

Research Article

Research on Feature Extraction of Indicator Card Data for Sucker-Rod Pump Working Condition Diagnosis

Yunhua Yu, Haitao Shi, and Lifei Mi

College of Information & Control Engineering, China University of Petroleum (Huadong), Qingdao 266580, China

Correspondence should be addressed to Yunhua Yu; upcyh1040@gmail.com

Received 8 June 2013; Revised 23 October 2013; Accepted 13 November 2013

Academic Editor: Jason Gu

Copyright © 2013 Yunhua Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Three feature extraction methods of sucker-rod pump indicator card data have been studied, simulated, and compared in this paper, which are based on Fourier Descriptors (FD), Geometric Moment Vector (GMV), and Gray Level Matrix Statistics (GLMX), respectively. Numerical experiments show that the Fourier Descriptors algorithm requires less running time and less memory space with possible loss of information due to nonoptimal numbers of Fourier Descriptors, the Geometric Moment Vector algorithm is more time-consuming and requires more memory space, while the Gray Level Matrix Statistics algorithm provides low-dimension feature vectors with more time consumption and more memory space. Furthermore, the characteristic of rotational invariance, both in the Fourier Descriptors algorithm and the Geometric Moment Vector algorithm, may result in improper pattern recognition of indicator card data when used for sucker-rod pump working condition diagnosis.

1. Introduction

The sucker-rod pump system is the most widely used form of artificial lift for the onshore oil well production [1–3]. Approximately, 80% of the oil wells in the world, 90% of those in China, are being produced by the sucker-rod pumps [4, 5]. The maintenance and optimization of a sucker-rod pump system is a costly and time-consuming operation. The indicator card is the relation curve between the load and the displacement of a sucker-rod pump in an intact suck cycle, in which x -axis represents displacement and y -axis represent load [6]. The indicator card is helpful to analyze the down-hole working condition of the sucker-rod pump wells [7], which can judge the operation condition of the sucker-rod pump well and provide reliable proof of high efficiency, reasonable exploitation for the oil well production. While the system is operating, the card can indicate such shape that might be a normal operation or a fault situation. According to different kinds of real-time indicator card data, the pattern recognition and fault diagnosis techniques are used to identify some different curve shapes, locate which kind of abnormal situation is, and interpret why the fault occurs [8]. Therefore, the correct and quick identification of

the sucker-rod pump indicator card is essential to the fault diagnosis of down-hole working condition. The automatic fault diagnosis of sucker-rod pump working condition is a visual interpretation process [9]. Nowadays, the traditional methods of interpretation are not suitable for the automatic fault diagnosis of the down-hole conditions. And several signal processing methods, such as artificial neural network (ANN) [10] and fuzzy support vector machine (FSVM) [11], have been studied and applied to pattern recognition of indicator cards to improve the accuracy and efficiency of sucker-rod pump system fault diagnosis. Because there are more fault patterns and less fault samples in the sucker-rod pump working conditions, the above approaches have their limits, respectively. In recent years, a method called biomimetic pattern recognition (BPR) [12] has been proposed, which is based on the principle of homology continuity and is suitable to recognize and classify those objects with more pattern types and less samples.

This paper is focused to study, compare, and select the proper feature extraction methods used to analyze the sucker-rod pump indicator card data before pattern classification. The optimal feature extraction algorithm of sucker-rod indicator card data followed by the matching pattern recognition

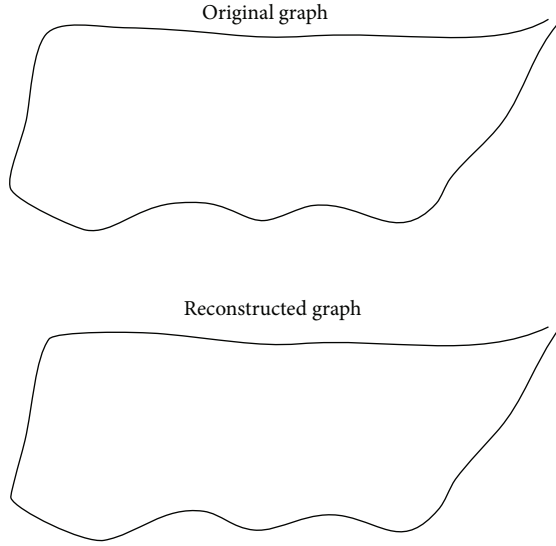


FIGURE 1: Result of indicator card feature extraction.

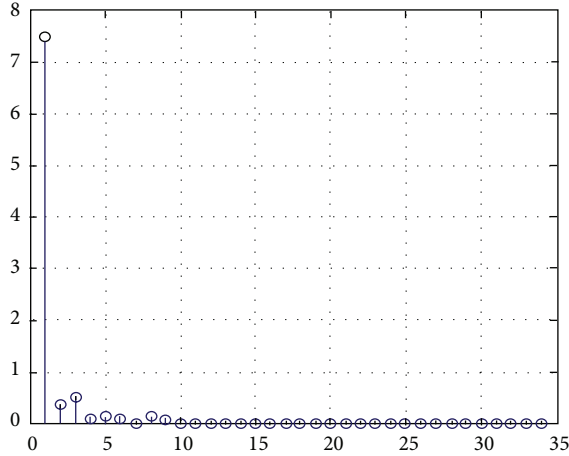


FIGURE 2: Amplitude spectrum of indicator card.

is helpful to locate the exact fault type of working sucker-rod pump, which is of great significance in improving crude oil production as well as preventing some possible safety accidents. The different feature extraction algorithms of indicator card are based on Fourier Descriptors, Geometric Moment Vector, Gray Level Matrix Statistics, Area and Difference Curve, respectively [13].

The area-based method can only determine some limited fault types of narrow-band distribution indicator card, such as pump spraying fault [14]. The difference curve-based method could not determine some greatly dangerous pump failures, such as stuck pump fault [15]. In this paper, we focus on three feature extraction algorithms of the indicator card, which are based on Fourier Descriptors (FD), Geometric Moment Vector (GMV), and Gray Level Matrix Statistics (GLMS), respectively. With numerical experimental simulation and analysis, three different algorithms are compared in terms of the consumption of time and the complexity of memory space.

2. Feature Extraction Methods of Indicator Card Data

2.1. Method I: Algorithm Based on Fourier Descriptors. Fourier transformation is a popular method for reconstruction and classification of image. It generates a complete set of complex numbers-the Fourier Descriptors (FD), which represent the object shape in a frequency domain [16].

To reduce the computational complexity, the polygonal approximation method is used to cross out those redundant indicator card points (data). The procedures are as follows.

Firstly, according to a given value, which is called D , we traverse all the digital pixels (data) of the indicator card curve in order to choose the feature pixels (data) of the polygon, which meet the condition with maximum curvature of a certain length curve.

Secondly, we store those feature data to an array.

The given value (D) is assigned as 0.008 according to comparison of different computation procedures. Take the pump-on-touch fault as an example; 34 feature pixels (data) are extracted from the original 702 pixels. Figure 1 shows the reconstructed graphic curve of the pump-on-touch fault compared with the original one.

It is reasonable to employ Discrete Fourier Transforms (DFT) to obtain FD [17]. However, some sample errors may be produced during the sampling process.

After the feature extraction of polygonal approximation, the Fourier Transform (FT) of each polyline is employed to avoid the possible error caused by sampling and to improve the speed and accuracy of calculation.

The Fourier Descriptors, denoted by $Z(k)$, are generated by such math formula as follows [18]:

$$p(l) = x(l) + jy(l) = \sum_{k=-\infty}^{+\infty} z(k) e^{j(2\pi kl/L)},$$

$$z(k) = \frac{1}{L} \int_0^L p(l) e^{-j(2\pi kl/L)} dl \quad (1)$$

$$= \frac{1}{L} \sum_{n=0}^{N-1} \int_{l_n}^{l_{n+1}} p(l) e^{-j(2\pi kl/L)} dl,$$

where L denotes the polygons perimeter. (x_n, y_n) denotes the coordinates of vertex \mathbf{P}_n . l_n denotes the accumulated length sum of those short polylines between \mathbf{P}_0 and \mathbf{P}_n .

The amplitude spectrum of the indicator card obtained after Discrete Fourier Transforms (DFT) is shown in Figure 2. The lower frequency elements of Fourier Descriptors contain the most important information of the indicator card while the higher frequency ones contain less information. Therefore, a subset of the Fourier Descriptors can be used to discriminate different shapes of curves. In this paper, k is set as 10, which means the first 10 FDs are used.

The Fourier Descriptors are affected by the location of vertex \mathbf{P}_n , the scale and direction of curve. To eliminate these effects, the normalization is employed. If the given curve object is firstly magnified by r times, secondly its starting position is shifted by a , and then it is rotated by φ degree, finally it is translated by the displacement (x_0, y_0) . We get a

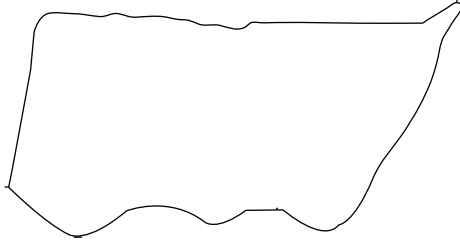


FIGURE 3: Image Binarization and refinement.

new curve object, and the new Fourier Transform coefficient of the new curve is

$$\begin{aligned}
 z'(k) &= F \left[(x' + iy') r e^{j\varphi} + (x_0 + iy_0) \right] \\
 &= r e^{j\varphi} F(x' + iy') + F(x_0 + iy_0) \\
 &\Rightarrow \frac{r \left\| e^{j\varphi} e^{j(2\pi/L)ka} z(k) \right\|}{r \left\| e^{j\varphi} e^{j(2\pi/L)ka} z(1) \right\|} = \frac{\|z(k)\|}{\|z(1)\|}.
 \end{aligned} \quad (2)$$

The effect of modulus and phase changes on Fourier Descriptors is eliminated due to the identically equal ratio as shown in the formula (2).

2.2. Method II: Algorithm Based on Geometric Moment Vector. In image processing field, Geometric Moment Vector (GMV) can be used as an important feature to represent objects due to its translation invariance, rotation invariance, and scale invariance.

Since the scanned image of indicator card is grayscale and the image outline is not smooth, it is necessary to preprocess the image by using binarization and refinement technology. Figure 3 is a single-pixel binary image after the binarization and refinement processing.

The seven two-dimensional moment invariants are as follows, any of which is not sensitive for translation, scaling, mirroring, and rotation [18]:

$$\begin{aligned}
 f_1 &= \eta_{20} + \eta_{02}, \\
 f_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\
 f_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \\
 f_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \\
 f_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\
 &\quad \times [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
 &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\
 &\quad \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2],
 \end{aligned}$$

$$\begin{aligned}
 f_6 &= (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
 &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \\
 f_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \\
 &\quad \times [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
 &\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \\
 &\quad \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].
 \end{aligned} \quad (3)$$

2.3. Method III: Algorithm Based on Gray Level Matrix Statistics. The grid method is one of the traditional image feature extraction methods. The processing steps of the grid method are as follows.

Firstly, we divide the image of indicator card into a number of small grids with same size and shape, in the horizontal and vertical direction, respectively. Then, we mark the grids which are traversed by the curve of indicator card. Finally, we can obtain the feature parameters of the indicator card image.

In this paper, we use a Gray Level Matrix Statistics (GLMS) feature extraction method which is based on grid method.

Before the GLMS feature extraction, the indicator card curve should be converted to grayscale graphic matrix. The steps of GLMS feature extraction algorithm are as follows.

(1) The mesh of grayscale matrix is initialized: if a grid is traversed by the indicator card curve, then the gray value is assigned as "1".

(2) According to the gray contour principle, other grids are assigned as such gray value: if the grid is located at the inside of the curve, then the gray value is equal to the initial value plus n for its n grids distance away from the curve; if the grid is located at the outer region of the curve, then the gray value is equal to the initial value minus n for its n grids distance away from the curve.

(3) Finally, we can get 6 statistic parameters of the gray level matrix [19].

3. Numerical Experiments and Results

Three typical fault indicator cards are shown in Figures 4, 5, and 6.

According to the Method I (Fourier Descriptors), we compute 12 normalized subsets of the Fourier Descriptors (FDs) for 12 different typical fault of indicator cards. Each set includes 10 main FDs, which is called a feature vector denoted as $F = [FD_1, FD_2, FD_3, FD_4, FD_5, FD_6, FD_7, FD_8, FD_9, FD_{10}]$. Table 1 shows the numerical result.

According to the Method II (Geometric Moment Vector), we compute 12 subsets of the Geometric Moment Vectors (GMV) for 12 different typical fault of indicator cards. Each set includes 7 GMVs, which is called a feature vector denoted as $F = [f_1, f_2, f_3, f_4, f_5, f_6, f_7]$. Table 2 shows the numerical result.

TABLE 1: Normalized Fourier Descriptors of 12 typical fault indicator cards.

	FD ₁	FD ₂	FD ₃	FD ₄	FD ₅	FD ₆	FD ₇	FD ₈	FD ₉	FD ₁₀
Top pump bumping	1	0.04825	0.06673	0.00921	0.01938	0.01228	0.01789	0.00373	0.00615	0.00368
Bottom pump bumping	1	0.04190	0.06150	0.01079	0.02105	0.00580	0.02075	0.00018	0.00525	0.00382
Broken-dropped of sucker rod	1	0.00256	0.07137	0.04122	0.00181	0.00423	0.01041	0.00278	0.00612	0.00312
Insufficient fluid supply	1	0.09135	0.07921	0.03970	0.02756	0.00268	0.00274	0.02133	0.00656	0.00519
Leakage of traveling valve	1	0.02729	0.08138	0.01280	0.02071	0.02071	0.01073	0.00803	0.00936	0.00458
Leakage of standing valve	1	0.04080	0.08172	0.01214	0.02129	0.01990	0.00740	0.00619	0.00890	0.00371
Leakage of both valves	1	0.00256	0.08158	0.01256	0.02517	0.01928	0.00683	0.00725	0.00912	0.00369
Stuck piston	1	0.00823	0.07259	0.01193	0.00158	0.00925	0.00957	0.00458	0.00349	0.00344
Gas lock	1	0.08365	0.06759	0.02358	0.02135	0.00245	0.00256	0.02325	0.00589	0.00623
Liquid hammer	1	0.09195	0.06064	0.02082	0.02073	0.00283	0.00255	0.04195	0.00764	0.00082
Sand inflow	1	0.00925	0.07436	0.00564	0.02577	0.00556	0.01728	0.00058	0.00278	0.00166
Plunger out pump barrel	1	0.04255	0.08212	0.01223	0.02019	0.01625	0.00722	0.00539	0.00693	0.00316

TABLE 2: Geometric Moment Vectors of typical faults indicator card.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
Top pump bumping	0.4392	0.0658	0.0175	0.0977	0.0038	0.0250	0.0015
Bottom pump bumping	0.4936	0.0928	0.0302	0.1365	0.0084	0.0415	0.0027
Broken-dropped of sucker rod	0.5877	0.1857	0.1261	0.2652	0.0480	0.1142	0.0067
Insufficient fluid supply	0.7061	0.0615	0.1619	0.3834	0.0095	0.0939	0.0951
Leakage of traveling valve	0.4278	0.0646	0.0117	0.0920	0.0030	0.0234	0.0005
Leakage of standing valve	0.4549	0.0827	0.0371	0.1151	0.0073	0.0331	0.0018
Leakage of both valves	0.4997	0.1340	0.0671	0.1636	0.0171	0.0599	0.0016
Stuck piston	3.9352	5.6349	8.3123	49.690	1003.6	11792	111.89
Gas lock	0.9168	0.1321	0.6305	0.7222	0.1330	0.2230	0.4689
Liquid hammer	0.4799	0.0552	0.0031	0.1165	0.0019	0.0273	0.0011
Sand inflow	0.4677	0.0792	0.0314	0.1161	0.0064	0.0326	0.0028
Plunger out pump barrel	0.4730	0.1149	0.0732	0.1509	0.0156	0.0511	0.0028

TABLE 3: Gray Level Matrix Statistics of typical faults indicator card.

	f_1	f_2	f_3	f_4	f_5	f_6
Top pump bumping	2.7046	29.6603	-0.3710	2.9569	0.0589	0.9699
Bottom pump bumping	2.0381	29.292	-0.14445	2.4059	0.0568	0.9796
Broken-dropped of sucker rod	0.0078	20.3339	-0.1684	2.4799	0.0710	0.9634
Insufficient fluid supply	-4.1172	59.7812	-0.7084	2.6101	0.0483	0.9753
Leakage of traveling valve	2.4658	40.9686	-0.5122	3.1043	0.0505	0.9742
Leakage of standing valve	3.8604	23.7471	-0.0216	2.2812	0.0618	0.9684
Leakage of both valves	0.3887	35.5648	0.0830	2.4249	0.0518	0.9735
Stuck piston	-9.0381	67.0591	-0.6385	2.2935	0.0454	0.9768
Gas lock	-5.2549	60.6294	-0.6209	2.2908	0.0473	0.9759
Liquid hammer	0.3276	43.7525	-0.6760	3.4674	0.0530	0.9729
Sand inflow	2.6519	25.0238	-0.1072	2.3713	0.0620	0.9682
Plunger out pump barrel	1.3003	37.3400	-0.1966	2.6507	0.0521	0.9734

TABLE 4: Comparison of three types of algorithms.

	Algorithm based on FDs	Algorithm based on GMV	Algorithm based on GLMS
Running time (second)	0.698414	5.707594	3.292005
Memory space units for each sample	10	7	6

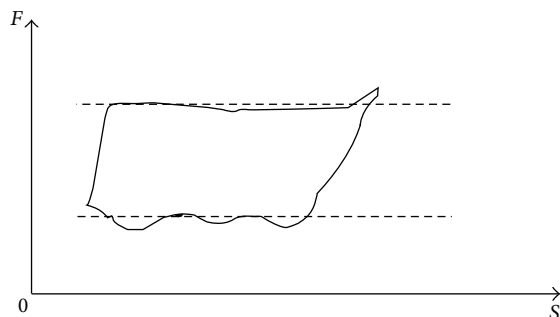


FIGURE 4: Top Pump bumping.

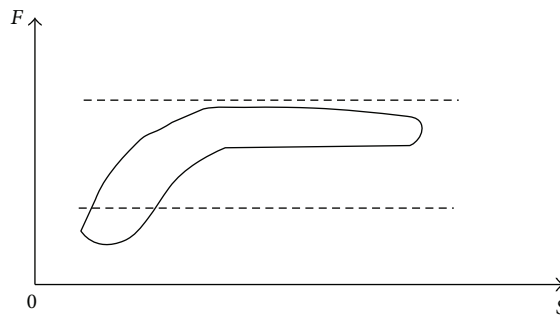


FIGURE 6: Gas lock.

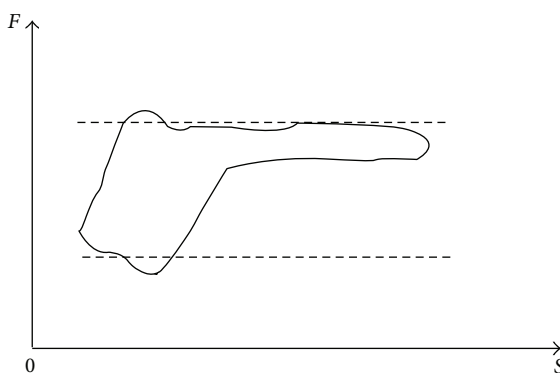


FIGURE 5: Insufficient fluid supply.

According to the Method III (Gray Level Matrix Statistics), we compute 12 subsets of the Gray Level Matrix Statistic (GLMS) for 12 different typical fault of indicator cards. Each set includes 6 GLMSs, which is called a feature vector denoted as $F = [f_1, f_2, f_3, f_4, f_5, f_6]$. Table 3 shows the numerical result.

4. Conclusion

In this paper, three different feature extraction methods, which are based on Fourier Descriptors, Geometric Moment Vector, and Gray Level Matrix Statistic, respectively, have been analyzed and simulated. The computing speed and memory consuming of these 3 algorithms are compared as shown in Table 4.

Numerical experiments show that the FD algorithm is with high computing speed and more memory space but possible loss of information; because of different numbers of FDs, the GMV algorithm is more time-consuming and less memory consuming, while the GLMS algorithm provides low-dimension vectors with good performance of speed and space.

The characteristic of rotational invariance, both in the FD algorithm and the GMV algorithm, may cause improper pattern recognition of indicator card data when used for sucker-rod pump working condition diagnosis. Further research on feature extraction of indicator card data should continue for better performance.

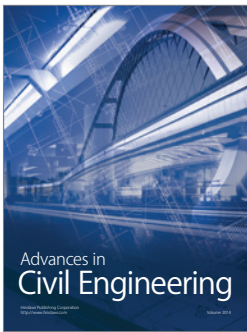
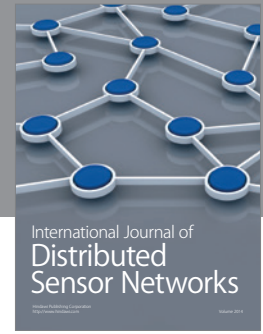
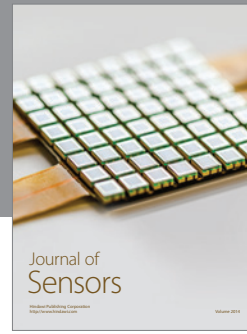
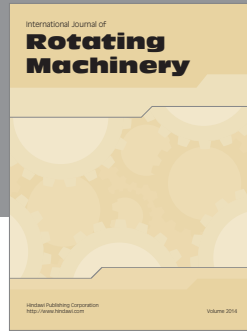
Acknowledgment

This work is fully supported by Scientific Research Foundation for the Excellent Young Scientist of Shandong Province (no. BX2010DX036).

References

- [1] G. P. Schirmer, J. C. P. Toutain, and J. C. Gay, "Use of advanced pattern recognition and knowledge-based system in analyzing dynamometer cards," *SPE Computer Application*, vol. 3, no. 6, pp. 21–24, 1991.
- [2] L. Alegre and C. Morooka, "Intelligent diagnosis of rod pumping problems," in *Proceedings of the 68th Annual Technical Conference*, vol. 3, pp. 97–108, SPE, 1993.
- [3] G. Han and X. Wu, "Application of dynamometer card identification in diagnosis of working-condition for suck rod pump," *Oil Drilling & Production Technology*, vol. 25, no. 5, pp. 70–74, 2003.
- [4] H. Tripp, "A review: analyzing beam-pumped wells," *Journal of Petroleum Technology*, vol. 41, no. 5, pp. 457–458, 1989.
- [5] G. Nazi and J. Lea, "Application of artificial neural network to pump card diagnosis," *SPE Computer Application*, vol. 6, no. 6, pp. 9–14, 1994.
- [6] M. J. Gao, "Portable intelligent instrument for pumping data collection," *Journal of Transducer Technology*, vol. 24, no. 9, pp. 58–62, 2005.
- [7] F. Barreto and M. Tygel, "Automatic down-hole card generation and classification," in *Proceedings of the SPE Annual Technical Conference*, pp. 311–318, 1996.
- [8] Z. Pan and J. Ge, "An adaptive neural network for identification of dynamometer card," *Acta Petrolei Sinica*, vol. 17, no. 3, pp. 104–109, 1996.
- [9] R. Dickinson and J. Jennings, "Use of pattern-recognition techniques in analyzing down-hole dynamometer cards," *SPE Production Engineering*, vol. 5, no. 2, pp. 187–192, 1990.
- [10] D. Baodong and L. Mingshan, "Analyzing dynamometer card of well pumping by using artificial nerve network method," *Well Testing*, vol. 7, no. 1, pp. 27–29, 1998.
- [11] W. Shi and J. Dai, "FSVM based recognition approach for down-hole pump dynamometer cards," *Oil Field Equipment*, vol. 33, no. 4, pp. 46–48, 2004.
- [12] W. Shoujue, "Bionic (topological) pattern recognition—a new model of pattern recognition theory and its applications," *Acta Electronica Sinica*, vol. 30, no. 10, pp. 1417–1420, 2002.
- [13] S. C. Yuan, "Simple distinguish way for dynamometer card," *Well Testing*, vol. 14, no. 1, pp. 49–50, 2005.

- [14] B. D. Du and M. S. Li, "Analyzing dynamometer card of well pumping by using ANN method," *Well Testing*, vol. 1, no. 7, pp. 27–29, 1998.
- [15] S. H. Yang, "Applications of difference curve methods in diagnosis of indicator diagram," *Petroleum Drilling Techniques*, vol. 30, no. 2, pp. 63–64, 2002.
- [16] A. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1989.
- [17] H. Kauppien and T. Seppanen, "An experiment comparison of auto regressive and Fourier-based descriptors in 2D shape classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, no. 2, pp. 201–207, 1995.
- [18] R. C. Gonzalez, *Digital Image Processing*, Electronic Industrial Press, 2005.
- [19] W. Wu and G. D. Chen, "Fault diagnosis system for pump work indicating diagram based on neural network and gray-level matrix," *Journal of Xi'an Shiyou University*, vol. 22, no. 3, pp. 119–121, 2007.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

