

*ICIAM 2011*

# Exemplar-based Image Inpainting and Applications

By *V. Caselles*

Image inpainting, also known as “image completion” or “disocclusion,” is the recovery of missing or corrupted parts of an image in a given region  $O$  so that the reconstructed image looks natural (see Figure 1). It has become a standard tool in digital photography for image retouching (e.g., removal of scratches from old photographs), and intensive research is under way to convert image inpainting into a key tool for video and 3D cinema post-production (e.g., for object elimination or replacement). Besides its numerous applications to image and video editing, the problem is of theoretical interest in that its analysis involves an understanding of the self-similarity visible in the repetitive geometric and texture patterns that appear in almost any natural image.

Although it is impossible to do justice to all of the many research efforts under way, most inpainting methods in the literature can be divided into two groups: geometry- and texture-oriented methods. Techniques of the two types were introduced at almost the same time [13,21]; texture-oriented methods have proved better adapted to the production of natural-looking images, although ideas from both have contributed to the production of state-of-the-art results.

This classification also represents the two main ideas underlying most work in image inpainting. With geometry-oriented methods, images can be 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. Taking advantage of this structure, these methods interpolate the inpainting domain by continuing the geometric structure of the image (its level lines or edges), usually as the solution of a (geometric) variational problem or by means of a partial differential equation. Such a PDE can be derived from variational principles, as in [5,10,20,21], or inspired by phenomenological modeling [7,8]. The methods are local in the sense that the associated PDEs involve only interactions between neighboring pixels on the image grid. This implies that only the data around the boundary of the inpainting domain is used. These methods perform well in propagating smooth level lines or gradients, but fail in the presence of texture.

Texture-oriented (also called exemplar-based) inpainting was initiated by Efros and Leung [13] in their work on texture synthesis using non-parametric sampling techniques. This method exploits the self-similarity prior by directly sampling the desired texture to perform the synthesis. The value of each target pixel  $x \in O$  is copied from the center of a (square) patch in the sample image, chosen to match the available portion of the patch centered at  $x$ . Unlike geometry-oriented inpainting, exemplar-based approaches are non-local: To determine the value at  $x$ , the whole image might be scanned in the search for a matching patch. The self-similarity prior is one of the most influential ideas underlying recent progress in image processing, especially in image denoising [9].

The synthesis of methods of these two types is a trend in current research. Variational analysis in particular is an appropriate tool for a unified treatment of local and non-local methods. As pointed out in [12], the problem of exemplar-based inpainting can be stated as that of finding a correspondence map  $\Gamma$  that assigns to each location  $x$  in the inpainting domain  $O$  a corresponding location  $\Gamma(x) \in O^c$ , the domain in which the image  $u$  is known. The unknown part of the image  $u|_O$  is then synthesized using the correspondences  $\Gamma$  by  $u|_O(x) = u(\Gamma(x))$ ,  $x \in O$ . This basic idea has led to a set of powerful algorithms, some based on well-founded heuristics [11], others formulated variationally [4,12]. In the latter, the energy is usually highly non-convex and difficult to minimize [4]. Some authors, as in [19], have approached the problem of finding a correspondence map by looking for a simpler optimization problem, formulated as probabilistic inference on a graphical model. One of the most effective optimization strategies relies on a relaxation of the problem, in which the unknown variables are the image (to be reconstructed) and the correspondence map, which appears as an auxiliary variable. The resulting algorithm can be regarded as an alternate optimization of the two variables. This is the approach pursued in [2,3,18,23,25] in work combining the advantages of non-local variational models and an improved new algorithm that produces state-of-the-art results.



**Figure 1.** *Inpainting problem. Left, the given image. Middle, the region  $O$  to be inpainted. For the inpainting problem, this region is provided to the user, so that its computation is a separate problem. Right, inpainting result.*



**Figure 2.** *Left, the given image. Right, the inpainting result.*

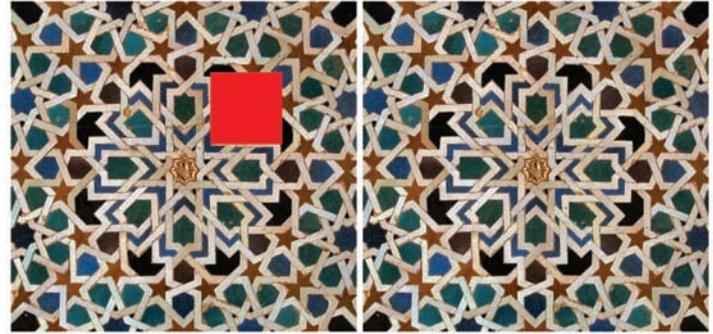
We can formulate the inpainting problem in a statistical mechanics framework, in which we maximize the self-similarity of the reconstructed image with a uniform prior for selecting equally similar patches, expressed as a maximization of the entropy [2,3,14]. This leads to a Gibbs energy functional for the similarity weights  $w(x, x')$  between points  $x \in O$  and points  $x' \in O^c$ . This general formulation covers inpainting as well as image reconstruction from a set of sparse samples, in which case annealing is recommended (letting the temperature go to 0 in the Gibbs energy) [14]. The correspondence map formulation corresponds to the limit of the Gibbs energy when the temperature is allowed to go to 0. By choosing different similarity measures, we generate different functionals, which copy image intensities, gradients, or some combination of them. Features from geometry-based models can easily be incorporated, e.g., in the continuation of image gradients. Figure 2 illustrates the behavior of the algorithm.

For the correspondence map, without the introduction of a particular regularization, the correspondences obtained exhibit piecewise-regular behavior; this is supported by our theoretical results [1]. This behavior is illustrated in Figure 3. The PatchMatch algorithm [6], which exploits the correlation between patches so that they can be found collectively, speeds up the computation of optimal correspondences. The exponential behavior of the algorithm is presented in [1]. In some cases, there are no patches in the image that can be used to naturally continue the structures of the image beyond the boundary of the inpainting mask, as illustrated in Figure 4 (a wall in the Alhambra, Granada). It was by generating symmetric versions of the image that we were able to create the patches that permit the natural extension of the image.

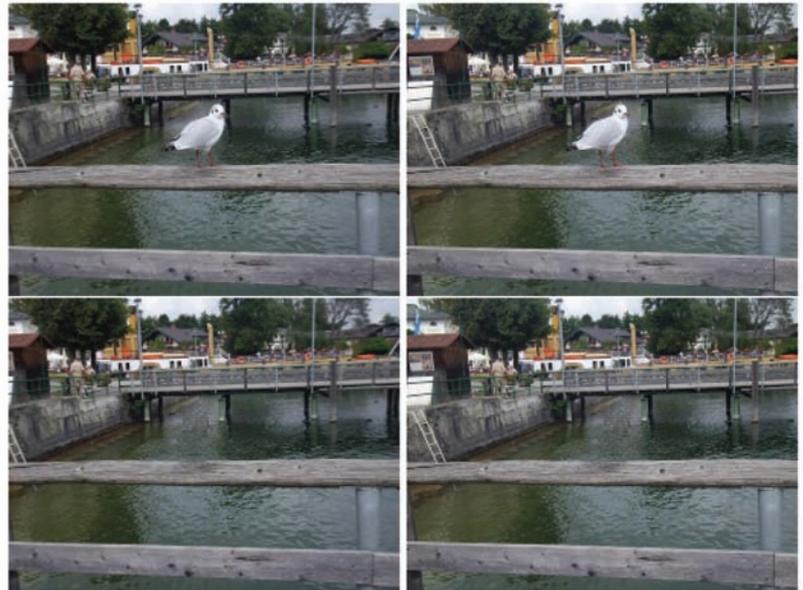
The recent commercial interest in showing movies and sports or music events in 3D has motivated the development of post-production tools that assist in their acquisition or that eliminate unwanted objects, such as rigs or cables, that may be unavoidable during filming. In this context, stereo image inpainting [17] addresses the reconstruction of missing information in pairs of stereo images; the reconstructed parts of both images have to look like the projection of real 3D objects and must produce a depth-consistent perception. The first step in the process is the estimation of the depth of the scene and its reconstruction in the regions to be inpainted. This permits the simultaneous and depth-coherent inpainting of stereo images where corresponding points in the two images are inpainted with the same color. Examples of stereo image inpainting are shown in Figures 5 and 6.



**Figure 3.** Copy regions. Left, the given image. Middle, the region  $O$  to be inpainted. Right, the emergence of copy regions.



**Figure 4.** Left, the given image. Right, the inpainting result. The patches permitting a natural continuation of the image into the hole were generated by symmetrizing the images.



**Figures 5 and 6.** Stereo image inpainting. The given pairs of stereo images (top) and the results of simultaneous and depth-consistent image inpainting (bottom).

A final interesting problem is that of inpainting in video sequences. The enormous variety of situations makes this problem both extremely difficult and extremely interesting. The automatic or semiautomatic choice of the inpainting region, and the development of techniques for imposing time coherence in the inpainted result and for creating illumination effects consistent with the surrounding parts of the image [15, 22] are just some of the difficulties that arise. Although we have made progress on this problem, a lot of work remains to be done if we are to provide an effective tool for the video and cinema post-production industry.

## References

- [1] P. Arias, V. Caselles, and G. Facciolo, *Analysis of a variational framework for exemplar-based image inpainting*, preprint, July 2011.

- [2] P. Arias, V. Caselles, G. Facciolo, and G. Sapiro, *Variational framework for exemplar-based image inpainting*, Int. J. Comput. Vis., 93:3 (2011), 1–29.
- [3] P. Arias, V. Caselles, and G. Sapiro, *A variational framework for non-local image inpainting*, Proceedings of the 7th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition, EMMCVPR, Springer LNCS, Bonn, August 2009.
- [4] J.F. Aujol, S. Ladjal, and S. Masnou, *Exemplar-based inpainting from a variational point of view*, SIAM J. Math. Anal., 42:3 (2010), 1246–85.
- [5] C. Ballester, M. Bertalmío, V. Caselles, G. Sapiro, and J. Verdera, *Filling-in by joint interpolation of vector fields and gray levels*, IEEE Trans. IP, 10:8 (2001), 1200–11.
- [6] C. Barnes, E. Shechtman, A. Finkelstein, and D.B. Goldman, *PatchMatch: A randomized correspondence algorithm for structural image editing*, Proceedings of SIGGRAPH, ACM, New York, 2009, 1–11.
- [7] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester, *Image inpainting*, Proceedings of SIGGRAPH, ACM, 2000.
- [8] F. Bornemann and T. März, *Fast image inpainting based on coherence transport*, J. Math. Imag. and Vis., 28:3 (2007), 259–78.
- [9] A. Buades, B. Coll, and J.M. Morel, *A non local algorithm for image denoising*, Proceedings of the IEEE Conference on CVPR, Volume 2, 2005, 60–65.
- [10] T. Chan, S.H. Kang, and J.H. Shen, *Euler’s elastica and curvature based inpaintings*, SIAM J. App. Math. 63:2 (2002), 564–92.
- [11] A. Criminisi, P. Pérez, and K. Toyama, *Region filling and object removal by exemplar-based inpainting*, IEEE Trans. IP, 13:9 (2004), 1200–12.
- [12] L. Demanet, B. Song, and T. Chan, *Image inpainting by correspondence maps: A deterministic approach*, Appl. and Comput. Math., 1100 (2003), 217–50.
- [13] A. Efros and T. Leung, *Texture synthesis by non-parametric sampling*, Proceedings of the IEEE Conference on CVPR, 1999, 1033–38.
- [14] G. Facciolo, P. Arias, V. Caselles, and G. Sapiro, *Exemplar-based interpolation of sparsely sampled images*, Proceedings of the 7th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition, EMMCVPR, Springer LNCS, Bonn, August 2009.
- [15] G. Facciolo, R. Sadek, A. Bugeau, and V. Caselles, *Temporally consistent gradient domain video editing*, Proceedings of the 8th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition, EMMCVPR, Saint Petersburg, July 25–27, 2011.
- [16] G. Gilboa and S.J. Osher, *Nonlocal linear image regularization and supervised segmentation*, SIAM Mult. Mod. and Sim., 6:2 (2007), 595–630.
- [17] A. Hervieu, N. Papadakis, A. Bugeau, P. Gargallo, and V. Caselles, “Stereoscopic Image Inpainting Using Scene Geometry,” IEEE International Conference on Multimedia and EXPO, July 11–15, 2011, Barcelona.
- [18] N. Kawai, T. Sato, and N. Yokoya, *Image inpainting considering brightness change and spatial locality of textures and its evaluation*, in Advances in Image and Video Technology, Springer, Berlin/Heidelberg, 2009, 271–82.
- [19] N. Komodakis and G. Tziritas, *Image completion using efficient belief propagation via priority scheduling and dynamic pruning*, IEEE Trans. IP, 16:11 (2007), 2649–61.
- [20] S. Masnou, *Disocclusion: A variational approach using level lines*, IEEE Trans. IP, 11:2 (2002), 68–76.
- [21] S. Masnou and J.M. Morel, *Level lines based disocclusion*, Proceedings of IEEE ICIP, 1998.
- [22] P. Pérez, M. Gangnet, and A. Blake, *Poisson image editing*, Proceedings of SIGGRAPH, ACM, 2003.
- [23] G. Peyré, *Manifold models for signals and images*, Comp. Vis. and Im. Unders., 113:2 (2009), 249–60.
- [24] G. Peyré, S. Bougleux, and L. Cohen, *Non-local regularization of inverse problems*, ECCV 2008, Springer-Verlag, Berlin/Heidelberg, 2008, 57–68.
- [25] Y. Wexler, E. Shechtman, and M. Irani, *Space-time completion of video*, IEEE Trans. PAMI, 29:3 (2007), 463–76.

V. Caselles ([vicent.caselles@upf.edu](mailto:vicent.caselles@upf.edu)) is a senior professor in the Image Processing Group, Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu-Fabra, in Barcelona, Spain.