



Image Inpainting Based on Local Patch Search Supported by Image Segmentation

Sarah Almeida Carneiro^{1,2}, Helio Pedrini²,
and Silvio Jamil Ferzoli Guimarães^{1(✉)}

¹ Computer Science Department, Pontifical Catholic University of Minas Gerais,
Belo Horizonte, MG, Brazil

sjamil@pucminas.br

² Institute of Computing, University of Campinas, Campinas, SP, Brazil

Abstract. Image inpainting can be defined as a restoration process in which damaged or selected regions are repaired by taking into account the image content. In this work, we employ a local-based strategy instead of a global one to identify the best existing patch with information to replace the damaged/selected patch. In order to properly identify the most representative patches, we propose a method based on the (i) creation of a local graph using a similarity of patches in the original image and (ii) partition of the image into regions according to hierarchical image segmentation to support the local patch identification. The experimental results demonstrate that our local search outperformed the results of image inpainting in terms of both qualitative and quantitative aspects, when compared to global search of patches.

Keywords: Image inpainting · Hierarchical image segmentation · Patch graph

1 Introduction

Image inpainting, also known as image restoration, is the term used to describe the restoration process in order to repair a damaged or selected region of a given image [9, 10, 18, 19]. According to Bertalmio et al. [3], most of the inpainting approaches are either based on the extraction of fractions of the image, known as patch interpolation, or partial differential equation (PDE) approaches.

In the patch-based category, techniques for texture expansion by recursively filling pixel by pixel of the masked region of the image were performed by Efros and Leung [8]. Unfortunately, the size of the neighborhood and, in some cases, the tendency of reproducing incoherent textures are two problems related to their method. Another similar approach was proposed by taking into account

This work received funding from CAPES, FAPESP (17/12646-3), FAPEMIG (PPM 00006-16), and CNPq (Universal 421521/2016-3, PQ 307062/2016-3 and PQ 305169/2015-7).

changes in the filling order by (i) prioritizing border pixels without information and (ii) copying entire image patches to cover the flaws, making it possible to work with linear structures [6]. In this sense, this method aims to be able to remove large objects from the image by considering the principles of patch priority, propagating texture and structure information, and updating the values of the filled patches. A disadvantage of the method proposed by Criminisi [6] is the difficulty in working with images in perspective.

In the non-local patch graph category, some methods have been proposed in the literature [4, 7, 11–13, 17]. Partial differential equation (PDE) [16] approaches have also been conducted through the construction of a Gaussian pyramid based on convolutions and samples of the original image [13]. Therefore, a region selected for inpainting can be filled by linear interpolation using samples from the Gaussian pyramid. Other studies have been developed by interpolating gray scale images, thus extending the lines of constant intensity [12]. On the other hand, studies to reproduce the technique used by manual painting restorers have been accomplished in order to propagate information of the areas near the flaws along the border [4]. One of the drawbacks related to this technique is the difficulty in producing content for regions with large failures. Extensions of the study have used the idea of dynamic fluid propagation in order to scatter the padding information from the external regions towards the fault [2].

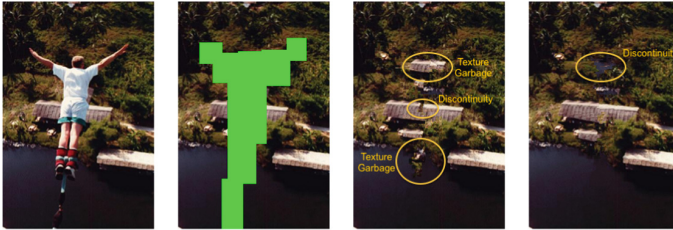


Fig. 1. Example of image inpainting. From left to right, we illustrate the original image, the image mask, the result obtained with the method described in [6], and the result obtained with our approach by taking into account local constraints. (Color figure online)

To deal with some problems in images with perspectives and some kind of texture, mainly due to the computation of patch priority, we propose an image inpainting method by considering patch locality in conjunction to hierarchical image segmentation since we look for patches close to the missing patch to be restored. In Fig. 1, we illustrate an example of inpainting in which the selected region represented by the green mask is restored by the method developed by Criminisi et al. [6] and our local-based approach. As we can observe, the result obtained in [6] when applied to the original image presents some problems, such as texture garbage and discontinuities. On the other hand, our local patch search strategy better reconstructed some parts of the original image, as illustrated

in the rightmost image. It is worth observing that, compared to [6], the main difference is related to the local patch search which is supported by an image segmentation and, of course, this new local search will affect the patch priority computation. Considering that, in this work, our main assumption is that the probability of similar patches to the blank patch is higher if the patch with information belongs to a closer region to the mask, we propose (i) the creation of a local graph using a similarity of patches in the original image and (ii) the partition of the image into regions according to a hierarchical image segmentation to support the local patch identification.

Due to the use of a local patch graph and image partitions, our main contributions are two-folds: (i) possibility of controlling the size of the patch search space preserving the homogeneous region properties; and (ii) only patches completely inside the patch search space are considered and consequently patches with texture due to different regions are ignored.

This work is organized as follows. In Sect. 2, we define some important concepts and we describe two methods graph-based image inpainting to be studied and compared. In Sect. 3, we describe our method for coping with image inpainting. In Sect. 4, some experimental results are presented and analyzed. We have showed that the proposed method outperformed, in terms of PSNR and MSE values, the compared approaches. Some final remarks and directions for future work are outlined in Sect. 5.

2 Theoretical Background

In this section, we define some important concepts and we present the two inpainting methods that will be explored.

2.1 Patch Graph

Let I be an image. We define a patch extraction operator R who returns the pixels into the squared $\sqrt{d} \times \sqrt{d}$ around a pixel location x , thus $R(x) = (I(x_1), \dots, I(x_d))$, in which d is the number of pixels of the patch. Let x and y be two pixel locations, given two patches centered at x and y , we define the distance between them as the usual squared Euclidean distance. According to [7], the visual similarity between two patches normalized in $[0, 1]$ can be computed given the distance between them by a simple exponential filtering defined as

$w(x, y) = e^{-\frac{d(x, y)}{h^2}}$, where $d(x, y)$ is the squared Euclidean distance between x and y , and the parameter h controls the decay of the exponential function.

Let $G = (V, E)$ be a patch graph which models the relationship between the patches extracted from the image I , in which V is the set of pixels of the image I and E is the set of edges. (G, W) be a weighted patch graph in which W is a weight function that assigns a positive weight to each edge. Each pixel at location x of the image I is represented by a vertex u whose data is given by the patch $R(x)$. Each pair of vertices u and v , that represents the pixels

at locations x and y , is connected by an edge (u, v) whose weight is computed by the similarity function $w(x, y)$. This weight is positive and tends to zero when patches are highly dissimilar. In practice, we only connect the K nearest neighbors (KNN) to reduce the number of edges. As discussed in [7], similar patches should be well connected together while weakly connected to the others, while the patches located at the transition between textures are ideally mildly connected to both clusters. Vertices representing patches who contain unknown pixels are not connected to any other vertex during the patch graph construction, thus this graph is initially disconnected and will ultimately be fully connected as these vertices are sequentially inserted.

2.2 Image Inpainting

The following subsections briefly describe two image inpainting approaches used for comparison purpose in this work.

Defferrard’s Method. The method proposed by Defferrard [7] could be defined as follows. A non-local patch graph from the known patches of the source region defined as the entire image minus the damaged/selected region. Thus, for computing the graph G , all pixels can be considered for computing the weight edges. Unknown patches are completely disconnected. Thus, the method is a sequential process with the following steps: (i) those among the disconnected patches who possess a sufficient amount of information (*i.e.*, some fraction of the pixels they represent have a value) are compared against all known; and (ii) connected patches are inserted into the graph with appropriate weights. The graph is thus sequentially completed, while no pixels have actually been inpainted. To allow further graph completion, we must indeed inpaint unknown pixels to recover enough information about a patch to compare it to the others.

Criminisi et al.’s Method. The method proposed in [6] is an exemplar-based image inpainting defined as follows. The original image is decomposed into two components: (i) one of which is processed by inpainting; and (ii) the other by texture synthesis. The result is the sum of the two processed components. Let I be the image to be inpainted. As in the Defferrard’s method, the source region may be defined as the entire image minus the damaged/selected region to be inpainted. Next, as with all exemplar-based texture synthesis, the size of the patch must be specified. In this method, each pixel maintains a colour value (or “empty”, if the pixel is unfilled) and a confidence value, which reflects our confidence in the pixel value, and which is frozen once a pixel has been filled. During the course of the algorithm, patches along the fill front are also given a temporary priority value, which determines the order in which they are filled. Then, the algorithm iterates the following three steps until all pixels have been filled: (i) computing patch priorities; (ii) propagating texture and structure information; and (iii) updating confidence values.

3 Local-Based Strategy for Image Inpainting

Instead of computing patch similarity taking into account the entire image as in the original method [6, 7], we reduce the patch search space by using segmentation strategy for maintaining homogeneous patches in that search. This kind of strategy will influence the patch priority computation. Thus, our assumption is that the probability of similar patches to the blank patch is higher if the patch with information belongs to a closer region to the mask. Closer regions are obtained through a hierarchical segmentation algorithm. Further details about this method can be found in [5, 14].

Algorithm 1. Image inpainting algorithm based on local patch graph.

Input: Original image f
Input: Image mask m
Parameter : Radius r
Parameter : Method $method$
Output: Inpainted image

```

1  $segImage \leftarrow Watershed(m)$ ;
2  $adjList \leftarrow Neighborhood(segImage, r)$ ;
3  $(G, W) \leftarrow CreatePatchGraph(m, method, adjList)$ ;
4  $patches \leftarrow ComputeBlankPatches(method, m)$ ;
5 while  $patches$  is not empty do
   | // Identify patches with high filling priority
6   |  $p \leftarrow PatchPriority(method, patches)$ ;
7   | Inpaint the missing patch  $p$ ;
8   |  $patches \leftarrow patches \setminus \{p\}$ ;
9   | Update  $(G, W, method)$ ;
10 end
```

The main steps of the proposed method for image inpainting is illustrated in Algorithm 1. The main goal of our work is to study the behavior of different inpainting methods, for instance, Criminisi et al. [6] and Defferrard [7], when we limit the patch search space for nearest regions to the blanked patches to be restored. Instead of partitioning the images in fixed regions, we have applied a watershed method (Line 1) to produce a set of regions. Considering that the watershed by area will compute a hierarchy of partitions, an ideal cut could be identified by Mumford-Shah energy [1]. After that, we compute the neighboring regions (Line 2) of the blanked patches that are at most r regions farther. Lines 3, 4 and 6 depend on the method in which we are interested to improve. In our proposal, function `CreatePatchGraph` is the most important one since it will compute the patch graph based on the two studied inpainting methods taking into account our local-based search strategy, in which only the patches inside the r regions farther are considered as candidates to restore the blanked patch. At the end (Line 9), the patch graph G must be updated. It is worth observing that if the adjacency list has all image patches, instead of local one, the result is the same as the original method.

4 Experiments

In order to analyze the results achieved with our approach compared to the original ones, we will illustrate some examples to show the quality of our results. Moreover, we present some quantitative values in terms of peak signal-noise ratio (PSNR) and mean square error (MSE). Here, we compare the following methods: Defferrard [7], local-based Defferrard, Criminisi et al. [6], and local-based Criminisi et al.

4.1 Qualitative Analysis

For a qualitative analysis, we illustrate some examples in which the restoration process is done through the compared methods. It is worth observing that, when dealing with homogeneous colors, such as water and sky, the Defferrard method [7] presents some disagreements when compared to homogeneity and texture. Such problems have been solved when local-based analysis is applied (for instance, the water is better restored in the Boat image). The local-based approach in conjunction with the method by Criminisi et al. [6] clearly exhibited a better construction compared to the original. For instance, the horse in Fig. 2 is properly removed through our inpainting method.

4.2 Quantitative Analysis

For a quantitative analysis, we compute the most widely used measure for image inpainting according to [15], the peak signal-noise ratio (PSNR) and mean square error (MSE). The PSNR is used to measure the quality of the reconstructed compressed data, whereas the MSE is a liability function. Higher PSNR values usually correspond to the composition with better quality, whereas smaller MSE values mean that the restoration is superior.

Since, to the best of our knowledge, there is no benchmark dataset for image inpainting, we decide to use a set of images commonly employed in this type of application. According to our experiments (illustrated in Fig. 2), the MSE values for Defferrard [7], local-based Defferrard, Criminisi et al. [6] and local-based Criminisi et al. approaches are 43.41, 31.19, 54.38 and 33.24, respectively, whereas the PSNR values are 34.43, 35.87, 35.09 and 36.56, respectively. As we can observe from these results, our inpainting method based on local strategy has improved both measures. More specifically, the MSE for Criminisi et al. [6] is 54.38, while the value for our local-based proposed method is 33.24.

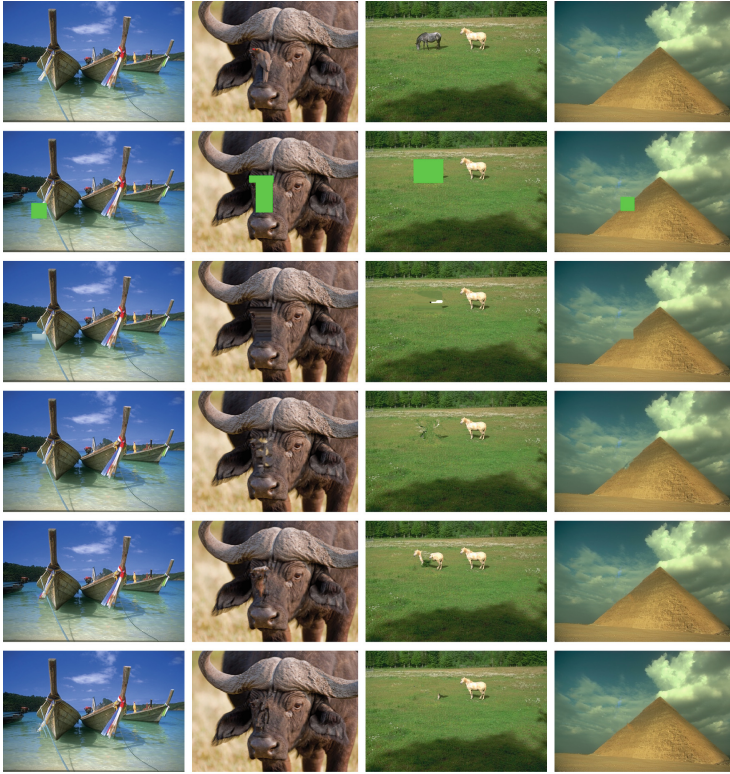


Fig. 2. Some results obtained with inpainting methods, more specifically, with the original methods Defferrard [7] and Criminisi et al. [6], and our proposed methods, local-based Defferrard and local-based Criminisi et al., based on these original ones in which we modified the search strategy in order to take into account the nearest regions to the one to be restored. Results for the compared methods are shown from top to bottom: original image; image mask; (third row) Defferrard [7]; (fourth row) local-based Defferrard; (fifth row) Criminisi et al. [6]; and (sixth) row local-based Criminisi et al. In terms of quantitative results, the MSE values are 43.41, 31.19, 54.38 and 33.24, respectively, whereas the PSNR values are 34.43, 35.87, 35.09 and 36.56, respectively.

5 Conclusions

Image inpainting could be defined as a restoration process in which damaged or selected regions are repaired taking into account the image content. In this paper, we have used local-based strategy instead of a global one to identify the best existing patch with information to replace the damaged/selected patches. Considering that our main assumption is that the probability of similar patches to the blank patch is higher if the patch with information belongs to a closer region to the mask, the (i) creation of a local graph using a similarity of patches in the original image and (ii) the partition of the image into regions according to

hierarchical image segmentation to support the local patch identification could be considered as our main contributions.

In order to properly identify the most representative patches, we proposed a method based on a partition of the image into regions according to hierarchical image segmentation to help in the identification of the local patches. According to our experiments, the local patch analysis have contributed to improve the quality of the results in both qualitative and quantitative assessment.

As directions for future work, we intend to investigate the use of a learning structure based on predetermined image data sets. Moreover, we will consider more sophisticated measure for evaluating image inpainting methods, some examples of these measures are summarized in [15].

References

1. Ambrosio, L., Fusco, N., Hutchinson, J.E.: Higher integrability of the gradient and dimension of the singular set for minimisers of the Mumford-Shah functional. *Calc. Var. Partial. Differ. Equ.* **16**(2), 187–215 (2003)
2. Bertalmio, M., Bertozzi, A.L., Sapiro, G.: Navier-stokes, fluid dynamics, and image and video inpainting. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1. IEEE (2001)
3. Bertalmio, M., Caselles, V., Masnou, S., Sapiro, G.: Inpainting. Technical report, Duke University (2011)
4. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: *27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 417–424. ACM Press/Addison-Wesley Publishing Co. (2000)
5. Cousty, J., Najman, L., Kenmochi, Y., Guimarães, S.: Hierarchical segmentations with graphs: quasi-flat zones, minimum spanning trees, and saliency maps. *J. Math. Imaging Vis.* **60**, 479–502 (2017)
6. Criminisi, A., Pérez, P., Toyama, K.: Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. Image Process.* **13**(9), 1200–1212 (2004)
7. Defferrard, M.: Graph-based image inpainting. Technical report, École Polytechnique Fédérale de Lausanne (EPFL) (2014)
8. Efros, A., Leung, T.: Texture synthesis by non-parametric sampling. In: *Seventh IEEE International Conference on Computer Vision*, p. 1033. IEEE, September 1999
9. Fan, Q., Zhang, L.: A novel patch matching algorithm for exemplar-based image inpainting. *Multimedia Tools Appl.* **77**, 1–15 (2017)
10. Guillemot, C., Le Meur, O.: Image inpainting: overview and recent advances. *IEEE Signal Process. Mag.* **31**(1), 127–144 (2014)
11. Karaca, E., Tunga, M.A.: An interpolation-based texture and pattern preserving algorithm for inpainting color image. *Expert Syst. Appl.* **91**, 223–234 (2018)
12. Masnou, S., Morel, J.-M.: Level Lines based disocclusion. In: *International Conference on Image Processing*, pp. 259–263. IEEE (1998)
13. Ogden, J.M., Adelson, E.H., Bergen, J.R., Burt, P.J.: Pyramid-based computer graphics. *RCA Eng.* **30**(5), 4–15 (1985)
14. Perret, B., Cousty, J., Guimarães, S.J.F., Maia, D.S.: Evaluation of hierarchical watersheds. *IEEE Trans. Image Process.* **27**(4), 1676–1688 (2018)

15. Qureshi, M.A., Deriche, M., Beghdadi, A., Amin, A.: A critical survey of state-of-the-art image inpainting quality assessment metrics. *J. Vis. Commun. Image Represent.* **49**, 177–191 (2017)
16. Roubířek, T.: *Nonlinear Partial Differential Equations with Applications*, vol. 153. Springer, Basel (2013). <https://doi.org/10.1007/978-3-0348-0513-1>
17. Wang, H., Jiang, L., Liang, R., Li, X.-X.: Exemplar-based image inpainting using structure consistent patch matching. *Neurocomputing* **269**, 90–96 (2017)
18. Yang, C., Lu, X., Lin, Z., Shechtman, E., Wang, O., Li, H.: High-resolution image inpainting using multi-scale neural patch synthesis. In: *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, p. 3 (2017)
19. Zhao, G., Liu, J., Jiang, J., Wang, W.: A deep cascade of neural networks for image inpainting, deblurring and denoising. *Multimed. Tools Appl.* **77**, 1–16 (2017)