The Journal of GIS Association of Korea, Vol. 15, No. 4, pp. 385-395, December 2007

Vector Map Data Watermarking Method using Binary Notation

Jung-yeop, Kim* · Soo-hong Park**

ABSTRACT

As the growth of performance of the computer and the development of the Internet are exponential, sharing and using the information illegally have also increased to the same proportion. In this paper, we proposed a novel method on the vector map data among digital contents. Vector map data are used for GIS, navigation and web-based services etc. We embedded watermark into the coordinate of the vector map data using bit operation and extracted the watermark. This method helps to protect the copyright of the vector map data. This watermarking method is a spatial domain method and it embeds the watermark within an allowable error. Our experiment shows that the watermark produced by this method is resistant to simplification and translation.

Keywords : Watermarking, copyright protection, vector data, binary notation, GIS

요 약

컴퓨터의 사용과 인터넷이 발달로 인해 데이터의 사용과 공유가 매우 증가하고 있으며, 그에 따라 불법적인 데이터의 보급도 발생하고 있다. 본 연구에서는 이러한 불법적인 데이터의 복제 문제를 해결하기 위해 디지털 워터마킹 기법을 제안한다. 특히, GIS에서 많이 사용하고 있는 데이터인 벡터데이터에 워터마크를 삽입하고, 소유권을 주장할 수 있는 워터마킹 방법을 제안한다. 연구에서 제안한 방법은 벡터데이터의 좌표에 이진 연산을 이용하여 워터마크를 직접 삽입을 하고, 워터마크를 삽입하는 역과정을 통해 워터마크를 추출하는 것이다. 실험 결과를 통해 제안한 방법이 벡터데이터에 대한 다양한 공격에 대해 강인함을 알 수 있었다.

주요어 : 워터마킹, 저작권 보호, 벡터 데이터, 이진 연산, GIS

* Ph. D. Candidate, Department of Geoinformatic Engineering, Inha University (jyfloo@empal.com)
** Associate Professor, Department of Geoinformatic Engineering, Inha University (shpark@inha.ac.kr)

1. Introduction

Today in the computer era, the usage of computer networks is a significant amount. Internet has become the world library of information and it has allowed people to share different kinds of information such as images and sounds. In particular, the digitized data can be transferred more reliably without any loss. Since the digital information can be transmitted without any loss of data, the need for the security of the data becomes undeniable.

Even though there have been more research and results in the security of the data, watermarking is regarded as a reliable security technique. This technique has been standardized, however, there were other techniques which provide the same protection. There were various watermark algorithms developed, which were meant to change the tone in the images, videos and audios. Those algorithms are mostly frequency domain watermarking technique such as DFT (Discrete Fourier Transform), DCT (Discrete Cosine Transform), and DWT (Discrete Wavelet Transform). While previous studies were mainly on the multimedia data, relatively there were few studies on vector map data used in GIS (Geographic Information System). Also, previous watermarking methods cannot be directly applied to the vector map data because the vector map data consists of many points, polylines, and polygons and also the data structure is different from the other data [3].

Vector map data have become quite popular in the last few years. These data are used not only in GIS, but also in various fields. For example, they are used in car navigation systems (Telematics) and web-based map services. GIS-data represent a high material value due to its high efforts and costs in the acquisition and maintenance of point coordinates [1].

A GIS is a system for capturing, storing, analyzing and managing the data and the associated attributes which are spatially in reference with the earth using computer techniques. In Korea, many databases for GIS communication were constructed by several local bodies and enterprises, and the Korean government too have begun NGIS (National GIS) project. Even though there are more advantages in the construction of the data and the protection of the data using this algorithm, there is little awareness regarding the copyright protection of the data in GIS field. In order to enhance the copyright protection in GIS, applying watermarking for vector map data is important.

The vector and the raster structure are used to represent the spatial objects relative to the geographic information in a GIS. Raster structure is a geographic data set in which values are assigned to a rectangular array of object in two dimensions the plane is covered with a rectangular array. Vector is a data structure, used to store the spatial data. Vector map data is comprised of lines or arcs, defined by a beginning and an end point, which meet at nodes. The locations of these nodes and the topological structure are usually stored explicitly. Each structure has its own advantages and disadvantages. However, the vector map structure, which has accuracy and a good topology, can be used to represent complex realistic objects, therefore they are used for GIS.

There are many watermarking methods for the image data. These methods are applied for raster data in GIS. However, there are few watermarking studies for vector data. In this paper, a novel watermarking method for the vector map data is presented. The method embeds a bit using azimuth and binary notation, and the bit embedded evenly. Our proposed method is evaluated experimentally in terms of its robustness towards attacks. And any distortion after embedding watermark is also evaluated using RMSE.

2. Related Works

Digital watermarking algorithms for vector data were studied less when compared to the algorithm for other data. However, because of the increase in the use of the vector data, it is necessary to do research on watermark method for vector data [2].

A method, proposed by V.Solachid is et al [12], is a blind method comprises the Fourier descriptors for watermarking of vector graphic images. The 2D coordinate (x, y) of each vertex is combined to construct the complex value z = x + iy. After that, DFT can be applied to the vertices of a polyline or polygons to obtain the Fourier descriptors. The watermark can be embedded with the Fourier descriptors and could be extracted with the correlation between the embedded watermark and the watermarked Fourier descriptors. This paper shows that the properties of the Fourier descriptors ensure that the novel watermarking algorithm endures rotation, translation, scaling, reflection, change of traversal starting point/direction and smoothing. The methods [13, 14] that are using complex value for vector data and as well as watermarking methods for vector data using complex value were proposed [12].

Previous works([12], [13], [14]) changed coordinates into complex value to embed watermark. These works were regarded as the modification of the image watermarking methods.

In reference [15], a digital watermarking algorithm for vector digital map was proposed. First, the original data implemented three area subdivision methods. After the subdivision, a watermark is embedded into the map by each bit 0 or 1 in the rectangles. Prior to the extraction, a similarity transformation applied to the watermarked map is removed. The same rectangles used during the embedding are created through this preprocessing. A message bit is extracted using the algorithm which compares the averaged vertex coordinates among a corresponding pair of rectangles. This paper experimented and analyzed to show the robustness of the attacks such as translation, enlarge, shrink, vertex deletion, additive random noise and cropping. This method has defect that depends on the validity of the original map. And the method hasn't robustness of the algorithm.

Jong-Uk Choi et al. [16] proposed a novel vector map watermarking algorithm. As a first step, the image is partitioned into blocks of equal area and the vertex in a block is changed to a new coordinate value. The value of the watermark is changed as an area is inserted into the watermark. To extract the embedded watermark, the whole image is divided into blocks, and the data of the image in bits are obtained in the triangular region. The result from this experiment is robust against the noise attack.

Watermarking methods applying frequency domain method are also proposed [17, 18]. Michael Voigt et al. [17] proposed a method based on a reversible coefficient modification in the 8-point integer DCT domain. This paper confers the improvement in the coefficient modification to reduce the visual distortion for the watermarked data. This method shows good results in watermark embedding. [17] reduced the robustness to decrease the distortion of the data. Therefore, the method can be fragile on attacks. A method, proposed by Ryutarou Ohbuchi et al [18], embeds message bits into a 2D vector map by modifying a "frequency" domain representation of the map. To calculate the "frequency" domain representation, the method uses the Delaunay triangulation. Modification of the frequency domain coefficients changes the vertex coordinates in the spatial domain. The proposed method is resistant to some attacks such as additive random noise, addition of vertices, rotation, scaling, and cropping of the map. The method uses landmarks to extract watermark. Because the landmarks can be eliminate as it is attacked, this method has weakness on extracting watermark.

3. Proposed Watermarking Method

3.1 The Selection of the Objects

In proposed methods, we used azimuth to select the objects to embed watermark. Because every objects has azimuth, we could azimuth on embedding watermark.

The direction of one object from another is usually expressed as an angle in degrees relative to the true north. Azimuths are usually measured in the clockwise direction, thus an azimuth of 90 degrees indicates that the second object is due east of the first (Figure 1). In this paper, we selected objects which are embedded with watermark using azimuth. Also,

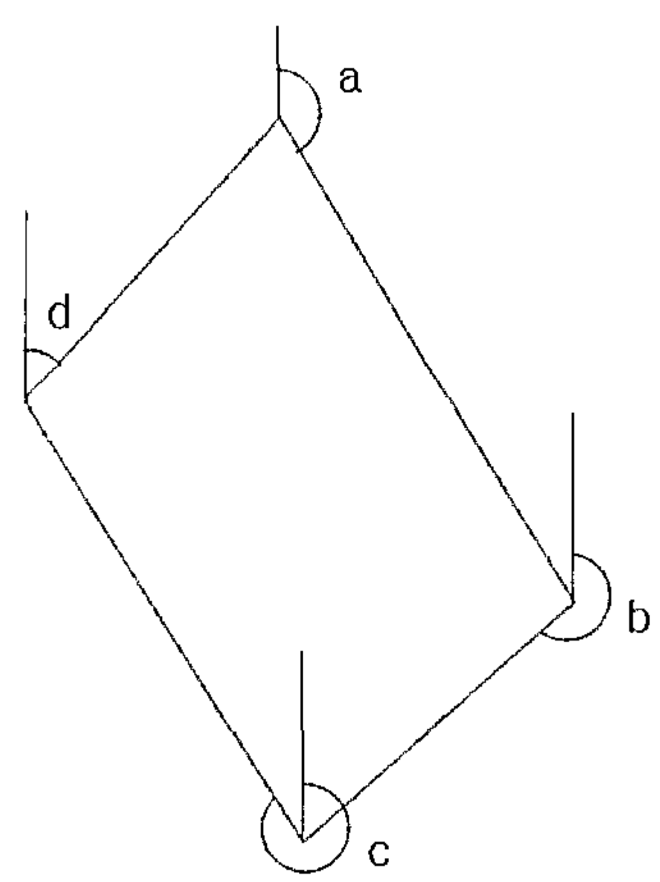

Figure 1. The concept of the azimuth (Sum of azimuth = a + b + c + d)

the azimuth can be a watermark key in our method.

3.2 Experiment Data

We used NBI map data that is served in NGII (National Geographic Information Institute). NGII make public map data of scale such as 1:1,000, 1:5,000, 1:25,000, and 1: 250,000. NBI vector map data could be supplied and utilized for users to solute the disadvantage of the DXF. The Internal Regulation of the NGII describes that the numerical map should represent coordinates as TM coordinate and record coordinates to the second decimal place. We used a numerical map to a scale of 1:1,000 NBI data. The map consists of building, soil boundary, datum point, drainage network, and road layer etc. The data describes the format as signed 64-bit IEEE double-precision floating point number.

Proposed watermarking method is to embed watermark into vector map data within an allowable error of the numerical map. Table 1 shows the allowable error of the numerical map. The vector map of the scale of 1:1,000 has an allowable error of ±0.70 m. We proposed a novel watermarking method in the spatial domain, and embedded the watermark into the coordinate points within an allowable-error. After we attack watermarked vector data, we extract watermark and calculate the extraction rate.

Table 1. Allowable error of map data

Scale	Plane position error	Altitude position error
1:500	±0.25m	±0.25m
1:1,000	±0.70m	±0.33m
1:2,500	±1.75m	±0.66m
1:5,000	±3.50m	±1.66m
1:10,000	±7.00m	±3.33m

The following is a simple description of NBI file format.

- The NBI file consists of the sections of file head, layer, head and data.
- The section of the data consists of point, linestring, polygon, multipoint, multilinestring, multipolygon, text, and network chain.

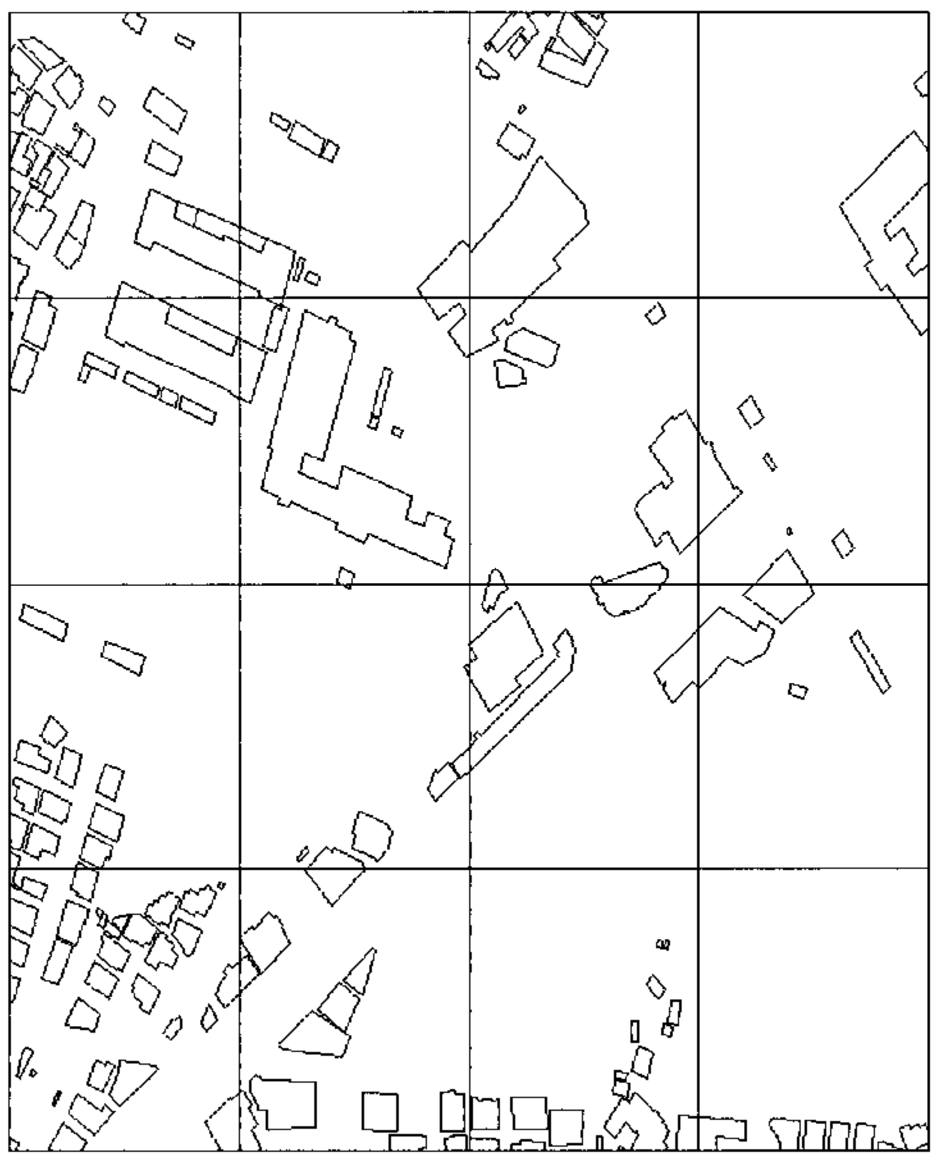

Figure 2. NBI vector map (1:1,000 scale)

3.3 Watermark Embedding

In the digital watermarking, the ideal case should be the original vector data, the water-

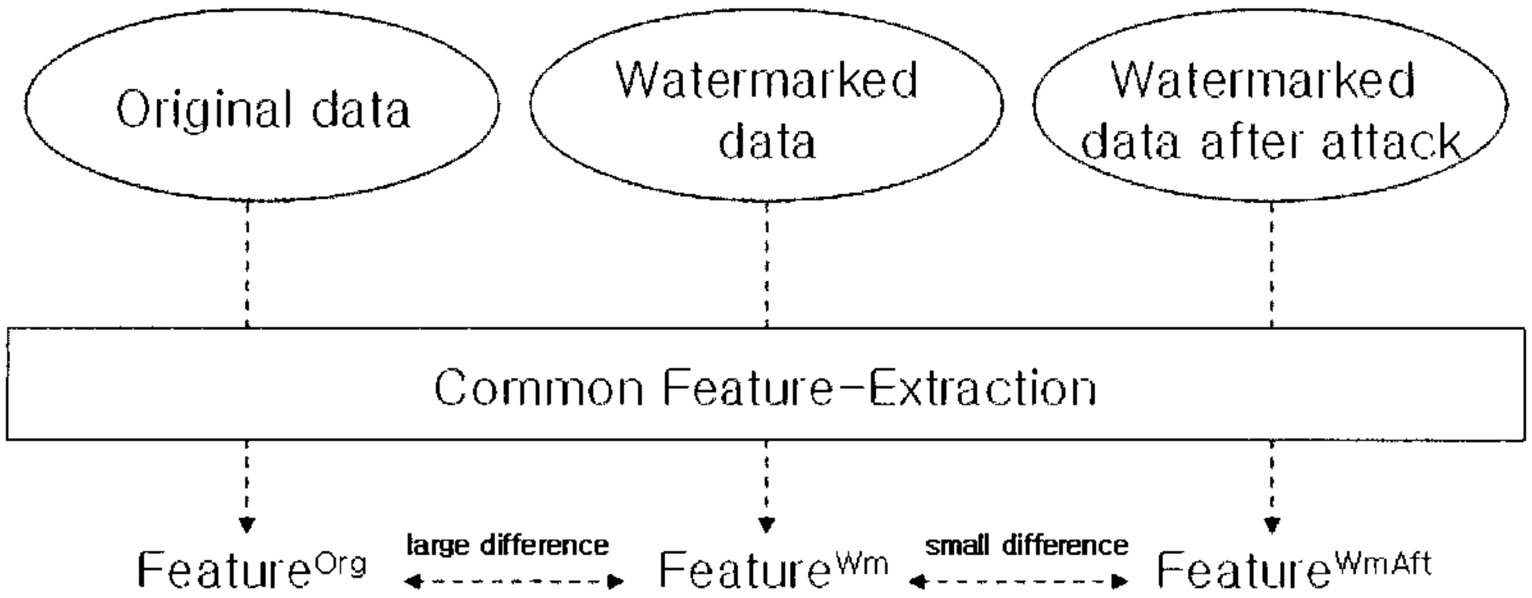

Figure 3. Overview of an ideal feature-extraction [1]

marked data, and the watermarked data after attack. In other words, there should be a significant distinction between the feature of the original data and the sets which are watermarked and attacked [1]. It is efficient to extract common feature for vector data watermarking.

In general, frequency-based approaches have more robustness because of the significant coefficients which usually remain stable even after the image manipulation. However, modifying significant coefficients will result in large distortions and also serious quality degradations [3]. Our method does not embed watermark in frequency domain, but embed directly the watermark into the original vector map data within an allowable error in spatial domain. Using azimuth, the objects, within which the watermark has to be embedded, were selected. In the first step, the sum of the azimuth (S) in each object is calculated, and all we compute is the mean-azimuth (M). According to the normal distribution, we can obtain the standard deviation. Finally, we choose objects O, which is at a range of $M \pm 1\sigma$. In the next step, the coordinate of the selected objects is changed into a binary number. And then we add one bit at the specific position of the binary number.

The embedding procedure is as follows:

① Figure out the sum of azimuth of the objects in vector data
② Calculate mean-azimuth (M) and standard deviation in vector data
③ Select objects (O) to embed watermark. i.e., select objects within $M \pm 1\sigma$
④ Change the coordinate values into binary numbers.
⑤ Adds one bit at the first decimal place if the objects are in $M + 1\sigma$, or else subtract one bit from the first decimal place if the objects are in $M - 1\sigma$.

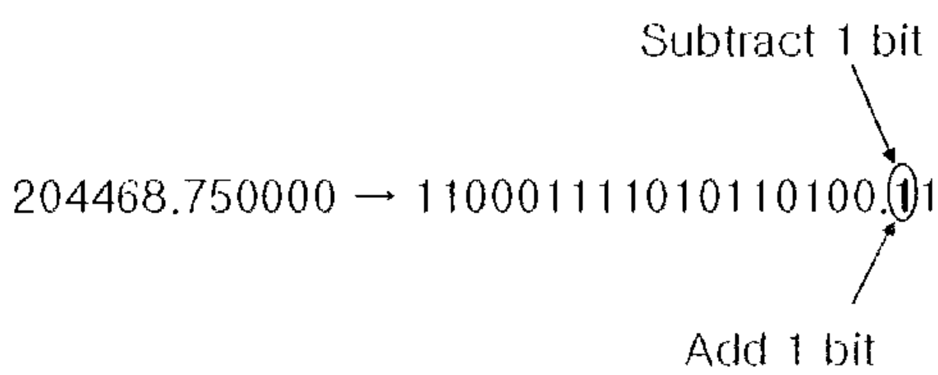

Figure 4. The example of watermark embedding

① and ② procedures are necessary in our

watermarking method. Namely, the azimuth and the standard deviation are calculated in order to choose the objects to be embedded with watermark. Watermark key is one of the features of the watermarking. Even though the watermarking algorithm is known, if the key is not know then the watermark that was embedded cannot be eliminated. In this paper, when objects were selected to embed watermark, median azimuth, specific object or ± 2 σ can be used instead of mean azimuth. For example, we can choose objects within median azimuth ± 1σ, mean azimuth ± 2σ or the azimuth of the specific object ± 1σ. As these things become watermark key, the proposed method can be robust, and the location of the watermark will be obvious. In experiment, we used mean azimuth, and the selected objects.

3.4 Watermark Extraction

To extract watermark, we use the watermark key generated in the course of watermark embedding. According to the watermark key, objects in original vector data are extracted and the same objects in attacked vector data are also extracted. Because we embed and subtract 1 bit from the first decimal place of the objects, which are within M + σ azimuth and M - σ azimuth, the second decimal place is not changed.

Therefore, the coordinate of the object extracted in original vector data (M - σ) are sorted, the coordinates of the same object in attacked vector data are also sorted. Finally, we calculate the coordinates which are the same as the second decimal places. We can apply the same extraction method about objects in case of M - σ. The following (figure 5) are simple example which shows how one can operate the coordinate of two objects.

The extraction procedure is as follows:

① Select objects using watermark key in

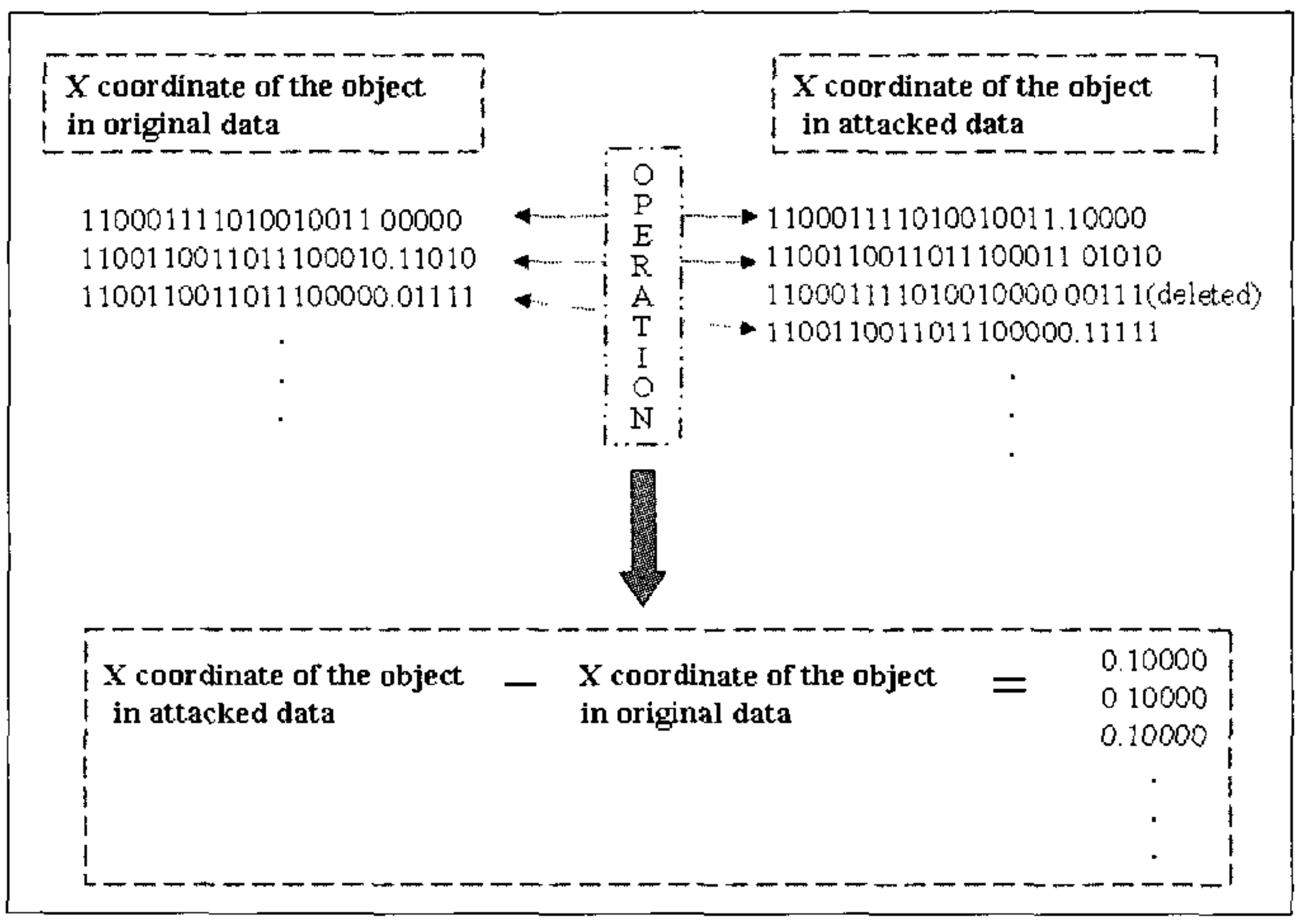

Figure 5. Example of 1 bit embedding and extraction (M + σ)

the original vector data and attacked data.

② Sort each coordinate and operate the same coordinate.

③ Check the result and count the number of correspondence coordinate.

3.5 Distortion and Extraction Rate of Watermark

Our method is able to extract the watermark from the vector map data. However, embedding watermark should be extracted after attacks. We evaluated the robustness of the watermark for simplification and translation, and the RMSE statistic is used to compare the accuracy of the original vector data with the watermarked vector data.

In general, if watermark is embedded, then distortion is generated in the watermarked data. Therefore, distortion due to the watermark should not deteriorate vector map data. We estimate RMSE, which is a measure of the accuracy of the watermarked data. Table 5 shows RMSE statistic after embedding watermark.

$$RMSE(x-axis)=\sqrt{\frac{\sum_{i=1}^{n}e_i^2}{n-1}},\quad RMSE(y-axis)=\sqrt{\frac{\sum_{i=1}^{n}e_i^2}{n-1}} \tag{1}$$

$$RMSE(coordinate)=\sqrt{(RMSE\_x)^2+(RMSE\_y)^2} \tag{2}$$

Where, $e_i = y_i' - y_i$ is error, and y_i' is coordinate value after embedding watermark, and y_i is coordinate value of original data.

In digital watermarking method, the rate of watermark extraction needs to certify whether extracted watermark owner embeds it or not. CR (Correspondence Ratio) is a simple ratio of index, which certifies the rate of watermark extraction. Here the denominator means the number of coordinate embedded watermark into original vector data. And the numerator means the number of coordinate extracted watermark. When the index is closer to 1, it represents that the watermarking method is robust.

$$CR=\frac{the\ number\ of\ coordinate\ extracted\ watermark}{the\ number\ of\ coordinate\ embedded\ watermark} \tag{3}$$

4. Experiment

We embedded watermark into NBI vector map (1:1,000 scale) and extracted the watermark. We verified the validity of extracted watermark. To test the robustness of the proposed method, we extracted watermark in attacked data after attack such as simplification and translation, which are used in GIS. We simplified vector data using Douglas-Peuker algorithm, and translated all the vertices in the vector data by the same amount.

We made watermark key using mean azimuth and standard deviation instead of median azimuth or specific object. In our experiment mean azimuth is 638.587458, and standard deviation is 342.486872. Vector data used in experiment has 169 objects, and then has 29 objects and 4901 coordinates within mean azimuth (M) +1 σ. The vector data has 97 objects and 16393 coordinates within mean azimuth (M) - σ. We embedded the watermark into coordinates and RMSE is below 0.391831. (Table 5) Figure 6 shows watermarked vector map data.

Table 5. Experiment result

RMSE	0.391831
Mean azimuth	638.587458
Standard deviation	342.486872
The number of objects	169
The number of coordinates	13446
The number of eliminated coordinates	153
The number of objects within M+1 σ	29
The number of objects within M-1 σ	97
The number of coordinates within M+1 σ	4901
The number of coordinates within M-1 σ	16393
CR after simplification	0.92510
CR after translation	1

Watermarked data could be attacked by users intentionally or not. We simplified and translated to test the robustness of the method. The rate of watermark extraction(CR) is shown in table 5.

In table 5, RMSE is 0.391831. That is, the vector data after being embedded with watermark has distortion within the allowable error. Therefore, we can infer from the results that this is a good method. CR is 0.929510 because of the simplification of the vector data. This means that although coordinates are eliminated by attack, most of the coordinates that are watermarked survived.

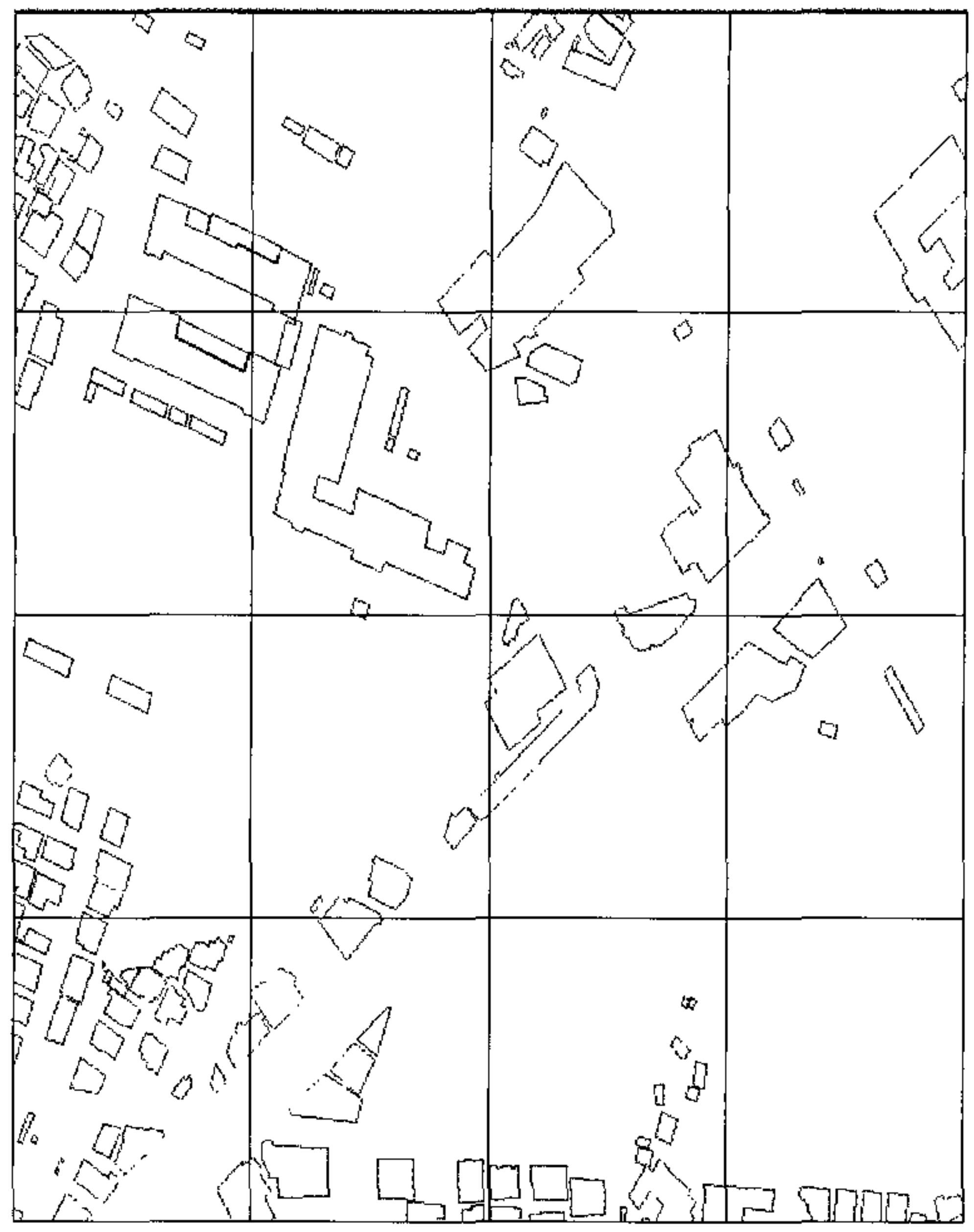

Figure 6. Watermarked vector map data

5. Conclusion

We proposed a novel watermark method in spatial domain using the binary notation. We selected specific objects to embed watermark into vector map data using azimuth, and then directly embedded watermark into the coordinates. Because watermarks are embedded within allowable error, neither problem nor deterioration is generated after embedding watermark. Also, the advantage of the method is to use simple binary operation instead of a complicated formula. Therefore, it is easy to embed and extract watermark. However, we generated watermark key and the key is used for embedding and extracting watermark, so the proposed method is robust method. The results of this experiment strengthen the validity of our model and method. Our future work will include experiment after various attacks and when watermark is embedded, we will improve watermarking method which the vector map data has not distortion. And then it needs to extract watermark without original vector map data.

References

Michael Voigt, Christoph Busch, "Feature-based Watermarking of 2D-Vector Data," Proceedings of SPIE-the international society for optical engineering, Vol. 5020, pp. 359-366, 2003.

Michael Voigt, Christoph Busch, "Watermarking 2D-Vector Data for Geographical Information Systems", Proceedings of SPIE-the international society for optical engineering, Vol. 4675, pp.621-628, 2002.

I. Kitamura, S. Kanai and T. Kishinami, "Copyright Protection of Vector Map using Digital Watermarking Method based on Discrete Fourier Transform", Geoscience and Remote Sensing Symposium, IGARSS '01, Vol. 3, pp. 1191-1193, 2001.

D-C Lou, C-H Sung, "Robust Watermarking Technique for Digital Images utilizing the Invariant Relation of Vector Quantization Indices", The Imaging Science Journal, Vol. 50, No. 4, pp. 291-302, 2002.

Jung-Soo Lee, Jong-Weon Kim, Jong-Uk Choi, "A Study On Image Watermarking Technique against Geometrical Deformation", Proceedings of The 30th KISS Spring Conference, pp. 440-442, 2003.

Jae-Wook Shin, "A new watermarking method using entropy-based region segmentation," INHA University, Master paper, 1999.

Myung-Hwan Jang, "A Study of Image Watermarking Schema using the Features of Wavelet Transform Coefficients," TAEJON University, Master paper, 2001.

Hsien-Chu Wu, Chin-Chen Chang, "A Novel Digital Image Watermarking Scheme based on the Vector Quantization Technique," Computers & Security, pp. 460-471, 2005.

Zhe-Ming Lu, "Multipurpose Image Watermarking Algorithm Based on Multistage Vector Quantization," IEEE TRANSACTIONS ON IMAGE PROCESSING, pp. 822-831, 2005.

Anamitra Makur, S. Sethu Selvi, "Variable Dimension Vector Quantization based Image Watermarking," Signal Processing 81, pp. 889-893, 2001.

Zhe-Ming Lu, Wen Xing, Dian-Guo Xu, "Digital Image Watermarking Method Based on Vector Quantization with Labeled Codewords", IEICE TRANS, pp.2786-2789, 2003.

Hyung-Do, Kim, "Blind Image Watermarking using Codebook Grouping in Vector Quantization", Telecommunications Review, Vol. 12, No. 5, pp. 663-669, 2002.

V.Solachidis, N.Nikolaidis, I.Pitas, "Fourier Descriptors Watermarking of Vector Graphics Images," In proc. of XIII Brazilian Symposium on Computer Graphics and Image Processing, Vol. 3, pp. 9-12, 2000.

Li Yuanyuan, Xu Luping, "A Blind Watermarking of Vector Graphics Images," ICCIMA 2003 Proceedings, pp. 424-429, 2003.

I. Kitamura, S. Kanai, T. Kishinami, "Copyright Protection of Vector Map using Digital Watermarking Method based on Discrete Fourier Transform", Geoscience and Remote Sensing Symposium IEEE 2001 International, Vol. 3, pp. 1191-1193, 2001.

Endoh. S, Ohbuchi. R, Ueda. H, "Robust watermarking of vector digital maps", In pro. of IEEE International Conference on Multimedia and Expo, Vol. 1, pp. 577-580, 2002.

Jong-Uk Choi, Hwan-Il Kang, Kab-Il Kim, "A map data watermarking using the generalized square mask", In proc. of IEEE International Symposium on Industrial Electronics, Vol. 3, pp. 1956-1958, 2001.

Michael Voigt, Bian Yang, Christoph Busch, "Reversible Watermarking of 2D-Vector Data", Proceedings of the 2004 workshop on Multimedia and Security, pp. 160-165, 2004.

Ryutarou Ohhuchi, Hiroo Ueda, Shuh Endoh, "Watermarking 2D vector map in the mesh spectral domain," International Conference on Shape Modeling and Applications 2003, pp. 216- 225, 2003.