Optik 126 (2015) 3978–3981

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Exemplar based inpainting in a multi-scaled space

Baek-Sop Kim [a], JunSeong Kim [b], Jaehwa Park [b,*]

[a] Department of Computer Engineering, Hallym University, Chuncheon, Republic of Korea
[b] School of Engineering, Chung-Ang University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 4 July 2014
Accepted 26 July 2015

Keywords:
Image inpainting
Exemplar based
Image pyramid

ABSTRACT

A novel exemplar based inpainting approach running on multi-scaled space is presented. The presented approach focuses on designing a novel processing architecture that works on a multi-resolution hierarchy. An inpainting process is designed to run recursively on the Laplacian image pyramid from coarse to fine level. Through the recursive process running on the image pyramid, both of structural features and texture informations near the target regions are permeated in the restored region. It reduces the artifact effects caused by exhaustive patch searching of conventional methods which work on a single layer. More plausible inpainting results and improvement on the processing speed are achieved when the presented method is applied in the experiments.

1. Introduction

Image inpainting is a technique that plausibly restores area of removed objects or corrupted region of an image [1]. The plausibility of inpainting results is very subjective and objective references or metrics are difficult to be defined. To produce good results, an inpainting technique restores target regions to be corresponded to the neighboring regions as smoothly as possible. Both geometrical and color information of the known region should be propagated into the target region in a given image.

Most of previous works have been focused on how target regions can be extrapolated with the local features of the neighboring regions. They are categorized in *synthesis* and *exemplar* schemes, according to how the target region is filled. The synthetic techniques fill the target region with artificially generated texture patterns. These approaches are developed from texture synthesis, which seeks to replicate texture from a small source sample of pure texture, inspired by mathematical approaches [2,3].

In exemplar scheme, known part of image serves as the source of target region filling [4]. The properties of known regions near the target regions are extracted and restoration priority along with the border line of the target region is determined from the extracted properties. Then patches are selected from the known region of image and the target region is restored with them.

If target region is small, most of conventional techniques produce satisfactory results. However, if target region is large, the restoration results are not so plausible. In case that global structures are overlaid throughout the large target regions, a successful result is seldom obtained. A method which captures the global structural features is necessary to enhance the restoration plausibility.

Successful works varied some part of the steps with supplementary algorithms, such as reorder the filling to ensure linear structures [4], using Bezier curve to reconstruct the line structures [5], alternating the priority function into a form of linear combination rather than multiplication [6], adaptive sizing or generating patch by weighted linear combinations [7].

Generally two features, global structure and localized texture, are considered to be the major factors for successful restoration. Structural features are known to be more sensitive to the human visual perception. But the texture features play also significant roles to generate successful results. If structural features are not aligned, incongruous contents are generated in the restored region. If texture features are not match, the boundary lines are exposed along the patch borders. In most cases, structural features are concentrated on low frequency range. Contrarily, texture features are distributed on high frequency range.

A recursive processing architecture working on a multi-scaled space for exemplar based inpainting is presented. In the architecture, inpainting processes works recursively on the Laplacian image pyramid from coarse to fine level. Through the recursive process, both of structural features and texture informations near the target regions are permeated in the restored region. The presented approach focus on designing a novel processing architecture

* Corresponding author. Tel.: +82 10 2461 5257.
E-mail addresses: bskim@hallym.ac.kr (B.-S. Kim), junkim@cau.ac.kr (J. Kim), jaehwa@cau.ac.kr (J. Park).

http://dx.doi.org/10.1016/j.ijleo.2015.07.168
0030-4026/

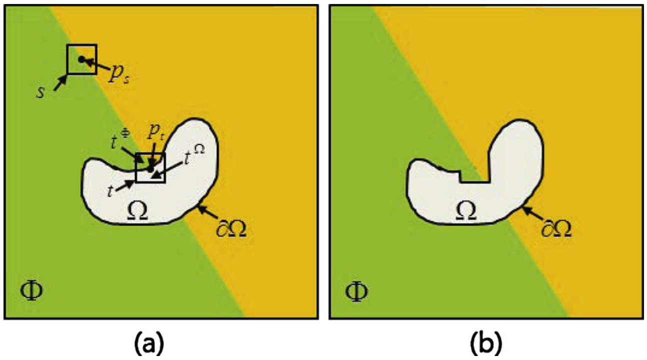

Fig. 1. Typical procedure of exemplar based inpainting.

that works on a multi-resolution hierarchy. Most of conventional inpainting methods are directly applicable to the architecture.

2. Exemplar based inpainting

In exemplar based image inpainting, the known part of the image serves as a source of target region filling [4]. As shown in Fig. 1(a), let Φ be a known region and Ω be a target region to be restored in an image. $\partial\Omega$ denotes the border line between Φ and Ω. Also let t be a target patch centered at a pixel p_t on $\partial\Omega$ with the pre-given size. t^{Φ} and t^{Ω} denote the known region and actual target region to be restored in t respectively. A source patch s centered at a pixel p_s on Φ which has the maximum similarity to t^{Φ} is selected and its corresponding pixels are copied onto t^{Ω} as shown in Fig. 1(b). This procedure keeps running iteratively till Ω is completely filled.

In most cases the iterative inpainting procedure is controlled by two functions, $P(p)$ and $S(p)$. $P(p)$ computes the restoration priority of all the pixels $p \in \partial\Omega$. $S(p)$ measures the similarity of a patch centered by $p \in \Phi$ to t^{Φ} which is the known part of t selected by P previously.

For an example, in [4] a restoration priority P on p is defined as

$$P(p) = C(p) \cdot D(p) \tag{1}$$

where

$$C(p) = \frac{\sum_{p \sqsubseteq t^{\Phi}} \rho(p)}{|t|} \tag{2}$$

$$D(p) = \frac{|\nabla I_p^{\perp} \cdot n_p|}{\alpha} \tag{3}$$

$C(p)$, known as the confidence term, is the ratio of known data in t where $\rho(p)$ is the confidence of p. If the pixel is initially known $\rho = 1$ otherwise $0 \le \rho < 1$. $|t|$ denotes the patch size. $D(p)$, known as the data term, is a gradient measurement on p where $\nabla I_p^{\perp}$ is the orthogonal vector of gradient vector, and n_p is orthonormal vectors of $\partial\Omega$ on p. α is a normalize parameter, usually 255 for 8-bit gray scale.

In Fig. 1(a), p_t is chosen to maximize P and the patching region t is determined. Then p_s and the patch s are selected from Φ which maximize (or minimize) S. Usually $S(p)$ is given by a patch distance between $t_{p_t}^{\Phi}$ and corresponding region in s. And its corresponding pixels of s are copied onto $t_{p_t}^{\Omega}$ as shown in Fig. 1(b). A patch distance between t and s is defined as,

$$d(t, s) = \frac{\sum_{r \in t^{\Phi}} d(r)}{|t^{\Phi}|} \tag{4}$$

where $d(r)$ is the difference between two corresponding pixels located on r in t and s. Euclidean distance on intensity of RGB values between two pixels is often used as a primary distance metric.

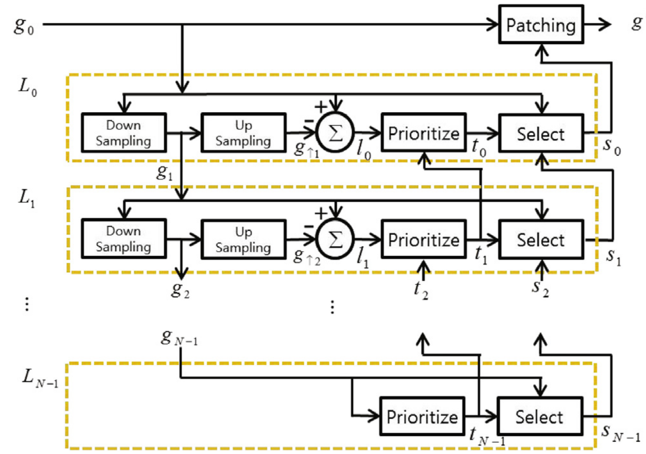

Fig. 2. Block diagram of recursive inpainting process in multi-scale.

Most of the previous studies have been focused on how P and S are defined to maximize the plausibility. Various features, such as gradient, contour, Hessian matrix, belief propagation, sparsity, etc. have been studied for P and S [8,9]. But all the formulations are somewhat rigid on balancing the structural and texture features and are bounded in single resolution.

Approaches using concepts of multi-resolution are presented in [10]. But the inpainting procedure is a conditional combination of multiple processes running on different resolutions independently. Most of restorations are done by the main process, which may be replaced by other processes for exceptional cases.

An hierarchical processing architecture working on a multi-scaled space is designed. In the architecture, the P and S select patches recursively on the Laplacian image pyramid from coarse to fine level. The presented approach does not focus on formulating alternative P and S but focuses on designing a hierarchical approach that works on a multi-resolution. Most of conventional Ps and Ss are directly applicable to the architecture.

3. Recursive inpainting in multi-scaled space

Let g_0 be an input image and g_1 be the down-sampled image. The low frequency information is remained in g_1, since the down sampling has an effect of low pass filtering. Let $g_{\uparrow 1}$ be the up-sampled image of g_1. The image size of $g_{\uparrow 1}$ is recovered but the high frequency information is not. If g_0 is subtracted by $g_{\uparrow 1}$, the Laplacian image, l_o is obtained as shown in Fig. 3, which has only the high frequency information of g_0 compared to $g_{\uparrow 1}$. If the same process is applied to g_1, l_1 is obtained, which has the band limited information in between lower frequency than l_0 but higher frequency than $g_{\uparrow 2}$. By recursion of $N-1$ times, a N-layered Laplacian image pyramid can be built [11]. If the down-sampling rate is 1/2, the image size of bottom layer reduces $2^{N-1} \times 2^{N-1}$ times.

Fig. 2 shows the recursive inpainting process running on a Laplacian image pyramid, which has N layers of $L_0 \sim L_{N-1}$ from top to bottom. Let t_n be the target patch on L_n which has the highest priority selected by P and let s_n be the source patch on L_n which is the most similar to t_n^{Φ} selected by S. Let $t_{\uparrow n}$ and $s_{\uparrow n}$ be the 2×2 times enlarged regions of L_{n-1}, corresponding to t_n and s_n of L_n, respectively.

The patch selection starts from the bottom layer. Since the bottom layer does not have the Laplacian image, the selection process starts with g_{N-1}. t_{N-1} and s_{N-1} are chosen by P and S on the entire $\partial\Omega_{N-1}$ and Φ_{N-1} of g_{N-1}. Then the process moves to L_{N-2}. The region of $t_{\uparrow N-1}$ and $s_{\uparrow N-1}$ corresponding to t_{N-1} and s_{N-1} are extracted from l_{N-1} and g_{N-1}, respectively. New patch regions t_{N-2}