A GRID-BASE LINE ANALYSIS FOR STREET OCCLUSION REMOVAL AND BUILDING FACADE TEXTURING

Yawen Liu
Zhen Guan
School of Remote Sensing and Information Engineering
Wuhan University
Wuhan, 430079
liuyawen70@gmail.com
vieri_21@tom.com

ABSTRACT

The building façade model is the most prominent streetscape feature for rendering virtual street models. This paper aims to develop a method to identify street obstacles which hinder photo-realistic texturing of building façade. In order to achieve this, the presented method focuses on two research components. Firstly, we develop an effective grid-based line analysis method to differentiate image elements including building façade from the ones containing street obstacles. The image is divided into a number of grids with a regular size. The dominant façade's geometry plays a significant role as scene knowledge to filter street obstacle by analyzing the numbers of parallel lines and the density of parallel line per each grid. Secondly we present a method to generate an occlusion-free texturing of building façade based-on planar affine transformation and bilinear interpolation. We evaluate the effect of the presented method using real street-view images.

INTRODUCTION

The Geospatial Web browser such as Google Earth and Microsoft Virtual Earth is an emerging technology which enables to provide 3D photorealistic street scenes over a number of megacities. As the city's most prominent feature, virtual streetscape is made in 3D object space for a wide variety of geospatial applications. For virtual streetscape, 3D visual sensation is often more important than the precision of model's coordinates. Therefore, the model is considerably enhanced by adding detail in the form of photo-realistic texture on the surface of the object model(van den Heuvel,1998). Terrestrial images for texture mapping streetscapes, which are captured along a narrow street, are hardly free of the occlusions which hinder the realistic façade texture. Thus, the full automation of texture occlusion removal has been regarded as a valuable research for rendering street scene.

Some relative researches were presented including background estimation method (Dongsheng et al.,2001;Tao et al.,2004;Kapje et al.,2008), multi-image fusion approach((Bo¨hm,2004;Yi,2008; Ortin and Remondino, 2005) and synthesizing image texture method(Qinghong,2008;Criminisi,2004; Ashikhmin,2001).These methods can be classified two classes according to the steps of occlusion process. One needs two stages to realize occlusion-free image generation: occlusion extraction and missing texture recovery. The other is integration analysis of image occlusion in order to obtain the optimal lost texture. For the former, no unified extraction algorithm is robust to all diversified shapes of occlusions without prior knowledge. Thus, some algorithms label the occlusion region manually or interactively. The latter needs to build evaluating model for estimating the optimal missing pixel based on multi-images fusion.

This paper aims to develop an effective method to detect occlusion based on the numbers of parallel line and grid line density analysis which highlights façade dominant feature and is free of the constraint of occlusion object shape. At the same time, this method presented provides a photo-realistic texture fast generation by planar affine transformation.

Section 2 introduces city streetscape and its dominant features- building façade and parallel line. Section 3 details our algorithm for occlusion extraction and removal including sensitive analysis and quality evaluation of occlusion grid detection, and fast occlusion-free image generation. Experimental results and conclusion are separately presented in section 4 and 5.

BUILDING FACADE AND PARALLELITY

In most cases, streetscape includes the area from building face to building face on both sides of the street and all the elements in between. It typically consists of the following zones: shy zone, clear zone and landscape& furniture zone. The shy zone accommodates any building protrusions and the clear zone is largely unobstructed area. Most street side landscaping is located in the landscape and furniture zone including street trees, parking meters, utility poles, bicycle racks, streetlights, garbage cans, bus shelters, signage, and benches(Department of Community Planning,2007). Amongst them, Building façade is the outstanding component of streetscape which gives prominence to the character of the city or town.

The perception of lines, in particular the perception of horizontal and vertical lines, plays a dominant role in the visual evaluation of house facades and streetscapes(Chalup, Clement et al. 2007). A lot of lines in the scene are parallel and a lot of edges in the scene are orthogonal. In an outdoor environment e.g. streets, buildings and pavements satisfy this assumption(Rother,2000).

Parallelism and perpendicularity, commonly presented in building facades have been successfully applied for camera calibration and building reconstruction. Vanishing point is the intersection of the lines in image, which are parallel lines before projecting into one image(Wenhan,2004). The camera constant and principal point can be recovered using vanishing points of three orthogonal directions(Grammatikopoulosa,2004). When these camera calibration methods are applied in three dimension reconstruction of regular buildings, it can also realize on-line camera calibration(Wenhan,2004).Parallelism, perpendicularity constraints are combined with line-photogrammetric measurement of a single image for building reconstruction(van den Heuvel,1998).These constraints lead to very simple and geometrically intuitive methods to calibrate the intrinsic and extrinsic parameters of cameras and to recover Euclidean models of the scene from only two images from arbitrary positions(Cipolla et al.1999).

In this paper, parallelity, which is a prominent geometry feature of building façade, is considered to be a unique cue for filtering building façade from occlusions existing in an image. This algorithm based on parallel line analysis avoids the dependency on different attributes of occlusion objects that are usually used to extract occlusions from building facades in many researches presented. In additional, it can accommodate the shape and attribute of diversified occlusions simultaneously during the process of occlusion detection.

OCCLUSION REMOVAL AND TEXTURE MAPPING

In literature (Zhizhong et al,2005), tree occlusions are removed from the facades by color cue analysis assisted with line density based on grid. In our approach**,** more types of occlusions are considered such as pedestrian, cars, parking sign and so on. Therefore the parallel line is focused as a unique cue for filtering the facades from occlusions. At the same time, the algorithm proposed pays more attention to the sensitivity analysis of line density threshold for occlusion detection. Firstly an occlusion image is divided into a number of grids with regular size. For a grid, parallel line density is calculated. Then according to scenic knowledge described in section 2, building façade can be differentiated from occlusions by analyzing the numbers of parallel lines and the parallel line density of a grid. In this section, we will detail the significant processes including grid size determination, parallel line density threshold selection and occlusion grid detection.

Occlusion Detection Based On Grid Line Density

Grid size determination. Building façade represents the distinctive repetition of structure and texture. Dividing an image to regular grids is detecting the repetitions and further analyzing the attribute of similar tiles. The literature(Mu¨ler, Zeng et al. 2007) derives 3D models of high visual quality from single façade image by hierarchical façade subdivision such as floor, tile and small rectangles.

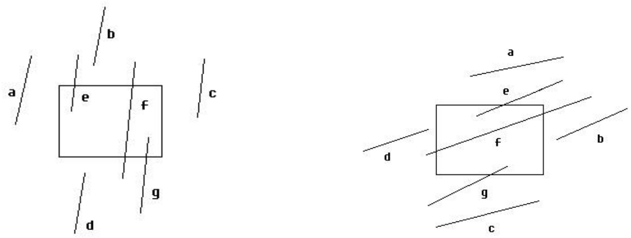
In our approach, the grid is considered to be the measurement unit of occlusions. Therefore in theory, the size of grid is as small as possible. At the same time the grid is determined by combining the similar smallest components of building facades in shape and structure and containing the line features. In practical application, the image resolution is also one factor related to the size of grid. The image could be divided evenly by the grid.

The criterion to choose the grid is that the size of grid falls in between the size of one tile and two tiles of the building façade, and is one minimum integer division of the image. For a building façade, the typical tile is a window including the surrounding wall. If line features exist in the wall around the window, the tile can be reduced to the rectangle of the window that doesn't contain lines except its structure edges.

Grid line density. The density of parallel line for a grid can be calculated according to the equation shown as(Zhizhong et al,2005):

$$\rho = \sum_{i=1}^{n} L_i \Big/ S \tag{1}$$

L_i expresses the number of pixels in a line, S is the sum of pixels in a grid and n is the number of lines in a grid. For the computation of the length of a line L_i, a simple and fast strategy is designed. Firstly, parallel lines are grouped into two classes according to their directions. If the angle of parallel line is more than 45 degree, these lines are considered to be vertical lines shown in figure 1(a). Otherwise these lines are horizontal lines shown in figure 1(b). Then for a grid, the vertical or horizontal lines which are cross this grid need to be judged separately according to the coordinates of two end points of these lines and the range of the grid. In figure 1, the lines labeled e,g and f are related to the gird and the other lines can be ignored.



(a) (b)

Figure 1. Line and gird.

At last, these pixels in the grid generated by the projection of these related lines on vertical or horizontal direction are considered to be the length of L_i .

Occlusion detection. In this process, primitive occlusion detection is firstly implemented based on the analysis of the numbers of the parallel lines. The complex or mixed occlusions usually appear the bottom of terrestrial images. For instant, the low objects in the front of the building facade will be projected on the part of an image. In most case, this part of a terrestrial image is composed by the doors and shopping windows of commercial stores, entrance hall or the zone in front of building facade. The numbers of vertical parallel lines are limited or less in this part comparing with the middle or top of a terrestrial image covered by the building facade. Thus, for a row in an image, the numbers of parallel lines in vertical directionare less than pre-defined threshold, this row is marked as an occlusion row.

After this step, the existing occlusions are the projection of high parts of objects around the building such as the top of trees or the top of poles. Their covering areas on the image are the part of the building facade which possesses the repetitive texture structure. These occlusion detection needs to be judged grid by grid based on the line density. The threshold to differentiate the occlusion grids from facade grids is determined by statistically analyzing the numbers of correct detectable occlusion grids and the undetectable occlusion grids.

At last, the occlusion region can be formed by grouping the occlusion grids. However, not all occlusions can be detected successfully. Especially when the occlusion occupies a little part of a grid and there are a lot of line features in an occlusion grid, this grid usually passes the examination of line density threshold and becomes to be an undetectable occlusion grid. At the same time, it is possible for a facade grid to be misinterpreted as an occlusion grid due to less line features. A wrong occlusion grid could be differentiated from an occlusion grid by comparing the similarity between the pixels of the grid and the corresponding grid in adjacent images. If the similarity is high, the wrong occlusion grid is identified.

In order to evaluate the performance of occlusion detection based on grid line density analysis, the quality measures chosen are completeness and correctness which are defined in the literature(Heipke,et al., 1997) for evaluation of automatic road extraction and are used for evaluation of building detection in some researches presented(Rottensteiner et al.,2007).

$$\text{Completeness}=\frac{TP}{TP+FN} \qquad (2)$$

completeness $\in [0;1]$

$$\text{Correctness}=\frac{TP}{TP+FP} \qquad (3)$$

correctness $\in [0;1]$

In this paper, the detected the occlusion grids and the reference results obtained by manual interpretation are compared grid by grid. In Eqs.(2) and (3), TP(True Positive) describes the number of occlusion grids detected correctly. FN(False Negative) is the number of undetectable occlusion grids, for example, the number of occlusion grids identified to facade grids. FP(False Positive) denotes the numbers of error detectable occlusion grids, for example, the number of facade grids identified to occlusion grids.

3.2 Texture Mapping Based On Planar Affine Transformation

Once the occlusion regions are formed, the lost texture can be recovered from adjacent terrestrial images. Planar affine transformation (Jianqing et al, 2003) is used to realize the geometry transformation between the target image and its corresponding adjacent image. Eqs.(4) describes the transformation.

$$\begin{aligned} x &= a_0 + a_1 \cdot x' + a_2 \cdot y' \\ y &= b_0 + b_1 \cdot x' + b_2 \cdot y' \end{aligned} \qquad (4)$$

x' and y' are the 2-D pixel coordinates in the target image while x, y are the corresponding coordinates of the adjacent image. The corresponding points between the target image and the adjacent images for planar affine transformation can be obtained by image matching technique.

After planar affine transformation, the grey value of the pixel at a position will not exactly locate the center of the pixels in adjacent image, so its estimated value of color should be interpolated from its neighbor pixels. The bilinear interpolation is used in this proposed algorithm.

EXPERIMENTAL RESULTS

The aim of this section is to evaluate the feasibility and effectiveness of the proposed occlusion detection technique. Figure 2 shows original images that depict four views of the facades partially occluded by cars, pedestrian, traffic sign and trees. Among these occlusion objects, trees occlude a significant part of facade due to their height. Figure2(a) shows the image which is 4500x3000 pixels. Other images are 2592x3888 pixels shown in Figure2(b,c,d).

Parallel lines of the scenery were detected from these original images. From figure 2, it is found that most of detected lines come from the facades and a little attribute to occlusion objects such as car body and lamp pole. In figure 2, the numbers of parallel lines shown in four images are 528,535,527 and 769 separately.

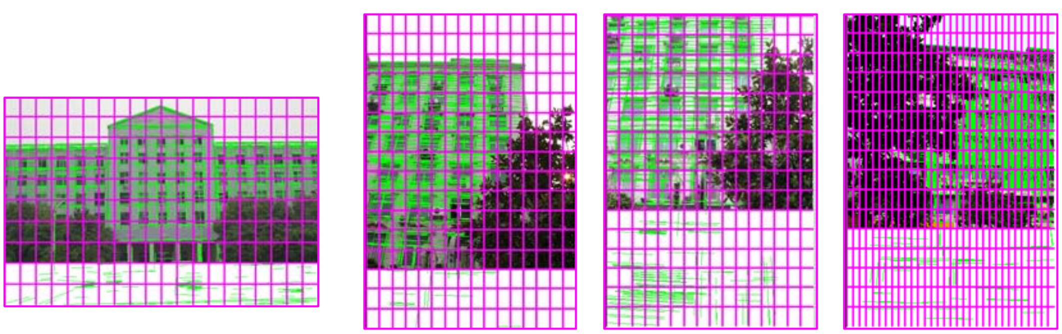
(a) Image1 (b) Image2 (c) Image3 (d) Image4

Figure 2. Parallel line detection.

The tiles of facades for determining the girds of images are shown in figure 3. Among them, figure 3(b,c,d) depict the facade tile which consists of a window and walls with many line features. Therefore, these tiles are reduced to the size of a rectangle which approximately contains a window pane in oblique terrestrial image. The size of tiles and corresponding grids of images are shown in figure 6. Figure 4 depicts the results of the occlusion row detection. The thresholds of vertical parallel lines are 8,6,16 and 12 corresponding to the four images. From figure 4, the mixed occlusions at the bottom of images are mainly removed. The rest is the occlusions almost caused by trees.

(a) Image1 (b) Image2 (c) Image3 (d) Image4

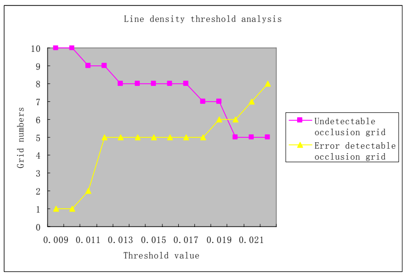
Figure 3. Facade tile.



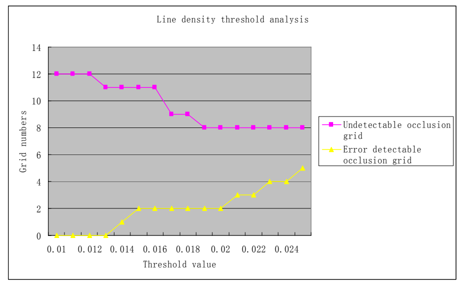
(a) Image1 (b) Image2 (c) Image3 (d) Image4

Figure 4. Occlusion row detection based on vertical lines.

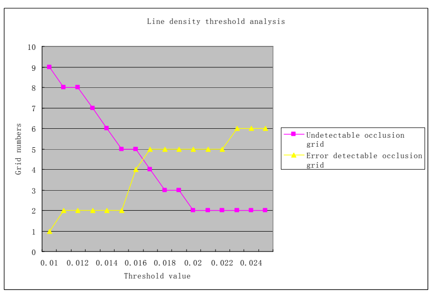
Figure 5 shows the quantity change of undetectable occlusion grids and error detectable occlusion grids under the different line density threshold. According to minimizing the numbers of undetectable and error detectable occlusion grids, the line density threshold can be determined for four images shown in figure 6.



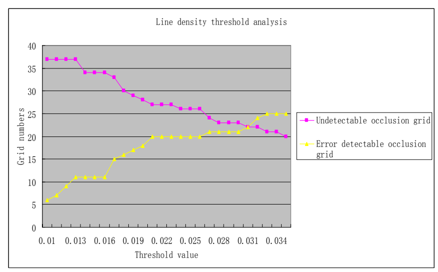
(a) Image1



(b) Image2



(c) Image3



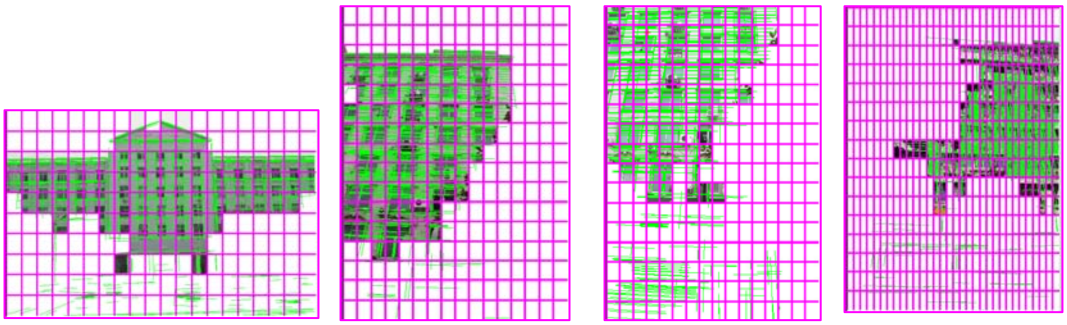
(d) Image4

Figure 5. Line density threshold and occlusion grid detection.

Grid size and line density threshold				
	Image1	Image2	Image3	Image4
Image size(pixel)	4500x3000	2592x3888	2592x3888	2592x3888
Extract line number	528	535	527	769
Tile size (pixel)	252x274	114x114	135x171	81x106
Grid size (pixel)	225x300	162x243	162x243	81x243
Line density threshold	0.018	0.019~0.02	0.015	0.014~0.016

Figure 6. Grid size and optimal line density threshold.

Figure 7 shows the results of occlusion detection under the optimal line density threshold. Most of tree occlusions are successfully detected from the facades. Their correctness and completeness are evaluated and shown in figure 8.



(a) Image1 (b) Image2 (c) Image3 (d) Image4

Figure 7. Occlusion detection.

the completeness and the correctness of occlusion extraction					
	TP(grid)	FN(grid)	FP(grid)	correctness	completeness
Image1	81	7	5	0.94	0.92
Image2	109	8	2	0.98	0.93
Image3	151	5	2	0.99	0.97
Image4	325	34	11	0.97	0.90

Figure 8. Occlusion detection evaluation.

Figure 9 depicts the process of the proposed approach from original image, occlusion detection to texture recovery. An original target image is shown in figure 9(a).And from figure 9(b), two occlusion regions are detected. One expresses the pedestrian, and the other is caused by lacking line features. A part of the pedestrian's hair keeps in the image. The reason is the occlusion covers a part of grid and parallel lines in the grid are more than the threshold of line density. Therefore it is necessary to research further for accurate occlusion detection such as hierarchical gird occlusion analysis. Figure 9(c) depicts the result of texture recovery from the adjacent image. The left bottom of the image is filled by white pixels because of no enough grey information in adjacent image.

(a)Original (b) Occlusion grid detection (c)Lost texture recovery

Figure 9. Occlusion detection and texture recovery.

CONCLUSIONS

In this paper we presented a method for occlusion detection from building facades by analyzing the numbers of parallel lines and the grid line density, generation of occlusion-free textures by planar affine transformation and bilinear interpolation. The approach doesn't detect the real boundary of different shape of occlusions, which avoids the process of extracting and tracing the edges of complex occlusions from images. And it is simple and common, especially is suitable for removing large occlusions which cover a large area in an image. As experimental results show, most of occlusion has been successfully detected including mixed occlusions in the bottom of an image without considering the attributes and shapes of occlusions. However there are some imperfections in this method. First, the occlusion region is needed to be refined. For instant, the occlusion occupies a small part of grids, and these grids are usually identified as occlusion-free grids and locate the boundary of occlusion region. Therefore these grids should be divided for accurate occlusion detection. Second, for error detectable occlusion grids, for example, facade grids are labeled as occlusion grids, it is necessary to add auxiliary information to correct these error judges and delete them from the occlusion region.

The future research will focus on improving these imperfections, And at the same time pay more attention to the method of lost texture filling for more high visual impression.

REFERENCES

Ashikhmin, M., 2001.Synthesizing natural textures, *Proceedings of the 2001 symposium on Interactive3D graphics*, pp:217-226.

Bo¨hm, J., 2004. Multi-image fusion for occlusion-free facade texturing. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(5): 867-872.

Chalup, S. K., R. Clement, et al., 2007. Representations of streetscape perceptions through manifold learning in the space of hough arrays, *Proceedings of the 2007 IEEE Symposium on Artifical Life*,pp.362-369.

Cipolla, R., Drummond, T., Robertson, D., 1999. Camera Calibration from Vanishing Points in Images of Architectural Scenes *Proc. British Machine Vision Conf.* , pp. 382-391.

Criminisi, A., P. Perez, et al., 2004. Region filling and object removal by exemplar-based image inpainting, *IEEE Transactions on Image Processing*, 13(9): 1200-1212.

Department of Community Planning, Housing and Development, 2007. *Planning Division.Rosslyn-Ballston Corridor Streetscape Standards*.

Dongsheng, W., Tao, F., Heung-Yeung, S., Songde, M.,2001.A novel probability model for background maintenance and subtraction, *Proceedings of the International Conference on Vision Interface*, pp. 109-117.

Grammatikopoulosa, L., Karrasa, G., Petsab, E., 2004, Camera Calibration Combining Images with Two Vanishing Points, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 35(5):99-104.

Heipke, C., Mayer H., Wiedemann C., 1997. evaluation of automation road extraction, *International Archives of Photogrammetry and remote Sensing*, 32(Part3-2W3),47-56.

Jianqing Z., Li P., Shugeng W.,2003. *Photogrammetry*, Wuhan University.

Kapje,S., Youngbae,H., In-So,K., 2008. Robust background maintenance for dynamic scenes with global intensity level changes, *The 5th International Conference on Ubiquitous Robots and Ambient Intelligence*, pp.759-762.

Mu¨ler, P., G. Zeng, et al., 2007. Image-based procedural modeling of facades, *ACM Transactions on Graphics*, **26**(3): 85.

Ortin, D. and Remondino, F., 2005. Occlusion-free image generation for realistic texture mapping, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(5/W17): 7 pages (onCD-ROM).

Qinghong, S., 2008. Research on hillside forest survey based on ground images, *Ph.D. Dissertation,* Wuhan University.

Rother, C., 2000. A new approach for vanishing point detection in architectural environments, *Proceedings of the 11th British Machine Vision Conference (BMVC'00),* pp:382-391.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2007.Building detection by fusion of laser scanner data and multi-spectral images: performance evaluation and sensitivity analysis, *ISPRS Journal of Photogrammetry & Remote Sensing,* 62(2007):135-149.

Tao,Y., Quan, P., Stan Z.L., Jing, L., 2004. Multiple layer based background maintenance in complex environment, *Proceedings of the Third International Conference on Image and Graphics*, pp. 112–115.

Van den Heuvel, F., 1998. 3D reconstruction from a single image using geometric constraints, *ISPRS Journal of Photogrammetry & Remote Sensing*, 53(1998): 354-368.

Wenhan, X., 2004.Camera calibration based on vanishing points of multi-image, *Ph.D. Dissertation*, Wuhan University.

Yi, D., 2008. Close-range image sequences based regular architecture 3D reconstruction, *Ph.D. Dissertation, Wuhan Universit*.

Zhizhong, K., Zuxun, Z.,Jianqing, Z., 2005. The rapidly reconstructing texture for building facades from vehicle-based image sequence, *Journal of Wuhan University(natural science edition)*, 30(11):1-5.