform the analysis we suppose that the spectral decay of the target images $\hat{u}$ is known. For natural images it is commonly accepted that the spectral decay is $1/|k|$ (where k

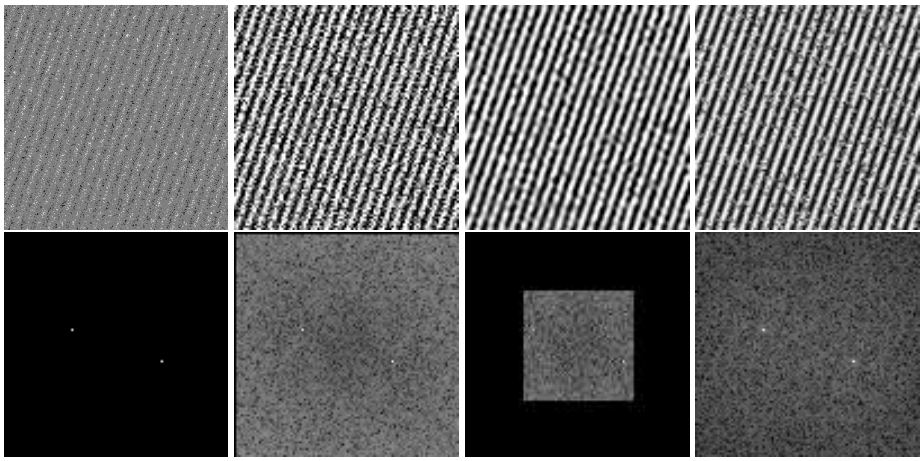


Figure 4.3: Reconstruction of a randomly sampled image using the pseudo Fourier transform. In the first row from left to right: the samples from the original image, the image reconstructed using the pseudo Fourier transform (that is, $F^{-1}S^*Wz$), the reconstruction obtained by projecting half the frequencies (because the maximal gap of the sampling is 2), and the restoration obtained with the ACT algorithm. In the second row from left to right: the Fourier transform of the original image before sampling, the remaining images correspond to the pseudo Fourier transforms (S^*Wz) of the corresponding image in the first row.

is the frequency), we take a slightly different decay of the expected value of the energy $E\{|\hat{u}(k)|^2\} = \phi(k)^2 = \left((1+|k|)^{-1.6}\right)^2$. Assuming independence of the frequencies, the total energy of the signal will be:

$$E = \sum_k \phi(k)^2.$$

Now, if we consider the pseudo Fourier transform $\tilde{u}$ such that $\tilde{u} = D\hat{u}$, then we can express each pseudo Frequency as:

$$\tilde{u}(l) = D(l,l)\hat{u}(l) + \sum_{k \neq l} D(l,k)\hat{u}(k).$$

When the sampling is regular then $D = DFT \circ IDFT$ so $D(l,l) = 1$ and $D(l,k) = 0 \quad \forall k \neq l$. But for an irregular sampling the sum $\sum_{k \neq l} D(l,k) \neq 0$, and carries the alias contributions coming from other frequencies.

We propose the following heuristic for determining the alias-free frequencies: the frequency l is recoverable if its expected alias energy $\sum_{k \neq l} D(l,k)$ is "less" than the expected energy in the l-th component, that is

$$C_2(l) = H\left(\tau \|D(l,l)\|^2 \phi(l)^2 - \sum_{k \neq l} \|D(l,k)\|^2 \phi(k)^2\right),$$

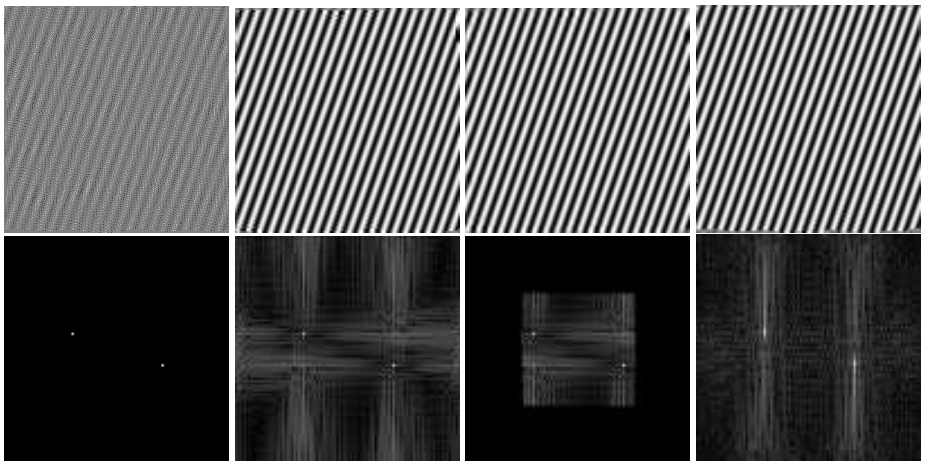


Figure 4.4: Reconstruction of an image acquired with a perturbed sampling using the pseudo Fourier transform. In the first row from left to right: the samples of the original image, the image reconstructed using the pseudo Fourier transform (that is, $F^{-1}S^*Wz$), the reconstruction obtained by projecting half of the frequencies (because the maximal gap of the sampling is 2), and the restoration obtained with the ACT algorithm. In the second row from left to right: the Fourier transform of the original image before sampling, the remaining images correspond to the pseudo Fourier transforms (S^*Wz) of the corresponding image in the first row.

Where $H(\cdot)$ is the Heaviside step function, $\tau = 1$ and $\phi^2(k)$ is the expected power of the frequency k. Computing C_2 for all the frequencies results in a binary matrix which tells us when the sampling is insufficient for a certain frequency. In Figure 4.5 we show some examples of projectors C_2 obtained with several types of samplings. Although this procedure is not theoretically justified, we observe that the obtained projectors C_2 are sound for all the samplings considered in Figure 4.5.

As a final experiment we apply the projector C_2 to stabilize the restoration with the ACT algorithm (described in Section 3.2.1) when the sampling is insufficient:

$$C_2S^*WSC_2\hat{u} = C_2S^*Wz.$$

In Figure 4.6 we show the restorations obtained using the ACT algorithm with the C_2 projector.

4.4 Image fusion experiments

To illustrate the use of the spectral projectors in an image fusion problem we present here a simple application: the fusion of a regularly sampled image with a second image that was sampled on a shifted grid.

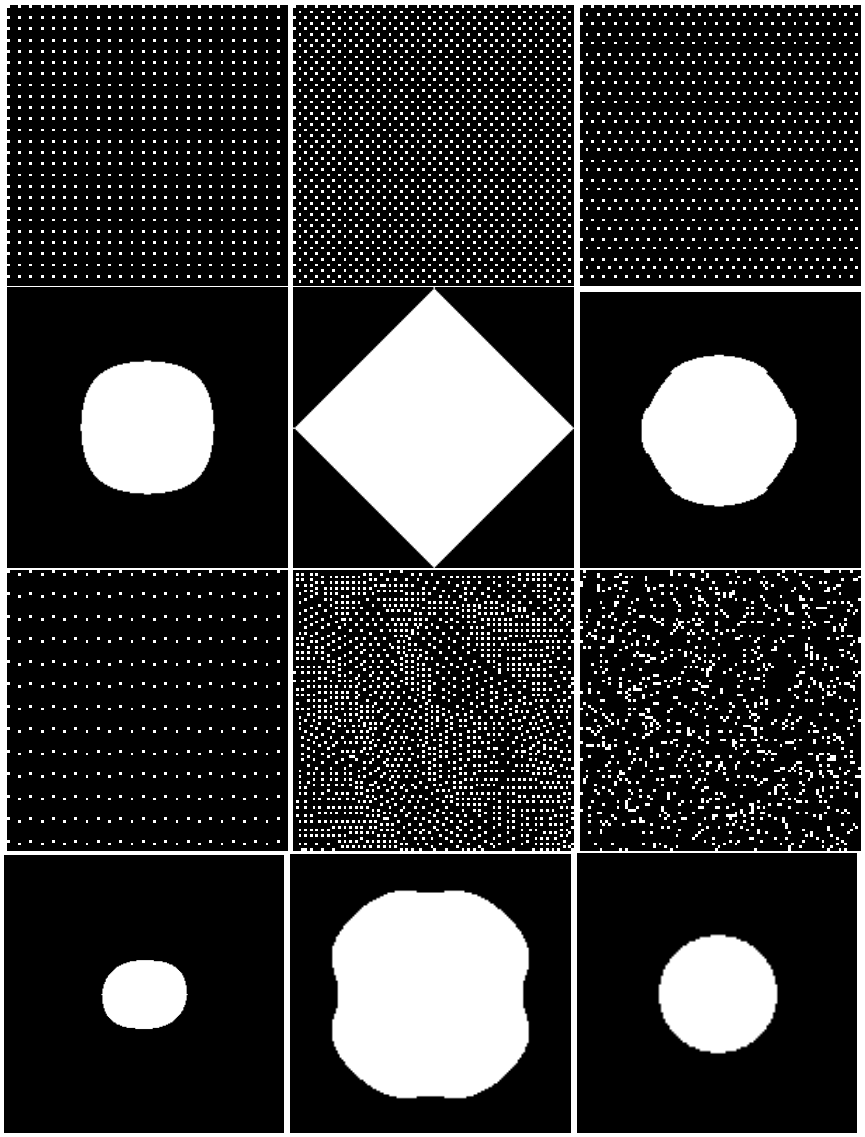


Figure 4.5: Examples of empirical cell C_2. We show some cells for some samplings grids: squared, quinconx, hexagonal, superposition of two squared samplings, perturbed and random sampling. The samplings are shown zoomed 4 times, while the cells are computed in the domain $[-2\pi, 2\pi]^2$, corresponding to a zoom of factor 2.

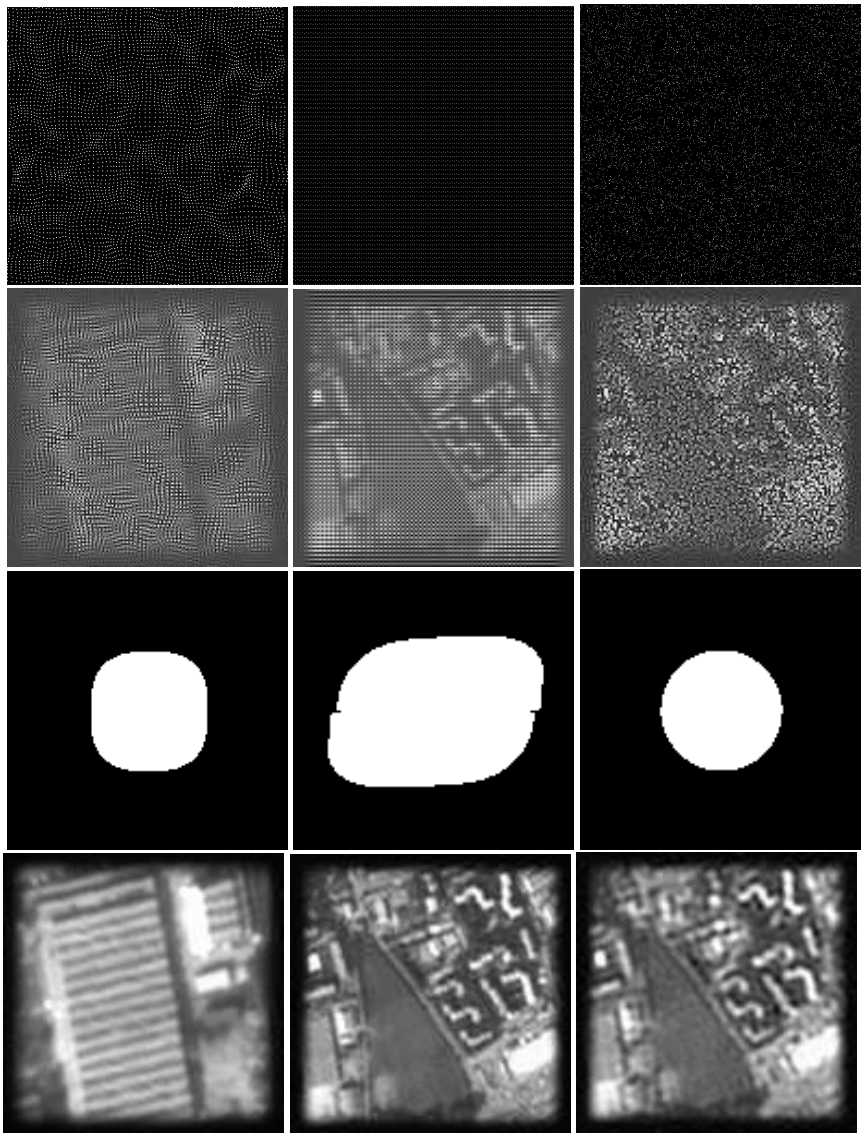


Figure 4.6: Restoration using the empirical cell C_2. From the first row to the last: sample positions, result of the classical ACT algorithm performing a 2x zoom, the computed projector C_2 ($\tau = 1$) and the result of the modified ACT using the C_2 projector. The experiment of the first column used $\sim$ 5000 perturbed samples for restoring a 149×149 image , the second column $\sim$ 8100 random samples for restoring a 129×129 image, and the third column shows a sampling grid ($\sim$ 8100 samples) formed by the union of two regular grids used to restore a 129×129 image.

(a) Reference Image (b) Left sub-sampled image

Figure 4.7: Fusion experiment: reference image (257×257 pixels) and the simulation of a subsampled ($2x$) image corrupted by gaussian additive white noise with $SNR = 20$ (or 10% of the signal power).

Generalized supermode. The objective of the first experiment is to compare the performance of the generalized supermode fusion (described in Section 4.2) against the original supermode proposed by Latry and Rougé [LR00]. To do so, we simulated an ill posed case of image fusion where the two grids $2\mathbb{Z}^2$ are relatively translated by a vector $v = (1, 0.05)^t$, which results in an insufficient vertical sampling. In addition, the simulated images are also corrupted by a gaussian additive white noise of 10% of the signal power ($SNR = 20$) (see Figure 4.7).

In Figure 4.8 we compare the results of restoring the image using the supermode projector and the adapted projector of generalized supermode described in Section 4.2. The later allows to remove high frequency artifacts associated to the poor vertical sampling, see Figure 4.8.

Pessimistic reciprocal cell. A restoration computed with the pessimistic reciprocal cell is shown in Figure 4.10. We used the same images as in the previous experiment (see Figure 4.7).

Figure 4.8: Supermode spectral projector (left) and the generalized supermode projector corresponding to two shifted grids with a relative translation $v = (1, 0.05)^t$ and considering a signal to noise ratio of $SNR = 20$ (right).

(a) Fusion using the supermode projector. ($MSE = 19.7$)

(b) Fusion using the spectral projector computed with the generalized supermode. ($MSE = 21.0$)

Figure 4.9: Fusion results obtained by applying the restoration model (4.2) using the spectral projectors of Figure 4.8.

(a) Pessimistic reciprocal cell.

(b) Fused image using the pessimistic RC. ($MSE = 16.5$)

Figure 4.10: Fusion using the pessimistic reciprocal cell. Notice that some details are more blurred with respect to the result obtained using the generalized supermode.

Part II

Interpolation from sparse samples

Chapter 5

Non local methods for interpolation of sparsely sampled images

In this part we consider the problem of reconstructing an image from a set of sparse samples taken on a regular grid. We will address this problem using an exemplar-based variational approach which exploits the self-similarity of the image. For that we first propose and analyze a framework for exemplar based inpainting, and then generalize it to handle interpolation problems.

5.1 Introduction

We consider the problem of image reconstruction from a set of sparse samples. The samples may form a subgrid of the regular grid or be a random subset of it (see Figure 5.1). In the first case, the problem can be considered as a zooming problem. The case of samples taken on a non-integer grid will not be considered in this document.

Let Ω be the standard regular grid, and let $\Lambda \subseteq \Omega$ be the set of given samples. We want to reconstruct the image $u \in \Omega$ from the data given on Λ. Let us remark that when we refer to a sparse set of samples Λ we implicitly assume that the samples are spread uniformly across the domain Ω.

Existing interpolation approaches consider priors based on smoothness or regularity assumptions, which can be imposed by restricting the solution to be, for instance, band limited [GS00], of bounded variation [CS01], expanded over a basis of functions (*e.g.* splines [ASHU05a], radial basis functions [She68a]), or as the result of anisotropic diffusion process [GWW$^+$08], among others. All these approaches are local in the sense that only the neighboring pixels influence the interpolation at a given point.

Recently, non-local techniques applied to general image restoration tasks have gained considerable attention in the image processing community. As opposed to local methods, these non-local techniques use information from other parts of the image or even other images. Its recent development was triggered in part by the texture synthesis works of Efros and Leung [EL02] and Wei and

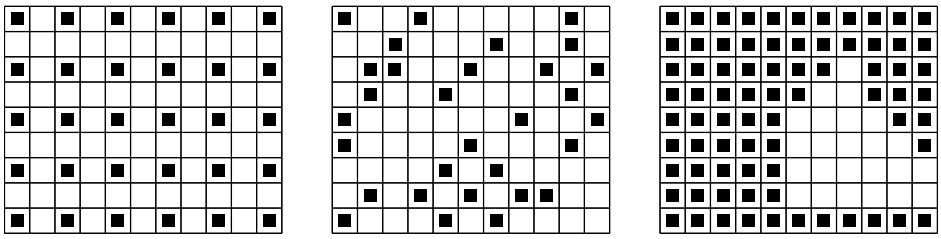


Figure 5.1: Example of different samplings. From left to right, a regular subgrid corresponding to a zoom, a random subset (interpolation from irregular samples) and a subgrid that forms a connected component (inpainting).

Levoy [WL00] using non-parametric sampling techniques. Non-local regularizers, as the one proposed [PBC09], model the set of image patches with their similarity relations as a weighted graph and the interpolation is done by imposing regularity in this graph [LEB07, GO07a]. Other non-local techniques are based on the *sparseland* model [CW08, AEB06]. According to this model the regularized image has a sparse representation in an overcomplete basis or *dictionary* [MSE08, AEB06, ESQD05]. The dictionary is learned from the same image or from a training set, and the regularity is imposed by sparsity over the dictionary.

The advances on texture synthesis of [EL02, WL00] also influenced the more related field of image inpainting. The objective of inpainting is, as in the case of interpolation, to obtain a visually plausible interpolation of the missing data, the difference is the organization of the known data. Applying the local regularity priors mentioner before does not produce plausible results, since the interpolated areas become visually apparent as texture-less artifacts. This is why texture synthesis approaches are so relevant to this application. The priors used for inpainting are generally founded upon structure propagation [BSCB00] and copy of exemplars from the same image [CPT04]. Exemplar-based methods are the quite successful since they also incorporate to some extent the propagation of structures.

5.2 Overview

In this part we will address the problem of image interpolation from non-uniformly sparsely sampled data using an exemplar-based variational approach which exploits the self-similarity of the image. In particular our formulation can be regarded as a non-local regularizer for sparsely sampled data. For that, we first present and analyze a framework for exemplar based inpainting, and then generalize it to handle interpolation problems.

Note that in the context of exemplar-based methods, the main difference between the inpainting and interpolation is the organization of the available

data: sparse samples for interpolation and large regions in the case of inpainting (Figure 5.1). Indeed in an usual inpainting problem the inpainting domain or "hole" (denoted as $O \subset \Omega$) is generally a region of the image. The exemplars (patches) that are used for completing the hole O are extracted from its complement $O^c = \Omega \setminus O^c$, which is composed by large known regions of Ω. In the interpolation case, on the other hand, the image is known only at some isolated positions $\Lambda \equiv O^c$ distributed across the image, this corresponds to a very big inpainting domain O.

For the inpainting problem we propose an energy that promotes the self-similarity of the image $u \in \Omega \subset \mathbb{R}^{N \times N}$, which is unknown at O and known at $O^c = \Omega \setminus O$. The functional is:

$$E(u, w) := \frac{1}{h} \sum_{x \in O} \sum_{\hat{x} \in O^c} w(x, \hat{x}) \|p_u(x) - p_u(\hat{x})\|^2 + \sum_{x \in O} H_w(x), \qquad (5.1)$$

where $H_w(x) := \sum_{\hat{x} \in O^c} \log(w(x, \hat{x}))w(x, \hat{x})$, $p_u(x)$ denotes a patch of u centered at x and $\| \cdot \|$ denotes the L^2 norm of a patch. The similarity weight function $w : O \times O^c \to \mathbb{R}$ is an auxiliary variable that accounts for the similarity of the pair of patches $p_u(x)$ and $p_u(\hat{x})$. The *selectivity parameter* h is generally set $h \to 0$ in the inpainting problem. The energy (5.1) contains two terms, the first one measures the coherence between patches centered in the unknown region O and those centered in the known part of the image O^c. This term permits the estimation of the image u from the weights w. The second term allows us to compute the weights w given the image u. The trivial minima of (5.1) when $w = 0$ is avoided by constraining $w(x, \cdot)$ to be a probability: $\sum_{\hat{x} \in O^c} w(x, \hat{x}) = 1 \quad \forall x \in O$.

The minimization of (5.1) is carried out by alternate minimization of u and w. Let us observe that for a fixed image u the minimization of $E(u, w)$ with respect to w yields exponential weights $w(x, \hat{x}) \propto \exp(\frac{1}{h}\|p_u(x) - p_u(\hat{x})\|^2)$, assigning higher weight to very similar patches. These are the familiar non-local weights used in denoising algorithms [BCM05, AW06]. And, for a fixed w, the minimization on u induces a non local averaging of patches (attention, it is an averaging of patches not of pixels as in non-local means). In conclusion, model (5.1) incorporates in an elegant way a sort of non-local means iteration both for the image and the weights update. Moreover, the model (5.1) can easily be extended to other patch norms such as L^1 or norms of gradients. These extensions will be studied in the case of inpainting.

To deal with the interpolation of sparsely sampled images we generalize the inpainting model (5.1). The assumption that there are complete patches $p_u(\hat{x})$ contained in O^c is false in the case of sparse samples interpolation, since the known samples may not be aggregated. In order the cope with this situation, we tailor the data attachment term to consider only the known pixels of one or both patches under comparison. Additionally to avoid bad local minima, we propose a coarse to fine scheme that progressively reduces the selectivity parameter h.

In Chapter 6 we describe a variational framework for non local image inpainting. The variational formulation of self-similarity proposed in that framework will be the basis for our energy formulation of the interpolation of sparsely

sampled images. This will be the subject of Chapter 7, where we introduce the interpolation model, explain its properties, develop an optimization algorithm for solving it, and present the results of some experiments.

Chapter 6

A Variational Framework for Exemplar-Based Image Inpainting

Non-local methods for image denoising and inpainting have gained considerable attention in recent years. This is in part due to their superior performance in textured images, a known weakness of purely local methods. Local methods on the other hand have demonstrated to be very appropriate for the recovering of geometric structures such as image edges. The synthesis of both types of methods is a trend in current research. Variational analysis in particular is an appropriate tool for a unified treatment of local and non-local methods. In this chapter we present a variational framework for non-local image inpainting. This framework relates to previous inpainting schemes, and also allows to derive novel ones.

6.1 Introduction

Image inpainting, also known as *image completion* or *disocclusion*, is an active research area in the image processing field. The purpose of inpainting is to obtain a visually plausible image interpolation in a region in which data are missing due to damage or occlusion. Usually, to solve this problem, the only available data is the image outside the region to be inpainted. In addition to its theoretical interest, image inpainting has applications to image and video editing and restoration.

Most inpainting methods found in the literature can be classified into two groups: *geometry-* and *texture-oriented* methods. We now briefly review the developments in both types of approaches, with emphasis in texture-oriented methods. This review will be helpful for motivating the proposed formulation.

Geometry-oriented methods. Images are modeled as functions with some degree of smoothness, expressed for instance in terms of the curvature of the level lines or the total variation of the image. The interpolation is performed by continuing and imposing this model inside the inpainting domain, usually by means of a partial differential equation (PDE). Such PDEs can be derived from

variational principles, as for instance in [BBC+01, CKS02, CS01, ES02, Mas02, MM98], or inspired by physical processes [BSCB00, BM07, TD05]. These methods show good performance in propagating smooth level lines or gradients, but fail in the presence of texture. They are often referred to as *structure* or *cartoon* inpainting.

Geometry-oriented methods are *local* in the sense that the associated PDEs only involve interactions between neighboring pixels on the image grid. An implication of this is that, among all the data available from the image, these methods only use the information around the boundary of the inpainting domain.

Texture-oriented methods. Texture-oriented inpainting was born as an application of texture synthesis, e.g., [EL02, IP97]. Its recent development was triggered in part by the works of Efros and Leung [EL02] and Wei and Levoy [WL00] using non-parametric sampling techniques (parametric models have also been considered, *e.g.* [LZW03]). In these works texture is modeled as a two dimensional probabilistic *graphical model*, in which the value of each pixel is conditioned by its neighborhood. These approaches rely directly on a sample of the desired texture to perform the synthesis.

In practice these methods work progressively by expanding a region of synthesized texture. The value of a target pixel x is copied from the center of a (square) patch in the sample image, chosen among those that best match the available portion of the patch centered at x. This philosophy has been extensively used for inpainting [BVSO03, BLLC02, CPT04, DCOY03, EL02, PGB04]. As opposed to geometry-oriented inpainting, these so-called *exemplar-based* approaches, are *non-local*: To determine the value at x, the whole image may be scanned in the search for a matching patch.

Since these texture approaches are *greedy* procedures (each hole pixel is visited only once), the results are very sensitive to the order in which pixels are processed [CPT04]. This issue was addressed in [KT07, WSI07] where the inpainting problem is stated as a probabilistic inference problem in a graphical model. In both cases, the image is modeled using a pair-wise Markov Random Field (MRF). In [KT07] an energy considering the overlap error of adjacent patches is minimized using *belief propagation*. The method in [WSI07] can be seen as an approximated Expectation-Maximization (EM) method. Finally, the iterative algorithm explored in [KSY09] corresponds to a particular case of our framework.

A different variational justification for texture-based methods was proposed in [DSC03], where the inpainting problem is reformulated as that of finding a *correspondence map* $\Gamma : O \to O^c$, O being the inpainting domain and O^c its complement w.r.t. the image domain. Denoting the image by u, the inpainted value at position $x \in O$ is then given by $u(x) = u(\Gamma(x))$. The authors proposed a continuous energy functional in which the unknown is the correspondence map itself:

$$E(\Gamma) = \int_O \int_{\Omega_p} (u(\Gamma(x-y)) - u(\Gamma(x)-y))^2 dy\, dx,$$

where Ω_p is the patch domain (centered at $(0,0)$). Thus Γ should map a pixel x and its neighbors in such a way that the resulting patch is close to the one centered at $\Gamma(x)$. This model has been the subject of further analysis by Aujol *et al*. [ALM10], where extensions are proposed and the existence of a solution in the set of piecewise roto-translation maps Γ is proved. These approaches are theoretical and no numerical optimization scheme is available so far.

Let us finally note that the texture synthesis works [EL02, WL00] have also influenced the development of non-local methods for other applications, such as denoising [AW06, BCM05], superresolution [PETM09] and regularization of inverse problems [GO07b, PBC09]. As opposed to the case of inpainting, in these contexts the estimation of a pixel value may involve many locations in the image. The resulting correspondence is not one-to-one, but rather one-to-many, usually encoded as a weight function $w : \Omega \times \Omega \to \mathbb{R}$, with Ω being the image domain. For each x, $w(x, \cdot)$ weights the contribution of each image location to the estimation of x.

Exemplar-based methods provide excellent results in recovering textures and repetitive structures. However, their ability to recreate the geometry without any example is limited and not well understood. Therefore, different strategies have been proposed which combine geometry and texture inpainting [BVSO03, CGMP09, DCOY03, JT04]. These methods usually decompose the image in some sort of structure and texture components. Structure is reconstructed using some geometry-oriented scheme, and this is used to guide the texture inpainting.

Contributions. Despite these combined methods, geometry and texture inpainting are still quite separate fields, each one with its own analysis and implementation tools. Variational models as the one introduced in this paper can provide common tools allowing a unified treatment of both trends. We therefore propose a variational framework for non-local image inpainting as a contribution to the modeling and analysis of texture-oriented methods.

Like the *non-local means* denoising algorithm [AW06, BCM05] we encode the image redundancy and self-similarity (measured as patch similarity) as a non-local weight function $w : O \times O^c \to \mathbb{R}$. This function serves as a fuzzy correspondence, and differs from the works [ALM10, DSC03], although a (eventually multivalued) correspondence map can be approximated as a limit of our model. The proposed formulation is rather general and different inpainting schemes can be derived naturally from it, via the selection of the appropriate patch similarity criterion. In this work we present three of them, *patch NL-means*, *-medians* and *-Poisson*, corresponding to similarity criterions based on L^2- and L^1-norms between patches and L^2-norm of their gradients. A fourth inpainting scheme (*patch NL-gradient medians*) is obtained by taking the L^1-norm of the gradients as patch similarity criterion. Although it fits our framework, we will not discuss this method in this chapter, we refer to [AFCS10] for details.

The *patch NL-means* is related to the method of [KSY09, WSI07, KEBK05]. In the context of texture synthesis, the authors of [KEBK05] also proposed a func-

tional that incorporates the gradient of the patch. For inpainting the gradient of the patch, as used in *Patch NL-Poisson*, permits to combine the exemplar-based interpolation with PDE-based diffusion schemes. This results in a smoother continuation of the information across the boundary and inside the inpainting domain and better propagation of the structures. We provide an comprehensive empirical comparison on real and synthetic problems, showing the benefits and limitations of each model.

Although the focus of this work lies on inpainting, the framework we are introducing can be adapted for its application to other contexts. In particular, it is related to graph-based non-local regularizers. Inspired by graph regularization techniques [ZS05], these approaches model the image as a graph characterized by the similarity weights [GO07b, LEB07]. Recently, our formalism [ACS09] was extended by Peyré *et al*. in [PBC09] to provide a variational justification to the case in which the graph is adaptive. On the other hand the patch NL-means scheme provides a model for an iterative version of the non-local means algorithm with adaptive similarity weights. Similar approaches have been applied to denoising [AW06, PMD$^+$10], super-resolution [PETM09], texture denoising [BKC08], and demosaicing [BCMS09] among others. We will discuss the relation with these models in the text.

The rest of this chapter is organized as follows. In Section 6.2 we introduce the proposed variational framework, together with the derivation of the different inpainting schemes. These are analyzed and compared in Section 6.3. The links with related work are discussed in Section 6.4. In Section 6.5 we present and discuss the multiscale approach. And in Section 6.6 we present experimental results on real images allowing to compare our results with the state of the art. Concluding remarks and future work is discussed in Section 6.7.

Notation. Images are denoted as functions $u : \Omega \to \mathbb{R}$, where Ω denotes the image domain, usually a rectangle in $\mathbb{R}^2$. Pixel positions are denoted by x, $\hat{x}$, z, $\hat{z}$ or y, the latter for positions inside the patch. A patch of u centered at x is denoted by $p_u(x) = p_u(x, \cdot) : \Omega_p \to \mathbb{R}$, where Ω_p is a rectangle centered at $(0, 0)$. The patch is defined by $p_u(x, y) = u(x + y)$, with $y \in \Omega_p$. $O \subset \Omega$ is the hole or inpainting domain, and $O^c = \Omega \setminus O$. We still denote by u the part of the image u inside the hole, while $\hat{u}$ is the part of u in O^c: $\hat{u} = u|_{O^c}$. Additional notation will be introduced in the text.

6.2 Variational framework

6.2.1 Preliminaries

Our variational framework is inspired by the non-local functional

$$F_w(u) = \int_O \int_{O^c} w(x, \hat{x})(u(x) - \hat{u}(\hat{x}))^2 \, d\hat{x} \, dx, \tag{6.1}$$

where $w : O \times O^c \to \mathbb{R}^+$ is a weight function that measures the similarity between patches centered in the inpainting domain and in its complement.

Let us assume for the moment that the weights are known. The minimum of (6.1) should have a low pixel error $(u(x) - \hat{u}(\hat{x}))^2$ whenever the similarity $w(x, \hat{x})$ is high. In this way the similarity weights drive the information transfer from known to unknown pixels. A similar functional was proposed in [GO07a] as a non-local regularization energy in the context of image denoising. It models the non-local means filter [AW06, BCM05] when the weights are *Gaussian*

$$w(x, \hat{x}) \propto \exp\left(-\frac{1}{h}\|p_u(x) - p_{\hat{u}}(\hat{x})\|^2\right),$$

where $\|\cdot\|$ is a weighted L^2-norm in the space of patches and h is a parameter that determines the selectivity of the weigths w.

In [GO07a] the weights w are considered to be known and remain fixed through all the iterations. While this might be appropriate in applications where they can be estimated from the noisy image, in the image inpainting scenario here addressed, the weights are not available and have to be inferred together with the image (as in [PBC08, PETM09]). One of the novelties of the proposed framework is the inclusion of adaptive weights in a variational setting.

6.2.2 The proposed formulation

For this reason, we will consider the weight function w as an additional unknown. Instead of prescribing explicitly the Gaussian functional dependence of w w.r.t. u, we will do it implicitly, as a component of the optimization process. In doing so, we obtain a simpler functional, avoiding to deal with the complex, non-linear dependence between w and u. In our formulation, $w(x, \cdot)$ is a probability density function, and can be seen as a relaxation of the one-to-one correspondence map of [ALM10, DSC03], providing a fuzzy correspondence between each $x \in O$ and the complement of the inpainting domain.

In this setting, we propose an energy which contains two terms, one of them is inspired by (6.1) and measures the coherence between the pixels in O and those in O^c, for a given similarity weight function w. This permits the estimation of the image u from the weights w. The second term allows us to compute the weights given the image. The complete proposed functional is

$$\begin{aligned} E(u, w) &= \frac{1}{h}\widetilde{F}_w(u) - \int_{\widetilde{O}} H_w(x)\, dx, \\ \text{subject to } &\int_{\widetilde{O}^c} w(x, \hat{x})\, d\hat{x} = 1 \quad \forall x \in \widetilde{O}, \end{aligned} \tag{6.2}$$

where

$$\widetilde{F}_w(u) = \int_{\widetilde{O}} \int_{\widetilde{O}^c} w(x, \hat{x}) \varepsilon(p_u(x) - p_{\hat{u}}(\hat{x}))\, d\hat{x}\, dx, \tag{6.3}$$

$\varepsilon(\cdot)$ is an error function for image patches, and

$$H_w(x) = -\int_{\widetilde{O}^c} w(x, \hat{x}) \log w(x, \hat{x})\, d\hat{x},$$

is the entropy of the probability $w(x, \cdot)$.

We take $\widetilde{O}$, the *extended inpainting domain*, as the set of centers of patches that intersect the hole, i.e. $\widetilde{O} = O + \Omega_p = \{x \in \Omega : (x + \Omega_p) \cap O \neq \emptyset\}$. Thus, patches $p_{\hat{u}}(\hat{x})$ centered at $\hat{x} \in \widetilde{O}^c$ are entirely outside O (Figure 6.1). Accordingly, we consider that the weight function w is defined over $\widetilde{O} \times \widetilde{O}^c$ and $\int_{\widetilde{O}^c} w(x, \hat{x})\, d\hat{x} = 1$. For a simplified presentation, we assume that $\widetilde{O} + \Omega_p \subseteq \Omega$, i.e.every pixel in $\widetilde{O}$ supports a patch centered on it and contained in Ω. Analogously, we also shrink $\widetilde{O}^c$ to have $\widetilde{O}^c + \Omega_p \subseteq \Omega$.

Let us now make some additional comments on the functional. We observe that the term $(u(x) - \hat{u}(\hat{x}))^2$ in (6.1), that penalizes differences between pixels, is substituted in (6.3) by the patch error function $\varepsilon(p_u(x) - p_{\hat{u}}(\hat{x}))$. This has two consequences. First, minimizing (6.3) with respect to the image will force patches $p_u(x)$ to be similar to $p_u(\hat{x})$ for each pair x, $\hat{x}$ such that $w(x, \hat{x})$ is large. The other implication has to be understood together with the inclusion of the second term, which integrates the entropy of each probability $w(x, \cdot)$ over $\widetilde{O}$. For a given completion u, and for each $x \in \widetilde{O}$, the optimum weights minimize the mean patch error for $p_u(x)$, given by

$$\int_{\widetilde{O}^c} w(x, \hat{x}) \varepsilon(p_u(x) - p_{\hat{u}}(\hat{x}))\, d\hat{x},$$

while maximizing the entropy. For instance taking ε as the squared L^2-norm of the patch, then the resulting weights are Gaussian. This can be related to the *principle of maximum entropy* [Jay57], widely used for inference of probability distributions. The parameter h controls the trade-off between both terms and is also the selectivity parameter of the Gaussian weights. The trivial minima of E with $w(x, \hat{x}) = 0$ everywhere is discarded by restricting $w(x, \cdot)$ to be a probability.

The patch error function. Patches are functions defined on Ω_p, and the error function $\varepsilon : \mathbb{R}^{\Omega_p} \to \mathbb{R}^+$ is defined as the weighted sum of pixel-wise errors $\mathrm{e} : \mathbb{R} \to \mathbb{R}^+$

$$\varepsilon(p_u(x) - p_{\hat{u}}(\hat{x})) := \int_{\Omega_p} g_a(y) \mathrm{e}(u(x + y) - \hat{u}(\hat{x} + y))\, dy, \tag{6.4}$$

where the *intra-patch* weight function g_a is a Gaussian centered at the origin with standard deviation a. We will consider the L^1- and the squared L^2-norms as particular cases of $\varepsilon(\cdot)$, with $\mathrm{e}(\cdot) = |\cdot|$ and $\mathrm{e}(\cdot) = |\cdot|^2$ respectively.

We will also consider a patch error function involving the gradient of the image. As an abuse of notation we will denote the gradient's patch and pixel-wise error functions as $\varepsilon : \mathbb{R}^{2^{\Omega_p}} \to \mathbb{R}^+$ and $\mathrm{e} : \mathbb{R}^2 \to \mathbb{R}^+$, respectively. It will be clear from the argument which case is intended, as in this example $\varepsilon(p_{\nabla u}(x) - p_{\nabla u}(\hat{x})) := \int_{\Omega_p} g_a(y) \mathrm{e}(\nabla u(x + y) - \nabla \hat{u}(\hat{x} + y))\, dy$.

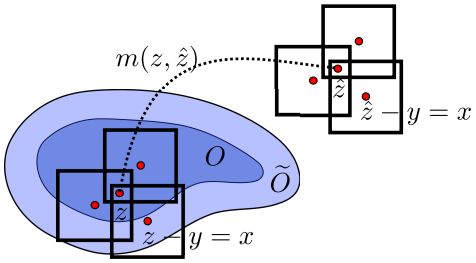


Figure 6.1: Patch-wise non-local inpainting. The value at $z \in O$ is computed using contributions from all the patches that overlap it, these are patches centered at $x \in \widetilde{O}$ such that $z = x + y$ with $y \in \Omega_p$. The influence function $m(z, \hat{z})$ accumulates all the contributions $w(z-y, \hat{z}-y)$ from patches centered at $\hat{z}-y$ to $z-y$.

As in [KSY09] we can modify the patch error function to account for the distance between the patch centers $\varphi(x-\hat{x})$ just by adding it to the patch error function $\varepsilon_d(p_u(x) - p_u(\hat{x}), x, \hat{x}) = \varepsilon(p_u(x) - p_u(\hat{x})) + \varphi(x-\hat{x})$, so that the use of distant patches is penalized. We will not consider this modification in the present document.

As it will be discussed below, the patch error function determines not only the similarity criterion but also the image synthesis, and thus is a key element in the proposed framework.

6.2.3 Minimization of E

We have formulated the inpainting problem as the constrained optimization

$$(u^*, w^*) = \arg\min_{u,w} E(u,w) \qquad (6.5)$$
$$\text{subject to } \int_{\widetilde{O}^c} w(x,\hat{x})\, d\hat{x} = 1 \quad \forall x \in \widetilde{O},$$

where E is the inpainting energy defined in (6.2).

To minimize the energy E, we use an alternate minimization algorithm. At each iteration, two optimization steps are solved: The constrained minimization of E with respect to w while keeping u fixed; and the minimization of E with respect to u with w fixed. This procedure yields the following iterative scheme:

In the weights update step, the minimization of E w.r.t. w yields:

$$w_k(x,\hat{x}) = \frac{1}{q(x)} \exp\left(-\frac{1}{h}\varepsilon(p_{u_k}(x) - p_{\hat{u}}(\hat{x}))\right), \qquad (6.6)$$

where $q(x) = \int_{\widetilde{O}^c} \exp\left(-\frac{1}{h}\varepsilon(p_{u_k}(x) - p_{\hat{u}}(\hat{x}))\right)$ is a normalization factor that makes $w(x,\cdot)$ a probability. For patch NL-poisson the weight update equation is analogous to (6.6) replacing the patch error function with $\varepsilon(p_{\nabla u_k}(x) - p_{\nabla \hat{u}}(\hat{x}))$.

Algorithm 1 Alternate minimization of $E(u, w)$.

Require: *Initial Condition* $u_0(x)$ with $x \in O$.
1: **repeat**
2: *Update Weights*: $w_k = \arg\min_w E(u_k, w)$, s.t. $\int_{\widetilde{O}^c} w(x, \hat{x})\, d\hat{x} = 1$.
3: *Update Image*: $u_{k+1} = \arg\min_u E(u, w_k)$.
4: **until** *Stop Criterion*: $\|u_{k+1} - u_k\| <$ tolerance.

The parameter h determines the selectivity of the similarity. If h is large, maximizing the entropy becomes more relevant, yielding weights which are less selective. In the limit, when $h \to \infty$, $w(x, \cdot)$ becomes a uniform distribution over $\widetilde{O}^c$. On the other hand, a small h yields weights which concentrate on the patches similar to $p_u(x)$. When $h \to 0$, we compute the weights as $\lim_{h\to 0} w(x, \hat{x}) = \frac{1}{|n(x)|}\delta(\hat{x} - n(x))$, where $n(x) \subseteq \widetilde{O}^c$ is the set of nearest neighbors of x, defined as

$$n(x) = \arg\min_{\hat{x}\in\widetilde{O}} \varepsilon(p_u(x) - p_{\hat{u}}(\hat{x})).$$

That is $w(x, \cdot)$ can be considered as an approximation to a multivalued correspondence. For simplicity, in practice we assume that $|n(x)| = 1$, i.e.the nearest neighbor is unique.

The image update step deserves more attention and is described next.

Image update step

In this section we present the derivation of the image update step corresponding to the three patch error functions mentioned earlier. First we will present the cases when image patches are compared using the squared L^2-norm and the L^1-norm. We refer to the resulting algorithms as *patch-wise non-local means* (patch NL-means), and *medians* (patch NL-medians). Then we consider a functional involving the the squared L^2-norm of the gradients of the patches, which will be referred as *patch-wise non-local Poisson* (patch NL-Poisson).

Before moving to the derivation of the these schemes, let us remark that with the change of variables $z = x + y$, $\hat{z} = x + y'$, the image energy term can be expressed as an accumulation of pixel-wise errors:

$$\begin{aligned}\widetilde{F}_w(u) &= \int_{\widetilde{O}} \int_{\widetilde{O}^c} w(x, \hat{x}) \int_{\Omega_p} g_a(y) \mathrm{e}(u(x+y) - \hat{u}(\hat{x}+y))\, dy\, d\hat{x}\, dx \\ &= \int_O \int_{O^c} m(z, \hat{z}) \mathrm{e}(u(z) - \hat{u}(\hat{z}))\, d\hat{z}\, dz + \mathrm{C}, \end{aligned} \tag{6.7}$$

where C is a constant term. We have introduced the *pixel-wise influence weights* $m : O \times O^c \to \mathbb{R}^+$ defined as

$$m(z, \hat{z}) := \int_{\Omega_p} \chi_{\widetilde{O}^c}(\hat{z} - y) g_a(y) w(z - y, \hat{z} - y)\, dy. \tag{6.8}$$

The function $\chi_{\widetilde{O}^c}$ takes the value 1 on $\widetilde{O}^c$ and 0 on $\widetilde{O}$. An analogous expression can be computed for gradient patch error functions. This rewriting simplifies the following derivations and provides some insights on the implications of using patch-wise errors.

For each pair of pixels $(z, \hat{z}) \in O \times O^c$, $m(z, \hat{z})$ weights the effective contribution of the pixel-wise error between $u(z)$ and $\hat{u}(\hat{z})$ in the total value of the energy. The quantity $m(z, \hat{z})$ is computed by integrating the similarity $w(z - y, \hat{z} - y)$ between all patches that overlap $\hat{z}$ and those that overlap z in the *same relative position* (shown in Figure 6.1). It tells us how much evidence there is supporting a correspondence between the locations z and $\hat{z}$. Also observe that for each z, $\int_{O^c} m(z, \hat{z}) d\hat{z} = 1$, then $m(z, \cdot)$ can also be interpreted as a probability density function. Note that the energy (6.7) corresponds to (6.1), with the patch similarity weights w being substituted with the pixel-wise influence weights m, a sort of spatial convolution of w with kernel g_a.

Patch non-local means

If we use a weighted squared L^2-norm as a patch error function $\varepsilon(p_u(x) - p_{\hat{u}}(\hat{x})) := \int_{\Omega_p} g_a(y) |u(x + y) - \hat{u}(\hat{x} + y)|^2 dy$ in (6.3) then the image energy term (6.7) is quadratic on u, and its minimum for fixed weights w can be computed explicitly as a non-local average:

$$u(z) = \frac{1}{c(z)} \int_{O^c} m(z, \hat{z}) \hat{u}(\hat{z}) \, d\hat{z}, \tag{6.9}$$

for $z \in O$, where the normalization constant $c(z) := \int_{O^c} m(z, \hat{z}) d\hat{z}$. Although $c(z) = 1$ with the current definition of w and g_a, we will keep a generic notation in the following derivations.

The formal similarity with the non-local means equation hides some important differences, which are a direct consequence of the use of a patch error function in the image energy term. To obtain more insight about this let us expand m to obtain:

$$u(z) = \frac{1}{c(z)} \int_{\Omega_p} g_a(y) \int_{\widetilde{O}^c} w(z - y, \hat{x}) \hat{u}(\hat{x} + y) \, d\hat{x} \, dy.$$

There are two averaging processes involved in the synthesis. The outer integral goes through all patches $p_u(z - y)$ overlapping the target pixel z. Each patch *suggests* a value for z resulting from the inner sum: A non-local average of the pixel at position y in all patches $p_{\hat{u}}(\hat{x})$ in $\widetilde{O}^c$. This sum is weighted by the similarity between the patch $p_u(z - y)$ and each $p_{\hat{u}}(\hat{x})$.

Therefore, we can distinguish two types of pixel interactions. Interactions due to the patch overlap of nearby pixels in the image lattice and non-local interactions driven by the similarity weights. The latter can be controlled by the selectivity parameter h, but the extent of the overlap interactions is given by the

patch size. In particular when $h \to 0$ Equation (7.12) yields

$$u(z) = \int_{\Omega_p} g_a(y)\hat{u}(n(z-y)+y)\,dy$$

(recall we are assuming a unique nearest neighbor). For each $x \in \widetilde{O}$, the nearest neighbor of $p_u(x)$ is centered at x and all these patches are averaged according to g_a. This blending may cause some blur, which leads us to consider the L^1-norm in the search of a more robust image synthesis.

Patch non-local medians

The L^1-norm patch error function in the image energy term corresponds to taking $\mathrm{e}(x) = |x|$ in (6.7). The Euler equation for u, given the influence function m, can be formally written as

$$[\delta_u \widetilde{F}_w(u)](z) = \int_{O^c} \mathrm{sign}[u(z) - \hat{u}(\hat{z})]m(z,\hat{z})\,d\hat{z} \ni 0.$$

This expression is multivalued, since $\mathrm{sign}(r) = r/|r|$ if $|r| > 0$ and $\mathrm{sign}(r) \in [-1,1]$ if $r = 0$. Its solution for each $u(z), z \in O$ is obtained as a weighted median of the pixels of the complement O^c, with weights $m(z,\cdot)$.

Both schemes presented so far perform inpainting by transferring (by averages or medians) known gray levels into the inpainting domain. As we will see next, using a patch error function based on the gradient of the image yields a method which transfers gradients and compute the resulting image as the solution of a PDE. This results in better continuation properties of the solution, in particular at the boundary of the inpainting domain.

Patch non-local Poisson

The squared L^2-norm of the gradient in the image energy term (6.3) corresponds to taking $\mathrm{e}(\cdot) = \|\cdot\|^2$ in (6.7), where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^2$. The energy term becomes

$$\widetilde{F}_w(u) = \int_O \int_{O^c} m(z,\hat{z})\|\nabla u(z) - \nabla\hat{u}(\hat{z})\|^2\,d\hat{z}\,dz. \tag{6.10}$$

For this problem we will consider Dirichlet conditions at the boundary between the inpainting domain and the image, and Neumann conditions at the boundary of the image, i.e.$u(x) = \hat{u}(x)$ in $\partial O \setminus \partial\Omega$ and $\nabla u(x)\cdot n(x) = 0$ for $x \in \partial O \cap \partial\Omega$. The Euler equation w.r.t. u is given by a *non-local* Poisson equation, i.e.a Poisson equation with non-local coefficients:

$$\nabla\cdot[c(z)\nabla u(z)] = \nabla\cdot[c(z)\boldsymbol{v}(z)], \tag{6.11}$$

for all $z \in O$, where $c(z) = \int_{O^c} m(z,\hat{z})d\hat{z}$ and the field $\boldsymbol{v} : O \to \mathbb{R}^2$ is given by

$$\boldsymbol{v}(z) = \frac{1}{c(z)}\int_{O^c} m(z,\hat{z})\nabla\hat{u}(\hat{z})\,d\hat{z}. \tag{6.12}$$

Observe that the solutions of (6.11) are minimizers of

$$\int_O c(z)\|\nabla u(z) - \boldsymbol{v}(z)\|^2\, dz.$$

Therefore, u is computed as the image with the closest gradient (in the L^2 sense) to the *guiding* vector field $\boldsymbol{v}$: a non-local weighted average of the gradients in the complement. See [PGB03] for further uses of the Poisson equation in image editing and [SJZ05] for an application to exemplar-based inpainting.

Patch NL-Poisson computes the similarity weights w based only on the gradients (and thus also the pixel-wise influence weights m). In most cases however, the gradient is not a good feature for measuring the patch similarity, and it would be desirable to consider also the gray level data.

We can use the gradient-based energy in conjunction with the patch NL-means energy by considering a linear combination of both patch error functions. This yields

$$\widetilde{F}_\lambda(u) = \int_O \int_{O^c} m(z,\hat{z})[(1-\lambda)\|\nabla u(z) - \nabla \hat{u}(\hat{z})\|^2 + \lambda|u(z) - \hat{u}(\hat{z})|^2]\, d\hat{z}\, dz, \tag{6.13}$$

where the parameter $\lambda \in [0,1]$ controls the mixture. When $\lambda \in (0,1)$ the resulting algorithms update the weights w considering both the intensity and the gradient, therefore improving their selectivity. Notice that (6.13) can be rewritten as

$$\widetilde{F}_\lambda(u) \propto \int_O \int_{O^c} m(z,\hat{z})\|\nabla u(z) - \nabla \hat{u}(\hat{z})\|^2\, d\hat{z}\, dz + \frac{\lambda}{1-\lambda}\int_O c(z)|u(z) - f(z)|^2\, dz,$$

where $f(z) = c(z)^{-1}\int_{O^c} m(z,\hat{z})\hat{u}(\hat{z})\, d\hat{z}$ is the solution of the patch NL-means image update step. Thus we see that the combination with the squared L^2 patch error function translates into a patch NL-means attachment term.

The Euler equation w.r.t. u becomes:

$$\nabla\cdot[c(z)\nabla u(z)] - \frac{\lambda}{(1-\lambda)}c(z)u(z) = \nabla\cdot[c(z)\boldsymbol{v}(z)] - \frac{\lambda}{(1-\lambda)}c(z)f(z), \tag{6.14}$$

which is linear and can be solved with a conjugate gradient scheme.

6.3 Comparison of the proposed schemes

Let us advance some results on synthetic problems to highlight the main characteristics of the proposed methods. First we consider the inpainting of a regular texture (shown in Figure 6.2) with two different mean intensities, where the inpainting domain hides all patches on the boundary between the dark and bright textures. With this example we can test the ability of each method to *create* an interface between both regions. Situations like these are common in real inpainting problems, for instance due to inhomogeneous lighting conditions. We have

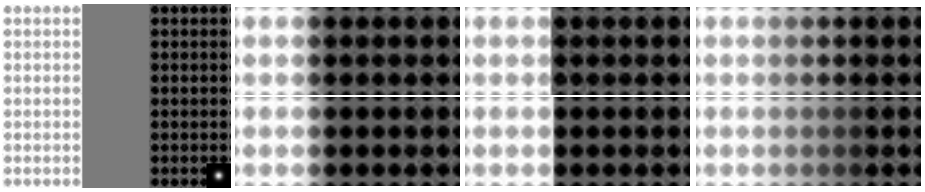


Figure 6.2: Inpainting of a synthetic texture. The initial condition is shown in the first column. The other three columns show a zoom (region in the red rectangle) of the results of patch NL-means, -medians and -Poisson. Top row, $h = 0$, bottom row $h = 200$, $h = 14$ and $h = 400$, respectively. The intra-patch weight kernel g_a is shown in the bottom right corner of the initial condition, it has a standard deviation $a = 5$ and the patch size is $s = 15$.

also added Gaussian noise with standard deviation $\sigma = 10$ to show the influence of the selectivity parameter h. Each column of Figure 6.2 shows the results of the three methods described in the previous section. We have tested each method with $h = 0$ (top row), and $h > 0$, chosen approximately to match the expected deviation of each patch error due to the noise.

The first notorious difference is on how the methods handled the transition between the dark and bright textures. Patch NL-means produces a smooth transition whereas a sharp step is obtained with the patch NL-medians. On the other hand, both gradient based methods yield a much smoother shading of the texture. This is due to the fact that the image update step is computed as the solution of a PDE which diffuses the intensity values present at the boundary of the inpainting domain. These PDEs are driven by a vector field estimated non-locally and therefore combine *non-local* exemplar based inpainting with *local* interpolation PDEs. For the case of patch NL-Poisson this interpolation is linear, since this is a solution of the homogeneous Poisson equation (i.e.Laplace equation).

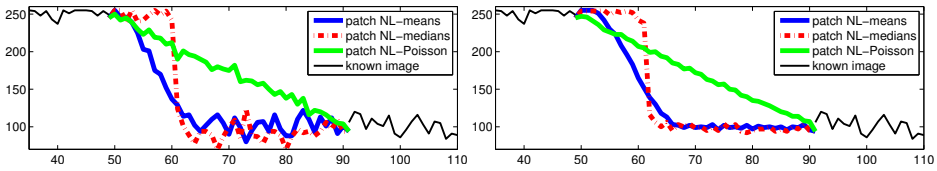


Figure 6.3: Profiles of the results in Figure 6.2. The profiles are taken from an horizontal line going between the circles in Figure 6.2. Left: results with $h = 0$ and right: results with $h > 0$.

As expected, the results using a higher h show some denoising, since for larger h the patch similarity weights are less selective. This effect can be better

appreciated in the profiles shown in Figure 6.3, which depict the image values for a horizontal line between the circles. In the usual context of inpainting, in which the available data is not perturbed by noise, this denoising translates into an undesirable loss of texture quality. For that reason we use $h = 0$.

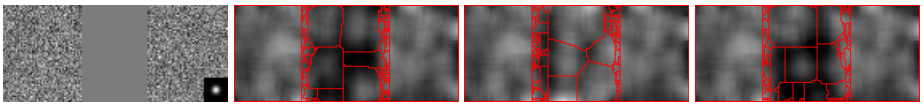


Figure 6.4: Inpainting of random texture. From left to right: initial condition (the gray square is the hole); images of the local variance superimposed with the boundaries between copy regions (red curves) for patch NL-means, -medians and -Poisson. Notice the attenuation of the variance over the red curves, specially for L^2-based methods.

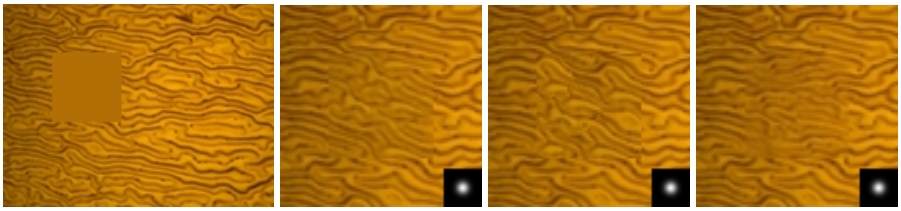


Figure 6.5: Inpainting of structured texture. From left to right: Initial condition, result of patch NL-means, -medians and -Poisson. Results with $s = 15$, $a = 5$ and $h = 0$. The treatment of color images is described in Section 6.6.1.

Recall that in the limit when $h = 0$, the weights $w(x, \hat{x})$ converge to a Dirac's delta function at the set $n(x)$ of nearest neighbors of $p_u(x)$. Even if we assume that the nearest neighbor is unique, the value of a pixel is still computed from a population obtained from those nearest neighbors of all patches that overlap the pixel. In the case of periodic patterns, once the minimization has reached a stable state, all values in the population will be basically the same: patches surrounding a pixel agree on its value, and in this case all schemes behave similarly. Differences arise when neighboring patches cannot agree on their suggested values. Such is the case of the step in Figure 6.2. For non-periodic patterns and random textures this disagreements will be common, which may affect the perceptual similarity of the synthesized texture with the original.

Transitions

Figure 6.4 studies this effect. The first image consists of Gaussian noise, and the remaining images show an estimate of the local variance (computed by smoothing the image of the squared differences w.r.t. the mean) for the completion with the different methods. The red curves require a brief explanation. Following

[BSFG09], we use the term *Nearest Neighbor Field* (NNF) to refer to the vector field $n(x) - x$, defined over $\widetilde{O}$, where $n(x) \in \widetilde{O}^c$ is the position of the (assumed) unique nearest neighbor of $p_u(x)$. In Figure 6.4 we show in red the boundaries of the regions with constant NNF. In those regions patches are translated rigidly from somewhere in the complement. We can observe that, for all inpainting schemes, the resulting completion consists basically of a *patchwork* of large regions of rigid translation. The interior of these regions, away from their boundaries (red curves), reproduces an exact copy of the source, and thus the variance is preserved. The variance decreases along the red curves, where pixel values are synthesized using incoherent contributions. Both for the intensity models or the gradient models, the L^2-norm causes a higher decay of the variance than with the L^1-norm. This effect can also be observed in Figures 6.2 and 6.3, where it can be seen that patch NL-medians and -GM preserve more of the Gaussian noise than the methods based on L^2-norms. We will refer to the regions of constant NNF as *copy fronts*. Recall however that copy occurs only away from their boundaries.

In Figure 6.5 we show results on a non-periodic, structured texture. We can see that patch NL-means and -Poisson show some smoothing, whereas patch NL-medians obtained sharper results. Both intensity based methods show discontinuities at the boundary of the inpainting domain. Notice also that for the patch NL-medians algorithm it is easy to identify the regions in the complement that have been replicated. Where two copy fronts meet a seam is produced. With the patch NL-means the spatial averaging of overlapping patches creates a smooth blending of the copy fronts.

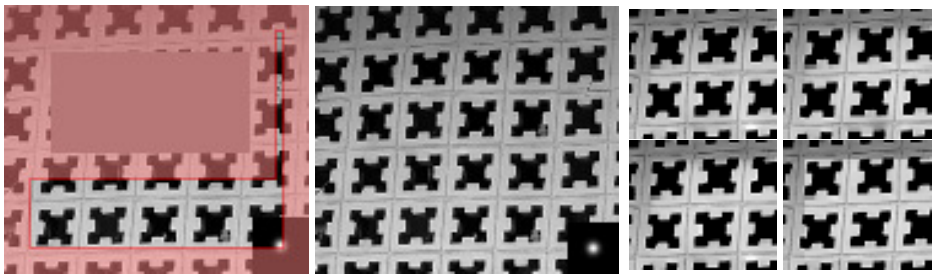


Figure 6.6: Linear combination of patch NL-Poisson with patch NL-means. First image starting from the left: Initialization. The gray rectangle is the inpainting domain. Only patches centered outside the red area are available. Second image: Result obtained with patch NL-Poisson, using $\lambda = 0$. Details: From left to right, top to bottom: results with patch NL-Poisson using $\lambda = 0.01, 0.05, 0.1, 1$ (the latter corresponds to patch NL-means).

Combined schemes

Patch NL-Poisson produces smooth interpolations and enforces the continuity of the image at the boundary of the inpainting domain, which are generally desirable features. For this to be done variationally the similarity weights w must be computed using patches of the gradient, which in most cases does not provide a reliable measure of patch similarity. In practice better results are obtained by combining it with patch NL-means, allowing to take the image values into account for the computation of the patch error, and to synthesize the image with a diffusion PDE. This adds a parameter λ, which controls the mixture.

In Figure 6.6 we show some results corresponding to the combination of patch NL-Poisson with patch NL-means while varying the mixture coefficient λ. The image shows a periodic pattern with an illuminance gradient. Most of the dark exemplars are incomplete, and thus only bright exemplars from the bottom of the image are available. The small images on the right show results of both gradient methods with different values of λ. One of the details corresponds to the result of patch NL-means: the image has been completed using bright patches and presents a discontinuity on the upper side of the hole.

The value of the mixing parameter λ should be carefully selected since it mixes two different magnitudes (norms of gradients and gray levels), with $\lambda \sim 0.1$ some of the good continuation properties are preserved and enough color information is added to the patch comparison criterion.

In general one would like to synthesize the image using gradients and to compute the patch similarity using only intensity. Since this is not feasible within our variational framework we had to introduce these schemes that combine both gradient and intensity patch error functions. However one can still envision a broader family of non variational methods in which the patch comparison criterion can be chosen independently from the synthesis method. Although these new algorithms do not have a variational justification, some of them result useful in practice, as explored in [PMD$^+$10] and [AFCS10].

Geometric interpolation

To evaluate the ability of each method to continue the geometry of the structures at the boundary of the inpainting domain, we consider a very simple image with a gap (shown in Table 6.1). The inpainting region is initialized with the background color. For this evaluation we fix the size of the patch and increase the width of the gap. For narrow gaps, the method will be able to join both ends of the green vertical line. When increasing the gap, at a certain width the gray initialization prevails and the method is no longer capable of recovering the vertical line. The first column of Table 6.1 shows the maximum width in pixels that produces a good reconstruction. Observe that the combined schemes propagate the line through wider gaps, and that the optimal mixing parameter λ for this purpose is around the values that we have proposed in the previous section. Basically these schemes have two propagation mechanisms: A *local* one, by diffusion of the intensity values by the PDE, and a *non-local* one by transference

Method	GAP	GAP ($t_c = 5$)
patch NL-Poisson with $\lambda = 0$	13	40
patch NL-Poisson with $\lambda = 0.1$	16	46
patch NL-Poisson with $\lambda = 0.5$	11	34
patch NL-Means	9	29
patch NL-Medians	7	42

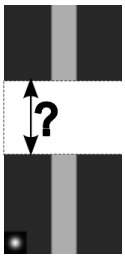


Table 6.1: Geometric interpolation. The inpainting domain is shown in white and the patch in the lower right corner (9×9 pixels). The table reports the maximum gap width (measured in pixels) for which the algorithm is capable of recovering the vertical line. The rightmost column shows the maximal gaps obtained with the use of a confidence masks with decay $t_c = 5$ (see Section 6.3.1).

of gradients from O^c. When $\lambda > 0$ these mechanisms reinforce each other: the diffused values allow a better estimation of the weights, and therefore the transference of more appropriate gradients, which will help to diffuse the intensity values further. On the other hand, intensity based methods depend only on the iteration of weights computation and image update to propagate information.

The second column of Table 6.1 shows the results obtained by incorporating the confidence mask later described in Section 6.3.1. An alternative way to prolong the geometric structures is to increase the patch size as will be discussed in Section 6.5.

6.3.1 Extensions

Color images. An energy for color images can be obtained by defining a patch error function for color patches as the sum of the error functions of the three scalar components:

$$\varepsilon(p_{\mathbf{u}}(x) - p_{\hat{\mathbf{u}}}(\hat{x})) = \sum_{i=1}^{3} \varepsilon(p_{u_i}(x) - p_{\hat{u}_i}(\hat{x})),$$

where $\mathbf{u} : \Omega \to \mathbb{R}^3$ is the color image, and u_i, with $i = 1, 2, 3$, its components (analogously for gradient-based errors). Given the weights, each channel is updated using the corresponding scheme for scalar images. All channels are updated using the same weights.

Confidence mask. For large inpainting domains, it is useful to introduce a mask $\kappa : \Omega \to (0, 1]$ which assigns a confidence value to each pixel, depending on the certainty of its information (see also [CPT04, KT07]). This will help in guiding the flow of information from the boundary towards the interior of the hole, eliminating some local minima and reducing the effect of the initial

condition. The resulting image energy term takes the form

$$\widetilde{F}_w(u) = \int_{\widetilde{O}} \int_{\widetilde{O}^c} \kappa(x) w(x, \hat{x}) \varepsilon(p_u(x) - p_{\hat{u}}(\hat{x}))\, d\hat{x}\, dx,$$

where κ modulates the penalization of the incoherences between w and the error ε between patches.

The effect of κ on the image update step can be seen on the pixel-wise influence weights m

$$m(z, \hat{z}) = \int_{\Omega_p} \chi_{\widetilde{O}^c}(\hat{z} - y) g_a(y) \kappa(z - y) w(z - y, \hat{z} - y)\, dy.$$

Thus, the contribution of the patch $p_u(z - y)$ to the evidence function is now weighted by its confidence. Patches with high confidence will have a stronger influence. Notice that with the inclussion of the confidence mask, the normalization coefficient $c(z)$ becomes:

$$c(z) = \int_{O^c} m(z, \hat{z})\, d\hat{z} = \int_{\Omega_p} g_a(y) \kappa(z - y)\, dy.$$

This does not affect intensity-based methods, but has implications on the image update step for gradient-based ones. For example, for patch NL-Poisson the errors with respect to the non-local guiding vector field $\|\nabla u(z) - \boldsymbol{v}(z)\|$ are penalized according to $c(z)$.

On the similarity weights, the confidence mask has the effect of modifying the selectivity parameter h by a locally varying $h/\kappa(x)$. If the confidence is high, the effective selectivity $h/\kappa(x)$ will be lower, thus increasing the selectivity of the similarity measure. When $h \to 0$ the weights tend to a Dirac's delta independently of κ. The same reasoning applies to the gradient based energies.

For the experiments shown in this paper, the confidence mask was set to

$$\kappa(x) = \begin{cases} (1 - \kappa_0) \exp\left(-\frac{d(x, \partial O)}{t_c}\right) + \kappa_0 & \text{if } x \in O, \\ 1 & \text{if } x \in O^c, \end{cases}$$

which shows an exponential decay w.r.t. the distance to the boundary inside the hole $d(\cdot, \partial O)$. Here $t_c > 0$ is the decay time and $\kappa_0 > 0$ determines the asymptotic value reached far away from the boundary. Setting $t_c = 0$ amounts to using a constant confidence mask. Table 6.1 shows the effect of using a confidence mask with $t_c = 5$ and $\kappa_0 = 0.1$, allowing the restoration of the vertical line for much wider gaps, and thus alleviating the dependence with the gray initial condition.

Other shapes of the confidence mask could be used for controlling different aspects of the dynamics of the completion algorithm. For instance controlling the decay of the mask from certain points of the boundary allows us to privilege the continuation of structures from them.

6.4 Discussion and connections

In this section we will briefly review the connections of our work with other inpainting algorithms and also with existing variational models of non-local regularization which have been proposed in contexts such as image denoising.

The method in [WSI07] is closely related to the patch NL-means scheme of Equation (7.12). The key difference lies in the underlying theoretical model. The problem is addressed as a MRF, where pixels outside the hole are observable variables, missing pixels in the hole are the parameters, and the hidden variables are given by the correspondence $\Gamma : O \to O^c$, which assigns a patch outside the hole to each x in O. The method can be seen as an approximate EM algorithm for maximizing the log-likelihood w.r.t. the pixels in O, and some approximations have to be taken to make the optimization tractable. Based on heuristics, the authors also propose to use more robust estimators than the mean for the synthesis of pixels. Within the framework here proposed, robust estimators (as the median) naturally result from particular choices of the patch error functions $\varepsilon(\cdot)$.

In [KSY09] an energy is presented which can be seen as a limit case of the patch NL-means energy when $h \to 0$. The authors propose modifications of the energy which improve the results, such as some spatial localization of the similarity weights and brightness invariance. The latter is achieved by introducing a multiplicative constant that matches the mean illuminance between each pair of patches.

In the following we will comment on the relation of this model with recent works on non-local regularization.

The UINTA algorithm, presented in [AW06] is a non-local denoising algorithm that minimizes the entropy of the patches in the image. Casting this idea to the context of inpainting the UINTA's entropy is estimated as the sample mean

$$E_U(u) = -\int_{\widetilde{O}} \log\left[\int_{\widetilde{O}^c} \exp(-\frac{1}{h}\|p_u(x) - p_{\hat{u}}(\hat{x})\|^2)\, d\hat{x}\right] dx,$$

where the inner integral is the probability of occurrence of the patch $p_u(x)$ obtained as a Parzen density estimate. The corresponding Euler-Lagrange equation can be solved with a fixed point iteration which coincides with the patch NL-means scheme (7.12). In [AW06] this energy is minimized by considering all patches as independent (disregarding the overlap between neighboring patches), and evolving each of them according to a gradient descent of E_U. After this, an image is formed with the centers of these new patches. The repetition of this process results in an iterative application of *pixel-wise* NL-means.

In [BKC08] the authors use a variational principle for deriving the iterated pixel NL-means regularizer, and show its application to the restoration of texture. The underlying energy corresponds to the quadratic penalty between the solution image u, and a pixel NL-means type average of the noisy input image $\hat{u}$. The weights for this average are computed using u. Due to the dependence of the weights with the regularized image u, the minimizer is no longer a weighted average as NL-means, but the solution of a nonlinear optimization problem. It is

shown that if the derivative of the nonlinear component is neglected, the resulting Euler-Lagrange equation matches the proposed fixed point algorithm: The iterated NL-means regularizer.

In [PMD+10] the authors presented a variational framework for image denoising consisting in non-local regularization and data adjustment terms. Inpainting could be performed by considering only the data term as follows:

$$E_P(u) = -\int_{\widetilde{O}} \int_{\widetilde{O}^c} \exp(\varepsilon(p_u(x) - p_{\hat{u}}(\hat{x}))) \, d\hat{x} \, dx.$$

This energy is the same as the one adapted from the UINTA algorithm E_U, without the logarithm. In [PMD+10] the Euler-Lagrange equation is solved with a fixed point iteration. This model has two differences with our framework. First it allows to use a more general nonlinearity for the computation of the weights other than the exponential. Second, even in the case of the exponential, the methods differ in the normalization, for instance, when ε is the squared L^2-norm, the resulting scheme is as the patch NL-means, with the unnormalized weights.

After its introduction in [ACS09], our model has been interpreted as a non-local self-similarity regularizer in [PBC09], where in conjunction with appropriate data fitting terms it has been applied to the solution of inverse problems, including inpainting, super-resolution and compressive sensing. In [PBC09] a different patch-error function ε is used, namely the L^2-norm between patches (without squaring it). This choice is motivated as a patch-wise version of their work [PBC08] on non-local Total Variation [GO07b, LEB07, ZS05] with adaptive weights. This patch-wise non-local TV is defined as the L^1-norm of the non-local gradient of the patch valued image $p_u : \Omega \to \mathbb{R}^{\Omega_p}$. The non-local gradient is defined as a function $\nabla_w p_u : \Omega \times \Omega \to \mathbb{R}^{\Omega_p}$ given by $\nabla_w p_u(x, \hat{x}) = w(x, \hat{x})(p_u(x) - p_u(\hat{x}))$. Thus, the patch-wise non-local TV reads

$$\|\nabla_w p_u\| = \int_\Omega \int_\Omega w(x, \hat{x}) \|p_u(x) - p_u(\hat{x})\|_2 \, d\hat{x} \, dx.$$

Note that in this sense, the model of the patch NL-medians corresponds to an anisotropic version of the non-local TV where the 2-norm in the integral is replaced by the 1-norm. Our work and the work of [PBC09] are complementary. In [PBC09] the regularization term is fixed, and the authors focus on the possibilities given by different data terms suited for different applications. On the other hand in this work we focus on the regularization term exploring its properties with different patch error functions ε, and applying them to a problem in which the data term plays no role at all, since there is no data to adjust to.

6.5 Multiscale scheme

Exemplar-based inpainting methods show a critical dependence with the size of the patch. In Figure 6.7, we show completions obtained with patch NL-means using different patch sizes: Two results with a small patch ($a = 4$) and one result with a large patch ($a = 19$). The latter is able to reproduce the periodic pattern of

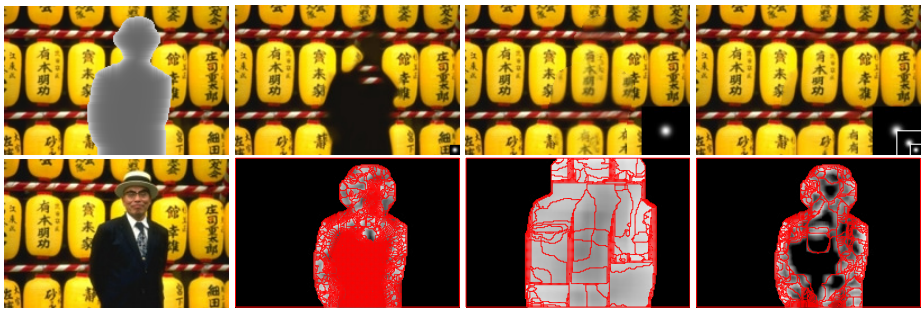


Figure 6.7: Single scale vs. multiscale. Left column: inpainting domain and initial condition. For the rest of the columns, from left to right: single scale inpainting with a 9×9 patch with $a = 4$; single scale inpainting with a 43×43 patch with $a = 19$; multi-scale inpainting with three scales, corresponding to patch sizes of 9×9 with $a = 4$, 21×21 with $a = 9$, and 43×43 with $a = 19$. All results have been computed with the patch NL-means scheme. The bottom row shows the boundaries between copy regions superimposed over the energy density image.

the lamps, but the completion is blurry due to the spatial overlap of the patches and presents many discontinuities at the boundary of the hole.

The results with the small patch do not show these artifacts, but one of them has failed to reproduce the lamps. The only difference between both is the initialization. One of them was initialized with the original image shown in the bottom left, whereas the other one with the result obtained with the multiscale approach described in this section.

As in many state of the art exemplar-based inpainting methods (*e.g.* [KT07, WSI07, KSY09]), we will incorporate a multiscale scheme. This is usually motivated as an heuristic to avoid local minima, to find a good initialization and/or to alleviate the computational cost. We believe however (following [HGY08]) that, as the example of Figure 6.7 suggests, inpainting is inherently a multiscale problem: Images have structures of different sizes, ranging from large objects to fine scale textures and edges. The multiscale scheme responds to the fact that several patch sizes are needed to reproduce all these structures properly.

6.5.1 Multiscale algorithm

In the following we describe the multiscale method we adopted, which goes along the lines of what is customary in the literature [FL09, KSY09, WSI07]. It consists on applying sequentially the inpainting scheme on a Gaussian image pyramid, starting at the coarsest scale. The result at each scale is upsampled and used as initialization for the next finer scale. The patch size is constant through scales.

Let us consider S scales, the finest denoted with $s = 0$. We will specify the size of the image at the coarsest level A_{S-1}. Denoting the size of the image at

the finest scale by A_0, we compute the sampling rate as $r = (A_0/A_{S-1})^{1/(S-1)} \in (0,1)$. The width of the Gaussian filtering is associated to the subsampling factor as in [MY08]. Let a_0 be the size of the patch and E_{a_0} the corresponding energy. We will add the superindex $s = 0, \dots, S-1$ to the variables u and w to denote the scale. As before, the subindex 0 refers the initial condition, i.e.u_0^s is the initial condition at scale s.

Algorithm 2 Multiscale scheme.

Require: u_0^S, S, a_0 and A_{S-1}
1: Initialize: $(u^{S-1}, w^{S-1}) = \arg\min_{(u,w)} E_{a_0}(u,w)$
2: **for** each scale $s = S-2, \dots, 0$ **do**
3: Upsample u^{s+1} to obtain u_0^s
4: $(u^s, w^s) = \arg\min_{(u,w)} E_{a_0}(u,w)$
5: **end for**

The upsampling from $s+1$ to s is obtained as in [WSI07]. The coarse weights w^{s+1} are first interpolated to the finer image size, yielding w_0^s. These weights are then used to solve an image update step at the new scale: $u_0^s = \min_u E_{a_0}(u, w_0^s)$. More conventional upsampling schemes by local interpolation (such as bilinear or splines) introduce a bias towards low-frequency non-textured regions. This exemplar-based upsampling avoids this bias.

Notice that keeping the patch size constant while filtering and reducing the image, is almost equivalent to enlarging the patch domain and filtering an image of constant size. The process can thus be seen as the sequential minimization of a series of inpainting energies with varying patch size given by $a_s = (1/r^s)a_0$, $s = 0, \dots, S-1$, over a corresponding series of filtered images. In the coarsest scale $S-1$, a larger portion of the inpainting domain is covered by partially known patches. This makes the inpainting task easier and less dependent on the initialization. The energy at this scale should have fewer local minima. The dependency of the minimization process on the initial condition ensures that each single scale solution remains close to the coarse scale initialization. The multiscale algorithm *exploits* this dependency to obtain an image u^0 which is approximately self-similar for all scales (or equivalently, for all patch sizes).

Figure 6.7 shows a comparison between single and multiscale results with the patch NL-means scheme. The multiscale result shows the benefits of large and small patch sizes. The missing lamps have been completed with the correct shape and spacing by the coarser stages, and the fine details are overall much less blurry and there are almost no discontinuities at the boundary of O. The bottom row shows the copy regions. The single scale results show a coarse partition with the large patch (the copying is more rigid), and one with many small regions with the smaller patch. The multiscale's NNF shows an intermediate partition, with some large regions inside of the hole and smaller ones around its boundary. The inpaintings at the finer scales work by refining the coarse partition obtained at coarser scales.

6.6 Experimental results

In this section we further will show the performance of the proposed schemes on real inpainting problems. The images used were obtained from Komodakis and Tziritas [KT07] and from the 100 images benchmark proposed by Kawai *et al*. [KSY09], available at `http://yokoya.naist.jp/research/inpainting/`. We refer to [KSY09, KT07] for a comparison with their approaches.

6.6.1 Experimental setting

We consider three inpainting methods, variations of our proposed framework, namely patch NL-means, -medians and -Poisson. The latter is always combined with patch NL-means with mixing parameter λ, as in (6.13). In all cases we use the multiscale approach. To prevent blurring we set h, the selectivity of the similarity weights w, to $h \to 0$. In this case, the weights select the nearest neighbor of each patch in $\widetilde{O}$. We use the CIE La*b* color space.

The calculation of the weights dominates the computational load of the algorithms. With an exhaustive search for the exact nearest neighbor, the cost of each iteration is $\mathcal{O}(\mathcal{A}(O) \times \mathcal{A}(O^c) \times s^2)$. However, a significant speed-up can be obtained with approximate searches, almost without any noticeable decrement in the quality of the results (see [BKC08] and references therein).

In our implementation we use a modified version of the PatchMatch algorithm introduced in [BSFG09]. PatchMatch is an iterative algorithm that estimates the NNF jointly for all patches in the inpainting domain by exploiting the coherence of natural images. The modification we implemented allows the estimation of a lists of the L first nearest neighbors of each patch. The algorithm has a computational cost of $\mathcal{O}(\mathcal{A}(O) \times s^2 \times L)$ per iteration. Typically between 5 and 10 iterations are sufficient to obtain results comparable with the exhaustive search algorithm when using lists of size $L \simeq 10$. In Appendix 6.A we describe the modified PatchMatch.

The results are shown in Figures 6.8 and 6.9. The first column shows the initialization. The rest of the columns show the results obtained with patch NL-means, -medians and -Poisson in that order. Obtaining good results requires fixing the following parameters:

Patch size. For almost all experiments we used patches of size s between 3×3 and 9×9. We did not use Gaussian intra-patch weights to shape the patch, since they need a larger support, which implies less available exemplars.

Multiscale parameters. The multiscale scheme has two parameters: the size of the coarsest image A_{S-1} and the number of scales S. From these, A_{S-1} is the most critical parameter. Smaller images with smaller inpainting domains are easier to complete. Care must be taken, however, since small images also imply less available exemplars to copy from. In these experiments A_{S-1} was set to a 20% of the original size when possible. Some cases required less subsampling. The number of scales S was set such that the subsampling rate $r = (A_0/A_{S-1})^{1/(S-1)} \approx (1/2)^{1/3} \approx 0.8$ as in [WSI07].

Confidence mask. The confidence mask has two parameters, the asymptotic value c_0 and the decay time t_c. For all experiments we fix $c_0 = 0.1$ and used a decay time $t_c = 5$ except for small inpainting domains, in which we set $t_c = 1$.

For the mixing coefficient λ of patch NL-Poisson we tested two configurations: Low-λ corresponding to $\lambda = 0.01$, and high-λ corresponding to $\lambda = 0.1$. Recall that lower values of λ give a higher weight to the gradient component of the energy. This is appropriate for structured images with strong edges. In almost each row of Figs. 6.8 and 6.9 we used the same λ for patch NL-Poisson.

The rest of the parameters are also the same for each row in Figure 6.9. Instead, for the images in Figure 6.8 we used different parameters for the image based methods (NL-means and NL-medians), in these cases, image-based methods required larger patches than NL-poisson.

6.6.2 Observations and comments

Gradient vs. intensity. Patch NL-Poisson shows a better reconstruction of structure and periodic patterns (Figure 6.8). Intensity-based methods present discontinuities at the boundary of the hole and between copy fronts. Furthermore, patch NL-Poisson facilitate the prolongation of structures and edges, due to the reinforcement of local PDE diffusion and non-local propagation (see Section 6.3). This allows to use smaller patches, which alleviates the computational load and reflects in more available exemplars and less blending due to patch overlap (either by averages or medians).

On the other hand, patch NL-Poisson shows problems with random textures (Figure 6.9) and in some occasions between copy fronts. In these cases, the nearest neighbors of overlapping patches disagree on the gradient of a certain pixel. This may result in the attenuation or the omission of gradients, causing a "spilling effect" as seen in rows 2 and 3 of Figure 6.8.

Some failures of patch NL-Poisson on random textures result from the use of gradients in the patch error function, as seen in Figure 6.9. In the second row of Figure 6.9, for instance, segments of the sky have been reproduced in the snow. Figure 6.10 shows results with an inpainting scheme in which the weights are computed based only on the image values (with the squared L^2-norm), and the image is updated using patch NL-Poisson with a low value of λ. Such scheme is non-variational and its convergence its not guaranteed.

We should point out that in many other cases, considering gradients in the comparison criterion improves the results. See for instance Figure 6.8.

Means vs. medians. It is notorious that patch NL-medians perform better at the reproduction of fine texture. The results of the L^2 methods are smoothed by the spatial averaging of overlapping patches. Patch NL-medians creates sharp discontinuities as in Figure 6.2 when different copy fronts meet (*e.g.* rows 2, 3 and 5 in Figure 6.8). These discontinuities are very noticeable and in these cases some smoothing is desirable.

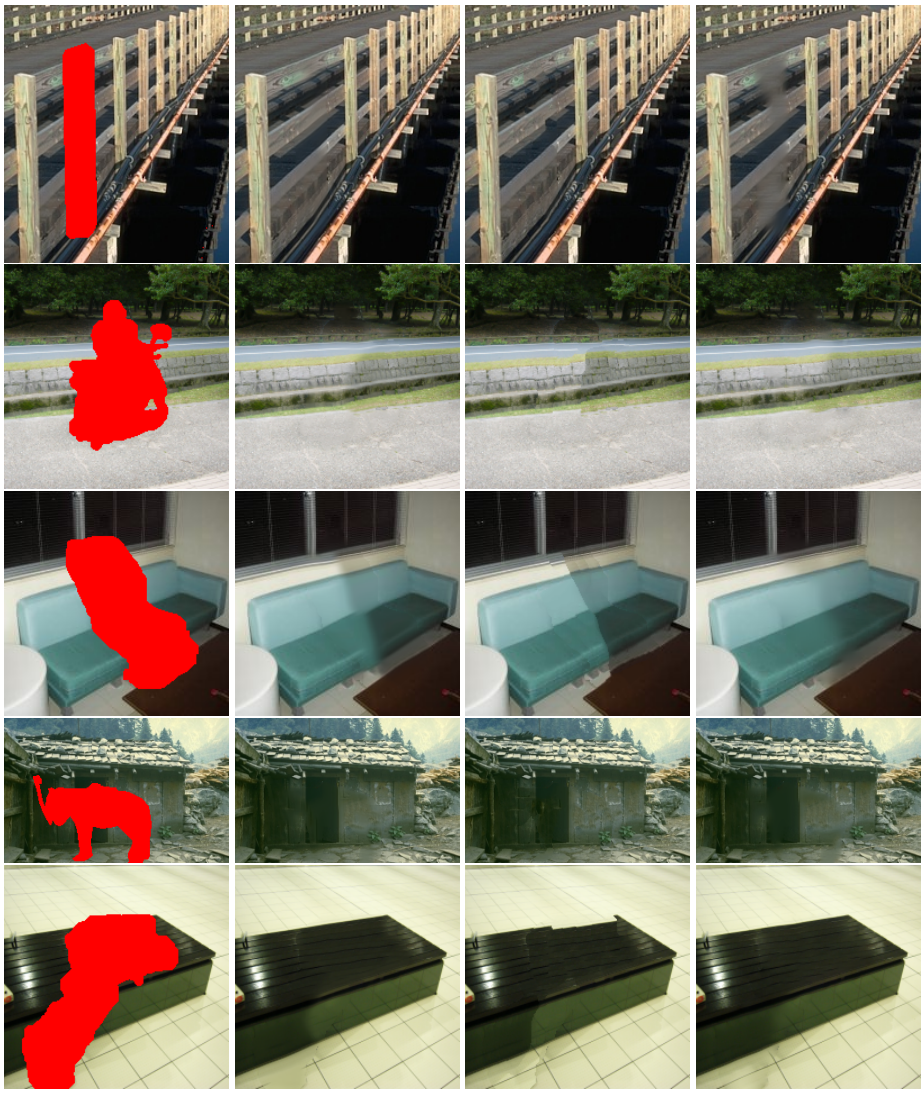


Figure 6.8: Results on structured images. From left to right: Original image and mask, patch NL-means, -medians, and -Poisson.

Figure 6.9: Results on random textures. From left to right: Original image and mask, patch NL-means, -medians and -Poisson. All images in each row have been generated with the same parameters.

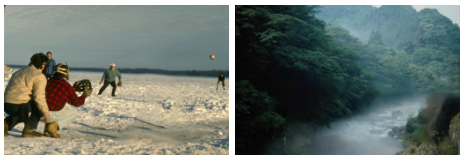


Figure 6.10: Results with a non variational algorithm. Patch NL-Poisson sometimes introduce blurry areas in the solution. Here we show the results using a non variational version of the patch NL-Poisson for some of those images.

We have observed that patch NL-medians is resistant to introduce changes in the solution during the minimization process. The same robustness that allows a better performance with textures makes it more greedy. Once a set of neighboring patches have settled on a locally stable solution (typically a region of constant NNF), it is hard for the algorithm to change that local decision.

Although the confidence mask diminishes this effect, patch NL-medians still shows more dependence on the initialization. A result of this are some misalignments in straight lines, due to subsampling artifacts of the multiscale scheme. Also generally requires the use of larger patches, particularly for structured images. This is not always possible, as in row 4 of Figure 6.8, where we could not find a proper set of parameters. However we found a good result by after 3 iterations of the multiscale scheme.

6.7 Conclusions and future work

In this work we presented a variational framework for exemplar-based non-local image inpainting. The proposed energy lends itself to intuitive interpretations in terms of probabilistic models and has connections with mean shift and statistical mechanics. We minimize it using a coordinate descent algorithm, alternating between similarity weights and the image updates. These processes are coupled through the patch error function, i.e. the criterion used to compare patches.

We derived from this framework three different inpainting schemes, corresponding to error functions based on L^1- and L^2-norms of image patches and the L^2-norm of the gradient patches.

The proposed functional shows a critical dependence with the patch size. Furthermore, it is nonconvex and has many local minima, particularly for small patch sizes. To tackle these issues we use a multiscale approach. This is customary in the literature, usually motivated as a way to obtain a good initialization and for computational reasons. We believe however that inpainting is inherently a multiscale problem, and that underlying the multiscale approach lies an inpainting criterion. We are currently working on a variational formulation of multiscale inpainting.

Although the focus of this work lies more on introducing and exploring the variational framework than in presenting a single inpainting algorithm, we would like to comment on which is the best method among the ones presented. Based on the experimental observations, it is clear that the answer to this question will depend on the characteristics of the inpainting problem. However, *a priori*, the combination of patch NL-Poisson with patch NL-means seems a reasonable compromise between quality and computational cost. It is also able to provide good results for a wider range of parameters.

Other topic of current research is the use of other patch error functions based on the comparison of structure tensors, which could provide a more robust estimation of the morphological structure of the image.

6.A Computation of the Nearest Neighbor Field

The computation of the nearest neighbor (or of the weight function w) is the most time consuming step of an exemplar-based algorithm. Two key observations allow to reduce the computational load for this task. The first one is that, the nearest neighbor search can be approximated without compromising the quality of the output, allowing to trade precision for speed (see for instance [BKC08] and references therein). The second one, as noted in [BSFG09], is that due to the coherence of natural images, nearest neighbors of patches centered on nearby pixels are likely to be located at nearby positions.

PatchMatch is a very efficient algorithm for computing approximate nearest neighbors proposed in [BSFG09]. The authors in [BSFG09] consider two regions O and O^c of the same or different images. The objective is to find for each patch centered in $x \in O$ the location of the best match $n(x) \in O^c$. The authors define

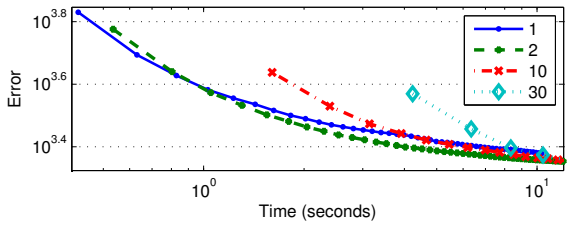


Figure 6.11: Performance of the modified PatchMatch. The graph shows the evolution of the nearest neighbor's error with the iterations of the algorithm. Each graph corresponds to a different size L of the sets ($L = 1$ corresponds to the original PatchMatch [BSFG09]). The dots indicate the completion of an iteration.

the Nearest Neighbor Field (NNF) as the function $x \mapsto (n(x) - x)$ defined over O. Instead of searching for the nearest neighbor of each query patch independently, PatchMatch computes the NNF simultaneously for all query patches. It exploits the fact that since query patches overlapp, the offset $n(x) - x$ of a good match for at x is likely to lead to a good match for the adjacent points of x as well. It is an iterative algorithm which starting from a random initialization, alternates between steps of propagation of good offsets and random search (see [BSFG09] for details). For most applications with natural images, a few iterations after a random initialization are often sufficient.

We extended this algorithm following a suggestion in [BSFG09]. Instead of storing a single offset for each query patch, we store queues of L offsets in an L-Nearest Neighbors Field (L-NNF). Thus the number of initial random guesses gets multiplied by L. During the propagation step adjacent queues are merged in to a temporal queue of length $2L$ from which only the best L offsets are kept. This allows to preserve and transfer non optimal offsets that may in turn be optimal for farther positions. Our implementation of the random search is essentially as in [BSFG09].

As expected using queues allows to achieve results similar to the original PatchMatch with less iterations, however each iteration takes considerably more time. Figure 6.11 shows the evolution of the nearest neighbor's error measured as $\sum_{x \in O} \|p_u(x) - p_u(n(x))\|^2$ with time, using queues of different lengths. In the long run using the queue pays back allowing to attain smaller errors in less time, even with $L = 2$ we observe that starting from the middle of the graph the errors are always smaller than $L = 1$. Since we are not interested in attaining a high precision, we do not use big values of L. In practice we set $L \leq 10$ and perform a small number of iterations (generally 5 or 10).

Let us remark that each queue is an approximated neighborhood of the query patch in the patch space. This neighborhood can be interpreted as a truncated representation of the function $w(x, \cdot)$ for each pixel x, which is useful in the non-degenerated case when $h > 0$ in (6.6).

Chapter 7

Exemplar-Based Interpolation of Sparsely Sampled Images

In this chapter we propose a nonlocal formulation for the problem of image interpolation from a sparse set of samples. This formulation is derived from the nonlocal inpainting framework presented in the previous chapter. Following the paradigm of exemplar-based methods, this formulation encourages the transfer of information between similar image patches, but contrary to the classical inpainting problem, no complete patches are available from the sparse image samples, hence the patch similarity criterion has to be redefined as here proposed. The results show the ability to recover textures at low sampling densities. We also explore some departures from the variational setting, showing a remarkable ability to recover textures at low sampling densities.

7.1 Introduction

The terms image *inpainting* and *interpolation* refer to the problem of recovering missing information in an image, in a visually plausible manner exploiting available image information. Both, *inpainting* and *interpolation*, are ill-posed inverse problems, and as such, some sort of prior knowledge is needed for its solution. We will distinguish between: *inpainting* when the available data is given on a (not necessarily connected) region of the image; and *interpolation* when the available data consists of a set of isolated samples (be regular or irregular). The literature on these topics is vast, since lies at the heart of many relevant applications such as: image editing and desocclusion for the former; zooming, demosaicing and super-resolution for the latter. Let us review the main trends in both areas.

For the inpaiting problem, the available information usually allows to determine the image derivatives on the boundary of the unknown region of the image. First approaches to inpainting took advantage of this, completing the image by means of variational methods [MM98] or PDEs [BSCB00, BBS01] that continued the image gradients or the level lines inside the inpainting domain. These schemes involve only interactions between *local* pixels, fail with textured images or large inpainting domains. Advances in the field of texture synthesis [EL02]

served as inspiration for new inpainting strategies. Based on the hypothesis that natural images are redundant, and self similar, these methods determine the value of a pixel from known pixels with similar neighborhoods. These methods are often refereed to as *non-local* or *exemplar-based* [LZW03, CPT04, SYJS05]. Exemplar-based methods excel in recovering textures but their ability to recover geometry is limited, while the local methods are good at recovering geometric texture-less features. Recent efforts aim at combining both, local and non-local strategies *e.g.* [BVSO03, CGMP09].

Interpolation approaches on the other hand are based on priors of smoothness or regularity assumptions. These can be imposed by restricting the solutions to be, for instance, band limited [GS00], of bounded variation [CS01], expanded over a base of functions (*e.g.* splines [ASHU05a], radial basis functions [She68a]), among others. PDE's for anisotropic diffusion are successfully applied to the interpolation of sparse samples [GWW$^+$08] by estimating a diffusion tensor from the available samples. These techniques are also local in the sense that they regard only local interactions between the pixels. And as in the inpainting case these techniques fail to recover textures. A recent front of activity is given by the techniques based on the *sparseland* model [CW08, AEB06], in which the image is restricted to have a sparse representation over an overcomplete basis or *dictionary* [MSE08, AEB06, ESQD05]. The dictionary based methods are very similarly to the exemplar-based methods for inpainting. The main difference between them is where the missing information is obtained from. Dictionary based methods look for the missing data in the dictionary (as a linear combination of a few atoms), whereas exemplar-based methods assume that the information needed is elsewhere in the image itself (or in a database of images [HE07]). Non-local priors have also been considered for the resolution of inverse problems [PBC09], and in particular for interpolation. In [PBC09] the set of image patches with their similarity relations is modeled as a weighted graph, and the interpolation is done by imposing regularity in this graph [LEB07, GO07a] which corresponds to a non-local regularization on the image.

Both inpainting and interpolation problems are very similar. Thus, it is natural to think of applying the inpainting techniques to interpolation problems. However, since in our case only a sparse set of samples is available, the adaptation of the inpainting techniques is not trivial because: 1. The gradients as well as the directions of the level lines are unknown; 2. There are no complete patches available on the image. Under these conditions, the PDE based methods for inpainting cannot be directly applied to interpolation and exemplar-based inpainting methods need to be adapted (as we will do).

In Chapter 6 we presented a framework for exemplar based inpainting, which models the self similarity of the image explicitly as a weight function (see Section 7.2 for a summary). These problems are solved by alternate minimization of the energy, with respect to the image and the weights. The main difference between inpainting and interpolation lies the available data and its geometric organization in the image. In a typical inpainting problem, a region of the image is missing (*the hole* or *inpainting domain*) but also large regions of

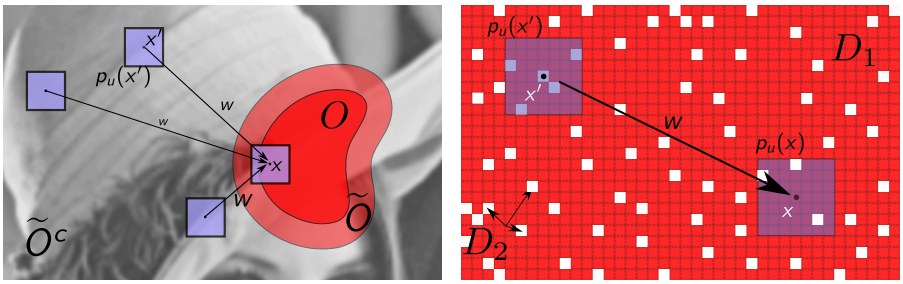
the image are known. From the known regions complete patches are copied and used to fill the inpainting domain (see Figure 7.1(a)). In the interpolation problem here addressed, we may still have entire regions of missing information, but in addition the image is only known at some isolated positions distributed through all the image (see Figure 7.1(b)), this means that there are no complete patches to copy from. This prevents the direct application of the inpainting energy from the previous chapter to the interpolation problem, since in the current case the set of complete patches is likely to be empty.

In this chapter we will generalize the framework of Chapter 6 to address the problem of image interpolation from non-uniformly sampled data. In this setting, we consider the simpler case where the samples are arranged on a discrete (but non regular) grid, and leave the sub-pixel case for future development. Hence we will formulate the interpolation problem in a discrete setting. Since in this context there are no complete patches to copy from, the patch error function is tailored to compare only the known pixel positions in one or both of the patches being compared. The interpolated image is computed by solving a sequence of problems each using a decreasing value of the *selectivity parameter* (see Section 7.3), this "coarse to fine" strategy permits to iteratively resolve details of the interpolated image [BCMS09].

Related work. Our work is related to the nonlocal techniques applied to demosaicing [BCMS09] and super-resolution [PETM09] problems which can be cast as image interpolation from a regular sampling set. These methods work by averaging known pixels according to the similarity of their neighborhoods, and are closely related with our approach. More detailed comments on them will be given in subsequent sections. The work [MBP+09] combines sparsity and non-local techniques. The regularity is imposed by sparsity over a dictionary, which is learned by assigning a common representative to similar patches. Similar ideas can be also found in the field of 3D tomographic imaging [BSL+08], where incomplete 3D volumes are reconstructed via grouping them by similarity and averaging the exemplars in each cluster.

Let us mention that the problem of interpolating an image from a set of sparse samples could be approached with the techniques of compressed sensing [CW08, CR08]. Even if the standard setting uses a set of random measurements (e.g. projections on a random basis, or noiselets) one could apply the corresponding reconstruction schemes with a random sampling of the image, as in our case. As far as we know, there is no detailed comparison between exemplar-based methods and compressed sensing in the context of image interpolation. On the other hand, as shown in this paper, exemplar-based methods can handle the problem of interpolating non uniformly sampled images with large unsampled regions.

Notation. Images are denoted as functions $u : \Omega \to \mathbb{R}$, where Ω denotes the image domain, usually a rectangle in $\mathbb{Z}^2$. Pixel positions are denoted by x, $\hat{x}$, z, $\hat{z}$ or y, the latter for positions inside the patch. A patch of u centered at x, is



(a) Inpainting (b) Interpolation

Figure 7.1: Domains for the inpainting (left), and for the interpolation problems (right). In the case of inpainting the red areas indicate the unknown regions O of the image; we have also shown the extended inpainting domain $\widetilde{O}$ its complement complement $\widetilde{O}^c$. In the case of interpolation, D_1 is the whole image domain and D_2 represents the unknown pixels (in red), the known pixels are displayed in white.

denoted by $p_u(x) = p_u(x, \cdot) : \Omega_p \to \mathbb{R}$, where Ω_p is a disk (or a square) centered at $(0, 0)$. The patch is defined by $p_u(x, y) = u(x + y)$, with $y \in \Omega_p$. $O \subset \Omega$ is the set of unknown image pixels or the domain to be interpolated, and $O^c = \Omega \setminus O$ is the known part of the domain. For simplicity we will assume that the image is defined on an extended domain $\widetilde{\Omega} = \Omega + \Omega_p$ (i.e. $\widetilde{\Omega}$ is a dilation of Ω) and we work in Ω, hence a patch can be centered at any pixel in Ω without escaping the image domain. Additional notation will be introduced in the text.

7.2 Variational framework

First let us briefly recall the non-local inpainting framework of Chapter 6, now in a discrete context. Then we will discuss the modifications for its application to the interpolation problem.

7.2.1 Review: Non-local functional for image inpainting

Let us define, $\widetilde{O}$ as an extension of O containing the centers of all patches intersecting O (see Figure 7.1(a)). In doing so, patches $p_u(\hat{x})$ centered at pixels $\hat{x} \in \widetilde{O}^c$ consist entirely of known pixels. The energy proposed in the previous chapter is composed of two terms, the first measures the coherence between the patches in $\widetilde{O}$ and those in $\widetilde{O}^c$, for a given the similarity function $w : \widetilde{O} \times \widetilde{O}^c \to \mathbb{R}^+$. This permits the estimation of the image from the weights w. And the second term

allows us to compute the weights w given the image u.

$$\widetilde{E}(u,w) = \frac{1}{h}\widetilde{F}_w(u) - \sum_{x\in\widetilde{O}} \widetilde{H}_w(x), \qquad (7.1)$$
$$s.t. \quad \sum_{\hat{x}\in\widetilde{O}^c} w(x,\hat{x}) = 1,$$

where

$$\widetilde{F}_w(u) = \sum_{x\in\widetilde{O}} \sum_{\hat{x}\in\widetilde{O}^c} w(x,\hat{x})\varepsilon(p_u(x) - p_u(\hat{x})), \qquad (7.2)$$

$\varepsilon(\cdot)$ is the *patch error function*, and

$$\widetilde{H}_w(x) = -\sum_{\hat{x}\in\widetilde{O}^c} w(x,\hat{x}) \log w(x,\hat{x})$$

is the entropy of the probability $w(x,\cdot)$ for $x \in \widetilde{O}$ and h is the *selectivity parameter* .

The term (7.2) promotes the similarity between the image patches centered at $x \in \widetilde{O}$ and $\hat{x} \in \widetilde{O}^c$. Indeed, minimizing $\widetilde{F}_w$ w.r.t. the image u, for a given fixed weight function w, forces pairs of patches for which $w(x,\hat{x})$ is high to be similar. Since $p_u(\hat{x})$ lies outside the inpainting domain, it is fixed and the similarity can only be enforced by modifying $p_u(x)$. Thus, incomplete patches *receive* information from outside the inpainting domain.

The similarity weight function $w : \widetilde{O} \times \widetilde{O}^c \to \mathbb{R}^+$ measures the similarity between patches centered in the inpainting domain and in its complement. To avoid trivial minima of F_w, $w(x,\cdot)$ is constrained to be a probability density function, $\sum_{\hat{x}\in O^c} w(x,\hat{x}) = 1$. For a given completion u, and for each $x \in \widetilde{O}$, the optimum weights minimize the mean patch error function for $p_u(x)$, given by $\sum_{\hat{x}\in\widetilde{O}^c} w(x,\hat{x})\varepsilon(p_u(x) - p_u(\hat{x}))$, while maximizing the entropy. The resulting weights are exponential, as can be confirmed easily by derivating the energy. Gaussian weights $w(x,\hat{x}) \propto \exp\left(-\frac{1}{h}\|p_u(x) - p_u(\hat{x})\|^2\right)$, are obtained by setting the patch error function as the squared L^2 norm of between the patches i.e.$\varepsilon(p_u(x)) = \sum_{y\in\Omega_p}(u(x+y))^2$, where the proportionality factor is a constant such $w(x,\cdot)$ adds up to one. The parameter h controls the trade-off between both terms and is also the *selectivity parameter* of the Gaussian weights. Since $w(x,\cdot)$ is a probability, we discard trivial minima of E with $w(x,\hat{x}) = 0$ everywhere.

Summarizing, the first term of (7.1) permits the estimation of the image u from the weights w, whereas the second one allows us to compute the weights given the image.

7.2.2 Generalization to interpolation

Let us now discuss the modifications needed to adapt the formulation (7.1) to the problem of image interpolation. We will focus our attention on the image term

(7.2). For that let us assume for the moment that we know a weight function w which measures the similarity of pairs of incomplete patches. Later we will detail the issues related with the computation of these weights.

The main difference between inpainting and interpolation is the available data and its geometric organization in the image. In a typical inpainting problem, large regions of the image are known, and information is transferred from the complete patches of the complement to the inpainting domain. Whereas in the interpolation case the image is known only at some isolated positions distributed through all the image and there are no complete patches available. This prevents the direct application of the inpainting energy (7.1) to the interpolation problem, since every patch contains unknown pixels, and thus needs information from other patches (which are also incomplete). Moreover in this case it is likely that $\widetilde{O}^c = \emptyset$. This suggests that the summation domains in equation (7.2), as well as the patch comparison metric, have to be modified. We address this here.

For the sake of generality we will use generic summation domains and denote them by D_1 and D_2. The distinction between summation domains D_i and known data O^c, is also useful for the understanding of the resulting algorithm. The corresponding definitions for the inpainting functional (7.2) are $D_1 = \widetilde{O}$ and $D_2 = \widetilde{O}^c$, while for all methods implemented below we used $D_1 = \Omega$ and $D_2 = O^c$, i.e. D_2 the set of known pixels (see Figure 7.1(b)). The weight function is thus defined over $D_1 \times D_2$ and is such that for each $x \in D_1$ $w(x, \cdot)$ is a probability over D_2.

A general description of the image term in the interpolation functional is the following:

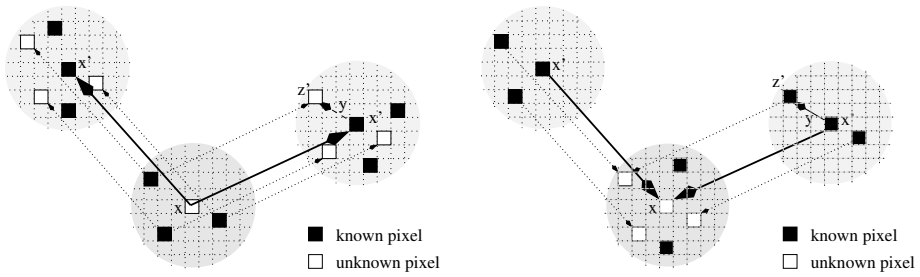
$$F_w(u) = \sum_{x \in D_1} \sum_{\hat{x} \in D_2} w(x, \hat{x}) V_{\mathrm{e}}(p_u(x), p_u(\hat{x})). \tag{7.3}$$

We have introduced a general *pair-wise patch similarity potential* V_{e}, substituting the patch error function $\varepsilon(\cdot)$. Since we deal with sparsely sampled patches, the pair-wise patch potential V_{e} is based only on the known pixels around x and $\hat{x}$:

$$V_{\mathrm{e}}(p_u(x), p_u(\hat{x})) = \sum_{y \in \Omega_p} \frac{g_\sigma(y)}{\rho(x, \hat{x})} (\alpha \mathcal{X}_{O^c}(x+y) + \beta \mathcal{X}_{O^c}(\hat{x}+y)) \mathrm{e}(u(x+y) - u(\hat{x}+y)), \tag{7.4}$$

where g_σ is a Gaussian centered at the origin with standard deviation σ, $\mathcal{X}_S$ denotes the characteristic function of the set S and $\mathrm{e}(r) = |r|^p$, $r \in \mathbb{R}$ and $p = \{1, 2\}$. Taking $p = 2$, leads to an algorithm analogous to *patch NL-means* and with $p = 1$ leads to *patch NL-medians* (see 7.3), also more general functions $\mathrm{e} : \mathbb{R} \to \mathbb{R}^+$ could be considered. The constant parameters $\alpha, \beta \in \{0, 1\}$ are set by the user. They control whether known positions around x or $\hat{x}$ are used in the computation of the similarity potentials (at least one of them has to be 1).

If $\alpha = 1$ the positions with known data around x are used for the computation of the similarity potential (7.4). This happens whether the corresponding locations arround $\hat{x}$ belong to the data set or not. If $\beta = 1$ the similar-



(a) $\alpha = 1, \beta = 0$: Transmit (b) $\alpha = 0, \beta = 1$: Receive

Figure 7.2: Visualization of transmission and reception processes due respectively to terms V_{e}^{α} and V_{e}^{β} in the patch similarity potential (see equation (7.6)). Recall that $\hat{x} \in D_2$ and $x \in D_1$. Thus, when $D_1 = \Omega$, x may be known or unknown (this is not reflected in the figure).

ity potential accounts for the known pixels around $\hat{x}$. If both of them are 1, then V_{e} accounts for the locations known in any of the two patches. This last case coincides with the patch comparison criterion defined in [BCMS09] in the context of demosaicing. The normalization factor $\rho(x, \hat{x})$ is such that $\sum_{y \in \Omega_p} \frac{g_\sigma(y)}{\rho(x,\hat{x})}(\alpha \mathcal{X}_{O^c}(x+y) + \beta \mathcal{X}_{O^c}(\hat{x}+y)) = 1$ for all $x \in D_1$, $\hat{x} \in D_2$.

Considering the overlap between known positions in both patches (see for instance [BSL$^+$08]) would also make sense for comparing patches with missing data. However, this cannot be applied to the current formulation since this eliminates the dependency of the energy on the unknown image (recall that the energy depends on the image though the similarity potential V_{e}).

The complete functional for the interpolation problem becomes:

$$E(u, w) = \frac{1}{h} F_w(u) - \sum_{x \in D_1} H_w(x), \tag{7.5}$$

where $H_w(x) = -\sum_{\hat{x} \in D_2} w(x, \hat{x}) \log w(x, \hat{x})$ is again the entropy of the probability $w(x, \cdot)$ for $x \in D_2$. And as in the case of inpainting, this last term allows to model the weight update together with the image update.

A similar functional for image super-resolution was proposed in [PETM09] without explicitly modeling the weight updating step. The image update in [PETM09] is related to the case where $\alpha = 0$ and $\beta = 1$.

Interpretation of the functional. The proposed functional can be easily understood by splitting the pairwise patch potential into two terms $V_{\mathrm{e}} = V_{\mathrm{e}}^{\alpha} + V_{\mathrm{e}}^{\beta}$, with

$$V_{\mathrm{e}}^{\alpha}(p_u(x), p_u(\hat{x})) = \alpha \sum_{y \in \Omega_p} \frac{g_\sigma(y)}{\rho(x, \hat{x})} \mathcal{X}_{O^c}(x+y) \mathrm{e}(u(x+y) - u(\hat{x}+y)), \tag{7.6}$$

and analogously for V_{e}^{β}. For a fixed weight function w, the energy F_w can be split accordingly in two terms. The first potential measures differences between known pixels in $p_u(x)$, with $x \in D_1$, and the corresponding pixels in $p_u(\hat{x})$, with $\hat{x} \in D_2$. Since known pixels are fixed, its minimization implies the modification of unknown pixels around $\hat{x}$, thus transferring information from $p_u(x)$ to $p_u(\hat{x})$. On the other hand, V_{e}^{β} considers differences between known pixels in $p_u(\hat{x})$ and the corresponding locations in $p_u(x)$. In this case known information flows from $p_u(\hat{x})$ centered at D_2 to $p_u(x)$ centered at D_1.

Since the weights $w(x, \cdot)$ are a probability over D_2 for each $x \in D_1$, we will adopt subsequently the point of view of the patch $p_u(x)$ centered at $x \in D_1$. We refer to these patches as *central* patches, and to patches centered in D_2 as *peripheral* patches. From this perspective, the minimization of the term with V_{e}^{α} implies the *transmission* of the information (the pixel values) of known positions in the central patch $p_u(x)$ towards the unknown positions in peripheral patches $p_u(\hat{x}) \in D_2$ (see Figure 7.2(b)). Whereas the minimization of the term with V_{e}^{β} implies *receiving* known pixel values from peripheral patches at D_2 (see Figure 7.2(a)). We refer to these processes as *transmission* and *reception*.

7.2.3 Rewriting of the energy term F_w

As in Chapter 6, we rewrite the energy term (7.3) so that the image values appear directly, and not as part of patches. This will be useful for the derivation of the methods. After the change of variables $z = x + y$, $\hat{z} = x + \hat{y}$, the energy can be rewritten by adding up the pair-wise pixels differences as

$$F_w(u) = \sum_{z \in \widetilde{\Omega}} \sum_{\hat{z} \in \widetilde{\Omega}} m(z, \hat{z})(\alpha \mathcal{X}_{O^c}(z) + \beta \mathcal{X}_{O^c}(\hat{z}))\mathrm{e}(u(z) - u(\hat{z})), \tag{7.7}$$

where $\widetilde{\Omega} = \Omega + \Omega_p$ (since $D_1, D_2 \subseteq \Omega$, we have that $D_1 + y, D_2 + y \subseteq \widetilde{\Omega}$ for all $y \in \Omega_p$), and we have defined the *pixel-wise* influence weights $m(z, \hat{z})$ as

$$m(z, \hat{z}) = \sum_{y \in \Omega_p} \mathcal{X}_{D_1}(z - y)\mathcal{X}_{D_2}(\hat{z} - y) w(z - y, \hat{z} - y) \frac{g_\sigma(y)}{\rho(z - y, \hat{z} - y)}.$$

These weights integrate the similarities of patches centered at $z - y \in D_1$ containing z and those centered at $\hat{z} - y \in D_2$ containing $\hat{z}$ for $y \in \Omega_p$.

The formulation given in (7.3) accumulates the pair-wise potentials for each pair of patches centered in D_1 and D_2, where the potentials are computed as the sum of pixel differences. In (7.7), the same energy is rewritten so that the contribution of each pixel difference $u(z) - u(\hat{z})$ is made explicit. If the characteristic functions $\mathcal{X}_{O^c}(z)$ and $\mathcal{X}_{O^c}(\hat{z})$ in (7.7) are both zero (i.e.neither z nor $\hat{z}$ are known), then the corresponding term is not considered in the energy. Therefore only those differences involving at least one known pixel are taken into account. It is also clear from (7.7) that its minimization encourage the similarity of $u(z)$ and $u(\hat{z})$ for all pairs of pixels z and $\hat{z}$ such that $m(z, \hat{z})(\alpha \mathcal{X}_{O^c}(z) + \beta \mathcal{X}_{O^c}(\hat{z}))$ is large (i.e.the current patches do not match).

7.3 Minimization of E

We have formulated the interpolation problem as the constrained optimization problem

$$(u^*, w^*) = \arg\min_{u,w} E(u, w) \quad \text{subject to} \tag{7.8}$$

$$\sum_{\hat{x} \in D_2} w(x, \hat{x}) = 1 \quad \text{for all} \quad x \in D_1. \tag{7.9}$$

To minimize the energy E, we use an alternate coordinate descent algorithm. At each iteration, two optimization steps are solved: The constrained minimization of E with respect to w while keeping u fixed; and the minimization of E with respect to u with w fixed. This procedure is summarized in Algorithm 3.

Algorithm 3 Alternate minimization of $E(u, w)$.

Require: *Initial Condition:* $u_0(x)$ defined $\forall x \in D_1$.

1: **repeat**
2: *Weights Update Step:* $w_k = \arg\min_w E(u_k, w)$, s.t. (7.9).
3: *Image Update Step:* $u_{k+1} = \arg\min_u E(u, w_k)$.
4: **until** *Stopping Criterion:* $\|u_{k+1} - u_k\| \leq \tau$

In the weights updating step, the minimization of E w.r.t. w yields

$$w_k(x, \hat{x}) = \frac{1}{q(x)} \exp\left[-\frac{1}{h} V_{\mathrm{e}}(p_{u_k}(x), p_u(\hat{x}))\right], \tag{7.10}$$

where $q(x)$ is a normalization factor such that $\sum_{\hat{x} \in D_2} w(x, \hat{x}) = 1$ for each patch $p_{u_k}(x)$. The parameter h determines the selectivity of the similarity potential. If h is large, maximizing the entropy becomes more relevant, yielding weights which are less selective. In the limit, when $h \to \infty$, $w_k(x, \cdot)$ becomes a uniform distribution over D_2. On the other hand, a small h yields weights more concentrated on the patches that are similar to $p_u(x)$. In fact, when $h \to 0$ the weights are given by $\lim_{h \to 0} w(x, \hat{x}) = \frac{1}{\#n(x)} \mathcal{X}_{n(x)}(\hat{x})$, where $n(x) \subseteq O^c$ is the set of minimizers of $V_{\mathrm{e}}(p_u(x), \cdot)$, i.e.$n(x) = \{\hat{x} \in O^c : V_{\mathrm{e}}(p_u(x), p_u(\hat{x})) = V_{\min}(x)\}$, where $V_{\min}(x)$ is the minimum potential w.r.t. $p_u(x)$. In other words, when $h \to 0^+$ the weights encode a multivalued assignment of patches with centers in D_2 for each $x \in D_1$.

The Euler equation of E w.r.t. u is

$$\sum_{\hat{z} \in O^c} (\alpha m(\hat{z}, z) + \beta m(z, \hat{z})) \mathrm{e}'(u(z) - u(\hat{z})) = 0 \qquad \text{for all } z \in O. \tag{7.11}$$

This equation specifies the information transferred from the datum $u(\hat{z})$ to the unknown $u(z)$. This information can be transferred in any of the two modes discussed previously, i.e. reception, by a patch in D_1 covering z, of data coming

from a patch in D_2 covering $\hat{z}$, and/or transmission, of data from a patch in D_1 covering $\hat{z}$, to a patch in D_2 covering z. The term $m(z,\hat{z})$ gathers all contributions by reception, whereas the term $m(\hat{z},z)$ considers all transmissions.

When $\mathrm{e}(t) = t^2$ we call the resulting method *patch NL-means*. And equation (7.11) becomes

$$u(z) = \frac{1}{C(z)} \sum_{\hat{z}\in O^c} (\alpha m(\hat{z},z) + \beta m(z,\hat{z}))u(\hat{z}), \tag{7.12}$$

for each $z \in O$, where the normalization constant $C(z)$ is given by $C(z) = \sum_{\hat{z}\in O^c}(\alpha m(\hat{z},z) + \beta m(z,\hat{z}))$. Let us comment that this image update differs from the one used in [BCMS09]. Here $u(z)$ is updated using contributions (captured by m) of many patches that overlap z, while in [BCMS09] only the central pixel of patch centered at $\hat{z}$ is used for computing $u(z)$.

Taking $\mathrm{e}(t) = |t|$, we get the *patch NL-medians*. In this case, the Euler equation (7.11) for u, given w, becomes

$$\sum_{\hat{z}\in O^c} (\alpha m(\hat{z},z) + \beta m(z,\hat{z}))\mathrm{sign}(u(z) - u(\hat{z})) = 0,$$

and its solution $u(z)$ is obtained as a weighted median of $u(\hat{z})$ with $\hat{z} \in O^c$.

7.4 A departure from the variational model

We have seen that three different schemes can be derived from the proposed variational model, by changing the values of α and β. We have interpreted them, by observing the effect over the unknown pixels of u, as transmission ($\alpha = 1$, $\beta = 0$), reception ($\alpha = 0$, $\beta = 1$), or combination of both ($\alpha = 1$, $\beta = 1$). But each scheme also forces the manner we compute w. However, if we abandon the variational framework, we can combine different strategies for computing w and updating u. This is equivalent to using a different analysis and synthesis for the inpainting problem (denoted as non variational algorithms in Section 6.6.2).

We now propose a new scheme that computes the weights w according to the transmission scheme ($\alpha = 1$, $\beta = 0$), and updates the image u using the combined scheme ($\alpha = 1$, $\beta = 1$). The resulting algorithm was experimentally found to be numerically stable, and for relatively high sampling densities to behave like the combined scheme ($\alpha = 1$, $\beta = 1$). However for low sampling densities it exhibits a remarkable ability to speed up the convergence. An intuitive reason that may explain this scheme relies on the fact that using the transmission potential ($\alpha = 1$, $\beta = 0$), the weights $w(x,\cdot)$ are always computed using coordinates around x, with known values. Adding known positions around $\hat{x}$ may provide a poorer estimate of the weights, specially if the current interpolation around x is bad.

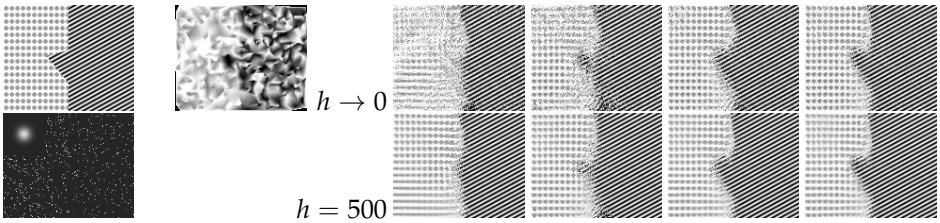


Figure 7.3: Synthetic. The first column shows the original image, and the random samples (5% of the image) with the window g_σ (see (7.4)) depicted in the upper left corner. The second column shows a linear interpolation over the Delaunay triangulation of the samples. The remaining columns (from left to right) correspond to the results of the schemes A, B, AB and O. The rows correspond to two different values of h.

7.5 Experimental results and discussion

We now present experimental results with both synthetic and natural images randomly sampled with densities from 20% to 5% of the image points. [1] We will show only results corresponding to patch NL-means (i.e.$p = 2$ in the definition of e from (7.4)). The four schemes derived from the potential (7.4) in Section 7.2, are referred here as A (for $\alpha = 1$, $\beta = 0$), B ($\alpha = 0$, $\beta = 1$), AB ($\alpha = 1$, $\beta = 1$), and O for the *departure from the variational model* (which is a variant of AB). All of them have a computational cost proportional to $\mathcal{A}(D_1) \times \mathcal{A}(D_2)$ (where $\mathcal{A}(D_i)$ is the number of pixels of D_i). Since D_2 is a fraction of D_1 (the density of the sampling) then the algorithm is $\mathcal{O}(T \times \mathcal{A}(D_1)^2)$, where T is the number of iterations (usually $T < 200$). A single iteration for a 256×256 pixels image takes about 3 min on a 3GHz processor. However, with the *coarse to fine* scheme described below, the convergence is generally attained with less than 40 iterations. This amounts to 120 minutes of computation.

Role of locality in the non-local algorithms. A common strategy to improve the computational performance of nonlocal methods is to reduce the size of the *search window* (subset of D_2 arround the central pixel x), thereby reducing the number of comparisons performed for each pixel. As a desirable side effect, this enforces the ergodicity assumption over the data. In other words, the patches needed to estimate the current point are assumed to be found in the vicinity of it, not far away. As a consequence, the size of the search window is a very important parameter, and it may be itself subject of optimization as in [KB06]. In our experiments we choose the search windows to have a reasonable size (containing 100 to 500 samples) with respect to the density of the image.

Choice of V_e. The experiments shown in Figure 7.3 are aimed to compare the performance of the different schemes. The best result for this experiment is ob-

[1] The experiments shown here are also available at: `http://gpi.upf.edu/static/vnli`.

tained with the scheme AB. Other experiments are also consistent with these results, so from now on we will mainly show results of AB and O. Notice in Figure 7.3 that the textures are recovered in great detail, while the interface between them, is very imprecise. This evidences the exemplar-based nature of the algorithms, since there are plenty of examples of textures, but only few of the interface.

Initial condition and h. If the initial condition has artifacts, then minimizing (7.5) with a small h tends to reinforce them. To reduce this dependence on the initial condition we adopt the *coarse to fine* scheme proposed in [BCMS09], where a decreasing sequence of h is used to recover first large scale structures and later refine them (as h decreases). Figures 7.7 and 7.9 show the results of applying the algorithms O and AB to natural images with sampling densities from 20% to 5%. For high densities the performances of both schemes is similar. For lower densities (5% for instance) O exhibits less dependence on the initial condition than AB. In particular, we can obtain with O results similar to those obtained with AB even without the coarse to fine scheme. Using $h > 0$ produces smooth results with blurred details, while using $h \to 0$ introduces a staircasing effect. In the first two columns of Figure 7.7 we display: a set of random samples, and an optimal dithered set of the same image (optimal for the Laplacian-based interpolation as described in [BBBW08]). Both sets contain 10% of the image points. The Laplacian interpolation from dithered samples takes advantage of the distribution of the samples along the edges and permits to recover a visually pleasant smooth image with clear edges (see [BBBW08]), while our method is less fit for this task (second column of Figure 7.7). However, for random samplings the results of the Laplacian interpolation are less convincing, while our algorithm manages to recover most of the edges and textures of the image (see Figure 7.9).

Interpolation of large holes. In Figure 7.8 we show a preliminary result using method B (only reception process) applied to the interpolation of a hole in a sampled image, this choice of the potential leads to a functional similar to the inpainting one shown in the previous chapter.

Let us remark that for this experiment the algorithm was applied without making any distinction between the hole and the sampled areas. Other methods that involve the transmission process (AB or O) fail to fill the large holes, although all manage to recover the sparsely sampled area. We attribute the non regularity of the solution to the low frequency of the texture, which implies less exemplars to copy from, evidencing again the main limitation of exemplar-based methods. To overcome this problem a local regularization could be used to impose smoothness on the result [PETM09].

7.5.1 Discussion: How does it work?

As we see from Figures 7.3 and 7.9 the proposed algorithm is able to recover high frequency textures even when the sampling density is well below the cor-

responding Nyquist frequency. There are tree factors that contribute to make it work:

1. The textures are stationarity;
2. The sampling is non-regular;
3. And the overlapping of the patches in the interpolated image.

Let us study a couple of toy models.

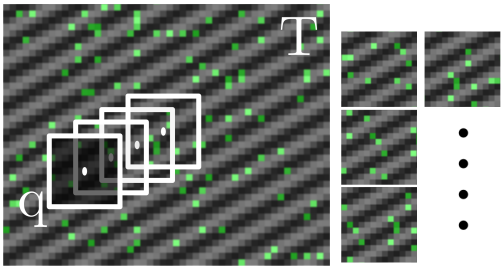


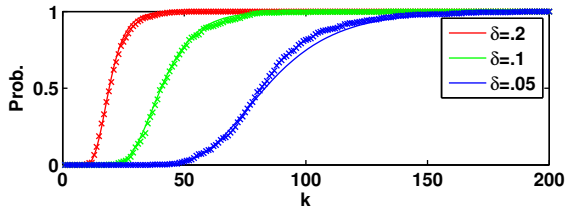
Figure 7.4: Texture T and a patch q. The (known) pixels are highlighted in green. Note that many incomplete exemplars of the same patch can be collected nearby q (shown on the right), and the positions of the known pixels are different in each exemplar.

Single texture. If the texture is stationary, then even if it is under-sampled we can obtain several exemplars of the same textured patch from a single image. And since the sampling is not regular then the exemplars will not exhibit the same alias. Consequently collecting a sufficient amount of randomly sampled exemplars and simply *piling* them permits to recover the texture.

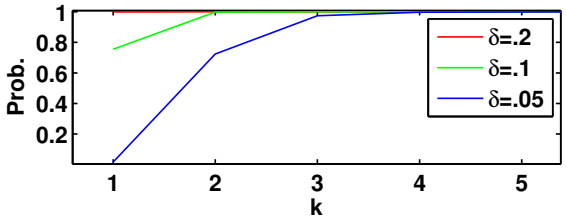
Let us study this with a simple example, consider the texture T as the one shown in Figure 7.4, and a patch $q \subset T$. We will assume that are known a set of k incomplete (randomly sub-sampled) exemplars of the same texture patch (as shown in Figure 7.4). It is easy to see that the probability of completely recover the patch $q \in \mathbb{R}^D$ (with D pixels) from a set of k incomplete exemplars is:

$$P\left(\text{recovering } q\right) = \left(1 - (1-\delta)^k\right)^D, \tag{7.13}$$

where δ is the sampling density (known/total pixels). $1-(1-\delta)^k$ is the probability of recovering a certain pixel by piling the k exemplars. While the probability of recovering the entire patch is the product of D pixel probabilities. The probability $P\left(\text{recovering } q\right)$ is plotted in Figure 7.5(a) with increasing k for the case of patches with $D = 49$ pixels and with sampling densities δ of 0.05, 0.1, and 0.2.



(a) Probability of recovering a patch without considering overlaps.



(b) Probability of recovering a patch considering the overlap (in the final image) of other 49 patches.

Figure 7.5: Probability of recovering a patch of size 7×7 ($D = 49$), considering k exemplars. The exemplars have densities δ: 20%, 10% and 5%. The graph in (a) shows the probability of recovering a single patch as function of k. In the plot the crosses are obtained by empiric verification of the formula (7.13). In (b) is shown the estimation of the probability (7.14) of recovering a patch considering the overlap of other 49 patches.

From Figure 7.5(a) we see that in order to recover a patch q with high probability, using exemplars with a density $\delta = 0.05$ we need $k > 160$ exemplars. For most textures it is easy to find many exemplars of the same pattern as shown in Figure 7.4. But 160 patches is a respectable quantity that in some cases (as in Figure 7.6) is not available, still the interpolation algorithm manages to recover most of the textures.

If we consider that the recovered patches are going to be overlapped for assembling the texture T, then we see that not all the patches need to be recovered completely. Since the interpolation algorithm assembles all the recovered patches in an image, then the lack of pixels in one patch may be compensated by the abundance in its neighbor. Therefore considering the overlapping of patches allows to adjust the probability of recovering the texture:

$$P(\text{recovering q considering overlap}) = (1 - (1 - \delta)^{k\,D})^{D}. \qquad (7.14)$$

That is, for each pixel of the texture we compute: the probability of not recovering a certain pixel considering k exemplars plus the overlap of other $k * (D - 1)$ exemplars. This probability grows very fast with k (as seen in Figure 7.5(b)), ob-

serve that with an average of 3 exemplars per patch is sufficient to recover the patch q with a high probability, independently of the density δ. The previous statement means that there are (in average) 3 exemplars for each one of the 49 overlapped patches.

Figure 7.6: Interpolation with only few exemplars. From left to right: the original image (140×25 pixels), the mask of known pixels (with 194 samples) and the interpolation obtained using the scheme O with $h = 100$.

Multiple textures. In the previous discussion we considered the recovery of a single textured patch, this hides the complexity of finding the correspondence relation between the exemplars. Let us not concentrate on the dynamic of the interpolation algorithm. A remarkable fact is that the proposed method is capable of discerning and recover several textures simultaneously. In part this is due to the locality of the search window (mentioned earlier); However following [ACS09] we can also justify this behavior (at least partially for patch NL-means) by interpreting the interpolation algorithm in terms of blurring mean-shift [Che95].

Mean-shift [FH75] is a procedure for estimating the modes of a density. Let us consider a dataset $\{\mathrm{z}_n\}_{n=1}^N \subset \mathbb{R}^D$ and define a kernel density estimate

$$p(\mathrm{z}) = \frac{1}{N}\sum_{n=1}^{N} K\left(\left\|\frac{\mathrm{z}-\mathrm{z}_n}{h}\right\|^2\right),$$

with bandwidth h and kernel $K(t) = e^{-t/2}$ for Gaussian. Mean-shift iteratively finds for the modes of the density p, it does so by arranging the equilibrium equation $\nabla p(\mathrm{z}) = 0$ into an iterative scheme $\mathrm{z}^{\tau+1} = f(\mathrm{z}^\tau)$ with

$$f(\mathrm{z}) = \sum_{n=1}^{N} \frac{K'\left(\left\|\frac{\mathrm{z}-\mathrm{z}_n}{h}\right\|^2\right)}{\sum_{n'=1}^{N} K'\left(\left\|\frac{\mathrm{z}-\mathrm{z}_{n'}}{h}\right\|^2\right)} \mathrm{z}^n. \tag{7.15}$$

In blurring mean-shift [Che95], each point z_m of the dataset actually moves to the point $f(\mathrm{z}_m)$. That is for each point z_m in the dataset we obtain a new point $\tilde{\mathrm{z}}_m \leftarrow f(\mathrm{z}_m)$. After one iteration the blurring mean-shift results in a new dataset $\{\tilde{\mathrm{z}}_n\}_{n=1}^N$ which is a blurred version of the previous one. This process is known [Che95, CPn06] to hierarchically collapse the datasets into clusters, where the final number of clusters depend on the support of the kernel (in the case of a Gaussian kernel it is a single cluster).

Going back to the analysis of our algorithm. For our case let us further simplify the situation by considering the potential that corresponds to taking $\alpha = 0$

and $\beta = 1$ in (7.4).Rewriting the corresponding image update (7.11) and undoing the variable change we get

$$\sum_{y\in\Omega_p} u(x+y) = \sum_{y\in\Omega_p}\sum_{\hat{x}\in D_2} \frac{w(x,\hat{x})\mathcal{X}_{O^c}(\hat{x}+y)}{\sum_{\hat{x}\in D_2} w(x,\hat{x})\mathcal{X}_{O^c}(\hat{x}+y)} u(\hat{x}+y), \quad \forall x \in D_1. \tag{7.16}$$

In the inner sum of (7.16) we recognize the structure of the mean-shift update rule (7.15)

$$[f(\mathrm{z}_x)](y) = \sum_{\hat{x}\in D_2} \frac{w(x,\hat{x})\mathcal{X}_{O^c}(\hat{x}+y)}{\sum_{\hat{x}\in D_2} w(x,\hat{x})\mathcal{X}_{O^c}(\hat{x}+y)} \mathrm{z}_{\hat{x}}(y), \quad y \in \Omega_p$$

where the patches are denoted as vectors $\mathrm{z}_m = (p_u(m+y))_{y\in\Omega_p} \in \mathbb{R}^{|\Omega_p|}$, and the indicator functions $\mathcal{X}_{O^c}$ are accounting for the possibility of unknown pixels in $p_u(\hat{x})$. The weights $w(x,\hat{x})$ are exponential defined in (7.10), they correspond to the kernel's derivative. The update rule (7.11) becomes an average of mean-shifted patches

$$u(x) = \sum_{y\in\Omega_p} [f(\mathrm{z}_x)](y). \tag{7.17}$$

In this case the dataset is a set of points correspond to the set of image patches. Modulo the average that takes place in (7.17) this procedure behaves as blurring mean-shift, and we expect the dataset to collapse into clusters. However, in our case the points are not free to move in all the directions, because several dimensions of each point are locked. These dimensions correspond to the known pixel values of each patch, which translate into hard constraints on the mean-shift process. Since the points are constrained to different subspaces, the evolution of the dataset will not converge to a single cluster. The evolution will drive the clusters to the intersection of the hyperplanar constraints, hopefully these will be nearby the original image patches.

7.6 Conclusions and future work

In this chapter we extended the variational framework for exemplar based image inpainting to image interpolation. Being based on the self-similarity prior, the resulting algorithms are capable of hallucinating textures even for severely undersampled images. And sometimes, when the sampling is sufficiently irregular and the texture is big enough, then the hallucinated textures are correct. These results suggest that imposing adequately a self-similarity prior may lead to restoration algorithms capable of recovering from aliasing artifacts. In the future we will extend the current model to consider samples located at non-integer positions.

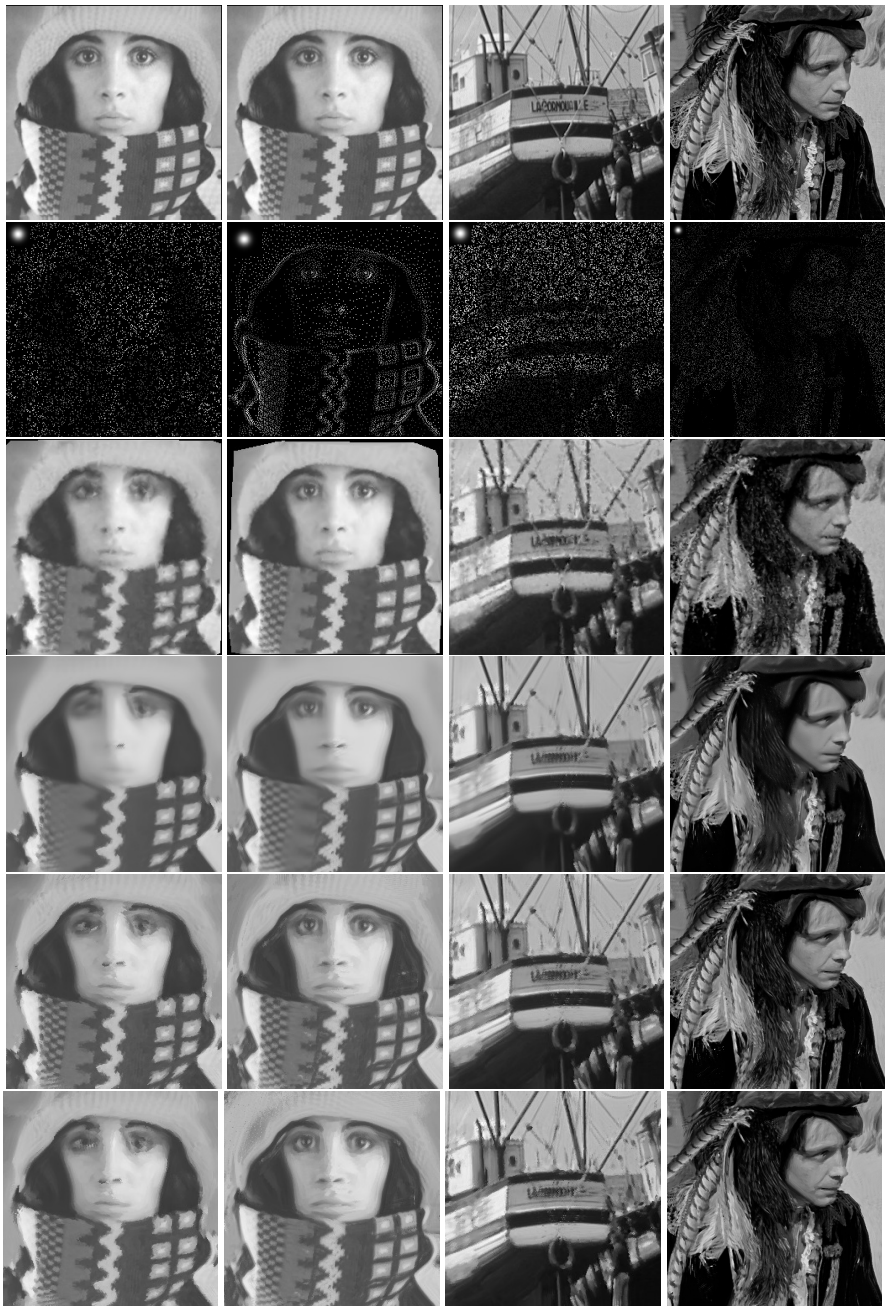


Figure 7.7: Sparse sampling interpolation. 1st row: original images. 2nd row: input data with sample densities of 10%, 10% dithered [BBBW08], 20% and 9%. 3rd row: linear interpolation over the Delaunay triangulation of the samples; PSNRs: 25.8, 30.6 (not considering the black frame), 25.0 and 22.74. 4rt row: results of method AB with $h = 100$; PSNRs: 22.5 ,22.7, 25.5 and 22.79. 5th row: results of AB with $h \to 0$; PSNRs: 22.6, 23.0, 25.5 and 22.56. 6th row: results of the method O with $h \to 0$; PSNRs: 22.6,21.7,25.1 and 22.69. (Details can be better appreciated by zooming on a computer screen)

Figure 7.8: Inpainting experiment. From left to right top to bottom: the original image *Mitama Matsuri* 200×113, the available samples (the sampling density outside the hole is 20% yielding a global sampling density of 14%), the result of linear interpolation over the Delaunay triangulation and the result of algorithm *B*. The algorithm makes no distinction between the hole and the sampled regions.

Figure 7.9: Interpolation with lower sampling density. From left to right top to bottom: original image *Barbara* 512×512, the available samples 5% of the samples, the result of linear interpolation over the Delaunay triangulation (PSNR 20.1), results of the Laplacian interpolation (PSNR 19.9) and the output of algorithm O with $h = 100$ (PSNR 22.8).

Part III

Digital elevation model interpolation

Chapter 8

Range data interpolation

The acquisition of range data is generally associated, due to different factors, to some data loss. This produces holes in the obtained data and, therefore the elevation maps need to be interpolated. When a reference image of the scene is available it is possible to interpolate and to increase the spatial resolution of the underlying measurements by incorporating geometric information provided by the image. In this chapter we motivate the problem of digital elevation model interpolation from sparse or incomplete data. We explain under what circumstances the elevation maps contain holes and outline three methods for interpolation them which are described in detail in the following chapters.

8.1 Motivation: interpolation of elevation data

The demand for high resolution remote sensing products on urban areas has experimented a rapid growth in the last times. In the case of *Digital Elevation Models* (DEM) and High Resolution Imagery for civil and defense applications, not only has grown the need of accuracy but also the amount of data (see for instance Google Earth), making impractical any non-automatic technique.

At present, the urban elevation data is mainly harvested from cadastral data, acquired with LIDAR range scanners (light detection and ranging or laser imaging detection and ranging), radar interferometry or applying 3D reconstitution methods to aerial images. Except for the cadastral data, the other techniques generally require post-processing of the data to impose regularity of the reconstructed surface or to fill the gaps in the acquisition.

We will concentrate on data produced by a particular class of stereo algorithms called *small baseline* stereo [SAM08, SMA10]. The baseline or stereoscopic coefficient is the distance of the cameras between the two acquisitions. It is common to admit that the accuracy of the disparity measure (which is directly related to the distance of the object) depends only on the resolution of the images, so to obtain the maximal precision the baseline should be increased. But in [Del04] it is shown that if the images are correctly interpolated and the disparities are computed with sub-pixel precision, then the precision is not dependent on the stereoscopic coefficient.

Small baseline stereo is particularly attractive because it reduces the occlusions artifacts, still allowing to compute disparities under strong sub-pixel accuracy requirements [Del04, SMA10]. However disparities at such precisions cannot be computed everywhere in the image, leaving gaps in the model. Although data acquired by LIDAR and radar interferometry also requires post-processing because of occlusions, noise and limited spatial resolution, we will not focus on them.

Stereopsis is a process that allows to reconstruct the depth information of a scene from two images. It can be subdivided into four steps: *camera calibration*, *rectification*, *correspondence computation* and *reconstruction*. In the calibration and rectification steps the physical geometry of the camera system (position, orientation, focal length...) is estimated and the images are re-projected on planes that facilitate the correspondence task. The correspondence step computes depth of corresponding features observed in the two images. The reconstruction takes the correspondences and construct a feasible model that explains them. Even though the main subject of this part is the reconstruction step, we should first understand the artifacts and characteristics of the data provided by the correspondence methods. So, we comment here on the correspondence methods, and we will discuss the reconstruction in next section.

Stereo algorithms extract depth information from two or more images of a scene by finding corresponding features between the images. If the correspondence is correct we can effectively compute the depth of a feature (a point for instance) by computing its *disparity*, however not all the features can be matched faithfully. It is clear that, to overcome this uncertainty some hypothesis about the regularity of the scene must be made in the reconstruction step. Common hypotheses are based on reasonable physical and geometrical properties of the scene.

Correspondence methods are classically [BBH03] grouped into: *local correspondence methods* for those that rely on local information to determine the corresponding point and *global correspondence methods* for those that rely on information of a whole *scanline* or the entire image to compute the solution.

Global correspondence methods. The global methods produce dense disparity maps. The correspondence is computed using discrete optimization techniques such as graph cuts [KZ06], belief propagation [SZS03], dynamic programming [FKO$^+$04] or using continuous variational techniques [PSGB08, GKC]. These methods generally rely on smoothness constrains imposed on the disparity map that range from total-variation to piecewise smooth models.

In a certain way, the global correspondence methods also include the reconstruction phase, because the smoothness constraints translate into priors for the reconstructed model. However since the correspondence is generally computed as a minimum of a global objective function it is difficult to asses the precision of the individual measurements since the errors are diluted in the global optimization.

Local correspondence methods. The local correspondence methods match local features from one image to local features in the other image. Block matching methods are among the most used local correspondence methods, for which the feature is a window (block) surrounding the current point. The correspondence is computed by minimizing a distance measure of a neighborhood of the interest pixel. The blocks are usually compared with NCC (Normalized Cross Correlation), SSD (Sum of Squared Differences) or SAD (Sum of Absolute Differences). The choice of this function determines the performance of the method and depends on the application of the specific algorithm, for instance SSD and SAD are very fast but are very sensitive to radiometric changes between the images, see [SS02] for a compendium of methods.

Independently of the choice of the block distance measure, a minimum of this distance does not assure a correct match. In general only a proportion of the image can be reliably matched [SMA10]. Due to the so called "flattening effect" [SB97, Del04] block matching assigns the wrong disparity to points whose window touches some edge of the model. Some techniques alleviate this problem by adapting the shape of the window [BGGP07, FRT97, KSC01, KO94] or matching other geometric primitives [SZ00]. Another approach is to detect these situations. The techniques pioneered by Delon and Rougé [Del04] and further developed by Sabater *et al*. in [SMA10] allow to asses the quality of the matches by using an contrario rejection algorithm. As a result the disparity information provided by this type of algorithms is very precise but incomplete, particularly near the boundaries of the objects.

While local methods are very efficient they are sensitive to ambiguous regions (like uniform zones or occlusions), while global methods are less sensitive to local problems due to their global nature (see Scharstein and Szeliski taxonomy [SS02] for a comparison of algorithms). But also due to their global nature, their precision is harder to asses locally and are more computationally expensive. We will concentrate on the interpolation of incomplete disparity data provided by local methods.

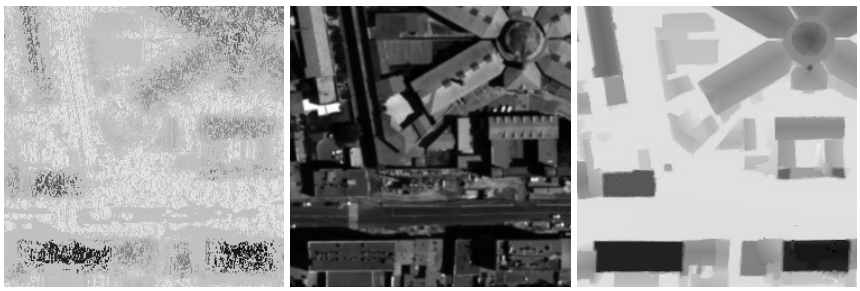


Figure 8.1: DEM data. From left to right: a sparse DEM to be interpolated, the reference image and the ground truth elevation data for the represented area. The elevation is encoded as a gray level image (higher elevation values are darker).

8.2 Model reconstruction

We have seen that the local correspondence methods generally produce sparse and not equally significant measurements over all the image. In Figure 8.1 we show an incomplete correspondence map obtained after rejecting some of the correspondences by the criterion proposed in [Del04] (the *correlation curvature*). To obtain a dense model, the measurements must be interpolated and/or regularized. The objective is to interpolate the known data according to some regularity prior, taking into account the shapes and structures present in an image of the scene. The regularity prior can be imposed by requiring a minimal surface, minimum curvature of the surface [MS05], as piecewise models such as the affine clusters in [ABGvG10], or affine regions obtained from a segmentation of the image [IPG$^+$07, CJDS03, KSK06]. Other approaches use higher level priors that impose the scene structure [LDZP08, KP09] (in this case roofs). Although some methods may interpolate the model using only the depth information [ABGvG10], when available, the information provided by an image of the scene is valuable and most methods will take advantage of it.

When the elevation data is very sparse it becomes harder to determine the boundaries between the different objects, and this is the case we are interested in. The aim is to obtain geometrically meaningful interpolation of the data by imposing the shapes observed in a reference image of the scene. This idea has been applied before for image fusion [BCBV06, Alv02]. In the current context it corresponds to the fusion of elevation data with the image.

8.3 Anisotropic diffusion and model interpolation

In the following chapters we discuss three models for diffusion and interpolation of digital elevation models, based on the *Lambertian hypothesis*. That is, a uniform surface (in the model) with a constant angle will be seen with a constant intensity in the image. This assumption allow us to extrapolate information across uniform regions of the image.

We must acknowledge that areas of constant intensity in the image do not necessarily correspond to the same surface in the model (these are false negatives of the edge detection), but this situation is unlikely since in most cases a discontinuity in the model corresponds to a change of intensity in the image. On the other hand, heavily textured objects will induce false boundaries in the model (false positives). Although the latter should not be a problem if there are enough samples for recovering each apparent part of the object. The problem is that using the gray level as a cue for determining a surface can be misleading. To alleviate this problem other (richer) descriptors, of texture for instance, can be used [SGT02, BS07], but we are not going to consider them.

Chapter 9 approaches the interpolation problem from a variational perspective, using an anisotropic regularizer based on a popular approximation of the

total variation. For simplicity the functional is formulated in the discrete setting

$$\min_{u \in U} \sum_{x \in N^2} \sqrt{\beta^2 + |A(x)\nabla u(x)|^2} + \lambda \sum_{x \in N^2} w(x)|u(x) - f(x)|^p,$$

where $U := \mathbb{R}^{N \times N}$, β is a small constant, $A(x) \in \mathbb{R}^{2\times 2}$ determines a tensor field that defines the anisotropy (when $A = Id$ we recover the isotropic case), $p = \{1, 2\}$ selects the data fitting term, $f, w \in U$ are respectively the initial data and a binary mask that indicates where this data is valid. The main property of the total variation is that preserves the discontinuities of the model u. The discontinuity preserving property is important for many Computer Vision problems, e.g. to preserve edges in the solution. And the approximation $\beta > 0$ avoids the difficulties associated with the non-differentiability of the total variation. Moreover the model with $\beta > 0$ can also be interpreted as a minimal surface regularization, which seems appropriate for this application. The tensor $A(x)$ is defined so that the diffusion is inhibited across the level lines of the reference image I. The direction of the level lines is approximated by $z(x)^\perp$, where $z(x) = \frac{\nabla I(x)}{\sqrt{1+|\nabla I(x)|^2}}$ and $z(x)^\perp$ denotes the counterclockwise rotation of $z(x)$ of angle $\frac{\pi}{2}$. Then the tensor $A(x) = Id - z(x) \otimes z(x) = z^\perp(x) \otimes z^\perp(x)$.

The second model (Chapter 10) is formulated in the context of anisotropic Cheeger sets (sets that minimize the quotient between an anisotropic perimeter and its area). The model we consider is based on the anisotropic total variation for a general class of anisotropies ϕ (to be made precise in Chapter 10)

$$\min_{u \in BV_\phi(\Omega) \cap L^2(\Omega, h\,dx)} \left\{ \int_\Omega |Du|_\phi + \frac{\lambda}{2} \int_\Omega (u - f)^2 h\,dx + \int_{\partial\Omega} \phi(x, \nu^\Omega)|u|\, d\mathcal{H}^{N-1} \right\},$$

where $\phi : \Omega \times \mathbb{R}^N \to \mathbb{R}$ is a metric integrand in Ω which is symmetric in ξ, i.e., $\phi(x, -\xi) = \phi(x, \xi)$ for any $\xi \in \mathbb{R}^N$ and any $x \in \Omega$; $h \in L^\infty(\Omega)$, $h(x) > 0$ a.e., with $\int_\Omega \frac{1}{h(x)}\,dx < \infty$; $f \in L^2(\Omega, hdx)$, $\lambda > 0$; and ν^Ω denotes the outer unit normal to Ω at points of $\partial\Omega$. We denote by $L^2(\Omega, hdx)$ the set of measurable functions $u : \Omega \to \mathbb{R}$ such that $\int_\Omega u^2\, h\,dx < \infty$. We apply this formalism to compute the global minima of the geodesic active contours model, and to the interpolation problem. We define a diffusion tensor similar to the previous case, that also captures the magnitude of the edge in addition to its orientation. Thus $\phi(x, \xi) = |A_x \xi|$, where for each $x \in \Omega$ A_x is a symmetric positive definite matrix defined as $A_x = z^\perp(x) \otimes z^\perp(x) + \frac{1}{\sqrt{1+|\nabla I(x)|^2}} z(x) \otimes z(x)$.

Finally, in Chapter 11 we propose a new interpolation algorithm that uses geodesic distances to propagate intensities to neighboring pixels (neighbors in the geodesic sense). The geodesic distance between two points $a, b \in \Omega$ is

$$d(a, b) = \min_{\gamma_{a,b}} \int_0^1 |\nabla I(\gamma_{a,b}(t)) \cdot \dot{\gamma}_{a,b}(t)| dt,$$

where $\gamma(t) : [0, 1] \to \Omega$ is a curve in Ω, $\gamma_{a,b}$ denotes the curve connecting a and b such that $\gamma_{a,b}(0) = a$ and $\gamma_{a,b}(1) = b$, and $\dot{\gamma}(t)$ denotes the tangent of the

curve at t. The distance accumulated by a curve crossing a uniform area of I or moving perpendicular to ∇I is 0, while the length increases when the curve crosses a steep gradient. This means that pairs of points in the same level line have small distances. This permits to define geodesic neighbours of each sample in f which are used in the interpolation step to define affine planes (see Chapter 11 for more details).

Chapter 9

Constrained Anisotropic Diffusion and some Applications

In this chapter we consider an anisotropic minimal surface regularizer coupled with a robust data fitting term for interpolation of sparse digital elevation models. The proposed regularizer constraints the diffusion of the elevation data to the isophotes of a reference image u. This allows to recover geometric features not present in the elevation data.

9.1 Introduction

Among the stereo matching methods [BBH03, SS02], the algorithms based on windowed correlation allow to achieve the highest accuracy in terms of depth estimation [Del04, SMA10]. This feature is particularly attractive for the acquisition of *digital elevation models* (DEM) using photogrammetric techniques. However, due to the local nature of correlation methods, the elevation data they provide is not equally accurate at all the points of the image. In Figure 9.2 we show an incomplete correspondence map obtained after rejecting the correspondences with high error as estimated by the criterion proposed in [Del04] (the *correlation curvature*).

In order to recover a dense model, the elevation data must be interpolated. The objective is then to interpolate the sparse data by diffusing them to unknown areas, while imposing certain regularity of the interpolated model. For urban scenes the target models are discontinuous, therefore the regularity prior must allow discontinuities. Since the relevant discontinuities are likely to occur also as edges in the image of the scene, this image should be taken into account in the definition of the model prior. Thus, we propose to interpolate the sparse digital elevation data by a diffusion process constrained by the geometry extracted from a reference image [BCBV06, Tsc02], coupled with a data fitting term.

For simplicity, we shall work in a discrete setting, and thus images are represented as vectors in $X = \mathbb{R}^{N\times N}$ (without loss of generality we consider square images), we denote by $\Omega = \{1, 2, \ldots, N\}^2$, the image domain. We state our

problem as the minimization of a functional:

$$\min_{p \in X} E(p) + \lambda D(p), \tag{9.1}$$

where $D(p)$ is a data term to be made precise later, and $E(p)$ is an anisotropic regularization term

$$E(p) = \sum_{x \in \Omega} \sqrt{\beta^2 + |A(x)\nabla p(x)|^2}. \tag{9.2}$$

For each $x \in \Omega$ $A(x)$ is a 2×2 matrix and ∇ denotes a suitable discretization of the gradient (to be made precise later). The matrix A is called the *structure tensor* and allows to embed information on the directions were we want to constrain the diffusion process. The regularizer (9.2) corresponds to an anisotropic minimal surface regularizer. The minimal surface regularizer [FK98] can be recovered by setting $A = Id$, $\beta = 1$ in (9.2). The gradient descent minimization of (9.1) is:

$$\frac{\partial p}{\partial t} = \operatorname{div}\left(\frac{A^\intercal A \nabla p}{\sqrt{\beta^2 + |A\nabla p|^2}}\right) - \lambda \frac{\partial D}{\partial p}(p), \tag{9.3}$$

plus Neumann boundary conditions, where $-\operatorname{div}$ denotes the adjoint of the gradient ∇ in X.

Diffusion equations of the form:

$$\partial_t p = \operatorname{div}(\mathcal{A}\nabla p),$$

where $\mathcal{A}$ is a *diffusion tensor*, i.e., a map which associates to each $x \in \Omega$ a positive definite matrix $\mathcal{A}(x)$, have been studied in many contexts. In particular, Weickert [Wei98] studied the continuous, the time discrete and spatio-temporal discrete diffusion models and its applications to image processing; Tschumperle considered this problem [Tsc02] applied to the diffusion of multi-valued data. In [WSW06] Welk et al. have developed the so called Locally Analytic Schemes (LAS) for this type of problems and analyzed its connections to wavelet shrinkage. Graph-cuts have also been applied to similar problems [BVZ01, KZ06], however for the precision we are considering here the memory requirements grow prohibitively as we increase the number of labels i.e. the number of quantized altitudes.

In next section we propose a model for DEM interpolation and we develop an explicit numerical method for solving (9.3). Then and we analyze it in Section 9.3. This explicit scheme, although slower (in terms of convergence) than an implicit one, allows to easily combine the regularizer with a wide range of data terms such as the one described in [FAP05]. In Section 9.4 we apply the proposed method to the interpolation of digital elevation models and present some experiments to illustrate its performance.

It is worth noting that the algorithm we developed was also applied in the context of diffusion of probability vector maps. *Vector probability diffusion* (VPD)

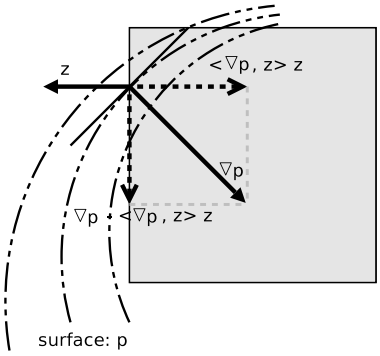


Figure 9.1: Regularization subject to level lines. The evolution of surface p (dotted level set) is constrained by the geometry (square) given by its normal z. The gradient of p is projected over the basis $\{z, z^{\perp}\}$ and $\nabla p - \langle \nabla p, z \rangle z$ is aligned with $z^{\perp}$.

[PS01] is a tool for improving classification that diffuses the probability maps of concrete features detected in an image. The anisotropic version of (VPD) was proposed in [Lec05].

9.2 An anisotropic diffusion model for DEM regularization and interpolation

Assume that $u \in X$ is a given reference image of a stereo pair from which we have computed the disparity map $m : D \to \mathbb{R}$. We denote by D the subset of Ω where m can be reliably computed. We define the mask function $w(x) = 1$ if $x \in D$ and 0 otherwise. We assume that, for each known point x, an estimate of the error's standard deviation $Err(x)$ is available (see [Del04, SMA10] for details on how this estimate could be computed). The height map m we are considering is incomplete, specially after considering the mask w.

We will interpolate m with a function p on the full $N \times N$ grid by minimizing the energy (9.1). Assuming that the geometry of the reference image u (expressed in terms of its edges) is correlated to the height p, we define as in [BCBV06, Tsc02] a structure tensor that inhibits the diffusion of p across the level lines of u. In this way, we correlate the geometry of u and p.

Structure tensor. The structure tensor is obtained as $A(x) = (I - z(x) \otimes z(x))$, where the vector field of unit normals to the level lines of the image u is defined as $z(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$ if $|\nabla u(x)| \geq b$ and $z(x) = \frac{\nabla u(x)}{\sqrt{b^2 + |\nabla u(x)|^2}}$ otherwise. The constant $b > 0$ is a threshold on the gradient which ensures that its direction is reliable, in our cases we select $b = 50$ for images with a range $[0, 3000]$.

The effect of the structure tensor (depicted in Figure 9.1) is a projection of the

gradient field ∇p in the direction of z:

$$A\nabla p = \begin{pmatrix} 1 - z_x^2 & -z_x z_y \\ -z_x z_y & 1 - z_y^2 \end{pmatrix} \nabla p = \nabla p - \langle z, \nabla p \rangle z.$$

Data fitting terms. We consider data terms based on the L^2- and L^1-norms. The weighted L^2-norm yields

$$D_2(p) = \sum_{x \in \Omega} w(x) |p(x) - m(x)|^2, \tag{9.4}$$

where w is the binary function described before.

To avoid bias due to outliers in m we also consider a term based on the L^1-norm

$$D_1(p) = \sum_{x \in \Omega} w(x) \sqrt{a^2 + |p(x) - m(x)|^2}, \tag{9.5}$$

$a > 0$. This term keeps the robustness of the L^1-norm while avoiding the numerical difficulties associated to the non-differentiability of the L^1-norm at 0.

When the measurements m have a known and uniform error variance σ^2 it is possible to impose the data fitting as a constraint $D_2(p) \leq \sigma^2 L$, where L is the number of pixels such that $w(x) \neq 0$ (or $D_1(p) \leq \sqrt{a^2 + \sigma^2} L$). Then the value of λ is chosen so that the constraint is satisfied (for instance, by means of Uzawa's method [BCRS03, GSZ06, ABCH08]). However in general the disparities computed with stereo correlation have a non uniform error variance. In this case a single Lagrange multiplier as in (9.1) will give an unfair weight to the values with high variance.

Indeed, in [Del04] and [SMA10] the authors suggest that the disparities computed near the image borders are more reliable than in flat areas and provide an estimate of the error variance (in the case of [SMA10] a quite precise one). Therefore we can assume that we know a local estimate $Err(x)^2$ of the error's variance at each point x, and that this error has 0 mean (the differences $(p - m)$ have 0 mean) .

Then, instead of adapting locally the value of λ as in [BCRS03, GSZ06, ABCH08], we include the error estimate directly into the data terms obtaining the two new terms:

$$D_2^{**}(p) = \sum_{x \in \Omega} w(x) \frac{|p(x) - m(x)|^2}{Err(x)^2}, \tag{9.6}$$

$$D_1^{**}(p) = \sum_{x \in \Omega} w(x) \sqrt{1 + \frac{|p(x) - m(x)|^2}{Err(x)^2}}. \tag{9.7}$$

For both terms the new expected error per pixel is 1, and we write the constraints as: $D_2^{**}(p) \leq L$ and $D_1^{**}(p) \leq \sqrt{2} L$. This observation allows us to use a unique value of the multiplier λ for the whole data term.

9.3 Numerical Analysis of Constrained Diffusion

We denote by X the Euclidean space $\mathbb{R}^{N\times N}$. The Euclidean scalar product and the norm in X will be denoted by $\langle\cdot,\cdot\rangle_X$ and $\|\cdot\|_X$, respectively. Then the image $u\in X$ is the vector $u=(u(x))_{x\in\Omega}$. If $u\in X$, the discrete gradient is a vector in $Y=X\times X$ given by $\nabla u(x)=(\nabla^1 u(x),\nabla^2 u(x))$, where

$$\nabla^1 u(i,j)=\begin{cases} u(i+1,j)-u(i,j) & \text{if} \quad 1\le i<N\\ 0 & \text{if} \quad i=N,\end{cases}$$
$$\nabla^2 u(i,j)=\begin{cases} u(i,j+1)-u(i,j) & \text{if} \quad 1\le j<N\\ 0 & \text{if} \quad j=N,\end{cases}$$

are the horizontal and vertical partial derivatives computed with forward differences (the pairs (i,j) represent the coordinates of $x\in\Omega$). Other choices of the gradient are possible, this one will be convenient for the developments below.

We denote the euclidean norm of a vector $v\in\mathbb{R}^2$ by $|v|$. We define the Euclidean scalar product in Y in the standard way by $\langle r,q\rangle_Y=\sum_{x\in\Omega}(r^1(x)\,q^1(x)+r^2(x)\,q^2(x))$ for every $r=(r^1,r^2)$, $q=(q^1,q^2)\in Y$. The norm of $q\in Y$ is, as usual, $\|q\|_Y=\langle q,q\rangle_Y^{1/2}$. We define the discrete divergence div as the dual operator of ∇, i.e., for every $p\in Y$ and $u\in X$ we have

$$\langle -\text{div } p,u\rangle_X=\langle p,\nabla u\rangle_Y.$$

One can easily check that div is a discrete divergence computed using backward differences.

Let us consider the discrete functional:

$$E(p)+\lambda D(p)=\sum_{x\in\Omega}\sqrt{\beta^2+|A(x)\nabla p(x)|^2} \tag{9.8}$$
$$+\lambda\sum_{x\in\Omega}w(x)\alpha(x)\sqrt{a^2+|p(x)-m(x)|^2},$$

where $w(x)$ is a binary mask defined in Section 9.2 and $\alpha(x)>0$ is a weighting function to be described later. We propose to minimize (9.8). Since the functional $E+\lambda D$ is strictly convex, there is a unique minimum of it in X, call it p^*. To minimize $E(p)+\lambda D(p)$ we use an explicit gradient descent method, i.e., an iterative scheme of the form $p^{n+1}=p^n-t_n\nabla(E+\lambda D)(p^n)$ with $t_n>0$ for all n, which in the present case can be written as:

$$p^{n+1}=p^n+t_n\left(\text{div}\left(A^t_{p^n}A_{p^n}\nabla p^n\right)-\lambda w\alpha\frac{p^n-m}{\sqrt{a^2+|p^n-m|^2}}\right), \tag{9.9}$$

plus Neumann boundary conditions, where $A_p=A_p^t=\frac{A}{(\beta^2+|A\nabla p|^2)^{1/4}}$.

Observe that, since we are in the discrete case, the term inside the parenthesis is bounded in X with a bound independent of n. Moreover, since the minimum of $E+\lambda D$ is unique, using Theorem 11 below (see [Sho85]) we have that $\{p^n\}$ converges to p^*

Theorem 11 (Shor 1985). *Let X be a vector space of finite dimension and J be a convex function defined on X which has a bounded set of minimum points X^*. Assume that t_n is a sequence of positive numbers satisfying the conditions*

$$t_n > 0, \quad \lim_{n\to+\infty} t_n = 0, \quad \sum_{n=0}^{+\infty} t_n = +\infty, \tag{9.10}$$

and let x^k be a sequence of vectors in X generated by the formula: $x^{k+1} \in x^k - t_k \partial J(x^k)$, where $\partial J(x_k)$ is a subgradient of J at x^k with initial condition given by some $x^0 \in X$. If the sequence $\partial J(x^k)$ is bounded, then there is a subsequence converging to some point $x^* \in X^*$. If the minimum of J is unique, then x^k converges to it.

The previous result applies to three cases since the formulation (9.8) includes:

- **Case 1:** $\alpha(x) = 0$. In this case we are only considering the energy $E(p)$, and its evolution is given by: $p^{n+1} = p^n + t_n \text{div}\left(A^t_{p^n} A_{p^n} \nabla p^n\right)$.
- **Case 2:** $\alpha(x) = 1$, a being a positive constant and $w \in X$ being a binary mask. It corresponds to the minimization of $E(p) + \lambda D_1(p)$ (Equation (9.5) as described in Section 9.2).
- **Case 3:** $\alpha(x) = \frac{1}{Err(x)}$; $a = Err(x)$. In this case the value of $\alpha = 1/a$ is not constant and can be interpreted as a weighting function to impose a local control on the errors. This applies to the minimization of $E(p) + \lambda D_1^{**}(p)$ (Equation (9.7)).

Now let us consider the case with a quadratic data fitting term in the energy, for instance, assume that we replace the term $D(p)$ in (9.1) by:

$$D_2(p) = \sum_{x\in\Omega} \alpha(x) w(x) |p(x) - m(x)|^2 \tag{9.11}$$

and we minimize $E(p) + \lambda D_2(p)$. In this case, $\frac{\partial D_2}{\partial p} = 2\alpha(x)w(x)(p^n - m)$ is not bounded uniformly in p^n and we cannot use Theorem 11. Let us denote

$$\mathcal{C}_n(p) = \text{div}\left(A^t_{p^n} A_{p^n} \nabla p\right), \qquad B_n(p) = \mathcal{C}_n(p) - 2\lambda\alpha w\, p,$$

where $\alpha w p(x) = \alpha(x)w(x)p(x)\ \forall x \in \Omega$. In this case, we may write the gradient descent equations for $E(p) + \lambda D_2(p)$ as:

$$p^{n+1} = p^n + dt\mathcal{C}_n(p^n) - dt\lambda\alpha w 2(p^n - m) = (Id + dtB_n)p^n + 2dt\lambda\alpha w m. \tag{9.12}$$

By iterating we may express p^{n+1} in terms of p^0 by the formula:

$$p^{n+1} = \underbrace{\prod_{i=0}^{n} [Id + dtB_i(p^i)]p^0}_{I} + \underbrace{2dt\ \lambda\alpha w m \left\{1 + \sum_{j=0}^{j=n-1}\left(\prod_{k=j+1}^{n}\Big[Id + dtB_k(p^k)\Big]\right)p^0\right\}}_{II}.$$

The sequence p^n is bounded and the above method is stable. This will follow as a consequence of Proposition 1 below which guarantees the boundedness of the term (I) and the convergence of the series in (II) as soon as $dt < \frac{\beta}{4+\lambda\alpha\beta}$. This condition allows to cover two cases:

- **Case A:** $\alpha(x) = 1$ and $w \in X$ being a binary mask. It corresponds to the minimization of $E(p) + \lambda D_2(p)$: Equation (9.4) in Section 9.2.
- **Case B:** $\alpha(x) = \frac{1}{Err(x)^2}$. In this case the value of α is not constant and can be interpreted as a weighting function to impose a local control on the errors. In this case, we minimize the functional $E(p) + \lambda D_2^{**}(p)$ (Equation (9.6)) .

For a linear operator $A : X \to X$, we denote by $\rho(A)$ its spectral radius, that is $\max\{|\lambda_i| : \lambda_i \text{ eigenvalue of } A\}$.

Proposition 1. *If $dt < \frac{\beta}{4+\lambda\alpha\beta}$, then $\|Id + dtB_n\|_X = \rho(Id + dtB_n) \leq 1 - \varepsilon$ for some $\varepsilon > 0$.*

Proof. Observe that $\mathcal{C}_n$ is a symmetric operator and thus its norm in X coincides with its spectral radius. Now, observe that λ_i is an eigenvalue of $Id + dtB_n$ if and only if $\lambda_i = 1 + dt\mu_i - 2dt\lambda\alpha$ for some eigenvalue μ_i of $\mathcal{C}_n$. Our statement follows if we prove that

$$-1 + \varepsilon \leq 1 + dt\mu - 2dt\lambda\alpha \leq 1 - \varepsilon \quad \forall \mu \text{ eigenvalue of } \mathcal{C}_n. \tag{9.13}$$

Now, we observe that the eigenvalues of $\mathcal{C}_n$ are contained in the interval $[-\frac{8}{\beta}, 0]$. Indeed, if x is an eigenvector of $\mathcal{C}_n$ corresponding to the eigenvalue μ, we have

$$\begin{aligned}\mu|x|_2^2 = \langle \mathcal{C}_n(x), x\rangle_X = \langle \text{div}\,(A_x^t A_x \nabla x), x\rangle_X = -\langle A_x^t A_x \nabla x, \nabla x\rangle_Y \\ = -\langle A_x \nabla x, A_x \nabla x\rangle_Y = -\|A_x \nabla x\|_Y^2 \leq 0\end{aligned}$$

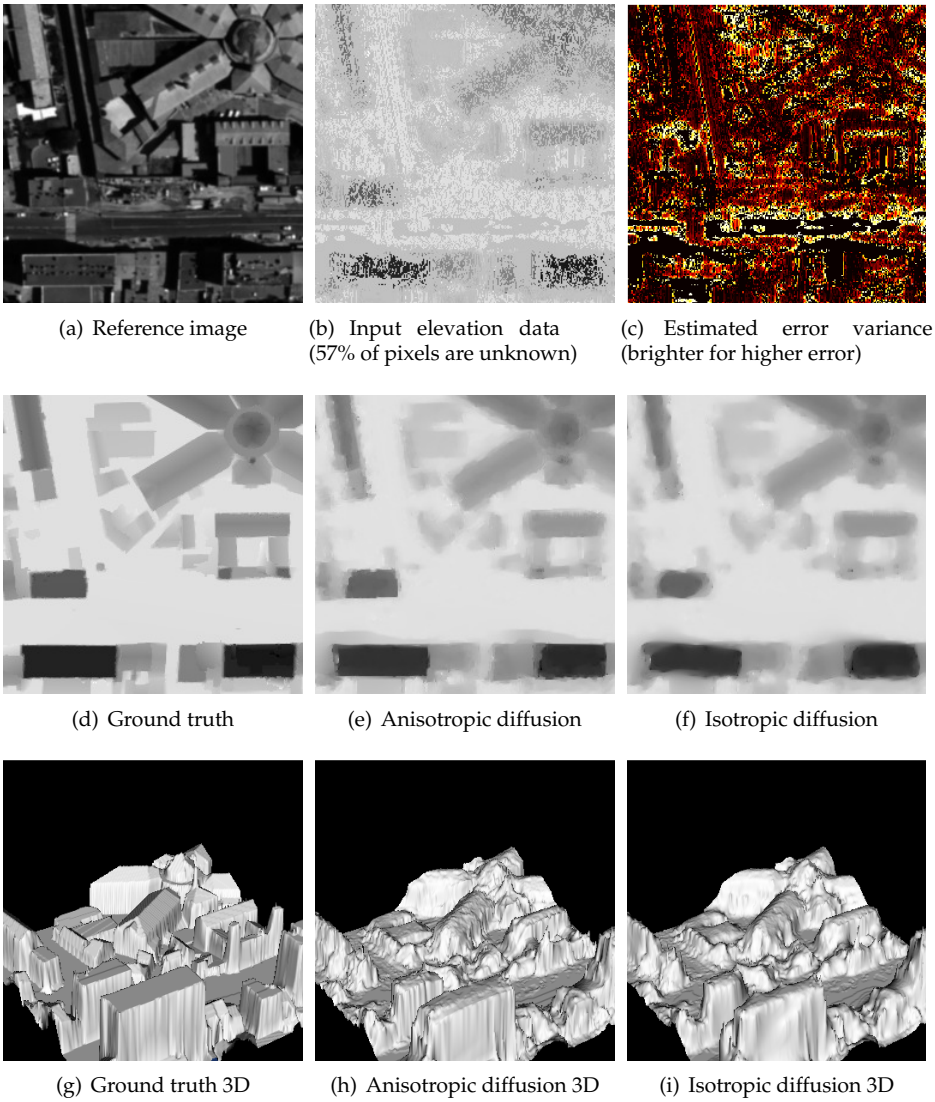
and $\|A_x \nabla x\|_Y^2 \leq \frac{1}{\beta}\|A\nabla x\|_Y^2 \leq \frac{8}{\beta}$ because $\|A_x\| \leq 1$ and $\|\nabla p\|_Y^2 \leq 8\|p\|_X^2$. Since the eigenvalues of $\mathcal{C}_n$ are negative, (9.13) can be written as

$$2 - \varepsilon \geq dt(|\mu| + 2\lambda\alpha) \geq \varepsilon. \tag{9.14}$$

If $\{\mu_i\}$ are the eigenvalues of $\mathcal{C}_n$, this condition is implied if we choose dt such that

$$\frac{\varepsilon}{\min_i |\mu_i| + 2\lambda\alpha} \leq dt \leq \frac{1-\varepsilon}{\max_i |\mu_i| + 2\lambda\alpha},$$

which is implied by a choice of dt such that $\frac{\varepsilon}{2\lambda\alpha} \leq dt \leq \frac{1-\varepsilon}{(8/\beta)+2\lambda\alpha}$. We can find a value of $\varepsilon > 0$ satisfying this as soon as $dt < \frac{\beta}{4+\lambda\alpha\beta}$.



(a) Reference image (b) Input elevation data (57% of pixels are unknown) (c) Estimated error variance (brighter for higher error)

(d) Ground truth (e) Anisotropic diffusion (f) Isotropic diffusion

(g) Ground truth 3D (h) Anisotropic diffusion 3D (i) Isotropic diffusion 3D

Figure 9.2: DEM interpolation. The first row shows the inputs of the method: the reference image, the elevation data to be interpolated and the error map. In the second row are shown the ground truth and two interpolation results obtained with anisotropic and isotropic regularization, both using the D_1^{**} data fitting. The elevations are encoded as gray level images. The third row displays the corresponding elevation models.

	ISOTROPIC		ANISOTROPIC	
DATA TERMS	Mean d_1	Mean d_2	Mean d_1	Mean d_2
D_1	1.0165	0.0367	1.0106	0.0236
D_1^{**}	1.0143	0.0321	1.0115	0.0254
D_2	1.0199	0.0438	1.0118	0.0259
D_2^{**}	1.0158	0.0358	1.0123	0.0274
initial condition	1.0241	0.0552	N/A	N/A

Table 9.1: Experiments with DEM interpolation. We report here different combinations of interpolation algorithm and data terms. The errors are measured against the ground truth information as the mean of $d_1(p(x), m(x)) = \sqrt{a^2 + |p(x) - m(x)|^2}$ and the mean of $d_2(p(x), m(x)) = |p(x) - m(x)|^2$.

9.4 Experimental results

The experiments with DEM interpolation where performed using a low disparity synthetic data set kindly provided to us by the CNES - *Centre National d'Etudes Spatiales* (the reference image is shown in Figure 9.2(a)). They also provided us with a disparity map and the map of estimated disparity errors $Err(x)$ (Figures 9.2(b) and 9.2(c)) . Both where obtained using the correlation technique mentioned in Section 9.2. As can be seen in Figure 9.2(b), the initial disparity map contains sparse information, mostly because some points have errors beyond a certain threshold, thus, the known data must be interpolated. To test the performance of the different interpolation strategies we compare them against the provided ground truth (Figure 9.2(d)).

All the experiments where performed with the same value of $\beta = 0.025$, the time step dt was determined according to the criterion described in Section 9.3, and the parameter λ was adjusted using Uzawa's algorithm [BCRS03, ABCH08] in order to satisfy the constraints corresponding to each data term (Section 9.2). In order to compare the results obtained using local and global constraints we compute the global error σ from the error map $Err(x)$ (which is the standard deviation at each point). The global standard deviation is given by $\sigma = \sqrt{\frac{1}{L}\sum_{x\in\Omega} Err(x)^2} = 0.032$.

In Figure 9.2 we compare the results of isotropic and anisotropic regularization (Figures 9.2(f) and 9.2(e) resp.) of the data given in Figure 9.2(b). As it can be seen, the shape driven regularization produces a more coherent result. This is confirmed by the reconstruction displayed in Figures 9.2(h) and 9.2(i) and by the errors in Table 9.1 (computed with respect to to the ground truth). We have also verified that the data fitting constraints are satisfied by our algorithm and that the data fitting terms have a better local behavior when using D_i^{**} that when using D_i. Finally, let us observe that the use of a robust norm (L^1 in this case) effectively increases the precision, but this improvement is generally marginal compared to the one obtained with the use of the precision information (Err). Indeed the results obtained with L^1- and L^2-norm are virtually indistinguishable

except for the apparition of some outlier points in the L^2 case.

9.5 Conclusions

We have considered a variational model coupling an anisotropic regularization term with a data fitting term and its applications to the interpolation of sparse digital elevation models. We have minimized the energy via a gradient descent algorithm using an explicit numerical scheme, and we have analyzed its stability. We proposed the use of an anisotropic diffusion that respects the geometry of a reference image. In combination with robust data fitting terms, this permits to interpolate urban digital elevation models while preserving discontinuities. Still, we observe an excessive diffusion effect because the minimal surface does not allow as sharp discontinuities as the total variation does.

Chapter 10

Anisotropic Cheeger Sets and Applications

This chapter is dedicated to the study of Cheeger sets in the context of image processing. In particular, we consider Cheeger sets computed with respect to different non-uniform anisotropic norms relevant for applications to image processing. We describe the applications of these Cheeger sets to: the computation of global minima of geodesic active contours, edge linking problems and anisotropic diffusion.

10.1 Cheeger sets and Total Variation

Given a nonempty open bounded subset Ω of $\mathbb{R}^N$, we call Cheeger constant of Ω the quantity

$$\mathcal{C}_\Omega := \min_{F \subseteq \Omega} \frac{P(F)}{|F|}. \tag{10.1}$$

Here $|F|$ denotes the N-dimensional volume of F, and $P(F)$ symbolize the perimeter of F. The minimum in (10.1) is taken over all nonempty sets of finite perimeter contained in Ω. A Cheeger set of Ω is any set $G \subseteq \Omega$ which minimizes (10.1). Observe that G is a Cheeger set of Ω if and only if $|G| > 0$ and G minimizes

$$\min_{F \subseteq \Omega} P(F) - \mathcal{C}_\Omega |F|. \tag{10.2}$$

Existence of Cheeger sets follows directly from the direct methods of calculus of variations. Uniqueness of Cheeger sets is a more delicate issue and is not true in general (a counterexample is given in [KLR06] when Ω is not convex), though it has recently been proved that it is generically true [CCN] (that is, true modulo a small perturbation of the domain Ω). However, uniqueness of Cheeger sets inside convex bodies of $\mathbb{R}^N$ was proved in [CCN07] when the convex body is uniformly convex and of class C^2 and in [AC09] in the general case. The case of convex bodies of $\mathbb{R}^2$ was studied in [ACC05b, KLR06].

The computation of Cheeger sets has recently been the object of several papers [BCC07, CCP09]. One of the possible algorithms consists in solving the

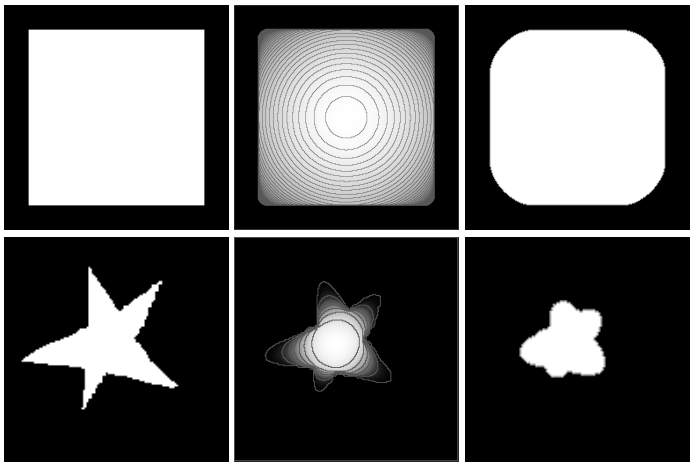


Figure 10.1: Euclidean Cheeger set examples for two domains (first column). For a better display the second column shows $\log(u)$ for the solution u of (10.3), with some level lines superimposed. The Cheeger set (right) is selected among the level sets of the second image as the one that minimizes the ratio (10.1).

variational problem

$$\min_{u\in BV(\Omega)\cap L^2(\Omega)} \left\{ \int_\Omega |Du| + \int_{\partial\Omega} |u|\, d\mathcal{H}^{N-1} + \frac{\lambda}{2}\int_\Omega (u-1)^2\, dx \right\}. \tag{10.3}$$

This strictly convex lower semicontinuous functional has a unique minimizer $u \in BV(\Omega)\cap L^2(\Omega)$ satisfying $0 \le u \le 1$. Moreover, for any $s \in (0,1]$ the level set $E_s := \{x \in \Omega : u(x) \ge s\}$ is a solution of

$$\min_{F\subseteq\Omega} P(F) - \mu|F|, \tag{10.4}$$

where $\mu := \lambda(1-s)$ and the infimum is taken over the sets $F \subseteq \Omega$ of finite perimeter in $\mathbb{R}^N$ [ACC05a, CCN09]. When taking $\lambda \in (0,+\infty)$ and $s \in (0,1]$ we are able to cover the whole range of $\mu \in [0,\infty)$ [ACC05a]. Since the family of level sets E_s is nested, the solution of (10.4) is unique for any $\mu \in (0,+\infty)$ up to a countable exceptional set. Moreover, when λ is big enough, the level set associated to the maximum of u, $\{x \in \Omega : u(x) = \|u\|_\infty\}$, is the maximal Cheeger set of Ω [ACC05a, CCN07]. Observe that this provides an algorithm for computing the maximal Cheeger set (and also the solution of the family of problems (10.4)). In particular, using Chambolle's algorithm [Cha04] to minimize (10.4), one passes to a dual variational problem which can be solved by a simple iterative scheme (see Figure 10.1).

Our purpose in this chapter is to study Cheeger sets in the context of image processing, in particular, their connections with active contours and edge linking. For that we use the theory of anisotropic perimeters developed in

[AB94, BBF00] to extend model (10.3) to general anisotropic perimeters. In the next section we present the mathematical background regarding general anisotropic perimeters. Then in Section 10.4 we will briefly describe the applications including as particular cases the geodesic active contour model with an inflating force [CKS95, CKS97, KKO+96], a model for anisotropic diffusion, and a model for edge linking. In Sections (10.5 and 10.6) discuss the applications in detail and present some experiments. Finally, in Section 10.7 we address the computational aspects related to the computation of anisotropic Cheeger sets.

10.2 Anisotropic Cheeger sets

Let us start defining the general notion of total variation with respect to a metric integrand.

Anisotropic metric integrand and total variation

Let Ω be an open subset of $\mathbb{R}^N$. Following [BBF00] we say that a function $\phi : \Omega \times \mathbb{R}^N \to [0,\infty)$ is a *metric integrand* if ϕ is a Borel function satisfying the following conditions:

$$\text{for a.e. } x \in \Omega, \text{ the map } \xi \in \mathbb{R}^N \to \phi(x,\xi) \text{ is convex,} \tag{10.5}$$

$$\phi(x,t\xi) = |t|\phi(x,\xi) \qquad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^N, \quad \forall t \in \mathbb{R}, \tag{10.6}$$

and there exists a constant $\Lambda > 0$ such that

$$0 \le \phi(x,\xi) \le \Lambda\|\xi\| \qquad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^N. \tag{10.7}$$

The polar function $\phi^0 : \Omega \times \mathbb{R}^N \to \mathbb{R}$ of ϕ is defined by

$$\phi^0(x,\xi^*) = \sup\{\langle \xi^*, \xi\rangle : \xi \in \mathbb{R}^N,\, \phi(x,\xi) \le 1\}. \tag{10.8}$$

Definition 12 ([CFM09, Definition 3.1]). *Let $u \in L^1(\Omega)$. We define* ***the ϕ-total variation*** *of u in Ω as*

$$\int_\Omega |Du|_\phi := \sup\left\{\int_\Omega u\, div\, \sigma\, dx : \sigma \in \mathcal{K}^\infty_\phi(\Omega)\right\}, \tag{10.9}$$

where for any $p \in [1,\infty]$, we define

$$\mathcal{K}^p_\phi(\Omega) := \{\sigma \in X_p(\Omega) : \phi^0(x,\sigma(x)) \le 1 \text{ for a.e. } x \in \Omega,\, [\sigma \cdot \nu^\Omega] = 0\},$$

with $X_p(\Omega) := \{z \in L^\infty(\Omega;\mathbb{R}^N) : div\, z \in L^p(\Omega)\}$.

We set $BV_\phi(\Omega) := \{u \in L^1(\Omega) : \int_\Omega |Du|_\phi < \infty\}$, which is a Banach space when endowed with the norm $|u|_{BV_\phi(\Omega)} := \int_\Omega |u| dx + \int_\Omega |Du|_\phi$.

Sets of finite ϕ-perimeter. We say that $E \subseteq \mathbb{R}^N$ has finite ϕ-perimeter in Ω if $\chi_E \in BV_\phi(\Omega)$. We set

$$P_\phi(E,\Omega) := \int_\Omega |D\chi_E|_\phi.$$

If $\Omega = \mathbb{R}^N$, we denote $P_\phi(E) := P_\phi(E,\mathbb{R}^N)$. By assumption (10.7), if $E \subseteq \mathbb{R}^N$ has finite perimeter in Ω, it also has finite ϕ-perimeter in Ω. The coarea formula for the ϕ-total variation was proved in [BBF00] (see also [AB94] in a slightly different formulation):

$$\int_\Omega |Du|_\phi = \int_{\mathbb{R}} P_\phi(\{u > s\},\Omega)\, ds \qquad \forall u \in BV_\phi(\Omega). \tag{10.10}$$

Anisotropic Cheeger set as solution of a variational problem

Similarly to the Euclidean case to compute anisotropic Cheeger sets we study the problem

$$\min_{u\in BV_\phi(\Omega)\cap L^2(\Omega,h\,dx)} \left\{ \int_\Omega |Du|_\phi + \frac{\lambda}{2}\int_\Omega (u-f)^2 h\,dx + \int_{\partial\Omega} \phi(x,\nu^\Omega)|u|\, d\mathcal{H}^{N-1} \right\}, \tag{10.11}$$

where $\phi : \Omega \times \mathbb{R}^N \to \mathbb{R}$ is a metric integrand in Ω which is symmetric in ξ, i.e., $\phi(x,-\xi) = \phi(x,\xi)$ for any $\xi \in \mathbb{R}^N$ and any $x \in \Omega$; $h \in L^\infty(\Omega)$, $h(x) > 0$ a.e., with $\int_\Omega \frac{1}{h(x)}\,dx < \infty$; $f \in L^2(\Omega, hdx)$, $\lambda > 0$; and ν^Ω denotes the outer unit normal to Ω at points of $\partial\Omega$.

We denote by $L^2(\Omega, hdx)$ the set of measurable functions $u : \Omega \to \mathbb{R}$ such that $\int_\Omega u^2\, h\, dx < \infty$.

Although for ϕ-Cheeger sets we need only the case $f = 1$, the general case where $f \neq 1$ is of interest in Section 10.6, where we discuss the application to anisotropic diffusion. To shorten the expressions inside the integrals we shall write h, u instead of $h(x), u(x)$, with the only exception being $\phi(x,\nu^\Omega)$.

In order to compute ϕ-Cheeger sets using the energy formulation (10.11) we have to assume that the metric integrand $\phi(x,\xi)$ is continuous and coercive near the boundary of Ω, which amounts to saying that for $x \in \Omega$ near $\partial\Omega$ we have

$$\alpha|\xi| \le \phi(x,\xi) \le \beta|\xi| \qquad \forall \xi \in \mathbb{R}^N,\ 0 < \alpha < \beta. \tag{10.12}$$

We note that in case that ϕ is coercive everywhere we can minimize (10.11) in the space of functions of bounded variation $BV(\Omega)$. In the present situation, where we do not assume that ϕ is coercive, we need the anisotropic extension of $BV(\Omega)$ given in Definition 12, namely, the space $BV_\phi(\Omega)$.

Notice that the use of degenerate (noncoercive) metric integrands ϕ comes from the applications, where it can be natural to assume that ϕ vanishes on a subset of Ω (e.g., for the geodesic active contour poblem it may vanish on arcs of curve if $N = 2$ or surface patches if $N = 3$).

The PDE associated to the variational problem (10.11)**.** In order to minimize (10.11) we solve its associated Lagrange equation which can be written as

$$hf \in hu - \lambda^{-1}\mathrm{div}\left(\partial_{\xi}\phi(x, \nabla u)\right) \tag{10.13}$$

with Dirichlet boundary conditions. We use the notion of solution defined in [CFM09, Definition 6.1]. As proved in [CFM09, Theorem 6.2], if $f \in L^{\infty}(\Omega)$, then there exists a unique solution of (10.13) which is also the unique solution of (10.11).

10.3 The maximal ϕ-Cheeger set inside Ω.

If u is a solution of (10.11), then the level sets $\{x \in \Omega : u(x) \geq s\}$, $s \in (0, 1]$, are solutions of

$$\min_{F \subseteq \Omega} P_{\phi}(F) - \mu |F|_h, \tag{10.14}$$

where $\mu := \lambda(1 - s)$, recall the $P_{\phi}(F)$ is the anisotropic perimeter of F, and $|F|_h = \int_F h(x)\, dx$. Let us state this result [CFM09, Proposition 6.3]:

Proposition 13 ([CFM09, Proposition 6.3]). *Let $u \in BV_{\phi}(\Omega) \cap L^2(\Omega, h\, dx)$ be the solution of the variational problem* (10.11) *with $f = 1$. Then $0 \leq u \leq 1$. Let $E_s := \{u \geq s\}$, $s \in (0, 1]$. Then for any $s \in (0, 1]$ we have*

$$P_{\phi}(E_s) - \lambda(1 - s)|E_s|_h \leq P_{\phi}(F) - \lambda(1 - s)|F|_h \tag{10.15}$$

for any $F \subseteq \Omega$.

As in the Euclidean case, the solution of (10.14) is unique for any $s \in (0, 1]$ up to a countable exceptional set. Moreover, when λ is big enough, the level set associated to the maximum of u, $\{x \in \Omega : u(x) = \|u\|_{\infty}\}$, is the maximal (ϕ, h)-Cheeger set of Ω. A (ϕ, h)-Cheeger set in Ω is a minimizer of the problem

$$\inf\left\{\frac{P_{\phi}(F)}{|F|_h} : F \subseteq \overline{\Omega} \text{ of finite perimeter}, |F|_h > 0\right\}, \tag{10.16}$$

and the value of this infimum is the (ϕ, h)-Cheeger constant, denoted by $\mathcal{C}_{\Omega}^{\phi,h}$. The computation of the maximal (ϕ, h)-Cheeger set (together with the solution of the family of problems (10.14)) can be computed using Chambolle's algorithm [Cha04]. To simplify our expressions, instead of (ϕ, h)-Cheeger set and constant, we will say ϕ-Cheeger set and constant.

Local ϕ-Cheeger sets in Ω

In the applications below we will need an extension of the notion of Cheeger set. Suppose that we want to detect several objects in an image and we have a model that enables us to identify them as Cheeger sets. The Cheeger set will minimize the perimeter/area in the whole image domain. This gives a first object in the

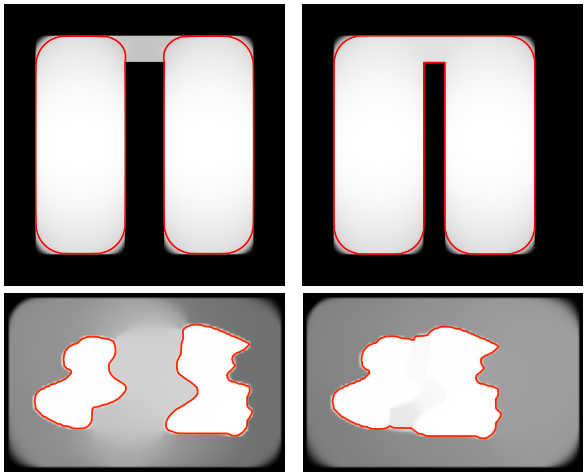


Figure 10.2: Local ϕ-Cheeger sets (red curves). In the first row are shown local ϕ-Cheeger sets in the Euclidean case. The second row shows local ϕ-Cheeger sets obtained on a domain that occupies the whole image but using a metric integrand $\phi(x,\cdot) \sim \frac{1}{1+|\nabla I(x)|}$, where I is an image containing two shapes (not shown). In all the figures the backgrounds show the corresponding solution of (10.11). For both: the Euclidean and the non-uniform metric cases, the second column show an experiment where, with a very similar setup, the local ϕ-Cheeger sets collapsed into a single set. In the Euclidean case the shape of the domain is changed, and in the non constant metric case the shapes in I have moved.

image. By extracting it, we may compute the Cheeger set inside the remaining domain and this will provide us with a second object. This requires to define the notion of local Cheeger set. For that we notice that Cheeger sets are computed as the maximum region of the solution u of (10.11) which we expect to be a flat region. By analyzing the level sets of u we discover that there may be several local maxima (see Figure 10.2) which are flat regions and correspond to Cheeger sets in its neighborhood, that is, they are local Cheeger sets.

To develop the notion of local ϕ-Cheeger set, we assume that ϕ is continuous and coercive in Ω. Let $E \subseteq \mathbb{R}^N$ be a set of finite perimeter. We say that E is decomposable if there exists a partition (A, B) of E such that $P_\phi(E) = P_\phi(A) + P_\phi(B)$ and both $|A|$ and $|B|$ are strictly positive. We say that E is indecomposable if it is not decomposable; notice that the properties of being decomposable or indecomposable are invariant modulo Lebesgue null sets and that, according to our definition, any Lebesgue negligible set is indecomposable.

The following result was proved in [ACMM01] for the Euclidean perimeter. The proof easily extends to cover the case where ϕ is continuous and coercive in Ω, but it also follows from the Euclidean case since the assumptions on ϕ imply

that

$$P_\phi(E) = \int_{\partial^* E} \phi(x, \nu^E(x))\, d\mathcal{H}^{N-1}$$

for any set $E \subseteq \mathbb{R}^N$ with finite perimeter.

Theorem 14 ([CFM09, Theorem 6.10]). *Let E be a set with finite perimeter in $\mathbb{R}^N$. Then there exists a unique finite or countable family of pairwise disjoint indecomposable sets $\{E_i\}_{i\in I}$ such that $|E_i| > 0$ and $P_\phi(E) = \sum_i P_\phi(E_i)$. Moreover, the sets E_i are maximal indecomposable sets; i.e., any indecomposable set $F \subseteq E$ is contained modulo a Lebesgue null set in some set E_i.*

In view of the previous theorem, we call the sets E_i the ϕ-connected components of E.

Proposition 15 ([CFM09, Proposition 6.11]). *Assume that ϕ is continuous and coercive in Ω. Let $u \in BV_\phi(\Omega) \cap L^2(\Omega, h\,dx)$ be the solution of* (10.11)*. Let $t \in (0, 1]$ and $E_t := \{u \geq t\}$. Let E'_t be a ϕ-connected component of E_t, and let $F_s = \{u \geq s\} \cap E'_t$, $s \geq t$. Then for any $s \in (0, 1]$ we have*

$$P_\phi(F_s) - \lambda(1-s)|F_s|_h \leq P_\phi(F) - \lambda(1-s)|F|_h \tag{10.17}$$

for any $F \subseteq E'_t$. If $s = \max_{x\in E'_t} u(x)$, then F_s is a maximal ϕ-Cheeger set in E'_t.

The sets F_s will be called local ϕ-Cheeger sets.

Recall that, when ϕ is coercive, by the isoperimetric inequality there is a constant $\alpha > 0$ (depending on the domain) such that any ϕ-Cheeger set has measure $\geq \alpha$. Moreover, the union and intersection of ϕ-Cheeger sets are ϕ-Cheeger [CCN]. In particular, there are minimal ϕ-Cheeger sets and there are finitely many of them [CCN].

10.4 Applications overview

We illustrate the previous formalism with three examples: (a) the geodesic active contour model; (b) a model for edge linking; and (c) an anisotropic diffusion model. In this section we will briefly discuss these examples, and in Section 10.5 and 10.6 we will develop them.

(a) The geodesic active contour model. Let $I : \Omega \to \mathbb{R}^+$ be a given image in $L^\infty(\Omega)$, G be a Gaussian function, and

$$g(x) = \frac{1}{\sqrt{1 + |\nabla(G * I)|^2}}, \tag{10.18}$$

where in $G * I$ we have extended I to $\mathbb{R}^N$ by taking the value 0 outside Ω. Observe that $g \in C(\overline{\Omega})$ and $\inf_{x\in\overline{\Omega}} g(x) > 0$. The geodesic active contour model [CKS95, CKS97, KKO+96] with an inflating force (see [Coh91]) corresponds to (10.14) when we take $\phi(x, \xi) = g(x)|\xi|$, $|Du|_\phi = g(x)|Du|$, and $h(x) = 1$, $x \in \Omega$.

The purpose of this model is to locate the boundary of an object of the image at the points where the gradient is large. The presence of the inflating term helps to avoid minima collapsing into a point. The analysis of this model was done in [CCCD93, CKSS97, KKO$^+$96] using the level set formulation of (10.14). In this case we write $P_g(F)$ instead of $P_\phi(F)$, and we have $P_g(F) := \int_{\partial^* F} g\, d\mathcal{H}^{N-1}$, where $\partial^* F$ is the reduced boundary of F [AFP00].

In this case the ϕ-Cheeger sets are a particular instance of geodesic active contours with a constant inflating force $\mu = \mathcal{C}^{g,1}_\Omega$. An interesting feature of this formalism is that, assuming that ϕ is continuous and coercive, it permits us to apply the result of Proposition 15, and obtain local ϕ-Cheeger sets as local (regional) maxima of the function u. They are ϕ-Cheeger sets in a subdomain of Ω. They can be identified with boundaries of the image, and the above formalism permits us to compute several active contours at the same time (the same holds true for the edge linking model).

A more general active contour model, based on Finsler metrics, was introduced in [MPAT08]. In this paper, the authors minimized the Finsler metric using dynamic programming. A different numerical approach based on graph cuts, and valid for submodular Finsler metrics, was proposed in [KB05].

(b) An edge linking model. Another interesting application of the above formalism is to edge linking. Given a set $\Gamma \subseteq \Omega$ (which may be the output of an edge detector formed by arcs of curve if $\Omega \subseteq \mathbb{R}^2$ or surface patches if $\Omega \subseteq \mathbb{R}^3$), we define $d_\Gamma(x) = \operatorname{dist}(x, \Gamma)$ and the metric integrand $\phi(x, \xi) = d_\Gamma(x)|\xi|$. In this case, we experimentally see that the ϕ-Cheeger set determined by this weighted metric has a boundary formed by a set of curves ($N = 2$) or surfaces ($N = 3$) linking Γ.

Let us mention the formulation of active contour models without edges proposed in [CV01] by Chan and Vese, whose solution can be related to the general formulation (10.14). Let $I : \Omega \to \mathbb{R}^+$ be a given image and $g \in C(\overline{\Omega})$ be such that $\inf_{x\in\overline{\Omega}} g(x) > 0$. The authors proposed minimizing

$$\min_{F\subseteq\Omega, c_1, c_2\in\mathbb{R}} \mathcal{E}_g(F, c_1, c_2) := P_g(F;\Omega) + \lambda \int_F (I(x) - c_1)^2\, dx + \lambda \int_{\Omega\setminus F} (I(x) - c_2)^2\, dx, \tag{10.19}$$

where the minimum is taken over the sets F of finite perimeter in Ω, with $c_1, c_2 \in \mathbb{R}$, $\lambda > 0$, and $P_g(F, \Omega) := \int_{\partial^* F\cap\Omega} g\, d\mathcal{H}^{N-1}$ is the weighted perimeter of F in Ω. Although in the initial proposal [CV01] the authors took $g = 1$ (in this case (10.19) is the restriction of the Mumford–Shah functional to a binary segmentation of the image), the extension to a weighted perimeter was natural and has been considered, for instance, in [BEV$^+$07]. If the set F is fixed, then the minimum of $\mathcal{E}_g(F, c_1, c_2)$ with respect to $c_1, c_2 \in \mathbb{R}$ gives us the values $\bar{c}_1 = \frac{\int_F I(x)\,dx}{|F|}$ and $\bar{c}_2 = \frac{\int_{\Omega\setminus F} I(x)\,dx}{|\Omega\setminus F|}$, and we may write

$$\min_{F\subseteq\Omega, c_1, c_2\in\mathbb{R}} \mathcal{E}_g(F, c_1, c_2) = \min_{F\subseteq\Omega} \mathcal{E}_g(F, \bar{c}_1, \bar{c}_2). \tag{10.20}$$

Observe that

$$\mathcal{E}_g(F,\bar{c}_1,\bar{c}_2) = P_g(F,\Omega) + \lambda \int_F ((I(x)-c_1)^2 - (I(x)-c_2)^2)\,dx + \\ + \lambda \int_\Omega (I(x)-c_2)^2\,dx$$

and that it suffices to minimize with respect to F the first two terms of this sum, the last one being a constant. Now, as proposed in [NEC06, BEV$^+$07], for any fixed values $c_1, c_2 \in \mathbb{R}$ the global minimizer of $\mathcal{E}_g(F,\bar{c}_1,\bar{c}_2)$ with respect to F can be found by solving the convex optimization problem [NEC06, Theorem 2]

$$\min_{0\le u\le 1} \int_\Omega g|Du| + \lambda \int_\Omega ((I(x)-c_1)^2 - (I(x)-c_2)^2)\,u\,dx \tag{10.21}$$

and then setting $F := \{u > t\}$ for a.e. $t \in (0,1)$. Thus, iterating between the solution of (10.21) with fixed values of c_1, c_2 and the updating of c_1, c_2 just described, we have a two step algorithm to solve the Chan–Vese model [NEC06, BEV$^+$07].

Let us also mention the interesting work [PSGB08] in which the authors propose a method for convexifying the total variation regularization of some nonlinear and nonconvex data attachment terms (e.g. the computation of disparity in rectified stereo pairs) which leads to the solution of an anisotropic total variation problem with Dirichlet-type boundary conditions.

(c) An anisotropic diffusion model. The model (10.11) contains the case $\phi(x,\xi) = |A_x\xi|$, where A_x is a symmetric positive definite matrix for each $x \in \Omega$. A particular instance when $N = 2$ is the anisotropic diffusion model given by $A_x = z^\perp(x) \otimes z^\perp(x) + g(x)z(x) \otimes z(x)$, where $z(x) = \frac{\nabla I(x)}{\sqrt{1+|\nabla I(x)|^2}}$ and $z(x)^\perp$ denotes the counterclockwise rotation of $z(x)$ of angle $\frac{\pi}{2}$. Notice that by the structure of A_x we could also take the clockwise rotation. Notice that this model uses Neumann boundary conditions. As in the Dirichlet case there are existence and uniqueness results for this problem [CFM09, Theorem 7.2].

10.5 Minimization of geodesic active contour model with an inflating force, and edge linking

Given an image $I : \Omega \to \mathbb{R}$, let us consider the following formulation of geodesic active contour with an inflating force:

$$\min_{E\subseteq\Omega} P_g(E) - \mu|E|_h, \qquad \mu > 0, \tag{10.22}$$

where $P_g(E)$ is a weighted perimeter with weights $g(x) = (1+|\nabla(G*I)|^2)^{-\frac{1}{2}}$, $|E|_h = \int_E h(x)\,dx$ is the weighted area, and μ is a parameter that controls the balloon force. In Proposition 13 we have shown that if u is the solution of problem (10.11), then $E_s := \{u \ge s\}, s \in (0,1]$, is a global minimum of (10.22) with

$\mu = \lambda(1-s)$. In this case ϕ is of the form $\phi(x,\xi) = g(x)|\xi|$ and we will use the expression g-Cheeger set. In particular, if $\lambda > \mathcal{C}_\Omega$ (the g-Cheeger constant), then the set $\{u = \|u\|_\infty\}$ is the g-Cheeger set of Ω. Therefore to compute the solutions of (10.22) it suffices to solve problem (10.11) for some $\lambda > \mu$ or, equivalently, to find u by solving

$$hu - \lambda^{-1}\mathrm{div}\left(g\frac{Du}{|Du|}\right) = h. \tag{10.23}$$

The solution of (10.23) is computed using the scheme proposed by Chambolle [Cha04] described in Section 10.7.1 (with $f = \chi_\Omega$). For λ big enough, for all values of $\mu > 0$, the solutions of (10.22) can be found as the level sets of u. In practice we select the g-Cheeger set as the upper level set of u that minimizes

$$\min_{\Gamma\subseteq\Omega} \frac{P_g(\Gamma)}{|\Gamma|_h},$$

where the minimum is taken over the upper level sets of u. The details about the computation of the weighted perimeters of the level sets and areas are given in Section 10.7.2.

Experiments

We have used the theory described above in two different ways, corresponding to different choices of the metric integrand g. The first choice is $g(x) = (\sqrt{1+|\nabla(G*I)|^2})^{-1}$, and the second choice is the distance function to the set of edge points detected by a preprocessing of the image, that is, $g = d_S$, where S is the set of edges of u. We label these two cases $\frac{1}{|\nabla I|}$-*Cheeger sets* and d_S-*Cheeger sets*, respectively. We observe that the convergence of the iterative scheme to solve the PDE is much faster for d_S-Cheeger sets, and the result is less likely to miss parts of the image. On the other hand, the computation of $\frac{1}{|\nabla I|}$-Cheeger sets gives smooth results after a long time and sometimes misses parts of the desired objects or fails to break at holes. The choice of a subdomain $B \subseteq \Omega$ allows for some flexibility: we can enforce hard restrictions on the result by removing from the domain some points that we do not want to be enclosed by the output surfaces.

Notice that, for a given choice of g, we actually find many *local* g-Cheeger sets, disjoint from the global minimum, that appear as local minima of the g-Cheeger ratio on the tree of connected components of upper level sets. The computation of those sets is justified by Proposition 15 (see also [CFM09, Proposition 6.11]), however the assumptions for this result not cover the case where g vanishes (ϕ must be coercive). These are the sets which we show in the following experiments.

2D images. In Figure 10.3, we display some local g-Cheeger sets of 2D images for different choices of metric g. These experiments are equivalent to applying the model (10.22) to edge linking problems. As in [Coh91] the inflating force allows us to link the pieces of the boundaries of the objects. We display

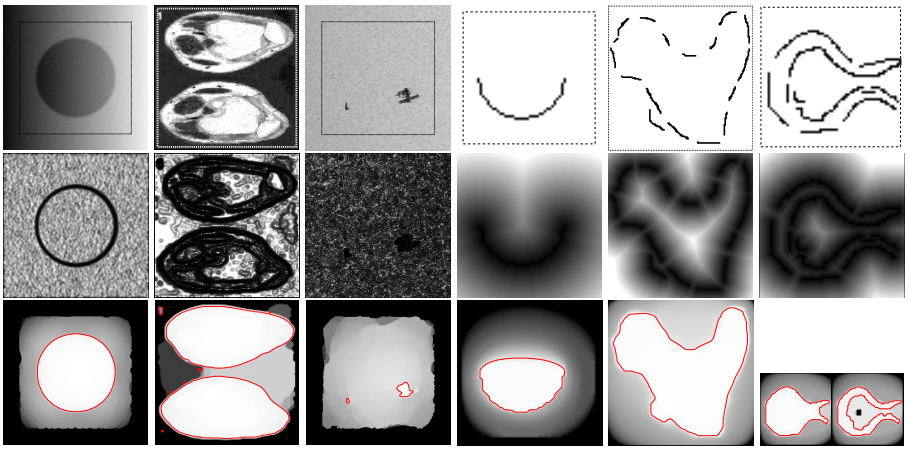


Figure 10.3: Geodesic active contours as g-Cheeger minimizers. The first row shows the images I to be processed. The second row shows the weights g used for each experiment (white is 1, black is 0): in the first two cases $g = (\sqrt{1+|\nabla(G*I)|^2})^{-1}$, for the third $g = 0.37(\sqrt{0.1+|\nabla(G*I)|^2})^{-1}$, and for the linking experiments $g = d_S$, the scaled distance function to the given edges. The third row shows the disjoint minimal local g-Cheeger sets extracted from u (shown in the background); there are $1, 7, 2, 1, 1$, and 1 sets, respectively. The last linking experiment illustrates the effect of introducing a barrier in the initial domain (black square).

in Figure 10.3 some 2D linking experiments, which show how the d_S-Cheeger set indeed links the edges. Let us point out here a limitation of this approach, which can be observed in the last subfigure. Even if this linking is produced, the presence of a bottleneck (bottom right subfigure) causes the d_S-Cheeger set to be a set with large volume. This limitation can be circumvented by adding barriers in the domain Ω.

*S*ynthetic 3D image. The first 3D example is a synthetic image built in the following way. We have taken the characteristic function of a slanted torus plus a linear function and then added some blurring and Gaussian noise to the result. Some slices and a level surface of this image are shown in the left subfigure of Figure 10.4. The first experiment with this synthetic image has been to segment it using the $\frac{1}{|\nabla I|}$-Cheeger set of the image domain. This gives a reasonable segmentation of the object, as shown in Figure 10.4. The second experiment with this synthetic image has been to perform edge linking. We have taken the output of an edge detector [DMM01, MZFC08] and used the distance function to the set of edges as a metric. The d_S-Cheeger set of the image domain is a surface that correctly interpolates the given patches. We can observe that the result of the edge linking has a ragged appearance. In Figure 10.5 we display the input edges, the corresponding metric, and the final result. In Figure 10.6 we display the graph of the $\frac{1}{|\nabla I|}$-Cheeger ratio and different level sets of u.

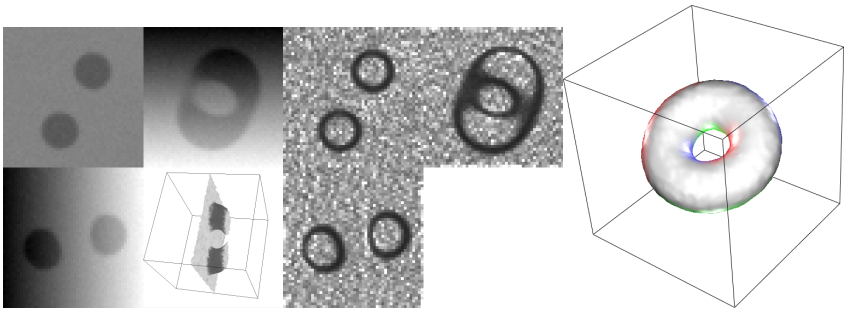


Figure 10.4: Pipeline for computing $\frac{1}{|\nabla I|}$-Cheeger sets, applied to a synthetic 3D image. From left to right: slices of the original image I, slices of the metric $g = \frac{1}{\|\nabla I\|}$, and $\frac{1}{|\nabla I|}$-Cheeger set of the image domain.

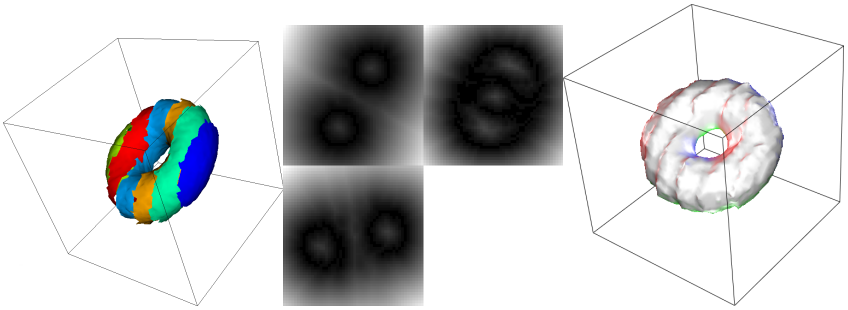


Figure 10.5: Pipeline for computing d_S-Cheeger sets, applied to the same synthetic image as in Figure 10.4. From left to right: detected 3D edges S, slices of the metric $g = d_S$, and d_S-Cheeger set of the image domain.

Real 3D computed tomography (CT) image. The first real 3D example is based on a CT of cerebral arteries containing an aneurysm. We have tried both $\frac{1}{|\nabla I|}$ and d_S metrics (where S is computed, as before, by an edge detector). The results are visually similar. Noticing that both methods give an incorrect segmentation on a small part of the image (at the neck of the aneurysm), we have forced a correct segmentation by manually marking some voxels, as in the rightmost column in Figure 10.3. Thus, instead of computing the ϕ-Cheeger set of the image domain, we have computed the ϕ-Cheeger set of the image domain minus some manually selected voxels. In Figure 10.7 we display the results, and in Figure 10.8, we display three different level surfaces of the solution u (the central one being the $\frac{1}{|\nabla I|}$-Cheeger set).

Real 3D magnetic resonance (MR) image. The second real 3D example is an edge linking experiment coming from an MR image. This is a very low-

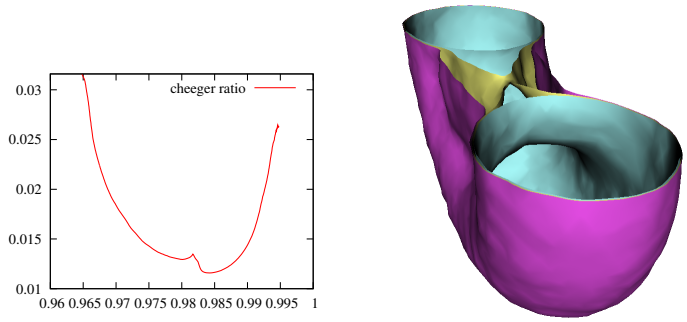


Figure 10.6: Left: graph of the $\frac{1}{|\nabla I|}$-Cheeger ratio $F(t)$ for the input image "torus." Right: superposition of the three level sets shown, corresponding to the interesting points of $F(t)$ (the two minima and the cusp). To see the inner surfaces, the display is clipped near the central singularity. Notice that these level surfaces are all local minima of the classical geodesic snakes functional with an inflating force, for different weights of the inflating force.

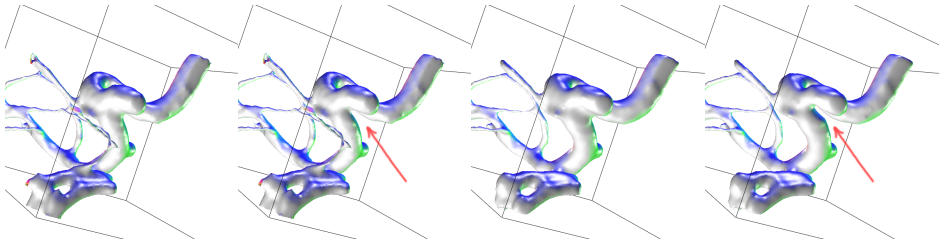


Figure 10.7: Computation of $\frac{1}{|\nabla I|}$-Cheeger sets of the CT image. From left to right: (1) d_S-Cheeger set of the whole image domain, (2) d_S-Cheeger set of the image domain minus some manually selected voxels at the neck of the aneurysm, (3) $\frac{1}{|\nabla I|}$-Cheeger set of the whole image domain, and (4) $\frac{1}{|\nabla I|}$-Cheeger set of the image domain minus some manually selected voxels at the neck of the aneurysm.

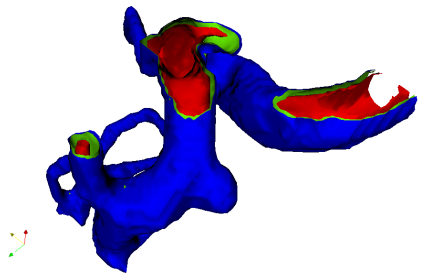


Figure 10.8: Three level surfaces of the solution u of $\frac{1}{|\nabla I|}$-total variation minimization. The central surface is the $\frac{1}{|\nabla I|}$-Cheeger set of the image domain according to this active–contours-like metric. The other two surfaces have a higher $\frac{1}{|\nabla I|}$-Cheeger ratio and appear as local extrema of an active contour with appropriate inflating force. In this figure, the image domain is split so that the innermost surfaces can be seen. Notice that the inner surface, having a higher level t, is separated from the other two. This indicates the concentration of values around the maximum 1.

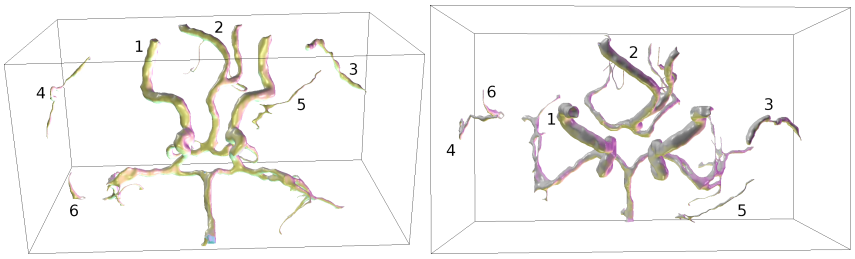


Figure 10.9: These two figures display the best six local d_S-Cheeger sets of the MR image, labelled and from different points of view.

resolution image, where the thin vessels have a width of one voxel. An edge detector correctly finds most of the vessels (in several different connected components). We show the best six local d_S-Cheeger sets of this image in Figure 10.9.

10.6 Anisotropic diffusion

Consider the anisotropic diffusion problem formulated as

$$\min_{u \in X} \frac{\|\pi_Z u - f\|^2}{2} + \lambda^{-1} J_\phi(u), \tag{10.24}$$

where $X = L^2(\Omega)$ and π_Z is the orthogonal projection onto a convex set $Z \subset X$, and $f \in Z$. The regularizer $J_\phi(u) = \int_\Omega \phi(x, \nabla u(x))$ is defined so that the diffusion is constrained to the geometry (given by the level lines) extracted from a reference image I. We define, for example, $\phi(x, \xi) = |A_x \xi|$, where A_x is a matrix that embodies knowledge about the boundaries of the objects in I. A common example in two dimensions corresponds to $A_x = z(x)^\perp \otimes z(x)^\perp$ with $z(x) = \frac{\nabla I(x)}{\sqrt{1+|\nabla I(x)|^2}}$. This example favors the diffusion along the level lines of I. In low gradient (flat) zones the previous definition can be relaxed to allow diffusion across the level lines (as depicted in Figure 10.10) in a way inversely proportional to the modulus of the gradient. In that case, we may take $A_x = z(x)^\perp \otimes z(x)^\perp + \frac{1}{\sqrt{1+|\nabla I(x)|^2}} z(x) \otimes z(x)$, where $z(x)^\perp$ denotes the counterclockwise rotation of $z(x)$ of angle $\frac{\pi}{2}$. Notice that by the structure of A_x we could also take the clockwise rotation.

We will solve (10.24) by adapting the zoom algorithm proposed in [Cha04]. Observing that $\|\pi_Z u - f\| = \min_{w \in Z^\perp} \|u - (f + w)\|$, (10.24) can then be reformulated as

$$\min_{u \in X, w \in Z^\perp} \frac{\|u - f - w\|^2}{2} + \lambda^{-1} J_\phi(u), \tag{10.25}$$

which is solved by alternate minimization with respect to u and w. The first minimization is done by the algorithm described in Section 10.7.1, $u_n = (f - w_n) - \pi_{\mathcal{K}_\phi}(f - w_n)$, and the second one consists in a projection over $Z^\perp$: $w_{n+1} = \pi_{Z^\perp}(u_n - f)$.

Experiments

The scheme presented above for solving (10.25) can be applied in a variety of diffusion problems, such as image colorization [LLW04], or to the interpolation of sparse height data in a digital elevation model [FLA+06]. In each of these cases, however, there are better algorithms for performing the task than the one we propose here, which is meant only as an illustration.

In the case of colorization and interpolation, Z is defined as $Z = \{\chi_\Gamma f : f \in X\}$, where $\Gamma \subseteq \Omega$ is a subdomain of the image where the values are known, and the reference image $I : \Omega \to \mathbb{R}$ is used to compute the field $z(x)$ to guide the

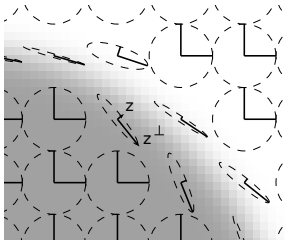


Figure 10.10: Anisotropic diffusion directions for a synthetic image (background). The plotted ellipses correspond to the tensor field $A_x = z(x)^\perp \otimes z(x)^\perp + \frac{1}{\sqrt{1+|\nabla I(x)|^2}} z(x) \otimes z(x)$ with $z(x) = \frac{\nabla I(x)}{\sqrt{1+|\nabla I(x)|^2}}$. The orientation and widths of the ellipses reflect the preferred direction of the diffusion.

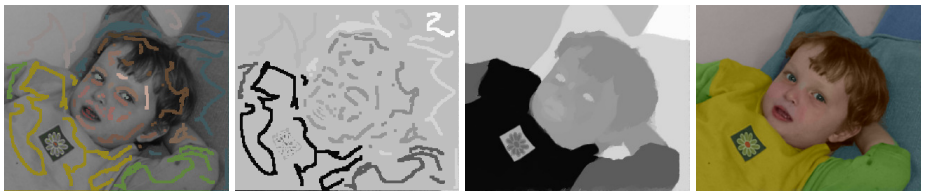


Figure 10.11: Colorization. From left to right, scribbled image, the known values of the U channel (20% of the image), diffusion in the U channel, and colorized result using the reference Y channel (not shown).

diffusion of these values. For the colorization experiment shown in Figure 10.11 the result is computed in YUV color space, where Y is the input luminance channel and the chromatic channels U and V are interpolated with (10.24), where the field $z(x) = \frac{\nabla I(x)}{\sqrt{1+|\nabla I(x)|^2}}$ restricts the diffusion to the geometry of I.

The last example concerns the interpolation of digital elevation models (see Figure 10.12). In this case the datum f is known only at sparse locations, and it is provided by a stereo subpixel correlation algorithm [SAM08] (which also provides an estimation of the measure's variance Err). The reference image of the stereo pair is used as a geometric constraint for the interpolation, and the variability Err is used to normalize the data fitting by adapting the spatial metric $h(x)^{1/2} = 1/Err(x)$. In Figure 10.12 we compare this method with the anisotropic minimal surface interpolation described in [FLA+06].

10.7 Numerical Computation of the ϕ-Cheeger set

The computation of the ϕ-Cheeger set requires first to compute the solution u of the PDE (10.13), and then to find the upper level set Γ of u that minimizes the ratio $P_\phi(\Gamma)/|\Gamma|$. In this section we will address these two issues.

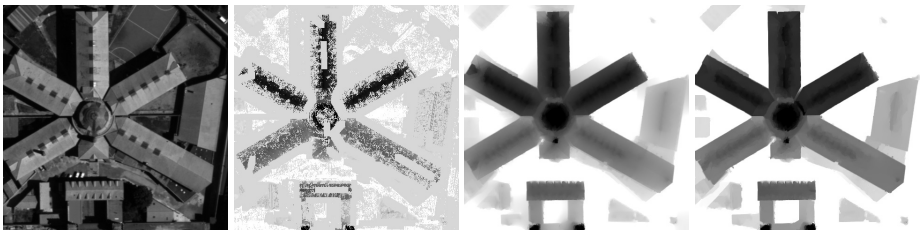


Figure 10.12: Disparity interpolation in an urban digital elevation model. From left to right: (1) the reference image of the stereo pair, (2) the incomplete data set computed with [SAM08] (30% of the image), where each point's gray level represents the height (darker is higher, mid-gray is unknown), (3) the interpolation obtained with the minimal surface interpolation [FLA$^+$06] (RMSE 0.239 when compared with the ground truth) and with the proposed algorithm (RMSE 0.190), (4) the minimal surface model [FLA$^+$06] recovers the slanted surfaces better than total variation; however, the latter is better at approximating the geometry near jumps.

10.7.1 Numerical solution of the PDE

In this subsection we present an adaptation of Chambolle's algorithm [Cha04] that permits us to solve a discrete version of (10.13) for some particular instances of $\phi(x,\xi)$. Our development will be restricted to the two-dimensional (2D) case, but it can be easily extended to higher dimensions. Let us give some notation that we use in what follows, keeping in mind that, for simplicity, we will denote the discrete functions we use like their continuous counterparts.

Let us consider the discrete domain $\Omega = \{0,1,\dots,N-1\}^2$ (more generally, we could assume that $\Omega \subseteq \{0,1,\dots,N-1\}^2$). For convenience, let us denote by Ω^e the extended domain $\{-1,0,\dots,N\}^2$. We denote by U the Euclidean space $\mathbb{R}^{(N+2)\times(N+2)}$. Let us give the definition of the discrete gradient which is adapted to problem (10.13) (which considers Dirichlet boundary conditions). In Section 10.7.1 we shall use Neumann boundary conditions with the definition of the gradient and divergence taken as in [Cha04]. Given $u \in U$, its discrete gradient ∇u will be a vector in $V := U \times U$ given by $\nabla u(i,j) = ((\nabla u)^1(i,j),(\nabla u)^2(i,j)),\ (i,j) \in \Omega^e$, where

$$(\nabla u)^1(i,j) = \begin{cases} u(i+1,j) - u(i,j) & \text{if } (i+1,j),(i,j) \in \Omega, \\ -u(i,j) & \text{if } (i+1,j) \notin \Omega,\ (i,j) \in \Omega, \\ u(i+1,j) & \text{if } (i+1,j) \in \Omega,\ (i,j) \notin \Omega, \\ 0 & \text{if } (i+1,j),(i,j) \notin \Omega, \end{cases} \tag{10.26}$$

$$(\nabla u)^2(i,j) = \begin{cases} u(i,j+1) - u(i,j) & \text{if } (i,j+1),(i,j) \in \Omega, \\ -u(i,j) & \text{if } (i,j+1) \notin \Omega,\ (i,j) \in \Omega, \\ u(i,j+1) & \text{if } (i,j+1) \in \Omega,\ (i,j) \notin \Omega, \\ 0 & \text{if } (i,j+1),(i,j) \notin \Omega. \end{cases} \tag{10.27}$$

The above case amounts to saying that $u(i,j) = 0$ when the indexes are in $\Omega^e \setminus \Omega$. These definitions of gradient embody the Dirichlet boundary conditions. The extension of the anisotropy ϕ to Ω^e will be made precise in the examples below. We will avoid the use of the indices (i,j) (unless necessary) and use x instead. The scalar product and the norm in U are defined as usual and denoted by $\langle \cdot, \cdot \rangle_U$ and $\| \cdot \|_U$, but in the absence of ambiguities the subindex will be omitted. In V the scalar product is denoted $\langle p, q \rangle_V = \sum_{x \in \Omega} p(x)^T \, q(x)$ and the norm $\|p\|_V = \langle p, p \rangle_V$. Finally, the divergence is defined so that it verifies $\langle p, \nabla u \rangle_V = -\langle \operatorname{div} p, u \rangle_U$:

$$(\operatorname{div} p)(i,j) = \begin{cases} p^1(i,j) - p^1(i-1,j) & \text{if } (i,j) \in \Omega, \\ 0 & \text{if } (i,j) \notin \Omega \end{cases} \tag{10.28}$$

$$+ \begin{cases} p^2(i,j) - p^2(i,j-1) & \text{if } (i,j) \in \Omega, \\ 0 & \text{if } (i,j) \notin \Omega. \end{cases} \tag{10.29}$$

Algorithm for the geodesic active contour type models

Let us consider the following generalization of the problem studied by Chambolle in [Cha04]:

$$\min_{u \in U,\, u=0 \text{ in } \Omega^e \setminus \Omega} \frac{\|(u-f)h^{1/2}\|_U^2}{2} + \lambda^{-1} J_g(u), \quad \text{where } \; J_g(u) = \sum_{x \in \Omega^e} g(x) |\nabla u(x)|, \tag{10.30}$$

$f, g, h \in U$ and $h(x) > 0$ for all $x \in \Omega^e$. We consider these functions defined originally in Ω and extended to Ω^e by specular symmetry. Observe that the Euler equation of (10.30) is a discretization of (10.13), where $\phi(x, \nabla u) = g(x)|\nabla u|$ (discrete case), and that where we write $\partial_\xi \phi(x, \nabla u) = g \frac{\nabla u}{|\nabla u|}$,

$$hu - \lambda^{-1} \operatorname{div} \left(g \frac{\nabla u}{|\nabla u|} \right) \ni hf. \tag{10.31}$$

As in [Cha04] let us derive the dual formulation for (10.30) by rewriting (10.31) as $h\lambda(f-u) \in \partial J_g(u)$, which is equivalent to $u \in \partial J_g^* \left(h\lambda(f-u) \right)$, where J_g^* is the Legendre–Fenchel transform of J_g. Writing $w = \lambda(f-u)$, we have

$$0 \in (w - \lambda f)h + h\lambda \partial J_g^* \left(hw \right), \tag{10.32}$$

which is the minimizer of the *dual* problem

$$\min_{w \in U} \frac{\|h^{1/2} w - b\|_U^2}{2} + \lambda J_g^* \left(hw \right) \quad \text{with} \quad b = h^{1/2} \lambda f. \tag{10.33}$$

Since J_g is homogeneous, J_g^* is the indicator function of a convex set $\mathcal{K}_g$ given by

$$J_g^*(w) = \begin{cases} 0 & \text{if } w \in \mathcal{K}_g \\ +\infty & \text{otherwise} \end{cases} \tag{10.34}$$

$$\text{with } \; \mathcal{K}_g = \left\{ -\operatorname{div} \xi : \xi \in V, |\xi(x)| \leq g(x) \; \forall x \in \Omega^e \right\}.$$

Therefore we may write (10.33) as

$$\min_{hw \in \mathcal{K}_g} \|h^{-1/2} h\, w - b\|_U^2. \tag{10.35}$$

Note that any solution $hw \in \mathcal{K}_g$, must satisfy $h(x)w(x) = -\mathrm{div}\,(g(x)p(x))$ with $|p(x)| \leq 1$. Hence we may write (10.35) as

$$\begin{aligned} &\min_{p \in V} \|h^{-1/2}\mathrm{div}\,(gp) + b\|_U^2 \\ &\text{subject to} \quad |p(x)|^2 - 1 \leq 0 \quad \forall x \in \Omega^e \end{aligned} \tag{10.36}$$

and, introducing the Lagrange multipliers $\alpha(x)$ for the constraint, we obtain the functional

$$\mathcal{F}(p, \alpha) = \sum_{x \in \Omega^e} |h(x)^{-1/2}\mathrm{div}\,(gp)(x) + b(x)|^2 + \sum_{x \in \Omega^e} \alpha(x)(|p(x)|^2 - 1),$$

with $\alpha \in U$, $p \in V$. Proceeding as in [Cha04] the solution of (10.36) satisfies

$$-[g\nabla(h^{-1}\mathrm{div}\,(gp) + \lambda f)](x) + \alpha(x)p(x) = 0 \quad \forall x \in \Omega^e. \tag{10.37}$$

The Karush–Kuhn–Tucker theorem yields the existence of the Lagrange multipliers $\alpha^*(x) \geq 0$ for the constraints in (10.37), which are either $\alpha^*(x) > 0$ if $|p(x)| = 1$ or $\alpha^*(x) = 0$ if $|p(x)| < 1$, but in this case also $[g\nabla(h^{-1}\mathrm{div}\,(gp) + \lambda f)](x) = 0$. In any case $\alpha^*(x) = |[g\nabla(h^{-1}\mathrm{div}\,(gp) + \lambda f)](x)|$, and substituting it into (10.37) and using a gradient descent we arrive at the following fixed-point algorithm:

$$p^{n+1} = \frac{p^n + \tau\left\{g\nabla[h^{-1}\mathrm{div}\,(gp^n) + \lambda f]\right\}}{1 + \tau|g\nabla[h^{-1}\mathrm{div}\,(gp^n) + \lambda f]|}, \tag{10.38}$$

where the maximum $\tau > 0$ will depend on the chosen discretization. For the present scheme, with a straightforward computation [Cha04, ACHR06], one can show that the method converges if $\tau < \frac{1}{8}\frac{1}{\max|g|^2}\frac{1}{\max|h^{-1/2}|^2}$. At convergence, the solution is obtained using the formula $u = f + \lambda^{-1}h^{-1}\mathrm{div}\,(gp)$. We summarized these steps in Algorithm 4.

Algorithm for anisotropic diffusion-type problem

In this case, we consider an anisotropic diffusion problem with Neumann boundary conditions. The discretization of the gradient and divergence are the same as in [Cha04]. Let us consider the anisotropic total variation with $\phi(x, \xi) = |A_x\xi|$, for all $x \in \Omega^e$, where A_x is a symmetric and positive definite (hence, invertible) matrix. As before, the solution of the minimization problem

$$\min_{u \in U} \frac{\|h^{1/2}(u - f)\|_U^2}{2} + \lambda^{-1} J_\phi(u) \quad \text{with} \quad J_\phi(u) = \sum_{x \in \Omega^e} \phi(x, \nabla u x)$$

Algorithm 4 Chambolle's algorithm with Dirichlet boundary conditions

Require: *Data and weights:* $f, g, h \in U$ and parameter λ.
Ensure: *u solution of* (10.30).
1: Initialize $p^0 = 0 \in V$, $q^0 = 0 \in U$, and $t = 0$
2: **while** p not converged **do**
3: Compute: $p^{t+1} \leftarrow \frac{p^t + \tau g \nabla q^t}{1 + \tau |g \nabla q^t|}$
4: Compute: $q^{t+1} \leftarrow h^{-1}\mathrm{div}\,(g p^{t+1}) + \lambda f$
5: **end while**
6: Recover the solution $u = \lambda^{-1} q^{t+1}$

In step 3: p^t is updated for $x \in \Omega^e$, while in step 4: q^t may be updated only for $x \in \Omega$ and $q^t(x) = 0 \; \forall x \in \Omega^e \setminus \Omega$.

is obtained via its dual formulation

$$\min_{w \in U} \frac{\|h^{1/2} w - h^{1/2} \lambda f\|_U^2}{2} + \lambda J_\phi^*(hw) \quad \text{with} \quad w = (f - u)\lambda. \tag{10.39}$$

Since J_ϕ is homogeneous, J_ϕ^* is the characteristic function of a set $\mathcal{K}_\phi$, which we characterize next. Following [AB94] we have $\mathcal{K}_\phi = \{-\mathrm{div}\,\xi^* : \phi^0(x, \xi^*) \leq 1\}$, where

$$\phi^0(x, \xi^*) = \begin{cases} 0 & \text{if } \xi^* = 0, \\ +\infty & \text{if } \xi^* \notin Z_x^\perp, \\ \sup\limits_{\xi : \phi(x,\xi) \leq 1} \langle \xi, \xi^* \rangle & \text{if } \xi^* \in Z_x^\perp \setminus \{0\}, \end{cases}$$

with $Z_x = \{\xi : \phi(x, \xi) = 0\} = \{\xi : |A_x \xi| = 0\}$. Since A_x is symmetric and invertible, $Z_x = \{0\}$ and $Z_x^\perp = \mathbb{R}^n$. Since the second condition is empty and $\sup_{\xi : |A_x \xi| \leq 1} \langle A_x^{-1} A_x \xi, \xi^* \rangle = |A_x^{-1} \xi^*|$, it holds that $\phi^0(x, \xi^*) \leq 1$ if and only if $\xi^* = A_x p$ with $|p| \leq 1$. We get that

$$\mathcal{K}_\phi = \{-\mathrm{div}\,\xi^*(x) : \xi^*(x) = A_x p(x), |p(x)| \leq 1, \forall x \in \Omega^e\}.$$

This allows us to write problem (10.39) as

$$\min_{p(x) : |p(x)| \leq 1} \|h^{-1/2} \mathrm{div}\,(A_x p(x)) + h^{1/2} \lambda f\|_U^2 \tag{10.40}$$

and derive the following fixed-point algorithm:

$$p^{n+1} = \frac{p^n + \tau \left\{ A_x \nabla [h^{-1} \mathrm{div}\,(A_x p^n) + \lambda f] \right\}}{1 + \tau |A_x \nabla [h^{-1} \mathrm{div}\,(A_x p^n) + \lambda f]|}. \tag{10.41}$$

At convergence, the solution is computed as $u = f + \lambda^{-1} h^{-1} \mathrm{div}\,(A_x p)$. For applications to image diffusion it is better to use Neumann boundary conditions, which are imposed by adapting the definitions of the gradient and divergence as in [Cha04].

10.7.2 Computation of the ϕ-Cheeger set and numerical aspects of the ϕ-perimeter computation

The numerical scheme described above produces a function u which is a solution of the PDE (10.13). By Proposition 13, the level sets of u are global minima of the anisotropic ϕ-perimeter with an inflating force. And the regional maxima of u are local ϕ-Cheeger sets in a suitable domain containing them.

In this subsection we describe a method for finding these local ϕ-Cheeger sets. We want to define local extrema of a function $P_\phi(\cdot)/|\cdot|$ which is defined on the set of all connected components of upper level sets of an image u. To fix ideas, let us assume that $N = 2$. Then we have to examine the connected components of the upper level sets $\{u > t\}$, $t \in (0,1]$, of the solution u of (10.13). The ϕ-Cheeger set is defined by $\{u = \|u\|_\infty\}$, but due to the floating-point operations we cannot proceed to a direct computation of this set. Instead we take the ϕ-Cheeger set as the minimum of $t \to P_\phi(\{u > t\})/|\{u > t\}|$ with a suitable discretization of the variable t. Similarly, to compute the local ϕ-Cheeger sets we use the tree of connected components of upper level sets of the image (see [MG00, CM09] and [Mei06] when $N = 3$) and look for the local minima of $P_\phi(cc\{u > t\})/|cc\{u > t\}|$, where $cc\{u > t\}$ denotes a connected component of $\{u > t\}$. Thanks to the topological structure of the tree of connected components of upper level sets, we can speak of local extrema of functions defined on that set. Intuitively, a neighborhood of $\Gamma_t = cc\{u > t\}$ consists in those connected components of upper level sets whose levels are slightly above or below the level of Γ_t.

Let us explain how to compute the weighted perimeter and volume of a given level set. Then we will show how to use this computation to obtain an efficient algorithm to find the connected components of the upper level sets which are local minimizers of the ϕ-Cheeger ratio. When ϕ is of the form $\phi(x,\xi) = g(x)|\xi|$, we will use the expression g-Cheeger set.

Subpixel computation of weighted perimeters and areas

Notice that it is not trivial to compute the perimeter of a set which is defined by pixels or voxels. The naive approach of counting the voxels which touch the boundary of the region does not work, mainly because this quantity is not invariant by rotations. There are two common solutions to this problem: approximate the perimeter using integral geometric measure techniques as in graph cuts [BK03] or approximate the ragged boundary of the set by a smoother surface and compute its perimeter. We found the second option best suited to our needs because, as the goal is to compute perimeters of level sets, we can produce high-resolution approximations of their boundaries by methods such as marching cubes or marching squares [LC87]. Once we have a triangulated surface, we can compute its weighted perimeter by adding the areas of all triangles, each one multiplied by the weight ϕ interpolated at the barycenter of each triangle.

To test the consistency and precision of this scheme, let us consider a spherical image $u(x) = f(|x - x_0|)$ whose profile f is an increasing function from

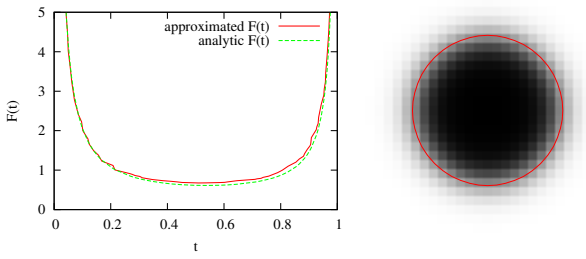


Figure 10.13: Numerical evaluation of $\frac{1}{|\nabla I|}$-Cheeger ratios for a synthetic image where they can also be computed analytically. The image is given by the model in (10.42) with $\kappa = 8$ and $s = 1$. Left: graphs of the exact and computed $F(t)$ for that image. Right: interpolated level curve at which the minimum of $F(t)$ is attained, overlaid on the original image. This example has a very low resolution (the original image has dimensions 31×31). For higher resolutions the approximation is nearly perfect, and the two curves $F(t)$ are visually identical.

$[0, +\infty)$ to $[0, 1)$. For each t in $(0, 1)$, the level set of value t is a sphere of radius $r(t) = f^{-1}(t)$ centered at x_0, and this surface is weighted by $\frac{1}{|\nabla u|} = \frac{1}{f'(r(t))} = r'(t)$. The $\frac{1}{|\nabla u|}$-Cheeger ratio is then

$$F(t) = \frac{NV_N r^{N-1} r'}{V_N r^N} = N\frac{r'(t)}{r(t)},$$

where V_N is the volume of the unit ball of $\mathbb{R}^N$. This function $F(t)$ is a real-valued function whose minimum can be evaluated numerically, or even analytically in some easy cases.

We can set, for example,

$$f(x) = \frac{1}{1 + \exp\frac{\kappa - x}{s}}, \tag{10.42}$$

where κ and s are parameters, such as $\kappa = 8$ and $s = 0.1$. Intuitively, the desired segmentation of this image is a circle of radius $\kappa \pm s$ or, equivalently, some level set of value near $t = \frac{1}{2}$. In Figure 10.13, we compare the graphs of the $\frac{1}{|\nabla u|}$-Cheeger ratio over t as computed analytically and with the numerical methods described above. The minima in both cases is attained very near $t = \frac{1}{2}$, which agrees with our intuition and suggests that the numerical approximations we use are consistent.

As another numerical test, we computed the Euclidean Cheeger set of a square and a cube (using the Euclidean metric). See Figures 10.15 and 10.14 for the plausible result we obtained.

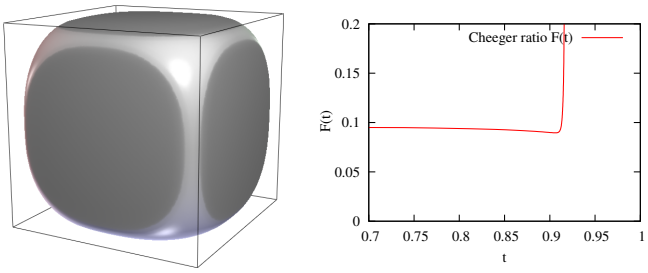


Figure 10.14: Cheeger set of a cube. This was computed by Euclidean total variation minimization of the characteristic function of the cube, followed by the selection of the level surface that minimized the Cheeger ratio. On the right, the graph of $F(t)$. Notice that for large values of t, the level set collapses to a central point of the domain, so that its Cheeger ratio is much larger than the optimum.

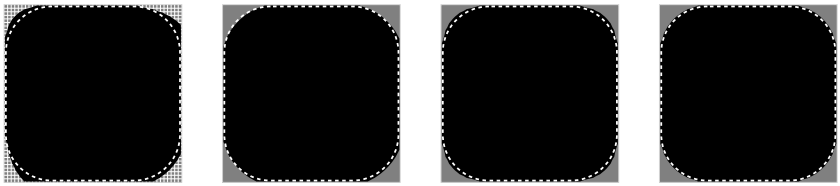
The discrete images u are obtained by an iterative numerical method, and they have floating-point values. Most of the values are concentrated around 1, with the interesting part of the range being often contained in the interval $[0.99, 1]$. Thus, it is important to conserve their floating-point values. This implies that there are as many different level surfaces as pixels, one for each different floating-point value. But it is not necessary to compute the ϕ-Cheeger ratio of all of these surfaces: via dichotomic search we can efficiently locate the minimum.

On the symmetry of the numerical scheme

In Figure 10.15 we show an example of the Cheeger set determined with the method described above. Observe that the first solutions are asymmetric; this effect is particularly clear (and annoying) for small images (the first square is 50 pixels high), and it is due to the forward/backward scheme adopted to discretize the gradient (10.26). In [Cha04] the author remarks that the finite difference scheme converges to the continuous formulation as the number of samples $N \to \infty$, but when applied to volumetric images increasing the sampling is not an affordable option. To maintain the symmetry of the solutions, we propose the *consensus algorithm*, which computes the mean solution of all the finite difference schemes (4 schemes in two dimensions, and 8 in three dimensions) at each iteration. The consensus algorithm outperforms the considered schemes while keeping the symmetry, but it is 4 times (or 8 for three dimensions) slower than the standard finite difference scheme.

10.8 Conclusion

We have introduced the anisotropic total variation for eventually degenerate metric integrands ϕ. As a particular case, we have considered the geodesic active



1/12.858 1/12.960 1/12.949 1/12.961

Figure 10.15: Computation of the Euclidean Cheeger sets (black regions) for a square of 50 pixels. For the square the exact Cheeger set is known (dashed line), and its constant is 1/13.253. The first column shows the results obtained with Chambolle's numerical scheme [Cha04]; observe that it is asymmetric. The second solution is obtained by increasing the sampling (as proposed by Chambolle); this considerably improves the solution, but the asymmetry persists. The third column shows a symmetric solution obtained by averaging all the numerical schemes (4 in two dimensions). The fourth solution is obtained using the *consensus algorithm*. All the results were obtained after 20000 iterations of (10.38).

contour model, which corresponds to $\phi(x, \xi) = g(x)|\xi|$, where $g(x)$ is a function that may vanish for some values of x. We have defined the notion of ϕ-Cheeger set and we have shown that, for suitable metric integrands ϕ, the maximal ϕ-Cheeger set can be computed as the level set associated to the maximum of the solution of a ϕ-total variation minimization problem with Dirichlet boundary conditions and datum $f = 1$. Thus, in the particular case of the geodesic active contour model with inflating force, we can compute a global minimum. Considering other definitions of ϕ permits to apply the model for edge linking or to interpolate data along the level lines of a reference image.

Chapter 11

Geodesic neighborhoods for piecewise affine interpolation of sparse data

Two widely used techniques for acquiring digital elevation models of urban areas are laser range scanning (LIDAR) and stereo photogrammetry. Both techniques have a common drawback, for a variety of reasons the information they provide is sparse or incomplete. But in both cases it is fair to assume that a high resolution image of the scene is available.

In this chapter we propose a different interpolation method for sparse data that incorporates the geometric information of a reference image u. This allows us to interpolate the data set while respecting the edges of u. The idea consists in defining for each sample a set of neighboring samples and then fit a model (affine for instance) for interpolating the current point. The core of the algorithm is a fast method for computing geodesic distances *between image points, which has been successfully applied to colorization and segmentation. The geodesic distance allows to determine the set of points that is used to interpolate a piecewise affine model in the current sample. This first interpolation is then refined by merging the obtained affine patches using a greedy Mumford-Shah like algorithm. The output is a piecewise affine interplation of the data set that respects both the given data and the radiometric information provided by the reference image u.*

11.1 Introduction

We consider the problem of interpolating a set of range measurements of a scene using the additional knowledge of the radiometric information of the same scene given by the image u.

This scenario is common in the case of LIDAR measurements, since a digital image has a higher density, and its acquisition is faster, when compared to the range data. We will take advantage of the information provided by the associated image to interpolate the sparser range measurements. The same applies to the case of stereo reconstruction of urban *Digital Elevation Models (DEM)*, since

in this case the height information can be accurately determined only at a sparse set of locations in the image [Del04, DR07, SMA10], and therefore using the reference image to interpolate them will provide a denser height map.

As in most works we adopt the Lambertian hypothesis. That is, a uniform surface (in the model) with a constant angle will be seen with a constant intensity in the image. This assumption allows us to extrapolate information across uniform regions of the image. Clearly not all uniform regions will have some data sample in it, especially if we consider textured surfaces. In those cases we extrapolate from the "nearest" sample, where nearest refers to the geodesic distance that takes into account the radiometric and edge information of the image. We have to acknowledge that areas of constant intensity in the image do not necessarily correspond to the same surface in the model, but this situation is unlikely since in most cases a discontinuity in the model corresponds to a change of intensity in the image.

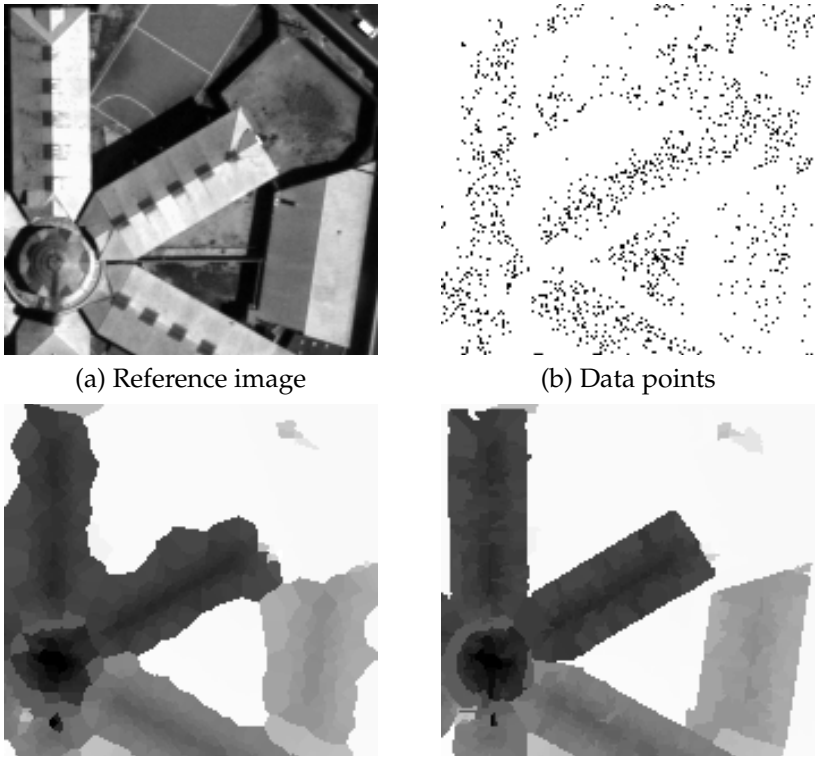
11.1.1 The basic idea

Let us denote by $u(x) : \Omega \to \mathbb{R}^+$ the monochromatic image of a scene defined on $\Omega \subset \mathbb{R}^2$, and let $G(\lambda) : \Lambda \to \mathbb{R}$, $\lambda \in \Lambda \subset \Omega$, be a given depth function which we assume to be known only at the samples Λ. As in [YS06, BS07, SGT02] the idea is to use the geodesic distance to incorporate the radiometric information provided by the image u in the interpolation of the sparse data G to produce a dense depth map $H(x) : \Omega \to \mathbb{R}$ that fits the values of G. This distance measures the minimum variation of u between two points. Therefore, the distance of two points along the same isophote is 0, while the distance of two points at both sides of a discontinuity of u will be proportional to the "jump" of u.

With the geodesic distance we define the geodesic Voronoi diagram for the sites in Λ, and then interpolate in each cell using an affine model. The geodesic Voronoi cells, as opposed to the Euclidean nearest sample interpolation, permit to recover sharp and meaningful boundaries of the model (see Figure 11.1 and Figure 11.2). Moreover, since the distances can be efficiently computed with a modified pixel queue algorithm, this method provides a fast alternative to iterative anisotropic diffusion algorithms.

In [YS06] the authors use the K-nearest geodesic samples to blend the chrominance information given by color the scribbles, while here we reinterpret those *geodesic neighborhoods* as samples belonging to the same surface as the current point (see Figure 11.2). The linear blending used in [YS06] cannot recover kinks in the model (see Figure 11.3(a)). Because, as pointed out in [She68b], they consider just the distance and not the spatial distribution of the nearby samples. It becomes clear that, in order to recover kinks and discontinuities, an affine model must be adjusted to each geodesic neighborhood. In Figure 11.3(b) is shown a result obtained by least squares fitting of a plane.

This procedure is suitable for noiseless measurements, however when considering noise or outliers in the measurements the quality of the results drops notably. To solve this issue we will resort to the robust selection and fitting of the neighborhoods, and the merging of regions with compatible models. It is possi-



(a) Reference image (b) Data points

(c) Euclidean Voronoi (d) Geodesic Voronoi

Figure 11.1: Piecewise constant geodesic Voronoi interpolation of height information (darker is higher). In (a) we show the reference image used to compute the geodesic distances, (b) shows the position of the samples with known depth (5% of the total pixels). Interpolating the samples of (b) with Euclidean Voronoi cells produces (c), while (d) corresponds to the geodesic Voronoi interpolation.

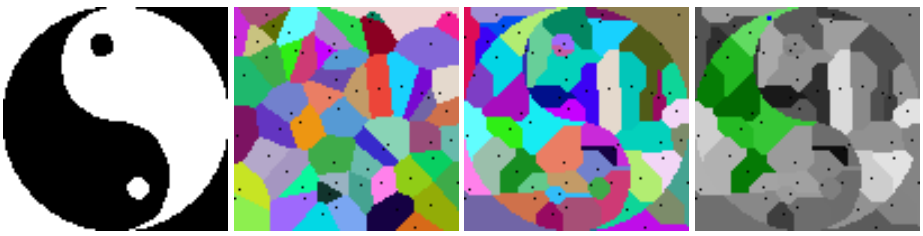
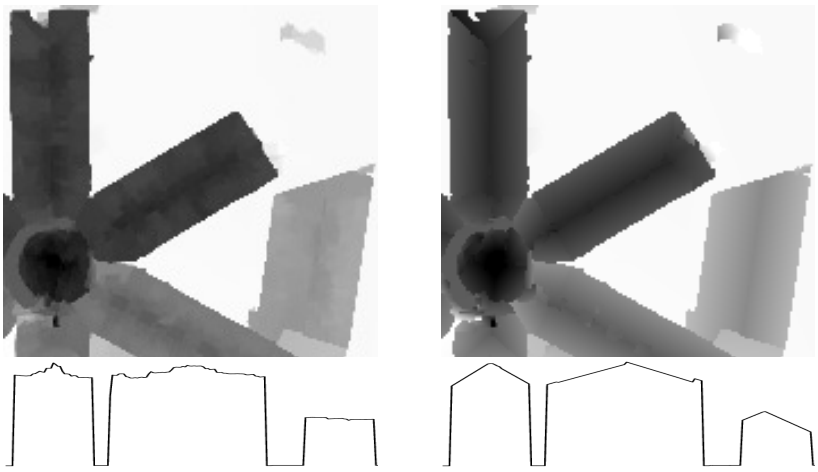


Figure 11.2: Geodesic Voronoi cells and geodesic neighborhoods. From left to right: the reference image, the samples with the corresponding euclidean Voronoi cells, the geodesic Voronoi cells and the geodesic neighborhoods corresponding to the blue dot in the forth image.



(a) Linear blend (b) Affine adjust

Figure 11.3: Interpolation of the samples in Figure 11.1(b) using the 5-nearest (geodesic) samples. Left: Weighted mean blending [YS06]. Right: interpolation by affine plane estimation using the same 5 samples. In the second line we compare a horizontal line section from each of the images.

ble to propose plenty of merging strategies to address this issue [IPG+07], which will eventually lead us to a higher level representation of the scene. To improve the estimation in case of noisy samples, we will merge adjacent geodesic Voronoi cells using a greedy criterion and an error based stopping condition.

The idea of using the information provided by the image u to guide the interpolation is not new. In [FLA+06] the authors solve an anisotropic minimal surface problem to interpolate sparse disparity data. But minimal surfaces are unable to resolve discontinuities and TV is unable to recover kinks in the model, moreover, this type of schemes are likely to be slow due to its iterative nature. In [IPG+07] the authors incorporate u through an initial segmentation, adjusting the data points of each region with a robust estimation and use a region merging strategy to merge similar regions. Ideally, with a "perfect" initial segmentation, this method gives the best possible result, but since each initial region must already contain the points needed for the estimation, a bad initialization could be catastrophic. In our method there is no initial segmentation, the Voronoi regions are induced by the samples, and the affine planes are estimated using the geodesic neighborhood relations.

The chapter is organized as follows: in next Section we define the geodesic distance, the geodesic neighborhoods and we use them in Section 11.3 to fit a piecewise affine model. In Section 11.4 we present a basic region merging algorithm to increase the robustness of the estimation. Section 11.5 is devoted to the discussion of the results, the limitations of the proposed method and their possible solution. In Section 11.6 we give some conclusions and future work.

11.2 Geodesic distances and neighborhoods

Let s and t be two points in Ω and let $C(p) : [0,1] \to \Omega$ be a curve in Ω. We write $C_{s,t}$ to refer to a curve connecting s and t such that $C_{s,t}(0) = s$ and $C_{s,t}(1) = t$, and $\dot{C}_{s,t}(p)$ denotes the tangent of the curve at p. Then we define the geodesic distance between s and t as:

$$d(s,t) = \min_{C_{s,t}} \int_0^1 |\nabla u(C_{s,t}(p)) \cdot \dot{C}_{s,t}(p)| + \varepsilon |\dot{C}_{s,t}(p)| \, dp. \tag{11.1}$$

This distance is a regularized anisotropic distance. When $\varepsilon \ll 1$ it behaves as the geodesic distance (at least for short curves)

$$\min_{C_{s,t}} \int_0^1 |\nabla u(C_{s,t}(p)) \cdot \dot{C}_{s,t}(p)| \, dp. \tag{11.2}$$

The curve with smallest geodesic length is the a curve that crosses the least level lines of the image u. For instance if s and t are located on the same isophote then the shortest geodesic path is along the isophote, since in this case $\dot{C}_{s,t}(p) \perp \nabla u(C_{s,t}(p))$. Then along the isophotes (and across regions where the image u is constant) all the paths have 0 geodesic length (11.2) no matter how long are the paths. This causes to the following problem: when choosing the nearest sample among two samples, both at the same geodesic distance, we still prefer the nearest one (in the Euclidean sense). This is why we add a small part of euclidean distance in (11.1). The distance (11.1) will behave like (11.2), as long as the Euclidean length of the shortest curve is less than ε^{-1}, and to make sure of it we select ε inversely proportional to the size of the domain and the range of u. A geodesic distance defined as (11.1) will respect the contrasted boundaries of the image and in case of a tie it will use the Euclidean criterion as tiebreaker.

Then the *geodesic neighborhood* $GN_K(p)$ is the set formed by the K-nearest (in the geodesic sense) samples of Λ to the point p. For $K = 1$ we get a correspondence $GN_1 : \Omega \to \Lambda$ that for each point in $p \in \Omega$ assigns the nearest sample $\lambda \in \Lambda$. We define the *geodesic Voronoi region* associated to of λ, as the set $\{p \in \Omega : GN_1(p) = \lambda\}$ (see the geodesic Voroni diagrams in Figures 11.1 and 11.2). With $K > 1$ we get a multi-correspondence map that associates several samples to each point p. In Figure 11.2 we highlighted a geodesic neighborhood of a point. Observe that all the points belong to the same object, we will take advantage of this in the next section.

Remark 9. We have seen that the geodesic distance (11.1) embodies information provided by the edges of the image u. This will prevent two points s and t, at both sides of an edge, from having a small distance. However under some circumstance (mainly due to noise or aliasing) the shortest path may indeed cross an edge, and the crossing will occur at the least contrasted point of the edge. To prevent this, when some high level information is available, it can be embedded in the distance. For instance, the edges obtained with an edge detector can be introduced as a hard constraint in the metric in order to penalize

curves crossing them (add infinite cost $|w(C_{s,t}(p)) \cdot \dot{C}_{s,t}(p)|$ where w is a map of the edge's normals).

11.2.1 Computation of the Geodesic Neighborhoods

To compute the geodesic neighborhoods we use an algorithm similar to the one proposed in [YS06, SGT02]. That is, for each data sample in Γ we run Dijkstra's algorithm over a 4-connected pixel lattice to compute the shortest paths to all the pixels in the image. And for all the pixels in the image we store the K-nearest geodesic distances to all samples in Γ. The computation of the geodesic lengths of the paths defined on the 4-connected pixel lattice, decomposes into lengths of path elements which are the edges of the lattice.

The overall performance of this method depends on the number of neighbors K in the geodesic neighborhood, and this depends on the density of samples. However since we only require the K-nearest neighbors, we can early stop the Dijkstra iterations. The algorithm is summarized in Algorithm 5.

Remark 10. Processing the data points in a random order improves the efficiency. Randomly selecting the current pixel $p \in \Omega \setminus \Lambda$ increases the speed of the algorithm by a factor 3.5 (for datasets that cover 30% of the pixels) and 2 (for datasets that cover 5% of the pixels).

11.3 Robust affine plane interpolation

For each data sample $\lambda \in \Lambda$ we have defined a geodesic Voronoi cell and a geodesic neighborhood. Figures 11.1 and 11.2 illustrate that the geodesic Voronoi diagram successfully accounts for discontinuities in the image. We propose to use the geodesic neighborhood $GN_K(\lambda)$ to fit a plane through its points and extend it to the whole geodesic Voronoi cell. This will give a piecewise affine model H.

To compute the interpolation H at $p = (p_x, p_y) \in \Omega$, we first determine its nearest (with respect to the geodesic distance) data sample $p^* \in \Lambda$, and its geodesic neighborhood $GN_K(p^*)$. Then compute

$$H(p) = w_p^T v_{p^*},$$

where $w_p = (p_x, p_y, 1)^T$ are the homogeneous (spatial) coordinates of p, and $v_{p^*} \in \mathbb{R}^3$ contains the affine parameters determined by the least squares regression

$$v_{p^*} = \arg\min_{v \in \mathbb{R}^3} err(v, GN_K(p^*)), \tag{11.3}$$

$$err(v, GN_K(p^*)) = \sum_{q \in GN_K(p^*)} |w_q^T v - G(q)|^2.$$

Observe that the samples of $GN_K(p^*)$ used in (11.3) are (except for p^* itself) all outside the geodesic Voronoi cell of p^*. This is the main advantage of this approach: using the information from the geodesic neighbors to extract knowledge

Algorithm 5 Geodesic neighborhood algorithm

Data structures. For each pixel in Ω define:

- status *label*: needed for the Dijkstra algorithm (ACCEPTED, FAR, TRIAL).
- *distance*: stores the current minimum distance.
- *nearest*: refers to the nearest pixel in $\Omega \setminus \Lambda$.
- *K-Nearest List*: stores the *K*-nearest distances to samples obtained up to the moment. This list is the output of the algorithm and only keeps the K smallest values.
- *maxDistance*: the value of the maximum distance stored in the list, if the list is not full then *maxDistance* is infinity.

Geodesic neighborhoods computation.

Construct a priority queue Q of pixels, sorted by the *fitness* field.
for all pixel $p \in \Omega \setminus : \Lambda$ **do**
  Clear the queue Q.
  Label all the pixels of Ω as FAR.
  Add p to the priority queue, setting $p.distance = 0$.
  while Q is not empty **do**
    Extract the top of the queue $Q \rightarrow q$.
    Update $q.label$ as ACCEPTED.
    Add the $q.distance$ to the *K-Nearest List* of q.
    for all non ACCEPTED neighbors of $q \rightarrow r$ **do**
      Compute the distance from p to r trough q, that is:
      $dNew = d(p,q) + d(q,r)$.
      if $r.label$ is TRIAL **and** $dNew < r.distance$ **then**
        Update $r.distance = dNew$.
      end if
      if $r.label$ is FAR **and** $p.distance < r.maxDistance$ **then**
        Update $r.label$ as TRIAL, set $r.distance = dNew$ and add r to the queue Q.
      end if
    end for
  end while
end for

about the current region. Exploiting the hypothesis that the geodesic neighbors belong to the same surface, allows to construct an affine model for the current region. Otherwise (using a single region) we could only determine a constant model.

The size K of the neighborhood is a critical parameter here. If it is too small, then the estimation of the plane will be poor. If it is too big, then we will not be able to recover small planar surfaces (with only a few samples on them). Moreover if the geodesic neighborhood contains outlier samples (that do not belong to the same object as the center p) then the result will be biased.

To remove the outliers we use a *RANdom SAmple Consensus* (RANSAC) [FB81], which we modified to ensure that the consensus set always contains the 3-nearest geodesic neighbors of p. This choice of the 3 neighbors is arbitrary, however this parameter depends on the density of the samples. Indeed with this selection we are supposing that each planar region contains at least 3 samples in it.

Remark 11. Let us note that some neighborhoods may be elongated because the geodesic neighborhoods are expected to be anisotropic. In this case the regression is ill conditioned, and a single plane cannot be estimated. We are not considering this possibility, however this should be taken into account for future developments.

11.4 Region merging

Fitting each individual Voronoi region with an independent plane represents an inconvenient when the height information is perturbed by noise (or small measurement errors). Even for adjacent regions, that share some samples, the different fittings may not coincide, resulting in an irregular model as shown in Figure 11.4(b).

To further improve the plane estimation we will merge the models of adjacent geodesic Voronoi regions obtained in the previous section. For that, we minimize a simplified Mumford-Shah functional

$$E(B, f) = \sum_{R \in \mathcal{P}(\Omega)} RErr(R, f) + \lambda \int_B g(s)\, ds, \quad \lambda \geq 0, \tag{11.4}$$

where $\mathcal{P}(\Omega)$ denotes a partition of Ω, B the set of boundaries of the regions of $\mathcal{P}(\Omega)$, $RErr(X, f) = \sum_{x \in X \cap \Lambda} |f_X(x) - H(x)|^2$, H denotes the output of the affine plane interpolation, and f_X is the affine model for the region X obtained as

$$f_X(x) = w_x^T \left(\arg\min_v \sum_{p \in X \cap \Lambda} err(v, GN_K(p)) \right).$$

The weight $g(s)$ of boundary length term is a function that is big at poorly contrasted boundaries and very small at the well contrasted ones. In practice, we

Algorithm	Noiseless Disp.	Noisy Disp. [SAM08]
Linear Blend [YS06]	0.0902	0.0983
Affine (no RANSAC)	0.0810	0.0829
Affine	0.0687	0.0708
Affine+Merge	0.0649	0.0637

Table 11.1: Mean Square Errors (MSE) of the results obtained using the ground truth information. All the experiments were performed with the 25 geodesic nearest neighbors. The first column corresponds to the interpolation of noiseless data (the values of samples are exact), and in the second the samples are computed with a stereo correlation algorithm [SAM08].

take $g(x) = \exp\left(-\frac{|\nabla u(x)|^2}{\sigma^2}\right)$, with $\sigma^2 = \mathrm{Var}(u)$, so that the length of a curve along a contrasted boundary is almost zero. As in [MS95], a greedy minimization of (11.4) will merge two regions A and B when

$$\lambda > \frac{RErr(A \cup B, f) - RErr(A, f) - RErr(B, f)}{\ell_g(\partial(A, B))},$$

where $\ell_g(\partial(A, B)) := \int_{\partial(A,B)} g(s)\, ds$ is the weighted length of the common boundary between A and B.

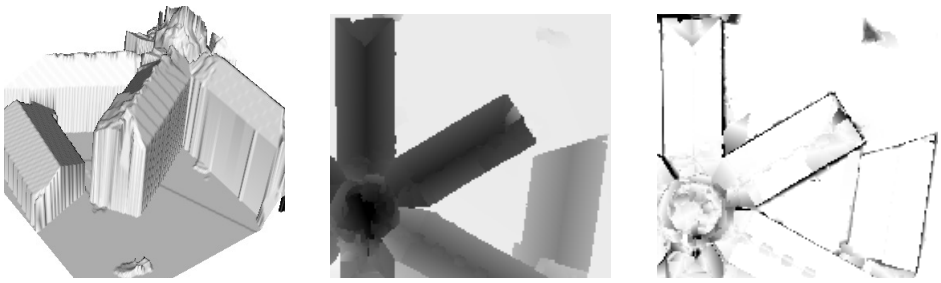
In the context of urban landscape interpolation there may be clear boundaries that should not be removed by the merging. This could be the case of the segments detected by the LSD algorithm [vGJMR10] (see Figure 11.5(a)), which reflect the geometry of the urban landscape. If the edge crosses a line segment, then its length will be set to zero, and this will avoid merging these two regions. Moreover, we will keep track of all the edges in this situation and forbid any merge that removes them from the segmentation.

Finally, let us mention that the selection of λ is critical for obtaining reasonable results, because if set too low it will produce an over-segmentation, and if too high it will merge different objects. Instead of selecting a particular value of λ we will control it indirectly by monitoring the merging error $RErr(A \cup B, f)$ at each step and stopping the algorithm if it exceeds a threshold proportional to the measurement's error.

Remark 12. Note that at each step of the region merging loop the affine model f_R for the merged region must be updated, however it is easy to show that the merging of two affine models and the the re-evaluation of the error is a constant time operation.

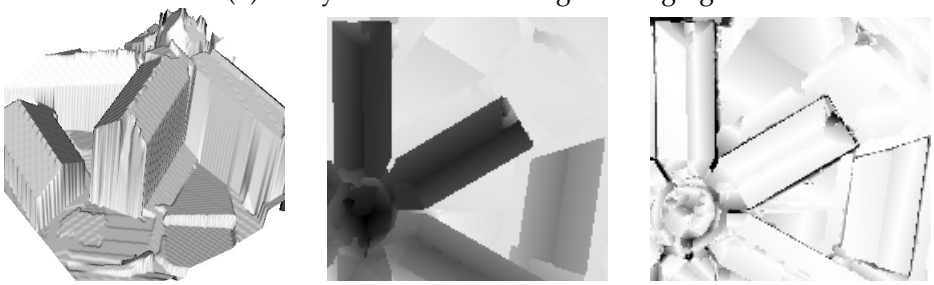
11.5 Experiments

In our first experiment we will use the stereo data set kindly provided to us by CNES. We displayed it in Figure 11.1. We consider two cases: exact depth measurements (obtained from the ground truth), and noisy depth measurements



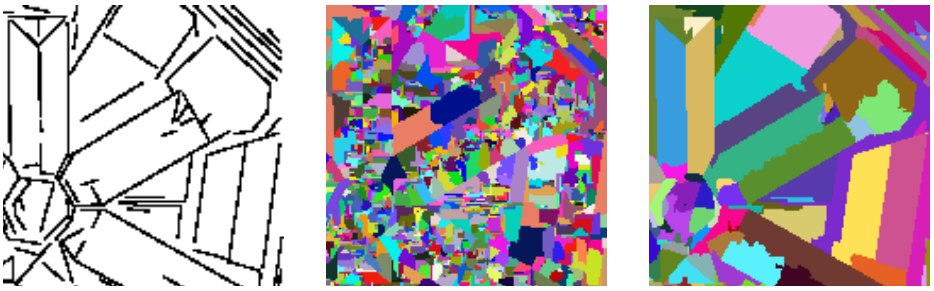
(a) Noiseless dataset without region merging

(b) Noisy dataset without region merging



(c) Noisy dataset after region merging

Figure 11.4: Left: interpolated map. Middle: interpolated map represented as a gray level image. Right: error map. In (a) we display the interpolation of the noiseless data without region merging. (b) is the same for the case of noisy data. In (c) we display the region merging applied to (b).



(a) LSD segments (b) Before merging (c) After merging

Figure 11.5: Region merging. The region merging algorithm do not merge regions across detected segments (a), and stops if the error is too big. (c) shows the regions obtained after merging the regions of (b).

computed with the subpixel stereo algorithm described in [SAM08]. For the noisy case the variance of the error is estimated to be 0.013, and we use it to select the parameters for RANSAC and the region merging stopping condition.

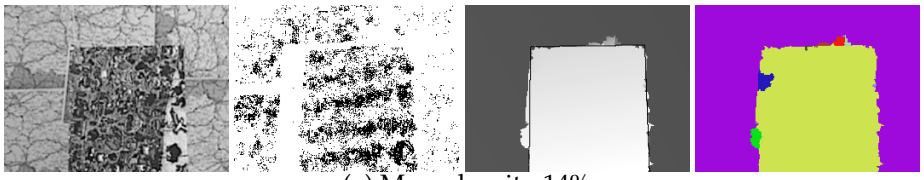
Both datasets were decimated to 5% of the image pixels (240×240 pixels), and the depth information along the detected segments was erased. Such a low density is unlikely in a real application, however this may prove the performance of the method. For these images the computation of the 25 geodesic nearest neighborhoods takes 14 seconds on a 1.6 Ghz CPU with non optimized C code. All the RANSAC trials are computed in less than 5 seconds and the region merging takes less than 2 seconds.

In Table 11.1 we summarize the performance (in terms of Mean Square Error) of each step of the proposed method. First, let us consider the noiseless case. In Figure 11.4(a) we display the result of the affine interpolation without merging. The error map, shown in the rightmost column of Figure 11.4(a) (darker means higher error with respect to the ground truth information), shows that most of the errors come from small planes, non-planar surfaces, and discontinuities. Note that, despite of the low sample density, kinks are recovered with great precision and planes are well recovered.

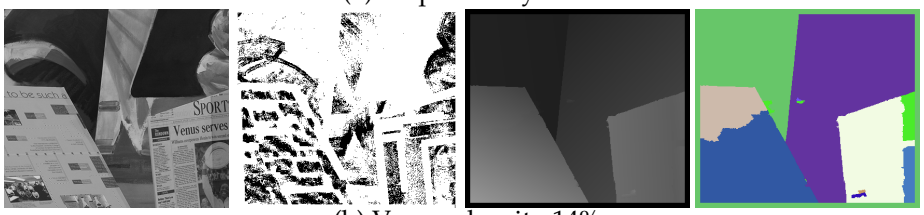
In the noiseless case we can skip the merging step since all the height information is consistent across neighborhoods. However in the noisy case this is not possible since the result of the neighborhood interpolation is very irregular (see Figure 11.4(b)). The result shown in Figure 11.4(c) is obtained after merging the region models in Figure 11.4(b). As mentioned in Section 11.4 the algorithm avoids merging regions that are separated by LSD segments (shown in Figure 11.5(a)), and stops when the merged error variance exceeds 0.013. It is interesting to see that the resulting partition (see Figure 11.5(c)) also resembles a segmentation of the model.

In our second experiment we considered the images from the Middlebury dataset[1] shown in Figure 11.6, the disparity was computed using [SAM08] and

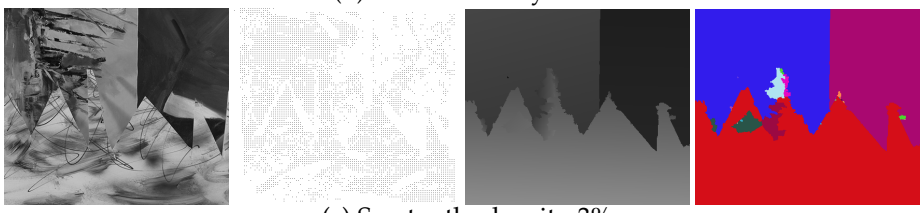
[1] `http://vision.middlebury.edu/stereo/` .



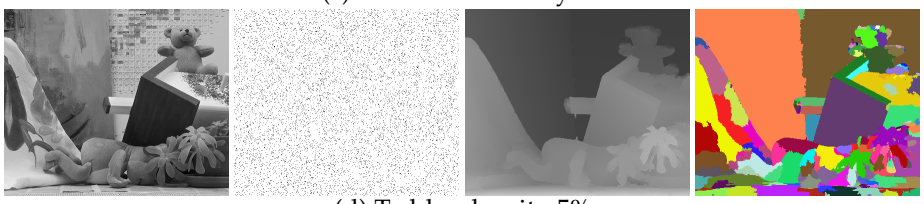
(a) Map: density 14%.



(b) Venus: density 14%.



(c) Sawtooth: density 3%.



(d) Teddy: density 5%.

Figure 11.6: Middlebury dataset. From leftmost to rightmost columns: the reference image, the depth samples, the obtained interpolation (after merging) and the merged regions. The depth data in the first three rows was computed with a stereo correlation algorithm [SAM08] and subsampled (density specified in each case). For the Teddy image the data is a random subsampling of the ground truth.

subsampled to different factors. The results are also shown in Figure11.6 together with the final segmentation. Observe that in presence of poorly contrasted edges between strongly textured regions the method tends to produce artifacts as seen in Figure 11.6(a,c). These artifacts are consequence of our definition of geodesic distances, which is based on pixel graylevel differences. Hence, complex textures are likely to produce false boundaries while poorly contrasted boundaries are not distinguished by the geodesic distance. To alleviate this type of problem, in [SGT02, BS07], the authors considered the use of higher level features in the definition of the geodesic distance.

Finally, let us point out the main limitations of this method. Since we do not construct a geometric description of the scene, there is no distinction between kinks and discontinuities in the model. Note that in Figure 11.4(c) many roof planes do not coincide at the top.

11.6 Conclusions and future directions

The geodesic Voronoi cells adapt to the shapes present in the reference image, contributing to the interpolation of sparse data. We have also seen that, using the information obtained from the nearby (in the geodesic sense) samples, it is possible to locally infer a model from the data. We presented here an algorithm that uses an affine interpolation model. However the methodology can be easily adapted to other (non-affine) local models [She68b].

A region merging criterion that distinguishes kinks from discontinuities, would improve the results. A step in this direction could be to weight the length of the boundary between two regions by a factor of "coplanarity", favoring the elimination of boundaries between similar planes.

A different possibility is to merge the concept of geodesic Voronoi cells with the region growing method for detecting planar patches, as proposed in [ABGvG10]. In [ABGvG10] the authors considered dense samplings, but it is a straightforward extension to use the geodesic Voronoi cells for interpolating sparse data.

Finally, one could also interpret the region merging as a clustering problem. Consider a graph where the nodes are the Voronoi regions and the edges connect the geodesic neighborhoods (with weights proportional to the geodesic distance). An example of such graph is depicted in Figure 11.7. Then we can see the region merging as a spectral clustering problem, where the issue is separate the graph into disjoint subgraphs by cutting the least number of edges. Techniques such as Normalized Cuts [SM02] compute such segmentations by iteratively splitting the graph.

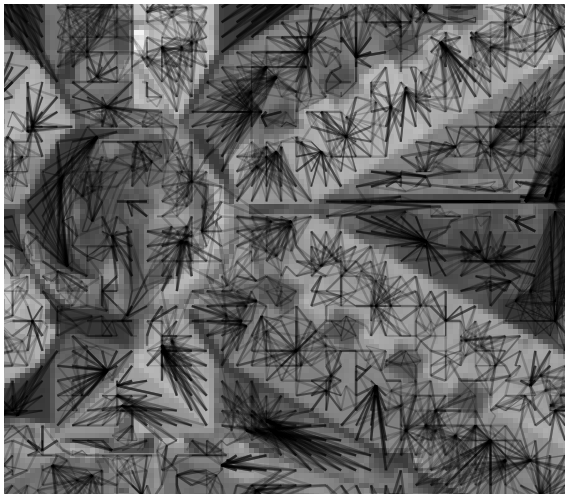


Figure 11.7: Geodesic neighborhood coherence. The black lines show the geodesic nearest neighbors (four of them) corresponding to several points chosen on a regular subgrid of the image. The width of the line is inversely proportional to the distance (a small geodesic distance is represented by a thick line). Observe that the samples belonging to the same object form strongly connected subgraphs, while different objects are connected by thin edges or remain unconnected.

Chapter 12

Conclusions

As we mentioned in the introduction the three parts of this thesis are motivated by problems related to the acquisition of satellite images and the obtention of digital elevation models from them. This led us to the study of several interpolation problems.

In the first part of the thesis we considered the restoration (denoising and deconvolution) of irregularly sampled images as a constrained regularization problem using a band limited model. We propose a method that imposes the regularity of the solution using a frequency adaptive regularizer, and enforces the noise constrains locally. The combination of both allows us to obtain very competitive results. We also studied the problem of image fusion, we discussed the restoration of aliased images, and developed two methods for determining the spectral region of an image that is "alias-free".

A complementary approach to the interpolation of irregularly sampled images is to hallucinate the features that are not captured by the sampling. In the second part we explored the use of a self-similarity prior as a regularity model for the interpolation of sparsely sampled images (restricted to the case where the samples are at integer positions). Our study starts with the introduction of a new exemplar-based inpainting framework which we subsequently adapted to the interpolation of sparsely sampled images. The interpolation results obtained with the self-similarity prior are surprisingly good allowing to recover textured patterns even at very low sampling densities.

Finally, in the third part, we considered the interpolation of sparse elevation data as provided by the small baseline stereo correlation algorithm in [SMA10]. For this problem we explored several anisotropic regularizers that allowed us to impose the geometric structure of a reference image in the interpolated result. Indeed, we studied three different regularization models: an anisotropic minimal surface regularizer, the anisotropic total variation and a new piecewise affine interpolation algorithm. Let us mention that the study of the anisotropic total variation was primarily motivated by the computation of ϕ-Cheeger sets for solving exactly the geodesic active contours model with inflating force. The piecewise affine interpolation algorithm associates to each pixel a set of geodesic

nearest neighbors that is used for fitting a plane.

Summarizing, the main contributions of this thesis are:

- An algorithm for restoring irregularly sampled images, with local constraints and using a spectral adaptive regularizer.
- An exemplar based framework for image inpainting. The extension of the previous framework to the interpolation of sparsely sampled images.
- The application of anisotropic total variation for computing ϕ-Cheeger sets and its use to exactly solve the geodesic active contour model with an inflating force.
- An algorithm for interpolating and regularizing urban digital elevation models that uses a geodesic distance to impose the geometry of a reference image.

12.1 Future directions

Irregular sampling restoration in presence of antialias. Removing the alias from an irregularly sampled image has proven to be an ambitious task. We see the image fusion problem as the restoration from irregularly sampled and aliased images with a prescribed bandwidth (which is indicated by the spectral projector). But, as we have shown in Chapter 4, using the spectral projector (in the irregular case) leads to a bias in the estimation of the noise which is difficult to fix.

Having approached the restoration problem using a band limited interpolation model all the samples are needed to determine the restoration of each pixel. This has two drawbacks, first the performance, and second the impossibility to adapt locally the modulation transfer function and the spectral projector. The second is particularly annoying because, previous to the fusion, the images should be rectified and the rectification of the terrain distortions may induce changes in the modulation transfer function [ACF07].

Some steps have already been taken in order to build "local" restoration algorithms for irregularly sampled images [ACD10, Van06]. Hopefully, this work will continue in the direction of including local modulation transfer functions, ultimately leading to the development of high quality image fusion algorithms.

Image fusion using exemplar-based interpolation. Extending the framework of Chapter 7 to handle images sampled at non integer positions will allow to inject new samples into the restoration process. These samples can taken either from the image being restored or from other images of the same scene.

We anticipate that this kind of processing will be computationally expensive because of the nearest neighbor search. Since in this case it is not possible to take advantage of search tree algorithms, this task is currently performed by a brute force search on the neighborhood of the pixel. However we believe that it is

possible to overcome these limitations by using collaborative random search algorithms such as in [BSFG09, PELS10] combined with candidate lists [TZL+02]. In Section 6.A, in the context of image inpainting, we already described an extension of [BSFG09] that consider lists of candidates.

Better uses for the geodesic neighborhoods. As we mentioned in Chapter 11 geodesic neighborhoods provide useful information regarding the similarity of the geodesic Voronoi cells with its neighbors. But this information needs to be understood globally in order to extract some useful relations about the regions and build clusters. Future research will concentrate on determining the applicability of graph clustering techniques [SM02] to the graph of geodesic neighborhoods.

Exemplar-based urban digital elevation model interpolation. Clearly the features that define the urban environment are recurrent, so it seems reasonable to define dictionaries of common features which can be used for assembling a digital elevation model that fits the known data.

Extension of exemplar-based inpainting to video. The extension of the exemplar based inpainting framework to video is not trivial, the main challenge being the temporal consistency and the handling of occlusions.

Publications

The work presented in this thesis has been published in the following peer reviewed journals:

- G. Facciolo, A. Almansa, J. F. Aujol, and V. Caselles, "Irregular to regular sampling, denoising, and deconvolution," Multiscale Modeling and Simulation, vol. 7, no. 4, pp. 1574-1608, 2009.
- V. Caselles, G. Facciolo, and E. Meinhardt, "Anisotropic cheeger sets and applications," SIAM Journal on Imaging Sciences, vol. 2, no. 4, pp. 1211-1254, 2009.
- P. Arias, G. Facciolo, V. Caselles and G. Sapiro,. "A Variational Framework for Exemplar-Based Image Inpainting", Accepted for publication in Springer International Journal of Computer Vision, 2010.

and peer reviewed conferences:

- G. Facciolo, F. Lecumberry, A. Almansa, A. Pardo, V. Caselles, and B. Rougé, "Constrained Anisotropic Diffusion and some Applications," in British Machine Vision Conference, vol. 3, Edinburgh, U.K., September 2006, pp. 1049-1058.
- G. Facciolo and V. Caselles, "Geodesic neighborhoods for piecewise affine interpolation of sparse data," in 2009 16th IEEE International Conference on Image Processing (ICIP). IEEE, November 2009, pp. 365-368.
- G. Facciolo, P. Arias, V. Caselles, and G. Sapiro, "Exemplar-based interpolation of sparsely sampled images," in Energy Minimization Methods in Computer Vision and Pattern Recognition, D. Cremers, Y. Boykov, A. Blake, and F. R. Schmidt, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, vol. 5681, ch. 25, pp. 331-344.

Bibliography

[AB94] M. Amar and G. Belletini. A notion of total variation depending on a metric with discontinuous coefficients. *Annales de l'Institut Henri Poincaré. Analyse non linéaire*, 11(1):91–133, 1994.

[ABCH08] A. Almansa, C. Ballester, V. Caselles, and G. Haro. A tv based restoration model with local constraints. *Journal of Scientific Computing*, 34(3):209–236–236, March 2008.

[ABGvG10] A. Almansa, E. Bughin, and R. Grompone von Gioi. Fast plane detection in disparity maps. In *2010 17th IEEE International Conference on Image Processing (ICIP)*, 2010.

[AC09] F. Alter and V. Caselles. Uniqueness of the cheeger set of a convex body. *Nonlinear Analysis: Theory, Methods & Applications*, 70(1):32–44, January 2009.

[ACC05a] F. Alter, V. Caselles, and A. Chambolle. A characterization of convex calibrable sets in r^n. *Mathematische Annalen*, 332(2):329–366, June 2005.

[ACC05b] F. Alter, V. Caselles, and A. Chambolle. Evolution of characteristic functions of convex sets in the plane by the minimizing total variation flow. *Interfaces Free Bound*, 7(1):29–53, 2005.

[ACD10] A. Almansa, J. Caron, and S. Durand. Deblurring of irregularly sampled images by tv regularization in a spline space. In *2010 17th IEEE International Conference on Image Processing (ICIP)*, pages 1181–1184, September 2010.

[ACF07] A. Almansa, V. Caselles, and G. Facciolo. Echantillonnage irrégulier et fusion multivues, rapport sur : Stb echantillonnage irrégulier, restitution du 3d par stéréoscopie à très faible b/h. Technical report, CMLA, ENS Cachan, 2007.

[ACHR06] A. Almansa, V. Caselles, G. Haro, and B. Rougé. Restoration and zoom of irregularly sampled, blurred, and noisy images by accurate total variation minimization with local constraints. *Multiscale Modeling & Simulation*, 5(1):235–272, 2006.

[ACMM01] 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 European Mathematical Society*, 3:39–92, 2001. 10.1007/PL00011302.

[ACS09] P. Arias, V. Caselles, and G. Sapiro. A variational framework for non-local image inpainting. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 345–358. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[ADR04] A. Almansa, S. Durand, and B. Rouge. Measuring and improving image resolution by adaptation of the reciprocal cell. *Journal of Mathematical Imaging and Vision*, 21(3):235+, November 2004.

[AEB06] M. Aharon, M. Elad, and A. Bruckstein. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representatio. *IEEE Trans. on Signal Processing*, 54(11):4311–22, 2006.

[AFCS10] P. Arias, G. Facciolo, V. Caselles, and G. Sapiro. A variational framework for exemplar-based image inpainting. *Int. J. Comput. Vision*, 2010. Accepted for publication.

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems (Oxford Mathematical Monographs)*. Oxford University Press, USA, May 2000.

[AG06] J.-F. Aujol and G. Gilboa. Constrained and snr-based solutions for tv-hilbert space image denoising. *J. Math. Imaging Vis.*, 26(1-2):217–237, 2006.

[AK06] G. Aubert and P. Kornprobst. *Mathematical problems in image processing: partial differential equations and the calculus of variations*. Springer-Verlag New York Inc, 2006.

[Alm02] A. Almansa. *Échantillonnage, interpolation et détection. Applications en imagerie satellitaire*. PhD thesis, Ph. D. thesis, ENS de Cachan, 2002.

[ALM10] J.-F. Aujol, S. Ladjal, and S. Masnou. Exemplar-based inpainting from a variational point of view. *SIAM J. Math. Anal.*, 42(3):1246–85, 2010.

[Alv02] L. Alvarez. Dense disparity map estimation respecting image discontinuities: A pde and scale-space based approach. *Journal of Visual Communication and Image Representation*, 13(1-2):3–21, March 2002.

[AM01] P. Azmi and F. Marvasti. Comparison between several iterative methods of recovering signals from nonuniformly spaced samples. In *in Proc. IEEE Conference on Sampling Theory and Applications SampTA2001*, volume 1, pages 207–224, Florida-USA, May 2001. Sampl. Theory Signal Image Process.

[ASHU05a] M. Arigovindan, M. Suhling, P. Hunziker, and M. Unser. Variational image reconstruction from arbitrarily spaced samples: a fast multiresolution spline solution. *IEEE Trans. on IP*, 14(4):450–460, 2005.

[ASHU05b] M. Arigovindan, M. Sühling, P. Hunziker, and M. Unser. Variational Image Reconstruction from Arbitrarily Spaced Samples: A Fast Multiresolution Spline Solution. *IEEE Transactions on Image Processing*, 14(4):450–460, 2005.

[AV94] R. Acar and C. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse problems*, 10:1217, 1994.

[AVCM04] F. Andreu-Vaillo, V. Caselles, and J. Mazón. *Parabolic quasilinear equations minimizing linear growth functionals*. Birkhäuser Verlag, 2004.

[AW06] S. Awate and R. Whitaker. Unsupervised, information-theoretic, adaptive image filtering for image restoration. *IEEE Trans. on PAMI*, 28(3):364–376, 2006.

[BBBW08] Z. Belhachmi, D. Bucur, B. Burgeth, and J. Weickert. How to choose interpolation data in images. Tech. Rep. No. 205, Saarland University, Saarbrücken, Germany, 2008.

[BBC$^+$01] C. Ballester, M. Bertalmío, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. *IEEE Trans. on IP*, 10(8):1200–11, 2001.

[BBF00] G. Bellettini, G. Bouchitté, and I. Fragalà. Bv functions with respect to a measure and relaxation of metric integral functionals. *J. Convex Anal.*, 6(2):349–366, 2000.

[BBH03] M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):993–1008, 2003.

[BBS01] M. Bertalmío, A. Bertozzi, and G. Sapiro. Navier-stokes, fluid-dynamics and image and video inpainting. In *Proc. of the IEEE Conf. on CVPR*, 2001.

[BCBV06] C. Ballester, V. Caselles, R. Bernard, and J. Verdera. A variational model for p+xs image fusion. *Int. J. Comp. Vision*, 69:143–158, 2006.

[BCC07] G. Buttazzo, G. Carlier, and M. Comte. On the selection of maximal cheeger sets. *Differential and integral equations*, 20(9):991–1004, 2007.

[BCM05] A. Buades, B. Coll, and J. Morel. A non local algorithm for image denoising. In *Proc. of the IEEE Conf. on CVPR*, volume 2, pages 60–65, 2005.

[BCMS09] A. Buades, B. Coll, J. M. Morel, and C. Sbert. Self-similarity driven color demosaicking. *Image Processing, IEEE Transactions on*, 18(6):1192–1202, 2009.

[BCRS03] M. Bertalmío, V. Caselles, B. Rougé, and A. Solé. TV based image restoration with local constraints. *Journal of Scientific Computing*, 19(1-3):95–122, 2003.

[BEV+07] X. Bresson, S. Esedoḡlu, P. Vandergheynst, J.-P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, June 2007.

[BFCAB95] L. Blanc-Feraud, P. Charbonnier, G. Aubert, and M. Barlaud. Nonlinear image processing: modeling and fast algorithm for regularization with edge detection. In *Proceedings of the 1995 International Conference on Image Processing (Vol. 1)-Volume 1-Volume 1*, volume 1, pages 474 –477. IEEE Computer Society, oct. 1995.

[BGGP07] T. Bekaert, S. Gautama, R. Goossens, and W. Philips. Dense and reliable stereo matching in urban areas using adaptive windows. In *Proceedings of 18th Annual Workshop on Circuits, Systems and Signal Processing ProRISC 2007*, 2007.

[BK03] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *Proceedings Ninth IEEE International Conference on Computer Vision*, pages 26–33 vol.1. IEEE, 2003.

[BKC08] T. Brox, O. Kleinschmidt, and D. Cremers. Efficient nonlocal means for denoising of textural patterns. *IEEE Trans. on IP*, 17(7):1057–92, 2008.

[BLLC02] R. Bornard, E. Lecan, L. Laborelli, and J.-H. Chenot. Missing data correction in still images and image sequences. In *Proc. ACM Int. Conf. on Multimedia*, 2002.

[BM07] F. Bornemann and T. März. Fast image inpainting based on coherence transport. *J. of Math. Imag. and Vis.*, 28(3):259–78, 2007.

[BS07] X. Bai and G. Sapiro. A geodesic framework for fast interactive image and video segmentation and matting. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8, 2007.

[BSCB00] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proc. of SIGGRAPH*, 2000.

[BSFG09] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. PatchMatch: a randomized correspondence algorithm for structural image editing. In *Proc. of SIGGRAPH*, pages 1–11, New York, NY, USA, 2009. ACM.

[BSL+08] A. Bartesaghi, P. Sprechmann, J. Liu, G. Randall, G. Sapiro, and S. Subramaniam. Classification and 3d averaging with missing wedge correction in biological electron tomography? *Journal of Structural Biology*, 162(3):436–450, June 2008.

[BVSO03] M. Bertalmío, L. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture inpainting. *IEEE Trans. on Image Processing*, 12(8):882–89, 2003.

[BVZ01] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.

[CBFAB97] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *Image Processing, IEEE Transactions on*, 6(2):298 –311, feb. 1997.

[CCCD93] V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31, 1993.

[CCN] V. Caselles, A. Chambolle, and M. Novaga. Generic uniqueness of the cheeger set. preprint, 2008.

[CCN07] V. Caselles, A. Chambolle, and M. Novaga. Uniqueness of the Cheeger set of a convex body. *Pacific J. Math.*, 232(1):77–90, 2007.

[CCN09] V. Caselles, A. Chambolle, and M. Novaga. Some remarks on uniqueness and regularity of cheeger sets. *Rend. Sem. Mat. Univ. Padova*, 2009.

[CCP09] G. Carlier, M. Comte, and G. Peyré. Approximation of maximal cheeger sets by projection. *Mathematical Modelling and Numerical Analysis*, 43(1):139–150, 2009.

[CFH91] C. Cenker, H. Feichtinger, and M. Herrmann. Iterative algorithms in irregular sampling: a first comparison of methods. In *Computers and Communications, 1991. Conference Proceedings., Tenth Annual International Phoenix Conference on*, pages 483 –489. IEEE, mar. 1991.

[CFM09] V. Caselles, G. Facciolo, and E. Meinhardt. Anisotropic cheeger sets and applications. *SIAM Journal on Imaging Sciences*, 2(4):1211–1254, 2009.

[CGMP09] F. Cao, Y. Gousseau, S. Masnou, and P. Pérez. Geometrically guided exemplar-based inpainting. HAL, 2009.

[Cha04] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, 2004.

[Che95] Y. Cheng. Mean shift, mode seeking and clustering. *IEEE Trans. on PAMI*, 17(8):790–99, 1995.

[Cia98] P. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Dunod, Paris, 1998.

[CJDS03] N. Chehata, F. Jung, M. P. Deseilligny, and G. Stamon. A region-based matching approach for 3d-roof reconstruction from hr satellite stereo pairs. In *DICTA*, pages 889–898, 2003.

[CKS95] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *Computer Vision, IEEE International Conference on*, 0:694+, 1995.

[CKS97] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, February 1997.

[CKS02] T. Chan, S. H. Kang, and J. H. Shen. Euler's elastica and curvature based inpaintings. *SIAM J. App. Math.*, 63(2):564–92, 2002.

[CKSS97] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. Minimal surfaces: A geometric three dimensional segmentation approach. *Numerische Mathematik*, 77(4):423–451, 1997.

[CL97] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188–188, April 1997.

[CM99] T. F. Chan and P. Mulet. On the convergence of the lagged diffusivity fixed point method in total variation image restoration. *SIAM J. Numer. Anal.*, 36(2):354–367, 1999.

[CM09] V. Caselles and P. Monasse. *Geometric description of images as topographic maps*. Springer-Verlag New York Inc, 2009.

[Coh91] L. D. Cohen. On active contour models and balloons. *CVGIP: Image Underst.*, 53(2):211–218, 1991.

[CPn06] M. Carreira Perpiñan. Fast nonparametric clustering with Gaussian blurring mean-shift. In *Proceedings of the 23rd international conference on Machine learning*, ICML '06, pages 153–160, New York, NY, USA, 2006. ACM.

[CPT04] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based inpainting. *IEEE Trans. on IP*, 13(9):1200–1212, 2004.

[CR08] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *CoRR*, abs/0805.4471, 2008.

[CS01] T. Chan and J. H. Shen. Mathematical models for local nontexture inpaintings. *SIAM J. App. Math.*, 62(3):1019–43, 2001.

[CV01] T. F. Chan and L. A. Vese. Active contours without edges. *Image Processing, IEEE Transactions on*, 10(2):266–277, 2001.

[CW08] E. J. Candes and M. B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[CZC95] T. F. Chan, H. M. Zhou, and R. H. Chan. Continuation method for total variation denoising problems. Technical Report 95-18, University of California, Los Angeles, 1995.

[DB03] G. Dahlquist and A. Björck. *Numerical Methods*. Dover Publications, Incorporated, 2003.

[DCOY03] I. Drori, D. Cohen-Or, and H. Yeshurun. Fragment-based image completion. *ACM Trans. on Graphics. Special issue: Proc. of ACM SIGGRAPH*, 22(3):303–12, 2003.

[Del04] J. Delon. *Fine comparison of images and other problems.* PhD thesis, École Normale Supérieure de Cachan, 2004.

[DMM01] A. Desolneux, L. Moisan, and J.-M. Morel. Edge detection by helmholtz principle. *Journal of Mathematical Imaging and Vision*, 14(3):271–284–284, May 2001.

[DMR00] S. Durand, F. Malgouyres, and B. Rougé. Image deblurring, spectrum interpolation and application to satellite imaging. *Control, Optimisation and Calculus of Variations*, 5:445–475, 2000.

[DR07] J. Delon and B. Rougé. Small baseline stereovision. *Journal of Mathematical Imaging and Vision*, 28(3), 2007.

[DSC03] L. Demanet, B. Song, and T. Chan. Image inpainting by correspondence maps: a deterministic approach. Technical report, UCLA, 2003.

[DV97] D. Dobson and C. Vogel. Convergence of an iterative method for total variation denoising. *SIAM Journal on Numerical Analysis*, 34(5):1779–1791, 1997.

[EG92] L. Evans and R. Gariepy. *Measure theory and fine properties of functions*. CRC Press, Boca Raton, 1992.

[EL02] A. Efros and T. Leung. Texture synthesis by non-parametric sampling. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1033–1038. IEEE, 2002.

[ES02] S. Esedoglu and J. H. Shen. Digital image inpainting by the Mumford-Shah-Euler image model. *European J. App. Math.*, 13:353–70, 2002.

[ESQD05] M. Elad, J. Starck, P. Querre, and D. Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *Applied and Computational Harmonic Analysis*, 19(3):340–358, 2005.

[FAAC09] G. Facciolo, A. Almansa, J. F. Aujol, and V. Caselles. Irregular to Regular Sampling, Denoising, and Deconvolution. *Multiscale Modeling & Simulation*, 7(4):1574–1608, 2009.

[FAP05] G. Facciolo, A. Almansa, and A. Pardo. Variational approach to interpolate and correct biases in stereo correlation. In *20eme Colloque sur le traitement du signal et des images*, volume 2, pages 513–516, Louvain-La-Neuve, Belgium, sept 2005. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images.

[Fau88] P. Faurre. *Analyse numérique. Notes d'optimisation*. École Polytechnique, 1988. Ed. Ellipses.

[FB81] M. Fischler and R. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6), 1981.

[FH75] K. Fukunaga and L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 21(1):32–40, January 1975.

[FK98] O. Faugeras and R. Keriven. Variational principles, surface evolution, pde's, level set methods, and the stereo problem. *Image Processing, IEEE Transactions on*, 7:336–344, 1998.

[FKO+04] S. Forstmann, Y. Kanou, J. Ohya, S. Thuering, and A. Schmitt. Real-time stereo by using dynamic programming. In *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04. Conference on*, 2004.

[FL09] C.-W. Fang and J.-J. J. Lien. Rapid image completion system using multiresolution patch-based directional and nondirectional approaches. *IEEE Trans. on IP*, 18(12):2769–2779, 2009.

[FLA+06] G. Facciolo, F. Lecumberry, A. Almansa, A. Pardo, V. Caselles, and B. Rougé. Constrained anisotropic diffusion and some applications. In *British Machine Vision Conference*, volume 3, pages 1049–1058, Edinburgh, Scotland, September 2006.

[FRT97] A. Fusiello, V. Roberto, and E. Trucco. Efficient stereo with multiple windowing. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 0, pages 858–863, Los Alamitos, CA, USA, 1997. IEEE Comput. Soc.

[GKC] L. Garrido, M. Kalmoun, and V. Caselles. Contrast invariant approach for dense motion estimation using the mg/opt algorithm. In preparation 2010.

[GM01] Y. Gousseau and J. M. Morel. Are natural images of bounded variation? *SIAM Journal on Mathematical Analysis*, 33(3):634–648, 2001.

[GO07a] G. Gilboa and S. Osher. Nonlocal linear image regularization and supervised segmentation. *SIAM Mult. Mod. and Sim.*, 6(2):595–630, 2007.

[GO07b] G. Gilboa and S. Osher. Nonlocal operators with application to image processing. Technical report, UCLA, 2007.

[GR92] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(3):367–383, 1992.

[GS00] K. Gröchenig and T. Strohmer. Numerical and theoretical aspects of non-uniform sampling of band-limited images. In F. Marvasti, editor, *Theory and Practice of Nonuniform Sampling*. Kluwer/Plenum, 2000.

[GSZ03] G. Gilboa, N. Sochen, and Y. Y. Zeevi. Texture preserving variational denoising using an adaptive fidelity term. In *VLSM 2003*, pages 137–144, Nice, France, 2003.

[GSZ06] G. Gilboa, N. Sochen, and Y. Zeevi. Variational denoising of partly textured images by spatially varying constraints. *Image Processing, IEEE Transactions on*, 15(8):2281 –2289, aug. 2006.

[GW90] C. Gasquet and P. Witomski. *Analyse de Fourier et applications: filtrage, calcul numérique, ondelettes*. Masson, 1990.

[GWW+08] I. Galić, J. Weickert, M. Welk, A. Bruhn, A. Belyaev, and H.-P. Seidel. Image compression with anisotropic diffusion. *Journal of Mathematical Imaging and Vision*, 31(2):255–269, July 2008.

[HE07] J. Hays and A. Efros. Scene completion using millions of photographs. In *SIGGRAPH*, New York, NY, USA, 2007. ACM.

[HGY08] M. Holtzman-Gazit and I. Yavneh. A scale-consistent approach to image completion. *Int. J. Multiscale Comput. Eng.*, 6(6):617–628, 2008.

[HML94] O. Hagolle, J. M. Martinuzzi, and C. Latry. How to double the spatial resolution of a push-broom instrument. In *Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, volume 3, pages 1553–1555 vol.3, 1994.

[IP97] H. Igehy and L. Pereira. Image replacement through texture synthesis. In *Proc. of the IEEE Conf. on CVPR*, October 1997.

[IPG+07] L. Igual, J. Preciozzi, L. Garrido, A. Almansa, V. Caselles, and B. Rougé. Automatic low baseline stereo in urban areas. *Inverse Problems and Imaging*, 1(2), 2007.

[Jay57] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–30, 1957.

[JBFZ02] A. Jalobeanu, L. Blanc-Feraud, and J. Zerubia. Estimation of blur and noise parameters in remote sensing. In *IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS SPEECH AND SIGNAL PROCESSING*, volume 4, pages IV–3580 – IV–3583 vol.4, 2002.

[JT04] J. Jia and C.-K. Tang. Inference of segmented color and texture description by tensor voting. *IEEE Trans. on PAMI*, 26(6):771–86, 2004.

[Kad64] M. Kadec. The exact value of the paley-wiener constant. *Sov.Math.Doklady*, 5:559–561, 1964.

[KB05] V. Kolmogorov and Y. Boykov. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 1, pages 564–571. IEEE, 2005.

[KB06] C. Kervrann and J. Boulanger. Optimal spatial adaptation for patch-based image denoising. *IEEE Trans. on IP*, 15(10):2866–2878, 2006.

[KEBK05] V. Kwatra, I. Essa, A. Bobick, and N. Kwatra. Texture optimization for example-based synthesis. *ACM Trans. Graph.*, 24(3):795–802, 2005.

[KKO+96] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Conformal curvature flows: from phase transitions to active vision. *Archive for Rational Mechanics and Analysis*, 134(3):275–301, 1996.

[KLR06] B. Kawohl and T. Lachand Robert. Characterization of Cheeger sets for convex subsets of the plane. *Pacific journal of mathematics*, 225(1):103, 2006.

[KO94] T. Kanade and M. Okutomi. A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9):920–932, 1994.

[KP09] K. Karantzalos and N. Paragios. Variational model-based 3d building extraction from remote sensing data. In *2009 16th IEEE International Conference on Image Processing (ICIP)*, pages 545–548. IEEE, November 2009.

[KSC01] S. B. Kang, R. Szeliski, and J. Chai. Handling occlusions in dense multi-view stereo. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, pages I–103–I–110. IEEE Comput. Soc, 2001.

[KSK06] A. Klaus, M. Sormann, and K. Karner. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *18th International Conference on Pattern Recognition (ICPR'06)*, pages 15–18. IEEE, 2006.

[KSY09] N. Kawai, T. Sato, and N. Yokoya. Image inpainting considering brightness change and spatial locality of textures and its evaluation. In T. Wada, F. Huang, and S. Lin, editors, *Advances in Image and Video Technology*, volume 5414 of *Lecture Notes in Computer Science*, chapter 24, pages 271–282. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[KT07] N. Komodakis and G. Tziritas. Image completion using efficient belief propagation via priority scheduling and dynamic pruning. *IEEE Trans. on IP*, 16(11):2649–61, 2007.

[KZ06] V. Kolmogorov and R. Zabih. Graph cut algorithms for binocular stereo with occlusions. *Handbook of Mathematical Models in Computer Vision*, pages 423–437, 2006.

[Lan67] H. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Mathematica*, 117(1):37–52, July 1967.

[LC87] W. Lorensen and H. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In *Proceedings of the 14th annual conference on Computer graphics and nteractive techniques*, pages 163–169. ACM, 1987.

[LDZP08] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrotdeseilligny. Automatic building extraction from dems using an object approach and application to the 3d-city modeling. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(3):365–381, May 2008.

[LEB07] O. Lezoray, A. Elmoataz, and S. Bougleux. Graph regularization for color image processing. *Comput. Vis. Image Underst.*, 107(1-2):38–55, 2007.

[Lec05] F. Lecumberry. Cálculo de disparidad y segmentación de objetos en secuencias de video. Master's thesis, Facultad de Ingenieria, Universidad de la Republica, Uruguay, 2005.

[LLW04] A. Levin, D. Lischinski, and Y. Weiss. Colorization using optimization. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers*, pages 689–694, New York, NY, USA, 2004. ACM.

[LM04] S. Lintner and F. Malgouyres. Solving a variational image restoration model which involves l^∞ constraints. *Inverse Problems*, 20(3):815–831, June 2004.

[LM08] C. Louchet and L. Moisan. Total Variation denoising using posterior expectation. In *Proceedings of the 16th European Signal Processing Conference (EUSIPCO)*, August 2008.

[LR00] C. Latry and B. Rouge. Optimized sampling for ccd instruments: the supermode scheme. *Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International*, 5:2322–2324 vol.5, 2000.

[LS04] M. Lassas and S. Siltanen. Can one use total variation prior for edge-preserving bayesian inversion? *Inverse Problems*, 20(5):1537+, October 2004.

[LZW03] A. Levin, A. Zomet, and Y. Weiss. Learning how to inpaint from global image statistics. In *Proc. of ICCV*, 2003.

[Mal00] F. Malgouyres. *Increase in the Resolution of Digital Images: Variational Theory and Applications*. PhD thesis, Ph. D. thesis, Ecole Normale Superieure de Cachan, 2000, Cachan, France, 2000.

[Mas02] S. Masnou. Disocclusion: a variational approach using level lines. *IEEE Trans. on IP*, 11(2):68–76, 2002.

[MBP+09] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. Submitted, 2009.

[Mei06] E. Meinhardt. The tree of shapes of a 3d image. Master's thesis, Universitat Pompeu Fabra, Barcelona, 2006.

[MG00] P. Monasse and F. Guichard. Fast computation of a contrast-invariant image representation. *Image Processing, IEEE Transactions on*, 9(5):860–872, 2000.

[MG02] F. Malgouyres and F. Guichard. Edge direction preserving image zooming: A mathematical and numerical analysis. *SIAM Journal on Numerical Analysis*, 39(1), 2002.

[MM98] S. Masnou and J.-M. Morel. Level lines based disocclusion. In *Proc. of IEEE ICIP*, 1998.

[Moi07] L. Moisan. How to discretize the total variation of an image? *PAMM*, 7(1):1041907–1041908, 2007.

[MPAT08] J. Melonakos, E. Pichon, S. Angenent, and A. Tannenbaum. Finsler active contours. *IEEE T Pattern Anal*, 30(3):412–423, 2008.

[MS95] J.-M. Morel and S. Solimini. *Variational methods in image segmentation*. Birkhauser Boston Inc., Cambridge, USA, 1995.

[MS05] A. Mohan and G. Sapiro. Constrained regularization of digital terrain elevation data. IMA Preprint Series # 2049, UMN, May 2005.

[MSE08] J. Mairal, G. Sapiro, and M. Elad. Learning multiscale sparse representations for image and video restoration. *SIAM MMS*, 7(1):214–241, 2008.

[MY08] J. M. Morel and G. Yu. On the consistency of the SIFT Method. *Preprint, CMLA*, 26, 2008.

[MZFC08] E. Meinhardt, E. Zacur, A. Frangi, and V. Caselles. 3d edge detection by selection of level surface patches. *Journal of Mathematical Imaging and Vision*, October 2008.

[NEC06] M. Nikolova, S. Esedoḡlu, and T. F. Chan. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal on Applied Mathematics*, 66(5):1632–1648, 2006.

[Nik07] M. Nikolova. Model distortions in bayesian map reconstruction. *Inverse Problems and Imaging*, 1(2):399–422, May 2007.

[PBC08] G. Peyré, S. Bougleux, and L. Cohen. Non-local regularization of inverse problems. In *ECCV '08*, pages 57–68, Berlin, Heidelberg, 2008. Springer-Verlag.

[PBC09] G. Peyré, S. Bougleux, and L. D. Cohen. Non-local regularization of inverse problems. *Preprint Hal-00419791*, 2009.

[PELS10] P. Panareda Busto, C. Eisenacher, S. Lefebvre, and M. Stamminger. Instant Texture Synthesis by Numbers. *Vision, Modeling & Visualization 2010*, pages 81–85, 2010.

[PETM09] M. Protter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the non-local-means to super-resolution reconstruction. *IEEE Trans. on IP*, 18(1):36–51, 2009.

[PGB03] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. In *Proc. of SIGGRAPH*, 2003.

[PGB04] P. Pérez, M. Gangnet, and A. Blake. PatchWorks: Example-based region tiling for image editing. Technical report, Microsoft Research, 2004.

[PMD+10] L. Pizarro, P. Mrázek, S. Didas, S. Grewenig, and J. Weickert. Generalised nonlocal image smoothing. *IJCV*, 90:62–87, 2010.

[PS01] A. Pardo and G. Sapiro. Vector probability diffusion. *IEEE Signal Processing Letters*, 8(4):106–109, April 2001.

[PSGB08] T. Pock, T. Schoenemann, G. Graber, and D. Bischof, Horst abd Cremers. A convex formulation of continuous multi-label problems. *Computer Vision–ECCV 2008*, pages 792–805, 2008.

[PST01] D. Potts, G. Steidl, and M. Tasche. *Fast Fourier transforms for nonequispaced data: A tutorial*, chapter 12, pages 247–269. Birkhäuser, 2001.

[Rau98] M. Rauth. *Gridding of geophysical potential fields from noisy scattered data*. PhD thesis, University of Vienna, 1998.

[RO94] L. I. Rudin and S. Osher. Total variation based image restoration with free local constraints. In *Proceedings of 1st International Conference on Image Processing*, volume 1, pages 31–35. IEEE Comput. Soc. Press, 1994.

[ROF92] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica*, D(60):259–268, 1992.

[SAM08] N. Sabater, A. Almansa, and J.-M. Morel. Rejecting wrong matches in stereovision. CMLA Preprint 2008-28, ENS Cachan, CNRS, Cachan, France., 2008.

[SB97] R. Sára and R. Bajcsy. On occluding contour artifacts in stereo vision. In *CVPR*, pages 852–857, 1997.

[Sch07] J. Schott. *Remote sensing: the image chain approach*. Oxford University Press, USA, 2007.

[SGG+08] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Applied Mathematical Sciences, vol. 167. Springer, 1 edition, October 2008.

[SGT02] E. Sifakis, C. Garcia, and G. Tziritas. Bayesian Level Sets for Image Segmentation. *Journal of Visual Communication and Image Representation*, 13(1-2):44–64, 2002.

[She68a] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proc. of 23rd ACM national conf.*, pages 517–524, New York, NY, USA, 1968. ACM.

[She68b] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM national conference*, New York, NY, USA, 1968. ACM.

[Sho85] N. Shor. *Minimization Methods for Non-differentiable Functions*. Springer, 1985.

[SJZ05] J. Shen, X. Jin, and C. Zhou. Gradient based image completion by solving poisson equation. *Advances in Mulitmedia Information Processing-PCM 2005*, pages 257–268, 2005.

[SM02] J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2002.

[SMA10] N. Sabater, J.-M. Morel, and A. Almansa. Block matching reaches theoretical accuracy bounds. *CMLA Preprint*, 2010.

[SS02] D. Scharstein and R. Szeliski. A taxomomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42, April 2002.

[SYJS05] J. Sun, L. Yuan, J. Jia, and H. Y. Shum. Image completion with structure propagation. In *Proc. of SIGGRAPH*, 2005.

[SZ00] C. Schmid and A. Zisserman. The geometry and matching of lines and curves over multiple views. *Int. J. Comput. Vision*, 40(3):199–233, December 2000.

[SZS03] J. Sun, N. N. Zheng, and H. Y. Shum. Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7):787–800, July 2003.

[TA77] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977.

[TD05] D. Tschumperlé and R. Deriche. Vector-valued image regularization with PDE's: a common framework for different applications. *IEEE Trans. on PAMI*, 27(4), 2005.

[Tru09] N. Trudinger. On the analyticity of generalized minimal surfaces. *Bulletin of the Australian Mathematical Society*, 5(03):315–320, 2009.

[Tsc02] D. Tschumperlé. *PDE-Based Regularization of Multivalued Images and Applications*. PhD thesis, University of Nice-Sophia Antipolis, December 2002.

[TZL+02] X. Tong, J. Zhang, L. Liu, X. Wang, B. Guo, and H. Y. Shum. Synthesis of bidirectional texture functions on arbitrary surfaces. *ACM Trans. Graph.*, 21:665–672, July 2002.

[Van06] P. Vandewalle. *Super-resolution from unregistered aliased images*. PhD thesis, École polytechnique Fédérale de lausanne, 2006.

[vGJMR10] R. G. von Gioi, J. Jakubowicz, J. M. Morel, and G. Randall. LSD: A Fast Line Segment Detector with a False Detection Control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):722–732, April 2010.

[VO96] C. Vogel and M. Oman. Iterative methods for total variation denoising. *SIAM Journal on Scientific Computing*, 17(1):227–238, 1996.

[Wei98] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner-Verlag, Stuttgart, Germany, 1998.

[WL00] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. In *Proc. of the SIGGRAPH*, 2000.

[WSI07] Y. Wexler, E. Shechtman, and M. Irani. Space-time completion of video. *IEEE Trans. on PAMI*, 29(3):463–476, 2007.

[WSW06] M. Welk, G. Steidl, and J. Weickert. Locally analytic schemes: A link between diffusion filtering and wavelet shrinkage. Technical Report No. 2100, IMA, University of Minnesota, Minneapolis, USA, February 2006.

[YS06] L. Yatziv and G. Sapiro. Fast image and video colorization using chrominance blending. *Image Processing, IEEE Transactions on*, 15(5):1120–1129, 2006.

[ZS05] D. Zhou and B. Schölkopf. Regularization on discrete spaces. In *Proceedings of the 27th DAGM Symposium*, pages 361–368. Springer, 2005.