PERCEPTUALLY GUIDED GEOMETRICAL PRIMITIVE LOCATION METHOD FOR 3D COMPLEX BUILDING SIMPLIFICATION

Qing Zhu, Junqiao Zhao*, Zhiqiang Du, Xiaochun Liu and Yeting Zhang

State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, 430079, P.R.China – zhuqing@lmars.whu.edu.cn, johnzjq@gmail.com, duzhiqiang@lmars.whu.edu.cn, lxc_shd@126.com, zhangyeting@263.net

KEY WORDS: 3D complex building, Geometrical primitive location, Human vision system, Ray-casting

ABSTRACT:

Because of detailed geometrical components based description, 3D complex building contains the most elaborated perceptual and comprehensive semantic information. However, since the lack of optimal simplification method, the automatic LOD generation of such kind of model becomes a bottle neck which prohibited the high-fidelity 3D city applications. This paper proposed a perceptually guided geometrical primitive location method for the optimal simplification of 3D complex buildings. Firstly, the rendered image is snapped and a 2D discrete wavelet transform based human vision system filtering approach is adopted to extract the imperceptible details in the image, and then a kind of visual difference image is generated with sufficient perceptual information. Secondly, a ray-casting like method is proposed to precisely map the perceptual information from the image onto the geometric primitives. The statistics is carried out to determine whether a traced primitive is to be preserved or simplified. The results show that this method is able to efficiently locate the perceptible primitives and leave the imperceptible and undisplayable primitives to be further handled by simplification operations which enable a strong perceptual feature preserved simplification of 3D complex building models.

1. INTRODUCTION

Because of the limitation of rendering capability of modern display hardware, the simplification of 3D complex building model with highly detailed geometrical components is necessary in interactive applications (Li et al., 2006). However, this kind of 3D models is not only rich in coplanar, perpendicular and parallel features (Kada, 2007), but also contains components of various size which make it too difficult to be handled using existing simplification methods.

The essential problem of the automatic simplification of 3D complex building models is the optimal extraction of the geometric primitives that to be preserved (or simplified). Researchers of building generalization adopt the ground plan or the cell-decomposition to find trivial geometry from walls and roof of a building for reduction (Glander and Döllner, 2007; Kada, 2008; Sester, 2000), and methods based on scale space as well as using CSG tree aggregate the adjoined planes or small faces (Forberg, 2007; Thiemann and Sester, 2004). In these methods, geometric primitives are extracted in order to preserve the building structure and semantics. A detail review of these methods is done by (Sester, 2007). However, only the simple model is adopted in these researches rather than the 3D complex model. In computer graphics, mesh simplification is a well researched topic (Luebke et al., 2003). But the proposed polygon reduction algorithms are suitable for continuous surface rather than component based building model. The main reasons are the strong structured features of building and the aggregate detail of components (Kada, 2007; Cook et al., 2007). To overcome these defects, human perception is introduced into polygon reduction in order to allocate more geometry to more visual important details, vice versa (O'Sullivan et al., 2004). The proposed methods can be divided into two kinds: one is implemented in model space and the other in screen space. The former is usually achieved by evaluating the perception of geometry using projected errors or curvatures (Luebke and Hallen, 2001; Lee et al., 2005). However, these methods adopt a simpler implementation of human vision system (HVS) model which may not be precise. Otherwise, screen is the general intermedium for people to observe 3D models, extracting perceptual information by means of screen space is therefore preferable. Approaches based on screen space errors are able to evaluate the perceptibility of image on screen by introducing more rigorous HVS models derived from the researches on image processing (Winkler, 2000). However, existing perceptually guided simplification methods treat the perceptual information as weights to adjust the sequence of simplification operations such as edge collapse etc. (Qu and Meyer, 2008), which is obviously not suitable for aggregate details.

In conclusion, current simplification methods are proposed either for simple building models or continue surface models rather than 3D complex building models. In this case, an optimal extraction of the geometric primitives of a 3D complex building model for simplification is needed. This paper adopts a 2D discrete wavelet transform (DWT) based HVS filtering technique to extract the imperceptible details in the rendered image snapped from the model. Then, a visual difference image contains perceptual information is precisely mapped to the model using a ray-casting like method to locate all the perceptible geometry. Such idea has been firstly proposed by (Du et al., 2008) and the main contribution of this paper is a fast and precise output sensitive perceptually guided method to locate the perceptible geometric primitives.

* Corresponding author.

2. PERCEPTUALLY GUIDED GEOMETRICAL PRIMITIVE LOCATION ALGORITHM

Adopting perceptual information in screen space to guide simplification, a connection between 2D screen space and 3D model space should be created. Based on the principle of rasterization rendering, pixels are generated by rasterization of geometric primitives (Shreiner et al., 2004). Thus, there are two paths to build this connection. One is the projection of geometric primitives onto the screen. The perceptual information of the projected primitive is then decided by the projected coordinates of primitive. For example, Qu adopted a projected texture like method to find the vertex index in a vision important image (Qu and Meyer, 2008). However, this kind of method could not directly locate perceptible primitives and it can hardly deal with occlusion of primitives.

The other way to bridge 2D screen and 3D model is the ray-casting technique. The tracing back from pixels to its corresponding geometric primitives is the inverse process of rasterization rendering, as illustrated in Figure 1**.** Based on the correct rasterization rule, the emitted ray from pixel would be able to directly find the exact geometric primitive that contributes to the pixel. Perceptual information carried by the ray is then transmitted to the primitive and finally perceptible primitives are able to be located. There are several advantages of this method comparing to the former: Firstly, the perceptual information could be fully transmitted from 2D pixels to 3D primitives. Secondly, ray-casting automatically eliminates the occlusion which wipes out the inner structure naturally. Thirdly, when simplifying model for using in distance, this kind of method is able to eliminate the primitives that are smaller than the resolution of screen. So it is output sensitive.

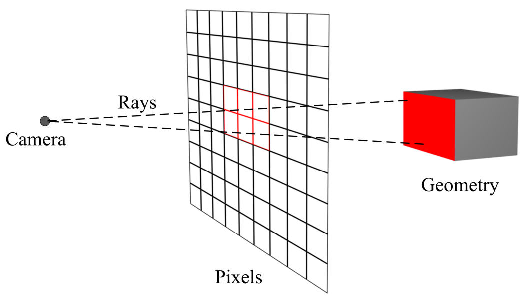

Figure 1. The tracing back from pixels to geometry

The flow chart of the algorithm is illustrated in Figure 2**.** The first step is the extraction of perceptual information in the rendered image of 3D model. Various researches on image processing had adopted HVS based perception models, such as evaluating the visual similarity of two images (Bradley, 1999; Wang et al., 2004), and image compression (Nadenau, 2000). The former requires two input images. Existing methods based on it are either too costly (Lindstrom and Turk, 2000), or not precise (simplification adopting visual discrimination metric (VDM) method (Qu and Meyer, 2008)). The latter adopts the HVS to compress the imperceptible details in an image more aggressively which is suitable for our application.

After rendering the image of 3D model at a given viewing distance, a contrast sensitivity function (CSF) based HVS model is introduced to simulate the approximate result of vision perception. Then a visual difference image indicates the imperceptible detail in the original image is generated by subtracting the vision simulated image and the original image. The vision simulation procedure includes three key steps: 2D DWT decomposition, HVS filtering and 2D DWT reconstruction. To precisely implement HVS, the image has to be transformed from the spatial domain to the frequency domain. Among many existing transform approaches, the 2D DWT better fits the HVS model because 2D DWT decomposition is similar to the multiple-channel model of the HVS, which allows the processing to be acted on each spatial frequency channel independently (Bradley, 1999). After decomposition, the CSF is implemented as finite impulse response (FIR) filters, which is more precise and adaptive (Nadenau et al., 2003).

Finally, the inverse 2D DWT transform is carried out to reconstruct the vision simulated image. The visual difference image is then generated. In this image, the imperceptible details are represented by pixels which have a value lager than 0 and the perceptible details are represented by black pixels which mean the lossless details. The perceptible primitives are then located by precisely mapping the visual difference image onto the model using a ray-casting like method which is to be detail discussed in the following section.

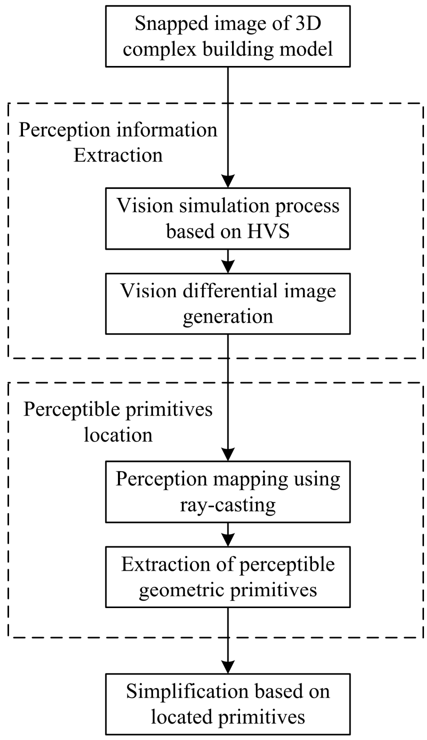

Figure 2. The flow chart of geometrical primitive location method

3. FAST LOCATION OF PERCEPTIBLE PRIMITIVES

There are two functions of the ray-casting like method. One is to extract the displayable primitives of the model. The other is to map the pixel value of the visual difference image onto the primitive that contributes to it. In order to guide the simplification, all the displayable primitives are measured based on the pixel values transmitted from the image. The perceptible ones are to be preserved and the rest that are undisplayable or imperceptible are to be simplified.

Although this method is implemented off-line, a proper index structure is needed for efficiently locating the primitives because of the huge data size and complex geometry. The most popular index structures for ray-tracing application are the BSP-tree, KD-tree and Octree (Havran, 2000). Considering the regular structure and relative simplicity of components of 3D complex building model, the Octree is chosen. The leaf node of the index contains the intersected minimum bounding boxes (MBB) of components which are not further split for quickly constructing the index structure. The ray is generated employing the position of camera and the projective coordinates of pixels in the near clipping plane.

```
//Mapping perceptual information using Ray-casting
Input Vision Differential Image (VDimage);
Get the start pixel (Pix) in VDimage;
Do{
    //Primitive locating
    Generate a Ray (R) based on Pix and Camera;
    Get the intersection point (Pt) of R and the MBB of Octree;
    While (Pt is inside the MBB of Octree){
    Find the intersection leaf node (Lnode) based on Pt;
    Record the intersected triangles (Tris) between R and the Comps
    belong to Lnode;
    Compute the other point (Pt_out) intersected between R and Lnode;
    Pt=Pt_out;
    }
    Rank the list of recorded Tris of to find the closest one (Tri_close);
    //Perceptual information mapping
    IF(the value of Pix >0){
        The traced pixel number of Tri_close++;
        The imperceptible pixel number of Tri_close++;
    }ELSE
        The traced pixel number of Tris_close++;
    Pix =the next pixel of VDimage;
}While (Pix belongs to VDimage)

//Extracting the perceptible geometic primitives (triangles)
Get the first Tri of the model;
While (Tri belongs to the model){
    The simplification mark of Tri= TRUE; // initialed to be simplified
    IF (The traced pixel number of Tri!=0)
        IF((imperceptible pixel number / traced pixel number) <0.5)
            The simplification mark of Tri= FALSE; //to be preserved
}
```

Figure 3. The pseudo code of primitive location algorithm

It should be noted that the screen coordinates of pixels for rasterization are half-integer so the ray should be adjusted by adding 0.5 to the X and Y value of screen coordinates of pixels (Segal and Akeley, 2008) .

A fast ray intersection method is adopted to efficiently locate the geometric primitive, in which the intersection test is carried out directly with the leaf nodes of the Octree along the ray. Then, pixel values which represent the perceptual information are recorded in the data structure of primitives. The result contains two types of numbers. One is the number of pixels that have a value larger than 0 indicating the quantity of the imperceptible details and the other is the number of pixels that have a value equal to 0 indicating the quantity of the perceptible details. These two numbers are accumulated during ray-casting and finally each primitive record, firstly, the number of traced pixels, which is the total number of pixels contributed by this primitive; secondly, the number of imperceptible pixels contributed by this primitive. If the number of traced pixels is equal to 0, then this primitive is marked as undisplayable and if the number of imperceptible pixels takes up half or more of the number of traced pixels, this primitive is marked as imperceptible. The rest are marked as perceptible primitives. The pseudo code of the algorithm is shown in Figure 3. After this process, all the primitives of the model are marked and the simplification is implemented based on the results.

4. EXPERIMENTAL ANALYSIS

A 3D complex building model of Chinese style built up by 98265 components and 706978 triangles was selected in the experiment, as illustrated in Figure 4. The observation equipment was a 19 inch TFT LCD monitor with the resolution of 86.27 dpi (1280x1024) and a luminance of 200 (cd/m^2). The observation distance was set at 0.5m which is a typical visual environment for desktop applications.

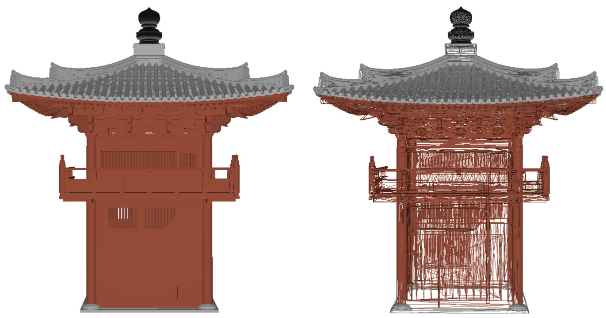

Figure 4. The 3D complex building model

The maximum spatial frequency perceived by human form the monitor was calculated to be 14.82 cycles/deg (cpd) by equation (1)(Nadenau et al., 2003):

$$f_{\max} = 0.5 f_s = 0.5 * (2v \tan(0.5^{\circ}) r / 0.0254) \qquad (1)$$

where f_{max} is the maximum frequency, which is down sampled from f_s, the sampling frequency of the screen, at Nyquist rate of 0.5 cycles/pixels. The adopted CSF expression is widely used in many other works (Luebke et al., 2003):

$$A(\alpha)=2.6(0.0192+0.144\alpha)\exp(-(0.144\alpha)^{1.1}) \qquad (2)$$

Then, Daubechies 9/7 wavelet was selected to transform the image into frequency domain because of its fitness for the image processing (Antonini et al., 1992). The CSF was implemented as FIR filters to convolute with the wavelet coefficients and the vision simulated image was generated by the inverse DWT transform, as illustrated in Figure 5 (middle). Figure 5 (right) shows the visual difference image generated by subtracting the above two images, in which black pixels inside the model represent the perceptible detail and the grey pixels represent the imperceptible details. It can be seen that most of

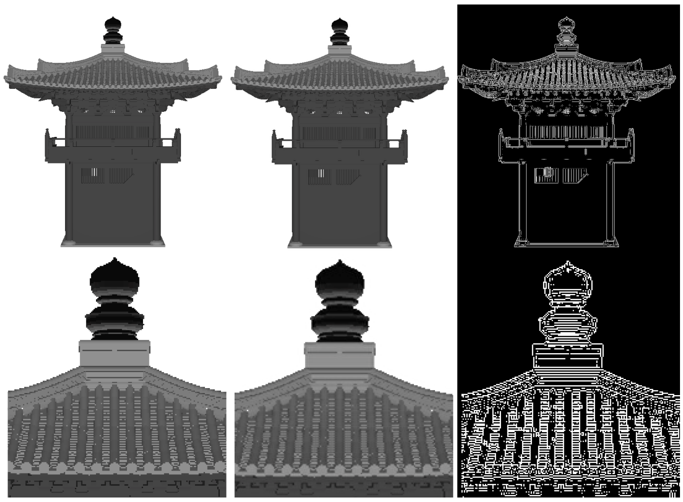

Figure 5. The first row: the rendered image (left); the vision simulated image (middle) and the visual difference image (right). The second row: the 300% zoom in images of the first row

the imperceptible details located at the regularly aligned shape edges and the trivial structures which represent high spatial frequency and low contrast sensitivity signals. The result is in conformity with results of perceptual-based image processing researches (Nadenau et al., 2003).

Figure 6 shows the index structure of the test model, the depth of Octree was 9 and the maximum number of components contained in the leaf node was 40. The visual difference image was then adopted to generate rays to intersect with the index structure and finally map the perceptual information onto the geometric primitives. The result of the location process at viewing distance of 56 m is illustrated in Figure 7, where green triangles represent the perceptible primitives; red triangles represent the imperceptible primitives and blue triangles represent the undisplayable primitives. It can be seen that at the fixed viewing distance, almost all the triangle primitives of the model were marked as perceptible (green) as expected. But, moving the camera closer to the model, many trivial triangles that cannot be displayed at the fixed viewing distance were revealed, as well as the imperceptible ones, as shown in Figure 7 (right). It should also be noted that some of the partly occluded triangles may also be marked as imperceptible (red) because of the perceptibility of its uncovered parts, such as the red triangle on the left side of the model in Figure 7 (right).

The algorithm was implemented on four facades of the model respectively and all the marked triangles were then sent to simplification process using half edge collapse operator based on quadric error metric (Garland and Heckbert, 1997). All the undisplayable and imperceptible triangles were to be decimated by collapsing one of its edges. The edge that connected to the vertex of the perceptible triangle was collapsed to that vertex. So the perceptible triangles are preserved in the simplification.

The result was shown in Figure 8. The simplified model contained 6175 components and 159020 triangles which is equal to only 22.5% data size of the original model but preserved the similar appearance. Simplification examples at other viewing distances were also given, as shown in Figure 9. Notice that observing at the fixed viewing distance, the aggressively simplified model looks similar to the original model which proved the effectiveness of this approach.

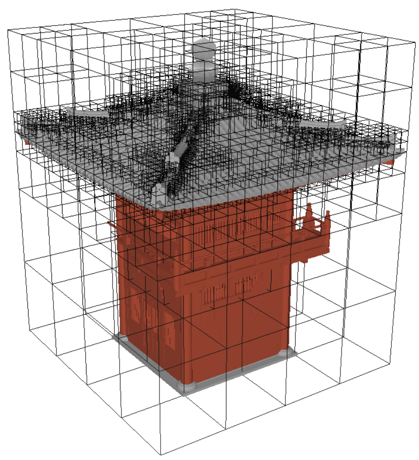

Figure 6. The generated Octree index structure

5. CONCLUDING REMARKS

Aiming the LOD model generation of 3D complex building model, this paper proposed a perceptually guided geometrical primitive location method which can precisely locate the perceptible primitives and provide useful perceptual information, which guarantees the simplification to preserve the perceptual features of final LOD models. Future research topics include the efficient perceptually guided simplification methods of 3D complex building models as well as the simplification combined with texture.

ACKNOWLEDGEMENT

The work described in this paper is supported by the National Natural Science Foundation of China (No. 40871212, No. 40721001) and the National High Technology Research and Development Program (No.2006AA12Z224 and No. 2008AA121600).

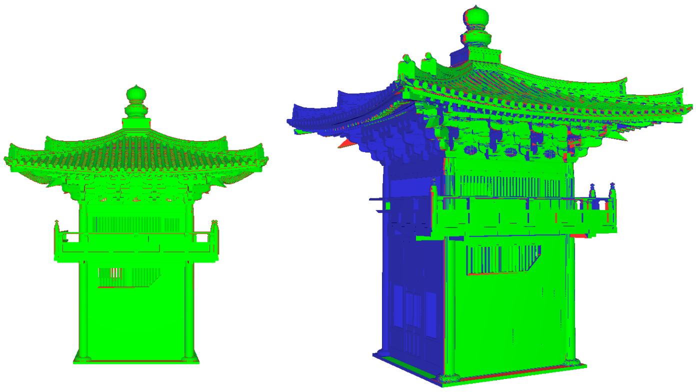

Figure 7. The result of perceptually guided primitive location at 56 m
(Left: observation at 56 m; Right: closer observation)

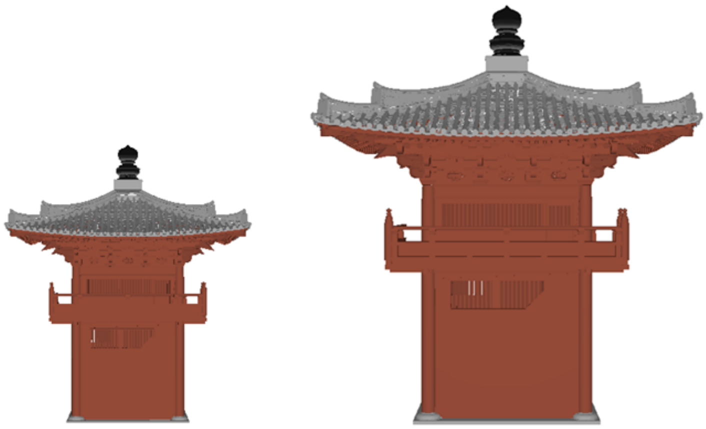

Figure 8. The result of simplification
(Left: observation at 56 m; Right: closer observation)

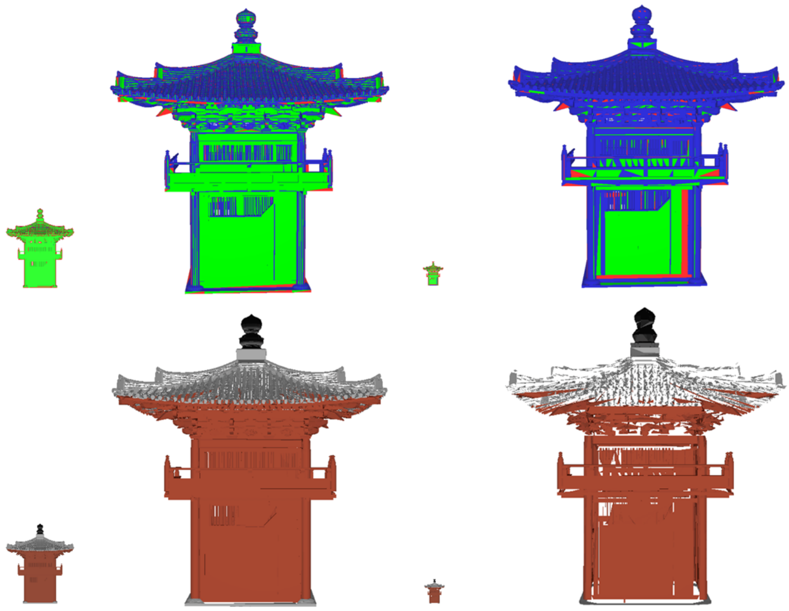

Figure 9. Results of primitive location and simplification at farther viewing distances
(Left: observation at 132 m and closer observation; Right: observation at 427 m and closer observation)

REFERENCES

Antonini, M., Barlaud, M., Mathieu, P. and Daubechies, I., 1992. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2): 205-220.

Bradley, A.P., 1999. A wavelet visible difference predictor. *IEEE Transactions on Image Processing*, 8(5): 717-730.

Cook, R.L., Halstead, J., Planck, M. and Ryu, D., 2007. Stochastic simplification of aggregate detail. *ACM Transactions on Graphics*, 26(3).

Du, Z.Q., Zhu, Q. and Zhao, J.Q., 2008. Perception-driven simplification methodology of 3D complex building models, In: *ISPRS2008*, Beijing.

Forberg, A., 2007. Generalization of 3D building data based on a scale-space approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(2): 104-111.

Garland, M. and Heckbert, P.S., 1997. Surface simplification using quadric error metrics, In: *Proceedings of SIGGRAPH 1997*, Los Angeles, pp. 209-216.

Glander, T. and Döllner, J., 2007. Cell-based generalization of 3D building groups with outlier management, In: *Proceedings of the 15th International Symposium on Advances in Geographic Information Systems (ACM GIS)*. ACM, New York, Seattle, Washington, pp. 1-4.

Havran, V., 2000. Heuristic ray shooting algorithms, Czech Technical University, Prague.

Kada, M., 2007. A Contribution to 3D Generalisation, In: *51st Photogrammetric Week 2007*, Stuttgart.

Kada, M., 2008. Generalization of 3D building models for map-like presentations, In: *ISPRS2008*, Beijing.

Lee, C.H., Varshney, A. and Jacobs, D.W., 2005. Mesh saliency. *ACM Transactions on Graphics*, 24(3): 659-666.

Li, D.R., Zhu, Y.X., Du, Z.Q. and Hong, T., 2006. Virtual Tang-Style Timber-frame building complex. *Lecture Notes in Computer Science*, 4282:880-888.

Lindstrom, P. and Turk, G., 2000. Image-Driven Simplification. *ACM Transactions on Graphics*, 19(3): 204-241.

Luebke, D. et al., 2003. *Level of Detail for 3D Graphics*. Morgan Kaufmann, San Francisco.

Luebke, D. and Hallen, B., 2001. Perceptually Driven Simplification for Interactive Rendering, In: *Proceedings of the Eurographics Workshop. Rendering Techniques 2001*. Springer, London, United Kingdom, pp. 223-234.

Nadenau, M., 2000. Intergration of Human Color Vision Models into High Quality Image Compression, Swiss Federal Inst. Technol, Lausanne, Switzerland.

Nadenau, M.J., Reichel, J. and Kunt, M., 2003. Wavelet-based color image compression: exploiting the contrast sensitivity function. *IEEE Transactions on Image Processing*, 12(1): 58-70.

O'Sullivan, C., Howlett, S., McDonnell, R., Morvan, Y. and O'Conor, K., 2004. Perceptually adaptive graphics. *Eurographics State of the Art Reports*: 141-164.

Qu, L. and Meyer, G.W., 2008. Perceptually guided polygon reduction. *IEEE Transactions on Visualization and Computer Graphics*, 14(5): 1015-1029.

Segal, M. and Akeley, K., 2008. *The OpenGL Graphics System:A Specification (Version 3.0)*. The Khronos Group Inc.

Sester, M., 2000. Generalization based on least squares adjustment, In: *International Archives of Photogrammetry and Remote Sensing*, Amsterdam, pp. 931-938.

Sester, M., 2007. 3D Visualization and Generalization, In: *51st Photogrammetric Week*, Stuttgart, Germany, pp. 285-295.

Shreiner, D., Woo, M., Neider, J. and Davis, T., 2004. *OpenGL Programming Guide: The Official Guide to Learning OpenGL*, Version 1.-4th ed. Addison-Wesley Professional.

Thiemann, F. and Sester, M., 2004. Segmentation of Buildings for 3D-Generalisation, Working Paper of the ICA Workshop on Generalisation and Multiple Representation, Leicester, UK.

Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600-612.

Winkler, S., 2000. Vision models and quality metrics for image processing applications, Swiss Federal Inst. Technol, Lausanne, Switzerland.