HOKKAIDO UNIVERSITY

Title	Missing Image Data Reconstruction Based on Adaptive Inverse Projection via Sparse Representation
Author(s)	Ogawa, Takahiro; Haseyama, Miki
Citation	IEEE Transactions on Multimedia, 13(5), 974-992 https://doi.org/10.1109/TMM.2011.2161760
Issue Date	2011-10
Doc URL	http://hdl.handle.net/2115/47216
Rights	© 2011 IEEE. Reprinted, with permission, from Ogawa, T.; Haseyama, M.; Missing Image Data Reconstruction Based on Adaptive Inverse Projection via Sparse Representation, IEEE Transactions on Multimedia, Oct. 2011. This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Hokkaido University products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.
Type	article (author version)
File Information	ToM 13-5.pdf

Missing Image Data Reconstruction Based on Adaptive Inverse Projection via Sparse Representation

Takahiro Ogawa, *Member, IEEE* and Miki Haseyama, *Senior Member, IEEE,*

Abstract—In this paper, a missing image data reconstruction method based on an adaptive inverse projection via sparse representation is proposed. The proposed method utilizes sparse representation for obtaining low-dimensional subspaces that approximate target textures including missing areas. Then, by using the obtained low-dimensional subspaces, inverse projection for reconstructing missing areas can be derived to solve the problem of not being able to directly estimate missing intensities. Furthermore, in this approach, the proposed method monitors errors caused by the derived inverse projection, and the low-dimensional subspaces optimal for target textures are adaptively selected. Therefore, we can apply adaptive inverse projection via sparse representation to target missing textures, i.e., their adaptive reconstruction becomes feasible. The proposed method also introduces some schemes for color processing into the calculation of subspaces on the basis of sparse representation and attempts to avoid spurious color caused in the reconstruction results. Consequently, successful reconstruction of missing areas by the proposed method can be expected. Experimental results show impressive improvement of our reconstruction method over previously reported reconstruction methods.

Index Terms—Image reconstruction, image texture analysis, interpolation, inverse projection, sparse representation.

I. INTRODUCTION

Reconstruction of missing image data affords numerous applications in image processing, such as image enlargement and restoration of missing areas. Image enlargement is achieved by various kinds of missing intensity interpolation, and several interpolation-based resolution enhancement methods are included in commercial digital imaging software [1]–[4]. Missing image data reconstruction also realizes restoration of missing areas in digital images [5]–[18], and it is applied to removal of unnecessary objects, error concealment for video communications, etc. Compared to image enlargement, it is difficult to obtain natural reconstruction results from restoration of missing areas. Thus, only the restoration of missing areas in digital images is discussed in this paper.

T. Ogawa is with Graduate School of Information Science and Technology, Hokkaido University, Sapporo, 060-0814 JAPAN E-mail:ogawa@lmd.ist.hokudai.ac.jp

M. Haseyama is with Graduate School of Information Science and Technology, Hokkaido University, Sapporo, 060-0814 JAPAN E-mail:miki@ist.hokudai.ac.jp

Recently, many missing image data reconstruction methods have been proposed for missing area restoration. Generally, they can be broadly classified into two categories, structural and textural reconstruction approaches, and many papers on these approaches have been published. Attractive methods that perform simultaneous reconstruction of missing structures and textures in images have also been proposed [19]–[13]. Most of the algorithms so far reported are based on structural inpainting techniques for accurate reconstruction of missing edges [5]–[7]. These techniques are effective for pure structure images. However, since general images also contain many textures, different methods work better in these areas. Thus, several methods have been proposed for accurate reconstruction of missing textures [8]–[17]. The remainder of this paper focuses on the texture reconstruction approach with discussion of its details.

Traditionally, missing texture reconstruction is realized as one of the applications of texture synthesis. Efros et al. firstly proposed a pioneered method for the texture synthesis [8], [9]. Their approach models textures by using the MRF (Markov Random Field) model and enables missing texture reconstruction by copying pixels of a target image itself, i.e., non-parametric sampling in synthesis. Furthermore, Wei et al. proposed a fast algorithm for the searching step in the texture synthesis by using multi-resolution concepts [10]. Then many methods which perform the exemplar-based inpainting are mainly inspired by the non-parametric sampling in [8]. Drori et at. proposed a fragment-based algorithm for image completion which could preserve structures and textures [11]. Furthermore, the exemplar-based image inpainting method proposed by Criminisi et al. is a representative one based on the texture synthesis [12], [13]. This method adopts a patch-based greedy sampling scheme similar to the fragment based completion, but it is simpler and faster. In recent years, faster and more accurate reconstruction methods which improve the approach shown in [12], [13] have been proposed [14], [16]. In [14], Kwok et al. introduce decomposition of exemplars into the frequency coefficients and select fewer coefficients that are the most important for evaluating the matching score. Furthermore, a local gradient-based algorithm to interpolate missing areas in a query patch is also proposed. Then these approaches enable much faster exemplar-based image completion and provide better results in several cases since selecting the most significant coefficients in the frequency domain has a denoising effect. In [16], Xu et al. newly introduce structure sparsity to determine priorities for inpainting. Furthermore, sparse

representation for reconstructing missing areas from selected exemplars is performed to achieve accurate performance. A good review of the exemplar-based inpainting methods based on [8] is shown in [15].

Reconstruction of missing areas is an ill-posed problem, and it is difficult to directly estimate missing intensities. Thus, most reconstruction methods perform approximation of textures within target images in lower-dimensional subspaces and derive inverse projection for the corruption. Missing areas can then be reconstructed by using the obtained inverse projection. Various algorithms including PCA, CCA, and sparse representation have been used for obtaining low-dimensional subspaces. Specifically, Amano et al. proposed an effective PCA-based method for reconstructing missing textures using back projection for lost pixels and achieved accurate reconstruction performance [18]. Furthermore, kernel methods have recently been developed and their achievements have been reported in a number of papers [21]–[23]. Subspaces constructed on the basis of kernel methods are also suitable for approximating nonlinear texture features in target images. Thus, several missing texture reconstruction methods that utilize projection schemes onto nonlinear subspaces obtained by kernel PCA and CCA have been proposed [24], [25]. In these methods, known textures are previously clustered, and the subspaces of the optimal clusters are adaptively selected for reconstructing target missing textures. This scheme improves the performance of missing texture reconstruction compared to the performance of traditional methods using only one type of subspace. However, it should be noted that the number of subspaces utilized in these methods is only the same as the number of clusters. Thus, since it is difficult to adaptively generate the optimal subspace for reconstructing each missing texture, approximation performance in lower-dimensional subspaces is not always satisfactory.

Recently, image reconstruction that utilizes sparse representation has been intensively studied [26], [27]. Sparse representation enables the selection of optimal signal-atoms from a dictionary for approximating target signals. Thus, by using sparse representation, subspaces that are suitable for restoring missing areas can be provided more adaptively than by conventional methods. Therefore, several missing area reconstruction methods based on sparse representation have been proposed [16], [28]–[30]. However, it should be noted that conventional methods select optimal signal-atoms for representing target textures from only known neighboring areas. Thus, there is no guarantee for providing subspaces optimal for reconstructing missing textures. Furthermore, these methods estimate missing textures based on the sparse representation that is derived from only the known neighboring areas. This means they assume the sparse representation coefficients become the same between the target textures including missing areas and the neighboring known textures. As the size of missing areas becomes larger, there is a tendency for the assumption not to be satisfied, and reconstruction performance of these methods tends to become worse.

In this paper, a novel missing image data reconstruction method based on an adaptive inverse projection via sparse representation (IPVSR) is proposed. The proposed method performs sparse representation of target textures including missing areas and the corresponding original unknown textures for obtaining their optimal subspaces. Then these two textures can be approximated into lower-dimensional subspaces, and derivation of the inverse projection for the corruption becomes feasible. Since projection from the corrupted textures to their original ones can be directly derived, our method does not rely on the above conventional assumption. Furthermore, in order to realize this approach, the optimal signal-atoms must be selected for obtaining the optimal subspaces. Therefore, the proposed method monitors errors caused by the derived inverse projection and realizes adaptive selection of the optimal signal-atoms, i.e., the subspaces utilized for estimating the inverse projection are adaptively generated for each texture. This approach provides a solution to the problem of conventional methods not being able to correctly select optimal signal-atoms suitable for reconstruction of the target textures. In addition, the proposed method introduces some novel schemes that avoid spurious color in reconstruction results into the whole procedures of sparse representation. Consequently, the proposed method performs reconstruction of missing areas more successfully than do conventional methods.

This paper is organized as follows. First, in Section II, we briefly explain sparse representation of signals. Next, a missing image data reconstruction method based on adaptive IPVSR is proposed in Section III. Experimental results that verify the performance of the proposed method are shown in Section IV. Finally, conclusions are given in Section V.

II. SPARSE REPRESENTATION

Sparse representation of signals is explained in this section. Given an overcomplete dictionary $\mathbf{D} \in \mathbf{R}^{n \times K}$ whose columns are prototype signal-atoms $\mathbf{d}_j \in \mathbf{R}^n$ ($j = 1, 2, \cdots, K$), a target signal $\mathbf{y} \in \mathbf{R}^n$ can be represented as a sparse linear combination of these signal-atoms. Specifically, $\mathbf{y}$ can be approximated as

$$\mathbf{y} \cong \mathbf{D}\mathbf{x} \qquad \left(\mathbf{x} \in \mathbf{R}^K \right), \tag{1}$$

where $\mathbf{x}$ is a vector containing the representation coefficients of signal $\mathbf{y}$, and it satisfies

$$\|\mathbf{y} - \mathbf{D}\mathbf{x}\|_p \leq \epsilon. \tag{2}$$

In this paper, we set p to two.

If $n < K$ and $\mathbf{D}$ is a full-rank matrix, an infinite number of solutions are available for the above representation problem. Thus, new constraints are introduced into this problem, and the sparsest solution is obtained by solving

$$(P_0) \quad \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\mathbf{x} \tag{3}$$

or

$$(P_{0,\epsilon}) \quad \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \epsilon, \tag{4}$$

where $\| \cdot \|_0$ represents the l^0-norm, counting the non-zero entries of a target vector. It is well known that the calculation of the optimal solution in Eqs. (3) or (4) is an NP-hard problem [31]. Thus, there have been several conventionally proposed methods that approximately provide solutions of the above problems, and their simplest ones are matching pursuit (MP)

[32] and orthogonal matching pursuit (OMP) algorithms [33]–[36]. The basis pursuit (BP) algorithm is also a representative algorithm solving the problems in Eqs. (3) and (4) by replacing the l^0-norm with an l^1-norm [37]. The focal underdetermined system solver (FOCUSS) is a similar algorithm using l^p-norm ($p \leq 1$) [38]–[40].

Next, given a set of signal vectors $\mathbf{y}_i$ ($i = 1, 2, \cdots, N$), there exist dictionary matrices providing the sparse solution $\mathbf{x}_i$. The K-SVD algorithm [26] can provide the optimal dictionary matrix $\mathbf{D}$ and coefficient vectors $\mathbf{x}_i$ ($i = 1, 2, \cdots, N$) by solving

$$\min_{\mathbf{D}, \mathbf{X}} \{\|\mathbf{A}(\mathbf{Y} - \mathbf{DX})\|_F^2\} \quad \text{subject to} \quad \forall i, \quad \|\mathbf{x}_i\|_0 \leq T, \tag{5}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N]$ and $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_N]$, and $\|\cdot\|_F$ represents the Frobenius norm. Furthermore, T determines the sparsity of the signals. The K-SVD algorithm approximately calculates the optimal solution of Eq. (5) by iterating the calculation of $\mathbf{x}_i$ based on the OMP algorithm and renewal of the dictionary matrix $\mathbf{D}$. Note that in the conventional K-SVD algorithm, $\mathbf{A}$ is set to the identity matrix.

III. MISSING IMAGE DATA RECONSTRUCTION BASED ON ADAPTIVE IPVSR

In this section, the missing image data reconstruction method based on adaptive IPVSR is presented. The proposed method clips a patch $\hat{f}$ including missing areas from a target image and estimates its missing intensities based on IPVSR. The known and unknown areas within the target patch $\hat{f}$ are denoted as $\bar{\Omega}$ and Ω, respectively. Figure 1 shows the stage-diagram of the proposed reconstruction method. In order to perform accurate reconstruction for the target image containing several kinds of textures, we have to calculate dictionary matrices suitable for the target image and adaptively select the optimal prototype signal-atoms for each target patch including missing areas. Therefore, in this section, we first explain the calculation of the dictionaries for sparse representation (III-A). Specifically, we prepare two types of dictionaries for representing corrupted patches including missing areas and their corresponding original patches. It should be noted that the dictionary of the original patches cannot be obtained directly. Therefore, we estimate it by using known areas within the target image. From the obtained dictionaries, the proposed method adaptively selects the optimal signal-atoms and enables the generation of subspaces for the target patch $\hat{f}$. Then the IPVSR can be derived by using the obtained subspaces to reconstruct the missing area Ω. Its details are shown in III-B. Finally, in III-C, we discuss the effectiveness of the proposed method equipped with the above novel approaches.

A. Calculation of Dictionary Matrices

In this subsection, the algorithm for designing dictionaries is presented. First, we clip known patches f_i ($i = 1, 2, \cdots, N$) whose size is $w \times h$ pixels from the target image in the same interval. Next, for each patch f_i, a vector $\mathbf{y}_i \in \mathbf{R}^{3wh}$, whose elements are raster-scanned pixel values, is defined. Note that the first wh elements of this vector are raster-scanned R (red) component values, and the second wh values are G (green) component values, followed by B (blue) values. Using an
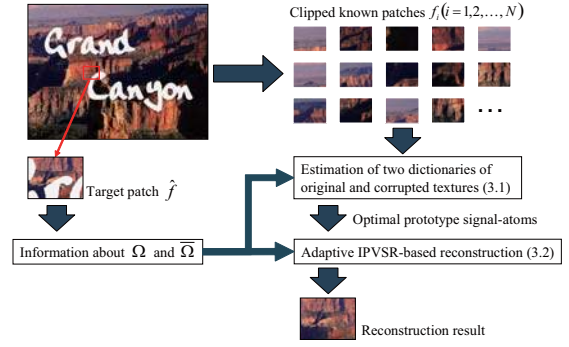

Fig. 1. Stage-diagram of the proposed missing texture reconstruction method.

overcomplete dictionary matrix $\mathbf{D} \in \mathbf{R}^{3wh \times K}$ that contains K prototype signal-atoms $\mathbf{d}_j \in \mathbf{R}^{3wh}$ ($j = 1, 2, \cdots, K$), each vector $\mathbf{y}_i$ is represented as a sparse linear combination of these atoms

$$\mathbf{Gy}_i \cong \mathbf{GDx}_i, \tag{6}$$

where $\mathbf{y}_i$ satisfies

$$\|\mathbf{G}(\mathbf{y}_i - \mathbf{Dx}_i)\|^2 \leq \epsilon \tag{7}$$

for a fixed value ϵ. The vector $\mathbf{x}_i \in \mathbf{R}^K$ contains the representation coefficients of $\mathbf{y}_i$. In the above equations,

$$\mathbf{G} = \mathbf{I} + \gamma \mathbf{K}, \tag{8}$$

where $\mathbf{I}$ is the $3wh \times 3wh$ identity matrix and

$$\mathbf{K} = \begin{bmatrix} \mathbf{J} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J} \end{bmatrix}. \tag{9}$$

In the above equation, $\mathbf{J}$ is a $wh \times wh$ matrix filled with ones. Furthermore, γ in Eq. (8) is a positive constant. As shown in [28], the matrix $\mathbf{G}$ in Eqs. (6) and (7) is introduced for avoiding spurious color in the sparse representation. The matrix $\mathbf{G}$ defined in Eq. (8) is introduced for avoiding spurious color in the sparse representation. This matrix contains two terms $\mathbf{I}$ and $\gamma \mathbf{K}$. If we ignore the second term, i.e., $\mathbf{G} = \mathbf{I}$, the left part in Eq. (7) becomes a simple Euclidean distance between $\mathbf{y}_i$ and $\mathbf{Dx}_i$. Thus, by focusing on this point, the left part of Eq. (7) can be rewritten as follows:

$$\begin{aligned} &\|\mathbf{G}(\mathbf{y}_i - \mathbf{Dx}_i)\|^2 \\ &= \|(\mathbf{I} + \gamma \mathbf{K})(\mathbf{y}_i - \mathbf{Dx}_i)\|^2 \\ &= \|(\mathbf{y}_i - \mathbf{Dx}_i) + \gamma(\mathbf{Ky}_i - \mathbf{KDx}_i)\|^2 \\ &= \left\|(\mathbf{y}_i - \mathbf{Dx}_i) + wh\gamma\left(\frac{1}{wh}\mathbf{Ky}_i - \frac{1}{wh}\mathbf{KDx}_i\right)\right\|^2. \end{aligned} \tag{10}$$

From the above equation, the first wh elements, the second wh elements, and the third wh elements in the vector $\frac{1}{wh}\mathbf{Ky}_i$ are respectively the average values of RGB components of $\mathbf{y}_i$. Furthermore, the first wh elements, the second wh elements, and the third wh elements in the vector $\frac{1}{wh}\mathbf{KDx}_i$ are respectively the average values of RGB components of $\mathbf{Dx}_i$. Therefore, $\left(\frac{1}{wh}\mathbf{Ky}_i - \frac{1}{wh}\mathbf{KDx}_i\right)$ in the above equation outputs

the differences of the average values for RGB components between $\mathbf{y}_i$ and $\mathbf{D}\mathbf{x}_i$. The proposed method takes into account the average color and enables the sparse representation which can avoid spurious color.

If $3wh < K$ and $\mathbf{D}$ is a full-rank matrix, an infinite number of solutions are available for the representation problems. Thus, in our method, the solution of

$$\min_{\mathbf{x}_i} \|\mathbf{x}_i\|_0 \text{ subject to } \left\| \mathbf{G}\left(\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\right)\right\|^2 \le \epsilon \qquad (11)$$

is adopted. The dictionary matrix $\mathbf{D}$ satisfying Eq. (11) is obtained by solving the minimization problem in Eq. (5), where $\mathbf{A} = \mathbf{G}$. In the proposed method, the optimal dictionary matrix $\mathbf{D}$ is estimated by using the K-SVD algorithm [26].

In the proposed method, we have to estimate another dictionary matrix. i.e., that of corrupted patches. For each patch f_i, the proposed method adds missing areas Ω at the same position in $\hat{f}$, and a corrupted set of patches following the pattern in $\hat{f}$, $\hat{f}_i$ $(i = 1, 2, \cdots, N)$ is obtained. Then, from $\hat{\mathbf{y}}_i$ $(i = 1, 2, \cdots, N)$ obtained in the same way as $\mathbf{y}_i$, a dictionary matrix $\hat{\mathbf{D}}$ can be also estimated by solving

$$\min_{\hat{\mathbf{D}},\hat{\mathbf{X}}} \left\{ \left\| \hat{\mathbf{G}}\left(\hat{\mathbf{Y}} - \hat{\mathbf{D}}\hat{\mathbf{X}}\right)\right\|_F^2 \right\}$$
$$\text{subject to } \|\hat{\mathbf{x}}_i\|_0 \le T \quad (i = 1, 2, \cdots, N) \qquad (12)$$

based on the K-SVD algorithm, where $\hat{\mathbf{Y}} = [\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2, \cdots, \hat{\mathbf{y}}_N]$ and $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \cdots, \hat{\mathbf{x}}_N]$. Furthermore,

$$\hat{\mathbf{G}} = (\mathbf{I} + \hat{\gamma}\mathbf{K})\,\mathbf{\Sigma} \qquad (13)$$

and $\hat{\gamma}$ is a positive constant defined by

$$\hat{\gamma} = \frac{wh}{N_{\bar{\Omega}}}\gamma. \qquad (14)$$

The value $N_{\bar{\Omega}}$ is the number of pixels in $\bar{\Omega}$. Furthermore, $\mathbf{\Sigma}$ is a diagonal matrix whose diagonal elements are zero or one. The matrix $\mathbf{\Sigma}$ removes the intensities of the corresponding pixels in Ω.

Note that if the size of the patches becomes larger, the computational complexity also becomes higher. Thus, in such a case, we calculate the dictionary matrix $\hat{\mathbf{D}}$ in the following way:

1) Calculate $\tilde{\mathbf{D}}$ from $\mathbf{D}$ by using the matrix $\mathbf{\Sigma}$ as follows:

$$\tilde{\mathbf{D}} = \mathbf{\Sigma}\mathbf{D}. \qquad (15)$$

2) Normalize each prototype signal-atom $\tilde{\mathbf{d}}_j \in \mathbf{R}^{3wh}$ in $\tilde{\mathbf{D}}$ as

$$\hat{\mathbf{d}}_j = \frac{\tilde{\mathbf{d}}_j}{\|\tilde{\mathbf{d}}_j\|} \quad (j = 1, 2, \cdots, K) \qquad (16)$$

and obtain $\hat{\mathbf{D}}$ whose columns are $\hat{\mathbf{d}}_j$ $(j = 1, 2, \cdots, K)$. Then, by using the above two procedures, we can simply obtain the dictionary matrix $\hat{\mathbf{D}}$ from $\mathbf{D}$.

As described above, we can obtain the dictionary matrices for representing original patches and their corresponding corrupted patches. By selecting optimal signal-atoms in the dictionary matrices, the subspaces for representing original and corrupted patches can be generated, and reconstruction on the basis of IPVSR becomes feasible. The details of the proposed reconstruction algorithm are shown in the following section.
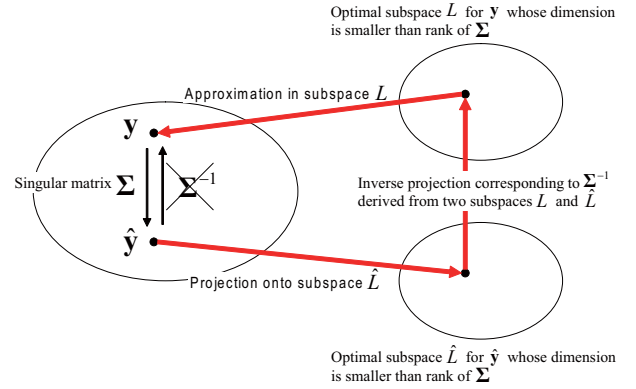

Fig. 2.   Visual illustration of the essential idea in the proposed reconstruction algorithm based on adaptive inverse projection via sparse representation.

B. Missing Area Reconstruction Algorithm

The adaptive IPVSR-based missing area reconstruction algorithm using the two dictionary matrices obtained in the previous subsection is presented in this section. The proposed method reconstructs the missing area Ω in $\hat{f}$ and estimates its original image f. First, we define vectors of the original patch f and its corrupted image $\hat{f}$ as $\mathbf{y}$ and $\hat{\mathbf{y}}$, respectively, which are obtained in the same way as $\mathbf{y}_i$ and $\hat{\mathbf{y}}_i$. These two vectors satisfy the following equation:

$$\hat{\mathbf{y}} = \mathbf{\Sigma}\mathbf{y}, \qquad (17)$$

and we assume

$$\hat{\mathbf{G}}\hat{\mathbf{y}} \cong \mathbf{\Sigma}\mathbf{G}\mathbf{y}. \qquad (18)$$

In the above equations, $\mathbf{\Sigma}$ is singular and its inverse matrix cannot be calculated directly to estimate the missing intensities in Ω. Thus, in order to solve the problem due to the singular matrix, the proposed method tries to estimate a smaller full-rank matrix corresponding to $\mathbf{\Sigma}$ in Eq. (17), whose dimension is smaller than the rank of $\mathbf{\Sigma}$. A visual illustration of the proposed reconstruction algorithm is given in Fig. 2. Furthermore, we define the projection of $\hat{\mathbf{y}}$ onto lower-dimensional subspace $\hat{L}$ and also obtain the approximation relationship of $\mathbf{y}$ in lower-dimensional subspace L. The details of the calculation of $\hat{L}$ and L are shown later. Then, through the projection onto $\hat{L}$, the inverse projection based on the smaller full-rank matrix, and the approximation in L, we can recover $\mathbf{y}$ from $\hat{\mathbf{y}}$ as shown in Fig. 2. In order to realize this idea, we have to provide the lower-dimensional subspaces suitable for deriving the inverse projection. In [18], they utilize eigenspaces obtained from training patches as the lower-dimensional subspaces in Fig. 2, and missing area reconstruction is realized. On the other hand, the proposed method utilizes the sparse representation for providing the two lower-dimensional subspaces suitable for the two vectors $\mathbf{y}$ and $\hat{\mathbf{y}}$. Since the adaptive estimation of the optimal subspaces can be expected by using the sparse representation, we adopt it for the derivation of the inverse projection. In the rest of this subsection, the specific algorithms of the proposed reconstruction method are shown.

Using the dictionaries $\mathbf{D}$ and $\hat{\mathbf{D}}$, $\mathbf{y}$ and $\hat{\mathbf{y}}$ satisfy the following equations:

$$\mathbf{Gy} \cong \mathbf{GDx} \quad \text{subject to} \quad \|\mathbf{x}\|_0 \leq T, \tag{19}$$

$$\hat{\mathbf{G}}\hat{\mathbf{y}} \cong \hat{\mathbf{G}}\hat{\mathbf{D}}\hat{\mathbf{x}} \quad \text{subject to} \quad \|\hat{\mathbf{x}}\|_0 \leq T. \tag{20}$$

Furthermore, these two equations can be respectively rewritten as

$$\mathbf{Gy} \cong \mathbf{GDEp}, \tag{21}$$

$$\hat{\mathbf{G}}\hat{\mathbf{y}} \cong \hat{\mathbf{G}}\hat{\mathbf{D}}\hat{\mathbf{E}}\hat{\mathbf{p}}, \tag{22}$$

where $\mathbf{p}$ and $\hat{\mathbf{p}} \in \mathbf{R}^T$ respectively contain non-zero elements of $\mathbf{x}$ and $\hat{\mathbf{x}}$. Furthermore, $\mathbf{E}$ and $\hat{\mathbf{E}} \in \mathbf{R}^{K \times T}$ are extraction matrices for obtaining signal-atoms $\mathbf{d}_j$ and $\hat{\mathbf{d}}_j$ ($j = 1, 2, \cdots, K$) utilized for approximating $\mathbf{y}$ and $\hat{\mathbf{y}}$, respectively. Specifically, the elements of the extraction matrices $\mathbf{E}$ and $\hat{\mathbf{E}}$ are one or zero. If ξ-th ($\xi = 1, 2, \cdots, T$) nonzero element exists in ζ-th ($\zeta = 1, 2, \cdots, K$) element of $\mathbf{x}$, (ζ, ξ)-th element of $\mathbf{E}$ becomes one, and the other elements become zero. The extraction matrix $\hat{\mathbf{E}}$ is also defined in the same way as the above explanation. Then they satisfy the following equations:

$$\mathbf{x} = \mathbf{Ep}, \tag{23}$$

$$\hat{\mathbf{x}} = \hat{\mathbf{E}}\hat{\mathbf{p}}. \tag{24}$$

Furthermore, $\mathbf{DE}$ and $\hat{\mathbf{D}}\hat{\mathbf{E}}$ represent the matrices whose columns are the signal-atoms only utilized for the sparse representation of $\mathbf{y}$ and $\hat{\mathbf{y}}$, respectively. Note that the subspaces spanned by the columns of $\mathbf{DE}$ and $\hat{\mathbf{D}}\hat{\mathbf{E}}$ respectively correspond to L and $\hat{L}$ in Fig. 2. For the following explanation, we define $\mathbf{S} = \mathbf{GDE}$ and $\hat{\mathbf{S}} = \hat{\mathbf{G}}\hat{\mathbf{D}}\hat{\mathbf{E}}$. Note that from Eq. (21), the following equation is obtained:

$$\mathbf{Gy} = \mathbf{Sp}. \tag{25}$$

Thus,

$$\mathbf{S}'\mathbf{Gy} = \mathbf{S}'\mathbf{Sp}, \tag{26}$$

and

$$\mathbf{p} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'\mathbf{Gy}. \tag{27}$$

Then since $\mathbf{Gy}$ can be approximated as

$$\mathbf{Gy} \cong \mathbf{GDE}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'\mathbf{Gy}, \tag{28}$$

the following equation can be obtained:

$$\mathbf{y} \cong \mathbf{DE}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'\mathbf{Gy}. \tag{29}$$

Therefore, from the above derivation, $\mathbf{DE}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'\mathbf{G}$ becomes the projection operator for L. Similarly, $\hat{\mathbf{D}}\hat{\mathbf{E}}(\hat{\mathbf{S}}'\hat{\mathbf{S}})^{-1}\hat{\mathbf{S}}'\hat{\mathbf{G}}$ becomes the projection operator for $\hat{L}$.

Next, by substituting Eqs. (21) and (22) into Eq. (18), the following relation is obtained:

$$\hat{\mathbf{S}}\hat{\mathbf{p}} \cong \mathbf{\Sigma}\mathbf{Sp}. \tag{30}$$
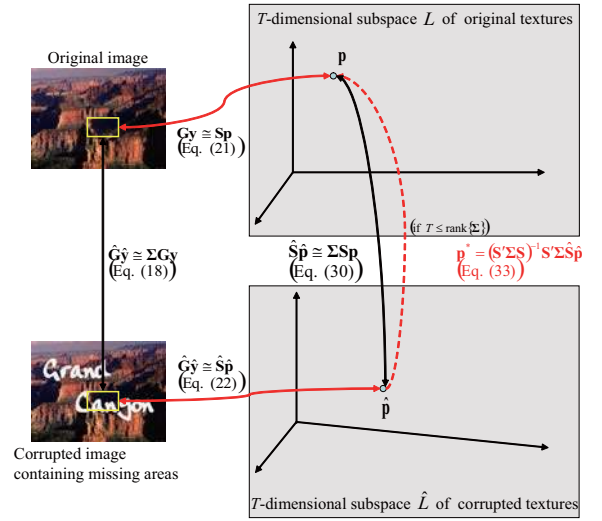
Fig. 3. Relationships shown in Eqs. (17)–(33) for deriving adaptive inverse projection via sparse representation.

In the above equation, if $T \leq \text{rank}\{\mathbf{\Sigma}\}$ is satisfied, the pseudo inverse matrix can be calculated. Specifically, from the above equation,

$$(\mathbf{\Sigma}\mathbf{S})' \mathbf{\Sigma}\mathbf{Sp} \cong (\mathbf{\Sigma}\mathbf{S})' \hat{\mathbf{S}}\hat{\mathbf{p}}, \tag{31}$$

where vector/matrix transpose is denoted by the superscript $'$ in this paper. Then Eq. (31) is rewritten as

$$\mathbf{S}'\mathbf{\Sigma}\mathbf{Sp} \cong \mathbf{S}'\mathbf{\Sigma}\hat{\mathbf{S}}\hat{\mathbf{p}}. \tag{32}$$

It should be noted that we utilize $\mathbf{\Sigma}' = \mathbf{\Sigma}$ and $\mathbf{\Sigma}\mathbf{\Sigma} = \mathbf{\Sigma}$. In Eq. (32), the rank of $\mathbf{E}$ becomes T, and those of $\mathbf{D}$ and $\mathbf{G}$ are larger than T. Therefore, if the above condition, i.e., $T \leq \text{rank}\{\mathbf{\Sigma}\}$, is satisfied, the rank of $\mathbf{S}'\mathbf{\Sigma}\mathbf{S}$ becomes T. Then since it becomes a full rank matrix, its inverse matrix can be obtained to calculate the estimation result $\mathbf{p}^*$ of $\mathbf{p}$ as

$$\mathbf{p}^* = (\mathbf{S}'\mathbf{\Sigma}\mathbf{S})^{-1} \mathbf{S}'\mathbf{\Sigma}\hat{\mathbf{S}}\hat{\mathbf{p}}, \tag{33}$$

where $(\mathbf{S}'\mathbf{\Sigma}\mathbf{S})^{-1} \mathbf{S}'\mathbf{\Sigma}$ is the pseudo-inverse matrix. Furthermore, by substituting Eq. (33) into Eq. (21),

$$\begin{aligned}\mathbf{Gy}^* &= \mathbf{Sp}^* \\ &= \mathbf{S}(\mathbf{S}'\mathbf{\Sigma}\mathbf{S})^{-1} \mathbf{S}'\mathbf{\Sigma}\hat{\mathbf{S}}\hat{\mathbf{p}}. \end{aligned} \tag{34}$$

Then estimation result $\mathbf{y}^*$ of $\mathbf{y}$ can be obtained by using IPVSR. In Fig. 3, we show the relationships shown in the above equations for deriving IPVSR.

In order to obtain the estimation result $\mathbf{y}^*$ by the above equation, we have to calculate the non-zero coefficient vector $\hat{\mathbf{p}}$ and the extraction matrices $\hat{\mathbf{E}}$ and $\mathbf{E}$. Note that the calculation of $\hat{\mathbf{p}}$ and $\hat{\mathbf{E}}$ corresponds to the calculation of $\hat{\mathbf{x}}$ as shown in Eqs. (20) and (22). Therefore, the proposed method calculates the solution of the problem in Eq. (12) for the fixed dictionary matrix $\hat{\mathbf{D}}$. In this scheme, we use the OMP algorithm[1] and obtain the non-zero coefficient vector $\hat{\mathbf{p}}$ and the extraction

---

[1]In this paper, we utilize the OMP algorithm for its simplicity. Several approaches for solving this problem can be adopted, and better estimation results can be expected.

Initialization:

$\mathbf{y}_0 = \mathbf{0}$, $\mathbf{R}_0\mathbf{y} = \hat{\mathbf{y}}$, $D_0 = \{\}$, $k = 0$.

(I) Compute $\left\{E_{j_k}; \mathbf{d}_j \in D \backslash D_k \right\}$, where

$$E_{j_k} = \left\| \hat{\mathbf{G}}\mathbf{R}_k\mathbf{y} - \mathbf{\Sigma}\mathbf{G}\mathbf{y}_j^* \right\|^2,$$

and $\mathbf{y}_j^*$ satisfies

$$\mathbf{G}\mathbf{y}_j^* = \frac{\mathbf{G}\mathbf{d}_j\mathbf{d}_j{}'\mathbf{G}'\mathbf{\Sigma}\hat{\mathbf{S}}\hat{\mathbf{p}}}{\mathbf{d}_j{}'\mathbf{G}'\mathbf{\Sigma}\mathbf{G}\mathbf{d}_j}.$$

(II) Find $\mathbf{d}_{j_{k+1}} \in D \backslash D_k$ such that

$$\mathbf{d}_{j_{k+1}} = \arg \min_{\mathbf{d}_j \in D \backslash D_k} E_{j_k}.$$

(III) If $k = T$, then stop.

(IV) Set $D_{k+1} = D_k \cup \{\mathbf{d}_{j_{k+1}}\}$ and compute the matrix $\mathbf{DE}$ from D_{k+1}.

(V) Compute $\mathbf{y}_k^*$ satisfying

$$\mathbf{G}\mathbf{y}_k^* = \mathbf{S}\left(\mathbf{S}'\mathbf{\Sigma}\mathbf{S}\right)^{-1}\mathbf{S}'\mathbf{\Sigma}\hat{\mathbf{S}}\hat{\mathbf{p}}$$

and update

$$\mathbf{R}_{k+1}\mathbf{y} = \hat{\mathbf{y}} - \mathbf{y}_k^*.$$

(VI) Set $k \leftarrow k + 1$, and repeat (I)–(VI).

Fig. 4. Procedure of the algorithm solving Eq. (36) in our method.

matrix $\hat{\mathbf{E}}$. Next, we have to calculate the extraction matrix $\mathbf{E}$. Unfortunately, it is difficult to directly calculate $\mathbf{E}$ from the solution of

$$\min_{\mathbf{x}} \|\mathbf{G}(\mathbf{y} - \mathbf{Dx})\|^2 \text{ subject to } \|\mathbf{x}\|_0 \le T \quad (35)$$

since the original vector $\mathbf{y}$ is unknown. Thus, in order to solve this problem, the proposed method utilizes the solution of

$$\min_{\mathbf{E}} \|\hat{\mathbf{G}}\hat{\mathbf{y}} - \mathbf{\Sigma}\mathbf{S}\left(\mathbf{S}'\mathbf{\Sigma}\mathbf{S}\right)^{-1}\mathbf{S}'\mathbf{\Sigma}\hat{\mathbf{S}}\hat{\mathbf{p}}\|^2$$
$$\text{subject to } \|\mathbf{x}^*\|_0 \le T \quad (36)$$

as a substitute for Eq. (35), where

$$\mathbf{x}^* = \mathbf{E}\mathbf{p}^*. \quad (37)$$

It should be noted that Eq. (36) can be also rewritten as the following simple equation from Eq. (34):

$$\min_{\mathbf{E}} \|\hat{\mathbf{G}}\hat{\mathbf{y}} - \mathbf{\Sigma}\mathbf{G}\mathbf{y}^*\|^2 \quad \text{subject to } \|\mathbf{x}^*\|_0 \le T. \quad (38)$$

Then the solution of Eq. (36) can be obtained in a way similar to that of the OMP algorithm. Its specific algorithm is shown in Fig. 4. In Eq. (36), the criterion utilized for obtaining $\mathbf{E}$ corresponds to the error caused in the known neighboring areas of Ω by the derived inverse projection. Thus, by minimizing Eq. (36), the optimal estimation of $\mathbf{E}$ can be expected. In this way, we can obtain the non-zero coefficient vector $\hat{\mathbf{p}}$ and the extraction matrices $\mathbf{E}$ and $\hat{\mathbf{E}}$, and restoration of the missing area Ω from the estimation result $\mathbf{y}^*$ becomes feasible.

As described above, we can reconstruct missing textures in the target patch $\hat{f}$. The proposed method clips patches ($w \times h$ pixels) including missing areas in a raster scanning order 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 patches. In such a case, we regard the result minimizing the cost function in Eq. (36) as the final result.

C. Discussion of The Effectiveness

By using the above schemes based on sparse representation, we can adaptively generate subspaces for approximating original and corrupted patches. This enables the derivation of adaptive inverse projection for reconstructing the missing area Ω. In the conventional methods in [18], [24], [25], one or several subspaces are generated and the optimal one is selected for reconstruction. However, the kinds of subspaces utilized for reconstruction are limited to one or a few clusters. Thus, it is difficult to adaptively obtain the optimal subspace for each target texture. On the other hand, the proposed method respectively selects the optimal T signal-atoms from the dictionaries containing K prototype signal-atoms for the original and corrupted patches. Therefore, it can effectively solve the conventional problems.

Furthermore, the conventional methods based on the sparse representation [16], [28]–[30] assume that the optimal sparse representation of target textures and those of the known neighboring areas are the same. Thus, they select optimal signal-atoms for representing target textures from only the known neighboring areas. It should be noted that the proposed method does not rely on the above assumption. In the basic idea of the conventional methods, $\hat{\mathbf{x}}$ is utilized for $\mathbf{x}$, i.e., the relationship of $\mathbf{E}$ and $\mathbf{p}$ becomes the same as that of $\hat{\mathbf{E}}$ and $\hat{\mathbf{p}}$. This tends to be satisfied if the missing area Ω is much smaller than the other known area $\bar{\Omega}$ in f. Otherwise, the assumption tends not to be satisfied. On the other hand, the proposed method regards that $\mathbf{E}$ and $\mathbf{p}$ of the original patch are respectively different from $\hat{\mathbf{E}}$ and $\hat{\mathbf{p}}$ and enables the derivation of $\mathbf{p}$ from $\hat{\mathbf{p}}$ by the inverse projection shown in Eq. (33). In these procedures, $\hat{\mathbf{E}}$ and $\hat{\mathbf{p}}$ (i.e., $\hat{\mathbf{x}}$) are obtained almost in the same way as the conventional methods. However, in order to obtain $\mathbf{E}$ that is necessary for estimating $\mathbf{p}$, the proposed method newly defines the criterion in Eq. (36). Specifically, the optimal signal-atoms for representing the original vector $\mathbf{y}$ are selected by monitoring errors caused by the derived IPVSR. This means the selected signal-atoms minimize the errors caused in the reconstruction process.

In this way, our method enables the adaptive generation of the two subspaces of the selected signal-atoms to derive the IPVSR, and thus this effectively solves the conventional problem which needs the above assumption. Therefore, the most novel approaches in the proposed method, i.e., the biggest differences between our method and the conventional methods are twofold:

1) Our method does not rely on the assumption utilized in the conventional methods, and the calculation of $\mathbf{p}$ is realized by the inverse projection.

2) The new criterion in Eq. (36) is defined for obtaining $\mathbf{E}$ to estimate $\mathbf{p}$ by the inverse projection in 1).

In order to realize the first approach, we introduce the criterion in Eq. (36) (i.e., the second approach). Thus, the key innovation in our method is the introduction of Eq. (36). From the above discussion, we conclude that the proposed method can provide subspaces for representing original patches more accurately than the conventional methods. This is also confirmed in the experiments shown in the following section.

IV. EXPERIMENTAL RESULTS

In this section, we verify the performance of the proposed method. First, in IV-A, we show the conventional methods, which are utilized for the comparisons of the proposed method, and the experimental conditions. Next, in IV-B, the results of reconstruction based on the proposed method and the conventional methods are shown, and the effectiveness of the proposed method is discussed. Finally, in IV-C, the performance limitation of the proposed method is discussed.

A. Experimental Conditions

In the subjective evaluation, we utilized the five conventional methods [13], [18], [28], [14], [16]. The method in [13] is one of the most influential works in the field of exemplar-based texture reconstruction, and we utilized this method for comparison with the proposed method. Furthermore, the methods in [14], [16] are proposed as improved methods of [13], and we regard them as the state-of-the-art approaches in this study field. The method in [18] is a representative method that reconstructs missing areas by using the eigenspace obtained by PCA. Furthermore, by introducing kernel methods into the PCA-based approach, nonlinear subspaces can capture the nonlinear texture features [23]. Its improvement has been achieved in [24], [25] by novel texture classification approaches. Since it has been shown in [25] that its performance is superior to that of [24], we utilized the conventional methods in [23] and [25] as comparative methods in quantitative evaluation shown later. In recent studies, sparse representation has also been applied to the reconstruction of missing areas. Thus, in this experiment, we utilized the conventional method in [28], which is a representative method using sparse representation, for comparison with the proposed method. Furthermore, since the exemplar-based method in [16] uses the sparse representation to fill missing areas, it also becomes the comparison based on the sparse representation. In this way, we used the methods in [13], [18], [28], [14], [16] for comparison with the proposed method in subjective evaluation. It should be noted that in the experiments, we implemented both the proposed method and the conventional methods by using Matlab, i.e., all of the conventional methods in [13], [18], [23], [25], [28], [14], [16] were based on our own implementation.

Next, we discuss the conditions of the experiments. The parameters of the proposed method were set as $w = 30$, $h = 30$, $K = 3700$, $T = 40$ and the clipping intervals of the patches were 10 in width and 8 in height. Furthermore, we performed the reconstruction by the conventional methods in [13], [18], [23], [25], [28], [14], [16] in the following conditions:

- Conventional method in [28]:
  This method is a representative approach using the sparse representation, which is similar to the proposed method, and it can be seen that our method is its improved version. Thus, in order to directly compare the performance between the proposed method and this method, we set the size of patches, the number of training patches, and the order of the reconstruction to the same values as those of the proposed method, i.e., the same conditions were used.
- Conventional methods in [18], [23], [25]:
  These methods use eigenspaces and subspaces obtained by kernel canonical correlation analysis to calculate inverse projection for reconstructing missing areas, and their schemes are similar to the proposed method. Therefore, we performed the reconstruction by using these conventional methods in the same conditions as those of the proposed method.
- Conventional methods in [13], [14], [16]
  These methods adopt the exemplar-based inpainting approaches, and they need many training patches. Therefore, the clipping intervals of training patches were set to 2 in width and 2 in height. Then the number of the obtained patches is much larger than that of the proposed method. Next, the size of patches was set to 15×15 pixels (half size of the proposed method). In [13], [14], [16], the patch sizes were set to smaller values than those in this experiment. If the proposed method and the conventional methods in [18], [23], [25], [28] use smaller patches, the expression abilities of textures become much better, and the reconstruction performance is also improved. However, in order to clearly show the difference of the performance between these methods, we used larger size patches. It should be noted that if the size of patches is set to the same value as that of our method, it becomes difficult for the exemplar-based reconstruction methods to find the optimal training patches suitable for the reconstruction of missing areas. Therefore, the half size patches were used in this experiment. Furthermore, since the order of reconstruction is the most important point in these methods, it is determined in the same way as each method.

In this way, we used the same conditions as those of our method in [18], [23], [25], and [28]. In [13], [14], and [16], their conditions are better than those of the other methods.

B. Discussions of Experimental Results

We utilized a test texture image (480×360 pixels, 24-bit color levels) as shown in Fig. 5(a) and added missing areas to this image for obtaining its corrupted image shown in Fig. 5(b). In Fig. 5(b), the text regions "Chain of Mountain" correspond to missing areas (Fig. 5(c))[2]. For the target test image, we applied the proposed method and obtained its reconstruction result as shown in Fig. 5(d). Furthermore, Figs. 5(e)–(i) respectively show results obtained by the conventional

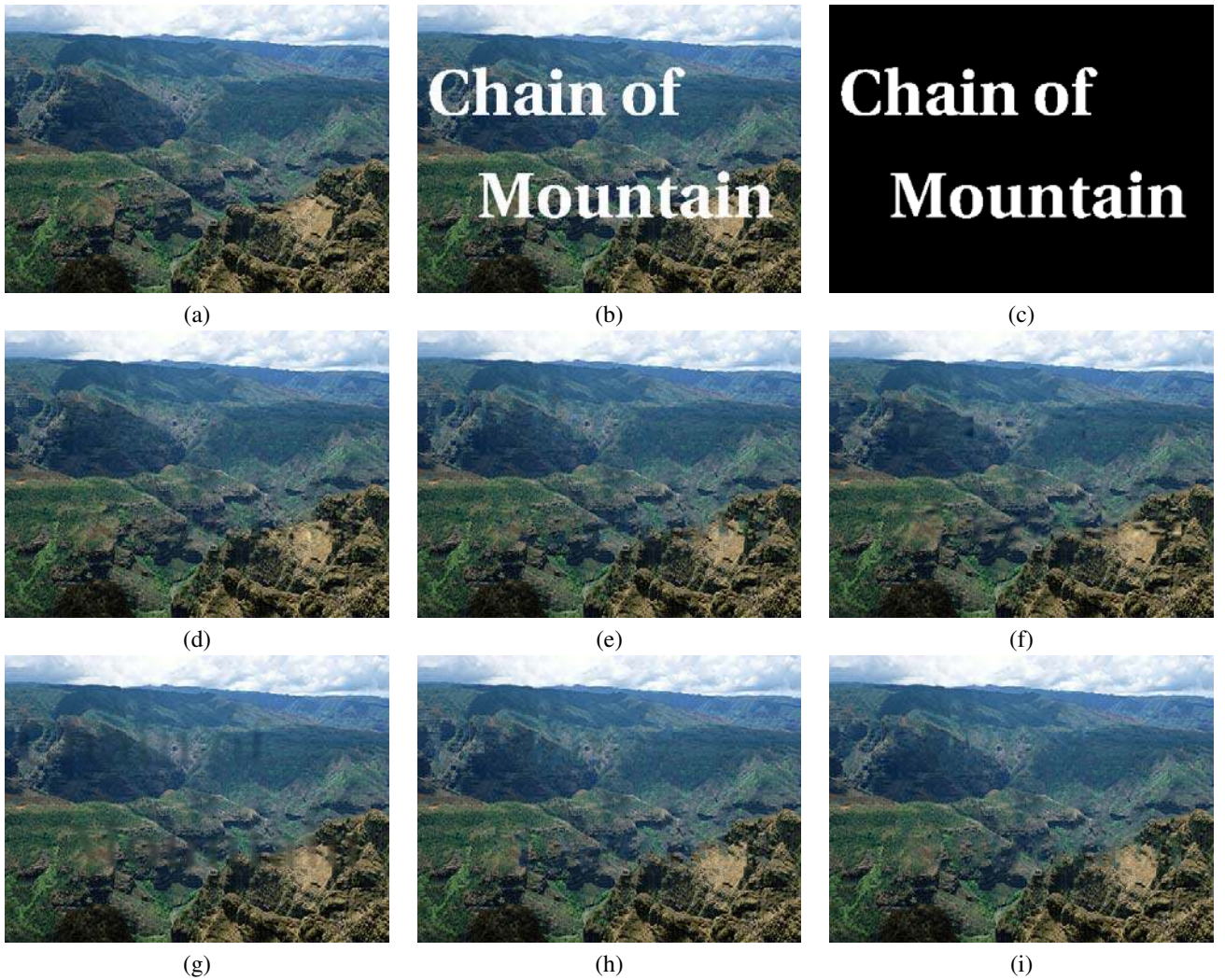[2]Note that in this experiment, positions of the missing areas are previously provided.

Fig. 5. (a) Original image (480×360 pixels, 24-bit color levels), (b) Corrupted image including text regions (8.9% loss), (c) Flag image whose white regions correspond to missing areas, (d) Reconstructed image obtained by the proposed method, (e) Reconstructed image obtained by the conventional method in [13], (f) Reconstructed image obtained by the conventional method in [18], (g) Reconstructed image obtained by the conventional method in [28], (h) Reconstructed image obtained by the conventional method in [14], (i) Reconstructed image obtained by the conventional method in [16].

methods in [13], [18], [28], [14], [16]. For better subjective evaluation, enlarged portions around the upper left of each image are shown in Fig. 6. It can be seen that the use of the proposed method has achieved noticeable improvements compared to the conventional methods. Different experimental results are shown in Figs. 7–12. Compared to the results obtained by the conventional methods, it can be seen that various kinds of textures are accurately restored by using the proposed method. Therefore, high performance of the proposed method was verified by the experiments.

In order to quantitatively evaluate the performance of the proposed method, we show the PSNR and the SSIM index [41] of the reconstruction results in Tables I and II. The SSIM index is one of the representative measures in the field of image quality assessment. It can be seen that our method has achieved improvement over the conventional methods in the SSIM index. Although the PSNR of the proposed method tends to become worse than those of the conventional methods,

we can see that the results for the PSNR cannot correctly reflect the visual quality in subjective evaluation. Here, we discuss the difference of the performance between the subject evaluation and MSE and PSNR-wise quantitative evaluation. As shown in [41]–[43], the MSE, PSNR and their variants cannot generally reflect perceptual distortions, and their values become higher for images altered with some distortions such as mean luminance shift, contrast stretch, spatial shift, spatial scaling, and rotation, yet negligible loss of subjective image quality. Furthermore, blurring severely deteriorates the image quality, but their values become lower than those of the above alternations. It is well known that most images contain more low-frequency components than high-frequency components. Thus, the MSE and PSNR-based criteria focus on these low-frequency components and output lower values even if target images severely suffer from some blurring artifacts. The results obtained by several conventional methods suffer from such artifacts, and their performance seems to be the worst,
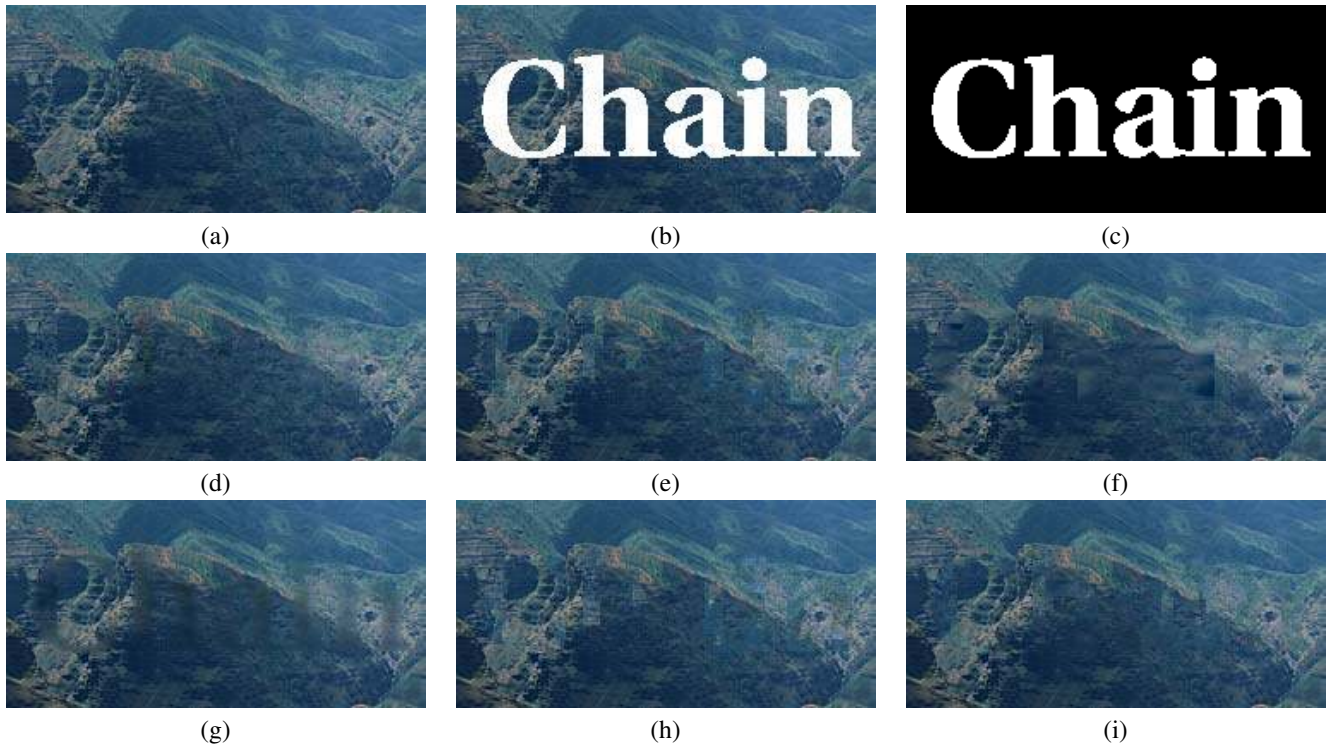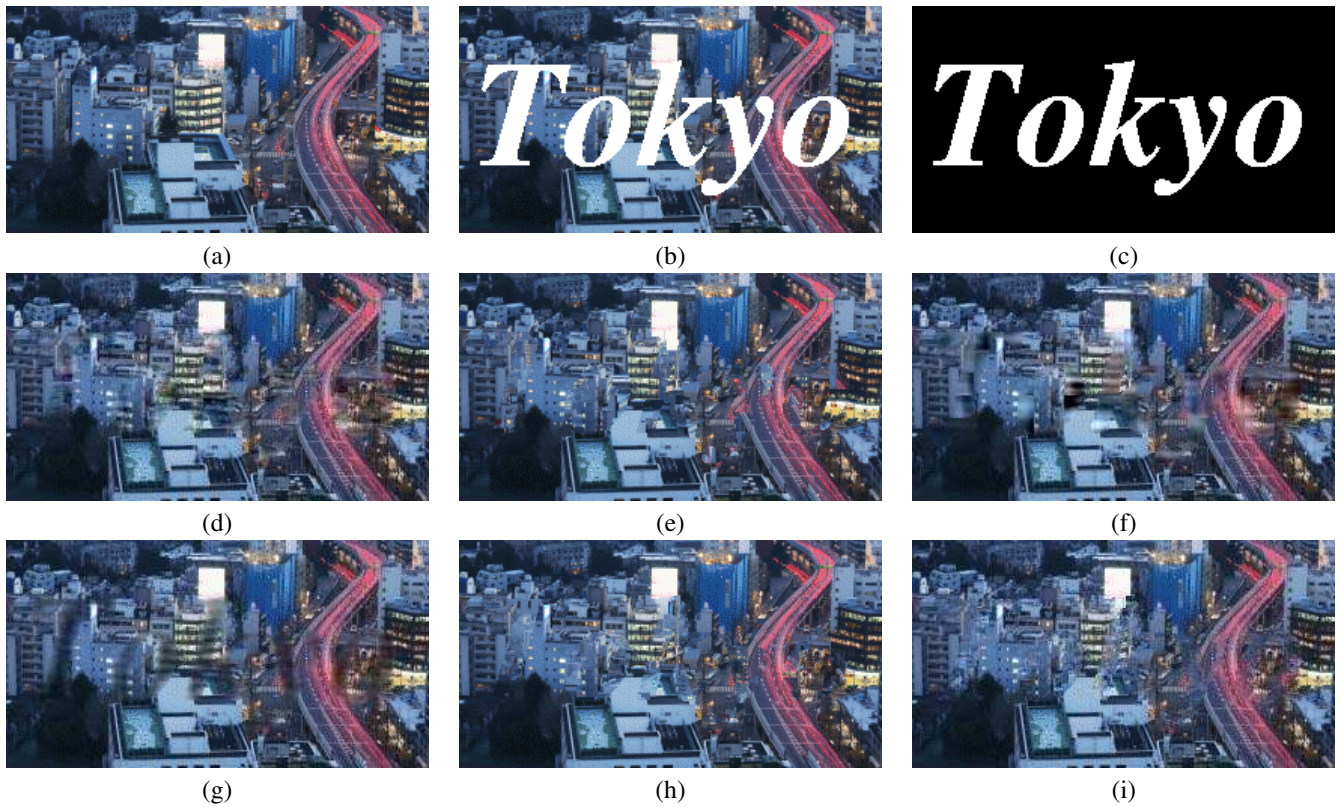
Fig. 6. (a) Zoomed portion of Fig. 5(a), (b) Zoomed portion of Fig. 5(b), (c) Zoomed portion of Fig. 5(c), (d) Zoomed portion of Fig. 5(d), (e) Zoomed portion of Fig. 5(e), (f) Zoomed portion of Fig. 5(f), (g) Zoomed portion of Fig. 5(g), (h) Zoomed portion of Fig. 5(h), (i) Zoomed portion of Fig. 5(i).

TABLE I
PERFORMANCE COMPARISON (PSNR) OF THE PROPOSED METHOD AND CONVENTIONAL METHODS.

Method	Figure 5	Figure 7	Figure 9	Figure 11
Reference [13]	28.76	20.94	24.63	26.35
Reference [18]	29.56	21.86	25.88	27.69
Reference [23]	29.98	22.20	26.23	27.39
Reference [25]	**30.96**	**23.33**	**27.00**	**28.06**
Reference [28]	30.39	22.43	25.85	27.28
Reference [14]	28.56	21.10	24.86	25.74
Reference [16]	29.54	21.18	25.76	27.01
Proposed method	29.27	22.15	26.12	26.93

TABLE II
PERFORMANCE COMPARISON (SSIM) OF THE PROPOSED METHOD AND CONVENTIONAL METHODS.

Method	Figure 5	Figure 7	Figure 9	Figure 11
Reference [13]	0.9539	0.9096	0.9292	0.9265
Reference [18]	0.9318	0.9042	0.9134	0.9137
Reference [23]	0.9309	0.8959	0.9100	0.9055
Reference [25]	0.9386	0.9114	0.9201	0.9138
Reference [28]	0.9384	0.9063	0.9202	0.9119
Reference [14]	0.9520	0.9098	0.9312	0.9204
Reference [16]	0.9579	0.9196	0.9369	0.9345
Proposed method	**0.9583**	**0.9208**	**0.9438**	**0.9364**

but the PSNR-wise quantitative evaluation does not reflect suitable performance. On the other hand, the SSIM index can represent the visual quality more accurately. This has also been pointed out by many researchers [41]–[43]. Therefore, we can conclude that the use of the SSIM index as a visual quality measure is appropriate for texture reconstruction.

We discuss the difference between the proposed method and the conventional methods. The conventional methods in [18], [23]–[25] generate one or several subspaces for reconstructing missing areas. In these methods, the number of subspaces available for reconstructing each texture is maximally limited to the number of clusters obtained by previously performing clustering of known local textures. This means these methods cannot generate the optimal subspace for each texture including missing areas. Therefore, by using sparse representation, the optimal subspace can be provided for each texture. However, it should be noted that the selection of the optimal signal-atoms is important for obtaining the subspaces, and its correctness affects the reconstruction performance. The conventional method in [28] based on sparse representation performs selection of the optimal signal-atoms from only the known neighboring areas. Furthermore, it needs the assumption that optimal sparse representation of target textures and those of the known neighboring areas are the same. This problem also occurs in [16] since they utilize similar schemes. However, since the patch size is smaller than that of the proposed method and the method in [28], it is less represented. On the other hand, the proposed method enables adaptive generation of the two subspaces for the target texture including missing areas and its corresponding original texture. In addition, since our method does not need the assumption by deriving the IPVSR based on the obtained subspaces, it effectively solves the problems of the conventional methods. In order to justify the above discussion, we show some experimental results. In this

Fig. 7.   (a) Original image (480 × 360 pixels, 24-bit color levels), (b) Corrupted image including text regions (11.3% loss), (c) Flag image whose white regions correspond to missing areas, (d) Reconstructed image obtained by the proposed method, (e) Reconstructed image obtained by the conventional method in [13], (f) Reconstructed image obtained by the conventional method in [18], (g) Reconstructed image obtained by the conventional method in [28], (h) Reconstructed image obtained by the conventional method in [14], (i) Reconstructed image obtained by the conventional method in [16].

experiment, we focus on similarities of subspaces to measure the accuracy of the signal-atoms selected by the proposed method and the conventional methods. Even when the selected signal-atoms are different from those obtained by using original patches, missing areas can be accurately reconstructed if the subspace of the selected signal-atoms is sufficiently similar to the subspace spanned by the signal-atoms obtained from original patches. Therefore, since the subspaces spanned by the signal-atoms are the most important factors which affect the accuracy of the reconstruction, we show their similarities for measuring the accuracy of the selected signal-atoms. In this experiment, we focus on the smallest canonical angle which is often used for Mutual Subspace Method (MSM) [44], [45]. Given two subspaces, one is the subspace spanned by the signal-atoms selected by our method or the conventional methods, and the other is the subspace spanned by the signal-atoms obtained from original patches[3], these T-dimensional subspaces are respectively denoted as L^{est} and L^{org}, where L^{est} corresponds to L of Fig. 2 in our method. In MSM, the similarity between the two subspaces L^{est} and L^{org} is based on their minimum canonical angle θ_1, and it is defined as

$$\cos^2 \theta_1 = \max_{\mathbf{u} \in L^{est}, \mathbf{v} \in L^{org}, \|\mathbf{u}\| \neq 0, \|\mathbf{v}\| \neq 0} \frac{(\mathbf{u}'\mathbf{v})^2}{\|\mathbf{u}\|^2 \|\mathbf{v}\|^2}. \tag{39}$$

Note that T canonical angles can be defined between T-dimensional subspaces, and θ_1 in the above equation corresponds to the minimum one, i.e., $\cos^2 \theta_1$ in the above equation represents the maximum similarity. Furthermore, $\cos^2 \theta_1$ in Eq. (39) is obtained as the maximum eigenvalue of $\mathbf{\Phi}^{est}\mathbf{\Phi}^{org}\mathbf{\Phi}^{est}$ or $\mathbf{\Phi}^{org}\mathbf{\Phi}^{est}\mathbf{\Phi}^{org}$, where

$$\mathbf{\Phi}^{est} = \sum_{m=1}^{M} \boldsymbol{\phi}_m^{est} \boldsymbol{\phi}_m^{est\prime}, \tag{40}$$

[3]The signal-atoms suitable for original patches are obtained by using OMP algorithm.

Fig. 8.   (a) Zoomed portion of Fig. 7(a), (b) Zoomed portion of Fig. 7(b), (c) Zoomed portion of Fig. 7(c), (d) Zoomed portion of Fig. 7(d), (e) Zoomed portion of Fig. 7(e), (f) Zoomed portion of Fig. 7(f), (g) Zoomed portion of Fig. 7(g), (h) Zoomed portion of Fig. 7(h), (i) Zoomed portion of Fig. 7(i).

$$\boldsymbol{\Phi}^{\text{org}} = \sum_{m=1}^{M} \boldsymbol{\phi}_m^{\text{org}} \boldsymbol{\phi}_m^{\text{org}\prime}. \qquad (41)$$

In the above equation, $\boldsymbol{\phi}_m^{\text{est}}$ and $\boldsymbol{\phi}_m^{\text{org}}$ ($m = 1, 2, \cdots, M$) are respectively orthonormal bases of the subspaces L^{est} and L^{org}. We calculate M eigenvectors of the subspaces L^{est} and L^{org}, whose eigenvalues are larger than the other ones, and use them as $\boldsymbol{\phi}_m^{\text{est}}$ and $\boldsymbol{\phi}_m^{\text{org}}$, respectively. Then we change the value of M and show the average values of the results in Eq. (39), which are calculated from the subspace L^{est} obtained by the proposed method and the conventional method [28], in Tables III and IV. In this experiment, the similarities in Eq. (39) are calculated for patches whose missing areas are more than 50% within patches in order to clearly show the difference between the two methods. In addition, the clipping interval was set to 8 in both width and height for calculating the average values from many examples. From the obtained results, it can be seen that the proposed method has higher similarities compared to the conventional method. Thus, we can conclude the proposed method provides more accurate subspaces than

TABLE III
SIMILARITIES IN EQ. (39) OBTAINED FROM SUBSPACE L^{EST} IN THE PROPOSED METHOD.

Test image	$M = 5$	$M = 10$	$M = 15$	$M = 20$	$M = 25$
Figure 5	0.4830	0.6465	0.7437	0.8120	0.8996
Figure 7	0.6026	0.7411	0.8113	0.8622	0.9269
Figure 9	0.4175	0.6371	0.7594	0.8517	0.9062
Figure 11	0.5515	0.6939	0.7772	0.8467	0.8971

the conventional methods. Furthermore, Figs. 13–16 show the reconstruction results whose similarities in Eq. (39) are quite different between the two methods compared to the other areas, where $M = 20$. Note that the patch size is 30×30 pixels, and it is difficult to see the difference between such small patches. Therefore, these figures show areas of size 90×90 pixels which contain target patches (30×30 pixels). From the obtained results, the difference between the two methods is significant, and the effectiveness of our method can be also confirmed.

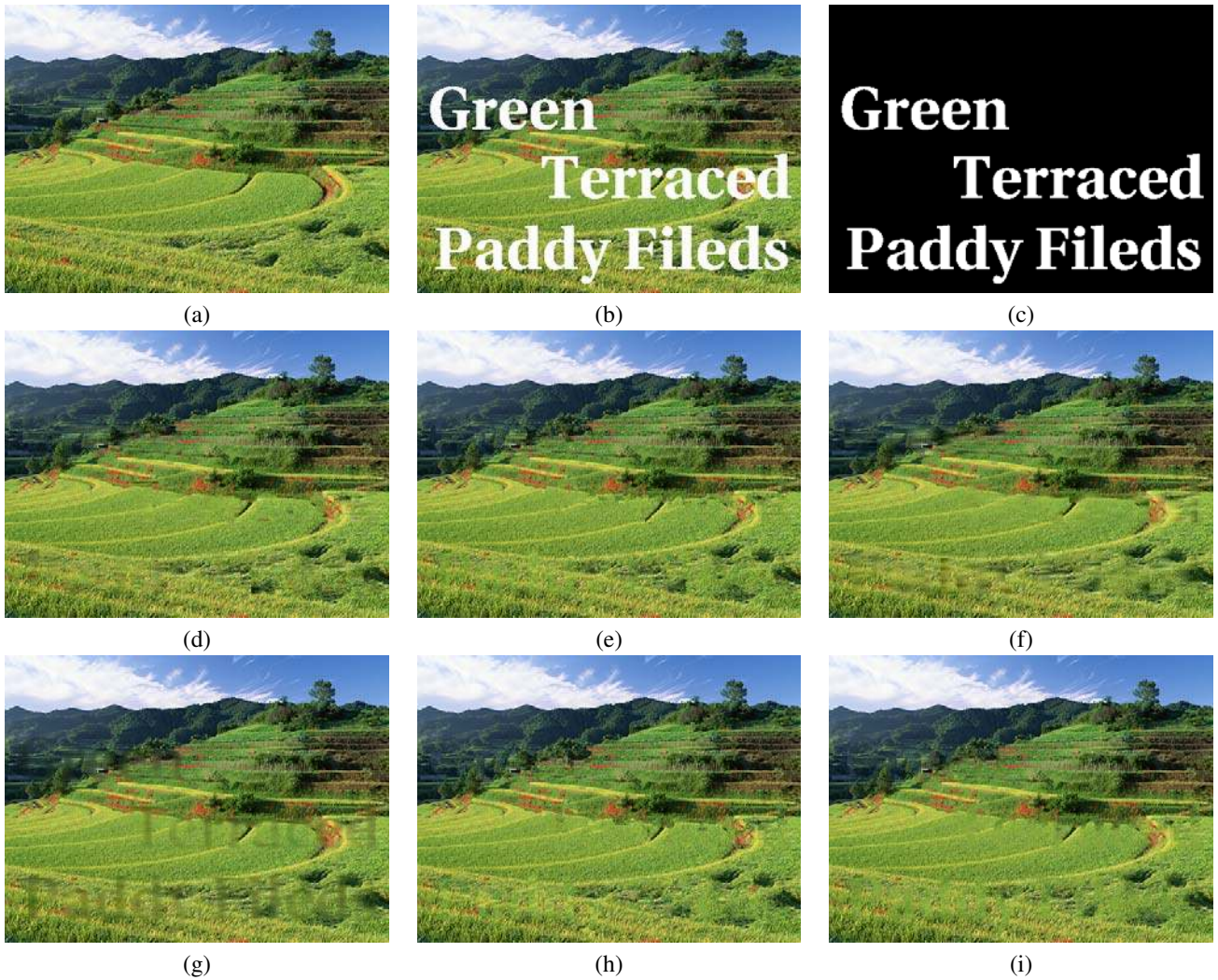Finally, in Figs. 17 and 18, we show some examples of

Fig. 9. (a) Original image (480 × 360 pixels, 24-bit color levels), (b) Corrupted image including text regions (10.7% loss), (c) Flag image whose white regions correspond to missing areas, (d) Reconstructed image obtained by the proposed method, (e) Reconstructed image obtained by the conventional method in [13], (f) Reconstructed image obtained by the conventional method in [18], (g) Reconstructed image obtained by the conventional method in [28], (h) Reconstructed image obtained by the conventional method in [14], (i) Reconstructed image obtained by the conventional method in [16].

TABLE IV
SIMILARITIES IN EQ. (39) OBTAINED FROM SUBSPACE L^{EST} IN THE CONVENTIONAL
METHOD [28].

Test image	$M = 5$	$M = 10$	$M = 15$	$M = 20$	$M = 25$
Figure 5	0.4373	0.5772	0.6931	0.7713	0.8816
Figure 7	0.4742	0.5941	0.6923	0.7707	0.8746
Figure 9	0.3882	0.5386	0.6646	0.7749	0.8704
Figure 11	0.4245	0.5656	0.6763	0.7784	0.8627

restoration by the proposed method for images including text regions in the whole areas. When missing areas exist all over the target image, we cannot obtain sufficient training patterns, and accurate sparse representation becomes difficult. In such a case, the proposed method again uses the results of reconstruction obtained by using our method for calculating the dictionaries and iterates the reconstruction procedures. From these figures, we can see that the proposed method

restores several kinds of missing areas, and many applications such as removal of unnecessary objects and error concealment can be expected.

C. Performance Limitation

In this subsection, we discuss the performance and its limitation of the proposed method. First, we focus on the percentage of missing areas within target images. In the proposed method, the number of training examples f_i ($i = 1, 2, \cdots, N$) must be lager than the number of the signal atoms $\mathbf{d}_j$ and $\hat{\mathbf{d}}_j$ ($j = 1, 2, \cdots, K$). This means $N > K$ must be satisfied for constructing dictionaries from the known patches f_i. Thus, although the percentages of missing areas within target images are large, the proposed method can perform the calculation of the dictionary matrices (Sec. III-A) when $N > K$ is satisfied. It should be noted that even if we cannot obtain sufficient training patterns, we can reconstruct missing areas by using some alternative schemes, and their examples are shown in

Fig. 10. (a) Zoomed portion of Fig. 9(a), (b) Zoomed portion of Fig. 9(b), (c) Zoomed portion of Fig. 9(c), (d) Zoomed portion of Fig. 9(d), (e) Zoomed portion of Fig. 9(e), (f) Zoomed portion of Fig. 9(f), (g) Zoomed portion of Fig. 9(g), (h) Zoomed portion of Fig. 9(h), (i) Zoomed portion of Fig. 9(i).

Figs. 17 and 18. In order to perform accurate reconstruction, we should calculate the dictionaries from various kinds of textures. Therefore, it is desirable that known patches be obtained from various texture parts to guarantee the performance of our method. Next, we focus on the number of known pixels within the target patch f. For calculating the inverse projection shown in Eq. (33), $T \leq \text{rank}\{\Sigma\}$ must be satisfied for obtaining the pseudo inverse matrix. Note that the rank of the matrix Σ, i.e., rank$\{\Sigma\}$, corresponds to the number of the known pixels in $\bar{\Omega}$. Therefore, in order to perform the reconstruction based on IPVSR, the number of the known pixels within the target patch f must be larger than T. Otherwise, it becomes difficult to directly calculate the inverse projection in Eq. (33), and some regularization terms must be introduced into the calculation of $\mathbf{p}^*$. In the experiments of this section, if the above condition is not satisfied, we skip the target patch and reconstruct its missing area from the other patches. In the proposed method, the clipping interval is smaller than the size of patches. Then each reconstructed pixel has multiple estimation results, and thus we can reconstruct all missing areas within target images even if some patches do not satisfy the above condition. Generally, when the size of missing areas becomes larger, the above condition tends not to be satisfied. Thus, we have to set the size of patches ($w \times h$ pixels) to larger values or set T to a smaller value. In the former case, rank$\{\Sigma\}$ tends to become larger. However, in both cases, the representation abilities of textures become lower, and it leads the degradation of the reconstruction performance. Finally, the clipping interval of

patches for reconstructing missing areas must be equal to or smaller than the patch size as described above. Otherwise, all missing intensities may not be reconstructed. In the proposed method, the clipping interval is smaller than the size of patches in order to address the problem of the above condition not being satisfied.

As described above, the number of training examples utilized for the proposed method is much smaller than those of the conventional methods [13], [14], [16]. Therefore, there is a limitation for accurately reconstructing missing textures by using the proposed method. This means by increasing the number of the training examples, our method can achieve more accurate reconstruction. It should be noted that the computation time of the proposed method is still larger than other methods, and thus the number of the training examples should not be increased for saving the computation time. The discussion of the computation time of our method is shown later. Furthermore, if we adopt several schemes of these conventional methods, such as the determination of the order (priority) of reconstruction pixels, the performance of our method becomes better. The current method only clips patches including missing areas in a raster scanning order and reconstructs them based on IPVSR. Therefore, in order to further improve the reconstruction performance, we should introduce the schemes used in the conventional exemplar-based methods into our method. However, since this is not the main focus of our paper, this will be addressed in the future work.

Fig. 11.   (a) Original image (480×360 pixels, 24-bit color levels), (b) Corrupted image including text regions (11.9% loss), (c) Flag image whose white regions correspond to missing areas, (d) Reconstructed image obtained by the proposed method, (e) Reconstructed image obtained by the conventional method in [13], (f) Reconstructed image obtained by the conventional method in [18], (g) Reconstructed image obtained by the conventional method in [28], (h) Reconstructed image obtained by the conventional method in [14], (i) Reconstructed image obtained by the conventional method in [16].

Finally, from the above discussion, we show some results whose missing areas cannot be successfully reconstructed by the proposed method in Fig. 19. In this example, the size of the missing areas within patches becomes larger, and thus it becomes difficult for the proposed method to reconstruct those missing areas, accurately. Thus, it is also difficult to perfectly avoid spurious color even if the matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$ are adopted. Furthermore, in this example, the target object contains many structural components, i.e., edges. Since the proposed method only focuses on the reconstruction of missing textures, the successful restoration of those components cannot be realized. In order to solve these problems, we have to introduce some approaches which consider the accurate reconstruction of structural components into our method. Some conventional methods [13], [14], [16] determine the priorities for reconstructing corrupted patches to correctly restore both of structural and texture components, and this scheme will be easily implemented in the proposed method. Furthermore,

Bertalmio et al. [19] adopt an approach that determines whether target areas are structural parts or texture parts and adaptively select suitable reconstruction methods. Then this also becomes the solution for solving the above problem.

In addition, we discuss the computation cost of the proposed method. The experiments shown in Figs. 5–11 were performed on a personal computer using Intel(R) Core(TM) i7 950 CPU 3.06 GHz with 8.0 Ggytes RAM. The proposed method was implemented by using Matlab. The average computation time to perform our algorithm for reconstructing missing areas is 9.86×10^2 sec. From the obtained result, we can see it is necessary to reduce the computational complexity of the proposed method for its practical use.

Here, we show some rough estimation of the computational complexity of the proposed method. In III-A, the proposed method performs the calculation of the dictionary matrices $\mathbf{D}$ and $\hat{\mathbf{D}}$. First, we focus on the calculation of $\mathbf{D}$ based on the K-SVD algorithm which is composed of two parts: sparse
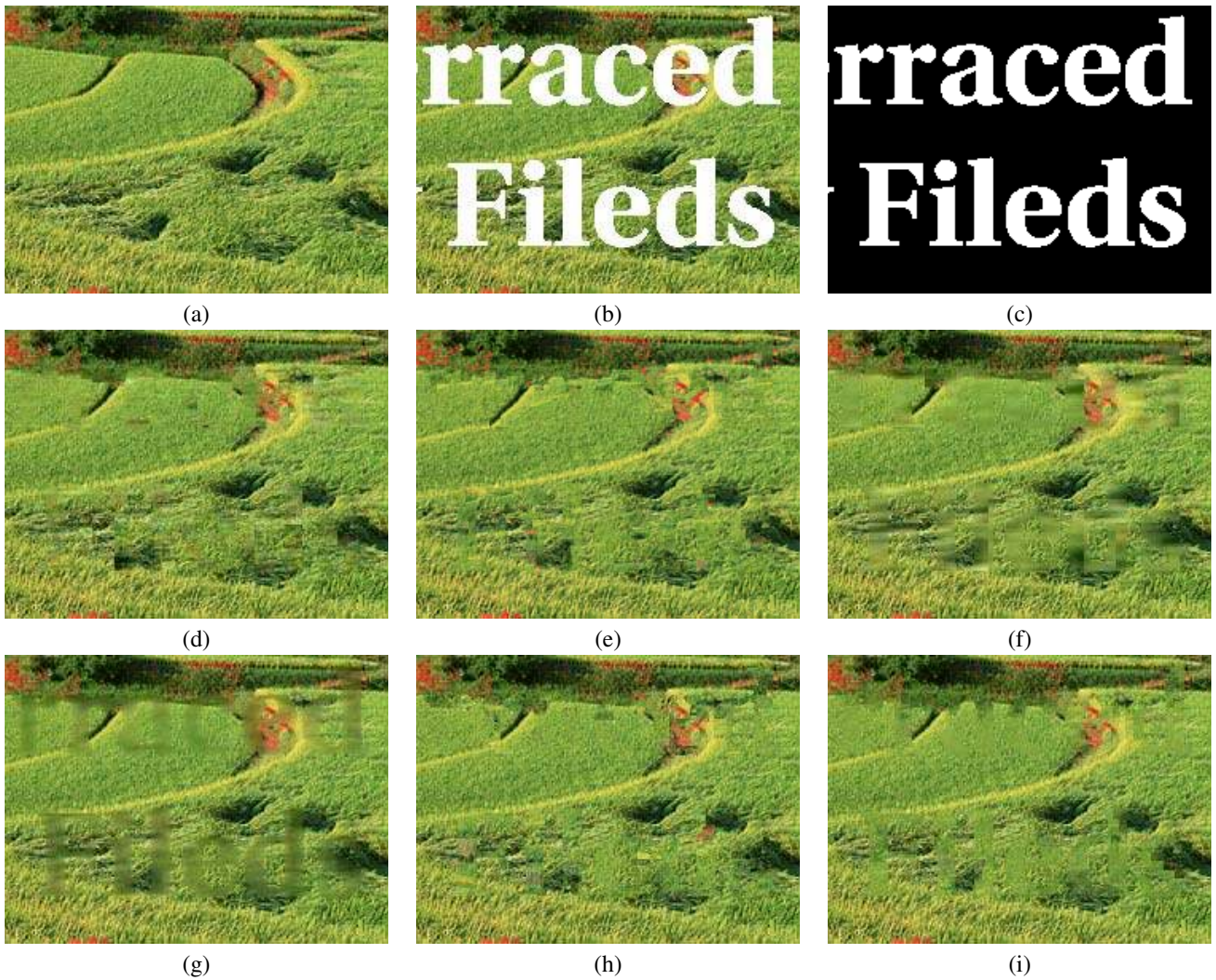
Fig. 12. (a) Zoomed portion of Fig. 11(a), (b) Zoomed portion of Fig. 11(b), (c) Zoomed portion of Fig. 11(c), (d) Zoomed portion of Fig. 11(d), (e) Zoomed portion of Fig. 11(e), (f) Zoomed portion of Fig. 11(f), (g) Zoomed portion of Fig. 11(g), (h) Zoomed portion of Fig. 11(h), (i) Zoomed portion of Fig. 11(i).

coding part and dictionary renewal part. The sparse coding part is dominated by the selection of the best matched signal-atoms and the calculation of the optimal sparse representation for each $\mathbf{x}_i$ ($i = 1, 2, \cdots, N$), and their time complexities are respectively $O(K\zeta T)$ and $O(k\zeta)$. Note that $\zeta = 3wh$, and k ($k = 1, 2, \cdots, T$) means that the selection of k-th optimal signal atom is performed. Furthermore, the dictionary renewal part is dominated by the singular value decomposition, and this procedure is performed for each signal-atom $\mathbf{d}_j$ ($j = 1, 2, \cdots, K$). Thus, its complexity is $O(\zeta \xi_j^2)$, where ξ_j represents the number of $\mathbf{x}_i$ which use the signal-atom $\mathbf{d}_j$ for its sparse representation. In addition, the complexity of the calculation of the dictionary matrix $\hat{\mathbf{D}}$ is $O(\zeta K)$. Next, the proposed method performs the reconstruction of the missing area Ω within the target patch f in III-B, and it is mainly dominated by the calculation of $\hat{\mathbf{E}}$ and $\hat{\mathbf{p}}$ and that of $\mathbf{E}$. These procedures are almost the same as the sparse coding part, i.e., they are composed of the selection of the best matched signal-atoms ($O(K\zeta T)$) and the calculation of the optimal sparse representation ($O(k\zeta)$).

Note that in the calculation of the dictionary matrix $\mathbf{D}$, the

sparse coding part and the dictionary renewal part have high computational complexities, and they are also iterated. Figure 20 shows the mean square error converged in the calculation of the dictionary matrix $\mathbf{D}$. From this figure, we can see the error tends to gradually reduce as the number of the iteration increases. However, in the proposed method, the number of the iteration is set to 20 in order to reduce the computation cost. Thus, by reducing the number of the iteration, the further reduction of the computation time can be expected although some performance degradation may occur.

In recent methods such as [14] and [16], they reported that their methods could perform faster reconstruction. Particularly, in [14], they introduce a DCT-domain decomposition scheme, a local gradient-based algorithm, and parallel implementation on GPU, and it achieves quite fast performance compared to the other conventional methods and our method. Therefore, in order to realize the practical use of the proposed method, we have to introduce some techniques for reducing the computation cost. This will be addressed in the future work.

Original patch					
Corrupted patch					
Reconstructed patch by our method					
Similarity in Eq. (37)	0.8600	0.7793	0.9427	0.7557	0.7783
Reconstructed patch by reference [28]					
Similarity in Eq. (37)	0.5780	0.5027	0.6929	0.5173	0.5673
Difference between two similarities	0.2820	0.2766	0.2498	0.2404	0.2110

Fig. 13. Reconstruction results whose similarities in Eq. (39) are quite different between the proposed method and the conventional method [28]. The target image is shown in Fig. 5.

Original patch					
Corrupted patch					
Reconstructed patch by our method					
Similarity in Eq. (37)	0.8500	0.8039	0.7509	0.9374	0.8591
Reconstructed patch by reference [28]					
Similarity in Eq. (37)	0.5555	0.5128	0.4683	0.6642	0.6279
Difference between two similarities	0.2945	0.2911	0.2826	0.2732	0.2312

Fig. 14. Reconstruction results whose similarities in Eq. (39) are quite different between the proposed method and the conventional method [28]. The target image is shown in Fig. 7.

V. CONCLUSIONS

In this paper, we have proposed a missing image data reconstruction method based on an adaptive inverse projection via sparse representation. In order to calculate the inverse projection for reconstructing missing areas, the proposed method performs sparse representation of target textures including missing areas and their corresponding original textures and enables the generation of subspaces. In this approach, the error caused in the known neighboring textures by the inverse projection is newly introduced as a new criterion, and adaptive generation of subspaces becomes feasible. Then, by using the obtained subspaces, the proposed method can apply the adaptive inverse projection via sparse representation to the reconstruction of target textures. The proposed method also introduces some schemes for color processing into the selection of optimal signal-atoms for obtaining subspaces. Consequently, since the optimal subspaces providing the inverse projection can be

Original patch					
Corrupted patch					
Reconstructed patch by our method — Similarity in Eq. (37)	0.9059	0.8476	0.9514	0.8619	0.7816
Reconstructed patch by reference [28] — Similarity in Eq. (37)	0.6295	0.6099	0.7176	0.6283	0.5727
Difference between two similarities	0.2764	0.2377	0.2338	0.2336	0.2089

Fig. 15. Reconstruction results whose similarities in Eq. (39) are quite different between the proposed method and the conventional method [28]. The target image is shown in Fig. 9.
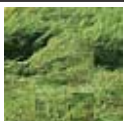
Original patch					
Corrupted patch					
Reconstructed patch by our method — Similarity in Eq. (37)	0.9742	0.8844	0.9478	0.8020	0.8700
Reconstructed patch by reference [28] — Similarity in Eq. (37)	0.6754	0.6064	0.6749	0.5315	0.6051
Difference between two similarities	0.2988	0.2780	0.2729	0.2705	0.2649

Fig. 16. Reconstruction results whose similarities in Eq. (39) are quite different between the proposed method and the conventional method [28]. The target image is shown in Fig. 11.

adaptively obtained, impressive improvements by the proposed method have been achieved.

In this study, we manually set parameters such as the size of patches and the dimension of subspaces. It is desirable that these values can be adaptively determined from the observed images. Thus, we need to complement this determination algorithm. We would like to study these ideas for reconstruction in video data. These topics will be the subject of subsequent reports.
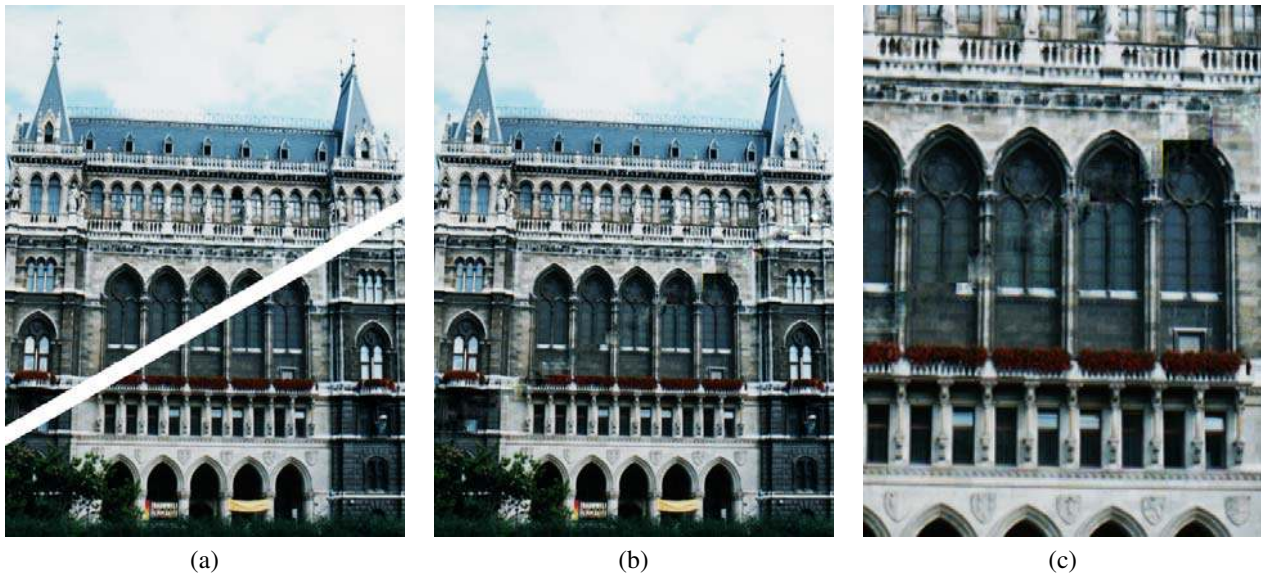
VI. ACKNOWLEDGEMENT

Fig. 19.   Examples which cannot be successfully reconstructed by our method: (a) Target image, (b) Reconstructed image by the proposed method, (c) Zoomed portion of (b).

REFERENCES

[1] H.S. Hou, H.C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. on Acoustics, Speech, Signal Processing* ASSP-26 (6) pp.508–517, 1978.

[2] R.G. Keys, Cubic convolution interpolation for digital image processing, *IEEE Trans. on Acoustics, Speech, Signal Processing*, 29 (6) pp.1153–1160, 1981.

[3] F. Arandiga,R. Donat, P. Mulet, "Adaptive interpolation of images," *Signal Processing*, vol.83, no.2, pp.459-64, 2003.

[4] S. Battiato, G. Gallo, F. Stanco, "A locally adaptive zooming algorithm for digital images," *Image and vision computing*, vol.20, no.11, pp.805–812, 2002.

[5] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, "Filling-In by Joint Interpolation of Vector Fields and Gray Levels," *IEEE Trans. on Image Processing*, vol.10, no.8, pp.1200–1211, 2001.

[6] T.F. Chan and Jianhong Shen, "Nontexture inpainting by curvature-driven diffusions," *Journal of Visual Communication and Image Representation,* vol.12, no.4, pp.436–449, 2001.

[7] A. Rares, M.J.T. Reinders, and J. Biemond, "Edge-based image restoration," *IEEE Trans. on Image Processing,* vol.14, no.10, pp.1454–1468, 2005.

[8] A. A. Efros and T. K. Leung, "Texture synthesis by nonparametric sampling," *IEEE Int. Conf. Computer Vision,* Corfu, Greece, pp.1033–1038, Sept. 1999.

[9] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," *Proceedings of SIGGRAPH 2001, ACM SIGGRAPH*, pp. 341–346, 2001.

[10] L.-W. Wey and M. Levoy, "Fast texture synthesis using tree-structured vector quantization," *in SIGGRAPH 2000, Computer Graphics Proceedings*, K. Akeley, Ed. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, pp. 479–488, 2000.

[11] I. Drori, D. Cohen-Or, and H. Teshurun, "Fragment-based image completion," *in SIGGRAPH 2003: ACM SIGGRAPH 2003 Papers.* New York, USA: ACM Press, pp. 303–312, 2003.

[12] A. Criminisi, P. Perez, and K. Toyama, "Object removal by exemplar-based inpainting," *Proceedings of IEEE Computer Vision and Pattern Recognition*, Jun. 2003.

[13] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. on Image Processing*, vol. 13, no. 9, pp. 1200–1212, 2004.

[14] T. H. Kwok, H. Sheung, and C. C. L. Wang, "Fast query for exemplar-based image completion," *IEEE Transactions on Image Processing*, vol. 19, no. 12, 2010.

[15] I. B. Fidaner, "A survey on variational image inpainting, texture synthesis and image completion," Available online: http://www.scribd.com/doc/3012627/A-Survey-on-Variational-Image-Inpainting-Texture-Synthesis-and-Image-Completion

[16] Z. Xu and J. Sun, "Image Inpainting by Patch Propagation Using Patch Sparsity," *IEEE Transactions on Image Processing*, vol. 19, no. 5, 2010.

[17] A. Kokaram, "A statistical framework for picture reconstruction using 2D AR models," *Image and Vision Computing*, vol.22, no.2, 1, pp.165–171, Feb. 2004.

[18] 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.

[19] M. Bertalmio, L. Vese, G. Sapiro, and S. Ssher, "Simultaneous structure and texture image inpainting," *IEEE Trans. on Image Processing*, vol. 12, no. 8, pp. 882–889, 2003.

[20] S. D. Rane, G. Sapiro, M. Bertalmio, "Structure and texture filling-in of missing image blocks in wireless transmission and compression applications," *IEEE Trans. on Image Processing*, vol. 12, no.3, 2003.

[21] 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. on Neural Networks*, vol.10, no.5, pp.1000–1017, 1999.

[22] 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.

[23] K.I. Kim, M.O. Franz , B. Schölkoph, "Iterative kernel principal component analysis for image modeling," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.27, no.9, 2005.

[24] T. Ogawa, 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.

[25] T. Ogawa, M. Haseyama, "Adaptive missing texture reconstruction method based on kernel canonical correlation analysis with a new clustering scheme," *IEICE Trans. Fundamentals*, vol. E92-A, no. 8, pp. 1950–1960, Aug. 2009.

[26] M. Aharon, M. Elad, A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. on Signal Processing*, vol.54, no.11, 2006.

[27] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. on Image Processing*, vol. 15, no. 12, 2006.

[28] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. on Image Processing*, vol.17, no.1, 2008.

[29] B. Wohlberg, "Inpainting with sparse linear combinations of exemplars," *in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing 2009 (ICASSP 2009)*, pp. 689–692, Apr 2009.

[30] B. Shen, W. Hu, Y. Zhang, and Y.-J. Zhang, "Image inpainting via sparse representation," *in Proceedings of IEEE International Conference*
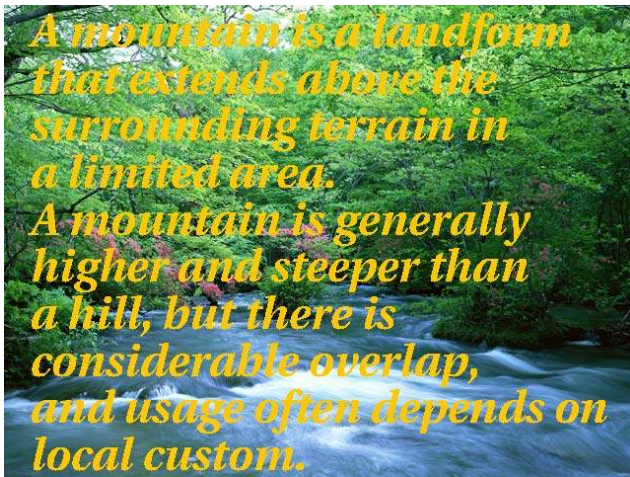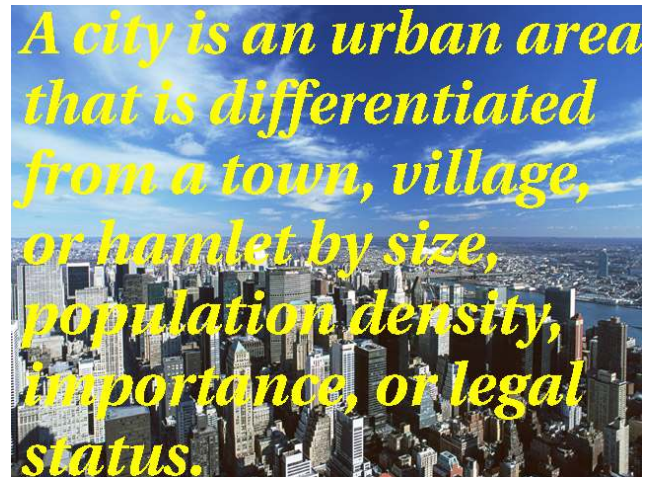
(a)

(b)

Fig. 17.   Example of restoration by the proposed method for an image including text regions in the whole area: (a) Corrupted image (640×480 pixels, 18.5% loss), (b) Image reconstructed by the proposed method.

(a)

(b)

Fig. 18.   Another example of restoration by the proposed method for an image including text regions in the whole area: (a) Corrupted image (640×480 pixels, 17.6% loss), (b) Image reconstructed by the proposed method.

on Acoustics, Speech, and Signal Processing 2009 (ICASSP 2009), pp. 697–700, Apr 2009.

[31] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *J. Construct. Approx.*, vol. 13, pp.57–98, 1997.

[32] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp.3397–3415, 1993.

[33] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their applications to non-linear system identification," *Int. J. Contr.*, vol. 50, no. 5, pp. 1873–1896, 1989.

[34] G. Davis, S. Mallat, and Z, Zhang, "Adaptive time-frequency decompositions," *Opt. Eng.*, vol. 33, no. 7, pp. 2183–2191, 1994.

[35] Y.C. Pati, R. Rezaiifar, and P.S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," *in Conf. Rec. 27th Asilomar Conf. Signals, Syst. Comput.*, vol.1, 1993.

[36] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Info. Theory*, vol. 50, pp. 2231–2242, Oct. 2004.

[37] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Automatic decomposition by basis pursuit," *SIAM Rev.*, vol.43, no. 1, pp.129–159, 2001.

[38] I. F. Gorodnitsky and B.D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted norm minimization algorithm," *IEEE Trans. Signal Process.*, vol.45, pp. 600-616, 1997.

[39] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Trans. Signal Process.*, vol. 47, pp. 187–200, 1999.

[40] B. D. Rao, K. Engan, S. F. Cotter, J. Palmer, and K. Kreutz-Delgado, "Subset selection in noise based on diversity measure minimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 760–770, 2003.

[41] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[42] B. Girod, "What's wrong with mean-squared error?," in *Digital Images and Human Vision*, A.B. Watson, Ed. Cambridge, MA: MIT Press, pp.207–220, 1993.

[43] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. on Image Processing*, novl. 15, no. 11, Nov. 2006.

[44] O. Yamaguchi, K. Fukui, and K. Maeda, "Face recognition using temporal image sequence," *in Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition 1998*, pp.318–323, 1998.

[45] D. Veljkovic, K. A. Robbins, D. Rubino, and N. G. Hatsopoulos, "Extension of mutual subspace method for low dimensional feature projection," *in Proceedings of IEEE International Conference on Image Processing 2007 (ICIP 2007)*, pp. II449–II452, 2007.
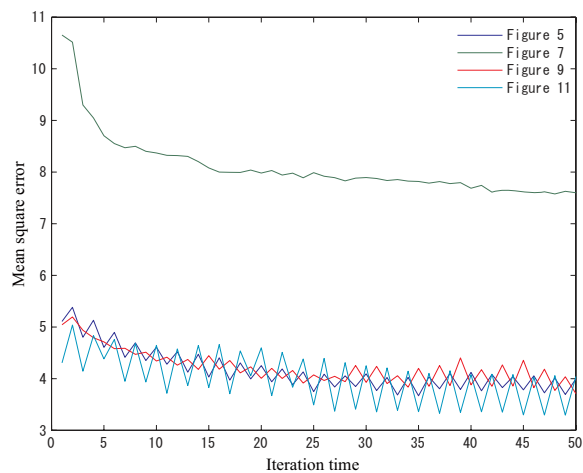
Fig. 20. Relationship between iteration time and mean square error converged in the calculation of the dictionary matrix **D**.

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 digital image processing and its applications. He is a member of the EURASIP, IEEE, IEICE, and Institute of Image Information and Television Engineers (ITE).

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 2005 to 2006. 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 is a member of the IEEE, IEICE, Institute of Image Information and Television Engineers (ITE) and Acoustical Society of Japan (ASJ).