Abstract—Corrosion causes many failures in chemical process installations. These failures generate high costs, therefore an effective corrosion monitoring system is needed. This paper focuses on the classification of the most important corrosion processes: pitting, stress corrosion cracking (SCC) and general corrosion. The computations and algorithms involved in the classification of the corrosion time series are presented. A technique for trend removal of the time series is proposed. Features to extract the characteristics of the corrosion time series are designed. A genetic algorithm for the selection of features is described in which the correlations between features are exploited. The combination of the proposed techniques leads to a new high performing pattern recognition system for corrosion time series classification.

Keywords—Genetic algorithm, Haar wavelet transform, minimum message length, subspace decomposition.

I. INTRODUCTION

In the literature [1]-[3] it is mentioned that about 3 to 5% of the Gross National Product in the United States is lost due to the costs caused by corrosion. Furthermore it is estimated [1], [2] that 35% of these costs can be avoided by the application of existing technologies (coatings, inhibitors, cathodic protection...) and by appropriate services (consulting, testing, monitoring...). Pitting, stress corrosion cracking (SCC) and general corrosion cause most of the corrosion damage in chemical process installations [4]. Pitting causes pits that lead to leaks, stress corrosion cracking causes cracks and general corrosion leads to a uniform decrease of thickness of the material. Each of the corrosion processes calls upon different actions to be taken by the service engineers of the chemical process installation. Therefore it is important to develop a pattern recognition system that is able to detect and classify the different corrosion processes automatically.

II. ELECTROCHEMICAL NOISE (ECN) TIME SERIES

Time series used in the analysis for this paper are acquired from electrochemical noise measurements on a reference probe of stainless steel in laboratory conditions [4]. In the final application the probe will be inserted in chemical plant installations. The corrosion behavior of the installation is then monitored from the corrosion of the reference probe. Time series are acquired at a sampling rate of 10Hz. A detailed description of the experimental set-up and different environmental conditions to induce the different types of corrosion can be found in [4]. For the analysis, 207 time series in absence of corrosion processes are selected, 207 time series with general corrosion, 75 time series with SCC and 90 time series with pitting. Fig. 1-Fig. 4 show examples for electrochemical noise time series in absence of corrosion, for general corrosion, for pitting and for SCC time series, respectively.

 Manuscript received November 4, 2004. This work was supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

G. Van Dijck is with the Computational Neuroscience Research Group of the Laboratory of Neurophysiology and the Department of Metallurgy and Materials Engineering (MTM), K.U. Leuven, 3000 Leuven, Belgium (phone: +32 16 32 12 47; fax: +32 16 32 19 90; e-mail: gert.vandijck@mtm.kuleuven.ac.be).

M. Wevers is with the Department of Metallurgy and Materials Engineering (MTM), K.U. Leuven, 3000 Leuven, Belgium, (e-mail: martine.wevers@mtm.kuleuven.ac.be).

M. Van Hulle is with the Computational Neuroscience Research Group of the Laboratory of Neurophysiology, K.U. Leuven, 3000 Leuven, Belgium, (e-mail: marc@neuro.kuleuven.ac.be).

Fig. 1. Example of electrochemical voltage time series in absence of corrosion.

Fig. 2. Example of electrochemical voltage time series for general corrosion.

Fig. 3. Example of electrochemical voltage time series for pitting.

Fig. 4. Example of electrochemical voltage time series for SCC.
III. DESIGN OF THE PATTERN RECOGNITION SYSTEM

Different stages in the design of the pattern recognition system will be presented in this paragraph. Firstly, a trend removal technique is proposed to facilitate the identification of time series in absence of corrosion. Secondly, features are derived from the time series. In the extraction of features the Haar wavelet transform and the power spectral density (PSD) are used. Thirdly, a Bayesian classifier is designed using the minimum message length (MML) criterion for density estimation. Finally, a genetic algorithm is described to search for a subset of features that leads to a high classification performance.

A. Trend Removal From Time Series

The observed time series contain voltage drifts, both linear and nonlinear, which are not related to the active corrosion process. These drifts are related to changes in system influences: change in temperature, change in flow of the liquid... [4]. These trends cause the different corrosion processes to have similar autocorrelations. This makes the separation between different corrosion processes more difficult. The proposed technique for trend removal operates in time domain by subtracting the estimated trend from the original time series. A description of the subspace decomposition based trend removal algorithm follows. First the signal is decomposed into 50% overlapping frames of length L. All samples within a frame are then multiplied by the corresponding coefficients of a Hanning window of size L. The resulting data samples are arranged in an \( M \times N \) Hankel matrix structure:

\[
\mathbf{H}(x_n) = \begin{pmatrix}
  x_0 & x_1 & x_2 & \cdots & x_{N-1} \\
  x_1 & x_2 & x_3 & \cdots & \vdots \\
  x_2 & x_3 & x_4 & \cdots & \vdots \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  x_{M-1} & \cdots & x_{L-2} & x_{L-1}
\end{pmatrix}
\]

(1)

where \( L = M + N - 1 \).

Next, this matrix is decomposed by singular value decomposition:

\[
\mathbf{H} = (\mathbf{U}_T \mathbf{U}_S \mathbf{U}_0) \begin{pmatrix}
  \mathbf{\Sigma}_T & 0 & 0 \\
  0 & \mathbf{\Sigma}_S & 0 \\
  0 & 0 & \mathbf{\Sigma}_0
\end{pmatrix} \begin{pmatrix}
  \mathbf{V}'_T \\
  \mathbf{V}'_S \\
  \mathbf{V}'_0
\end{pmatrix} = \begin{pmatrix}
  \mathbf{V}_T \\
  \mathbf{V}_S \\
  \mathbf{V}_0
\end{pmatrix}
\]

(2)

where \( \mathbf{U}_T \in \mathbb{R}^{M \times K}, \mathbf{U}_S \in \mathbb{R}^{M \times P}, \mathbf{U}_0 \in \mathbb{R}^{M \times (N-K-P)}, \mathbf{\Sigma}_T \in \mathbb{R}^{K \times K}, \mathbf{\Sigma}_S \in \mathbb{R}^{P \times P}, \mathbf{\Sigma}_0 \in \mathbb{R}^{(N-K-P) \times (N-K-P)}, \mathbf{V}'_T \in \mathbb{R}^{K \times T}, \mathbf{V}'_S \in \mathbb{R}^{P \times T}, \mathbf{V}'_0 \in \mathbb{R}^{N-K-P \times T} \).

Equation (2) models the electrochemical voltage signal as a trend (referred to by subscript T) on which the relevant part of the signal (referred to by subscript S) is superposed. Furthermore it is assumed that noise, not related to a corrosion process (referred to by subscript 0), is present and superposed on the relevant signal part and the trend. In this paper only the subtraction of \( \mathbf{H}_T \) from \( \mathbf{H} \) is addressed. The separation of noise from the relevant signal part is not considered. Observation of the singular values in (2) from the time series shows one large singular value and several smaller singular values. According to the model this singular value is identified as belonging to the trend. Therefore it is further assumed that \( K = 1 \).

Matrix \( \mathbf{H}_T \) is not of Hankel structure. Averaging the diagonal elements of \( \mathbf{H}_T \) for which \( i+j \) is constant, where \( i \) indicates the row index and \( j \) the column index, repairs the Hankel structure. The trend for the particular frame is then found on the main diagonal. The complete trend for the whole time series is found by adding the 50% overlapping trends from each frame. The use of the Hanning window in the time series decomposition guarantees a smooth reconstruction. Fig. 5 shows the estimated trend superposed on an electrochemical noise time series, with \( L \) set to 1500.

![Fig. 5. Estimated trend. The curve in white is the estimated trend from the subspace decomposition technique. Fig. 1 is obtained from Fig. 5 after subtracting the trend from the signal.](image_url)
B. Feature design from power spectral density (PSD) and Haar wavelet transform

Features are extracted from the time series to characterize each corrosion process in a few parameters. In this paragraph the choice of features is motivated. Three sets of features will be computed: one set derived from the PSD and two sets from the Haar wavelet transform. A GA selects among these sets of features the features that are relevant to distinguish between the different corrosion processes.

1) Features from the power spectral density

It is common practice to consider the power spectral density in analysis of electrochemical noise corrosion time series [5], [6]. Electrochemical noise voltages show power law behavior [7]. This means the power spectrum can be interpolated by an $f^{-\beta}$ function:

$$\left| A(f) \right|^2 = c f^{-\beta}.$$  \hspace{1cm} (3)

When power law behavior is observed for at least two decades, the time series are considered fractal [8]. However power law behavior is not present in all time series for the complete range of two decades. Therefore the frequency scale is divided in different decades. The 1st decade ranges from 0.012Hz-0.12Hz, the second from 0.12 Hz-1.2Hz and the last part of a decade from 1.2Hz-5Hz (half the sampling rate). To each frequency range and all combinations of ranges a straight line is fitted to the log. PSD vs. log. frequency scale using the least squares criterion. The PSD is estimated by means of the periodogram method [9]. The slopes of the lines are considered as parameters. By making all combinations of frequency ranges, what leads to 7 parameters, potentially irrelevant parameters are derived. The selection of relevant parameters is dealt with in a feature subset selection algorithm.

2) Features from the Haar wavelet transform

The choice of the Haar wavelet transform is motivated by observation of transients in pitting and SCC data. The CWT is applied in this case as a correlator with the Haar wavelet as a template. One can consider the Haar wavelet as a mathematical simplification of a corrosion transient occurring in the electrochemical noise time series. At positions where transients occur, large wavelet coefficients (5) are generated. By making comparisons in standard deviations, energy… of positive and negative wavelet coefficients one can quantize the ‘asymmetry in transient behaviour’ in parameters. With ‘asymmetry in transient behaviour’ the tendency to contain voltage transients, as observed in pitting, vs. negative voltage transients, as observed in SCC, is meant. One set of parameters compares standard deviations of positive vs. negative accumulated wavelet coefficients:

$$x(k) = \frac{\text{std}_{F(b)_{acc,k < 0}}(F(b)_{acc,k})}{\text{std}_{F(b)_{acc,k > 0}}(F(b)_{acc,k})}$$  \hspace{1cm} (6)

where $F(b)_{acc,k}$, from $F(b)_{acc,k}$ and comparing the power in the extremes:

$$y(k) = \frac{\text{mean}(F(b)^2_{acc,k \text{,max}})}{\text{mean}(F(b)^2_{acc,k \text{,min}})}$$  \hspace{1cm} (7)

Note that x(k) and y(k) represent sets of parameters, which depend upon the upper limit of accumulation. It is not a prior known which $k$ values lead to good features. This is related to the duration of transients. Selection of good $k$ values is addressed in the feature subset selection algorithm. Note that features from x(k) are strongly correlated, as well as the features from y(k). This is due to the fact that a small change in $k$ generates very similar features.

C. Classifier

In this paper a Bayesian classifier is opted for. In a Bayesian classifier feature $x$ is assigned to the class with maximum posterior class probability:

$$\psi(t) = 1, \forall t \in [0,0.5]; \quad \psi(t) = -1, \forall t \in [0.5,1]; \quad \psi(t) = 0, \forall t \notin [0,1]$$  \hspace{1cm} (4)

The continuous wavelet transform (CWT) is defined by [10]:

$$F(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt$$  \hspace{1cm} (5)

where $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$, with $a \neq 0$ and $a,b \in \mathbb{R}$.
\[ j = \arg \max_i P(C_i \mid \mathbf{x}) = \frac{P(\mathbf{x} \mid \mathbf{\theta}, C_i) P(C_i)}{P(\mathbf{x})} \]  \quad (8)

where \( C_i \) represents the class.

The class conditional densities \( P(\mathbf{x} \mid \mathbf{\theta}, C_i) \) are estimated by means of a Gaussian mixture model (GMM):

\[ P(\mathbf{x} \mid \mathbf{\theta}, C_i) = \sum_{m=1}^{k} \alpha_m p(\mathbf{x} \mid \theta_m, C_i) \]  \quad (9)

where \( \mathbf{\theta} = \{ \theta_1, \ldots, \theta_k, \alpha_1, \ldots, \alpha_k \} \). \( \theta_m \) are the parameters of the Gaussian probability function \( p(\mathbf{x} \mid \theta_m, C_i) \). \( \alpha_1, \ldots, \alpha_k \) are the mixing probabilities that must satisfy:

\[ \alpha_m \geq 0, \ m = 1, \ldots, k, \text{ and } \sum_{m=1}^{k} \alpha_m = 1. \]

The parameters in (9): \( k, \alpha_m, \theta_m \), are optimized by minimization of the MML criterion for GMM’s [11].

\[ L(\mathbf{\theta}, \mathbf{Y}) = \frac{N}{2} \sum_{m, n, m \neq n} \log(n \alpha_m^{2}) + \frac{k_{nz}}{2} \log(n) \]

\[ + \frac{k_{nz}(N+1)}{2} - \log p(\mathbf{X} \mid \mathbf{\theta}) \]  \quad (10)

where \( k_{nz} \) denotes the number of nonzero probability components, \( N \) is function of the dimensionality of the data, \( n \) the number of data points. \( \log p(\mathbf{X} \mid \mathbf{\theta}) \) is the log-likelihood of the data:

\[ \log p(\mathbf{X} \mid \mathbf{\theta}) = \log \left( \prod_{i=1}^{n} p(\mathbf{x}^{(i)} \mid \mathbf{\theta}) \right) = \sum_{i=1}^{n} \log \left( \sum_{m=1}^{k} \alpha_m p(\mathbf{x}^{(i)} \mid \theta_m) \right). \]  \quad (11)

Minimization of (10) is performed by means of the component-wise expectation maximization (CEM) algorithm [11], [12]. While the maximum-likelihood (ML) criterion has been extensively used in literature it cannot be used to estimate the number of components \( k \).

D. Feature subset selection

A two stage genetic algorithm (GA) [13] with variable chromosome length is designed to search for a subset of features with high classification performance. In the GA the correlations between features are exploited. It is shown from different simulations that this leads to subsets with higher classification performance compared to a standard GA without exploitation of the correlations [14]. A detailed description of the GA in combination with the classifier can be found in [14]. A short description follows.

In a first step the features are clustered with the correlations as a measure of distance between features. An agglomerative hierarchical clustering method [15] is used. For each cluster of features a representative feature is chosen. For this the feature closest to the cluster center is chosen. Next a GA with variable length of chromosomes searches for a combination of clusters that leads to a high classification performance using the representative feature from each cluster. Starting from the best combination of clusters a 2nd local search with a GA with fixed length of chromosomes is performed in which the features within the clusters are considered. As fitness function, a cosh function on the classification performance is chosen, averaged over 5 runs of 10-fold crossvalidation. The Bayesian classifier is used in the prediction of the class labels.

IV. RESULTS

From 4 different runs of the GA algorithm the best subset contained 5 features. 2 features are selected from set \( x(k) \), 2 from set \( y(k) \) and 1 from the set of PSD slopes. For this selection of features a very high classification performance is obtained.

The classification performance averaged over 10 runs of 10-fold crossvalidation equals 96.41%. Observation from confusion matrices shows that most confusion occurs between: general corrosion – SCC and general corrosion – pitting. Fig. 7 shows the decision boundaries computed by the Bayesian classifier for a good selection of 2 parameters.

V. CONCLUSION

A new high performing pattern recognition system is designed for electrochemical noise time series classification of different types of corrosion of stainless steel. A detrending technique based on subspace decomposition is proposed. It was shown that the detrending technique leads to an easier identification of noise. It was shown that a combination of features based on the PSD and the Haar wavelet transform leads to a high classification performance by means of a Bayesian classifier.
REFERENCES


