



## Technical Section

# Quantitative analysis of discrete 3D geometrical detail levels based on perceptual metric

Qing Zhu <sup>\*</sup>, Junqiao Zhao, Zhiqiang Du, Yeting Zhang

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China

## ARTICLE INFO

## Article history:

Received 30 October 2008

Received in revised form

12 October 2009

Accepted 15 October 2009

## Keywords:

Levels of detail

Quantitative analysis

Perceptual metric

Human vision system

## ABSTRACT

Aiming at the fundamental issue of optimal design of discrete levels of detail (LOD) for the visualization of complicated 3D building façades, this paper presents a new quantitative analytical method of perceptible 3D details based on perceptual metric. First, the perceptual metric is defined as the quantitative indicator of the visual perceptibility of façade details at a given viewing distance. Then, according to the human vision system, an algorithm employing 2D discrete wavelet transform and contrast sensitivity function is developed to extract the value of perceptual metric from the rendered image of the façade. Finally, a perceptual metric function is defined, based on the perceptual metric values extracted at equal interval viewing distances. The minimum detail redundancy model is then proposed for the optimal design of discrete LODs. This method provides a quantitative instruction for generating discrete LODs. The experimental results prove the effectiveness and great potential of this method.

© 2009 Elsevier Ltd. All rights reserved.

## 1. Introduction

Online virtual globe tools, led by Google Earth, are dramatically changing the way we interact with spatial data from 2D maps to the 3D Virtual Geographic Environments (VGEs) [1,2]. Up to now, 3D representations of geographic information in computers have been known as VGEs, which are of increasing importance in urban areas and provide more accurate and flexible mathematical models, effective tools, and user interfaces for geospatial communication. Therefore, 3D models in VGEs do not just have photorealistic appearances, but they are also accurate, with credible geometry and topology information of the built environment, achieved by making full use of advanced 3D measurement (such as imaging, LIDAR, GIS, and CAD) [3].

When representing a building, façades are often the most important elements, as they contain the exterior geometric details of the building. However, the number of polygons that represents the 3D geometry of complicated façades always exceeds the rendering capability of the hardware; a trade-off between complexity and performance is required. A preferred solution is the level of detail (LOD) technique, which increases the efficiency of rendering by reducing insignificant geometry for visualization [4]. One of the bottlenecks of this technology for complicated 3D

façades is the automatic generation of discrete levels of detail (LODs). Given the target number of detail levels, a collection of LODs are produced based on manually assigned control factors, such as a target number of polygons with equal interval partitions, or exponential interval partitions, or employing trial and error based on plenty of pre-simplified models [5]. However, assigning various factors manually to generate LODs for the original model is usually not optimal. Inevitable detail redundancy between LODs brings about excessive rendering cost. This paper proposes an optimal design approach for discrete LODs of complicated 3D façades based on analyzing the perceptibility of façade details at various viewing distances, which prevents the costly pre-simplification stage and provides quantitative instructions for the construction and management of discrete LODs.

This paper is organized as follows. The related work on discrete LOD is reviewed in Section 2. Section 3 introduces the definition of the perceptual metric and perceptual metric function. Section 4 presents the algorithm based on the perceptual metric function for LOD optimization, and Section 5 discusses the experimental results. Finally, the conclusions are presented in Section 6.

## 2. Related work

Discrete LOD is a traditional approach of the LOD technique, first proposed by Clark in 1976 [6]. The most important advantages of this kind of LOD are its simplicity, high efficiency, and the support of offline preprocessing [4], which make it suitable for visualization of complicated 3D façades. The

<sup>\*</sup> Corresponding author. Tel.: +86 27 68778322; fax: +86 27 68778229.

E-mail addresses: [zhuqing@lmars.whu.edu.cn](mailto:zhuqing@lmars.whu.edu.cn), [zhuq66@263.net](mailto:zhuq66@263.net) (Q. Zhu), [johnzjq@gmail.com](mailto:johnzjq@gmail.com) (J. Zhao), [duzhiqiang@lmars.whu.edu.cn](mailto:duzhiqiang@lmars.whu.edu.cn) (Z. Du), [zhangyeting@lmars.whu.edu.cn](mailto:zhangyeting@lmars.whu.edu.cn) (Y. Zhang).

generation of a collection of discrete LODs for such models comprises the following steps: first, determine the target number of detail levels; second, determine the proper control factor for each detail level, such as the number of polygons; and third, generate the LODs [3].

The proper target number of detail levels depends on various factors including the rendering capacity of the visualization system, the complexity of model, etc. As a result, it is hard to determine fully automatically. Moreover, the visualization system usually supports a fixed number of LODs, so the target number has to be selected manually.

Determining the control factor for each detail level is very important to the discrete LOD since models are often too complex to be simplified in real-time or for the simplification operations to be pre-recorded by progressive meshes [7]. As a result, most of the existing approaches for determining the detail level are based on pre-simplification processes. The original model is simplified to produce plenty of candidate LODs according to certain criteria, such as screen-space errors or object-space errors. Then proper LODs are selected by evaluating the errors of the candidates [4,8]. Although more sophisticated approaches based on the human vision system (HVS) have been proposed [9–11], it is too costly to pre-simplify a complicated 3D façade into a number of candidates. On the other hand, the target number of polygons and other control factors are also adopted to define each detail level, such as generating a collection of LODs based on pre-estimated runtime

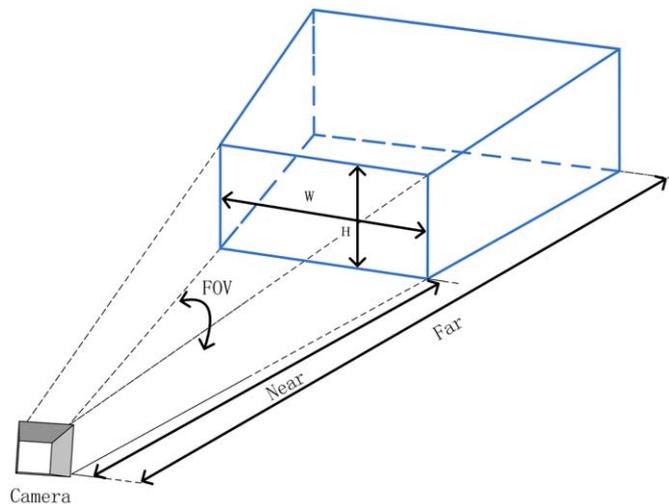


Fig. 1. The viewing frustum.

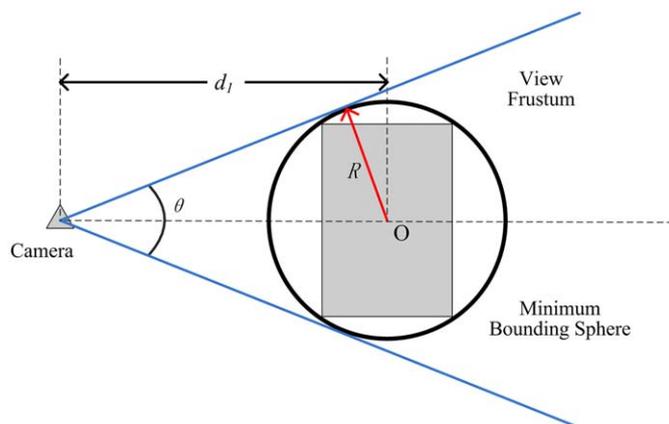


Fig. 2. The computing of  $d_1$ .

budgets [12], but it is hard to estimate the budgets accurately in an offline fashion.

To produce the LODs, various automatic or semi-automatic simplification algorithms have been comprehensively reviewed by David et al. and Cohen and Manocha [4,13].

In conclusion, there is still no good solution to the precise analysis and design of LODs for complex 3D models such as component-structured building façades, which hinders the practical use of discrete LOD. In the following sections, a new idea about the quantitative analysis of the perceptibility of geometric details is presented; such perceptual knowledge is extracted from the rendered images of a 3D façade, which is useful to guide the design of the LODs.

### 3. Perceptual metric and perceptual metric function

#### 3.1. Definition of perceptual metric

The process of human visual perception of 3D models consists of two successive stages. First, based on given parameters, a virtual camera samples the 3D model. This is called the 3D projection. Its result is a rendered image on the screen in which primitives less than the pixel resolution of the display are invisible to the viewers. Second, visual perception of the rendered image is generated through the retina, visual pathways, and finally the visual cortex [14]. This stage is modeled by the human vision system (HVS). Details displayed on the screen that are beyond the capacity of human visual perception are filtered out in this stage.

In order to quantify the visual perceptibility of the details in the 3D model, a perceptual metric (PM) is defined as

$$PM = VPs / TPs \quad (1)$$

where VPs is the number of pixels representing the perceptible details of the model in the rendered image and TPs is the number of pixels representing the total details of the model in the rendered image. These two numbers of pixels indicate the quantitative difference between the two stages of visual perception. The ratio between these two numbers, called PM, therefore indicates the perceptibility of geometric details. By means of PM, the geometric LODs of a 3D model can thus be quantitatively and intuitively analyzed.

#### 3.2. Calculation of perceptual metric

In order to obtain a reliable and accurate PM, the two stages of perception should be quantified. Before rendering, a view frustum is defined as illustrated in Fig. 1 [15]. The 3D model is then projected onto the projection plane according to the perspective projection transformation. Placing the model in the center of the sight, the number of non-background pixels can be counted as the number of pixels that represent the object in the rendered image.

However, restricted by the viewing frustum, this evaluation is not accurate when the model is near to the camera because the clipping plane will eliminate that part of the model outside the view frustum. Therefore,  $d_1$ , where the model is shown maximally on the screen, is set as the beginning of the evaluation, as illustrated in Fig. 2. The equation to calculate  $d_1$  is as follows:

$$d_1 = R / \sin(\theta/2) \quad (2)$$

where  $R$  is the radius of the minimum bounding sphere of the model measured in meters and  $\theta$  is the field of view (FOV) of the virtual camera. The details of the visible surface of the model are considered to be entirely displayed at  $d_1$ . Then,

$$TPs = Pn_1 \quad (3)$$

where  $Pn_1$  is the number of pixels that represent the object in the image rendered at the viewing distance  $d_1$ .

As shown in Fig. 3, after rendering the image of the 3D model, a HVS model is introduced to simulate the result of visual perception. Then the number of pixels represented the perceptible details in the rendered image can be calculated based on the result. In HVS, the contrast sensitive function (CSF) is one of the most important issues for the generation of vision because it concerns the decreasing sensitivity for higher spatial frequencies [16]. This property fits the CSF well for the LOD analysis because the imperceptible details on the surface of the model are usually small concave or convex shapes, which represent relatively high spatial frequencies in the rendered image. There are both luminance CSF and color CSF. However, color CSF has proved to be related to specific applications, and there is still no good definition of it [17]. Moreover, color contributes little to the identification of details [18]. As a result, luminance CSF is employed to construct the HVS filter.

The vision simulation procedure includes three key steps: 2D discrete wavelet transform (DWT) decomposition, HVS filtering, and 2D DWT reconstruction. To implement CSF precisely, 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 the 2D DWT decomposition is similar to the multiple channel models of the HVS, which allows the processing to act on each spatial frequency channel independently [16]. It also retains the frequency as well as the spatial information of the signal, which supports the further extension of other perceptual effects such as masking. The Daubechies 9/7 bi-orthogonal wavelet was selected because of its fitness for image processing [19].

After the DWT composition, CSF is implemented to process the wavelet coefficients. As a more precise approach, the *adaptive*

approach is adopted [20], which makes use of the finite impulse response (FIR) filter rather than employing simple weighting factors to implement CSF. Taking the observing distance from the human eyes to the screen and the screen resolution into account, the CSF is mapped as a modulation function to the corresponding frequency band in the down-sampled domain to construct the filters. Then the filters are employed to convolute with the wavelet coefficients.

Finally, the inverse DWT transform is carried out to reconstruct the filtered wavelet coefficients. The vision-simulated image is then generated, in which the imperceptible details are blurred and the perceptible details are retained. The number of pixels having different values between the rendered image and the vision-simulated image indicates the quantity of the imperceptible details. So VPs is calculated by

$$VPs = Pn - IPn \tag{4}$$

where  $Pn$  is the number of pixels of the object in the rendered image and  $IPn$  is the number of pixels showing different values between the rendered image and the corresponding vision-simulated image.

The PM is then computed as

$$PM = (Pn - IPn) / Pn_1 \tag{5}$$

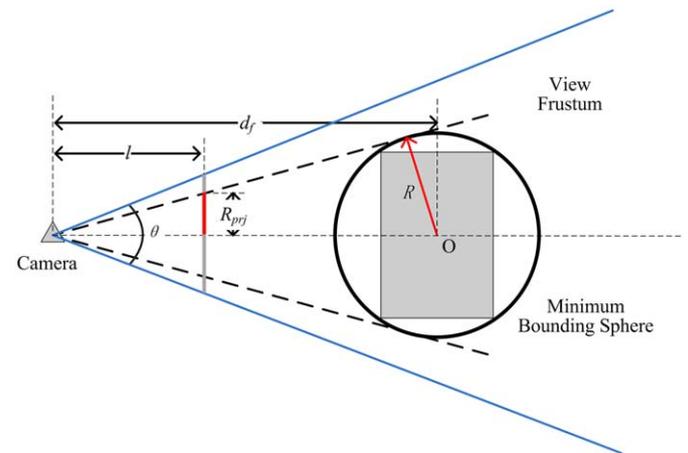


Fig. 4. The computing of  $d_f$ .

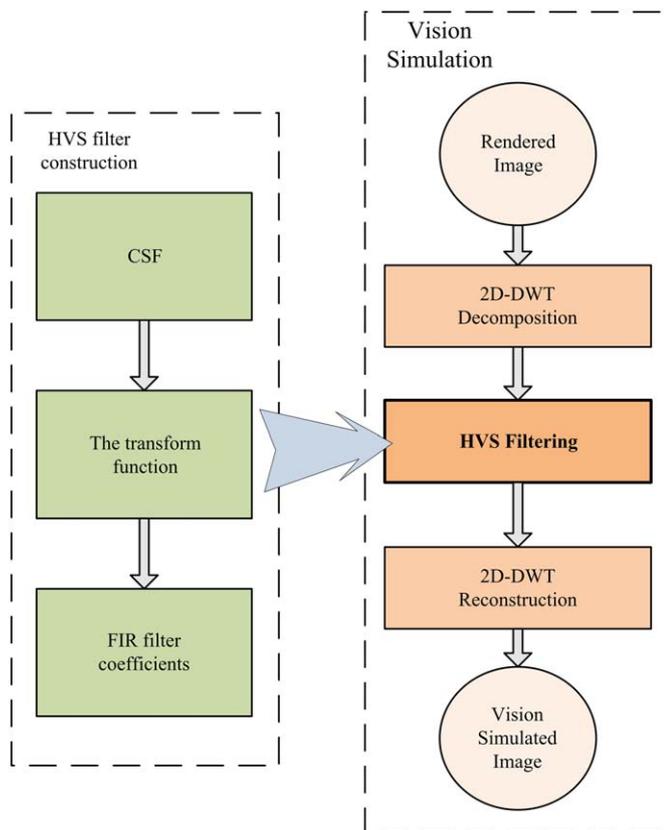


Fig. 3. The vision-simulation process.

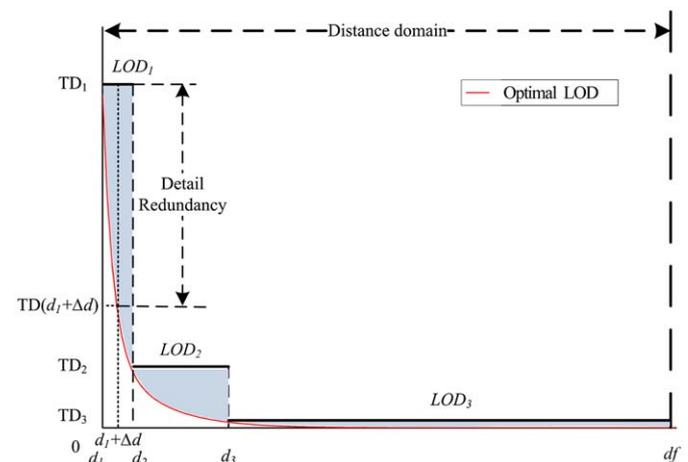


Fig. 5. Illustration of LODs design.

### 3.3. Perceptual metric function

Viewing distance is the most important parameter to the discrete LOD [21]. Therefore, the quantitative analysis of LODs

should be based on exploring the relationship between viewing distance and the perceptibility of model details, which is defined by the perceptual metric function  $PM(d)$ . A scheme in which  $d$  increases from near to far in order to render the model with equal

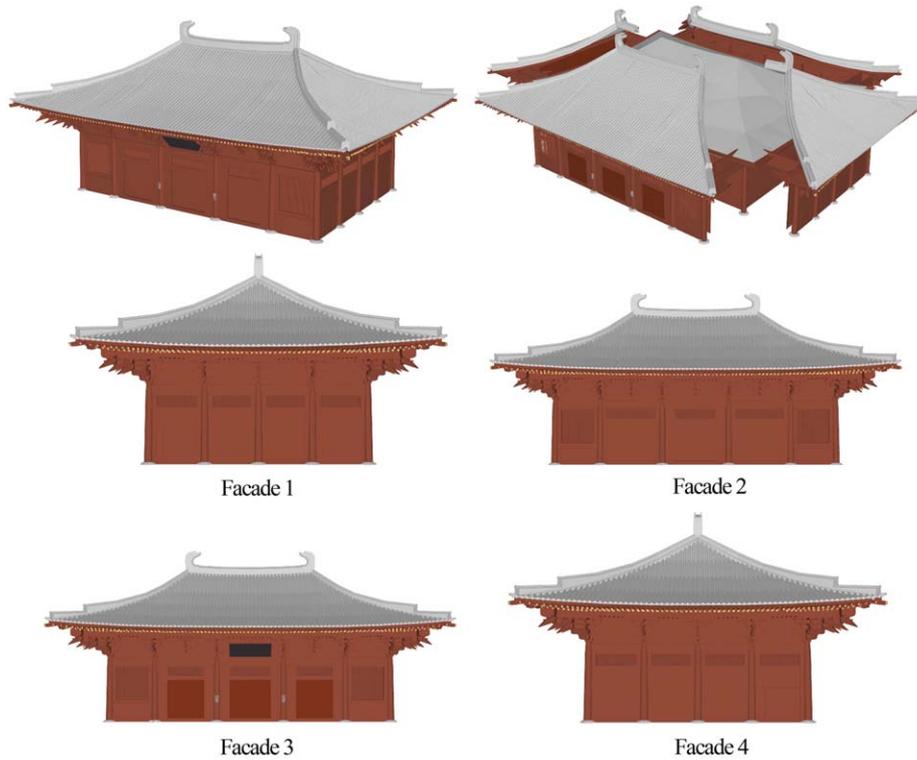


Fig. 6. Four façades of the great shrine hall.

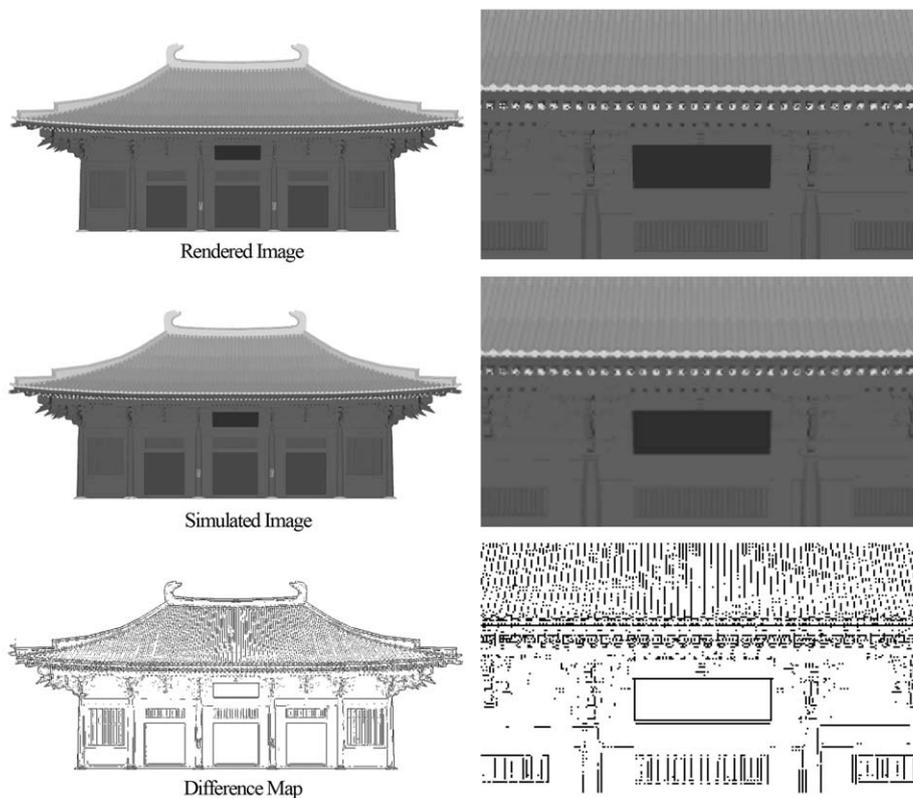


Fig. 7. The rendered image and its 300% zoom-in image (top); the vision simulated image and its 300% zoom-in image (middle); the visual difference map and its 300% zoom-in image (bottom).

intervals is employed to obtain a series of PM values. The expression for  $PM(d)$  is then obtained by the least squares fitting of these values.

The domain of *viewing distance* is defined as  $[d_1, d_f]$ , where  $d_1$  has already been obtained by Eq. (2) and  $d_f$  is the furthest boundary of *viewing distance* when the model projected on the screen reaches the size of one pixel in the screen. When  $d$  is larger than  $d_f$ , the model will not be displayed. Based on the perspective projection, the projection size of the model is calculated by the following equation, as illustrated in Fig. 4.

$$R_{prj} = (l \cdot R) / \sqrt{d_f^2 - R^2} \tag{6}$$

where  $R_{prj}$  is half of the length of the projected model in the projection plane and  $l$  is the distance from the projection plane to the camera center, all measured in meters.  $R_{pix}$ , the number of pixels that represent the  $R_{prj}$  on the screen, is calculated by equation

$$R_{pix} = r \cdot R_{prj} / 0.0254 \tag{7}$$

where  $r$  is the resolution of the screen measured in dots per inch (dpi), 0.0254 is the conversion factor between meter and inch. When  $R_{pix}$  is equal to 0.5 pixels, the projected length of the model

**Table 1**

The results of the evaluation for the vision simulation where the VDP shows that 0.1249% of pixels in the image have a probability of 75% of being noticed as different, whereas 0.0648% of pixels have a 95% chance of being detected.

Metrics	Results	
VDP	0.1249% (75%)	0.0648% (95%)
SSIM	99.5637%	

The SSIM returned a value of 99.5637% to indicate the level of perceivable differences between these two input images. If the value is equal to 1, the two images are perceptually equivalent.

is equal to one pixel. The furthest boundary  $d_f$  is then calculated by

$$d_f = R \sqrt{1 + (2r \cdot l / 0.0254)^2} \tag{8}$$

For a 3D model, the value of  $PM(d)$  decreases with the increase of *viewing distance*. The reasons are as follows. First, due to the perspective projection, the projection size of the model decreases when the model moves further away. Second, the spatial frequency of details in the model perceived by human eyes will grow higher when the model becomes smaller on the screen. Therefore, more details will be filtered out by HVS, which results in the decrease in the number of pixels that represent the perceptible detail.

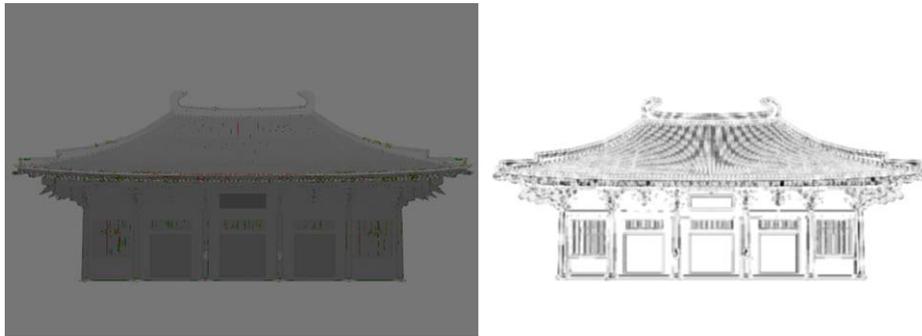
#### 4. Optimal design approach of LODs

Based on the perceptual metric function, the relationship between *viewing distance* and the perception of details is mathematically described. The LODs of the 3D model can thus be analyzed by this function. For visualization applications, optimized LODs should be output sensitive; thus the redundancy of details between detail levels in discrete LOD should be minimized.

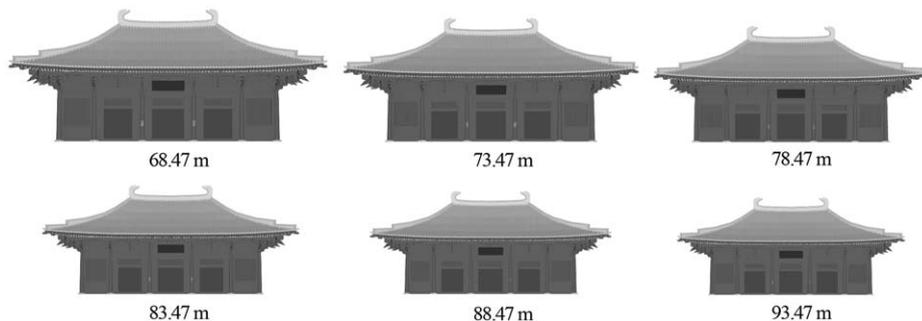
The total quantity of details (TDs) of discrete LODs can be represented by a piecewise function of  $d$

$$TD_{(d)} = \begin{cases} TD_0, & d \in (0, d_1) \\ TD_1, & d \in [d_1, d_2) \\ \dots & \dots \\ TD_{n-1}, & d \in [d_{n-1}, d_n) \\ TD_n, & d \in [d_n, d_f] \end{cases} \tag{9}$$

where  $TD_i$ ,  $\{i \in Z, i \in [0, n]\}$  is the quantity of details in  $LOD_i$ .  $d_i$ ,  $\{i \in Z, i \in [1, n]\}$  is the viewing distance for switching to  $LOD_i$ . As



**Fig. 8.** The VDP result (left) and the SSIM result (right). (For VDP, the green pixels show a 75% probability of difference detection while red pixels indicate a 95% probability that differences will be noticeable. For SSIM, the darker areas show greater visual differences.)



**Fig. 9.** The first six images of sampling.

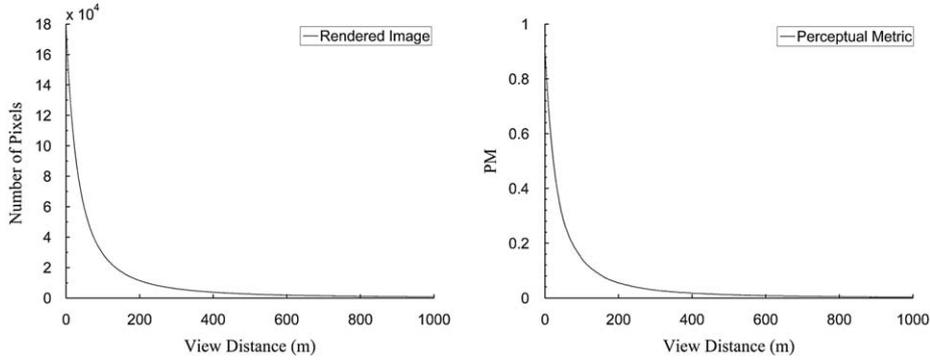


Fig. 10. The statistics of the number of pixels of the object in the rendered image (left); the PM value (right).

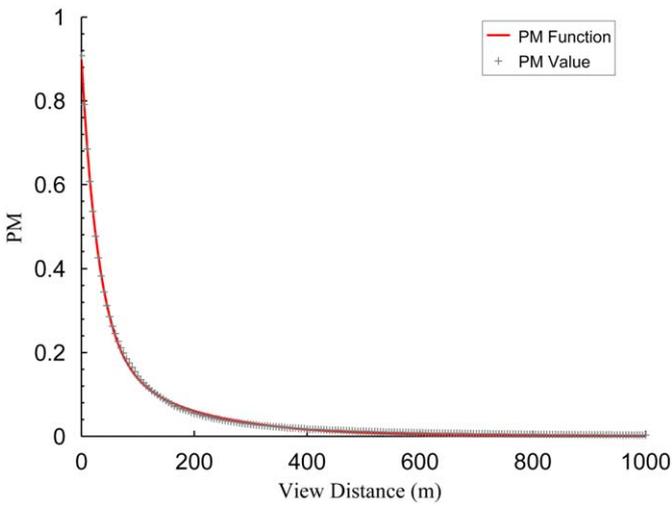


Fig. 11. The perceptual metric function generated by curve fitting.

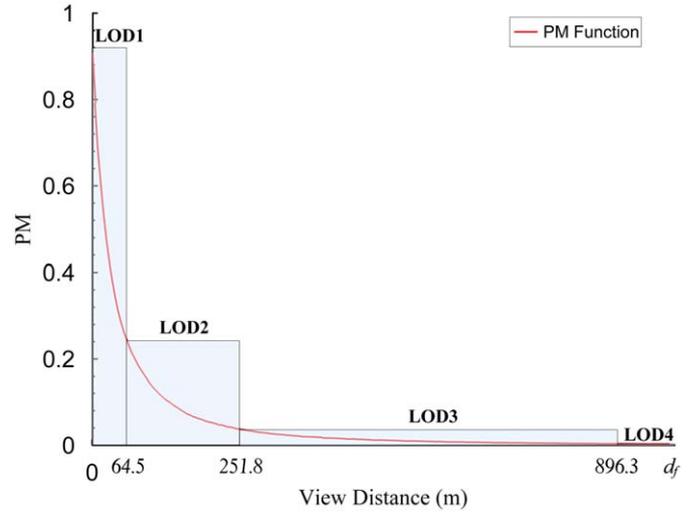


Fig. 12. The result of LODs design.

discussed before,  $d_1$  is the start point of *viewing distance*. Within  $(0, d_1)$ , part of the model occupies most of the screen, and is the dominating object for visual perception. The typical application under this circumstance is the observation of the details or the internal structure, which needs a model of the highest resolution. Also, since the render cost for displaying part of the model is not huge, it allows fast visualization. Therefore, the quantity of details of the original object is defined as  $TD_0$ .

After defining  $LOD_0$ , the remaining work is the defining of  $LOD_1$  to  $LOD_n$ . In order to keep the continuity of perception, the quantity of details in  $LOD_i$  must constantly be more than the minimum needed for representing the same perceptibility as in the original model. As shown in Fig. 5, the red solid line represents the minimum quantity of details that exactly meets the minimum needs of perception, which is termed the *optimal LOD*. The horizontal black line represents  $TD_i$ . At a given viewing distance  $d_i + \Delta d$ , the difference between the quantity of details of discrete LOD and *optimal LOD* is the redundancy detail (RD), as shown in Fig. 5, which is calculated by equation

$$RD(d_i) = TD(d_i) - TD(d_i + \Delta d) \quad (10)$$

where  $TD(d)$  is the quantity of the details of the *optimal LOD* at the viewing distance  $d$ , and  $TD_i = TD(d_i)$ .

Based on the previous definition, perceptual metric can be employed as a ratio that indicates the proportion of the perceptible details in the total details of the object. The redundancy can thus be

calculated by the product of  $TD_0$  and  $PM(d)$

$$RD(d_i + \Delta d) = TD_0 \cdot PM(d_i) - TD_0 \cdot PM(d_i + \Delta d) \quad (11)$$

The gray region in Fig. 5 illustrates the total detail redundancy, which can be calculated by the integral of RD over the distance domain

$$\int_{d_1}^{d_f} RD(d) dx = \sum_{i=1}^{n-1} \left( TD_0 \cdot (PM(d_i) \cdot (d_{i+1} - d_i) - \int_{d_i}^{d_{i+1}} PM(x) dx) \right) \quad (12)$$

Thereby, designing discrete LODs is equivalent to obtaining of a set of viewing distances that minimizes the total RD. It is obvious that when the total redundancy is minimized, the discrete LODs will have the least data size while keeping the same visual fidelity as the *optimal LOD*. Therefore, the discrete LODs are optimally designed.

## 5. Experimental analysis

### 5.1. Experimental description

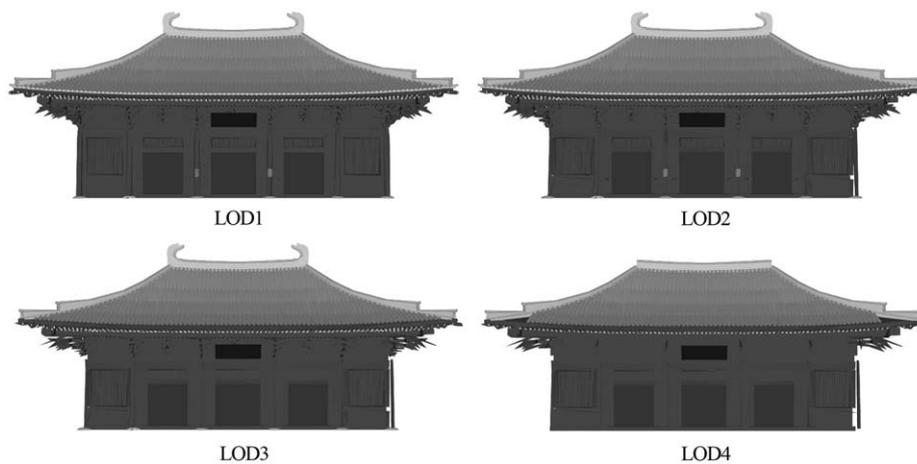
Because of the complex geometry, a complicated building model (the great shrine hall) in Chinese classical style composed of 4398 components and 196,037 triangles was selected as the test model. To reveal fully the details, the model was rendered with the flat shade mode using OpenGL. The scene adopted static light with a white background, and the FOV was 33.33 degrees.

**Table 2**  
The designed-viewing distances of the LODs for the great shrine hall (unit: meter).

Façade	LOD <sub>0</sub>	LOD <sub>1</sub>	LOD <sub>2</sub>	LOD <sub>3</sub>	LOD <sub>4</sub>
1	[0, 56.39)	[56.39, 112.91)	[112.91, 267.14)	[267.14, 836.49)	[836.49, 24500.1]
2	[0, 68.47)	[68.47, 133.28)	[133.28, 321.95)	[321.95, 973.79)	[973.79, 29747.6]
3	[0, 68.48)	[68.48, 132.94)	[132.94, 320.26)	[320.26, 964.79)	[964.79, 29750]
4	[0, 56.13)	[56.13, 109.6)	[109.6, 266.68)	[266.68, 836.07)	[836.07, 24387.2]

**Table 3**  
The comparison plan, in which Equalequal interval I is 100 m, equal interval II is 200 m, and equal interval III is 300 m; exponential interval I is 2 and exponential interval II is 3 (unit: meters).

	LOD <sub>0</sub>	LOD <sub>1</sub>	LOD <sub>2</sub>	LOD <sub>3</sub>	LOD <sub>4</sub>
Our method	[0, 68.48)	[68.48, 132.94)	[132.94, 320.26)	[320.26, 964.79)	[964.79, 29750]
Equal interval I	[0, 68.48)	[68.48, 168.48)	[168.48, 268.48)	[268.48, 368.48)	[368.48, 29750]
Equal interval II	[0, 68.48)	[68.48, 268.48)	[268.48, 468.48)	[468.48, 668.48)	[668.48, 29750]
Equal interval III	[0, 68.48)	[68.48, 368.48)	[368.48, 668.48)	[668.48, 968.48)	[968.48, 29750]
Exponential interval I	[0, 68.48)	[68.48, 136.92)	[136.92, 273.92)	[273.92, 547.84)	[547.84, 29750]
Exponential interval II	[0, 68.48)	[68.48, 205.44)	[205.44, 616.32)	[616.32, 1848.96)	[1848.96, 29750]



**Fig. 13.** The rendered images of the simplified model (Façade 3).

**Table 4**  
The number of triangles—Façade 3.

	LOD <sub>1</sub>	LOD <sub>2</sub>	LOD <sub>3</sub>	LOD <sub>4</sub>
Our method	49,396	47,776	27,917	8801
Equal interval I	49,396	41,226	29,307	26,681
Equal interval II	49,396	29,281	21,063	12,543
Equal interval III	49,396	26,667	12,631	8783
Exponential interval I	49,396	47,780	29,313	15,729
Exponential interval II	49,396	30,299	12,747	4974

The four façades of the building model, extracted manually from the whole model based on the orientation as illustrated in Fig. 6, were analyzed in the experiment in which Façade 3, the front façade, was selected to illustrate the detailed process.

The experiments were carried out using a 19 in TFT-LCD monitor with a resolution of 86.27 dpi (1280 × 1024) and a luminance of 200 (cd/m<sup>2</sup>). The observing distance *v* from the observer's eyes to the monitor was 0.5 m. This is a typical visual environment for desktop applications. The results are also suitable for a visual environment with monitors of lower resolution or observation at greater distances. For monitors of lower resolution, the quantity of the displayed details on the model will be reduced. For observation at a greater distance, on the other hand, the

increasing spatial frequency perceived by human eyes will decrease the quantity of perceptible details. As a result, if a more conservative solution is needed, a higher resolution monitor and a closer observing distance should be chosen. For another commonly used display device, namely a projector, though its resolution is lower than the desktop monitor (from 5 to 30 dpi), the observing distances are usually greater (from 5 to 10 m). Therefore, the spatial frequency perceived by human eyes is similar to that of the desktop monitors.

5.2. Vision simulation

To construct the filters, the CSF is mapped to the frequency intervals. The maximum spatial frequency perceived by humans from the monitor was calculated to be 14.82 cycles/deg (cpd) by Eq. (13)

$$f_{max} = 0.5f_s = 0.5(2v \tan(0.5^\circ)r/0.0254) \tag{13}$$

where *f*<sub>max</sub> is the maximum frequency, which is down-sampled from *f*<sub>s</sub>, 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 [4]

$$A(\alpha) = 2.6(0.0192 + 0.144\alpha)\exp(-0.144\alpha^{1.1}) \tag{14}$$

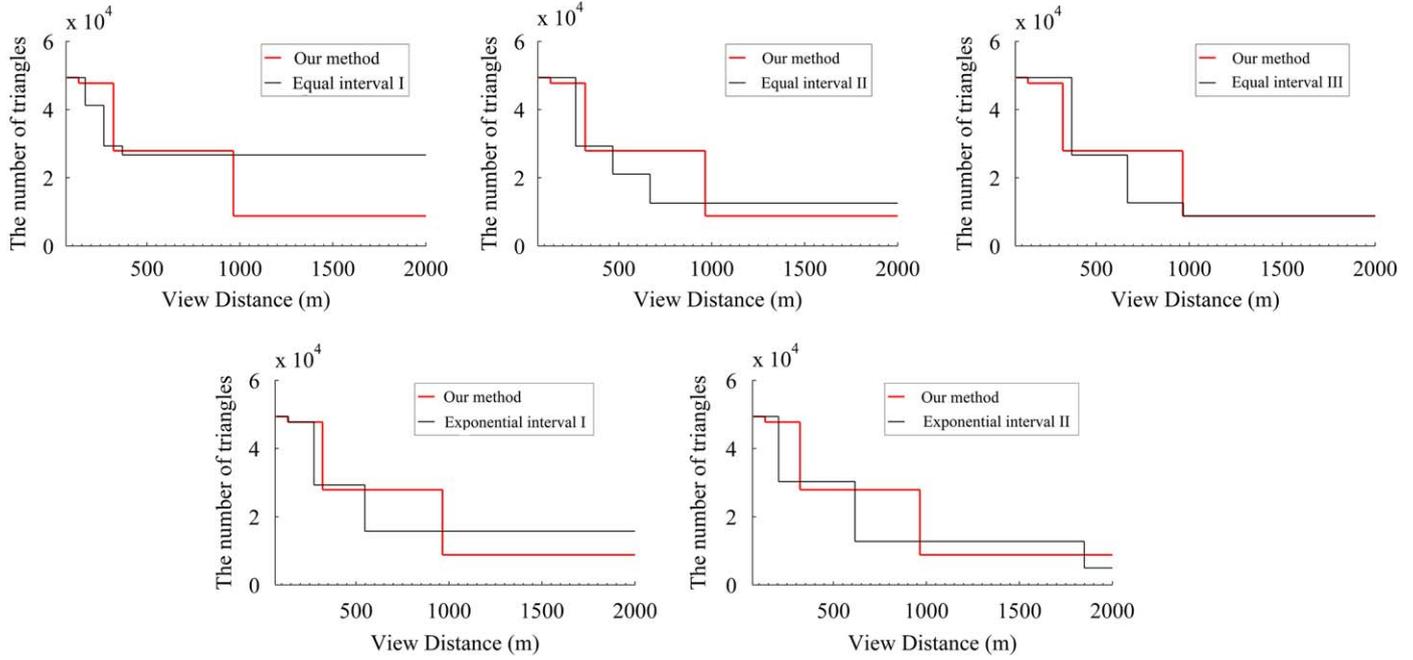


Fig. 14. Comparison of the results of different design plans.

Table 5  
Cost—Façade 3.

	Cost
Our method	308643443.7
Equal interval I	815921235.1
Equal interval II	404717505.4
Equal interval III	309396290.2
Exponential interval I	497820713.8
Exponential interval II	187402764.3

Based on a similar approach to remove the influence of low-frequency dip of the CSF [20], the coefficients with the spatial frequency of less than 6.2 cpd, where the CSF reaches its maximum, were not filtered. This approach can also be employed to decide the proper level of DWT decomposition. When the maximum frequency of level  $i$  of the decomposition was below 6.2 cpd,  $i-1$  levels decomposition was adequate for CSF filtering. As a result, two levels of DWT decomposition were adopted in the experiment because the minimum spatial frequency,  $f_{\max}/4$ , in the second level of decomposition is smaller than 6.2 cpd.

Fig. 7 shows the results. It can be seen that most of the differences between the original image and the processed image either occur at sharp edges, which represent high spatial frequencies such as the silhouette of sharp components, or at gentle changes of pixel values, which represent low contrast such as tiles. In general, there was little visible loss occurring.

The visible difference predictor (VDP) [22] and structural similarity (SSIM) [23] were adopted to evaluate objectively the performance of the vision simulation. The original rendered image and the vision-simulated image were input and the indices are shown in Table 1. Fig. 8 shows the output maps. The results indicate that these two images are visually equivalent, which proves the validity of the vision simulation.

### 5.3. LODs design

The domain of viewing distance was computed using Eqs. (2) and (8);  $d_1$  was equal to 68.48 m and  $d_j$  was equal to 29,750 m. The

sampling range of viewing distance was set to [68.48, 1068.48 m] because the projection size of the model decreases drastically at the start of the domain, and when the distance is greater than 1000 m, the projection of the test model hardly changes any more. The distance interval was set as 5 m to obtain sufficient samples for the curve fitting of  $PM(d)$ , as illustrated in Fig. 9. Five levels of LOD were chosen as an example.

Fig. 10 (left) shows the statistical result of the number of non-background pixels in the rendered images. The PM values, calculated by (5) through the visual simulation process, are shown in Fig. 10 (right). (If there are no special notes,  $d_1$  was set as the start point of the x-axis in the subsequent descriptions.)

According to the shape of the sample data, an exponential function as shown in (15) was adopted to implement the least squares fitting of the PM values

$$f(x) = a \exp(bx) + c \exp(dx) \quad (15)$$

The results were as follows:  $a=0.5588$ ,  $b=-0.09722$ ,  $c=0.1908$  and  $d=-0.02061$ . The R-square was 0.9994 and the RMSE was 0.001761, which indicate the fitting is credible. The  $PM(d)$  curve is shown by the red line in Fig. 11.

Assuming the switch distance for  $LOD_i$  is  $d_i$ , the total RD was then calculated as follows:

$$\begin{aligned} \sum RD = & TD_0 \cdot (PM(0) \cdot (d_2 - 0) - \int_0^{d_2} PM(x) dx + PM(d_2) \cdot (d_3 - d_2) \\ & - \int_{d_2}^{d_3} PM(x) dx + PM(d_3) \cdot (d_4 - d_3) \\ & - \int_{d_3}^{d_4} PM(x) dx + PM(d_4) \cdot (29750 - d_4) - \int_{d_4}^{29750} PM(x) dx) \end{aligned} \quad (16)$$

The simplex method was chosen to find the minimum of the above function and finally  $d_2$ ,  $d_3$ , and  $d_4$  were obtained as 64.5, 251.8, and 896.3 m, respectively, as illustrated in Fig. 12. The output viewing distances of LODs for the four façades are shown in Table 2. The results show that the optimized LODs were not equally partitioned by viewing distance. The higher the detail

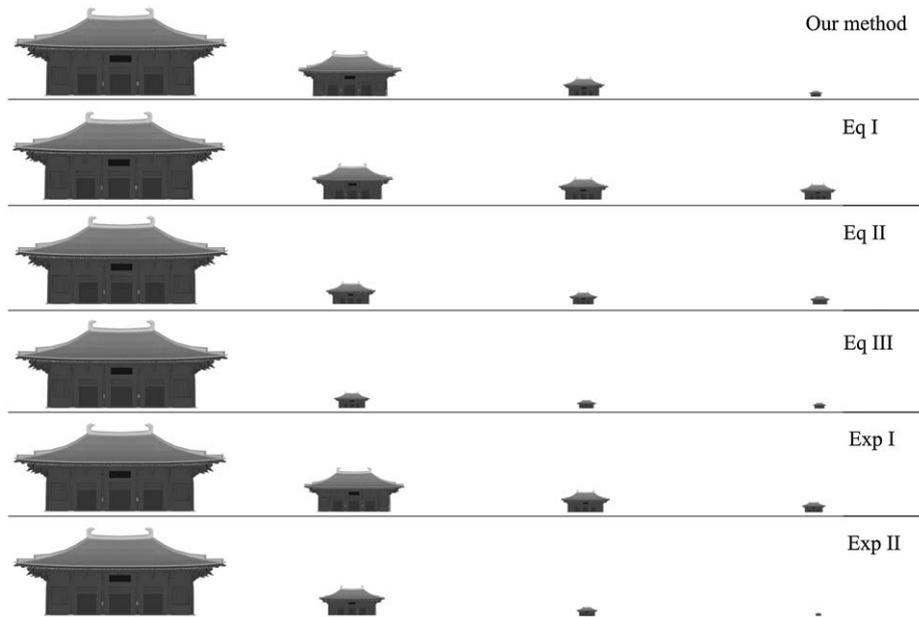


Fig. 15. The comparison results of Façade 3 (Eq represents equal interval and exp represents exponential interval).

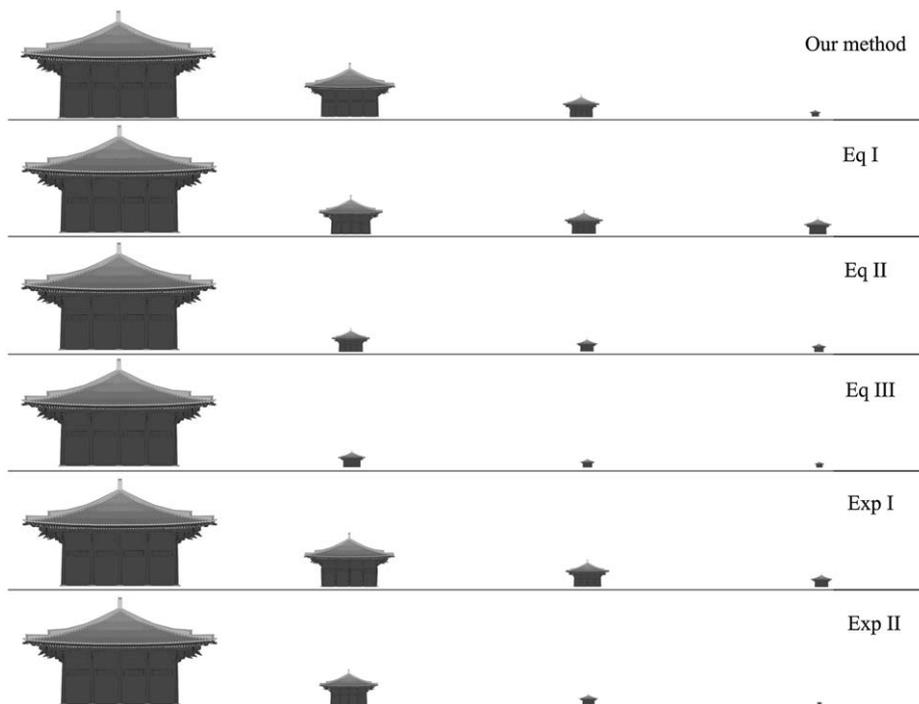


Fig. 16. The comparison results of Façade 1 (Eq represents equal interval and exp represents exponential interval).

resolution, the shorter is the action range for the LOD model. The rendering cost could thus be reduced based on common sense.

#### 5.4. Comparative study

Several groups of typical viewing distances with equal intervals as well as exponential intervals were used as contrasts, as shown in Table 3. The widely used simplification method QEM was selected to simplify the model according to the LOD design [24], in which the half edge collapse operation was adopted in order to keep the perpendicularity and parallelism features in the façades. For generating LODs that fit the observation at a given viewing distance, the screen space error size of one pixel was

chosen to calculate the proper simplification threshold at the viewing distance. Therefore, the LODs that keep the approximate perceptibility could be generated. As illustrated in Fig. 4, if the viewing distance  $d$  is given, the vertex–vertex distance  $\Delta d$ , which is the threshold for simplification, can then be conservatively calculated by

$$\Delta d = 0.0254d / (2r \cdot l) \quad (17)$$

Façade models were then simplified based on the calculated threshold at each viewing distance. Fig. 13 illustrates the rendered images of simplified Façade 3 based on our method. Table 4 shows the statistics of the number of triangles for each LOD model in different planes.

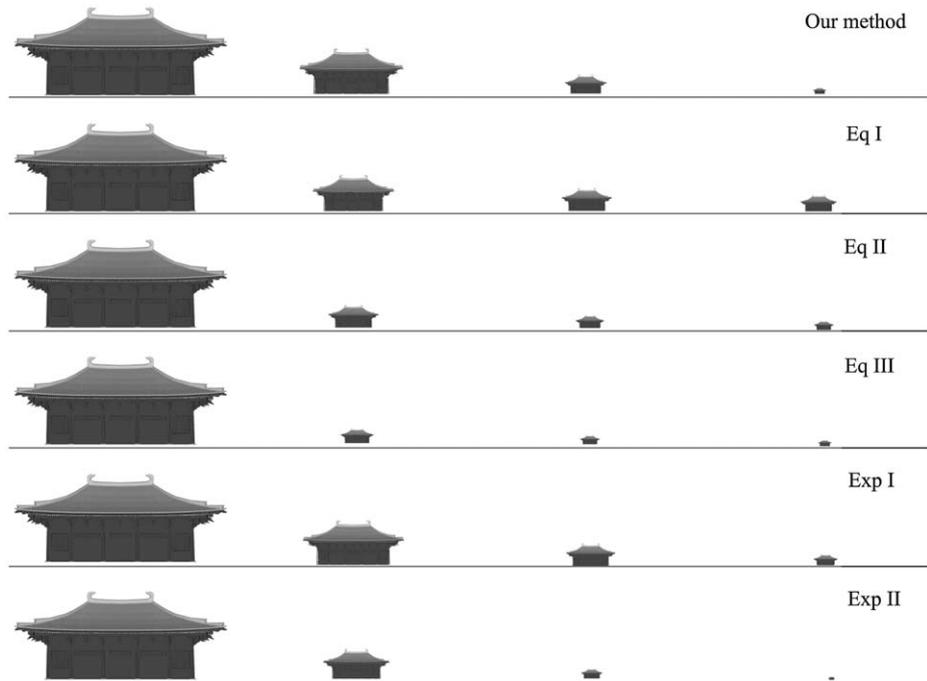


Fig. 17. The comparison results of Façade 2 (Eq represents equal interval and exp represents exponential interval).

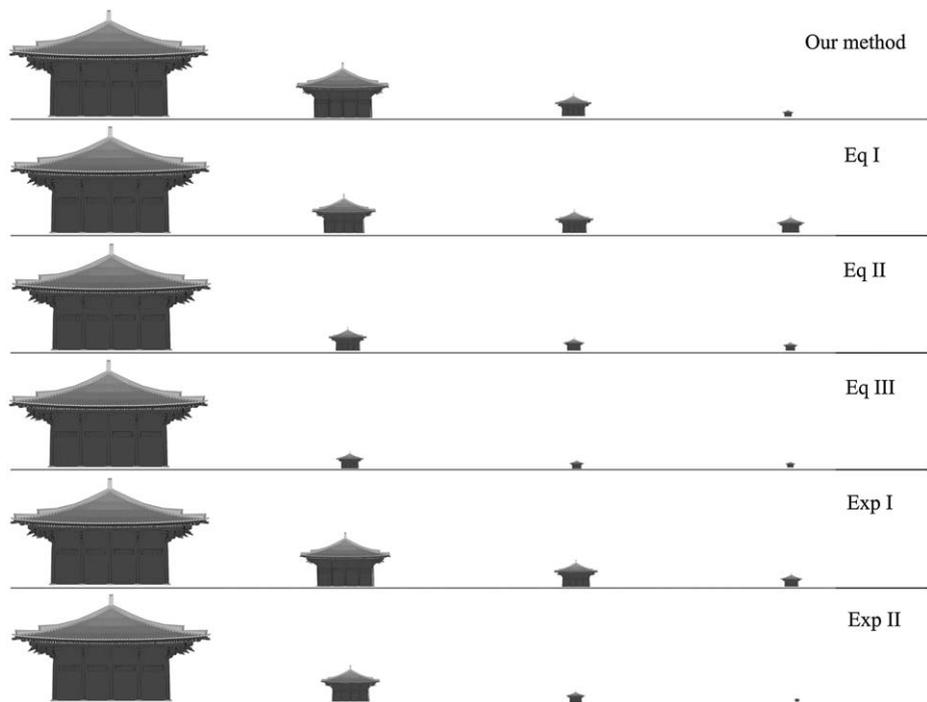


Fig. 18. The comparison results of Façade 4 (Eq represents equal interval and exp represents exponential interval).

Fig. 14 shows the comparison of the results, which illustrates the allocation of triangles along the viewing distance. An approximate approach for estimating the rendering cost of LODs was adopted, in which the cost of a LOD model is represented by the production of its rendering cost and the active viewing distance range. So the total cost of a collection of LODs is calculated by the following equation:

$$Cost = \sum_{i=0}^{i=n} (d_{i+1} - d_i) C_i \quad (18)$$

where  $d_i$  is the switch viewing distance for  $LOD_i$  and  $C_i$  is the cost for  $LOD_i$ , which is represented by the number of triangles. Then, the rendering cost of results was calculated and the results are shown in Table 5.

It is seen that the cost of our method is generally lower than other cases, except for the last one. On the one hand, when the model is under close observation, the complexity should be reduced quickly when model moves further away because the data size of  $LOD_0$  and  $LOD_1$  is often too large to be efficiently rendered. On the other hand, when the model is far from the

**Table 6**  
The data size statistics—Façade 1.

	LOD <sub>1</sub>	LOD <sub>2</sub>	LOD <sub>3</sub>	LOD <sub>4</sub>	Cost
Our method	136843	135045	20884	6766	179485647.4
Equal interval I	136843	30857	20986	19058	468999425.2
Equal interval II	136843	20976	15686	9556	242551492.8
Equal interval III	136843	19052	9554	6498	172621727.6
Exponential interval I	136843	135043	21338	15690	388171221.8
Exponential interval II	136843	29899	15646	3216	104048049.2
The data size statistics—Façade 2					
Our method	149867	148369	27534	8982	314099783.5
Equal interval I	149867	141809	29902	27220	831857446.4
Equal interval II	149867	29886	21246	12754	411074896.5
Equal interval III	149867	27222	12846	8984	315532114.1
Exponential interval I	149867	148357	29904	15912	503399751.7
Exponential interval II	149867	130910	12956	4980	229219396.2
The data size statistics—Façade 4					
Our method	136726	135022	20875	6511	172692418.7
Equal interval I	136726	30820	20973	19011	465706571.8
Equal interval II	136726	20959	15635	9503	240180358.2
Equal interval III	136726	19005	5905	6243	164770970
Exponential interval I	136726	134998	21331	15643	385244178.9
Exponential interval II	136726	29862	15599	3248	104227327.9

camera, the complexity should be reduced more slowly because the current data size of LODs is smaller. The costs of Equal interval III and Exponential interval II are similar or even smaller than the cost of the proposed method. However, the switching between the first two LODs in these results was much slower than with our result, which would lead to a worse performance if the data size of the model was relatively large. In addition, the simple practice of QEM may also bring about errors for this evaluation. Following the principle of human perception, new approaches should be proposed in the future to simplify the model according to the generated difference map, as illustrated in Fig. 7.

Results of the simplified façades are shown in Figs. 15–18. The statistics of number of triangles, as well as the costs, are shown in Table 6.

## 6. Conclusions and remarks

Keeping in mind the difficulty of precise quantitative design of 3D geometric LODs, this paper presents an analytical concept of perceptual metric; related novel algorithms for PM extraction and LODs design are also introduced. According to this kind of quantitative analysis, the redundancy of perceptual detail between LODs could be minimized, which facilitates the automatic generation of discrete LOD models. The result has great potential to play an important role in guiding the simplification of discrete LOD models. At a given viewing distance, the geometries contributing to the imperceptible details in the rendered image can be located for further simplification. Similar approaches have been proposed recently [25,26]. Future research will focus on the PM-driven model simplification for discrete and continuous LOD model generation of complicated textured 3D building models.

## Acknowledgements

Thanks are due to Professor Liqiu Meng, Professor Zhigen Pan, Professor Kun Zhou, Dr. Zhigao Yang and Dr. Dongbo Zhou for their

helpful discussions during the research. Thanks to Dr. Rafal Mantiuk for his help with VDP. This work is supported by the National Basic Research Program of China (973 Program, No. 2010CB731800), the National Natural Science Foundation of China (40871212, 40721001 and 40701144) and the National High Technology Research and Development Program of China (2008AA121600).

## References

- [1] Butler D. Virtual globes: the web-wide world. *Nature* 2006;439:776–8.
- [2] Lin H, Zhu Q. Virtual geographic environments. In: Zlatanova S, Prosperi D, editors. *Large scale 3D data integration—challenges and opportunities*. Boca Raton: Taylor & Francis (CRCpress); 2005.
- [3] Li DR, Zhu YX, Du ZQ, Hong T. Virtual tang-style timber-frame building complex. In: Pan ZG, Cheok AD, Haller M, Lau RW, Saito H, Liang RH, editors. *Advances in virtual reality and tele-existence*. Berlin: Springer-Verlag; 2006. p. 880–8.
- [4] David L, Reddy M, Cohen J, Varshney A, Watson B, Huebner R. *Level of detail for 3D graphics*. San Francisco: Morgan Kaufmann; 2003.
- [5] Pan ZG, Ma XH, Shi JY. The automatic generation algorithm for models at multiple levels of detail in virtual environment. *Journal of Software* 1996;7:526–31.
- [6] Clark JH. Hierarchical geometric models for visible surface algorithms. *Communications of the ACM* 1976;19:547–54.
- [7] Hoppe H. Efficient implementation of progressive meshes. *Computers & Graphics* 1998;22:27–36.
- [8] Erikson C, Manocha D, Baxter III WV. HLODs for faster display of large static and dynamic environments. In: *Proceedings of the 2001 symposium on interactive 3D graphics*. 2001. p. 111–20.
- [9] Cheng I, Shen R, Yang Y, Boulanger P. Perceptual analysis of level-of-detail: the JND approach. In: *Proceedings of the eighth IEEE international symposium on multimedia*. 2006. p. 533–40.
- [10] Watson B, Walker N, Hodges LF. Supra-threshold control of peripheral LOD. *ACM Transactions on Graphics (TOG)* 2004;23(3):750–9.
- [11] Scoggins RK, Moorhead RJ, Machiraju R. Enabling level-of-detail matching for exterior scene synthesis. In: *Proceedings of the conference on visualization*. 2000. p. 171–78.
- [12] Seo J, Kim GJ, Kang KC. Levels of detail (LOD) engineering of VR objects. In: *Proceedings of the ACM symposium on virtual reality software and technology*. 1999. p. 104–10.
- [13] Cohen JD, Manocha D. *Model simplification*. In: Hansen C, Johnson CR, editors. *Visualization Handbook*. Oxford: Elsevier Butterworth-Heinemann; 2005. p. 393–411.
- [14] Ferwerda J. Fundamentals of spatial vision. In: *Applications of visual perception in computer graphics*. Siggraph '98 Course Notes. 1998. p. 1–27.
- [15] Shreiner D, Woo M, Neider J, Davis T. *OpenGL programming guide: the official guide to learning OpenGL*, 4th ed.. Boston: Addison-Wesley Professional; 2004.
- [16] Bradley AP. A wavelet visible difference predictor. *IEEE Transactions on Image Processing* 1999;8:717–30.
- [17] Nadenau B. Integration of human color vision models into high quality image compression. Ph.D Thesis. Lausanne: EPFL; 2000.
- [18] Reddy M. Perceptually modulated level of detail for virtual environments. Ph.D Thesis. Edinburgh: University of Edinburgh; 1997.
- [19] Antonini M, Barlaud M, Mathieu P, Daubechies I. Image coding using wavelet transform. *IEEE Transactions on Image Processing* 1992;1(2):205–20.
- [20] Nadenau B, Reichel J, Kunt M. Wavelet-based color image compression: exploiting the contrast sensitivity function. *IEEE Transactions on Image Processing* 2003;12:58–70.
- [21] Reddy M. Specification and evaluation of level of detail selection criteria. *Virtual Reality* 1998;3(2):132–43.
- [22] Daly S. The visible differences predictor: an algorithm for the assessment of image fidelity. In: Watson A, editor. *Digital images and human vision*. Cambridge, MA: The MIT Press; 1993.
- [23] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 2004;13(4):600–12.
- [24] Garland M, Heckbert PS. Surface simplification using quadric error metrics. In: *Proceedings of SIGGRAPH 1997*. p. 209–16.
- [25] O'Sullivan C, Howlett S, McDonnell R, Morvan Y, O'Connor K. Perceptually adaptive graphics. In: *Eurographics state of the art reports*. 2004. p. 141–64.
- [26] Lee CH, Varshney A, Jacobs DW. Mesh saliency. In: *Proceedings of ACM SIGGRAPH 2005*. p. 659–66.