REGISTRATION OF SPATIAL DATA BASED ON PCA

Shang Tao, Wang Lili

School of Urban Design, Wuhan University, 430072

KEY WORDS: Registration; PCA; Coordinate Transformation; Spatial Data

ABSTRACT:

Registration of spatial data based on PCA is advanced in the paper. The principle is that principal components of spatial data set are changed with the coordinate system. For two spatial data sets of a certain object, we can estimate the coordinate system transformation matrix using the spatial transformation matrix between the corresponding principal components. The matrix of principal component is built up using the eigenvector and center of spatial data set. After obtaining the initial value of transformation matrix and transforming the spatial data, the transformation matrix can be obtained by using optimization algorithm. It is clearly the comparability between data sets is relative high. The experiment show the way is effective in obtaining the initial value of coordinate system in spatial data registration.

With the development of the spatial information gain technology, the researchers pay more attention on the matching spatial data including reconstruction of three-dimension, computer vision, and target identification. The matching is the way that seeks similar transformation relations to reference coordinate system from spatial data. Generally, the spatial data waiting for matching doesn't have strict homologue. At present, the matching algorithm first calculates approximate transformation between the coordinate system, and then finds its optimal solution based on either ICP or the smallest two multiplication method. The methods which calculate approximate value of coordinate transformation include: Using fixed reference marks; picking up feature points of spatial data automatically; giving the homologue coordinate artificially; using method of synchronization image matching and so on. The traditional method has already made great achievements in matching spatial data, but it still needed extra reference condition or manual intervention. However, in fact, the spatial data itself contains coordinate system's information. The relationship among principal components of the spatial data is related to spatial reference coordinate system. Therefore, speaking of the data acquisition with high-degree similarity, the transformation relations of principal components decides its relations of coordinate counterpart. Proposing the matching method of spatial data based on PCA transformation, the main subject is the matching which using different equipments or the same equipments with different spatial parameter condition to obtain the target spatial data. It also has the important reference value regarding the data splicing.

1 MATCHING SPATIAL DATA

Given the data $\{x_i\}_{i=1}^{N}$ and $\{y_j\}_{j=1}^{M}$ are different data point set of the same object. The spatial data is indicated by the homogeneous coordinate. The points in two sets could be great disparate. Data set maybe cannot be found the strict homologue; in addition, the order of data points in set is acquisition of transform parameters is critical. Excellent initial value can secure the right direction in its following exact matching

chaotic. So the aim for matching spatial data is to seek for spatial transform including the parameters of revolving matrix R、 Parallel vector t and proportion matrix s .Its objective function is:

$$H = \min \sum_{l}^{K} \left\| y_j - Tx_i \right\|_l^2 \qquad (1)$$

Among the x_i, y_i meets $\min \left\{ \left\| y_j - Tx_i \right\| \right\}_{i=1}^{M}$, $\left\| y_j - Tx_i \right\| < \varepsilon$, that is, the perigee after transformation as the approximate homologue. K is the calculated homologue point, T is the transform matrix , ε is the given limit value. Matching spatial data is the process searching continuously the most close homologue points, calculating transform relations of spatial data in the condition that restructuring the least error at the most close homologue points. Obviously, the initial value

2 ANALYSIS OF PRINCIPAL COMPONENTS

As the linear dimensionality reduction's typical method, the principal components analytic method has already known well for the scholars. Because of the simple concept, easy calculation, fine characteristics of reconstruction of linear error, it became one of most widespread use methods in the actual data dimensionality reduction.

We suppose data set $\{x_i\}_{i=1}^{N} \subset R^D$ which makes $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$, $S_N^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(x_i - \bar{x})^T$. Then we analyze S_N^2 . The result is: $S_N^2 = U^T \Lambda U$. Among it : $\Lambda = diag(\lambda_1, \lambda_2, \cdots \lambda_D)$,and , $\lambda_i \ge \lambda_{i+1} (i = 1,2,\cdots,D-1)$; $U = (u_1, u_2, \cdots u_D)$, u_i is the feature vector λ_i counterpart and U is the positive cross matrix.

The principal components of data sets is $y_i = U_d^T (x_i - \bar{x})$, $U_d = (u_1, u_2, \cdots, u_d)$. PCA transformation is to simplify for dimensions of spatial data in realizing reconstruction error by selecting the propel value d .

3 MATCHING ALGORITHM BASED ON PCA

After PCA transformation for those spatial data respectively, we conclude the principal components direction of spatial data; we receive structure matrix by principal components and spatial data's center coordinate; we calculate the transformation relations of structure matrix.

We analyze the principal components for data sets. The feature value is: $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$. Among them: λ_4 is the feature vector $[0,0,01]^T$ counterpart. Accordingly, the first, second and third principal components is the feature vector η_1, η_2, η_3 counterpart. They indicate vectors in spatial coordinate system. The center of spatial coordinate is $\overline{x}$, then structure matrix is $P_x = [\eta_1, \eta_2, \eta_3, \overline{x}]$ .As the same principle, we conclude structure matrix of data set $\{y_i\}$, that is $P_y = [\xi_1, \xi_2, \xi_3, \overline{y}]$. The structure matrix contains three principal components and center coordinate system of data sets. The value is related to spatial coordinate system. If two spatial data sets demonstrate the same complete object (or great overlap degree of object), the relationship among structure matrix is the mutual relationship of reference coordinate system counterpart. Thus, the transformation matrix shows below:

$$T = P_x^{-1} P_y = [\eta_1, \eta_2, \eta_3, \overline{x}]^{-1} [\xi_1, \xi_2, \xi_3, \overline{y}] \qquad (2)$$

The form of T is :

$$T = s \cdot \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \qquad (3)$$

After obtaining the initial value T, we transform spatial data and structure target function in the similar formula (1). In addition, we use the most optimal methods such as ICP(iterative Closest Point Algorithm), Heredity Algorithm, Least Squares, Levenberg-Marquardt to optimize transformation parameters. In obtaining initial value, more scholars proposed great researches in seeking for the most optimal solution. It isn't the important part in this article and over recounted here.

4 EXPERIMENT AND ANALYSIS

4.1 Experiment one

Firstly, we draw approximately 20,000 spatial points at random from spatial data in square-shape object. Then, we carry on coordinate system's revolving and translation, thus calculating a group of new coordinates. The given three rotation degree [8] (Y main axle) is: $\varphi = 0.3, \omega = -0.4, \kappa = 0.5$ .The translation vector is: $t = [45.102,-21.025,-70.222,1]^T$. Combining with two groups of data sets we get three-dimensional effect picture as Figure one shows. The structure matrix of primary data and transformation data after PCA transform for data set is :

$$P_1 = \begin{bmatrix} -0.7200 & -0.5145 & 0.4657 & 110.5306 \\ -0.0745 & -0.6099 & -0.7890 & 245.4561 \\ 0.6900 & -0.6028 & 0.4008 & 101.7791 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} -0.8046 & -0.0779 & 0.5887 & 28.5315 \\ -0.1095 & -0.9549 & -0.2760 & 265.8218 \\ 0.5837 & -0.2865 & 0.7597 & -86.6264 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

So the transform matrix is ：

$$T = P_2 P_1^{-1} = \begin{bmatrix} 0.8936 & -0.3570 & -0.2722 & 45.1019 \\ 0.4416 & 0.8083 & 0.3894 & -21.0249 \\ 0.0810 & -0.4682 & 0.8799 & -70.2220 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

According to analysis photogrammetry [8], we calculate the transform matrix is as same as the matrix T by using the rotation degree and translation vector which is given. The matching which is transformed spatial data by matrix shows in Figure two. Two groups of data achieved the complete superposition.

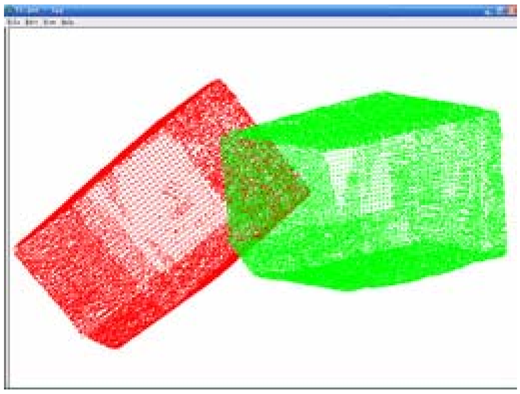

Figure 1. Non-matching data (1)

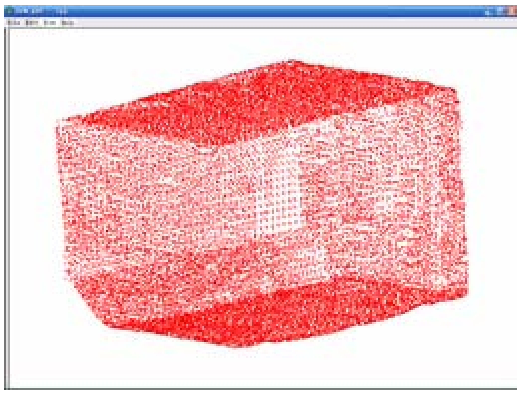

Figure 2 . Matching data clouds (2)

4.2 Experiment two

We draw approximately 10,000 spatial points at random from spatial data again. Then, we transform the spatial for data, using the same transformation parameters. Compared with primary data's three-dimensional display in Experiment one, as the figure 3 below shown, both of them exist space transformation.

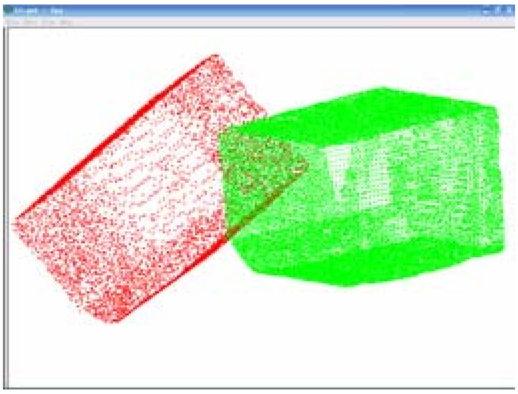

Figure3. Non-matching data (2)

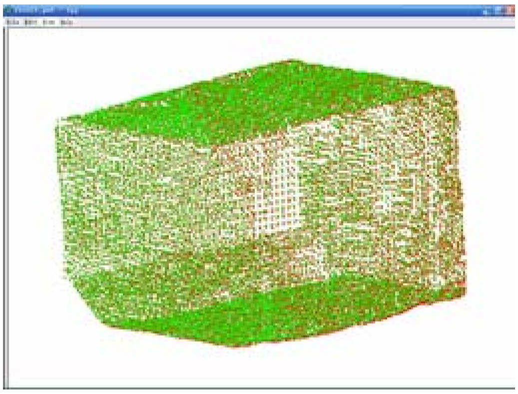

Figure 4. Matching data clouds (2)

According to our calculation, we obtain spatial data's structure matrix P_2' after transformation. We calculate spatial transform matrix T' according to the primary one. The rotation degree is: $\varphi_0 = 0.2548,\ \ \omega_0 = -0.4227,\ \ \kappa_0 = 0.5007$

$$P_2' = \begin{bmatrix} -0.7777 & -0.0957 & 0.6213 & 29.7760 \\ -0.0918 & -0.9604 & -0.2629 & 263.3917 \\ 0.6220 & -0.2616 & 0.7381 & -86.1862 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

$$T' = P_2' P_1^{-1} = \begin{bmatrix} 0.8985 & -0.3739 & -0.2299 & 45.6393 \\ 0.4378 & 0.8000 & 0.4102 & -23.1133 \\ 0.0305 & -0.4692 & 0.8826 & -64.2196 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

4.3. Analysis of result

In experiment one, we separately transform data to PAC after revolving translation for primary data. According to structure matrix, we conclude the data acquisition transformation matrix. For two groups of data is the strict homologue, the result is completely consistent with the theoretical value. In experiment two, the difference between two group's data point is obvious and strict homologue doesn't really exist. Based on PCA algorithm, we still could calculate steady the starting value of spatial transform relation. The essence of algorithm is based on the statistical information of spatial data. The requirement of statistics information is that the different space coordinates data is the relative stabilization, using statistics information to get transform matrix.

After the PCA transformation, it requires characteristic value size to establish structure matrix, namely, the vector order of structure matrix should be consistent with the order of the characteristic value. We need to judge the positive and negative direction of feature vector. As speak of symmetrical goal, it would be confused situation; we need to get further judge after transforming the data.

5 CONCLUSION

The article indicates algorithm of spatial data realization based on PCA transformation. This algorithm is aimed mainly at spatial data matching of different coordinate system but the same goal. It doesn't need to be selected artificially approximate seed spot, also avoiding complex algorithm of extracting automatically feature point. It provides an effective method for initial value's calculating of spatial parameter. The method needs spatial data to have bigger similarity based on data statistical information. In addtion, it exerts important reference value to splicing of multi-angle spatial data.

ACKNOWLEDGEMENTS

This paper is supported by National Science Foundation (50478058)and 863 Project（2006AA12Z151）

REFERENCES

[1] Chao cai, Mingyue Ding, Haolei, Jie Cao, etc. A Novel 3D Correspondence-less Method for MRI and Paxinos-Watson Atlas of Rat Brain Registration. Medical imaging and augmented reality third international workshop, Shanghai China, Aug.2006:269~276.

[2] Gustavo K Rohde, Akram A. The adaptive bases algorithm for intensity-based non-rigid image registration. IEEE Trans on Medical Imaging, 2003,22(11):1470~1480.

[3] Lisa Gottesfeld Brown. A Survey of Image Registration Techniques. ACM Computing Surveys, 1992,24(4):325~376.

[4] Besl P J, Mckay N D. A method for registration of 3-d shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239-256.

[5] Zai ruifang, mini cultural relic reconstruction research of laser dot clouds and digital film [D], dissertation of Wuhan University

[6] Zhu yanjuan, Zhou laishui, Zhang liyan, Scattered in disorder spot cloud data matching algorithm[C]. Computer-aided design and graphics journal, 2006, 18(4), 475-481.

[7]Luo xianbo, Zhong yuexian, Li renju, three dimensional scan system's data matching technology[C]. Journal of Tsinghua University (natural sciences version), 2004, 44(8), 1104-1106.

[8] Li deren, Zheng Zhaobao, Analysis photogrammetry[M]. Beijing: Survey Press. 1988