Exemplar-Based Interpolation of Sparsely Sampled Images

Gabriele Facciolo[†], Pablo Arias[†], Vicent Caselles[†], and Guillermo Sapiro[‡]

†Universitat Pompeu Fabra, DTIC, 08018 Barcelona, Spain
‡University of Minnesota, ECE, Minneapolis, MN 55455, USA
`{gabriele.facciolo,pablo.arias,vicent.caselles}@upf.edu,guille@umn.edu`

Abstract. A nonlocal variational formulation for interpolating a sparsely sampled image is introduced in this paper. The proposed variational formulation, originally motivated by image inpainting problems, encourages the transfer of information between similar image patches, following the paradigm of exemplar-based methods. Contrary to the classical inpainting problem, no complete patches are available from the sparse image samples, and the patch similarity criterion has to be redefined as here proposed. Initial experimental results with the proposed framework, at very low sampling densities, are very encouraging. We also explore some departures from the variational setting, showing a remarkable ability to recover textures at low sampling densities.

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. This is an ill-posed inverse problem, and as such, some sort of prior knowledge is needed for its solution. The literature on this topic is vast, since it lies in the heart of many relevant applications, such as zooming, demosaicing, super-resolution and image editing, among others.

For the purpose of this paper we distinguish two interpolation cases: when the available data consists of a set of isolated samples (be regular or irregular) and when it is given on a (not necessarily connected) region of the image. For the former we will use the term *interpolation*, reserving *inpainting* to denote the dense case.

In the case of inpaiting the available information usually allows to determine the image derivatives on the region with known data. First approaches to inpainting took advantage of this, completing the image by means of PDEs [1, 2] or variational methods [3] that continued the image gradients or the level lines inside the inpainting domain. These schemes involving only interactions between *local* pixels, fail with textured images or large inpainting domains. Advances in the field of texture synthesis [4] served as inspiration for new inpainting strategies, based on the hypothesis that natural images are redundant, and self similar: The value of a pixel is synthesized from known pixels with similar neighborhoods (*patches*). These methods are often refereed to as *non-local* or *exemplar-based*

(see for instance [5–7] and references therein). A current trend in research is the combination of both, local and non-local strategies *e.g.* [8, 9]. We refer to [9] for an account of this active area of research.

If the only available data consists of a nonuniform and sparse (as opposed to dense) set of samples then: 1. The gradients as well as the directions of the level lines are unknown, 2. There are no complete patches available on the image. In this setting PDE based methods cannot be directly applied and exemplar-based inpainting methods need to be adapted. This scenario appears in image super resolution, since after registering the low resolution images the overlapped grids may be seen as a non regular one.

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 [10], of bounded variation [11], expanded over a base of functions (*e.g.* splines [12], radial basis functions [13]), among others.

A recent front of activity is given by the techniques based on the *sparseland* model [14, 15], in which the image is restricted to have a sparse representation over an overcomplete basis or *dictionary* [16, 15, 17]. The main difference between dictionary-based and exemplar-based methods lies in 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 lies elsewhere in the image itself (or in a database of images [18]).

A non-local prior is used in [19]. In this work 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 [20, 21]. This corresponds to a non-local regularization on the image. A successful PDE approach using an anisotropic diffusion process was proposed in [22].

Our contribution. We address the problem of image interpolation from non-uniformly sparsely sampled data via a non-local exemplar-based variational approach that exploits the self-similarity of the image. In this approach, and just to prove the applicability of the self-similarity principle, we consider the simple case where the samples are arranged on a discrete (but non regular) grid, and leave the sub-pixel case for future development. The proposed variational formulation is a generalization of the inpainting framework presented in [23], which exhibits a good performance, but only for dense inpainting domains. As in [23], we set up a functional to model the nonlocal means iterations both for the image and the weights. Thus, besides the data attachment term, we include a regularization term for the weights given in terms of its entropy. The functional is then minimized with respect to both variables, the unknown image and the weights. The data attachment term is tailored to compare only the known pixel positions in one or both patches under comparison. Finally, both terms are balanced by a temperature parameter h and letting $h \to 0+$ (as in [24]) permits to iteratively improve the results. Let us mention that we have also explored a non variational model suggested by our approach that exhibits a faster convergence. The pre-

liminary experiments suggest that exemplar-based methods can be successfully applied to sparse data interpolation.

Related work. Our work is related to the nonlocal techniques applied to demosaicing in [24, 25] and super-resolution in [26], problems that 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. Similar ideas can be also found in the field of 3D tomographic imaging [27], 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 interpolation from a set of sparsely sampled images could be approached with the techniques of compressed sensing [14, 28]. Even if the standard approach 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 address the problem of interpolating non uniformly sampled images with large unsampled regions.

Finally, the work [25] combines sparsity and non-local techniques. There, the image self-similarity is used to obtain more robust sparse representations over a given dictionary, by assigning a common representation to similar patches.

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, x', z, 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 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 portion 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.

2 From inpainting to interpolation

The framework we present here is an adaptation of the non-local inpainting functional recently introduced in [23]. In this section we briefly review this work and discuss the modifications that have to be done to allow its application to the problem of image interpolation from sparse samples addressed in this paper.

2.1 Review: Non-local functional for image inpainting.

In [23] we proposed the functional

$$\widetilde{E}(u, w) = \frac{1}{h} \widetilde{F}_w(u) - \sum_{x \in \widetilde{O}} \widetilde{H}_w(x) \tag{1}$$

whose minimization yields a non-local exemplar-based inpainting method. The first term is given by

$$\widetilde{F}_w(u) = \sum_{x \in \widetilde{O}} \sum_{x' \in \widetilde{O}^c} w(x, x') \|p_u(x) - p_u(x')\|_{\Omega_p}, \tag{2}$$

and it is inspired by a functional presented in [21] in the context of non-local image denoising/regularization. $\widetilde{F}_w$ measures the coherence between the patches in $\widetilde{O}$ and those in $\widetilde{O}^c$, given the similarity weight function w and a patch norm-like function $\|\cdot\|_{\Omega_p}$. $\widetilde{O}$ is an extension of O containing the centers of all patches intersecting O. In doing so, patches $p_u(x')$ centered in $x' \in \widetilde{O}^c$ consist entirely of known pixels. The term (2) promotes the similarity between the image patches centered at $x \in \widetilde{O}$ and $x' \in \widetilde{O}^c$. Indeed, minimizing $\widetilde{F}_w$ w.r.t. the image u, for a given fixed weight function w, forces pairs of patches for which $w(x, x')$ is high to be similar. Since $p_u(x')$ 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 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. *Gaussian* weights are commonly used, i.e. $w(x, x') = \exp\left(-\frac{1}{h}\|p_u(x) - p_u(x')\|^2\right)$, where $\|\cdot\|$ is a weighted L_2 norm in the space of patches and h is the scale. In the frameworks described in [21] the weights are known and remain fixed through all the iterations. While this might be appropriate in case of denoising applications, where the weights can be estimated from the noisy image, in the image inpainting/interpolation scenario, weights are not available and have to be inferred together with the image. This idea has been applied before for super-resolution [26], denoising [29] and in a more general regularization framework [19]. None of these works present a variational justification for the weight update.

This issue was addressed in [30, 23]. In [23] we consider the weight function w as an additional unknown. Instead of prescribing explicitly the Gaussian functional dependence of w w.r.t. u we do it implicitly, as a component of the optimization process. This results in a simpler functional, avoiding to deal with the complex, non-linear dependence between w and u. To this end, $w(x, \cdot)$ is constrained to be a probability density function, $\sum_{x' \in O^c} w(x, x') = 1$, and a second term given by $\sum_{x \in \widetilde{O}} \widetilde{H}_w(x)$ is added (the second term in (1)), where

$$\widetilde{H}_w(x) = - \sum_{x' \in \widetilde{O}^c} w(x, x') \log w(x, x'), \tag{3}$$

is the entropy of the probability $w(x, \cdot)$ for $x \in \widetilde{O}$. Summarizing, the first term of (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.

2.2 Generalization to interpolation.

We will discuss in this section the modifications needed to adapt the inpainting formalism to the problem of image interpolation. The mechanism for adapting

the similarity weight function remains unchanged, thus we will focus our attention on the image energy term. Let us assume for the moment that we know a weight function w which measures the similarity of the pairs of incomplete patches. We will detail later 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 transfer occurs between the available information and the patches inside the interpolation domain. In the interpolation application here addressed the image is known only at some isolated positions distributed through all the image. We can still have entire continuous regions of missing information (in contrast with typical approaches addressed in compressed sensing), but we do not have at all entire patches of *available* information. This does not allow the direct application of the inpainting energy (1) to the interpolation problem, since every image patch contains unknown pixels, and thus needs information from other patches. At the same time any patch may have information to transfer to potentially all other patches. This suggests that the summation domains in Eq. (2), as well as the patch comparison metric, have to be modified. We address this next.

For the sake of generality we will use generic summation domains and denote them by D_1 and D_2. For instance, the corresponding definitions for the inpainting functional (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. The weight function is thus defined over $D_1 \times D_2$ 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:

$$F(u, w) = \sum_{x \in D_1} \sum_{x' \in D_2} w(x, x') V_\varphi(p_u(x), p_u(x')). \qquad (4)$$

We have introduced a general *pair-wise patch similarity potential* V_φ, substituting the patch norm-like function $\| \cdot \|_{\Omega_p}$. Since we deal with sparsely sampled patches, the pair-wise patch potential V_φ is based only on the known pixels around x and x':

$$V_\varphi(p_u(x), p_u(x')) = \sum_{y \in \Omega_p} \frac{g_\sigma(y)}{\rho(x, x')} (\alpha \mathcal{X}_{O^c}(x+y) + \beta \mathcal{X}_{O^c}(x'+y)) \varphi(u(x+y) - u(x'+y))$$

$$(5)$$

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 $\varphi(r) = |r|^p$, $r \in \mathbb{R}$, $1 \leq p < \infty$ (a more general function could be considered). For instance, taking $p = 1$ leads to an algorithm based on medians (see [23]), here due to space limitations we will restrict us to the case $p = 2$. The constant parameters $\alpha, \beta \in \{0, 1\}$ are set by the user. They control whether known positions around x or 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 (5). This happens whether the corresponding locations
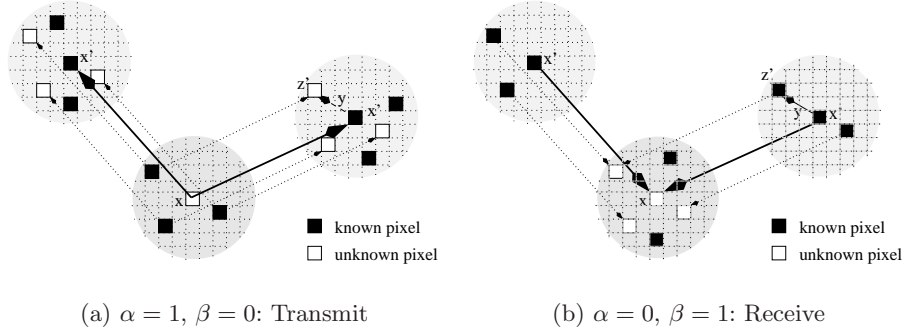
(a) $\alpha = 1$, $\beta = 0$: Transmit   (b) $\alpha = 0$, $\beta = 1$: Receive

Fig. 1. Visualization of transmission and reception processes due respectively to terms V_φ^α and V_φ^β in the patch similarity potential (see Eq. (6)).

around x' belong to the data set or not. If $\beta = 1$ the similarity potential accounts for the known pixels around x'. If both of them are 1, in which case V_φ is computed from the locations known in both patches. This last case coincides with the patch comparison criterion defined in [24] in the context of demosaicing.

The normalization factor $\rho(x, x')$ is such that $\sum_{y \in \Omega_p} \frac{g_\sigma(y)}{\rho(x,x')}(\alpha \mathcal{X}_{O^c}(x + y) + \beta \mathcal{X}_{O^c}(x' + y)) = 1$ for all $x \in D_1$, $x' \in D_2$. Considering the overlap between known positions in both patches (see for instance [27]) would also make sense for the 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_φ).

The proposed functional can be easily understood by splitting the pairwise patch potential into two terms $V_\varphi = V_\varphi^\alpha + V_\varphi^\beta$, with

$$V_\varphi^\alpha(p_u(x), p_u(x')) = \alpha \sum_{y \in \Omega_p} \frac{g_\sigma(y)}{\rho(x, x')} \mathcal{X}_{O^c}(x + y) \varphi(u(x + y) - u(x' + y)), \quad (6)$$

and analogously for V_φ^β. The energy F 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 $p_u(x')$, with $x' \in D_2$. Since known pixels are fixed, its minimization implies the modification of unknown pixels around x', thus transferring information from $p_u(x)$ to $p_u(x')$. On the other hand, V_φ^β considers differences between known pixels in $p_u(x')$ and the corresponding locations in $p_u(x)$. In this case known information flows from $p_u(x')$ 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_φ^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 $p_u(x') \in D_2$ (see Figure 1(b)). Whereas the minimization of the term with V_φ^β implies *receiving* known pixel values from peripheral patches at D_2 (see Figure 1(b)). We refer to these processes as *transmission* and *reception*.

The complete functional for the interpolation problem becomes:

$$E(u,w) = \frac{1}{h} F(u,w) - \sum_{x \in D_1} H_w(x), \tag{7}$$

where as before $H_w(x) = -\sum_{x' \in D_2} w(x,x') \log w(x,x')$ is again the entropy of the probability $w(x,\cdot)$ for $x \in D_2$. As in the case of inpainting, this term allows to model the estimation of the weights together with the image.

A similar functional for image super-resolution was considered in [26] without explicitly modeling the weight updating step. The functional in [26] is related to the case where $\alpha = 0$ and $\beta = 1$ (or to the case of a full patch comparison).

2.3   Reinterpretation of the image term F

Let us rewrite the energy term (4) in a different way, in which the image values appear directly, and not as part of patches. This formulation will be useful for posterior analysis. After the change of variables $z = x+y$, $z' = x+y'$, the energy can be rewritten by adding up the pair-wise pixels differences as

$$F(u,w) = \sum_{z \in \widetilde{\Omega}} \sum_{z' \in \widetilde{\Omega}} m(z,z')(\alpha \mathcal{X}_{O^c}(z) + \beta \mathcal{X}_{O^c}(z'))\varphi(u(z) - u(z')), \tag{8}$$

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 $m(z,z')$ as

$$m(z,z') = \sum_{y \in \Omega_p} \mathcal{X}_{D_1}(z-y)\mathcal{X}_{D_2}(z'-y)w(z-y,z'-y)\frac{g_\sigma(y)}{\rho(z-y,z'-y)}. \tag{9}$$

These weights integrate the similarities of patches centered at $z - y \in D_1$ containing z and those centered at $z' - y \in D_2$ containing z' for $y \in \Omega_p$.

The formulation given by Eq. (4) accumulates the pair-wise potentials for each pair of patches centered in D_1 and D_2. The potentials are given by the addition of pixel value differences. In (8), the energy is rewritten by explicitly computing the contribution of each pixel value difference. The characteristic functions $\mathcal{X}_{D_1}(z-y)$ and $\mathcal{X}_{D_2}(z'-y)$ in (8) are zero if neither z nor z' are known. Only those differences involving at least one known pixel are taken into account. It becomes clear that pixel differences for which we have a large value of $m(z,z')$ are penalized. This implies the modification of $u(z)$ or $u(z')$, depending on which of them is given and which is unknown. This shows again the difference with more frequently used patch distances, where only pixels available in both patches are considered for the computation. Certainly if such approaches are iterated, as sometimes done [27,6], pixels with originally only "one side" available start to influence the computation as well after the first iteration or the first time yet are "filled".

3 Minimization of E

We have formulated the interpolation problem as the constrained optimization

$$(u^*, w^*) = \arg\min_{u,w} E(u, w) \quad \text{subject to} \tag{10}$$

$$\sum_{x' \in D_2} w(x, x') = 1 \quad \text{for all} \quad x \in D_1. \tag{11}$$

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 yields the following iteration

1. [*Initial Condition*]   Given $u_0(x)$ with $x \in O$.
2. [*Weights Update Step*] $w_k = \arg\min_w E(u_k, w)$, subject to (11).
3. [*Image Update Step*] $u_{k+1} = \arg\min_u E(u, w_k)$.
4. [*Stopping Criterion*]   If $\|u_{k+1} - u_k\| > \tau$, go back to step 2.

In the weights updating step, the minimization of E w.r.t. w yields $w_k(x, x') = \frac{1}{q(x)} \exp\left[-\frac{1}{h} V_\varphi(p_{u_k}(x), p_u(x'))\right]$, where $q(x)$ is a normalization factor such that $\sum_{x' \in D_2} w(x, x') = 1$ for each patch $p_{u_k}(x)$. 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 $h \to \infty$, $w_k(x, \cdot)$ 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 $\lim_{h \to 0} w(x, x') = \frac{1}{\#n(x)} \mathcal{X}_{n(x)}(x')$, where $n(x) \subseteq O^c$ is the set of minimizers of $V_\varphi(p_u(x), \cdot)$, *i.e.* $n(x) = \{x' \in O^c : V_\varphi(p_u(x), p_u(x')) = V_{\min}(x)\}$, 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 image updating step deserves more attention and is described next.

3.1 Image updating step

The equilibrium equation for E results in

$$\sum_{z' \in O^c} (\alpha m(z', z) + \beta m(z, z')) \varphi'(u(z) - u(z')) = 0 \quad \text{for all } z \in O. \tag{12}$$

This equation specifies the information transferred from the datum $u(z')$ 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 z', and/or transmission, of data from a patch in D_1 covering z', to a patch in D_2 covering z. The term $m(z, z')$ gathers all contributions by reception, whereas the term $m(z', z)$ considers all transmissions.

When $\varphi(t) = t^2$ we call the resulting method *patch-wise non-local means*. In this case Eq. (12) can be written as

$$u(z) = \frac{1}{C(z)} \sum_{z' \in O^c} (\alpha m(z', z) + \beta m(z, z')) u(z'), \tag{13}$$

for each $z \in O$, where the normalization constant $C(z)$ is given by $C(z) = \sum_{z' \in O^c}(\alpha m(z', z) + \beta m(z, z'))$. Let us say in passing that due to our variational formulation, the image updating step is different from [24], since only the central pixel of the patch is updated in [24]. Taking $\varphi(t) = |t|$, we get the *patch-wise non-local medians*. In this case, the Euler equation (12) for u, given w, becomes $\sum_{z' \in O^c}(\alpha m(z', z) + \beta m(z, z'))\text{sign}(u(z) - u(z')) = 0$, and its solution $u(z)$ is obtained as a weighted median of the known values $u(z')$.

4   A departure from 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 ($\alpha = 0$, $\beta = 1$) and combined ($\alpha = 1$, $\beta = 1$). But each scheme also forces the manner to compute w. Now, if we abandon the variational framework, we can combine different update schemes of w and u.

We now propose a new scheme by updating the weights w according to the transmission scheme ($\alpha = 1$, $\beta = 0$), and 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 x' may provide a poorer estimate of the weights, specially if the current interpolation around x is bad.
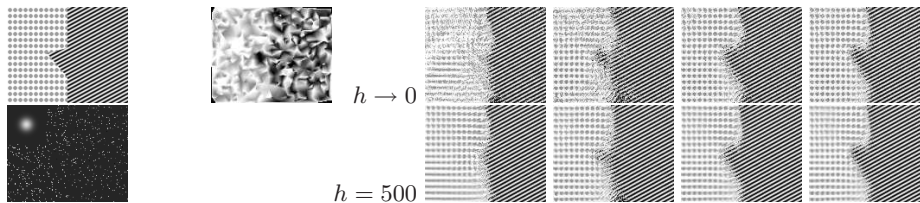
5   Experimental results

Fig. 2. Synthetic. The first column shows the original image, and the random samples (5% of the image) with the window g_σ (see (5)) depicted in the upper left corner. The second column from left shows a linear interpolation over the Delaunay triangulation of the samples. Remaining columns (from left to right) correspond to results of the schemes A, B, AB and O; The rows correspond to two different values of h.

We now present experimental results with both synthetic and natural images randomly sampled with densities from 20% to 5% of the image points. The four schemes derived from the potential (5) in Section 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 a computational time of 120 minutes.

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 around 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 [31]. 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_φ. The experiments shown in Figure 2 are aimed to compare the performance of the different schemes. The best results for this data are obtained with the scheme AB. Therefore, since the experiments are also consistent with these results, from now on we will mainly show AB and O. Also notice 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 for small h these methods tend to reinforce them. To reduce the dependence on the initial condition we adopt the *coarse to fine* scheme proposed in [24], where a decreasing sequence of h is used to recover first large scale structures and later refine them (as h decreases). Figures 3 and 4(b) 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 $h\rightarrow0$ introduces a staircase effect; we expect to improve these results by using $h>0$ in the median case (which corresponds to $p=1$ in φ (5)). In the first two columns of Figure 3 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 [22]). 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 [22]), while our method is less fitted for this task (second column of Figure 3). However, for

random samplings the results of the Laplacian interpolation are less convincing, while our method recovers most edges and textures of the image (see Figure 4).

Interpolation of large holes. In Figure 4(a) 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 [23]. Let us remark that the method was applied "as it is" to the problem, 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, showing the main limitation of exemplar-based methods. A local regularization term can be used to impose smoothness on the result [26]. The results shown here are also available at: `http://gpi.upf.edu/static/vnli`.

6   Conclusions and future work

A variational formulation for non-local example-based image interpolation was introduced in this paper. The obtained results show a promising performance. In subsequent work we will extend the present model to cover the case of samples located at non-entire positions and we will explore some variants of it.

References

1. Bertalmío, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: Proc. of SIGGRAPH. (2000)
2. Bertalmío, M., Bertozzi, A., Sapiro., G.: Navier-stokes, fluid-dynamics and image and video inpainting. In: Proc. of the IEEE Conf. on CVPR. (2001)
3. Masnou, S., Morel, J.M.: Level lines based disocclusion. In: Proc. of IEEE ICIP. (1998)
4. Efros, A., Leung, T.: Texture synthesis by non-parametric sampling. In: Proc. of the IEEE Conf. on CVPR. (September 1999) 1033–1038
5. Levin, A., Zomet, A., Weiss, Y.: Learning how to inpaint from global image statistics. In: Proc. of ICCV. (2003)
6. Criminisi, A., Pérez, P., Toyama, K.: Region filling and object removal by exemplar-based inpainting. IEEE Trans. on IP **13**(9) (2004) 1200–1212
7. Sun, J., Yuan, L., Jia, J., Shum, H.Y.: Image completion with structure propagation. In: Proc. of SIGGRAPH. (2005)
8. Bertalmío, M., Vese, L., Sapiro, G., Osher, S.: Simultaneous structure and texture inpainting. IEEE Trans. on Image Processing **12**(8) (2003) 882–89
9. Cao, F., Gousseau, Y., Masnou, S., Pérez, P.: Geometrically-guided exemplar-based inpainting. Submitted (2008)

10. Gröchenig, K., Strohmer, T.: Numerical and theoretical aspects of non-uniform sampling of band-limited images. In Marvasti, F., ed.: Theory and Practice of Nonuniform Sampling. Kluwer/Plenum (2000)
11. Chan, T., Shen, J.H.: Mathematical models for local nontexture inpaintings. SIAM J. App. Math. **62**(3) (2001) 1019–43
12. Arigovindan, M., Suhling, M., Hunziker, P., Unser, M.: Variational image reconstruction from arbitrarily spaced samples: a fast multiresolution spline solution. IEEE Trans. on IP **14**(4) (2005) 450–460
13. Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data. In: Proc. of 23rd ACM national conf., New York, NY, USA, ACM (1968) 517–524
14. Candes, E.J., Wakin, M.B.: An introduction to compressive sampling. Signal Processing Magazine, IEEE **25**(2) (2008) 21–30
15. Aharon, M., Elad, M., Bruckstein, A.: The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representatio. IEEE Trans. on Signal Processing **54**(11) (2006) 4311–22
16. Mairal, J., Sapiro, G., Elad, M.: Learning multiscale sparse representations for image and video restoration. SIAM MMS **7**(1) (2008) 214–241
17. Elad, M., Starck, J., Querre, P., Donoho, D.: Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). Applied and Computational Harmonic Analysis **19**(3) (2005) 340–358
18. Hays, J., Efros, A.: Scene completion using millions of photographs. In: Proc. of SIGGRAPH. (2007)
19. Peyré, G., Bougleux, S., Cohen, L.: Non-local regularization of inverse problems. In: ECCV '08, Berlin, Heidelberg, Springer-Verlag (2008) 57–68
20. Lezoray, O., Elmoataz, A., Bougleux, S.: Graph regularization for color image processing. Comput. Vis. Image Underst. **107**(1-2) (2007) 38–55
21. Gilboa, G., Osher, S.: Nonlocal linear image regularization and supervised segmentation. SIAM Mult. Mod. and Sim. **6**(2) (2007) 595–630
22. Belhachmi, Z., Bucur, D., Burgeth, B., Weickert, J.: How to choose interpolation data in images. Technical Report No. 205, Department of Mathematics, Saarland University, Saarbrücken, Germany (2008)
23. Arias, P., Caselles, V., Sappiro, G.: A variational framework for non-local image inpainting. In: Proc. of EMMCVPR, Springer (2009)
24. Buades, A., Coll, B., Morel, J., Sbert, C.: Self similarity driven color demosaicing. IEEE TIP **18**(6) (June 2009) 1192–1202
25. Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Non-local sparse models for image restoration. Submitted (2009)
26. Protter, M., Elad, M., Takeda, H., Milanfar, P.: Generalizing the non-local-means to super-resolution reconstruction. IEEE Trans. on IP **18**(1) (2009) 36–51
27. Bartesaghi, A., Sprechmann, P., Liu, J., Randall, G., Sapiro, G., Subramaniam, S.: Classification and 3d averaging with missing wedge correction in biological electron tomography? Journal of Structural Biology **162**(3) (June 2008) 436–450
28. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. CoRR **abs/0805.4471** (2008)
29. Awate, S., Whitaker, R.: Unsupervised, information-theoretic, adaptive image filtering for image restoration. IEEE Trans. on PAMI **28**(3) (2006) 364–376
30. Brox, T., Kleinschmidt, O., Cremers, D.: Efficient nonlocal means for denoising of textural patterns. IEEE Trans. on IP **17**(7) (2008) 1057–92
31. Kervrann, C., Boulanger, J.: Optimal spatial adaptation for patch-based image denoising. IEEE Trans. on IP **15**(10) (2006) 2866–2878
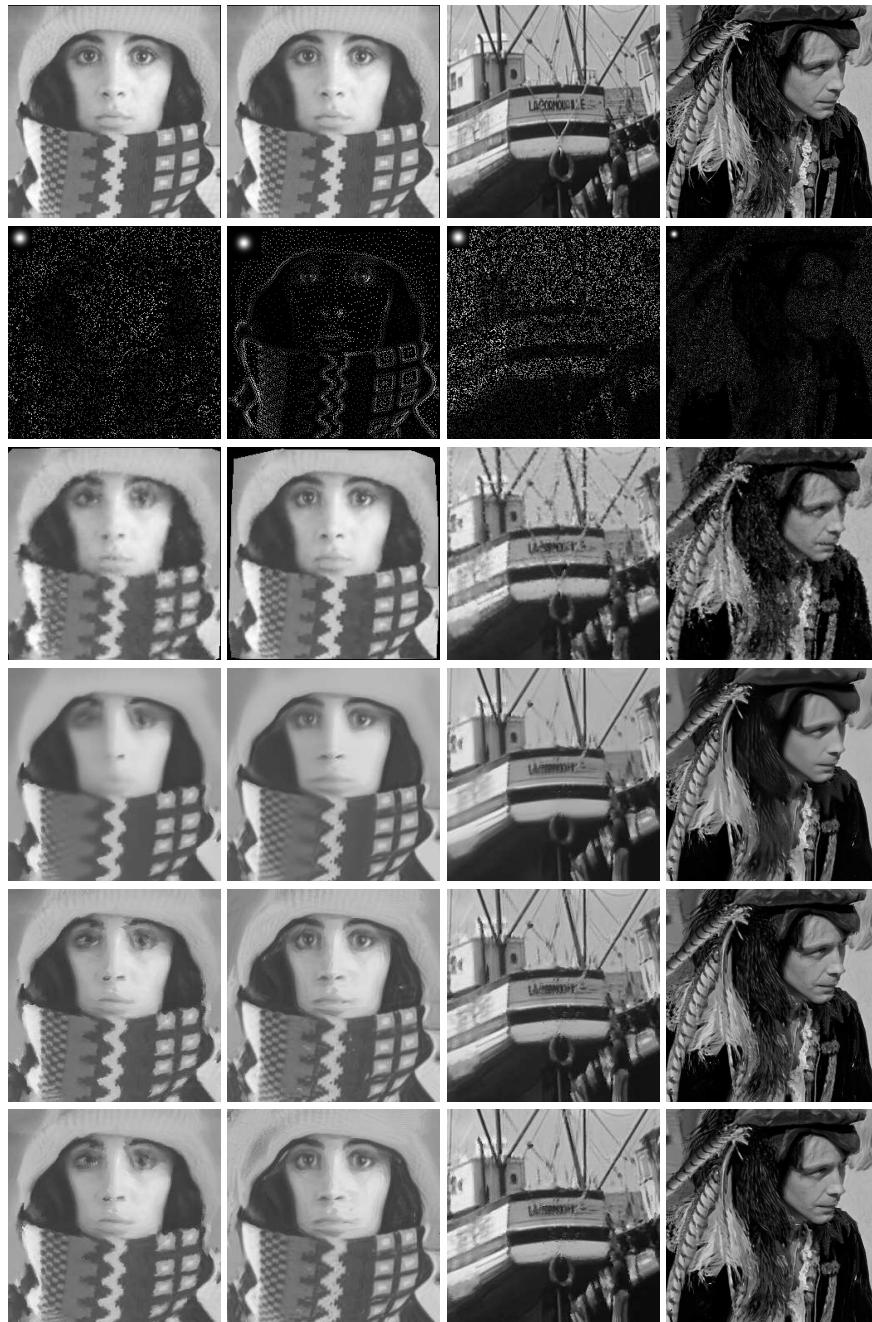
Fig. 3. *Sparse sampling interpolation.* 1st row: original images. 2nd row: input data with sample densities of 10%, 10% dithered [22], 20% and 9%. 3rd row: linear interpolation over the Delaunay triangulation of the samples; PSNRs: 25.8, 30.6 (not considering the black frame), 25.0 and 22.74. 4rt row: 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)

(a) *Cylinders,* 200×113. Inpainting a hole in a subsampled image with algorithm *B*. The algorithm makes no distinction between the hole and the sampled regions. The sampling density is 20% yielding a global sampling density of 14%.

(b) *Barbara,* 512×512 *with* 5% *of the samples.* 2nd line are results of: linear interpolation (PSNR 20.1), Laplacian interpolation (PSNR 19.9) and algorithm *O* with $h = 100$ (PSNR 22.8).

Fig. 4. Experiments with lower sampling densities. Each figure shows the original image (top left), the available samples (top right), the result of linear interpolation over the Delaunay triangulation (bottom left) and the output the algorithm specified in each figure (bottom right).