

# WATERMARKING GIS DATA FOR DIGITAL MAP COPYRIGHT PROTECTION

Shen Tao

Chinese Academy of Surveying and Mapping, Beijing 100039, China  
[shentao@casm.ac.cn](mailto:shentao@casm.ac.cn)

Xu Dehe

Institute of resources and environment, North China University of Water Resources and Electric Power, Zhengzhou 450011, China  
[xdhwzs@126.com](mailto:xdhwzs@126.com)

Li Chengming

Chinese Academy of Surveying and Mapping, Beijing 100039, China  
[cml@casm.ac.cn](mailto:cml@casm.ac.cn)

Sun Jianguo

College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China  
[sjg@hrbeu.edu.cn](mailto:sjg@hrbeu.edu.cn)

**Abstract:** In recent years, with the rapid development of computer communication and the Internet, it makes easy to duplicate and distribute the GIS data because of its digital forms. So the application of GIS data faces some issues such as digital copyright protecting, data authenticating, and data source tracing, etc, and these issues could be the main obstacles restricting the further usage of digital maps. In order to solve these problems, we should construct the GIS production copyright protection system. The technique of watermarking provides potential solutions for copyright protection. Considering the characteristics of digital watermarking for digital maps, the DFT (Discrete Fourier Transform) blind watermarking model is put forward in this paper. First, the vertex sequence that extracted from vector map is carried on DFT transform. Second, divide the phase of DFT according to the quantization step size. Finally, we introduce a new watermarking scheme to embed the watermark into the phase of DFT through quantification, which provides robustness for watermark attacking. Extensive experiments are conducted to validate the availability, invisibility and the robustness of our scheme.

**Keywords:** digital map; digital watermarking; DFT;

## 1 Introduction

Digital map is a kind of important and strategic information resource and is widely used in economic and social applications, it also closely related to national defense affairs. The acquisition and maintenance of the data is a high costs and efforts process, there are needed large amount of manual labor and analysis in order to acquire original geospatial data such as land survey, aerial photography, satellite image, etc. Furthermore, it is a hard work for people obtain the information through digitalization and vectorization from the original data<sup>[1][2]</sup>. So as a result, geospatial data especially the high precision GIS data normally cannot be used for free. In addition, some high precision data are only permitted to be used in military or government units and extra security requirements are also needed. But in recent years, with

the rapid development of computer communication and the Internet, it makes easy to duplicate and distribute the GIS data because of its digital forms. So there is an emergent need for copyright protection and authentication of vector maps.

Watermarking is a promising solution to address this issue. In addition to the general characteristics of digital products, vector data has the following features<sup>[3]</sup>: There is no fixed sequence of data; there is no clear concept of sampling rates; data contains not only the geometric information includes topology information; and vector geospatial data is required high precision, small redundancy, hierarchical organizations, which are an increase of watermark embedding and extraction more difficult. Vector geospatial data with digital watermark in addition to the digital watermark must be non-perception, and stability, delectability, but also to ensure its availability, the data embedded watermark should not affect its availability.

Over the past few years, the research on the digital watermark of raster data is more, but that of vector map in the literature is very few<sup>[4-9]</sup>, among which the original vector data is needed in the watermark extraction phase, the detection method is relatively easy to achieve, and the extracted results are highly dependent on the quality of the raw data. In practical applications, raw data can't be obtained easily on many occasions, so the blind watermarking model<sup>[10-15]</sup> which the raw data is not required when the watermark is extracted should be paid more attention.

In this paper, Considering both the characteristics of digital watermarking for vector maps and the high-accuracy data requirement for vector data, a new DFT-based blind watermarking method for the vector maps is proposed, by transforming the vertex sequence using DFT transformation, dividing the phase of DFT according to the quantization step size, the watermark sequence can be embedded into the phase of DFT, which ensures the availability, invisibility and robustness.

## 2 System model

The main purpose of watermark model in this paper is to achieve blind detection<sup>[16]</sup>. And the main idea is that the coordinate axis would be divided into the A set of intervals and the B set of intervals by the equal-space, interval size is  $\Delta$ . Here, A set represents 0, B set on behalf of 1. Adjust the coefficient value of the watermark to be embedded to causes it to be equal to one recent correspondence set in the intermediate quantity. When the watermark is detected, just to determine the range of the coefficient falls on the A or B set, you can get the corresponding watermark information 0 or 1.

Because the transform domain, the modulation of the object data is the range of transform domain or phase, may also be a real part or imaginary part. Transform domain coefficient of assumptions, embedded digital watermark for the principle as shown:

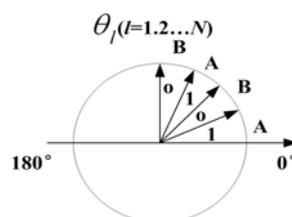


Figure 1. The phase quantification

Demarcation interval set: select the quantization step size will be divided into axes as shown in Figure1. A and B set interval. Those who belong to set the coordinates of A band represents bit "1", where a range of set B represents the position coordinates of "0".

The mathematical operation that seeks the quotient and remainder: select the quantization step size to quantify the parameters which is phase of DFT coefficients, integer values and remainder is operated. Assumptions  $a$  for the quotient,  $b$  for remainder, there are:

$$a = \left\lfloor \frac{360}{\Delta} \right\rfloor \quad b = 360 - a\Delta$$

Quantified embedded: The quantification of parameters  $\theta_i$  is closely related to the watermark bit  $w_i$ : When  $w_i = 1$ , enable the quantification results  $\theta_i$  equivalents to the intermediate coordinates which on a certain range of A set the nearest  $\theta_i'$ , When  $w_i = 0$ , enable the quantification results  $\theta_i$  equivalents to the intermediate coordinates which on a certain range of B set the nearest  $\theta_i'$ . The expression are as follows (for  $k=0,1,2,\dots$ ):

( 1 ) 当  $w_i = 1$  时 :

$$\theta_i' = \begin{cases} \theta_i & a = 2k \\ \theta_i + \Delta & a = 2k + 1 \end{cases}$$

( 2 ) 当  $w_i = 0$  时 :

$$\theta_i' = \begin{cases} \theta_i + \Delta & a = 2k \\ \theta_i & a = 2k + 1 \end{cases}$$

After quantification of  $\theta_i$ , the contained information of the watermark bit  $w_i$  is only determined by the range set in the quantization results  $\theta_i'$  : If  $\theta_i$  is in A set,  $\theta_i'$  on behalf of the watermark bit information  $w_i = 1$ ; the other hand,  $\theta_i$  is in the range of B set,  $\theta_i'$  will represent the digital watermark information  $w_i = 0$ , the above (1) belonging to A range set, (2) belonging to B range set. From where we can see that  $|\theta_i' - \theta_i| \leq \Delta$ , the maximum error caused by the quantization operation is quantify step  $\Delta$ .

As can be seen from the above description, the process of embedding the watermark based on the quantification is different from the previous embedding rules of addition or multiplication, it can be seen as appropriate modifications of parameters  $\theta_i$  and process which be given the watermark information. In order to reduce interference resulted from the random error in the process of quantify, improve the robustness of embedded watermark,  $\theta_i$  will be quantified to the nearest A or B areas.

When the watermark is detected, test data is carried on DFT, the corresponding watermark information  $w_i = 0$  or  $w_i = 1$  is get from the location of the embedded watermarks, the value of  $\theta_i'$  belonging to A or B set:

$$w_i = \begin{cases} 1 & \theta_i' \in A \\ 0 & \theta_i' \in B \end{cases}$$

The extracted watermark sequence is carried an anti-scrambling, if raw watermark is two-dimensional images, one dimensional sequence is transformed to two dimensional sequence.

In this article, the binary image is get from the mapped text, so the extracted watermark sequence is done with to get the two dimensional image.

### 3 The blind watermarking scheme

Based on the above system model, DFT as the transform tools, the watermark is embedded into the phase of DFT coefficients, the phase value of DFT coefficients is quantified by Appropriate step size , as well as the watermark is embedded in quantified value. The embedded watermark vector geospatial data is obtained after that the embedded phase value and the unmodified magnitude carries inverse DFT; when the watermark embedded in the vector data is extracted, first, the embedded vector data is carried on DFT, then the coefficient of DFT is calculated, the phase is get, the watermark is extracted according to the quantify program.

#### 3.1 The embedded phase

The embedded model of digital watermark of the vector geospatial data is shown in Figure 2.

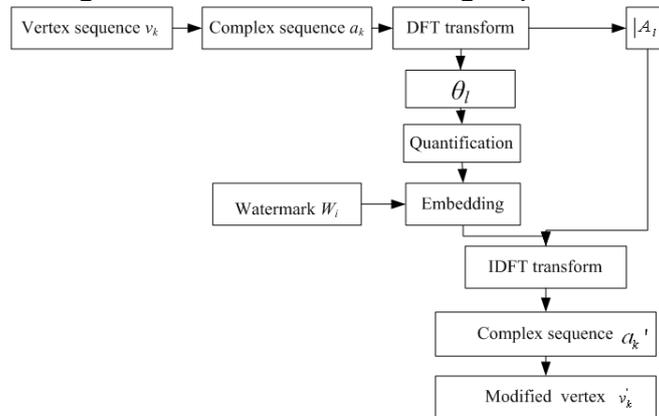


Figure 2. The embedded model

Concrete steps are as follows:

Step 1: The original watermark length is transformed to meet  $2^N$  (FFT algorithm conditions, a quick algorithm of DFT). In order to increase the security of the watermark, the watermark can be encrypted or scrambled.

Step2: Geospatial vector data  $V$  is read according to the storage order of coordinates, and then the coordinates of vertex sequence constitute a complex sequence in accordance with formula (1).

$$a_k = x_k + iy_k \quad k \in [0, N-1] \quad (1)$$

Here  $N$  is the Vertices number in  $V$ .

Step3: Sequences  $\{a_k\}$  is carried on DFT, the DFT coefficients is get by formula (2).

$$A_l = \frac{1}{N} \sum_{k=0}^{N-1} a_k (e^{-2\pi i/N})^{kl} \quad l \in [0, N-1] \quad (2)$$

The DFT coefficients of a sequence contain the amplitude  $|A_l|$  and phase  $\theta_l$ .

Step4: The phase of the coefficient is quantified by the selected appropriate step size, the quantified value equals to the median of the interval in accordance with the "0" or "1".

Step5: The embedded magnitude combined with the unmodified magnitude is carried inverse DFT, the embedded watermark vector geospatial data  $V'$  is obtained.

#### 3.2 The extracted phase

The extracted phase of digital watermark is shown in Figure.3.

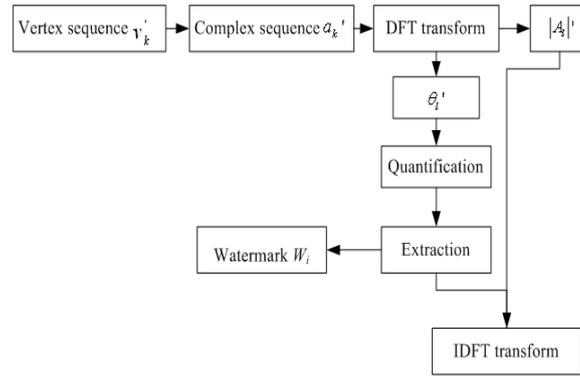


Figure 3. The extracted model

Step1: Vertex coordinates in  $V'$  is read, then a complex sequence is  $\{a_k'\}$  constructed in accordance with formula (1).

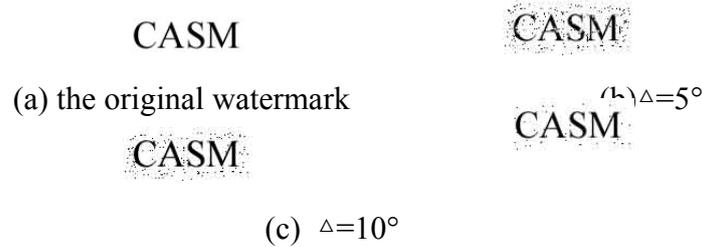
Step2: The sequence  $\{a_k'\}$  is carried on DFT by formula (2), the coefficient of DFT is calculated.

Step3: The phase of the DFT coefficient is quantified to mean values of the location of embedded watermark. The phase value is calculated to judge in which interval quantified, the watermark is extracted according to the quantization range where the phase values is in.

Step4: The extracted watermark bit sequence is carried on anti-scrambling, then and the self-correlation  $c$ , in order to enhance the objectivity of the result of extracted watermark, the self-correlation method is used, the self-correlation coefficient obtained  $sc|\theta_k'$  is used to help determine the effect of watermark extraction, if the raw watermark is two-dimensional images, the extracted watermark is processed to get the final two-dimension watermark.

#### 4 Experiments

The test data contains 11,080 vertex of river lines of 1:50 million. The watermark data is  $116 \times 32$  bit as Figure 4(a), according to FFT, the data is extend to the 4096 ( $2^{12}$ ), and the quantization step size is 5,10,15 respectively.



The quantization step size is variable parameters when the range is fixed, the different step size are used in experiment to contrast, the used quantization step size are the  $\Delta = 5^\circ$ ,  $\Delta = 10^\circ$  and  $\Delta = 15^\circ$ . The extracted watermark are shown in Figure 4(b), Figure 4(c) and Figure 4(d).

Figure 5, Figure 6 and Figure 7 respectively show the compared result of the original vector geospatial data and the embedded data when step size  $\Delta = 5^\circ$ ,  $\Delta = 10^\circ$ ,  $\Delta = 15^\circ$ , the dotted line for the embedded watermark data and the solid line for the raw data. We can find the embedded model is well invisibility, and the embedded data does not affect the display quality of

graphics, compared to three-step size, the invisibility of graphics displayed in figure 5 is best, Figure 6 follows, the Figure 7 is the worst.

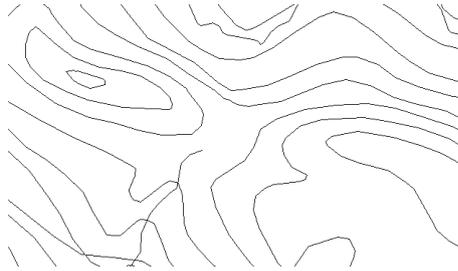


Figure 5. Comparison of before and after watermark embedded when  $\Delta=5^\circ$



Figure 6. Comparison of before and after watermark embedded when  $\Delta=10^\circ$



Figure 7. Comparison of before and after watermark embedded when  $\Delta=15^\circ$

#### 4.1 Availability

The results of the a bit error ratio(BER) analysis corresponding to  $\Delta=5^\circ, \Delta=10^\circ, \Delta=15^\circ$  are shown in figure8. As can be seen from the statistical results, the BER increases with the step length becomes large, but the largest BER is smaller that 2.5 percent, which validates the availability of the proposed scheme.

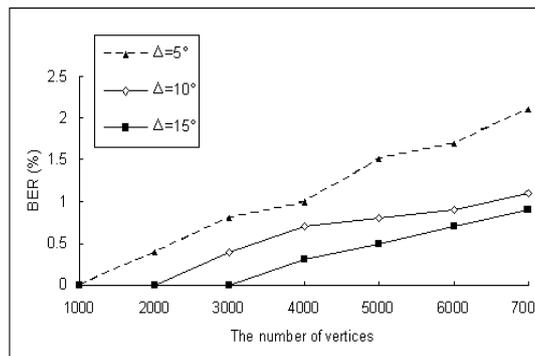


Figure 8. BER analysis corresponding to  $\Delta=5^\circ$ ,  $\square=10^\circ$  and  $\square=15^\circ$

## 4.2 Robustness

The robustness of the watermark is evaluated for different types: format conversion, deleting the vertices, rotate, translation, scale in experiment, and find that the watermark can not be extracted effectively after that the vertices are deleted randomly and the embedded data is scaled, so the following analysis of only a few attacks on the remaining cases and the results are shown in Tab.1, which exhibit the ability of resisting attacks, and the values of  $NC$  (self-correlation coefficient) calculated by formula 3.

$$NC = \frac{|w \cdot w'|}{\sqrt{w \cdot w} \sqrt{w' \cdot w'}} \quad (3)$$

Table 1 the robustness experiments

attacks		$NC$ ( $\Delta=5^0$ )	$NC$ ( $\Delta=10^0$ )	$NC$ ( $\Delta=15^0$ )
Format conversation		0.744552	0.985421	0.994852
Translation (units)	1	0.756772	0.979669	0.996942
	5	0.693524	0.982759	0.996552
	10	0.656602	0.980455	0.990432
Rotation (degree)	1	0.454451	0.870635	0.973293
	5	0.324081	0.864928	0.970950
	10	0.289436	0.855841	0.966949

## 4.3 Experimental results analysis

Through the experiment, we could analyze and find:

(1) The step size (that is  $\square$ ) is greater, the difference of the extracted watermark and original watermark is smaller, so the robustness is better. (2) If the step size selected is too large, the graphic quality displayed and visibility are affected. If the step size selected is too small, availability and visibility can be ensured, but the effect of watermark extracted is bad. Therefore step size used to quantify is determined in accordance with the robustness of the watermark and the used condition of vector geospatial data. (3) The error in the extracted watermark of integer data embedded vector geospatial data after translation, rotation, format conversion comes mainly from its own error and rounding error when the FFT computation, the effects of the extracted watermark are the same after the format conversion and translation of the embedded vector geospatial data. After the rotation, the watermark extraction error in the data mainly from the data is cut-off. The model is fragile after deleted vertices, scaled.

## 5 Conclusions

In this paper, a blind watermarking scheme based on DFT is proposed. By transforming the vertex sequence using DFT transformation, dividing the phase of DFT according to the quantization step size, the watermark sequence can be embedded into the phase of DFT. The experimental results show that the embedded vector geospatial data has good availability, invisibility, the watermark is robust after translation, rotation and format conversion of the

embedded watermark vector map, the computational complexity of the model is low and easy to implement. It is confirmed that the proposed model is effective in protecting the copyright of the vector GIS data. Future work will include the robustness against deleting vertices, clipping, scaling to the achievement of the copyright protection of the vector map and applying this model to the other types of vector data.

### Acknowledgement

This work was supported by Research and Development Program of CASM and Funded by Key Laboratory of State Bureau of Surveying and Mapping ( No. 200827 )

### References

- [1] Anbo Li, Bing-xian Lin, Ying Chen, Study on copyright authentication of GIS vector data based on Zero-watermarking. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B4. Beijing 2008.*pp.1783-1786.
- [2] Ohbuchi, Robust watermarking of vector digital maps, *Proc. of the IEEE International Conference on Multimedia and Expo., Lausanne, Switzerland, 2002* vol.1, pp.577-580.
- [3] XU De-he, ZHU Chang-qing, WANG Qi-sheng. “A Survey of the Research on digital watermarking for the Vector Digital Map”, *Geometrics World*. Editorial Dept. of Geomatics World, Harbin, China, 2007, pp.42-48.
- [4] Voigt M, Busch C. “Watermarking 2D-Vector data for geographical information systems”. *Proceedings of the SPIE, Security and Watermarking of Multimedia Content*, San Jose, USA, 2002, pp. 621-628.
- [5] Schulz G, Voigt M. “A high capacity watermarking system for digital maps”. *Proceedings of the 2004 Multimedia and Security Workshop on Multimedia and Security*, Magdeburg, Germany, 2004, pp.180-186.
- [6] Wang Xun, Lin Hai, Bao Hu-Jun. “A robust watermarking algorithm for vector digital mapping”, *Journal of computer-aided design & computer graphics*, Editorial Dept.of Journal of computer-aided design & computer graphics, Beijing, China, 2004, pp.1377-1381.
- [7] Nikolaidis N, Pitas I, Solachidis V, “Fourier descriptors watermarking of vector graphics images”, *Proceedings of the International Conference of Image Processing*, Vancouver, BC, Canada, 2000, pp.10-13.
- [8] Solachidis V, Nikolaidis N, Pitas I. “Watermarking polygonal lines using Fourier descriptors”, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Istanbul, Turkey, 2000, pp.1955-1958.
- [9] Nikolaidis N, Pitas I, Giannoula A, “Watermarking of sets of polygonal lines using fusion techniques”, *Proceedings of the 2002 IEEE International Conference on Multimedia and Expo*, Lausanne ,Switzerland, 2002, pp.26-29.
- [10] LI Yan-ping, TANG Zhen-min, QIAN Bo, “Audio Watermark Algorithm Robust to Desynchronization Based on Speech Parameter Model”, *Computer Engineering*, Editorial Dept. of Computer Engineering, Shanghai,China, 2008,pp.177-179.
- [11] Sullivan K, Bi Z, Madhow U, et al, “Steganalysis of Quantization Index Modulation Data hiding”, *IEEE International Conference on Image Processing*, Suntec City, Singapore, Oct.2004, pp.1165-1168.
- [12]Benedens, O. and C. Busch, Towards blind detection of robust watermarks in polygonal models, *Eurographics 2000 Proceedings, vol.19,.no.3,2000.* pp. C199-C208.

- [13] Brian C, Gregory W, “Pre processed and post processed quantization index modulation methods for digital watermarking”, *Processing of SPIE: Security and Watermarking of Multimedia Contents II*, San Jose, USA ,2000, pp. 48-59.
- [14] Chen B, Wornell G, “Quantization index modulation: A class of provably good methods for digital watermarking and information embedding”, *IEEE Transactions on Information Theory*, IEEE Inc, USA, 2001, pp. 1423-1443.
- [15] Ronald N, Bracewell. *The Fourier Transform and Its Applications*. Xi'an Jiaotong University Press, Xi'an, China, 2005, pp.203-213.
- [16] YANG Yi-xian, NIU Xin-xin. *Theory and Applications of Digital Watermarking*. Higher Education Press, Beijing, China,2006