Road Surface Modeling and Representation from Point Cloud Based on Fuzzy Clustering

ZHANG Yi  YAN Li

Abstract  A scheme for an automatic road surface modeling from a noisy point cloud is presented. The normal vectors of the point cloud are estimated by distance-weighted fitting of local plane. Then, an automatic recognition of the road surface from noise is performed based on the fuzzy clustering of normal vectors, with which the mean value is calculated and the projecting plane of point cloud is created to obtain the geometric model accordingly. Based on fuzzy clustering of the intensity attributed to each point, different objects on the road surface are assigned different colors for representing abundant appearances. This unsupervised method is demonstrated in the experiment and shows great effectiveness in reconstructing and rendering better road surface.

Keywords  surface modeling; point cloud; distance-weighted fitting; fuzzy clustering; normal vectors; intensity
CLC number  P237.4; TP753

Introduction

Modeling of road surface is a foundational work in design, reconstruction and visualization. The way the data is captured directly influences modeling the quality. Existing relief maps and constructing drawings, data from total-station and GPS and extracted road features from images are sources of road data[1].

The traditional surveying of a highway is extremely restrictive as it implies lane closures or even entire road closures. As long as lanes remain open, it becomes extremely time-consuming to measure points on the road as passing cars and trucks interfere with the clear line of sight required by traditional surveying equipment. Line-array laser scanning on vehicles contributes to the accurate and efficient measurement highways[2,3]. Although the results reveal the availability of road inspection and profile construction, the area where this technique can be used is limited and the data only represent those fixed regions that laser rays are focused on, which is inadequate for other analysis and visualization. LIDAR (light detection and ranging) has recently been widely used in the extraction of vehicles and road features[4-7]. The road model is described as a line element because of sparse sampling points that cannot provide much detail of the road surfaces.

The road point cloud is a good data source for surface modeling as the laser scanner almost immediately delivers real 3D dataset with no further complicated compilation processes. Point cloud is a collec-
tion of $XYZ$ coordinates in a common reference system that portrays to the viewer an understanding of the spatial distribution of a subject or a site. It may also include additional information such as intensity or RGB value\cite{8}. A point cloud can describe geometry and show spatial relationship between objects by its $XYZ$ coordinates, and represent the properties of objects surface as well as compute the reliability of the range data for each point\cite{9}.

Points corresponding to passing vehicles are the very problems in geometric modeling. In this study, we take it as noise for filtering. An effective algorithm is introduced based on normal vectors clustering analysis other than manual removing from the original point cloud. The correctness of normal vectors is fundamental to the clustering result. We use a distance-weighted fitting method to estimate the normal vector of every point. Texture modeling is traditionally processed with true color image based on ortho-rectification to model the precise texture\cite{10}. But for road representation, texture is simply represented as two distinguishable parts: road surface and road marking (white lines, zebra crossing, etc). An intensity clustering-based method is presented to distinguish them. Accordingly, appropriate gray and white color is given to corresponding geometric model avoiding complicated registration between color image and geometric model. The modeling results show high quality and nice visualization of road surface.

1 Estimation of normal vectors

Normal vector of a point cloud is the perpendicular direction of differential plane fitted by several points around it (Fig.1). For estimating the normal vector of a certain point $Q(x_i, y_i, z_i)$, setting a sphere centered at $Q$ with a radius $r$ to search those points inside the sphere, a plane equation (Eq.(1)) can be fitted. Vector $v(A, B, C)$ represents the normal vector of $Q$.

$$A(x - x_i) + B(y - y_i) + C(z - z_i) + D = 0 \quad (1)$$

The critical issue is the determination of $r$ because the differential plane of $Q$ is influenced by it. Smaller value $r$ leads to insufficient data to compute for an accurate normal vector. Larger value $r$ makes the fitting plane out of the differential property on point $Q$. So, a distance weight is considered to compute for the fitting parameters. It is according to the idea that the nearer a point is to $Q$, the more influence it will have. Eq.(1) can be transformed to a function model (Eq.(2)) and its corresponding error function (Eq.(3)):

$$Ax + By + Cz + D' = 0 \quad (2)$$

$$V = FX - F\tilde{X} \quad (3)$$

where $D' = D - Ax_i - By_i - Cz_i$ ; $F = [x \ y \ z \ 1]$ ; $X = [A \ B \ C \ D']^T$ ; $\tilde{X}$ is the approximate value calculated by Eq.(1) with sufficient data. Defining a distance weight as weight function:

$$P = 1/d^2 \quad (4)$$

where $d = \sqrt{(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2}$ . When $V^TPV = \min$ , the $X$ is expressed by Eq.(5):

$$X = (F^TPF)^{-1}F^TP\tilde{X} \quad (5)$$

The weight function assures fitting plane by reducing the influence of remote points and increasing the spatial property of local points as much as possible. All points assign vector $V$ with Eq.(5) and the estimation of normal vectors is accomplished.

2 Fuzzy clustering of normal vectors

In order to filter noisy point cloud, normal vectors are transformed to sphere coordinate, then $X$ axle and $Y$ axle, respectively, indicate latitude and longitude, while $Z$ axle calculates the frequencies corresponding to a certain point on the $X$-$Y$ plane. With the analysis of the normal vectors distribution of road surface points and noise points (Fig.2), it is concluded that normal vectors distribution of road surface points is comparatively concentrative and have high value opposite to that of the noisy points. This kind of diversity is coherent with the applicability of fuzzy
C-means (FCM). It is based on the minimization of the following objective function[11]:

$$ F_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \|v_i - c_j\|^m, \quad 1 \leq m < \infty $$

where $m$ is any real number greater than 1; $s_{ij}$ is the degree of membership of $v_i$ in the cluster $j$; $v_i$ is the $i$th of d-dimensional measured data; $c_j$ is the $d$-dimension center of the cluster; and $\|\|$ is any norm expressing the similarity between any measured data and the center.

Fig. 2 Point cloud and distribution of normal vectors

The point cloud is defined as two clusters. One belongs to the road surface, the other belongs to the noise.

$v_i$ is $i$th measurement of normal vectors, $v_i = (A_i, B_i, C_i)$. $c_j$ is $j$th cluster center of normal vectors, $c_j = (A_j, B_j, C_j)$.

$$ \|v_i - c_j\|^2 = \sqrt{(A_i - A_j)^2 + (B_i - B_j)^2 + (C_i - C_j)^2} $$

defined as the distance between them. The degree of membership $s_{ij}$ and center of the cluster $c_j$ are described as the following expressions:

$$ s_{ij} = \frac{1}{N} \sum_{k=1}^{N} \frac{\|v_i - c_j\|^m}{\|v_i - c_k\|^m} $$

(6)

$$ c_j = \left( \sum_{i=1}^{N} u_{ij}^m \cdot v_i \right) / \left( \sum_{i=1}^{N} u_{ij}^m \right) $$

(7)

The algorithm first initializes the degree of membership. Searching a point on $X$-$Y$ plane whose frequency is the maximum as a center, a circle is formed with a configured threshold of radium on latitude and longitude planes. Those points cloud have normal vectors inside the circle that belong to the road surface cluster and their degree of membership is initialized with $s_{ij} = 0.5, s_{12} = 0.9$. Subsequently, the algorithm repeatedly updates $s_{ij}$ and $c_j$ iteratively until it conforms to the equation $\max_y (|s^{(k+1)}_{ij} - s^{(k)}_{ij}|) < \varepsilon \cdot \varepsilon$ is a convergence value between 0 and 1, $k$ is the iteration times. Now, the degree of membership contains information of the clustering result. The higher $s_{ij}$ is, the more it is possible that corresponding cluster belongs to it. Because of the uncertainty of noisy point cloud, complete filtering is needed by fuzzy clustering over several occasions. So, a full filtering algorithm is designed. It recalculates normal vectors after every fuzzy clustering operation and ensures only two clusters before each operation. By iteration, the noisy points cloud are finally gradually removed.

1) Initializing the degree of membership $s_{ij} = 0.1, s_{12} = 0.9$.
2) Calculating $c_j$ according to Eq.(6).
3) Updating $s_{ij}$ according to Eq.(7).
4) If $\max_y (|s^{(k+1)}_{ij} - s^{(k)}_{ij}|) < \varepsilon \cdot \varepsilon$ ( $\varepsilon = 0.1$, go to next step. Else go to step 2).
5) Updating normal vectors according to Eq.(5).
6) Calculating angle $\theta$ of normal vectors before and after updating.
7) If $\theta \leq \delta$ ($\delta$ is a convergence threshold) then the iteration stops. Else go to step 1).

Pure road surface datasets include almost the same normal vectors, with which the perpendicular projective plane can be created. The entire surface points projected onto the plane construct the Delaunay triangular networks and then project it back to 3D space, then the road model is created. The model is able to represent the actual surface of the road since high-density points cloud amply generate full details. Tiny triangular units can directly render spatial continuity without further smoothness and interpolation.

3 Fuzzy clustering of intensity

The process of texture modeling involves color mapping of a spatial object after geometric modeling[12]. It is noticeable that the road, in appearance, is
simply represented as two distinguishable parts, the road surface and the road marking (white lines, zebra crossing, etc). No rich color information is needed, as gray and white colors are enough for road visualization. Therefore, as long as the road marking is distinguished, it can be rendered with white color. Considering the intensity information of points cloud, Liu Jingnan discussed laser intensity as used in the classification of LIDAR data\[^{13}\]. Different mediums have different values of laser reflectivity. The surface of the road consists of asphaltum (or concrete) and paintings on it as road marking. By analyzing the laser reflectivity of the two different mediums, asphaltum (or concrete) has a larger reflectivity than the road marking. So, the distribution of intensity is divisible. Similarly, on the distribution chart, the boundary is not clear for an abrupt classification (Fig.3) because there are other factors influencing statistics result. Meanwhile, the intensity distribution also has two concentrative and high peak values. With the FCM approach, it is effective to demarcate and gives appropriate color to them. The algorithm performs the same as the normal vectors clustering with a little difference in the measurement dimension. There are also two clusters, \( I_i \) is \( i \)th measurement of intensity, \( c_j \) is \( j \)th cluster center of intensity, where 
\[
\| I_i - c_j \| = \| I_i - c_j \|.
\]

**4 Experiment and results**

A scheme for surface modeling and representation of road point cloud is designed based on the introduction above (Fig.4). It has two sub flows separately indicating geometric modeling and texture modeling, and a common module indicating fuzzy clustering. We select raw data of road smoothness and slight wave painted with various roads marking, which consist of above 1 000 000 points. The spatial resolution of the points is superior to 2 cm. With normal vectors estimation and original included intensity, the road point cloud is ready for fuzzy clustering.

Fig.5 shows two kinds of road surfaces. Both of them consist of many noisy points above their surfaces. Clustering result depends on the correctness of the normal estimation. After normal vectors estimation and fuzzy clustering, clean surfaces remain in Fig.6. Through clustering operation, there are several points attributed to the road surface which are filtered as well, which makes the surface appear thinner than before, but this does not affect the data density for further modeling. In Fig.7, normal vectors distributions show concentrative and high values indicating that results are the point cloud of the surface without any noise. Fig.8 presents the road point cloud after intensity clustering and it shows distinguishable color according to their intensity attributes. Fig.9 indicates asphaltum and paintings in-
Classification is faster than normal vector because it needs no recalculating intensity. Fig.10 is the final model. The Delaunay triangular networks are created with clean road surface and appropriately rendering colors according to intensity clustering results.

Fig.5 Two kinds of noisy road point cloud

Fig.6 Two kinds of roads after fuzzy clustering

Fig.7 Normal vectors distributions

Fig.8 Intensity fuzzy clustering results

Fig.9 Intensity distributions
5 Conclusions

The scheme presented in this paper is mainly based on fuzzy clustering. It automatically processes point cloud with geometric attribute and spectral attribute in statistics, which provides distinguishable information for automatic modeling. The accurate calculation of the normal vectors is critical for modeling quality. So, the weight function is emphasized in the estimation of normal vectors in order to assure its correctness. The experiment and results demonstrate that the scheme is effective and feasible in modeling and representing road surfaces with huge noisy point cloud.

Although the results look promising, it has to be recognized that there are still a number of issues that should be addressed in future research. The convergence value in FCM algorithm directly influences the time costs of modeling. Also, how to determine an intelligent value to conform with different data needs more consideration. The gradient and fragmentation of the road surface also restrict the validity of normal vectors clustering. Intensity noise leads to promiscuous texture on the modeling as well. Resolving these problems will tremendously help in data processing in various situations and applications.

References

[1] Zuo Xiaqing, Li Qingquan, Tang Luliang (2004) 3D modeling and data organization of highway[J]. Geomatics and Information Science of Wuhan University, 29(2): 179-183 (in Chinese)
[2] Laurent J, Talbot M, Doucet M (1997) Road surface inspection using laser scanners adapted for the high precision 3D measurements of large flat surfaces[C]. The International Conference on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa, Ontario
[3] Kirchner A, Ameling C (2000) Integrated obstacle and road tracking using a laser scanner[C]. The Intelligent Vehicles Symposium, Detroit, USA
[4] Hutger C, Brenner M (2003) Extraction of road geometry parameters from laser scanning and existing databases[C]. ISPRS Working Group III/3 Workshop, Presden, Germany
[5] Rakusz A, Lovas T, Barsi A (2004) Lidar-based vehicle segmentation[C]. The 20th ISPRS Congress, Istanbul, Turkey
[6] Grejner-Brzezinska D, Toth C, Paska E, et al. (2004) Precise vehicle topology and road surface modeling derived from airborne LiDAR data[C]. The 60th Annual Meeting of ION 2004, Dayton, USA
[7] Grejner-Brzezinska D, Toth C, Moafipoor S, et al. (2006) Traffic flow monitoring from airborne LIDAR/CCD data[C]. The 3rd IAG and the 12th FIG Symposium, Baden
[8] Mills J, Barber D (2003) An addendum to the metric survey specifications for English heritage-the collection and archiving of point cloud data obtained by terrestrial laser scanning or other methods[OL]. http://www.english-heritage.org.uk/server/show/nav.9236
[9] Mota J G (2002) Localisation of mobile robots using scanned laser and reconstructed 3D models[D]. Portugal: Instituto Superior Técnico, Universidade Técnica de Lisboa
[10] Zhang Yi, Yan Li, Lu Xiuqin (2006) Precise texture modeling with 3D laser scanning[C]. The 14th International Conference on Geoinformatics, Wuhan, China
[11] Wang Xinzhou, Shi Wenzhong, Wan Shuliang (2003) Fuzzy geo-spatial information processing[M]. Wuhan: Publishing House of Wuhan University (in Chinese)
[12] Zhang Zuxun, Zhang Jianqing (2003) Solutions and core techniques of city modeling[J]. World Sci-Tech R & D, 25(3):23-29 (in Chinese)
[13] Liu Jingnan, Zhang Xiaohong (2005) Classification of laser scanning altimetry data using laser intensity [J]. Geomatics and Information Science of Wuhan University, 30 (3): 189-193 (in Chinese)