Evaluation of a Spectral Data Transformation Method for Meaningful Mesh Segmentation

Corentin DUMERY - April 2020

I. ABSTRACT

We define a data transformation method based on existing literature, and use well-known clustering algorithms to generate a simple 3D segmentation. A human-generated ground truth segmentation is used to evaluate the segmentation results. Consequently, we are able to measure the improvement brought by the data transformation. Experiments are then performed to determine in which conditions this spectral transformation yields the best results.

II. INTRODUCTION

The goal of 3D segmentation is to compute a representation of a 3D object that is easier to analyze and more meaningful. Dividing a mesh into relevant chunks is necessary to make data easier to understand and easier to handle. A relevant classification can also be used as a base for further computational analysis through artificial intelligence. In particular, 3D segmentation can be used to detect tumors or to identify objects in a 3D scene.

More generally in data science, clustering is used to divide a set into different clusters. However, well-known algorithms such as KMeans or DBSCAN may not give satisfactory results when directly applied to 3D meshes. Representing a 3D mesh as plain 3-dimensional data and using the Euclidean distance may not yield a meaningful clustering because the human intuition of segmentation relies on other factors than distance. For example, a constant curvature or alignment of the triangles' normals may be a sign of a satisfying clustering.

This study aims to identify how 3D data can be efficiently transformed to perform clustering. We describe a method to represent the input data a in k-dimensional space that encompasses relevant information about the mesh's curvature and distances. Experiments are then performed to analyse the efficiency of this transformation when used in collaboration with various clustering algorithms. Data is visualized in the appropriate form to help understand what makes the described method successful or not. Thanks to this data analysis, we are finally able to conclude on the circumstances in which this method is successful.

III. RELATED WORK

A. Segmentation methods

Segmentation can be performed through various methods. Many approaches have been proposed, based on a convex decomposition based approach [4], randomized cuts [5], watershed segmentation [6], or critical point analysis [7]. This study aims to complement these methods by extracting relevant information from the input data to make segmentation easier. Clustering can be seen as a naive and straightforward way to perform 3D segmentation, and improvements to segmentation quality using clustering methods will help make the data more meaningful to more elaborate approaches.

B. Clustering Metrics

Segmentation evaluation can be performed through comparison with a ground truth. [8] has defined objective metrics to assess clustering quality. The homogeneity score ensures each cluster contains member of a single class in the ground truth, and the completeness score shows all members of a given class are assigned to the same cluster. The harmonic mean of these two scores is called V-measure and aims to objectively evaluate clustering results. [9] describes

an alternative score called Rand index, that measures the similarity of the result and the ground truth.

C. Data transformation

Clustering vertices of 3D meshes may not be the most efficient way to use the 3D data. Concave areas on the mesh tend to be better borders between meaningful regions. Furthermore, the geodesic distance, which is the distance from one point to the other along the mesh's surface, is more meaningful than Euclidean distance. [2] describes a weight definition between two triangles that increases when the triangle normals are not aligned and when the geodesic distance increases. Based on this definition, [1] initializes an association and affinity matrix that are built to make clustering easier. After this transformation, clustering is performed to obtain a meaningful 3D segmentation.

IV. DATASET

In order to accurately evaluate clustering quality, ground truth segmentation is essential. The dataset proposed in [3] provides a wide range of 3D objects and each model has a human-generated ground truth segmentation. An example of human-generated segmentation is shown in Figure 1.
However, two humans may not always produce the same clustering given the subjective aspect of the task. As such, clustering results may be considered successful even without a perfect score on any of the metrics.

V. METHOD

The 3D vertices in the input mesh are not a meaningful representation of the surface. This study aims to describe a data transformation method that will make clustering more efficient.

FIG. 1: Ground Truth clustering for hand model.

A. Dual Graph

In order to represent the input mesh in a meaningful structure, a dual graph is defined. Each triangle of the input becomes a vertex in the dual graph, and each pair of neighboring triangles has an edge connecting them. The weight of these edges is taken to be the face-to-face distance defined in [2]. This definition relies on the geodesic distance on the mesh's surface and on an angular distance defined as follows, where α_{ij} is the angle between the triangles' normals:

$$Ang\_Dist(\alpha_{ij}) = \eta(1 - \cos\alpha_{ij})$$

The value of η can be chosen in $[0, 1]$. This parameter influences the importance of angle distance in the data representation. If η is close to 1 and the angle between the normals is large, meaning this area is concave, the angular distance will also be large. The intuition behind this representation is that neighboring triangles with non-aligned normals are most often in different meaningful regions. This angular distance is therefore used to add a penalty to such triangles and reduce the likeliness of them being clustered together. The weight is then defined as:

$$Weight(dual(f_i), dual(f_j)) = \delta\frac{Geod(f_i, f_j)}{avg(Geod)} + (1 - \delta)\frac{Ang\_Dist(f_i, f_j)}{avg(Ang\_Dist)}$$

The second parameter $\delta \in [0, 1]$ modulates the importance of angular distance over geodesic distance. A value close to 0 will give a bigger importance to concavity during clustering.

B. Transformation

In this study, the number of clusters desired k is assumed to be known. The dual graph is an efficient and meaningful representation of the input data. [1] describes a method to make this data easier to cluster in k regions by representing the original faces in k-dimensions.

First, the distances between every vertices in the dual graphs are computed using an all-par-shortest-path algorithm. We then define an affinity matrix W of size $n \times n$, with n the number of faces:

$$W(i, j) = e^{-dist(i,j)/2\sigma^2}$$

The parameter σ is chosen as the average of pairwise distances, as recommended in [1]. The values in the affinity matrix W vary between 0 and 1, and a value close to 1 indicates that the two faces are separated by a short distance and likely to be clustered in the same region. In particular, the diagonal values $W(i, i)$ are all equal to 1.

It would be possible to perform clustering on W by identifying each row to its corresponding face, thus representing a triangle in n-dimensions. However, it would be inefficient to cluster this data in k regions as n is much greater than k. Dimensionality reduction can be performed to avoid the curse of dimensionality associated with such data. [1] describes a method to compute an association matrix that will represent faces in k-dimensions. A diagonal matrix D is used to compute a normalized affinity matrix N:

$$D_{ii} = \sum_j W_{ij}$$
$$N = D^{-1/2} W D^{-1/2}$$

The method to obtain the best rank-k approximation of N is well known from linear algebra. We denote $\lambda_1 \geq ... \geq \lambda_n$ the eigenvalues of N in increasing order, and $e_1, ..., e_n$ their respective eigenvectors. These eigenvalues and eigenvectors are computed using the Singular Value Decomposition of N. The matrix V is then defined such that the i^{th} column of V is set to $\sqrt{\lambda_i} e_i$. Finally, the matrix $Q = VV^T$ is known to be the best rank-k approximation of N with respect to the Froebenius norm.

Alternatively, we define the columns of another matrix V' simply as e_i for each column instead of $\sqrt{\lambda_i} e_i$. Experiments have shown that V' most often leads to better results than V.

Therefore, the rows of $V' \in \mathbb{R}^{n \times k}$ can be seen as a k-dimensional representation of the n input faces described in the normalized affinity matrix N. This reduction in dimensionality will help to make data more separable. Other work such as [1] has further studied how to use $Q' = V'V'^T$ to generate an initial clustering for K-means, but computing a clustering on Q' of size $n \times n$ can take significantly more time than on V', which is only of size $n \times k$. As a result, we will simply perform clustering on the rows of V' in this study.

VI. EXPERIMENTS

A. Clustering and Scoring

In order to assess how meaningful of a representation V' is, we will perform clustering on its rows and compare the resulting regions with the ground truth. Since this study is not centered around the clustering method itself but rather on the data representation, we will compute several clustering results using well-known algorithms.

The selection of algorithms used in this study aims to include very different approaches to see how well they perform with our data representation. The sklearn project [10] provides various implementations for commonly used algorithms. The number of clusters is chosen to be the same as that of the corresponding ground truth. This study includes K-Means, the most common approach, which iteratively computes new centers for clusters. Agglomerative clustering begins with small clusters and gradually adds elements to these regions. Finally, the BIRCH algorithm is a hierarchical technique that locally

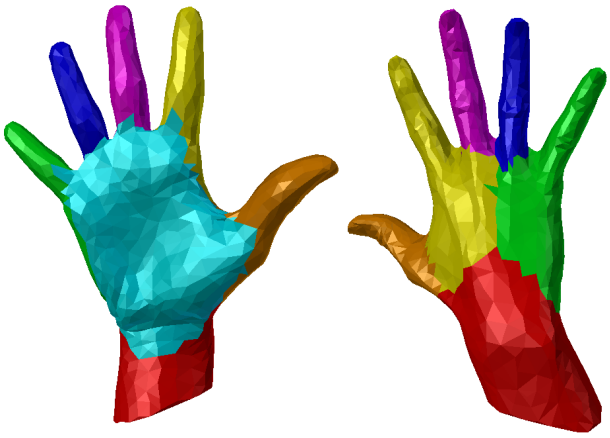


FIG. 2: Post-transformation Clustering results on hand model

computes clusters with a smaller memory cost.

Other clustering methods were also considered. DBSCAN has shown to perform very poorly in our experiments, both on the input and the transformed points. The threshold parameter can be chosen by dichotomy through trial and error, but the philosophy behind DBSCAN makes it unfit for segmentation. In particular, the output clustering will have outliers which will not belong to any cluster. DBSCAN is therefore not used in our experiments. Spectral clustering, on the other hand, was not selected for this study because it's approach is based on a data transformation technique similar to the one we described.

Sklearn [10] also provides useful tools to evaluate clustering quality. In particular, V-measure [8] is an objective score definition based on homogeneity and completeness. The Rand index [9] will provide an additional score to measure similarity between clustering results and a ground truth. Finally, the silhouette score will not be used because the dataset used in this study includes a ground truth segmentation, but this indicator would be interesting otherwise.

B. Results

It remains to evaluate the efficiency of the method we have described. This approach relies on a data transformation that encompasses geodesic and angular distance. Since this information may not always be useful to the segmentation based on the intrinsic meaning of the 3D mesh, we do not expect this method to be improve clustering on all input data. In this section, we will analyze the effect of our data transformation on meshes on which it was beneficial and some on which it was detrimental. This will help us understand in what circumstances and for what objective this clustering technique is useful.

The hand model which was shown in Figure 1 gives satisfying results with our approach. The resulting clusters are shown in Figure 2, and the precise V-Measure and Rand index scores in Figure 3. In particular, the V-Measure of K-means goes from 60% to 68% thanks to our method. We can observe from the segmentation that the strong concave angle between each fingers was accurately taken into account, as there is a border between each finger. However, the segmentation does not always identify the end of the finger. This is due to the fact that normals on the back of the hand are aligned to that of the finger and there is no strong angle separating them. Finally, the palm of the hand was not identified as the same cluster as the back of the hand, as opposed to the ground truth in Figure 1. This choice of segmentation should not be considered an error, but rather a segmentation choice.

Several other 3D segmentations were computed. The results can be seen in Figure 5. Overall, these regions seem more satisfying than a naive clustering. Figure 4 summarizes the V-measure results by showing the range of scores obtained by different algorithms on each model. As expected, the data transformation method improves results on some meshes and deteriorates results on others.

For the bird model, this can be explained by the fact that the top and bottom of each wing are clustered in different regions. This is due to the difference in alignment of the faces' normals, which result in a large distance in the dual graph. Similarly, in the bearing model, opposite sides of a cylinder are not clustered together.

The precise V-Measure and Rand index scores are provided in the appendix. It can be

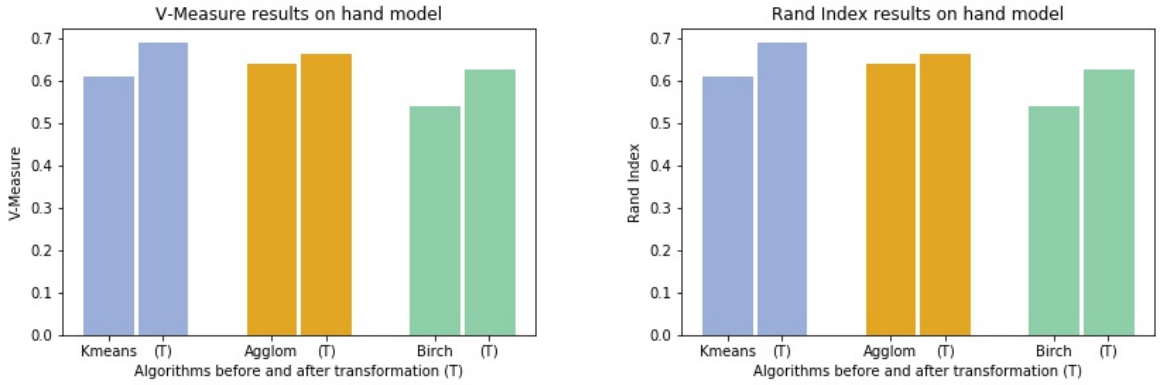


FIG. 3: Clustering results on hand model.

noted that the choice of algorithm matters little, indicating that it is the curvature of the mesh and the desired style of segmentation that will determine the efficiency of the described approach. To further analyse why this method works on some meshes and fails on others, we have represented the contents of some relevant matrices used during the transformation in Figure 6. On the first row, the distance matrix simply represents the initial distance between two faces used in our method. The affinity and association matrices W and Q', in the second and third rows, are defined in the Method section. Clustering on transformed points is performed on V' where $Q = V'V'^T$.

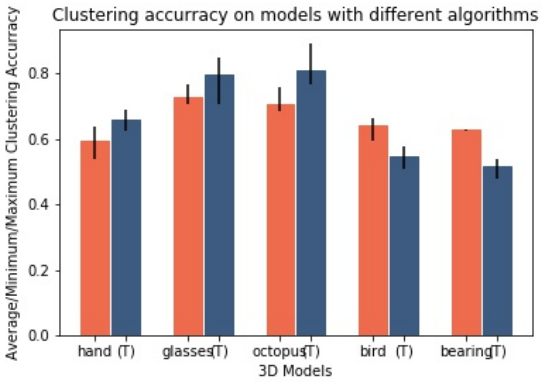


FIG. 4: Minimum/Maximum Clustering Results for each model, without and with transformation (T)

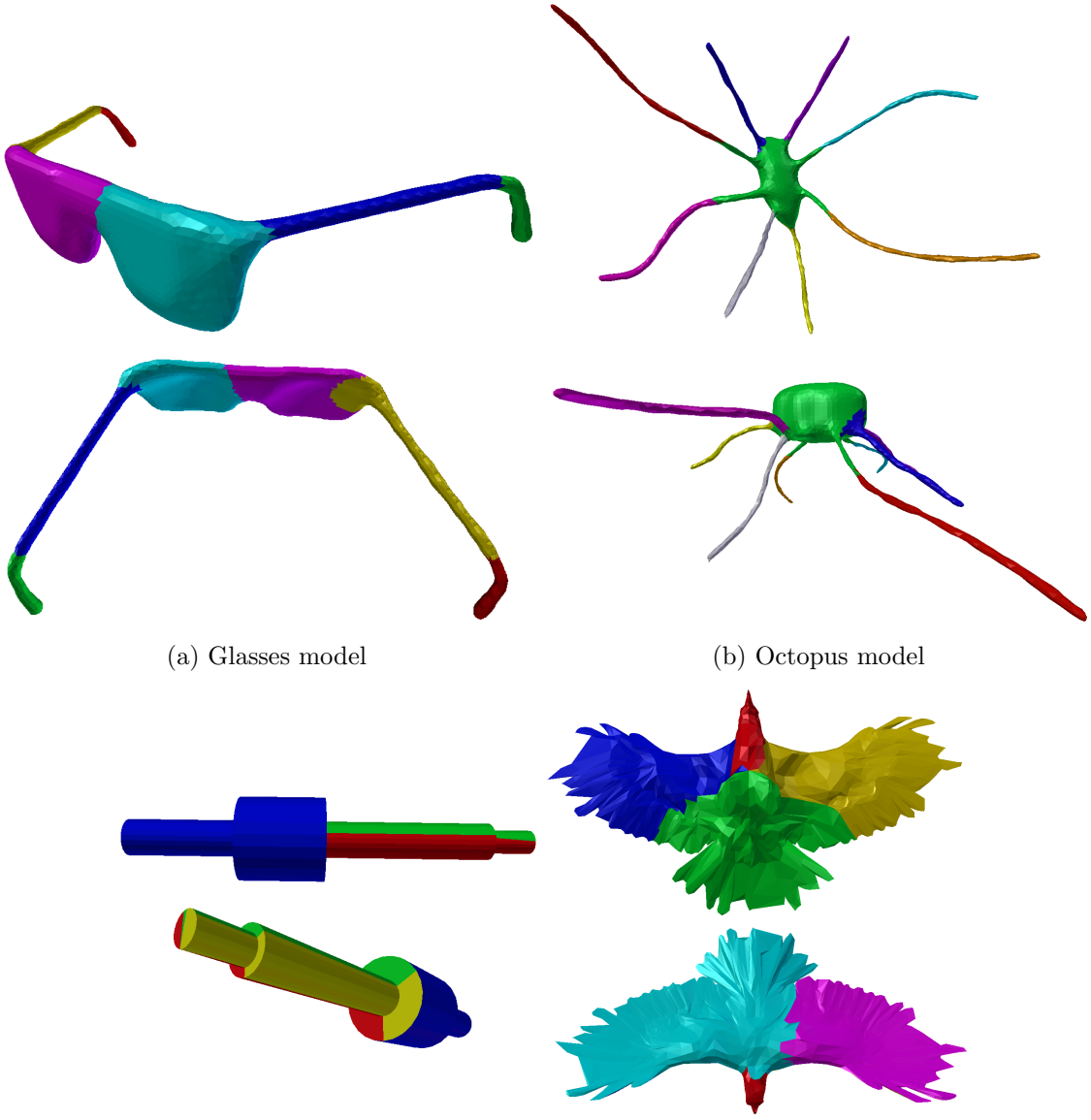
Intuitively, a set is easy to cluster if it is composed of some compact clusters distant from each other. In other words, a minority of pairwise distances should be very small, a majority should be large, and there should be very few values in between. We conclude from Figure 6 that computing a clustering on the association matrix is indeed easier than on the initial distance matrix, as its distribution corresponds to that description of an easy to cluster set. However, our study has shown that models which did not lead to an improvement in results with our transformation often had an association matrix that did not match this description. In particular, the association matrix of the bearing model shown in Figure 6 has many intermediate values. Furthermore, the frequency of elements close to 1, implying they are separated by a small distance and likely to be clustered together, is much larger than on the hand model.

These experiments have shown that the method previously described may not always be beneficial. However, we also notice that the efficiency of our approach can be evaluated with the distribution of the association matrix.

VII. CONCLUSION

In this study, we have described a data transformation method to represent 3D data in a k-dimensional space that encompasses information on the mesh's angular and geodesic distance. This approach is then evaluated by comparing the results of clustering algorithms with and without data transformation.

Our research shows the choice of algorithm



(a) Glasses model

(b) Octopus model

(c) Bearing model

(d) Bird model

FIG. 5: Post-transformation clustering results on example models

matters little, whereas the input mesh can have a strong impact on the efficiency. We identified in which situation the segmentation was not improved by our transformation, and showed how this seems to be characterized by an association matrix with a non-satisfying distribution.

Future research could focus on defining a criteria to determine an estimation of the efficiency of this approach using the distribution of the association matrix. This would be useful when no ground truth is available. Furthermore, it would be interesting to research how this transformation can be used in synergy with more advanced segmentation methods instead of naive clustering. The use of a convex decomposition or a randomized cuts approach would lead to smoother borders between regions.

VIII. REFERENCES

[1] R. Liu and H. Zhang, "Segmentation of 3D meshes through spectral clustering," in 2004, DOI: 10.1109/PCCGA.2004.1348360.

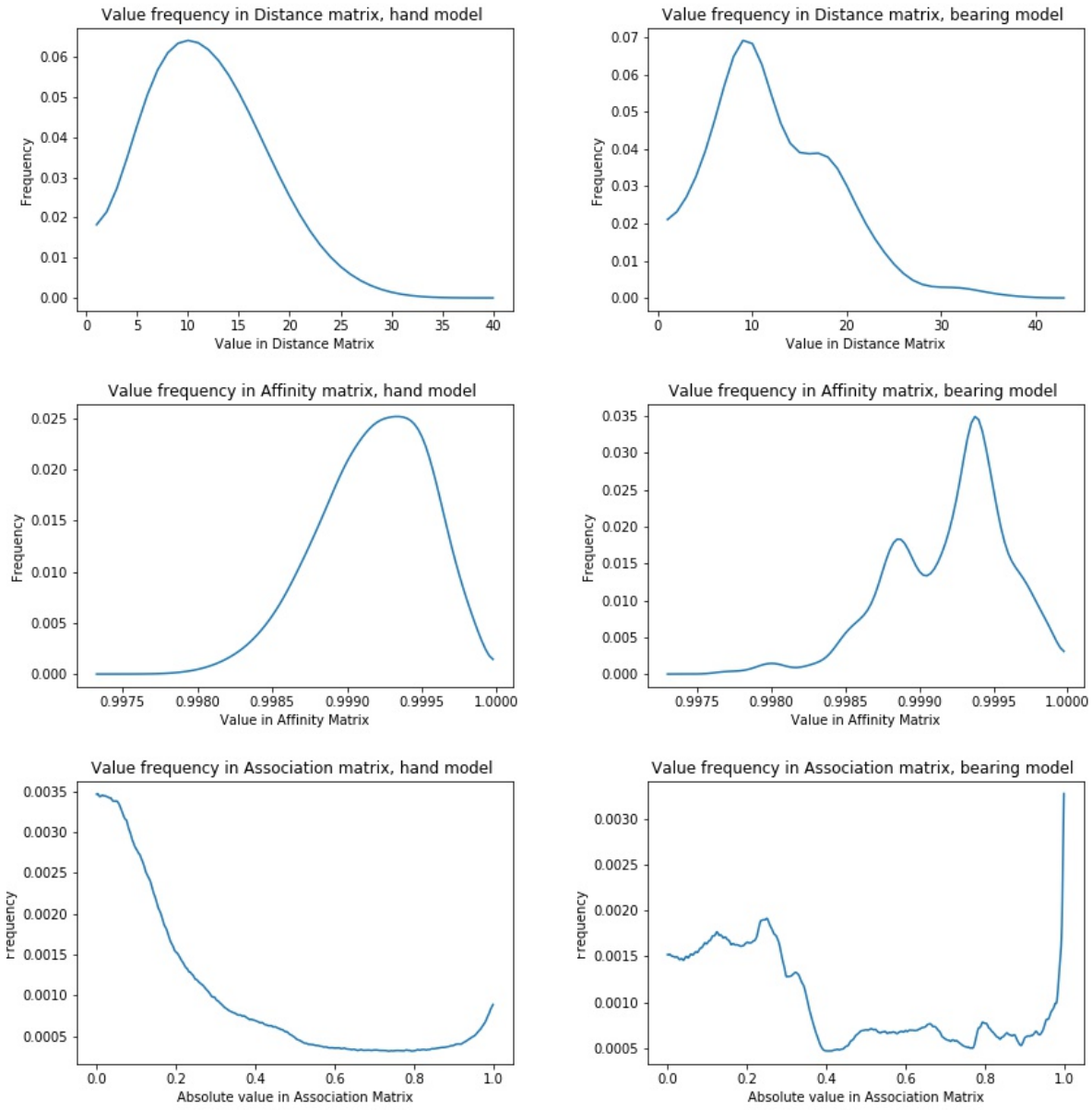


FIG. 6: Visualization of transformation matrices on a successful mesh (Hand, left column) and an unsuccessful mesh (Bearing, right column)

[2] S. Katz and A. Tal, "Hierarchical mesh decomposition using fuzzy clustering and cuts," ACM Transactions on Graphics, vol. 22, (3), pp. 954, 2003.

[3] X. Chen, A. Golovinskiy and T. Funkhouser, "A benchmark for 3D mesh segmentation," ACM Transactions on Graphics (TOG), vol. 28, (3), pp. 1-12, 2009.

[4] B. Chazelle et al, "Strategies for polyhedral surface decomposition: An experimental study," Computational Geometry: Theory and Applications, vol. 7, (5), pp. 327-342, 1997.

[5] A. Golovinskiy and T. Funkhouser, "Randomized cuts for 3D mesh analysis," ACM Transactions on Graphics, vol. 27, (5), pp. 1-12, 2008.

[6] A. P. Mangan and R. T. Whitaker, "Partitioning 3D surface meshes using watershed segmentation," IEEE Transactions on Visualization and Computer Graphics, vol. 5, (4), pp. 308-

321, 1999.

[7] H. S. Lin, H. M. Liao and J. Lin, "Visual Salience-Guided Mesh Decomposition," IEEE Transactions on Multimedia, vol. 9, (1), pp. 46-57, 2007.

[8] A. Rosenberg and J. Hirschberg, "V-measure: A conditional entropy-based external cluster evaluation measure," Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp. 410–420, 2007.

[9] W. M. Rand, "Objective Criteria for the Evaluation of Clustering Methods," Journal of the American Statistical Association, vol. 66, (336), pp. 846-850, 1971.

[10] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp; 2825-2830, 2011.

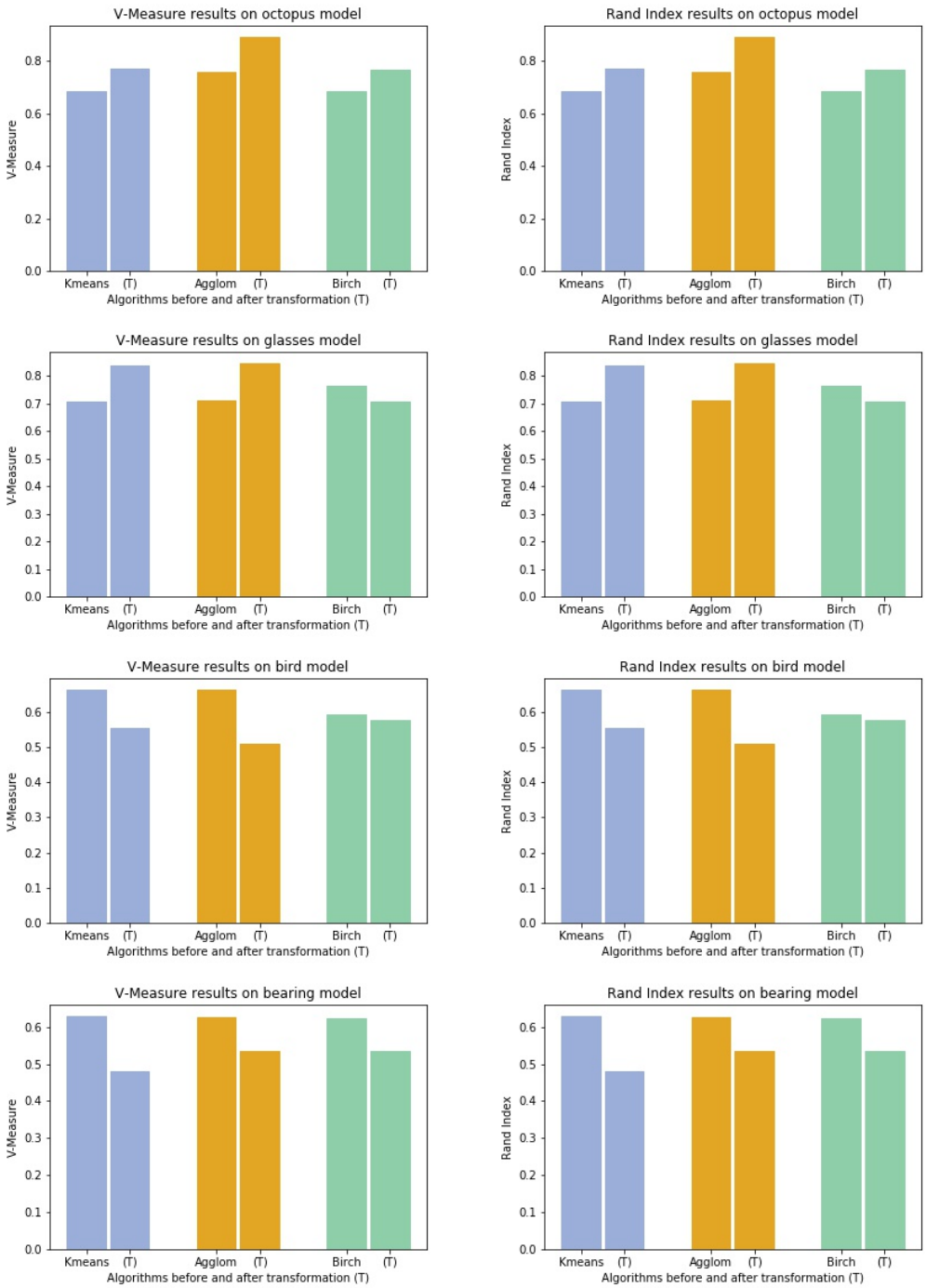


FIG. 7: Appendix - Detailed clustering results on mesh corpus