Received October 31, 2017, accepted November 28, 2017, date of publication December 18, 2017, date of current version February 14, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2784798

Procedural Restoration of Texture and Restructuring Geometry From Facade Image

J. DIVYA UDAYAN[1], (Member, IEEE), AND HYUNGSEOK KIM[2], (Member, IEEE)
[1]School of Information Technology and Engineering, VIT University , Vellore 632014, India
[2]Department of Software, Konkuk University, Seoul 143-701, South Korea

Corresponding author: Hyungseok Kim (hyuskim@konkuk.ac.kr)

This work was supported by the Bio-Synergy Research Project of the Ministry of Science, ICT and Future Planning, through the National Research Foundation under Grant NRF-2013M3A9C4078140.

ABSTRACT We present a novel methodology to procedurally restore facade texture from images, where severe occlusions and lighting variations occur, making the textures difficult to reconstruct effectively. A refinement strategy is designed, which includes an iterative weighted-average algorithm, to restore and update the high-quality consensus texture of the targeted building. Our approach combines component-based structure analysis and rule-based texture recovery techniques that are suitable for removing the occluded areas from a given building facade. We demonstrate our framework on several real-world buildings with varying amounts of occlusion, and show that our approach can be used to generate 3-D buildings with far more aesthetic quality than the previous approaches.

INDEX TERMS 3D reconstruction, feature extraction, image processing, image texture analysis.

I. INTRODUCTION

3D models of city buildings have been used in various applications such as movies, gaming and urban architectural design. All such applications demand 3D models, which have life-like appearances and provide enhanced immersion and reliable simulation results with lower costs.

Previous approaches for constructing 3D building models involve estimating 3D depth information with a foreground image and generating integrated model data with it. The main limitation with such approaches is caused by noises that are not easily predicted in advance, such as irregular lighting, smoke, signboards, trees, cars and pedestrians. These kinds of noises are not easily distinguishable from the image alone and tend to cause incorrect results in the 3D geometry of the building model.

In this paper, building patterns are extracted from architectural components in the image. Noise has a specification of low repeatability compared to architectural components. As a first step, the positions of architectural components are estimated and extracted. From the extracted components, structural clusters and repeated symmetries are determined. From the estimated data, candidate areas for the presence of noise are identified. Among candidate areas, noisy areas are iteratively selected and updated by replacing them with reasonable data according to patterns from the structural and repeatability information. The structural and symmetry information is iteratively updated with the newly filled areas.

II. PREVIOUS WORK

Currently, intensive research [1], [2] on refining texture from images is being conducted; occlusion is the main cause that prevents the correct discovery of the attributes of buildings from the input images.

In previous research [3], the correlations between the correctness of an original image and of the resulting 3D model generated by image-based approaches have been studied.

Facade texture restoration and restructuring geometry is the key to success for 3D building reconstruction and modeling. Interpolation methods using multiple-view geometry have been proposed for texture fusion/recovery [4] and inpainting techniques [5]. The main drawback of these methods is that they do not handle occlusions automatically.

In the literature [6], solutions for determining occlusions caused by regular, modeled structures but not unmodeled structures have been suggested. A solution for unmodeled occlusions was given by [7]. However, this method still has limitations that may cause blurred or disrupted boundaries of structures.

Another method for recovering facade texture and microstructure from real world images was proposed by Wang and Hanson. [6]. Their work describes a new

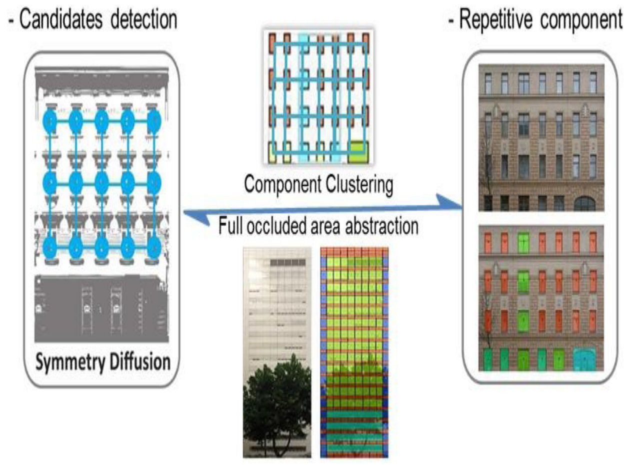

FIGURE 1. The combined process of repetitive component analysis and structural analysis. (Left) Symmetric structure analysis, (Right) Repeated component detection, (Middle) Component clustering.

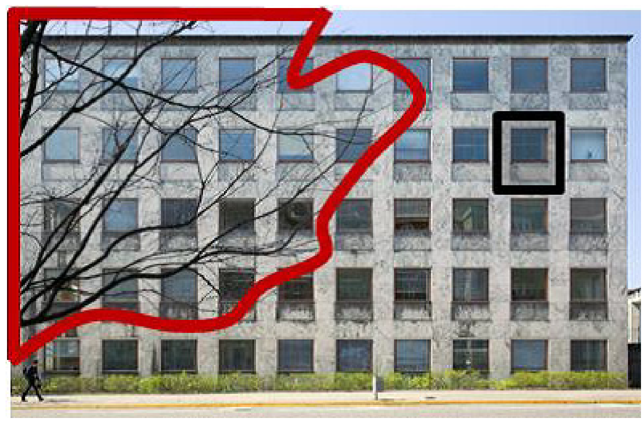

FIGURE 2. The Facade image showing noise (red outlier) and highlighted window pattern (black).

methodology to obtain a realistic facade texture map by eliminating occlusions and the effects of varying illumination from an image. This method takes as input a set of images of buildings taken from different sides of the buildings with reasonably accurate, but not exact, camera pose information and a coarse geometric model of the buildings.

Adding the assumption that building components follow repetitive patterns [8], these methods can provide convincing results. The remaining issues are the dynamic parts of the facade, including windows and balcony parts. Repetition and symmetric structure are not easily applied to the dynamic parts. Thus, those parts are often categorized as occluded parts. In this research, we propose a new method of identifying both dynamic and static parts of a façade as structural parts while also correctly identifying noise parts.

III. REFINEMENT STRATEGY

In this paper, we introduce a methodology for removing the noise from a building facade image. The repair process combines two analysis methods; Figure 1 shows this combined process of repetitive component analysis and structural analysis. The entire process consists of two steps. First, we analyze the facade components and classify the component and noise regions. Next, patterns are defined based on architectural structure.

There are two different kinds of occlusion states, and hence, different refinement strategies need to be devised for each of kind. The first occlusion state has a partially occluded building image from which a certain amount of outlier information can still be extracted. The second state has a severely occluded building image from which it is difficult to extract meaningful structural information. The refinement strategy for the first case is to perform color correction, and for the second case refinement proceeds by identifying and working with the components.

Compared with other texture reconstruction approaches [9], eliminating noise factors may cause poor 3D reconstruction results even though the appearance of the resulting model is enhanced. Furthermore, it is expected that minimizing, rather than eliminating, the noise factors will enhance accurate 3D building reconstruction.

IV. METHODOLOGY

In this section, we define patterns based on architectural structure, as shown in Figure 2. Through image analysis, candidate components are extracted. Among the extracted components, similar parts are collected. The part which is closest to the average of the collected parts is selected as the representative part. Part similarity is defined as image similarity in the regional margins (static region) and difference in the interior region (dynamic region). For texture reconstruction, the static and dynamic regions in each component are analyzed to identify partially occluded components. The structure between candidate components is analyzed to identify entirely occluded components.

A. PATTERN DEFINITION BASED ON ARCHITECTURAL STRUCTURE

We divide the facade into elements using the repetitive structures algorithm [10], as shown in Figure 3, and extend the area using structural information. The similarity relationship between each element and the rest of the area is analyzed by repetitive component similarity analysis [8]. Elements are clustered into larger parts. Each element is analyzed in a bidirectional manner.

Figure 3 shows the similarity relation of each element (input element) to all other elements (output elements). With these relations, we can divide the façade into a set of clusters, each of which has large correlations among its elements. We use intra-relation degrees of each node in a similarity graph as a weight for clustering. Candidates having the highest number of relations in each cluster are defined as seeds, and candidates having a lower number of relationships are classified as noise. Figure 3 shows an example of clustering.

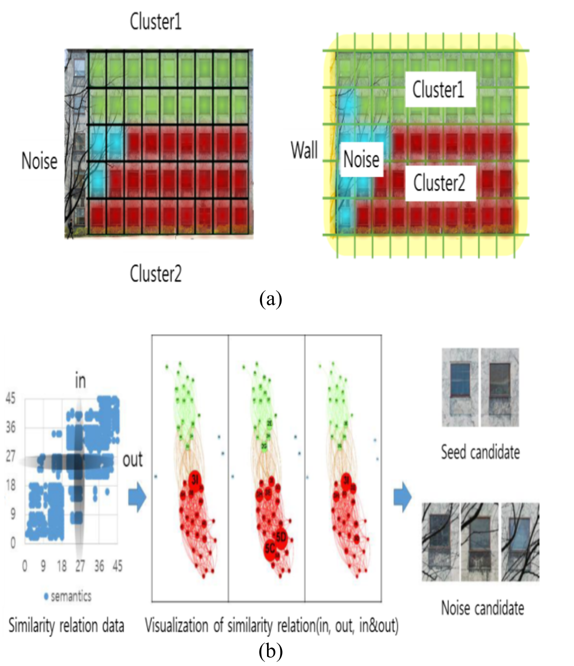

FIGURE 3. Component extraction process example. (a) Classification by clusters. (Left) Initial component division structure. (Right) Facade division structure by grid (b) Facade component analysis process.

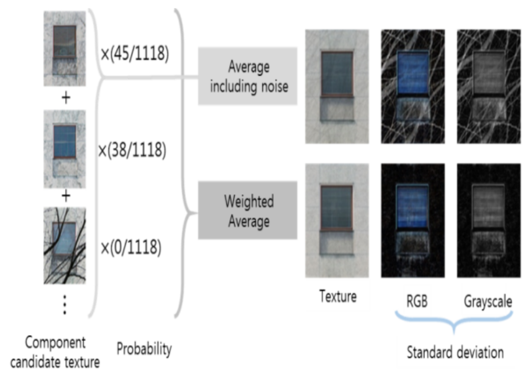

FIGURE 4. Generation of texture for a component (top) using uniform weight for averaging (bottom), applying less weight for noisy components.

B. SPECIFICATION DEFINITION THROUGH ANALYSIS OF COMPONENT CANDIDATES

Figure 4 shows the result of selecting one component after lowering the degree of cluster partitioning to define the component characteristics. Component characteristics are common properties, excluding noise and other diversity of the candidates in the cluster. Here, we divide the component parts based on their diversity determined according to the average texture of the cluster.

Figure 5 shows the process of separating regions into dynamic and static parts. It mixes two methods of thresholding. To detect dynamic parts, we use a locally adaptive thresholding method. Since the dynamic parts have large color differences, it is necessary to apply a local threshold for each pixel location.

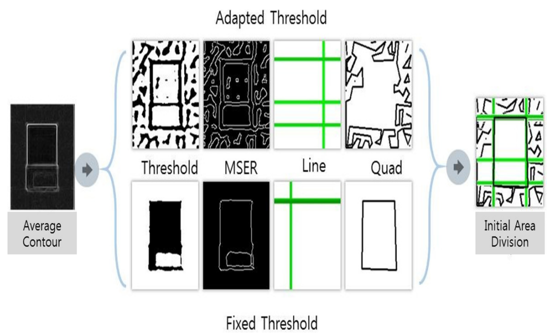

FIGURE 5. Component area division process by a combination of various threshold methods.

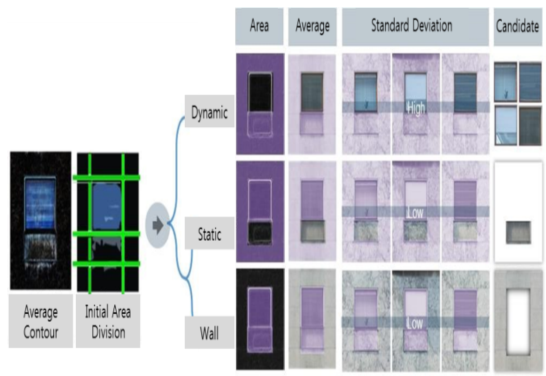

FIGURE 6. Component layer division.

A global fixed thresholding method is used to identify the static regions. Static regions have relatively regular color distributions and structures. Thus, it is natural to use a fixed threshold to correctly identify these regions.

The average of the standard deviations of each region of the initial area division is binarized to separate the static region and the dynamic region, as shown in Figure 6 In the static region, the area lying on the outer side of the component is divided into walls. Both the texture in the dynamic region of the candidates and the weighted average texture of the wall and the static region are used for restoring the partially occluded texture.

C. STUCTURE ANALYSIS BETWEEN COMPONENTS (PATTERNS)

We use a rule-based architectural modeling system [3], in which the tile unit is the minimum unit of the building component, to devise a grammar. Using the partitioning rule, it is possible to restore over the noisy area by extension, as shown in Figure 7. This methodology also analyzes the vertical and horizontal adjacency structure, which is the simplest architectural structure based on horizontal division of floor units and vertical division of component units. The adjacency relationship pattern as shown in Figure 8 is this analysis result excluding the noise. The number of adjacencies in the vertical (V) and horizontal (H) directions is used as the weight of the pattern.

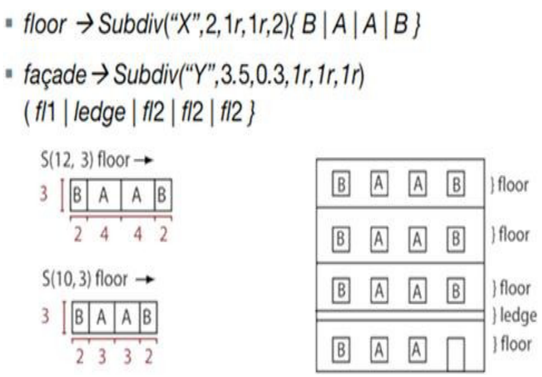

FIGURE 7. Rule expansion example.

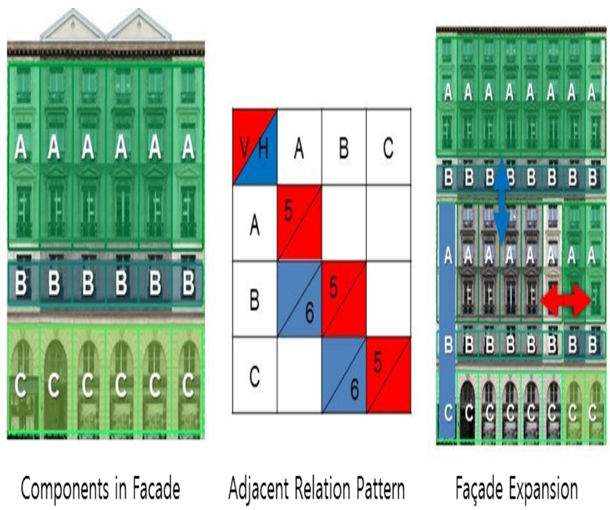

FIGURE 8. Adjacency relationship of components.

The extension process of the partition rule defined by the adjacency relationship is shown in Figure 8. If a pattern A is adjacent with another pattern A along its right vertical edge, the V value of the adjacency matrix will be increased by one. If a pattern A is adjacent with a pattern B along its bottom horizontal edge, the H value of the adjacency matrix will be increased by one. After adding the bottom, top, right and left cases for each component, the component with the highest probability is selected as a candidate for filling noise parts. In Figure 8, the noise part of the right-most image is the left bottom area. According to the relation weights, B could be filled in the center as B has always been adjacent with B. Next, C could be filled with relation to another component C and the filled component B. Finally, A is filled as it could satisfy both the relations with B below and A on its right.

D. TEXTURE RESTORATION

This section describes the process of removing environmental noise from the façade texture. The noise is a set of partially occluded components. There are two ways to restore noise depending on the degree of occlusion as shown in Figure 9 that illustrates the overall pipeline for texture reconstruction. Noise parts in façade (N in the left figure) are subdivided into regular size components (α, β). Based on adjacent relation, each occluded component is estimated with candidate symbols (A,B,C). For each candidate symbol, the occluded symbol is identified with the one that mostly matches with the image on the occluded components. When the component is compared with the image parts for matching, the component is divided into two parts; common static layer and dynamic layer. Dynamic layer is a part corresponds to windows or doors which could be different for each component such as half open, full open, or closed. When images are matched with symbol template, those dynamic parts are masked out from the matching. From the matching result, the occluded symbols are estimated with the best matching ones (C and A in the right figure).

The restoration progresses next to the elements having no relation or few relations in pattern. First, we first try to select a component using the element's texture, due to the probability of preserved information. Since a partially occluded component mainly has effects of changing color values, the similarity is defined using the feature points of the contour data. As previously defined, components have static and dynamic regions. Higher accuracy results from checking feature points in the static region of texture than in the dynamic region. If the highest similarity exceeds the standard occlusion probability, it is estimated to be an alternative component, and the pattern is updated. In other cases, the element is an entirely occluded component. In this second case, for entirely occluded components, we try to select a component using probabilistic inference about neighbor components. This matching procedure is detailed in Figure 10. The partially occluded part (C on the top left; green part is the occluded part) is computed its matching score with all candidate symbols with dynamic mask (A, B, C on the top image). The symbol with best matching number of features (C in Figure 10) is selected for the best matching one. When the best matching symbol is found, its adjacency relation is also updated. For the fully occluded components (bottom line of Figure 10), matching symbol is found by selecting from the adjacent relation table. In Figure 9, part β is adjacent vertically with B (B4) and horizontally with newly recovered C. From the table, both A and C could be vertical neighbor of existing symbol B. Among A and C, A is the only symbol that could adjacent with newly recovered C. Thus, A is selected as the best matching symbol for the part.

After selecting the closest matching components, the relation matrix is updated with the new components. Texture restoration is performed for elements with low weights in similarity relationships, as in Figure 3, which includes the noisy parts. For texture repair, we propose a method to maintain the minimum diversity and remove noise using the estimated information.

As shown in equation (1), each original element (O) has its own unique specification (U), compared with the average texture of component (Avg).

$$Avg + U = O \tag{1}$$

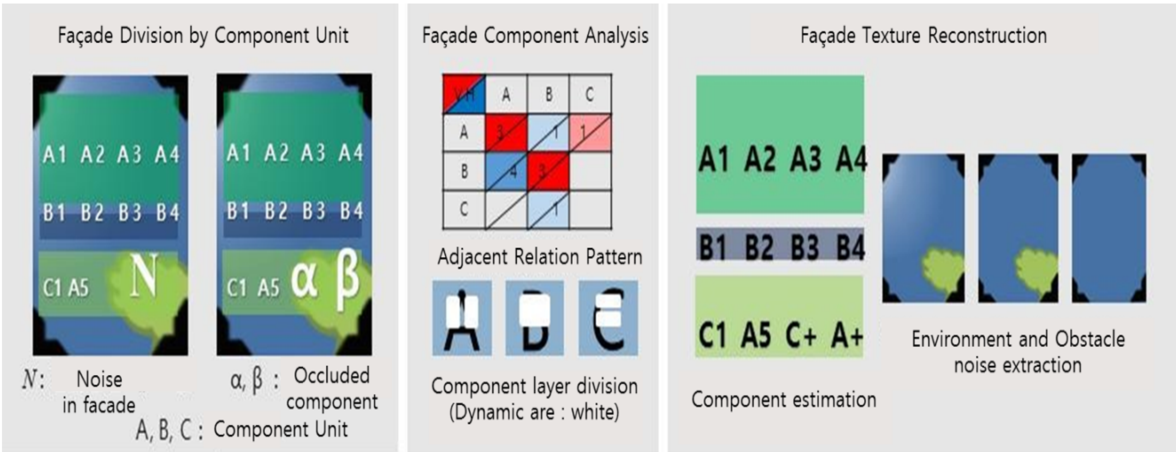

FIGURE 9. Pipeline of texture reconstruction process.

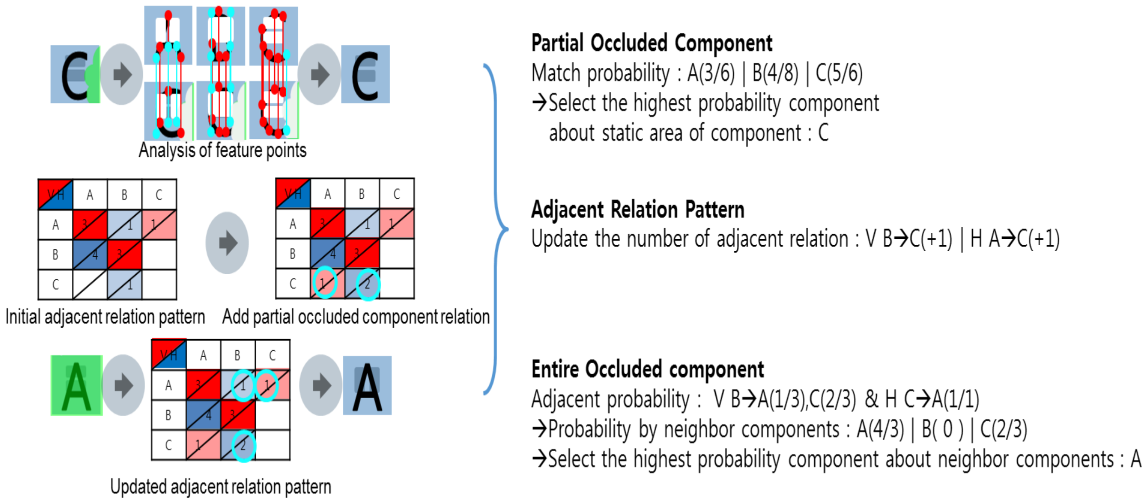

FIGURE 10. Component estimation process.

The input texture (I) includes noise on the original element. As previously defined, environment noise (EN) due to environmental causes has low transform of contour from O. Conversely, obstacle noise (ON) occludes the components by adding additional values. From this observation, we use the following simplified equation for the image.

$$I = O \times EN + ON \tag{2}$$

From equation (1), we can expand this equation as

$$(Avg \times EN) + (U \times EN) + ON = I \tag{3}$$

Environmental noises can be removed using an edge detector such as the Sobel edge detector. Detected edges can be modeled by the following equation, where D is defined as the difference between I and Avg.

$$|I - Avg| = D$$
$$\because (Sobel(D) = Sobel(I - Avg) = Sobel(U + ON) \tag{4}$$

The area of every element (X) is the sum of the dynamic area (DA) and the static area (SA), including the wall.

$$X = X_d\ (= X \cap DA)\ X_S\ (= X \cap SA)\ (DA \cap SA = \varphi) \tag{5}$$

Except for the DA, the standard deviation (Dev) of the contour of U for the entire texture is relatively less than the Dev of the contour of ON. If the result of Sobel applied to the SA is less than the standard value (Integer), it is regarded as the original element specificity and is stored ($Saves$). If it exceeds the standard value, it is extracted as the contour of the noise ($Extracts$).

$$Dev\,(Sobel\,(Us)) < Dev\,(Sobel\,(ONs)) \tag{6}$$

hence,

$$(Dev\,(Sobel\,(Ds)) < Fixedthres\ hold)?$$
$$Sobel\,(Us) \rightarrow Saves : (Sobel\,(ONs) \rightarrow Extracts)$$

Maximally stable extremal regions ($MSER$) are estimated to extend the stable region. It extends to high occupancy seg-

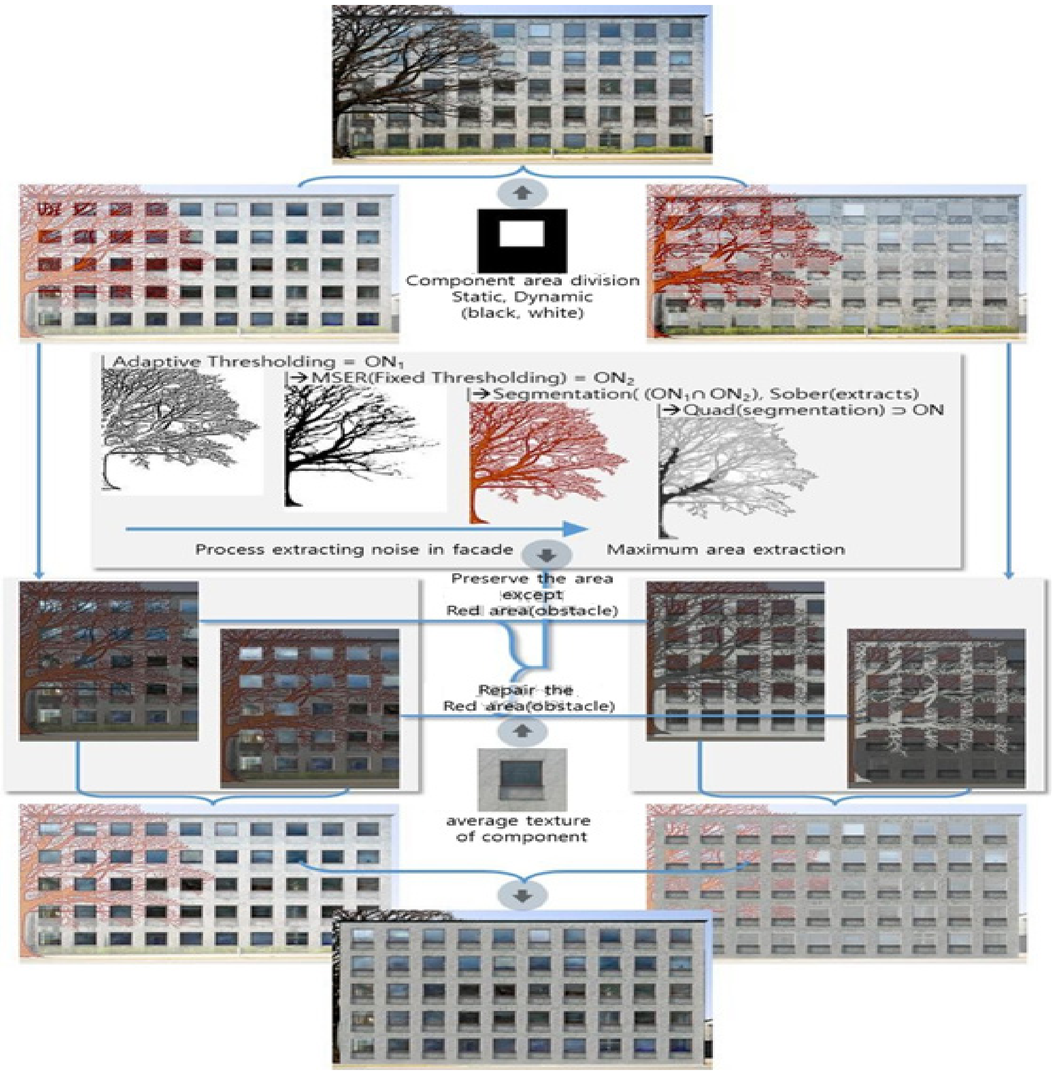

FIGURE 11. **Facade texture restoration example process.**

mentation in the area divided by the outline of *ON*. The area including the seed point of the obstacle noise (*Fill*) is defined as the expected obstacle noise area (M) including the obstacle noise. In equation (7), *FT(D)* is the fixed thresholding result of *D*.

If excluding regions from the input texture, the latter is included in the former. Only the *EN* of the *O* remains for the area excluding the estimated *ON:*

$$(ON) \subset (Fill\,(MSER\,(FT\,(D)) \cap Sobel\,(ON)) = M) \quad (7)$$

M is divided into area in *SA* (M_s) and area in DA (M_d):

$$Fill\,(MSER\,(FT\,(D_S)) \cap Extracts) = M_S$$
$$Fill\,(MSER\,(FT\,(D_d)) \cap MSER\,(FT\,(I_d))) = M_d$$

Thus, it could be stated that

$$\left(I \cap M^c\right) \subset \left(I \cap (ON)^c\right)$$
$$(I \cap M^c = (I - ON) \cap M^c = (O \times EN) \cap M^c) \times \left(EN_s \cap M^c\right)$$

The color information is replaced by the adjacent pixels for the detected singularities using adaptive thresholding (*AT*), applied to the dynamic region, which is the intersection of *D* and the complement of *SA* in the estimated M. The *EN* of the corresponding region is extracted by smoothing (*Sm*) the adjacent region. It is possible to restore the *SA* in *I* for the area excluding the estimated obstacle area due to the environmental noise extracted as in equation (8)

$$Sm\left(Fill\left(D \cap \left(AT\left(D \cap M_s^c\right)\right)\right)^c\right) = EN \cap M_s^c$$
$$\because (I_s \cap M_s^c) \div (EN \cap M_s^c) = O_s \cap M_s^c \quad (8)$$

For the I of *DA*, candidates (C) with the least standard deviation of the contour are selected except for the area of the obstacle noise region.

The candidate is the *DA* texture of the currently highest weighted element and "#" is a symbol representing the number of elements in the set.

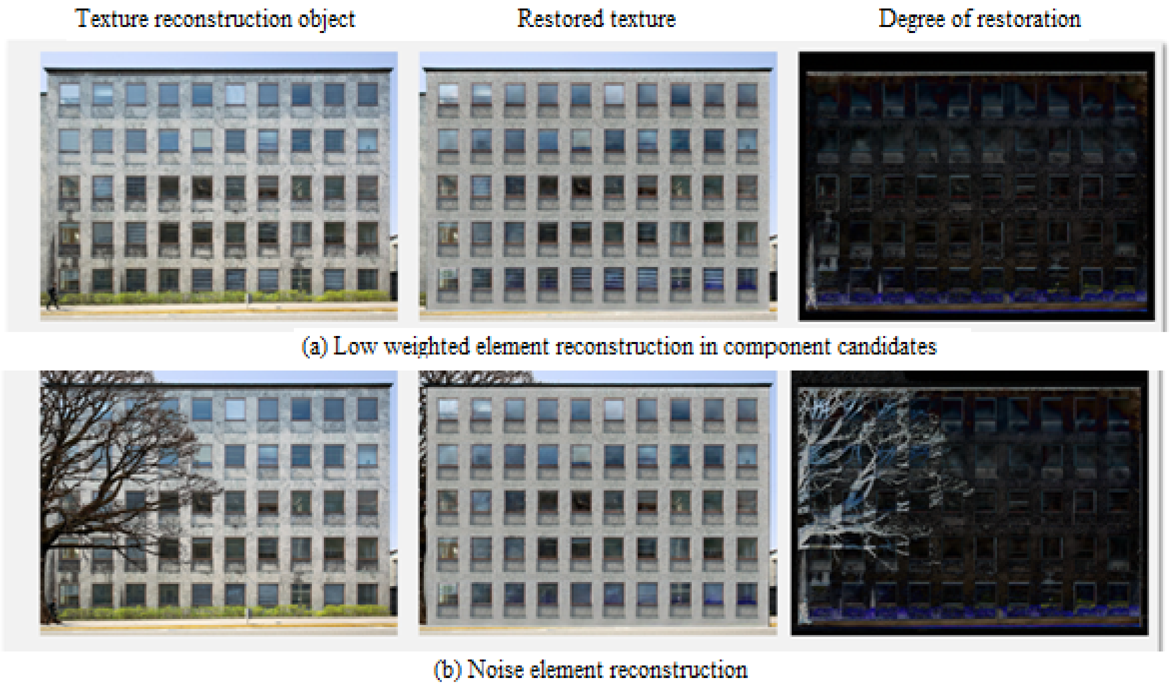

FIGURE 12. **(a) Restored result (middle image) for low noise, (b) Restored result for large noisy parts.**

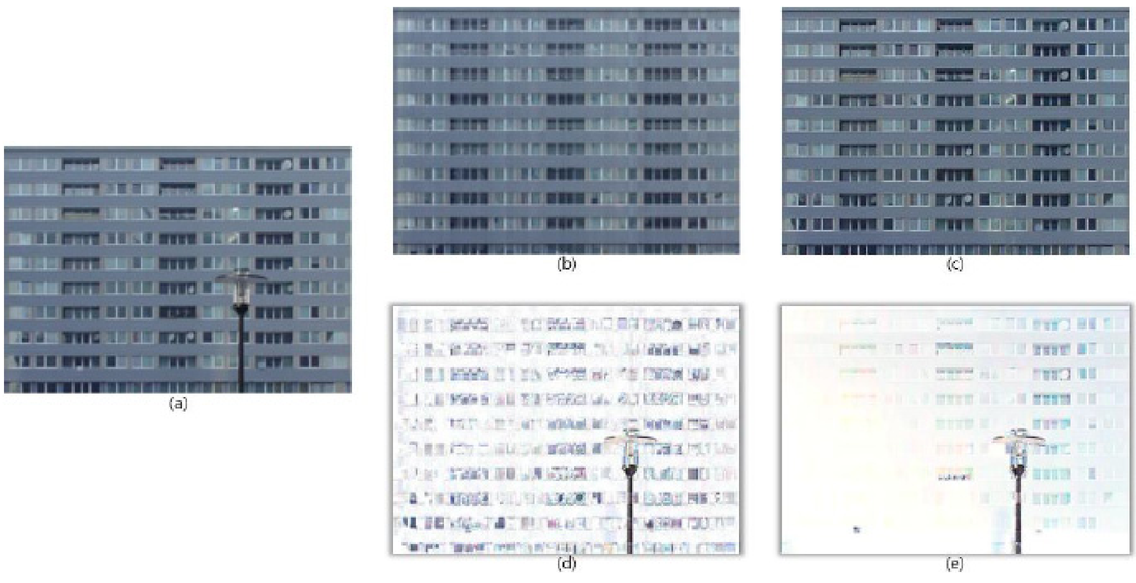

FIGURE 13. **Visual comparison with image in-painting methods [12]. (a) Input image (b) In-painting restoration result [12]. (c) restoration result with the proposed method. (d) image difference of (b) with (a). (e) image difference of (c) with (a).**

For the *SA*, the result of removing the environmental noise from the area outside the obstacle and the obstacle area are restored by combining the average texture added with the component element specificity. In the *DA*, the obstacle region is reconstructed first through neighboring content recognition correction, and a correction is made to the *EN* with *C* selected, as shown in equation (9). Figure 11 shows the results of facade texture restoration.

$$\begin{aligned}
&\min(Dev((Sobel(O_d) \cap M_d^c) \\
&\quad -(Sobel(Dev(C_i) \cap M_d)) = C_d \\
&\qquad\qquad\qquad (i = 0\ldots\#DA\ Candidates) \\
&\quad \because ((O_s \cap M_s^c) \cup (A_s \cap M_s \mathrm{xSaves})) \\
&\qquad\qquad \cup(Fill(O_d \cap M_d^c x(C_d)) = \mathrm{Output} \qquad (9)
\end{aligned}$$

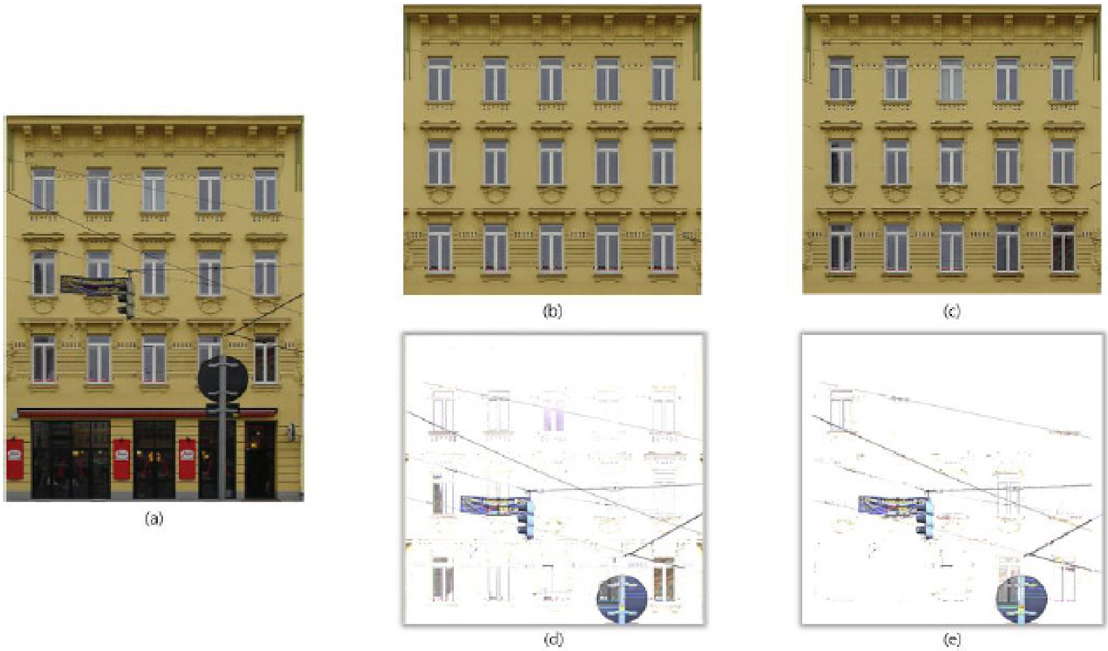

FIGURE 14. **Comparison of results with structural reconstruction method (symmetry based methods) [11]. (a) Input image (b) Result of structural reconstruction [11], (c) Result of the proposed method, (d) Image difference between (b) and (a), (e) Image difference between (e) and (b).**

V. EXPERIMENTAL RESULTS

In unprocessed real data, this methodology can rectify the occluded area in facade images. First, we analyzed the façade components not only to extract the noise but also to find each cluster's representative image, position of the dynamic area in the component and the position of the pattern between positions of components. Next, we focus on rectifying the facade texture, extracting noise and restoring the area based on the information from the other candidates. Figure 12 shows the result of the proposed procedure. A locally occluded component's rectification using the dynamic area of components is shown in Figure 12(a), and an example of a fully occluded component's repair using an adjacent component pattern is shown in Figure 12(b). Figure 13 and 14 shows that the proposed method keeps the original detail and its dynamic variances ((d) and (e) of Figure 13 and 14) while effectively removing occluding parts.

VI. CONCLUSION

In this paper, we have presented a methodology to remove noise from the texture image region of the building within the foreground image. Separating occlusion states into two different types, we have devised various strategies for each of them. We have combined the process of repetitive component analysis and structural analysis to suit our purpose. Compared with previous approaches, eliminating noise factors may possibly enhance the performance of the appearance of the resulting 3D building model, while it is expected that minimizing the noise factors enhances accurate reconstruction of the 3D building. We have experimentally demonstrated our methodology on several real-world buildings with varying amounts of occlusion and have shown that our approach can, in the future, generate 3D buildings with far more anesthetic quality than previous approaches.

REFERENCES

[1] J. Xiao, T. Fang, P. Tan, P. Zhao, E. Ofek, and L. Quan, "Image-based façade modeling," *ACM Trans. Graph.*, vol. 27, no. 5, 2008, Art. no. 161.
[2] N. J. Mitra, M. Pauly, M. Wand, and D. Ceylan, "Symmetry in 3D geometry: Extraction and applications," *Comput. Graph. Forum*, vol. 32, no. 6, pp. 1–23, 2013.
[3] J. D. Udayan and H. S. Kim, "Constrained procedural modeling of real buildings from single facade layout," *Int. J. Comput. Vis. Signal Process.*, vol. 6, no. 1, pp. 33–46, 2016.
[4] P. Debevec, C. J. Taylor, and J. Malik, "Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach," in *Proc. SIGGRAPH*, 1996, pp. 11–20.
[5] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. SIGGRAPH*, 2000, pp. 417–424.
[6] X. Wang and A. R. Hanson, "Surface texture and microstructure extraction from multiple aerial images," *Comput. Vis. Image Understand.*, vol. 83, no. 1, pp. 1–37, 2001.
[7] S. Coorg and S. Teller, "Extracting textured vertical facades from controlled close-range imagery," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1999, pp. 625–632.
[8] J. D. Udayan, D. E. Lee, H. S. Kim, and J.-I. Kim, "An approach for 3D building model retargeting based on structural inference," in *Proc. Korea Comput. Congr. (KCC)*, 2014, pp. 1300–1302.
[9] F. Wu, D.-M. Yan, W. Dong, X. Zhang, and P. Wonka, "Inverse procedural modeling of facade layouts," *ACM Trans. Graph.*, vol. 33, no. 4, pp. C121-1–C121-10, 2014.
[10] C. Wu, J.-M. Frahm, and M. Pollefeys, "Detecting large repetitive structures with salient boundaries," in *Proc. Comput. Vis.-ECCV*, 2010, pp. 142–155.
[11] P. Musialski, P. Wonka, M. Recheis, S. Maierhofer, and W. Purgathofer, "Symmetry-based façade repair," in *Proc. Modeling, Vis. Workshop*, 2009, pp. 3–10.
[12] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

J. DIVYA UDAYAN (M'12) received the Ph.D. degree in Internet and multimedia engineering from Konkuk University, South Korea. She is currently an Associate Professor with the Department of Information Technology, VIT University, India. Her research interests include semantic modeling, pattern recognition, augmented reality, and HCI.

HYUNGSEOK KIM (M'01) received the Ph.D. degree in computer science from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2003. He was a Senior Researcher with the MIRALab, University of Geneva, Geneva, Switzerland. He is currently a Professor with the Department of Software, Konkuk University, Seoul, South Korea. His research interests include real-time interaction in virtual environments and multimodal interaction mechanisms, shape modeling for real-time rendering, and evoking believable experiences in virtual environments.