HOKKAIDO UNIVERSITY

Title	Inpainting via Sparse Representation Based on a Phaseless Quality Metric
Author(s)	Ogawa, Takahiro; Maeda, Keisuke; Haseyama, Miki
Citation	IEICE transactions on fundamentals of electronics communications and computer sciences, E103A(12), 1541-1551 https://doi.org/10.1587/transfun.2020SMP0020
Issue Date	2020-12
Doc URL	https://hdl.handle.net/2115/80303
Rights	copyright©2020 IEICE
Type	journal article
File Information	Inpainting via Sparse Representation Based on a Phaseless Quality Metric.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

PAPER Special Section on Smart Multimedia & Communication Systems

Inpainting via Sparse Representation Based on a Phaseless Quality Metric

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

SUMMARY An inpainting method via sparse representation based on a new phaseless quality metric is presented in this paper. Since power spectra, phaseless features, of local regions within images enable more successful representation of their texture characteristics compared to their pixel values, a new quality metric based on these phaseless features is newly derived for image representation. Specifically, the proposed method enables spare representation of target signals, i.e., target patches, including missing intensities by monitoring errors converged by phase retrieval as the novel phaseless quality metric. This is the main contribution of our study. In this approach, the phase retrieval algorithm used in our method has the following two important roles: (1) derivation of the new quality metric that can be derived even for images including missing intensities and (2) conversion of phaseless features, i.e., power spectra, to pixel values, i.e., intensities. Therefore, the above novel approach solves the existing problem of not being able to use better features or better quality metrics for inpainting. Results of experiments showed that the proposed method using sparse representation based on the new phaseless quality metric outperforms previously reported methods that directly use pixel values for inpainting.

key words: *inpainting, phase retrieval, phaseless quality metric, phaseless features*

1. Introduction

Inpainting has attracted much attention since it affords a number of attractive applications. Unnecessary object removal from photographs, error concealment and restoration of old films, i.e., removal of blotches and scratches, are well-known applications. Various approaches for realization of image inpainting including an exemplar-based approach and a multivariate analysis-based approach have been proposed [1], [2]. Efros and Leung proposed texture synthesis, which is the most traditional method of the exemplar-based approach [3]. Their proposed method was improved by Criminisi et al. [4], and the improved method reported by Criminisi et al. became a representative and benchmarking method introducing "patch priority", a new criterion for determining the inpainting order within target images. In recent years, many researchers have improved the above exemplar-based method with extension of image representation performance and sophistication of the determination of patch priorities [5], [6].

Image inpainting methods using multivariate analysis-based approaches have been intensively studied for accelerating the image representation performance of the above-described exemplar-based methods. The simplest difference between a multivariate analysis-based approach and an exemplar-based approach is whether or not several patches or bases are used for approximating target patches including missing regions. Generally, methods based on the exemplar-based approach try to find the best matched patch and directly copies it to missing regions. On the other hand, since methods using the multivariate analysis-based approach approximate patches including missing regions by using several known patches or bases derived by multivariate analysis methods, the image representation performance becomes better. Various multivariate analysis methods including principal component analysis (PCA) series [7]–[9], sparse representation [10]–[15], neighboring embedding [16] and rank minimization [17], [18] have been used for recovering missing regions based on approximation in lower-dimensional subspaces.

Although the methods using the multivariate analysis-based approach try to realize accurate image inpainting by improving the image representation performance, there remains the following fundamental problem. Generally, pixel values are directly utilized as visual features, i.e., texture features of small patches, but these values cannot reflect visual characteristics. Therefore, for example, if two small patches are clipped from the same texture, but their clipping interval is different from the period of the texture, this mismatch leads to a large difference of visual features represented by pixel values between the two small patches. We discuss the above points in more detail below.

The methods using the multivariate analysis-based approach can represent textures including some variations within images more successfully compared to the methods using the exemplar-based approach. It is difficult for the exemplar-based approach to perfectly remove the effect of the mismatch even if it tries to find the best matched patches. Therefore, it is desirable to approximate patches by using several bases spanning low-dimensional subspaces in the multivariate analysis-based approach. However, when pixel values are used as visual features, these features of patches clipped from different positions are generally different from each other even if they are obtained from the same textures. Thus, it becomes difficult for the methods using the multivariate analysis-based approach to approximate patches in the low-dimensional subspace. Furthermore, if several textures are included in a target image, the problem becomes much more difficult.

Manuscript received February 4, 2020.
Manuscript revised May 6, 2020.

[†]The authors are with the Faculty of Information Science and Technology, Hokkaido University, Sapporo-shi, 060-0814 Japan.
[††]The author is with Office of Institutional Research, Hokkaido University, Sapporo-shi, 060-0808 Japan.
a) E-mail: ogawa@lmd.ist.hokudai.ac.jp

DOI: 10.1587/transfun.2020SMP0020

Copyright © 2020 The Institute of Electronics, Information and Communication Engineers

The most reasonable solution to this problem is the use of a new better quality metric obtained from visual features that are more suitable for representing texture characteristics. Although many accurate texture features [19], [20] have been proposed, it is generally difficult to use these features for image inpainting since the final output, i.e., the estimation results, must be pixel values, and it is difficult to convert these texture features to pixel values. In addition, many quality metrics [21]–[23] have recently been proposed, but it is difficult for them to solve the above problem since their functions tend not to be directly derived from pixel values, and acquisition of recovered pixel values therefore becomes difficult when using these metrics.

Traditionally, Fourier-based features have been commonly used for analyzing texture characteristics in images. These features such as Fourier transform magnitudes and power spectra include only phaseless components and provide original visual characteristics that are not affected by spacial gaps. We have also reported that texture characteristics can be successfully represented by using Fourier transform magnitudes and that missing texture reconstruction can be achieved on the basis of phase retrieval [24], [25]. Studies on phase retrieval have been carried out [26]–[28], and these methods enable reconstruction of signals from their phaseless components. Therefore, by using a phase retrieval technique, the problem of not being able to recover pixel values from better phaseless visual features can be solved. Studies on phase retrieval have been intensively carried out in recent years since new techniques based on sparse representation have been developed [29]–[31]. Thus, a more successful inpainting method is expected to be realized by introducing such newly proposed methods.

In this paper, we propose a new inpainting method using sparse representation based on a phaseless quality metric. As described above, since phaseless features can represent original visual texture characteristics, we adopt power spectra as the phaseless features. Furthermore, since the phaseless features are robust to some mismatches such as the difference between the patch clipping interval and the period of textures, they can solve the existing problem. In the proposed method, we use sparse representation for performing inpainting based on a new phaseless quality metric. This new quality metric can be derived by monitoring errors converged through a phase retrieval algorithm, which is based on GrEedy Sparse PhAse Retrieval (GESPAR) algorithm [29] but can be simplified for our inpainting method. This is the biggest contribution of our method. For finding the optimal dictionary and sparse representation coefficients, we apply phase retrieval to the optimization of problems determined on the basis of the new phaseless quality metric. Interestingly, the phase retrieval algorithm adopted in our method has the following two important roles: (1) derivation of the new phaseless quality metric and (2) recovery of original pixel values from phaseless features. Consequently, our novel approach achieves successful inpainting since the proposed method provides better image representation performance.

It should be noted that the proposed method can be regarded as an extended version of [32]. Our previous work is based on an example-based approach, and the proposed method newly introduces sparse representation using the new phaseless quality metric. Although several methods based on deep learning techniques such as the methods in [33], [34] have been proposed, this paper focuses only on improvement of image representation performance based on the new quality metric. This is the placement of this study.

The organization of this paper is shown as follows. First, in Sect. 2, characteristics of phaseless features, which are suitable for representing textures, are explained. In Sect. 3, the new signal recovery algorithm using sparse representation based on the phaseless quality metric is presented. In Sect. 4, the inpainting method based on our new signal recovery algorithm is shown. In Sect. 5, experimental results obtained by applying our inpainting method to images including missing pixels are shown. Finally, concluding remarks are presented in Sect. 6.

2. Characteristics of Phaseless Features

Unique characteristics of phaseless features that can be observed in images are described in this section. As explained above, the direct use of pixel values as visual features is not suitable since they are sensitive to several variations of images, which are negligible changes in human perception. Phaseless features, e.g., Fourier transform magnitude and power spectra, have better characteristics than these features. An interesting example is shown in Fig. 1 and Table 1 for confirming the superiority of phaseless features. In this example, two patches (Figs. 1(a) and (b)) were clipped from different positions within the same texture image, and one



(a) (b) (c)

Fig. 1 Examples of texture images. The two images in (a) and (b) include the same kind of texture. The texture of the image shown in (c) is different from that of (a) and (b).

Table 1 Characteristics of two different kinds of texture features (intensity values and power spectrum values). Images shown in (a)–(c) correspond to (a)–(c) in Fig. 1, respectively. "Dist. of Int." and "Dist. of Pow." represent "Distance of Intensities" and "Distance of Power Spectra", respectively. Furthermore, "same" and "different" mean that the pair of the two images include "the same textures" and "different textures", respectively.

Image pair	Dist. of Int.	Dist. of Pow.
(a) and (b) (same)	2.08×10^3	4.34×10^7
(b) and (c) (different)	1.78×10^3	7.92×10^7
(c) and (a) (different)	1.57×10^3	7.41×10^7

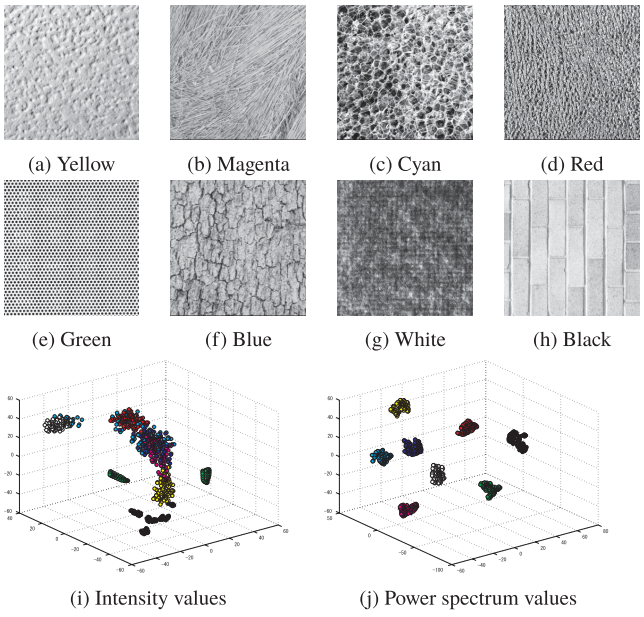


Fig. 2 Comparison of texture representation performance between two different kinds of features, "pixel (intensity) values" and "power spectrum values", obtained from clipped patches within eight images in (a)–(h). The two figures in (i) and (j) show dimensionality reduction results obtained for the above two kinds of features, where t-SNE [35] was adopted for the dimensionality reduction. Each dot in these figures corresponds to a patch obtained from the eight images. The color of each dot in (i) and (j) corresponds to the color shown in the captions of (a)–(h).

patch (Fig. 1(c)) was clipped from an image including a different kind of texture. Then we calculated two kinds of distances based on intensity values and power spectra for these clipped patches. From the obtained results, it can be seen that the distances based on power spectra are better than those based on pixel (intensity) values.

Next, we show another example. In this example, small patches were clipped from eight test images shown in Figs. 2(a)–(h), and their distances based on intensity values and power spectra were respectively calculated for all pairs of these small patches. Then we performed dimensionality reduction by using the calculated distances for visualizing the relationship of these small patches as shown in Figs. 2(i) and (j), where t-distributed Stochastic Neighbor Embedding (t-SNE) [35], which had been a benchmarking method, was used for the dimensionality reduction. As shown in these visualization results, power spectra can represent texture characteristics, i.e., visual characteristics, more successfully than can intensity values, and they can be thus used as phaseless texture features.

In image inpainting, we assume that several known regions exist around target missing regions. Thus, by using the phaseless features of those known regions, we can estimate texture features of the target missing regions. It should be noted that we can estimate the phaseless features, i.e., power spectra, of the missing regions from their surrounding regions, but the final outputs that we have to estimate are the pixel values within the target missing regions. Therefore, we adopt a phase retrieval algorithm for converting these phaseless features to pixel values, and effective use of these phaseless features can be expected.

3. Algorithm for Recovery of Missing Values

An algorithm for signal recovery of missing values is presented in this section. The proposed method approximates target signals, i.e., image signals, including missing values by using sparse representation based on the phaseless quality metric. In 3.1, we first perform dictionary learning from known signals based on the phaseless quality metric. Next, in 3.2, we show the algorithm for recovering the target signals including missing values based on the phaseless quality metric, which is newly derived by using the phase retrieval algorithm.

3.1 Dictionary Learning Based on Phaseless Quality Metric

An algorithm for dictionary learning based on the phaseless quality metric is shown in this subsection. Given training signals $\boldsymbol{x}_n \in \mathcal{R}^D$ ($n = 1, 2, \cdots, N$; N being the number of known training signals) whose dimension is D, the proposed method tries to approximate them on the basis of the following sparse representation:

$$\boldsymbol{x}_n \cong \boldsymbol{D}\boldsymbol{a}_n, \tag{1}$$

where $\boldsymbol{D} = [\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_K] \in \mathcal{R}^{D\times K}$ represents the dictionary matrix including K signal atoms, and $\boldsymbol{a}_n \in \mathcal{R}^K$ is a sparse vector including less than or equal to T non-zero elements, i.e., $\|\boldsymbol{a}_n\|_0 \le T$. Generally, for evaluating the approximation quality in Eq. (1), the quadratic distance between the two signals is traditionally used, and the following problem is commonly solved:

$$\min_{\boldsymbol{a}_n} \|\boldsymbol{x}_n - \boldsymbol{D}\boldsymbol{a}_n\|^2 \quad \text{subject to} \quad \|\boldsymbol{a}_n\|_0 \le T. \tag{2}$$

On the other hand, the proposed method adopts the following phaseless quality metric instead of using Eq. (2):

$$\min_{\boldsymbol{a}_n} \|\mathcal{P}(\boldsymbol{x}_n) - \mathcal{P}(\boldsymbol{D}\boldsymbol{a}_n)\|^2 \quad \text{subject to} \quad \|\boldsymbol{a}_n\|_0 \le T, \tag{3}$$

where $\mathcal{P}(\cdot)$ is an operator calculating a power spectrum from a given signal. Equation (3) can be rewritten as

$$\min_{\boldsymbol{a}_n} \sum_{d=1}^{D} \left(\boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n - \boldsymbol{a}_n^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D}\boldsymbol{a}_n\right)^2 \quad \text{subject to} \quad \|\boldsymbol{a}_n\|_0 \le T, \tag{4}$$

where $\boldsymbol{G}_d = \mathrm{Re}(\boldsymbol{F}_d)^\top \mathrm{Re}(\boldsymbol{F}_d) + \mathrm{Im}(\boldsymbol{F}_d)^\top \mathrm{Im}(\boldsymbol{F}_d)$, and $\boldsymbol{F}_d$ corresponds to the dth row of the discrete Fourier transform matrix. Thus, the dth elements of $\mathcal{P}(\boldsymbol{x}_n)$ and $\mathcal{P}(\boldsymbol{D}\boldsymbol{a}_n)$ are $\boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n$ and $\boldsymbol{a}_n^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D}\boldsymbol{a}_n$, respectively, in Eq. (4). Note that the transpose of vectors/matrices is represented by $^\top$. By solving the problem of Eq. (4), sparse representation based on the phaseless quality metric becomes feasible.

In the proposed method, we perform not only calculation of the coefficient vectors $\boldsymbol{a}_n$ $(n = 1, 2, \cdots, N)$ but also calculation of the dictionary matrix $\boldsymbol{D}$ that minimizes

$$\min_{\boldsymbol{D},\boldsymbol{A}} \sum_{n=1}^{N} \sum_{d=1}^{D} \left(\boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n - \boldsymbol{a}_n^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D} \boldsymbol{a}_n \right)^2$$
$$\text{subject to } \|\boldsymbol{a}_n\|_0 \le T \quad (n = 1, 2, \cdots, N), \tag{5}$$

where $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N]$. In the same manner as the K-SVD algorithm [36], the proposed method iteratively updates the coefficient vectors $\boldsymbol{a}_n$ $(n = 1, 2, \cdots, N)$ and the dictionary $\boldsymbol{D}$.

3.1.1 Update of Sparse Representation Coefficients

For each $\boldsymbol{x}_n$ $(n = 1, 2, \cdots, N)$, we iteratively perform selection of non-zero elements in $\boldsymbol{a}_n$ one-by-one and estimate the optimal values of these non-zero elements. In the following explanation, we show the lth non-zero element selection and calculation. First, according to Eq. (4), the lth non-zero element selection is performed by finding the signal atom minimizing

$$\begin{aligned}
&\sum_{d=1}^{D} \left\{ \boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n - \left(\boldsymbol{a}_{n,\bar{l}}^\top \boldsymbol{D}_{\bar{l}}^\top \boldsymbol{G}_d \boldsymbol{D}_{\bar{l}} \boldsymbol{a}_{n,\bar{l}} + a_{n,l}^2 \boldsymbol{d}_l^\top \boldsymbol{G}_d \boldsymbol{d}_l \right) \right\}^2 \\
&= \sum_{d=1}^{D} \left\{ \boldsymbol{d}_l^\top \boldsymbol{G}_d \boldsymbol{d}_l a_{n,l}^2 - \left(\boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n - \boldsymbol{a}_{n,\bar{l}}^\top \boldsymbol{D}_{\bar{l}}^\top \boldsymbol{G}_d \boldsymbol{D}_{\bar{l}} \boldsymbol{a}_{n,\bar{l}} \right) \right\}^2 \\
&= \sum_{d=1}^{D} \left(c_{l,d} a_{n,l}^2 - e_{n,l,d} \right)^2,
\end{aligned} \tag{6}$$

where $\boldsymbol{a}_{n,\bar{l}} \in \mathcal{R}^{K-1}$ is a vector that is obtained by removing the lth element from $\boldsymbol{a}_n$ and whose $l-1$ non-zero elements have been previously calculated in the previous $l-1$ iterations, and $\boldsymbol{D}_{\bar{l}} \in \mathcal{R}^{D \times K-1}$ is a matrix obtained by removing the lth signal atom. Furthermore,

$$c_{l,d} = \boldsymbol{d}_l^\top \boldsymbol{G}_d \boldsymbol{d}_l, \tag{7}$$
$$e_{n,l,d} = \boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n - \boldsymbol{a}_{n,\bar{l}}^\top \boldsymbol{D}_{\bar{l}}^\top \boldsymbol{G}_d \boldsymbol{D}_{\bar{l}} \boldsymbol{a}_{n,\bar{l}}. \tag{8}$$

The optimal result of $a_{n,l}^2$ minimizing Eq. (6) is easily obtained by

$$a_{n,l}^2 = \frac{\sum_{d=1}^{D} c_{l,d} e_{n,l,d}}{\sum_{d=1}^{D} c_{l,d}^2}, \tag{9}$$

and the lth optimal signal atom minimizing Eq. (6) can be selected. Note that from the definition of Eq. (6), we need only $a_{n,l}^2$ for finding the optimal l. Therefore, in our method, we need not to obtain the sign of $a_{n,l}$ in this procedure.

Next, we show the calculation of values of the l non-zero elements. First, given a vector including only the l non-zero elements as $\boldsymbol{\gamma}_{n,l} \in \mathcal{R}^l$ and a matrix including the corresponding signal atoms as $\hat{\boldsymbol{D}}_l \in \mathcal{R}^{D \times l}$, we define the following optimization problem to obtain the optimal coefficients $\boldsymbol{\gamma}_{n,l}$ for each $\boldsymbol{x}_n$:

$$\min_{\boldsymbol{\gamma}_{n,l}} \sum_{d=1}^{D} \left(\boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n - \boldsymbol{\gamma}_{n,l}^\top \hat{\boldsymbol{D}}_l^\top \boldsymbol{G}_d \hat{\boldsymbol{D}}_l \boldsymbol{\gamma}_{n,l} \right)^2. \tag{10}$$

The optimal solution in Eq. (10) is estimated by a simple iteration algorithm, and its details are shown below. We first define

$$\begin{aligned}
f_d(\boldsymbol{\gamma}_{n,l}) &= \boldsymbol{\gamma}_{n,l}^\top \hat{\boldsymbol{D}}_l^\top \boldsymbol{G}_d \hat{\boldsymbol{D}}_l \boldsymbol{\gamma}_{n,l} - \boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n \\
&= \boldsymbol{\gamma}_{n,l}^\top \boldsymbol{H}_{l,d} \boldsymbol{\gamma}_{n,l} - w_{n,d},
\end{aligned} \tag{11}$$

where

$$\boldsymbol{H}_{l,d} = \hat{\boldsymbol{D}}_l^\top \boldsymbol{G}_d \hat{\boldsymbol{D}}_l, \tag{12}$$
$$w_{n,d} = \boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n. \tag{13}$$

Given the estimation result of $\boldsymbol{\gamma}_{n,l}$ in the tth iteration as $\boldsymbol{\gamma}_{n,l,t}$, we approximate $f_d(\boldsymbol{\gamma}_{n,l})$ around $\boldsymbol{\gamma}_{n,l,t}$ as follows:

$$\begin{aligned}
\tilde{f}_d(\boldsymbol{\gamma}_{n,l}) &= f_d(\boldsymbol{\gamma}_{n,l,t}) \\
&\quad + \nabla f_d(\boldsymbol{\gamma}_{n,l,t}) (\boldsymbol{\gamma}_{n,l} - \boldsymbol{\gamma}_{n,l,t}) \\
&= \boldsymbol{\gamma}_{n,l,t}^\top \boldsymbol{H}_{l,d} \boldsymbol{\gamma}_{n,l,t} - w_{n,d} \\
&\quad + 2 (\boldsymbol{H}_{l,d} \boldsymbol{\gamma}_{n,l,t})^\top (\boldsymbol{\gamma}_{n,l} - \boldsymbol{\gamma}_{n,l,t}) \\
&= \left(2 \boldsymbol{\gamma}_{n,l,t}^\top \boldsymbol{H}_{l,d} \right) \boldsymbol{\gamma}_{n,l} \\
&\quad - \left(\boldsymbol{\gamma}_{n,l,t}^\top \boldsymbol{H}_{l,d} \boldsymbol{\gamma}_{n,l,t} + w_{n,d} \right).
\end{aligned} \tag{14}$$

Then, based on Eq. (14), Eq. (10) is rewritten as

$$\min_{\boldsymbol{\gamma}_{n,l}} \sum_{d=1}^{D} \tilde{f}_d(\boldsymbol{\gamma}_{n,l})^2, \tag{15}$$

and the cost function is a quadratic form of $\boldsymbol{\gamma}_{n,l}$. Thus, the updated solution $\boldsymbol{\gamma}_{n,l,t+1}$ can be easily obtained. Therefore, by iterating the above calculation, we can obtain the optimal non-zero l elements.

Finally, by substituting the elements in $\boldsymbol{\gamma}_{n,T}$ obtained by the above algorithm into the corresponding non-zero elements of $\boldsymbol{a}_n$, we can obtain the optimal solution of $\boldsymbol{a}_n$ in Eq. (5).

3.1.2 Update of Dictionary

As described above, we can update the sparse representation coefficients $\boldsymbol{a}_n$ $(n = 1, 2, \cdots, N)$. Next, we fix them and update the dictionary $\boldsymbol{D}$. Specifically, the signal atoms $\boldsymbol{d}_j$ $(j = 1, 2, \cdots, K)$ are updated one-by-one by solving the following problem:

$$\min_{\boldsymbol{d}_j} \sum_{n|a_{n,j} \neq 0} \sum_{d=1}^{D} g_{n,d} \left(\boldsymbol{d}_j \right)^2, \tag{16}$$

where $a_{n,j}$ is a jth element of $\boldsymbol{a}_n$, and

$$\begin{aligned}
g_{n,d}\left(\boldsymbol{d}_j \right) &= \boldsymbol{a}_n^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D} \boldsymbol{a}_n - \boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n \\
&= a_{n,j}^2 \boldsymbol{d}_j^\top \boldsymbol{G}_d \boldsymbol{d}_j + \boldsymbol{a}_{n,\bar{j}}^\top \boldsymbol{D}_{\bar{j}}^\top \boldsymbol{G}_d \boldsymbol{D}_{\bar{j}} \boldsymbol{a}_{n,\bar{j}} \\
&\quad - \boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n \\
&= \boldsymbol{d}_j^\top \boldsymbol{S}_{n,j,d} \boldsymbol{d}_j - z_{n,j,d},
\end{aligned} \tag{17}$$

Algorithm 1 Dictionary learning algorithm based on phaseless quality metric

Input: $\boldsymbol{x}_n$ $(n = 1, 2, \cdots, N)$, K and T.
Output: Dictionary $\boldsymbol{D} = [\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_K]$ and coefficient matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_N]$.

while $\boldsymbol{D}$ and $\boldsymbol{A}$ are not converged **do**
  for $n = 1, 2, \cdots, N$ **do**
    for $l = 1, 2, \cdots, T$ **do**
      Select the optimal lth non-zero element minimizing Eq. (6) for nth sample.
      Calculate the optimal $\boldsymbol{\gamma}_{n,l}$ in Eq. (15).
    end for
    Update $\boldsymbol{a}_n$ by substituting $\boldsymbol{\gamma}_{n,T}$ into the corresponding non-zero elements.
    for $j = 1, 2, \cdots, K$ **do**
      Update $\boldsymbol{d}_j$ minimizing Eq. (16) whose $g_{n,d}\left(\boldsymbol{d}_j\right)$ is approximated by $\hat{g}_{n,d}\left(\boldsymbol{d}_j\right)$.
    end for
  end for
end while

where

$$\boldsymbol{S}_{n,j,d} = a_{n,j}^2 \boldsymbol{G}_d, \tag{18}$$

$$z_{n,j,d} = -\boldsymbol{a}_{n,\bar{j}}^\top \boldsymbol{D}_{\bar{j}}^\top \boldsymbol{G}_d \boldsymbol{D}_{\bar{j}} \boldsymbol{a}_{n,\bar{j}} + \boldsymbol{x}_n^\top \boldsymbol{G}_d \boldsymbol{x}_n. \tag{19}$$

Then we can easily estimate the optimal solution of $\boldsymbol{d}_j$ by approximating $g_{n,d}\left(\boldsymbol{d}_j\right)$ as the quadratic form $\tilde{g}_{n,d}\left(\boldsymbol{d}_j\right)$ of $\boldsymbol{d}_j$ in the same manner as Eq. (14).

As shown in the above explanation, we can perform dictionary learning based on the new phaseless quality metric. Its specific algorithm is shown in Algorithm 1. By removing the phase components for updating the dictionary, the fundamental visual characteristics can be represented by fewer signal atoms. In order to perform missing data recovery, which is usually an ill-posed problem, approximation of target signals in lower-dimensional subspaces is necessary for obtaining unique solutions. Then since our method can realize more accurate approximation in the lower-dimensional subspaces (T-dimensional subspaces) obtained by the dictionary based on the phaseless quality metric, improvement of the recovery performance can be expected.

3.2 Signal Recovery Based on Phaseless Quality Metric

Given a target signal including missing values, the proposed method estimates the missing values by retrieving its phase. First, we define an original signal vector $\boldsymbol{x} \in \mathcal{R}^D$ for which signal values in the missing parts are known, and it is decomposed into two elements as

$$\boldsymbol{x} = \boldsymbol{B}_{\boldsymbol{\mu}} \boldsymbol{\mu} + \boldsymbol{B}_{\boldsymbol{\kappa}} \boldsymbol{\kappa}, \tag{20}$$

where $\boldsymbol{\mu} \in \mathcal{R}^{D_\mu}$ and $\boldsymbol{\kappa} \in \mathcal{R}^{D_\kappa}$ are vectors including unknown and known signal values, respectively, within $\boldsymbol{x}$, and $D = D_{\boldsymbol{\mu}} + D_{\boldsymbol{\kappa}}$. Furthermore, $\boldsymbol{B}_{\boldsymbol{\mu}} \in \mathcal{R}^{D \times D_\mu}$ and $\boldsymbol{B}_{\boldsymbol{\kappa}} \in \mathcal{R}^{D \times D_\kappa}$ are matrices, for which elements are binary values, for reconstructing $\boldsymbol{x}$ from the two decomposed elements.

In the proposed method, we perform estimation of the original vector $\boldsymbol{x}$ by the following sparse representation:

$$\boldsymbol{x} \cong \boldsymbol{D}\boldsymbol{a} \quad \text{subject to} \quad \|\boldsymbol{a}\|_0 \leq T, \tag{21}$$

where $\boldsymbol{a} \in \mathcal{R}^K$ is the sparse representation coefficient vector of $\boldsymbol{x}$. Then the solution of the above equation is obtained by solving

$$\min_{\boldsymbol{x},\boldsymbol{a}} \left\| \mathcal{P}(\boldsymbol{x}) - \mathcal{P}(\boldsymbol{D}a) \right\|^2 \quad \text{subject to} \quad \|\boldsymbol{a}\|_0 \leq T, \tag{22}$$

and it can be rewritten as

$$\min_{\boldsymbol{x},\boldsymbol{a}} \sum_{d=1}^{D} \left(\boldsymbol{x}^\top \boldsymbol{G}_d \boldsymbol{x} - \boldsymbol{a}^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D}\boldsymbol{a} \right)^2$$
$$\text{subject to} \quad \|\boldsymbol{a}\|_0 \leq T. \tag{23}$$

Furthermore, by substituting Eq. (20) into Eq. (23), the problem is rewritten as

$$\min_{\boldsymbol{\mu},\boldsymbol{a}} \sum_{d=1}^{D} g_d(\boldsymbol{\mu}, \boldsymbol{a})^2 \quad \text{subject to} \quad \|\boldsymbol{a}\|_0 \leq T, \tag{24}$$

where

$$\begin{aligned} g_d(\boldsymbol{\mu}, \boldsymbol{a}) &= \left(\boldsymbol{B}_{\boldsymbol{\mu}} \boldsymbol{\mu} + \boldsymbol{B}_{\boldsymbol{\kappa}} \boldsymbol{\kappa} \right)^\top \boldsymbol{G}_d \left(\boldsymbol{B}_{\boldsymbol{\mu}} \boldsymbol{\mu} + \boldsymbol{B}_{\boldsymbol{\kappa}} \boldsymbol{\kappa} \right) \\ &\quad - \boldsymbol{a}^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D}\boldsymbol{a} \\ &= \boldsymbol{\mu}^\top \boldsymbol{B}_{\boldsymbol{\mu}}^\top \boldsymbol{G}_d \boldsymbol{B}_{\boldsymbol{\mu}} \boldsymbol{\mu} + 2\boldsymbol{\kappa}^\top \boldsymbol{B}_{\boldsymbol{\kappa}}^\top \boldsymbol{G}_d \boldsymbol{B}_{\boldsymbol{\mu}} \boldsymbol{\mu} \\ &\quad + \boldsymbol{\kappa}^\top \boldsymbol{B}_{\boldsymbol{\kappa}}^\top \boldsymbol{G}_d \boldsymbol{B}_{\boldsymbol{\kappa}} \boldsymbol{\kappa} - \boldsymbol{a}^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D}\boldsymbol{a}. \end{aligned} \tag{25}$$

It should be noted that the problem of estimating unknown signal values $\boldsymbol{\mu}$ from the approximated power spectrum is equivalent to that of phase retrieval for the target signal $\boldsymbol{x}$. Then based on the GESPAR algorithm [29], we try to estimate the unknown vector $\boldsymbol{\mu}$ by estimating the sparse representation coefficients $\boldsymbol{a}$. Therefore, the cost function shown in Eq. (24) corresponds to our newly proposed phaseless quality metric derived by the phase retrieval algorithm. The proposed method finds the optimal solutions by iteratively updating $\boldsymbol{\mu}$ and $\boldsymbol{a}$, where $\boldsymbol{a}$ can be updated in the same way as the estimation of $\boldsymbol{a}_n$ in the previous subsection. Furthermore, by fixing $\boldsymbol{a}$, we define

$$g_d(\boldsymbol{\mu}) = \boldsymbol{\mu}^\top \boldsymbol{P}_d \boldsymbol{\mu} + 2\boldsymbol{q}_d^\top \boldsymbol{\mu} + r_d, \tag{26}$$

where

$$\boldsymbol{P}_d = \boldsymbol{B}_{\boldsymbol{\mu}}^\top \boldsymbol{G}_d \boldsymbol{B}_{\boldsymbol{\mu}}, \tag{27}$$

$$\boldsymbol{q}_d = \boldsymbol{B}_{\boldsymbol{\mu}}^\top \boldsymbol{G}_d \boldsymbol{B}_{\boldsymbol{\kappa}} \boldsymbol{\kappa}, \tag{28}$$

$$r_d = \boldsymbol{\kappa}^\top \boldsymbol{B}_{\boldsymbol{\kappa}}^\top \boldsymbol{G}_d \boldsymbol{B}_{\boldsymbol{\kappa}} \boldsymbol{\kappa} - \boldsymbol{a}^\top \boldsymbol{D}^\top \boldsymbol{G}_d \boldsymbol{D}\boldsymbol{a}. \tag{29}$$

The estimation of the optimal solution of $\boldsymbol{\mu}$ can be achieved in the same manner as shown in the previous subsection, i.e., the iteration algorithm. Given the tth estimation result of $\boldsymbol{\mu}$ as $\boldsymbol{\mu}_t$, we first approximate $g_d(\boldsymbol{\mu})$ in Eq. (26) around $\boldsymbol{\mu}_t$ as follows:

$$g_d(\boldsymbol{\mu}) \cong g_d(\boldsymbol{\mu}_t) + \nabla g_d(\boldsymbol{\mu}_t)^\top (\boldsymbol{\mu} - \boldsymbol{\mu}_t). \tag{30}$$

Algorithm 2 Signal recovery algorithm

Input: Known signal values $\boldsymbol{\kappa}$.
Output: Recovered signal values $\hat{\boldsymbol{\mu}}$ of $\boldsymbol{\mu}$.

while $\boldsymbol{a}$ is not converged **do**
  for $l = 1, 2, \cdots, T$ **do**
    Select the optimal lth non-zero element minimizing Eq. (6), where the subscript $_n$ is removed from this equation.
    Calculate the optimal $\boldsymbol{\gamma}_l$ in Eq. (15), where the subscript $_n$ is removed from this equation.
  end for
  Update $\boldsymbol{a}$ by substituting $\boldsymbol{\gamma}_T$ into the corresponding non-zero elements.
  while $\boldsymbol{\mu}_t$ is not converged **do**
    Update $\boldsymbol{\mu}_t$ using Eq. (31).
  end while
  $\boldsymbol{x} = \boldsymbol{B}_{\mu}\boldsymbol{\mu}_t + \boldsymbol{B}_{\kappa}\boldsymbol{\kappa}$.
end while
$\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}_t$.

Then the problem for estimating $\boldsymbol{\mu}$, i.e., the update of $\boldsymbol{\mu}_t$, is written as

$$\boldsymbol{\mu}_{t+1} = \arg\min_{\boldsymbol{\mu}} \sum_{d=1}^{D} \left\{ 2\left(\boldsymbol{P}_d\boldsymbol{\mu}_t + \boldsymbol{q}_d\right)^{\top}\boldsymbol{\mu} - \left(\boldsymbol{\mu}_t^{\top}\boldsymbol{P}_d\boldsymbol{\mu}_t - r_d\right) \right\}^2. \tag{31}$$

The right side of the above equation is a quadratic form of $\boldsymbol{\mu}$, and the optimal solution is easily found. As described above, our method can estimate the unknown values $\boldsymbol{\mu}$ within the target signal $\boldsymbol{x}$ by retrieving its phase. Our phase retrieval algorithm is based on the GESPAR algorithm [29] and can be regarded as its simplified version since the known values $\boldsymbol{\kappa}$ within $\boldsymbol{x}$ can be used as the signal constraint. Finally, we show the algorithm of the signal recovery in Algorithm 2.

In our method, we monitor the errors converged in phase retrieval of the target signal, i.e., the minimum distance between the power spectra of the target signal and its approximation result on the basis of sparse representation. As shown in Sect. 2, since the phaseless features are suitable for representing visual characteristics, the phaseless quality metric shown in Eq. (5) of the previous subsection and the metric derived in Eq. (23) are better than that directly obtained from signal values. Furthermore, by using the phase retrieval algorithm, our method enables the conversion of phaseless features to signal values, and this provides a solution for the existing problem of previously reported methods.

4. Image Inpainting Method

In this section, we show the specific procedures of our image inpainting method. Given a target image including missing regions, the proposed inpainting method performs their missing intensity estimation. Since it is necessary to determine their recovery order, i.e., patch priority, we determine it on the basis of a previously reported approach [16]. Then the proposed method clips patches from known areas within the target image and obtains their intensity vectors as $\boldsymbol{x}_n$ as described in the previous section to derive the optimal dictionary matrix $\boldsymbol{D}$. Then, according to the patch priority [16], the target patch including missing pixels is selected, and its intensity vectors $\boldsymbol{x}$ are recovered as shown in 3.2. This recovery procedure is iterated until all of the missing intensities are estimated.

5. Experimental Results

This section shows the effectiveness of the novel approach in the proposed method from experimental results obtained by applying the proposed method and comparative methods to image inpainting. First, in this experiment, we prepared seven images shown in the first column of Figs. 3 and 4. We used these images due to the following reason. From Figs. 1 and 2, we can confirm the effectiveness of the phaseless features, but the problem of inpainting these textures is too simple to show the difference of the performance between the methods. With increasing the kinds of textures within images, the difference becomes larger since the inpainting problem becomes difficult. Therefore, in this experiment, we used the real images including several textures. Next, as shown in the second column of Figs. 3 and 4, we added superimposed text regions corresponding to missing regions to the original images. The proposed method realizes high image representation performance based on the new quality metric. Thus, the proposed method corresponds to a new improved method among the methods that adopt multivariate analysis series. From this point of view, we focus only on verifying how our method can improve representation performance among the methods adopting multivariate analysis. Thus, we mainly focus on this and not on recovery of large missing regions. In the experiment, training patches were clipped from known regions within each target image, i.e., the dictionary was trained from the single input. Note that we assume that the pixel positions of these missing regions had been previously provided.

For these images including missing regions, we performed inpainting based on the proposed method shown in Sects. 3 and 4. For comparison, we used the recent and state-of-the-art methods [4], [13], [15], [16], [18], [25], [32]. The methods in [4], [13], [15], [16], [18] directly utilize intensity values for recovering missing regions, and we show Fig. 3 for verifying the effectiveness of using the phaseless features. Furthermore, the methods in [25], [32] shown in Figs. 5 and 6 use phaseless features but adopt exemplar-based approaches, i.e., they use only one-dimensional space for approximating textures. Since the proposed method adopts multi-dimensional approximation based on the sparse representation, Figs. 5 and 6 are shown to verify its effectiveness. The method in [25] uses Fourier transform magnitude as phaseless texture features, and the method in [32] performs exemplar-based image inpainting based on power spectrum values. In Figs. 5 and 6, we show the results obtained by not only the methods in [25], [32] but

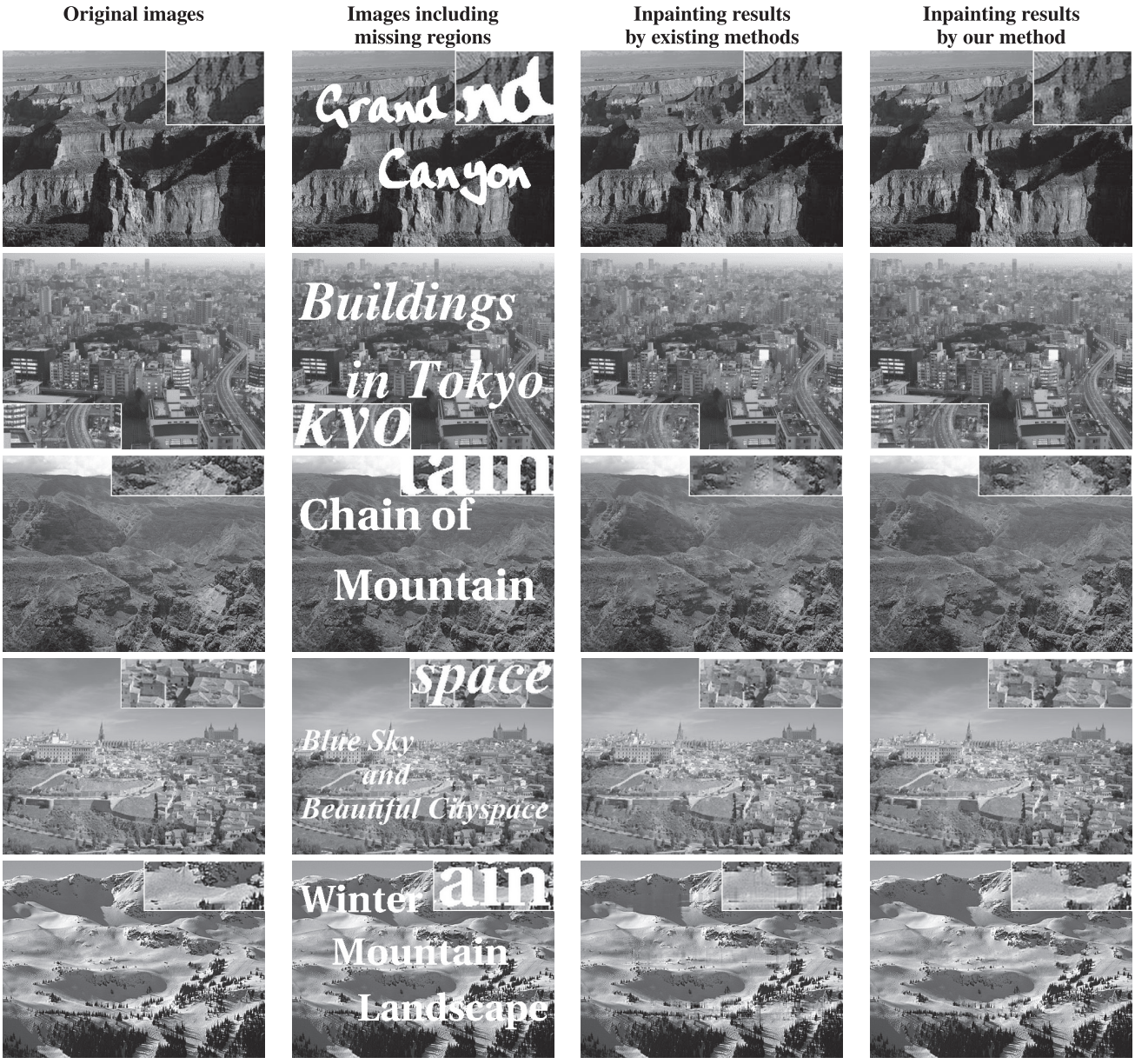


Fig. 3 Results of image inpainting by the existing methods in [4], [13], [15], [16], [18] that directly use pixel values and our method. From the top to bottom, we call these images "Images 1–5", and they correspond to the images shown in Table 2. The inpainting results shown in the third column are obtained by the methods in [4], [13], [16], [15] and [18], respectively. The sizes of Images 1, 2, 3 and 5 are 480 × 360 pixels, and the size of Image 4 is 640 × 480 pixels. All of the five test images are 8-bit gray levels. The ratios of missing pixels in Images 1–5 are 8.9%, 10.7%, 8.9%, 5.9%, and 10.2%, respectively.

also those in [4], [13], [15], [16], [18].

In our method, the size of patches was set to 11×11 pixels, and their clipping interval was set to half the size of the patches. Furthermore, $T = 10$, and K was set to $\frac{N}{2}$. Since it is desirable to set K in such a way that it satisfies $D < K < N$, we determined K as $\frac{N}{2}$ which tends to satisfy this condition in almost all cases. Next, T should be set in such a way that its value becomes lower than the dimension of κ, i.e., D_κ. Thus, T was set to 10 since it tends to satisfy this condition in almost all cases. It should be noted that when considering the recovery performance, it is desirable to set T to a larger value in the range satisfying the above condition, and K is also the same. With considering that the computational cost becomes larger when K and T become larger, we should determine their optimal values for each image. Therefore, their determination will be addressed in our future work.

Inpainting results obtained by our method and the com-



Fig. 4 Two other test images used in the experiment. From the top to bottom, we call these images "Images 6 and 7", and they correspond to the images shown in Table 2. The size of Image 6 is 640 × 480 pixels, and that of Image 7 is 480 × 360 pixels. Both of the test images are 8-bit gray levels. The ratios of missing pixels in Images 6 and 7 are 5.5% and 11.3%, respectively.

parative methods are shown in the third and forth columns of Fig. 3 and Figs. 5 and 6. From the obtained results, it can be seen that our method outputs better results than those of the comparative methods. Specifically, as shown in the enlarged portions of Fig. 5, if sufficient training patches are not obtained, the methods based on the exemplar-based approach cannot find suitable known patches recovering the missing regions. By using the phaseless features, it can be confirmed that the reconverting performance, e.g., in window parts, is improved by comparing the method in [4] and those in [25], [32]. This tendency can be confirmed from the enlarged portions of Fig. 6. Furthermore, from the enlarged potions shown in Fig. 6, our method using the phaseless features can preserve the important visual characteristics such as edges in leaves more successfully than the comparative methods [13], [15], [16], [18] adopting the multivariate analysis-based approach but not using the phaseless features. Table 2 shows the structural similarity (SSIM) index [21] obtained from these inpainting results, and the proposed method outperforms the other comparative methods.

As described above, methods that directly use the pixel values are affected by some mismatches. The biggest problem is the gap between the clipping interval of patches and the period of textures. As shown in Sect. 2, even if patches include the same kinds of textures, the characteristics of pixel values become different from each other due to the above gap. Therefore, methods that directly use pixel values such as the methods in [4], [13], [15], [16], [18] cannot successfully approximate textures in lower-dimensional subspaces. On the other hand, the methods in [25], [32] and the proposed method use phaseless features, for which the gap can be ignored, and it becomes feasible to approximate textures by the best matched samples or in lower-dimensional subspaces. Furthermore, since the methods in [25], [32] perform approximation of textures by only one known sample, i.e., they adopt the exemplar-based approach, their approximation performance is limited. Therefore, the proposed method newly introduces sparse representation based on the phaseless quality metric derived from such phaseless features. Then inpainting of missing pixels is realized by effectively using the phase retrieval algorithm, which enables both the derivation of the phaseless quality metric for sparse representation and the estimation of missing pixel values from only the phaseless features. This non-conventional approach solves the above-described problems that existed in previously reported methods.

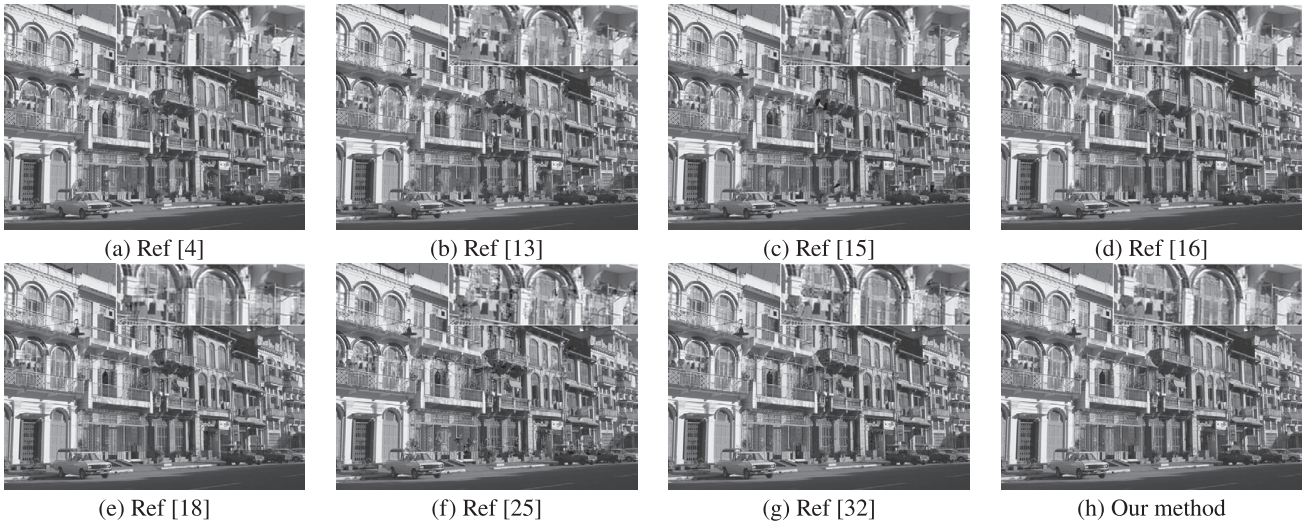
6. Conclusions

A new inpainting method via sparse representation based on a phaseless quality metric is presented in this paper. The proposed method realizes inpainting based on a phase retrieval algorithm, i.e., a simplified version of the GESPER algorithm, and newly derives the phaseless quality metric obtained by monitoring errors converged in this phase retrieval algorithm. Furthermore, dictionary learning for successfully approximating target signals is performed on the basis of the phaseless quality metric. Since the phaseless quality metric in our method is robust to various changes as shown in Sect. 2, sparse representation based on this quality metric can accurately approximate the unique characteristics of target images.

As described above, the phase retrieval algorithm used in our method has two important roles. The first one is derivation of the new quality metric for sparse representation of target images including missing pixel values. The second one is conversion of the phaseless features to signal values under the constraint of the neighboring known pixel values. In this way, inpainting based on the phase retrieval algorithm becomes feasible. Experimental results showed that the proposed method achieves successful recovery compared to previously reported methods that directly use pixel values.

In future work, we will extend the proposed method to inpainting of color images including three color channels. Furthermore, since it is necessary to train the dictionary for each image in the proposed method, the computational cost of the training becomes large. Actually, in the proposed method, it averagely took about 134 seconds for training the dictionary for each image. The computation time was measured on a personal computer using Intel(R) Core(TM) i7 980 CPU 3.33 GHz with 4.0 Gbytes RAM. The implementation of our method was performed by using Matlab 2016. From the above result, it is necessary to reduce its computational cost, and this problem should be addressed in our future work.

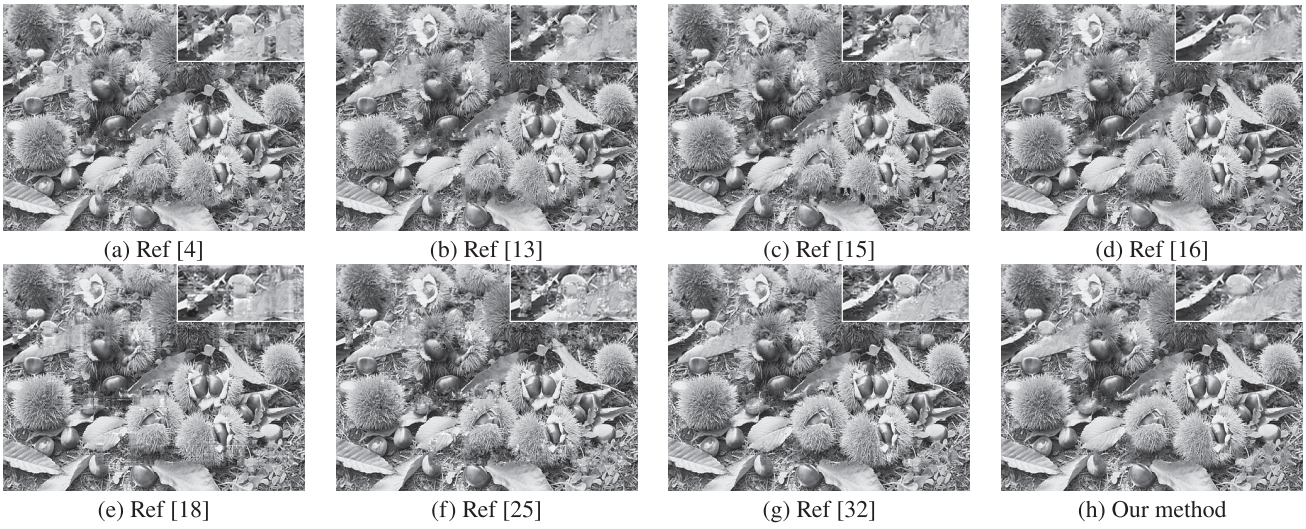
In recent years, inpainting of audio signals has also been studied [37]–[39] since it can also afford various fun-



(a) Ref [4] (b) Ref [13] (c) Ref [15] (d) Ref [16]

(e) Ref [18] (f) Ref [25] (g) Ref [32] (h) Our method

Fig. 5 Results of image inpainting for Image 6 by the existing methods and the proposed method. The methods in [4], [13], [15], [16], [18] correspond to the methods using pixel values as visual features, where the method in [4] is based on the exemplar-based approach, and the methods in [13], [15], [16], [18] are based on the multivariate analysis-based approach. The methods in [25], [32] correspond to the exemplar-based methods using phaseless features, and our method is based on sparse representation using the phaseless quality metric.



(a) Ref [4] (b) Ref [13] (c) Ref [15] (d) Ref [16]

(e) Ref [18] (f) Ref [25] (g) Ref [32] (h) Our method

Fig. 6 Results of image inpainting for Image 7 by the existing methods and the proposed method. The details of each method are the same as those shown in Fig. 5.

Table 2 Comparison of SSIM index between the proposed method and other existing methods.

	Ref [4]	Ref [13]	Ref [15]	Ref [16]	Ref[18]	Ref[25]	Ref [32]	Our method
Image 1	0.632	0.681	0.705	0.687	0.728	0.721	0.802	**0.835**
Image 2	0.716	0.739	0.756	0.752	0.796	0.775	0.824	**0.846**
Image 3	0.823	0.827	0.840	0.815	0.814	0.830	0.849	**0.866**
Image 4	0.791	0.790	0.816	0.801	0.825	0.801	0.851	**0.870**
Image 5	0.772	0.778	0.732	0.766	0.747	0.731	0.797	**0.838**
Image 6	0.727	0.718	0.795	0.734	0.832	0.763	0.853	**0.874**
Image 7	0.738	0.772	0.723	0.760	0.667	0.660	0.773	**0.801**
Average	0.743	0.758	0.767	0.759	0.773	0.754	0.821	**0.847**
Median	0.738	0.772	0.756	0.760	0.796	0.763	0.824	**0.846**

damental applications such as interpolation, extrapolation, concealment and declipping. As one of the applications of audio inpainting, audio declipping can be easily realized by modifying audio inpainting methods [40]–[42]. In the study of audio inpainting, there exists the same problem as that of image inpainting. The inpainting is performed by directly using audio signal values as audio features, but these values will not be able to reflect audio characteristics. Furthermore, even though better audio features can be adopted, it is difficult to recover unknown audio signals from these features. Therefore, since the same background exists in the study of audio inpainting, it is expected that our method can solve the common problem. This will be addressed in our future work.

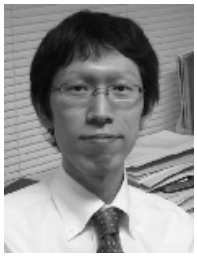
Acknowledgments

This work was partly supported by JSPS KAKENHI Grant Number JP18K11367.

References

[1] C. Guillemot and O. Le Meur, "Image inpainting: Overview and recent advances," IEEE Signal Process. Mag., vol.31, no.1, pp.127–144, 2014.
[2] T. Ogawa and M. Haseyama, "Missing texture reconstruction," Wiley Encyclopedia of Electrical and Electronics Engineering, pp.1–19, 2015.
[3] A. Efros and T. Leung, "Texture synthesis by nonparametric sampling," Proc. IEEE International Conference on Computer Vision, pp.1033–1038, 1999.
[4] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," IEEE Trans. Image Process., vol.13, no.9, pp.1200–1212, 2004.
[5] P. Buyssens, M. Daisy, D. Tschumperlé, and O. Lézoray, "Exemplar-based inpainting: Technical review and new heuristics for better geometric reconstructions," IEEE Trans. Image Process., vol.24, no.6, pp.1809–1824, 2015.
[6] T. Ružić and A. Pižurica, "Context-aware patch-based image inpainting using markov random field modeling," IEEE Trans. Image Process., vol.24, no.1, pp.444–456, 2015.
[7] T. Amano and Y. Sato, "Image interpolation using BPLP method on the eigenspace," Syst. Comput. Jpn., vol.38, no.1, pp.87–96, 2007.
[8] 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.
[9] T. Ogawa and M. Haseyama, "Missing intensity interpolation using a kernel PCA-based POCS algorithm and its applications," IEEE Trans. Image Process., vol.20, no.2, pp.417–432, 2011.
[10] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," IEEE Trans. Image Process., vol.17, no.1, pp.53–69, 2008.
[11] B. Wohlberg, "Inpainting with sparse linear combinations of exemplars," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp.689–692, 2009.
[12] M.J. Fadili, J.L. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," The Computer Journal, vol.52, no.1, pp.64–79, 2009.
[13] Z. Xu and J. Sun, "Image inpainting by patch propagation using patch sparsity," IEEE Trans. Image Process., vol.19, no.5, pp.1153–1165, 2010.
[14] T. Ogawa and M. Haseyama, "Missing image data reconstruction based on adaptive inverse projection via sparse representation," IEEE Trans. Multimedia, vol.13, no.5, pp.974–992, 2011.
[15] Z. Li, H. He, H.M. Tai, Z. Yin, and F. Chen, "Color-direction patch-sparsity-based image inpainting using multidirection features," IEEE Trans. Image Process., vol.24, no.3, pp.1138–1152, 2015.
[16] C. Guillemot, M. Turkan, O. Le Meur, and M. Ebdelli, "Object removal and loss concealment using neighbor embedding methods," Signal Processing: Image Communication, vol.28, no.10, pp.1405–1419, 2013.
[17] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," IEEE Trans. Pattern Anal. Mach. Intell., vol.35, no.1, pp.208–220, 2013.
[18] K.H. Jin and J.C. Ye, "Annihilating filter-based low-rank hankel matrix approach for image inpainting," IEEE Trans. Image Process., vol.24, no.11, pp.3498–3511, 2015.
[19] M. Tuceryan and A.K. Jain, "Texture analysis," Handbook of Pattern Recognition and Computer Vision, pp.235–276, World Scientific, 1993.
[20] X. Xie and M. Mirmehdi, "A galaxy of texture features," Handbook of Texture Analysis, pp.375–406, World Scientific, 2008.
[21] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol.13, no.4, pp.600–612, 2004.
[22] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," IEEE Trans. Image Process., vol.20, no.8, pp.2378–2386, 2011.
[23] W. Lin and C.C.J. Kuo, "Perceptual visual quality metrics: A survey," J Vis. Commun. Image R., vol.22, no.4, pp.297–312, 2011.
[24] T. Ogawa, M. Haseyama, and H. Kitajima, "Reconstruction method of missing texture using error reduction algorithm," Proc. IEEE International Conference on Image Processing, pp.II1026–II1029, 2005.
[25] T. Ogawa and M. Haseyama, "Missing texture reconstruction method based on error reduction algorithm using Fourier transform magnitude estimation scheme," IEEE Trans. Image Process., vol.22, no.3, pp.1252–1257, 2013.
[26] J.R. Fienup, "Reconstruction of an object from the modulus of its fourier transform," Opt. Lett., vol.3, no.1, pp.27–29, 1978.
[27] J.R. Fienup, "Phase retrieval algorithms: A comparison," Appl. Opt., vol.21, no.15, pp.2758–2769, 1982.
[28] J.A. Rodriguez, R. Xu, C.C. Chen, Y. Zou, and J. Miao, "Oversampling smoothness: An effective algorithm for phase retrieval of noisy diffraction intensities," J. Appl. Cryst., vol.46, no.2, pp.312–318, 2013.
[29] Y. Shechtman, A. Beck, and Y.C. Eldar, "GESPAR: Efficient phase retrieval of sparse signals," IEEE Trans. Signal Process., vol.62, no.4, pp.928–938, 2014.
[30] E.J. Candes, Y.C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," SIAM Rev., vol.57, no.2, pp.225–251, 2015.
[31] Y. Shechtman, Y.C. Eldar, O. Cohen, H.N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," IEEE Signal Process. Mag., vol.32, no.3, pp.87–109, 2015.
[32] T. Ogawa and M. Haseyama, "Exemplar-based image completion via new quality measure based on phaseless texture features," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp.1827–1831, 2017.
[33] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A.A. Efros, "Context encoders: Feature learning by inpainting," Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp.2536–2544, 2016.
[34] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent image completion," ACM Trans. Graph., vol.36, no.4, pp.107:1–107:14, 2017.
[35] L.v.d. Maaten and G. Hinton, "Visualizing data using t-SNE," J. Machine Learning Research, vol.9, pp.2579–2605, 2008.

[36] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Trans. Signal Process., vol.54, no.11, pp.4311–4322, 2006.
[37] A. Adler, V. Emiya, M.G. Jafari, M. Elad, R. Gribonval, and M.D. Plumbley, "Audio inpainting," IEEE Trans. Audio, Speech, Language Process., vol.20, no.3, pp.922–932, 2012.
[38] V. Mach and R. Ozdobinski, "Optimizing dictionary learning parameters for solving audio inpainting problem," International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems, vol.2, no.1, pp.39–44, 2013.
[39] K. Siedenburg, M. Dörfler, and M. Kowalski, "Audio inpainting with social sparsity," SPARS (Signal Processing with Adaptive Sparse Structured Representations), 2013.
[40] B. Defraene, N. Mansour, S. De Hertogh, T. van Waterschoot, M. Diehl, and M. Moonen, "Declipping of audio signals using perceptual compressed sensing," IEEE Trans. Audio, Speech, Language Process., vol.21, no.12, pp.2627–2637, 2013.
[41] K. Siedenburg, M. Kowalski, and M. Dorfler, "Audio declipping with social sparsity," Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, pp.1577–1581, 2014.
[42] M.J. Harvilla and R.M. Stern, "Efficient audio declipping using regularized least squares," Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, pp.221–225, 2015.



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 joined Graduate School of Information Science and Technology, Hokkaido University in 2008. He is currently an associate professor in the Faculty of Information Science and Technology, Hokkaido University. His research interests are AI, IoT and big data analysis for multimedia signal processing and its applications. He was a special session chair of IEEE ISCE2009, a Doctoral Symposium Chair of ACM ICMR2018, an organized session chair of IEEE GCCE2017-2019, a TPC Vice Chair of IEEE GCCE2018, a Conference Chair of IEEE GCCE2019, etc. He has been also an Associate Editor of ITE Transactions on Media Technology and Applications. He is a senior member of IEEE and a member of ACM, IEICE and ITE.



Keisuke Maeda received his B.S., M.S. and Ph.D. degrees in Electronics and Information Engineering from Hokkaido University, Japan in 2015, 2017 and 2019, respectively. He is currently a Specially Appointed Assistant Professor at Office of Institutional Research, Hokkaido University. His research interests are multimodal signal processing, machine learning and its applications. He is a member of the IEEE and IEICE.



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