ROBUST FEATURE LINE EXTRACTION ON CAD TRIANGULAR MESHES

Vincent Vidal, Christian Wolf
Université de Lyon, CNRS, INSA-Lyon, LIRIS, UMR5205, Villeurbanne F-69621, France
{*vincent.vidal, christian.wolf*}*@liris.cnrs.fr*

Florent Dupont
Université de Lyon, CNRS, Université Lyon 1, LIRIS, UMR5205, Villeurbanne F-69622, France
florent.dupont@liris.cnrs.fr

Keywords: Triangle meshes, Feature lines, Crest lines, Feature extraction, Robust detection, Potts model, SVMs.

Abstract: Feature lines are perceptually salient features on 3D meshes. They are of interest for 3D shape description, analysis and recognition. Their detection is a necessary step in several feature sensitive mesh processing applications such as mesh simplification, remeshing or non-photorealistic rendering.
In this paper, an estimator for the angle between tangent plane normals is introduced and a new automatic method is proposed for robust detection of crest lines on 2-manifold triangular meshes, in particular Computer-Aided Design models. The method integrates learning into a global minimization framework favoring geometrically coherent solutions. We study our method in detail and compare it with other methods for the detection of feature edges on 3D meshes. Our comparative results indicate that our method outperforms classical techniques especially in the presence of noise.

1 INTRODUCTION

Feature line extraction consists in finding perceptually salient lines over 3D meshes that a human eye will notice. Detection of feature lines on polygonal surfaces is currently an area of intensive research (Hildebrandt et al., 2005; Kim et al., 2009; Yoshizawa et al., 2005; Zhihong et al., 2009). Feature line extraction has several applications such as surface segmentation (Stylianou and Farin, 2004), shape analysis, matching, and retrieval with geometric query in 3D object databases. Feature line information is also of interest in mesh simplification, remeshing (Attene et al., 2005) or surface smoothing. In non-photorealistic rendering feature lines are highlighted (DeCarlo et al., 2003).

Similar to edge detection in images, feature line extraction on 3D surface meshes is an ill posed problem due to the lack of information on the sampling process, the noise process and the signal geometry. According to the situation at hand, the same geometrical discontinuity might be due to a "true" feature in the object, to insufficient sampling in a region of high curvature, or to noise in vertex positions. Additional difficulties in feature line extraction process are its sensitivity to irregular mesh connectivity when extracted feature lines are composed of mesh edges, the use of a parametric model (e.g snakes) when feature line segments are independent of the mesh connectivity (Lee and Lee, 2002; Pauly et al., 2003), and the selection of significant feature lines at the right scale.

Most of existing works tackle the problem of feature line extraction over dense triangulated meshes. Some works deal with multi-scale feature extraction (Lee et al., 2005; Pauly et al., 2003) but remain oriented more towards the detection of feature vertices than feature lines. For smoothly varying natural meshes, the majority of approaches are based on robust computation of principal curvature extrema (Hildebrandt et al., 2005; Ohtake et al., 2004; Yoshizawa et al., 2005; Zhihong et al., 2009) and require the estimation of curvature derivatives. For Computer-Aided Design (CAD) meshes, feature line extraction often involves identifying all feature edges using either the *tensor voting theory* (Kim et al., 2009), dihedral angles or angles between best polynomial fits (Hubeli et al., 2000).

In this paper, we propose a new robust method for feature line extraction over 2-manifold triangular CAD meshes. Our main contributions are the follow-

ing:

- introduction of a robust estimator for the angle between tangent plane normals (in section 3),
- design of relevant geometric features for feature edge learning (in section 4),
- globally consistent detection of feature edges over triangular surface meshes (in section 5).

The remainder of this paper is organized as follows. In section 2 relevant definitions are presented. Section 3 describes the computation of our estimator for the angle between tangent plane normals. Section 4 tackles feature edge learning and gives a feature vector characterizing feature edges. In section 5 our method for globally consistent feature line extraction is detailed. In section 6 experimental results are analyzed. Finally, section 7 concludes the paper and presents some future work.

2 DEFINITIONS

The following definitions hold for 2-manifold triangular meshes with borders.

Dihedral angles are angles between the two normals of two adjacent triangles to an edge. θ_e denotes the dihedral angle at edge e. The *angle between tangent plane normals* is defined as the angle between the two normals of two estimated tangent planes to an edge. The estimation of tangent planes depends on a scale parameter. $\hat{\theta}_e^r$ stands for an estimation of the angle between tangent plane normals at edge e and at scale r. Details on its computation are given in section 3. A mesh edge e is a *feature edge* if e is a boundary edge or if there is a high discontinuity of the normal direction along e. A *boundary edge* is shared by one triangle and its dihedral angle and estimator for the angle between tangent plane normals are equal to π. A mesh edge e is a *normal edge* if e is not a feature edge.

3 ESTIMATING THE ANGLE BETWEEN TANGENT PLANE NORMALS

This section introduces the computation of the angle between the two normals of two estimated tangent planes to an edge. A local tangent plane fitting (first order approximation) is proposed and the angle between the two unit normals of tangent planes is computed afterwards.

3.1 Robust Estimation of Tangent Planes

The estimation of the local tangent planes is obtained through region growing. However, it is not a classical region growing in sense that a new triangle can be added to one region even through it is not adjacent to any region triangles. Moreover, once a triangle is added to a local region, it remains in that region till the region growing ends.

Robustly estimating the two neighboring tangent planes involves:

- taking into account a scale parameter resulting in a bounding sphere (centered at the edge midpoint) in which a plane is estimated,
- the detection of outlier triangles due to noise,
- the identification of additional discontinuities in the bounding sphere due to multiple region changes (e.g. at additional feature lines close to the edge or at areas of high curvature).

As in classical fitting problems, the plane parameters (such as the normals) depend on the classification of the triangles into inliers, outliers and out of region triangles (see figure 1), and the latter classification depends on the plane parameters.

3.2 Our Approach

This section describes our approach for the computation of the two unit normal estimates of tangent planes to an edge. The final tangent planes are forced to go through the edge midpoint.

Algorithm Description. At the beginning of the algorithm, the two triangles adjacent to the edge are added to their corresponding regions. Thus, first estimates of tangent plane normals are set to corresponding first region triangle normals. Then, a greedy traversal of triangles within the bounding sphere is done: triangles closer to the edge separating plane (see figure 1) are tested first for addition (using a priority queue with euclidean point-to-edge separating plane distance as priority). That permits to get an estimation of tangent plane normal which depends on the close-to-far away triangle order of traversal. The left and right regions, and estimated normals of tangent planes are updated as follows:

- triangles, whose signed distance from its barycenter to the separating plane is positive (resp. strictly negative), are candidates for right (resp. left) region;
- if a candidate triangle has a normal close to the currently estimated plane normal (the angle in be-

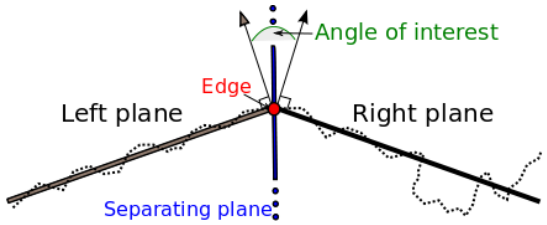


Figure 1: Estimator for the angle between tangent planes: dotted lines represent the mesh surface (with inliers, outliers and out of region triangles). The estimation of left and right tangent planes depends on a bounding sphere centered at the edge midpoint.

tween normals must be less than a threshold), it is added to the region;
- after each addition of triangle in a region, the estimated normal of the tangent plane associated with this region is updated as the area weighted averaging of normals over the associated triangles;
- when the region growing ends, all previously rejected triangles are tested again for acceptance in the same region: the plane normal has changed during the region growing and thus some previously rejected triangles may be accepted this time.

4 LEARNING FEATURE EDGES

To make the edge classification process more robust, it is based on several measures and on complex interactions. Learning a statistically representative model of feature edges from ground-truth data permits to do that. In this paper, Support Vector Machines (SVMs), which is a supervised learning method, are used. More details on SVMs can be found in (Vapnik, 1998). In the following, the feature vector associated with an edge for learning is presented.

Feature Vector Construction. Let F_e be a feature vector associated with an edge e, i.e. it characterizes e in terms of geometrical and contextual measures. A small number of features for F_e is selected, e.g. less than 50 features, according to their F-score defined in (Chen and Lin, 2006). The larger the F-score is, the more likely this feature is discriminative.

The four first features in feature vectors F_e are the following measures:

- cosine of the dihedral angle $cos(\theta_e)$,
- cosine of the angle between the unit normals of the vertices opposite to the edge (-1 is taken for border edges),
- mean of the two edge vertices' minimum of cosines of their adjacent edges' dihedral angles,
- mean of the two edge vertices' *normal variation* (Lee and Lee, 2002).

Then, three scale-dependent features are added into F_e (depends on a radius r):

- cosine of the angle between tangent plane normals $cos(\hat{\theta}_e^r)$ ($\hat{\theta}_e^r$ is defined in section 2),
- five curvature measures such as the mean of edge vertices curvature (*Kmin* and *Kmax* principal curvatures (Cohen-Steiner and Morvan, 2003), Gaussian curvature: $Kmin * Kmax$, mean curvature: $0.5(Kmin + Kmax)$, $Kmax - Kmin$),
- mean of the two edge vertices' *similarity measure* (Lee et al., 2005).

r is set as a percentage of the mesh bounding box minimal dimension (min BB). The following scales are worth identifying feature edges in the presence of noise and are thus included in F_e:

- for the estimator for the angle between tangent plane normals: r = 4, 7, 10, 13, 16, 19, 22, 25, 28 % of min BB,
- for principal curvatures and similarity measure: r = 1, 5, 9, 13, 17 % of min BB.

F_e is composed of 43 features. Each feature in F_e is normalized to $[-1,1]$.

5 GLOBALLY CONSISTENT FEATURE EDGE DETECTION

In the previous section, feature edge detection has been improved by selecting more measures for characterizing them and by learning a prediction model capable of taking into account all informative features in the training set. However, all final decisions are local for each edge. Local decisions suffer from not taking into account the label distribution of neighboring edges. Conversely, global decisions over all mesh edges may improve the detection result quality by favoring consistent labelling.

Post-processing can introduce a slight dependence on the neighboring edges' results. However, this is far from optimal, as the true circular dependence is not taken into account: the result for each edge depends on the results of neighboring edges and vice versa.

A more powerful solution is to integrate the full set of dependencies directly into the decision process and to combine it with the local geometrical measures. This can be achieved by minimizing a global objective function (energy) over all edge labels, which gives lower scalar values for better solutions (labelling), both in terms of geometry and consistency:

$$\tilde{\omega} = \underset{\omega}{\operatorname{argmin}}\, E(\omega, \Theta)$$

$$E(\omega,\Theta) = \sum_{e} \underbrace{E_d(\omega_e)}_{unary\ term} - \mu \sum_{\{e,e'\}\in N} \underbrace{E_h(\omega_e,\omega_{e'},\Theta)}_{pairwise\ term} \quad (1)$$

ω (resp.Θ) is the set of all edge label variables (resp. edge geometrical measures). N is the set of ordered pairs of adjacent edges. The unary term $E_d(\omega_e)$ (cf. equation 2) is a data term which evaluates the local performance of assigning the binary label ω_e to the edge e. Binary labels 0 and 1 correspond to normal edge and feature edge. The pairwise term $E_h(\omega_e,\omega_{e'},\Theta)$ (cf. equation 3) is devoted to region homogeneity and is based on an improved Ising/Potts model (Geman and Geman, 1984) which locally encourages consistent region labellings. The positive weight μ sets the relative strength of the regularizing pairwise term compared to the unary term.

The data term is given as follows:

$$E_d(\omega_e) = exp((-1)^{\omega_e} d(e)) \quad (2)$$

$E_d(\omega_e)$ directly incorporates classification and confidence of the SVM — which depends on the obtained prediction model — through the signed distance d(e) of the mapped feature vector associated with edge e to the separating hyperplane (in feature space). It is positive for predicted feature edges and negative otherwise.

The regularizing pairwise term is given as:

$$E_h(\omega_e,\omega_{e'},\theta_e,\theta_{e'},T_{ee'}) = \begin{cases} 0 & \text{if } w_e \neq w'_e \\ \beta & \text{if } w_e = w'_e = 0 \\ \alpha_{ee'}(\theta_e,\theta_{e'},T_{ee'}) & \text{else} \end{cases} \quad (3)$$

β is a positive constant which controls the homogeneity of normal edges. $\alpha_{ee'}$ (cf. equation 4) is a positive functional which favors consistent labelling of aligned (within a tolerance) neighboring edges if they have similar dihedral angles.

$$\alpha_{ee'}(\theta_e,\theta_{e'},T_{ee'}) = exp\left\{ -\lambda \left(\underbrace{\frac{|cos(\theta_e)-cos(\theta'_e)|}{\sigma}}_{similarity} + \underbrace{1-cos(T_{ee'})}_{alignment} \right) \right\} \quad (4)$$

In equation 4, λ is a positive constant which monitors the neighboring edge alignment tolerance. The greater λ is, the more two neighboring edges must be aligned. σ handles the tolerated variance of $\lambda|cos(\theta_e)-cos(\theta'_e)|$ along feature lines. $T_{ee'}$ is the *turning angle* (tangential angle) between e and e'. The similarity term avoids detecting adjacent feature edges with a very different dihedral angle. The alignment term predominates over similarity term, for favoring non-jagged feature line extraction. Note that the 00 and 11 labelling cases are favored compared to 01 and 10 cases: therefore homogeneous labellings are preferred in region composed of a majority of normal edges and along feature lines to fill in some gaps. Mathematically speaking, this intuitively explained notion is called sub-modularity. It allows to calculate the exact global minimum of (1) using graph cuts technique (Kolmogorov and Zabih, 2004).

6 EXPERIMENTATION AND DISCUSSION

In this section, the mesh data set is presented as well as the chosen method for comparing edge classification approaches. Then, experiment parameters are defined. Finally, experimental results are analyzed.

6.1 CAD Mesh Data Set

Our data set is made of 181380 edges grouped into 18 CAD mesh models (9 normalized meshes with their 9 noisy versions obtained by adding 0.5% Gaussian noise intensity to mesh vertex positions). The normalization process consists in a mesh scaling to make the maximal dimension of a mesh bounding box equal to one. 0.5% refers to the bounding box maximal dimension.

Among 181380 edges there are 4468 ground-truth feature edges (cf. table 2). All meshes have feature lines composed of mesh edges. Ground-truth for the 9 non-noisy meshes has been manually set, and noisy meshes ground-truth has been deduced from it. Ground-truth is needed for learning feature edges and permits to quantitatively compare all methods presented in this paper.

6.2 Comparison by ROC Curves

Comparison between two feature-normal edge classification methods is done using Receiver Operating Characteristic (ROC) curves. A ROC graph depicts relative trade-offs between benefits (true positive on the Y axis) and costs (false positives on the X axis) (Fawcett, 2006). To determine the best method over an interval, the Area Under Curve (AUC) of the ROC curve in that interval is computed.

Table 1: Minimum, maximum, mean and standard deviation of average of f-scores (Chen and Lin, 2006) computed as follows: Firstly, features' f-scores are computed for each mesh model (except sphere and torus). Then, the mean of these features' f-scores are computed for non-noisy and noisy mesh models. Finally, some statistics (min, max, mean and standard deviation) are computed on similar subsets of features (i.e. estimator for the angle between tangent planes, curvatures, and first four features in F_e).

Features in F_e	Models	Min	Max	Mean	sdev
$\tilde{\theta}_e^r$	non-noisy	9.196	1249.950	277.934	523.248
$\hat{\theta}_e^r$	noisy	2.618	3.202	2.910	0.208
curv.	non-noisy	0.143	0.699	0.360	0.152
curv.	noisy	0.056	0.456	0.238	0.127
4 first	non-noisy	0.533	1428.810	357.705	714.068
4 first	noisy	0.405	2.569	0.969	1.068

6.3 Experimental Settings

Tangent Plane Estimation. A triangle is accepted into a region during the region growing process if the angle between its normal and the currently estimated normal of the tangent plane is less than 23 degrees.

Feature Vector. The estimator for the angle between tangent planes $\hat{\theta}_e^r$ has a quite high informative power as can be seen in table 1. It is much more significant than curvature measures according to f-scores (Chen and Lin, 2006) and than the four first features in the feature vector for noisy meshes. However, for non-noisy mesh models, the four first features are slightly more informative. By increasing the size of the support, the angle between tangent plane normals has become more robust to the presence of noise, at the cost of being less sensitive to small features in non-noisy data.

Learning Feature Edge. For learning with SVMs, the LIBSVM library (Chang and Lin, 2001) has been used with RBF kernel. The best model hyper-parameters have been selected using grid search with cross-validation and maximization of the AUC which is able to cope with unbalanced training datasets.

All duplicated entries in the training set have been removed and a subsampling taking at maximum 5000 training samples per mesh model has been done. The subsampling process keeps the class distributions unchanged. To compensate for unbalanced training data, the error weighting factor associated with the feature edge class is set 9 times greater than the one used for normal edge class.

Globally Consistent Feature Edge Detection. To evaluate the benefits of the global minimization of equation 1 (Potts model) and those of the learning term separately, a data term depending on the edge dihedral angle alone is proposed:

Table 2: Statistics for mesh models used for edges classification (a: non-noisy models ; b: models with Gaussian noise GN): nb. of edges, nb. of feature edges, Area Under Curve (%) for 4 methods: thresholding, hysteresis thresholding, globally consistent edge detection with data term based on dihedral angle, and with data term based on SVM.

Mesh	#t.e.	#f.e.	thres.	Hys.	dih+G	SVM+G
1232_joint	9024	660	100.0	100.0	100.0	100.0
cone	14850	50	100.0	100.0	99.8	100.0
cup	17010	381	99.7	99.7	99.7	98.2
cut_cone	864	144	100.0	100.0	100.0	100.0
cylinder	3540	40	100.0	100.0	100.0	100.0
fandisk	19479	743	99.9	99.9	100.0	97.3
screw	3723	216	100.0	100.0	100.0	100.0
sphere	14700	0	-	-	-	-
torus	7500	0	-	-	-	-
total	90690	2234	99.9	99.9	99.9	99.4

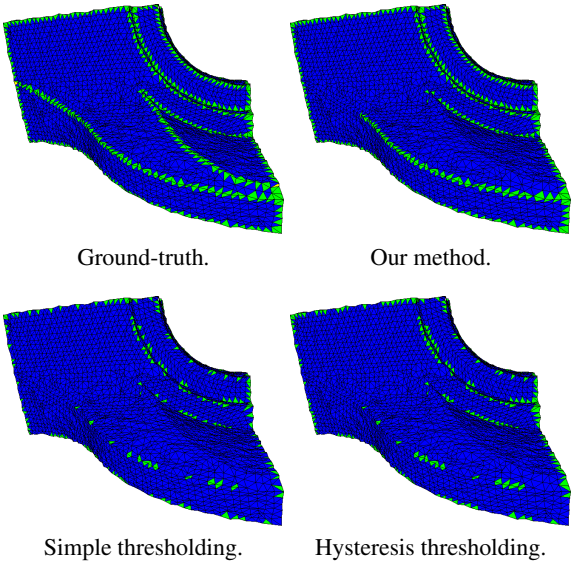
a) Non-noisy mesh models.

Mesh+GN	#t.e.	#f.e.	thres.	Hys.	dih+G	SVM+G
1232_joint	9024	660	93.1	92.5	90.3	96.9
cone	14850	50	93.0	92.8	94.1	93.9
cup	17010	381	98.2	98.8	97.3	95.0
cut_cone	864	144	100.0	100.0	99.3	100.0
cylinder	3540	40	99.8	99.8	99.8	100.0
fandisk	19479	743	99.4	99.6	99.1	98.3
screw	3723	216	97.8	97.9	100.0	98.6
sphere	14700	0	-	-	-	-
torus	7500	0	-	-	-	-
total	90690	2234	97.3	97.3	97.1	97.5

b) Noisy mesh models.

$$E_d(w_e,\theta_e) = \begin{cases} (2\theta_{true} - |\theta_e|)^2 & \text{if } w_e = 1 \\ \theta_e^2 & \text{otherwise} \end{cases} \quad (5)$$

$E_d(w_e,\theta_e)$ is an even positive smooth function over θ_e. Note that $E_d(0,\theta_e) = E_d(1,\theta_e)$ when $|\theta_e| = \theta_{true}$ ($\theta_{true} \in [0,\pi[$). The chosen parameters for the pairwise term have been set as a good trade-off between all models parameters using grid search. For the SVM based data term (cf. equation 2), we experimentally set μ to 0.1 (cf. equation 1), β to 0 (cf. equation 3), λ to 15 and σ to 10 (cf. equation 4). For dihedral angle based data term (cf. equation 5), we experimentally set μ to 2 (cf. equation 1), β to 10^{-3} (cf. equation 3), λ to 15 and σ to 10 (cf. equation 4). To generate the ROC curves of equation 1 with SVM based data term, the bias/constant term of the prediction model vary in $[-2,1]$ (200 samples at all). To generate the ROC curves of equation 1 with dihedral angle based data term (cf. equation 5), θ_{true} vary in $[0,\pi]$ (200 samples at all).



Ground-truth. Our method.

Simple thresholding. Hysteresis thresholding.

Figure 2: Noisy fandisk model: feature line extraction obtained for a false positive rate of 10^{-4}: as can be seen our method is capable of extracting most of feature lines with almost no false positives.

6.4 Experimental Results and Discussion

Results are presented in table 2 and in figure 2.

Thresholding. As can be observed, thresholding techniques based on dihedral angle work quite well for non-noisy classical CAD models. However, they are very sensitive to the presence of noise as can be seen for noisy cone and noisy 1232_joint models in table 2. Indeed, due to noise the false positive rate rises quickly and the AUC is decreased.

Globally Consistent Feature Edge Detection. Results for the global model (cf. equation 1) with data term set from equation 5 are similar to those obtained for thresholding techniques (cf. table 2). Results with data term set from learning (cf. equation 2) are better than all other methods for noisy mesh models. However, they are slightly worse than other techniques for non-noisy mesh models. The robustness against noise results in less sensitivity too small geometric features.

7 CONCLUSIONS AND FUTURE WORK

We have introduced the estimator for the angle between tangent planes which has a high informative power, in particular for noisy mesh models. Moreover, we described a globally consistent feature line extraction technique over 2-manifold triangular meshes: it globally finds the best solution taking into account all edge dependencies with the edge local features (either based on dihedral angle or learnt). This globally homogeneous decision makes the classification results less sensitive to the presence of noise. As future work, we will investigate automatic parameter learning for parameters in equation 1.

ACKNOWLEDGEMENTS

This work has been supported by the French National Research Agency (ANR) through MDCO program (project MADRAS No. ANR-07-MDCO-015).

REFERENCES

Attene, M., Falcidieno, B., Rossignac, J., and Spagnuolo, M. (2005). Sharpen & Bend: Recovering curved sharp edges in triangle meshes produced by feature-insensitive sampling. *IEEE TVCG*, 11(2):181–192.

Chang, C.-C. and Lin, C.-J. (2001). *LIBSVM: a library for support vector machines*. Software available at http://www.csie.ntu.edu.tw/˜cjlin/libsvm.

Chen, Y.-W. and Lin, C.-J. (2006). *Combining SVMs with various feature selection strategies*, volume 207 of *Studies in Fuzziness and Soft Computing*, pages 315–324. Springer.

Cohen-Steiner, D. and Morvan, J.-M. (2003). Restricted delaunay triangulations and normal cycle. In *SCG '03: Proceedings of the nineteenth annual Symposium on Computational Geometry*, pages 312–321. ACM.

DeCarlo, D., Finkelstein, A., Rusinkiewicz, S., and Santella, A. (2003). Suggestive contours for conveying shape. *ACM TOG*, 22(3):848–855.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letter*, 27(8):861–874.

Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE TPAMI*, 6(6):721–741.

Hildebrandt, K., Polthier, K., and Wardetzky, M. (2005). Smooth feature lines on surface meshes. In *SGP '05: Proceedings of the third Eurographics Symposium on Geometry Processing*, page 85. Eurographics Assoc.

Hubeli, A., Meyer, K., and Gross, M. (2000). Mesh edge detection. *Workshop Lake Tahoe*.

Kim, H. S., Choi, H. K., and Lee, K. H. (2009). Feature detection of triangular meshes based on tensor voting theory. *Computer-Aided Design*, 41(1):47–58.

Kolmogorov, V. and Zabih, R. (2004). What Energy functions can be Minimized via Graph Cuts? *IEEE TPAMI*, 26(2):147–159.

Lee, C. H., Varshney, A., and Jacobs, D. W. (2005). Mesh saliency. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*, pages 659–666. ACM.

Lee, Y. and Lee, S. (2002). Geometric snakes for triangular meshes. *Computer Graphics Forum*, 21(3):229–238.

Ohtake, Y., Belyaev, A., and Seidel, H.-P. (2004). Ridge-valley lines on meshes via implicit surface fitting. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers*, pages 609–612. ACM.

Pauly, M., Keiser, R., and Gross, M. (2003). Multi-scale feature extraction on point-sampled surfaces. In *Computer Graphics Forum*, volume 22, pages 281 – 289.

Stylianou, G. and Farin, G. (2004). Crest lines for surface segmentation and flattening. *IEEE Transactions on Visualization and Computer Graphics*, 10(5):536–544.

Vapnik, V. N. (1998). *Statistical Learning Theory*. Wiley-Interscience.

Yoshizawa, S., Belyaev, A., and Seidel, H.-P. (2005). Fast and robust detection of crest lines on meshes. In *SPM '05: Proceedings of the 2005 ACM Symposium on Solid and Physical Modeling*, pages 227–232. ACM.

Zhihong, M., Guo, C., and Mingxi, Z. (2009). Robust detection of perceptually salient features on 3d meshes. *The Visual Computer*, 25:289–295.