Signal Processing 103 (2014) 127–141

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

A perceptual image completion approach based on a hierarchical optimization scheme

T.T. Dang [a], A. Beghdadi [a,*], M.-C. Larabi [b]

[a] L2TI, Université Paris 13, Sorbonne Paris Cité, France
[b] XLIM, Dept. SIC, Université de Poitiers, France

ARTICLE INFO

Article history:
Received 18 July 2013
Received in revised form
14 October 2013
Accepted 29 November 2013
Available online 20 December 2013

Keywords:
Image inpainting
Image completion
Hierarchical representation
Greedy algorithm
Graph cuts

ABSTRACT

This paper aims at introducing a novel efficient approach for high-quality and fast image restoration by combining both a greedy strategy and a global optimization strategy, based on a pyramidal representation of the image. A coarse version of the input image is first restored by exemplar-based method using a greedy strategy. From the low-resolution inpainted image, higher resolutions are interpolated and refined by a global optimization strategy. Experimental results on natural images demonstrate the effectiveness of the proposed method. Moreover, a comparison with some methods of the state-of-the-art confirms its superiority in terms of image quality and computational time.

1. Introduction

Image inpainting, also known as blind image completion, refers to the action of filling missing parts or objects in an image. During the last decade, it becomes a very important research topic in the field of computer vision and image processing because of its suitability for plenty of professional and consumer applications such as (i) application of digital effects (e.g. removing undesired objects or logos), (ii) image restoration (e.g. deleting scratches or blotches in old photographs), and (iii) image coding and transmission (e.g. recovering missing blocks, error concealment).

Since, the targeted completion is performed blindly and without any cue about what would be the original content, the focus is put on restoring the damaged parts/objects by maintaining as much as possible the naturalness of the image. Moreover, the restored parts should not be visible or perceptually annoying to human viewers and the used algorithm needs to be robust, efficient and requiring minimal user interactions or quick feedbacks.

To date, several approaches of image inpainting have been proposed in the literature [1]. The pioneering inpainting approach is the diffusion based in which the image is modeled as a physical medium where the diffusion process takes place [2–5]. The completion is performed by interpolating the image information from the known region into the missing region at the pixel level. This approach tends to produce a blurring effect in textured and structured regions. In their work, Bertalmio et al. [2] proposed an algorithm for object removal that inwardly propagates information from the boundaries of the selected object, in a smoothly manner. This approach tries to reproduce real techniques performed by professional restorers. It is based on fluid dynamics models using a third order partial differential equations (PDE) [5]. Bornemann and Marz proposed in [6] an improvement of the Bertalmio's approach by modifying the weight function and replacing the edge-oriented transport direction method by the coherence direction. In the same vein, Chan and Shen proposed an inpainting model relying on a variational framework based on the total variation to recover the missing information [4]. They also introduced a new PDE-based inpainting approach exploiting curvature driven diffusion. The literature on inpainting algorithms involving variational or PDE approaches is relatively rich and various. Nevertheless,

* Corresponding author.
E-mail address: azeddine.beghdadi@univ-paris13.fr (A. Beghdadi).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.11.036

it is commonly agreed that the one proposed by Tschumperlé, based on an efficient second-order anisotropic diffusion model for multi-valued image regularization [3], is among the most efficient when parameters are carefully selected. However, there are no criteria for tuning the parameters or objective measure for performance evaluation. Finally, even though the methods falling in this approach are very efficient for untextured and relatively small region, they show some important drawbacks due to their incapacity to restore texture and tend to introduce blurring effect in the case of large missing regions.

The second category of approaches is the exemplar-based algorithms, in which texture is modeled through probability distribution of the pixel brightness values. This approach is inspired from texture synthesis techniques proposed by Efros and Leung [7] and improved by Ashikhmin [8] with the aim of reducing the computational cost of patches matching based on the notion of coherence, or Kwatra et al. [9,10] for patch synthesis based on graph-cuts. However, natural images are composed of structures and textures, in which the structures refer to edges or contours and the textures are image regions with homogeneous patterns or feature statistics (including flat patterns). This is why pure texture synthesis techniques cannot be efficiently applied to missing regions/objects with composite textures and structures. Authors in [11] decomposed the image into structure and texture components. Then, the restoration is performed simultaneously and independently on each component by means of geometry and texture oriented methods, respectively. The developed approach relies on a number of parameters and constraints such as contour preservation, structure anisotropy and the number of iterations. Indeed, this approach allows avoiding blurring effect observed with the diffusion approaches. However, recovering missing large regions is still a challenging issue.

In such a process, neighboring known pixels around the missing region or the object to be removed is an important source of the most relevant information regarding the target region. So, the natural way suggests to start filling the target region inwardly in an onion-peel fashion. Instead, a patch propagation based on patch priority is proposed in [12] to encourage the filling-in of patches on the structures. Several improvements have been introduced for patch priority such as cross-isophotes patch priority in [13], color distribution priority in [14], patch sparsity in [15] as well as for patch synthesis such as non-local means [15,16]. Generally, these approaches are known as greedy strategies and they have acceptable computation time in comparison to diffusion approaches. Moreover, these approaches try to exploit some perceptual features of the human visual system (HVS) [17]. For instance, priority is designed based on the salient structures considered as important for human perception. It is hence known that one of them is the geometrical information that plays an important role in image homogeneity and coherence [18,19]. However, these approaches show some common problems related to local optimization, patch priority, patch selection and so on.

Besides the aforementioned technique, inpainting can be considered as a global optimization problem that can be solved by minimizing a coherence measure [20] or energy functions of smoothness [21,22] or bidirectional similarity [23,24]. Global optimization strategies often provide better results in comparison to other strategies but at the cost of a higher computational complexity. The latter is mainly due to the fact that time complexity increases linearly with both the number of source pixels and unknown pixels.

Recently, more general sparse image representations using dictionaries [25,26] have proven their efficiency in the context of inpainting [27,28]. The use of dictionary learning based approaches for image enhancement is one of the promising approaches. The idea of this approach is to represent an image by a sparse combination of an over complete set of transforms. Then, missing pixels are inferred by adaptively updating this sparse representation [29]. However, similar to the diffusion-based approach, this one may fail in recovering structures or may introduce a smoothing effect when filling large missing regions.

Although tremendous progress has been achieved on image inpainting over recent years, there are still significant challenges. For instance, image completion of large missing regions is one of the major open problems. The computational complexity is also one of the main issues to consider in image inpainting. In this study, we propose a novel approach for high-quality and fast image completion by combining and leveraging the benefits of both greedy and global optimization strategies using a pyramidal representation of the image.

The proposal is directed by the observation that the human visual system is more sensitive to salient structures being stable and persistent at different scales. Therefore, a hierarchical image inpainting scheme is developed in order to control and preserve salient features during the completion process. This scheme allows restoring the missing regions in a visually plausible way. A top-down completion is implemented from the top level (the lowest resolution) to the bottom level (the original resolution). Following this strategy, a greedy algorithm, based on the idea introduced in [30], is developed for the lowest resolution to restore the damaged region. It provides a good initialization, accounting for human perception, at the higher resolution. It is worth noticing that, since the low-resolution inpainted image has a critical impact on the quality at the final output, caution should be taken when using the algorithm proposed in [12]. Therefore, the inpainting algorithm in [12] is first improved by evaluating the filling-in priority with all pixels in the patch. This makes it different from simply using the gradient-based or cross-isophotes-based priority as proposed in [12,13] while taking advantage of a lower complexity than those proposed in [14,15].

The impact of patch synthesis terms on the quality of the inpainted images is also studied by using recent Image Inpainting Quality Assessment (IIQA) metrics [31,32]. The similarity measurement based only on color channels is insufficient to propagate accurate linear structures into the target region and leads thus to garbage growing. This comes from the observation that the HVS is sensitive to not only the intensity of a spectral color but also to the context in which it is observed [33]. To maintain this variation, a new term representing image gradient is introduced as a weighting parameter in the computation