

AUTOMATIC RETRIEVAL OF NEAR PHOTO-REALISTIC TEXTURES FROM SINGLE GROUND-LEVEL BUILDING IMAGES

M. Turker^{a,*}, E. Sumer^b

^a Hacettepe University, Department of Geodesy and Photogrammetry Engineering, 06800-Beytepe, Ankara, TURKEY
- mturker@hacettepe.edu.tr

^b Baskent University, Department of Computer Engineering, Eskisehir Road 20.km, 06530 Ankara, TURKEY-
esumer@baskent.edu.tr

Commission IV, WG IV/4

KEY WORDS: Building Façade Texture, 3D Models, Occlusion Removal, Watershed Segmentation, Rectification

ABSTRACT:

An approach is presented for the automatic retrieval of near photo-realistic textures from single ground-level building images. First, the building facade texture is extracted using the Watershed segmentation technique, which is carried out repetitively until the most successful segment is extracted. Next, the retrieved building image is geometrically rectified in an automated way. After that the occlusions are removed using an image-based approach, which includes marking the occluded region, searching the marked region in the remaining parts of the facade image, selecting a candidate patch with the highest correlation, and copying it into the marked region. The developed concept was tested using two distinct data sets. The first data set contains fifteen rectilinear buildings selected from the residential area of the Batikent district of Ankara, Turkey. The second dataset includes five images selected from eTRIMS database, which contains over a hundred buildings captured in major European cities. The qualitative results obtained are quite promising. For the assessment of façade image extraction, both datasets provided a quantitative accuracy of above 80%. In the rectification results, 15 out of 20 buildings produced the positional mean errors of below 9 pixels. The subjective assessment of the occlusion removal yielded the mean rating scores of 2.58 and 2.28 for the Batikent and eTrims datasets, respectively. While the rating score of 2.58 can be categorized in the middle of the criterions “Fine” and “Passable”, the score 2.28 would be “Fine”.

1. INTRODUCTION

For decades researchers in the field of computer graphics, image processing, and remote sensing have aimed towards methods that will enable the modelling and visualization of various objects. This is due to the fact that 3D visualization greatly improves the ability of human perception. The geographic objects, such as buildings, vegetation, terrain, etc. can be examined through walking or flying using the virtual reality technology. The vital components of a virtual 3D urban model, such as terrain textures, digital terrain models, and building geometry, can be generated by processing the geodata. However, one challenge is the acquisition and mapping of the texture information that is one of the most important steps for generating 3D urban models. Particularly, for texturing the buildings in an urban model, photorealism of the façade texture is an essential requirement. Moreover, since the occlusions may cause lack of quality in the texturing, the extraction of occlusion-free façade textures becomes a necessity.

In previous studies, several approaches have been developed for extracting the building façade textures. These include car-mounted real-time mapping systems, manual or semi-automated mapping from texture libraries and the mapping systems utilizing aerial or space imagery. Früh and Zakhor (2003) created the textured 3D city models using a mobile scanning system. In a study conducted by Früh *et al.* (2004) and Wang *et al.* (2008), the invisible portions of the texture were captured and mapped from oblique aerial images. Lorenz and Döllner

(2006) extracted the façade textures of large-scale city models from overlapping aerial images. Mohan and Murali (2007); Tan *et al.* (2008); and David (2008), extracted the building facades from single view perspective terrestrial imagery. Ripperda (2008); Song and Shan (2004); Ding *et al.* (2007); and Laycock *et al.* (2007) developed methodologies enabling the generation and mapping of façade textures from single ground level images due to low acquisition cost and high resolution. Tsai and Lin (2007); Ortin and Remondino (2005); Hoegner and Stilla (2007); Tsai *et al.* (2006); and Poullis *et al.* (2008) utilized multiple ground level images for extracting occlusion-free building facade textures.

In the present study, we present an approach for automatically extracting photo-realistic façade textures from single ground-level building images. The level of automation has been highly increased by employing an iterative Watershed image segmentation procedure and the novel approaches for rectification and occlusion removal. The markers needed for the segmentation are automatically initialized. The rectification of the extracted building façade texture is also performed in an automated way. In the final step, the occlusions that obscure the buildings are removed. The proposed concept was tested using two distinct data sets, one including 15 rectilinear buildings selected from the residential area of the Batikent district of Ankara, Turkey and the other including 5 images selected from eTRIMS database. The preliminary results reveal that a considerable progress will be made towards the automated construction of the virtual cities.

* Corresponding author.

2. THE METHODOLOGY

2.1 Façade Image Extraction

The facade textures are automatically extracted from the ground-level single building photographs using watershed segmentation, which is carried out repetitively until a stable foreground (building) segment is reached. The initial markers are automatically seeded for both the foreground and background regions. For the foreground region, the initial locations of the markers fall in the middle of the image frame, while the background markers are located at the edges of the image frame. In the present case, a repetitive segmentation procedure is proposed such that for each run, new foreground markers are automatically positioned inside the newly segmented regions. The number of the markers, except from the initial step, is computed as 1 percent of the total image area. That is, if the size of the original image is 480x480 pixels, then the marker count will be 2304 (480 x 480 x 0.01). The execution is terminated either by providing a stable texture amount or by reaching a maximum repetition count. The proposed repetitive segmentation is illustrated in figure 1. In Figure 1a, the initial positions of the foreground pixels are shown in red colour while the background pixels located at the edges are illustrated in green colour. The segmented region after the first iteration is shown in Figure 1b. The foreground and background markers after the fourth iteration are illustrated in red and green colours in Figure 1c, while the resulting segmented image is shown in figure 1d.

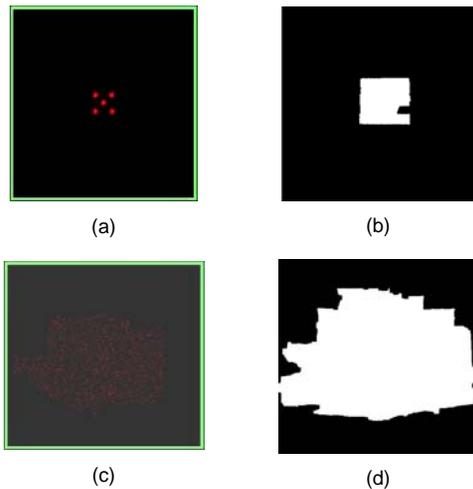


Figure 1. (a) The initial markers, (b) the segmented region after the first iteration, (c) the markers in the fourth iteration, and (d) the segmented region after the fourth iteration.

To reduce over-segmentation, the segmentation is re-executed from the beginning and the new segments are stored separately. This process is repeated until reaching a specified number of iterations. Then, the produced segments are overlaid and the intersections are taken. In this manner, the over-segmented results are likely reduced to a reasonable level, shown in Figure 2a.

2.2 Façade Image Rectification

After extracting the building façade image, the automatic rectification is performed next. In the first step of the automatic rectification, the edges of the building façade image are

detected through the Canny edge detection algorithm. This is followed by detecting the strongest vertical façade edges through Hough transform. For a building image, the final segmented region and the vertical lines detected through Hough transform are illustrated in Figures 2a and 2b, respectively.



Figure 2. (a) The final segmented region and (b) the vertical lines detected through Hough transform

In the second step, the façade corners are detected automatically. To do that the end-points are located using the Harris corner detector (Harris and Stephens, 1988), which defines a corner to be a point with low self-similarity. The similarity is measured by taking the weighted sum of squared differences (SSD) between two image patches. In Figure 3a, the detected end-points of the Hough lines are illustrated. There may be a number of cases where more than four end-points may be detected. Therefore, the end-points that are closest to the image corners are selected. Since we consider the rectangular façade textures four end-points that are marked in yellow colour are needed inherently (Figure 3b). Because of the curvilinear nature of the Hough lines, the selected end-points may not yet be the final points for the geometric rectification. To solve this problem and identify façade corner points, a trend line (shown in yellow colour in Figure 3b) is drawn automatically.

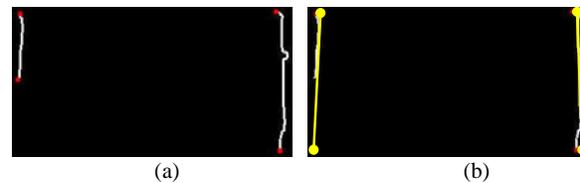


Figure 3. (a) The detected vertical Hough lines and (b) the selected end-points and the trend lines drawn automatically

The third step of the rectification procedure is the projective transformation, which maps lines to lines but does not necessarily preserve parallelism. It is carried out by means of the original and aligned positions of the above detected façade corner points. Any plane projective transformation can be expressed by a 3x3 non-singular homogenous coordinates. In other words, any non-singular 3x3 matrix (homogenous matrix) defines a projective transformation of the plane. The procedure of a projective transformation is illustrated in Figure 4.

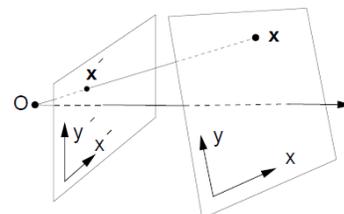


Figure 4. A projective transformation with center O, mapping xy plane to x'y' plane.

After locating the final positions of the façade corners, the aligned positions of these points are found. To do that the angle between the trend line and the vertical direction is calculated and the aligned positions are computed using this angle. For the horizontal direction, an imaginary line is drawn between the upper end-point pair. Then, the angle between the imaginary line and the horizontal direction is calculated and the aligned positions are computed. In Figure 5, the locations of the final and aligned end-points are shown. The yellow dashed lines are the trend lines, while the yellow end-points (x_1 - x_4) refer to final end-points. The green dashed lines correspond to vertical lines while the green end-points (x_1' - x_4') refer to aligned end-points. The pink dashed line is the imaginary horizontal line and the blue corresponds to horizontal line.



Figure 5. The locations of the final and aligned end-points

Once the locations of the end-points are found, the projective transformation of the façade image is carried out using the MATLAB image processing toolbox functions. With the geometric rectification, the perspective distortions are reduced to a reasonable level. For an extracted building façade, before and after rectifications are shown in figures 6a and 6b, respectively.

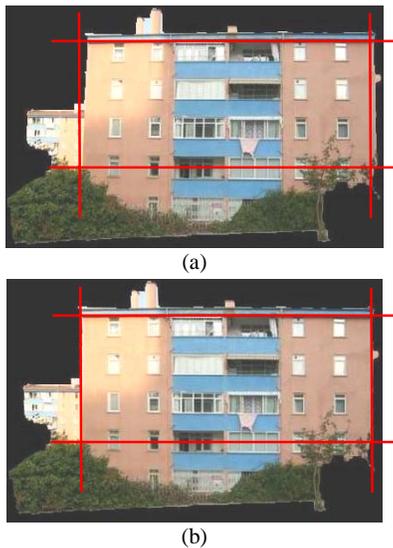


Figure 6. For a building façade, (a) before geometric rectification and (b) after geometric rectification

2.3 Occlusion Removal

An occlusion is a foreground object that partially or completely obstructs a building standing in the background. They are generally trees, bushes, lamp posts, parked vehicles or

pedestrians. Therefore, the occlusions should be removed before the acquisition of the final building façade textures. The first step of the proposed occlusion removal procedure is the automatic cropping of the rectified façade image. To do that the corner points of the rectified façade image are computed using the Harris corner detection algorithm. Then, two of the corner points that are closest to the upper left and upper right corners of the image are selected as the upper façade corners. By locating these two points, the amount of cropping from the top and the sides are determined. The rectified and the cropped façade images are illustrated in figures 7a and 7b, respectively.



Figure 7. The rectified façade image (a) before cropping and (b) after cropping

After the cropping operation, the user manually selects the occlusion to be removed by drawing a quadrilateral enclosing the occlusions (Figure 8a) and the selected patch is stored separately (Figure 8b). Then, this patch is segmented into two segments, one representing the occlusions (black) and the other representing the building (white) (Figure 8c).

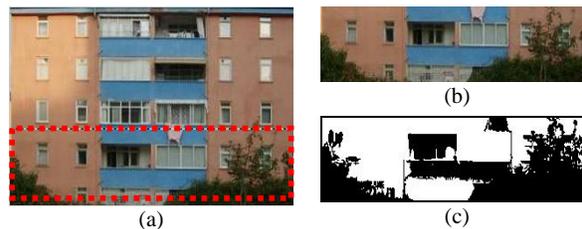


Figure 8. (a) The selection of an occluded area, (b) the selected patch, and (c) the segmented patch.

The extraction of the occlusion is followed by the texture mending procedure, in which the most similar patch to the selected patch (all but itself) is searched within the cropped façade image. For measuring the similarity, the correlation metric is used only for the non-occluded areas. In the present case, the correlation coefficient is computed for each of the R-G-B bands separately and the average is taken as the final correlation value. Then, the patch having the highest correlation with the occluded patch (candidate patch) is pasted into the cropped façade image (Figure 9).



Figure 9. (a) An occluded patch, (b) the candidate patch, and (c) the mended image after pasting the patch in (b).

2.4 Post Processing

After removing the occlusions from the façade image, two more issues are to be handled before reaching the final building façade texture. The first issue is the sharp transitions (seams) at the borders of the pasted patches. We solve this problem by applying a motion blurring filter that creates a linear-, radial, and zoom-movement blur. The size and direction of blurring can be altered by the parameters length and angle. In this study, we use a linear type of filter, in which blurring occurs in either horizontal or vertical direction. For the parameter length, which represents the blur intensity, we use the value of 5. For the parameter blurring angle, we use 90^0 for horizontal and 0^0 for vertical seams. In many image processing applications, the motion blurring is known to have negative effects. In this case, we utilize it for generating a seamless texture. Before and after the motion blurring of a part of a building façade image are illustrated in Figures 10a and 10b, respectively.

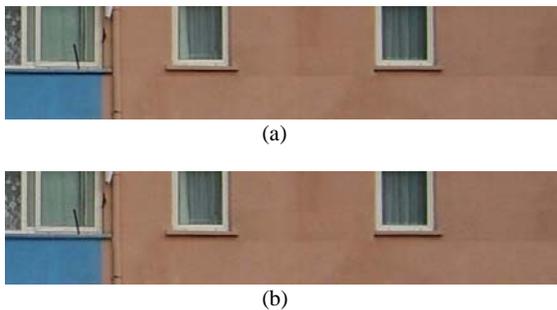


Figure 10. A part of a building façade image (a) before and (b) after the motion blurring

The second issue is the non-uniform illumination such that the candidate patch can have different illumination characteristics with respect to its neighbourhood (Figure 11). To reduce this effect the occluded patch is subtracted from the candidate patch. Then, a threshold is defined and those pixels to be assigned from the occluded patch into the candidate patch are determined. That is if, for a specified location (x,y) , the difference is less than the threshold then, the corresponding pixel in the occluded patch is assigned into the same position in the candidate patch. Otherwise, the original pixel value in the candidate patch remains unchanged. The difference image and the updated candidate patch are illustrated in Figure 12.

Although the level of illumination is normalized considerably, the noisy pixels may arise. Therefore, we remove the noisy pixels by employing a weighted image averaging procedure. The candidate and the updated candidate images are summed up with the equal weight of 0.5 and thus, the balance between the noise level and the non-uniformity of the illumination is preserved to a certain extent. The image patch after removing the noisy pixels is shown in Figure 13.

As mentioned earlier, the extraction of the occlusions is a semi-automated procedure. The user can stop the execution when the results are satisfactory or resume the process to obtain better results. We are aware that removal of all occlusions in a single iteration may not be possible. Therefore, to remove the remaining occlusions the same steps must be repeated until extracting the final building façade texture. In Figure 14, the extracted final building façade image is illustrated.



Figure 11. The abrupt illumination change indicated by the red circle

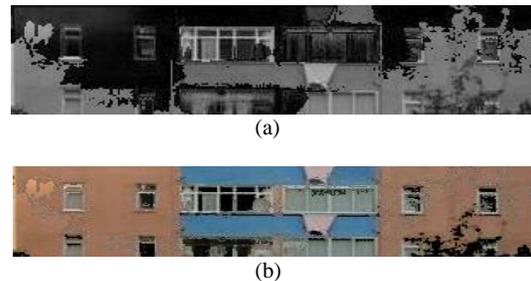


Figure 12. (a) The difference image in gray scale, in which dark areas illustrate minor or no change and (b) the updated candidate image



Figure 13. The image produced after the weighted averaging



Figure 14. The final building façade image

3. THE EXPERIMENTAL RESULTS

The experimental data comprises 15 building façade images from the Batikent, Ankara dataset and 5 images from the eTrims dataset. We made the assessment of the segmented façade images by a quantitative evaluation metric (Shufelt and McKeown 1993). This metric is based on labelling the pixels by comparing the output of the proposed methodology with the reference data. During labelling, four possible categories that are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) may occur for each pixel. In the case of TP, both the analysis results and the reference data label the pixels as belonging to foreground (building façade). In the case of TN, the background is labelled both by the analysis results and the reference data. In the case of FP, the analysis results label the pixels as belonging to foreground, while the reference data labels them as background. The FN case is the exact opposite of the FP case.

To evaluate the performance, the counts of TP, TN, FP, and FN were calculated, and then the metrics *building detection percentage (bdp)*, *branching factor (bf)* and *quality percentage (qp)* were computed. The quantitative evaluation results for the eTrims and Batikent datasets are given in Tables 1 and 2, respectively.

$$bdp = \frac{100 \times TP}{TP + FN}, \quad bf = \frac{FP}{TP}, \quad qp = \frac{100 \times TP}{TP + FP + FN} \quad (1)$$

Table 1. The quantitative results of the eTrims dataset

| Bld # | TP | TN | FP | FN | BDP (%) | BF | QP (%) |
|-------|-------|-------|-----|-------|---------|-------|--------|
| 1 | 28237 | 22476 | 127 | 9847 | 74,1 | 0,005 | 73,9 |
| 2 | 28669 | 22307 | 320 | 8595 | 76,9 | 0,011 | 76,3 |
| 3 | 29989 | 11804 | 54 | 14920 | 66,8 | 0,002 | 66,7 |
| 4 | 41336 | 24909 | 153 | 2808 | 93,6 | 0,004 | 93,3 |
| 5 | 40716 | 18051 | 117 | 1301 | 96,9 | 0,003 | 96,6 |

Table 2. The quantitative results of the Batikent dataset

| Bld # | TP | TN | FP | FN | BDP (%) | BF | QP (%) |
|-------|--------|--------|-------|-------|---------|------|--------|
| 1 | 65769 | 136515 | 9067 | 2290 | 96,6 | 0,13 | 85,3 |
| 2 | 68801 | 139742 | 13886 | 687 | 99,0 | 0,20 | 82,5 |
| 3 | 68749 | 149612 | 1296 | 240 | 99,7 | 0,01 | 97,8 |
| 4 | 70822 | 147042 | 4751 | 3 | 100 | 0,06 | 93,7 |
| 5 | 80136 | 120719 | 19617 | 465 | 99,4 | 0,24 | 80,0 |
| 6 | 104676 | 84655 | 2898 | 26291 | 79,9 | 0,02 | 78,2 |
| 7 | 67038 | 143627 | 10874 | 14 | 100 | 0,16 | 86,0 |
| 8 | 78006 | 74032 | 504 | 67404 | 53,6 | 0,00 | 53,5 |
| 9 | 68683 | 142598 | 3647 | 6401 | 91,5 | 0,05 | 87,2 |
| 10 | 45743 | 169453 | 710 | 6465 | 87,6 | 0,01 | 86,4 |
| 11 | 81688 | 111670 | 3219 | 24358 | 77,0 | 0,03 | 74,8 |
| 12 | 110865 | 54080 | 476 | 52855 | 67,7 | 0,00 | 67,5 |
| 13 | 65731 | 146194 | 3683 | 6383 | 91,1 | 0,05 | 86,7 |
| 14 | 86805 | 113873 | 19801 | 180 | 99,8 | 0,22 | 81,3 |
| 15 | 84725 | 129815 | 5428 | 207 | 99,8 | 0,06 | 93,8 |

In Table 1, the façade detection percentage is varying between about 74% and 97%. The average “*bdp*”, “*bf*” and “*qp*” values were found to be 81.7%, 0.005, and 81.4%, respectively. The “*qp*” values were computed to be close to “*bdp*” values since the branching factors are relatively small. In Table 2, the “*bdp*” values are alternating between 53.6% and 100%. In this dataset, the average façade detection percentage was found to be 89.5%, while the average values were 0.09 and 82.3% for “*bf*” and “*qp*”, respectively. Because of the high branching factor values, a significant gap between the average “*bdp*” and “*qp*” values is evident. Due to high false positive pixels, buildings #8 and #12 have considerably low detection percentages. If these buildings are excluded from the dataset, the average detection percentage of 93.9% can be reached.

The results of the automatic rectification of the segmented façade images were evaluated by a quantitative metric (Aronoff, 1989). For each façade image, a reference dataset was prepared by doing manual geometric rectification using the GIMP-2 image processing software. Next, four test points were determined both on the reference image and the rectified image and, for each test point, the pixel distance errors were computed from the corresponding (x,y) pair. When computing the errors a

normal distribution model was used and, for each building, the accuracies were computed at the confidence levels of 80%, 85%, 90%, and 95%.

For the eTrims dataset, buildings #2 and #3 provided the mean errors of 1.83 and 1.66 pixels, respectively. The mean errors computed for buildings #1 and #5 were 3.28 and 3.63 pixels, respectively. Of the five test buildings, the highest mean error of 9.76 pixels and the standard deviation of 5.02 pixels were computed for building #4. This is due to the fact that at the right façade border of building #4, the Hough transform generated too many short line segments with different orientations. Hence, drawing an optimal coinciding trend line from the Hough line segments was not possible, therefore, highly affecting the rectification process and boosting the positional error. The lowest accuracies were computed for buildings #2 and #3, which were followed by buildings #1 and #5.

In the Batikent dataset, buildings #1, #2, #4, #10, #14, and #15 appear to be the most successful buildings with the mean errors ranging from 2.19 to 5.03 pixels. The mean errors of the buildings #5, #6, #8, #11, and #13 were computed between 6.15 and 8.87 pixels. For the remaining buildings #3, #7, #9, and #12, the errors were quite high staying between 10.25 and 17.33 pixels. For these buildings, the reasons of the failures were investigated and it was found that most of the errors arise from the deficiencies or redundancies of the extracted façade border lines. In either case, the rectification accuracy is negatively affected.

The results of occlusion removal were assessed by a subjective metric. Although the objective methods offer a simple and convenient way, we found it to be more appropriate to make the assessment of the final image subjectively. Thus, we presented an occlusion-free image to a cross-section of viewers and took the average of their evaluations. In the present case, the evaluations were carried out using an absolute rating scale, which is illustrated in Table 3 (Gonzales and Woods, 2008).

Table 3. Rating scale of the Television Allocations Study Organization (Frendendall and Behrend, 1960)

| Value | Rating / Description |
|-------|---|
| 1 | Excellent / An image of extremely high quality, as good as you could desire |
| 2 | Fine / An image of high quality, providing enjoyable viewing. Interference is not objectionable |
| 3 | Passable / An image of acceptable quality. Interference is not objectionable. |
| 4 | Marginal / An image of poor quality; you wish you could improve it. Interference is somewhat objectionable. |
| 5 | Inferior / A very poor image, but you could watch it. Objectionable interference is definitely present. |
| 6 | Unusable / An image so bad that you could not watch it. |

In the present case, the occlusion-free building façade images of the two datasets were evaluated by 10 viewers, who are senior students in computer science, research assistants and professors working in the field of image processing. For the eTrims dataset, the average ratings of the façade images stayed between 1.6 and 2.9. The average of the overall dataset was computed to be 2.28, which can be identified as “Fine”. For the Batikent dataset, the average rating values ranged from 1.7 to 3.9. The average of the whole dataset was computed to be 2.58 staying in the middle of the criterions “Fine” and “Passable”. When the reasons for the failures were investigated we found that the

problems were encountered with the excessive blurring of the seams, discontinuities between the objects (e.g. shifted windows), and the occlusions that were not able to be removed completely.

4. THE CONCLUSIONS

The automated extraction of the façade images were carried out by applying pre-processing, marker initialization and watershed segmentation steps, respectively. Watershed segmentation procedure was performed repetitively until a stable foreground segment was obtained. This step was repeated until reaching a specified number of iterations and the produced segments were superimposed and then intersected to minimize the over-segmentation. The quantitative results were found to be quite satisfactory. For the Batikent dataset, the average detection and quality percentages were computed as 89.5% and 82.3%, respectively. Similarly, eTrims dataset turned out the same percentages of 81.7% and 81.4%.

In façade image rectification, the perspective distortions were reduced in an automated way based on the detection of façade corner points. To do that, a series of operations that include the detection of the edges, the extraction of the strongest vertical lines, the detection of the end-points, the adjustment of the vertical façade edges, and the projective transformation were employed. The positional accuracies obtained were promising. For the Batikent dataset, 11 façade images out of 15 produced positional mean errors under 9 pixels. In eTrims dataset, 4 of 5 images yielded positional mean errors no more than 9 pixels.

The qualitative results of the occlusion removal were found to be successful. For the Batikent dataset, the average of the ratings was computed to be 2.58 staying in the middle of the criterions “Fine” and “Passable”. In parallel, eTrims dataset was included into “Fine” category having an average rating of 2.28. In general, the results reveal that a considerable progress will be made towards the automated construction of the virtual cities

References

Aronoff, S., 1989. *Geographic Information Systems: A Management Perspective*. WDL Publications, pp. 133-148, Ottawa, CANADA.

David, P., 2008. Detection of Building Facades in Urban Environments. *Proceedings of SPIE*, Vol. 6978, pp. 139-148.

Ding, W., Zhu, F., Hao, Y., 2007. Interactive 3D City Modeling using Google Earth and Ground Images. *Fourth International Conference on Image and Graphics (ICIG 2007)*, pp. 849-854, Sichuan, China.

Friendendall, G.L., Behrend, W.L., 1960. Picture Quality - Procedures for Evaluating Subjective Effects of Interference. *Proceedings of Institute of Radio Engineers (IRE)*, Vol. 48, pp. 1030-1034.

Frueh, C., Sammon, R., Zakhor, A., 2004. Automated Texture Mapping of 3D City Models With Oblique Aerial Imagery. *Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04)*, Thessaloniki, Greece.

Frueh, C., Zakhor, A., 2003. Constructing 3D City Models by Merging Aerial and Ground Views. *IEEE Transactions on Computer Graphics and Applications*, 23(6), pp. 52-61.

Gonzalez, R.C. and Woods, R.E., 2008. *Digital Image Processing, 3rd Edition*, pp. 407-413, New Jersey, USA

Harris, C., Stephens, M.J., 1988. A combined corner and edge detector. *Proceedings of the 4th Alvey Vision Conference*, pp. 147-152, Manchester, England.

Hoegner, L., Stilla, U., 2007. Texture Extraction for Building Models from IR sequences of Urban Areas. *Urban Remote Sensing Joint Event*, Paris, France.

Laycock, R.G., Ryder, G.D.G., Day, A.M., 2007. Automatic Generation, Texturing and Population of a Reflective Real-Time Urban Environment. *Computers and Graphics*, Vol. 31, pp. 625-635.

Lorenz, H., Döllner, J., 2006. Towards Automating the Generation of Facade Textures of Virtual City Models. *ISPRS Commission II, WG II/5 Workshop*, Vienna, Austria.

Mohan, S., Murali, S., 2007. Automated 3D Modeling and Rendering from Single View Images. *International Conference on Computational Intelligence and Multimedia Applications (ICCI 2007)*, Vol.4, pp.476-480, Tamil Nadu, India.

Ortin, D., Remondino, F., 2005. Occlusion-Free Image Generation for Realistic Texture Mapping. *Proceedings of the ISPRS WG V/4 Virtual Reconstruction and Visualization of Complex Architectures*, Venice, Italy.

Poullis, C., You, S., Neumann, U., 2008. Rapid Creation of Large-scale Photorealistic Virtual Environments. *Proceedings of IEEE Virtual Reality Conference (VR'08)*, Vol. 153-160, Reno, Nevada, USA.

Ripperda, N., 2008. Determination of Façade Attributes for Façade Reconstruction. *Proceedings of International Society for Photogrammetry and Remote Sensing (ISPRS'08) Congress*, Beijing, China.

Shufelt, J.A., McKeown, D.M., 1993. Fusion of monocular cues to detect man-made structures in aerial imagery. *Computer Vision, Graphics and Image Processing: Image Understanding*, 57(3), pp. 307-330.

Song, Y., Shan, J., 2004. Photorealistic Building Modeling and Visualization in 3-D Geospatial Information System. *Proceedings of International Society for Photogrammetry and Remote Sensing (ISPRS'04) Congress*, Istanbul, Turkey.

Tan, Y.K.A., Kwok, L.K., Ong, S.H., 2008. Large Scale Texture Mapping of Building Facades. *Proceedings of International Society for Photogrammetry and Remote Sensing (ISPRS'08) Congress*, Beijing, China.

Tsai, F., Lin, H.C., 2007. Polygon-based Texture Mapping for Cyber City 3D Building Models. *International Journal of Geographical Information Science*, 21(9), pp. 965-981.

Tsai, F., Liu, J-K., Hsiao, K-H., 2006. Morphological Processing of Video for 3-D Building Model Visualization. *Proceedings of 27th Asian Conference on Remote Sensing (ACRS2006)*, Ulaanbaatar, Mongolia.

Wang, M., Bai, H., Hu, F., 2008. Automatic Texture Acquisition for 3D Model Using Oblique Aerial Images. *First International Conference on Intelligent Networks and Intelligent Systems (ICINIS 2008)*, pp. 495-498, Wuhan, China.