Digitization of Well-Logging Parameter Graphs Based on Gridlines-Elimination Approach

Bo Yuan, Qing Yang*

Southwest Petroleum University, School of Electrical Engineering and Information and No. 8, Xindu Avenue, Chengdu, Sichuan, China.

* Corresponding author. Tel.: 18382085073; email: 201722000106@stu.swpu.edu.cn
Manuscript submitted October 10, 2019; accepted December 22, 2019.
doi: 10.17706/jsw.14.12.573-578

Abstract: Most of the well-logging parameter graphs were drawn on the paper before digital logging instruments were widely used in oilfield. How to keep the parameter graphs permanently and use the data implied in those curves efficiently are still open problems in the petroleum industry. In this paper, the gridlines-elimination approach is proposed to digitize well logging curves according to the characteristics of well-logging parameter graphs. Theory of morphological image processing and pixel statistics method are employed to eliminate gridlines, which isolates the curves and the gridlines with incredible effect. The remaining tiny grid lines and noise point are cleared according to the characteristics of the small size of their connected components. The X-Y coordinates of curves are obtained according to the middle position of each row of pixels in the pixel matrix of well-logging parameter graphs. Comparison with Neuralog demonstrates that the approach based on gridlines-elimination is valid and superior to Neuralog.

Key words: Digitization, gridlines-elimination, morphological image processing, well-logging parameter graphs.

1. Introduction

Well-logging [1] is the process of taking measurements of various rock properties along the length of the well down into the ground by drilling tools. Log responses [2] are functions of lithology, porosity, fluid content, and textural variation of formation. The well-logging parameter is being widely used in reservoir description, such as the classification of well log data into different lithofacies groups, facies-by-facies description of rock properties.

The current world oil and gas industry is facing the main challenges brought by low oil prices, and cost control has become an considerable factor affecting the competitiveness of oil companies. In this era of digitization and intelligence, the oil and gas industry expects to achieve cost reduction and efficiency through digital and intelligent transformation to meet this challenge. For example, machine learning is used to analyze and predict lithologic facies, which extremely reduces uncertainty and improves the accuracy of conventional and unconventional reservoir description. But one impediment here is that before digital logging instruments were widely used, well-logging data was drawn on the parameter graph in curve format. Well-logging parameter graphs have many disadvantages: large size, large memory space, and interference information such as gridlines. So it is necessary to convert well-logging parameter graphs into X-Y coordinates, where X represents parameter values, and Y represents depth values.

To digitize the curves embedded in the binary image, several methods have been proposed and implemented. These methods are roughly divided into two categories: pixel-based and non-pixel-based.

Pixel-based methods include the thinning method and the Global Curve Vectorization (GCV) method [3]. The thinning method [4]-[7] is to reduce the width of a line to only one pixel, leaving only the skeleton that can characterize its features. The main disadvantage of the thinning method is that it has high time complexity, loses line width information, and is prone to deformation and wrong branches in the intersection area. GCV is an improvement of Sparse Pixel Tracking (SPV) method [8]. It traces the curve according to the local slope of the curve, adjusts the direction step by step in the tracking process, and adopts the dynamic step size, so that it not only can cross the intersection and adhesion parts of the line correctly, skip the gap of the curve, but also be applicable to any curved curve. However, the cross-recovery ability of GCV is limited, and it can't solve all cross-over problems. GCV method is suitable for line processing, but poor for point line processing.

Non-pixel-based methods mainly fall into two categories: contour-based and adjacency graph- based. The contour-based approach [9] is to extract the contour of the image first, and then find the matched contour pairs. The skeleton of curves can be obtained from the midpoint of the corresponding contour pair. The method based on adjacency graph firstly apply run-length encoding to graphs, then analyses the segments, and generates various adjacency graph structures, such as line adjacency graph (LAG) and block adjacency graph (BAG) [10]. Chunsheng Li *et al*. [11] proposed an SCTR (Scanning, Compressing, Tracing, and Rectifying) approach by employing LAG data structure. Yang et al. improved the SCTR method and put forward the PCTR (Preprocessing, Compressing, Tracing, and Rectifying) method [12]. Rao *et al*. proposed an algorithm for erasing grid-lines and reconstructing strokes in Chinese handwriting based on BAG [13]. Such methods are difficult to deal with the intricate situation in well-logging parameter graphs, especially the analysis of nodes.

Now the most famous logging curve digitization software is Neuralog (http://www.neuralog.com/pages/NeuraLog.html). Due to the interference of the background grid, this software frequently pauses during curve tracking. Therefore, removing the background grid is very important in the digitization process of the well-logging curve.

Given the shortcomings of the existing methods, this paper puts forward an approach to digitizing well-logging parameter graphs with gridlines-elimination based.

2. Workflow

Fig. 1 illustrates a workflow of the system for digitizing well-logging parameter graphs. The system may receive a well-logging parameter graph that has one or more measurements, and least one measurement has a set of values in which those values wrap around the curve. Once the curve part is received, the method may process the well-logging parameter graph. The processing may include eliminating gridlines, digitization, curve rebuilding. Once gridlines are eliminated, the method automatically generates X-Y coordinates of curves. Finally, the X-Y coordinate generates a new curve, which is compared with the original curve to evaluate the method.

Fig. 1. Digitization workflow.

3. Preliminaries

In this section, we will introduce the notion of morphological image processing which will be used in the following discussions. Dilation [14] is fundamental to morphological processing. There are some Preliminaries about dilation. The reflection of a set B, denoted $\hat{B}$, is defined as:

$$\hat{B} = \{w \mid w = -b, b \in B\} \tag{1}$$

The translation of a set B by point $z = (z_1, z_2)$, denoted $(B)_z$, is defined as:

$$(B)_z = \{c \mid c = b + z, b \in B\} \tag{2}$$

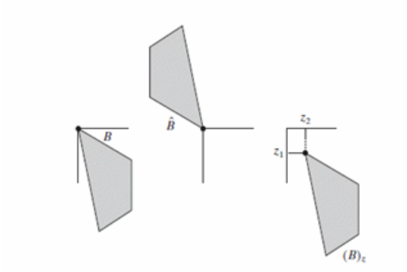


Fig. 2. Set B and its reflection.

Fig. 2 shows a simple set B and its reflection: $\hat{B}$ and $(B)_z$, With A and B as sets in Z^2, the dilation of A by B, denoted $A \oplus B$, is defined as:

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \varnothing\} \tag{3}$$

This equation is based on reflecting B about its origin, and shifting this reflection by z. The dilation of A by B then is the set of all displacements, z, such that $\hat{B}$ and A overlap by at least one element. Based on this interpretation, Eq. (3) can be written equivalently as

$$A \oplus B = \{z \mid [(\hat{B})_z \cap A] \subseteq A\} \tag{4}$$

Here B is a structuring element, and A is the set (image objects) to be dilated. Fig.3 shows the process of dilation. B in Fig. 3 is a structuring element (in this case $\hat{B} = B$ because the SE is symmetric about its origin). The five-point star is the central element of the structural element. The dashed line in Fig.3 shows the original set for reference, and the solid line shows the limit beyond which any further displacements of the five-point star by z would cause the intersection of $\hat{B}$ and A to be empty. Therefore, all points on and inside this boundary constitute the dilation of A by B. It's worth noting that dilation is the expansion of the highlight area of the image.

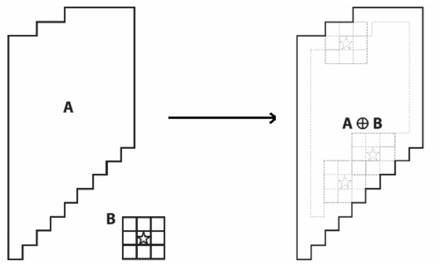


Fig. 3. The principle of dilation.

To define the connected component, the connectivity needs to be defined first [15]. It is sufficient to say that a pixel, A, with image pixel coordinates (r, c) is adjacent to four pixels, those with image pixel coordinates $(r-1, c)$, $(r+1, c)$, $(r, c+1)$, and $(r, c-1)$. In other words, each image pixel A (except those at the edges of the image) has four neighbors: the pixel directly above, directly below, directly to the right and directly to the left of pixel A. This relationship is referred to as 4-connectivity, and the two pixels are 4-connected if they are adjacent by this definition. If the definition of adjacency is expanded to include those pixels that are diagonally adjacent (i.e., the pixels with coordinates $(r-1, c-1)$, $(r-1, c+1)$, $(r+1, c-1)$, and $(r+1, c+1)$), then the adjacent pixels are 8-connected. A connected component is a collection of pixels, S, such that for any two pixels, say P and P', in S, there is a 4-connected path or an 8-connected path between them, and this path is contained in S [16].

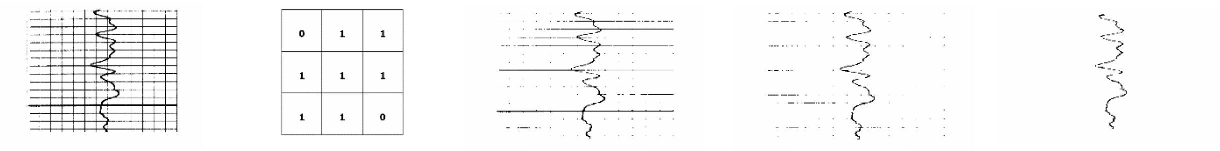
4. Gridlines Elimination and Digitization

It is observed that many uneven grids are mixed with curves in well-logging parameter graphs. Fig. 4(a) shows an extract of well-Logging parameter graph. In this section, three steps will be used to remove gridlines in well-logging parameter graphs.

Step 1: Vertical gridlines are removed by dilation. The result is shown in Fig. 4 (c). In the experiments, optimal result is achieved when the structuring element is (b) in Fig. 4.

Step 2: Pixel statistics method is employed to remove horizontal gridlines. As can be seen from the Fig. 4 (c), in the horizontal direction, the number of pixels in the gridlines is larger than the curve. Pixel statistics method was designed according to this feature. First of all, we count the number of black pixels per line of the well-Logging graphs, denoted counter. If the counter is less than the threshold, it is a well-log curve. Otherwise it is a horizontal gridline, and the horizontal gridline will be removed. The result after step2 is shown in Fig. 4(d), which has almost removed all the horizontal grid lines.

Step3: the method based on the size of the connected component is used to clean tiny grid lines, and noise point left over from step 1 and step 2. First of all, we reverse the color of the graph after roughly removing the gridlines and calculate the size of each connected component in the image, denoted sizer. If the size is less than the threshold, the connected component will be removed. Finally, step 1 is repeated. The result after step3 is shown in Fig. 4(e), which has removed all tiny grid lines and noise point.



(a) (b) (c) (d) (e)

Fig. 4. (a) Extract of well-Logging parameter graph; (b) Adopted structuring; (c) The graph after step 1; (d) The 135 graph after step 2; (e) The graph after step 3.

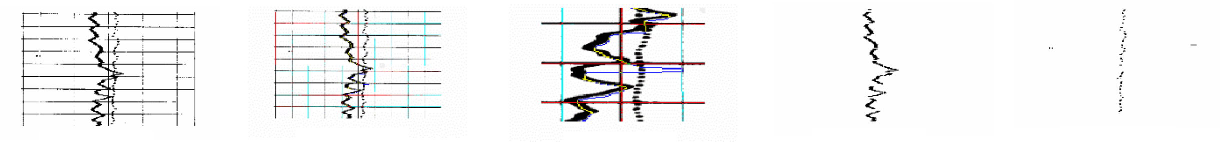
Since pixel-width of the curve in well-logging parameter graphs is greater than one, the X-Y coordinate of the middle axis of the curve is used to represent well-logging data. Assume that the top and bottom depths of the curve are d_1 and d_2 ($d_2 > d_1$), and the left and right parameter ranges are s_1 and s_2 ($s_2 > s_1$). The horizontal and vertical position of the middle axis of the curve is l and v in well-logging parameter graphs. The X-Y coordinates of the curve:

$$x = s_2 - (((s_2 - s_1) \times (l - l_{min}) / (l_{max} - l_{min}) + s_1) - s_1) \tag{5}$$

$$y = (d_2 - d_1) \times (v - v_{min}) / (v_{max} - v_{min}) + d_1 \tag{6}$$

5. Comparison and Verification

In this section, the proposed method and Neuralog are compared from the result of the curve extraction. One well-Logging parameter graph embedded in the cross line was used as a test-case for comparison. Fig.5 (a) shows extract of the test-case. In Fig.5 (c), the yellow and blue lines are the result of extraction, and the red and cyan are the background gridlines that are manually labeled. Obviously, the blue curve has deviated from the target curve, and the dotted line is not automatically extracted at all, and can only be extracted manually by point. The result after the proposed method is shown in Fig. 5(d-e) ， which has accurately extracted two curves. The result of the curve extraction in Fig.5 (b-e) shows that Neuralog has low automation and low precision, and the quality of curve extraction by the proposed method is better than that of Neuralog.



(a) (b) (c) (d) (e)

Fig. 5 (a). Extract of well-Logging parameter graph embedded in cross line; (b) The result of the curve extraction by Nerualog; (c) Partial enlargement of (b); (d) (e) The result of the curve extraction by the proposed method.

6. Conclusion

The gridlines-elimination based approach in this paper have been applied to precisely and automatically digitize well-logging parameter graphs. The proposed method and Neuralog are compared in this paper, and the experiments show that the method can successfully remove the gridlines in well-logging graphs, which makes it easy to automatically trace curves and greatly improves the working efficiency of digitization. In the future, more algorithms will be used to improve the automation and speed of digitization of well-logging curves.

Acknowledgment

This study was funded by National Science Foundation (No. 51504211 and 51204139), Sichuan Science and Technology Program (18GJHZ0195), Southwest Petroleum University, underground intelligent measurement and control youth research and innovation team (2018CXTD04) and Oil and Gas Equipment Key Laboratory Fund of the Ministry of Education (OGE201702-19). The authors thank the managers and staff workers for their hospitability, time and opinions.

References

[1] Purohit, P. *et al*., (2017). Preserving and securing legacy well log data. *Abu Dhabi International Petroleum Exhibition & Conference,13, Society of Petroleum Engineers*. Abu Dhabi, UAE

[2] Hong, Y. (1979). *Application of Logging Data*. Beijing: Petroleum industry press.

[3] Chen, B. F. (2007). *Research and Implementation of Logging Curve Vectorization Method and Rebuilding Technique*, Harbin Institute of Technology.

[4] Ding, W. D. *et al*., (2006). Study on vectorization algorithm of engineering drawing. *Computer Engineering, 24,* 157-158+161.

[5] Zheng, Y., Li, H., & Doermann, D. (2005). A parallel-line detection algorithm based on HMM decoding.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(5), 777-792.
[6] Xavier, H., & Karl, T. (2002). Improving the accuracy of skeleton-based vectorization. *Graphics Recognition Algorithms and Applications, 2390*, 273-288.
[7] Vijay, N., & Noshir A. (1990). Langrana engineering drawing processing and vectorization system. *Computer Vision, Graphics and Image Processing. 49,* 379-397.
[8] Liu, W., & Dori. D. (1996). Sparse pixel tracking: a fast vectorization algorithm applied to engineering drawings, *Proceedings of the International Conference on Pattern Recognition*.
[9] Hori, O., & Tanigawa, S. (1993). Raster-to-vector conversion by line fitting based on contours and skeletons, *Proceedings of the 2nd International Conference on Document Analysis and Recognition,* Tsukuba Science City.
[10] Pavlidis T. (1982). *Algorithms for Graphics and Image Processing*. Heidelberg:Computer Science Press.
[11] Li, C. S. *et al*., (2003). SCTR: An Approach to digitizing well-logging graph, *Proceedings of the 6th IASTED International Conference on Computer Graphics and Imaging*
[12] Er-Long, Y. (2009). PCTR: An improved approach to digitizing well-logging parameter graph. *Proceedings of the 2009 International Conference on Machine Learning and Cybernetics*.
[13] Rao, X. B. *et al*., (2006). An Algorithm for erasing grid-lines and reconstructing strokes in Chinese handwriting based on B lock adjacency graph. *Journal of Image and Graphics*.
[14] Rafael C. Gonzalez. (2017). *Digital Image Processing Third Edition*. Beijing: Electronic Industry Press.
[15] Wilson, M., & Spong, M. W. (2006). Robot modeling and control. *Industrial Robot an International Journal, 17(5),* 709-737.
[16] Wang, Z. Z. (2007). Automated Monitoring of High -speed Metal transfer for real-time control dissertation. Graduate School University of Kentucky.

Bo Yuan was born in Hejiang County, Sichuan Province in May 1995. He is a graduate student of Southwest Petroleum University. He has worked as an assistant management and teaching assistant at the School of Electrical Information, Southwest Petroleum University. Bo Yuan already owns a software copyright. Bo Yuan's research interests are graphic image processing, artificial intelligence, intelligent detection and control.

Qing Yang was born in Luoshan County, Henan Province in May 1980 and received a doctorate degree from China University of Petroleum. She is an associate professor in the School of Electrical Engineering and Information of Southwest Petroleum University and serves as the director of the Automation Teaching and Research Office. Qing Yang has published more than 20 conference / journal papers. Qing Yang's research interests are intelligent detection and control, theories and methods of predictive control, artificial intelligence, and digital image processing.