Sensors & Transducers, Vol. 178, Issue 9, September 2014, pp. 209-213

Sensors & Transducers

Researches on Key Algorithms in Analogue Seismogram Records Vectorization

[1, 2] Maofa WANG, [1] Qigang JIANG, [2] Jilin FENG, [2] Shichen FENG
[1] College of Geoexploration Science and Technology, Jilin University, Jilin, Changchun, China, 130026
[2] Department of Disaster Information Engineering, Institute of Disaster Prevention, Hebei, Sanhe, China, 065201
E-mail: wangmaofa2008@126.com

Received: 22 June 2014 /Accepted: 29 August 2014 /Published: 30 September 2014

Abstract: History paper seismograms are very important information for earthquake monitoring and prediction, and the vectorization of paper seismograms is a very import problem to be resolved. In our study, a new tracing algorithm for simulated seismogram curves based on visual filed feature is presented. We also give out the technological process to vectorizing simulated seismograms, and an analog seismic record vectorization system has been accomplished independently. Using it, we can precisely and speedy vectorize analog seismic records (need professionals to participate interactively). *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Analog seismic record, Vectorization, Digital image processing, Visual field feature.

1. Introduction

Historical paper earthquake records are important information for earthquake monitoring and prediction [1-2]. Although the construction of digital seismic stations has been undertaken in many countries at present, those massive saved simulated seismogram records over the past decades still have a great value, which is important original information associated with the earthquake-generating process for analyzing and studying earthquake.

As is known to all, with the development of computer technology, a large number of digital seismic records across time and space are need in seismic research. Unfortunately, in a data explosion era, analog seismic records still cannot be processed quickly by computer, although they are an important part of human seismic records. This means that the analog seismic record will lose its precious historical value without digitized and vectorized.

In Harvard University, a seismograph station located at the Adam Dziewonski Observatory (Oak Ridge, MA) was deployed in 1933 which is one of the oldest seismograph stations in the USA. The recordings of this station after 1972 are digital and accessible online through IRIS. However, the earlier seismograms were not digitized or cataloged in a systematic manner. Starting around mid 2012, a named HRV seismogram Archival Project is established, and a significant amount of resources are devoted to clean and digitize the old seismograms from the HRV station [3]. After the processed, simulated earthquake records are stored and released in digital image form, which is more convenient to store and read, but still cannot be processed such as seismic data mining and extraction easily and quickly by computer without further vectorizing.

In China, many seismic stations such as Chengde seismic station have a great deal of analog seismic records, and a lot of recording drawings are piled up in warehouses, long-term suffering from rat biting, air mildew, erosion of various natural phenomena. In today's digital world, simply digitization such as taking pictures by camera and scanning by scanner, also just to prolong its storage period, and no further reflect or extend its using value, can say is alive but no significance.

So it is an important and urgent task for vectorizing paper analog seismic records. In this paper, we present a new tracing algorithm for simulated seismogram curves based on visual filed feature(VFF), give out the whole technological process to vectorizing simulated seismograms, and an analog seismic records vectorization system has been accomplished independently. Using it, we can precisely and speedy vectorize analog seismic records (some time need professionals to participate interactively).

2. Key Algorithms Analysis and Discussion

2.1. The Whole Technological Process to Vectorizing Simulated Seismograms

We want to vectorize paper simulated seismograms with digital image processing technique. Fig. 1 is the whole technological process to vectorizing simulated seismograms system, throughout it, there are six approaches: loading images, preprocessing, vectorizing image, waveform and time coordinate mosaic, Warehousing waveform and time coordinate, Waveform inversion. In the paper, we give the detailed discuss and depict of the above key approaches.

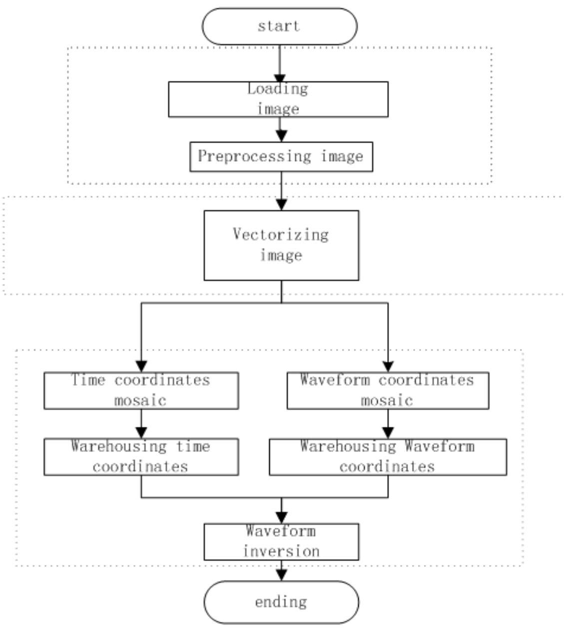

Fig. 1. Technological process to vectorizing simulated seismograms.

2.2. Image Preprocessing

Firstly, all the simulated seismogram drawings are digitally scanned using industrial scanner, and different formats of scanning images are converted into PNG format in loading image stage.

To reduce the noise of the subsequent vectorization, image binarization is in need. Binarization is a digital image processing technique, by which one image can be divided into two parts: object part and background part, and here Otsu algorithm is used [4-6]. Fig. 2 (a) is an original image, and Fig. 2 (b) is the last binarization processing result.

In order to reduce the effective waveform information (entropy) damage because of pretreatment to a minimum, we do not adopt further preprocessing methods, especially line thinning. And all the subsequent processing steps are based on binary images.

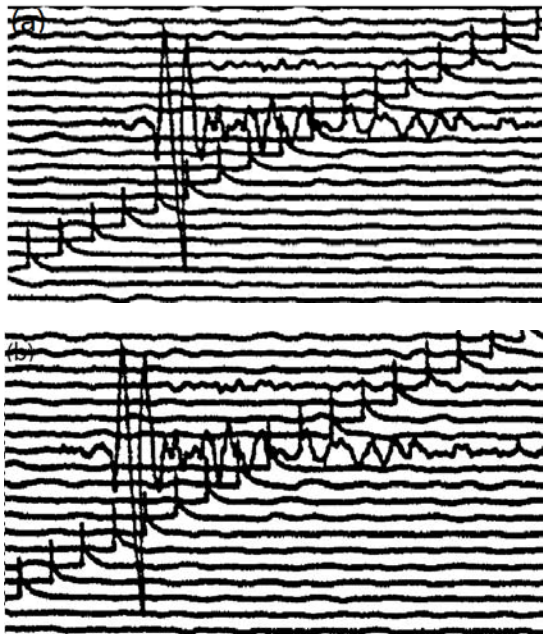

Fig. 2. (a) An original image of simulated seismogram; (b) The last binarization processing result.

2.3. Image Vectorizing Technological Process

For every one curve, firstly we need to acquire its starting point. Then we start to trace one curve waveform from the starting point, and record all the key points on it which will be stored in the database. The whole tracing process is like a man climbs a lot of mountains, sometimes the mountains (waveforms) are gentler, and sometimes the mountains (waveforms) are very much steep. For different waveforms, we use different tracing algorithms which will be discussed in following paragraphs. When the tracing result of one curve is not so satisfactory, we can choose a back-roll point. And from the back-roll point, we can pick out key points manually or automatically tracing points within in a smaller local area. When the tracing precision meets requirement, automatically tracking is continuing again. A technological process to vectorizing seismogram curves is shown in Fig. 3.

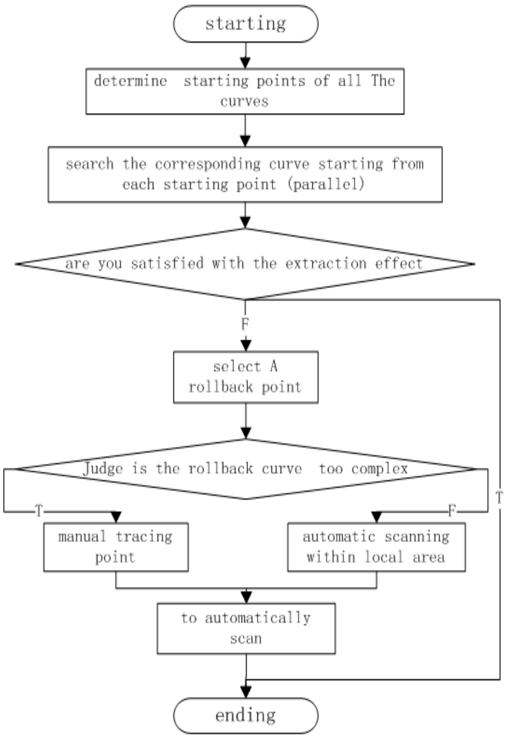

Fig. 3. Technological process to vectorizing seismogram curves.

2.4. A Tracking Algorithm for Smooth Waveforms of one Seismogram Curve

According to the characteristics of seismogram curves, the starting point $C(x_0, y_0)$ of one curve is taken as its first key point, and the whole searching direction is horizontal right. Horizontal searching step length is $step_x$, and vertical searching width is $step_y$ which is the average width of adjacent curves and is set dynamically. Fig. 4 is the sketch map of tracking algorithm for smooth waveforms. The specific algorithm is as follows:

Horizontal direction

$$x_1 = x_0 + step_x,$$

$$x_2 = x_1 + step_x,$$

$$x_n = x_{n-1} + step_x, \tag{1}$$

Vertical direction

If $C_n(x_n, y_n)$ is one known key point, we start to judge whether point $C_n'(x_n + step_x, y_n)$ is black, if it is white, turn to complex waveform searching which will discussed in following paragraph. If point $C_n'(x_n + step_x, y_n)$ is black, we search black pixels in two directions to within the range $[y_n - step_y, y_n + step_y]$ of point $C_n'(x_n + step_x, y_n)$ up and down, until the emergence of discoloration, and get two discolor points: $C_L(x_n + step_x, y_{nL})$ and $C_H(x_n + step_x, y_{nH})$.

$$y_{n+1} = \frac{y_{nH} + y_{nL}}{2}, \tag{2}$$

That is to obtain $C_{n+1}(x_{n+1}, y_{n+1})$:

$$\begin{cases} x_{n+1} = x_n + step_x = x_0 + (n+1) * step_x \\ y_{n+1} = \dfrac{y_{nH} + y_{nL}}{2} \end{cases}, \tag{3}$$

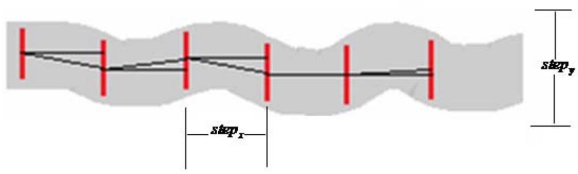

Fig. 4. Sketch map of tracking algorithm for smooth waveforms.

Fig. 5(a) is an original image simulated seismogram, and Fig. 5 (b) is the last smooth waveforms processing result.

2.5. A Tracking Algorithm for Complicated Waveforms of one Seismogram Curve

In our research, we have tried many times using different ways to trace complicated waveforms, including of tracking algorithm based on sparse pixel [7], tracing method with anti-noise analysis [8], homotopic method following solution curve [9], searching negative curvature points [10], and the like. But all the results do not look so good. Through a lot of analysis, we found that some waveform segment of one seismogram curve may be so complicated which includes many steep wave crests, irregular cross sections, smudge, breakpoints and so on. So it is very difficult to find a way to deal with all the interference factors for tracing complicated waveforms.

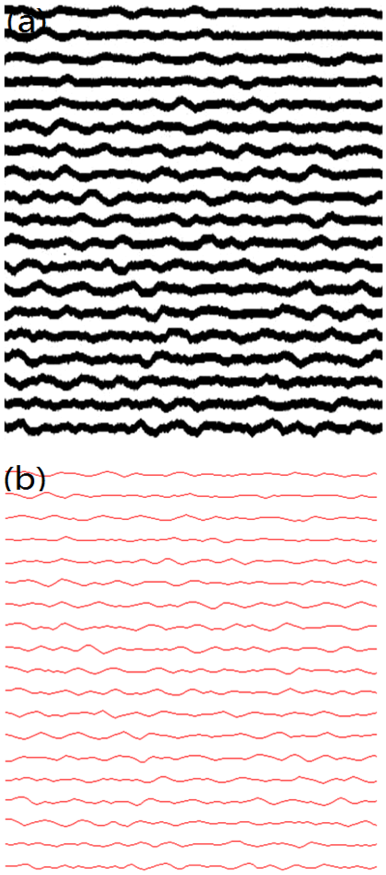

Fig. 5. (a) One original seismogram curve image, (b) The last smooth waveforms processing result.

On further analysis, however, it is noticed that seismic wave energy will increase and decrease with periodic and has equilibrium position of its own. If we will shift an observation point (a starting position for searching next one key point) slightly further from next one key point, not nearby in next one key point, and relatively fixed such as wave energy equilibrium position, errors and noise will be isolated in a very small range, and will not stack and influence each other. Based on the above analysis and thinking, finally we propose a tracking algorithm for complicated waveforms based on visual filed features (VFF). It takes the steps as follows:

1) Suppose $C_n(x_n, y_n)$ is a known key point, we can judge whether Point $C_n'(x_n + step_x, y_n)$ is black, if it is black, turn to smooth waveforms which has been discussed in the above paragraph.

2) If $C_n'(x_n + step_x, y_n)$ is white, we search black pixels in two directions to within the range $[y_0 - step, y_0 + step]$ of Point $C_n''(x_n + step_x, y_0)$ up and down, and Point $C_n''(x_n + step_x, y_0)$ is an energy equilibrium position. Then saving y coordinate values of all the lower boundary points of black connected area searched above into set $a_{down}[i](\mathrm{i} = 0,1,\ldots,\mathrm{m})$, and y coordinate values of all the Upper bound points of black connected area searched above into set $a_{up}[i](\mathrm{i} = m+1,\ldots.)$,. So that I can get a points set $C_{n+1}'(x_{n+1}, a[i])(\mathrm{i} = 0,1,\mathrm{m},\mathrm{m}+1\ldots)$.

3) We can establish many groups of four points: $(x_n - \Delta, \mathrm{y}_n)$, $(x_n + \Delta, \mathrm{y}_n)$, $(x_{n+1} - \Delta, a[i])$, and $(x_{n+1} + \Delta, a[i])$ $(\mathrm{i} = 0,1,\ldots)$ where Δ is an experience value decided by dynamically. For each a candidate key point $C_{n+1}'(x_{n+1}, a[i])$, we can get a parallelogram as connected monitoring area with the four points, and whose area can be expressed as Eq. (4).

$$S_i = \Delta * |a[i] - y_n|, \quad (4)$$

4) Then we calculate out the number of black pixels in each parallelogram cited above denoted by n_i, and the ratio of n_i to the area of the Corresponding parallelogram denoted by η_i can be computed by Eq. (5).

$$\eta_i = \frac{n_i}{S_i}, \quad (5)$$

The larger the value of η_i corresponding to one candidate Point $C_{n+1}'(x_{n+1}, a[i])$ is, the better the connectivity between it and the previous key point $C_n(x_n, y_n)$ becomes. So point $C_{n+1}'(x_{n+1}, a[i])$ which owns the maximum value of η_i is fixed as the next key point.

5) If $C_{n+1}(x_{n+1}, y_{n+1})$ is one new key point searched by the above steps, we start to judge whether point $C_{n+1}'(x_{n+1} + l, y_{n+1})$ $(l < \Delta)$ is black, if it is white, turn to smooth waveform searching discussed on the above paragraph. If point $C_{n+1}'(x_{n+1} + l, y_{n+1})$ is white, Point $C_{n+1}''(x_{n+1} + l, y_0)$ is taken as one new energy equilibrium position. And start searching from which as by steps (2-4), lastly we can get the next key point $C_{n+2}(x_{n+2}, y_{n+2})$.

6) Repeat steps (1-5), until all the key points of the whole curve are detected out.

Fig. 6 is the sketch map of tracking algorithm for complicated waveforms based on VFF.

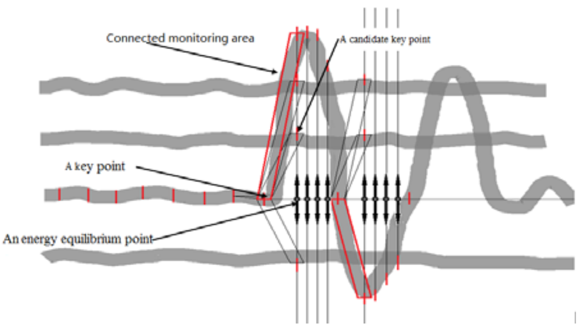

Fig. 6. Sketch map of tracking algorithm for complicated waveforms based on visual field features (VFF).

Fig. 7(a) is one complicated seismogram curve image; (b) is the detected result of one corresponding complicated waveform; (c) is the overall detected result.

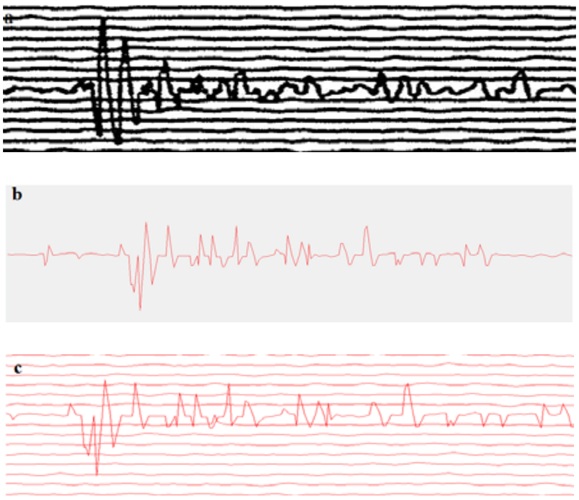

Fig. 7. (a) One complicated waveforms image; (b) The detected result of one corresponding complicated waveform; (c) The overall detected result.

3. Summary

The paper focuses on key algorithms in analogue seismogram records vectorization. From the detected results of smooth and complicated waveforms as shown in Fig. 5 and Fig. 7 we can see that the topology of the tracing result is consistent with o original curves.

We also write the correlative software with c# codes, which applying the above digital image processing algorithms. The analogue seismogram records which come from Chengde seismic station in 1991 and excess of 10 G have been vectorized using the software.

And Mosaic Algorithm of waveforms and time coordinates by vectorized will be studied and discussed on in a subsequent article.

Acknowledgment

Two of Authors (Maofa Wang, Jilin Feng) thank the financial support from Special Fund of Fundamental Scientific Research Business Expense for Higher School of Central Government (Projects for creation teams)(Research on the Acquisition and rapid Processing of earthquake disaster information ZY20120104).

References

[1]. Jiansheng Xu, Kangsheng Xu, Yonggang Wei, Xiangyun Guo, Saving and data sharing of historical seismogram in Beijing National Earth Observatory, *Seismological and Geomagnetic Observation and Research*, 2008, Vol. 29, Issue 3, pp. 100-104.

[2]. Zhian Pan, Jilin Feng, Maofa Wang, A base-point searching algorithm in the digitization of seismograms, *Communications and Information Processing*. 2012, Vol. 289, pp. 699-706.

[3]. Seismology Harvard http://www.seismology.harvard.edu/resources.html#HRV

[4]. N. Otsu, A Threshold Selection Method from Gray-Level Histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, Vol. 9, Issue 1, pp. 62-66.

[5]. Maofa Wang, Jilin Feng, Recognition of amount of carbon atoms of graphite STM images, *Journal of Computational and Theoretical Nanoscience*, 2011, Vol. 8, Issue 11, pp. 2204-2208.

[6]. Maofa Wang, Xiaoping Zou, Recognition of atomically-resolved STM images of graphite, *Journal of Computational and Theoretical Nanoscience*, 2010, Vol. 7, pp. 404-407.

[7]. Baifang Chen, Research and implementation of log curve vectorization methods and rebuilding technique, *Harbin Institute of Technology*, 2007.

[8]. Hang Zuo, Automatic recognition and tracing of curves on well log graphs, *University of Sichuan*, 2000.

[9]. Zhou Zhou, Xiaoping Zhu, Zhaozhi Zhang, Qiangang Liu, A new homotopic algorithm Following Solution Curve, *Journal of Northwestern Polytechnical University*, 1998, Vol. 16, Issue 2, pp. 246-250.

[10]. Danhua Cao, Jinkuan Tang, Yubin Wu, Algorithm for fast extracting human limb contours using searching NCM points, *J. Huazhong Univ. of Sci. &Tech. (Nature Science Edition)*, 2007, Vol. 35, Issue 5, pp. 16-18.

____