MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Superpixel Based Inpainting for Interactive 3D Indoor Modeler

Kalaivani Thangamani[1, 2]
[1]University of Tsukuba
[2]Center for Service Research, 1-1-1 Umezono, AIST, Tsukuba 305-8568 Japan.
thangamani.kalaivani@aist.go.jp

Takeshi Kurata[1, 2]
[1]University of Tsukuba
[2]Center for Service Research, 1-1-1 Umezono, AIST, Tsukuba 305-8568 Japan.
t.kurata@aist.go.jp

Tomoya Ishikawa[2]
[2]Center for Service Research, 1-1-1 Umezono, AIST, Tsukuba 305-8568 Japan.
tomoya-ishikawa@aist.go.jp

Abstract

This paper discusses the new inpainting algorithm which uses the superpixel constraints with the Exemplar-based inpainting technique for bringing the realistic view in the 3D model developed by the interactive 3D indoor modeler. The inpainting algorithm is exclusively developed for the interactive 3D 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. Inpainting techniques may be employed 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 the new inpainting technique that combines the Exemplar-based inpainting with the superpixel constraints in order to improve the structure propagation since the structures are important for indoor images.

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. The interactive modeler [1, 2] enables 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. This paper tested the Exemplar-based inpainting technique [3] which is a simpler and faster method for texture filling the untextured regions. 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. The paper proposes the usage of superpixel constraints in improving the structure propagation part in the Exemplar-based inpainting method.

The second section briefly discusses the scenario of the interactive indoor modeler, the need for inpainting and the automatic generation of the inpainting mask. The third section explains the proposed algorithm and the subsections explain the concept of superpixels and their role in improving the structure propagation. The fourth section explains the Exemplar-based inpainting, its application in the interactive modeler. The later part of this section summarizes the related works in the area of structure propagation. The fifth section experiments the proposed algorithm in the modeler images and compares the Exemplar-based inpainting.

2. Interactive 3D indoor modeler and Inpainting

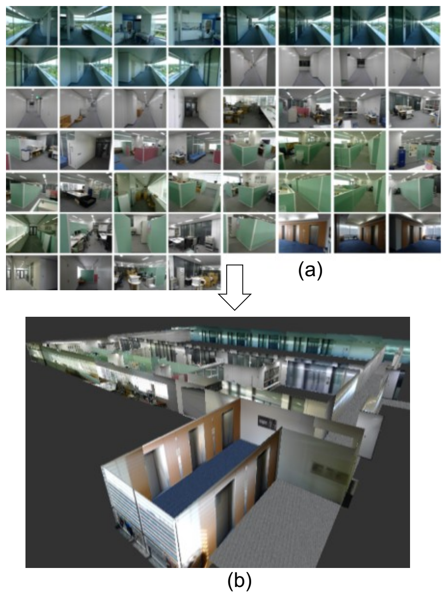

Figure 1. Interactive 3D indoor modeling [1] a. Multiple photos b. Indoor model made from the multiple photos

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 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. 1a shows the multiple photos taken for the indoor environment, and are used to model the indoor environment which is shown in Fig. 1b. 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. Need for Inpainting

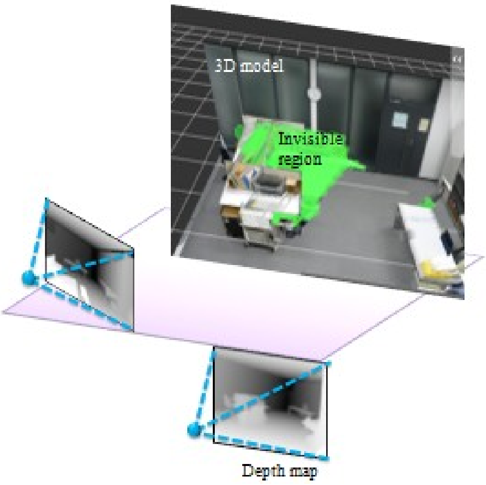

Figure 2. Inpainting mask generation

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.

3. Algorithm overview

The aim of the superpixel based inpainting is to guarantee the structure propagation and reduce the computation time. The overview of our proposed superpixel constraints for structure propagation is shown in Fig. 3. The block diagram in figure 3 explains the Exemplar-based inpainting method and the introduction of superpixel constraints. 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. 4c shows the superpixels that are in contact with the mask boundary for the test image in Fig. 4a. The textures for the contact superpixels are shown in Fig. 4d. The superpixel that is in contact with the highest priority boundary point is propagated inside the mask region based on their category. Categorizing the superpixel is explained briefly in the subsection. If the superpixel is found to be of type primary, the edges are smoothly joined by spline curves. Also the spline curves are used for smoothing the superpixel boundary in order to calculate the exact orientation of the edges. If the superpixel is found to be of type secondary, the edges are extrapolated along its direction with respect to the weight function which determines the distance of the structure propagation from the real texture. Also the weight function handles the intersection of two or more structures inside the mask region, in some cases. Once the stucture skeleton is formed, then the skeletons are filled with the texture flesh. These skeletons get the texture from the corresponding superpixel by texture synthesis [4, 5]. The tedious patch search is improved by means of search by hashing. Including the superpixel constraints reduces the number of iterations and speed up the inpainting process.

Exemplar-based method fills a single patch in each iteration, whereas the superpixel based boundary extension fills the wide region that is monitored by the superpixel. The process is repeated till the entire mask gets inpainted.

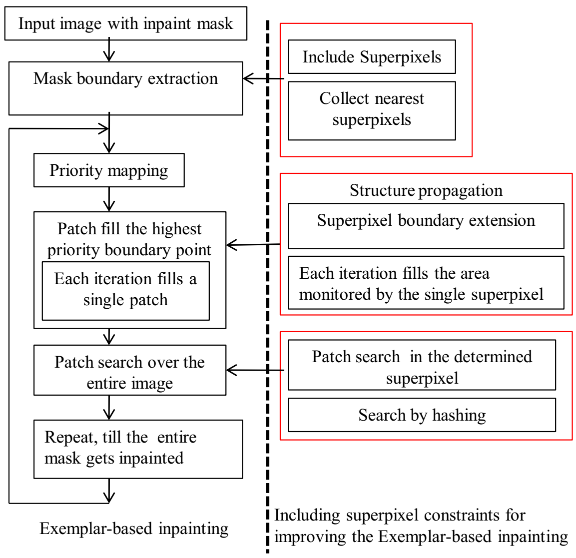

Figure 3. Including superpixel constraints for improving the Exemplar-based inpainting technique

3.1. Superpixels

Superpixels correspond to small, nearly-uniform segmented regions in the image. The idea of superpixels is originally developed by X. Ren and J. Malik [6], 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 of using the superpixels in the modeling is obtained from Derek Hoiem et al. [7] from their work Automatic photo pop-up where they use the superpixels for categorizing the horizontal and vertical planes in the given image. The efficient graph based segmentation [8] 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 methods in traditional segmentation algorithms. Fig. 4b shows the superpixels for the test image shown in Fig. 4a.

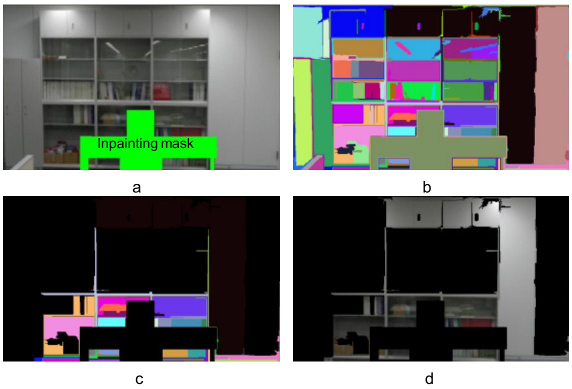

Figure 4. Superpixels for inpainting a. Image with inpaint mask b. Superpixels for image in 4a c. Superpixels that are in contact with the mask boundary d. corresponding textures for 4c

3.2. Categorizing the superpixels

The boundary of the superpixel edges that are in contact with the inpaint mask is denoted by the term edge couple. The nature of the edge couple is the key measure in structure propagation. There are three types of edge couples observed in the experiment. If the superpixel boundary encloses the inpainting mask, they are called the primary superpixels. If the superpixel edges just touch the mask, there is a possibility for the superpixel to get continued inside the mask region. They are called the secondary superpixels. And there are discarded superpixels which are very small, do not hold any meaningful textures are merged with the neighboring superpixels.

4. Exemplar-based Inpainting

The Exemplar-based inpainting proposed by Criminisi et.al [3], is tested 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. Fig. 6b shows the 3D model which consists of a floor, walls and other objects (table and a chair) created by the modeler from the input image which is shown in Fig. 6a. 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. The inpainting mask for one of the plane is shown in Fig. 5a and the application of the Exemplar-based method is shown in detail in Fig. 5b. The inpainting handles the patch size of 40x40 with 82 iteration steps to fill the inpainting mask present in the image. The inpainted result is shown in Fig. 5c**.**

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. The Exemplar-based method follows the greedy way of sampling for filling the textures. The structures are not treated globally. Treating the structures in the global aspect and fixing the structure skeleton before filling the textures leads to the guaranteed structure propagation. There are quite a good number of research studies in the area of structure propagation in inpainting. Jason C. Hung et al. [9] extend the Exemplar-based inpainting by incorporating Bezier curves to construct missing edge information. Jian Sun et al. [10] referred the user drawn lines for the possible structure propagation. Andrei Rares et al. [11] 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. [12] proposed an alternate image completion method for the Exemplar based method with the Markov Random Field (MRF). There are good references on tensor voting [13] for the robust image synthesis. Treating the image in the global aspect recommends the good structure propagation [14].

5. Experiments

Fig. 5 compares the inpainting process flow between the Exemplar-based method and the superpixel based inpainting. The Exemplar-based inpainting follows the tedious patch search procedure. For example, the example image shown in Fig. 5a contains 358776 number of texture patches of size 40x40. For every iteration, the successful patch is chosen from checking the SSD of all the 358776 texture patches from the image. This consumes a lot of computation time.

The patch size determines the quality of the inpainted result and it depends on the texture and structure nature of the particular image. The user has to adjust the patch size to be slightly larger than the largest distinguishable texture element, or “texel”, in the source region. When the patch size is not adjusted properly, it leads to the vis ible patch boundary which altogether makes the inpainted region odd from the rest of the image.

To overcome these drawbacks, the superpixel constraints are included in the Exemplar-based inpainting technique. Fig. 5e shows the priority superpixel which is of type secondary got propagated inside the mask region and the result after iteration 1 is shown in Fig. 5f. The boundary extrapolation is guided by the orientation of the edge couples and also by the weight function. The entire image is inpainted in just 10 iterations.

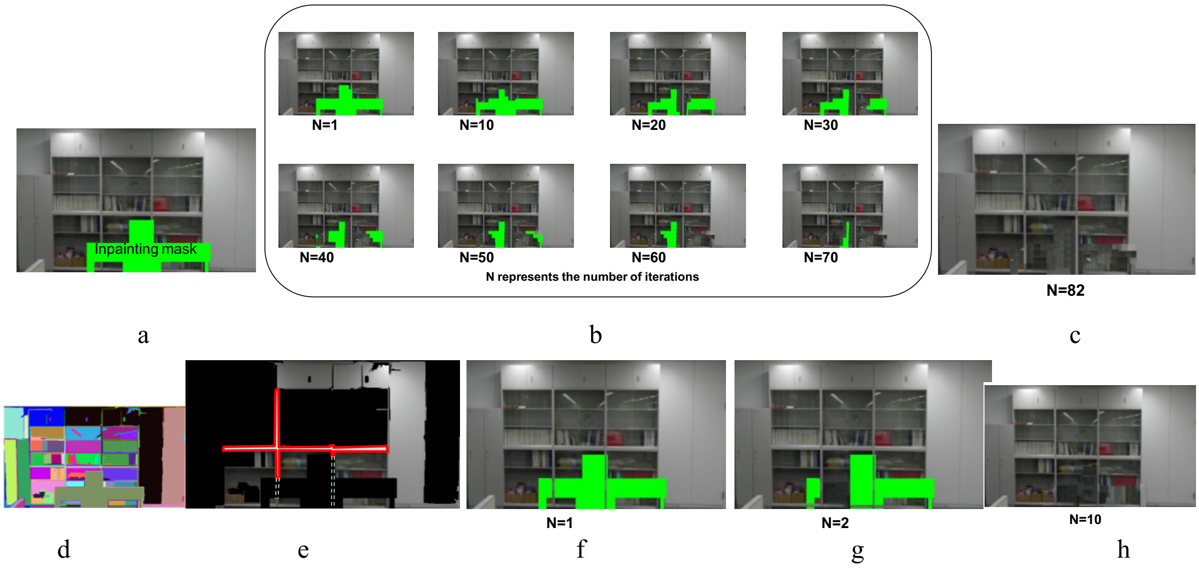

Figure 5. Comparison between the Exemplar-based inpainting (top row) and the superpixel based inpainting (bottom row) a. Plane in the indoor model (inpaint mask shown in green) b. Iteration steps in Exemplar-based inpainting c. Result of Exemplar-based inpainting (in 82 iterations) d. Superpixels e. Priority superpixel for first iteration and the boundary extension f. After first iteration g. After second iteration h. Result of Superpixel based inpainting (in ten iterations)

6. Conclusion and Future works

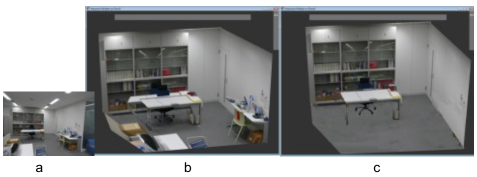

Figure 6. a. Input Image b. 3D Indoor model c. Inpainted 3D model

To include the probabilistic, stochastic and statistical elements for evaluating the efficiency of the superpixel based structure propagation. The length, width and the texture of each structure are the good evidences for improving the structure propagation. Also including the user interaction to handle the inpainting data will improve the quality of the inpainted images. Testing the structure propagation in a variety of images would help to improve the performance of the proposed work.

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 proceedings of INDIN 2010, pp. 967-972.

[3] A. Criminisi, P. Perez and K. Toyama, "Object Removal by Exemplar-Based Inpainting", Proceedings of IEEE CVPR, vol. 2, pp. 721-728, 2003.

[4] A. A. Efros et al., "Texture synthesis by non-parametric sampling", In International conference on computer vision, Corfu, Greece, pp. 1033-1038, September 1999.

[5] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer", SIGGRAPH 2001, pp. 341-346.

[6] X. Ren and J. Malik. "Learning a classification model for segmentation", In Proc. 9th Int. Conf. Computer Vision, vol. 1, pp. 10-17, 2003.

[7] Derek Hoiem, Alexei A. Efros, Martial Hebert, " Automatic Photo pop-up", ACM Transactions on Graphics(TOG), vol. 24, no. 3, pp. 577-584.

[8] Pedro F. Felzenszwalb and Daniel P. Huttenlocher, "Efficient Graph-Based Image Segmentation", International Journal of Computer Vision, vol. 59, no. 2, 2004.

[9] 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, pp. 57-64, 2008.

[10] Jian Sun, Lu Yuan, Jiaya Jia and Heung-Yeung Shun, "Image completion with structure propagation", SIGGRAPH2005, vol. 24, pp. 861-868.

[11] Andrei Rares, Marcel J. T. Reinders and Jan Biemond, "Edge-based image restoration", IEEE Transactions on Image Processing, vol.14, no.10, pp. 1454-1468, 2005.

[12] Nikos Komodakis and Georgios Tziritas, "Image completion with global optimization", Proceedings of IEEE CVPR, pp. 442-452, 2006.

[13] Jiaya Jia and Chi-Keung Tang "Image Repairing: Robust image synthesis by Adaptive ND Tensor voting", Proceedings of IEEE CVPR, pp. 643-650, 2003.

[14] A. Levin, A. Zomet and Y. Weiss, "Learning how to inpaint from global image statistics", Proceedings of the ninth IEEE conference on computer vision, vol. 2, pp. 305, 2003. Journal of Computer Vision, vol. 59, no. 2, 2004.