Computers, Environment and Urban Systems 35 (2011) 485–492

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

A key points-based blind watermarking approach for vector geo-spatial data

Haowen Yan [a,b,*], Jonathan Li [b], Hong Wen [c]

[a] School of Mathematics, Physics and Software Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China
[b] Department of Geography & Environmental Management, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
[c] National Key Laboratory of Science and Technology on Communications, UESTC, Chengdu 611731, China

ARTICLE INFO

Article history:
Received 30 April 2010
Received in revised form 25 October 2010
Accepted 26 October 2010
Available online 30 November 2010

Keywords:
Blind watermarking
Vector data
Key points
Similarity degree

ABSTRACT

This paper presents a blind watermarking approach to protecting vector geo-spatial data from illegal use. By taking into account usability, invisibility, robustness, and blindness, the approach firstly determines three feature layers of the geo-spatial data and selects the key points from each layer as watermark embedding positions. Then it shuffles the watermark and embeds it in the least significant bits (LSBs) of the coordinates of the key points. A similar process for selecting the feature layers and the key points in the watermark embedding process is carried out to detect the watermark followed by obtaining the embedded watermark from the LSBs of the coordinates of the key points. Finally, the similarity degrees of three versions of the watermark from three feature layers are calculated to check if the data contains the watermark. Our experiments show that the method is rarely affected by data format change, random noise, similarity transformation of the data, and data editing.

1. Introduction

Generally speaking, vector geo-spatial data is of great value because the acquisition of such data is a high cost process (López, 2002), in which high precision instruments and a large amount of physical labour resources are required, and the digitization and vectorization of original data also need hard work. Consequently, vector geo-spatial data normally can not be directly used without the owner's permission. Nevertheless, the rapid development of computer communication and Internet techniques make it easy to duplicate and distribute digital data via networks, which brings a lot of trouble for the data owners to protect the data from piracy.

Digital watermarking, coming from steganography, provides a viable solution for data security. A digital watermark is defined as an imperceptible but identifiable digital signal or mode permanently embedded in other data, namely host data, while it does not affect the host data's usability (Ahmed, 2004). There are four important rules that should be obeyed in any successful watermarking techniques (Cox & Miller, 2002; Zhou, Ren, & Pan, 2006). First of all, the embedded watermark should not degrade the quality of the host data. Secondly, the watermark should be perceptually invisible to data users to maintain its protective secrecy. Next, the technique must be robust enough to resist common data processing attacks and not be easily removable by illegal users, but only the data owners ought to be able to extract the watermark. Finally, the watermark extraction process should be blind if possible, i.e. the watermark can be detected by the data owner without the original data and the original watermark at hand.

Have not only the techniques of digital watermarking received a great deal of attention to ensure copyright protection for multimedia message, such as video data (Hartung & Girod, 1998; Langelaar & Lagendijk, 2001), audio data (Kirovski & Malvar, 2003; Seok, Hong, & Kim, 2002; Wang, Niu, & Qi, 2008) and image data (Aslantas, 2008; Cox, Kilian, Leighton, & Shamoon, 1997; Langelaar & Lagendijk, 2001), and been a focus in network information security (Cox & Miller, 2002), but also it has become a hot issue in the community of Geo-Science for protecting vector geo-spatial data from piracy (Lafaye, Béguec, Gross-Amblard, & Ruas, 2007; López, 2002; Niu, Shao, & Wang, 2006). Generally, there are two categories of watermarking algorithms for two-dimensional (2D) geo-spatial data, i.e. space domain and frequency domain.

Many frequency domain algorithms (e.g., Solachidis & Pitas, 2004; Zhu, Yang, & Wang, 2008) embed the watermarks in the Fourier descriptors of the curves or polygonal lines, causing invisible distortions of the vertices coordinates. Ohbuchi, Ueda, and Endoh (2003) presented a method embedding watermarks in the frequency-domain representation of a 2D vector digital map. The method treats vertices on the map as a point set, and imposes connectivity among the points using Delaunay triangulation, then computes the mesh-spectral coefficients from the mesh created. Modifications of the coefficients according to the message bits, and inverse transforming the coefficients back into the coordinate domain produces the watermarked map. Voigt, Yang, and Busch

* Corresponding author at: School of Mathematics, Physics and Software Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China.
E-mail address: haowen2010@gmail.com (H. Yan).

0198-9715/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2010.10.004

(2004) addressed a reversible watermarking scheme that exploits the high correlation among points in the same polygon in a map and achieves the reversibility of the whole scheme by an 8-point integer discrete cosine transform, which ensures that the original 2D vector data can be watermarked during the watermark embedding process and perfectly restored during the watermark extraction process, with the watermark accurately extracted at the same time. The watermarks generated by these frequency domain algorithms can generally withstand rotation, translation, scaling, and reflection, and their detection process is blind. However, the spatial precision change (e.g. spatial topological relations among objects) in the watermarked data is difficult to control because the watermark is not directly embed in spatial data. In addition, such algorithms are generally fragile to the attacks from the noise and the revision of the data (e.g. insertion and removal of points).

Most of the existing space domain algorithms (e.g. Kang, 2001; Li & Xu, 2004; Ohbuchi, Ueda, & Endoh, 2002) are based on the idea of changing the positional relations of the points in vector maps. The principle of these algorithms are as follows: subdivide a vector map into some rectangular blocks of adaptive sizes according to the density of vertices, and embed each bit of the watermark by displacing the coordinates of a set of the vertices in the block. There are also some algorithms (Park, Kim, Kang, & Han, 2002; Sonnet, Isenberg, Dittmann, & Strothotte, 2003) that insert new points into the original data and take them as the watermark embedding positions. Moreover, an algorithm proposed by Jia, Chen, Ma, and Zhu (2006) inserts the bits of the watermark in the least significant bits (LSBs) of the coordinates to make the watermark capable of resistant to the data revision. The advantage of space domain algorithms is the precision of the watermarked data is controllable, and the watermarks generated by these algorithms are generally resistant against additive random noise, similarity transformation, and vertices revision, to some extent. However, none of such algorithms are blind in detection process.

To critically sum up the above review on the watermarking techniques in space domain: (1) the space domain algorithms prevail over the frequency domain ones in preserving the precision of watermarked data. This should be emphasized because the precision is of most importance to geo-spatial data. (2) The existing space domain algorithms do not differentiate among point, linear and areal features, and take into little consideration of the spatial characteristics of geo-spatial data. This is important for our new watermarking approach and will be further discussed in Section 2.1. (3) None of the existing space domain algorithm is blind in the watermark detection process.

For these reasons, our research will make improvements at the above three points and aim at proposing a blind watermarking approach in space domain that allows copyright protection of vector geo-spatial data.

The remainder of this paper is organized as follows. Section 2 details an algorithm that allows watermark embedding in vector geo-spatial data. Section 3 describes an algorithm for the watermarking detection. These two sections comprise the main parts of the work. Section 4 describes two examples to test the validity of the proposed watermarking approach. Finally, conclusions are drawn in Section 5.

2. Watermark embedding algorithm

2.1. General approach

To overcome the shortcomings of the existing watermarking algorithms in the space domain, the following strategies have been employed in our new approach:

(1) *Inserting the watermark multiple times in the feature layers*. It is a common sense in the community of cartography and geographic information science that vector geo-spatial data are divided into multiple feature layers for the purpose of storage and management. If the watermark can be multiply embedded in different feature layers, it must become more difficult to be removed. Therefore, its robustness can be improved. More importantly, multiple embedding means multi-versions of the watermarks can be detected from different feature layers. This provides a potential for judging if the data contains the watermark by comparing the extracted watermarks without the original data and original watermark, i.e. blind detection of the watermark.

(2) *Utilizing the key points as the watermark embedding positions.* Key points are more important in their geometric and/or attribute aspects than the other ones; so they own less probability to be removed and/or edited by the users. Without doubt to embed the watermark in the key points are favorable to the robustness of the watermarking technique.

(3) *Using the LSBs of the key point coordinates to embed the watermark*. The number of LSBs used for watermarking in each coordinate can be adjusted according to the data precision requirements from practical applications; hence, the data precision change caused by watermark embedding is controllable. In this sense, the quality of the watermarked data can be maintained.

Based on the above strategies, an algorithm for embedding watermarks in vector geo-spatial data is proposed. Fig. 1 demonstrates its principles and procedures: determination of the feature layers for watermark embedding, selection of the key points, preparation of the watermark, and watermark embedding.

2.2. Determination of embedding feature layers

There are generally multiple feature layers in a vector geo-spatial database. For example, China's topographic map database at scale 1:1000,000 comprises of 14 layers, including control points, hydrography, settlements, roads, topography, boundaries, vegetations, etc. Deciding how many and which feature layers should be used for watermark embedding is the key in this procedure. Our experiences and experiments identify the following four rules for the selection:

Rule 1: The ideal number of the feature layers for watermark embedding should be three for the purpose of the robustness

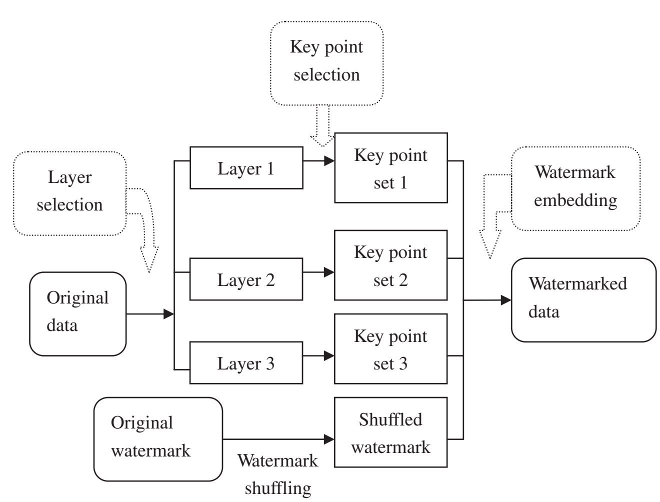

Fig. 1. Procedures of watermark embedding.