Linear Feature Simplification Using Wavelets in GIS*

Chen LIANG Chung-Ho LEE Jae-Hong KIM Hae-Young BAE
Department of Computer Science and Engineering, Inha University, Inchon, 402-751, KOREA
E-Mail: liangchen76@hotmail.com

ABSTRACT

Feature Simplification is an essential method for multiple representations of spatial features in GIS. However, spatial features are various, complex and of a large size. Among spatial features which describe spatial information, linear feature is the most common. Therefore, an efficient linear feature simplification method is most critical for spatial feature simplification in GIS.

This paper proposes an original method, by which the problem of linear feature simplification is mapped into the signal processing field. This method avoids conventional geometric computing in existing methods and exploits the advantageous properties of wavelet transform. Experimental results are presented to show that the proposed method outperforms the existing methods and achieves the time complexity of O(n), where n is the number of points of a linear feature. Furthermore, this method is not bound to two-dimension but can be extended to high-dimension space.

1. Introduction

Feature Simplification is an essential method for multiple representations of spatial features in GIS or automatic cartography. As computers become increasingly involved in automated cartography, efficient algorithms are needed for the tasks of extraction and simplification [7]. Likewise, many researches in GIS realm such as [3,4,8,9] stress the indispensability of multi-representations of spatial features. In GIS spatial features are associated with scale naturally [8]. Different representations of spatial features should be presented to users according to various scales. However, practical application of feature simplification faces a big challenge. It is well-known that a map consists of a large number of features and these features are various and complex and of a large size, due to the complex geographic phenomena. As a consequence, a good performance is very important for feature simplification methods in processing a large number of complex features. Among a variety of features, linear feature is the most common features to describe spatial information [9]. Informally, a linear feature is a term for a group of line-like features, such as point, polyline and polygon, whose data structure is usually a sequence of points. Thus, an efficient linear feature simplification method is an unavoidable and most critical issue in spatial feature simplification in GIS.

This paper proposes a novel method, by which the problem of linear feature simplification is mapped into the signal processing field. This method avoids going through conventional methods — geometric computing in existing simplification algorithms, and exploits the advantageous properties of wavelet transform. Experimental results are presented to show that the proposed method outperforms the existing methods and achieves the time complexity of O(n), where n is the number of points of a linear feature. Furthermore, this method is not bound to two-dimension but can be extended to high-dimension space.

The reminder of this paper is organized as follows. In section 2 we revisit some existing methods. In section 3 we give a detailed description to the proposed method. In section 4 we discuss its performance and demonstrates experimental results. In section 5 we state the conclusion and future work.

2. Related Work

The task of feature simplification is to erode the minor spatial variations (details) and extract the major ones (approximations), that is, to characterize the shape according to a given scale. The large number of complex spatial features also inhibit linear feature simplification algorithm to achieve good efficiency. Over decades, many studies have been carried on in the linear feature simplification algorithms and their performance improvement.

Perhaps the well-known one is the line simplification algorithm proposed by Douglas and Peucker in 1973 [9]. Douglas-Peucker Line Simplification Algorithm is simple and easy to implement but suffers a high time complexity, quadratic in the number of points in the worst case [7]. Recently, some variations [7] on this algorithm were made to enhance the performance of simplification. Although they improve the complexity up to $O(n\log_2 n)$, where n is the number of points of a linear feature, it is still costly to apply in a large number of complex linear features in GIS. In fact, the common of these methods is to use geometric computing and establish a mathematic framework to decide which points are removed or preserved. Such methods inevitably lead to low performance because they directly deal with the complex shape of a linear feature in GIS.

The proposed method avoids conventional routines in existing methods, i.e., geometric computing. It models the problem of linear feature simplification into signal processing, in order to take advantage of wavelet transform. Consequently, the time complexity of this method reaches O(n). Another advantage of this method is that it can be extended to high dimension space. To our best knowledge, this is the first time to apply wavelet transform into linear feature simplification in GIS.

3. Linear Feature Simplification Using Wavelets

This section discusses the proposed method — Wavelet-Based Linear Feature Simplification — in detail.

3.1 Wavelet Transform

In signal processing field, wavelets are mathematical functions that cut up a signal into different frequency components, and then study each component with a resolution matched to its scale [5]. According to their definition, wavelets are well-suited for approximating data with sharp discontinuities. Wavelet transform is a very powerful analytical tool for signal processing. Moreover, wavelet transform has been applied in a variety of fields beyond signal processing, such as image processing, patter recognition, arbitrary intelligence and, more recently, spatial data handling [6,10]. Although wavelet transform is involved in complex mathematic analyses and theoretical proof, in this paper we would rather concentrate on the application of wavelet transform on linear feature simplification than on theoretical mathematic proof.

The property of wavelet transform used by this method is detail filtering or signal denoising. Figure 1 illustrates the effect of wavelet transform to filter details out of a signal. The original signal is at left side and the filtered one is at right side. Comparison between these two charts shows that wavelet transform approximates a signal while suppressing its details.

* The research is supported by Ministry of Information and Communication under work of university S/W research center.

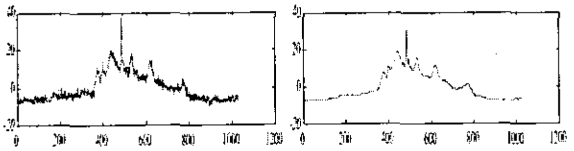


Figure 1. "Before" and "After" Illustrations of Wavelet Transform on a Signal

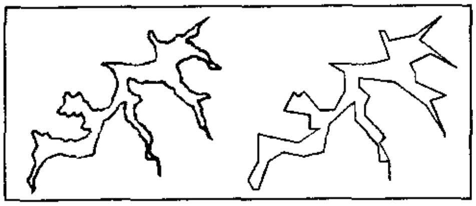


Figure 2. "Before" and "After" Illustrations of *Simplification on a Linear Feature*

3.2 Wavelet-based Simplification

From Figure 1 and Figure 2, we can observe the shared properties. Firstly, high-frequency parts of a signal correspond to details of a linear feature while low-frequency parts to major spatial variations. Secondly, the details are removed while those major spatial variations are preserved during feature simplification or signal denosing.

By analogy, a linear feature (2D) can be regarded as two noisy signals (1D). Therefore, its major spatial variations describe the cleared-up signal while spatial details are noises. Correspondingly, scale is regarded as the noise level or amplitude, which therefore provides us to decide which points should be removed or preserved. Thus, the problem of linear feature simplification could become the problem of denosing a noisy signal. Then, we could exploit wavelet transform for linear feature simplification because it does a very good job in signal denosing. *If the details are small enough, they might be omitted without* substantially affecting the main characteristics of a feature. Since some of the resulting wavelet coefficients correspond to details in the dataset [5], wavelet transform provides us a useful technique to distinguish details from spatial sharps with wavelet coefficients during simplification.

Algorithm 1 describes Wavelet-Based Linear Feature Simplification. Correspondingly, Figure 3 illustrates the simplification procedure. In Algorithm 1, the principle of wavelet *shrinkage and thresholding* method is employed, whose mathematic proof given in [5], and Fast Wavelet Transform (FWT) algorithm (its source code in [1]) is used to implement wavelet transform, whose time complexity is O(n) [1].

***Algorithm 1.* SimplifyObj(Points, n, S)**

Input: Points: a sequence of points describing a linear feature, each consists of x (x_i) and y (y_i) coordinates.
n: the number of points in the linear feature.
S: the scale for simplifying the linear feature.

Output: the simplified linear feature.

01. Take the x-coordinates and y-coordinates of a linear feature as two individual input data. Fast Wavelet Transform in [1] can process a signal whose range is within (0, 1). To take *advantage of the algorithm, x and y coordinates* are to be mapped into the range (0, 1) and preprocess those signals by $\mathbf{sig}_{forward} = (x_i / \max(x)) \big/ \sqrt{n}$ or $(y_i / \max(y)) \big/ \sqrt{n}$.
02. *Perform the forward wavelet transform (FWT) to the result* signals $\mathbf{sig}_{forward}$. As a result, a set of coefficients is generated.
03. Apply the soft-threshold nonlinearity [2] to the generated coefficients and set zero to those coefficients less than the threshold $t = \sqrt{2\log(n)} * S' \big/ \sqrt{n}$, where S' is the result of the specific scale mapped into (0, 1). Through this step, those signals corresponding to the zero wavelet coefficients are filtered according to the specific scale.
04. Perform the inverse wavelet transform (inverse FWT) with the filtered coefficients to reconstruct the signals $\mathbf{sig}_{inverse}$. *Only those signals corresponding to the nonzero wavelet* coefficients are reconstructed.
05. Map the reconstructed signals to the original spatial coordinate system by the inverse transforms in line 01. Then, simplified linear features can be constructed from those processed coordinates.

End of Algorithm 1

Between two wavelet transforms we use wavelet shrinkage and soft-thresholding [2] to filter a signal, where wavelet coefficients and threshold provide us with a flexible means to control the level of detail of simplified linear features.

Note that because wavelet transform is different from conventional geometric computing, we avoid handling geometrically complex linear features directly.

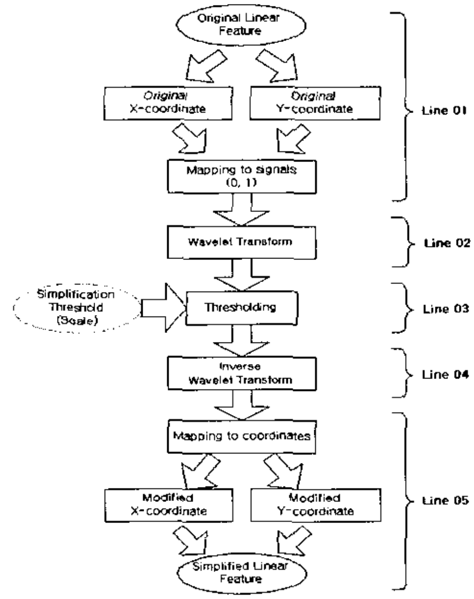


Figure 3. The Procedure of Wavelet-based Linear Feature *Simplification*

4. Performance Evaluation

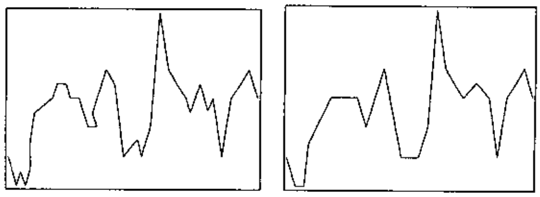
4.1 Time Complexity Analysis

First let us consider a single linear feature in two-dimensional space. Let n be the number of points in a linear feature, where n may be a very large number. We analyze the time complexity by going through Algorithm 1 in section 3.2. Complexity of mapping x or y coordinates (line 01) is O(2*n). Complexity of applying the forward wavelet transform (line 02) is O(l*2*n) = O(n), where l is a small constant representing the length of filter used in the forward wavelet transform. Wavelet transform generates 2*n coefficients, so complexity of thresholding these coefficients (line 03) is O(2*n) = O(n). And like the forward wavelet transform, the inverse wavelet transform (line 04) causes O(l*2*k) complexity, where k (k <= n) is the number of the remaining points and l represents the length of filter used in inverse wavelet transform. Complexity of mapping the *results into the original spatial space (line 05) is O(2*n) at most.* Altogether, the overall complexity of wavelet-based linear feature simplification is O(2*n) = O(n). Compared with Douglas-Peucker algorithm and its improved ones, the complexity of this method is improved much. Its good performance will contribute to the practical application of this method in a large number of complex linear features. And this method treats x and y coordinates separately so that it is ready for parallel processing.

Besides its advantage on complexity, this method can be extended to high dimensional space because wavelet transform treats one-dimensional data at one time. In the d dimensional space, the complexity is $O(d*n)=O(n)$.

4.2 Experiment Results

We did experiments on arbitrary and real data. DS1 is arbitrarily generated shown in Figure 4(a), while DS2 is the real data describing the coarse boundary of Korea mainland in Figure 5(a). DS1 has 50 points and DS2 is made of 984 points. We separately apply the proposed method and improved Douglas-Peucker algorithm [7] to those two linear features. Figure 4(b) shows the simplified feature DS1. Figure 5(b) and (c) show the simplified feature D2. Table 1 demonstrates the comparison between wavelet-based method and improved Douglas-Peucker algorithm in complexity and processing time. The experiment results demonstrate good efficiency of the proposed method in simplifying linear features, especially when dealing with a large number of complex ones.

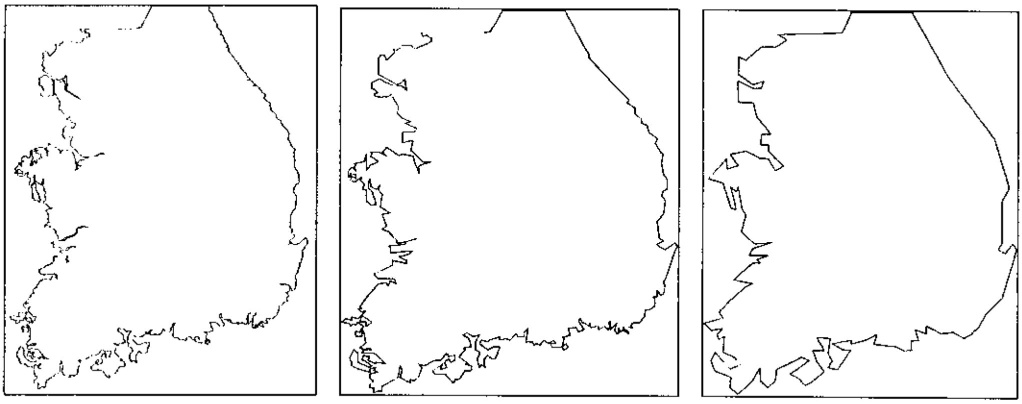


(a) (b)

Figure. 4. (a) is the original feature DS1 for the experiment and (b) is the simplified one

Number of points in a linear feature	Complexity	DS1 (50 pts)	DS2 (984 pts)
Wavelet-Based Linear Feature Simplification	O(n)	155	3051
Improved Douglas-Peucker Algorithm	$O(n\log_2 n)$	237	7826

Table 1. The Required Time (in millisecond) of Simplification



(a) (b) (c)

Figure 5. (a) is the original feature DS2 (Korea Mainland) for the experiment, (b) is the simplified one at scale 1:200 and (c) is the simplified one at scale 1:500

5. Conclusion and Future Work

In this paper, we proposed an original and efficient method for the linear feature simplification. It maps the problem of linear feature simplification into the signal processing field and applies wavelet transform which preserves the major characteristics of a linear feature and removes minor spatial variations according to a given scale. It avoids conventional geometric computing, and therefore has achieved the best time complexity of O(n), where n is the number of points of a linear feature. As a result, it is particularly attractive for simplifying a large number of complex linear features in GIS. Our experimental results demonstrated that the proposed method outperforms existing methods. Moreover, this method is not bound to two-dimension but can be easily extended for processing high dimensional features.

Since the application of wavelet transform in spatial feature simplification is at the beginning, there are a lot of things that need us to reconsider. For instance, which one among a variety of wavelets in what situation can make the best contribute to simplify various spatial features in GIS.

References

[1] Mac A. Cody, The Fast Wavelet Transform, http://www.acm.org/pubs/articles/journals/drdobbs/1992/9204/9204a/9204a.htm

[2] David L. Donoho and Iain M. Johnstone, Ideal Spatial Adaptation by Wavelet Shrinkage, http://www-stat.stanford.edu/~donoho/Reports/ 1992/isaws.pdf

[3] Clodoveu A. Davis Jr. and Alberto H. F. Laender, Multiple Representations in GIS: Materialization Through Map Generalization, Geometric, and Spatial Analysis Operations, ACM GIS '99

[4] Paolino Di Felice and Illiseo Clementini, Towards a Standard for SQL-Based Spatial Query Languages, ACM 1992

[5] Amara Graps, An Introduction to Wavelets, Institute of Electrical and Electronics Engineers, Inc., 1995

[6] Markus H. Gross, Oliver G. Staadt and Roger Gatti, Efficient Triangular Surface Approximations Using Wavelets and Quadtree Data Structures, IEEE Transactions On Visualization and Computer Graphics, Vol. 2, No. 2, June 1996

[7] John Hershberger and Jack Snoeyink, An O(nlogn) implementation of the Douglas-Peucker algorithm for line simplification, Proceedings of the tenth annual symposium on Computational geometry, June 1994

[8] Chen Liang, Chung-Ho Lee, Zu-Kuan Wei and Hae-Young Bae, Efficient Data Transmission Using Map Generalization On Client-side Web GIS, Kiss Fall Conference, 2000

[9] Robert Laurini and Derek Thompson, Fundamentals of Spatial Information Systems, Academic Press, 1992

[10] Gholamhosein Sheikholeslami, Surojit Chatterjee and Aidong Zhang, WaveCluster: A Multi-Resolution Clustering Approach for Very Large Spatial Databases, Proceedings of the 24th VLDB Conference, 1998