Robust Watermarking of Digital Vector Maps for Copyright Protection

Tawfiq A. Abbas, Majid Jabbar Jawad
College of Information Technology
University of Babylon
Iraq
{tawfiqasadi63, majid_al_sirafi}@yahoo.com

Sud Sudirman
School of Computing and Mathematical Sciences
Liverpool John Moores University
United Kingdom
s.sudirman@ljmu.ac.uk

Abstract— **Digital vector maps are used in several applications, such as car navigation system, web-based map services, geographical information system (GIS), and computer-aided design (CAD). As with any other digital data, digital maps are easy to update, duplicate, and distribute. As a result, illegal operations such as forgery and redistribution of the maps can be done equally easily. Therefore, the application of digital vector maps faces a number of challenges including copyright protection, data authentication, and data source tracing. These issues are problematic and limit the use of digital vector maps. Digital watermarking has been proposed as the solution to these problems. However, recent literature review in this field suggests that most of the watermarking schemes for digital vector maps in the literature do not take into account the preservation of the map topology which is important in preserving the integrity of the map. Furthermore, most techniques lack the features to control the distortions. In this paper, we propose a novel watermarking algorithm for copyright protection of digital vector maps that takes into account the preservation of the map topology after embedding the watermark operation as well as provides the users a means to control the distortions. The technique works by inserting additional vertex coordinates between adjacent vertices which locations are governed by the value of the watermark data. Our experiments have shown that our proposed technique is robust against all geometric attacks, format exchange attacks, and additive noise attacks.**

Keywords: Digital Watermarking, Vector Map, GIS,CAD, Topology, Parametric equation

I. INTRODUCTION

In general, vector data structure produces smaller file size than raster image because a raster image needs space for all pixels while only point coordinates are stored in vector representation. This is especially true in the case when the graphics or images have large homogenous regions and the boundaries and shapes are the primary interest. Besides the size issue, vector data is easier than raster data to handle on a computer because it has fewer data items and it is more flexible to be adjusted for different scale. This makes vector data structure the apparent choice for most mapping, GIS (Geographic Information System) and CAD (Computer Aided Design) software packages. CAD has been traditionally used to assist in engineering design and modelling for representation, analysis and manufacturing. So, because of vector is very accurate according to its structure, CAD depends on it in representing the objects. Also, GIS depends on the vector maps, because in vector maps the data can be represented at its original resolution, accurate geographic location of data is maintained, and it allow for efficient encoding of topology.

The acquisition of vector maps is a high cost process. High precision instruments are needed for land measuring when producing geographical information. This is in addition to other costs such as labour cost that must be expended to employ expert staff to perform the data measurement and collection. As a result, vector maps are very expensive to produce and their intellectual value must be protected against illegal use.

The rest of the paper is organized as follows. In section 2, some information about watermarking terminology and related works is presented. In section 3, a brief description about characteristics of digital vector map is illustrated. In Section 4, a digital watermarking algorithm for vector map is presented. Experimental results and capability analyses are shown in Section 5. Conclusions are drawn in Section 6.

II. DIGITAL WATERMARKING TERMINOLOGY

In general, digital watermarking means digitally adding a small amount of data(referred to as watermark) in a digital object (host). The information encoded in the watermark can be used to identify the copyright owner of the object or to detect any tampering performed onto the object. Digital watermarking has several features as described below.

The first feature is the type of extraction process. Digital watermarking techniques can be blind, semi-blind or non-blind. Blind watermarking does not require the host object or the watermark data. On the other hand, non-blind watermarking does require both whereas semi-blind watermarking requires only one of them. Ideally, the extraction process should be blind, since sometime we don't have the original host.

The second feature is fidelity. This is defined as the relative similarity between the un-watermarked host object and the one after the watermarking operation. A good watermarking technique will produce a high fidelity, i.e., it introduces very little distortion to the host object.

The third feature is robustness which is the resilience of the inserted watermark data to any processes (attacks) aimed at either removing or distorting it. A robust watermarking

technique should produce watermarks that are able to withstand different types of attacks and its watermark extraction process should be able to retrieve the watermark from the attacked host with little difficulty.

The security of a watermarking technique is defined as the level of difficulty in identifying what algorithms used to perform the watermarking process. A highly secure watermarking process would produce output that does not contain any specific signatures that can be used to identify the algorithm.

The fifth characteristic is the watermark payload which is the amount of information that the watermark signal carries. The ideal size of a watermarking technique depends on the application in which the technique is used. Some applications such as copyright protection requires very small payload whereas watermarking for broadcast monitoring requires high payload.

The sixth characteristic is reversibility which determines whether the reverse of the watermarking technique can be applied to reconstruct the un-watermarked data from its watermarked counterpart.

Another important feature is topology preservation. This feature is only applicable to vector map watermarking in which the topology of vector map can change after a watermarking operation in a high degree and makes the watermarked map invalid to use. Most of proposed digital vector map watermarking schemes, do not taken in account this situation. For example, if there are two parallel rivers in the map, may be after watermarking operation these rivers may intersect with each other, so at this situation the map will be invalid.

The technique of digital watermarking can be classified into several criteria. Firstly, it can be classified according to the representation domain in which the watermark is embedded. There are two mostly used types of representation domains namely frequency and spatial domains. In frequency domain, the watermark is embedded into coefficients data rather than the original data. That means the original data is transformed into frequency domain and the watermark is embedded into the coefficients of that data. There are several frequency domain transformations that can be used in the embedding methods namely Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Discrete Fourier Transform (DFT) [1, 2, 6]. In spatial domain, the watermark is embedded directly into the original data [5, 7, 8]. In [1], a watermarking algorithm is proposed for vector map based on discrete cosine transform. Watermark is hidden DCT coefficients with middle frequency. Experiments show that the watermarking algorithm is robust against many attacks such as rotation, scaling, translation, and simplification. In [2], considering the characteristic of vector geo-spatial data for digital watermarking, this paper presents a new watermarking algorithm for vector geo-spatial data based on integer wavelet transform. By the proposed algorithm, the line and surface coordinate points are decomposed by integer wavelet transform firstly, then watermark is embedded in the low frequency coefficients of integer wavelet transform, and the

vector geo-spatial data with watermark is obtained by inverse integer wavelet transform finally. The experiments show that the proposed algorithm can effectively resist the attack of noise, data compressing, points deleting and format exchanging. For integer vector geo-spatial data, the algorithm can restitute the data well. In [5], based on improved difference expansion (IDE), a reversible watermarking scheme for 2D CAD engineering graphics is proposed. First, all vertices of the 2D CAD engineering graphics are extracted, and then the relative coordinates of these vertices are calculated. Finally, the encrypted watermark is embedded into the scale factor of the relative coordinates by using IDE technology. Experimental results show that the proposed scheme has a large capacity, good imperceptibility, and is robust against operations such as transformation, rotation, and equal scaling.

III. CHARACTERISTICS OF VECTOR MAP

Data in a vector map consists of points and lines, in a form of floating point data sequence that represents all the objects in the map. This makes vector map data is more complex compared to other multimedia data such as image, video, and audio.

Vector graphics is composed of a series of metadata which is described by a set of instructions; these instructions describe the location, dimension and shape properties of the graphic content. In order to display vector graphics on the screen, special software is need, which convert the instructions to the shape and color, that is, according to different structures of vector data and different algorithms, the graphics is "drawing" on the screen. Through vector graphics software, the various elements of vector graphics can be created and edited, and vector graphics can be moved, zoomed, rotated, and distorted. The vector graphics is suitable for such transformation, and it also has the advantages of high precision, good quality, as well as a small amount of storage. Vector graphics is used widely in computer-aided design, geographic information systems (GIS), art design, animation, and so on. The majority software of GIS, CAD and 3D modelling use vector graphics as the basic graphics storage format.

Vector Graphics has its own characteristics in data expression and use, and it is difficult to apply the existing digital watermarking technology to deal with vector graphics data [4].

Most of algorithms of vector map watermarking have not considered the topology. Topological spatial relationship is used to describe the spatial relations between spatial objects, such as adjacency, inclusion, intersection and so on. It has important implications in data processing and spatial analysis:

- Topological relationship can clearly reflect the logical structure of relations between entities, and it is more stable than geometric relationship as it doesn't change with the projection transformation.

- The use of topological relations can benefit the query/search process by adding spatial elements into the query.

- It can be used to rebuild geographical entities.

As we often modify the redundant bits of data or use the spatial features of points in the process of embedding watermark, it may result in the changes of topological relations. So it is necessary to inspect the topological relations of watermarked vector data [3].

IV. THE PROPOSED WATERMARKING SCHEME

The proposed watermarking scheme consists of three steps namely pre-processing of the host object (map), embedding operation, and extracting operation.

A. Pre-processing of the host object

Before embedding operation, the original data that will be used in the embedding operation must be pre-processed by creating a new vertex between every pair of connected vertices. Figure 1, illustrates the above operation.
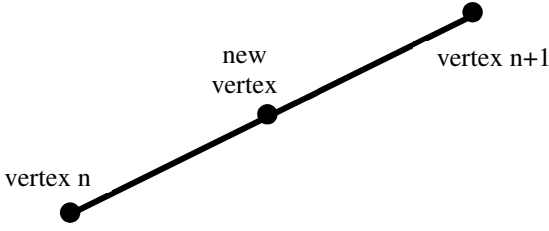

Fig. 1. Adding a new vertex to adjacent vertices

B. Watermark Embedding Process

The process of embedding the watermark is carried out on the pre-processed host image. The embedding is done by translating each of the newly added vertices by an amount depending on the watermark value. Let's denote the vector containing the vertex coordinates of the ith original new vertex and the embedded new vertex, given the ith watermark wm_i, as nv_i and nvm_i respectively. The values for nvm_i is calculated as:

$$nvm_i = \begin{cases} nv_i + d & if \quad wm_i = 1 \\ nv_i & if \quad wm_i = 0 \end{cases} \quad (1)$$

where d is the displacement vector which value calculated by using parametric equation [10] and depending on the coordinates of the adjacent vertices.

Let's denote the vector containing the vertex coordinates of the two adjacent vertices, a and b, to the ith new vertex nv_i as v_a and v_b respectively. The displacement vector d to be added to nv_i is calculated as:

$$d = (1 - t).v_a + t.v_b \quad (2)$$

where the variable t, $0 < t < 1$, controls the distance of the displacement. The smaller the value of t the smaller the displacement and vice versa. The embedding operation is illustrated in Figure 2.
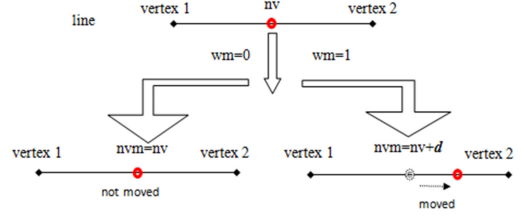

Fig. 2. Watermark embedding process

C. Watermark Extraction Process

The watermark wm_i can be extracted from the watermarked vector map by observing position of the interpolated vertex in relation to the geometrical centre of the line connecting the two adjacent vertices.

The difference value d_i is calculated by measuring the distance between the extracted position p_i of the ith interpolated vertex and the centre position of line pc_i.

$$d_i = \sqrt{(p_{ix} - pc_{ix})^2 + (p_{iy} - pc_{iy})^2} \quad (3)$$

$$wm_i = \begin{cases} 1 & |d_i| > T \\ 0 & |d_i| \le T \end{cases} \quad (4)$$

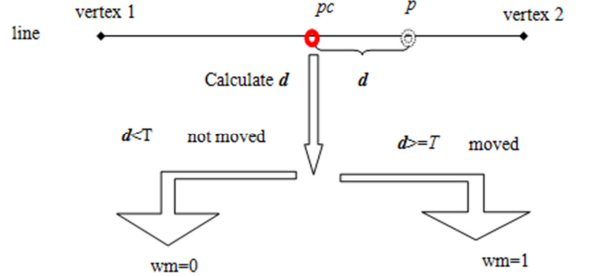The process of extracting of the watermark is illustrated in Figure 3.

Fig. 3. Watermark extraction process

To evaluate the similarity between the original watermark and extracted watermark, *Normalized Correlation (NC)* is calculated using Eq. 5 [9].

$$NC = \sum_{x=0}^{n-1} \frac{wm_i(x)wm'_i(x)}{n} \quad (5)$$

Where *n* denotes the length of the watermark and $wm_i(x)$ and $wm_i'(x)$ are the original and extracted watermark respectively.

V. EXPERIMENTAL RESULTS

An experiment has been conducted to test the performance of the proposed technique. A set of watermark images and a set of vector maps are used. For each experiment, each watermark image is embedded to all the vector maps. There are four watermark images and fourteen vector maps used in total. In the interest of brevity, only one of each category is shown in this paper as shown Figure 4. The complete set of the maps can

be obtained in [11]. The watermark images are 70x70 pixels in dimension and are black and white. The vector maps are in ESRI shape file. The value of parameter *t* used is 0.03.

The performance of the proposed algorithm is measured using two criteria namely imperceptibility and robustness against attack.
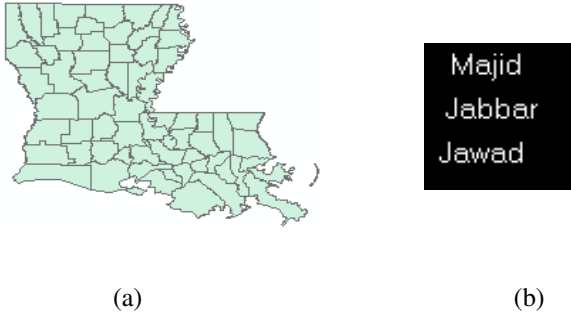

(a)   (b)

Fig. 4.   A sample of the (a) vector maps and (b) watermark images that used in the experiments

A. Imperceptibility

Figure 5, shows part of the watermarked and original map. As we see, visually, there no distortion after embedding operation. That is means, the fidelity (imperceptibility)is very good.
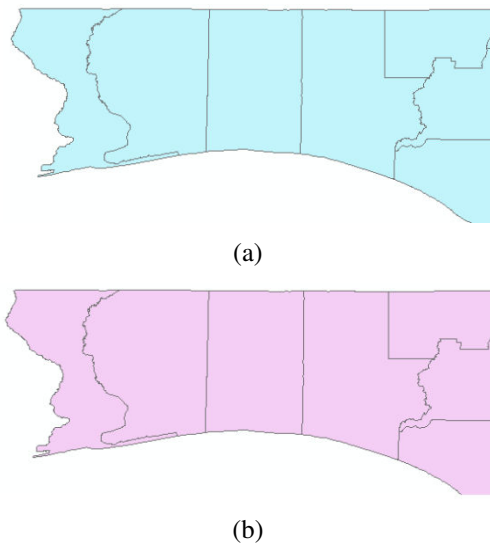

(a)

(b)

Fig. 5.   Comparison between the (a) Original and (b) Watermarked map

B. Testing the robustness against attacks

To measure the ability of the proposed technique in withstanding attacks we have conducted an experiment that applied a number of attacks to the watermarked maps. From these attacked watermarked vector maps we attempt to extract the watermark images. To quantify the amount of similarity between the original watermark image and the extracted watermark image we calculate the Normalization Correlation (NC) as described in Eq. 5.

The types of attacks can be categorised into four types namely Geometric attacks, Format Exchange attacks, Additive Noise attacks and Vertex Deletion attacks. The results for Geometric and Format Exchange attacks are summarised in Table 1. The results show that the proposed technique can produce perfect reconstruction of the original watermark image after these types of attacks.

TABLE I.   RESULTS FOR GEOMETRIC AND FORMAT EXCHANGE ATTACKS

Attacks		NC
Geometric	Rotation	100%
	Scaling	100%
	Translation	c100%
Format Exchange	From SHP format to DXF format and vice versa	100%
	From SHP format to DWG format and vice versa	100%

Additive noise attack is applied on the watermarked vector map by adding noise randomly to any vertex in the map. Table 2 shows some of extracted watermarks after applying noise attack on the watermarked vector map. The overall performance of the proposed technique on this particular criteria can be seen Figure 6. It shows that the low threshold results in a higher NC value at low noise strength region, while the high threshold gives higher NC value at high noise strength region.
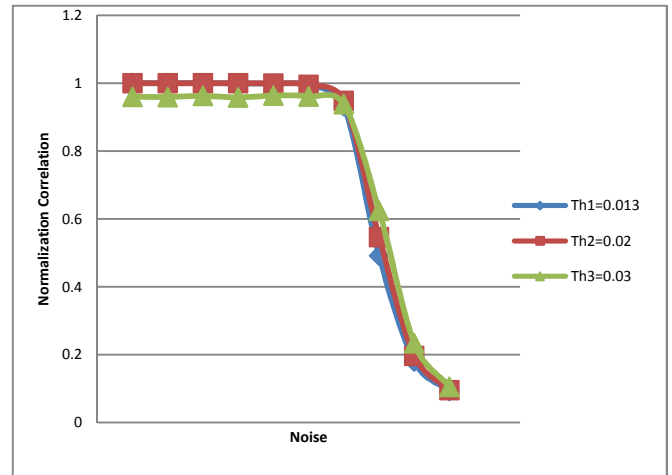

Fig. 6.   Noise vs. Normalization Correlation

TABLE II. SUMMARY OF THE WATERMARK EXTRACTION RESULTS AFTER NOISE ATTACKS

Noise Strength	Threshold	Extracted Watermark	Normalization Correlation
$5 * 10^{-11}$	0.015	copyright LJMU	1
$5 * 10^{-10}$		copyright LJMU	1
$5 * 10^{-5}$		copyright LJMU	0.86
$5 * 10^{-11}$	0.02	copyright LJMU	1
$5 * 10^{-10}$		copyright LJMU	1
$5 * 10^{-5}$		copyright LJMU	0.9
$5 * 10^{-11}$	0.03	copyright LJMU	0.92
$5 * 10^{-10}$		copyright LJMU	0.918
$5 * 10^{-5}$		copyright LJMU	0.88

Another attack, which is deletion of vertex noise attacks. It is applied on the watermarked vector map by deleting the vertices randomly from map. Table 3, shows some of extracted watermarks after applying deletion of vertex attack on the watermarked vector map.

TABLE III. SUMMARY OF THE WATERMARK EXTRACTION RESULTS AFTER DELETION ATTACKS

Strength of deletion vertex	Extracted Watermark	Normalization Correlation
.01%	copyright LJMU	0.61
.02%	copyright LJMU	0.53
1%	copyright LJMU	0.49

C. Topology preservation

The change of the coordinates become after watermarks are embedded into the original data. Although the ideal method is that there is no change of the coordinates, it is impossible. So, after embedding watermarks, if the change is within the allowable error, we can say it is not to damage to the vector map.

In vector map data, the positional accuracy is very important unlike in general multimedia data. That is, if the positional accuracy is wrong, the map is useless. And if the topology of the vector map is changed after embedding watermarks, it makes a big problem. For example, if two lines are in parallel case and they are intersected after embedding operation, some problems are generated. Figure 7, shows the possible problem. Many watermarking on vector map have not considered the topology. The proposed algorithm is preserved the topology since the vector map is watermarked by moving a watermarked vertex between two vertices on the straight line, as shown in Figure 8.
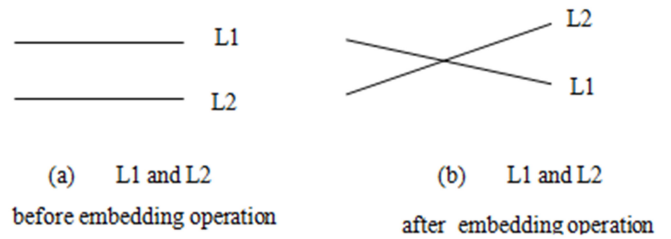

(a) L1 and L2 before embedding operation

(b) L1 and L2 after embedding operation

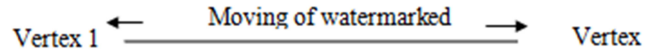Fig. 7. Possible problem after embedding process

Fig. 8. Moving watermarked vertex

VI. CONCLUSION AND FUTURE WORKS

In this paper, a reversible watermarking scheme for 2D vector map drawings has been proposed. The proposed scheme is also blind which means that the extraction of watermark does not need the original cover. The important thing is that the scheme taken in account preserving of the map topology .The embedding and extracting of watermark depend on the moving of one the vertex between two vertices. The experimental results shows that this scheme is robust to several attacks such as scaling, translation, rotation, adding noises, format exchanging. Future research work should include applying this scheme without adding a new vertex, also the robustness of the proposed algorithm should also be improved against more attacks.

REFERENCES

[1] Bin Liang, Jiang Rong, Chuanjian Wang, "A Vector Maps Watermarking Algorithm Based On DCT Domain", Canadian Geomatics Conference and Symposium of Commission I, International Society Photogrammetry and Remote Sensing(ISPRS) Convergence in Geomatics – Shaping Canada's Competitive Landscape , Calgary,Alberta,Canada,June,15-16, 2010.

(http://www.isprs.org/proceedings/XXXVIII/part1/)

[2] Chang-Qing Zhu, Cheng-Song Yang, Qi-Sheng Wang, "A Watermarking Algorithm For Vector Geo-Spatial Data Based On Integer Wavelet Transform", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII. pp. 15-18,Part B4. Beijing 2008.

[3] Li Huang, Wei Zhou, Rui Jiang, Anbo Li, "Data Quality Inspection of Watermarked GIS Vector Map", 18th International Conference on Geoinformatics, pp. 1 – 5, 18-20 June 2010.

[4] LiangbinZheng, Ruqi Chen, Linhong Li, YeliLi, "Study on Digital Watermarking for Vector Graphics", The Second International Workshop on Education Technology and Computer Science, Vol. 2, pp. 535 - 538, 6-7 March 2010.

[5] Peng Fei, Lei Yu-Zhou, Long Min, Sun Xing-ming, "A reversible watermarking scheme for two-dimensional CAD engineering graphics based on improved difference expansion", Computer-Aided Design Journal, Vol. 43, Issue 8, pp. 1018 – 1024, August 2011.

[6] Shen Tao, XuDehe, Li Chengming, Sun Jianguo, "Watermarking GIS Data For Digital Map Copyright Protection", In Proceedings of 24th International Cartographic Conference, Santiago, Chile, Nov. 15-21, pp.1-9, 2009.

[7] Stefan Bird, Chris Bellman, Ron van Schyndel, "A Shape-based Vector Watermark for Digital Mapping", Conference on Digital Image Computing: Techniques and Applications, pp. 454 – 461, 1-3 Dec. 2009.

[8] Xiao Tong Wang, Cheng Yong Shao, Xiao Gang Xu, and Xia Mu Niu, "Reversible Data-Hiding Scheme for 2-D Vector Maps Based On Difference Expansion", IEEE Transactions On Information Forensics And Security, Vol. 2, No. 3, September 2007.

[9] Zope-Chaudhari, Sangita, and P. Venkatachalam. "Robust Watermarking for Protection of Geospatial Data." In Proc. of International Conference on Security Science and Technology, Hong Kong, pp. 34-38. 2012.

[10] http://www.sonoma.edu/users/w/wilsonst/papers/geometry/parametric/default.html (Accessed: 25 March 2013)

[11] http://www.census.gov/geo/www/cob/cbf_counties.html (Accessed: 25 March 2013)