3D Mesh Saliency from local spiral hop descriptors

Olivier Lézoray, Anass Nouri

3D Mesh Saliency from local spiral hop descriptors

Olivier Lézoray; Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, Caen, France
Anass Nouri; Laboratoire des Systèmes Électroniques, Traitement de l'Information, Mécanique et Énergétique, Ibn Tofail University, Kenitra, Morocco

Abstract

Mesh saliency, the process of detecting visually important regions in 3D meshes, is a significant component in computer graphics, that can be used in various applications such as denoising and simplification. In this paper, we propose a new 3D mesh saliency measure that can identify sharp geometric features in meshes. A local normal-based descriptor is built for each vertex thanks to a spiral path within a k-hop neighborhood. First, a geometric-based saliency is computed as the mean local alignment between the spiral descriptors. Second, a spectral-based saliency is computed from the structure tensor total variation of each vertex structure tensor with the gradient defined from the spiral descriptor alignments. The final saliency is then defined as the vertex roughness weighted by the sum of geometric ans spectral saliencies. This single-scale saliency can be extended to a multi-scale saliency by decimating the mesh at several scales. The approach presents competitive results with the state-of-the-art.

Introduction

Nowadays, thanks to 3D scanners, it is easy to create digital 3D content. As such, 3D meshes are widely used in application fields such as computer graphics and games. Many problems in computer graphics can take advantage of knowing which regions from a mesh are the most important. For example, in a mesh simplification algorithm, the importance of the vertices can be used to guide the decimation. Mesh saliency is such a measure that wants to capture the importance of a region on a 3D mesh, hopefully, close to human visual perception.

Our objective is to define a measure of saliency that can operate on 3D meshes. To do so, we locally study the geometry of the mesh vertices' normals. To have a robust measure, the study should not be based on measures made by the sole vertex normals. Indeed, to well describe the geometry at a given vertex, its neighborhood has to be taken into account: a flat region being less salient that an abrupt one. This requires the definition of a local descriptor extracted at each vertex and that accounts for the normals' variations within a k-hop neighborhood similarly to patches for images. Most descriptors of the literature have adapted features from the 2D case (LBP, histograms) to the 3D case, and used it either in the spatial or spectral domains. Here we propose a descriptor specifically designed for 3D meshes. This descriptor is a local normal-based spiral descriptor built for each vertex thanks to a spiral path within a k-hop neighborhood. This descriptor serves as a building block for comparing two vertices. This enables to identify sharp geometric areas in the mesh that are considered as being salient. From this, saliency measures are derived in both the spatial and spectral domains, and are combined into a global saliency measure.

Local spiral hop descriptors

A mesh is represented by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that consists in a finite set $\mathcal{V} = \{v_1, \ldots, v_m\}$ of vertices and a finite set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ of edges. We assume $\mathcal{G}$ to be undirected, with no self-loops and no multiple edges. Let (v_i, v_j) be the edge of $\mathcal{E}$ that connects two vertices v_i and v_j of $\mathcal{V}$. The notation $v_i \sim v_j$ is used to denote two adjacent vertices. The set $\mathcal{N}(v_i) = \{v_j, v_j \sim v_i\}$ gives the set of all the adjacent vertices to v_i within a 1-hop (vertices that can be reached in one walk). In [8], Lim *et al.* have proposed local spiral hop operators. The principle is that the surrounding vertices of one vertex can be enumerated by following a spiral, as illustrated in Figure 1.

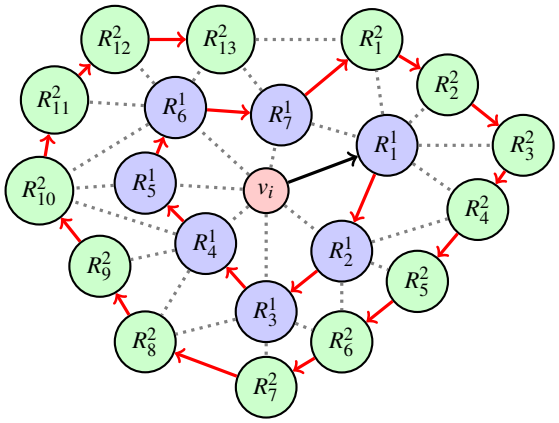

Figure 1. *Example of a local spiral 2-hop descriptor for a node v_i of a mesh.*

To define local spiral hop descriptors, we introduce the following definitions for local spiral hop operators:

$$
\begin{aligned}
R^0(v_i) &= \{v_i\} \\
R^{(k+1)}(v_i) &= \mathcal{N}(R^{(k)}(v_i)) \setminus k\text{-disk}(v_i) \\
k\text{-disk}(v_i) &= \bigcup_{l=0,\ldots,k} R^l(v_i) \\
k\text{-rings}(v_i) &= k\text{-disk}(v_i) \setminus \{v_i\}
\end{aligned}
$$

Given a vertex v_i and a radius size k, a k-disk(v_i) is the set of vertices that can be reached from v_i in 0 to k walks. As this includes v_i, the set k-rings(v_i) is the same set without v_i. $R^k(v_i)$ is the k-ring (also called k-hop): an ordered set of vertices whose shortest path to v_i is exactly k-hops long. Then, $R^{(k+1)}(v_i)$ is the set of vertices that can be reached in 1 walk from $R^k(v_i)$ without going through its k-disk (that contains vertices that can be reached from v_i in 0 to k walks). $R^k_j(v_i)$ denotes the j-th element in the k-ring. Obviously one has $R^1(v_i) = \mathcal{N}(v_i)$ and $R^0_1(v_i) = v_i$. From these

operators, we can define the local spiral hop operator $\mathrm{Sp}(v_i,k)$ as an ordered sequence from the concatenation of the ordered rings:

$$\begin{aligned}\mathrm{Sp}(v_i,k) &= (v_i, 1\text{-ring}(v_i),\ldots,k\text{-ring}(v_i)) \\ &= \left(R_1^0(v_i),R_1^1(v_i),R_2^1(v_i),\ldots,R_{|R^k|}^k(v_i)\right)\end{aligned}$$

This operator has two degrees of freedom: the direction (clockwise or counterclockwise) of the rings and the first chosen vertex $R_1^1(v_i)$ (shown by the black arrow in Figure 1). The rest of vertices are ordered inductively. To suppress both freedom degrees, we fix the orientation clockwise and choose the initial vertex $R_1^1(v_i)$ as the one in the direction of the shortest geodesic path to v_i:

$$R_1^1(v_i) = \arg\min_{v_j\in\mathcal{N}(v_i)} d_{\mathcal{G}}(v_i,v_j)$$

where $d_{\mathcal{G}}$ is the geodesic distance between two vertices on the graph $\mathcal{G}$. This enables the operator to become invariant to rotations of the neighborhoods. Such a choice has also been considered in [2] for the definition of spiral convolution on graphs. The local hop spiral operator has been used in [8, 2] and provided competitive results for shape correspondence. However, as the size of the operator $\mathrm{Sp}(v_i,k)$ varies for the vertices (as all the vertices do not have the same number of neighbors), both these approaches have considered only fixed-size spiral. They either truncate or zero-pad each spiral depending on its size. This obviously does not enable to well capture the similarities between two different spirals. Therefore, we proceed differently. Given a graph signal $\mathbf{F}$ on $\mathcal{G}$, with $\mathbf{F}:\mathcal{V}\to\mathbb{R}^d$, we define the difference between the spiral descriptors of two vertices as the sum of the differences between their respective k-rings:

$$d(Sp(v_i,k),Sp(v_j,k)) = \sum_{l=0}^{k} d(R^l(v_i),R^l(v_j))$$

Two k-rings are compared by mapping the vertices of the largest ring to the smallest one:

$$d(R^l(v_i),R^l(v_j)) = \sum_{n=0}^{|R^l(v_i)|} d(R_n^l(v_i),R_{n'}^l(v_j))$$

with $n' = \left\lfloor \frac{n\cdot|R^l(v_j)|}{|R^l(v_i)|} \right\rfloor$ and $|R^l(v_i)| > |R^l(v_j)|$. The distance between two vertices is then the distance between their graph signal vectors:

$$d(R_n^l(v_i),R_m^l(v_j)) = \|\mathbf{F}(R_n^l(v_i)) - \mathbf{F}(R_m^l(v_j))\|_2$$

The whole process is illustrated in Figure 2: vertices of the largest k-hop are mapped to these of the smallest k-hop. The graph signal $\mathbf{F}$ can be spatial coordinates or normal vectors. In this work we will consider normal vectors: $\mathbf{F}(v_i)=\mathbf{n}(v_i)$. Faces' normals $\hat{\mathbf{n}}_j$ are defined on each triangle t_j of a mesh as the vector orthogonal to triangle's plane. Vertices' normals $\mathbf{n}(v_i)$ are computed as a weighted average of incident face normals with angle-based weighting [1]:

$$\mathbf{n}(v_i) = \frac{\sum_{j\in\mathcal{N}_f(v_i)} \alpha_j\hat{\mathbf{n}}_j}{\|\sum_{j\in\mathcal{N}_f(v_i)} \alpha_j\hat{\mathbf{n}}_j\|_2}$$

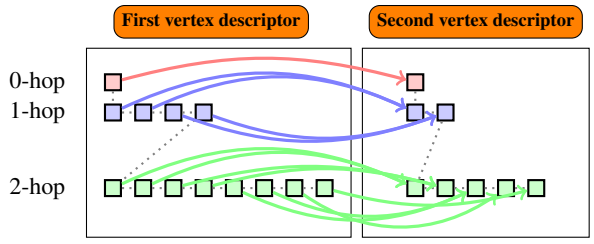

Figure 2. *Comparison of two local spiral 2-hop descriptors.*

with $\mathcal{N}_f(v_i)$ the faces incident with v_i and α_i the incident angle.

Figure 3 shows an example where the local spiral 2-hop descriptor of a vertex v_s is compared with the one of all the other vertices. The distance $d(Sp(v_s,2),Sp(v_j,2))$ is then affected to each vertex v_j and displayed. As it can be seen this enables to well capture the similarity between similar vertices. We will use this descriptor as the basis for comparing vertices to define geometric and spectral saliencies.

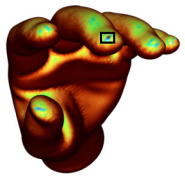

Figure 3. *Comparison of the spiral 2-hop descriptors of a selected vertex v_s (shown in the black rectangle) with all the other vertices. Warm colors mean high values of $d(Sp(v_s,2),Sp(v_j,2))$.*

Roughness

In [3], it was shown that curvature plays an important role for the detection of saliency. In [6], Lee *et al.* proposed the first mesh saliency approach and used differences between Gaussian-weighted mean curvatures. Indeed, for saliency detection, it is preferable to consider the variation of the curvature instead of the curvature itself [5]. This enables to avoid detecting high-curvature smooth areas as salient regions. Therefore, we consider instead the notion of roughness proposed by Wang *et al.* [15] that we extend to rings instead of local neighborhoods.

First, for each vertex v_i of the mesh, we compute its mean curvature. The two principal curvatures $(k_1(v_i),k_2(v_i))$ at a vertex on a mesh measure how much the surface bends in different directions. They correspond to the maximal and minimal curvature at a given vertex of the surface. The approach of Panozzo *et al.* [11] is used to compute the two principal curvatures. In the 2-rings(v_i) neighborhood around every vertex v_i, a best-fit quadric is found and principal curvature values are analytically computed on this quadric. Then, the mean curvature is the average of the principal curvatures [9]:

$$\kappa(v_i) = \frac{1}{2}(k_1(v_i)+k_2(v_i))$$

Values are close to 0 for flat surface, positive in convex areas and negative in concave areas. The roughness at a vertex v_i is defined as the difference between the curvature of a vertex and its neighbors (weighted by its Laplacian). In [15], the definition considered only the direct neighbors $\mathcal{N}(v_i)$. To have a more precise

measure, we consider larger neighborhoods γ-rings(v_i) and take the γ power of the Laplacian to create weights between vertices that are γ walks far. The roughness is defined as:

$$R(v_i) = \left| \kappa(v_i) - \frac{\sum_{v_j \in \gamma\text{-rings}(v_i)} L_{ij}^{\gamma} \cdot \kappa(v_j)}{\sum_{v_j \in \gamma\text{-rings}(v_i)} L_{ij}^{\gamma}} \right|$$

with L the cotangent Laplacian:

$$L_{ij} = \begin{cases} \cot\alpha_{ij} + \cot\beta_{ij} & v_j \in \mathcal{N}(v_i), \\ 0 & j \notin \mathcal{N}(v_i), \\ -\sum_{k \neq i} L_{ik} & i = j, \end{cases}$$

α_{ij} and β_{ij} are the angles opposite to edge (v_i, v_j).

Geometric and Spectral saliencies

We define the gradient at a given vertex v_i as the (nonlocal) vector of all the distances between the spiral descriptors of v_i and its neighbors within its γ-rings(v_i):

$$\nabla \mathbf{f}(v_i) = \left[d(Sp(v_i,k), Sp(v_j,k)), v_j \in \gamma\text{-rings}(v_i) \right]^T$$

From this nonlocal gradient, we define geometric and spectral saliencies. The geometric saliency is the normalized L_1 norm of the gradient:

$$\begin{aligned} GS(v_i) &= \frac{1}{|\gamma\text{-rings}(v_i)|} \|\nabla \mathbf{f}(v_i)\|_1 \\ &= \frac{1}{|\gamma\text{-rings}(v_i)|} \sum_{v_j \in \gamma\text{-rings}(v_i)} d(Sp(v_i,k), Sp(v_j,k)) \end{aligned}$$

This definition of geometric saliency is similar to the one proposed in [10] where a patch descriptor is used instead of our local spiral hop descriptor. The geometric saliency will have high values for vertices that are very different from their neighbors as the spiral distances will be high.

The structure tensor $\mathbf{J}$ is the outer product of the gradient: $\mathbf{J}(v_i) = \nabla^T \mathbf{f}(v_i) \cdot \nabla \mathbf{f}(v_i)$. It is a $|\gamma\text{-rings}(v_i)| \times |\gamma\text{-rings}(v_i)|$ positive semi-definite matrix: $\mathbf{J}(v_i) \in \mathbb{S}_+^{|\gamma\text{-rings}(v_i)|}$. It summarizes the dominant directions of the gradient in the γ-rings(v_i) of a vertex v_i. The importance of the structure tensor lies in its eigenvalues that provide a rich and discriminative description of the local geometry by summarizing the distribution of the gradients in the γ-rings(v_i) neighborhood. Its spectral decomposition is $\mathbf{J}(v_i) = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ with $\mathbf{U}$ its eigenvectors and $\mathbf{\Lambda}$ its eigenvalues. The eigenvalues of the structure tensor can be more informative of the local structures of the mesh than the gradient magnitude, that is very smooth (see Figure 4). This leads us to define the spectral saliency from the Structure Tensor Total Variation (STTV) [7]. It is defined as:

$$SS(v_i) = \sqrt{\sum_{j=1}^{|\gamma\text{-rings}(v_i)|} \lambda_j^2}$$

with $\lambda_j = \mathbf{\Lambda}(j,j)$. Since the structure tensor can capture first-order information in the γ-rings(v_i) neighborhood, the spectral saliency based on SSTV can provide more robust measures of variation. Figure 4 shows the robustness and accuracy of both the geometric and spectral saliencies. As it can be seen they can identify sharp geometric features of the mesh even under strong noise (on the mesh vertices' coordinates).

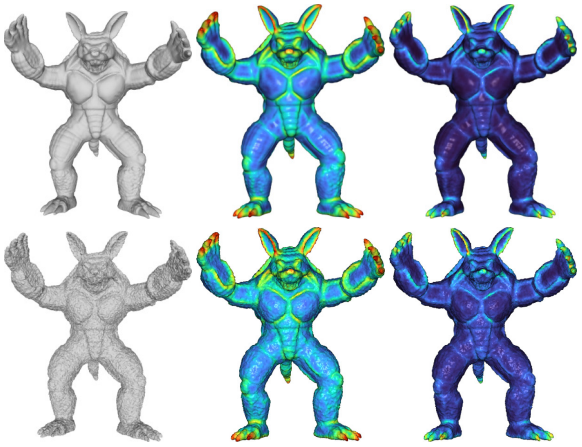

Figure 4. *Robustness to noise. Original (top) and noisy (bottom) mesh with their geometric and spectral saliencies (from left to right) on 3-rings with 2-hop spiral descriptors.*

Given the roughness, the geometric and spectral saliencies of a given vertex, we now introduce the computation of the saliency. As the roughness is based on differences of curvatures, it tends to be very sensitive to small changes. In contrast the geometric saliency is smooth, and the spectral saliency very localized. They both provide different salient informations on the mesh. To merge these informations altogether into a single saliency measure, we weight the roughness by the sum of the normalized geometric and spectral saliencies:

$$S(v_i) = K_\sigma \left(R(v_i) \cdot \left(\overline{GS(v_i)} + \overline{SS(v_i)} \right) \right)$$

This definition enables to have a saliency measure that puts forward areas with high values of roughness, geometric and spectral saliencies. K_σ is a Gaussian smoothing in the γ-rings(v_i) neighborhood with $\sigma = 0.002 \cdot B$ where B is the length of the diagonal of the bounding box of the mesh (as in [12]). The geometric and spectral saliencies are normalized between 0 and 1 beforehand with $\overline{S(v_i)} = \frac{S(v_i) - \min(S)}{\max(S) - \min(S)}$.

Results

To illustrate the benefits of our proposed mesh saliency method, several experiments are led. In Figure 5, we show the resulting mesh roughness $R(v_i)$, mesh geometric saliency $\overline{GS(v_i)}$, mesh spectral saliency $\overline{SS(v_i)}$, and the final saliency $S(v_i)$ combining them altogether. The roughness is a good starting point for saliency. However, it tends to perform a lot of over-detection in areas that are not salient, and to be saturated in areas that are highly salient. On the contrary, the geometric saliency is very smooth and detects more finely some sharp geometric features (such as the eyes of the centaur or the dress of the angel). The spectral saliency also detects finely sharp geometric features but is much less smooth and more localized on strong variations. The final combinaison enables to put forward only the most salient parts of the mesh, while smooth or flat areas are not detected as salient.

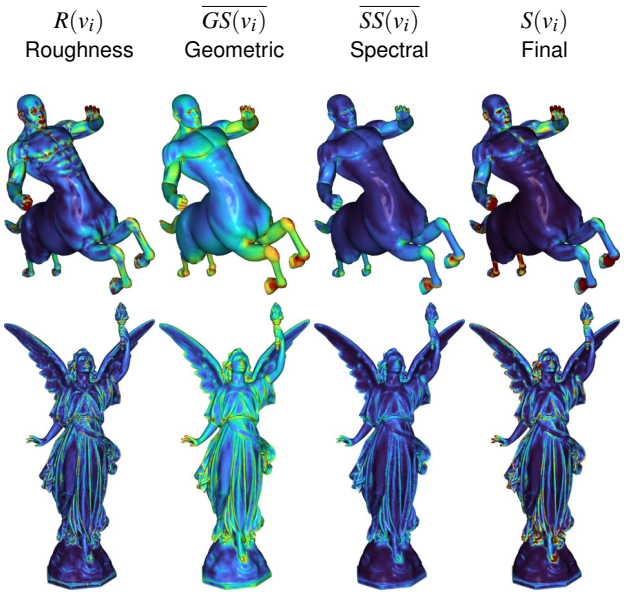

Figure 5. *Saliency computation on 3-rings with 2-hop spiral descriptors $Sp(v_i,2)$. Warm colors mean high values of saliency.*

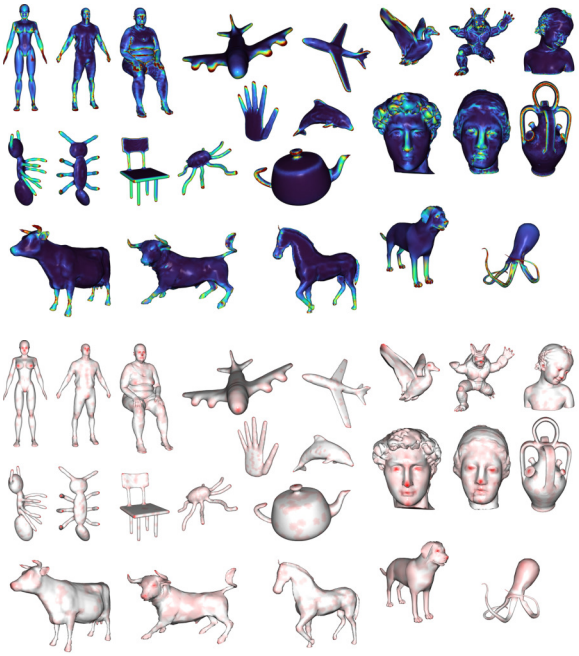

Figure 6. *A gallery of mesh saliency where warm color denotes high saliency. Above results computed by our method. Below corresponds to pseudo ground truths provided in [3]. These models are courtesy of the Watertight Models of SHREC 2007. Saliency computation was made on 2-rings with 1-hop spiral descriptors $Sp(v_i,1)$.*

To further test our approach, we have considered the meshes from Chen *et al.* [3] where a pseudo ground truth is provided from eye fixation maps. Figure 6 shows a gallery of our saliency detection results. We use the same parameter settings for all the meshes. As the latter are of relatively small size (between 5K and 15K vertices), we consider 2-rings with 1-hop spiral descriptors. As it can be seen, our results are largely consistent with the ground truths. However, on some of the meshes, the ground truth salient regions are in smooth areas, and are not detected by our approach (such as the head of the octopus) as it is based on the principle that salient regions should be different from their surrounding.

Finally, we compare our method with the state of the art. Figure 8 compares our results with those from [12, 13, 14]. Our approach produces much more fine results than those of [14] and [12] that are very coarse and often fail to detect small salient areas. The approach of [13] produces more fine results but tends to over-detect saliency in smooth areas. This is not the case with our approach. In addition, some approaches using spectral decomposition (such as [14]) compute the saliency on a decimated mesh to accelerate the processing, this is not the case of our approach. Nevertheless, we can also benefit from computing the saliency at different scales. To do so, we use the Qslim algorithm [4] to decimate the mesh into three scales (the original one and two other versions with respectively 50% and 25% of the vertices). Once the saliency is computed on a decimated mesh, it is mapped back to the original one by mapping the saliency of each decimated vertex onto their birth vertex. The other vertices get their saliency from averaging those of the birth vertices in their neighborhood. On these decimated meshes we run exactly the same algorithm but diminish the size of the rings for the computation of the saliency (from 3 to 2 and 1 rings), as the size of the salient areas becomes rougher. Results are shown in Figure 8. Importantly, the saliencies (although simplified) remain consistent across the scales (see the columns 5 to 7) and close to the eye tracker ground truth. Finally, we have tried to combine these three saliency maps into a single one with a basic max rule. The result is shown in the last column of Figure 8. Although very basic, this enables to recover in the fusion salient regions that have disappeared during the mesh decimation (such as the navel of the man or the eye of the horse). In the future we plan to investigate more elaborate fusion schemes.

As a last experiment, we show the result of our (multi-scale) saliency detection on 3D meshes obtained from the scanning of real objects. Figure 7 shows these results. Whatever the scanned object (the drill, the shoe or the persons), the most salient parts of the objects are well captured.

Conclusion

In this paper we have proposed a new 3D mesh saliency approach. To accurately detect salient regions (that are different from their surrounding), we consider a local normal-based descriptor that is built for each vertex according to a spiral path within a k-hop neighborhood. This descriptor is then be used to define a nonlocal gradient for each vertex. The normalized L_1 norm of this gradient defines a geometric saliency measure while the structure tensor total variation defines a spectral saliency measure. Both are combined altogether with a vertex roughness measure. Experimental results show the interest of the approach.

References

[1] J. A. Bærentzen and H. Aanæs. Signed distance computation using the angle weighted pseudonormal. *IEEE Trans. Vis. Comput. Graph.*, 11(3):243–253, 2005.

[2] G. Bouritsas, S. Bokhnyak, S. Ploumpis, S. Zafeiriou, and M. M. Bronstein. Neural 3d morphable models: Spiral con-

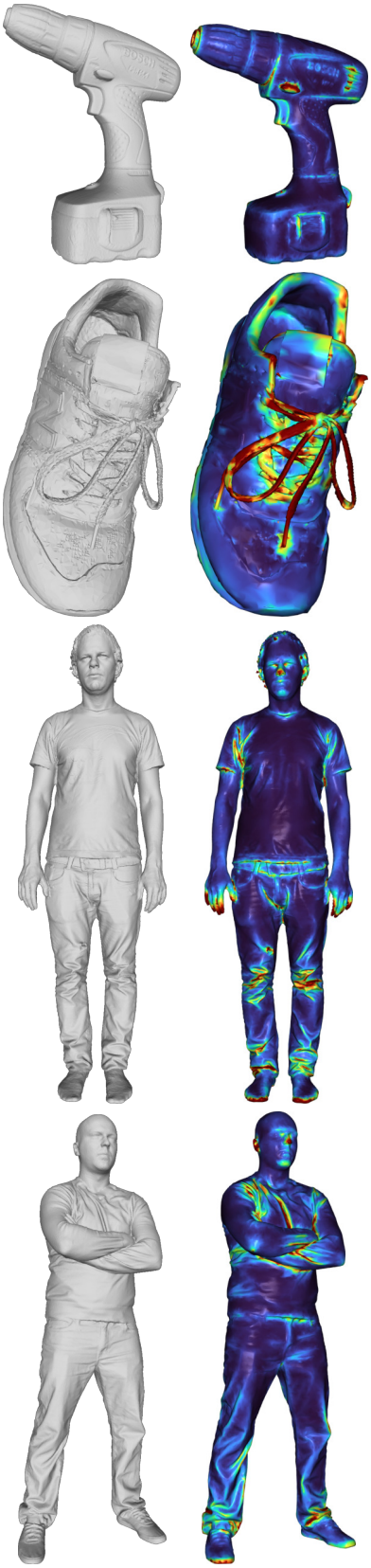

Figure 7. *Saliency detection on 3D meshes obtained from 3D scanners. Saliency computation was made on 3-rings with 2-hop spiral descriptors.*

volutional networks for 3d shape representation learning and generation. In *ICCV*, pages 7212–7221, 2019.

[3] X. Chen, A. Saparov, B. Pang, and T. A. Funkhouser. Schelling points on 3D surface meshes. *ACM Trans. Graph.*, 31(4):29:1–29:12, 2012.

[4] M. Garland and P. S. Heckbert. Surface simplification using quadric error metrics. In *SIGGRAPH*, pages 209–216. ACM, 1997.

[5] G. Lavoué. A local roughness measure for 3D meshes and its application to visual masking. *ACM Trans. Appl. Percept.*, 5(4):21:1–21:23, 2009.

[6] C. H. Lee, A. Varshney, and D. W. Jacobs. Mesh saliency. *ACM Trans. Graph.*, 24(3):659–666, 2005.

[7] S. Lefkimmiatis and S. J. Osher. Nonlocal structure tensor functionals for image regularization. *IEEE Trans. Computational Imaging*, 1(1):16–29, 2015.

[8] I. Lim, A. Dielen, M. Campen, and L. Kobbelt. A simple approach to intrinsic correspondence learning on unstructured 3d meshes. In *ECCV*, volume 11131 of *LNCS*, pages 349–362, 2018.

[9] M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. In *Visualization and Mathematics*, pages 35–57, 2002.

[10] A. Nouri, C. Charrier, and O. Lézoray. Multi-scale mesh saliency with local adaptive patches for viewpoint selection. *Signal Process. Image Commun.*, 38:151–166, 2015.

[11] D. Panozzo, E. Puppo, and L. Rocca. Efficient multi-scale curvature and crease estimation. In *Computer Graphics, Computer Vision and Mathematics*, 2010.

[12] R. Song, Y. Liu, R. R. Martin, and K. Rodriguez-Echavarria. Local-to-global mesh saliency. *Vis. Comput.*, 34(3):323–336, 2018.

[13] R. Song, Y. Liu, R. R. Martin, and P. L. Rosin. Mesh saliency via spectral processing. *ACM Trans. Graph.*, 33(1):6:1–6:17, 2014.

[14] J. Sun, M. Ovsjanikov, and L. J. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Comput. Graph. Forum*, 28(5):1383–1392, 2009.

[15] K. Wang, F. Torkhani, and A. Montanvert. A fast roughness-based approach to the assessment of 3d mesh visual quality. *Comput. Graph.*, 36(7):808–818, 2012.

Author Biography

Olivier Lézoray *received the Ph.D and Habilitation degrees in Computer Science from the University of Caen Normandy (UNICAEN), France, in 2000 and 2007. In 2000, he joined UNICAEN as an Associate Professor of Computer Science. In 2010, he became a Full Professor and was department head from 2015 to 2021. Since 2022 he is deputy director of the GREYC CNRS research lab. His research interests are in graph-based signal processing and machine learning.*

Anass Nouri *received the Ph.D degree in Computer Science in 2016 at the GREYC Laboratory in the National School of Engineering of Caen (ENSICAEN), Normandie University (France). He gaves lectures as an assistant and associate professor of computer science in ENSICAEN and ISEN-Brest and was a postdoctoral fellow in the INSERM laboratory in Nantes. Currently, he is an associate professor of Computer Science in Ibn Tofail University (Morocco).*

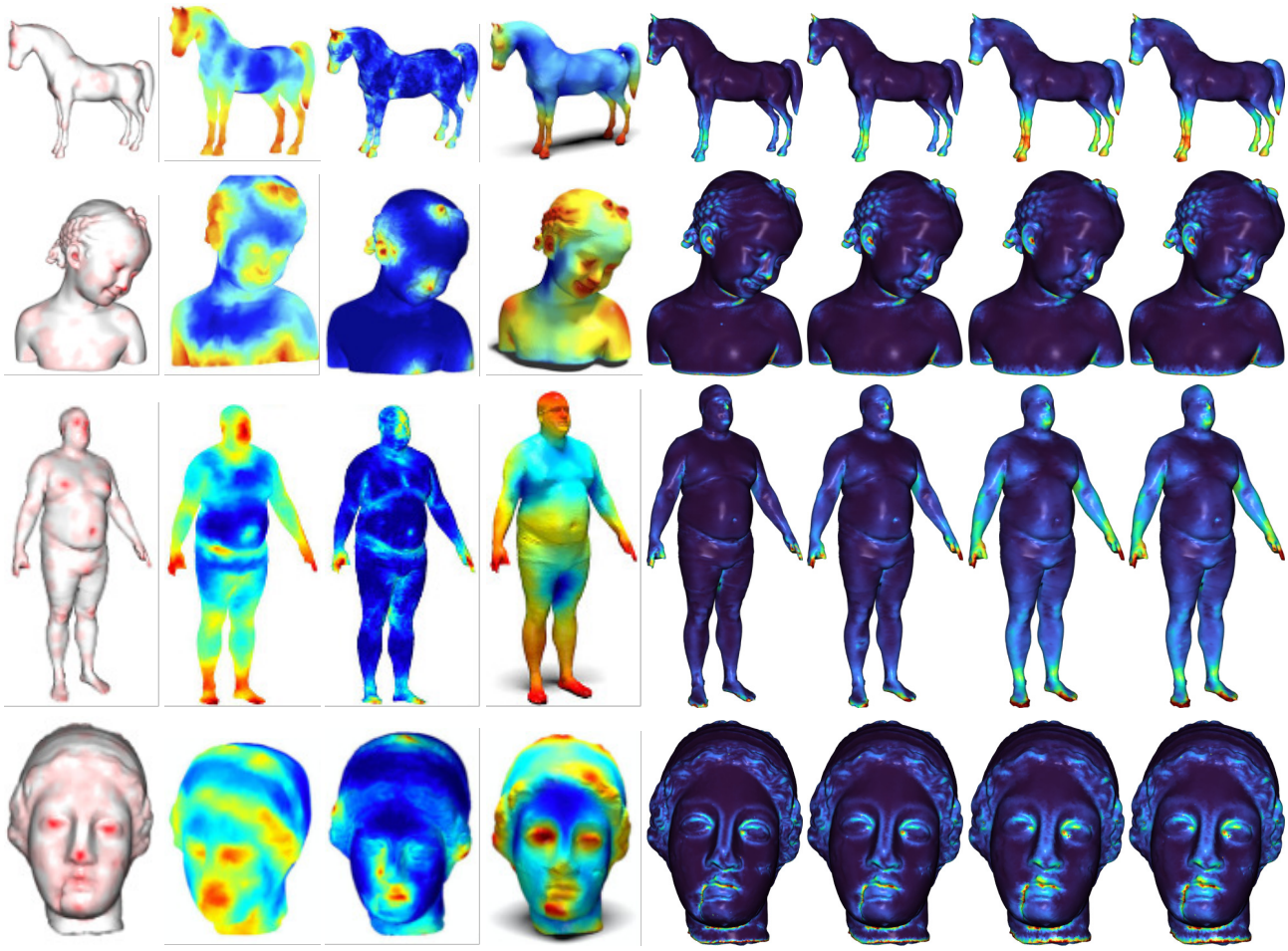

Figure 8. *From left to right : Ground truth [3], saliency detection methods of [14], [13], [12], Ours (first scale), Ours (second scale), Ours (third scale), Ours (fusion of the scales by max)*