MORPHOLOGICAL PROCESSING OF VIDEO FOR 3D BUILDING MODEL VISUALIZATION

Fuan Tsai[1], Jin-Kim Liu[2] and Kuo-Hsing Hsiao[2]

[1]Center for Space and Remote Sensing Research, National Central University
300 Zhong-Shan Rd, Zhong-Li, Taoyuan 32001, Taiwan Tel: +886-3-4227151 ext. 57619 Email: ftsai@csrsr.ncu.edu.tw
[2]Energy and Resources Laboratory, Industrial and Technology Research Institute, Hsin-Chu Taiwan

KEY WORDS: Texture Mapping, Building Model, Cyber City, Video Mosaic, Image Morphology

ABSTRACT: This study developed a highly automated video processing system for photo-realistic visualization of three-dimensional (3D) building models from close-range video sequences. The objective was to produce complete and seamless photo-realistic building facade images and to map them onto correct building model facets effectively and efficiently. Emphases were placed on two important tasks. The first was to merge overlapped video frames to generate mosaicked texture images that were continuous in both geometric outlines and in color shadings. The second was to remove unwanted texture regions, including shadows and areas blocked by road trees and other foreign objects, and to reconstruct the removed texture blocks from near by texture information. The developed system first identified interest points on overlapped frames extracted from video sequences. The interest points were then filtered and matched using a developed algorithm based on the Normalization Cross Correction operation to generate seamless texture mosaics. Then, a combination of morphological image processing and other techniques were applied to the generated texture images to automatically identify and remove unwanted texture regions. Video morphology was also used to reconstruct the removed texture blocks by mirroring nearby texture information to produce complete facade textures. Finally, the texture mosaics were mapped onto corresponding building model facets with parametric linear transformation to for a photo-realistic visualization of 3D building models.

1. INTRIDUCTION

Three dimensional modeling of urban environments (cyber city) is an emerging topic in the researches and applications of geoinformatics. Building model is one of the most important objects in cyber city visualization. Remote sensing and other advanced technologies have been used to effectively reconstruct 3D models of buildings in built-up areas (Chen et al., 2005). However, realistic texture (appearance) attributes of building exteriors is not a common function in most building model implementations.

Beck (2003) used color shading and attached generic pseudo texture to furnish bare building models with facade textures for visual effects, but it did not provide true texture attributes of the buildings. Mapping field collected photographs onto model facets was a valid approach (Lee et al., 2002; Tsai et al., 2005; Tsai and Lin, 2006). For the texture generation and mapping of large scale cyber city implementations, using digital video (DV) should be a more efficient approach. However, few of the existing video processing algorithms have truly designed for accurately visualizing building models. Therefore, this study developed a highly automated video processing system to generate photo-realistic texture attributes of building exteriors from video sequences acquired with commodity DV camcorders and map them onto corresponding 3D building model facets for cyber city visualization.

2. TEXTURE MAPPING OF BUILDING MODELS

There are three primary issues when using DV for texturing building models. First, it is necessary to rectify video data to establish the correct relationships between texture images and corresponding model facets. Ideally, if camera parameters are known, photogrammetric processes can be applied to achieve high precision correction results. However, camera parameters are usually not available or incomplete for most commodity DV cameras. An effective approach to address this issue was to register a group of images or texture frames consisting of the same building facade to a common image (texture) space (Kim et al., 2003; Tsai et al., 2005; Tsai and Lin, 2006).

Secondly, it is usually necessary to merge multiple video frames to generate a complete texture mosaic for a (long, wide) building facade. The mosaic should be continuous in both geometry and color shading. A number of algorithms were developed for image or video mosaicking. Among them, weighted blending based algorithms (e.g. Levin et al., 2004; Uyttendaele et al., 2001) are appropriate for creating seamless texture mosaics from digital photographs or video sequences.

Thirdly, occlusions of generated texture mosaics should be removed. A polygon-based approach was developed to semi-automatically identify non-interested texture regions from individual images and exclude them from being merged into the mosaic (Tsai et al., 2005; Tsai and Lin, 2006). This study further developed an automated operation to detect and remove and patch up occlusions in texture mosaics generated from video sequences.

3. MATERIAL AND METHODOLOGY

The study site of this research was a portion of a university campus. Raw texture data were acquired using a Sony DCR-DVD201 DV and were converted to video sequences. Frames with 40%-60% overlap were extracted from the video sequences to generate complete building facade texture attributes of the study site. Each extracted video frame was corrected using algorithms developed in a previous study (Tsai et al., 2005; Tsai and Lin, 2006). An interest point based procedure was used to stitch overlapped frames into texture mosaics. The procedure first employed Harris Corner Detector (Harris and Stephens, 1998) to identify corner points on the extracted images. Then non-maximum suppression was used to eliminate non-corner points. Because adjacent extracted video frames had almost identical viewing parameters, the overlapped regions of the two frames should have a high correlation. In addition, the relative displacement of each matched pair should also be identical. Therefore, the corner points were matched using these criteria. In this study 20 (pixels) was specified as the threshold to check the image displacement. A Normalized Cross Correlation (NCC) operation was also implemented to examine correlations of detected corner points, using the ending 40% of the left frame as the target window and the whole image of the right frame as the search window. The highest correlated corner points were identified as the tie points for stitching the video frames to generate texture mosaics.

The next step was to identify occlusions on the mosaicked images. This study used a combination of Greenness Index (GI) and image morphology to automatically remove occlusions. GI was useful for identifying non-interested regions caused by road trees or other (green) vegetation, which is one of the most common occlusions in building texture. For other types of non-interested areas, morphological "closing" (i.e. dilation then erosion) and "bottom-hat" (subtracting closed image from the original) operations were employed. Figure 1 displays an example of the morphological operations. After GI operation, the blocking road tree was removed. Subtracting the closing of GI result from the original, a preliminary areas for removal were identified (Figure 1d). Combined with the GI result, the non-interested regions were removed from the original texture image.

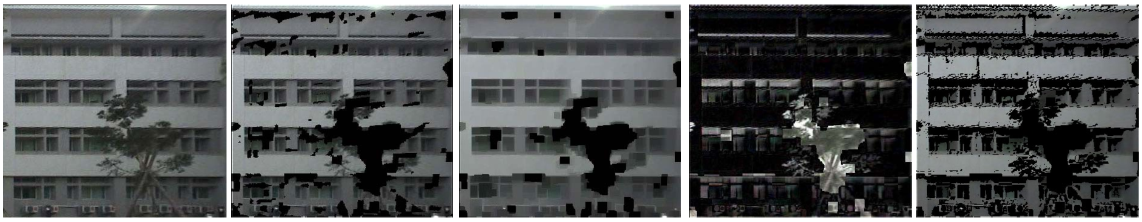

Figure 1: Removal of occlusions using GI and image morphology. From left to right: (a) original; (b) GI operation result; (c) closing of GI result; (d) bottom-hat ((a) – (c)); (e) final result.

After the removal of occlusions, the next step was to mend the removed texture blocks. Since most building texture has symmetric features (for example, window frames in the example of Figure 1 and for most buildings), an effective method to patch up the occlusions would be finding the mirroring axis and applying mirroring operations to fill in the blank. Finding the mirroring axis was also based on morphological operations in the developed system. As illustrated in Figure 2, a binary image was produced from the closing image of the original, and region-growing was applied to segment the binary image. The segmentation were filtered based on the histogram of segment areas to identify possible areas of the most common pattern (building windows, in this case). The identified regions were further shaped with a homogeneity (in area and shape) test to merge blocks that were located closely and exclude fragmented small blocks. Finally, the mirroring axis and the effective mirroring region were determined from the remaining blocks and based on their locations for self-mending. A example of the removal and mending of occlusions is displayed in Figure 3. One thing to note is that areas outside the effective mirroring region were not processed, so the road trees on the bottom portion of the image in Figure 3 were not removed or corrected.

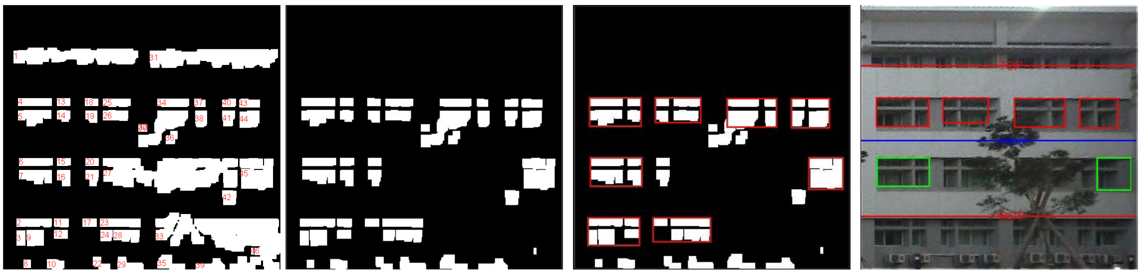

Figure 2: Automatic determination of mirroring axis. From left to right: (a) region-growing segmentation of binary image; (b) identification of possible target areas; (c) homogeneity test to merge segments; (d) calculating the mirroring axis.

Figure 3: Removal and mending of occlusions (top: original; bottom: result).

The final step was mapping generated facade texture mosaics onto corresponding building model facets. Depending on the type of registration performed in the first step, the mapping was done using linear transformation in either object space or parametric space. For regular-shaped facade textures, because the mosaics were already registered to the model (object) space. The mapping was a direct linear transformation. However, for irregular-shaped facades, texture mosaics were in a common image (texture) space but not necessarily the model space. Therefore, the transformation was carried out in a parametric coordinate system.

4. RESULTS

Figure 4 shows two perspective views of the 3D building model visualization in the study site. A simulation of a "street view" is also displayed in Figure 5. As demonstrated in these figures, the developed video processing system successfully generated complete and seamless facade textures from video sequences and mapped them onto model facets correctly to create photo-realistic visualization of 3D building models.

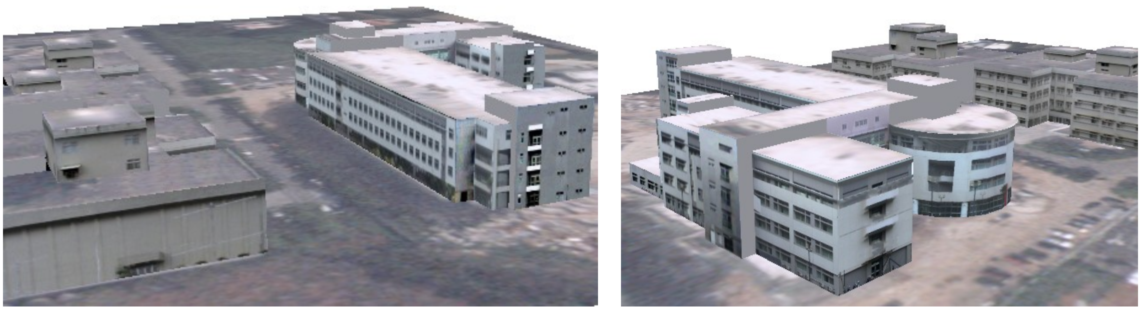

Figure 4: Two perspective views of the (near) photo-realistic visualization result.

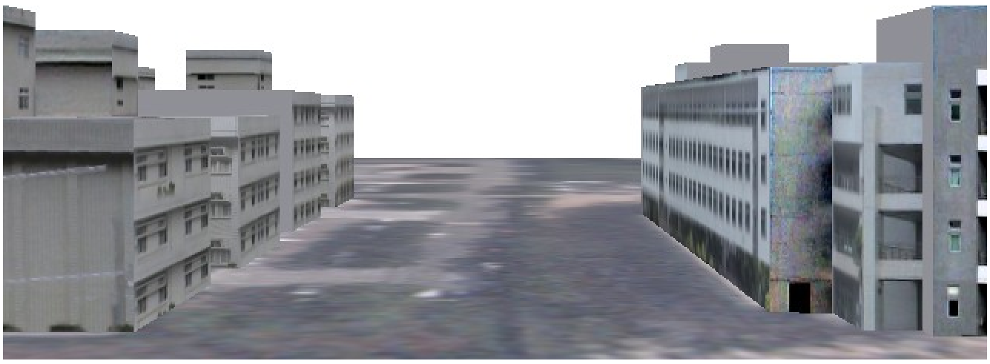

Figure 5: Simulation of a perspective "street view" in the testing areas. Facade textures for the visualization of the buildings were produced from extracted video frames. The ground texture was generated from aerial photographs.

As demonstrated in these figures, the developed video processing system successfully created seamless and texture features of buildings from extracted video frames. Especially, the morphological processing procedure has been proved effective in removing occlusions and automatically patching them up by mirroring operations.

One thing to note is that some minor structures of the building were not presented in the building models. Therefore, they were treated as parts of a facet texture instead of independent texture blocks. In addition, the raw texture videos were not orthorectified, so the registration did not eliminate all distortions and displacements in projection geometry.

5. CONCLUSION REMARKS

This study developed a highly automated video processing system for photo-realistic visualization of 3D building models. Corner detection algorithms were used to identify possible interest points for matching multiple frames. The matching process was based on image-displacement and normalized cross correlation measures to connect tie points. Combined with an efficient blending algorithm, the mosaicking procedure was effective in generating geometrically continuous and seamless texture mosaics.

Non-interested texture regions (occlusions) were identify with GI and morphological operations and mended by mirroring operations. The generated facade texture mosaics were mapping onto corresponding model faces using linear or parametric transformations to create the visualization of 3D building models. Test examples shown in this paper demonstrated that the developed video processing system can produce coherent and photo-realistic appearance of 3D building models effectively and efficiently. The developed system should be a useful contribution to the improvement in the reality and practicality of cyber city implementations and applications.

In particular, the developed system was fully automatic and with high efficiency, except the identification of initial roof-ground tie points pairs and the polynomial fitting for the registration of irregular-shaped facade images, which required human interactions. This provides a significant advantage for visualizing large-scale urban environments.

ACKNOWLEDGMENTS

This study was partially supported by the National Science Council of Taiwan (Project No. NSC-95-2211-E-008-104-MY2) and by the NCU-ITRI Joint Research Center under Project No. NCU-ITRI 960304.

REFERENCES

Beck, M., 2003. Real-time visualization of big 3D city models. International Archives Photogrammetry, Remote Sensing, Spatial Information Sciences, Vol. XXXIV-5/W10.

Chen, L.-C., Teo, T.-A., Rau, J.-Y. Liu, J.-K., Hsu, W.-C., 2005. Building reconstruction from LIDAR data and aerial imagery. IGARSS'05, Vol. 4, pp. 2846-2849.

Harris, C., Stephens, M., 1998. A combined edge and corner detector. Proc. 4th Alvey Vision Conference, pp. 189-192.

Kim, D. H., Yoon, Y. I., Choi, J. S., 2003. An efficient method to build panoramic image mosaics. Pattern Recognition Letters, Vol. 24, pp. 2421-2429.

Lee, S. C., Jung, S. K., Nevatia, R., 2002. Automatic integration of facade textures into 3D building models with projective geometry based line clustering. EUROGRAPHICS 2002, Vol. 21, pp. 259-273.

Levin, A., Zoment, A., Peleg, S., Weiss, Y., 2004. Seamless image stitching in the gradient domain. ECCV 2004, LNCS, Vol. 3024, pp. 377-389.

Tsai, F., Lin, H.-C., Liu, J.-K., Hsiao, K.-H., 2005. Semiautomatic texture generation and transformation for cyber city building models. IGARSS'05, Vol. 7, pp. 4980-4983.

Tsai, F. and Lin, H.-C., 2006. Polygon-based texture mapping for cyber city 3D building models. International Journal of Geographic Information Science, accepted.

Uyttendaele, M., Eden, A., Szeliski, R., 2001. Eliminating ghosting and exposure artifacts in image mosaics. CVPR 2001, Vol. 2, pp. 509-516.