Interactive Image Inpainting of Large-Scale Missing Region

IRAWATI NURMALA SARI, EMIKO HORIKAWA, AND WEIWEI DU (Member, IEEE)
Department of Information Science, Kyoto Institute of Technology, Kyoto 606-8585, Japan
Corresponding author: Weiwei Du (duweiwei@kit.ac.jp)

ABSTRACT Image inpainting is a challenging reconstruction of the damaged image in photography, especially for more valued artwork than before. The damages are mostly caused by scratches and worn out, so they cannot be easily fixed physically. Thus, many scientists proposed sophisticated methods for restoring the damaged image into a new one similar to an original image. However, these methods have not solved the problem effectively if the missing region is large. In this paper, we focus on how to restore a large missing region in image inpainting. This algorithm is composed of two steps: structure propagation and color propagation. In structure propagation, we segment a large region (non-homogeneous) into several small regions (homogeneous) based on the salient structure of missing region. Then, we applied a simple pixel-based inpainting method called the Fast Marching Method (FMM) to fill in the missing homogeneous regions by color propagation. In the experimental section, we applied several kinds of missing regions, such as irregular and regular missing regions, in large sizes. The results show that our proposed method performs well in various conditions.

INDEX TERMS Image inpainting, user interaction, structure propagation, color propagation, fast marching method (FMM).

I. INTRODUCTION

Many technologies preserved images and artworks a long time ago. The spots and scratches caused many damages that are not easily removed. Scientists convert them into digital images and try to restore them to original images. Image inpainting is one of the techniques to restore damaged images. Some previous methods worked well if the damaged region is small in size. However, if the size of a missing region is too large, it will cause a blurry result.

In general, a missing region can be filled in by using several interactive operations manually, such as Adobe Photoshop. Since interactive tools require accurate work driven by a skilled professional artist to complete the restored image smoothly, it will be a high challenge for the beginner.

Qureshi et al. [32] explained three kinds of traditional image inpainting methods: Partial Differential Equation (PDE)-based, pixel-based, and exemplar-based. The PDE-based method used texture information in isotropic, anisotropic, linear, or nonlinear directions. Local structure information is diffused from known to unknown regions. This method is employed for extended narrow areas such as lines and cracks because it provides more reliable confidence accuracy. The pixel-based method extracts information using pixels around an unknown region and then fills in the damaged part. However, it does not apply to large unknown texture regions because it only works well in the surrounding boundary unknown area so that a blurry area will exist in the center. In contrast to PDE-based and pixel-based methods, the exemplar-based method will match the similarity between known and unknown patches. This method is suitable for restoring the large unknown texture region. Unfortunately, some unwanted artifacts from disconnected edges and inconsistent texture areas will be visible in a particular condition.

Bertalmio et al. [1], and Oliveira et al. [3] proposed a PDE-based image inpainting method for pre-processing propagated information to an unknown region from its surrounding pixels in the isophotes direction. They used isophotes to find the minimum change by tracing contour. After that, the information of known region will be spread to an unknown region for maintaining continuity of linear sketch.

Telea [5] employed a pixel-based method, which transmitted the pixel information of surrounding unknown regions along the image gradient to improve the accuracy value. It is a fast and straightforward method for restoring small and homogeneous regions.
Reconstruction in a large missing region is more challenging than in a small region since it should deal with non-homogeneous image information. The exemplar-based method can be an alternative way of reconstructing large missing regions. Drori et al. [2] estimated a large missing region and filled in it using robust fragments. A fragment is selected from the most similar and constant examples of the known region in different radius scales. Therefore, performance depends on the total of the available fragments. However, it will result in a blurred artifact if the missing region is on object boundaries. Based on the concept of Drori et al. [2], Criminisi et al. [4] proposed an improved version. The gradient of an unknown region assigned the order of inpainting to reduce the running time.

Learning-based methods approved techniques to solve image inpainting, which consider two aspects: image generation and structured prediction. In image generation, Nazeri et al. [7] used an edge generator and image completion. Edge information is an essential part of first object initialization. It is followed by image completion to analyze the unknown region. Structured prediction defined the model of structured similarity between known and unknown regions. Yang et al. [6] proposed a multi-scale neural patch synthesis that retains the structure and matching patches with the most similar features that produced the high-frequency of unknown regions.

This paper designs an unsupervised and fast novel algorithm for solving image inpainting with a large missing region. First, our method segments a large missing region (non-homogeneous) into some small missing regions (homogeneous) by using structure propagation (an initial process of Sun et al. [8]). Then, we apply a simple pixel-based method, which is FMM [5] to fill in each pixel intensity of unknown homogeneous regions. Those two existing methods are not simply used. The basic idea of our approach is how to find homogenous regions. There are three points of our contributions:

1) An interactive and simple image inpainting technique based on the traditional method.
2) Segmentation of the large missing region (non-homogeneous) into some small missing regions (homogeneous).
3) Completion of large missing regions using a fast method that reduces compilation time.

This paper is organized as follows. After the explanation of related works (Section II), we detail our proposed method (Section III). Next, we show the experimental results (Section IV) and discuss about implementation (Section V). Finally, we present the conclusion and interesting future works (Section VI).

II. RELATED WORK

Most researchers used exemplar-based methods to fill in large missing regions because it effectively generated a restored region by sampling and copying information from known areas. This method used a preference function, which selected a patch of an unknown region first and then found a patch with similar details in known regions. Exemplar-based methods can only be applied to images that have a simple structure and texture. Criminisi et al. [4] was the first method to propose patch-based or exemplar-based for recovering large missing regions by simultaneously reconstructing structure and texture in missing regions. This approach employed the best-first filling strategy that depended on each patch, which has a substantial edge and was surrounded by high-confidence pixels. Then it was selected as the priority choice. The highest patch values of known region will be propagated via structure and texture diffusion to fill in the missing patch with similar information. However, this method cannot restore large missing regions based on structure and texture.

Many methods proposed to improve Criminisi et al. [4] in restoring large missing regions. The improved approaches were categorized into two groups. The first group focused on improving texture reconstruction [10]–[12], and [13]. Cheng et al. [10] generalized a function of [4] to provide more robust performance. It selected the higher weight of data and value used for the best matching patch from known regions. Hesabi and Mahdavi-Amiri [13] presented a modified patch, which measured the similarities between fill-front and candidate patches to analyze the information of surrounding pixels. However, Cheng et al. [10], and Hesabi and Mahdavi-Amiri [13] could not handle the complex texture of unknown regions. Besides, the second group to restore the structure from known region to unknown region more accurately such as [12], [14]–[17], and [18]. Anupam et al. [12] proposed a technique to segment a large missing region into small regions and then search the optimal patches in the known region based on Minimum-mean Squared Error (MSE).

Sun et al. [8] presented structure propagation for optimization problem by accomplishing the consistency structure. It had a similar proposed technique with Anupam et al. [12] to segment large missing regions into several small missing regions. After structure propagation was completed, it used an exemplar-based method [4] to fill in small unknown regions. This method had a superior idea of solving the inpainting problem by segmenting a large missing region and using the exemplar-based method to fill in the remaining areas. Nonetheless, the artifact area will be visible caused by an exemplar-based method that could not obtain the optimal fit patches in the complicated texture and large size between known and unknown regions.

Efros and Leung [9] employed a non-parametric model based on the assumption of spatial locality. This method filled in missing regions, pixel by pixel. All pixels in a rectangle window around a defined pixel are used as samples. This method first finds all neighbor samples in a known region similar to an unknown region for synthesizing an unknown pixel. Then, the center of a sample will be the newly synthesized pixel. Moreover, Telea [5] also presented a simple pixel-based and fast inpainting method. This method propagated pixel gradient from the known region to fill in unknown areas. It considered some factors such as directional component,
More recently, learning-based methods solved inpainting problems and then obtained acceptable results such as [19]–[24], and [25]. Yu et al. [20], and Jo and Park [22] generated images with an interactive technique that can apply user sketches as direction to produce more user-desired results. These methods have shown that systems were excellent for restoring large regions in one pass and obtained realistic results. Zheng et al. [24] and Zeng et al. [23] trained models both regular and irregular regions. However, training models for irregular regions take about twice the time compared to regular regions.

All of the learning-based methods can work to restore large missing regions. Nonetheless, training models cannot effectively transmit information to missing inner areas and then produces a large blurry result between an object and its boundary edge. The learning-based method also needs more time to train a model than an unsupervised learning method, faster until it obtains the final restored image.

The goal of our proposed method is to restore a large missing region by generating a fast and straightforward technique based on an unsupervised method, as shown in Figure 1. We improved the result of image inpainting from previous methods that only employ structure propagation [8] or FMM [5]. Our work purposes a user interaction to segment a large missing region (non-homogeneous) into several small missing regions by structure propagation [8]. Because small missing regions are homogeneous, we employ FMM [5] for filling in unknown regions based on their surrounding pixels information. Therefore, our method performs well compared with other methods, even in various conditions such as regular and irregular missing regions. However, our approach has a failure case when something is missing at the peak of a curved object. It only depends on sample patch information along a user-defined curve of the known region. So, our method cannot work well in a missing region with significantly different structures, as seen in Figure 2.

III. OUR PROPOSED METHOD

In this section, we introduce the concept of our proposed method: (1) Description of the algorithm in a simple flowchart, (2) Segmentation of the large missing region into several small missing regions by structure propagation [8], (3) Explanation to fill the small regions based on the pixel-based method using FMM [5], and (4) Detailed explanation of our proposed method.

A. FLOWCHART

Flowchart 3 describes a short explanation of our method. In structure propagation, input data is an image with a large missing region.
missing region. Users will be asked to draw the curve or line manually to detect salient structures in a large missing region. The method will segment that region based on a user-defined curve into several small missing regions using structure propagation.

The output of structure propagation will be input for color propagation. In color propagation, small missing regions will be filled in by spreading the surrounding pixel information from known to unknown regions using FMM. For more details, this method will be explained in section III-D.

B. STRUCTURE PROPAGATION

The problem of structure propagation is how to estimate a structure through a user-defined curve \( C \) in an unknown region by spreading patches from the known region. According to the theory of structure propagation [8], this process calculates the similarity of both unknown and known patches using global optimization. It also considers consistency between two adjacent patches along a user-defined curve based on their overlapped regions by Sum of normalized Squared Difference (SSD).

Dynamic programming [26] computes minimal cost among known patches. Afterward, this method develops the minimum cumulative cost from center points of patch \( x_{1} \) to \( x_{k} \) through the unknown region curve for all possible to obtain optimal unknown labels [8]:

\[
M_{1-1,1} (x_{i}) = \min_{x_{i-1}} \{ E_{1} (x_{i-1}) + E_{2} (x_{i-1}, x_{i}) + M_{1-2,1-1} \}
\]

(1)

where \( M \) is total number of patches, \( M_{0,1} = 0 \) and \( L \) is current center point of patch.

\( E_{1} \) calculates the similarity patches of the unknown to known regions which uses this equation:

\[
E_{1} (x_{i}) = w_{s} \cdot E_{S} (x_{i}) + w_{l} \cdot E_{L} (x_{i}), \quad \text{where } w_{s} \text{ and } w_{l} \text{ are the related weights.}
\]

\( E_{S} (x_{i}) \) represents as the structure similarity on the curve \( C \) at each point of an unknown region \( p_{i} \), \( E_{L} (x_{i}) \) describes the similarity between patches on the boundary of unknown region \( \Omega \) and known patches \( \Gamma - \Omega \). \( E_{2} (x_{i}, x_{j}) \) determines the consistency between two adjacent patches. \( E_{1} (x_{i}) \) and \( E_{2} (x_{i}, x_{j}) \) are calculated as Sum of normalized Squared Differences (SSD) between their overlapped regions.

Completely, the converged optimal label of patch \( L \) is computed by:

\[
x_{L} = \arg \min_{x_{L}} \{ E_{1} (x_{L}) + M_{L-1,1} \}
\]

(2)

C. FAST MARCHING METHOD (FMM)

Fast Marching Method is an analytical method for solving boundary problems that is proposed by Sethian [27]. It uses a numerical technique for computing the position of propagating values. Based on Telea [5], FMM can also solve the pixel-based problem of image inpainting. This method uses travel time \( \sqrt{T} \) and distance map \( T \) of unknown \( \Omega \) pixel to boundary \( \partial \Omega \).

Borgefor [28], [29] and Meijster et al. [30] purposed distance transform (DT) to compute the distance map \( T \). It was also similar to FMM. However, FMM has an advantage that provided a narrow band such as a stack for dividing between known and unknown regions. Moreover, it will decide which next pixel should be inpainted. The pixel of the narrow band is the inpainting boundary \( \partial \Omega \). Every pixel has some \( T \) value, pixel value \( I \), and flag \( f \). The flag has three kinds of values such as BAND, KNOWN, and UNKNOWN.

Since FMM-based image inpainting is a method to solve the boundary problem of missing region, it could lead to more robust results due to propagation from the closest known pixel and determine which the next pixel should be processed. FMM calculates the weighting function \( w (p, q) \) for determining how to fill in the unknown pixel \( p \) by propagating the pixel value of the known region \( q \). This function considers about directional component \( \text{dir} (p, q) \), geometric distance component \( \text{dst} (p, q) \), and level set distance component \( \text{lev} (p, q) \).

D. THE PROPOSED METHOD

In the following section, we will introduce more details of our proposed method, as shown in Algorithm 1. Given an image \( I \) that has a large missing region. Firstly, we segment a large missing region (non-homogeneous) into several small missing regions (homogeneous) using structure propagation [8]. Since small missing regions are homogeneous regions, we use a simple image inpainting method such as FMM [5] to repair them and then produce remarkable results.
Algorithm 1: Our Proposed Method

**Structure Propagation**

**Data:** An image with the large missing region (non-homogeneous).

**Result:** An image with the small missing regions (homogeneous).

- Initialize a user-defined curve $C$ from known to unknown region $\Omega$.
- Spread a set of patches $Y$ from known to the unknown region along curve $C$ on each center point of patches $x_i$.
- Determine optimal label of patches $Y(x_i)$ by using Dynamic Programming.

**Fast Marching Method (FMM)**

**Data:** An image with small missing regions (homogeneous).

**Result:** A restored image.

**Initialization**

- $T = 0$ for pixel in the boundary of unknown region. Otherwise, $T = \text{INF}$.
- Flag $f$ to all pixels of image (KNOWN, BAND, and UNKNOWN).
- Boundary pixels are stored to narrow band as BAND and be sorted in ascending order of $T$ values.

**Propagation**

```plaintext
while narrow band != empty do
    Decide the smallest $T$ value of narrow band
    Change the flag of $(x, y)$ to be KNOWN for $(a, b)$ in $(x - 1, y), (x, y - 1), (x + 1, y), (x, y + 1)$
    if (flag $(a, b) = \text{UNKNOWN}$) then
        Determine flag $(a, b)$ to be BAND;
        Propagate neighbor pixels of a known region inside radius $r$;
        Update a narrow band by inserting $(a, b)$ that has minimum $T$ value;
    else
        Continue;
    end
end
```

IV. EXPERIMENTS

Applied dataset of experiments approved that our approach is better than other image inpainting methods. It is available at https://github.com/openimages/dataset and the size of each image is roughly $500 \times 500$ pixels.

The experiment employs two kinds of missing regions [19]: regular (rectangular shaped holes) and irregular (random shaped holes). We apply these missing regions to images that have line and curve objects. Sizes of regular region are $50 \times 50$, $100 \times 100$, and $150 \times 150$ (height $\times$ width) in Figure 6. Otherwise, irregular regions are 30, 50, and 75 pixels (width) in Figure 7.

The process of color propagation will work simultaneously for every small missing region. FMM fills in for each small missing region, pixel-by-pixel. Therefore, it outperforms the result of previous methods. Given pixel $p$ on the boundary $\partial \Omega$ of the unknown region $\Omega$ as shown in Figure 5(a). For filling in $p$, we use the information of known pixels inside radius area $q(p)$. This technique uses the propagation process based on the weighting function, assisted by four neighboring pixels, as shown in Figure 5(b). Suppose that $p_1$ is an initial pixel position of boundary in an unknown region. By determining the order of the minimum distance between $p_1$ and its adjacent pixels $p_i$ in boundary, it proposes to fill in the closest unknown pixel. This method works for all boundary pixels iteratively until there are no more unknown pixels.

The experiment employs two kinds of missing regions [19]: regular (rectangular shaped holes) and irregular (random shaped holes). We apply these missing regions to images that have line and curve objects. Sizes of regular region are $50 \times 50$, $100 \times 100$, and $150 \times 150$ (height $\times$ width) in Figure 6. Otherwise, irregular regions are 30, 50, and 75 pixels (width) in Figure 7.

Result of our method will be compared with other previous methods: original FMM [5], structure propagation and texture propagation (SP-TP) [8], and EdgeConnect [7].

![Figure 5. Color propagation](image)

**FIGURE 5.** Color propagation : (a) Information of known pixels inside radius $q_p$. (b) Using four neighbor pixels to assist fill in pixel $p$. 

![Figure 6.](image)

**FIGURE 6.** Input images : regular missing regions in different sizes (height $\times$ width). (a)-(c) Line objects, (d)-(f) Curve objects.

Result of our method will be compared with other previous methods: original FMM [5], structure propagation and texture propagation (SP-TP) [8], and EdgeConnect [7].
Moreover, we consider visual and quantitative comparisons for the validity of measurement.

A. VISUAL COMPARISON

As shown in Figure 9, 10, 11, and 12, original FMM [5] generates some artifacts and disconnected salient structure of missing region. FMM is more difficult to recover missing regions in large size because it only propagates the information from the nearest neighbor pixels of missing regions and is more challenging than when only a small missing region.

Moreover, we compare our results with those obtained from SP-TP [8]. The results show that our method is slightly better than the SP-TP result. SP-TP also uses structure propagation to restore the missing region’s complex salient structure, similar to our method. However, SP-TP uses exemplar-based propagation for filling in missing regions, which depends on their size. If a missing part is large, this method will take a large exemplar from known regions to restore it. Therefore, this method produces failed propagation when the missing region’s exemplar does not match or exist in known regions.

Afterward, we also compare with a deep learning method such as EdgeConnect [7]. This method produces a large blurry region and disconnected salient structure of missing region because the training model cannot effectively propagate the information to a missing inner region. The edge generating model sometimes fails to accurately depict the boundaries in the large missing region.

B. QUANTITATIVE COMPARISON

Many evaluation metrics can be used to compute the accuracy of image processing outcomes. However, image inpainting cannot use simple traditional metrics such as Mean Squared Error (MSE) and Peak Signal to Noise Ratio (PSNR) because they are not well correlated with perceptual quality assessment [31], [32].

To evaluate our method, we use three types of Image Inpainting Quality Assessment (IQQA) metrics: Perceptual SSIM [34], Perceptual MSSSIM [35], and Perceptual FSIM [33]. The perceptual metric includes two evaluations: (1) objective metric based on the equation, (2) subjective assessment that is human judgments of perceived quality images. Moreover, the objective metric should be statistically consistent with subjective evaluation.

In our experiment, we have assisted by six observers with the normal or corrected vision to judge the subjective evaluation. Observers rate each image at a time and then give their score. Optional scores which could be selected by observers are “not noticeable”, “noticeable”, “not acceptable”, “acceptable”, “salient”, “very salient”, “good”, or “perfect”. The score starts from 1 to 8, respectively.

1) PERCEPTUAL SSIM COMPARISON

Structural Similarity Index Measure (SSIM) evaluates the similarities of images from three aspects: brightness, contrast, and structure. Range of SSIM value is [0, 1]. The higher SSIM value represents more similarities between the two images.

\[
SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1) + (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}
\]

where \(\alpha\), \(\beta\), and \(\gamma\) are parameters to define brightness, contrast, and structure. If we set \(\alpha = \beta = \gamma = 1\), then SSIM is given by

\[
SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1) + (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}
\]

where \(\mu_x\) and \(\mu_y\) are average grayscale values of image X and Y, \(\sigma_x\) and \(\sigma_y\) represent standard deviations of image X and Y. \(\sigma_{xy}\) is the covariance between image X and Y. \(C_1\) and \(C_2\) are small constants to prevent the denominator from being zero.

2) PERCEPTUAL MSSSIM COMPARISON

Multi-scale SSIM (MSSSIM) constructs an image pyramid between two images and then calculates the SSIM at each layer separately. MSSSIM evaluation is obtained by combining the calculation on a different scale by

\[
SSIM(x, y) = \left[ I_M(x, y) \right]^{2\gamma} \prod_{j=1}^{M} \left[ (c_j(x, y))^\beta \cdot \left[ s_j(x, y) \right]^{\gamma} \right]
\]

where \(j\) is scale index, and \(M\) is the total of scales. This equation is similar to Equation (3), which \(\alpha_M\), \(\beta_j\), and \(\gamma_j\) are importance of three components (brightness, contrast, and structure).

3) PERCEPTUAL FSIM COMPARISON

Feature Similarity Index Measure (FSIM) proposes Human Visual System (HVS) that perceives images based on its low-level features. FSIM uses two kinds of elements: high Phase Congruency (PC) and Gradient magnitude (G). \(S_{PC}\) is used as a primary feature to compute the contrast invariant and the

\[
\text{FIGURE 7. Input images : irregular missing regions in different sizes (width). (a)-(c) Line objects, (d)-(f) Curve objects.}
\]
TABLE 1. Objective metrics.

| Mask  | Object | Size | SSIM | MSSSIM | FSIM |
|-------|--------|------|------|--------|------|
|       |        |      | FMM  | SP-TP  | Edge | Ours | FMM  | SP-TP  | Edge | Ours | FMM  | SP-TP  | Edge | Ours |
| Irregular | Line   | 30   | 0.981| 0.992 | 0.97 | 0.992| 0.97 | 0.995 | 0.985 | 0.994| 0.979| 0.996 | 0.952 | 0.995 |
|        | Curve  | 75   | 0.996| 0.973 | 0.94 | 0.979| 0.963| 0.975 | 0.953 | 0.984| 0.956| 0.973 | 0.953 | 0.984 |
| Regular | Line   | 50x50| 0.997| 0.998 | 0.989| 0.998| 0.997| 0.999 | 0.994 | 0.998| 0.997| 0.999 | 0.995 | 0.999 |
|        |       | 100x100| 0.982| 0.994 | 0.974| 0.994| 0.977| 0.996 | 0.979 | 0.996| 0.983| 0.995 | 0.949 | 0.995 |
|        | Curve  | 150x150| 0.997| 0.984 | 0.949| 0.988| 0.959| 0.986 | 0.936 | 0.99 | 0.966| 0.988 | 0.935 | 0.992 |
|        |       | 150x150| 0.95 | 0.945 | 0.937| 0.958| 0.955| 0.953 | 0.953 | 0.962| 0.964| 0.956 | 0.948 | 0.966 |

4) OUR PERFORMANCE

The total number of evaluated images is 192 images in different missing region sizes, missing region types, and object types. We employ mean evaluations in the similar categories in Table 1, 2, 3, and 4.

As shown in Table 1, the accuracy of our proposed method is higher than other methods when the missing region is larger, although using different objective metrics such as SSIM, MSSSIM, and FSIM. Meanwhile, when a missing region is smaller, the score of our method is similar to SP-TP because a small missing region can be restored easily, even if the propagation applies the pixel-based or exemplar-based method. We also perform the ranking between objective and subjective metrics in Table 2 consistently.

TABLE 2. Ranking of objective and subjective metrics.

| Mask  | Object | Size | Metric | SSIM | MSSSIM | FSIM |
|-------|--------|------|--------|------|--------|------|
|       |        |      | FMM  | SP-TP  | Edge | Ours | FMM  | SP-TP  | Edge | Ours | FMM  | SP-TP  | Edge | Ours |
| Irregular | Line   | 30   | 3   | 1 | 4 | 2 | 4 | 1 | 3 | 2 | 4 | 1 | 3 | 2 |
|        | Curve  | 75   | 3 | 2 | 4 | 1 | 4 | 2 | 3 | 1 | 4 | 2 | 3 | 1 |
| Regular | Line   | 50x50| 3 | 2 | 4 | 1 | 4 | 2 | 3 | 1 | 4 | 2 | 3 | 1 |
|        |       | 100x100| 3 | 2 | 4 | 1 | 4 | 2 | 3 | 1 | 4 | 2 | 3 | 1 |
|        | Curve  | 150x150| 3 | 2 | 4 | 1 | 4 | 2 | 3 | 1 | 4 | 2 | 3 | 1 |
|        |       | 150x150| 3 | 2 | 4 | 1 | 4 | 2 | 3 | 1 | 4 | 2 | 3 | 1 |

4) OUR PERFORMANCE

The total number of evaluated images is 192 images in different missing region sizes, missing region types, and object types. We employ mean evaluations in the similar categories in Table 1, 2, 3, and 4.

As shown in Table 1, the accuracy of our proposed method is higher than other methods when the missing region is larger, although using different objective metrics such as SSIM, MSSSIM, and FSIM. Meanwhile, when a missing region is smaller, the score of our method is similar to SP-TP because a small missing region can be restored easily, even if the propagation applies the pixel-based or exemplar-based method. We also perform the ranking between objective and subjective metrics in Table 2 consistently.

TABLE 3. Comparison between line and curve objects.

| Mask  | Object | Size | SSIM | MSSSIM | FSIM | SROCC | MEAN SROCC |
|-------|--------|------|------|--------|------|-------|------------|
|       |        |      | FMM  | SP-TP  | Edge | Ours | FMM  | SP-TP  | Edge | Ours | FMM  | SP-TP  | Edge | Ours |
| Irreg | Line   | 30   | 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
|       | Curve  | 75   | 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
|       | Line   | 50x50| 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
|       | Curve  | 150x150| 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
| Reg   | Line   | 30   | 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
|       | Curve  | 75   | 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
|       | Line   | 50x50| 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
|       | Curve  | 150x150| 0.8  | 0.8  | 1 | 0.86 | 0.86 | 0.93 |
TABLE 4. Comparison between irregular and regular missing regions.

| Mask   | MEAN SROCC FSIM | SSIM | MSSIIM |
|--------|-----------------|------|--------|
| Irregular | 0.76            | 0.73 | 0.73  |
| Regular  | 0.965           | 0.93 | 0.63  |

Spearman Rank Order Correlation Coefficient (SROCC) is operated to measure the correlation between two variables’ ranks (objective and subjective metric). The range is from $-1$ to $+1$. Suppose a value is close to $+1$, which represents a strong correlation. On the other hand, the $-1$ value has a strong disagreement. SROCC value is zero when there is no correlation between both variables. As shown in Table 3, our method is close to $+1$, explaining the strong correlation between objective and subjective metrics. The highest values are highlighted by boldface. It proves that our approach has better performance to recover large missing regions in the line object than a curved object because structure propagation does not solve a salient dynamic structure with a changed shape frequently. Meanwhile, Table 4 shows that a regular missing region is better than an irregular missing region because it has a random and more complicated pattern. This case generates a significant failure when spreading the patch from a known to an unknown region, as shown in Figure 8.

FIGURE 8. Comparison between irregular and regular missing regions in similar images.

C. TIME COMPARISON

In the following section, we compute the average compilation time of similar image categories. As shown in Table 5, Edge-Connect has the fastest running time. It is a learning-based method that the testing time does not take a long time to fill in the missing region. However, training time takes around two days.

SP-TP matches the exemplar between unknown and known regions fastly. It can run straightforward and fast to solve the uncomplicated patch. FMM needs running time longer than the exemplar-based method because it recovers pixel by pixel based on the unknown region’s surrounding information. Furthermore, our computational time is faster than the FMM but slower than SP-TP. It is caused by operating structure propagation and color propagation based on SP-TP and FMM. However, our visual and quantitative results outperform than theirs.

V. DISCUSSION

In the experiment section, we apply an open-access dataset that contains many image categories. We employ missing regions between foreground and background with different colors and textures. Therefore, if there is a large missing region, it will be difficult to restore the center of a missing region because of less information in its known surrounding area. The deep learning method cannot work effectively for this problem because it will produce a large blurry in the restored missing region’s center area. We employ the traditional approach that uses structure propagation to segment a large missing region into many homogeneous regions with almost similar color and texture to solve this problem. After that, we operate an easy, simple, and fast method for completing all of the missing homogeneous regions simultaneously using a pixel-based method, FMM. Because it is a fast and straightforward method, our approach has a shorter running time than deep learning methods.

Moreover, we propose an unsupervised or traditional approach rather than a deep learning method because a traditional approach is faster for running time and more straightforward for implementation. The deep learning method needs a training process that takes long times, at least 2-3 days, to create a model from many dataset images for testing an input image. Sometimes the high specification of computer is required. However, our method only takes 3-32 seconds to run the program. Besides, it does not need many dataset images to create a training model and high computer specification to restore missing regions. Results also show our traditional approach is better than image inpainting based on deep learning methods.

Most comparison results display that our method is better than the original FMM, starting from small to large missing regions. FMM is appropriate to recover the small area because it works pixel-by-pixel. It has a significant failure when propagating information in a large missing region. SP-TP works to restore a large missing region using the
I. N. Sari et al.: Interactive Image Inpainting of Large-Scale Missing Region

FIGURE 9. First to third row: irregular missing regions with different sizes (width = 30, 50, and 75 pixels, respectively).

FIGURE 10. First to third row: regular missing regions with different sizes (height × weight) = 50 × 50, 100 × 100, and 150 × 150, respectively.

exemplar-based method. However, if there is no proper and fit exemplar between unknown and known regions, it will be unworkable to synthesize the desired structure and pixel information.

Considering image inpainting in different missing regions, we show that a regular missing region is better than an irregular missing region. It has a more complicated shape and generates complexity when structure propagation is applied.
FIGURE 11. Irregular missing regions with different sizes (width = 30, 50, and 75 pixels, respectively).
FIGURE 12. Regular missing regions with different sizes (height × width = 50 × 50, 100 × 100, and 150 × 150, respectively).
Moreover, we also compare line and curve objects of images. Our proposed method demonstrates that an image with the missing region in a line object would result better than a curve object. It is produced by structure propagation does not consider dynamic propagation that has a changed structure regularly.

Since structure propagation is a method to restore large missing regions, it needs a user-defined curve from known to unknown regions to determine the salient structure of a missing region. In this case, image inpainting works well and obtains better results properly. However, structure propagation that employs user interaction will take more time to process. The user is required to draw the curve precisely. If a missing region is large, a user needs to analyze more structure areas in the known region. It is used for initializing samples and then propagate to an unknown region.

VI. CONCLUSION AND FUTURE WORKS

This paper has presented an interactive approach for restoring the damaged image with a large missing region using a fast and straightforward method. We operate structure propagation to segment a large missing region into small homogenous regions. Due to the homogeneous area, we employ a fast and straightforward method for color propagation by filling in the small missing regions using FMM. Our method effectively outperforms in various large missing regions such as regular and irregular shapes. Moreover, most comparisons show that our proposed approach is better than previous unsupervised methods, even in the learning-based techniques.

In the future, we plan to implement an automatic process for structure propagation that demonstrates a better result than user interaction. We also consider an uncomplicated texture propagation to produce a more natural-looking result. Furthermore, we will apply a complex content image as input and present interesting opportunities.

REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” in Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn. SIGGRAPH, 2000, pp. 417–424. [Online]. Available: https://www.csie.ntu.edu.tw/~bbehis/CS474/StudentPaperPresentations/imageinpainting.pdf
[2] I. Drori, D. Cohen-Or, and H. Yeshurun, “Fragment-based image completion,” ACM Trans. Graph., vol. 22, no. 3, pp. 305–312, Jul. 2003.
[3] M. Oliveira, B. Bowen, R. McKenna, and Y. Chang, “Fast digital image inpainting,” in Proc. Int. Conf. Vis., Imag. Image Process., Sep. 2001, pp. 106–107. [Online]. Available: https://www.inf.ufrgs.br/~oliveira/pubs_files/inpainting.pdf
[4] A. Criminisi, P. Perez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” IEEE Trans. Image Process., vol. 13, no. 9, pp. 1200–1212, Sep. 2004.
[5] A. Telea, “An image inpainting technique based on the fast marching method,” J. Graph., GPU, Game Tools, vol. 9, no. 1, pp. 23–24, 2004, doi: 10.1016/S0734-189X(04)10487-5.
[6] C. Yang, X. Lu, Z. Lin, E. Shechtman, O. Wang, and H. Li, “High-resolution image inpainting using multi-secQale neural patch synthesis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 6721–6729. [Online]. Available: https://arxiv.org/pdf/1611.09969.pdf
[7] K. Nazeri, E. Ng, T. Joseph, F. Z. Qureshi, and M. Ebrahimi, “EdgeConnect: Generative image inpainting with adversarial edge learning,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2019, pp. 1–17. [Online]. Available: https://arxiv.org/abs/1901.00212.
[8] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum, “Image completion with structure propagation,” ACM Trans. Graph., vol. 24, no. 3, pp. 861–868, Jul. 2005, doi: 10.1145/1073204.1073274.
[9] A. Efros and T. Leung, “Texture synthesis by non-parametric sampling,” in Proc. Int. Conf. Comput. Vis., Sep. 1999, pp. 1033–1038, doi: 10.1109/ICCV.1999.790383.
[10] W. H. Cheng, C. W. Hsieh, S. K. Lin, C. W. Wang, and J. L. Wu, “Robust algorithm for exemplar-based image inpainting,” in Proc. Int. Conf. Comput. Graph., Imag. Vis., Jul. 2005, pp. 64–69.
[11] M. Desai, “Modified fast and enhanced exemplar based inpainting algo- rithm for solving unknown row filling problem,” Int. J. Comput. Applit., vol. 56, no. 9, pp. 20–24, Oct. 2012, doi: 10.5120/8199-2977.
[12] Anupam, P. Goyal, and S. Diwakar, “Fast and enhanced algorithm for exemplar based image inpainting,” in Proc. 4th Pacific–Rim Symp. Image Video Technol., Nov. 2010, pp. 325–330, doi: 10.1109/PSVT.2010.61.
[13] S. Hesabi and N. Mahdavi-Amiri, “A modified patch propagation-based image inpainting using patch sparsity,” in Proc. 16th CSI Int. Symp. Artif. Intell. Signal Process. (AISP), May 2012, pp. 43–48.
[14] K. mangeeth, P. Sengottuvelan, and E. Balanmurugan, “A novel exemplar- based image inpainting algorithm for natural scene image completion with improved patch prioritizing,” Int. J. Comput. Applit., vol. 36, no. 4, pp. 1–6, 2011. [Online]. Available: https://citeeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.734.8379&rep=rep1&type=pdf
[15] M. Waykule and M. Patil, “Region filling and object removal by exemplar-based image inpainting,” Int. J. Sci. Eng. Res., vol. 3, no. 1, pp. 6–12, 2012.
[16] Z. Xu and J. Sun, “Image inpainting by patch propagation using patch sparsity,” IEEE Trans. Image Process., vol. 19, no. 5, pp. 1153–1165, May 2010, doi: 10.1109/TIP.2010.2042098.
[17] M. J. Abdollahifar and S. Kalantari, “Gradient based search space reduc- tion for fast exemplar-based image inpainting,” in Proc. Int. Conf. New Res. Advic. Electr. Comput. Eng., 2016, pp. 1–6.
[18] L. J. Deng, T. Z. Huang, and X. L. Zhao, “Exemplar-based image inpainting using a modified priority definition,” PLoS One, vol. 10, no. 10, 2015, Art. no. e0141199, doi: 10.1371/journal.pone.0141199.
[19] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, “Image inpainting for irregular holes using partial convolutions,” in Proc. Eur. Conf. Comput. Vis. (ECCV), Sep. 2018, pp. 85–100. [Online]. Available: https://arxiv.org/abs/1804.07723
[20] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, “Free-form image inpainting with gated convolution,” in Proc. IEEE Int. Conf. Comput. Vis., Oct. 2019, pp. 4471–4480. [Online]. Available: https://arxiv.org/pdf/1804.07723.pdf
[21] X. Hong, P. Xiong, R. Ji, and H. Fan, “Deep fusion network for image completion,” in Proc. 27th ACM Int. Conf. Multimedia, Oct. 2019, pp. 2033–2042. [Online]. Available: https://arxiv.org/pdf/1904.08060.pdf
[22] M. Waykule and J. Park, “SC-FEGAN: Face editing generative adversarial network with user’s sketch and color,” in Proc. IEEE Int. Conf. Comput. Vis., Oct. 2019, pp. 1745–1753. [Online]. Available: https://arxiv.org/pdf/1902.06838.pdf
[23] Y. Zeng, J. Fu, H. Chao, and B. Guo, “Learning pyramid-context encoder network for high-quality image inpainting,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2019, pp. 1486–1494. [Online]. Available: https://arxiv.org/pdf/1904.07475.pdf
[24] C. Zheng, T. J. Cham, and J. Cai, “Plurisestyle image completion,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2019, pp. 1438–1447. [Online]. Available: https://arxiv.org/pdf/1903.04227.pdf
[25] F. Altilcin, M. Ozay, and T. Okatani, “Deep structured energy-based image inpainting,” in Proc. 24th Int. Conf. Pattern Recognit. (ICPR), Aug. 2018, pp. 423–428. [Online]. Available: https://arxiv.org/pdf/1801.07939.pdf
[26] R. E. Bellman, “Dynamic Programming,” 1st edition, Princeton University Press, Princeton, NJ, USA, 1957.
[27] J. A. Sethian, Level Set Methods and Fast Marching Methods, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1999.
[28] G. Borgefors, “Distance transformations in arbitrary images,” Comp. Vis., Graph., Image Process., vol. 27, pp. 321–345, Jun. 1984, doi: 10.1016/0734-189X(84)90035-5.
[29] G. Borgefors, “Distance transformations in digital images,” Comp. Vision, Graph., Image Process., vol. 32, pp. 344–371, Jun. 1986, doi: 10.1016/S0734-189X(86)80047-0.
[30] A. Meijster, J. Roerdink, and W. Hesselink, “A general algorithm for putting distance transforms in linear time,” in Mathematical Morphology and Its Applications to Image and Signal Processing, Dordrecht, The Netherlands: Springer, 2000, pp. 331–340, doi: 10.1007/3-540-47025-X_36.
[31] Z. Wang, A. C. Bovik, and L. Lu, “Why is image quality assessment so difficult,” in Proc. IEEE Int. Conf. Acoust. Speech Signal Process., vol. 4, May 2002, pp. 3313–3316, doi: 10.1109/ICASSP.2002.1004620.

[32] M. A. Qureshi, M. Deriche, A. Beghdadi, and A. Amin, “A critical survey of state-of-the-art image inpainting quality assessment metrics,” J. Vis. Commun. Image Represent., vol. 49, pp. 177–191, Nov. 2017, doi: 10.1016/j.jvcir.2017.09.006.

[33] 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, Aug. 2011, doi: 10.1109/TIP.2011.2109730.

[34] I. Molodetskikh, M. Erofeev, and D. Vatolin, “Perceptually motivated method for image inpainting comparison,” 2019, arXiv:1907.06296 [Online]. Available: http://arxiv.org/abs/1907.06296

[35] P. Tiefenbacher, V. Bogischef, D. Merget, and G. Rigoll, “Subjective and objective evaluation of image inpainting quality,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2015, pp. 447–451, doi: 10.1109/ICIP.2015.7350838.

IRAWATI NURMALA SARI received the B.S. degree from the Sepuluh Nopember Institute of Technology (ITS), Indonesia, in 2012, and the master’s degree from the National Taiwan University of Science and Technology (NTUST), Taiwan, in 2015. She is currently pursuing the Ph.D. degree with the Department of Information Science, Kyoto Institute of Technology (KIT). She is also a Research Student with the Visual Information Laboratory. Her research interests include image processing, computer vision, pattern recognition, and machine learning.

EMIKO HORIKAWA received the B.S. degree from the Kyoto Institute of Technology, in 2020, where she has been a Graduate Student, since 2020. Her research interests include image inpainting, computer vision, and machine learning.

WEIWEI DU (Member, IEEE) received the B.S. degree from the Tianjin Institute of Urban Construction, China, in 2002, the master’s degree from The University of Aizu, Japan, in 2005, and the Ph.D. degree from Kyushu University, Japan, in 2008. In 2008, she joined the Kyoto Institute of Technology, as an Assistant Professor. She has been a Full Associate Professor in information and human science with the Kyoto Institute of Technology, since 2019. Her research interests include pattern recognition, machine learning, neural networks, image processing, and computer vision.

***