Automatic Assembling of Cadastral Maps Based on Generalized Hough Transformation

Fei Liu, Wataru Ohyama, Tetsushi Wakabayashi, and Fumitaka Kimura

Mie University, Faculty of Engineering,
1577 kurimamachiya-cho, Tsu-shi, Mie, 5148507, Japan
{liufei, ohyama, waka, kimura}@hi.info.mie-u.ac.jp

Abstract. There are numerous cadastral maps generated by past land surveying. The raster digitization of these paper maps is in progress. For effective and efficient use of these maps, we have to assemble the set of maps to make them superimposable on other geographic information in a Geographic Information System. The problem can be seen as a complex jigsaw puzzle where the pieces are the cadastre sections extracted from the map. We present an automatic solution to this geographic jigsaw puzzle, based on the generalized Hough transformation that detects the longest common boundary between every piece and its neighbors. The experiments have been conducted using the map of Mie Prefecture, Japan and the French cadastral map. The results of the experiment with the French cadastral maps show that the proposed method, which consists of extracting an external area and extracting and regularizing the north arrow, is suitable for assembling the cadastral map. The final goal of the process is to integrate every piece of the puzzle into a national geographic reference frame and database.

1 Introduction

In the process of application development of Geographic Information System (GIS), the core issue is the maintenance of national spatial databases. Cadastral maps are always made individually in each district. When the cadastral maps are made in each district, then they are finally combined in correspondence with the connected components of information and the surveyed ground information. The cadastral maps are chiefly assembled according to symbols given to the cadastral map and using prior knowledge available. Automatic map assembling is important because it saves labor power and time. Techniques for treating the separated cadastral maps as one piece of a huge jigsaw puzzle has been proposed by other researchers in the automation of the cadastral map assembling [1],[3],[4]. These techniques are based on the choice of some characteristics on the boundary of each segment [2], and combining segments by the matching results of the characteristics. Though the computational complexity is low because only a number of characteristics are used, when the shape of the segment becomes complex, the characteristic cannot be accurately extracted. Since this problem arises, in this paper, we propose automatic cadastral maps assembling technique based on generalized Hough transformation (GHT). The features of this technique is as follows.

- (i) Because the common longest boundary of two segments is determined based on GHT, segments can be accurately combined.
- (ii) It use all the points on the boundary of the segment, hence the extraction of the characteristics of the boundary becomes unnecessary.
- (iii) A smaller error margin of the assembled maps could be realized.

When a feature is detected from an image, the Hough transformation is effective when the feature is expressed by the algebraic equation. In this paper, to detect the curve of arbitrary shape, we propose the method of the common longest boundary detection based on generalized Hough transformation.

This paper is organized as follows. The next section presents the principle of cadastral map assembling. Section 3 describes the assembling process of blank map. In section 4 we present the preprocessing and the assembling by generalized Hough transformation of actual cadastral map. The experiments and the results are given in section 5.

2 Principle of Cadastral Map Assembling

2.1 Longest Common Boundary

When two segments share and parts are connected to each other, the border is called a common boundary. In general there is a possibility that two or more common boundaries exist between two segments. It is necessary to decide the common boundary to connect two segments correctly from these two or more common boundaries. For instance, two or more common boundaries can exist between segment A and segment B as shown in figure 1. A true common boundary is a common boundary shown by the arrow of figure 1 (c). In the proposed technique, a true common boundary was presumed based on the length of the common boundary and the longest common boundary is detected by generalized Hough transformation (GHT), then segments are automatically assembled.

Fig. 1. Two segments and the common boundary

2.2 The Flow of GHT

The flow of GHT with two segments is shown in figure 2. The longest common boundary detection technique that using GHT can be described as follows.

- (1) The boundary of each segment is traced.
- (2) The origin in border B is randomly defined and a symmetrical border B' with the point of origin is generated.

- (3) GHT is performed by sweeping the border B' with the origin on border A. Pixel values of GHT are called crossing counts.
- (4) The point where the crossing count is the highest is detected by the image of GHT. This point becomes the origin of segment B when each segment is connected by the common longest boundary.

Figure 3 illustrates the GHT between segment A and segment B. In this figure, the gray level is higher as much as the crossing count. The brightest point P in figure 3, at which the crossing count is the largest, becomes the origin of segment B.

Fig. 2. The flow of generalized Hough transformation

Fig. 3. Generalized Hough transformation image

3 Assembling of a Blank Map

The procedure of assembling two or more segments from the blank map by generalization Hough transformation (GHT) hill climbing is described as follows. The outermost shape when all segments are assembled is assumed to be known and given as a frame.

- (1) GHT is done between the frame and each segment, and the highest crossing count of each segment is found.
- (2) The segment with the highest crossing count is added to the frame and a new frame is generated.
- (3) Same procedure to the new frame and the remaining segment is done.
- (4) Until all segments are added to the frame, the procedure described above is repeated.

Table 1. The process of the assembling of the blank map

Crossing count	Segment						
	A	B	C	D	E	F	G
frame 1	502	117	74	38	53	127	93
frame 2	×	190	74	97	53	127	93
frame 3	×	×	93	126	105	127	93
frame 4	×	×	166	126	105	×	93
frame 5	×	×	×	126	207	×	93
frame 6	×	×	×	181	×	×	149
frame 7	×	×	×	×	×	×	149

Table 1 shows the procedure of assembling seven segments and outermost frame 1 of Mie Prefecture. The numerical values in the table show the highest crossing count of each segment. First of all, in order to assemble a blank map using GHT, the highest crossing counts of frame 1 against segment A to G are computed. Because the highest crossing count of segment A is the largest in this example, segment A is added to frame 1 and frame 2 is created. Then, GHT is done between frame 2 and the remaining segment, the highest crossing count of each segment is calculated. The segment with the highest crossing count is added to the frame and a new frame is generated. The white background map is assembled again repeating the above described procedure.

In this example, each segment was assembled in order of A→B→F→C→E→D→G.

4 Assembling of Actual Cadastral Map

In this section, the actual cadastral map assembling procedure, the needed preprocessing and the assembling by generalized Hough transformation (GHT) is described. There are some differences in an actual cadastral map and the blank map.

- (1) Various characters and symbols exist together in an actual cadastral map. Especially, the characters and symbols are chiefly included in an internal area of the cadastral map. Because these might become noises in the common boundary detection of the cadastral map, it is important to remove them.

(a) Original image

(b) Labeling of external area

(c) Color reversing

Fig. 4. Painting procedure of internal area

- (2) Because each segment was made individually in each district, the direction of the north arrow of each segment is different. Every segment has a north arrow, it is necessary to extract this north arrow, and to regularize direction of the north arrow of each segment.
- (3) Because the shape of the frame is unknown, it is necessary to employ GHT between each segment.

The procedure of the assembling of an actual cadastral map can be summarized as:

- Painting out of characters and symbols included in internal area.
- Extraction of north arrow.
- Regularization of the direction of the north arrow.
- GHT.

4.1 Painting Out of the Internal Area

The following are the main information included in the cadastral maps (Refer to figure 4(a)).

- The borderlines.
- Map symbols(chiefly, house sign and broken line).
- Coordinate values(numbers).
- Name of a places(characters).

These information are chiefly located in an internal area of the cadastral maps. Because small connected component that can become noise in the assembling process which included in an external area of the cadastral map, erosion and dilation of pixels of the original image are done to remove them. Next, an external area is painted out with the labeling of 4-neighborhood (figure 4(b)). In addition, the painted background is reversed, and painted out image is obtained (figure 4(c)).

4.2 Extraction of North Arrow

The north arrow is extracted by using the image generated in 4.1. The north arrow is independently located in an external area of each segment. hence, connected components of the map images are detected by the labeling process. The connected components are extracted as a candidate of connected components of the north arrow if its area is between two thresholds that are predefined. The candidate of connected components often is straight lines and curve lines excluding the north arrow. Therefore, a ratio R is given in expression (1). Where S is the area of connected components and C is the length of the surroundings of each connected components. The straight line and the curve, neither the area nor the surroundings length change so much, but because there is a width in the north arrow, the area of north arrow is larger than the surroundings length. That is, the value of R of the north arrow is larger in general than the straight line and the curve, the candidate of connected components with the largest value of R is presumed as a north arrow and extracted.

$$R = \frac{S}{C} \quad (1)$$

4.3 Regularization of North Arrow

The direction of north arrow is regularized by using the extracted north arrow. The direction of the principal axis of the north arrow is shown in figure 5, in other words, the angle of the north arrow is obtained as a solution of the expression (2).

$$\tan^2 \theta + \frac{M(2,0) - M(0,2)}{M(1,1)} \tan \theta - 1 = 0 \quad (2)$$

Fig. 5. Angle of north arrow

Here, $M(p, q)$ is a moment feature, and it is defined by the expression (3).

$$M(p, q) = \sum (i - i_g)^p (j - j_g)^q f(i, j) \quad (3)$$

Here, $f(i, j)$ is a binary image, and i_g, j_g is center of gravity. When the angle of the north arrow is obtained, the cadastral map is rotated around the center, and each cadastral map was regularized so that the angle of the north arrow is 90 degree.

4.4 Generalized Hough Transformation

After painting out the internal area, the extraction and the regularization of the north arrow, GHT was done. Here, this section explains the process of the map assembling with an example of the cadastral map of France (Fig.7). Table 2 shows the process of map assembling with five segments.

Table 2. The assembling process of the French cadastre map

Crossing count	Segment				
	A	B	C	D	E
A	×	811	888	467	1064
A ∪ E	×	811	886	720	×
A ∪ E ∪ D	×	1033	×	1191	×
A ∪ E ∪ D ∪ B	×	1033	×	×	×

An arbitrary segment (i.e segment A) is selected instead of the frame in the assembling of blank map. GHT is done between segment A and each segment, and the highest crossing count was calculated. The segment with the highest crossing count is connected to segment A and it is assumed to be a new segment. Because the highest crossing count of segment A and segment E is the highest in this example, segment A and segment E are connected. Next, GHT was done between the new segment and the rest of segment. The above procedure is repeated, and the cadastral maps are assembled.

The order of assembling was $A \rightarrow E \rightarrow C \rightarrow D \rightarrow B$ in this example.

5 Experiments and Results

5.1 Experiment of Assembling Blank Map of Mie Prefecture

The experiments of map assembling were done using blank map of Mie Prefecture (7 segments, 14 segments, and 28 segments). Figure 6 shows the seven segments and frame (figure 6(a)) in the blank map, and the assembling results (figure 6(b)).

The figure in figure 6(b) shows the order through which the assembling was done. Similarly, experiments for assembling to 14 segments and 28 segments show that all segments were correctly assembled.

Fig. 6. Seven segments, the frame and the result of assembling

5.2 Experiment of Assembling French Cadastral Map

5.2.1 Experimental Data

The experimental data included images of the cadastral maps generated by a scanner. A total of 51 segments of France were used from ten cities of the cadastral map for the experiments. Table 3 shows the number of segments of each city.

Table 3. The data used in French cadastral map assembling

City	Segment number
Bailly	5
Buc	8
Bourg	2
Bios	7
Bougival	5
Boul	3
Carrieres	7
chesnay	5
A	4
Villiers	5
Total	51

5.2.2 Extraction of North Arrow

The extraction experiments of the north arrow were done by using the above-mentioned 51 segments, and all the north arrow were correctly extracted.

5.2.3 Experiments for Map Assembling by Generalized Hough Transformation

The experiments for map assembling were done using cadastral maps of ten cities. Figure 7 shows the experimental results for five segments of Villiers city. Table 4 shows all the experimental results. The success rate in this table was visually judged to be considerably high.

Table 4. The result of the cadastral map assembling

Average success rate		Average processing time
Extraction of the north arrow	Assembling	
100%(51/51)	90.0%(9/10)	720s

In the experiment, the map assembling was failed for a city out of ten cities. It is because the combination of two segments with the longest common boundary found in generalized Hough transformation image was a wrong combination. It is scheduled to introduce mechanisms of error detection and recovery as a topic in the future work.

5.3 Summary of Experiment

Based on the experimental results, it can be summarized as follows:

- (1) It was possible to automatically assemble the cadastral maps by using generalized Hough transformation.

Fig. 7. The result of map assembling of Villiers city

- (2) The assembling of the cadastral maps in which various characters and symbols existed together became possible by painting out an internal area of the segments.
- (3) Automatic map assembling was done even when there was an error margin in the segment shape and the shape of the frame was unknown.

6 Conclusion

In this paper, a technique of the cadastral map assembling that uses generalized Hough transformation was proposed, and experiments were done with blank maps of Mie Prefecture, Japan and French cadastral maps. It was shown that generalized Hough transformation is an effective technique for cadastral map assembling.

The following areas might be included in the future work.

(1) Performance evaluation for large amount of cadastral maps. The performance of the proposed technique will be evaluated by using large amount of cadastral map, and practical use will be examined. (2) Introduction of recognition systems of other information included in cadastral maps such as characters and numbers.

References

- [1] Jean-Marc Viglino, Laurent Guigues: "Cadastre Map Assembling: a Puzzle Game Resolution ", Sixth International Conference on Document Analysis and Recognition 2001 pp.1235-1239(2001)
- [2] Davis, L.S.: "Shape matching using relaxation techniques." IEEE-PAMI 1,1, January 1979, 60-72 (1979).
- [3] L.L. Scarlatos (1999). "Puzzle Piece Topology: Detecting Arrangements in Smart Object Interfaces", Proceedings of the 7th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media, pp.456-462(1999).
- [4] Alex Drogoul and C. Dubreuil. "A distributed approach to n-puzzle solving". In Proceeding of the Distributed Artificial Intelligence Workshop (1993).
- [5] D.H. Ballard. "Generalizing the Hough transform to detect arbitrary shapes." Pattern Recognition 13, 2, 1981, 111-122