COMPUTER ANIMATION AND VIRTUAL WORLDS
Comp. Anim. Virtual Worlds 2016; **27**:351–357
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cav.1698

SPECIAL ISSUE PAPER

Efficient view selection by measuring proxy information

Tianhao Gao[1,2], Wencheng Wang[1,2]* and Honglei Han[3]

[1] State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Faculty of Art, Communication University Of China, Beijing, China

ABSTRACT

View selection aims at finding good views that can watch model information as much as possible. However, existing view evaluation methods used to measure model properties are generally complex and very time-consuming. In this paper, we address this challenge by generating succinct proxy information to measure the view quality. The proxy information is generated by principal component analysis of the model to obtain its six principal viewing directions and consists of the representative information computed by the principal viewing directions. We discuss and validate the effectiveness of the proxy information for view evaluation. Thus, measuring the proxy information for view evaluation, the time complexity can be reduced as it is independent of the facet number of the model. This is superior to existing methods by straight measurement of model properties and provides acceleration for view selection. Experimental results show that we can obtain good views as state-of-the-art methods and speed up view selection by at least two orders of magnitudes and achieve more acceleration when the models have more facets. Copyright © 2016 John Wiley & Sons, Ltd.

KEYWORDS
viewing algorithm; view selection; view evaluation; proxy information

***Correspondence**
Wencheng Wang, State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing, China.
E-mail: whn@ios.ac.cn

1. INTRODUCTION

In many applications such as computer animation and virtual reality, it is required to process 3D models efficiently. Here, efficient investigation of 3D models is very important for promoting working efficiency. For example, in identifying or comparing protein models, the characteristic views with minimum and maximum crossings are generated respectively to facilitate the process of identification [1]. Thus, view selection has attracted much attention from scientists in the world. In general, it tries to find good views to watch model information as much as possible, and so reducing the time on investigation. The pipeline is always by sampling many viewpoints around the model and then evaluating their related views to obtain high-quality views that can watch much more information about the model. Here, the critical work is view evaluation. For this, many methods have been proposed by measuring visible model properties in a view, such as the number of visible faces [2], visibility [3], depth maps [4], curvatures [5,6] and interesting surface regions [7]. Although these properties can be used to represent model information in view evaluation, their computation is always time-consuming. Moreover, viewpoints are always sampled densely to avoid missing good views. Thus, evaluating so many views is a very expensive work. The low efficiency of view selection seriously prevents its use in applications. For fast finding good views, it was ever proposed in [5] to build a bounding box of the model by principal component analysis (PCA) and then use the eight corners of the bounding box as good viewpoints. Although these views can have the main shape of the model recognized, they are very possible to miss many important contents of the model, because PCA can only catch the overall low-frequency information. Because of this, these views are not always in high quality.

In this paper, we present a new view evaluation method to fast find good views. This is based on our observation that when the bounding box of the model is produced through PCA, the important contents of the model can be well viewed from at least one of the six principal viewing directions (called PVs in short), which refer to the six viewing directions aligned with the normal of the six faces of the bounding box. Thus, by these PVs, we can generate proxy information to represent important contents about the model and then measure the proxy information to evaluate views. In our treatment, we adopt two geometric properties to generate our proxy information, the projection area and the number of visible facets, which have been shown to very effective for view evaluation in existing works. In implementation, the projection areas of the model and the numbers of visible facets are computed by these PVs respectively. Moreover, to have the proxy information effectively cover the visible contents from neighbouring views of the PVs, which will improve the effectiveness of the proxy information to represent model information in view evaluation, we investigate the projection areas by these views and count all the visible facets to these views. To save computation for acceleration, we even adopt approximation computation in generating the proxy information. This will not lower the measurement quality for view evaluation, to be discussed in Section 3. Therefore, we can obtain good views as existing methods to watch many important contents of the model. Clearly, view evaluation by measuring the proxy information is independent of the model resolution. This is helpful to speed up view evaluation, especially in treating large models. We made tests on several models, and experimental results show that our method can obtain quality views as the state-of-the-art methods and speed up view selection by at least two orders of magnitudes, obtaining more acceleration when the models have more facets.

In the remainder of the paper, we will discuss related works in Section 2 and then introduce our new view evaluation method by measuring proxy information in Section 3. Afterwards, results are discussed in comparison with existing methods in Section 4, and a summary is drawn in Section 5 in the final.

2. RELATED WORKS

Till now, view evaluation methods have been proposed in a large quantity. They try to develop view descriptors to well measure the visible model information in a view. As 3D models are generally investigated through their geometric contents, many view descriptors are based on some geometric features. Considering that a model can be better watched when its projection on the screen is larger, the method in [2] measures the projection areas in sampled views to find the view with the largest projection area. As degenerated faces may reflect the change of model contents, the method in [8] performs view evaluation by the degenerated faces, expecting the view with many more model contents. Besides these, there are many geometric features taken into account for view evaluation, including depth maps [4,9], silhouette curvature [10], mean curvature [11] and mesh saliency (MS) [12]. As measuring local geometric features may be much interfaced by model details, which will prevent evaluation of global information in the model, some methods employed the information theory to design view descriptors, such as viewpoint entropy (VE) [13], relative entropy [14], viewpoint mutual information [15] and viewpoint saliency Kullback–Leibler (vSKL) distance [16]. Of course, their computation is still based on measuring some local geometric properties, including counting visible facets and computing mesh saliency.

In some cases, measuring only geometric information is not sufficient to obtain good views, according to the theory of object recognition [17]. Thus, many view descriptors were proposed by measuring some semantic meaning. Here, the important work is to find the semantic visual cues for object recognition, such as distinctive regions [18], upright direction [19], meaningful segments [20], Schelling point [21], interesting regions [7] and even the ability of imagination [22]. As human beings generally have similar viewing preferences for similar models, some methods proposed to employ machine learning to obtain the correspondence between geometric features and semantic meaning, and so assign semantic meaning to the model for view evaluation [23–25]. For example, the support vector machine is adopted for adjusting term weights [10,26,27]. To obtain enough samples to guarantee the quality of the obtained results from data training, it is suggested to collect the images for 3D models in the Internet [24]. Although the methods by measuring semantic meaning can obtain very good results, semantic computation is much trouble and expensive, and the results may suffer from the limited samples for data training.

As discussed earlier, existing view evaluation methods are generally very time-consuming as they need complicated computation to derive important information about the model. Different from them, our method takes simple computation to obtain proxy information about the model, and the proxy information consists of only six pairs of numbers about the projection area and visible facets. Thus, our view evaluation method by measuring proxy information reduces the time complexity. As for the results, because our proxy information can well represent model information for view evaluation, we can obtain good views as the state-of-the-art methods, as shown in Section 4. As another benefit, our method is easy to implement, helpful to spread its use in applications.

3. VIEW EVALUATION BY MEASURING PROXY INFORMATION

3.1. Proxy information

As we know, PCA is effective to obtain low-frequency information of the model. Thus, we still use PCA to build the bounding box of the model to generate our proxy information for view evaluation. As discussed in Section 1, the corners of the bounding box are not suitable candidates for good views because they may miss many important contents of the model. Fortunately, we observe that the important contents of the model can be well watched from at least one of the six principal viewing directions, as illustrated in Figure 1.

Therefore, we generate proxy information by the six faces of the bounding box, respectively, to represent the

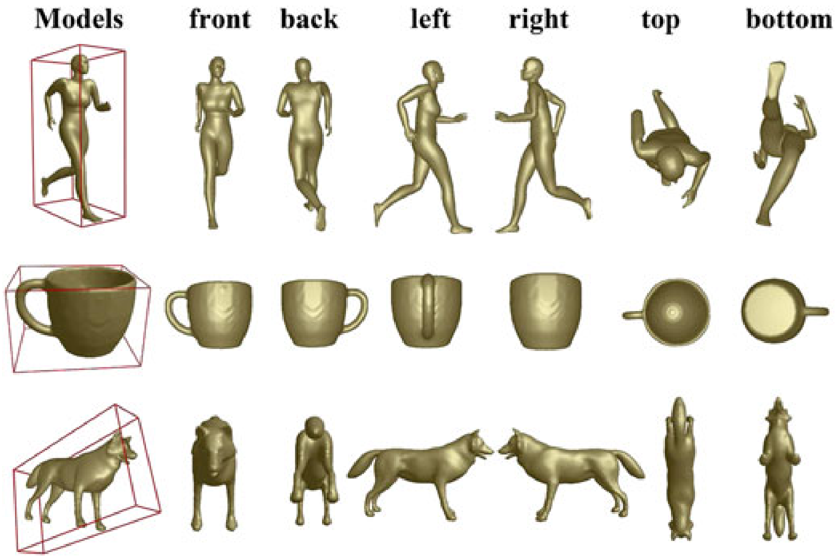

Figure 1. Important contents of the model can be well watched from at least one of the six principal viewing directions.

Table I. Statistics about approximation differences for the projection areas of PVs.

Model	Human (%)	Cup (%)	Dog (%)
Front	< 5.3	< 6.5	< 41.2
Back	< 6.4	< 6.4	< 17.4
Left	< 3.5	< 8.7	< 4.5
Right	< 5.1	< 8.0	< 4.1
Top	< 30.2	< 6.3	< 4.6
Bottom	< 9.5	< 12.7	< 1.7

For the nearby views to a PV, their related percentages are computed by $|s' - s|/s$, where s is the projection area by the PV and s' is the projection area by a nearby view. The largest one among these percentages is selected as their corresponding percentage for the PV in the table.

viewing effectiveness of important contents from the six principal viewing directions. Afterwards, the proxy information is taken to evaluate views. As studied in existing methods, the projection area and the number of visible facets are very effective properties for measuring viewing effectiveness. Thus, we compute the proxy information along the ith principal viewing direction, Inf_i, in the following equation,

$$Inf_i = (1 - \lambda) A_i + \lambda N_i \tag{1}$$

where A_i is the ratio of the orthogonal projection area of the model along the ith PV against the surface area of the model, N_i is the ratio of the number of the visible facets from the ith PV against the number of all the facets of the model and $\lambda \in (0, 1.0)$ is the weight to adjust the effect of A_i and N_i. In our tests, we generally set $\lambda = 0.5$ and can always obtain very good results.

As we know, nearby views may watch very different model contents. To improve the effectiveness of the proxy information to represent model information in view evaluation, we investigate the projection areas of the model by the views near a PV and count all the visible facets by these views to generate the proxy information related to this PV. Thus, in evaluating a view different from the PVs, its visible facets will be more possibly taken into account to improve the evaluation quality. Considering visibility computation is very expensive, we decide to adopt approximation computation to obtain the proxy information.

We investigate the differences between the projection area along a PV and the projection area by its nearby views and find the differences between them are not large in general, as illustrated in Table I for some models. Thus, it is not necessary to compute the projection areas by the nearby views, and it is enough to use the projection area by the ith PV in computing A_i. As for counting the visible facets for nearby views of the ith PV, we found that the count can be approximated by the number of the facets whose normal differ from the ith PV is not much. This is because the difference between them takes a low percentage of the count, as illustrated in Table II, where we set the angle threshold

Table II. Statistics about approximation differences for counting visible facets of PVs.

Model	Human (%)	Cup (%)	Dog (%)
Front	< 4.1	< 14.4	< 44.8
Back	< 6.3	< 13.9	< 23.8
Left	< 10.1	< 9.6	< 5.4
Right	< 6.5	< 20.5	< 5.9
Top	< 23.7	< 13.4	< 3.9
Bottom	< 40	< 11.9	< 6.8

The percentage for a PV in the table is computed by $|n' - n|/n$, where n is the number of the visible faces for the nearby views of the PV including the PV, and n' is the number of the faces whose normals differ from the PV by less than 60°.

between the facet normal and the ith PV be 60° to collect facets. In this way, N_i can be computed quickly without visibility computation.

From the statistics in Tables I and II, it is known that the differences for the top view of the human model are a little large, and so do for the front view of the dog model. With a simple thought, this may cause problems in applying our method. However, for a view different from the PVs, it can actually watch more than one faces of the bounding box, so that the proxy information related to these faces will all contribute in evaluating this view. Thus, the view can still be effectively evaluated. This will be attested by the results in Section 4.

3.2. View descriptor

We design our new view descriptor by measuring the generated proxy information. Because the proxy information is computed by the respective principal viewing directions and the viewing effectiveness of a same content may be different from different viewing directions, we use weighted computation to suitably represent the influence of proxy information in evaluating a view. In other words, when the view under investigation differs from the ith PV much more, the proxy information of the ith PV will take less effect in evaluating this view. For example, when the ith PV is back facing from the view, its proxy information should not be considered in evaluating this view. For computation convenience, we use the cosine value of the angle between the view and the ith PV as the weight to enhance the effect of the proxy information of the ith PV in view evaluation. Therefore, our descriptor $E(v)$ to evaluate the view v is given in the following,

$$E\left(\vec{v}\right) = \sum_{i=1}^{6} w\left(\vec{v}, \vec{n}_i\right) \cdot Inf_i \tag{2}$$

where

$$w = \begin{cases} \vec{n}_i \cdot \vec{v}, & if\ \vec{n}_i \cdot \vec{v} \geq 0 \\ 0, & if\ \vec{n}_i \cdot \vec{v} < 0 \end{cases} \tag{3}$$

Clearly, using our descriptor, the visible important contents for a view can be effectively taken into account in view evaluation. Benefited from this, we avoid the shortcoming of using the corners of the bounding box to recognize only the main shape of the model and so obtain good views to well watch many important contents of the model. As a result, we can obtain good views as with existing methods, as shown in the results in Section 4.

From the aforementioned discussion, our proxy information is easy to generate and our view descriptor only needs to treat the proxy information from six principal viewing directions. Therefore, we can speed up view selection considerably.

3.3. Algorithm

From the aforementioned discussion, our algorithm for view selection is listed in Algorithm 1.

Algorithm 1 Best view search by measuring proxy information

Require: the 3D object, M
  candidate viewpoint set, V
Ensure: the best view $v \in V$
  Perform PCA of the model, obtain six PVs;
  Build the bounding box;
  for each PV_i, $i = 1, 2, ..., 6$ **do**
    calculate A_i, N_i;
  end for
  for each $v_j \in V$ **do**
    for each PV_i, $i = 1, 2, ..., 6$ **do**
      calculate $w\left(\vec{v}_j, \vec{n}_i\right)$ in Eq2;
    end for
    calculate $E\left(\vec{v}_j\right)$ in Eq2;
    update current max E, and viewpoint id v_{max}
  end for
  return the Best View, v_{max}

4. RESULTS AND DISCUSSION

We implemented our method in MS Visual Studio2012 and collected the statistics on a Lenovo personal computer equipped with an Intel i7-2600 CPU, 4 GB RAM and an nVidia GT 420 GPU. The tested models are from the Princeton Benchmark of 3D Shapes [28]. For comparison, we also implemented some existing methods, including VE [13], MS [12], Linear5 [26] and vSKL [16]. In implementing these methods, the semantic computation for Linear 5 was processed by user interaction as suggested in [26], and CUDA was employed to efficiently compute mesh saliency for the MS and vSKL methods.

4.1. Effectiveness

To compare the quality of the obtained good views with these methods, we conducted a user study. For this, 10 volunteers were invited, who are graduate students with their major in computer graphics, and 85 models were used, which are evenly from the 17 categories of the Princeton Benchmark of 3D Shapes. For every tested model, five good views with these methods were provided, and the volunteers were asked to rank these views respectively from 5 to 1, corresponding to from 'the best' to 'the worst'. Afterwards, the ranked scores for the models of these categories are averaged respectively as the rank scores for these categories, which are illustrated in Figure 2. In Figure 3, it is listed good views of a few models with these methods as an illustration, because of the limit of paper length. From the statistics, it is known that our method could obtain very

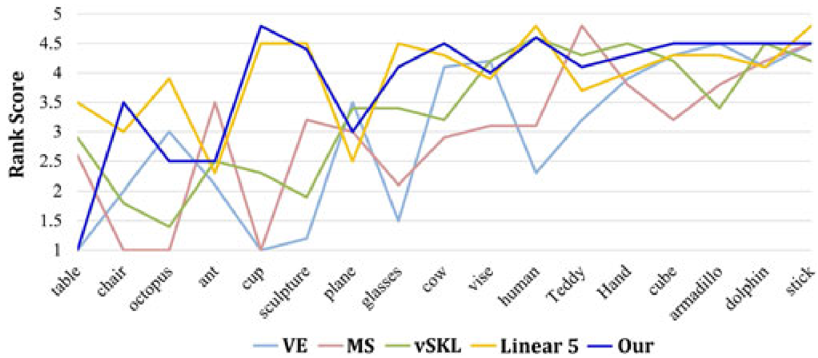

Figure 2. Comparison of the good views obtained with the five methods in comparison. Here, for the five methods, it is plotted their averaged rank scores for their good views of the models from the 17 categories.

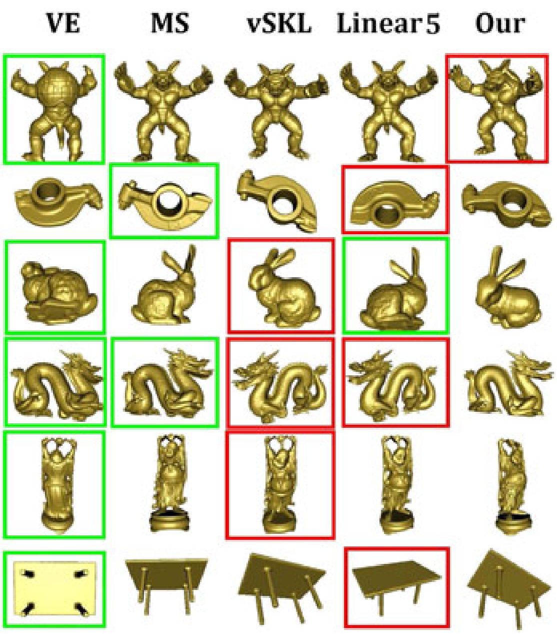

Figure 3. The good views of some tested models obtained with the methods in comparison. Here, the views with the highest and lowest averaged rank scores are marked in red and green boxes, respectively.

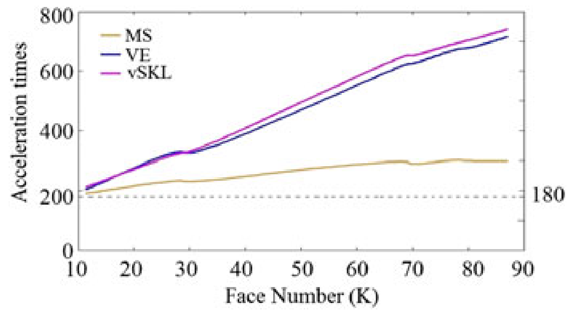

Figure 4. Acceleration times for our method against view entropy (VE), mesh saliency (MS) and viewpoint saliency Kullback–Leibler distance (vSKL) to find best views for the tested 85 models, respectively, arranged by the facet numbers of these models.

good views except for the table models, although they are not always ranked the best. For the table models, human beings prefer to watch them from the above, while our obtained views are from the bottom. Although our views can watch more geometric contents, they are different from human beings' preference. This is related to semantic measurement, not considered in our current implementation. In sum, our view evaluation method by measuring proxy information is very effective for view selection, especially when the geometric contents are much cared for.

4.2. Efficiency

We compared the view selection efficiency between our method and VE, MS and vSKL because they are known to be fast among the existing methods. As for other methods such as Linear5 and the methods via data training, they are generally very time-consuming, and so not considered in this comparison. In Figure 4, it is listed the acceleration times for our method against VE, MS and vSKL methods respectively in treating the 85 tested models, which are arranged by the facet numbers of these models from small to large gradually. Clearly, our method can be faster than these methods by at least 180 times and achieve more acceleration when the models have more facets. This is benefited from our proxy information that is very simple. When it is used for evaluating views, the time complexity is independent of the facet numbers of the models. As for the other methods, their time complexities are much dependent on the facet numbers of the models.

We made another test to know the efficiency of our method to treat the models with various facet numbers. As shown in Figure 5, the time cost increases linearly with the facet numbers. This is because in generating the proxy information, all the facets should be processed several times. After that, the time cost for view evaluation is independent of the facet numbers. From the statistics in Figure 5, it is known that we can take less than 40 milliseconds to treat the models with fewer than 140K facets and suggest the best view within 0.5 seconds for a model

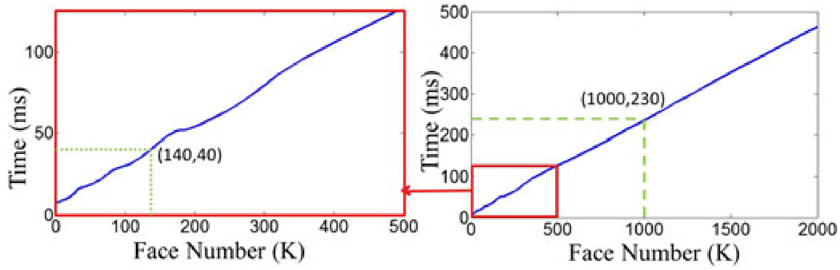

Figure 5. The time cost for our method to treat the models with various facet numbers.

with two million facets. As a result, our method can be employed for interactive model investigation in many applications, where the used models seldom have a very large number of facets.

Limitation In our method, the proxy information for a model is produced by its visible contents from outside. When the model has its important contents hidden in its concave regions very much, our method would fail in finding good views for it. Fortunately, such a case seldom occurs in applications.

5. SUMMARY

Existing view selection methods always evaluate views by measuring model properties, whose time complexity is much dependent on the facet numbers of the models. This is expensive, preventing the view selection efficiency. In this paper, we propose to produce a kind of proxy information to represent model information for view evaluation. The proxy information is generated by PCA of the model to obtain the principal viewing directions and consists of succinct representation information computed by the principal viewing directions. We discuss and validate its effectiveness for view evaluation. As measuring the proxy information is independent of the facet numbers of the model, the time complexity for view evaluation can be reduced much. Thus, we can considerably speed up view selection to obtain good views and obtain more acceleration over existing methods when the models have more facets.

In our current implementation, we only consider the geometric information of the model. This is useful in many applications, because watching geometric information of the model is always required for model investigation. However, when the semantic meaning for model investigation is not strongly related with the geometric information, our implementation may miss the good views as expected. In this paper, we have shown that using proxy information is an efficient way to improve view selection. Therefore, an interesting future issue is to study how to generate proxy information with semantic meaning considered, and so improving view selection in various applications.

ACKNOWLEDGEMENTS

This work is supported in part by the National Natural Science Foundation of China (61379087), the Knowledge Innovation Program of the Chinese Academy of Sciences, and the European Unions Seventh Framework Programme (FP7/2007-2013) under grant agreement (612627).

REFERENCES

1. Au C, Cai Y, Zheng J, Woo T. Orienting a protein model by crossing number to generate the characteristic views for identification. *Computer Modeling in Engineering and Sciences* 2010; **68**(3): 221–237.
2. Plemenos D, Benayada M. Intelligent display in scene modeling. new techniques to automatically compute good views. In *Proceedings of International Conference on Computer Graphics & Vision*: St Petersbourg (Russia), 1–5 of July 1996.
3. González F, Sbert M, Feixas M. Viewpoint-based ambient occlusion. *IEEE Computer Graphics and Applications* 2008; **28**(2): 44–51.
4. Vázquez P-P. Automatic view selection through depth-based view stability analysis. *The Visual Computer* 2009; **25**(5-7): 441–449.
5. Polonsky O, Patané G, Biasotti S, Gotsman C, Spagnuolo M. What's in an image? *The Visual Computer* 2005; **21**(8-10): 840–847.
6. Biedl TC, Hasan M, López-Ortiz A. Efficient view point selection for silhouettes of convex polyhedra. *Computational Geometry* 2011; **44**(8): 399–408.
7. Leifman G, Shtrom E, Tal A. Surface regions of interest for viewpoint selection. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. IEEE: Rhode Island, USA, 2012; 414–421.
8. Kamada T, Kawai S. A simple method for computing general position in displaying three-dimensional objects. *Computer Vision, Graphics, and Image Processing* 1988; **41**(1): 43–56.
9. Blanz V, Tarr M, Bülthoff HH, Vetter T. What object attributes determine canonical views? *Perception-London* 1999; **28**(5): 575–600.

10. Vieira T, Bordignon A, Peixoto A, Tavares G, Lopes H, Velho L, Lewiner T. Learning good views through intelligent galleries. *Computer Graphics Forum* 2009; **28**(2): 717–726.
11. Meyer M, Desbrun M, Schröder P, Barr AH. Discrete differential-geometry operators for triangulated 2-manifolds. In *Proceedings of Visualization and Mathematics III*. Springer: Berlin, Germany, 2003; 35–57.
12. Lee CH, Varshney A, Jacobs DW. Mesh saliency. *ACM Transactions on Graphics* 2005; **24**(3): 659–666.
13. Vázquez P-P, Feixas M, Sbert M, Heidrich W. Viewpoint selection using viewpoint entropy. In *Proceedings of Vision, Modeling and Visualization Conference*. Aka GmbH: Stuttgart, Germany, 2001; 273–280.
14. Sbert M, Plemenos D, Feixas M, González F. Viewpoint quality: measures and applications. In *Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*. Eurographics Association: Girona, Spain, 2005; 185–192.
15. Feixas M, Sbert M, González F. A unified information-theoretic framework for viewpoint selection and mesh saliency. *ACM Transactions on Applied Perception* 2009; **6**(1): 1:1–1:23.
16. Serin E, Sumengen S, Balcisoy S. Representational image generation for 3D objects. *The Visual Computer* 2013; **29**(6-8): 675–684.
17. Hoffman DD, Richards WA. Parts of recognition. *Cognition* 1984; **18**(1): 65–96.
18. Shilane P, Funkhouser T. Distinctive regions of 3D surfaces. *ACM Transactions on Graphics* 2007; **26**(2): 7:1–7:15.
19. Fu H, Cohen-Or D, Dror G, Sheffer A. Upright orientation of man-made objects. *ACM Transactions on Graphics* 2008; **27**(3): 42:1–42:8.
20. Mortara M, Spagnuolo M. Semantics-driven best view of 3D shapes. *Computers & Graphics* 2009; **33**(3): 280–290.
21. Chen X, Saparov A, Pang B, Funkhouser T. Schelling points on 3D surface meshes. *ACM Transactions on Graphics* 2012; **31**(4): 29:1–29:12.
22. Sokolov D, Plemenos D, Herder J. High level methods for scene exploration. *Journal of Virtual Reality and Broadcasting* 2006; **3**(12): 1860–2037.
23. Laga H. Semantics-driven approach for automatic selection of best views of 3D shapes. In *Proceedings of the 3rd Eurographics Conference on 3d Object Retrieval*. Eurographics Association: Norrk, Sweden, 2010; 15–22.
24. Liu H, Zhang L, Huang H. Web-image driven best views of 3D shapes. *The Visual Computer* 2012; **28**(3): 279–287.
25. Varhaníková I, Hučko M, Černeková Z, Kučerova J. Study of the best view selection based on a human observer. In *Proceedings of the 2013 XXIV International Symposium on Information, Communication and Automation Technologies*. IEEE: Sarajevo, Bosnia and Herzegovina, 2013; 1–7.
26. Secord A, Lu J, Finkelstein A, Singh M, Nealen A. Perceptual models of viewpoint preference. *ACM Transactions on Graphics* 2011; **30**(5): 109:1–109:12.
27. Lienhard S, Specht M, Neubert B, Pauly M, Müller P. Thumbnail galleries for procedural models. *Computer Graphics Forum* 2014; **33**(2): 361–370.
28. Chen X, Golovinskiy A, Funkhouser T. A benchmark for 3D mesh segmentation. *ACM Transactions on Graphics* 2009; **28**(3): 73:1–73:12.

AUTHORS' BIOGRAPHIES

Tianhao Gao received his BS degree from Shandong University at Weihai in 2009. He is currently a PhD candidate at the State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences. His research interests include human perception, shape analysis and computer graphics.

Wencheng Wang is a professor of the State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, where he leads a research group on Computer Graphics and Image Processing. He received his PhD degree in software from the Institute of Software, Chinese Academy of Sciences in 1998. His research interests include computer graphics, virtual reality, image editing and visualization. He has published papers on many high-level journals and conferences; please refer to http://lcs.ios.ac.cn/~whn. He is a member of the IEEE and the ACM.

Honglei Han joined the Communication University of China after receiving his MS degree in Computer Science. He received his PhD degree with the same major from University of Chinese Academy of Sciences. He is teaching various undergraduate and graduate courses in computer graphics, programming, and computer game development at the School of Animation and Digital Arts, and these are also his research interests.