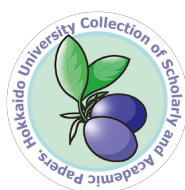




HOKKAIDO UNIVERSITY

Title	Perceptually Optimized Missing Texture Reconstruction via Neighboring Embedding
Author(s)	OGAWA, Takahiro; HASEYAMA, Miki
Citation	IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E98.A(8), 1709-1717 https://doi.org/10.1587/transfun.E98.A.1709
Issue Date	2015-08
Doc URL	https://hdl.handle.net/2115/59618
Rights	copyright©2015 The Institute of Electronics, Information and Communication Engineers
Type	journal article
File Information	Perceptually Optimized Missing Texture Reconstruction via Neighboring Embedding.pdf



Perceptually Optimized Missing Texture Reconstruction via Neighboring Embedding

Takahiro OGAWA^{†a)} and Miki HASEYAMA[†], *Members*

SUMMARY Perceptually optimized missing texture reconstruction via neighboring embedding (NE) is presented in this paper. The proposed method adopts the structural similarity (SSIM) index as a measure for representing texture reconstruction performance of missing areas. This provides a solution to the problem of previously reported methods not being able to perform perceptually optimized reconstruction. Furthermore, in the proposed method, a new scheme for selection of the known nearest neighbor patches for reconstruction of target patches including missing areas is introduced. Specifically, by monitoring the SSIM index observed by the proposed NE-based reconstruction algorithm, selection of known patches optimal for the reconstruction becomes feasible even if target patches include missing pixels. The above novel approaches enable successful reconstruction of missing areas. Experimental results show improvement of the proposed method over previously reported methods.

key words: missing texture reconstruction, quality assessment, SSIM index, neighboring embedding

1. Introduction

Missing area reconstruction has been intensively studied since it can afford a number of fundamental applications. For example, missing area reconstruction can be applied to restoration of old films corrupted by blotches and scratches, removal of unnecessary objects and error concealment that recovers missing blocks in videos transmitted in error-prone environments. As shown in [1], most previous works can be broadly categorized into two groups that aim at successful recovery of two important visual components, i.e., structure reconstruction approaches [2]–[5] and texture reconstruction approaches [6]–[26]. It should be noted that structure components correspond to edges included within images. In this paper, we focus on accurate reconstruction of missing textures and show its details.

In the field of missing texture reconstruction, Efros et al. proposed a pioneering method that uses texture synthesis for the recovery of missing textures [6]. This method is based on the Markov random field model, and missing intensity estimation becomes feasible by copying known pixels within a target image to missing pixels. Recently, Drori et al. and Criminisi et al. have proposed a fragment-based reconstruction method [7] and an exemplar-based reconstruction method [8], [9], respectively. The method in [8], [9] has become a benchmarking method in the field of missing texture

reconstruction, and faster versions have also been proposed by many researchers [10], [11]. In these previously reported methods, since a target patch whose textures are unknown is approximated by only one known patch, it becomes difficult to perform accurate texture representation when these textures include some changes within the target image.

Several reconstruction methods that approximate each local patch within a target image by low-dimensional subspaces have been proposed to improve the texture representation performance. From the characteristic that missing area reconstruction is one of the ill-posed inverse problems, the derivation of its solution becomes feasible on the basis of approximation using such low-dimensional subspaces. Therefore, the reconstruction performance depends on multivariate analysis algorithms adopted for obtaining these low-dimensional subspaces. A successful PCA-based method for reconstruction of missing areas based on the eigenspace of known patches was proposed [12]. Nonlinearly extended versions of that method using kernel PCA [27] have also been proposed [13], [14], and they include adaptive selection of optimal nonlinear eigenspaces. Furthermore, sparse representation-based image restoration has been intensively studied by many researchers [28], [29]. Based on the selection of optimal signal-atoms, adaptive generation of optimal subspaces becomes feasible and, thus, many missing area reconstruction methods based on sparse representation have been proposed [15]–[19]. Recently, several reconstruction methods that adopt the concept of neighboring embedding (NE) algorithms have also been proposed [20], [21]. These methods focus on the characteristic that target patches including missing areas can be linearly presented by their known nearest neighbor patches, and reconstruction is performed on the basis of non-negative matrix factorization (NMF) [30] and locally linear embedding (LLE) [31] using the selected nearest neighbors. In this paper, the nearest neighbors represent patches similar to target patches including missing areas. It should be noted that these similar known patches are not always close to the target patches spatially.

In previously reported methods, the representation of target patches including missing areas is performed on the basis of minimization of their approximation errors. Since the mean square error (MSE) is one of the simplest metrics that can be used as quality measures, it has been adopted to monitor the approximation error. However, it is known that the MSE and its variants tend not to reflect perceptual qualities [32], [33], and their use for image restoration is not al-

Manuscript received November 30, 2014.

Manuscript revised March 6, 2015.

[†]The authors are with the Graduate School of Information-Science and Technology, Hokkaido University, Sapporo-shi, 060-0814 Japan.

a) E-mail: ogawa@lmd.ist.hokudai.ac.jp

DOI: 10.1587/transfun.E98.A.1709

ways suitable. In particular, when the MSE is applied to reconstruction of missing textures, the obtained results suffer from some degradations such as over-smoothness since textures contain more high-frequency components than those in other regions. Another important point should also be noted. The previously reported methods perform the representation of target patches including missing areas by using several bases, which span low-dimensional subspaces. In most methods, these bases are selected by only monitoring differences or similarities in the other known areas within target patches. However, the use of only these known areas does not always provide the optimal ones for the following reconstruction procedures since the criterion for selection is independent of the reconstruction procedures.

In this paper, we present a perceptually optimized missing texture reconstruction method using an NE algorithm. Recently, many image quality assessment algorithms have been proposed [33]–[35], and the structural similarity (SSIM) index is a representative measure [35]. Due to its superiority to the MSE and its variants, this measure is widely used for several kinds of applications [24], [36]. Therefore, the proposed method tries to perform missing texture reconstruction on the basis of the SSIM-based NE algorithm, for which we adopt an LLE-based algorithm. As described above, there have been proposed many missing texture reconstruction methods based on various kinds of multivariate analysis techniques. Specifically, these methods perform dimensionality reduction using the multivariate analysis such as PCA and sparse representation to obtain a unique solution to the ill-posed problem estimating unknown image data via the approximation in the low-dimensional subspaces. In addition, it is known that nonlinear image representation is more suitable for approximating visual features compared to linear representation [13], [14], [27]. Thus, the approximation based on nonlinear multivariate analysis, i.e., kernel methods is desirable for our problem. Unfortunately, since it is not easy to perform nonlinear multivariate analysis whose object function is the SSIM, the calculation of its subspace is difficult. On the other hand, it has been reported in [31] that NE-based approaches enable linear representation in limited narrow ranges even if manifolds embedded in high-dimensional space have nonlinear structures. Therefore, the proposed method adopts the NE-based approach for realizing successful approximation in the low-dimensional subspace which focuses on this characteristic.

Furthermore, in our SSIM-based NE algorithm, a new scheme for selection of the known nearest neighbor patches, i.e., the known similar patches, is introduced by monitoring the SSIM index observed through the reconstruction algorithm. Since this SSIM-based criterion corresponds to the reconstruction performance observed from the other known areas within the target patch, the conventional problem of not being able to perform optimal basis selection closely related to the reconstruction procedures can be effectively solved. Consequently, based on the above non-conventional approaches, successful reconstruction of missing textures can be expected by the proposed method. The proposed

method can be also regarded as an improved version of the NE-based method [20], [21] by introducing the new image quality measure and the new nearest neighbor selection scheme.

It should be noted that several researchers and we have previously proposed missing texture reconstruction methods using the SSIM index [22]–[26]. In these methods, references [23] and [26] are PCA-based methods, and references [24] and [25] are sparse representation-based methods, where the quality measure of these methods is the SSIM. As described above, it is desirable that visual features are represented nonlinearly. Therefore, different from the above four methods, our method introduces the NE-based approach. Since the method in [22] performs linear approximation of target patches by using their neighboring similar patches, it corresponds to the most similar method to our method. However, this method monitors SSIM-based approximation performance only in known areas of the target patch, and the second problem of the previously reported methods cannot be solved. In this way, the proposed method is different from the previously reported SSIM-based methods in terms of the above points.

This paper is organized as follows. In Sect. 2, the definition of the SSIM index and its characteristics are shown as preliminaries. In Sect. 3, the new missing texture reconstruction method using the SSIM-based NE algorithm is presented. Then, in Sect. 4, the effectiveness of the proposed method is verified from several experimental results. Finally, conclusions are shown in Sect. 5.

2. Structural Similarity Index

Details of the SSIM index [35] are shown in this section. The SSIM index is a similarity measure between two signal vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ ($\in \mathcal{R}^n$), and its definition is shown as follows:

$$\text{SSIM}(\mathbf{x}_1, \mathbf{x}_2) = [l(\mathbf{x}_1, \mathbf{x}_2)]^\alpha \cdot [c(\mathbf{x}_1, \mathbf{x}_2)]^\beta \cdot [s(\mathbf{x}_1, \mathbf{x}_2)]^\gamma, \quad (1)$$

where the first two terms $l(\mathbf{x}_1, \mathbf{x}_2)$ and $c(\mathbf{x}_1, \mathbf{x}_2)$ in Eq. (1) respectively compare the mean and variance of the two signal vectors $\mathbf{x}_1$ and $\mathbf{x}_2$. Furthermore, $s(\mathbf{x}_1, \mathbf{x}_2)$ is the structural correlation between these two vectors. Specifically, the three terms in Eq. (1) are respectively calculated as

$$l(\mathbf{x}_1, \mathbf{x}_2) = \frac{2\mu_{\mathbf{x}_1}\mu_{\mathbf{x}_2} + C_1}{\mu_{\mathbf{x}_1}^2 + \mu_{\mathbf{x}_2}^2 + C_1}, \quad (2)$$

$$c(\mathbf{x}_1, \mathbf{x}_2) = \frac{2\sigma_{\mathbf{x}_1}\sigma_{\mathbf{x}_2} + C_2}{\sigma_{\mathbf{x}_1}^2 + \sigma_{\mathbf{x}_2}^2 + C_2}, \quad (3)$$

$$s(\mathbf{x}_1, \mathbf{x}_2) = \frac{\sigma_{\mathbf{x}_1, \mathbf{x}_2} + C_3}{\sigma_{\mathbf{x}_1}\sigma_{\mathbf{x}_2} + C_3}. \quad (4)$$

In the above equations, $\mu_{\mathbf{x}_i}$ and $\sigma_{\mathbf{x}_i}^2$ ($i = 1, 2$) are the mean and variance of $\mathbf{x}_i$, and $\sigma_{\mathbf{x}_1, \mathbf{x}_2}$ is the cross covariance between the two vectors. The constants C_1 , C_2 and C_3 are necessary for avoiding instability when the denominators are very close to zero. The parameters $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ determine the relative importance of the three components in

Eq. (1). By setting the parameters as $\alpha = \beta = \gamma = 1$ and $C_3 = \frac{C_2}{2}$, a simplified formulation of the SSIM index can be obtained as

$$\text{SSIM}(\mathbf{x}_1, \mathbf{x}_2) = \frac{(2\mu_{\mathbf{x}_1}\mu_{\mathbf{x}_2} + C_1)(2\sigma_{\mathbf{x}_1, \mathbf{x}_2} + C_2)}{(\mu_{\mathbf{x}_1}^2 + \mu_{\mathbf{x}_2}^2 + C_1)(\sigma_{\mathbf{x}_1}^2 + \sigma_{\mathbf{x}_2}^2 + C_2)}. \quad (5)$$

It is shown in [33], [35] that the SSIM index is superior to the MSE and its variants. Since the MSE tends to monitor the difference between low-frequency components, reconstruction based on the MSE outputs results suffering from over-smoothness. On the other hand, since the SSIM index, which is derived on the basis of the human visual system, monitors the contrast of the target image, it is a better quality measure. Therefore, it can be expected that the use of this similarity measure provides successful reconstruction of missing textures.

3. Missing Texture Reconstruction via the SSIM-based NE Algorithm

A missing texture reconstruction method using the SSIM-based NE algorithm is presented in this section. In our method, we clip a patch f ($w \times h = N$ pixels) including missing areas from the target image and estimate its missing intensities from the other known areas. For the following explanation, we respectively denote missing and known areas within the target patch f as Ω and $\bar{\Omega}$, respectively. Furthermore, two vectors, whose elements are intensities within f and $\bar{\Omega}$, are defined as $\mathbf{y} \in \mathcal{R}^N$ and $\mathbf{y}_{\bar{\Omega}} \in \mathcal{R}^{N_{\bar{\Omega}}}$, respectively, where $N_{\bar{\Omega}}$ is the number of pixels included in the area $\bar{\Omega}$.

In order to realize missing texture reconstruction using the SSIM-based NE algorithm, the proposed method must select optimal known patches suitable for representing the target patch f . It should be noted that the optimal patches correspond to the nearest neighbors of the target patch f , and these neighbors must be determined according to the proposed NE-based reconstruction algorithm. The algorithm for selection of optimal known nearest neighbors is shown in 3.1, and the SSIM-based NE reconstruction algorithm is presented in 3.2.

3.1 Algorithm for Selection of Optimal Known Nearest Neighbors

The algorithm for selection of optimal known nearest neighbors suitable for representing the target patch f is shown in this subsection. In the proposed method, we first clip known patches f_i ($i = 1, 2, \dots, L$), the size of which is the same as f , in a fixed interval and define their intensity vectors $\mathbf{x}_i \in \mathcal{R}^N$. From L patches, we select K known nearest neighbor patches suitable for reconstructing the target patch f .

In the following subsection, the target patch f is represented by a linear combination of the selected nearest neighbor patches maximizing the SSIM index. Therefore, the proposed method searches K patches providing the optimal linear combination. It should be noted that simultaneous selection of these K patches is an NP-hard problem. Therefore,

we adopt the simplest algorithm which is similar to several matching pursuit algorithms [37], [38] to select the optimal patches one-by-one. In the rest of this subsection, the details of k th ($k = 1, 2, \dots, K$) optimal known patch selection are shown.

In k th optimal patch selection, we perform approximation of $\mathbf{y}$ as follows:

$$\mathbf{y} \cong \mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)}, \quad (6)$$

where

$$\mathbf{X}_i^{(k)} = \begin{bmatrix} \mathbf{X}^{(k-1)} & \mathbf{x}_i \end{bmatrix} \quad (\in \mathcal{R}^{N \times k}), \quad (7)$$

$$\mathbf{a}_i^{(k)} = \begin{bmatrix} \mathbf{a}^{(k-1)} \\ a_i \end{bmatrix} \quad (\in \mathcal{R}^k), \quad (8)$$

and $\mathbf{X}^{(k-1)} (\in \mathcal{R}^{N \times (k-1)})$ is a matrix that contains $k-1$ vectors previously selected from $\mathbf{x}_i$ ($i = 1, 2, \dots, L$) in $k-1$ iterations. Furthermore, $\mathbf{a}_i^{(k)}$ is the coefficient vector for representing $\mathbf{y}$. As shown in the above equations, the proposed method selects the optimal one from $\mathbf{x}_i$ ($i = 1, 2, \dots, L$) in k th iteration. Specifically, we solve the following maximization problem:

$$\begin{aligned} \{\hat{\mathbf{y}}_i^{(k)}, \hat{\mathbf{a}}_i^{(k)}\} &= \arg \max_{\mathbf{y}_i^{(k)}, \mathbf{a}_i^{(k)}} \text{SSIM}(\mathbf{y}_i^{(k)}, \mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)}), \\ \text{subject to } \mathbf{M} \mathbf{y}_i^{(k)} &= \mathbf{y}_{\bar{\Omega}} \text{ and } \mathbf{1}_k' \mathbf{a}_i^{(k)} = 1, \end{aligned} \quad (9)$$

where $\mathbf{y}_i^{(k)}$ corresponds to an estimation result of $\mathbf{y}$, $\mathbf{M} (\in \mathcal{R}^{N_{\bar{\Omega}} \times N})$ is a binary matrix extracting only known intensities from the target patch f , $\mathbf{1}_k (\in \mathcal{R}^k)$ is a vector whose elements are all one, and vector/matrix transpose is represented by the superscript $'$. The proposed method performs optimization of Eq. (9) and selects the optimal nearest neighbor vector $\mathbf{x}_i$ ($i = 1, 2, \dots, L$) maximizing $\text{SSIM}(\mathbf{y}_i^{(k)}, \mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)})$, where $\text{SSIM}(\mathbf{y}_i^{(k)}, \mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)})$ is obtained according to Eq. (5), and it can be regarded as a function of $\mathbf{y}_i^{(k)}$ and $\mathbf{a}_i^{(k)}$ as follows:

$$\begin{aligned} \text{SSIM}(\mathbf{y}_i^{(k)}, \mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)}) &= \left(\frac{2\mu_{\mathbf{y}_i^{(k)}}\mu_{\mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)}} + C_1}{\mu_{\mathbf{y}_i^{(k)}}^2 + \mu_{\mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)}}^2 + C_1} \right) \left(\frac{2\sigma_{\mathbf{y}_i^{(k)}, \mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)}} + C_2}{\sigma_{\mathbf{y}_i^{(k)}}^2 + \sigma_{\mathbf{X}_i^{(k)} \mathbf{a}_i^{(k)}}^2 + C_2} \right) \\ &= S(\mathbf{y}_i^{(k)}, \mathbf{a}_i^{(k)}). \end{aligned} \quad (10)$$

The proposed method estimates the optimal vector $\hat{\mathbf{a}}_i^{(k)}$ with $\hat{\mathbf{y}}_i^{(k)}$ by simply applying the constrained cyclic coordinate ascent algorithm to Eq. (9). Although the cyclic coordinate ascent algorithm does not necessarily provide the global optimal solution in Eq. (9), we adopt it for reducing the complexity. By iterating the above procedures K times, we can select the K nearest neighbor patches based on the SSIM index. We denote K vectors of the patches selected from f_i ($i = 1, 2, \dots, N$) as $\mathbf{z}_k$ ($k = 1, 2, \dots, K$) and define $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K]$.

It is shown in [19] that the nearest neighbors obtained by monitoring only differences of intensities in the known areas $\bar{\Omega}$ tend not to find the true nearest neighbors optimal for reconstruction. This means that the nearest neighbors

selected from only the known information may not be optimal for reconstruction of the target missing areas. Furthermore, it is also shown in [19] that by monitoring the errors caused in the reconstruction, better nearest neighbors can be selected. Therefore, by introducing this idea into the SSIM-based reconstruction method, the proposed method enables selection of the nearest neighbor patches suitable for the target patch f .

3.2 Reconstruction Algorithm

In the previous section, it was shown that we can select the K nearest neighbor patches for the target patch f . In this subsection, we show the algorithm for reconstruction of the missing areas Ω within the target patch f based on the SSIM index. In the same manner as Eq. (9), the proposed method estimates the optimal linear combination $\mathbf{Za}$ approximating $\mathbf{y}$ by calculating the optimal vector of $\mathbf{a}$ ($\in \mathcal{R}^K$) as follows:

$$\begin{aligned} \{\hat{\mathbf{y}}, \hat{\mathbf{a}}\} &= \arg \max_{\mathbf{y}, \mathbf{a}} \text{SSIM}(\mathbf{y}, \mathbf{Za}), \\ &\text{subject to } \mathbf{M}\mathbf{y} = \mathbf{y}_{\bar{\Omega}} \text{ and } \mathbf{1}_K' \mathbf{a} = 1. \end{aligned} \quad (11)$$

In the above equation, $\text{SSIM}(\mathbf{y}, \mathbf{Za})$ is defined as

$$\begin{aligned} \text{SSIM}(\mathbf{y}, \mathbf{Za}) &= \left(\frac{2\mu_y \mu_{\mathbf{Za}} + C_1}{\mu_y^2 + \mu_{\mathbf{Za}}^2 + C_1} \right) \left(\frac{2\sigma_{y, \mathbf{Za}} + C_2}{\sigma_y^2 + \sigma_{\mathbf{Za}}^2 + C_2} \right) \\ &= \left(\frac{2\mu_y \mu_{\mathbf{Za}} + C_1}{\mu_y^2 + \mu_{\mathbf{Za}}^2 + C_1} \right) \\ &\quad \times \left[\frac{\frac{2}{N} (\mathbf{y} - \mu_y \mathbf{1}_N)' (\mathbf{Za} - \mu_{\mathbf{Za}} \mathbf{1}_N) + C_2}{\frac{1}{N} \|\mathbf{y} - \mu_y \mathbf{1}_N\|^2 + \frac{1}{N} \|\mathbf{Za} - \mu_{\mathbf{Za}} \mathbf{1}_N\|^2 + C_2} \right] \\ &= \left(\frac{2\mu_y \mu_{\mathbf{Za}} + C_1}{\mu_y^2 + \mu_{\mathbf{Za}}^2 + C_1} \right) \left(\frac{2\mathbf{y}' \mathbf{H}_N \mathbf{Za} + C_2}{\mathbf{y}' \mathbf{H}_N \mathbf{y} + \mathbf{a}' \mathbf{Z}' \mathbf{H}_N \mathbf{Za} + C_2} \right), \end{aligned} \quad (12)$$

where $\mu_y = \frac{1}{N} \mathbf{1}_N' \mathbf{y}$ and $\mu_{\mathbf{Za}} = \frac{1}{N} \mathbf{1}_N' \mathbf{Za}$ are the means of $\mathbf{y}$ and $\mathbf{Za}$, respectively. Furthermore, $\mathbf{H}_N \in \mathcal{R}^{N \times N}$ is the centering matrix defined as

$$\mathbf{H}_N = \mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N', \quad (13)$$

where $\mathbf{I}_N \in \mathcal{R}^{N \times N}$ is the identity matrix. The centering matrix $\mathbf{H}_N$ satisfies $\mathbf{H}_N' = \mathbf{H}_N$ and $\mathbf{H}_N^2 = \mathbf{H}_N$.

The similarity measure $\text{SSIM}(\mathbf{y}, \mathbf{Za})$ defined in Eq. (12) is a non-convex function of $\mathbf{y}$ and $\mathbf{a}$, and the derivation of its global optimal solution is difficult. In order to solve this problem, the proposed method introduces the calculation scheme used in [36] into the estimation of their optimal solutions $\hat{\mathbf{y}}$ and $\hat{\mathbf{a}}$, where this calculation scheme is the transformation of the original problem into a quasi-convex problem. First, we note that the first term in Eq. (12) depends only on μ_y and $\mu_{\mathbf{Za}}$. Therefore, we fix these values as $\mu_y = \rho$ and $\mu_{\mathbf{Za}} = \omega$. Then the problem of Eq. (11) is rewritten as

$$\max_{\mathbf{y}, \mathbf{a}} \left(\frac{2\mathbf{y}' \mathbf{H}_N \mathbf{Za} + C_2}{\mathbf{y}' \mathbf{H}_N \mathbf{y} + \mathbf{a}' \mathbf{Z}' \mathbf{H}_N \mathbf{Za} + C_2} \right)$$

$$\begin{aligned} &\text{subject to } \mathbf{M}\mathbf{y} = \mathbf{y}_{\bar{\Omega}}, \quad \mathbf{1}_K' \mathbf{a} = 1, \quad \frac{1}{N} \mathbf{1}_N' \mathbf{y} = \rho \text{ and } \frac{1}{N} \mathbf{1}_N' \mathbf{Za} = \omega. \end{aligned} \quad (14)$$

Therefore, the overall problem is to find the highest SSIM index by searching over ranges of ρ and ω . Furthermore, since the optimization problem in Eq. (14) is still non-convex, it is converted into a quasi-convex optimization problem as follows:

$$\begin{aligned} &\max_{\mathbf{y}, \mathbf{a}} \left(\frac{2\mathbf{y}' \mathbf{H}_N \mathbf{Za} + C_2}{\mathbf{y}' \mathbf{H}_N \mathbf{y} + \mathbf{a}' \mathbf{Z}' \mathbf{H}_N \mathbf{Za} + C_2} \right) \\ &\text{subject to } \mathbf{M}\mathbf{y} = \mathbf{y}_{\bar{\Omega}}, \quad \mathbf{1}_K' \mathbf{a} = 1, \quad \frac{1}{N} \mathbf{1}_N' \mathbf{y} = \rho \text{ and } \frac{1}{N} \mathbf{1}_N' \mathbf{Za} = \omega \\ &\Leftrightarrow \\ &\min : \tau \\ &\text{subject to} \\ &\left[\begin{array}{l} \max : \left(\frac{2\mathbf{y}' \mathbf{H}_N \mathbf{Za} + C_2}{\mathbf{y}' \mathbf{H}_N \mathbf{y} + \mathbf{a}' \mathbf{Z}' \mathbf{H}_N \mathbf{Za} + C_2} \right) \leq \tau \\ \text{subject to } \mathbf{M}\mathbf{y} = \mathbf{y}_{\bar{\Omega}}, \quad \mathbf{1}_K' \mathbf{a} = 1, \\ \frac{1}{N} \mathbf{1}_N' \mathbf{y} = \rho \text{ and } \frac{1}{N} \mathbf{1}_N' \mathbf{Za} = \omega \end{array} \right] \\ &\Leftrightarrow \\ &\min : \tau \\ &\text{subject to} \\ &\left[\begin{array}{l} \min : [\tau (\mathbf{y}' \mathbf{H}_N \mathbf{y} + \mathbf{a}' \mathbf{Z}' \mathbf{H}_N \mathbf{Za} + C_2) - (2\mathbf{y}' \mathbf{H}_N \mathbf{Za} + C_2)] \geq 0 \\ \text{subject to } \mathbf{M}\mathbf{y} = \mathbf{y}_{\bar{\Omega}}, \quad \mathbf{1}_K' \mathbf{a} = 1, \\ \frac{1}{N} \mathbf{1}_N' \mathbf{y} = \rho \text{ and } \frac{1}{N} \mathbf{1}_N' \mathbf{Za} = \omega \end{array} \right]. \end{aligned} \quad (15)$$

The first equivalence relationship holds since minimizing τ is the same as finding the least upper bound of Eq. (14). The second equivalence relationship holds since the denominator in Eq. (14) is strictly positive, allowing us to multiply through and rearrange terms. Then τ becomes a true upper bound if $\tau (\mathbf{y}' \mathbf{H}_N \mathbf{y} + \mathbf{a}' \mathbf{Z}' \mathbf{H}_N \mathbf{Za} + C_2) - (2\mathbf{y}' \mathbf{H}_N \mathbf{Za} + C_2)$ has a non-negative optimal value. Therefore, for the above quasi-convex problem, the proposed method estimates the solution by simply using the Lagrange multiplier approach. Note that the proposed method estimates the optimal value of τ based on a standard bisection procedure shown in Fig. 1. Furthermore, we have to estimate the optimal vectors $\hat{\mathbf{y}}$ and $\hat{\mathbf{a}}$ by searching over ranges of ρ and ω . In the proposed method, we change their values as $\mu_{y_{\bar{\Omega}}} - R\delta, \dots, \mu_{y_{\bar{\Omega}}} - 2\delta, \mu_{y_{\bar{\Omega}}} - \delta, \mu_{y_{\bar{\Omega}}}, \mu_{y_{\bar{\Omega}}} + \delta, \mu_{y_{\bar{\Omega}}} + 2\delta, \dots, \mu_{y_{\bar{\Omega}}} + R\delta$, where $\mu_{y_{\bar{\Omega}}}$ is obtained as $\frac{1}{N_{\bar{\Omega}}} \mathbf{1}_{N_{\bar{\Omega}}} \mu_{y_{\bar{\Omega}}}$. Then the optimal values of ρ and ω maximizing Eq. (12) are selected, and $\hat{\mathbf{y}}$ and $\hat{\mathbf{a}}$ obtained from the optimal values are adopted.

As shown in the previous subsection, since the proposed method can select the nearest neighbor patches accurately, the target patch is represented by these patches based on the SSIM index. Since it is shown in [21] that small patches of textures can be linearly represented by their nearest neighbors, the proposed method performs representation of the target patch based on the NE algorithm, i.e., the LLE-based reconstruction algorithm. In order to improve the perceptual quality, the proposed method introduces the SSIM index into the NE-based reconstruction algorithm. Then perceptually optimized reconstruction becomes feasible.

Finally, in the proposed method, we iteratively clip

- (i) An initial value of τ , which is defined as τ_0 , is provided between zero to one. Furthermore, $U_\tau = 1.0$ and $L_\tau = \tau_0$, where U_τ and L_τ represent the upper and lower limits of τ , respectively. We set $\tau_0 = 0.2$.
- (ii) The quasi-convex problem in Eq. (15) is solved by using τ .
- (iii) Two criteria V_τ and W_τ are respectively calculated as

$$V_\tau = \tau(\mathbf{y}'\mathbf{H}_N\mathbf{y} + \mathbf{a}'\mathbf{Z}'\mathbf{H}_N\mathbf{Z}\mathbf{a} + C_2) - (2\mathbf{y}'\mathbf{H}_N\mathbf{Z}\mathbf{a} + C_2),$$

$$W_\tau = U_\tau - L_\tau.$$
- (iv) From the obtained criteria V_τ and W_τ , the following steps are performed:
 - a) If $V_\tau \geq 0$ and $W_\tau < \epsilon$, the final optimal solution of τ is output, where $\epsilon = 0.05$.
 - b) If $V_\tau \geq 0$ but $W_\tau \geq \epsilon$, $\tau = \frac{U_\tau + L_\tau}{2}$ and $U_\tau = \tau$.
 - c) Otherwise, $\tau = \frac{U_\tau + L_\tau}{2}$ and $L_\tau = \tau$.
- (v) The procedures (ii)–(iv) are iterated.

Fig. 1 Algorithm for searching the optimal value of τ in the proposed method.

patches including missing areas from the target image and perform their reconstruction to estimate the whole missing intensities. In this procedure, we have to determine the order of reconstruction, i.e., the patch priority. Therefore, the proposed method adopts the most representative patch priority proposed by Criminisi et al. [9] and realizes reconstruction of the target image.

4. Experimental Results

This section shows some experimental results for verifying the effectiveness of the proposed method. In this experiment, we used six test images shown in Figs. 2–4 and intentionally added missing areas to the original images. We call these six images Images 1–6. We performed reconstruction of the obtained images including missing areas by using the proposed method. Note that in this experiment, we simply set the parameters of our method as follows: $w = h = 15$, $\delta = 5$, $R = 6$ and $K = 10$. Furthermore, C_1 and C_2 are determined according to the original work [35].

For comparison, we adopted five previously reported methods in [9], [12], [14], [18], [21]. Specifically, we used the representative exemplar-based method [9] and its improved version [18]. Reconstruction methods using multivariate analysis based on PCA [12], kernel PCA [14] and NE (LLE-based algorithm) [21] were also selected. The exemplar-based methods [9], [18] select optimal patches based on the MSE, and the multivariate analysis-based methods [12], [14], [21] also perform MSE-based missing area reconstruction. The NE-based method [21] performs selection of the nearest neighbor patches, and its scheme tries to solve the conventional problem, but its scheme is not directly related to the following reconstruction algorithm. Therefore, there is a difference from the selection algorithm in the proposed method. Furthermore, the method in [21] corresponds to the method which adopts the NE algorithm but uses the MSE as the quality measure. Therefore, this MSE+NE method is suitable for the comparison of our

SSIM+NE method.

For the above reasons, we selected these five previously reported methods for comparison with the proposed method. As stated above, we set the patch size to 15×15 pixels in this experiment, and the number of training patches became smaller. Since this comparison scheme was adopted in several papers, we also adopted such difficult conditions in order to make the difference in the performance of our method and the performance of previously reported methods clearer.

The results obtained from Image 1 are shown in Fig. 2. For better subjective evaluation, Fig. 3 shows their zoomed portions. From the obtained results, we can see that the proposed method successfully reconstructs missing textures. Other reconstruction results are also shown in Fig. 4. We only show the results of one previously reported method for each test image due to the limitation of space. Similar to Fig. 2, we can see high performance of the proposed reconstruction method.

In order to quantitatively evaluate the performance of the previously reported methods and our method, we calculated the SSIM index from the reconstruction results of Images 1–6. The results are shown in Table 1. Note that the SSIM values shown in this table were calculated from only the reconstructed areas. From this table, we can see that the proposed method also outperforms the previously reported methods in terms of the SSIM index. In this table, we also show results of our previously reported method [22]. The biggest difference between our method and this method is the introduction of the second novel scheme, i.e., the new selection scheme of the nearest neighbors. Thus, the comparison with the method in [22] is suitable for confirming the effectiveness of this novel selection scheme.

From the obtained results, it can be seen that reconstruction of missing textures by the proposed method is more accurate than that by the previously reported methods. It has been reported in [32], [33], [35] that the MSE and its variants cannot reflect perceptual qualities and that MSE-optimal algorithms do not necessarily produce images of high visual quality. Then the reconstruction results suffer from some degradations such as over-smoothness. It is known that general images contain more low-frequency components than high-frequency components. Therefore, since the MSE tends to monitor the differences in the low-frequency components, the MSE-optimal algorithms particularly represent the low-frequency components compared to their corresponding high-frequency components. Reconstruction results based on the MSE therefore tend to suffer from over-smoothness. On the other hand, since the SSIM index considers the contrast, it can compare the texture roughness, i.e., this similarity measure can compare to what degree high-frequency components are included. This point is also shown in [33], and Wang et al. confirmed that the SSIM index of blurred images that were perceptually degraded severely became lower. Thus, the proposed method can avoid the problem of over-smoothness of reconstruction results by using the SSIM-based reconstruction algorithm. By comparing our SSIM+NE method with the MSE+NE



Fig. 2 Reconstruction results obtained from Image 1 by the previously reported methods and our method: (a) Original image (480×360 pixels), (b) Flag image whose white regions correspond to missing areas, (c) Corrupted image including missing areas (8.9% loss), (d) Results by [9], (e) Results by [18], (f) Results by [12], (g) Results by [14], (h) Results by [21], (i) Results by the proposed method.

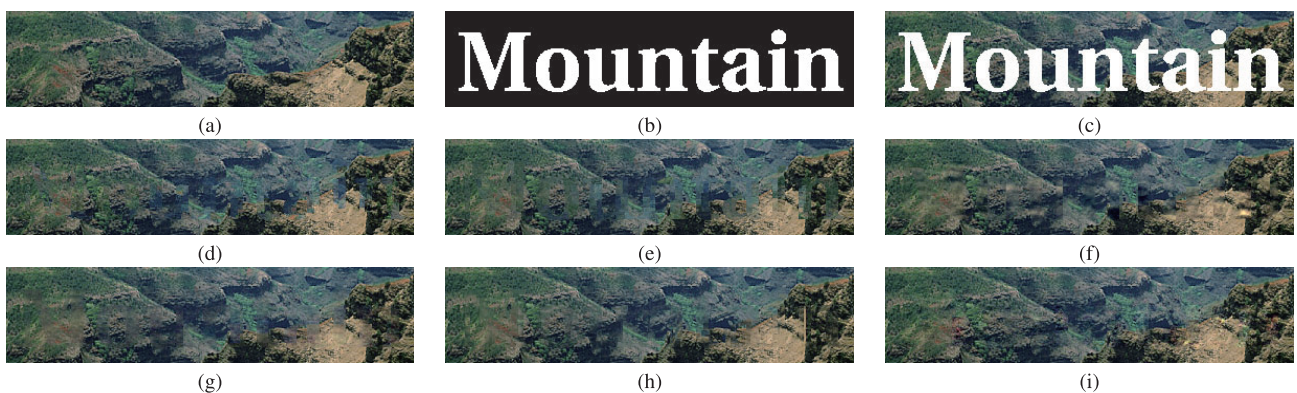


Fig. 3 Zoomed portions of Figs. 2(a)–(i).

method [21] in Table 1, the effectiveness of our method can be also confirmed.

In addition, from the characteristic that each patch can be linearly represented by its nearest neighbor patches, the

proposed method introduces the NE-based algorithm. In this approach, selection of the nearest neighbors from the target patch including missing areas becomes the most important problem. According to our previously reported method



Fig. 4 Reconstruction results obtained by the previously reported methods and our method. Five test images are used, and they respectively correspond to Images 2–6. Note that the previously reported methods used for reconstructing Images 2–6 are respectively Ref [9], Ref [18], Ref [12], Ref [14] and Ref [21]. The sizes of Images 2–6 are 640×480 pixels, 640×480 pixels, 480×360 pixels, 480×360 pixels and 640×480 pixels, respectively. The percentages of missing areas are 5.5%, 5.4%, 11.3%, 10.7% and 6.2% in Images 2–6, respectively.

Table 1 Performance comparison (SSIM index) between the previously reported methods and the proposed method.

Test image	Reference [9]	Reference [18]	Reference [12]	Reference [14]	Reference [21]	Reference [22]	Our method
Image 1	0.7456	0.7497	0.6081	0.6883	0.7587	0.7653	0.7959
Image 2	0.7031	0.6608	0.6695	0.7670	0.6935	0.7310	0.7789
Image 3	0.7383	0.7419	0.6668	0.7130	0.7324	0.7352	0.7526
Image 4	0.6648	0.6891	0.5678	0.6274	0.6308	0.6435	0.6593
Image 5	0.6722	0.6948	0.5887	0.6544	0.6929	0.7239	0.7510
Image 6	0.6998	0.7019	0.6200	0.7320	0.6674	0.7180	0.7405
Average	0.7040	0.7064	0.6202	0.6970	0.6959	0.7195	0.7464
Median	0.7015	0.6984	0.6141	0.7007	0.6932	0.7275	0.7518

[19], selection of nearest neighbors based on errors observed through the reconstruction procedures can outperform selection based only on differences in the known parts within the target patch. Therefore, the proposed method realizes selection of the nearest neighbors by using the SSIM index observed through the SSIM-based NE reconstruction algorithm. By comparing the performance of the proposed method with that of [22], the effectiveness of this selection scheme can be also confirmed. In this way, the proposed method can successfully perform reconstruction of missing textures by introducing the above non-conventional approaches.

5. Conclusions

A perceptual missing area reconstruction method using a neighboring embedding algorithm was presented in this paper. In the proposed method, the SSIM index is adopted for evaluating texture representation performance, and reconstruction of missing textures is performed on the basis of the neighboring embedding algorithm, i.e., the LLE-based algorithm. In this approach, the nearest neighbors can be selected by monitoring the SSIM index observed through the SSIM-based NE algorithm. Based on the above non-conventional approaches, the proposed method can reconstruct missing textures accurately. Experimental results show that the proposed method outputs better results than do previously reported methods.

Acknowledgement

This work was partly supported by Grant-in-Aid for Scientific Research (B) 25280036, Japan Society for the Promotion of Science (JSPS).

References

- [1] C. Guillemot and O. Le Meur, "Image inpainting: Overview and recent advances," *IEEE Signal Process. Mag.*, vol.31, no.1, pp.127–144, Jan. 2014.
- [2] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," *Proc. 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'00)*, pp.417–424, 2000.
- [3] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE Trans. Image Process.*, vol.10, no.8, pp.1200–1211, 2001.
- [4] A. Rares, M.J.T. Reinders, and J. Biemond, "Edge-based image restoration," *IEEE Trans. Image Process.*, vol.14, no.10, pp.1454–1468, 2005.
- [5] M. Bertalmio, "Strong-continuation, contrast-invariant inpainting with a third-order optimal PDE," *IEEE Trans. Image Process.*, vol.15, no.7, pp.1934–1938, 2006.
- [6] A.A. Efros and T.K. Leung, "Texture synthesis by non-parametric sampling," *Proc. Seventh IEEE International Conference on Computer Vision*, vol.2, pp.1033–1038, Corfu, Greece, 1999.
- [7] I. Drori, D. Cohen-Or, and H. Yeshurun, "Fragment-based image completion," *Papers on ACM SIGGRAPH 2003, SIGGRAPH'03*, pp.303–312, 2003.
- [8] A. Criminisi, P. Perez, and K. Toyama, "Object removal by exemplar-based inpainting," 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., pp.II-721–II-728, 2003.
- [9] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol.13, no.9, pp.1200–1212, 2004.
- [10] T.-H. Kwok and C.C.L. Wang, "Interactive image inpainting using DCT based exemplar matching," *Advances in Visual Computing, Lecture Notes in Computer Science*, vol.5876, pp.709–718, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [11] T.-H. Kwok, H. Sheung, and C.C.L. Wang, "Fast query for exemplar-based image completion," *IEEE Trans. Image Process.*, vol.19, no.12, pp.3106–3115, 2010.
- [12] T. Amano and Y. Sato, "Image interpolation using BPLP method on the eigenspace," *Syst. Comput. Jpn.*, vol.38, no.1, pp.87–96, 2007.
- [13] T. Ogawa and M. Haseyama, "POCS-based texture reconstruction method using clustering scheme by kernel PCA," *IEICE Trans. Fundamentals*, vol.E90-A, no.8, pp.1519–1527, Aug. 2007.
- [14] T. Ogawa and M. Haseyama, "Missing intensity interpolation using a kernel PCA-based POCS algorithm and its applications," *IEEE Trans. Image Process.*, vol.20, no.2, pp.417–432, 2011.
- [15] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol.17, no.1, pp.53–69, 2008.
- [16] B. Wohlberg, "Inpainting with sparse linear combinations of exemplars," 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, pp.689–692, 2009.
- [17] B. Shen, W. Hu, Y. Zhang, and Y.-J. Zhang, "Image inpainting via sparse representation," 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, pp.697–700, 2009.
- [18] Z. Xu and J. Sun, "Image inpainting by patch propagation using patch sparsity," *IEEE Trans. Image Process.*, vol.19, no.5, pp.1153–1165, 2010.
- [19] T. Ogawa and M. Haseyama, "Missing image data reconstruction based on adaptive inverse projection via sparse representation," *IEEE Trans. Multimedia*, vol.13, no.5, pp.974–992, 2011.
- [20] M. Turkan and C. Guillemot, "Locally linear embedding based texture synthesis for image prediction and error concealment," *Proc. 2012 19th IEEE International Conference on Image Processing*, pp.3009–3012, 2012.
- [21] C. Guillemot, M. Turkan, O.L. Meur, and M. Ebdelli, "Object removal and loss concealment using neighbor embedding methods," *Signal Processing: Image Communication*, vol.28, no.10, pp.1405–1419, 2013.
- [22] T. Ogawa and M. Haseyama, "Missing texture reconstruction method based on perceptually optimized algorithm," *EURASIP Journal on Advances in Signal Processing*, vol.2010, no.1, ID 208976, 13 pages, 2010.
- [23] T. Ogawa and M. Haseyama, "Perceptually optimized subspace estimation for missing texture reconstruction," 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.1141–1144, 2012.
- [24] A. Rehman, M. Rostami, Z. Wang, D. Brunet, and E.R. Vrscay, "SSIM-inspired image restoration using sparse representation," *EURASIP J. Advances in Signal Processing*, vol.2012:16, 2012.
- [25] T. Ogawa and M. Haseyama, "Image inpainting based on sparse representations with a perceptual metric," *EURASIP J. Advances in Signal Processing*, vol.2013:179, 2013.
- [26] T. Ogawa and M. Haseyama, "Missing intensity restoration via perceptually optimized subspace projection based on entropy component analysis," 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.175–179, 2014.
- [27] B. Schölkopf, S. Mika, C.J.C. Burges, P. Knirsch, K.-R. Müller, G. Rätsch, and A.J. Smola, "Input space versus feature space in kernel-based methods," *IEEE Trans. Neural Netw.*, vol.10, no.5, pp.1000–1017, 1999.
- [28] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol.54, no.11, pp.4311–4322, 2006.

- [29] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol.15, no.12, pp.3736–3745, 2006.
- [30] H.S. Seung and D.D. Lee, "Learning the parts of objects with nonnegative matrix factorization," *Nature*, vol.401, no.6755, pp.788–791, 1999.
- [31] S.T. Roweis and L.K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol.290, no.5500, pp.2323–2326, 2000.
- [32] B. Girod, "What's wrong with mean-squared error?," *Digital Images and Human Vision*, A.B. Watson, ed., pp.207–220, MIT Press, Cambridge, MA, 1993.
- [33] Z. Wang and A.C. Bovik, "Modern image quality assessment," *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol.2, no.1, pp.1–156, 2006.
- [34] H.R. Sheikh, M.F. Sabir, and A.C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol.15, no.11, pp.3440–3451, 2006.
- [35] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol.13, no.4, pp.600–612, 2004.
- [36] S.S. Channappayya, A.C. Bovik, C. Caramanis, and R.W. Heath, "Design of linear equalizers optimized for the structural similarity index," *IEEE Trans. Image Process.*, vol.17, no.6, pp.857–872, June 2008.
- [37] S.G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol.41, no.12, pp.3397–3415, 1993.
- [38] J.A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol.50, no.10, pp.2231–2242, 2004.



Miki Haseyama received her B.S., M.S. and Ph.D. degrees in Electronics from Hokkaido University, Japan in 1986, 1988 and 1993, respectively. She joined the Graduate School of Information Science and Technology, Hokkaido University as an associate professor in 1994. She was a visiting associate professor of Washington University, USA from 1995 to 1996. She is currently a professor in the Graduate School of Information Science and Technology, Hokkaido University. Her research interests include image and video processing and its development into semantic analysis. She has been a Vice-President of the Institute of Image Information and Television Engineers, Japan (ITE), an Editor-in-Chief of *ITE Transactions on Media Technology and Applications*, a Director, International Coordination and Publicity of The Institute of Electronics, Information and Communication Engineers (IEICE). She is a member of the IEEE, IEICE, Institute of Image Information and Television Engineers (ITE) and Acoustical Society of Japan (ASJ).



Takahiro Ogawa received his B.S., M.S. and Ph.D. degrees in Electronics and Information Engineering from Hokkaido University, Japan in 2003, 2005 and 2007, respectively. He is currently an assistant professor in the Graduate School of Information Science and Technology, Hokkaido University. His research interests are multimedia signal processing and its applications. He has been an Associate Editor of *ITE Transactions on Media Technology and Applications*. He is a member of the IEEE, EURASIP,

IEICE, and Institute of Image Information and Television Engineers (ITE).