EXEMPLAR-BASED IMAGE COMPLETION VIA NEW QUALITY MEASURE BASED ON PHASELESS TEXTURE FEATURES

Takahiro Ogawa and Miki Haseyama

Graduate School of Information Science and Technology, Hokkaido University
N-14, W-9, Kita-ku, Sapporo, Hokkaido, 060-0814, Japan
E-mail: ogawa@lmd.ist.hokudai.ac.jp, miki@ist.hokudai.ac.jp

ABSTRACT

This paper presents an exemplar-based image completion via a new quality measure based on phaseless texture features. The proposed method derives a new quality measure obtained by monitoring errors caused in power spectra, i.e., errors of phaseless texture features, converged through phase retrieval. Even if a target patch includes missing pixels, this measure enables selection of the best matched patch including most similar texture features for realizing the exemplar-based image completion. Furthermore, since the phaseless texture features are robust to various changes such as spatial gaps and luminance changes, the new quality measure successfully provides the best matched patch from few training examples. Then, by solving an optimization problem that retrieves the phase of the target patch from the phaseless texture features of the best matched patch, its missing areas can be reconstructed. Consequently, accurate image completion using the new quality measure becomes feasible. Subjective and quantitative experimental results are shown to verify the effectiveness of our method using the new quality measure.

Index Terms— Image completion, image quality measure, phase retrieval, phaseless texture feature.

1. INTRODUCTION

Image completion, which is also called image inpainting, has been intensively studied by many researchers, and a number of methods have been proposed [1, 2]. The basic idea of image completion is to recover missing areas included within a target image from the other known areas. The most traditional approach is based on texture synthesis [3], and unknown intensities within missing areas can be estimated by copying known intensities within the target image. Furthermore, Criminisi et al. proposed a representative method called an “exemplar-based approach” [4]. Although the basic strategy of the exemplar-based approach is the same as that of the texture synthesis-based approach, the biggest difference between these two approaches is introduction of a new criterion, “patch priority” for determining image completion order. In recent years, the exemplar-based approach has been extended by improving the performance of image representation and patch priority determination [5–8].

Traditionally, the exemplar-based approach tries to find the best matched patch, whose intensity values are the most similar to those of the target patch, from known parts within the target image. If we assume that the target image consists of several kinds of small textures, the intensity values are regarded as texture features. Then the best matched patch is selected by monitoring distances between these texture features. However, when a clipping interval of known training patches is different from the periods of the textures, this mismatch makes the distance of texture features larger even if two patches include the same kinds of textures [2]. Thus, it is necessary to prepare a tremendous number of training examples, i.e., known patches, from the target image. The same problem also occurs due to some other changes such as luminance changes. Furthermore, even if the number of the training examples increases, the above problem cannot be perfectly avoided. Thus, there have been proposed many alternative methods that adopt low-dimensional approximation of patches based on multivariate analysis such as Principal Component Analysis (PCA) [9], kernel PCA [10, 11], sparse representation [5, 12, 13], neighbor embedding [6, 14] and rank minimization [15, 16]. Although these methods can improve the image representation performance, they cannot directly solve the above problem.

The underlying problem is that the distance directly obtained from intensity values cannot reflect the difference of textures. One solution for solving this problem is introduction of new quality measures which are derived from better visual features accurately representing texture characteristics. Although many accurate texture features [17, 18] or accurate quality measures [19–22] have existed, they cannot be generally applied to the image completion since the final estimation results are intensity values of missing areas. On the other hand, we have reported that Fourier-based features can represent texture characteristics, and missing texture reconstruction becomes feasible by using a phase retrieval technique [23, 24]. It should be noted that the study of the phase retrieval was traditionally carried out [25–27], and in recent years, it has developed rapidly since new different techniques were discovered in the field of sparse representation [28–30]. Therefore, by introducing these state-of-the-art techniques, realization of a new extended image completion method is expected.

In this paper, we present a novel exemplar-based image completion method via a new quality measure based on phaseless texture features. Figure 1 shows interesting examples, results of dimensionality reduction using two kinds of texture features. As shown in Fig. 1, since power spectrum values can represent texture characteristics more successfully than intensity values, we adopt them as phaseless texture features in this paper. Since these phaseless texture features solve the underlying problem included in the existing methods, they are robust to various changes such as spatial gaps and luminance changes. Then this indicates realization of the image completion from fewer training examples. Given a power spectrum, we enable the phase retrieval of a target patch including missing areas from its known intensities. This phase retrieval algorithm can be derived as a simplified version of a recently reported GrEedy Sparse PhAse Retrieval (GESPAR) algorithm [28]. Then the error converged through
where the discrete Fourier transform matrix. In this paper, vector/matrix transpose is denoted by a superscript ‘\(^{\top}\)’.

First, we decompose \(x\) into two elements as

\[
x = B_{\mu} \mu + B_{\kappa} \kappa,
\]

where \(\mu \in \mathbb{R}^{D}\) and \(\kappa \in \mathbb{R}^{D}\) are vectors respectively including unknown and known intensities within the target patch \(f\), where \(\mu\) is the estimation target in our method, and \(D_{\mu} + D_{\kappa} = D\). Furthermore, \(B_{\mu} \in \mathbb{R}^{D \times D_{\mu}}\) and \(B_{\kappa} \in \mathbb{R}^{D \times D_{\kappa}}\) are binary matrices which reconstruct \(x\) from the two decomposed elements. The estimation of the unknown intensity vector \(\mu\) from the original power spectrum \(y\) is equivalent to the phase retrieval of the target patch \(f\). Therefore, according to the idea of the GESPAR algorithm [28], the proposed method tries to reconstruct the original intensity vector \(x\), i.e., the unknown vector \(\mu\), by solving the following optimization problem:

\[
\min \left\{ g(x) = \sum_{i=1}^{D} (x^\top A_i x - y_i)^2 : x \in \mathbb{R}^{D} \right\}.
\]

Since \(\kappa\) in \(x\) is already known, the above problem can be rewritten as

\[
\min \left\{ g(\mu) = \sum_{i=1}^{D} \left( (B_{\mu} \mu + B_{\kappa} \kappa)^\top A_i (B_{\mu} \mu + B_{\kappa} \kappa) - y_i \right)^2 : \mu \in \mathbb{R}^{D_{\mu}} \right\}
\]

by using Eq. (1). In Eq. (3), \(g(\mu)\) is rewritten as

\[
\begin{align*}
 g(\mu) &= \sum_{i=1}^{D} \left( \mu^\top B_{\mu}^\top A_i B_{\mu} \mu + 2 \kappa^\top B_{\mu}^\top A_i B_{\kappa} \mu + (\kappa^\top B_{\mu}^\top A_i B_{\kappa} \kappa - y_i) \right)^2 \\
 &= \sum_{i=1}^{D} \left( \mu^\top C_i \mu + 2 d_i^\top \mu + e_i \right)^2 \\
 &= \sum_{i=1}^{D} |h_i(\mu)|^2,
\end{align*}
\]

where \(C_i\), \(d_i\), and \(e_i\) are respectively defined as follows:

\[
\begin{align*}
 C_i &= B_{\mu}^\top A_i B_{\mu}, \\
 d_i &= B_{\mu}^\top A_i B_{\kappa}, \\
 e_i &= \kappa^\top B_{\mu}^\top A_i B_{\kappa} - y_i.
\end{align*}
\]

Furthermore, \(h_i(\mu)\) is defined as

\[
h_i(\mu) = \mu^\top C_i \mu + 2 d_i^\top \mu + e_i.
\]

The proposed method estimates the optimal solution in Eq. (3) by the following iteration algorithm, where \(i\)th estimation result of \(\mu\) is denoted as \(\mu_{i-1}\). First, we approximate \(h_i(\mu)\) in Eq. (8) around \(\mu_{i-1}\) as follows:

\[
h_i(\mu) \approx h_i(\mu_{i-1}) + \nabla h_i(\mu_{i-1})^\top (\mu - \mu_{i-1}).
\]

Then the problem shown in Eq. (3) is rewritten as

\[
\min_{\mu} \sum_{i=1}^{D} \left( (C_i \mu_{i-1} + d_i)^\top \mu - (\mu_{i-1}^\top C_i \mu_{i-1} - e_i) \right)^2.
\]

Finally, Eq. (10) can be also rewritten as the following linear least-squares problem:

\[
\mu_i = \arg \min_{\mu} \left\| P_{i-1} \mu - q_{i-1} \right\|^2,
\]

where

- \(A_i = \text{Re}(F_i)^\top \text{Re}(F_i) + \text{Im}(F_i)^\top \text{Im}(F_i)\), and \(F_i\) is the \(i\)th row of
- \(\text{Re}(F_i)\) and \(\text{Im}(F_i)\) denote the real and imaginary parts of \(F_i\), respectively.

Fig. 1. Comparison between two kinds of textures features, “intensity values” and “power spectrum values” obtained from patches clipped from six test images shown in (a)–(f). The two graphs in (g) and (h) correspond to results of dimensionality reduction using the two kinds of textures features. These results are obtained on the basis of the benchmarking dimensionality reduction method, t-distributed Stochastic Neighbor Embedding (t-SNE) [31]. The dots shown in these graphs correspond to patches clipped from the six images. The color shown in each caption of (a)–(f) corresponds to the color shown in (g) and (h).

2. MISSING AREA RECONSTRUCTION ALGORITHM BASED ON PHASE RETRIEVAL

This section presents the missing area reconstruction algorithm based on the phase retrieval. The algorithm shown in this section is necessary for both deriving the new quality measure and estimating missing intensities in our exemplar-based image completion method shown in the following section (Sec. 3).

Given a target patch \(f\) including missing areas and its corresponding original power spectrum, the proposed method reconstructs the missing areas by retrieving the phase of \(f\). First, we define an original intensity vector of the target patch \(f\) and its original power spectrum as \(x \in \mathbb{R}^{D}\) and \(y \in \mathbb{R}^{D}\), respectively, where \(D\) is the number of pixels within the target patch. It should be noted that \(i\)th \((i = 1, 2, \ldots, D)\) element \(y_i\) of \(y\) is obtained by \(y_i = x^\top A_i x\), where \(A_i = \text{Re}(F_i)^\top \text{Re}(F_i) + \text{Im}(F_i)^\top \text{Im}(F_i)\), and \(F_i\) is the \(i\)th row of

\(\text{Re}(F_i)\) and \(\text{Im}(F_i)\) denote the real and imaginary parts of \(F_i\), respectively.
Table 1. Quantitative evaluation of image completion results using SSIM index.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>0.8225</td>
<td>0.8266</td>
<td>0.8397</td>
<td>0.8146</td>
<td>0.8137</td>
<td>0.8489</td>
</tr>
<tr>
<td>Image 2</td>
<td>0.7162</td>
<td>0.7388</td>
<td>0.7558</td>
<td>0.7522</td>
<td>0.7958</td>
<td>0.8240</td>
</tr>
<tr>
<td>Image 3</td>
<td>0.7910</td>
<td>0.7896</td>
<td>0.8155</td>
<td>0.8009</td>
<td>0.8250</td>
<td>0.8509</td>
</tr>
<tr>
<td>Image 4</td>
<td>0.7275</td>
<td>0.7179</td>
<td>0.7950</td>
<td>0.7341</td>
<td>0.8324</td>
<td>0.8534</td>
</tr>
<tr>
<td>Image 5</td>
<td>0.7379</td>
<td>0.7720</td>
<td>0.7232</td>
<td>0.7598</td>
<td>0.6674</td>
<td>0.7725</td>
</tr>
<tr>
<td>Average</td>
<td>0.7590</td>
<td>0.7690</td>
<td>0.7858</td>
<td>0.7723</td>
<td>0.7869</td>
<td>0.8299</td>
</tr>
<tr>
<td>Median</td>
<td>0.7379</td>
<td>0.7720</td>
<td>0.7950</td>
<td>0.7598</td>
<td>0.8137</td>
<td>0.8489</td>
</tr>
</tbody>
</table>

where

\[
\mathbf{p}_{n-1} = \begin{bmatrix}
2(C_n \mu_{n-1} + d_1) \\
2(C_n \mu_{n-1} + d_2) \\
\vdots \\
2(C_n \mu_{n-1} + d_N)
\end{bmatrix},
\mathbf{q}_{n-1} = \begin{bmatrix}
\mu_{n-1}^y C_n \mu_{n-1} - e_1 \\
\mu_{n-1}^y C_n \mu_{n-1} - e_2 \\
\vdots \\
\mu_{n-1}^y C_n \mu_{n-1} - e_N
\end{bmatrix},
\] (12)

Since the cost function of the above problem becomes a quadratic form of \( \mu \), its optimal solution can be obtained.

In this way, the proposed method estimates the missing intensities \( \mu \) within the target patch \( f \), and the phase retrieval of \( f \) also becomes feasible, simultaneously. The algorithm shown in this section is a simplified version of the GESPAR algorithm [28], whose image constraint is the known intensities \( x \) within the target patch \( f \). Since we do not have to estimate the support set, our phase retrieval problem can be simplified as shown in the above explanation.

3. IMAGE COMPLETION VIA NEW QUALITY MEASURE BASED ON PHASELESS TEXTURE FEATURES

This section shows the image completion method using the new quality measure based on the phaseless texture features. First, we clip known training patches \( f_n (n = 1, 2, \cdots , N; N \) being the number of clipped patches) from the target image in the same interval, and the power spectrum \( y_s \) is calculated as the texture features for each \( f_n \). Furthermore, a target patch \( f \) including missing pixels is selected from the target image, and its missing intensities are estimated by procedures shown below. Note that the selection of the target patch, i.e., the determination of the image completion order, is performed on the basis of the well known patch priority proposed by Criminisi et al. [4].

The proposed method tries to select the best matched patch \( f_n \), whose power spectrum \( y_{n,\text{est}} \) is the most similar to that of the target patch \( f \), among \( f_n (n = 1, 2, \cdots , N) \). It should be noted that since the target patch \( f \) includes missing pixels, we cannot find the best matched patch by directly comparing the difference of the power spectrum. Therefore, we newly derive the alternative measure. Specifically, the proposed method regards \( y_s \) of each known training patch \( f_n \) as \( y \) and performs the optimization shown in Eq. (3) of the previous section. Then the converged result of \( g(\mu) \) is obtained as the new quality measure \( g_s \) (\( n = 1, 2, \cdots , N \)). This measure is the minimum square error of the power spectrum caused by retrieving the phase of the target patch \( f \) from \( y_s \). Since we regard \( y_s \) as the texture features, the new criterion \( g_s \) becomes the minimum distance between the texture features of the target patch \( f \) and those of each known training patch \( f_n \). Therefore, by finding the patch minimizing \( g_s \), the selection of the best matched patch \( f_n \) becomes feasible. Finally, from \( y_{\text{est}} \) of the best matched patch \( f_n \), we can estimate the missing intensities within the target patch \( f \) by retrieving its phase based on the algorithm shown in the previous section.

In this way, the proposed method realizes the exemplar-based image completion by using the new quality measure based on the phaseless texture features. It has been reported in [2,23,24] that compared to intensity values, phaseless features such as Fourier transform magnitude and power spectrum successfully represent texture characteristics since they are robust to various changes such as spatial gaps and luminance changes. Therefore, the new quality measure derived in our method can successfully reflect the difference of textures. Furthermore, based on the phase retrieval algorithm, the proposed method realizes the estimation of the missing intensities from the phaseless texture features.

4. EXPERIMENTAL RESULTS

In this section, experimental results are shown to verify the performance of the proposed method. As shown in Fig. 2, we added missing areas to five test images. For the target images including missing areas, we performed the image completion based on the proposed method and some comparative methods [4–6, 8, 13]. Since the method in [4] performs the exemplar-based image completion directly using intensity values, it corresponds to the benchmarking method removing the novel points in our method. The method in [5] improves the performance of image representation and patch priority determination. Furthermore, we introduced the methods in [6], [8] and [13] for the comparison since they are state-of-the-art methods. All of these comparative methods directly use intensity values for the image completion, and they are suitable for the comparison of our method. In this experiment, we used gray scale images, and the patch size was set to 15 x 15 pixels in the exemplar-based methods, where the clipping interval of known training patches was set to the half size of the patches. We adopted such difficult conditions for making the difference between the performance of our method and that of the comparative methods clearer.

As shown in Fig. 2, the proposed method realizes more accurate image completion compared to the previously reported methods. If sufficient number of training examples cannot be obtained, it becomes difficult for the comparative methods to accurately estimate the missing intensities. On the other hand, the proposed method can successfully grasp the whole texture characteristics within the target image from few training examples by newly adopting the phaseless texture features. Furthermore, by using the new quality measure based on the phaseless texture features, our method can not only find the best matched patches but also estimate the missing intensities through the phase retrieval algorithm, accurately.

Furthermore, we show results of quantitative evaluation. In this experiment, we calculated the SSIM index [19], which is one of the most representative image quality measures, for the image comple-
Fig. 2. Image completion results obtained by the comparative methods [4–6, 8, 13] and the proposed method. These five images correspond to Images 1–5 of Table 1. The comparative methods used for obtaining the results in Images 1–5 are [4], [5], [6], [8] and [13], respectively. The sizes of Images 1–5 are 480 × 360 pixels, 480 × 360 pixels, 640 × 480 pixels, 640 × 480 pixels and 480 × 360 pixels, respectively. All of these five test images are 8-bit gray levels. The percentages of missing areas are 8.9%, 10.7%, 5.9%, 5.5% and 11.3% in Images 1–5, respectively.

5. CONCLUSIONS

This paper has presented the exemplar-based image completion via the new quality measure based on the phaseless texture features. The proposed method newly derives the new quality measure, which represents the minimum distance between the phaseless texture features, based on the phase retrieval algorithm. This measure enables the selection of the best matched patch even if the target patch includes missing pixels. Furthermore, the proposed method estimates the missing intensities by retrieving the phase of the target patch from the phaseless texture features of the best matched patch. From the experimental results, it is verified that our method achieves the successful image completion and outperforms the state-of-the-art methods.
6. REFERENCES


