

AUTOMATIC DIVISION OF CENSUS DISTRICT BASED ON CONSTRAINT DELAUNAY TRIANGULATION

LIU D., MA W., CHENG J.

Chinese Academy of surveying and Mapping, BEIJING, CHINA

ABSTRACT

Division of census district is an import step in population census, which costs a lot of man power, finance and resource. Based on the theory and technology of delannay triangulation dividing, the census district is automatically divided with line feature as constraint condition. The paper introduces the characteristic of delaunay triangulation, discusses the technologies, measures and processes of census district division. The application example is given finally.

Key words: population census; census district; delaunay triangulation

1 INTRODUCTION

Population census is held in a country in the uniform method, item, question form and time, asking everyone living and working on the country to information about their personal lives [1]. The population census is taken once every 10 years in China and the sixth nationwide census started in November 1, 2010. About 90 percent of the people will be asked to fill in an 18-item form, covering their name, sex, ethnic group, household registration, and education; the other 10 percent, chosen randomly, will be asked to fill in a longer 45-question form. In the population census, division of census district is an important step in the work flow of census. Census district is a foundation for organization and implementation of census, reliable basis to ensure the integrity of enumeration area, efficient way for establish a quality control system of census data and a data source to build a population GIS to analysis the census data[2]. Division of census district had been done manually and costs a lot of man power, finance and resource in the previous censuses in China. Because of the difference in their quality, enumerator's map sometimes is hard to accurately depict and position the responsible district for an enumerator and the data source is hard for the further geo-spatial analysis. Therefore, automatic division of census district by digital support has an important meaning in the census process and application of census material.

The description and division of census district has different way in different countries. In Japan, regular grid is used to describe the census district, which has a simple structure, is easy to manage but is difficult to convert to other district system precisely, such as political area system or other management area system[3]. In United States, a Tiger(Topological Integrated Geographical Encoding Reference System) is used to match the house address and census question form, to fulfill the task of question form positioning and census district division[4]. Many developing countries is still manually depicts the census map and divides the census district.

Many work related to the application of Delaunay Triangulation in the field of surveying and mapping in map generalization and construction of Delaunay Triangulation has been done. Prof. Ai presents a method which applies constrained Delaunay TIN built in street regions to retrieve center-lines and further to generate street network structure. Three kinds of triangles are distinguished and corresponding different center-line extracting ways are offered. The spatial relations among street edges, street junctions and street blocks are described by graph element edge, node and loop. The method also supports to find the matching neighbor between center-lines and street block arcs. The provided data model unified the street network and block polygon [5]. Prof. Guo discusses the algorithm of polygon automatic generalization under digital environment and takes into account the orthogonal characteristics of urban building [6]. Mr. Liu discusses the arithmetic applied Delaunay triangle networks in polygons generalization and punctate objects generalization of land data type, describes the progress and the experimentation result of the arithmetic. He also analyzes and compares the several arithmetic of applying in punctate objects and introduces the realized progress [7]. For the Delaunay triangulation dividing, Mr. Yi gives a method to improve performance of Delaunay triangulation dividing with constraint line. Firstly, it is to generate Delaunay triangulation grid without constraint by parallel algorithm. Secondly, it is to insert new points by getting intersection point of circle defined by the triangulation's three vertices and constraint line [8]. Prof. Li presents a multiple diagonal exchanging algorithm for inserting constraint boundary in constraint delaunay triangulation [9].

2 METHODS

Based on the theory of Delaunay triangulation, its characteristics are that the area is unique, fully coverage of all area and less impact of one area modification to another area. It proves that the technology is suitable for the dividing of census district. In practice, the boundary of census district is divided by the line feature such as river, road, street, fence, etc, in order for the enumerator to distinguish his responsible enumeration area. The technology of Delaunay Triangulation with constraint condition is introduced for the dividing of census district. Based on the theory and technology of Constraint Delaunay Triangulation, GIS can play an important role in the automatic division of census district. The methods and processes are as follows:

(1) Extraction of point data

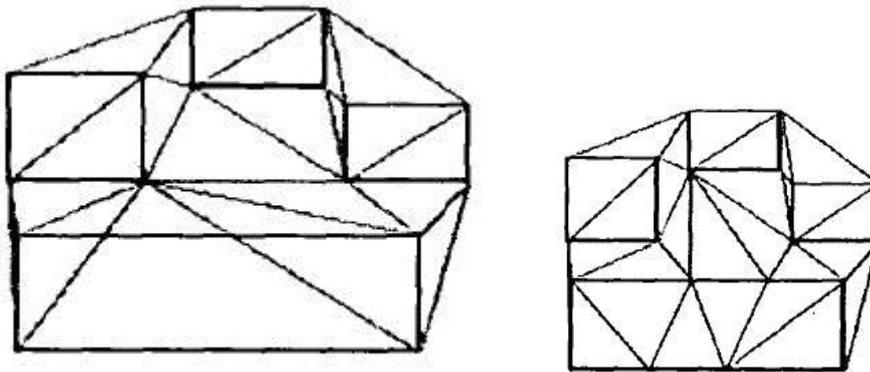
The data for division of census district includes building, administrative boundary, road and railway, river and fence, etc. The building has the attribute of number of household which is obtained in the pre-census. The point data from this feature are the generatrix to create the triangulation network and data from the administrative boundary, road and railway, river and fence are used as the constraint conditions. When the length of line from these points is too long, the side of triangulation will intersect with the original line feature and can not reflect the correct neighborhood relationship of the features (as shown in Figure 1a). In this case, intensification of point in the line is need and the intensified points are used to construct the triangulation network (as shown in Figure 1b). The method is as follow:

Suppose the points in polygon as $\{P_i\}$, when $L_{i,i+1} > w$ (w as the average width of street[5]), the points in a line must be intensified. The points intensified are $\{Q_k\}$:

$$X(Q_k) = X(P_i) + \{[X(P_{i+1}) - X(P_i)] * (k-1) / (n-1)\}$$

$$Y(Q_k) = Y(P_i) + \{[Y(P_{i+1}) - Y(P_i)] * (k-1) / (n-1)\}$$

$$\in [1, n], n=2+\text{int}(|P_i P_{i+1}|/w).$$



a. Construction of network from original points
b. Construction of network from intensified points

Figure 1 Construction of triangulation network

(2) Construction of constraint Delaunay triangulation network

After the point is intensified, triangulation network is constructed based on the technology of Delaunay triangulation.

the triangulation network can be built. The steps are described as follow:

- 1) Define a elementary polygon contained all points;
- 2) Select a point from a point set, insert to elementary polygon to form a elementary triangulation network;
- 3) Insert the next point P, search the triangle contained the point P, join the P and the 3 vertexes of the triangle;
- 4) Perform the step 3) repeatedly as all points are processed.

The data structure is a key issue of Delaunay triangulation. Considering the retrieve of boundary of census district, the paper design the data structure for the triangulation, vertex and side are as follows:

1) Data structure of triangle vertexes:

```
class t_point
{
public:
double x;
```

```

double y;
double z;
int isroad;
int parented
t_point();
virtual ~ t_point();
};

```

2) Data structure of triangle sides:

```

class t_line
{
public;
long p0;
long p1;
int useCount;
bool friend operator=(t_line line0,t_line line1)
t_line();
virtual ~ t_line();
};

```

3) Data structure of triangle:

```

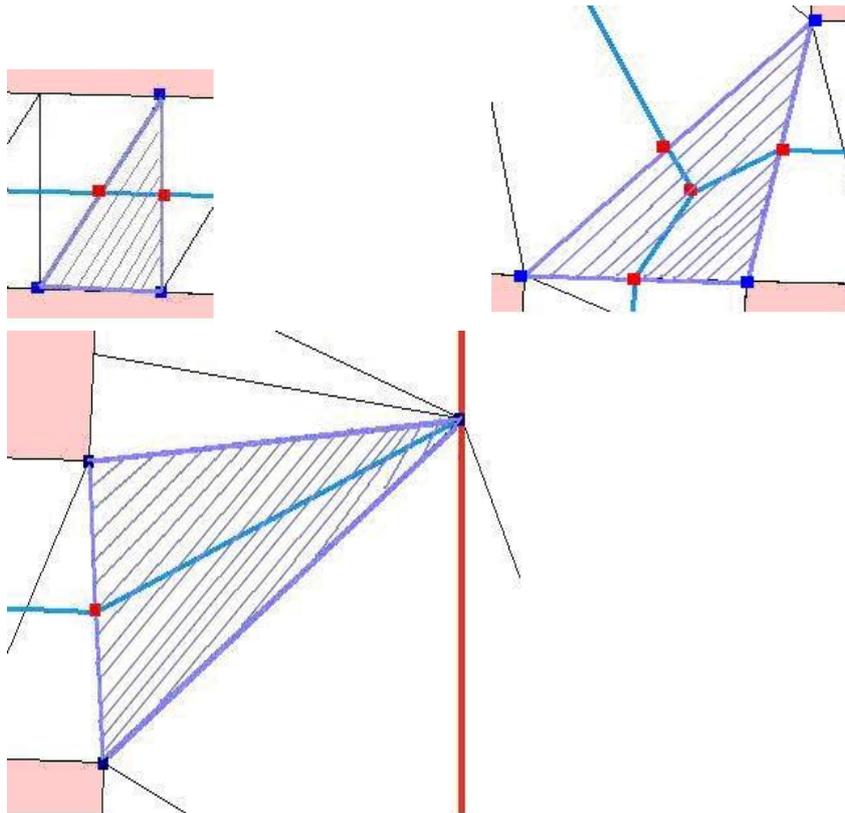
class t_tri
{
public;
long l0;
long l1;
long l2;
t_tri();
virtual ~ t_tri();
};

```

(3) Extraction of area boundary

There are 3 situations when area boundary is extracted, which are three vertex to construct a triangle from two different buildings (A class), three vertex to construct a triangle from three different buildings (B class), three vertex to construct a triangle which two point are from the different buildings and one point is from the constraint feature (C class), as shown in Figure 2. The rules and processes are described as follows:

- 1) A class: join the middle point from two lines connected two buildings;
- 2) B class: join the center point of triangle and 3 middle points of sides of triangle;
- 3) C class: join the point in constraint feature and middle point of line connected other 2 points. 4) In other cases, no process is done.



A class triangle B class triangle C class triangle

Figure 2 Extraction of area boundary

(4) Merge of area

After the above processes, the original census district is formed which contained only one building. According as the rule for the division of census district in the sixth national population census of China is about 80 households, In many cases, the original census district must be merged (as shown in Figure 3). The steps and processes are as follows:

Loop the all area, for each one:

- 1) if number of household > 80, no process;
- 2) if number of household < 80, search the neighbor area and merge but must meet the conditions: (a) no intersection with constraint feature; (b) closer to the number of 80 household.



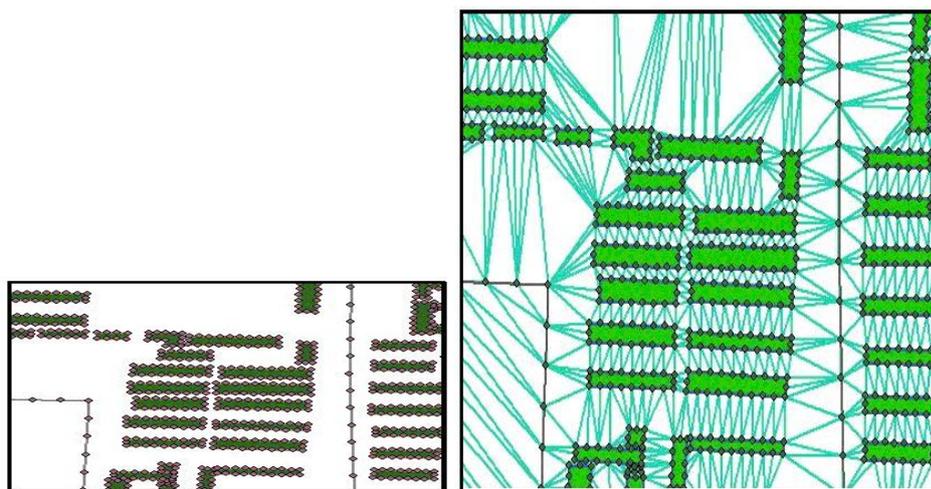
a Before merging

b after merging

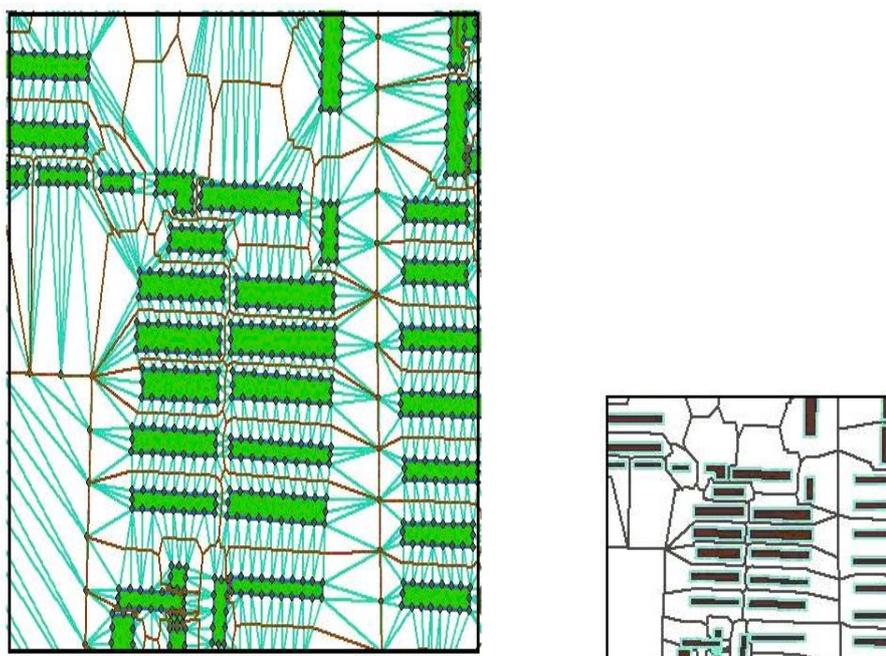
Figure 3 Merge of area

3 EXAMPLE

In the test city of Panjin, Liaoning Province in northeast China, a test site for division and mapping of census district of sixth National Population Census, the technology has been processed. The platform for software development is Visual Studio 2005, C++. The program has realized the functions such as input of the vector data, editing and process of data, division of census district, display and print out of enumerator's map, etc. The example of process and result is shown in Figure 4.



(a) Extraction of point data (b) Construction of constraint Delaunay triangulation network



(c) Extraction of area boundary (d) Merge of area

Figure 4 example of census district division

4 CONCLUSIONS

According to the principle of census district division of sixth National Population Census in China, the paper presents a method of automatic division of census district based on the theory of Delaunay triangulation, with the line features such as administrative boundary, road, street, river, fence, etc as constraint conditions. In practice, the method can ensure the no overlay and no miss of enumeration area, as an effective mean to guarantee total coverage of enumeration. It can also eliminate the error in the drawing enumeration map made by subjective carelessness of enumerator. Therefore, it is greatly alleviate the work load of enumerator. The data source is a great help to establishment a Census Geographic Information System, which will broaden the application and elevate potential value of population data.

REFERENCES

- [1] Qiao xiaochun. Research on China Population Census[M]. Chinese Population Publication House, 1995
- [2] Cheng gang. An approach on Division of Census District[J]. Study of Population, 2002,(3)

- [3] Wang zhonglian. Introduction of Japanese Population Census[J]. Study of Japan, 1990,(2)
- [4] Chen wanqing. Introduction of Population Census in the United States[J]. Study of Statistics, 2009,(10)
- [5] Ai Ting-hua. Extracting Center-lines and Building Street Network Based on Constrained Delaunay Triangulation[J]. ACTA GEODAETICA et CARTOGRAPHICA SINICA, 2000,(11)
- [6] Guo renzhong. Simplification and Aggregation of Building Polygon in Automatic Map Generalization [J]. Journal of Wuhan Technical University of Surveying and Mapping, 2000,(2)
- [7]Liu xonglin. Delaunay Triangulation Applying in the Polygons Generalization and Punctate Objects Generalization [J]. Journal of Zhengzhou Institute of Surveying and Mapping, 2007,(8)
- [8] Yi faling. Delaunay Triangulation Dividing Optimal Algorithm with Constraint Line[J]. Computer Engineering, 2001,(6)
- [9] Li li-Xin. Multiple Diagonal Exchanging Algorithm For Inserting Constrained Boundary In Constrained delaunay Triangulation[J]. CHINESE J.COMPUTERS, 1999,(10)