An Anti-compression Fragile Watermark Scheme for Vector Geographical Data

ZHU Chang-qing*, WANG Qi-sheng*, REN Na*, WU Wei*

*Institution 1, Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Jiangsu Province, Nanjing, 210046

Abstract: The growth of applications and the drastic increase in release of geographical data makes security as a mandatory requirement. In this paper, based on Douglas-peukeer algorithm, a fragile watermark algorithm of vector geographical data is proposed. The proposed scheme simplifies the map using the Douglas Peucker algorithm to obtain the feature points. The fragile watermark is embedded in the feature points and non-feature points of the line and area using different watermark embedding and detection rules. The experimental results on vector geographical data show that the proposed watermark algorithm has good invisibility, and is robust to compression, additive noise attacks. Besides, the maximum watermarking error can be strictly controlled.

Keywords: Vector Geographical Data, Douglas Peukeer Algorithm, Fragile Watermark

1. Introduction

The demand for vector geographical data authemtication and for effective means to control vector geographical data integrity has been steadily increasing in the last years. Such a demand is due to the ease with which the data can be tampered with thus compromising their credibility as faithful data of the scene they represent. Sevel techniques have been developed in order to prevent or at least detect unwanted alteration of vector geographical data. Among them, as an advanced technology for copyright protecting of digital products, digital watermarking receives a wide attention in the protection of vector geographic data.

Digital watermark can be divided into robust watermark and fragile watermark. Robust watermark is mainly for copyright protection, while, fragile watermark is mainly used to identify the authenticity and integrity of data, and it is also known as certification.

Generally, according to the embedding position, the watermark algorithm for vector geographical data can be divided into the transform domain and the spatial domain. The literature [1] makes use of the geometric invariance in the transform of DualTree complex wavelet. The literaturesuse the geometric invariance in the transform of discrete Fourier to achieve a watermark algorithm against geometric transformation [2-4]. The watermark is directly embeded in the data coordinate in the spatial domain, so it is more flexibly and easier to detect the attacks of anti-delete points, cutting and so on [5-8]. The literature [9] splits the curve of vector geographic data in accordance with the requirements, then embeds watermark. The algorithm is a kind of non-blind watermarking algorithm which can resist a certain degree of tailoring, D-P compression or the attack of geometric transformation. However, these studies above pay little attention to the blind watermarking for the compression of vector geographic data.

Since unavoidable reasons in producing and updating the data, the line and polygon of the vector geographic data may face the problem of some redundant points. In order to reduce the store and analysis data more conductively, the vector geographical data is often compressed. Especially, in the current network environment, more data compression is needed to increase the efficiency of data transmission.

In the compression of vector geographic data, the characteristics of the data points will normally be retained. In order to achieve the resistance of watermarking algorithm on the data compressed, we can embed the fragile watermark information in the feature points of linear or planar data, which is the popose of this paper.

The work is origanized as follows. In section 2, a brief overview of Douglas-peukeer scheme is given, by paying attention to introduce the steps of the algorithm that will be used in the following sections. In section 3, the fragile watermark algorithm based on the Douglas-peukeer is discussed, and watermark detection is described in section 4. Section 5 is devoted to the presentation of experimental results. Finally, some conclusions are drawn in section6.

2. A Fragile Watermark Scheme Based on the Douglas Peukeer Algorithm

2.1 Feature Points Extraction

The initial form of the Douglas Peukeer algorithm was put forward by D. Douglas and T. Peueker in 1973. D-P algorithm for short is now recognized as the classical scheme for reducing the number of points in a curve that is approximated by a series of points, which is filter the redundant vertices of curve with advantages of translation and rotation invariances. When curve and threshold are given, the sampled results are certain.

The basic idea of the featrure points extraction based on D-P algorithm is drawing straight virtual lines for each of a first and last point of the curves, seeking the distance of all points with a straight line and find out the maximum distance value called *dmax*, and compare *dmax* with the given precision tolerance threshold *D*.

If $dmax < D$, then delete all intermediate points in this curve, otherwise, if $dmax >= D$, then retain *dmax* corresponding vertex, and set the point as a boundary, divide the curve into two parts, then use repeatedly this method on the two parts untile *dmax* less than the given precision tolerance threshold *D*.

2.2 Watermark Embedding

The steps of embedding the fragile watermark informatiion based on D-P algorithm for vector geographical data are listed.

(1) The feature points of the original data are extracted by Douglas Peukeer algorithm. Then, get the data after identify feature points. Using the Douglas - Peukeer compression to extract and identify

(2) Embed watermark with the method of quantizing the feature points. *N* is the quantization step, the value range $[0,9]$ which precision bits have are divided into n

according to $N, n=\lceil 10/N \rceil$, embed the fragile watermark according to the quantization interval of x_{bj}, y_{bj}, the rules of embedding watermark are listed.

$$\begin{cases} x_{bj} = w_j, & x_{bj} < N \\ x_{bj} = N + w_j, & N \le x_{bj} < 2N \\ \cdots & \cdots \\ x_{bj} = (n-2)N + w_j, & (n-2)N \le x_{bj} < (n-1)N \\ x_{bj} = (n-1)N + w_j, & (n-1)N \le x_{bj} \le 9 \end{cases}$$

$$\begin{cases} y_{bj} = w_{j+1}, & y_{bj} < N \\ y_{bj} = N + w_{j+1}, & N \le y_{bj} < 2N \\ \cdots & \cdots \\ y_{bj} = (n-2)N + w_{j+1}, & (n-2)N \le y_{bj} < (n-1)N \\ y_{bj} = (n-1)N + w_{j+1}, & (n-1)N \le y_{bj} \le 9 \end{cases}$$

(3) Embed watermark in the other points using the follow fragile algorithm method. N is the quantization step, the value range $[0,9]$ which precision bits have are divided into n according to N, $n=\lceil 10/N \rceil$, $\lceil\ \rceil$ means ceiling, $\lceil 10/N \rceil$ gets the maximum among integer greater than $10/N$. Then embed watermark according to the quantization interval of x_{bj}, y_{bj}, the rules are as follows:

$$\begin{cases} x_{bi} = w_{xi}, & x_{bi} < N \\ x_{bi} = N + w_{xi}, & N \le x_{bi} < 2N \\ \cdots & \cdots \\ x_{bi} = (n-2)N + w_{xi}, & (n-2)N \le x_{bi} < (n-1)N \\ x_{bi} = (n-1)N + w_{xi}, & (n-1)N \le x_{bi} \le 9 \end{cases}$$

$$\begin{cases} y_{bi} = w_{yi}, & y_{bi} < N \\ y_{bi} = N + w_{yi}, & N \le y_{bi} < 2N \\ \cdots & \cdots \\ y_{bi} = (n-2)N + w_{yi}, & (n-2)N \le y_{bi} < (n-1)N \\ y_{bi} = (n-1)N + w_{yi}, & (n-1)N \le y_{bi} \le 9 \end{cases}$$

3. Watermark Detection

While detecting watermark, the algorithm proposed in the paper does not need the original data and is a blind one. Corresponds to the watermark embedding algorithm, the basic steps for watermark detection is as follows.

(1) In accordance with the inverse method of embedding to get the watermark information of each point from the data. The extracted watermark information is recorded as w'_{xj}, w'_{yj}, which represents the extracted watermark information on abscissa

and ordinate respectively, the accuracy bit data of latitude and longitude coordinates is (x'_{bj}, y'_{bj}).

(2) If $|x'_{bj} - y'_{bj}| = 0$ or N, the corresponding data points are non-feature points (the feature points may change, but the probability is too small, here regarded as impossible). If $w'_{xj} = x'_{bj}$ and $w'_{yj} = y'_{bj}$, the corresponding data points should not be changed, otherwise, if $w'_{xj} \neq x'_{bj}$ or $w'_{yj} \neq y'_{bj}$, the corresponding data points should be changed.

(3) In the circumstances that $|x'_{bj} - y'_{bj}| \neq 0$ or N, the corresponding data points may be feature points, if $w'_{xj} = x'_{bj}$ and $w'_{yj} = y'_{bj}$, the corresponding data points are feature points, and they are not changed, so that the feature points which are adjacent in the rear are also not change. If $w'_{xj} \neq x'_{bj}$ or $w'_{yj} \neq y'_{bj}$, it cannot be guaranteed that the corresponding data points are feature points, it can be seen that the data must be attacked in addition to compression. Therefore, it can only focus until the positioning data points are changed. If the data points are the feature points, whether the subsequent feature points change or not cannot be informed, it can only make a judgment through the watermark information from subsequent data points.

4. Experiments and Analysis

In this section we show a comparative example of processing an original and a watermarked data. Our aim is to assess the effect of watermark on some attacks.

Experiment was taken on a 1:250,000 vector geographic data (6213 data points) with 177 linear features using the proposed algorithm, and the original data shown in Figure 1.

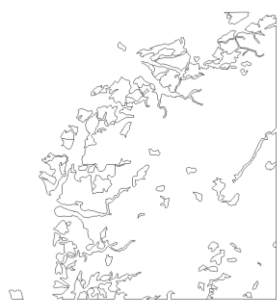

Figure 1. The original vector geographic data

It is assumed that the maximum change on the precision bit of the above data is $M = 5$.

Otherwise, the data changes should be minimized, so that we make $N = 2$, for embedding watermark will affect the extraction of feature poingts among vector geographic data. Therefore, the overall margin of data error caused by the watermark embedding is $[0, \sqrt{2}]$, and it will not destroy the data accuracy, and also do not affect the use of the data.

4.1 Invisualbility

Comparing the data before and after the watermark embedding, it is as shown in Figure 2. Figure (a) is the original data, and figure (b) is the watermarked data.

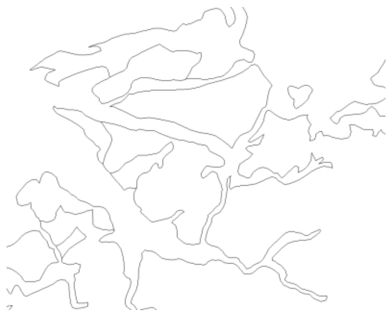

Figure (a) Origin data

Figure (b) Watermaked data

Figure 2. The visualization comparation on vector geographic data between watermark embedding before and after (local to enlarge)

From the above figure (a) and figure (b), we can see that the embedded watermark does not affect the visualization of data. So, the invisibility of this algorithm is verified well.

4.2 Error Analysis

Comparing the error of data before and after the watermark embedding, the results are shown in Table 1.

Table 1. The results of error analysis

Magnitude of the error	The number of data points	Percentage (%)
0	6101	98.20
0-1	106	1.71
1-2	6	0.09
More than 2	0	0

In Table 1, we can draw the conclusion that the error range of watermark consistent with the rage $[0, \sqrt{2}]$ of which is discussed theoretically, and it does not affect the use of the data.

4.3 Anti-compression Analysis

Data compression is sometimes used to reduce the amount of data to be transmitted to the ground station. While lossless compression is preferred, in some cases near-lossless compression can be accepted. High-rate lossy compression can also be used to preserve quality. Suitable watermark algorithm should thus be robust towards data compression.

Compress the vector geographic data afer watermark embedding, and detect the watermark from the compressed data, the results shown in Figure 3 and 4.

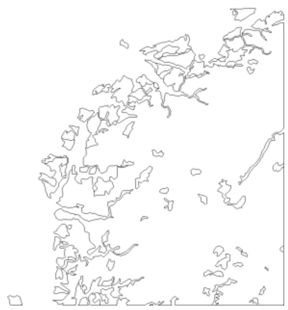

Figure 3. Detecting result after data compression (compression of 18%)

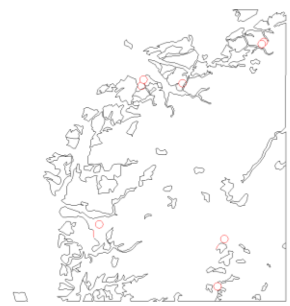

Figure 4. Detecting result after data compression (compression of 50%)

It is shown form the above figures that the algorithm can effectively resist a certain ratio of compression attacks on vector geographic data, the watermark detection results do not show tampeis under compression attacks, so that the selective authentication on the compressed data has achieved. When the compression ratio of vector geographic data is too large, it will have inevitable damage on the data, the feature point may lose. Therefore, in the case of the large compression ratio, the watermark can detect which feature point has been deleted. Further, it could be seen that if the data is compressed, we can embed watermark by extracting the feature points of data according to the ratio ranges of compression. If the ratio exceeds, regard it as tampering of the data, and selective authentication on compressing data is needed.

5. Conclusions

A fragile algorithm for generating and embedding watermark according to the difference between feature and non-feature points based on the Douglas Peukeer algorithm, we propose an semi-fragile anti-conpression watermark algorithm for vector geographic data, and make a visual comparison and error analysis on the data for the algorithm.then experiment is taken to show the the effect on selectivity certification of the algorithm. The results show that the algorithm can recognize the selectivet authentication under a certain compression ratio for vector geographic data, and detect other attacks on watermark.

6. Acknowledgements

The work was supported by the National Natural Science Foundation of China (Grant No. 41071245), Open Foundation of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing of Wuhan University (Grant No. 12I01), and the University Science Research Project of Jiangsu Province (Grant No. 12KJB420002)

7. References

[1] ZHANG Qin, XIANG Hui, MENG Xiang-xu. Watermarking Vector Graphics Based on Complex Wavelet Transform [J]. Journal of Image and Graphics, 2005，4：494-498.

[2] WANG Qisheng, ZHU Changqing, XU Dehe. Watermarking Algorithm for Vector Geo-spatial Data Based on DFT Phase [J]. Geomatics and Information Science of Wuhan University，

2011，5：523-526.
[3] MA Taolin, GU Chong, ZHANG Liangpei. Watermarking Algorithm on 2D Vector Digital Maps[J]. Geomatics and Information Science of Wuhan University，2006，9：792-794.
[4] MIN Lian-quan. A Robust Digital Watermarking in Cartographic Data in Vector Format [J]. Acta Geodaetica et Cartographica Sinica，2008，5：262-267.
[5] ZHANG Hong-sheng, LI Yan, CAO Yang. A Watermarking Algorithm for Vector Graphics Using Curve Partition[J]. Journal of Image and Graphics, 2009，8：1516-1522.
[6] FAN Yong-liang, ZHANG Hui-min, ZHANG Yun-long. Self-adaptive semi-fragile digital watermarking algorithm [J]. Application Research of Computers, 2011, 28(5): 1937-1939.
[7] Han-ling Zhang, Ming-ming Gao. A Semi-fragile Digital Watermarking Algorithm for 2D Vector Graphics Tamper Localization[C]. Multimedia Information Networking and Security, 2009. MINES '09. 2009: 549–552.
[8] Liangbin Zheng, Fucheng You. A Fragile Digital Watermark Used to Verify the Integrity of Vector Map [C]. E-Business and Information System Security, 2009. EBISS '09. 2009: 1–4.
[9] ZHENG Liang-bin, JIA Yu-lu, WANG Qun. Fragile Digital Watermark Used to Verify Integrity of Vector Map [J]. Computer Engineering and Applications, 2010, 46(26): 99-101.