



HAL
open science

Multi-Scale Saliency of 3D Colored Meshes

Anass Nouri, Christophe Charrier, Olivier Lézoray

► **To cite this version:**

Anass Nouri, Christophe Charrier, Olivier Lézoray. Multi-Scale Saliency of 3D Colored Meshes. International Conference on Image Processing, IEEE, Sep 2015, Québec, Canada. hal-01162996v1

HAL Id: hal-01162996

<https://hal.science/hal-01162996v1>

Submitted on 12 Jun 2015 (v1), last revised 25 Jul 2015 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MULTI-SCALE SALIENCY OF 3D COLORED MESHES

Anass Nouri, Christophe Charrier, Olivier Lézoray

Normandie Université, UNICAEN, ENSICAEN, GREYC UMR CNRS 6072, Caen, France

ABSTRACT

Mesh surface saliency detection is an important preprocessing step for many 3D applications. This paper proposes a novel saliency computation method by the use of a local vertex descriptor in the form an adaptive patch. This descriptor is used as a basis for similarity measurement and integrated into a weighted multi-scale saliency computation. Experimental results demonstrate that the proposed approach achieves competitive and innovative results, especially for 3d colored meshes.

Index Terms— Saliency, colored mesh, patch, multi-scale.

1. INTRODUCTION

Visual saliency is the perceptual information that makes some objects in the scene stand out from their surrounding regions and thus capture human visual attention. A large number of saliency models have been proposed in the past decades for 2D images [1]. In the meantime, there has recently been a strong development of 3D acquisition techniques, and this has led to the widespread acquisition of large amounts of 3D data in the form of 3D meshes. As 2D saliency models have facilitated different saliency-based applications such as salient object detection/segmentation and content-aware image/video retargeting, it is natural to want to dispose of saliency measures for 3D meshes. Indeed, this can be of high value in several human centered visual computing applications, such as adaptive compression [2, 3], smoothing [4], resizing [5], and viewpoint selection [6].

Mesh saliency is a measure that attempts to capture the importance of a point on a 3D surface mesh in a similar way to human visual perception. The literature is much less important than that of 2D saliency, but some important works can be quoted. Lee et al. [2] computed mesh saliency using a center-surround operator on Gaussian-weighted curvatures in a DoG scale space. Leifman et al [6] defined surface regions of interest by combining vertex distinctness (with similarities between SpinImage descriptors [7]) and mesh shape extremities. Wu et al [4] defined mesh saliency by considering both local contrast (with multi-scale similarities between local height maps [8]) and global rarity (using clustering on local contrast features). Tao et al [9] have proposed to over-segment

the mesh into regions using also local height maps [8], and to use manifold ranking to define each region saliency. Song et al [10] have proposed to estimate the saliency in the spectral domain by considering the properties of the log-Laplacian spectrum of the mesh and analyzing its deviations.

In this paper we propose a new multi-scale method for the estimation of the saliency of 3D meshes. It is based on a new local vertex descriptor in the form an adaptive patch descriptor per vertex. This descriptor is used as a basis for similarity measurement and integrated into a weighted multi-scale saliency computation. In addition, the proposed approach can be easily applied to colored 3D meshes, that has never been investigated before. The approach is compared to state-of-the-art methods and shows competitive results.

2. MESH SALIENCY

To define our 3D mesh saliency measure, we first need to construct, for each vertex, a local descriptor (named a patch) that characterizes its local geometric (or colorimetric) configuration on the mesh. Once local patches are defined, similarities between the local features of adjacent vertices are computed and the degree of a vertex provides its single-scale saliency. By varying the size of the local patch, several scales are obtained that are merged using entropy-based coefficients. A mesh \mathcal{M} is represented by a non-oriented graph [11] $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{v_1, \dots, v_N\}$ is the set of N vertices and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ the set of edges. The set of edges is deduced from the mesh faces that connect vertices. To each vertex v_i are associated its 3D coordinates $\mathbf{p}_i = (x_i, y_i, z_i)^T \in \mathbb{R}^3$ and eventually its color $\mathbf{c}_i = (r_i, g_i, b_i)^T \in \mathbb{R}^3$. The notation $v_i \sim v_j$ is also used to denote two adjacent vertices in \mathcal{G} (i.e., $(v_i, v_j) \in \mathcal{E}$).

2.1. Local Surface Patch Descriptor

We describe a definition of adaptive patches that account for local neighborhood configurations. This patch feature descriptor is a height map lying on the vertex tangent plane. It extends recent works [12, 8, 13] with an adaptive patch size to better reflect local geometric configurations. The interest of using 2D patches as 3D local surface descriptors comes from the success of self-similarity with patches in image processing [14]. To build patches at each vertex, we have to construct a

bounded sub-region of the tangent plane and we need to first estimate the latter. Then, the patch is filled with projection heights from neighbor vertices.

2.1.1. Tangent Plane Estimation

For every vertex v_i on the mesh surface, we compute its normal vector $\mathbf{z}(v_i)$ and the directional vectors $\mathbf{x}(v_i)$ and $\mathbf{y}(v_i)$ that correspond to the estimation of the 2D tangent plane at vertex v_i on the mesh. Classically, the PCA of the covariance matrix of the neighbors of v_i in a local sphere around v_i is considered [15]. The vertices contained in a sphere $S_\varepsilon(v_i) = \{v_j \mid \|\mathbf{p}_j - \mathbf{p}_i\|_2^2 \leq \varepsilon\}$ centered at v_i within a radius ε are considered. Let $\bar{\mathbf{p}} = \frac{1}{|S_\varepsilon(v_i)|} \sum_{v_j \in S_\varepsilon(v_i)} \mathbf{p}_j$ and $C(v_i) = \frac{1}{|S_\varepsilon(v_i)|} \sum_{v_j \in S_\varepsilon(v_i)} (\mathbf{p}_j - \bar{\mathbf{p}})(\mathbf{p}_j - \bar{\mathbf{p}})^T$ be respectively the center of gravity of $S_\varepsilon(v_i)$ and the associated covariance matrix. From this matrix, eigenvectors of the covariance matrix are used to estimate both the normal vector $\mathbf{z}(v_i)$ and the 2-directional vectors $\mathbf{x}(v_i)$ and $\mathbf{y}(v_i)$ that form an orthogonal basis of the tangent plane $\mathbf{P}(v_i)$. However the normal vectors of all the vertices can have different directions (outwards and inwards). To guide these normals outwards, we propagate the orientation of the neighboring normals using the minimum spanning tree of the graph [16].

2.1.2. Patch Construction

To construct the patch, points \mathbf{p}_j contained in the sphere $S_\varepsilon(v_i)$ are projected onto the 2D plan $\mathbf{P}(v_i)$ defined by the associated vectors, giving rise to projected points \mathbf{p}'_j . We then construct an adaptive patch, the size of which is not related to the radius ε of the sphere $S_\varepsilon(v_i)$, but to the maximum distance between all the 2D projected points along both x and y axis. The dimensions of the patch are determined by

$$T_d(v_i) = \max_{(\mathbf{p}'_j, \mathbf{p}'_k) \in \mathbf{P}(v_i)} (\|\mathbf{p}'_j - \mathbf{p}'_k\|_2^2) \quad (1)$$

where d represents the x or y coordinates, \mathbf{p}'_j^d the d -th coordinate of \mathbf{p}'_j , and $\|\cdot\|_2$ the Euclidean norm. Thus, the patch at vertex v_i is represented by a rectangle of size $T_x(v_i) \times T_y(v_i)$ centered at v_i . It should be noted that usually a patch is represented by a square of fixed size [12, 8]. This local patch is then divided into $l \times l$ cells and each projected point is affected to the cell the center of which is the closest. Finally, each cell \mathcal{P}_i of the patch is filled with the absolute value of the sum of the projections heights $\sum_{\mathbf{p}'_j \in \mathcal{P}_i} \|(\mathbf{p}_j - \mathbf{p}'_j)\|_2^2$ within one cell. The result is a $l \times l$ patch image describing the shape surrounding the vertex v_i : each vertex is then described by a local feature vector $\mathcal{P}(v_i)$ composed of all the patch cells' values. This approach is robust to poorly shaped triangles, non-manifoldness and surfaces with holes.

2.2. Single Scale Saliency

To compute the single-scale saliency of a vertex, a measure of similarity between its associated patch and the patches associated to its neighbors is required. The similarity assigned to the weight of the edge $(v_i, v_j) \in \mathcal{E}$ is:

$$w_{\mathcal{P}}(v_i, v_j) = \exp \left[\frac{\kappa(v_j) \cdot \|\mathcal{P}(v_i) - \mathcal{P}(v_j)\|_2^2}{\sigma_{\mathcal{P}}(v_i) \cdot \sigma_{\mathcal{P}}(v_j) \cdot \|\mathbf{p}_i - \mathbf{p}_j\|_2^2} \right] \quad (2)$$

where $\mathcal{P}(v_i) \in \mathbb{R}^{l \times l}$ is the vector of accumulated heights into the cells of the patch, $\kappa(v_j)$ is the curvature at the vertex v_j , and $\|\mathbf{p}_i - \mathbf{p}_j\|_2^2$ represents the Euclidean distance between the points of vertices v_i and v_j . We propose to locally compute the scale parameters $\sigma_{\mathcal{P}}(v_i)$. Indeed, using a specific scale parameter for each node allows us to better take into account the local distribution around each grid node. The scale parameter is defined as $\sigma_{\mathcal{P}}(v_i) = \max_{v_k \sim v_i} (\|\mathbf{p}_i - \mathbf{p}_k\|_2)$. We have also tested having the scale parameter based on the difference between the patch vectors and it has been found that the difference between 3D coordinates vertices leads to better results. Finally, the visual saliency of the vertex v_i on the 3D mesh is defined by its degree in \mathcal{G} as:

$$\text{ss-saliency}_{\mathcal{P}}(v_i) = \frac{1}{|v_j \sim v_i|} \sum_{v_i \sim v_j} w_{\mathcal{P}}(v_i, v_j) \quad (3)$$

2.3. Multi-scale saliency

To better consider the hierarchical aspect of human vision, we propose to estimate the saliency at different scales. Indeed, saliency at primary scales will detect finest and cramped details, while higher scales will highlight large regions (see Figure 1 (b-d)). The aim is to detect saliency at different scales to cope with noise, since it is only noticeable at some scales. To apply the concepts of multi-scale to our patch local feature, we observe that the size of the neighborhood $S_\varepsilon(v_i)$ can be used as a scale parameter and increasing the size of the local neighborhood is similar to applying a smoothing filter [15]. So, we vary the radius ε of the sphere S_ε to define the local patch descriptor, and we consider three different neighborhoods. Then, we calculate the single-scale map saliency for each considered neighborhood. Before combining the obtained saliency maps, we calculate the entropy that measures the disorder and the disparity of the saliency information on each map. To do so, at a given scale k , we compute a histogram H^k of the saliency values of the vertices to obtain the probability to have the saliency value i : $P_i^k = H^k(i)/|V|$, where H_i^k gives the number of vertices of saliency i at scale k . Then, the scale-entropy is defined by $E_k = - \sum_i P_i^k * \log P_i^k$. By weighting the saliency of each node by the scale-entropy in the combination of the different scales, we can obtain a robust multi-scale saliency map that

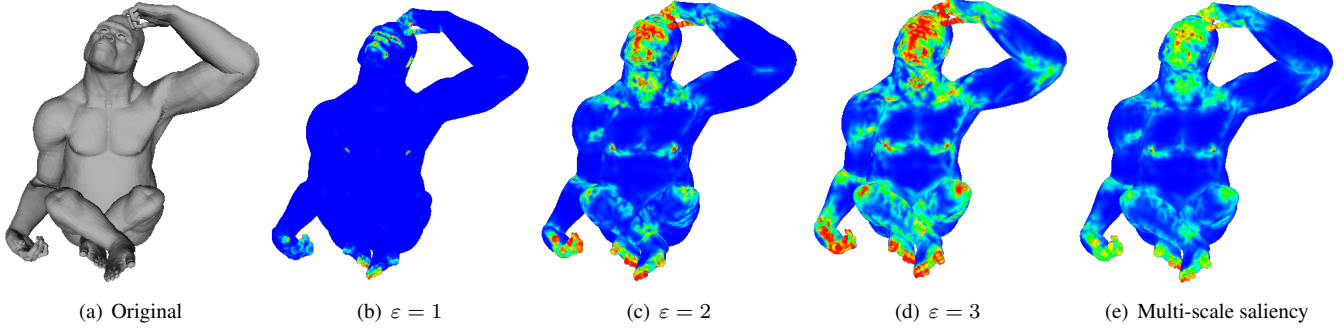


Fig. 1. Influence of the radius ε on single scale saliency computation (b-d), and the multi-scale result (e).

considers the disparity of the saliency at each scale. It is computed as follows:

$$\text{ms-saliency}_{\mathcal{P}}(v_i) = \frac{\sum_{k=1}^K \text{ss-saliency}^k(v_i) \cdot E_k}{\sum_{k=1}^K E_k} \quad (4)$$

where k is the scale index (corresponds to different values of the radius ε in the patch definition), and K the number of scales.

2.4. Extension to 3D colored meshes

In all the previous sections, we have considered 3D meshes: only points \mathbf{p}_i are associated to vertices v_i . However, with recent 2D scanners is now possible to also simultaneously acquire a color per vertex and a RGB color vector \mathbf{c}_i is also associated to the vertices. To extend our multi-scale saliency to 3D colored meshes, we proceed in the following way. A patch is constructed in the same way but patches' cells \mathcal{P}_i are filled-in with the average RGB colors of the projected points $\frac{1}{|\mathbf{p}'_j \in \mathcal{P}_i|} \sum_{\mathbf{p}'_j \in \mathcal{P}_i} \mathbf{c}'_j$, defining a local color patch vector $\mathcal{C}(v_i)$ for each vertex. Edges are then weighted with

$$w_{\mathcal{C}}(v_i, v_j) = \exp \left[-\frac{\|\mathcal{C}(v_i) - \mathcal{C}(v_j)\|_2^2}{\sigma_{\mathcal{C}}(v_i) \cdot \sigma_{\mathcal{C}}(v_j) \cdot l^2} \right] \quad (5)$$

with $\sigma_{\mathcal{C}}(v_i) = \max_{v_k \sim v_i} (\|\mathcal{C}_i - \mathcal{C}_k\|_2)$. From this we can define, similarly than with patch height maps,

$$\text{ss-saliency}_{\mathcal{C}}(v_i) = \frac{1}{|v_j \sim v_i|} \sum_{v_i \sim v_j} w_{\mathcal{C}}(v_i, v_j) \quad (6)$$

and deduce $\text{ms-saliency}_{\mathcal{C}}(v_i)$ as color saliency at one vertex.

3. EXPERIMENTAL RESULTS

In this section we present the benefit of our approach on 3D meshes, we provide a comparison with state-of-the-art approaches and innovative results on 3D colored meshes. In all the experiments, a patch is divided into 27×27 cells. The

number of scales K in the multi-scale saliency computation is fixed to 3. These 3 scales are obtained with spheres $S_{\varepsilon}(v_i)$ of radii ε_0 , $2\varepsilon_0$, and $3\varepsilon_0$ where ε_0 is manually fixed, depending on the considered meshes. To present the computed saliency, we use a color map. Warmer colors (reds and yellows) show high saliency and cooler colors (greens and blues) show low saliency.

First we begin by presenting single-scale saliency results. Figure 1 presents, for an original gorilla mesh, the computed saliencies with three different values of ε . As expected, with small values of ε , only small regions very different of their surrounding ones are considered as salient. When ε increases, the saliency is much more spread out and larger regions are considered as salient. However, in the same time, the saliency also saturates in very salient regions (e.g., eyes of the gorilla). The multi-scale saliency permits to cope with all these problems by efficiently blending all the single-scale results together.

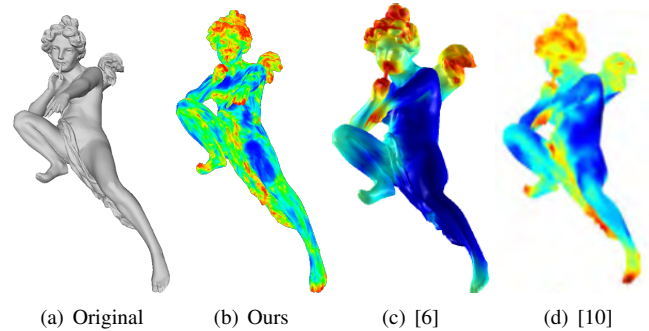


Fig. 2. Comparison of our saliency with state-of-the-art approaches on the 3D 'Angel' mesh.

Second we compare our approach to state-of-the-art ones. We have compared our approach with the reference approach of [6] and the recent one of [10]. Figure 2 presents these comparisons. The 3D object Angel's surface is complex as it contains many extremities. It has also both rough and smooth surfaces. One can see that the extremity of the scarf presents a lot of fluctuations. This one is considered salient by the ap-

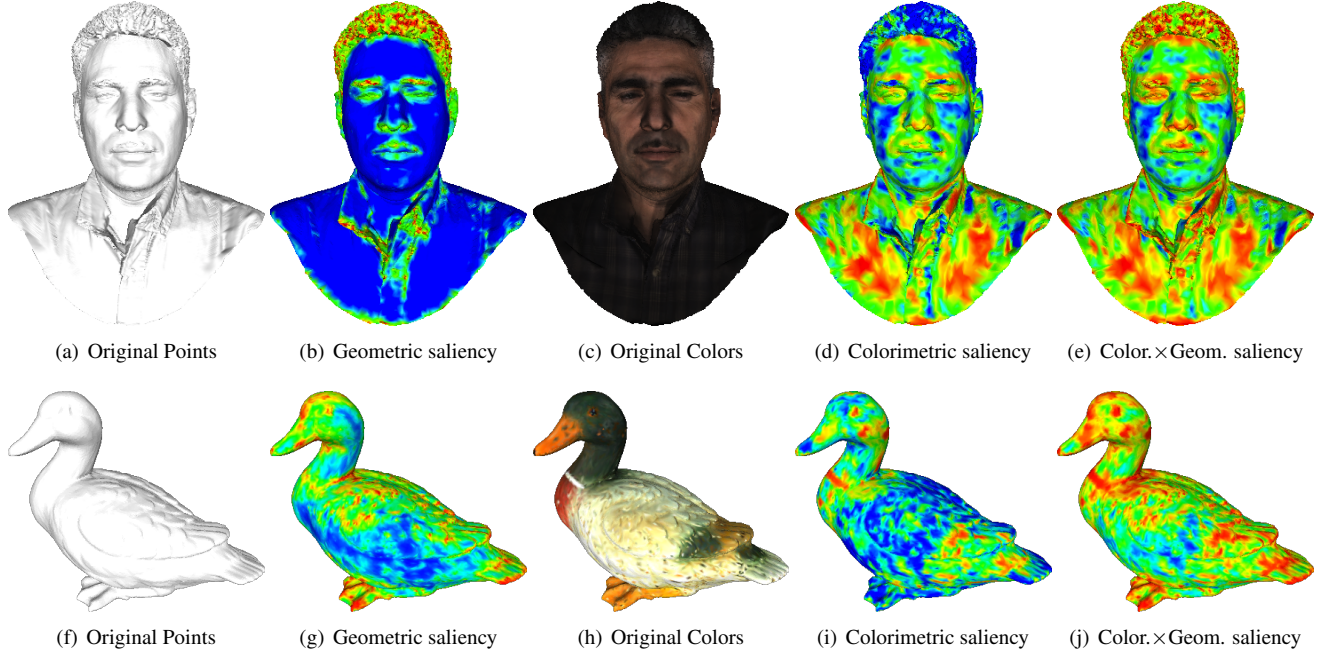


Fig. 3. Saliency of 3D colored meshes.

proach of [10] and ours while the approach of [6] considers it as non salient. The discontinuities on the eyes, the arms, the hip and the stomach are represented as salient by the approach of [10] and ours, contrary to the method of [6]. This comparison shows that our approach provide much more detailed saliency estimation than reference ones that provide a very rough estimation. We believe this is a strong benefit of our approach, and similar results with the other two reference approaches can simply be obtained by strongly smoothing our saliency map.

Third, to assess the accuracy of our approach and its fine detection of saliency, we compare the results we obtain on objects from the SHREC 2007 Watertight Models [17]. In [18], a pseudo saliency ground-truth is provided that was obtained from an analytical model derived from psycho-visual experiments. Figure 4 presents the comparison. One can notice that our detected saliency corresponds very well to the pseudo ground-truth salient regions and this shows the benefit of our approach.

Fourth, we show innovative results that have, to the best of our knowledge, never been investigated before: saliency computation on 3D colored meshes. Figure 3 shows such results on a 3D scan of a human head and of a stuffed duck. Figures 3(b),(g) present the obtained geometric saliencies that take into account only the vertices' coordinates to build patches. As expected flat regions appear as non salient whereas fluctuant ones are considered as salient. Figures 3(d),(i) present the obtained colorimetric saliencies that takes into account only the vertices' colors to build patches. These results are very different from the geometric saliencies since areas hav-

ing strong color variations appear as salient (the very visually salient white collar of the duck is now much salient). Figures 3(e),(j) present a merging of both results as a simple product of both geometric and colorimetric saliencies.

4. CONCLUSION

In this paper we have proposed a simple, yet effective, approach to the difficult problem of estimating saliency in 3D meshes. In order to describe salient regions, local adaptive patches describing the geometric or colorimetric neighborhood configurations are computed at each vertex. This descriptor is used as a basis for similarity measurement and integrated into a weighted multi-scale saliency computation. The approach goes beyond state-of-the-art by quantifying the saliency of 3D *colored* meshes, that was never investigated before.

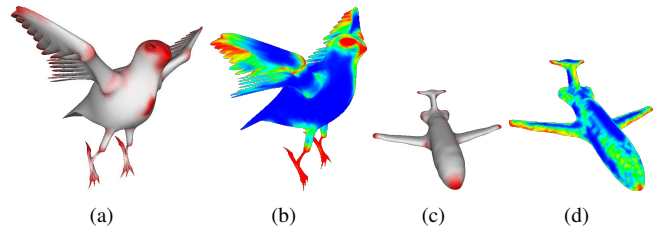


Fig. 4. Comparison with the pseudo ground-truth saliency in [18]. (a),(c) ground-truth saliency, (b),(d) our approach.

5. REFERENCES

- [1] Zhi Liu, Wenbin Zou, and Olivier Le Meur, “Saliency tree: A novel saliency detection framework,” *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 1937–1952, 2014.
- [2] Chang Ha Lee, Amitabh Varshney, and David W. Jacobs, “Mesh saliency,” *ACM Trans. Graph.*, vol. 24, no. 3, pp. 659–666, 2005.
- [3] Yitian Zhao, Yonghuai Liu, Ran Song, and Min Zhang, “A saliency detection based method for 3d surface simplification,” in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2012, pp. 889–892.
- [4] Jinliang Wu, Xiaoyong Shen, Wei Zhu, and Ligang Liu, “Mesh saliency with global rarity,” *Graphical Models*, vol. 75, no. 5, pp. 255 – 264, 2013.
- [5] Shixiang Jia, Caiming Zhang, Xuemei Li, and Yuanfeng Zhou, “Mesh resizing based on hierarchical saliency detection,” *Graphical Models*, vol. 76, no. 5, pp. 355 – 362, 2014.
- [6] George Leifman, Elizabeth Shtrom, and Ayellet Tal, “Surface regions of interest for viewpoint selection,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 414–421.
- [7] Andrew Edie Johnson and Martial Hebert, “Using spin images for efficient object recognition in cluttered 3d scenes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433–449, 1999.
- [8] A. Maximo, R. Patro, A. Varshney, and R. Farias, “A robust and rotationally invariant local surface descriptor with applications to non-local mesh processing,” *Graphical Models*, vol. 73, no. 5, pp. 231 – 242, 2011.
- [9] Pingping Tao, Junjie Cao, Shuhua Li, Xiuping Liu, and Ligang Liu, “Mesh saliency via ranking unsalient patches in a descriptor space,” *Computers & Graphics*, vol. 46, no. 1, pp. 264 – 274, 2015.
- [10] Ran Song, Yonghuai Liu, Ralph R. Martin, and Paul L. Rosin, “Mesh saliency via spectral processing,” *ACM Transactions on Graphics*, vol. 33, no. 1, pp. 6:1–6:17, 2014.
- [11] Olivier Lézoray and Leo Grady, *Image Processing and Analysis with Graphs: Theory and Practice*, Digital Imaging and Computer Vision. CRC Press / Taylor and Francis, 2012.
- [12] Julie Digne, “Similarity based filtering of point clouds,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 73 –79.
- [13] Francois Lozes, Abderrahim Elmoataz, and Olivier Lezoray, “Partial difference operators on weighted graphs for image processing on surfaces and point clouds,” *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 3896–3909, 2014.
- [14] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel, “Image denoising methods. A new nonlocal principle,” *SIAM Review*, vol. 52, no. 1, pp. 113–147, 2010.
- [15] Mark Pauly, Richard Keiser, and Markus H. Gross, “Multi-scale feature extraction on point-sampled surfaces,” *Computer Graphics Forum*, vol. 22, no. 3, pp. 281–290, 2003.
- [16] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle, “Surface reconstruction from unorganized points,” *ACM SIGGRAPH Computer Graphics*, vol. 26, no. 2, pp. 71–78, 1992.
- [17] Daniela Giorgi, Silvia Biasotti, and Laura Paraboschi, “Shrec: Shape retrieval contest: Watertight models track,” <http://watertight.ge.imati.cnr.it/>.
- [18] Xiaobai Chen, Abulhair Saparov, Bill Pang, and Thomas Funkhouser, “Schelling points on 3d surface meshes,” *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 29:1–29:12, 2012.