

A Hierarchical Approach for High-Quality and Fast Image Completion

Thanh Trung DANG, Azeddine BEGHDAI, and Mohamed-Chaker LARABI*

L2TI, Institut Galilée, Université Paris 13, France

*XLIM, Dept. SIC, Université de Poitiers, France

{dang.thanhtrung, azeddine.beghdadi}@univ-paris13.fr, chaker.larabi@univ-poitiers.fr

Abstract. Image inpainting is not only the art of restoring damaged images but also a powerful technique for image editing e.g. removing undesired objects, recomposing images, etc. Recently, it becomes an active research topic in image processing because of its challenging aspect and extensive use in various real-world applications. In this paper, we propose a novel efficient approach for high-quality and fast image restoration by combining a greedy strategy and a global optimization strategy based on a pyramidal representation of the image. The proposed approach is validated on different state-of-the-art images. Moreover, a comparative validation shows that the proposed approach outperforms the literature in addition to a very low complexity.

Keywords: inpainting, image completion, image restoration, hierarchical restoration, pyramidal representation

1 Introduction

Image inpainting, also known as blind image completion, is not only the art of restoring damaged images; but also a powerful technique in many real-world applications, such as image editing (removing undesired objects, restoring scratches), film reproduction (deleting logos, subtitles, and so on), or even creating artistic effects (reorganizing objects, smart resizing of images, blending images). Recently, it becomes an active research topic in image processing because of its challenging aspect and extensive use in various real-world applications. This topic began by skillful and professional artists in museum to manually restore the old painting.

Digital image inpainting tries to mimic this very precise process in an automatic manner on computers. Because the completion is performed blindly without reference to original images, the aim of digital image completion is only restoring the damaged image by maintaining its naturalness, i.e undetectable by viewers. However, this task is extremely difficult in the case of high resolution and structured images. On the one hand, the restored parts should not be visible or perceptually annoying to human viewers when filled; on the other hand, the used algorithm needs to be robust, efficient and requiring minimal user interactions and quick feedbacks.

An image inpainting algorithm often works in two stages. First the missing or damaged regions are identified (inpainting regions or target regions). Second, these regions are filled in the most natural manner possible. Up to now, there is no approach for automatically detecting damaged regions to be restored. For the sake of simplicity, they are usually marked manually using image editing softwares. Several approaches have been proposed in the literature and they may be categorized into two main groups [1]: geometry-oriented methods and texture-oriented methods.

The methods of the first group are designed to restore small or thin regions such as scratches or blotches, overlaid text, subtitles, etc. In this group, the image is modeled as a function of smoothness and the restoration is solved by interpolating the geometric information within the adjacent regions into the target region. Approaches falling in this category show good performance in propagating smooth level lines or gradient but they have the tendency to generate synthesis artifacts or blur effects in the case of large missing regions [2–4].

Whereas, the objective of the methods in the second group is to recover larger areas where the texture is assumed to be spatially stationary. Texture is modeled through probability distribution of the pixel brightness values. The pixel intensity distribution depends on only its neighborhood. This group could be further subdivided into two subgroups named: greedy strategy [5–8] and global optimization strategy [10–12]. Greedy strategies have acceptable computation time and take into account human perception features (priority is designed based on the salient structures considered as important for human perception). However, some problems such as local optimization and patch selection may limit the efficiency of these approaches. In contrast, global optimization strategies often provide better results. But, they are computationally expensive. This is mainly due to the fact that time complexity increases linearly both with the number of source pixels and unknown pixels.

In this study, we propose a novel approach for high-quality and fast image completion by combining both greedy and global optimization strategies based on a pyramidal representation of the image [13]. The use of pyramidal representation is twofold: first it allows accounting for the multi-scale characteristics of the HVS; second it offers a good way to accelerate the completion process. It is worth noticing that a perceptual pyramidal representation [16] would be better but at the expense of increased computational complexity.

The proposal is directed by the observation that the human visual system is more sensitive to salient structures being stable and repetitive at different scales. Also, a hierarchical completion is a suitable solution for preserving high frequency components in a visually plausible way, and thus generates high-quality outputs. Namely, a top-down completion is implemented from top level (the lowest resolution) to the bottom level (the original resolution). A greedy algorithm is applied for the lowest resolution to complete the damaged regions and create a good initialization accounting for the human perception for the next level. At each higher level, a relation map, called shift-map, is interpolated from adjacently lower level and then optimized by a global optimization algorithm,

i.e. multi-label graph-cuts [12, 14]. Experimental results highlight a noticeable improvement in both implementation performance and quality of the inpainted image. To affirm the performance of our implementation, the running time is calculated in comparison with some typical inpainting methods. To confirm the quality of our results, the viewer can visually evaluate outputs of inpainting approaches in conjunction with some objective inpainting quality metrics [17, 18].

The rest of the paper is organized as follows. More details of our framework are introduced in section 2. Section 3 is dedicated to experimental results and comparison with the state-of-the-art methods. Finally, this paper ends with some conclusions and future works.

2 Our proposal

The inpainting problem could be considered as an optimal graph labeling where a shift-map represents the selected label for each unknown pixels and it could be solved by optimizing an energy function using multi-label graph cuts. Because an unknown pixel in the damaged regions could originate from any pixel in the source regions, the global optimization strategies can be computationally infeasible. Moreover, they consider fairly possible label assignments but this does not fit with human perception. In term of inpainting quality, fair assignments may lead to unexpected bias for optimization. In terms of speed, a huge label set requires high computational load.

Our method is designed to overcome these limitations. In order to reduce the memory and computational requirements, a hierarchical approach for optimizing the graph labeling is developed. This hierarchy could provide enough-good results for inpainting problem, even though optimality cannot be guaranteed. In order to take into account human perception, a greedy strategy is applied at the lowest resolution to generate a suitable initialization for the next pyramidal levels. The priority of greedy strategy is designed based on the salient structures considered as one of the most important features for the HVS. An algorithmic description of our framework is given in the Fig. 1

For details, some notations that are similar to those in paper [7] are adopted. The whole image domain, I , is composed of two disjoint regions: the inpainting region (or target region) Ω , and the source region Φ ($\Phi = I - \Omega$). According to the above idea, a set of images G_0, G_1, \dots, G_N with various levels of details is generated using pyramidal operators, where $G_0 = I$ is the input or original image [13]. The inpainting regions are also reduced to the eliminated areas level by level.

2.1 Lowest level completion

In order to take into account HVS properties, a greedy strategy is applied for the lowest resolution. In our framework, an extension of [7] is developed to complete the reduced inpainting image. The algorithm for a single resolution image repeats the following steps (Fig. 2):

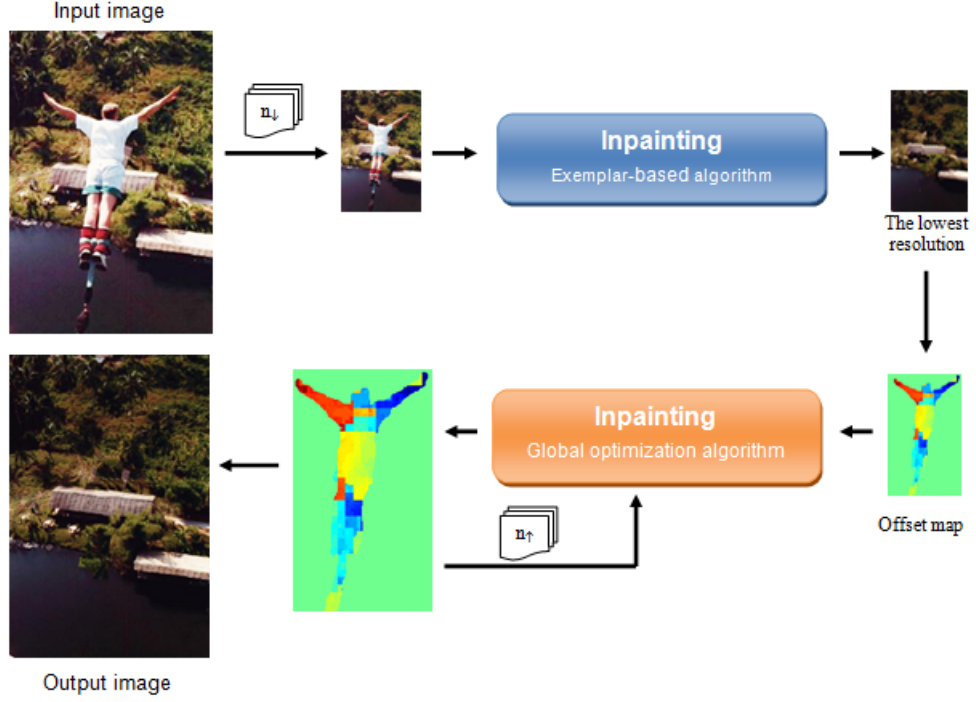


Fig. 1. Our scheme

1. *Initialization*: Identify inpainting boundary, $\delta\Omega$. If there is no pixel on the boundary, the algorithm is terminated.
2. *Priority estimation*: Compute the priority, $P(p)$, for all pixels on boundary, $p \in \delta\Omega$ and select randomly a pixel p with the highest priority.
3. *Patch match*: Find the patch or window Ψ_q that is most similar to Ψ_p thus minimizing mean squared error with existing pixels.
4. *Patch filling*: Fill the missing information in patch Ψ_p by copying the corresponding pixels from patch Ψ_q .
5. *Update*: Update the shift-map, SM_N , defining the relation between filled pixels and their sources and return to the step 1 for next iteration.

In this strategy, a good priority definition is very important because a decision taken based on it could not be changed anymore. Many models for priority have been proposed in the literature [5–9]. In this work, we used the priority model proposed in [7], namely window-based priority, which is more robust than the others. After inpainting the image at the lowest resolution, a complete shift-map is generated and used as an initialization for the completion of next levels.

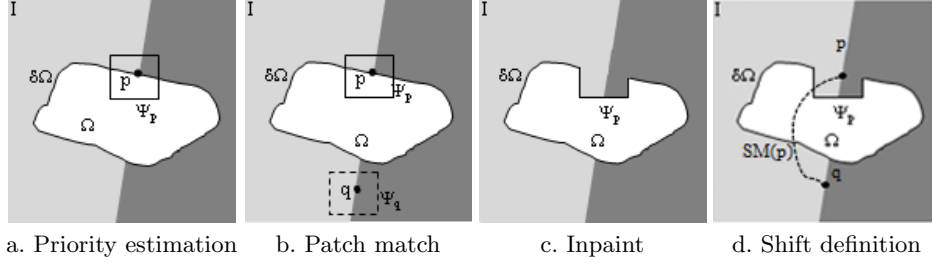


Fig. 2. The greedy strategy

2.2 Higher level completion

Since the principle of inpainting is to fill in unknown pixels $(p(x_p, y_p) \in \Omega)$ using the most plausible source pixels $(q(x_q, y_q) \in \Phi)$, a relationship between them needs to be defined. This relation can be characterized by a shift-map determining an offset from known pixel to unknown one for each coordinate in the image (Fig. 3b). The shift-map can be formulated by eq. (1). Then the output pixel $O(p)$ is derived from the input pixel $I(p + SM(p))$.

$$SM(p) = \begin{cases} (\Delta x, \Delta y) & p(x, y) \in \Omega \\ (0, 0) & \text{otherwise} \end{cases} \quad (1)$$

The naturalness of the resulting image is one of the most important issue of inpainting. Therefore, the used shift-map has to comply with such a requirement. In [12], authors proposed a solution to evaluate the shift-map by designing an energy function and optimizing it by a graph-cut algorithm. The energy function is defined as follows:

$$EM = \alpha \sum_{p \in \Omega} E_d(SM(p)) + (1 - \alpha) \sum_{(p, q) \in NB} E_s(SM(p), SM(q)) \quad (2)$$

Where E_d is a data term providing external requirements and E_s is a smoothness term defined over a set of neighboring pixels, NB . α is a user defined weight balancing the two terms fixed to $\alpha = 0.5$ in our case. Once the graph and energy function are given, the shift-map labeling is computed using multi-label graph-cuts algorithm [14, 15].

A. Data Term

The data term E_d is used to include external constraints. Because the unknown pixels are filled thanks to the known ones, the data term assumes that no pixels in the hole are used in the output image. The detail of the data term is given by Eq. (3):

$$E_d(SM(p)) = \begin{cases} \infty & (x + \Delta x, y + \Delta y) \in \Omega \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

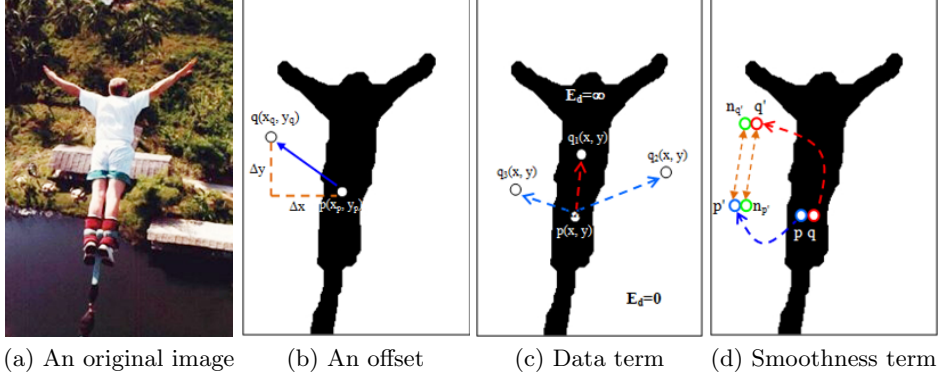


Fig. 3. Algorithm Operators

In some cases, the specific pixels in the input image can be forced to appear or disappear in the output image by setting $E_d = \infty$. For example, saliency map can be used to weight the data term. Therefore, a pixel with a high saliency value should be kept and a pixel with a low saliency value should be removed (Fig. 3c).

B. Smoothness Term

The smoothness term represents discontinuity between two neighbor pixels $p(x_p, y_p)$ and $q(x_q, y_q)$. In paper [12], the authors proposed an effective formula for smoothness term which takes into account both color differences and gradient differences between corresponding spatial neighbors in the output image and in the input image to create good stitching. This treatment is represented as eq. (4) (Fig. 3d):

$$E_s(SM(p), SM(q)) = \begin{cases} 0 & SM(p) = SM(q) \\ \beta \delta M(SM(p)) + \gamma \delta G(SM(p)) & otherwise \end{cases} \quad (4)$$

where β and γ are weights balancing these two terms, set to $\beta = 1$, $\gamma = 2$ in our experiment. δM and δG denote the differences of magnitude and gradient and they are defined as the follows:

$$\begin{aligned} \delta M(SM(p)) &= \|I(n_{p'}) - I(q')\| + \|I(n_{q'}) - I(p')\| \\ \delta G(SM(p)) &= \|\nabla I(n_{p'}) - \nabla I(q')\| + \|\nabla I(n_{q'}) - \nabla I(p')\| \end{aligned} \quad (5)$$

where, I and ∇I are the magnitude and gradient at these locations. $p' = p + SM(p)$ and $q' = q + SM(q)$ are locations used to fill pixels p and q , respectively. $n_{p'}$ and $n_{q'}$ are two 4-connected neighbors of p' and q' , respectively (Fig. 3d).

2.3 Shift-map Interpolation

A full shift-map is first inferred from a completion at the lowest level of pyramid. Then it is interpolated to higher resolutions using a *nearest neighbor interpolation*, and the shift-map values are *doubled* to match the higher image resolution.

At the higher level, only small shifts relative to the initial guess are examined. It means that only some parent neighbors are considered instead of all possible labels. In our implementation, the shift relative for each coordinate varies in range $[-a, a]$, so it takes $(2a + 1)^2$ labels for both direction. It is important to note that the data and smoothness terms are always computed with respect to the actual shifts and not to the labels (Fig. 4).

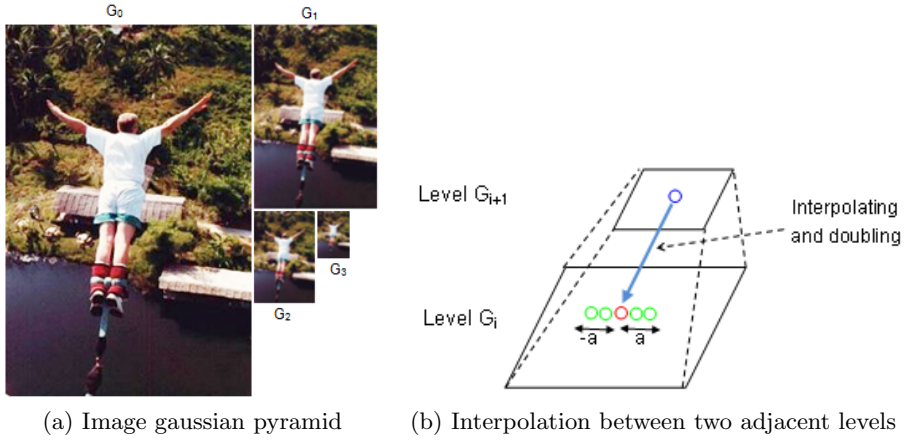


Fig. 4. Interpolation of Shift-Map.

3 Experimental Results

This section is dedicated to the study of performance of the proposed algorithm using some typical real images that cover several major challenges for inpainting. In order to try and cover all inpainting methods would be infeasible. For the sake of comparison with literature, three inpainting methods corresponding to algorithms proposed by *A. Criminisi et al* [5] and *T. T. Dang et al* [7] for greedy strategy and *Y. Pritch et al* [12] for global optimization strategy have been implemented. Five images, given on Fig. 6 were chosen for this experiment (including *bungee* (206×308), *angle* (300×252), *silenus* (256×480), *boat* (300×225) and *seaman* (300×218)).

Figure 6 illustrates the results obtained with the proposed approach in comparison to the others. Fig. 6a gives images to be inpainted where damaged areas cover respectively 12.6%, 5.83%, 7.74%, 10.73% and 14.87% of the whole image.

To evaluate the quality of inpainting output, some objective inpainted image quality metrics [17, 18] are considered and the metric in [18] is developed because all used images in our experiment are color. The metric values are shown in the table 1 and compared more visually in figure 5.

Table 1. The inpainted image quality metrics

Image	<i>bunjee</i>	<i>angle</i>	<i>silenus</i>	<i>boat</i>	<i>seaman</i>
Size	(206 × 308)	(300 × 252)	(256 × 480)	(300 × 225)	(300 × 218)
Damaged Area	12.6%	5.83%	7.74%	10.73%	14.87%
A. Criminisi [5]	0.0685	0.0817	0.0358	0.061	0.0449
T. T. Dang [7]	0.1157	0.0898	0.0407	0.065	0.0572
Y. Pritch [12]	0.0343	0.0805	0.0289	0.0597	0.0407
Our proposal	0.107	0.087	0.0407	0.069	0.0592

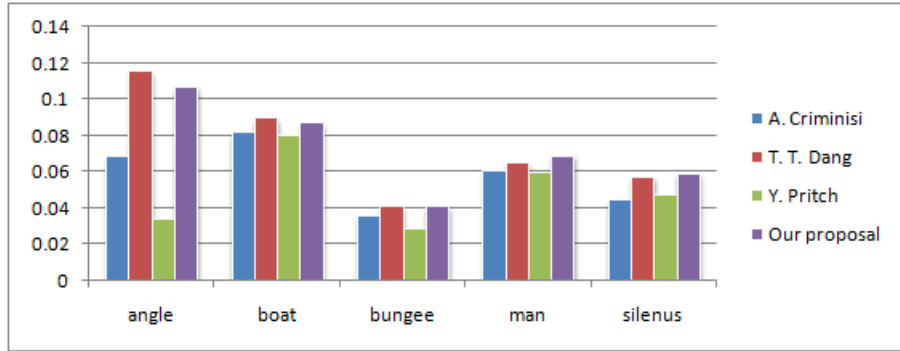


Fig. 5. A chart of quality performance.

The performance of the proposed approach is quantitatively evaluated by implementation time in comparison with the other approaches. In order to avoid bias, all approaches are programmed by the same programming language, C/C++ programming language, and implemented on the same PC with the configuration of Intel Core i5, 2.8GHz CPU and 4GB RAM. The running time in seconds of each methods is given in table 2 and shown visually in figure 7. As it can be seen from these results, our method provides an acceptable visual quality, often outperforming the others, with a much faster implementation. Indeed, visual inspection of results shows that the completion performed by our approach looks more natural and more coherent than the other approaches.

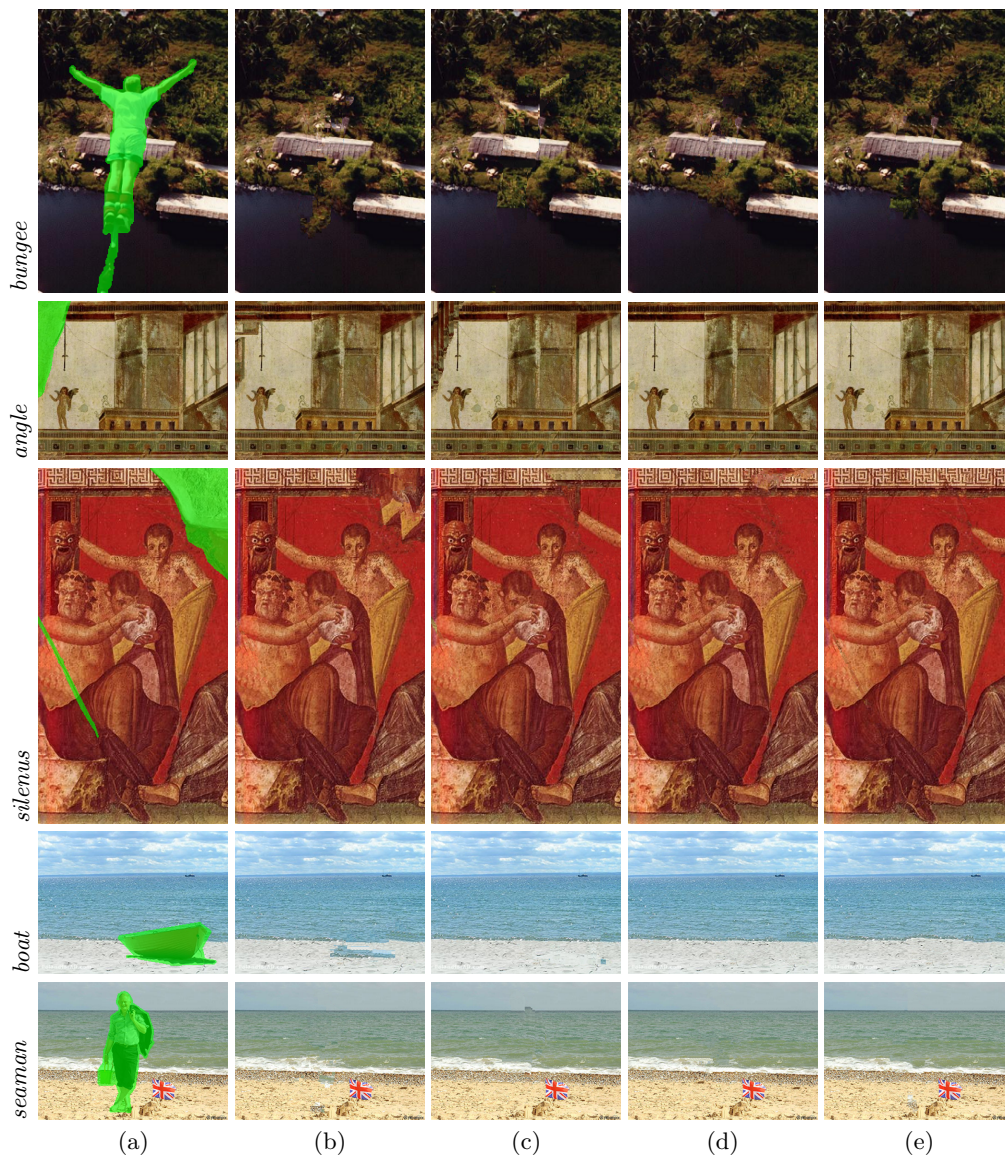


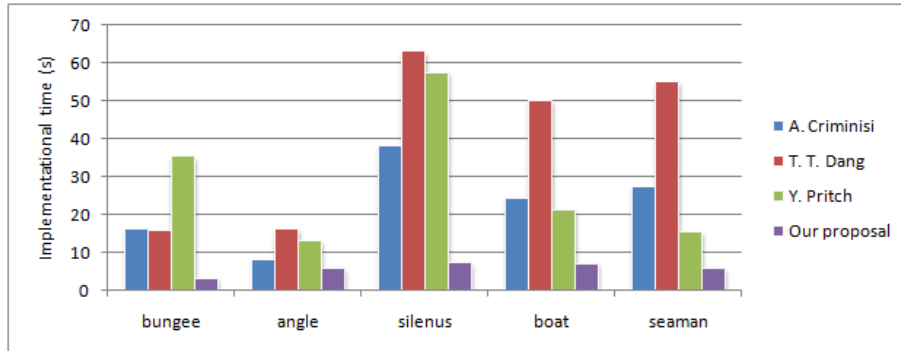
Fig. 6. The experimental results. (a) Image to be inpainted; The outputs when using the methods in (b) [5]; (c) [12]; (d) [7]; (e) our proposal;

4 Conclusions

In this paper, a novel framework of image completion is introduced by combining both greedy and global optimization strategies based on a pyramidal representation of the image. The greedy strategy is applied at the lowest resolution

Table 2. Computational time (in second) for implemented approaches and the set of used images

Image	<i>bungee</i>	<i>angle</i>	<i>silenus</i>	<i>boat</i>	<i>seaman</i>
Size	(206 × 308)	(300 × 252)	(256 × 480)	(300 × 225)	(300 × 218)
Damaged Area	12.6%	5.83%	7.74%	10.73%	14.87%
A. Criminisi [5]	16.30	8.20	38.29	24.54	27.31
T. T. Dang [7]	15.92	16.36	63.18	50.18	55.16
Y. Pritch [12]	35.39	13.24	57.68	21.18	15.50
Our proposal	3.32	5.81	7.53	7.25	5.97

**Fig. 7.** A chart of implementation performance.

in order to generate a good initialization accounting for human perception. At higher resolutions, the shift map is refined by a global optimization algorithm and multi-label graph-cuts. A comparison with some representative approaches from literature belonging to the second group (i.e. global optimization) is carried out and results show that our approach not only produces better quality of output images but also implements noticeably faster.

The obtained results are very encouraging and a more thorough evaluation procedure, including both objective and subjective evaluation, will be engaged as a future work. Computational complexity issues will be also addressed.

References

1. Arias, P., Facciolo, G., Caselles, V., Sapiro, G. : A Variational Framework for Exemplar-Based Image Inpainting. *International Journal of Computer Vision*, pp. 1-29, (2011).
2. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C. : Image inpainting *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 417 - 424, (2000).
3. Chan, T.F., Shen, J. : Non-texture inpainting by Curvature-Driven Diffusions (CCD). *Journal of Visual Communication and Image Representation*, vol 4. pp. 436 - 449, (2001).

4. Tschumperle D. : Fast anisotropic smoothing of multi-valued images using curvature-preserving pdes. *International Journal of Computer Vision*, vol. 68, pp. 65 - 82, (2006).
5. Criminisi, A., Perez, P., Toyama, K. : Region filling and object removal by exemplar-based image inpainting. *IEEE Transaction of Image Process. IEEE*, vol. 13 (9), pp. 1200 - 1212, (2004).
6. Wu, J., Ruan Q. : Object removal by cross isophotes exemplar based image inpainting. *Proceeding of International Conference of Pattern Recognition*, pp. 810 - 813, (2006).
7. Dang, T. T., Larabi M. C., Beghdadi, A. : Multi-resolution patch and window-based priority for digital image inpainting problem. *3rd International Conference on Image Processing Theory, Tools and Applications*, pp. 280 - 284, (2012).
8. Zhang, Q., Lin, J. : Exemplar-based image inpainting using color distribution analysis. In *Journal of Information Science and Engineering*, (2011).
9. Cheng, W., Hsieh, C., Lin, S., Wang, C., Wu, J. : Robust algorithm for exemplar-based image inpainting. *Proceeding of International Conference on Computer Graphics, Imaging and Visualization.*, (2005).
10. Wexler, Y., Shechtman, E., Irani, M. : Space-time video completion. *IEEE Transactions Pattern Analysis and Machine Intelligence.*, vol. 29, pp. 463 - 476, (2007).
11. Komodakis, G.T. N., Tziritas, G. : Image completion using global optimization. *Proceeding of IEEE Computer Society Conference Computer Vision and Pattern Recognition*, pp. 442 - 452, (2006).
12. Pritch, Y., Kav-Venaki, E., Peleg, S. : Shift-map image editing. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 151 - 158, (2009).
13. Peter, J. B., Edward, H. A. : The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, vol. 31, pp. 532 - 540, (1983).
14. Boykov, Y., Veksler, O., Zabih, R. : Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23 (11), pp. 1222 - 1239, (2001).
15. Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D., Cohen, M. : Interactive Digital Photomontage *Proceedings of SIGGRAPH*, pp. 294 - 302 , (2004).
16. Iordache, R., Beghdadi, A., de Lesegno, P. V. : Pyramidal perceptual filtering using Moon and Spencer contrast *International conference on image processing, ICIP'2001*, pp. 146 - 149, (2001).
17. Dang, T. T., Beghdadi, A., Larabi, M.C : Perceptual evaluation of digital image completion quality *21st European signal processing conference, EUSIPCO 2013*, (2013).
18. Dang, T. T., Beghdadi, A., Larabi, M.C : Perceptual quality assessment for color image inpainting *IEEE International conference on image processing, ICIP'2013*, (2013).