GEOMETRIC AND TEXTURE INPAINTING BASED ON DISCRETE REGULARIZATION ON GRAPHS

Mahmoud Ghoniem, Youssef Chahir, Abderrahim Elmoataz

GREYC, CNRS UMR-6072
Campus 2, Boulevard du Marechal Juin, 14032, Caen, France

ABSTRACT

We present an inpainting method for images and videos based on nonlocal discrete p-Laplace regularization on weighted graphs. Our work has the advantage of unifying local geometric methods and nonlocal exemplar-based ones in the same framework. Our image inpainting benefits from local and nonlocal regularities within the image. In addition to that, our video inpainting exploits temporal and spatial redundancies in order to obtain high quality results by considering a video sequence as a volume and not as a sequence of still frames. However, our method does not employ any motion estimation for video inpainting. Experiments demonstrate that our nonlocal method outperforms the local one by completing missing data with finer and more consistent details for textured and non-textured images and videos.

Index Terms— Image and video inpainting, nonlocal discrete regularization, weighted graphs

1. INTRODUCTION

The inpainting process consists in filling in the missing parts of an image or a video with the most appropriate data in order to obtain harmonious and hardly detectable reconstructed zones. Recent works on image and video inpainting may fall under two main categories, namely, the geometric algorithms and the exemplar-based ones. The first category employs partial differential equations (PDE). Bertalmio et al. [1] proposed some preliminary work on frame-by-frame PDEs based video inpainting, whereby areas are filled by propagating information based on PDE from the outside of the masked region along level lines (isophotes). The PDE is applied spatially, and completes the video frame by frame. This approach does not take into account the temporal information that a video provides. The second group of inpainting algorithms is based on the texture synthesis. The work of Efros and Leung [2] on non-parametric sampling is utilized in these exemplar-based techniques. Wexler et al. [3] proposed a global optimization for space-time completion of holes in a video sequence. This method yields good results, but is computationally very expensive and over-smoothing is observed.

In this paper, we present a generic framework for nonlocal discrete regularization on graphs that exploits and judiciously adapts a variational formulation for inpainting purposes on image and video media. We propose to restore the missing data using a regularization approach that takes into account local and nonlocal information. The main advantage of our method is the unification of the geometric and texture-based techniques. It is worth mentioning that our method does not employ any motion estimation for video inpainting unlike some approaches like the work of Kokaram et al. [4].

In Section 2, we briefly present the principles of regularization on weighted graphs. Thereafter, we state the inpainting problem in terms of regularization on graphs. In Section 3, we report and discuss empirical results for our proposed methods for image and video inpainting. We conclude this paper in Section 4 with a summary of our findings.

2. P-LAPLACE REGULARIZATION

In this section, we describe our discrete regularization framework [5] for image and video inpainting, which we have recently extended for video denoising [6]. First of all, we will introduce some preliminary definitions of derivatives on graphs required for this work.

2.1. Preliminary definitions

Let $G = (V, E)$ be a graph representing a general discrete domain that could be an image or a video. $V = \{v_1, \dots, v_n\}$ is a finite set of vertices and $E \subseteq V \times V$ is a finite set of edges. G is assumed to be undirected. Two vertices u and v are said to be adjacent if the edge $(u, v) \in E$. A graph is weighted if we associate to it a weight function $w : V \times V \to \mathbb{R}^+$, that must be a positive symmetric function satisfying: $\forall u \in E, w(u, u) = 0$. Let $\mathcal{H}(V)$ be a Hilbert space of real-valued functions on vertices. A function $f : V \to \mathbb{R}$ in $\mathcal{H}(V)$ assigns a vector f_v to each vertex v in V. The *local variation of the weighted gradient* operator $\|\bigtriangledown\|$ of a function $f \in \mathcal{H}(V)$ at a vertex v is defined by:

$$\|\bigtriangledown f(v)\| = \sqrt{\sum_{u \sim v} w(u, v)(f(v) - f(u))^2}.$$

This can be viewed as a measure of the regularity of a function around a vertex.
The *weighted p-Laplace operator*, with $p \in]0, +\infty[$, at a vertex v is defined on $\mathcal{H}(V)$ by:

$$(\Delta_p f)(v) = \frac{1}{p} \sum_{u \sim v} \gamma(u,v) \left(f(v) - f(u)\right) \text{ , where,}$$
$$\gamma(u,v) = w(u,v) \left(\|\bigtriangledown_f(v)\|^{p-2} + \|\bigtriangledown_f(u)\|^{p-2} \right)$$

2.2. Inpainting on weighted graphs

Consider a function f^0 that could be an image or a video. This function is defined over the vertices V of a weighted graph $G_w = (V, E, w)$ by $f^0 : V \to \mathbb{R}^m$. Let $V_0 \subset V$ be the subset of the nodes corresponding to the missing parts. The inpainting purpose is to interpolate the known values of f^0, $V - V_0$, to V_0. We formalize the inpainting problem as a discrete regularization using the weighted p-Laplace operator by the minimization of two energy terms:

$$f^* = \min_{f \in \mathcal{H}(V)} \left\{ \frac{1}{p} \sum_{v \in V} \|\bigtriangledown f(v)\|^p + \frac{\lambda(v)}{2} \|f - f^0\|^2_{\mathcal{H}(V)} \right\} \quad (1)$$

where,

$$\lambda(v) = \begin{cases} \lambda = constant, & \text{if } v \in V \backslash V_0 \\ 0, & \text{otherwise.} \end{cases}$$

$p \in [0, +\infty[$ is the smoothness degree, λ is the fidelity parameter, which specifies the trade-off between the two competing terms, and $\bigtriangledown f$ represents the weighted gradient of the function f over the graph. The solution of problem (1) leads to a family of nonlinear filters, parameterized by the weight function, the degree of smoothness, and the fidelity parameter. For missing parts, new values are computed with no initial values to be taken into account, hence, the fidelity parameter λ is set to 0.
The first energy in (1) is the smoothness term or regularizer, whereas the second is the fitting term. Problem (1) has a unique solution, for $p \geq 1$, which satisfies:

$$(\Delta_p f(v)) + \lambda(v)(f(v) - f^0(v)) = 0, \forall v \in V. \quad (2)$$

For further details on the solution, the more inquisitive reader is referred to [5].

From now on, we will consider an image as a particular video with only one frame. We consider a video sequence as a function f defined over the vertices of a weighted graph $G_{k_1,k_2,k_3} = (V, E, w)$, where $k_1, k_2, k_3 \in \mathbb{N}^3$. A vertex u is defined by a triplet (i, j, t) where (i, j) indicates the spatial position of the vertex and t, which is a frame number, indicates the temporal position of the vertex within the video sequence. We denote by $u \sim v$ a vertex u that belongs to the neighborhood of v which is defined as follows:

$$N_{k_1,k_2,k_3}(v) = \left\{ \begin{array}{l} u = (i', j', t') \in V \backslash V_0 : \\ |i - i'| \leq k_1, |j - j'| \leq k_2, |t - t'| \leq k_3 \end{array} \right\}$$

Similarly, we extend the definition of the patch to videos to obtain 3D patches. A patch around a vertex v is a box of size $r_x \times r_y \times r_t$, denoted by $B(v)$. Then, we associate to this patch a feature vector defined by:

$$F(f^0, v) = f^0(u), u \in B(v), u \in V \backslash V_0.$$

The weight function w associated to G_{k_1,k_2,k_3} provides a measure of the distance between its vertices that can simply incorporate local, semi-local or nonlocal features according to the topology of the graph and the image.

We consider the following two general weight functions:

$$w_L(u,v) = exp\Big(-\frac{|f(u) - f(v)|^2}{2\sigma_d^2}\Big)$$
$$w_{NL}(u,v) = w_L(u,v).exp\Big(-\frac{\|F(f^0,u) - F(f^0,v)\|^2}{h^2}\Big),$$

where σ_d^2 depends on the variations of $|f(u) - f(v)|$ over the graph. h can be estimated using the standard deviation depending on the variations of $\|F(f^0, u) - F(f^0, v)\|$ over the graph.

$w_L(u, v)$ is a measure of the difference between $f(u)$ and $f(v)$ values, and is used in the local approach of denoising. In addition to the difference between values, $w_{NL}(u, v)$ includes a similarity estimation of the compared features by measuring a $\mathcal{L}^2$ distance between the patches around u and v. It is the nonlocal approach.
The regularization problem is solved using the Gauss-Jacobi iterative algorithm presented in [5] which is specialized as follows: For all (u, v) in V_0,

$$\begin{cases} f^{(0)} = f^0 \\ \gamma^{(k)}(u,v) = w(u,v) \left(\| \bigtriangledown f^{(k)}(v)\|^{p-2} + \| \bigtriangledown f^{(k)}(u)\|^{p-2} \right) \\ f^{(k+1)}(v) = \dfrac{\sum_{u \sim v} \gamma^{(k)}(u,v) f^{(k)}(u)}{\sum_{u \sim v} \gamma^{(k)}(u,v)} \end{cases}$$

where $\gamma^{(k)}$ is the function γ at the k^{th} step. The weights $w(u, v)$ are typically computed from f^0, or could otherwise be explicitly given as an input.

Related Works
The aforementioned framework unifies and subsumes several special techniques that are explored in the literature. In fact, by considering particular parameter values, we recover results that have been established in image processing. For $p = 2$, w_{NL} and one iteration the nonlocal method is equivalent to the nonlocal means filter of Buades et al. [7] that has been adapted to inpainting by Wong et al. [8]. With $p = 1$ and $w = 1$, we obtain the local total variation (TV) inpainting of Chan and Shen [9]. With $p = 1$ and w_{NL}, it is the nonlocal TV inpainting. Our method could be considered as an extension of Efros and Leung's work [2]. In fact, if we construct the k-nearest neighbors graph with $k = 1$ and a patch distance between nodes, we obtain the same approach. However, we can consider in our algorithm different values for k.

2.3. Global Description of our Algorithm

Our method consists in filling that mask from its outer line to its center recursively in a series of nested outlines. The regularization process is applied iteratively on each node. This leads to an enhanced visual result. Once the entire outer line is processed, it is removed from V_0 and is considered as known data to take into account to process remaining holes. This means that we do not include the computed value of a pixel in the estimation of the other pixels on the same level. Thus, the risk of error propagation is reduced. As the inpainting process progresses, the mask gets dynamically smaller and eventually becomes empty. At this point, all the holes are filled in.

3. EXPERIMENTAL RESULTS

We present now some tests that demonstrate the efficiency of our proposed inpainting procedure. To this end, several media were selected. The patch distance was determined based on the intensity.

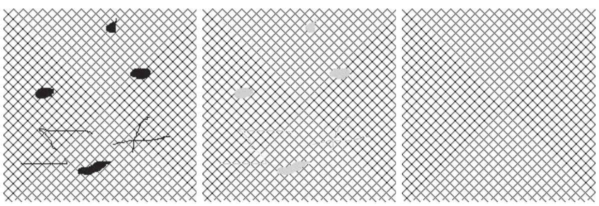

Fig. 1. Inpainting a synthetic video. Each frame of this video is a translation of the previous one. From left to right: the corrupted image, the result of local inpainting and the result of our nonlocal inpainting method.

In our first example, our algorithm was applied to a synthesis image. Figure 1 indicates that the image was perfectly reconstructed using our nonlocal algorithm, whereas the local inpainting approach produced poor results. This test serves as a basic verification of the nonlocal algorithm.

The result on a texture image is reported in Figure 2. Observe that the reconstructed zones are not detectable and merge harmoniously with the uncorrupted data.

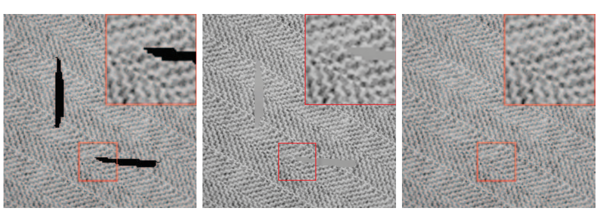

Fig. 2. Texture inpainting on weave image using a window 61 and patch 31. From left to right: the corrupted image, the result of a local inpainting and the result of our proposed nonlocal inpainting method.

The results of the inpainting on complex images containing zones of variable homogeneity are also very encouraging. Figure 3 reports the restoration of the reference image of Barbara. We purposely removed zones in a homogeneous region, textured zones, and zones of variable homogeneity to test the robustness of our method.

Fig. 3. Inpainting on barbara image using a window 21 and patch 5. From left to right, the original image, the corrupted image, the result of a local inpainting and our nonlocal inpainting result.

In Figure 4, we present the result of our algorithm for inpainting large zones. In this example, we removed all persons on the image to keep the beautiful landscape with the pyramid only. The removed area is very large. We observe on the down right corner that the bricks texture is extended to fill in the adjacent hole as well as the gradient blue color of the sky in the top left corner.

In Figure 5, it is possible to observe the contribution of temporal redundancy in the inpainting process. We corrupted one frame in the well-known video corpus of Suzie. The results demonstrate the high efficiency of our nonlocal approach in coping with both texture and structures in the video. For instance, thin structures such as Suzie's hair or homogeneous zones like the background are equally well recovered whereas the local approach shows its limits.

To test our algorithm on object removal applications from videos, we masked the ball in the tennis sequence. We can observe in Figure 6 that the ball is removed correctly and the inpainting yields very good results.

4. CONCLUSION

In this paper, we introduced a new algorithm for image and video inpainting based on nonlocal p-Laplace regularization on graphs. We take advantage of local and nonlocal regularities to complete the missing parts in a way that is harmonious with the uncorrupted parts of the damaged media. For video inpainting, temporal and spatial redundancies are combined to enhance inpainting results for videos. Our work unifies geometric and texture-based inpainting methods.

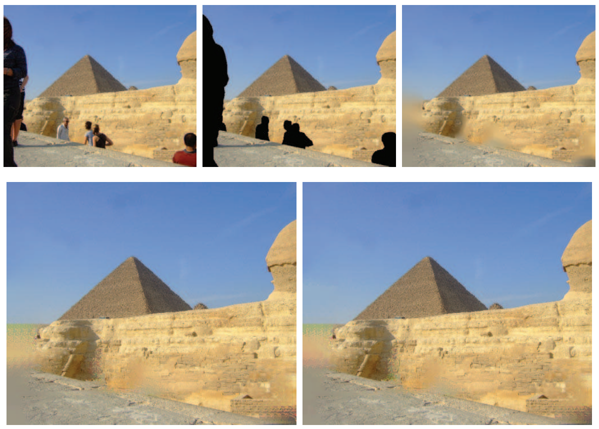

Fig. 4. Large holes inpainting on an image of Gizeh pyramid using a window 61 and patch 13 and ten iterations. From left to right, the original image, the corrupted image, the result of a local inpainting and our nonlocal inpainting results for $p = 2$ and $p = 1$.

Fig. 5. From left to right, the corrupted frame, the local inpainting and the nonlocal video inpainting results using a window size of 21 and a patch size of 6. The second line shows a zoomed in region of the first line.

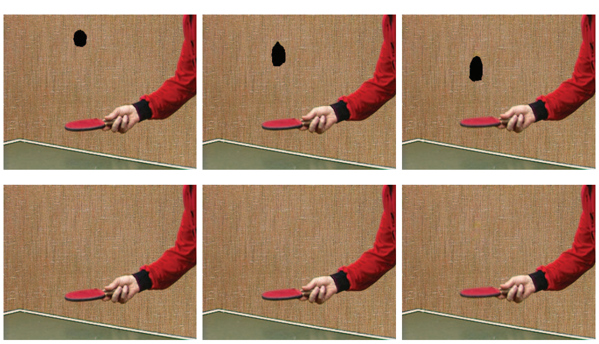

Fig. 6. Object removal from a color video. From left to right, frames from the original video, the corrupted video, the video without the tennis ball.

Our tests demonstrate the efficiency of our algorithm for completing the missing parts harmoniously with the data in the neighborhood so that no sharp edges appear between the existing data and the reconstructed portions. Moreover, we empirically observe that our results do not exhibit pronounced blur effects. Finally, it is worth mentioning that our algorithm can be advantageously applied to a broad spectrum of image and video editing applications.

5. REFERENCES

[1] M. Bertalmio, A.L. Bertozzi, and G. Sapiro, "Navier-stokes, fluid dynamics, and image and video inpainting," in *CVPR01*, 2001, pp. 355–362.

[2] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *IEEE International Conference on Computer Vision*, Corfu, Greece, September 1999, pp. 1033–1038.

[3] Y. Wexler, E. Shechtman, and M. Irani, "Space-time video completion," in *CVPR04*, 2004, pp. 120–127.

[4] A.C. Kokaram, B. Collis, and S. Robinson, "Automated rig removal with bayesian motion interpolation," vol. 152, no. 4, pp. 407–414, August 2005.

[5] A. Elmoataz, O. Lezoray, and S. Bougleux, "Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1047–1060, 2008.

[6] M. Ghoniem, Y. Chahir, and A. Elmoataz, "Video denoising via discrete regularization on graphs," in *ICPR08*, 2008, pp. 1–4.

[7] A. Buades, B. Coll, and J.-M. Morel, "Nonlocal image and movie denoising," *Int. J. Comput. Vision*, vol. 76, no. 2, pp. 123–139, 2008.

[8] Alexander Wong and Jeff Orchard, "A nonlocal-means approach to exemplar-based inpainting," in *ICIP*, 2008, pp. 2600–2603.

[9] Tony F. Chan and Jianhong Shen, "Variational image inpainting," *Comm. Pure Applied Math*, vol. 58, pp. 579–619, 2005.