

# Supplementary Material for “Mesh Saliency: An Independent Perceptual Measure or A Derivative of Image Saliency?”

Ran Song<sup>1,2</sup>

Wei Zhang<sup>1,2,\*</sup>

Yitian Zhao<sup>3</sup>

Yonghuai Liu<sup>4</sup>

Paul L. Rosin<sup>5</sup>

<sup>1</sup> School of Control Science and Engineering, Shandong University, China

<sup>2</sup> Institute of Brain and Brain-Inspired Science, Shandong University, China

<sup>3</sup> Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, China

<sup>4</sup> Department of Computer Science, Edge Hill University, UK

<sup>5</sup> School of Computer Science and Informatics, Cardiff University, UK

{ransong, davidzhang}@sdu.edu.cn, yitian.zhao@nimte.ac.cn, liuyo@edgehill.ac.uk, RosinPL@cardiff.ac.uk

## 1. Overview of the supplementary material

This supplementary material provides additional experimental results for the main paper. The experiments were carried out on the 3D visual attention dataset (3DVA) [3] and the Schelling dataset [1], respectively.

## 2. Additional results on the 3DVA dataset

We first compare our MIMO-GAN with 6 mesh saliency methods, namely Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5] and CfS-CNN [8] on the 3DVA dataset. The 3DVA dataset contains a variety of 32 3D objects and provides ground-truth human fixation maps for 3 designated views of each object, which thus enables a comparison of view-dependent mesh saliency over 96 views. Tables 1 and 2 of the main paper already show the overall performances of the mesh saliency methods on the 3DVA dataset. In this supplementary material, Figs. 1-11 not only show the visualised saliency maps of all 96 views produced by each competing method, but also report the corresponding linear correlation coefficients (LCC) as the measure for a quantitative evaluation. It can be seen from these figures that in terms of LCC, the proposed method accomplishes the best results for 65 views of 30 objects over a total number of 96 views of 32 objects.

## 3. Additional results on the Schelling dataset

We also provide extra results based on the Schelling dataset which collected 3D interest points by asking hu-

man subjects to “select points on the surface of a 3D object likely to be selected by other people”. Unlike the 3DVA dataset for evaluating mesh saliency methods in a view-dependent manner, the Schelling dataset facilitates an evaluation of view-independent mesh saliency. We generate the ground-truth saliency maps from the scattered interest points using the scheme depicted in Section 4.3 of the main paper. Then, in this supplementary material, we quantitatively evaluate the performance per object category of our MIMO-GAN based on LCC as the Schelling dataset which contains 400 objects belonging to 20 categories is much larger than the 3DVA dataset. According to Figs. 12-17, our method achieves LCC higher than 0.4 for most object categories when predicting view-independent mesh saliency.

## References

- [1] Xiaobai Chen, Abulhair Saparov, Bill Pang, and Thomas Funkhouser. Schelling points on 3d surface meshes. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):29, 2012. 1, 13, 14
- [2] Tingbo Hou and Hong Qin. Admissible diffusion wavelets and their applications in space-frequency processing. *IEEE Trans. Vis. Comput. Graph.*, 19(1):3–15, 2013. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
- [3] Guillaume Lavoué, Frédéric Cordier, Hyewon Seo, and Mohamed-Chaker Larabi. Visual attention for rendered 3d shapes. *Comput. Graph. Forum (Proc. Eurographics)*, pages 414–421, 2018. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
- [4] Chang Ha Lee, Amitabh Varshney, and David W Jacobs. Mesh saliency. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 24(3):659–666, 2005. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
- [5] George Leifman, Elizabeth Shtrom, and Ayellet Tal. Surface regions of interest for viewpoint selection. *IEEE Trans. Pat-*

\*Corresponding author

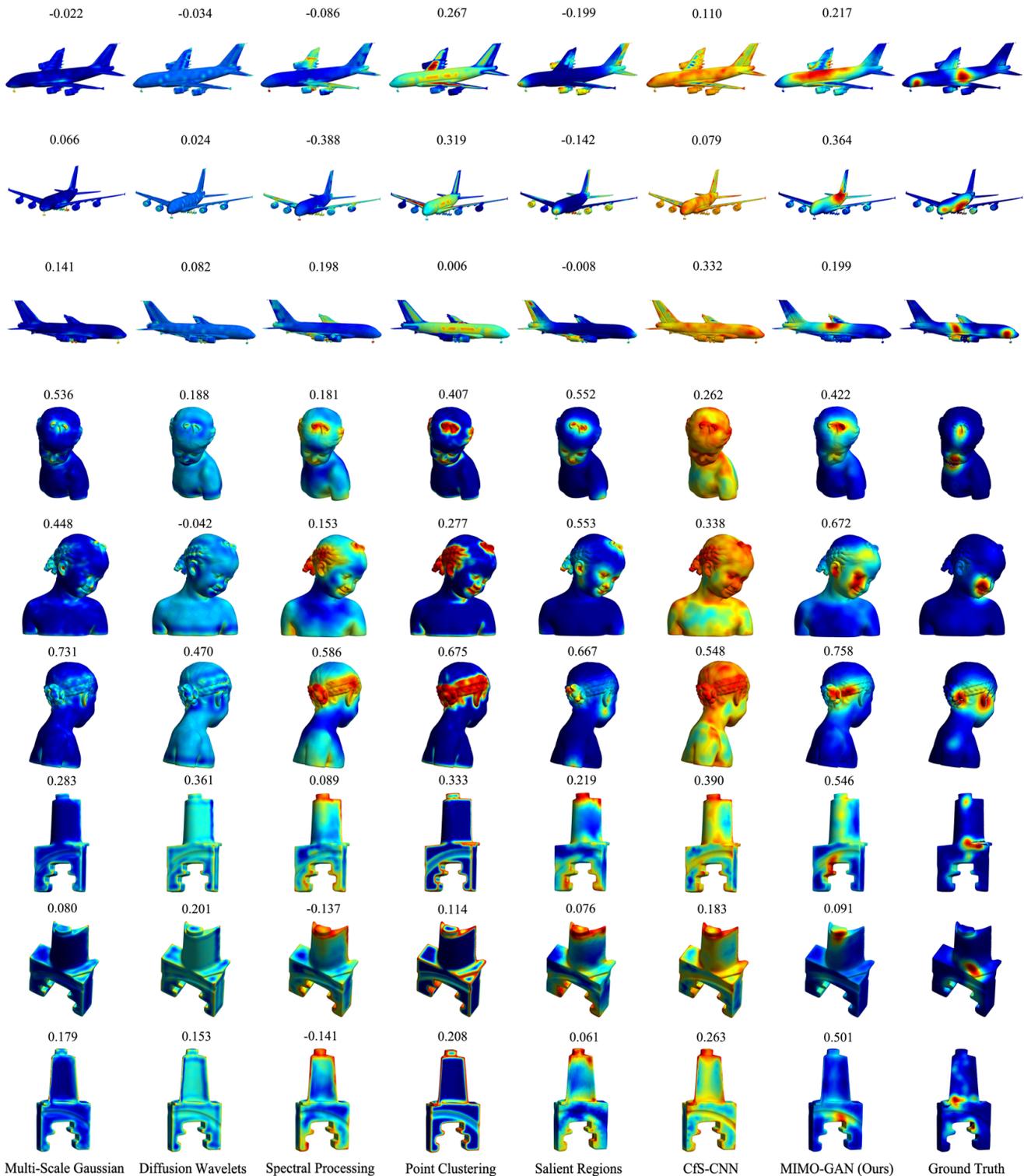


Figure 1. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

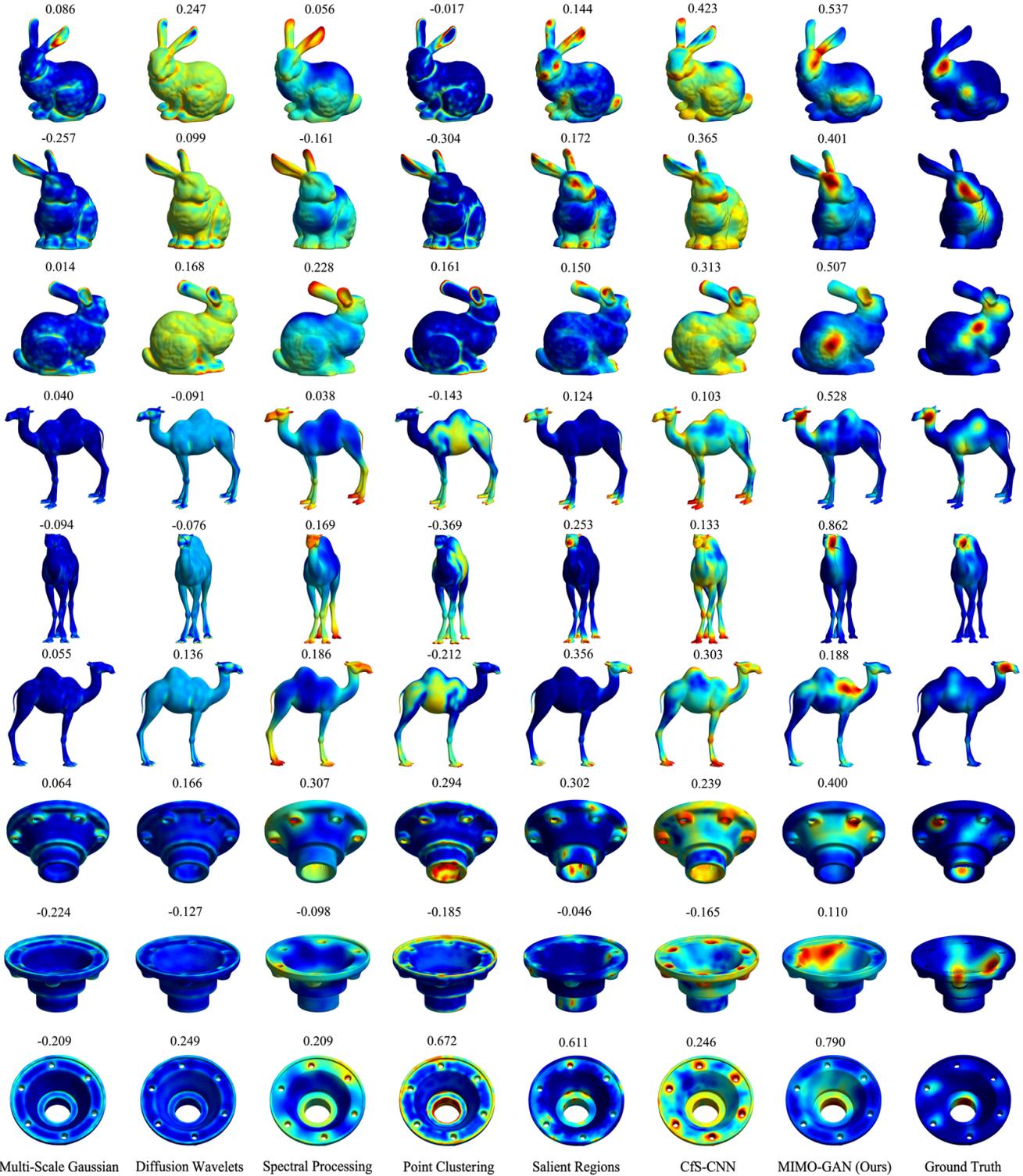


Figure 2. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

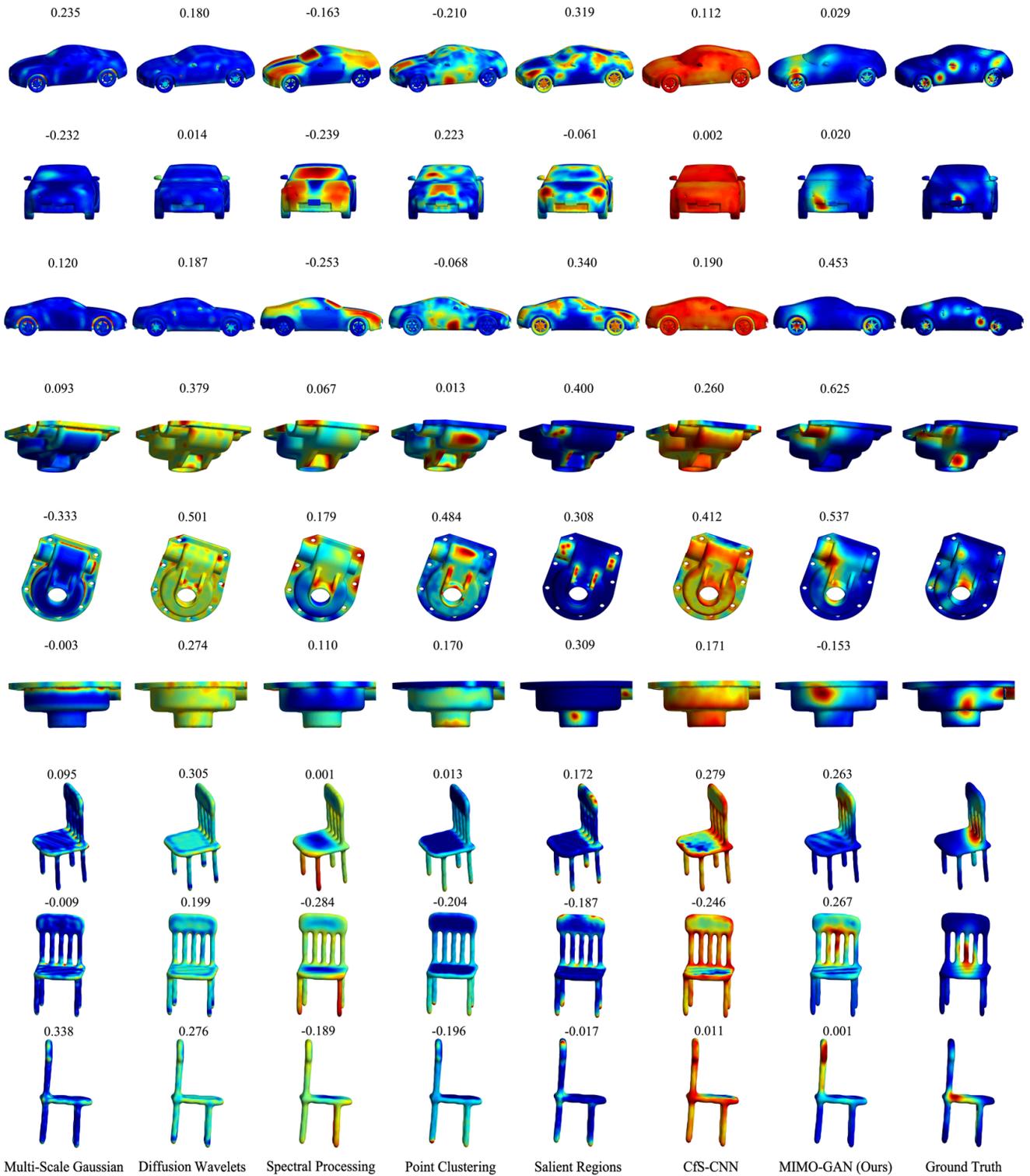


Figure 3. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

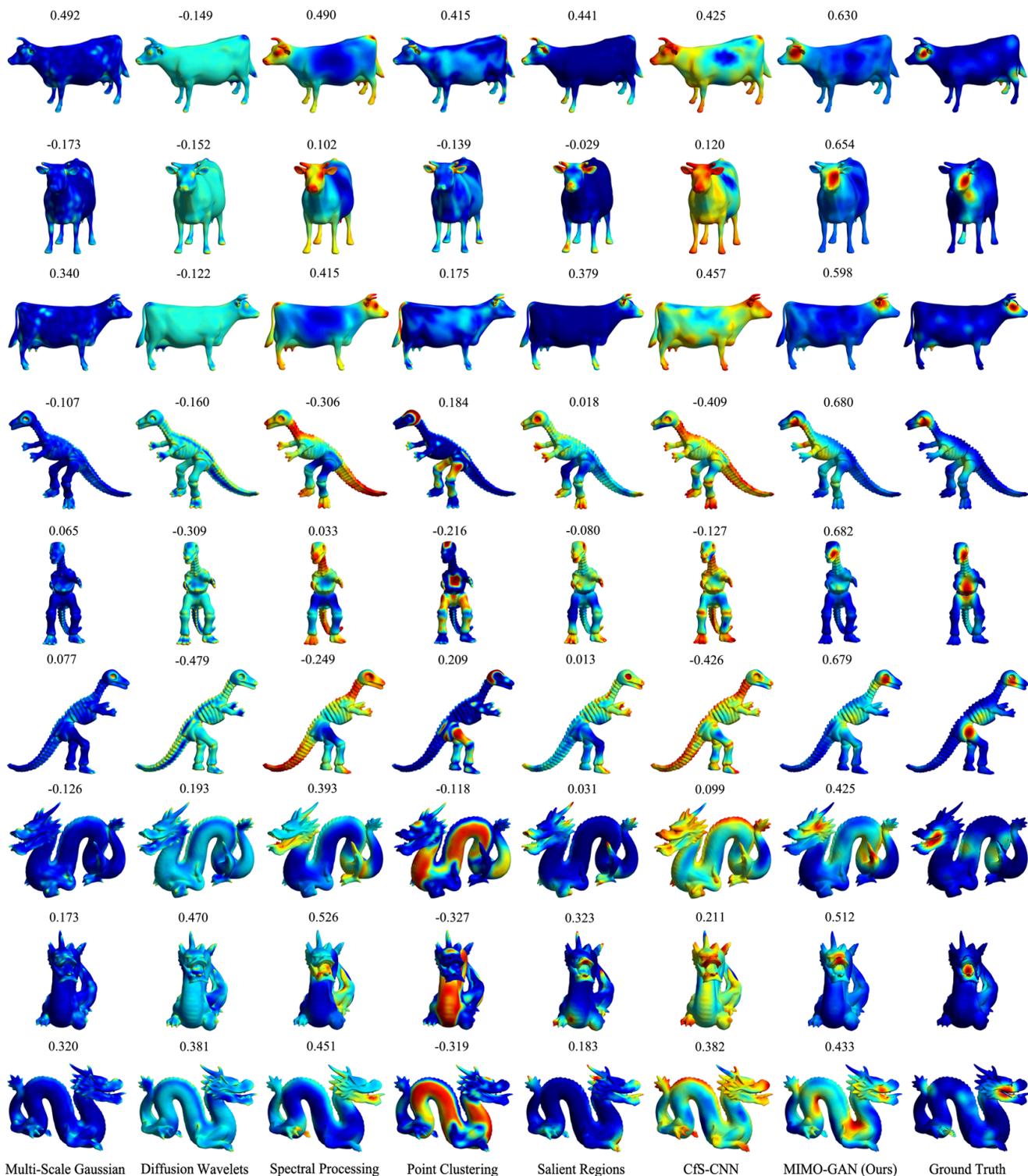


Figure 4. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

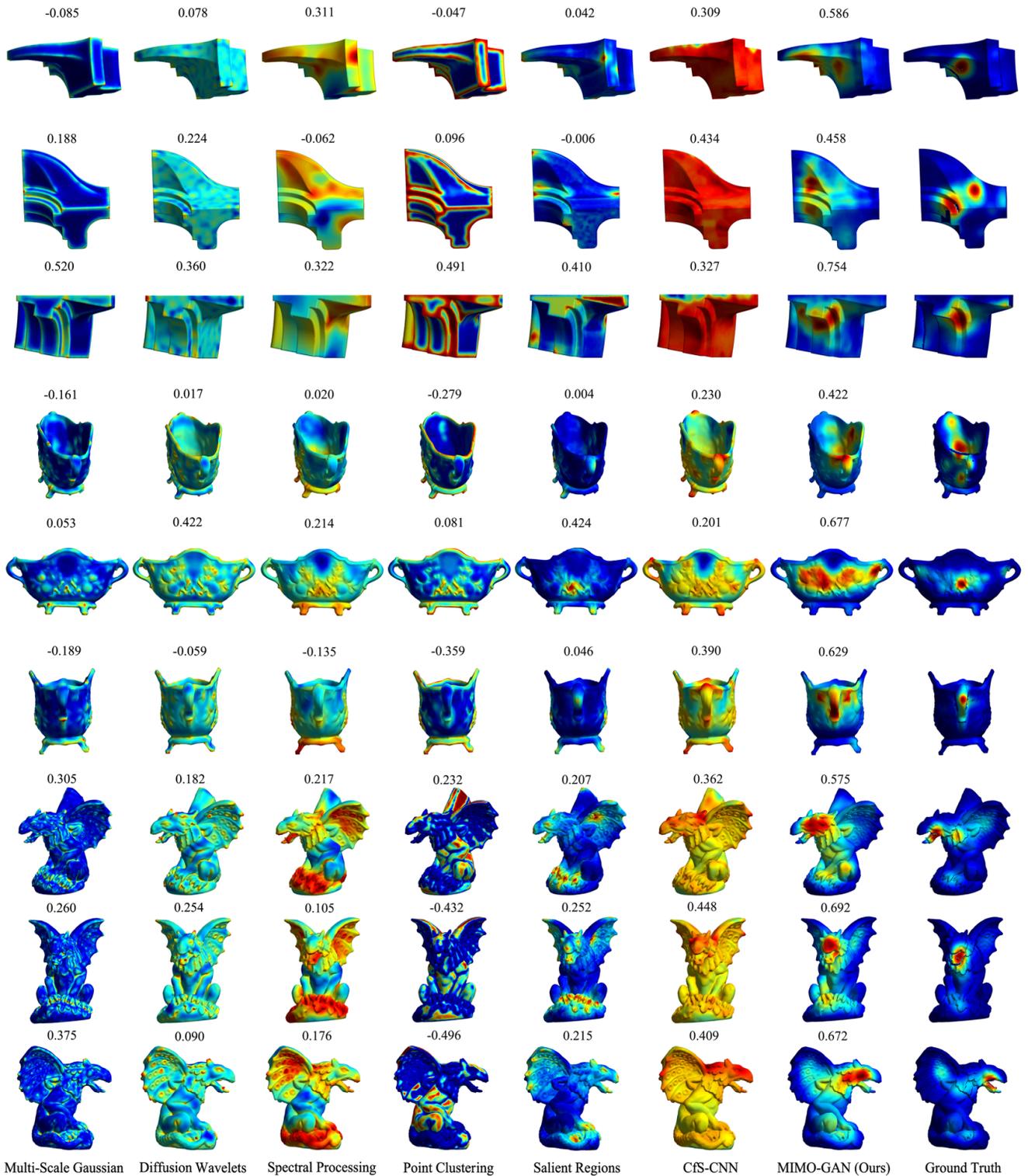


Figure 5. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

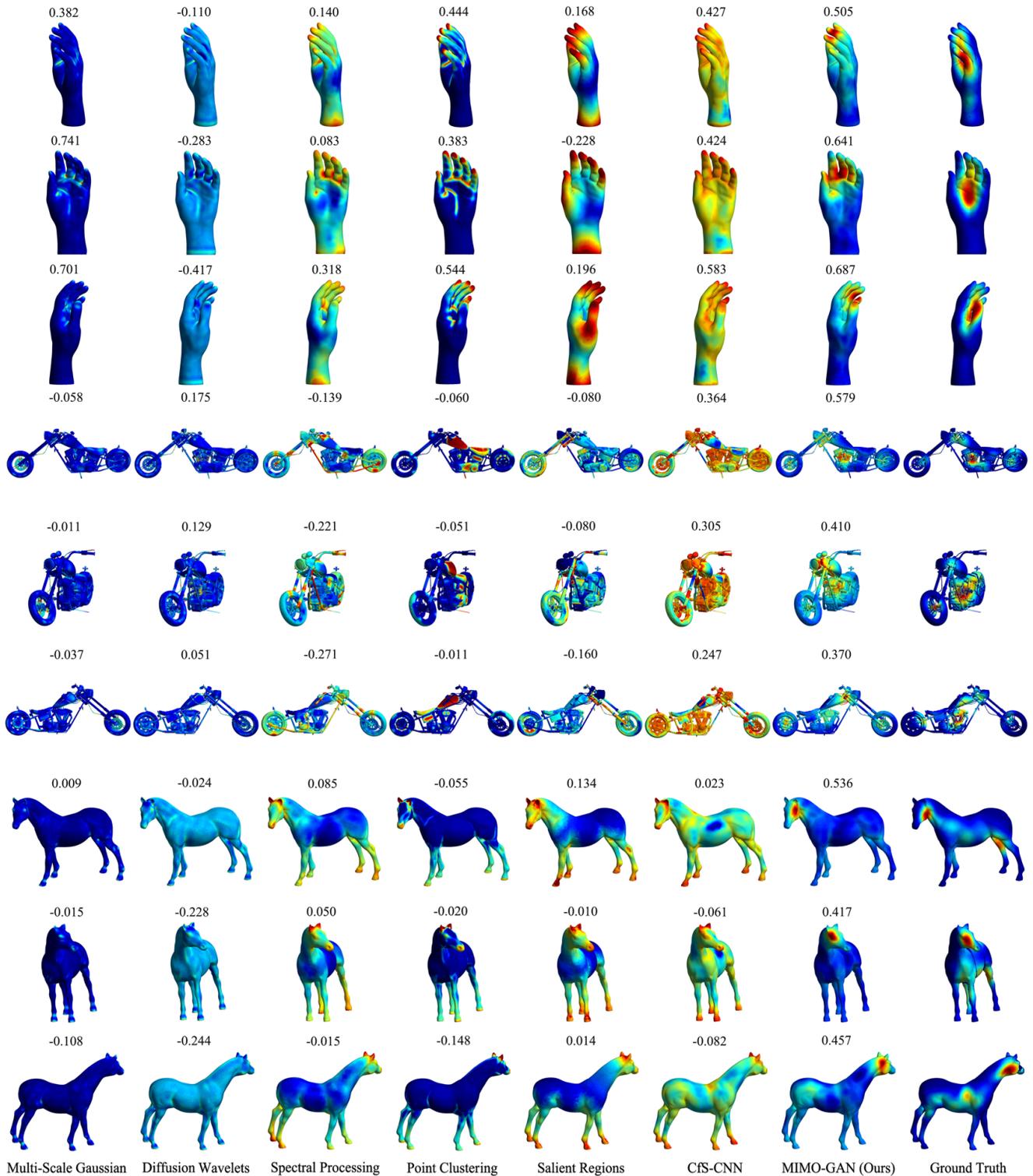


Figure 6. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

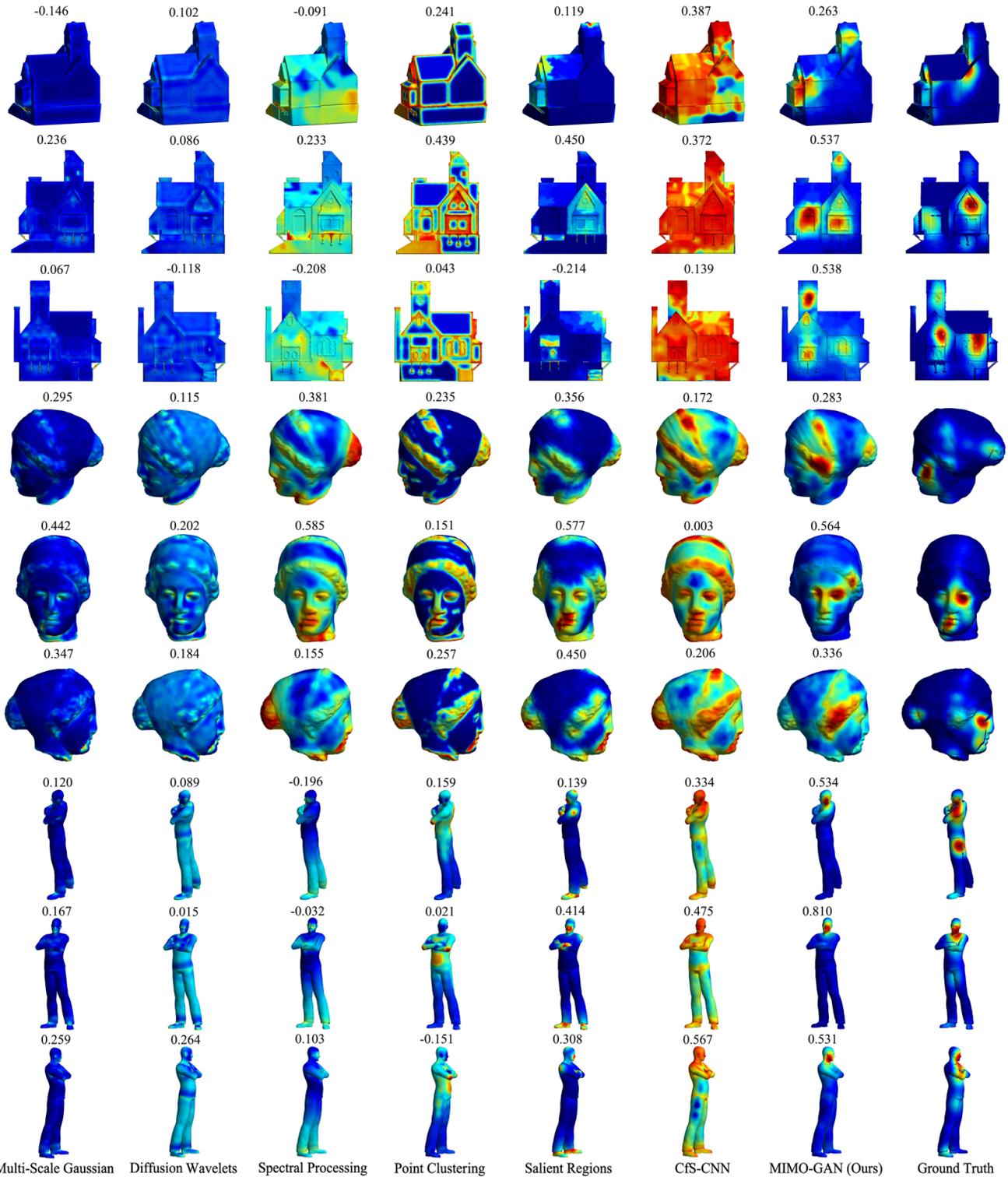


Figure 7. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

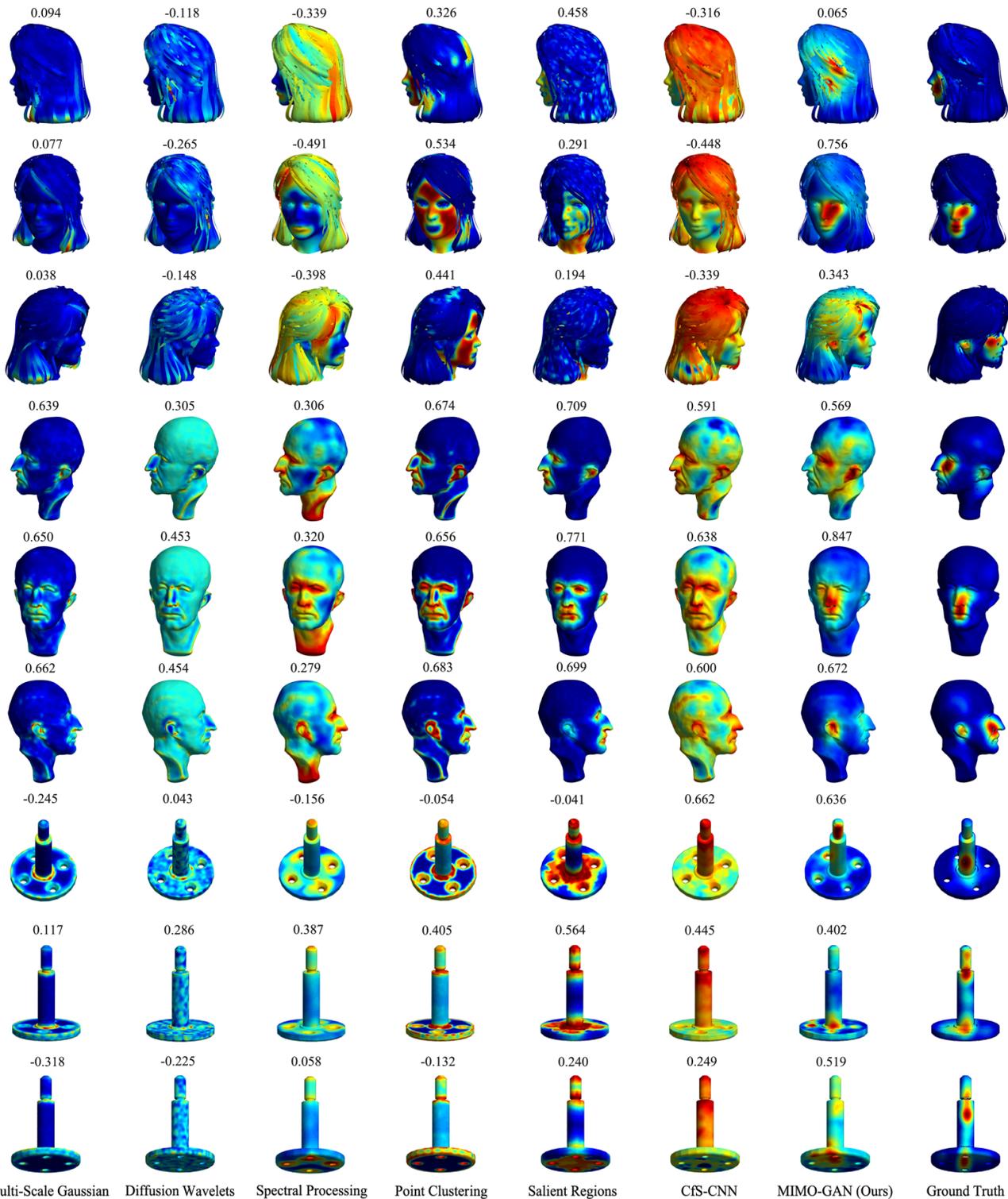


Figure 8. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

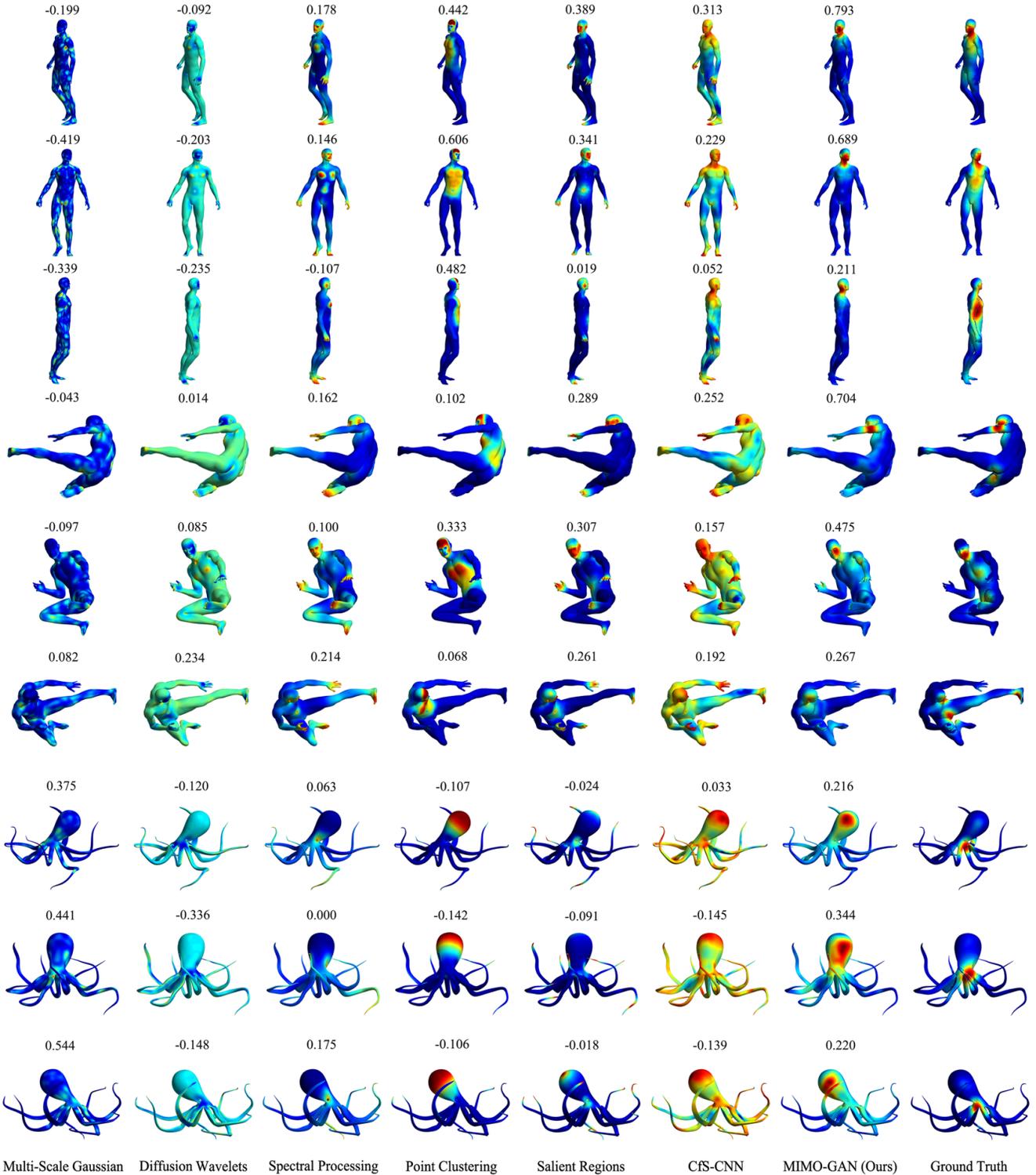


Figure 9. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

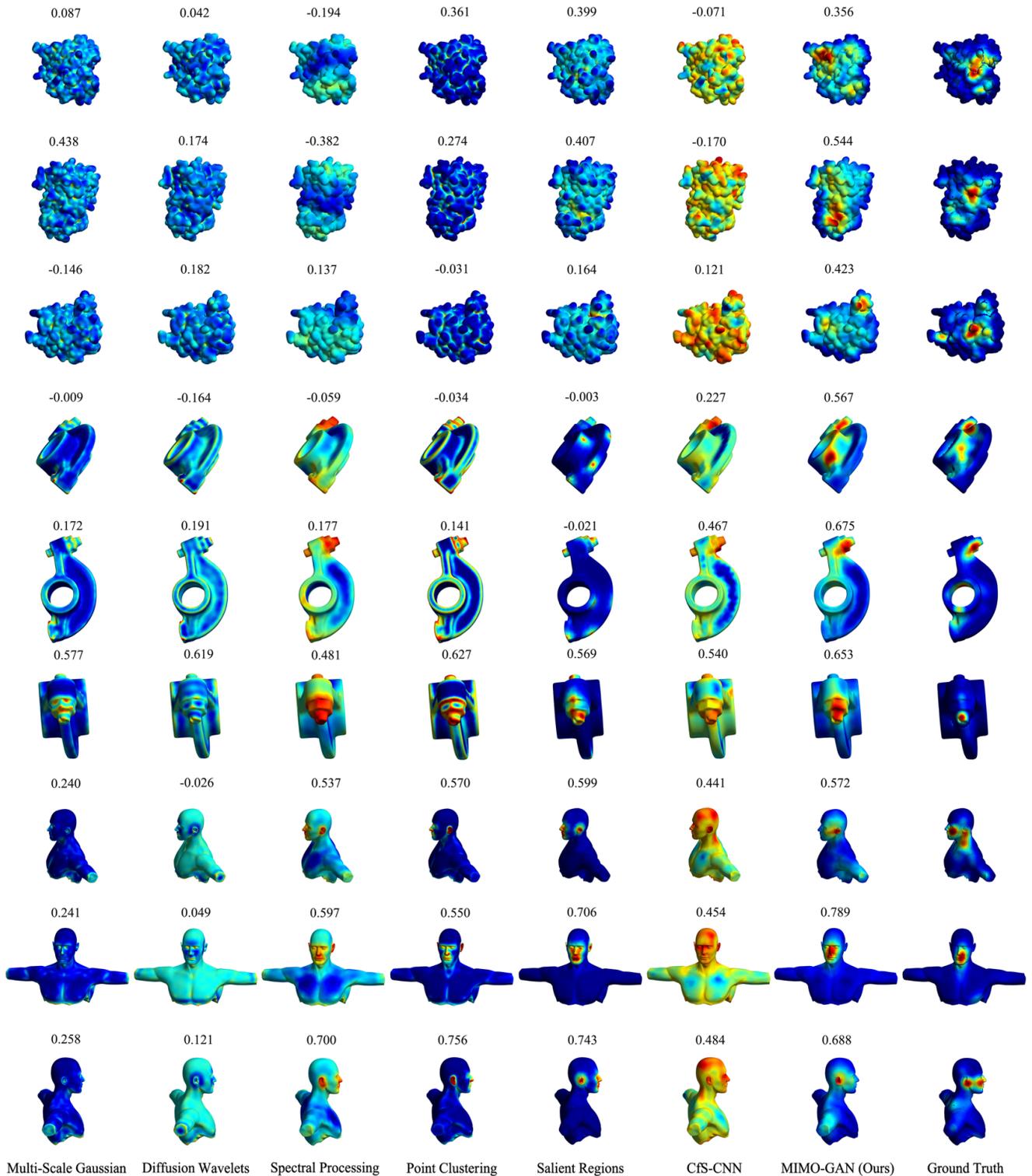


Figure 10. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

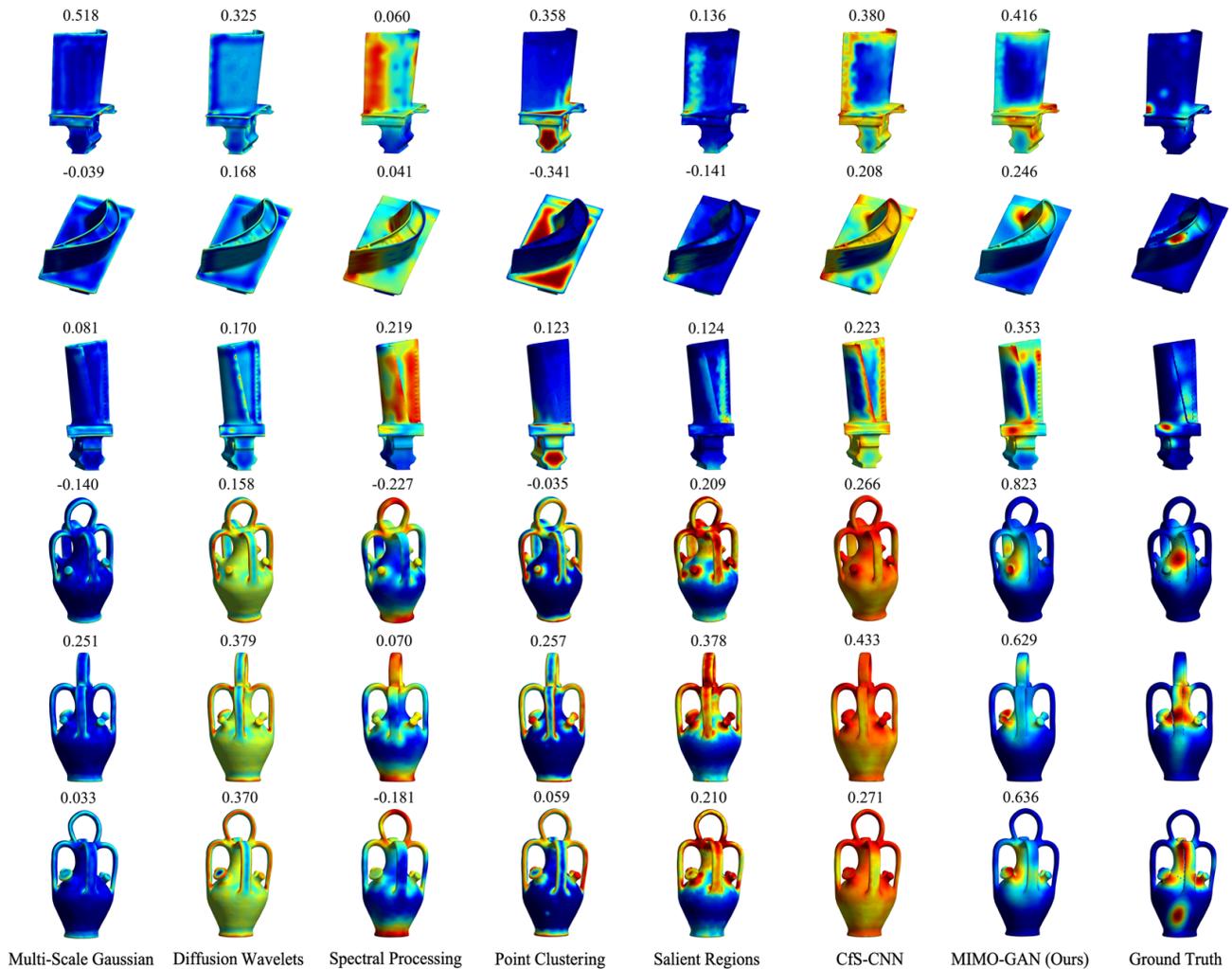


Figure 11. Qualitative and quantitative comparisons of view-dependent mesh saliency detected by different methods. From left to right: Multi-Scale Gaussian [4], Diffusion Wavelets [2], Spectral Processing [7], Point Clustering [6], Salient Regions [5], CfS-CNN [8], the proposed MIMO-GAN and the ground truth fixation maps provided by the 3DVA dataset [3]. Warmer colours show higher saliency and the corresponding LCC is shown on top of each saliency map.

*tern Anal. Mach. Intell.*, 38(12):2544–2556, 2016. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[6] Flora Ponjou Tasse, Jiri Kosinka, and Neil Dodgson. Cluster-based point set saliency. In *Proc. ICCV*, pages 163–171, 2015. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[7] Ran Song, Yonghuai Liu, Ralph R. Martin, and Paul L. Rosin. Mesh saliency via spectral processing. *ACM Trans. on Graph.*, 33(1), 2014. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[8] Ran Song, Yonghuai Liu, and Paul L. Rosin. Mesh saliency via weakly supervised classification-for-saliency CNN. *IEEE Trans. Vis. Comput. Graph.*, 21(1):151–164, 2021. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

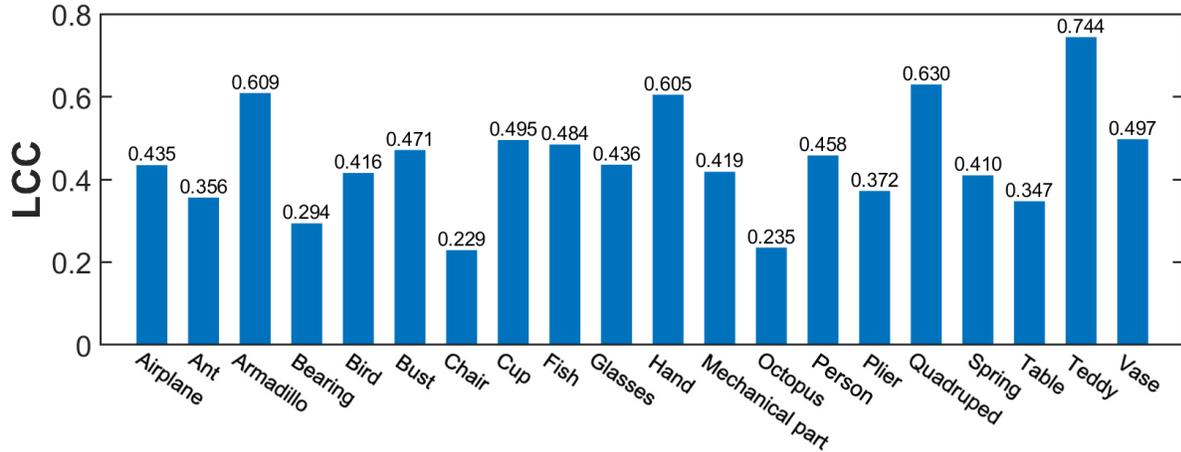


Figure 12. LCC per object category for our method ( $\sigma = 0.1B$ ).  $\sigma$  denotes the standard deviation of the Gaussian used to generate the pseudo ground truth from the Schelling dataset [1].  $B$  is the length of the diagonal of the bounding box of the mesh.

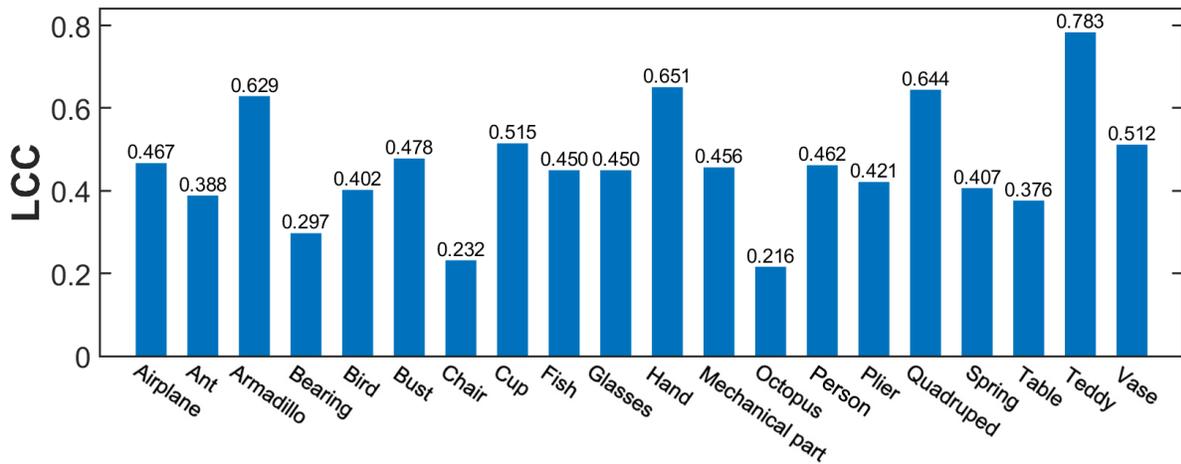


Figure 13. LCC per object category for our method ( $\sigma = 0.12B$ ).  $\sigma$  denotes the standard deviation of the Gaussian used to generate the pseudo ground truth from the Schelling dataset [1].  $B$  is the length of the diagonal of the bounding box of the mesh.

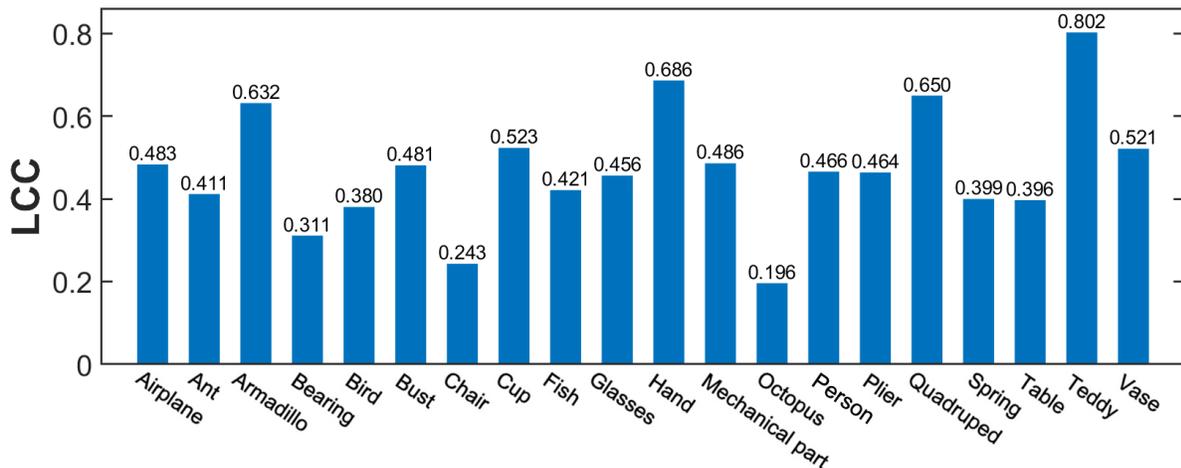


Figure 14. LCC per object category for our method ( $\sigma = 0.14B$ ).  $\sigma$  denotes the standard deviation of the Gaussian used to generate the pseudo ground truth from the Schelling dataset [1].  $B$  is the length of the diagonal of the bounding box of the mesh.

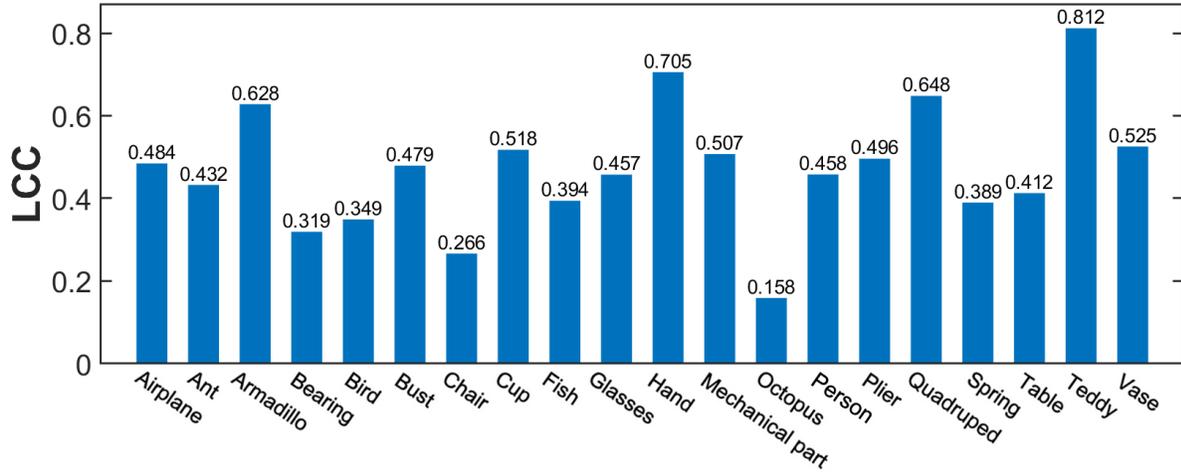


Figure 15. LCC per object category for our method ( $\sigma = 0.16B$ ).  $\sigma$  denotes the standard deviation of the Gaussian used to generate the pseudo ground truth from the Schelling dataset [1].  $B$  is the length of the diagonal of the bounding box of the mesh.

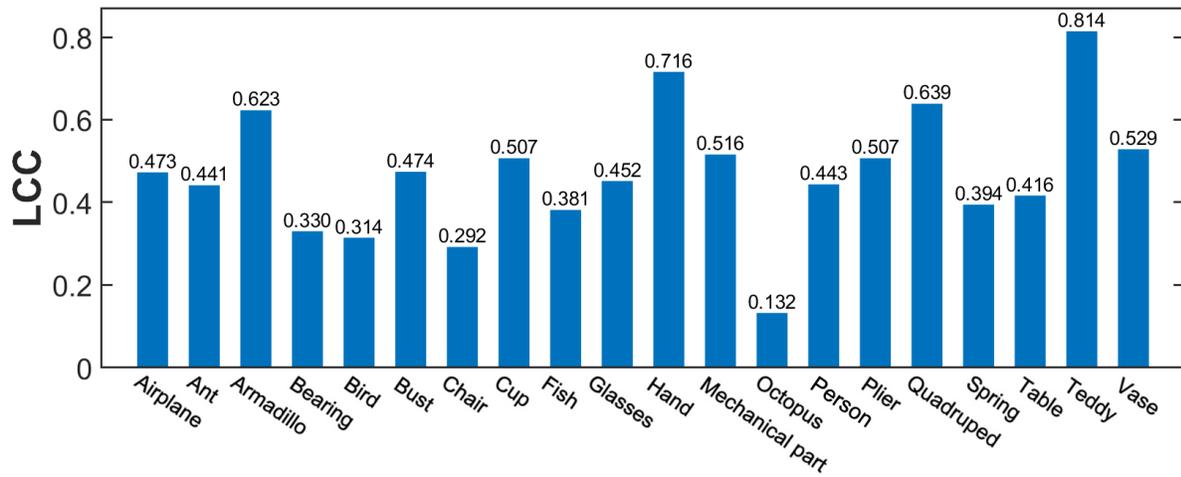


Figure 16. LCC per object category for our method ( $\sigma = 0.18B$ ).  $\sigma$  denotes the standard deviation of the Gaussian used to generate the pseudo ground truth from the Schelling dataset [1].  $B$  is the length of the diagonal of the bounding box of the mesh.

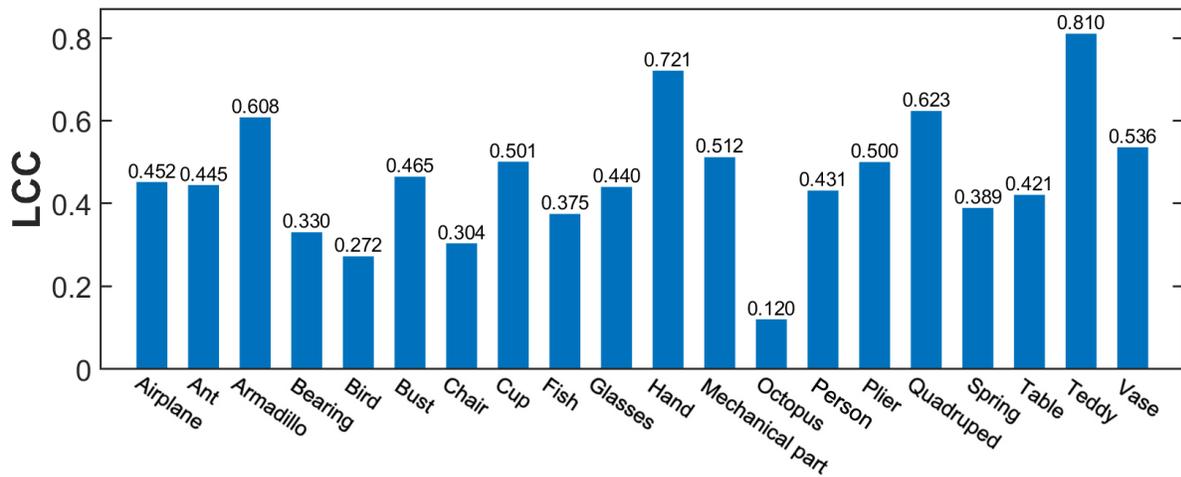


Figure 17. LCC per object category for our method ( $\sigma = 0.2B$ ).  $\sigma$  denotes the standard deviation of the Gaussian used to generate the pseudo ground truth from the Schelling dataset [1].  $B$  is the length of the diagonal of the bounding box of the mesh.