The International Journal of Engineering and Science (IJES)
|| Volume || 9 || Issue || 11 || Series I || Pages || PP 57-63 || 2020 ||
ISSN (e): 2319-1813 ISSN (p): 20-24-1805

An automatic picking algorithm for the best superposition speed based on edge extraction and peak detection

Zhengzhou[1], Jiang Murong[1], Wang Yunsheng[2], Chen Yuzhuo[1], Yang Zeli[1], Lu Shaofei[1]

[1]*School of Information, Yunnan University, Kunming, 650500, Yunnan*
[2] *Yunnan Aerospace Engineering Geophysical Inspection Co., Ltd., Kunming, 650000, Yunnan*
Corresponding Author: Zhengzhou

--ABSTRACT--

The optimal superimposed velocity value is an important parameter for velocity analysis and migration imaging. How to find the optimal superimposed velocity value from a large amount of signal data is a key issue in interpreting seismic data. In this paper, the signal data is processed by introducing hough transform edge extraction and vector analysis peak detection method, using hough change search to determine the maximum "energy sequence" and peak detection algorithm and cumulative calculation to determine a set of signal values that meet the minimum energy error criterion, and The best superimposing speed can be found through this set of signals. The result of the example shows that the algorithm in this paper realizes the automatic pick-up of the superimposed speed, and meets the accuracy requirements in actual engineering, and can quickly provide a reasonable and effective speed model for subsequent data analysis.

KEYWORDS: *superimposition speed, automatic picking, hough transform, peak detection*

Date of Submission: 24-11-2020 Date of Acceptance: 07-12-2020

I. INTRODUCTION

The propagation velocity of seismic waves in underground rock media is an important parameter for processing and interpreting geophysical data. Using velocity spectrum to pick up the best superimposed velocity value is one of the core issues in processing seismic data [1].

In the analysis of actual seismic wave data, in the face of hundreds of millions of data to be processed, the automatic pick-up algorithm for the optimal superimposed velocity becomes the key to the calculation of velocity spectrum parameters. Automatic picking is often achieved through the combination of objective function and optimization methods. The existing superimposition speed optimization methods use similar peak fitting methods combining linear constraints and nonlinear optimization [2], high-order dynamic correction optimization methods [3], based on Velocity analysis methods for undulating ground [4], non-hyperbolic-based velocity automatic analysis methods [5], etc.; in addition, some probability-based methods are constantly being proposed by scholars: Monte Carlo random disturbance method [6], viterbi algorithm automatic Picking and stacking speed method [7] and so on. The above-mentioned methods are more suitable for channel data with uniformly distributed sequences and high signal-to-noise ratio [8].

Since the superimposed velocity is affected by factors such as lithology and rock layer thickness, the distribution of velocity spectrum energy bands produced by different geological conditions is also different. For some channel data collected under complex terrain conditions, due to noise and terrain factors affecting part of the channel data distortion, only using random probability methods such as Monte Carlo is not enough to obtain a more accurate superimposed velocity value. The method used in this paper is a combination of the contour detection method of hough transform and the peak detection algorithm of vector analysis [9]. It first detects the maximum energy sequence in the energy spectrum, and then calculates the number of energy points corresponding to the hyperbola passing signal data peaks, and the energy The sequence and the peak number sequence are used as two conditions for estimation, and the joint distribution of the two can better simulate the channel distribution sequence in line with a large amount of actual data, which can effectively improve the accuracy of the superimposed speed pickup [10]. The experimental results show that the optimal superposition velocity value calculated by this method is close to the true value, and the accuracy meets the requirements.

II. PRINCIPLE OF METHOD

In velocity signal processing, the average amplitude energy velocity spectrum is often used to indirectly find the best superimposed velocity value, as shown in Figure 2-1. $X = (X_1, X_2, \dots, X_N)$ is expressed as N channel data collected， t_0 is the two-way reflection time with zero shot spacing, Δt is the amount of dynamic correction (normal time difference), H is the best time delay curve for dynamic correction, $\tilde{V}$ is the superimposition velocity, and$\bar{A}$is the average amplitude energy of N channels.

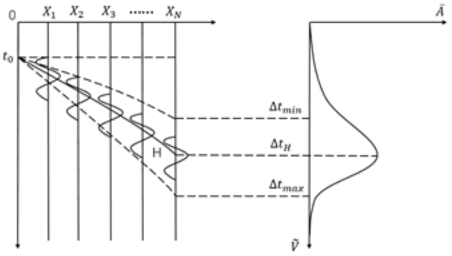


Figure 2-1 Schematic diagram of average amplitude energy of velocity spectrum

2.1 Define the objective function

The average amplitude energy velocity spectrum not only reflects the strength of the reflected wave energy, but also can determine the quality of each relevant channel in the channel data. Therefore, the average amplitude energy velocity spectrum is the basis for the calculation of the superimposed velocity value. According to the seismic wave hyperbolic travel time formula (2-1), the travel time information (including the superimposed velocity value) is calculated for each moment, and the amplitude energy of the same phase on different channels is determined from the travel time information and compared with the total energy to obtain the objective function (2 -2), the minimum energy error criterion [1]:

$$t_{x_i} = t_0 + \Delta t_i = \sqrt{t_0^2 + \frac{S_i^2}{(\tilde{V})^2}} \quad （2\text{-}1）$$

$$E = \arg\min\left[\sum_{j=0}^{M}\sum_{i}^{N} X_{i,j+r_i}^{2} - \frac{1}{N}\sum_{j=0}^{M}\left(\sum_{i}^{N} X_{i,j+r_1}\right)^2\right] \quad （2\text{-}2）$$

In the formula,$X_i (i = 1,2,3, \dots, N)$is N channel data, $j = (0,1,2, \dots, M)$ is the sampling point number in each channel, t_{xi} is the round trip corresponding to X_i For reflection time, δt is the sampling interval, and $r_i = t_{x_i}/\delta t$ is the delay amount.

The velocity spectrum can be regarded as a two-dimensional series. The rows and columns of this two-dimensional series respectively represent the two-way reflection time t_0 and the superimposed velocity $\tilde{V}$ with zero shot spacing. Each point in the series contains an average amplitude energy value. The energy extreme point on the velocity spectrum is an important basis for selecting a reasonable superposition velocity, and the distribution of its energy band conforms to the law of change with time t Therefore, the processing of velocity spectrum and channel data can be transformed into searching for the optimal energy value for a two-dimensional sequence, that is, for a given time t_0, select an appropriate superimposition velocity $\tilde{V}$, and obtain a set of optimal (real time of arrival) The value $r_i^* \Delta t$ corresponds to a set of $X_1^*, X_2^*, \dots, X_N^*$ (the amplitude value of the intersection of the delay curve and each channel), so that the sum of the squares of this set of amplitude values is the same as the recorded amplitude of N channels The difference between the average energy sum after the squared value is the smallest.

2.2 The edge extraction of hough transform and the peak detection method of vector analysis to solve the best speed value

The algorithm first solves the peak detection of known signal data, and calculates its energy value when the peak may pass. The point under the condition that two conditions are met is used as the optimal solution instead of the derivative (gradient-based (Optimization) process optimization, which can better adapt to the actual situation of various complex distributions.

Using this algorithm to solve the problem can be divided into two steps: (1) use hough transform to perform edge detection to extract the maximum energy cluster information; (2) detect the number of peaks of the hyperbola corresponding to the energy cluster point and perform cumulative calculations. Through a certain method, delete the points with small energy values and single irregular points to find the best speed value.

2.2.1Hough transform search to determine the largest "energy cluster sequence"

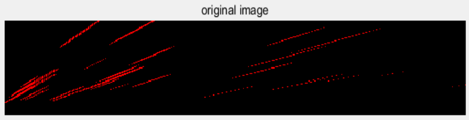
The essence of the problem of detecting straight lines in an image is to find all the pixels that make up a straight line. Then the problem is from finding a straight line to finding all(x,y) points conforming to y=mx+c. A straight line is a collection of a series of discrete points in the image. Through a straight line discrete polar coordinate formula, The discrete point geometric equation of a straight line can be expressed （3-1）,

X *cos(theta) + y * sin(theta) = r （3-1）

The angle thetarefers to the angle between r and the X axis, and r is the geometric vertical distance to the straight line. Any point on a straight line, x, y can be expressed, where r and theta are constants.

In the realm of image processing, the pixel coordinates P(x, y) of the image are known, and r, theta are the variables we are looking for. If we can plot each (r, theta) value according to the pixel coordinate P(x, y) value, then transform from the image Cartesian coordinate system to the polar coordinate Hough space system, this transformation from point to curve It is called the hough transform of straight lines. The transformation is performed by quantizing the Hough parameter space into a finite value interval equally divided or accumulated grid. When the hough transform algorithm starts, each pixel coordinate point P(x, y) is converted to the curve point of (r, theta) and added to the corresponding grid data point. When a white point appears, it means that there is a straight line.

The original energy map can be obtained from the original signal data (as shown in Figure 3-1). After the Canny factor edge detection (as shown in Figure 3-2), after edge detection, the contour of the data point can be determined by the rectangular coordinate system. Then transform it into the polar coordinate system by hough transform (as shown in Figure 3-3), and the white dots in the figure are straight lines in rectangular coordinates. We take the first 25 points with the most intersections in polar coordinates and transform them into 25 straight lines corresponding to the contours of the energy clusters in the original image (as shown in Figure 3-4) to search and determine the largest energy clusters.



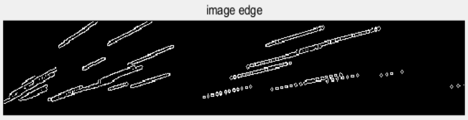
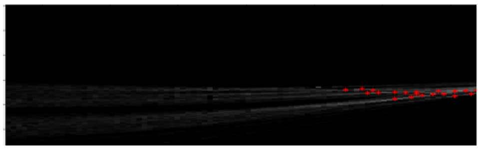


Figure 3-1 Energy DiagramFigure 3-2 Canny edge detection result



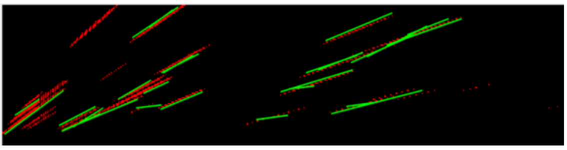


Figure 3-3 Polar coordinate diagramFigure 3-4 Hough transform detection result diagram

2.2.2 Search and determine the largest "wave crest sequence"

The vector-based analysis algorithm is applied to the peak detection principle of ABR waves. The specific steps to solve the peak sequence are as follows:

1. Suppose the sampling time is x, the sampling amplitude is y, that is, the sampling point is expressed as (x, y). As far as the curve of the collected signal wave is concerned (as shown in Figure 3-5), we take the window area and set it as A , The values are $y_1, y_2, \dots, y_m$, the sampling time is $x_1, x_2, \dots, x_m$, a total of m points are taken, m is an odd number, and the coordinates of each point are B(x_n, y_n), where n=1,2…m, then the midpoint coordinate in the window area is C($x_{\frac{m+1}{2}}, y_{\frac{m+1}{2}}$), the vector between other points and the midpoint of the window area can be expressed as $\bar{a}$ =(x_n, y_n)- ($x_{\frac{m+1}{2}}, y_{\frac{m+1}{2}}$), the amplitude difference between the two points is $\Delta y = \left| y_n -, y_{\frac{m+1}{2}} \right|$ now take the vector $(0, -\Delta y)$, denoted as $\vec{a}$,

Then calculate the cosine of the angle between the vector $\vec{a}$ and the angle $\vec{b}$ as:

$$\cos\theta = \frac{\vec{a}\vec{b}}{|\vec{a}||\vec{b}|} = \frac{\Delta y}{\sqrt{\Delta x^2+\Delta y^2}},\ （3\text{-}2）$$

Calculate the cosine $\cos\theta$ of the angle between vector$\vec{a}$ and vector $\vec{b}$ between each point and the midpoint in the window area, and accumulate the $\cos\theta$ value in window A and calculate SUM=$\sum\cos\theta$;

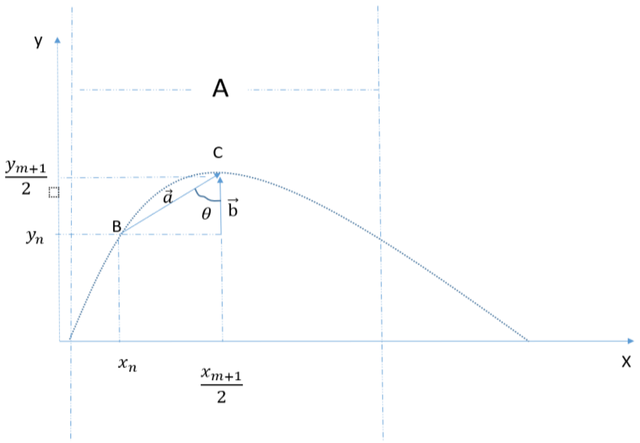


Figure 3-5 Vector analysis diagram

2. Suppose there are N data points in the data, the window size is m points, the window is moved with the step length of 1 point, and the window size remains unchanged, the number of windows obtained is N-m+1, denoted as $A_1, A_2, \ldots, A_{N-m+1}$;accumulate $\cos\theta$ in each window in turn, and calculate the corresponding $SUM_1, SUM_2, \ldots, SUM_{N-m+1}$, And record the coordinate value of the midpoint of each window; from the characteristics of the cosine function, when $SUM_n > SUM_{n-1}$, and $SUM_n > SUM_{n+1}$, Then there is a maximum value in the corresponding window A_n, and the maximum value is the midpoint of the window A_n;

3. Sort the cosine value and SUM of each window from largest to smallest, and the window with larger SUM in the front contains all extreme points are selected; among these points, the peaks of waves I, II, III, IV, and V are selected. The following are the peaks detected in one of the waveforms of the data (as shown in Figure 3-6).

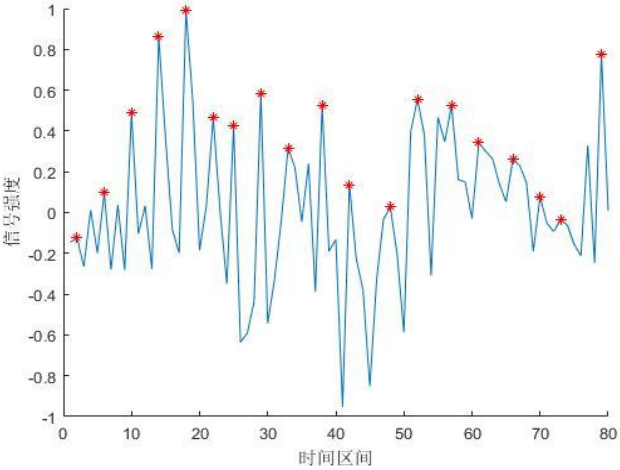


Figure 3-6 Wave peak detection diagram

2.3 Find the best speed value

First, the energy cluster obtained by hough transform is known, the sequence of energy points isK_n (A_n, H_n, V_n), n=1,2,..., Among them, A_n is the energy value,H_n is the depth, and V_n is the superimposition speed. Then, the above-mentioned wave peak detection algorithm records the number m of points where the corresponding dynamic correction curve passes the wave peak, and arranges them from large to small according to m to form a new array K_n $(m_n, A_n, H_n, \mathrm{V}_n)$. Without considering the interference of the external environment such as noise, the more crests passed, the better, so here we have taken the first 40 points with more crests, that is, the new array $K_n{}'(m_n, A_n, H_n, \mathrm{V}_n)$. That is to mark these points on the energy map (as shown in Figure 3-7),

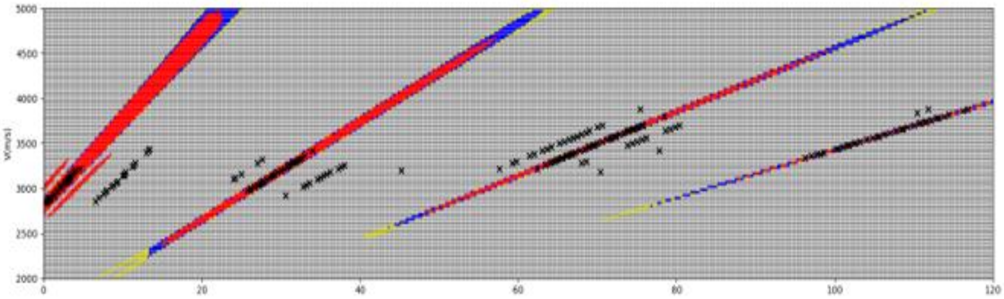


Figure 3-7

For the new array $K_n{}'(m_n, A_n, H_n, \mathrm{V}_n)$
1.Determine the size ofA_n, and delete it from the array when A_n<0.13.
From the above, the greater the average amplitude energy of the objective function, the closer the value to the theoretical correct value is. Therefore, the threshold is set and the point with the smaller energy is directly eliminated.
2. Due to external interference such as noise interference, some energy values obtained are single and irregular within a certain range.
Calculate the value of $H_n/10$ and group books with equal quotients into a group. If the number in the group is less than or equal to 2, delete it from the array; if it is greater than 2, find the average of H and V.
3.The point in the new array $K_n{}'\ '$ $(m_n, A_n, H_n, \mathrm{V}_n)$ obtained at the end is the actual value. Marked on the energy spectrum (as shown in Figure 3-8),

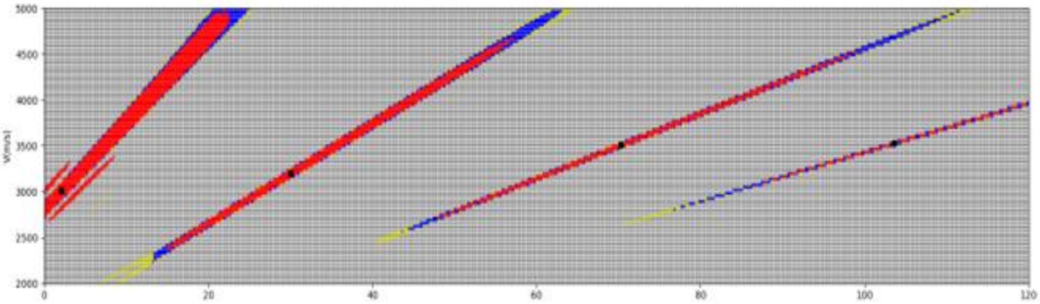
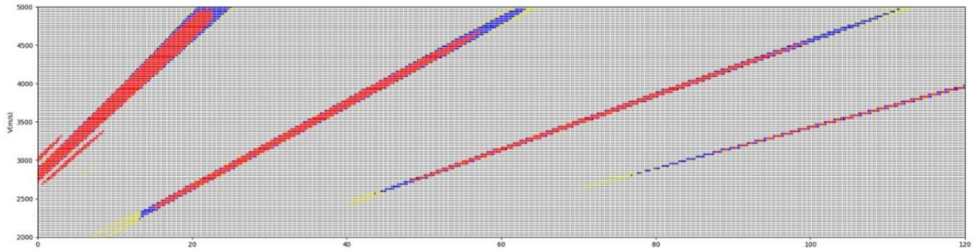


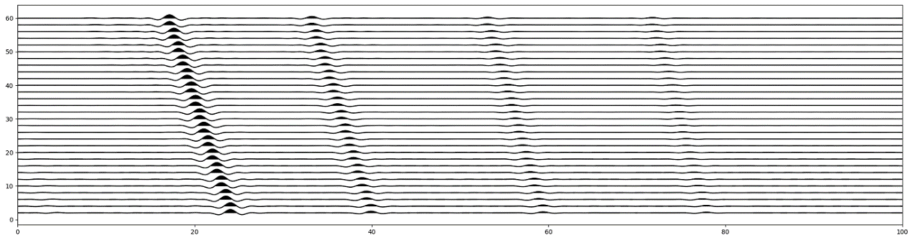
Figure 3-8

III. INSTANCE VERIFICATION

This paper uses the tunnel geological advanced prediction problem in the field of geophysical prospecting to verify the effectiveness of the proposed algorithm. The tunnel geological advance forecast is the forecast of the surrounding rock and stratum in front of the tunnel face and its surroundings. The important link is the simultaneous selection of speed and tunnel depth coordinates in the speed scan coordinate system, that is, selecting appropriate points and curves in the energy diagram and channel diagram shown in Figure 4-1.



(a)



(b)

Figure 4-1 Speed scan energy diagram (a) and channel signal diagram (b)

Actual model: channel data (including large noise and interference) during tunnel construction in Caoba area, the tunnel face is located 35m in front of the geophone, the length of the tunnel to be excavated is 120m, the width is 12m, and the height is 7m. The calculation results and errors of this algorithm are as follows As shown in Table 4-1, the velocity and interface errors in the first three layers are very small and close to the theoretical values. The fourth layer is greatly affected by noise and interference signals due to the long spatial distance from the detector, resulting in an increase in errors; The superimposed velocity value obtained by the algorithm in this paper is located inside the optimal energy band, and the corresponding dynamic correction curve passes through the peak or trough of each channel (both points with larger amplitude energy), and the optimal superimposed velocity value is better picked up (As shown in Figure 4-2).

Table 4-1 Comparison of measured data in Caoba area and calculated values in this paper

Number of layers	Actual measurement speed（m/s）	Actual measured depth（m）	The speed of my algorithm（m/s）	The depth of my algorithm（m）	Speed inaccuracy error（%）	Depth inaccuracy error（%）
1	2475	42.0	2500	42.4	1.01	0.95
2	2600	62.0	2600	62.2	0	0.32
3	2900	102.0	2900	101.8	0	0.20
4	2675	123.0	2600	116.8	2.80	5.04

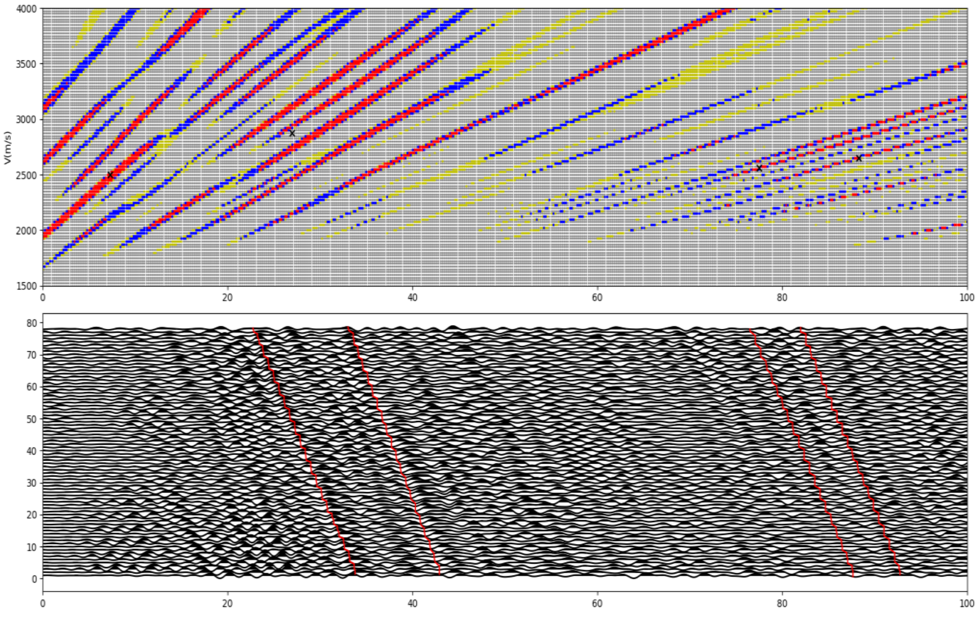


Figure 4-2 Actual excavation speed spectrum and energy spectrum in a certain area

IV. CONCLUSION

The graph databases and relational database both performed well. In general, graph databases performed better when objective tests were performed. The Implementation shows that graph databases retrieve the results of the set of predefines query faster than relational databases.Alsograph databases are more flexible than relational databases as we can add new relationships to graph databases without the need to restructure the schema again. This paper proposes an automatic pick-up algorithm for optimal superposition speed based on hough transform and vector analysis wave peak detection, and obtains good results in the theory of tunnel advance prediction and the measured model. In this paper, the hough transform is used to detect the contour of the energy cluster and the vector analysis wave peak detection algorithm to determine the optimization direction and path, which achieves the purpose of automatically picking up the best superimposed velocity value more accurately, which is useful for reducing possible economic losses (due to high computational cost In the speed analysis, only part of the speed information may be picked up, resulting in the economic loss caused by the loss of effective geological information), reducing exploration costs and improving production efficiency, which has positive significance.

However, because this method is more dependent on the initial model (the quality of the energy cluster distribution directly affects the picking accuracy), and is limited by the speed versatility of the energy map and the low signal-to-noise ratio of the actual data, it may be better if other related algorithms are applied Related research needs to be discussed in depth.

REFERENCE

[1]. Mu Yongguang et al. Seismic data processing method[M]. Beijing: Petroleum Industry Press, 2007: 130

[2]. Zhang L. Automatic picking and its application[J]. SEG,1991,70:275-292

[3]. You Jianjun, Chen Qiang, Zhang Chaomo, et al. Motion correction of seismic event axis with long offset in isotropic media[J]. Progress in Geophysics, 2008(04):209-214.

[4]. Pan Hongxun, Fang Wubao. Superimposed velocity analysis based on undulating surface[J]. Petroleum Geophysical Prospecting,2008, 043(001):29-33.

[5]. Sun Xiange, Zhong Benshan. Research on non-hyperbolic velocity scanning method[C]// Proceedings of the 22nd Annual Meeting of the Chinese Geophysical Society. 2006.

[6]. David E L. Monte Carlo automatic velocity picks[J]. SEG Mtg,75:1-20

[7]. Lin Niantian, Liu Hong, Li Guihua, et al. Path integral optimization method and smoothing processing technology for automatic speed pickup[J]. Chinese Journal of Geophysics, 2013, 056(001):246-254.

[8]. Pan Hongxun,FangWubao.Summary of seismic velocity analysis methods[J].Progress in Exploration Geophysics,2006,29(005):305-311.

[9]. Sheng Zou, Ke Shen. Image segmentation method of belt conveyor idler based on Hough transform. Energy and Environmental Protection, 2020, 42(10),

[10]. Gao Ruijing, Sun Ying. A new algorithm for ABR signal peak detection[J]. Chinese Journal of Geophysics, 2011, 053(011):2603-2612.

Zhengzhou, et. al. "An automatic picking algorithm for the best superposition speed based on edge extraction and peak detection." *The International Journal of Engineering and Science (IJES)*, 9(11), (2020): pp. 57-63.