Proceedings of the 2020 IISE Annual Conference
L. Cromarty, R. Shirwaiker, P. Wang, eds.

Automatic Feature Selection for Shape Registration in Additive Manufacturing

Abstract ID: 790259

Weizhi Lin, Peng Dai, Qiang Huang
University of Southern California
Daniel J. Epstein Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90007 USA
Email: qiang.huang@usc.edu

Abstract

There is a growing importance in characterizing 3D shape quality in additive manufacturing (a.k.a. 3D printing). To accurately define the shape deviation between the designed product and actual build, shape registration of scanned point cloud data serves as a prerequisite for a reliable measurement. However, manual registration is currently heavily involved, for example, in obtaining initial matching of the design and the scanned product based on landmark features. The procedure can be inefficient, and more importantly, introduce potentially large operator-to-operator variations for complex geometries and deformation. Finding a sparse shape correspondence before refined registration would be meaningful to address this problem. In that case, automatic landmark selection has been a challenging issue, particularly for complicate geometric shapes like teeth. In this work we present an automatic landmark selection method for complicated 3D shapes. By incorporating subject matter knowledge (e.g., dental biometric information), a 3D shape will be first segmented through a new density-based clustering method. The geodesic distance is proposed as the distance metric in the revised clustering procedure. Geometrically informative features in each segment are automatically selected through the principal component analysis and Hotelling's T^2 statistic. The proposed method is demonstrated in dental 3D printing application and could serve as a basis of sparse shape correspondence.

Keywords

shape registration, feature extraction, clustering, geodesic distance, Gaussian curvature

1. Introduction

A significant challenge in additive manufacturing (AM) is to predict and control the inevitable shape deviation or distortion between the designed product and the actual print [1]. A critical pre-condition for establishing a credible predictive model is to register the two digital objects so that we can measure the true deviation between them. In other words, the potentially randomly positioned shapes should be properly aligned and then their shape similarity or difference could be identified and further analyzed [2]. In practice, shape registration in AM still involves tedious manual work to a great extent, which is ineffective and expensive when the to-be-registered sample set is large. In addition, different operators and registration methods may introduce appreciable amount of variabilities. It will become even more difficult when it comes to objects with complex geometries and deformations, such as dental models. Different from the shape registration in computer vision, which usually refers to register two partial scans of the same object, shape registration in AM is more about shape alignment between two digital objects of the same type. However, deformations often occur between the designed shape and printed shape due to printing process or material quality. Essentially the problem belongs to the shape correspondence between shapes with non-rigid deformation [3].

In many shape correspondence methods, feature extraction, also called landmarks selection, is the first step to determine a sparse feature correspondence before finding the refined pairwise matching [4], [5]. Some shape correspondence methods are based on given landmarks or manually labeled landmarks [6], [7]. However, unlike human faces, tooth shapes are lack of clearly identifiable landmarks. The structures and shapes of teeth also vary among patients and even within the same patient over different time periods. Various automatic landmarks selection methods have been reported, such as methods based on Gaussian curvature [8], average geodesic distance[9], and heat kernel signature[10]. In dental 3D printing, integrating dental biometrics into feature selection is expected to enhance

the performance of shape correspondence. For instance, an adult commonly has 32 teeth, including 8 incisors, 4 canines (cuspids), 8 premolars (bicuspids) and 12 molars (including 4 wisdom teeth). Each group of teeth have different types of geometries. The shape of incisors is like small chisels, with sharp edges and relatively smaller size than premolars and molars. The tooth edges often have the highest Gaussian curvature on the whole tooth surface. The Canines have sharp and pointy surfaces, and their sizes are similar to incisors. Premolars are bigger than canines and incisors, having flatter surfaces than incisors and with ridges. The molars are the biggest teeth, with complex ridges on their surfaces. Ignoring the intrinsic dental biometric information will lead to the landmark selection biased towards to the group of teeth with high Gaussian curvatures.

In order to increase the efficiency, as well as to reduce the variability of manual landmarks selection in AM, our interest lies in the study of automatically selecting landmarks on tooth surface to ensure a reliable result of shape correspondence afterwards. Following the Introduction, Section 2 presents a landmark select method by integrating dental biometrics and the intrinsic surface geometric characteristics. We propose a new clustering method for shape segmentation, which is based on a density-based clustering method [11] with the Euclidean distance replaced with the more suitable geodesic distance. Section 3 illustrates the developed method with an actual dental model. To verify the effectiveness of the developed methods, simulation study is conducted to compare the informativeness of the selected points to alternative methods. Conclusions are summarized in section 4.

2. Landmarks Selection via Density-Based Geodesic Segmentation

Our landmark selection method includes a novel shape segmentation, principal component analysis (PCA) and Hotelling T^2 statistics to extract points with geometric properties. This method could automatically select points containing adequate geometric information within segments of different shape characteristics and work well for intricate surfaces such as teeth. Being able to automatically determine the number of cluster centers on the surface manifold, the shape segmentation method is built up on a density-based clustering method [11] with the innovation of proposing geodesic distance as the distance metric. Among each cluster, we apply PCA to reduce the data dimension and use Hotelling T^2 statistics to select points with highest value of geometric properties.

2.1. Clustering by Fast Search and Find of Density Peaks on 3D Surfaces

Directly selecting landmarks from the whole dental model surface may lose information in certain areas. For example, if we select points according to local maxima of Gaussian curvature, we will lose representative points for molars due to the difference in tooth types. Thus, it is natural to segment the shape before feature extraction to ensure the selected points equally representing the whole surface. Clustering based segmentation techniques have been widely used in image processing, especially in medical area [12]. For 3D cases, the shape segmentation methods include watershed algorithm [13], hierarchical mesh decomposition[14], and clustering based methods[15]. The ability to automatically segment the shape makes the clustering-based methods the most desirable method for our case. We first tested five clustering algorithms on teeth shape, including K-means, Fuzzy C-means, K-medians, K-centers and Gaussian mixture. The results are shown in Figure 1.

Figure 1: The clustering results on the same teeth shape via different clustering methods with 5 clusters. (a) Fuzzy C-means; (b) Gaussian mixture method; (c) K-centers; (d) K-means; and (e) K-medians algorithm.

Most of the clustering algorithms, for example K-means algorithm[16], require empirical knowledge of the number of cluster centers, which is hard for 3D shape segmentation, especially in dental case. Another concern of these clustering algorithms is that the lack of ability to detect non-spherical clusters [17]. Considering the natural structure

of teeth, which could be categorized to molars, premolars, canine and incisors (the last two share similar structure), as well as the symmetric structure, we set the number of cluster centers to be 5. As shown in the Figure 1, typical clustering methods failed to segment the teeth surface into geometrically meaningful segments. The cluster results of Gaussian mixture method and K-medians (Figure 1(b) and 1(e)) led to confusing boundaries of different clusters. Fuzzy C-means (Figure 1(a)) and K-means algorithm (Figure 1(d)) merely segmented the teeth into 5 even parts, which may be reasonable for the complete teeth set. However, for those who lose one or several teeth, this result may not be fair enough for the purpose of decomposition to equally important parts. And K-centers (Figure 1(c)) method did not even segment the teeth into even segments.

In order to detect non-spherical clusters and automatically choose the correct number of clustering centers, here we adopt the algorithm proposed by Rodriguez and Laio [11], named clustering by fast search and find of density peaks of data points (CFSFDP). The basic assumptions are that cluster centers are surrounded by neighbors with lower local density and are at a relatively large distance from points with higher local densities, which fit the nature of teeth shapes. The groove areas and biting surfaces have more vertices and separate naturally when describe in triangular meshes.

There are two critical variables should be clarified in advance, the first one is local density ρ_i of data point i:

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \tag{1}$$

Where $\chi(x)$ is an indicate function, and d_c is a cutoff distance. According to Rodrigues and Laio [11], local density ρ_i represents the number of points that are closer than d_c to point i. The second important variable in this algorithm measuring the minimum distance between the point i and any other point with higher density, defining as:

$$\delta_i = \min_{j:\rho_j > \rho_i} (d_{ij}) \tag{2}$$

Suppose point k is the point with maxima density, the nearest neighbor distance defined by $\delta_k = \max_j(d_{kj})$, at the aim of making it as large as possible. Therefore, for points whose density is local or global maxima, δ_i is much larger than the typical nearest neighbor distance. The cluster centers are recognized as points whose δ_i are extremely high. The number of the cluster centers is determined according to a so-called "decision graph", which shows the plot of δ_i as a function of ρ_i for each point. We choose the points on the upper right on the graph as the cluster centers with both high value of δ_i and ρ_i. Figure 3(a) gives an example of the decision graph. After determining the cluster centers, remaining points are clustered according to the nearest neighbor of higher density. Removing the uneven border region could be achieved by the robust assignment of the CFSFSP, which distinguishes the cluster core and cluster halo based on the border region of each cluster [11]. The gray margin in Figure 4(a) is the distinguished cluster noise.

2.2. Measuring Geodesic Distance on Meshes

However, in 3D shape processing, the widely used distances in clustering algorithms, such as Euclidean distance, may fail to segment the surface into geometrically meaningful segments (as shown in Figure 1). The occlusal surface, mesial surface and distal surface are full of grooves, simply using Euclidean distance can only capture the relations of points in a small area. To overcome this shortcoming and make a full use of the surface structure, we propose to use geodesic distance as the distance metric for clustering on the surface manifold *S* instead of the Euclidean distance in $\mathbb{R}^3$. The geodesic distance between two points on surface *S* is defined as the shortest path between them conforming to the surface, and the length of the shortest path is the geodesic distance. As Huang et al. [6] pointed out geodesic distance between two points on a dental model is a more meaningful measure than the common Euclidean distance.

After adjusting the distance metric to geodesic distance, the result of the CFSFDP method applied on teeth would be more rigorous since the distance between the points with high density become larger and the found cluster centers contain more geometrical information. As Figure 3(a) shows, the cluster centers can naturally separate different kinds of teeth, which are located at the "pits" of the teeth. To compute the geodesic distance on the mesh, we adopted the method proposed by Vitaly Surazhsky et al. [18].

2.3. Automatically selecting landmark among each Cluster

Among each cluster, we searched for points containing more geometric information, such as curvatures. It's natural that points on some specific regions of the surfaces among certain category will carry more information, such as the cusps and the sulcus for teeth shapes. Inspired by the method of choosing local maximums of the Gaussian curvature as landmarks [8], we want to choose points with the most distinguishing feature values as our landmarks. Geometrically, those points should have large distances from the center of the data set, in which the Mahalanobis distance is a commonly used distance measure and the Hotelling's T^2 statistic [19] can be utilized for feature selection. Base on PCA and Hotelling's T^2 measure, we here propose landmark selection method.

To begin with, we define a feature vector for each vertex in the surface at the aim of fully capturing the geometric information. We choose the lumped surface area of each vertex, the mean curvature, the Gaussian curvature and the discretized Laplacian vector as features, being written as feature vector $[a, m, G, l_1, l_2, l_3] \in \mathbb{R}^6$. Before computing the distance from each point to the center, we conduct PCA to reduce the dimension and extract the first two columns of the score matrix. Then we compute the Hotelling's T^2 value for each vertex after dimension reduction by: [20]

$$T^2_{A=2,\alpha=0.97} = \frac{t_1^2}{\lambda_1} + \frac{t_2^2}{\lambda_2} \tag{3}$$

where t_1 and t_2 are the scores of the first two PCs, λ_1 and λ_2 are the two largest eigenvalues or variances of the first two PCs, A is the number of PCs we choose, and α is the confidence limit of T^2. We here choose points outside the 97% confidence ellipse since we want to choose points with large distances from the central point. Shown on Figure 4(b), selected by this method, both the cusps and the incisal edge are captured.

3. Computational Experiments for Methodology Illustration and Validation

In this section, computational experiments are conducted to illustrate the proposed clustering method, feature selection through PCA, and the verification of the selected landmark point set. The tooth data from the actual print contains 4613 vertices. The geodesic distances are computed in MATLAB by exact geodesic for triangular meshes [1]. The feature vectors for point selection are generated by PyMESH, and the shape registrations were implemented by library Open3D [21].

3.1. Clustering and Points Selection Result

After obtaining the geodesic distances of each vertex to other vertices on the meshes, we compute the local density and the defined distance δ for every point. Figure 3(a) shows the decision graph to determine the number of the cluster centers, according to which 10 cluster centers are suggested. The locations of the cluster centers are visulized in Figure 3(b). We can observe that the cluster ceters separate the teeth naturlly and symetrically. Figure 4 shows the clustering results and the point selection results. In Figure 4(a), different clusters are colored differently, in particular, the gray part stands for the cluster noise which is the indifferent part of our analysis. Figure 4(b) and Figure 4(c) show the locations of the selected landmarks. From the side elevation, we can observe that the edges between teeth are also well captured by this method. The number of the selected points is 3% of the size of whole points set, in this specific case, we select 131 points out of 4613 vertices of this data set.

3.2. Representation of the informativeness of the selected points

The basic idea to assess the informativeness of the selected points is based on a stable shape registration method: if we align two surfaces with the exact same shape but on different position in the space, the better the optimal transformation generated merely by the selected points set align the whole shape, the more information we claim the selected points set contains. We here apply iterative closest point (ICP) registration algorithm [22] to exam the informativeness of the selected points, since it has been a mainstay of geometric registration and it is robust if certain conditions are meet, including a good initial transformation. With ICP, the proposed evaluation procedure is as follows: (1) apply an initial noise transformation to the original shape and the extracted points set; (2) calculate the initial root-mean-square error (RMSE) before registration; (3) use ICP algorithm to align the two selected point sets, return the transformation; (4) apply the transformation from partial registration result to the complete shape and calculate the RMSE between the complete source and target shape after registration; (5) compare the initial RMSE with the partial registered result. The smaller the partial registration root-mean-square error ($RMSE_p$), the more informative the selected points set is. The control groups to this experiment is directly registering the whole shape with the same initial noise transformation, denoting the error as $RMSE_c$. The initial noise transformations are translations with parameters

randomly generated from $[-2.5,2.5]$, ensuring that the identity matrix is a good estimation of the initial matrix. The $RMSE_i$ is the original error after applying the translations without registrations. $K = \frac{RMSE_c}{RMSE_p}$ is computed to evaluate the performance of the selected points. If $K \geq 1$, it indicates that the overall error of using selected landmarks set is even smaller than the error from directly registering the entire shapes, it further evaluates that the selected landmarks set contains the most information of the whole shape. In total, we conducted 10000 experiments, and there are 98.52% of the them with $K \gg 1$.

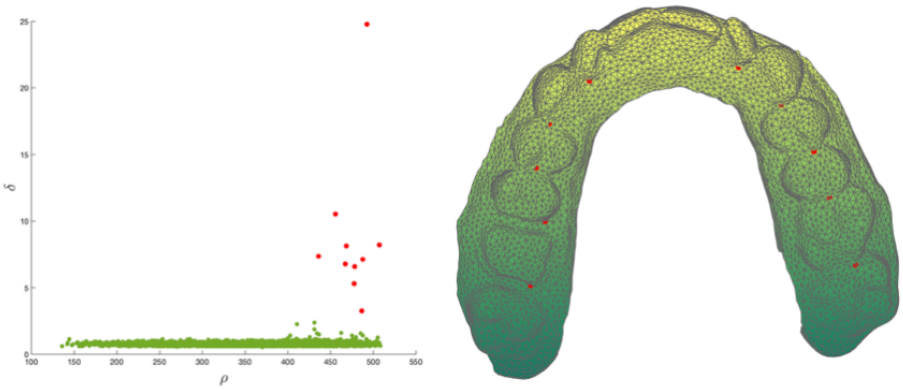

Figure 3: (a) decision graph of the CFSFDP; (b) shows the locations of the cluster centers chose by the CFSFDP algorithm.

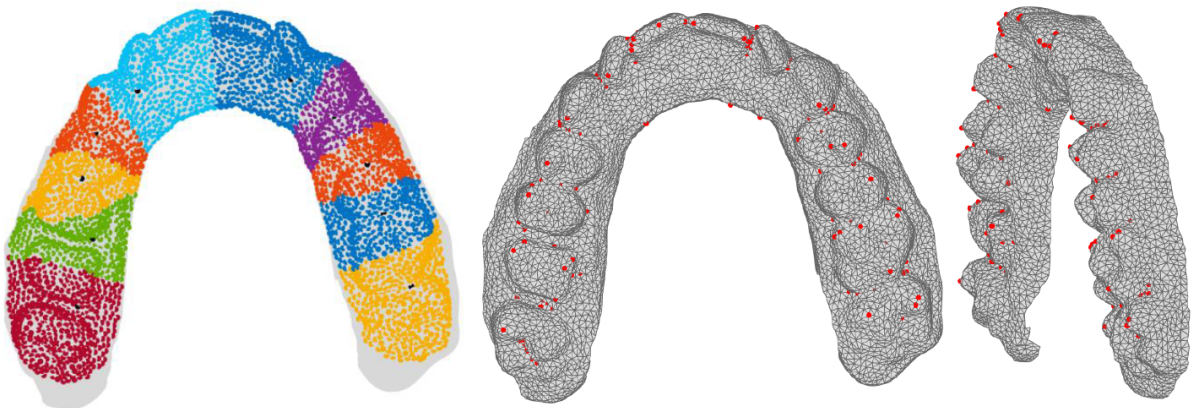

Figure 4: (a) Visualization of the clustering result on the teeth surface. The gray area is the cluster noise. (b) The selected points among all clusters. (c) Side elevation of the selected landmarks over the surface.

4. Conclusions

In this work, we develop an automatic landmark selection method for complex 3D printed objects such as dental models. To incorporate biometric information, we apply a novel clustering method for shape segmentation and revise it to be more geometrically meaningful by applying the geodesic distance. The boundary of the 3D object is treated as a manifold, in this case, surface, instead of directly in the Euclidean $\mathbb{R}^3$space. Without prior knowledge of the number of cluster centers, the proposed algorithm successfully selects cluster centers with geometric importance, which can naturally separate different kinds of teeth.

In each cluster, we use PCA to concentrate the intrinsic geometry information, including Gaussian curvature, mean curvature, surface area, etc. Choosing points outside the 97% confidence ellipse determined by Hotelling T^2 measure,

we successfully select points representative on both concave and convex areas among each cluster. We also conducted experiments to exam the informativeness of the selected point set. The result shows that the 3% points we chose from the whole shape are informative at the sight of the shape alignment. Future research effort will be devoted to fast non-rigid shape correspondence based on the selected point sets.

Acknowledgement

This work is supported by NSF grant CMMI-1901514.

References

[1] Q. Huang, J. Zhang, A. Sabbaghi, and T. Dasgupta, "Optimal offline compensation of shape shrinkage for three-dimensional printing processes," *IIE Trans. (Institute Ind. Eng.*, vol. 47, no. 5, pp. 431–441, 2015.

[2] J. Huang, T. H. Kwok, and C. Zhou, "V4PCS: Volumetric 4pcs algorithm for global registration," *Proc. ASME Des. Eng. Tech. Conf.*, vol. 1, pp. 1–10, 2017.

[3] J. Huang, H. Sun, T.-H. Kwok, C. Zhou, and W. Xu, "Geometric Deep Learning for Shape Correspondence in Mass Customization." 10-Jun-2019.

[4] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios, "Dense non-rigid surface registration using high-order graph matching," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 382–389, 2010.

[5] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or, "A survey on shape correspondence," *Eurographics Symp. Geom. Process.*, vol. 30, no. 6, pp. 1681–1707, 2011.

[6] J. Huang, H. Sun, T.-H. Kwok, C. Zhou, and W. Xu, "Geometric Deep Learning for Shape Correspondence in Mass Customization by 3D Printing."

[7] W. Zeng and X. D. Gu, "Registration for 3D surfaces with large deformations using quasi-conformal curvature flow," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2457–2464, 2011.

[8] Y. Lipman and T. Funkhouser, "Möbius voting for surface correspondence," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 1–12, 2009.

[9] V. G. Kim, Y. Lipman, and T. Funkhouser, "Blended Intrinsic Maps," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 1–12, 2011.

[10] J. Sun, M. Ovsjanikov, and L. Guibas, "A concise and provably informative multi-scale signature based on heat diffusion," *Eurographics Symp. Geom. Process.*, vol. 28, no. 5, pp. 1383–1392, 2009.

[11] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science (80-.).*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[12] P. Sharma and J. Suji, "A Review on Image Segmentation with its Clustering Techniques," *Int. J. Signal Process. Image Process. Pattern Recognit.*, vol. 9, no. 5, pp. 209–218, 2016.

[13] D. L. Page, A. F. Koschan, and M. A. Abidi, "Perception-based 3D triangle mesh segmentation using fast marching watersheds," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, pp. 3–8, 2003.

[14] S. Katz and A. Tal, "Hierarchical mesh decomposition using fuzzy clustering and cuts," *ACM SIGGRAPH 2003 Pap. SIGGRAPH '03*, pp. 954–961, 2003.

[15] R. Liu and H. Zhang, "Segmentation of 3D meshes through spectral clustering," *Proc. - Pacific Conf. Comput. Graph. Appl.*, pp. 298–305, 2004.

[16] A. Hartigan and M. A. Wong, "A K-Means Clustering Algorithm," *J. R. Stat. Soc.*, 1979.

[17] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.*, 2010.

[18] V. Surazhsky, T. Surazhsky, D. Kirsanov, S. J. Gortler, and H. Hoppe, "Fast exact and approximate geodesics on meshes," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 553–560, 2005.

[19] and D. W. W. RICHARD A. JOHNSON, *Applied Multivariate Statistical Analysis*, vol. 64, no. 6. 1974.

[20] K. G. Dunn, "Process Improvement using Data," no. 294-34b8, p. 381, 2014.

[21] Q.-Y. Zhou, J. Park, and V. Koltun, "Open3D: A Modern Library for 3D Data Processing," 2018.

[22] P. J. Besl and N. D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2. pp. 239–256, 1992.