IEICE TRANS. FUNDAMENTALS, VOL.E92–A, NO.8 AUGUST 2009

PAPER Special Section on Signal Processing

Adaptive Missing Texture Reconstruction Method Based on Kernel Canonical Correlation Analysis with a New Clustering Scheme

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

SUMMARY In this paper, a method for adaptive reconstruction of missing textures based on kernel canonical correlation analysis (CCA) with a new clustering scheme is presented. The proposed method estimates the correlation between two areas, which respectively correspond to a missing area and its neighboring area, from known parts within the target image and realizes reconstruction of the missing texture. In order to obtain this correlation, the kernel CCA is applied to each cluster containing the same kind of textures, and the optimal result is selected for the target missing area. Specifically, a new approach monitoring errors caused in the above kernel CCA-based reconstruction process enables selection of the optimal result. This approach provides a solution to the problem in traditional methods of not being able to perform adaptive reconstruction of the target textures due to missing intensities. Consequently, all of the missing textures are successfully estimated by the optimal cluster's correlation, which provides accurate reconstruction of the same kinds of textures. In addition, the proposed method can obtain the correlation more accurately than our previous works, and more successful reconstruction performance can be expected. Experimental results show impressive improvement of the proposed reconstruction technique over previously reported reconstruction techniques.

key words: *image restoration, texture, kernel canonical correlation analysis, nonlinear estimation*

1. Introduction

In the field of image restoration, reconstruction of missing areas in digital images is a very important issue since it has a number of fundamental applications. For example, it is applied to removal of unnecessary objects, restoration of corrupted old films, and error concealment for video communications. In order to realize these applications, many methods for reconstruction of important visual features, such as structural and texture features have been proposed. Structural reconstruction approaches focus on accurate restoration of missing edges in images and are effective for pure structure images [1]–[3]. However, since ordinary images also contain many textures, texture reconstruction approaches [4]–[6] work better for those parts. In this paper, we focus on the reconstruction of texture features.

Much progress has been made in recent studies on missing texture reconstruction [4]–[6]. Most algorithms reported in the literature reconstruct missing areas by utilizing statistical features of known textures within the target image as training patterns. Kernel principal component analysis (PCA) [7], [8] is a suitable method for extraction of nonlinear features in observed data. By utilizing projection schemes onto the nonlinear subspace obtained in the kernel PCA [9], the missing textures can be reconstructed from known textures. However, such approaches assume that arbitrary local textures within the target image are similar to each other; that is, the target image contains only one type of texture. Thus, if the target image consists of various textures, the missing textures should be adaptively reconstructed from only the same kinds of textures. Unfortunately, such textures cannot be selected by traditional methods since the distance between the target missing textures and the other ones cannot be calculated.

In order to solve the problem in traditional methods of not being able to perform adaptive texture reconstruction, we first proposed an error reduction (ER) algorithm [10] based adaptive texture reconstruction method in [11]. In this method, missing textures are reconstructed by retrieving their phases from the Fourier transform magnitudes estimated from known textures within the target image. Furthermore, the same kinds of textures could be adaptively provided for the estimation of the Fourier transform magnitudes by monitoring errors converged in the ER algorithm. However, in this ER algorithm, the constraint utilized in the Fourier domain was not convex, and the convergence of the optimal solution was not guaranteed. Therefore, in order to modify this method, we proposed a projection onto convex sets (POCS) [12] based adaptive reconstruction methods of missing textures by using the kernel PCA in [13] and [14]. Furthermore, in [15], we also proposed an improved method, which introduced a renewal approach of the constraints into the POCS algorithm. These methods in [13]–[15] tried to estimate correlation of intensities between known and unknown areas in the target image by using the kernel PCA to reconstruct missing textures. However, the correlation utilized in these methods contains not only the desired one but also the correlation of intensities between pixels within known areas and that between pixels within unknown areas. In the point view concerning the correlation between known and unknown areas directly, the kernel canonical correlation analysis

In this paper, an adaptive texture reconstruction method based on kernel CCA with a new clustering scheme is proposed. Since the kernel CCA is a useful method for finding underlying relationships between two different data sets, we utilize this method for finding the correlation of intensities between missing areas and the other known areas. Specifically, the proposed method estimates the correlation be-

Manuscript received December 7, 2008.
Manuscript revised March 4, 2009.

[†]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.E92.A.1950

tween two areas, which respectively correspond to a missing area and its neighboring area, from known parts within the target image and reconstructs the missing intensities. In this procedure, the kernel CCA is applied to each cluster containing the same kind of textures, and the optimal result is selected for the target missing area based on errors caused in the above reconstruction scheme. This approach provides a solution to the problem of the traditional methods and is the biggest advantage of the proposed method. Since each missing texture is adaptively reconstructed by the optimal correlation obtained from only the same kind of textures, successful restoration of the missing areas can be expected. Furthermore, since the proposed method estimates more accurate correlation than those obtained by our previous works [13]–[15], improvement of the reconstruction performance can be also expected.

This paper is organized as follows. The kernel CCA is explained in Sect. 2, and the kernel CCA-based missing texture reconstruction method is presented in Sect. 3. Experimental results that verify the performance of the proposed method are shown in Sect. 4. Finally, conclusions are given in Sect. 5.

2. Kernel CCA

The kernel CCA is explained in this section. Suppose there is a pair of multi-variates $\mathbf{x} \in \mathrm{R}^{n_x}$ and $\mathbf{y} \in \mathrm{R}^{n_y}$, and they are respectively transformed into the feature space via the nonlinear map $\phi_{\mathbf{x}}$ and $\phi_{\mathbf{y}}$. Furthermore, from the mapped results $\phi_{\mathbf{x}}(\mathbf{x})$ and $\phi_{\mathbf{y}}(\mathbf{y})^{\dagger}$, the kernel CCA seeks to maximize the correlation

$$\rho = \frac{\mathrm{E}\{st\}}{\sqrt{\mathrm{E}\{s^2\}\mathrm{E}\{t^2\}}} \tag{1}$$

between

$$s = \mathbf{a}'\phi_{\mathbf{x}}(\mathbf{x}) \tag{2}$$

and

$$t = \mathbf{b}'\phi_{\mathbf{y}}(\mathbf{y}) \tag{3}$$

over the projection directions $\mathbf{a}$ and $\mathbf{b}$. This means that the kernel CCA finds the directions $\mathbf{a}$ and $\mathbf{b}$ that maximize the correlation $\mathrm{E}\{st\}$ of corresponding projections subject to $\mathrm{E}\{s^2\} = 1$ and $\mathrm{E}\{t^2\} = 1$. Note that vector/matrix transpose is denoted by the superscript $'$ in this paper.

The optimal directions $\mathbf{a}$ and $\mathbf{b}$ can be found by solving the Lagrangean

$$\mathcal{L} = \mathrm{E}\{st\} - \frac{\lambda_1}{2}(\mathrm{E}\{s^2\} - 1) - \frac{\lambda_2}{2}(\mathrm{E}\{t^2\} - 1) + \frac{\eta}{2}(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2), \tag{4}$$

where η is a regularization parameter, and this computation scheme is called the regularized kernel CCA [17]. By taking the derivatives of Eq. (4) with respect to $\mathbf{a}$ and $\mathbf{b}$, $\lambda_1 = \lambda_2$ $(= \lambda)$ is derived, and the directions $\mathbf{a}$ and $\mathbf{b}$ maximizing the correlation ρ $(= \lambda)$ can be calculated.

3. Kernel CCA-Based Missing Texture Reconstruction

A kernel CCA-based missing texture reconstruction method is presented in this section. In Fig. 1, the outline of the proposed method is shown. In the the proposed method, a local image f ($w \times h$ pixels) including missing areas is clipped from the target image, and the missing intensities are estimated based on the kernel CCA. For the following explanation, we respectively denote two areas whose intensities are unknown and known within the target local image f as Ω and $\bar{\Omega}$, respectively.

Note that in the target image, there are many known local images whose textures are quite different from that of the target local image f. Such local images should not affect the reconstruction of the target local image f. Therefore, the proposed method applies the kernel CCA to each cluster of local images containing the same kind of texture, and the optimal result is adaptively utilized for reconstruction of the target local image f. In order to realize this scheme, clustering of the known local images within the target image must be performed before the reconstruction process. Thus, we firstly show the kernel CCA-based clustering algorithm of the known local images in 3.1. The adaptive reconstruction algorithm of the missing textures based on the kernel CCA is shown in 3.2.

3.1 Texture Clustering Algorithm

In this subsection, clustering of known textures within the target image as preprocessing for reconstruction of the missing textures is described. First, we clip known local images f_i $(i = 1, 2, \cdots, N)$ whose size is $w \times h$ pixels from the

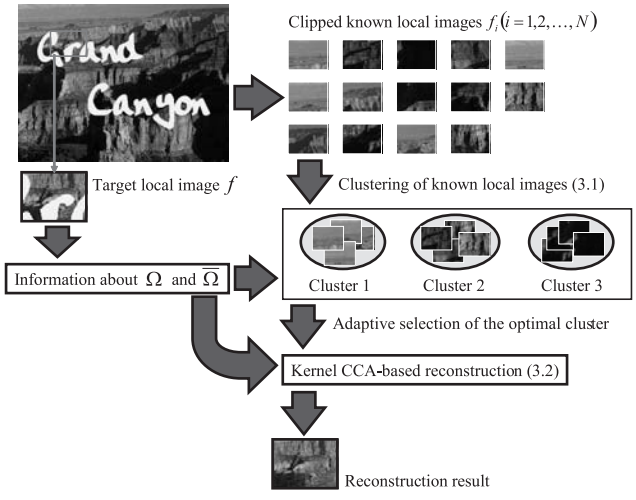

Fig. 1 Outline of the proposed texture clustering and reconstruction algorithms.

†In this section, we assume that $\mathrm{E}\{\phi_{\mathbf{x}}(\mathbf{x})\} = \mathbf{0}$ and $\mathrm{E}\{\phi_{\mathbf{y}}(\mathbf{y})\} = \mathbf{0}$ for brief explanation, where $E\{\cdot\}$ denotes the sample average of the random variates.

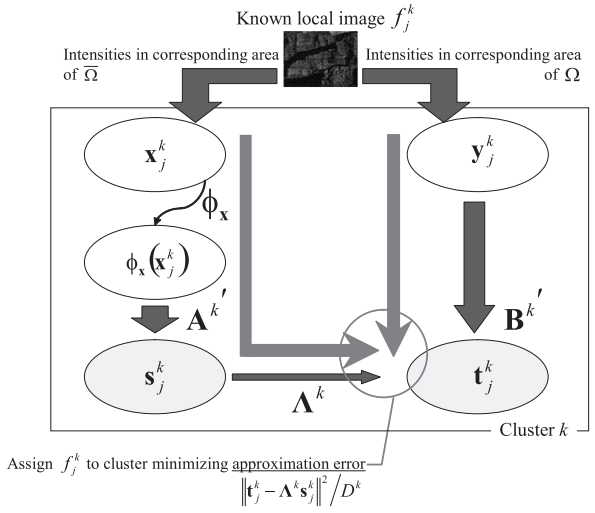

Fig. 2 Overview of the clustering algorithm of known local images.

target image in the same interval (width $\tilde{w}_1$ pixels, height $\tilde{h}_1$ pixels). Next, for each local image f_i, two vectors $\mathbf{x}_i$ ($\in \mathbf{R}^{wh-N_\Omega}$) and $\mathbf{y}_i$ ($\in \mathbf{R}^{N_\Omega}$), whose elements are respectively raster scanned intensities in the corresponding areas of $\bar{\Omega}$ and Ω, are defined, where N_Ω represents the number of pixels in Ω. Furthermore, the proposed method maps $\mathbf{x}_i$ into the feature space via the nonlinear map $\phi_{\mathbf{x}}$:$\mathbf{R}^{wh-N_\Omega} \rightarrow \mathbf{F}$ [16]. In this paper, we use the nonlinear map whose kernel function is a Gaussian kernel. The Gaussian kernel is a default general purpose kernel in the kernel methods community [9]. Therefore, in the proposed method, we utilize the Gaussian kernel. Note that an exact pre-image, which is the inverse mapping from the feature space back to the input space, typically does not exist [19]. Therefore, the estimated pre-image includes some errors. Since the final results estimated in the proposed method are the missing intensities in Ω, we do not utilize the nonlinear map for $\mathbf{y}_i$.

From the obtained results $\phi_{\mathbf{x}}(\mathbf{x}_i)$ and $\mathbf{y}_i$ ($i = 1, 2, \cdots, N$), the proposed method performs their clustering that minimizes the following criterion, and its overview is shown in Fig. 2:

$$C = \sum_{k=1}^{K}\sum_{j=1}^{N^k}\left\|\mathbf{B}^{k\prime}\left(\mathbf{y}_j^k - \bar{\mathbf{y}}^k\right) - \mathbf{\Lambda}^k\mathbf{A}^{k\prime}\left(\phi_{\mathbf{x}}(\mathbf{x}_j^k) - \bar{\phi}_{\mathbf{x}}^k\right)\right\|^2 / D^k, \quad (5)$$

where K is the number of the clusters. The vectors $\mathbf{x}_j^k$ and $\mathbf{y}_j^k$ ($j = 1, 2, \cdots, N^k$) respectively represent $\mathbf{x}_i$ and $\mathbf{y}_i$ ($i = 1, 2, \cdots, N$) assigned to cluster k. Furthermore, $\bar{\phi}_{\mathbf{x}}^k$ and $\bar{\mathbf{y}}^k$ are respectively the mean vectors of $\phi_{\mathbf{x}}(\mathbf{x}_j^k)$ and $\mathbf{y}_j^k$ ($j = 1, 2, \cdots, N^k$) and are defined below.

$$\bar{\phi}_{\mathbf{x}}^k = \frac{1}{N^k}\mathbf{\Xi}_{\mathbf{x}}^k\mathbf{1}^k, \quad (6)$$

$$\bar{\mathbf{y}}^k = \frac{1}{N^k}\mathbf{Y}^k\mathbf{1}^k, \quad (7)$$

where

$$\mathbf{\Xi}_{\mathbf{x}}^k = [\phi_{\mathbf{x}}(\mathbf{x}_1^k), \phi_{\mathbf{x}}(\mathbf{x}_2^k), \cdots, \phi_{\mathbf{x}}(\mathbf{x}_{N^k}^k)], \quad (8)$$

$$\mathbf{Y}^k = [\mathbf{y}_1^k, \mathbf{y}_2^k, \cdots, \mathbf{y}_{N^k}^k], \quad (9)$$

and $\mathbf{1}^k = [1, 1, \cdots, 1]'$ is an $N^k \times 1$ vector. The matrices $\mathbf{A}^k$ and $\mathbf{B}^k$ are coefficient matrices whose columns respectively correspond to the projection directions $\mathbf{a}$ and $\mathbf{b}$ in Eqs. (2) and (3), and $\mathbf{\Lambda}^k$ is a correlation matrix whose diagonal elements correspond to the correlation in Eq. (1). In addition, the value D^k is the dimension of $\mathbf{A}^k$, $\mathbf{B}^k$, and $\mathbf{\Lambda}^k$. Thus, the D^k-dimensional matrices $\mathbf{A}^k$, $\mathbf{B}^k$, and $\mathbf{\Lambda}^k$ are obtained by applying the kernel CCA to $\phi_{\mathbf{x}}(\mathbf{x}_j^k)$ and $\mathbf{y}_j^k$ ($j = 1, 2, \cdots, N^k$). Specifically, given $\{(\phi_{\mathbf{x}}(\mathbf{x}_j^k), \mathbf{y}_j^k) | j = 1, 2, \cdots, N^k\}$,

$$\mathbf{A}^k = [\mathbf{a}_1^k, \mathbf{a}_2^k, \cdots, \mathbf{a}_{D^k}^k] \quad (10)$$

and

$$\mathbf{B}^k = [\mathbf{b}_1^k, \mathbf{b}_2^k, \cdots, \mathbf{b}_{D^k}^k] \quad (11)$$

in Eq. (5) are matrices that maximize the correlation λ_d^k ($d = 1, 2, \cdots, D^k$) between the d-th elements in the following two vectors:

$$\mathbf{s}_j^k = \mathbf{A}^{k\prime}\left(\phi_{\mathbf{x}}(\mathbf{x}_j^k) - \bar{\phi}_{\mathbf{x}}^k\right), \quad (12)$$

$$\mathbf{t}_j^k = \mathbf{B}^{k\prime}\left(\mathbf{y}_j^k - \bar{\mathbf{y}}^k\right). \quad (13)$$

Then, $\mathbf{\Lambda}^k$ becomes a diagonal matrix whose diagonal elements are the correlation coefficients λ_d^k ($d = 1, 2, \cdots, D^k$). The details of the calculation of $\mathbf{A}^k$, $\mathbf{B}^k$, $\mathbf{\Lambda}^k$, and D^k are shown as follows.

In order to obtain $\mathbf{A}^k$, $\mathbf{B}^k$, and $\mathbf{\Lambda}^k$, we utilize the regularized kernel CCA shown in the previous section. Note that the optimal matrix $\mathbf{A}^k$ is given by

$$\mathbf{A}^k = \mathbf{\Xi}_{\mathbf{x}}^k\mathbf{H}^k\mathbf{E}^k, \quad (14)$$

where $\mathbf{E}^k = [\mathbf{e}_1^k, \mathbf{e}_2^k, \cdots, \mathbf{e}_{D^k}^k]$ is an $N^k \times D^k$ matrix. Furthermore,

$$\mathbf{H}^k = \mathbf{I}^k - \frac{1}{N^k}\mathbf{1}^k\mathbf{1}^{k\prime} \quad (15)$$

is a centering matrix, where $\mathbf{I}^k$ is the $N^k \times N^k$ identity matrix. From Eq. (14), the following equation is satisfied

$$\mathbf{a}_d^k = \mathbf{\Xi}_{\mathbf{x}}^k\mathbf{H}^k\mathbf{e}_d^k \qquad (d = 1, 2, \cdots, D^k). \quad (16)$$

Then, by calculating the optimal solution $\mathbf{e}_d^k$ and $\mathbf{b}_d^k$ ($d = 1, 2, \cdots, D^k$), $\mathbf{A}^k$ and $\mathbf{B}^k$ can be obtained. In the same way as Eq. (4), we calculate the optimal solution $\mathbf{e}^k$ and $\mathbf{b}^k$ that maximizes

$$\mathcal{L} = \mathbf{e}^{k\prime}\mathbf{L}^k\mathbf{b}^k - \frac{\lambda^k}{2}\left(\mathbf{e}^{k\prime}\mathbf{M}^k\mathbf{e}^k - 1\right) - \frac{\lambda^k}{2}\left(\mathbf{b}^{k\prime}\mathbf{P}^k\mathbf{b}^k - 1\right), \quad (17)$$

where $\mathbf{e}^k$, $\mathbf{b}^k$, and λ^k respectively correspond to $\mathbf{e}_d^k$, $\mathbf{b}_d^k$, and λ_d^k. In the above equation, $\mathbf{L}^k$, $\mathbf{M}^k$, and $\mathbf{P}^k$ are calculated as follows:

$$\mathbf{L}^k = \frac{1}{N^k}\mathbf{H}^k\mathbf{K}_{\mathbf{x}}^k\mathbf{H}^k\mathbf{H}^k\mathbf{Y}^{k\prime}, \quad (18)$$

$$\mathbf{M}^k = \frac{1}{N^k}\mathbf{H}^k\mathbf{K}_{\mathbf{x}}^k\mathbf{H}^k\mathbf{H}^k\mathbf{K}_{\mathbf{x}}^k\mathbf{H}^k - \eta_1\mathbf{H}^k\mathbf{K}_{\mathbf{x}}^k\mathbf{H}^k, \quad (19)$$

$$\mathbf{P}^k = \frac{1}{N^k}\mathbf{Y}^k\mathbf{H}^k\mathbf{H}^k\mathbf{Y}^{k\prime} - \eta_2\mathbf{I}^k. \tag{20}$$

Furthermore, $\mathbf{K}^k_{\mathbf{x}}$ $(= \mathbf{\Xi}^{k\prime}_{\mathbf{x}}\mathbf{\Xi}^k_{\mathbf{x}})$ is the matrix whose (p,q)-th $(p = 1,2,\cdots,N^k,\ q = 1,2,\cdots,N^k)$ element is defined as $\kappa_{\mathbf{x}}(\mathbf{x}^k_p, \mathbf{x}^k_q)$ by using the Gaussian kernel function $\kappa_{\mathbf{x}}(\cdot,\cdot)$ [18]. The values η_1 and η_2 in Eqs. (19) and (20) are regularization parameters. By taking derivatives of Eq. (17) with respect to $\mathbf{e}^k$ and $\mathbf{b}^k$, the optimal $\mathbf{e}^k$, $\mathbf{b}^k$, and λ^k can be obtained as the solutions of the following eigenvalue problems:

$$\mathbf{M}^{k-1}\mathbf{L}^k\mathbf{P}^{k-1}\mathbf{L}^{k\prime}\mathbf{e}^k = \lambda^{k2}\mathbf{e}^k, \tag{21}$$

$$\mathbf{P}^{k-1}\mathbf{L}^{k\prime}\mathbf{M}^{k-1}\mathbf{L}^k\mathbf{b}^k = \lambda^{k2}\mathbf{b}^k, \tag{22}$$

where λ^k is obtained as an eigenvalue, and the vectors $\mathbf{e}^k$ and $\mathbf{b}^k$ are respectively obtained as eigenvectors. Then, the d-th $(d = 1,2,\cdots,D^k)$ eigenvalue of λ^k becomes λ^k_d, where $\lambda^k_1 \geq \lambda^k_2,\cdots,\lambda^k_{D^k}$. Note that the dimension D^k is set to the value whose cumulative proportion obtained from λ^k_d $(d = 1,2,\cdots,D^k)$ becomes larger than Th. Furthermore, the eigenvectors $\mathbf{e}^k$ and $\mathbf{b}^k$ corresponding to λ^k_d become $\mathbf{e}^k_d$ and $\mathbf{b}^k_d$, respectively.

From Eqs. (6) and (7) and the obtained matrices $\mathbf{B}^k$, $\mathbf{E}^k$, and $\mathbf{\Lambda}^k$, Eq. (5) can be rewritten as follows:

$$\begin{aligned} C &= \sum_{k=1}^{K}\sum_{j=1}^{N^k}\left\|\mathbf{B}^{k\prime}\left(\mathbf{y}^k_j - \frac{1}{N^k}\mathbf{Y}^k\mathbf{1}^k\right)\right. \\ &\quad \left. -\mathbf{\Lambda}^k\mathbf{E}^{k\prime}\mathbf{H}^k\mathbf{\Xi}^{k\prime}_{\mathbf{x}}\left(\phi_{\mathbf{x}}(\mathbf{x}^k_j) - \frac{1}{N^k}\mathbf{\Xi}^k_{\mathbf{x}}\mathbf{1}^k\right)\right\|^2/D^k, \\ &= \sum_{k=1}^{K}\sum_{j=1}^{N^k}\left\|\mathbf{B}^{k\prime}\left(\mathbf{y}^k_j - \frac{1}{N^k}\mathbf{Y}^k\mathbf{1}^k\right)\right. \\ &\quad \left. -\mathbf{\Lambda}^k\mathbf{E}^{k\prime}\mathbf{H}^k\left(\kappa^k_j - \frac{1}{N^k}\mathbf{K}^k_{\mathbf{x}}\mathbf{1}^k\right)\right\|^2/D^k, \end{aligned} \tag{23}$$

where κ^k_j is an $N^k\times 1$ vector whose p-th element is $\kappa_{\mathbf{x}}(\mathbf{x}^k_j, \mathbf{x}^k_p)$.

From Eq. (23), the mapped result

$$\hat{\mathbf{t}}^k_j = \mathbf{\Lambda}^k\mathbf{E}^{k\prime}\mathbf{H}^k\left(\kappa^k_j - \frac{1}{N^k}\mathbf{K}^k_{\mathbf{x}}\mathbf{1}^k\right) = \mathbf{\Lambda}^k\mathbf{s}^k_j \tag{24}$$

is the optimal approximation result of $\mathbf{t}^k_j$ in Eq. (13) as shown in Fig. 2. Therefore, $\|\mathbf{t}^k_j - \hat{\mathbf{t}}^k_j\|^2$ corresponds to the minimum distance between the new variate $\mathbf{t}^k_j$ of $\mathbf{y}^k_j$ and $\hat{\mathbf{t}}^k_j$ obtained from the new variate $\mathbf{s}^k_j$ of $\phi_{\mathbf{x}}(\mathbf{x}^k_j)$. Then, by using criterion C, we can effectively perform the clustering of the local images f_i $(i = 1,2,\cdots,N)$.

In the conventional methods [13]–[15], we utilized different criterion with that of the proposed method to perform clustering of known local textures. Their criterion represents the sum of approximation errors between known local textures and those projected onto the nonlinear eigenspaces obtained from their belonging clusters. These errors contained both of two errors respectively calculated from two areas corresponding to known and unknown areas of target local images. Note that the errors we must minimize are those in the unknown areas Ω of the target local image f. Therefore, we utilize the new criterion representing the sum of the minimum errors between new variates of areas corresponding to Ω and those optimally estimated from areas corresponding to $\bar{\Omega}$ by the kernel CCA.

3.2 Texture Reconstruction Algorithm

In this section, we present an algorithm for reconstruction of the missing texture in the target local image f from the clustering results presented in the previous section. First, we denote the vectors of the raster scanned intensities in $\bar{\Omega}$ and Ω as $\mathbf{x}$ and $\mathbf{y}$, respectively. As shown in Fig. 3, the estimation result $\hat{\mathbf{y}}^k$ of the unknown vector $\mathbf{y}$ by cluster k is obtained as follows:

$$\hat{\mathbf{y}}^k = \mathbf{T}^k_{\mathbf{B}}\mathbf{\Lambda}^k\mathbf{A}^{k\prime}\left(\phi_{\mathbf{x}}(\mathbf{x}) - \bar{\phi}^k_{\mathbf{x}}\right) + \bar{\mathbf{y}}^k. \tag{25}$$

In the above equation, the matrix $\mathbf{T}^k_{\mathbf{B}}$ satisfies

$$\mathbf{T}^k_{\mathbf{B}}\mathbf{B}^{k\prime}\mathbf{Y}^k\mathbf{H}^k = \mathbf{Y}^k\mathbf{H}^k, \tag{26}$$

and it is obtained by calculating the pseudo-inverse matrix of $\mathbf{B}^{k\prime}\mathbf{Y}^k\mathbf{H}^k$ as follows:

$$\mathbf{T}^k_{\mathbf{B}} = \mathbf{Y}^k\mathbf{H}^k\mathbf{H}^k\mathbf{Y}^{k\prime}\mathbf{B}^k\left(\mathbf{B}^{k\prime}\mathbf{Y}^k\mathbf{H}^k\mathbf{H}^k\mathbf{Y}^{k\prime}\mathbf{B}^k\right)^{-1}. \tag{27}$$

Furthermore, by utilizing the calculation scheme of Eq. (23), Eq. (25) is rewritten as follows:

$$\hat{\mathbf{y}}^k = \mathbf{T}^k_{\mathbf{B}}\mathbf{\Lambda}^k\mathbf{E}^{k\prime}\mathbf{H}^k\left(\kappa^k - \frac{1}{N^k}\mathbf{K}^k_{\mathbf{x}}\mathbf{1}^k\right) + \frac{1}{N^k}\mathbf{Y}^k\mathbf{1}^k, \tag{28}$$

where κ^k is an $N^k\times 1$ vector whose p-th element is $\kappa_{\mathbf{x}}(\mathbf{x}, \mathbf{x}^k_p)$.

By calculating $\hat{\mathbf{y}}^k$ in Eq. (28), the missing intensities in Ω can be estimated from cluster k. In the proposed method, the matrices $\mathbf{A}^k$ and $\mathbf{B}^k$, which maximize the correlation between the new variates in Eqs. (12) and (13), are calculated from $\phi_{\mathbf{x}}(\mathbf{x}^k_j)$ and $\mathbf{y}^k_j$ $(j = 1,2,\cdots,N^k)$ by the kernel CCA. Then, from the obtained matrices $\mathbf{A}^k$, $\mathbf{B}^k$, and $\mathbf{\Lambda}^k$, $\mathbf{t}^k_j$ of cluster k in Eq. (13) can be optimally approximated by $\hat{\mathbf{t}}^k_j$ in Eq. (24) as shown in Fig. 2. Therefore, if we can classify the target local image f into the optimal cluster k^{opt}, the proposed method accurately estimates the unknown vector $\mathbf{y}$ from the known vector $\phi_{\mathbf{x}}(\mathbf{x})$ in Eq. (28). However, since the target local image f contains the missing area Ω, it cannot be classified by criterion C in Eq. (23). Thus, in order to achieve the classification of f, the proposed method utilizes the following novel criterion as a substitute for Eq. (23):

$$\tilde{C}^k = \left\|\mathbf{A}^{k\prime}\left(\phi_{\mathbf{x}}(\mathbf{x}) - \bar{\phi}^k_{\mathbf{x}}\right) - \mathbf{\Lambda}^k\mathbf{B}^{k\prime}\left(\hat{\mathbf{y}}^k - \bar{\mathbf{y}}^k\right)\right\|^2\Big/D^k. \tag{29}$$

By utilizing the calculation scheme of Eqs. (23) and (28), the above equation is rewritten as follows:

$$\begin{aligned} \tilde{C}^k &= \left\|\mathbf{E}^{k\prime}\mathbf{H}^k\left(\kappa^k - \frac{1}{N^k}\mathbf{K}^k_{\mathbf{x}}\mathbf{1}^k\right)\right. \\ &\quad \left. -\mathbf{\Lambda}^k\mathbf{B}^{k\prime}\left(\hat{\mathbf{y}}^k - \frac{1}{N^k}\mathbf{Y}^k\mathbf{1}^k\right)\right\|^2\Big/D^k. \end{aligned} \tag{30}$$

As shown in the previous section, the mapped result

$$\hat{\mathbf{s}}^k = \mathbf{\Lambda}^k \mathbf{B}^{k\prime}\left(\hat{\mathbf{y}}^k - \frac{1}{N^k}\mathbf{Y}^k\mathbf{1}^k\right) = \mathbf{\Lambda}^k\hat{\mathbf{t}}^k \tag{31}$$

becomes the optimal approximation result of the known vector

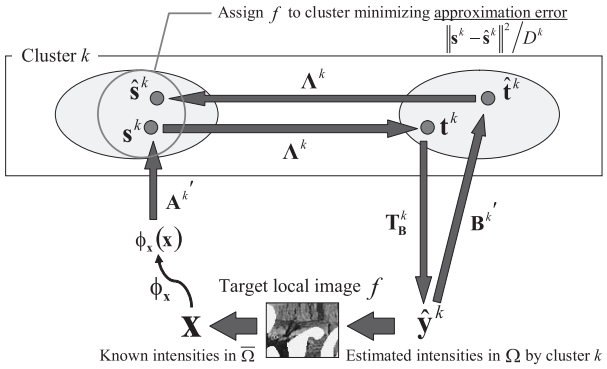

Fig. 3 Overview of the reconstruction algorithm of missing textures.

$$\mathbf{s}^k = \mathbf{E}^{k\prime}\mathbf{H}^k\left(\kappa^k - \frac{1}{N^k}\mathbf{K}_\mathbf{x}^k\mathbf{1}^k\right), \tag{32}$$

when the target local image f belongs to cluster k. This means that criterion $\tilde{C}^k$ corresponds to the minimum distance between the new variate $\mathbf{s}^k$ of the known vector $\phi_\mathbf{x}(\mathbf{x})$ and $\hat{\mathbf{s}}^k$ obtained from the new variate $\hat{\mathbf{t}}^k$ of the estimation result $\hat{\mathbf{y}}^k$ as shown in Fig. 3. Therefore, this criterion is applicable for the classification of the target local image f. Selection of the optimal cluster k^{opt} minimizing Eq. (30) for the target local image f then becomes possible. Furthermore, the proposed method regards the result $\hat{\mathbf{y}}^{k^{\text{opt}}}$ obtained by the optimal cluster k^{opt} as the output. Consequently, by performing the non-conventional approach, which adaptively selects the optimal cluster for the target missing area, we can reconstruct all of the missing textures in the target image accurately.

Finally, we compare the proposed method with our previous works. Note that in our previous works [13]–[15], we tried to estimate correlation of intensities between known and unknown areas in the target image by using the kernel PCA to reconstruct missing textures. However, the correlation utilized in these methods contains not only the de-

Fig. 4 (a) Original test image 1, (b) Corrupted image including text regions "Fall Harvest Sweet Chestnut" (11.3% loss), (c) Zoomed portion of (a), (d) Image reconstructed by reference [9], (e) Image reconstructed by reference [15], (f) Image reconstructed by the proposed method, (g) Zoomed portion of (d), (h) Zoomed portion of (e), (i) Zoomed portion of (f).

sired one but also the correlation of intensities between pixels within known areas and that between pixels within unknown areas. On the other hand, the proposed method can obtain the correlation of intensities between known and unknown areas more successfully than those methods. Therefore, more accurate performance can be obtained by the proposed method.

As described above, we can reconstruct the missing texture in the target local image. The proposed method clips local images ($w \times h$ pixels) including missing pixels in the same interval (width $\tilde{w}_2$ pixels, height $\tilde{h}_2$ pixels) and reconstructs them by using the above approach. Note that each restored pixel has multiple estimation results if the clipping interval is smaller than the size of the local images. In this case, the proposed method regards the result minimizing Eq. (30) as the final one.

4. Experimental Results

The performance of the proposed method is shown in this section. Figure 4(b) is a test texture image (480×360 pixels, 24-bit color levels) that includes the text regions "Fall Harvest Sweet Chestnut" from Fig. 4(a). We utilized the eigenspace method using projection of the nonlinear subspace obtained by the kernel PCA in [9] as a traditional method, and its reconstruction results are shown in Fig. 4(d). In addition, we also utilized another conventional method [15], which was the best published method in [11] and [13]–[15], and its reconstruction results are shown in Fig. 4(e). Figure 4(f) shows the results of reconstruction by the proposed method. In this simulation, we set the parameters of the proposed method as follows: $w = 40$, $h = 30$, $\tilde{w}_1 = 6$, $\tilde{h}_1 = 5$, $\tilde{w}_2 = 20$, $\tilde{h}_2 = 15$, $\eta_1 = 5.0 \times 10^{-7}$, $\eta_2 = 0.0$, Th $= 0.5$ and $K = 2$. The details of the determination of these parameters are shown in Appendix. For better subjective evaluation, the enlarged portions around the lower right of the images are shown in Figs. 4(g)–(i). It can be seen that noticeable improvements have been achieved by using the proposed method. In the traditional method, the reconstructed textures suffer from some degradation due to different kinds of textures. On the other hand, the proposed method can adaptively reconstruct the missing textures from only the reliable ones by selecting the optimal cluster including the same kinds of textures. Furthermore, in

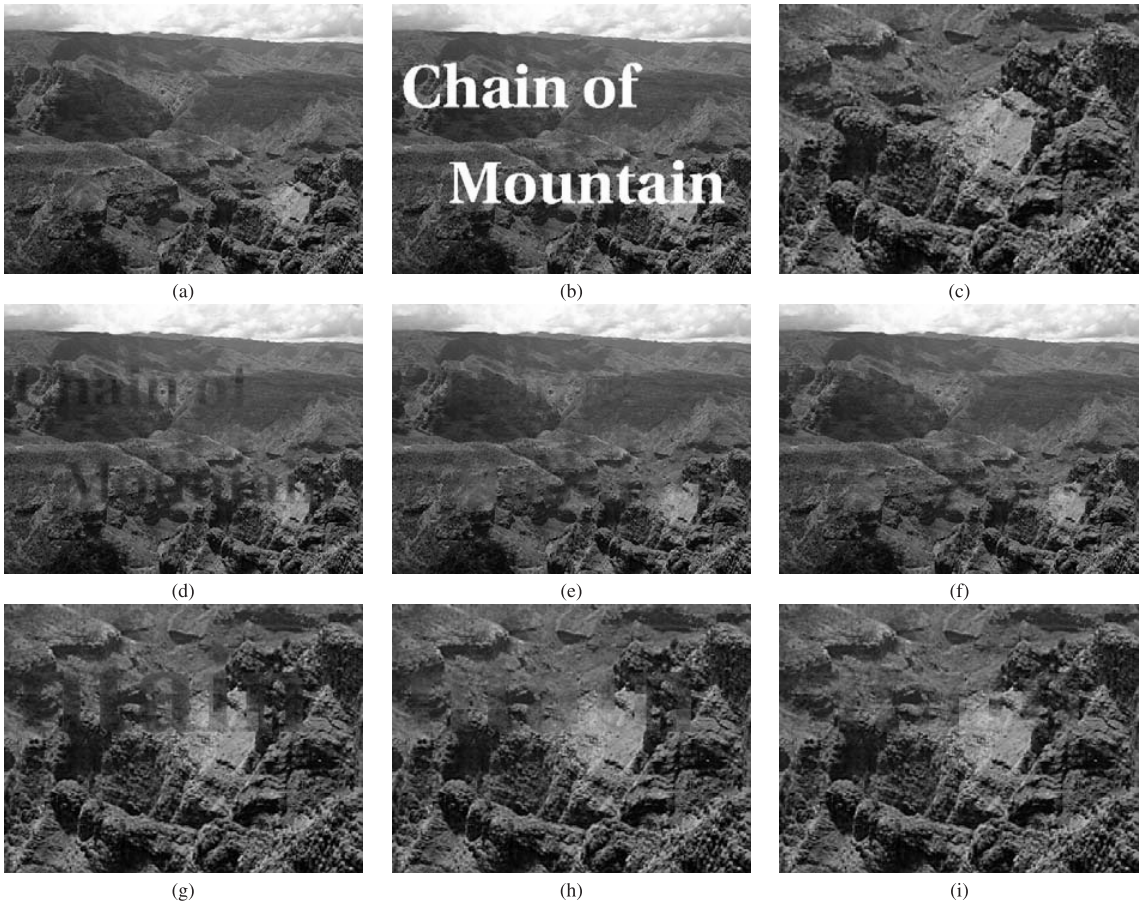

Fig. 5 (a) Original test image 2, (b) Corrupted image including text regions "Chain of Mountain" (8.9% loss), (c) Zoomed portion of (a), (d) Image reconstructed by reference [9], (e) Image reconstructed by reference [15], (f) Image reconstructed by the proposed method, (g) Zoomed portion of (d), (h) Zoomed portion of (e), (i) Zoomed portion of (f).

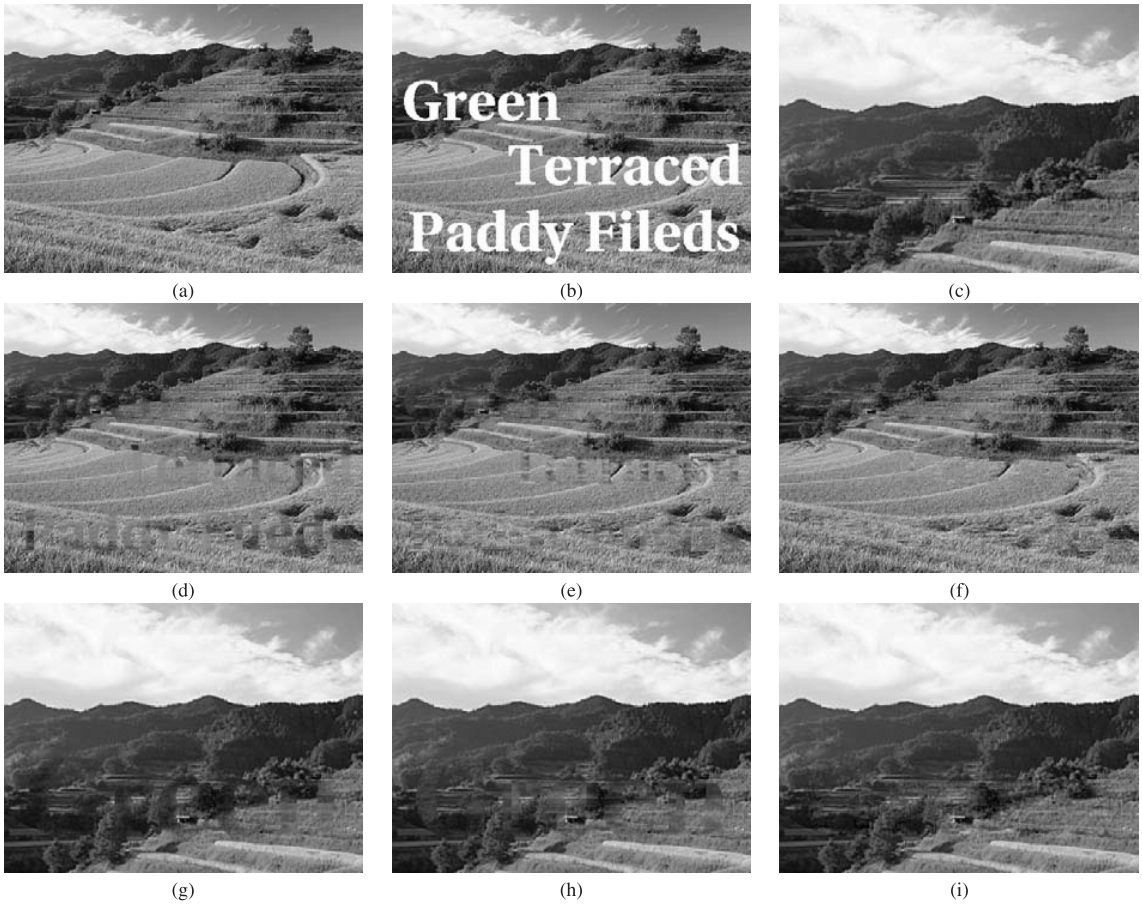

Fig. 6 (a) Original test image 3, (b) Corrupted image including text regions "Green Terraced Paddy Fields" (11.9% loss), (c) Zoomed portion of (a), (d) Image reconstructed by reference [9], (e) Image reconstructed by reference [15], (f) Image reconstructed by the proposed method, (g) Zoomed portion of (d), (h) Zoomed portion of (e), (i) Zoomed portion of (f).

the proposed method, the correlation of intensities between known and unknown areas are obtained more successfully than the conventional methods as shown in the previous section. Therefore, the missing textures can be reconstructed more accurately by the proposed method than by the conventional methods.

Different experimental results are shown in Figs. 5 and 6. Compared to the results obtained by the conventional methods, it can be seen that various kinds of textures can be accurately restored by using the proposed method. Furthermore, in order to quantitatively evaluate the performance of the proposed method, we show the PSNR of the reconstruction results in Table 1. The PSNR is defined as follows:

$$\mathrm{PSNR} = 10 \log_{10} \frac{\mathrm{MAX}^2}{\mathrm{MSE}}, \tag{33}$$

where MAX denotes the maximum value of intensities and MSE is the mean square error between the original image and the reconstructed image. It can be seen that our method has achieved an improvement of 0.55–2.06 dB over the traditional method [9]. However, some images reconstructed by reference [15] have higher PSNR values than those of the proposed method. Although the performance of the proposed method is really superior to that of the conventional method in subjective evaluation, the PSNR values cannot reflect those results. Thus, new quantitative evaluation methods should be considered in our future work.

Finally, we verify the computational loads of the proposed method and the conventional methods. Table 2 shows the comparison of the computation time between the proposed method and the conventional methods in [9] and [15]. This simulation was performed on a personal computer using Intel(R) Core(TM)2 Quad CPU Q6700 2.66 GHz with 2.0 Gbytes RAM. Both the proposed method and the conventional methods were implemented by using Matlab. As shown in this table, the computation time of the proposed method is much larger than those of the conventional methods. In the proposed method, we perform the clustering of known textures within the target image for each target local image f. On the other hand, the conventional method in [15] performs the clustering only once before the missing texture reconstruction process. Therefore, by applying this approach to our method, its computation cost will become much smaller. From this experiment, we can see that the reduction of the computational cost in the proposed method is

Table 1 Performance comparison (PSNR) of the proposed method and the conventional methods.

Image	Reference [9]	Reference [15]	**Proposed method**
Fig. 4	21.04 dB	21.84 dB	**21.59 dB**
Fig. 5	27.79 dB	30.22 dB	**29.85 dB**
Fig. 6	25.82 dB	26.13 dB	**26.46 dB**

Table 2 Comparison of computational time (sec) between the proposed method and the conventional methods.

Image	Reference [9]	Reference [15]	**Proposed method**
Fig. 4	4.27×10^3	2.51×10^2	5.65×10^4
Fig. 5	4.86×10^3	2.79×10^2	6.47×10^4
Fig. 6	4.13×10^3	2.56×10^2	5.82×10^4

necessary for practical use. This issue will be addressed in a future work.

5. Conclusions

In this paper, we have proposed a kernel CCA-based texture reconstruction method. The proposed method applies the kernel CCA to each cluster containing the same kind of textures and adaptively estimates the missing intensities from the optimal correlation. In order to select the optimal one, the error caused in the estimation scheme is introduced as a new criterion. This approach provides a solution to the problem in the traditional method of not being able to adaptively select the optimal textures for the missing area, and reliable textures can be utilized for reconstruction of the target textures. In addition, since the proposed method enables more accurate estimation of the correlation by using the kernel CCA than that by our previous works, impressive improvements in subjective measures have been achieved.

In the proposed method, we manually set parameters such as the size of local images and number of clusters. It is desirable that these parameters can be adaptively determined from the target image. Thus, we need to complement this determination algorithm. Furthermore, the proposed method takes more computational time than that of traditional methods. Therefore, reduction of computational cost is needed for practical use of the proposed method. These topics will be the subjects of subsequent studies.

Acknowledgment

This work was supported under project Regional Innovation Creation R&D Programs.

References

[1] C. Ballester, M. Bertalmio, V. Caselles, and G. Sapiro, "Filling-in by joint interpolation of vector fields and gray levels," IEEE Trans. Image Process., vol.10, no.8, pp.1200–1211, 2001.
[2] T.F. Chan and J. Shen, "Nontexture inpainting by curvature-driven diffusions," J. Vis. Commun. Image Represent., vol.12, no.4, pp.436–449, 2001.
[3] 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.
[4] A.A. Efros and T.K. Leung, "Texture synthesis by nonparametric sampling," IEEE Int. Conf. Computer Vision, pp.1033–1038, Corfu, Greece, Sept. 1999.
[5] A. Kokaram, "A statistical framework for picture reconstruction using 2D AR models," Image Vis. Comput., vol.22, no.2, pp.165–171, Feb. 2004.
[6] T. Amano and Y. Sato, "Image interpolation using BPLP method on the eigenspace," Systems and Computers in Japan, vol.38, no.1, pp.87–96, Jan. 2007.
[7] 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.
[8] S. Mika, B. Schölkoph, A. Smola, K.-R. Müller, M. Scholz, and G. Rätsch, "Kernel PCA and de-noising in feature spaces," Advances in Neural Information Processing Systems 11, pp.536–542, 1999.
[9] K.I. Kim, M.O. Franz, and B. Schölkoph, "Iterative kernel principal component analysis for image modeling," IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.9, pp.1351–1366, 2005.
[10] F.R. Fienup, "Phase retrieval algorithms: A comparison," Appl. Opt., vol.21, pp.2758–2769, 1982.
[11] T. Ogawa, M. Haseyama, and H. Kitajima, "Reconstruction method of missing texture using error reduction algorithm," Proc. ICIP 2005, vol.2, pp.1026–1029, Sept. 2005.
[12] D.C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1 — Theory," IEEE Trans. Med. Imaging, vol.MI-1, pp.81–94, Oct. 1982.
[13] T. Ogawa and M. Haseyama, "Adaptive reconstruction method of missing texture based on projection onto convex sets," Proc. ICASSP 2007, vol.1, pp.697–700, April 2007.
[14] 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.
[15] T. Ogawa and M. Haseyama, "POCS-based iterative reconstruction algorithm of missing textures," Proc. ICIP 2007, vol.3, pp.101–104, Sept. 2007.
[16] J. Shawe-Taylor and N. Cristianini, Kernel Methods for Pattern Analysis, Cambridge University Press, 2004.
[17] S. Akaho, "A kernel method for canonical correlation analysis," IMPS2001, 2001.
[18] J. Mercer, "Functions of positive and negative type and their connection with the theory of integral equations," Trans. Lond. Phil. Soc. (A), vol.209, pp.415–446, 1909.
[19] J.T.Y. Kwok and I.W.H. Tsang, "The pre-image problem in kernel methods," IEEE Trans. Neural Netw., vol.15, no.6, pp.1517–1525, 2004.

Appendix: Determination of Parameters

In Sect. 4, we set the parameters of the proposed method as follows: $w = 40$, $h = 30$, $\tilde{w}_1 = 6$, $\tilde{h}_1 = 5$, $\tilde{w}_2 = 20$, $\tilde{h}_2 = 15$, $\eta_1 = 5.0 \times 10^{-7}$, $\eta_2 = 0.0$, Th $= 0.5$ and $K = 2$. The details of the determination of those parameters are shown as follows.

(i) Size of local images: $w(= 40)$ and $h(= 30)$
In order to determine the optimal value, we changed the size of local images as 40×30, 60×45, and 80×60 pixels and calculated the PSNR values of the reconstruction results of three test images. These test images are gray-scale images of Figs. 4(b), 5(b), and 6(b) (480×360 pixels and 8-bit gray levels), and are also

utilized to determine other parameters in (iv)–(vi). Figure A· 1 shows the relationship between the size of local images and the PSNR values of the reconstruction results. From this figure, we can see the relationship between the size of local images and the PSNR values is different for each test image, and the quantitative performance is not sensitive for w and h. On the other hand, if the size of local images becomes larger, the subjective performance of the reconstruction results tends to become worse as shown in Fig. A· 2. Thus, it seems that $w = 40$ and $h = 30$ are the optimal. Furthermore, if the size of local images is set to a smaller value, the size of the matrices $\mathbf{L}^k$, $\mathbf{M}^k$, and $\mathbf{P}^k$ in Eqs. (18)–(20) becomes much larger, and the computation of Eqs. (21) and (22) becomes difficult. Therefore, we determined $w = 40$ and $h = 30$ in the experiments.

(ii) Sliding interval of local images for reconstructing missing textures: $\tilde{w}_2(= 20)$ and $\tilde{h}_2(= 15)$
In the proposed method, if the target local image f is clipped from the target image without any overlaps, several textures may be included within f. In such a case, the reconstruction performance of the proposed method tends to become worse. Therefore, the proposed method clips overlapped local images as f in the same interval $\tilde{w}_2(= 20)$ and $\tilde{h}_2(= 15)$, that is a half size of the target local image f. Note that if we set the clipping interval to a smaller value, the performance of the

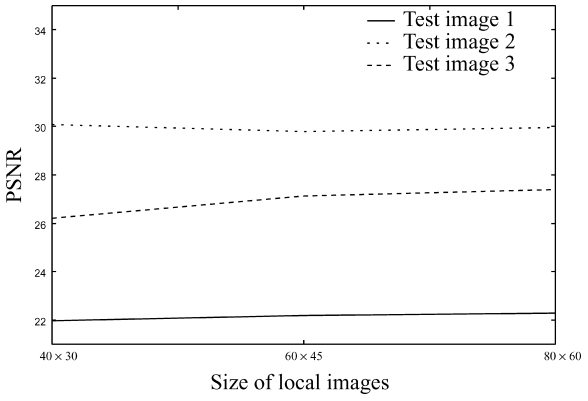

Fig. A· 1 Relationship between the size of local images and the PSNR values of the reconstruction results.

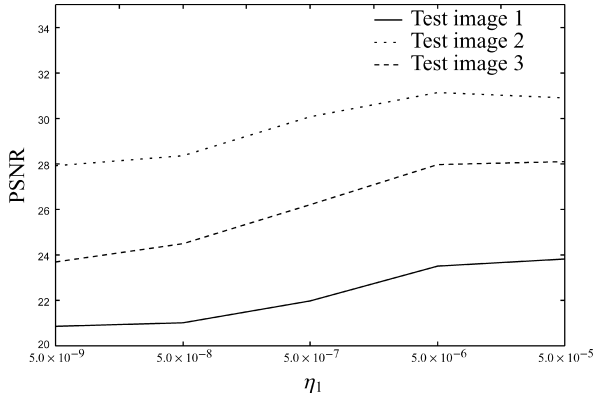

Fig. A· 3 Relationship between η_1 and the PSNR values of the reconstruction results.

(a) (b) (c)

Fig. A· 2 (a) Zoomed portion of Test image 3, (b) Zoomed portion of reconstruction results ($w = 40$ and $h = 30$), (c) Zoomed portion of reconstruction results ($w = 80$ and $h = 60$).

(a) (b) (c)

Fig. A· 4 (a) Zoomed portion of reconstruction results ($\eta_1 = 5.0 \times 10^{-8}$), (b) Zoomed portion of reconstruction results ($\eta_1 = 5.0 \times 10^{-7}$), (c) Zoomed portion of reconstruction results ($\eta_1 = 5.0 \times 10^{-6}$).

proposed method may become better, but its computation time becomes much larger. Therefore, we simply set $\tilde{w}_2 = 20$ and $\tilde{h}_2 = 15$ in the experiments.

(iii) Clipping interval for obtaining training local images: $\tilde{w}_1(= 6)$ and $\tilde{h}_1(= 5)$
In order to reconstruct missing textures, the proposed method clips training local images f_i $(i = 1, 2, \cdots, N)$ not including missing areas from the target image. It is desirable that the clipping interval $\tilde{w}_1$ and $\tilde{h}_1$ are respectively set to the same values as $\tilde{w}_2$ and $\tilde{h}_2$. However, if many missing areas are contained in the target image, it is difficult to obtain enough training local images f_i. Therefore, the proposed method sets the clipping interval in such a way that the number of the clipped local images f_i, N becomes larger than 2000 for all of the three test images shown in Figs. 4–6. Specifically, we set $\tilde{w}_2 = \lfloor \frac{w}{6} \rfloor$ and $\tilde{h}_2 = \lfloor \frac{h}{6} \rfloor$ in the experiments.

(iv) Regularization parameters: $\eta_1(= 5.0 \times 10^{-7})$ and $\eta_2(= 0.0)$
First, we changed the value of η_1 $(= 5.0 \times 10^{-9}, 5.0 \times 10^{-8}, 5.0 \times 10^{-7}, 5.0 \times 10^{-6}, 5.0 \times 10^{-5})$ and calculated the PSNR values of the reconstruction results of the three test images. Figure A·3 shows the relationship between η_1 and the PSNR values of the reconstruction results. From this figure, we can see the PSNR values almost monotonically increase with increasing η_1. However, when the regularization parameter is set to

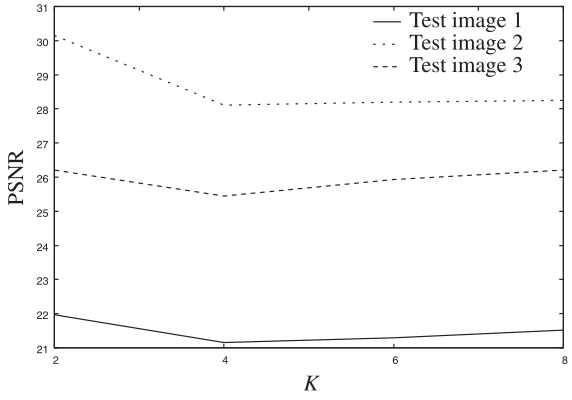

Fig. A·5 Relationship between K and the PSNR values of the reconstruction results.

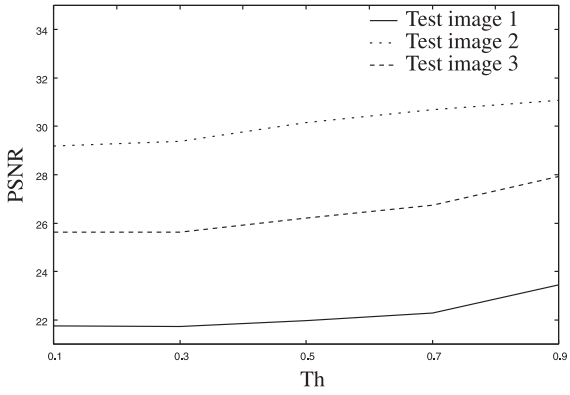

Fig. A·7 Relationship between Th and the PSNR values of the reconstruction results.

(a) (b) (c)

Fig. A·6 (a) Zoomed portion of Test image 1, (b) Zoomed portion of reconstruction results ($K = 2$), (c) Zoomed portion of reconstruction results ($K = 8$).

(a) (b) (c)

Fig. A·8 (a) Zoomed portion of Test image 2, (b) Zoomed portion of reconstruction results (Th = 0.5), (c) Zoomed portion of reconstruction results (Th = 0.9).

a larger value, the expression ability of textures tends to become worse. On the other hand, if it is set to a smaller value, the overfitting tends to occur. From the subjective evaluation shown in Fig. A· 4, these phenomena can be confirmed. Therefore, from this figure, we set the regularization parameter $\eta_1 = 5.0 \times 10^{-7}$ in the experiments.

Furthermore, since the proposed method did not utilize the nonlinear map for the vectors $\mathbf{y}_i$ $(i = 1, 2, \cdots, N)$ and $\mathbf{y}$, we thought the regularization parameter for those vectors was not necessary. Thus, we set $\eta_2 = 0.0$ in the experiments.

(v) Number of clusters: $K(= 2)$

In order to determine the number of clusters, we changed the value of K $(= 2, 4, 6, 8)$ and calculated the PSNR values of the reconstruction results of the three test images. Figure A· 5 shows the relationship between K and the PSNR values of the reconstruction results. From this figure, we can see the PSNR of the proposed method becomes the highest value when $K = 2$. This is because the training local images f_j^k $(j = 1, 2, \cdots, N^k)$ in cluster k becomes fewer when K becomes larger. In order to avoid this problem, we have to clip local images f_i $(i = 1, 2, \cdots, N)$ in a smaller interval, and then the computation time becomes much larger. Furthermore, we found the subjective performance of the proposed method was not severely affected by the value of K as shown in Fig. A· 6. Thus, we set the number of clusters $K = 2$ in the experiments. If we can known the number of textures contained in the target image, K can be adaptively determined. This will be the subject of the subsequent reports.

(vi) Threshold utilized for the determination of D^k: Th(= 0.5)

We changed the value of Th $(= 0.1, 0.3, 0.5, 0.7, 0.9)$ and verified the reconstruction performance of the three test images. Figure A· 7 shows the relationship between Th and the PSNR values of the reconstruction results by the proposed method. From this figure, we can see the PSNR values monotonically increase when Th becomes larger, and this can be easily expected. However, from subjective evaluation shown in Fig. A· 8, the reconstruction results of Th = 0.9 suffer from some blurring. On the other hand, the reconstruction performance of Th = 0.5 tends to be adequate in the proposed method. Therefore, we set Th = 0.5 in the experiments.

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 in image restoration and its applications.

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 is currently a professor in the Graduate School of Information Science and Technology, Hokkaido University. Her research interests are in digital signal processing and its applications.