TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.7, July 2014, pp. 5546 ~ 5551
DOI: 10.11591/telkomnika.v12i7.5203

Fuzzy c-Means and Mean Shift Algorithm for 3D-Point Clouds Denoising

Tongguang Ni[1], Xiaoqing Gu[2], Hongyuan Wang*[3]
School of Information Science and Engineering, Changzhou University, Changzhou 213164, China
Telp 86-0519-86330558, Fax 86-0519-86330284
*Corresponidng author, e-mail: hbxtntg-12@163.com[1], guxqcczu@gmail.com[2], tiddyddd@163.com[3]

Abstract

In many applications, denoising is necessary since point-sampled models obtained by laser scanners with insufficient precision. An algorithm for point-sampled surface is presented, which combines fuzzy c-means clustering with mean shift filtering algorithm. By using fuzzy c-means clustering, the large-scale noise is deleted and a part of small-scale noise also is smooth. The cluster centers are regarded as the new points. After acquiring new point sets being less noisy, the remains noise is smooth by mean shift method. Experimental results demonstrate that the algorithm can produce a more accurate point-sample model efficiently while having better feature preservation.

Keywords: *point-sampled model, mean-shift procedure, fuzzy c-means*

1. Introduction

Point-sampled models are normally generated by sampling the boundary surface of physical 3D-scanning devices. Despite the improvement of scanning accuracy, the data is invariably noisy. Moreover, the increasing use of 3D scanners has implied a growth in the complexity of the scanned models. Therefore, it is crucial that noisy models need to be denoised or smoothed before performing any subsequent geometry processing such as simplification, reconstruction and parameterization. There is a challenge to remove the inevitable noise while preserving the underlying surface features in computer graphics. In particular, fine features are often lost if no special treatment is provided [1, 2].

In recent year, a variety of denoising methods have been intruduced, such as the Laplacian operator [3], anisotropic diffusion [4, 5], diffusion of the normal field [6], and locally adaptive Wiener filtering [7]. A mean-shift-based anisotropic denoising algorithm [8] is proposed for point-sampled surfaces. Taking into account the vertex normal and curvature as the range component, the algorithm extend mean shift filtering to 3D surface smoothing. By clustering adjacent sample points of similar local modes, the method also provides a meaningful segmentation of the point model. The neighbors of each sample point are collected under spatial and range constraints. Finally, the proposed trilateral point filtering algorithm can remove noise while preserve geometric features. Although the method is efficient, oversmoothing will be produced when the mesh is suffered from large-scale noise.

In this paper, a two-stage point clouds denoising method is proposed, which combines fuzzy c-means with the mean-shift-based anisotropic denoising of point-sampled surfaces. This algorithm can handle the large-scale and small-scale noise and be formulated for mesh-based geometry and even for general 3D geometry.

2. Fuzzy c-means

Fuzzy c-means (FCM) is a famous method of clustering which allows one piece of data to belong to two or more clusters. It is based on minimization of the following objective function:

$$J_m(\mu, c) = \sum_{j=1}^{N} \sum_{k=1}^{C} \mu_{jk}^{m} \left\| p_j - c_k \right\|^2, 1 \le m \le \infty \qquad (1)$$

Received November 25, 2013; Revised February 25, 2014; Accepted March 11, 2014

Where m is any real number greater than 1, μ_{jk} is the degree of membership p_j in the cluster k, p_j is the jth d-dimensional points, c_k is the cluster d-dimensional center of the cluster, $\|\cdot\|$ is any norm expressing the similarity between any measured data and the cencter.

This paper adopts the improved method of the fuzzy clustering [9, 10], which defined a fuzzy weighting coefficient, it makes short distances become much shorter and long distances become much longer. So the performance of the clustering becomes better clustering. The detail procedures of the fuzzy clustering algorithm are given in [11].

3. Fuzzy c-means Clustering with Mean Shift Filtering Algorithm

The main idea of our algorithm is that, firstly the noisy data points will be pre-processed by an improved fuzzy c-means clustering method. For each data point, we detect the number of the neighboring point in the given sphere in order to determine if it is noise or not. If the number of the neighboring point is less than the given threshold, the point is a noise; otherwise we will cluster the points in the sphere and regard the cluster center as the new point, which can filter the noise near the point sets. This process will preserve some small-scale noise. Using the vertex normal and the curvature as the range component and the vertex position as the spatial component, the local mode of each vertex on point-based surfaces is computed by a 3D mean shift procedure dependent on local neighborhoods that are adaptively obtained by a kdtree data structure. Clustering pieces of point-based surfaces of similar local mode can provide meaningful model segmentation. Then, a trilateral point filtering scheme is applied based on the adaptively clustered neighbors. The scheme can adjust the position of sample points along their normal directions. Finally the noise is reduced from point-sampled surfaces successfully while preserving geometric features.

3.1. The Fuzzy c-means Algorithm for Large Scale Noise

We define that s is the surrounding sphere, r is the radius of s , and size is the given threshold of number of close points in the surrounding sphere s and m_i is the number of close points in the surrounding sphere s of point p_i .

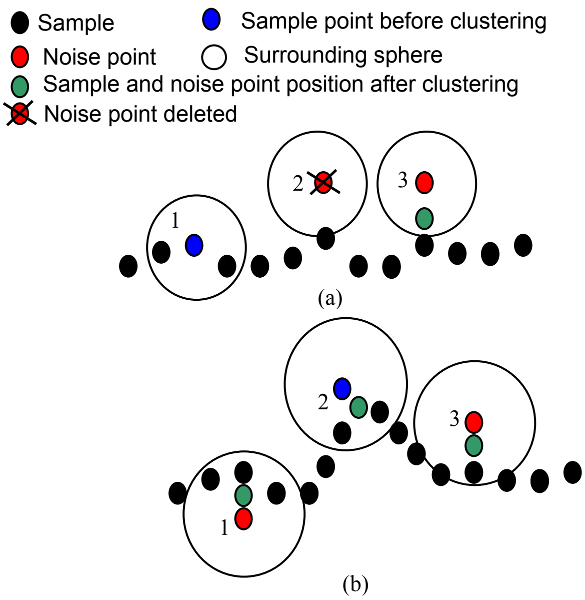


Figure 1. Small-scale noise partly filtered and Large-scale noise deleted by FCM. The parameter size is defined as 3. The red points are noise. (a) The second point is noise, m_2=2, m_2<size, we deleted it. The second point is noise, m_3=3=size, so it is moved to clustering center of points in the sphere by clustering method. (b) The first and third points are noise. Because the number of close points in the surrounding sphere is larger than size, we regard them as sample points

The following is the pseudo-code for applying fuzzy c-means clustering to a single point:

Large-scale DenoisePoint (point p_i)
{k_{ij}} = neighborhood (p_i)
For i: =1 to N
If (m_i<size)
Delete p_i
Else
Call FCM ( )
End
c_i=fuzzy c-means clustering center of p_i
Return new point q_i=c_i

In Figure 1, we can see that large-scale noise is deleted and small-scale noise partly filtered by FCM algorithm. But FCM can't delete small-scale noise, it only partly smooth them. In the next section we smooth them by bilateral filter.

3.2. Estimation of Normal Vector and Curvature

The normals and curvatures of the point-sampled geometry can be estimated by various methods to estimate [12,13]. Assume p_j, j=1, 2,…, m is a subset of the original measuring point set P={p_1, p_2,…, p_m}, based on the theory of principal component analysis (PCA), the 3×3 covariance matrix of p_j could be defined as follow:

$$C_i = [p_1 - \overline{p}_j, p_2 - \overline{p}_j, ..., p_n - \overline{p}_j][p_1 - \overline{p}_j, p_2 - \overline{p}_j, ..., p_n - \overline{p}_j]^T \quad (2)$$

Where C is a symmetric positive semi-definite matrix, and the centroid of p_j is $\overline{p}_j = \frac{1}{n}\sum_{i=1}^{n} p_i$. The normal of p_i is chosen to be the unit vector $e_{i,1}$, which corresponds to the minimal eigenvalue of C_i .

$$n_i = e_{i,1} \quad (3)$$

Pauly et al. [14] showed that surface variation is closely related to mean curvature, and here the curvature on p_i is taken as the surface variation,

$$H_i = \lambda_{i,1} / (\lambda_{i,1}+\lambda_{i,2}+\lambda_{i,3}) \quad (4)$$

Where $\lambda_{i,j}$, j = 1, 2, 3 are eigenvalues of C_i and satisfy $\lambda_{i,1} \le \lambda_{i,2} \le \lambda_{i,3}$.

3.3. Mean Shift Denoising

Since the spatial and range domains of 3D geometry are slightly different from those of images, 3D position of a vertex is usually regard as spatial information, but in this paper we regard the normal and curvature of the local surface as range information or feature space information. We extend the mean shift algorithm to 3D domain directly.

We assume that the data points p_i are the generalized points of the input raw point model $P \subset R^3$, and the spatial position information v_i=(x_i, y_i, z_i) and range information including the normal vector n_i and the mean curvature H_i of vertices are inclued in the vector components of p_i, which can be written as:

$$p_i=(v_i, n_i, H_i) \quad (5)$$

With i = 1, 2, . . . , n, and n being the number of points in P. Here the dimension of vector p_i is 7. The k nearest neighboring points of generalized points p_i are denoted by $N(p_i)$={$q_{i,1}$, $q_{i,2}$,…, $q_{i,k}$}.

Thus, the mean shift vector of p_i can be expressed as:

$$M_v(p_i)=\frac{\sum_{j=1}^{k} g\left(\left\|p_i^r-q_{ij}^r\right\|\right)\left(q_{ij}-M(p_i)\right)}{\sum_{j=1}^{k} g\left(\left\|p_i^r-q_{ij}^r\right\|\right)} \quad (6)$$

Where g(·) could be either a Gaussian kernel or an Epanechnikov kernel; $p_i^r=(n_i,H_i)$is the range part of p_i, and $M(p_i)$ is called the mean shift point associated with p_i, and $M(p_i)$ could be initialized to coincide with p_i, $M_v(p_i)$ is the mean shift vector associated with $M(p_i)$. Then we define the mean shift procedure as the repeated movement of data points to the sample means, written as:

$$M(p_i) := M(p_i) + M_v(p_i) \quad (7)$$

By running the procedure for all i = 1, 2,…,∞, each data point iterates to a local mode in the joint spatial range domain, and the mean shift procedure has provided a stable local mode detection for the point-sampled model.

3.4. Anisotropic Denoising Algorithm for Small Scale Noise

The detailed mean shift denoising algorithm consists of four processing stages.

Step 1. Initialization. Construct a kdtree structure for the point model and search neighbors $\{q_{ij}^s\}$ for each v_i in the spatial domain, then initialize range component p_i^r by principal component analysis of the spatial neighbors $\{q_{ij}^s\}$. The range bandwidths $h_r=\left(h_1^r, h_2^r\right)$ are usually defined as positive values related to normal and curvature, respectively; for formulation convenience, we write h_r directly.

Step 2. Mean shift procedure. Repeat the mean shift procedure discussed above until convergence.

Step 3. Clustering. Build clusters of points whose modes are similar. Generally, the neighborhood size k is an essential parameter for good shape smoothing results. Pauly et al. [15] suggest we select a k in the range from 6 to 20 in the spatial domain.

Step 4. Vertex estimation. After updating the range component of pi and its neighbors, we apply trilateral filtering in the influence region associated with a fixed local mode. Compared with bilateral filtering, we not only separate spatial and range signals to determine the local area with geometric coherence, but also introduce a curvaturerelated kernel to smooth high gradient regions efficiently.Our approach is slightly different from the trilateral normal filtering proposed in [16, 17].

The curvatures are considered as second-order properties of the 3D geometry, and the performance of the curvaturerelated kernel for the regions near the salient ridge and ravine structures is satisfactory.

$$\begin{cases} v_i = v_i + \omega_i n_i \\ w_i = \sum_{q_{ij}^S} G_i\left(n_i, v_i - q_{ij}^S\right) g_i \\ G_i = G_{\sigma_1}\left(d_i\right) G_{\sigma_2}\left(h_i\right) G_{\sigma_3}\left(e_i\right)\left(1+r_i\right) \end{cases} \quad (8)$$

Where $G_\sigma(\cdot)$ is a Gaussian kernel, d_i is the distance of $\|v_i - q_{ij}^s\|$, h_i is the projection of the vector ($v_i - q_{ij}^s$) onto the normal n_i, and e_i is the inverse of the curvature difference between v_i and q_{ij}^s.

Step 5. Adaptive neighbors. For large point data sets, the selection of neighbors is a trade-off between computational time and smoothing quality. When choosing a small and uniform spatial size of a neighborhood, it costs less in computational time, but smooths the model poorly; on the other hand, if we choose a large and uniformly spaced neighborhood, the model is oversmoothed. We use the kdtree to search for the k nearest neighbors, for instance k = 12, instead of a neighboring spatial ball, which is unfortunately dependent on the sampling density of the point model. Furthermore, our clustering scheme provides an adaptive neighbor searching method, where different influence regions are used adaptively to remove noise from vertices. Applying trilateral filtering to the adaptive neighborhoods greatly improves the smoothing capabilities of the mean shift filter in high gradient regions. Although the algorithm we present is ordinarily non-iterative whilst denoising, we can also use the resulting adaptive neighbors as inputs to estimate range information, then iteratively perform the four steps of the algorithm to make the point model smoother.

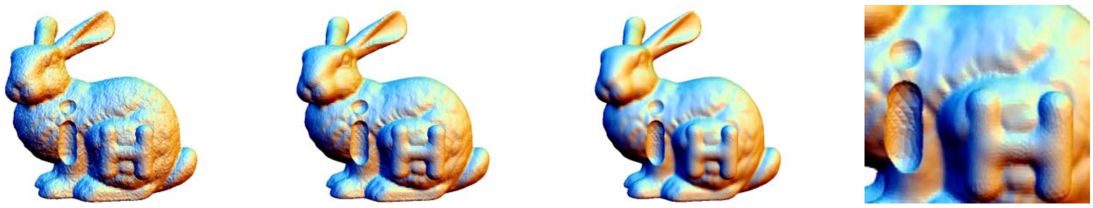
4. Experimental Results and Analysis

Using a PC of Intel core2 Q9550 and 8GB memory, this paper points of different models and predecessors of the denoising algorithms experiment and compared, We implemented the point clouds denoising algorithm as described in the previous section and gave our results, and our algorithm and previous algorithms were compared.

A comparison to the mean shift smoothing approach is shown in Figure 3. Figure 3(a) random noise is added to the moai model (the number of the noise is 6543, and the number of the point is 20000). Figure 3(b) is the results using the mean shift point clouds denoising method. In Figure 3(c) the noise deleted by FCM with parameter size=6. In Figure 3(d) small-scale noise smoothing by mean shift point clouds denoising method. We can see that our results of the two models are better than the mean shift point clouds denoising method while handling the large-scale noise and the mean shift point clouds denoising method will produce oversmoothing in sharp features.

5. Conclusion

This paper has presented a two-stage point cloud denoising method which combining fuzzy c-means with mean shift filtering approach. Our algorithms have a good result while working with unorganized and large-scale noisy point sets. But it has a disadvantage that the improved FCM will partly smooth sharp feature while clustering. In Figure 3(d), we may see hat the second point moves towards its close points. In the future, we hope to improve our approach in order to preserve the sharp features of models better.



(a) (b) (c) (d)

Figure 2. Noisy point deleted by our algorithm. (a) The noisy point clouds, (b) Large-scale noise deleted by FCM, (c) Smallscale noise smoothing by mean shift point clouds denoising method, (d) The model feature is preserved

(a) (b) (c) (d)

Figure 3. Comparison with mean shift point clouds denoising method for the moai. (a) is the noisy point clouds, (b) is the result using mean shift approach, (c) is Large-scale noise deleted by FCM, (d) shows noise smoothing by mean shift point clouds denoising

Acknowledgements

The models are of courtesy Y. Ohtake, and thanks to Y. Ohtake forproviding the basic platform, so we can implement our algorithm efficiently. This work was funded by the National Natural Science Foundation of China (No. 61070121).

References

[1] Zhang X, Li H, Cheng Z, Zhang Y. Robust Curvature Estimation and Geometry Analysis of 3d Point Cloud Surfaces. *Journal of Computational Information Systems*. 2009; 6(5): 1983-1990.
[2] Ni T, Ma Z. *Fast algorithm for point-sampled model denoising*. International Conference on Computer Engineering and Technology. 2010; 7: 719-722.
[3] Taubin G. *A signal processing approach to fair surface design*. Proc. SIGGRAPH. 95: 351–358.
[4] Bajaj C, Xu G. Anisotropic diffusion on surface and function on surfaces. *ACM Trans. on Graphics.* 2003; 22: 4–32.
[5] Clarentz U, Diewald U, Rumpf M. *Anisotropic geometric diffusion in surface processing*. Proc. IEEE Visualization. 2000: 397–405.
[6] Ohtake Y, Balyaev A, Seidel H, *Mesh smoothing by adaptive and anisotropic Gaussian filter applied to the mesh normal*. Proc. Vision, Modeling, and Visualization. 2002; 203–210.
[7] Alexa M. *Wiener filtering of meshes*. Proc. Shape Modeling International. 2002; 51–57.
[8] Hu G, Peng Q, Forrest A. Mean shift denoising of point-sampled surfaces.*The Visual Computer*. 2006; 22(3): 147-157.
[9] Liu W, Chen Z, Bai P, Fang S, Shi Y. *A kind of improved method of fuzzy clustering*. Proceedings of the Fourth International Conference on Machine Learning and Cybernetics. 2005; 5: 2646 - 2649.
[10] Ni T, Wang H, Gu X. A fast surface reconstruction algorithm by robust fuzzy C means and down sampling. *Journal of Computational Information Systems*. 2013; 9(3): 907-914.
[11] Wang L, Yuan B, Tang X. *Implicit Surface Reconstruction from Noisy 3D Scattered Data*. Proceedings of the Eighth International Conference on Signal Processing. 2006; 2: 134-142.
[12] Hu G, Peng Q. *Bilateral estimation of normal on point-sampled models*. International Conference on Computational Science and Its Applications. 2005; 758–768.
[13] Hoppe H, DeRose T, Duchamp T,McDonald J, Stuetzle W. *Surface reconstruction from unorganized points*. Proceedings of ACM SIGGRAPH. 1992; 71–78.
[14] Pauly M, Gross M, Kobblet L. *Efficient Simplification of Point-Sampled Geometry*. IEEE Visualization. 2002; 163–170.
[15] Pauly M, Keiser R, Kobblet L, Gross M. Shape modeling with point-sampled geometry. *ACM Trans. Graph*. 2003; 22(3): 641–650.
[16] Choudhury P, Tumblin J. *The trilateral filter for high contrast images and meshes*. Proceedings of Eurographics. 2003; 186–196.
[17] Louie H, Burns M, Lima C. *An introduction and user's guide to the IEEE Smart Grid Web Portal*. Innovative Smart Grid Technologies Conf. Europe. 2010; 1–5.