Designシンポジウム2010予稿集, dss10-0078 (2010)

インタラクティブ3D屋内モデリングのためのスーパー画素の拘束を利用した事例に基づく画像修復

ﾀﾝｶﾞﾏﾆ　ｶﾗｲﾊﾞﾆ**†　　石川　智也**　　蔵田　武志**†

Exemplar-based Inpainting Using Superpixel Constraints for Interactive 3D Indoor Modeling

Kalaivani THANGAMANI, Tomoya ISHIKAWA and Takeshi KURATA

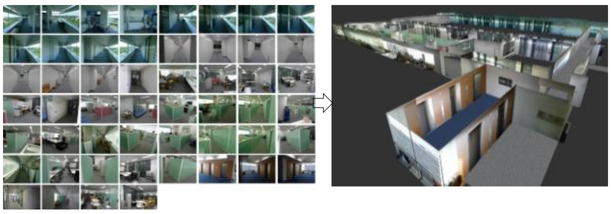
We have developed an interactive indoor modeler in which the user can create 3D indoor models from a single photo or multiple photos by simple interaction techniques based on computer vision principles. There are often invisible regions on some of the 3D planes since it is not easy to take a set of photos so that every region of the 3D models is included at least in one of those photos. We employ inpainting techniques for making up for the invisible regions with computer-generated texture patches and for merging the inpainted region with the neighboring visible regions. This paper describes a new inpainting technique that combines the Exemplar-based inpainting with the superpixel constraints in order to improve the structure propagation.

Key words: Inpainting, 3D modeling, Mixed Reality, Co-creative Intelligence Cycles, Service cooperation

1. INTRODUCTION

Virtualized real objects made from photos can make virtual environments more realistic. This reduces the gap between the real and virtual world for a number of applications such as personal navigation, visualization, and simulation. We have developed an interactive modeler [1, 2] to enable the users to easily create 3D indoor models from a single or multiple photos. However, some of 3D planes in the models often have imperfect or no texture due to occlusion by the target itself or other objects. We adopt the Exemplar-based inpainting technique [3] which is a simpler and faster method for texture filling the untextured regions [4]. The indoor modeler images contain many structures whose structure continuity needs to be maintained inside the untextured region. So there is a need to handle the inpainting which guarantees the good structure propagation. We propose the usage of superpixel constraints in improving the structure propagation part in the Exemplar-based inpainting method.

The paper is organized as follows. The second section briefly discusses the scenario of our interactive indoor modeler, the need for inpainting and the inpainted results through the Exemplar-based inpainting. The later part of this section summarizes the related works in the area of structure propagation. The third section introduces the concept of superpixels and their role in improving the structure propagation. The fourth section explains the overview of our proposed superpixel based structure propagation algorithm in combination with the Exemplar-based inpainting. The subsections explain in detail about the superpixel handling, the family of spline curves in curve completion and the subsequent texture propagation. The fifth section discusses our ongoing work, compares the Exemplar-based inpainting with the proposed superpixel based structure propagation and list out the future works in realizing the superpixel based structure propagation.



(a)　　(b)

Fig. 1 Interactive indoor modeling (a) multiple photos (b) Indoor model made from the multiple photos [1]

2. INTERACTIVE MODELER AND INPAINTING

The interactive modeler initiates 3D modeling by analyzing 2D input photos. The viewpoint and the rotation angles at which each photo was taken are estimated by using the vanishing points [5, 6] obtained from pairs of lines in the actual 3D world. The origin of the ground plane is interactively set by the user over which the texture-mapped 3D planes are added one by one for developing the 3D indoor model. The developed 3D planes can be translated, rotated, deformed or deleted using simple interaction techniques based on geometric constraints derived from the photos. Fig. 1 shows the multiple photos taken for our office and are used to model

＊ 原稿受付 平成22年10月15日
** 非会員　独立行政法人 産業技術総合研究所
サービス工学研究センター
（〒305-8568　茨城県つくば市梅園1-1-1）
† 非会員　筑波大学（〒305-8577
茨城県つくば市天王台1-1-1）

the indoor environment. Each 2D photo is used to make the local model in the local coordinate and then the local models are integrated to form the global model. For integrating the local model, the modeler estimates the transform parameters between the local and the global coordinate systems by means of sparse maps of landmarks generated in preprocessing and the result of image feature matching.

2.1 AUTOMATIC INPAINTING MASK GENERATION

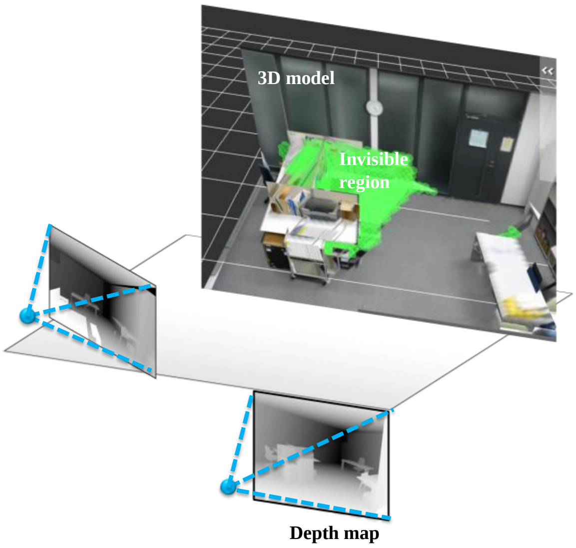


Fig. 2 Automatic inpainting mask detection

The regions which are invisible in any of the input photos inevitably hold textures of their frontal objects due to the projective texture mapping with GPU. The depth maps are computed for the developed 3D model. The invisible region is carved by checking the depths between the 3D planes. The dominant point clusters in every plane forms the inpainting mask. The green region in Fig. 2 denotes the invisible region which is detected by comparing the depths of the particular local model. The planes with the invisible region or in other words the untextured region are retrieved one after the other and the inpainting is applied for those planes. The inpainted planes are put back into the 3D model to produce the realistic indoor environment.

2.2 EXEMPLAR-BASED INPAINTING

The Exemplar-based inpainting proposed by Criminisi et.al, is applied for filling the textures in the untextured regions. The exemplar-based method fills the invisible region with the texture patches from the neighboring region in such a way that the structure of the texture is maintained. The exemplar-based algorithm sets the priority for the pixels in the mask boundary since the filling order is important in propagating the structure inside the mask region. The highest priority is given to the pixel which is surrounded by most of the data pixels and is in continuation of the strong edges. The suitable texture patch is selected by calculating the sum of the squared differences (SSD) of the pixel values between the texture patches. The priority mapping is explained in detail in section 4 where our proposed algorithm is aimed to improve the structure propagation part in the Exemplar-based inpainting method. Fig. 3(b) shows the 3D model which consists of a floor, a wall and the objects (table and a chair) created by the modeler from the input image which is shown in Fig. 3(a). The created 3D model contains the invisible region due to the objects in the room. The inpainting mask is automatically generated by the modeler by casting the projection shadows of the objects that causes the occlusion. Fig. 3(c) shows one such plane that contains the texture of the frontal objects that is to be removed. Fig. 3(d) shows the inpainting mask being set up in the region where the table and chair texture exist. Fig. 3(e) shows the inpainted plane by Exemplar-based inpainting. The inpainting handles the patch size of 40x40 with 82 iteration steps to fill the inpainting mask presents in the image of size 792x453. The iteration steps are shown in Fig. 3(f). Fig. 3(g) and 3(h) compares the before and after effect of inpainting.

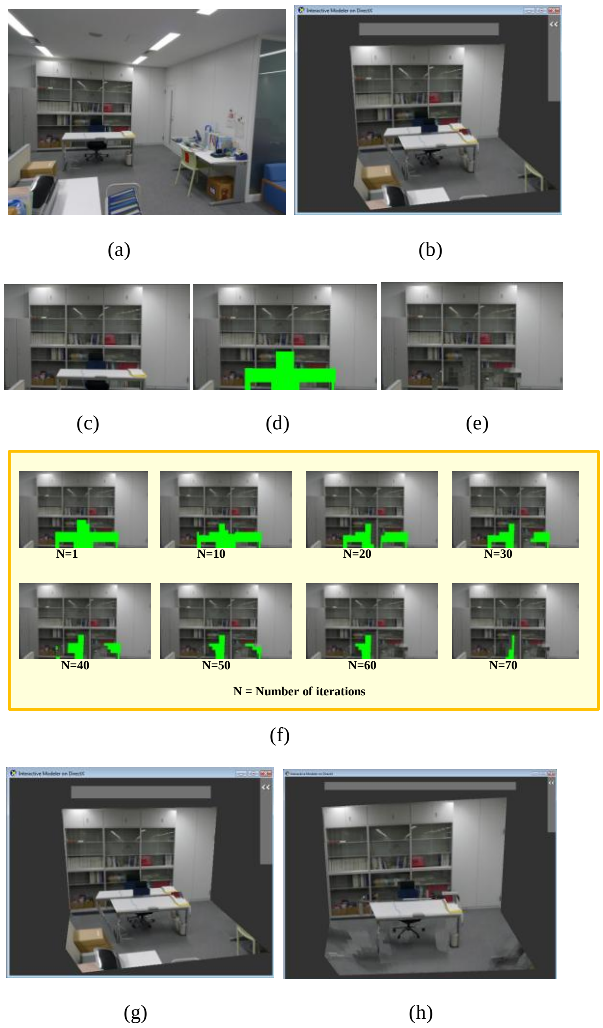


Fig. 3 Inpainting for indoor modeler.

(a) Input image, (b) 3D indoor model made from the image (a), (c) One of the 3D planes which holds the texture of the frontal object, (d) Inpainting mask shown in green, (e) Invisible region replaced by Exemplar-based inpainting, (f) Iteration steps in the Exemplar-based inpainting, (g) 3D model before inpainting and (h) 3D model after inpainting

We are keen in maintaining the structures in and out of the mask region since structures make an impact in every plane of the indoor environment. There are quite a good number of researches in the area of structure propagation in inpainting. Jason C. Hung et al. [7] extend the Exemplar-based inpainting by incorporating Bezier curves to construct missing edge information. Jian Sun et al. [8] referred the user drawn lines for the possible structure propagation. Andrei Rares et al. [9] proposed an inpainting algorithm that relies on explicit edge information. The edge information is used for the reconstruction of the structure in the missing areas, as well as for guiding the pixel interpolation. Nikos Komodakis et al. [10] proposed an alternate image completion method for the Exemplar based method with the Markov Random Field (MRF). They state that the Exemplar-based method follows the greedy way of sampling for filling the textures. There are good references on tensor voting [11] for the robust image synthesis. Treating the image in the global aspect recommends the good structure propagation [12].

3. SUPERPIXEL IN STRUCTURE PROPAGATION

Superpixels correspond to small, nearly-uniform segmented regions in the image. The idea of superpixels is originally developed by X.Ren and J.Malik [13], they mention that the superpixels are local, coherent and preserve most of the structure necessary for segmentation at the scale of interest. The concept of superpixels are of interest to many image processing and computer vision researchers. The inspiration for using the superpixels in 3D modeling is obtained from Derek Hoiem et al. [14] from their work Automatic photo pop-up where they use superpixels for categorizing the horizontal and vertical planes from the given image. The efficient graph based segmentation [15] is used for segmenting the image regions into superpixels. This graph based method segments the region based on the intensity, color, motion, location and other local attributes instead of the fixed threshold method in the traditional segmentation algorithms. Fig. 4 shows the superpixels for the image in Fig. 3(d).

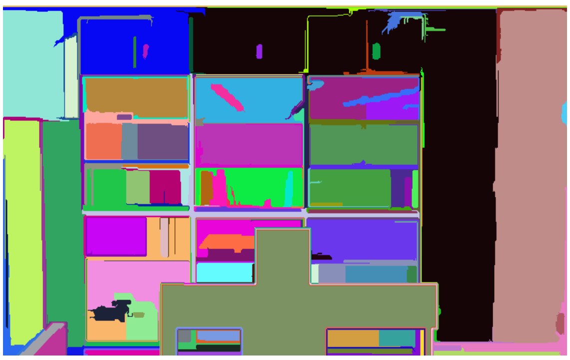


Fig. 4 Superpixels for the input image and the mask in Fig. 3(d)

4. ALGORITHM OVERVIEW

The overview of our proposed superpixel based structure propagation in the Exemplar-based inpainting is shown in Fig. 5. The boundary of the mask region is extracted and the superpixels that are in contact with the mask boundary are gathered and these are the only textures that are going to be referred for inpainting. Fig. 6 shows the superpixels that are in contact with the mask boundary for the test image in Fig. 3(d). The textures for the contact superpixels are shown in Fig. 7. The aim of the algorithm is to propagate the structures that are presented outside the mask region and not to miss any explicit structures. The priority is calculated for the mask boundary pixels to determine which superpixel is to be propagated first into the mask region.

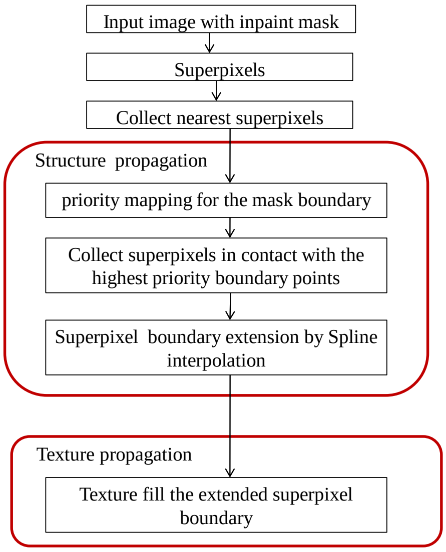


Fig. 5 Block diagram for the proposed superpixel based structure propagation

The edges belong to the same superpixel are grouped and the hermite interpolation is carried out for missing boundaries. The structure detection based on the superpixels reduces the tedious edge searching procedure in Edge-based image restoration [9] techniques. Superpixel based structure propagation is repeated till the explicit structures are propagated inside the mask region.

Once the stucture skeleton is formed, then the skeletons are filled with the texture flesh. These skeletons get the texture from the same superpixel. This is the step where the speed of the inpainting process is increased. In the original Exemplar-based method, for filling every patch, the algorithm searches the entire image for a successful patch.

Once all the major structures are propagated inside the mask region, the remaining regions are inpainted with the normal Exemplar-based inpainting method. The strategy in realizing the algorithm is explained in the following subsections.

4.1 PRIORITY MAPPING

Priority mapping is adopted from the Exemplar-based method. Priority computaion is biased toward the patches which are on the continuation of strong edges and which are surrounded by high confidence pixels. The following notations are followed same as that of the Exemplar-based inpainting original paper [3] with respect to the explanation diagram used.

Given a patch ψ_p centered at a boundary point p, its priority P(p) is defined as the product of two terms c(p) and D(p), where C(p) is the confident term and D(p) is the data term which is shown in Equation 1. The confidence term measures the amount of known pixels around the boundary point. The intension is to fill those patches which have more of their pixels already filled. The confidence term is represented by Equation 2.

Fig. 6 Superpixels that are in contact with the mask boundary

Fig. 7 The corresponding textures for the superpixels in Fig. 5

The data term D(p) is a function of the strength of the isophotes hitting the boundary pixel. There is a delicate balance maintained between the confidence and the data terms. The data term tends to push the isophotes rapidly inward, while the confidence term tends to supress preciesly this sort of incursion into the target region. The data term is represented by Equation 3.

$$\mathrm{P(p)} = \mathrm{C(p)}\,\mathrm{D(p)} \text{.........1}$$

$$\mathrm{C(p)} = \frac{\sum_{q \in \Psi_p \cap \bar{\Omega}} c(q)}{|\Psi_p|} \text{.....2}$$

$$D(p) = \frac{|\nabla I_p^{\perp} \bullet n_p|}{\alpha} \text{.......3}$$

Where,

C(p)	=	Confidence term
D(p)	=	Data term
C(q)	=	Number of pixels in the patch that are already filled
ψ_p	=	Area of the patch
$\nabla I_p^{\perp}$	=	Isophote(direction and intensity) at boundary point p
α	=	Normalization factor(255 for typical grey level image)
n_p	=	Normal to the boundary point p

4.2 SPLINE BASED CURVE COMPLETION

P_1	: Start point
P_2	: End point
∇p_1	: Tangent of p_1
∇p_2	: Tangent of p_2

Fig. 8 Hermite specification

$$s = \begin{vmatrix} s^3 \\ s^2 \\ s^1 \\ 1 \end{vmatrix} \text{...4} \qquad c = \begin{vmatrix} p_1 \\ p_2 \\ T_1 \\ T_2 \end{vmatrix} \text{...5} \qquad h = \begin{vmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} \text{...6}$$

$$p = s * h * c \text{...7}$$

Hermite curves is a type of spline curves family, are very easy to calculate but also very powerful. They are used to smoothly interpolate between key-points. The hermite curves work in any dimension; here the two dimensions are adopted for the curve completion. The start and the end points of the curves and their tangents (which are used to measure the direction and the speed) are the key points for the Hermite interpolation. Fig. 8 shows the curve completion example based on the hermite spline interpolation for the given points p_1 and p_2 having the direction vectors ∇p_1 and ∇p_2 respectively.

Equation 4 represents the vector S which denotes the interpolation point and its powers up to 3. Equation 5 represents the vector C, the parameters of the hermite curve. Equation 6 represents the matrix h, the matrix form of the 4 hermite polynomials. The points on the missing curve are found out by building the vector S and multiply the same with the matrix h and vector C as shown in Equation 7. J. C. Hung et al. [7] and Atzori et al. [16] make use of the Bezier curves, a type of spline curves for structure propagation in their inpainting works. Fig. 9 shows the mask region and the superpixel in contact with the highest priority boundary point. The edges of the superpixel in Fig. 9 are subjected to hermite interpolation for completing the curves inside the mask region. Once the missing curves are completed, they form the skeleton for the propagated structures. The next step is to fill textures for these structures. The extended superpixel boundary will get the textures from the original superpixels. Fig. 10 shows the extended structure skeleton filled by the texture patches from the original superpixel. The next subsection explains the texture filling process in detail.

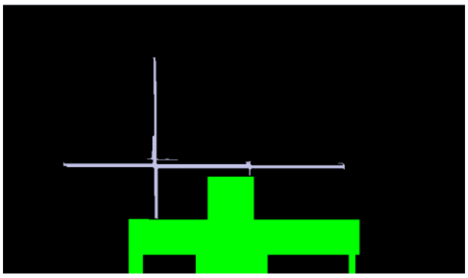


Fig. 9 Superpixel in contact with the highest priority boundary point

Fig. 10 Structure propagation using the superpixel shown in Fig. 5

4.3. TEXTURE FILLING THE STRUCTURE SKELETON

The extended boundaries of the superpixels will get the textures from the same superpixel by texture synthesis. Image quilting [17, 18] and graph cuts [19] have been the successful method for texture synthesis in many inpainting works. The exemplar-based method also adopts the texture synthesis based texture propagation for the sampling patches. In the Exemplar-based method, the patch with the minimum SSD value is chosen for filling the priority points. The extended boundary region will get the texture patches based on the priority calculation and the SSD values. In every propagated structure, the process of Exemplar-based inpainting is carried out with the original superpixel acting as the source and the extended region acts as the target. Even though the efficient graph based segmentation groups the superpixels based on their texture properties, there are cases where the single texture is separated into more than one superpixel and vice versa. The grouping of the superpixels or in other words grouping the groups once again is necessary for extrapolating the superpixel for structure and texture propagation.

5. DISCUSSION AND FUTURE WORKS

The proposed superpixel constraints improve the structure propagation part in the Exemplar-based inpainting and guarantee the structure propagation. When comparing the proposed work with the original Exemplar-based inpainting, in the Exemplar-based method, the success of the structure propagation is highly dependent on the order in which the

(a) (b)

Fig. 11 Comparison between the normal Exemplar-based.inpainting and the proposed superpixel based structure propagation (a) Inpainted region for the first iteration in Exemplar-based method (b) Inpainted region for the first iteration in superpixel based structure propagation

filling proceeds. There are chances for missing out the least priority structures. In every iteration, the mask boundary is revised and the priority values are updated. Hence the highest priority structure in the previous iteration will not be guaranteed the same priority value. The continuity of the structure is not maintained in the process flow. Whereas in the proposed work, in every iteration, a complete superpixel is extended inside the mask region and thus preserves the continuity in structures. Fig. 11 compares the first iteration step for the test image in the Exemplar-based inpainting method and the proposed superpixel based structure propagation. The Exemplar-based inpainting uses the CIE color space for the working but we have tried our algorithm in the normal RGB space.

The future work will take care of the priorities for the structure intersection inside the mask region. The weight function for maintaining the global approach of the structure intersection inside the mask region is necessary for the successful completion of this proposed inpainting work. Smoothening [20] the superpixel contour seems tobe an another important issue before applying the spline interpolation. Sometimes the boundaries of the superpixels are not regular and there are chances for getting the incorrect tangents for the superpixel edges that are in contact with the mask boundary. The superpixel boundary smoothing will help to get the correct orientation of the end points which will lead to the proper curve completion.

As stated in the subsection 4.3, the grouping of the superpixel groups is also an another important task for our proposed method. The texture properties of the superpixels are analyzed and the similar superpixels are grouped for defining a complete structure. On the whole, our future work will aim for trimming the algorithm to further dimensions in order to handle the variety of images.

REFERENCES

1) T. Ishikawa, K. Thangamani, M. Kourogi, Andrew P. Gee, Walterio Mayol-cuevas, Keechul Jung and Takeshi Kurata, "In-situ 3D indoor modeler with a camera and self contained sensors", HCII2009, LNCS **5622,** pp.454-464.
2) T. Kurata, M. Kourogi, T. Ishikawa, J. Hyun and A. Park, "Service Cooperation and Co-creative Intelligence Cycles Based on Mixed-Reality Technology", In Proc. INDIN 2010, pp. 967-972.

3) A. Criminisi, P. Perez and K. Toyama, “Object Removal by Exemplar-Based Inpainting”, In Proceedings of IEEE Computer Vision and Pattern Recognition, CVPR(2003), Vol **2**, pp. 721-728.
4) K. Thangamani, T. Ishikawa and T. Kurata, “Inpainting for Interactive 3D Indoor Modeling”, ISUVR2010.
5) R. Cipolla, T. Drummond and D. Robertson, “Camera calibration from vanishing points in images of architectural scenes”, BMVC99 http://www.bmva.ac.uk/bmvc/1999/papers/38.pdf
6) Carsten Rother, “A new approach to vanishing point detection in architectural environments”, Image and Vision Computing **20**(2002), pp. 647-655.
7) Jason C. Hung, Chun-Hong Hwang, Yi-Chun Liao, Nick C. Tang, Ta-Jen Chen , “Exemplar-based image inpainting base on structure construction”, Journal of software, Special issue on “Mobile System, Agent Technology, and Web Services”, Vol.**3**, No.**8**, November 2008, pp. 57-64.
8) Jian Sun, Lu Yuan, Jiaya Jia and Heung-Yeung Shun, “Image completion with structure propagation”, SIGGRAPH 2005, Vol. **24,** pp. 861-868.
9) Andrei Rares, Marcel J. T. Reinders and Jan Biemond, ”Edge-based image restoration”, IEEE Transactions on Image Processing, Vol.**14**, No.10, October 2005, pp. 1454-1468.
10) Nikos Komodakis and Georgios Tziritas, “Image completion with global optimization”, In Proceedings of IEEE Computer Vision and Pattern Recognition, CVPR 2006, pp. 442-452.
11) Jiaya jia and Chi-Keung Tang ”Image Repairing:Robust image synthesis by Adaptive ND Tensor voting”, In Proceedings of IEEE Computer Vision and Pattern Recognition, CVPR 2003, pp. 643-650.
12) Anat Levin, Assaf Zomet and Yair Weiss, “Learning how to inpaint from global image statistics”, Proceedings of the ninth IEEE conference on computer vision, Vol **2**(2003), pp. 305.
13) X. Ren and J. Malik. “Learning a classification model for segmentation”, In Proc. 9th Int. Conf. Computer Vision, 2003, Vol. **1**, pp. 10-17.
14) Derek Hoiem, Alexei A. Efros, Martial Hebert, " Automatic Photo pop-up", ACM Transactions on Graphics(TOG), Vol **24**, Issue **3**, pp. 577-584.
15) Pedro F. Felzenszwalb and Daniel P. Huttenlocher, ”Efficient Graph-Based Image Segmentation”, International Journal of Computer Vision, September 2004[2], Vol. **59**, No. **2**.
16) Luigi Atzori and Francesco G. B. De. Natale, ”Reconstruction of missing or occlusded contour segments using Bezeir interpolations”, Signal processing **80** (2000), pp. 1691-1694.
17) A. A. Efros and T. K. Leung, “Texture synthesis by non-parametric sampling”, In International conference on computer vision, Corfu, Greece, September 1999, pp. 1033-1038.
18) A. A. Efros and W. T. Freeman, “Image quilting for texture synthesis and transfer”, SIGGRAPH 2001, pp. 341-346.
19) Vivek Kwatra, Arno Schodl, Irfan Essa, Greg Turk and Aaron Bobick, “GraphCut textures: image and video synthesis using graph cuts”, In ACM Transactions on Graphics (SIGGRAPH), Vol. **24(3)**, July 2003, pp. 277-286.
20) Pornchai Mongkolnam, Thanee Dechsakulthorn and Chakarida Nukoolkit, “Image shape representation using curve fitting”, Proceedings of the 6th WSEAS International conference on signal, speech and image processing, Lisbon, Portugal, 2006, Sep 22-24.