We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,800
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Comparison of Border Descriptors and Pattern Recognition Techniques Applied to Detection and Diagnose of Faults on Sucker-Rod Pumping System

Fábio Soares de Lima, Luiz Affonso Guedes and Diego R. Silva
Universidade Federal do Rio Grande do Norte - UFRN
Brazil

1. Introduction

Due to high competition and the need to meet deadlines, modern industries with focus on market demand high availability and reliability of their equipment. With this view, in the last years the maintenance activity has undergone several changes which have led to an evolution in the standpoint of organization and planning of its execution. According to Kardec & Nascif (1998), the direct causes for this development are:

• The quick increase of the amount and the diversity of physical elements that compose varied equipment of process plants that must be kept available;
• Most complex engineering projects;
• New methods for maintenance activity;
• New approaches about the maintenance organization and their responsibilities.

The concept of predictive maintenance has emerged as a result of these demands. Predictive maintenance is the regular monitoring of operating condition (variables and parameters) - the performance - from a device or process that will provide the necessary data to ensure the maximum allowed interval between repair and better intervention planning.

Historically, the first method of artificial elevation that was used in the oil industry was the sucker-rod pumping. Its importance is reflected in the number of installations found on industry, being the most widely elevation method used around the world. Its popularity is related to low cost of investments and maintenance, flexibility of flow and depth, good energy efficient and the possibility of operating with fluids of different compositions and viscosity in a wide temperature range.

The main advantages of Sucker-rod Pumping are: the simplicity of operation, maintenance and design of new installations. From normal conditions can be used until the end of productive life of a well and the pumping capacity can be modified, depending on the changes of the well’s behavior. However, the main advantage of this method relates to the lower cost/production throughout the productive life of the well.

This method is the most common way of artificial elevation (Alegre, Morooka & da Rocha, 1993; Schirmer & Toutain, 1991). It is estimated that 90% of artificial elevation in the world use the system of mechanical pumping (Nazi & Lea, 1994; Tripp, 1989). In Brazil, 64% of the
total production is obtained through the mechanical pumping (de Oliveira Costa, 1994). In practice, the monitoring of the mechanical pumping system status is done by reading a card, called dynamometric card. With this card you can know about the condition of the pump located at the bottom of the well. The dynamometric card consists of a graph that relates the charge and the position. This graph reflects the current conditions of pumping (Barreto et al., 1996; Rogers et al., 1990). Thus, the card may take various formats during well’s production and represent situations of normal operation or indicate a possible irregularity in mechanical pumping system.

The process of identifying situations of abnormal operation of the mechanical pumping system becomes, thus, a problem of visual information interpretation (Dickinson & Jennings, 1990). In any case, this approach may be influenced by several factors, such as the behavior of the system itself, because of its complexity, which results in diverse forms of dynamometric cards, beyond the knowledge and experience of the engineer responsible for the well. Besides that, nowadays, each field petroleum engineer is responsible for over one hundred wells equipped with mechanical pumping. In this case, the traditional process of interpretation becomes impracticable in an acceptable time.

1.1 Objectives

This study aims to contribute to the predictive maintenance field through the development of intelligent computing techniques (Russell, 2003) based on digital image processing, capable of preventing damage in a particular equipment or industrial process in a predictive way.

In scientific terms, the main objective is to propose and analyze the performance of nonparametric patterns recognition techniques in the context of fault detection and diagnosis using boundary descriptors and metrics or statistics mathematical tools.

In technological terms, the objective of this study is to contribute to the area of fault automatic detection and diagnosis in dynamical systems, by proposing a new architecture based on visual similarity of signatures (images) that represent operating conditions. This, in turn, will bring benefits that may complement the tools that nowadays operate in industrial parks.

The proposed approach is the automation of fault analysis and diagnosis in dynamic systems, basing on the following points:

- A description model based on knowledge through system signatures, and;
- Fault recognition through metric distances or correlations.

In this study, it was used bottom hole dynamometer cards, because the surface cards incorporate various degenerative effects caused by the spread of the charge along the whole column of rods. These effects make surface cards represent only one well. When the bottom cards are used, it is possible to observe that the standards of the system operation are the same (de Almeida Barreto Filho, 1993).

Additionally, was chosen present a new approach of classification of the failures of sucker-rod pumping. As the fault diagnosis in mechanical pumping system is a recognition process of dynamometric cards references, various studies have been developed based on pattern recognition techniques using neural network or expert systems and have been proposed to improve accuracy and efficiency (Nazi & Lea, 1994) of this kind of diagnosis system (Alegre, A & Morooka, 1993; Alegre, Morooka & da Rocha, 1993; Barreto et al., 1996; Chacín, 1969; Dickinson & Jennings, 1990; Nazi & Lea, 1994; Rogers et al., 1990; Schirmer & Toutain, 1991; Schnitman et al., 2003; Xu et al., 2006). Thus, the approach used is based on pattern
recognition using a distance calculus technique (Euclidean Distance) or a similarity tool (Pearson Correlation).

1.2 Document structure

This study is divided into more five sections. In the following section, the tools used are theoretically justified, discussing the techniques of descriptors based on edges and some tools for calculating distance metrics and statistics. In Section 3, the oil Sucker-Rod Pumping System is presented, being a study case for implementation of the proposal work. Then, in the next section the methodology for detection and recognition of the failure is presented. Section 5 presents the results obtained and finally the conclusions are presented in Section 6.

2. Theoretical basis

2.1 Description of boundary descriptors

The boundary descriptors are mathematic methods that describe a object or a region of figure. The descriptors are separated in two groups (Gonzalez et al., 2003): Descriptors based in contour (border) and Descriptors based in region. The first, it describe the object shape basing in its contour. Now, the region descriptors describe the object inside. The proper descriptor ideal must show the following invariant features:

- Translation;
- Rotation;
- Scale;
- Start Point.

In the diagnose process of faults in the Sucker-rod pumping system through dynamometer cards, the rotation feature is unnecessary, because several faults show the same contour, but they are rotated.

2.1.1 Centroid

The contour descriptor by centroid have main focus to calculate the distance between the card geometric center to several points that compose the card to make a distance set \( D = \{D_0, D_1, ..., D_n\} \). The Equations 1 and 2 show the centroid calculus, where \( N \) are the quantity of points that compose the card and the ordered pair, \( x_i \) and \( y_i \), represents the centroid.

\[
\begin{align*}
xc &= \frac{1}{N} \sum_{i=1}^{N} x_i \\
yc &= \frac{1}{N} \sum_{i=1}^{N} y_i
\end{align*}
\]

(1)

(2)

The Equation 3 shows the distance calculus between the centroid and several points.

\[
D_i = \sqrt{(x_i - xc)^2 + (y_i - yc)^2}
\]

(3)

So, it can use the distance set like contour feature of dynamometer card.
2.1.2 Curvature descriptor

The curvature descriptor is a simple and easy algorithm to develop and the main purpose is calculate the distance between any point in relation to next point (clockwise or counter-clockwise). The Equations 4, 5 and 6 show the distance calculus.

\[ D_{x_i} = (x_i - x_{i+1})^2 \]  
\[ D_{y_i} = (y_i - y_{i+1})^2 \]  
\[ D_c = \sqrt{D_{x_i} - D_{y_i}} \]  

2.1.3 K-curvature

The K-curvature extractor shows the object contour through of the created angle relation between two vectors. From any initial point, \( p_i \), two points, \( p_{i+k} \) and \( p_{i+2k} \), are chosen with a spacing of \( k \) values with the purpose to eliminate contour noises. So, two vectors (\( v \) and \( w \)) are defined. The vector \( v \) is formed by points \( p_i \) and \( p_{i+k} \), while the vector \( w \) is formed by \( p_{i+k} \) and \( p_{i+2k} \). The Equation 7 shows the angle calculus between the vectors.

\[ \theta = \cos^{-1} \frac{v \cdot w}{|v| \cdot |w|} \]  

So, \( v \cdot w \) is the scalar product between vectors (Equation 8) and \( |v| \) and \( |w| \) are the norms of vectors (Equation 9 and 10).

\[ v \cdot w = v_1w_1 + v_2w_2 + ... + v_nw_n \]  
\[ |v| = \sqrt{v \cdot v} \]  
\[ |w| = \sqrt{w \cdot w} \]  

2.1.4 Fourier Descriptor

The Fourier Descriptor is compact and light algorithm. To develop this algorithm, consider the following points: \((x_k, y_k)\) that represent the object contour coordinates, where \( k = 0, 1, 2, ..., N - 1 \) and \( N \) is the points quantity of border. The Equation 11 indicates the complex function of object contour coordinates.

\[ z(k) = (x_k) + j(y_k) \]  

In spite of sequence not be important to this descriptor, in this work, \( x \) is the position of polish rod and \( y \) is the applied system force. The Fourier descriptors (Equation 12) are make applying the Discrete Fourier Transform (DFT) in the Equation 11.

\[ F_n = \frac{1}{N} \sum_{k=0}^{N-1} z(k)e^{-\frac{2\pi ink}{N}} \]
\[ N = 0, 1, 2, ..., N - 1 \] and \( F_n \) are transformation coefficient of \( z(k) \). The descriptors can be rotation invariants when used magnitudes of transformation, \( |F_n| \). The scale can be normalize to divide the magnitudes of coefficient by \( |F_1| \).

### 2.2 Mathematical tools for calculating of similarity

#### 2.2.1 Euclidean distance

The Euclidean distance between two points is the length of the line segment connecting them. In Cartesian coordinates, if \( p = (p_1, p_2, ..., p_n) \) and \( q = (q_1, q_2, ..., q_n) \) are two points in Euclidean space, then the distance from \( p \) to \( q \), or from \( q \) to \( p \) is given by Equation 13.

\[
D = d(p, q) = d(q, p) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + ... + (p_n - q_n)^2}
\]

Equation 14

\[
D = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}
\]

Equation 15

#### 2.2.2 Pearson correlation

The Pearson correlation (or “product-moment correlation coefficient”, or also “r of Pearson”) measure the correlation degree and the direction between two variables of metric scale. This coefficient is represented by \( r \) and can be between -1 and 1. So, \( r \) can be analyzed in the following manners:

+1: It means a perfect correlation and the variables are in the same direction;

−1: It means a perfect correlation too, but, in this analysis, the variables direction is opposite.

0: In this case, the variables does not have a linear dependence.

In other words, the signal of the result correlation shows if the correlation is positive or negative and the proportion variable shows the correlation force.

The Pearson correlation coefficient is calculated according to the next formula:

\[
r = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}
\]

Where \( x_1, x_2, ..., x_n \) e \( y_1, y_2, ..., y_n \) are measure values both of variables. Moreover, \( x_n \) can be written:

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i
\]

And \( y_n \) can be:

\[
\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i
\]

These variables \( x_n \) and \( y_n \) are arithmetics meanings, both of variables \( x \) and \( y \).

### 3. Sucker-rod pumping system

The first artificial lifting method was the sucker rod pumping, that appears after the birth of the oil industry. The importance of this method is showed in the number of installations now operating in the world. The Figure 1 presents a Sucker Rod Pumping Unit.

www.intechopen.com
The great success of the sucker rod pumping system is linked to low cost in investment, maintenance, flow and deep flexibility, good energetic efficiency and the possibility to operate several down-hole conditions. But, the main advantage is the lowest cost/production relationship during the field production life.

3.1 Components of the sucker rod pumping

3.1.1 Downhole pump

The Downhole Pump is a positive displacement pump - in other words, when the fluid gets in suction, it does not return.

3.1.2 Rod string

The string rod is responsible for providing the surface energy to the downhole pump.

3.1.3 Sucker rod pumping unit

The pumping unit changes the rotation movement of the electric motor in the applied reciprocate movement for polish rod, while the reduction box decreases the rotational speed of the electric motor to allow the pumping speed.

3.1.4 Dynamometer card

A dynamometer card is a graph of effects originated to active charge in the pump during a pump cycle. The Figure 2(a) presents a example of real card.
3.2 Reference card patterns

There are two types of dynamometer cards: the surface and the down-hole card. The charges are recorded in the surface through the dynamometers and in the down-hole through special devices or mathematical models. The dynamometer card is one of main tools to analyze and review the condition of the system. This card is a record of the charges along the rod string path. It is possible to view several conditions of pumping through the dynamometer card. In this subsection, some Sucker Rod Pumping System cards are presented. Each card showed was chosen based on the main problems of oil fields and they can be found in other previous papers. (dos Santos Côrrea, 1995)

These presented cards (Figure 2(b)) are some reference patterns for the proposed model in the forward section.

3.2.1 Normal operation

The normal pumping pattern is associated with the follow characteristics:

- High volumetric Efficiency;
- Low interference of gas;
- Low or medium suction pressure.

3.2.2 Fluid pound

These patterns are associated with the follow characteristics:

- Low suction pressure;
- Low interference of gas;
- Blocked Pump Suction.

3.2.3 Valve leak

This patterns happen when there is a leak in the down-hole valves (Traveling or Standing valve).
4. Methodology

In this part, it will be presented the used methodology in approximately 1500 dynamometric cards from the sucker-rod pumping system. The current methodology meet four reference cards. Next, the data flow and applied descriptors will be presented.

4.1 Data flow

The used model of data flow is based on selection and processing faults pattern cards. Thus, a border descriptor is generated for each pattern card. After this step, it is repeated for analyze each field card. After the border descriptor of the field card, the Euclidean Distance is calculated with each border descriptor of the pattern cards. The result of each distance calculus is compared in a minimum function. The lower value is linked to the closest pattern of the field card.

The Figure 2 presents an information flow in the proposed model.

4.2 Data acquisition

The data is obtained through the supervisory software. This software gathers the field variables like current, force, horse head position, head pressure, down-hole pressure, among others.

The dynamometer cards are a two dimension graph of force (ordinate axis) versus horse head position (abscissa axis). This card consists of one hundred points.
4.2.1 Patterns selection

The pattern cards are selected by an expert engineer. Next, the selected cards are available for system to process. From this moment the treatment is similar to the field card as the pattern cards.

4.2.2 Descriptors generator

After the data acquisition and patterns selection, the cards are processed using the mathematical tool showed in the Section 2.1. Thereby, each field card can be used to calculate the distance with a pattern card that has its own identity.

4.2.3 Calculating of similarity

Any signature can be compared with a pattern through their descriptors using some mathematical tools for calculating distance. Thus, one similarity value (coefficient) is generated for every calculation of distance between one signature and the different patterns.

4.2.4 Minimum classifier and fault recognition

It was necessary to use a classifier to recognize what pattern is closer to the field card. In this study, a simple classifier was used and it is just a maximum function. This function is applied in the generated table in Section 4.2.2. Thereafter, the Euclidean Distance calculus that has the lower value represents the fault.

5. Results

This section is divided in three subsection. The first subsection, the general results are showed and discussed. In the second subsection, each descriptor is tested for invariance of the features needed to Sucker-rod pumping (Scale, Translation and Start point). After, it is showed a modification to improve the performance of recognition. In the last, the consolidate results are presented.

5.1 General results of euclidean distance

The Figures 3, 4, 5 and 6 are the result of pattern recognition analysis with Euclidean Distance. The K-Curvature Descriptor and Fourier Descriptor are the best performance, but in the analysis of Fluid Pound and Gas Lock Pattern, both not show good results.

5.2 General results of pearson correlation

The Figures 7, 8, 9 and 10 are the result of pattern recognition analysis with Pearson Correlation. The Centroid Descriptor and Fourier Descriptor are the best performance, but in the analysis of Fluid Pound and Gas Lock Pattern, both also do not present good results.

5.3 Tests of invariant characteristics

In this subsection, the presented results are about the robustness tests of descriptors in relation to features that must be invariant and for fault recognize on Sucker-rod pumping system. In all the tests, the chosen card represents a fault pattern card. The results are present in the Table 1.
Fig. 3. Euclidean Distance - Results to Showed Cards of Leaking Standing Valve

Fig. 4. Euclidean Distance - Results to Showed Cards of Leaking Traveling Valve
Comparison of Border Descriptors and Pattern Recognition Techniques Applied to Detection and Diagnose of Faults on Sucker-Rod Pumping System

Fig. 5. Euclidean Distance - Results to Showed Cards of Gas Lock

<table>
<thead>
<tr>
<th></th>
<th>K-Curvature</th>
<th>Centroid</th>
<th>Curvature Descriptor</th>
<th>Fourier Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>1.11</td>
<td>1.11</td>
<td>1.11</td>
<td>1.11</td>
</tr>
<tr>
<td>Gas Lock</td>
<td>33.33</td>
<td>25.56</td>
<td>48.67</td>
<td>5.56</td>
</tr>
<tr>
<td>Fluid Pound</td>
<td>57.78</td>
<td>73.33</td>
<td>52.22</td>
<td>92.22</td>
</tr>
<tr>
<td>Standing Valve</td>
<td>7.78</td>
<td>0.00</td>
<td>0.00</td>
<td>1.11</td>
</tr>
<tr>
<td>Traveling Valve</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Fig. 6. Euclidean Distance - Results to Showed Cards of Fluid Pound

<table>
<thead>
<tr>
<th></th>
<th>K-Curvature</th>
<th>Centroid</th>
<th>Curvature Descriptor</th>
<th>Fourier Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>Gas Lock</td>
<td>26.09</td>
<td>12.56</td>
<td>48.07</td>
<td>6.76</td>
</tr>
<tr>
<td>Fluid Pound</td>
<td>54.83</td>
<td>87.20</td>
<td>51.69</td>
<td>83.33</td>
</tr>
<tr>
<td>Standing Valve</td>
<td>15.94</td>
<td>0.00</td>
<td>0.00</td>
<td>4.58</td>
</tr>
<tr>
<td>Traveling Valve</td>
<td>2.90</td>
<td>0.00</td>
<td>0.00</td>
<td>5.07</td>
</tr>
</tbody>
</table>

www.intechopen.com
Fig. 7. Pearson Correlation - Results to Showed Cards of Leaking Standing Valve

Fig. 8. Pearson Correlation - Results to Showed Cards of Leaking Traveling Valve
Comparison of Border Descriptors and Pattern Recognition Techniques Applied to Detection and Diagnose of Faults on Sucker-Rod Pumping System

Fig. 9. Pearson Correlation - Results to Showed Cards of Gas Lock

Fig. 10. Pearson Correlation - Results to Showed Cards of Fluid Pound
Table 1. Invariance Tests

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>Translation</th>
<th>Scale</th>
<th>Start Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid</td>
<td>OK</td>
<td>OK</td>
<td>FAIL</td>
</tr>
<tr>
<td>K-Curvature</td>
<td>OK</td>
<td>OK</td>
<td>FAIL</td>
</tr>
<tr>
<td>Curvature</td>
<td>OK</td>
<td>OK</td>
<td>FAIL</td>
</tr>
<tr>
<td>Fourier</td>
<td>OK</td>
<td>OK</td>
<td>FAIL</td>
</tr>
</tbody>
</table>

5.3.1 Translation invariance

Translation tests, the chosen card was translated as shown in Figure 11. As expected, all descriptors were successful in recognizing.

5.3.2 Scaling invariance

Scaling tests, the chosen card was scaled as shown in Figure 12. And also, all descriptors were successful in recognizing.

5.3.3 Start point invariance

Start point tests, the start point was changed. The Figure 13 shows the both start points (Original Start Point and Modified Start Point) from data acquisition. But now, as can be seen, all descriptors were not successful.

The Fourier descriptor is able to solve this problem. Modifying the Equation 12 to 16 through to calculate of absolute value of Fourier Transform. But, this procedure insert a new problem to recognize of fault that are equal when rotated. It can be seen in the Figure 14 where different faults are recognized as the same when this process is used.

\[
F_N = \text{abs} \left\{ \frac{1}{N} \sum_{k=0}^{N-1} z(k) e^{-\frac{2\pi ik}{N}} \right\}
\] (16)
Fig. 12. Used Example to Teste Scale Invariance

Fig. 13. Used Example to Teste Start Point Invariance
5.4 Proposed modification

To improve the performance of pattern recognizing to Fluid Pound and Gas Lock faults, it is proposed a specialist system that analyze the curvature card. The better curvature that represents a third order function is identified as Gas Lock fault. Look at the results for Euclidean Distance in the Figure 15 and in the Figure 16.

It is possible to observe that the results as the others faults not change its values.

5.5 Consolidate results

The Table 2 presents the consolidate results. Thus, it can be observed that Fourier Descriptor was better than others when used with Euclidean distance and it is as good as the Centroid Descriptor when used with the Pearson Correlation.

<table>
<thead>
<tr>
<th>Border Descriptors</th>
<th>Sucess (%)</th>
<th>Euclidean</th>
<th>Pearson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid</td>
<td>45,91</td>
<td>84,55</td>
<td></td>
</tr>
<tr>
<td>K-Curvature</td>
<td>81,49</td>
<td>68,72</td>
<td></td>
</tr>
<tr>
<td>Curvature</td>
<td>45,91</td>
<td>62,11</td>
<td></td>
</tr>
<tr>
<td>Fourier</td>
<td>86,60</td>
<td>83,12</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Consolidate Results
Fig. 15. Modified System - Results to Showed Cards of Fluid Pound

Fig. 16. Modified System - Results to Showed Cards of Gas Lock

6. Conclusion

Nowadays, the quantity of onshore wells using Sucker-Rod Pumping System is higher and it stickles the engineer work. In addition, the difficulty in recognizing a specific card shape augment as the amount of noise increase mainly as a function of well depth.

This study, using the processing image, is suitable for fault diagnosis of Sucker-Rod Pumping System and could help to interpret the down-hole condition of oil well promptly and correctly.
The results presented a high efficiency for processed field cards and seemed to be very robust to inherent problems in the processing images, like rotation, translation and scale. In the future, it is desired to test other recognizing functions and to develop an capable system of identify all faults. It will permit the prediction and the planning of the maintenance, so the field engineer manages his equipments.

7. References

This book presents several recent advances that are related or fall under the umbrella of 'digital image processing', with the purpose of providing an insight into the possibilities offered by digital image processing algorithms in various fields. The presented mathematical algorithms are accompanied by graphical representations and illustrative examples for an enhanced readability. The chapters are written in a manner that allows even a reader with basic experience and knowledge in the digital image processing field to properly understand the presented algorithms. Concurrently, the structure of the information in this book is such that fellow scientists will be able to use it to push the development of the presented subjects even further.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
