

# IMAGE INPAINTING CONSIDERING BRIGHTNESS CHANGE AND SPATIAL LOCALITY OF TEXTURES

Norihiko Kawai, Tomokazu Sato, Naokazu Yokoya

*Graduate School of Information Science, Nara Institute of Science and Technology*

*8916-5 Takayama, Ikoma, Nara 630-0192, Japan*

*norihiko-k@is.naist.jp, tomokazu-s@is.naist.jp, yokoya@is.naist.jp*

Keywords: image inpainting, image completion, energy minimization

Abstract: Image inpainting is a technique for removing undesired visual objects in images and filling the missing regions with plausible textures. Conventionally, the missing parts of an image are completed by optimizing the objective function, which is defined based on pattern similarity between the missing region and the rest of the image (data region). However, unnatural textures are easily generated due to two factors: (1) available samples in the data region are quite limited, and (2) pattern similarity is one of the required conditions but is not sufficient for reproducing natural textures. In this paper, in order to improve the image quality of completed texture, the objective function is extended by allowing brightness changes of sample textures (for (1)) and introducing spatial locality as an additional constraint (for (2)). The effectiveness of these extensions is successfully demonstrated by applying the proposed method to one hundred images and comparing the results with those obtained by the conventional methods.

## 1 INTRODUCTION

Image inpainting is a technique for removing undesired visual objects in images and filling the missing regions with plausible textures. This research can be classified into two categories. One is a non-exemplar-based method and the other is an exemplar-based method. The non-exemplar-based methods (A. Levin et al., 2003; C. Ballester et al., 2001a; C. Ballester et al., 2001b; D. Tschumperlé, 2006; E. Villéger et al., 2004; M. Bertalmio et al., 2001; M. Bertalmio et al., 2000; S. Esedoglu and J. Shen, 2003; S. Masnou and J.M. Morel, 1998; T. Chan and J. Shen, 2001; T. Chan et al., 2002) are based on pixel interpolation considering the continuity of pixel intensity. These methods are effective for small image gaps like scratches in a photograph. However, the resultant image easily becomes unclear when the missing region is large. Therefore, recently many exemplar-based inpainting methods have been intensively developed because they can synthesize complex textures in the missing region.

Exemplar-based methods basically synthesize textures for the missing region based on pattern simi-

ilarity that is defined between the missing region and the rest of the image. Some of the exemplar-based methods use the distance in the feature space as a similarity measure. As the feature space, Fourier space, wavelet domain and eigenspace have been used (A.N. Hirani and T. Totsuka, 1996; S.D. Rane et al., 1996; T. Amano, 2004). Most of the other exemplar-based methods simply employ SSD (sum of squared differences)-based pattern similarity measures (A. Criminisi et al., 2004; A.A. Efros and T.K. Leung, 1999; B. Li et al., 2005; C. Allène and N. Paragios, 2006; I. Drori et al., 2003; J. Jia and C. Tang, 2003; J. Sun et al., 2005; N. Komodakis and G. Tziritas, 2006; R. Bornard et al., 2002; Y. Wexler et al., 2007). Efros et al. (A.A. Efros and T.K. Leung, 1999) have proposed a method that successively copies the most similar pattern from the data region to the missing region. Although this method can generate complex textures, the quality of resultant images is severely affected by the order of texture copy. To obtain good results with the successive texture copy, confidence maps such as the number of fixed pixels in a window, strength of isophotes around the missing regions and pattern similarity have been used to de-

termine the order of texture copy (A. Criminisi et al., 2004; B. Li et al., 2005; R. Bornard et al., 2002). Although the duplication of similar textures preserves the local texture continuity in these methods, discontinuous textures are easily synthesized in the completed image. To avoid the ordering problem, recent inpainting methods employ the iterative global optimization approach (C. Allène and N. Paragios, 2006; Y. Wexler et al., 2007; N. Komodakis and G. Tziritas, 2006). In these methods, the objective functions that evaluate the pattern similarity are defined and optimized by using the EM algorithm, Belief Propagation approach and graph cut approach.

Although the global optimization methods have obtained good results for many images, unnatural images are still generated due to two factors: (1) available samples in the data region are quite limited, and (2) pattern similarity is one of the required conditions but is not sufficient for reproducing natural textures. Thus, in order to improve the image quality, these two factors should be considered. There have already been some attempts at this. For (1), the scale and orientation of textures have been considered to obtain effective samples (I. Drori et al., 2003). For (2), Sun et al. (J. Sun et al., 2005) and Jia et al. (J. Jia and C. Tang, 2003) have proposed techniques that use explicit constraints for texture boundaries. These methods synthesize textures preserving the edges or boundaries of the texture. However, automatic and effective determination of these explicit constraints is still difficult.

In this paper, in order to obtain good results for many images, we employ a new approach different from conventional ones. For (1), brightness change of sample textures that has not been considered in the literature is allowed to obtain effective samples. For (2), the spatial locality of texture pattern is considered as an implicit constraint that is usually satisfied in a lot of real scenes. In this study, these ideas are implemented with the framework of energy minimization based on Wexler’s objective function (Y. Wexler et al., 2007). The effectiveness of our extensions is demonstrated with comparison of completed images and subjective evaluation by a questionnaire.

## 2 IMAGE INPAINTING BY ENERGY MINIMIZATION

Figure 1 shows the flow of the proposed method. First, a user manually selects regions to be repaired such as physically damaged regions and undesired object regions in an image (a). Next, initial values are given to the missing regions (b). Finally, selected re-

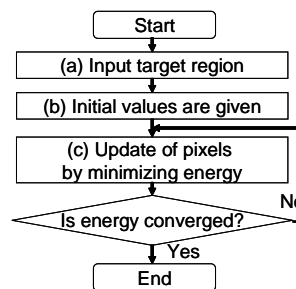


Figure 1: Procedure of the proposed method.

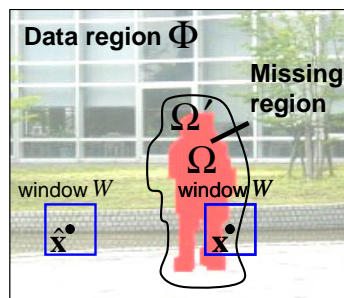


Figure 2: Missing and data regions in an image.

gions are completed by minimizing the energy function (c). In the following sections, first, we describe the conventional objective function which is defined based on pattern similarity SSD in Section 2.1. The new energy function considering brightness change and spatial locality is then defined in Section 2.2, and finally Section 2.3 describes the inpainting procedure that minimizes the energy function.

### 2.1 Energy function based on pattern similarity

This section briefly describes the SSD-based objective function for image inpainting originally proposed by Wexler et al. (Y. Wexler et al., 2007). Here, although the original objective function is defined as a probability density function, we redefine this objective function as an equivalent energy function.

As illustrated in Figure 2, first, an image is divided into region  $\Omega'$  including the missing region  $\Omega$  selected by a user and the data region  $\Phi$ , which is the rest of the image. The plausibility in the missing region  $\Omega$  is defined by using image patterns in the data region  $\Phi$ . Here,  $\Omega'$  is the expanded area of the missing region  $\Omega$  in which there is a central pixel of a square window  $W$  of size  $N_W$  (where  $N_W$  is a constant) overlapping the region  $\Omega$ . The energy function that represents the plausibility in the missing region is defined as the

weighed sum of SSD between the pixels around the pixel  $\mathbf{x}$  in region  $\Omega'$  and those around the pixel  $\hat{\mathbf{x}}_{org}$  in region  $\Phi$  as follows:

$$E_{org} = \sum_{\mathbf{x} \in \Omega'} w_{\mathbf{x}} SSD(\mathbf{x}, \hat{\mathbf{x}}_{org}), \quad (1)$$

where  $\hat{\mathbf{x}}_{org}$  in the data region  $\Phi$  denotes the pixel around which the pattern is the most similar to that around  $\mathbf{x}$  in the region  $\Omega'$ , and  $SSD(\mathbf{x}, \hat{\mathbf{x}}_{org})$  is defined as follows.

$$SSD(\mathbf{x}, \hat{\mathbf{x}}_{org}) = \sum_{\mathbf{p} \in W} \{I(\mathbf{x} + \mathbf{p}) - I(\hat{\mathbf{x}}_{org} + \mathbf{p})\}^2. \quad (2)$$

Here,  $I(\mathbf{x})$  represents the intensity of pixel  $\mathbf{x}$ . The pixel  $\hat{\mathbf{x}}_{org}$  for minimizing  $E_{org}$  is decided as follows.

$$\hat{\mathbf{x}}_{org} = f_{org}(\mathbf{x}) = \operatorname{argmin}_{\mathbf{x}' \in \Phi} SSD(\mathbf{x}, \mathbf{x}'). \quad (3)$$

Note that the weight  $w_{\mathbf{x}}$  is set as 1 if  $\mathbf{x}$  is inside of the region  $\Omega' \cap \Omega$  because pixel values in this region are fixed: otherwise  $w_{\mathbf{x}}$  is set as  $c^{-d}$  ( $d$  is the distance from the boundary of  $\Omega$  and  $c$  is a constant) because pixel values around the boundary have higher confidence than those in the center of the missing region.

In Wexler's work (Y. Wexler et al., 2007), the missing region is completed by calculating the pixel value  $I(\mathbf{x})$  in the missing region and the position of the pixel  $\hat{\mathbf{x}}_{org}$  that minimizes the energy function  $E_{org}$ .

## 2.2 Energy function extended by considering brightness change and spatial locality

In this study, we extend the original energy function  $E_{org}$  defined in Eq. (1) considering brightness change and spatial locality of texture patterns. Concretely, we introduce a modification coefficient to allow linear brightness change of the texture pattern. For considering spatial pattern locality, the cost function based on distance between the pixel in the missing region and the corresponding pixel in the data region is also added to the original energy function. The extended energy function is defined as follows:

$$E = \sum_{\mathbf{x} \in \Omega'} w_{\mathbf{x}} \{SSD'(\mathbf{x}, \hat{\mathbf{x}}) + w_{dis} SD(\mathbf{x}, \hat{\mathbf{x}})\}, \quad (4)$$

where  $SSD'(\mathbf{x}, \hat{\mathbf{x}})$  represents the pattern similarity considering brightness change, and  $SD(\mathbf{x}, \hat{\mathbf{x}})$  means the cost term for the spatial locality.  $w_{dis}$  is the weight representing the strength of spatial locality.  $\hat{\mathbf{x}}$  is determined as the pixel position that minimizes the extended energy function  $E$ :

$$\hat{\mathbf{x}} = f(\mathbf{x}) = \operatorname{argmin}_{\mathbf{x}' \in \Phi} (SSD'(\mathbf{x}, \mathbf{x}') + w_{dis} SD(\mathbf{x}, \mathbf{x}')). \quad (5)$$

In the following sections, definitions of  $SSD'(\mathbf{x}, \hat{\mathbf{x}})$  and  $SD(\mathbf{x}, \hat{\mathbf{x}})$  are detailed.

### 2.2.1 Pattern similarity considering brightness change

The similarity measure  $SSD'$  is defined as:

$$SSD'(\mathbf{x}, \hat{\mathbf{x}}) = \sum_{\mathbf{p} \in W} \{I(\mathbf{x} + \mathbf{p}) - \alpha_{\mathbf{x}\hat{\mathbf{x}}} I(\hat{\mathbf{x}} + \mathbf{p})\}^2. \quad (6)$$

To allow the brightness change, we introduce an intensity modification coefficient  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$ . In this paper, we employ the ratio of average pixel values around the pixels  $\mathbf{x}$  and  $\hat{\mathbf{x}}$  as the modification coefficient  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$ . However, an unnatural image is easily generated if large brightness change is approximated by linear transformation. Therefore, we limit the range of the value  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$  where brightness change can be approximated linearly as given in Eq. (7).

$$\alpha_{\mathbf{x}\hat{\mathbf{x}}} = \begin{cases} 1 - D & (\beta_{\mathbf{x}\hat{\mathbf{x}}} < 1 - D) \\ \beta_{\mathbf{x}\hat{\mathbf{x}}} & (1 - D \leq \beta_{\mathbf{x}\hat{\mathbf{x}}} \leq 1 + D) \\ 1 + D & (\beta_{\mathbf{x}\hat{\mathbf{x}}} > 1 + D), \end{cases} \quad (7)$$

where  $D$  is a constant ( $0 < D < 1$ ) and  $\beta_{\mathbf{x}\hat{\mathbf{x}}}$  is defined as follows:

$$\beta_{\mathbf{x}\hat{\mathbf{x}}} = \frac{\sqrt{\sum_{\mathbf{q} \in W} I(\mathbf{x} + \mathbf{q})^2}}{\sqrt{\sum_{\mathbf{q} \in W} I(\hat{\mathbf{x}} + \mathbf{q})^2}}. \quad (8)$$

The modification coefficient  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$  makes brightness of generated textures smooth while preserving texture patterns and enables utilization of the texture that has different brightness but the same pattern.

### 2.2.2 Spatial locality of texture

Spatial locality of a texture pattern is defined by using a sigmoid function:

$$SD(\mathbf{x}, \hat{\mathbf{x}}) = \frac{\|W\|}{1 + e^{\{-K(\|\mathbf{x} - \hat{\mathbf{x}}\| - X_0)\}}}, \quad (9)$$

where  $K$  and  $X_0$  are constant and  $\|W\|$  is the number of pixels in a window. This cost function is defined based on the assumption that the probability of similar texture existence for a certain pixel is uniformly high for the object region where the pixel exists. On the other hand, outside the object region, the probability can be assumed to be uniformly low. Note that a constant-sized object region is currently assumed in Eq. (9) because we could not know the range of the object in the missing region. By adding the constraint of spatial locality, even when the deformation of texture pattern exists around the target region, appropriate textures that exist near the target region are preferentially selected. Thus, textures are less likely to blur by not selecting blurry textures of low frequency far from the missing region.

### 2.3 Update of pixel values and window pairs for energy minimization

The energy function  $E$  defined in Eq. (4) is minimized by using a framework of greedy algorithm similar to Wexler's EM approach (Y. Wexler et al., 2007). In our method, we pay attention to the fact that the energy function  $E$  for each pixel can be treated independently if similar pattern pairs  $(\mathbf{x}, \hat{\mathbf{x}})$  calculated by Eq. (5) can be fixed and the change of coefficient  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$  in the iteration is very small. Thus, we repeat the following two processes until the energy converges: (I) update of pairs of windows for fixed pixel values, and (II) parallel update of all the pixel values in the missing region for fixed similar pairs of windows in the missing region and the data region.

In process (I), we update all the similar pattern pairs of windows in the missing region and in the data region fixing the pixel values calculated in the process (II). Concretely, the update of the pair of windows can be performed by calculating  $SSD'$  and  $SD$  that satisfy Eq. (5) and determining the position  $\hat{\mathbf{x}}$ .

In process (II), we update all the pixel values  $I(\mathbf{x})$  in the missing regions in parallel by minimizing the energy defined by Eq. (4). In the following, we describe the method for calculating the pixel values  $I(\mathbf{x})$  for fixed pairs of windows. First, the energy  $E$  is resolved into the element energy  $E(\mathbf{x})$  for each pixel in the missing region. As shown in Figure 3, the target pixel to be updated is  $\mathbf{x}$ , and the pixel position inside a window can be expressed as  $\mathbf{x} + \mathbf{p}$  ( $\mathbf{p} \in W$ ) and is corresponded to  $f(\mathbf{x} + \mathbf{p})$  by Eq. (5). Thus, the position of the pixel corresponding to the pixel  $\mathbf{x}$  is  $f(\mathbf{x} + \mathbf{p}) - \mathbf{p}$ . Now, the element energy  $E(\mathbf{x})$  can be defined in terms of the pixel values of  $\mathbf{x}$  and  $f(\mathbf{x} + \mathbf{p}) - \mathbf{p}$ , the coefficient  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$  and the Euclid distance between  $\mathbf{x}$  and  $f(\mathbf{x})$  as follows:

$$E(\mathbf{x}) = \sum_{\mathbf{p} \in W} w_{\mathbf{x}+\mathbf{p}} \{I(\mathbf{x}) - \alpha_{\mathbf{x}+\mathbf{p}f(\mathbf{x}+\mathbf{p})} I(f(\mathbf{x} + \mathbf{p}) - \mathbf{p})\}^2 + w_{dis} \frac{\|W\|}{1 + e^{\{-K(\|\mathbf{x}-\hat{\mathbf{x}}\| - X_0)\}}}. \quad (10)$$

The relationship between the energy  $E$  for all of the missing region and the element energy  $E(\mathbf{x})$  for each pixel can be written as follows:

$$E = \sum_{\mathbf{x} \in \Omega} E(\mathbf{x}) + C. \quad (11)$$

$C$  is the energy of pixels in the region  $\bar{\Omega} \cap \Omega'$ , and is treated as a constant because pixel intensities in the region and all pairs of windows are fixed here. By differentiating  $E$  with respect to  $I(\mathbf{x})$  in the missing region, the requirement for minimizing the energy  $E$

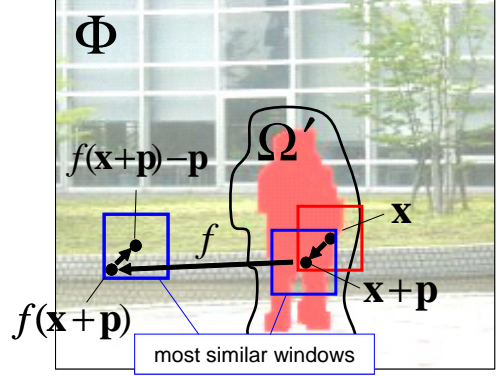


Figure 3: Relationship between pixels in energy calculation.

can be obtained as follows:

$$\frac{\partial E}{\partial I(\mathbf{x}_k)} = \sum_{\mathbf{x} \in \Omega} \frac{\partial E(\mathbf{x})}{\partial I(\mathbf{x}_k)} = 0. \quad (12)$$

Here, if it is assumed that the change of intensity modification coefficient  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$  is smaller than that of the pixel intensity  $I(\mathbf{x}_k)$ , we can obtain the following equation.

$$\frac{\partial \alpha_{\mathbf{x}_i \mathbf{x}'_j}}{\partial I(\mathbf{x}_j)} = 0 \quad (\forall \mathbf{x}_i, \mathbf{x}_j \in \Omega, \forall \mathbf{x}'_j \in \Phi). \quad (13)$$

From this equation, the equation  $\partial E(\mathbf{x}) / \partial I(\mathbf{x}_k) = 0$  ( $\mathbf{x} \neq \mathbf{x}_k$ ) is formed. Thus, we can minimize the energy  $E$  by calculating  $I(\mathbf{x}_k)$ , which satisfies the following equation.

$$\frac{\partial E}{\partial I(\mathbf{x}_k)} = \frac{\partial E(\mathbf{x}_k)}{\partial I(\mathbf{x}_k)} = 0. \quad (14)$$

By generalizing Eq. (14) to all the pixels in the missing region, each pixel value  $I(\mathbf{x})$  in the missing region can be calculated as follows:

$$I(\mathbf{x}) = \frac{\sum_{\mathbf{p} \in W} w_{\mathbf{x}+\mathbf{p}} \alpha_{\mathbf{x}+\mathbf{p}f(\mathbf{x}+\mathbf{p})} I(f(\mathbf{x} + \mathbf{p}) - \mathbf{p})}{\sum_{\mathbf{p} \in W} w_{\mathbf{x}+\mathbf{p}}}. \quad (15)$$

Eq. (15) gives an approximate solution when Eq. (13) is satisfied. We can obtain a good solution as the energy converges because the value of intensity modification coefficient  $\alpha_{\mathbf{x}\hat{\mathbf{x}}}$  converges as  $I(\mathbf{x})$  converges.

Additionally, in order to avoid local minima efficiently, a coarse-to-fine approach is also employed. Specifically, an image pyramid is generated and the energy minimization processes (i) and (ii) are repeated from higher-level to lower-level layers successively using a certain size of window. Good initial values are given to the lower layer by projecting results from the higher layer. This makes it possible to decrease computational cost and avoid local minima. In the lowest layer (original size), the energy minimization process is repeated while reducing the size of the window, and it enables reproduction of more detailed textures.

### 3 EXPERIMENTS

To demonstrate the effectiveness of our extensions of the objective function, we have applied five kinds of image inpainting methods including the conventional methods and the proposed method to one hundred images ( $200 \times 200$  pixels). First, the effectiveness of the proposed method is demonstrated by comparing the characteristic results of the proposed method and our implemented Wexler’s method (Y. Wexler et al., 2007). Next, by the subjective evaluation based on a questionnaire using our implemented Wexler’s method, Criminisi’s method (A. Criminisi et al., 2004) and the proposed method, the effectiveness of our method is objectively assessed.

In these experiments, we used a standard PC (CPU: Xeon 3.2 GHz, Memory: 8 GB) and each parameter in the energy function was set as shown in Table 1. Here, the missing region was manually specified, and the average pixel value of the boundary of the missing region is given as an initial value in the missing region.

#### 3.1 Comparison of inpainted images

In this section, four images that have different characteristics are selected from one hundred images as shown in Figure 4(a). The missing regions for these images are given in Figure 4(b). Figure 4(c) indicates the resultant images by our implemented Wexler’s method described in Section 2.1. Figure 4(d) shows the images completed by the proposed method.

Image (I) includes little brightness change under constant illumination and little pattern change in the same object region around the missing region. It can be confirmed that images generated by the conventional and proposed methods are natural. The subjective difference is very small.

Image (II) includes little pattern change without complex textures but large brightness changes under nonconstant illumination around the missing region. By the conventional method, the resultant image looks unnatural because sudden intensity changes appear at the seat and the seat back. By allowing bright-

ness changes of sample textures, the sudden intensity change is suppressed by the proposed method.

Image (III) includes little brightness change under constant illumination but large pattern change due to the various sizes and shapes of objects around the missing regions. In this image, although the same kinds of textures apparently exist around the missing regions, texture pattern greatly changes due to the different sizes of stones. A part of the missing regions is blurred in white by the conventional method because the SSD-based similarity is sensitive to the pattern changes especially for high-frequency components, and thus inappropriate textures are selected for the missing regions. It should be noted that there exists spatial locality of texture pattern such as water color and stones around the missing region in image (III). By considering spatial locality of texture pattern, neighboring textures are preferentially selected and thus the missing region is completed successfully by the proposed method.

Image (IV) includes large brightness change under nonconstant illumination conditions and texture pattern continuously changes due to the perspective projection effect. In the resultant image of the conventional method, an unnatural image is generated due to the blurs on the textured area with black squares and the discontinuous brightness changes at the wall and floor. On the other hand, in the proposed method, by using the constraint of the spatial locality of textures, neighboring textures are selected for completion of the missing region and windows of the post are reproduced in Figure 4(d). In addition, by allowing brightness changes of sample textures, brightness change inside the missing region becomes more natural than that by the conventional method.

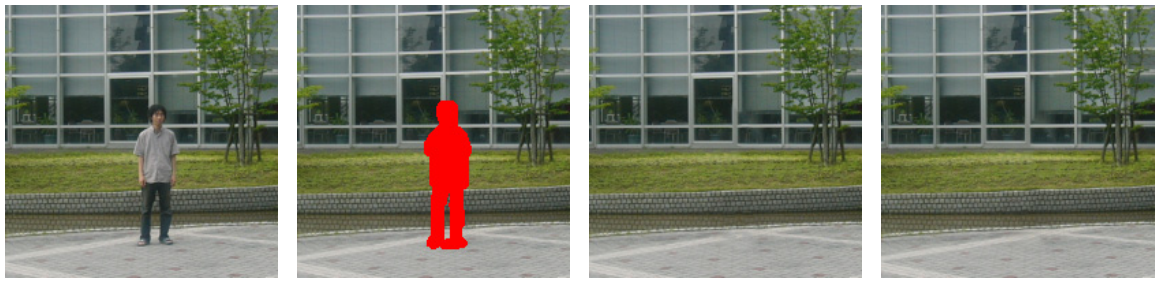
Next, we have compared the conventional and proposed methods with respect to computational cost. Table 2 shows the processing time of the conventional and proposed methods. The proposed method requires about three to five times as much time as the conventional method. This is because the computational cost for calculating intensity modification coefficients and cost function considering spatial locality is increased.

Table 1: Parameters in experiment.

Window size	$N_w$	max $9 \times 9$ min $3 \times 3$
Weight for distance	$w_{dis}$	120
Parameter in sigmoid function	$K$	0.4
	$X_0$	20
Range of coefficient $\alpha$	$D$	0.1

Table 2: Processing time.

	Conventional method	Proposed method
Image (I)	2'17"	8'45"
Image (II)	3'17"	12'15"
Image (III)	3'28"	18'59"
Image (IV)	5'25"	18'32"



(a)

(b)

(c)

(d)

Image (I): little change in brightness and pattern



(a)

(b)

(c)

(d)

Image (II): large change in brightness



(a)

(b)

(c)

(d)

Image (III): change in texture pattern



(a)

(b)

(c)

(d)

Image (IV): large change in brightness and continuous deformation of texture pattern

Figure 4: Image inpainting for four representative images: (a) original image, (b) missing region specified manually, (c) resultant image by conventional method, and (d) resultant image by our method.

### 3.2 Evaluation by a questionnaire

In this section, completed images for one hundred images by five kinds of inpainting methods are subjectively evaluated by 37 subjects. The subjects are men and women in their twenties and all of them often use computers. This experiment aims to illustrate the effectiveness of the proposed method objectively by evaluating the resultant images by subjects.

#### 3.2.1 Evaluation method

The subjects were requested to access the web page for questionnaire evaluation and evaluate the 500 resultant images for 100 input images by giving a score of 1 to 5. In this experiment, images were completed by five methods: our implemented Criminisi’s method (method A) that is the representative of successive synthesis methods, our implemented Wexler’s method (method B), proposed method allowing only brightness change (method C), proposed method considering only spatial locality (method D), and proposed method considering both brightness change and spatial locality (method E). On the evaluation web page, the resultant images generated by five methods were arranged in random order so that subjects could not know the relationship between image and method. The evaluation criteria were that the lowest score 1 was for an image that could not be used and the highest score 5 was for an image that was natural enough to use for personal homepages or magazines.

#### 3.2.2 Results and discussion

The average score for the 100 resultant images and the number of images that obtained the highest score are shown in Table 3 for each method.<sup>1</sup> Table 3 shows that the average score of the resultant images by the proposed method (method E) is higher than that by the conventional methods (methods A and B) and the inpainted images by the proposed method are scored as the best most frequently. In this experiment, scores of the proposed method (method E) and the conventional methods (methods A and B) were also compared by using the t-test with a 5% significant level. In the result, significant difference was observed between these scores. Therefore, the proposed method can be verified to be better than the conventional methods A and B. In addition, both method considering only brightness change (method C) and method considering only spatial locality (method D) obtained higher scores than methods A and B and the significant difference was also observed by using the t-test with a

<sup>1</sup>100 input images and resultant images are shown on the web page [<http://yokoya.naist.jp/research/visapp/>]

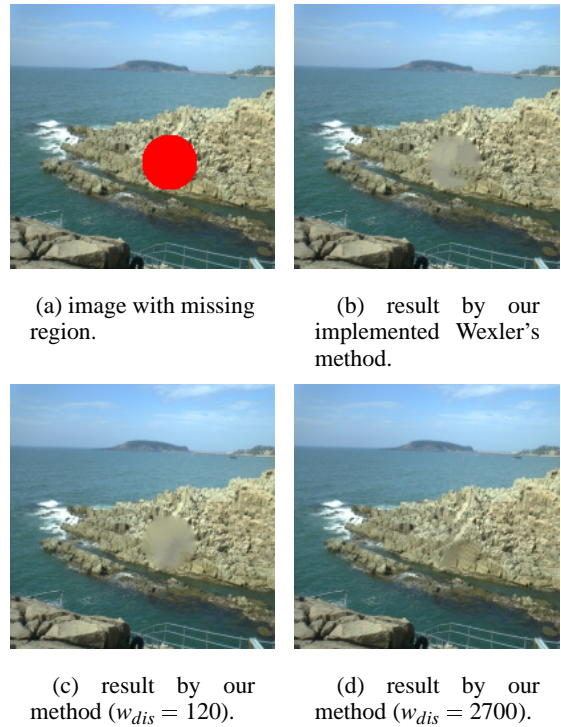


Figure 5: Example images for which both the proposed and conventional methods have a problem.

5% significant level. This means that the introduction of each factor to the energy function is clearly effective.

It should be noted here that there exist some images for which both the proposed (method E) and conventional (method B) methods have a problem as shown in Figure 5. The average scores for both methods are less than 2 points. In this image, due to the fairly large change in texture patterns, the texture in the missing region blurs despite the consideration of spatial locality. Here, when the weight for spatial locality was increased ( $w_{dis} = 2700$ ) a more natural image was generated as shown in Figure 5 (d). The com-

Table 3: Average score and the number of best-scoring images from 100 images.

	Average of score	Number of best-scoring images
Method A	2.21	7
Method B	3.24	7
Method C	3.39	21
Method D	3.42	20
Method E	3.60	45

pletion with a single weight coefficient does not always work well, and thus it is necessary to determine the parameter adaptively considering the characteristics of the image in order to obtain good results for many images containing complex textures.

## 4 CONCLUSION

In this paper, the objective function for image inpainting is extended to acquire natural images. To obtain good results, two factors were considered: (1) brightness change of sample textures was allowed, (2) spatial locality was introduced as a new constraint. By considering these two factors, the missing region was completed successfully for many images. In experiments, we have demonstrated the effectiveness of our method by comparing the resultant images of the conventional and proposed methods. In addition, by a questionnaire evaluation using 37 subjects, we have verified that the proposed method could obtain good results for more images than the conventional method. In experiments, parameters such as the size of window and the weight in the energy function were decided empirically. In future work, we should establish a method to decide optimum parameters.

## REFERENCES

- A. Criminisi, P. Pérez, and K. Toyama (2004). Region Filling and Object Removal by Exemplar-Based Image Inpainting. In *Trans. on Image Processing*, volume 13, No. 9, pages 1200–1212.
- A. Levin, A. Zomet, and Y. Weiss (2003). Learning How to Inpaint from Global Image Statistics. In *Proc. ICCV*, volume 1, pages 305–312.
- A.A. Efros and T.K. Leung (1999). Texture Synthesis by Non-parametric Sampling. In *Proc. ICCV*, pages 1033–1038.
- A.N. Hirani and T. Totsuka (1996). Combining Frequency and Spatial Domain Information for Fast Interactive Image Noise Removal. In *Proc. SIGGRAPH1996*, pages 269–276.
- B. Li, Y. Qi, and X. Shen (2005). An Image Inpainting Method. In *Proc. IEEE Int. Conf. on Computer Aided Design and Computer Graphics*, pages 531–536.
- C. Allène and N. Paragios (2006). Image Renaissance Using Discrete Optimization. In *Proc. ICPR*, pages 631–634.
- C. Ballester, M. Bertalmio, V. Sapiro, and J. Verdera (2001a). Filling-In by Joint Interpolation of Vector Fields and Gray Levels. In *Trans. on Image Processing*, volume 10, No. 8, pages 1200–1211.
- C. Ballester, V. Caselles, J. Verdera, M. Bertalmio, and G. Sapiro (2001b). A Variational Model for Filling-In Gray Level and Color Images. In *Proc. ICCV*, pages 10–16.
- D. Tschumperlé (2006). Curvature-Preserving Regularization of Multi-valued Images Using PDE's. In *Proc. ECCV*, volume 2, pages 295–307.
- E. Villéger, G. Aubert, and L. Blanc-Féraud (2004). Image Disocclusion Using a Probabilistic Gradient Orientation. In *Proc. ICPR*, volume 2, pages 52–55.
- I. Drori, D. Cohen-Or, and H. Yeshurun (2003). Fragment-Based Image Completion. In *Proc. SIGGRAPH2003*, pages 303–312.
- J. Jia and C. Tang (2003). Image Repairing: Robust Image Synthesis by Adaptive ND Tensor Voting. In *Proc. CVPR*, pages 643–650.
- J. Sun, L. Yuan, J. Jia, and H. Shum (2005). Image Completion with Structure Propagation. In *Proc. SIGGRAPH2005*, pages 861–868.
- M. Bertalmio, A. L. Bertozzi, and G. Sapiro (2001). Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting. In *Proc. CVPR*, pages 355–362.
- M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester (2000). Image Inpainting. In *Proc. SIGGRAPH2000*, pages 417–424.
- N. Komodakis and G. Tziritas (2006). Image Completion Using Global Optimization. In *Proc. CVPR*, pages 442–452.
- R. Bornard, E. Lecan, L. Laborelli, and J. Chenot (2002). Missing Data Correction in Still Images and Image Sequences. In *Proc. ACM Int. Conf. on Multimedia*, pages 355–361.
- S. Esedoglu and J. Shen (2003). Digital Inpainting Based on the Mumford-shah-euler Image Model. In *European J. of Applied Mathematics*, volume 13, pages 353–370.
- S. Masnou and J.M. Morel (1998). Level Lines Based Disocclusion. In *Proc. ICIP*, volume 3, pages 259–263.
- S.D. Rane, J. Remus, and G. Sapiro (1996). Wavelet-Domain Reconstruction of Lost Blocks in Wireless Image Transmission and Packet-Switched. In *Proc. ICIP*, volume 1, pages 309–312.
- T. Amano (2004). Image Interpolation by High Dimensional Projection Based on Subspace Method. In *Proc. ICPR*, volume 4, pages 665–668.
- T. Chan and J. Shen (2001). Non-texture Inpainting by Curvature-Driven Diffusions (CDD). In *J. of Visual Communication and Image Representation*, volume 12, No. 4, pages 436–449.
- T. Chan, S. Kang, J. Shen, and S. Osher (2002). Euler's Elastica and Curvature Based Inpaintings. In *SIAM J. of Applied Mathematics*, volume 63, No. 2, pages 564–592.
- Y. Wexler, E. Shechtman, and M. Irani (2007). Space-Time Completion of Video. In *Trans. on Pattern Analysis and Machine Intelligence*, volume 29, No. 3, pages 463–476.