SWITCH MECHANISM DIAGNOSIS USING A PATTERN RECOGNITION APPROACH

F. Chamroukhi[1], A. Samé[1], P. Aknin[1], M. Antoni[2]

[1]INRETS-LTN, 2 avenue du général Malleret-Joinville, 94114 Arcueil, France
{chamroukhi, same, aknin}@inrets.fr
[2]SNCF, Direction de l'Infrastructure, IMT-IM6, 17 rue d'Amsterdam; 75379 Paris Cedex 08
marc.antoni@sncf.fr

Keywords: Condition monitoring, diagnosis, switch mechanism, pattern recognition, Gaussian mixture model, EM algorithm

Abstract

An original pattern recognition approach for the diagnosis of switch mechanisms driven by an electric motor is presented in this paper. Its main advantage is that it does not require a physical model of the system and can easily be adapted to other complex systems. The available data for this task are the signals of the electrical power consumption during the switch actuation period and the proposed method consists of two steps: the feature extraction from the signals and the recognition of different operating states (class without defect, class with minor defect and class with critical defect) using Mixture Discriminant Analysis (MDA). This method assumes the classes to be represented by a Gaussian mixture distribution whose parameters are estimated by the maximum likelihood method, using the Expectation-Maximization (EM) algorithm. An experimental study performed on real measured signals covering a wide range of defects reveals some good performances of the proposed approach compared to others classification methods such as K-Nearest-Neighbors, Neural Networks and the classical Bayesian discriminant approach (with one Gaussian distribution per class).

1 Introduction

The remote monitoring of the railway infrastructure components and more particularly of the switch mechanism is of great interest for the operators. The problem consists in early detecting the presence of a defect in order to alert the maintenance service before the failure.

In this framework, a pattern recognition approach is proposed for the diagnosis of the switch mechanisms of the French network. The particular switches considered in this paper (see figure 1) are driven by an electric motor and equiped with a clamp locking system so called VCC ("Verrou Carter Coussinet" in french).

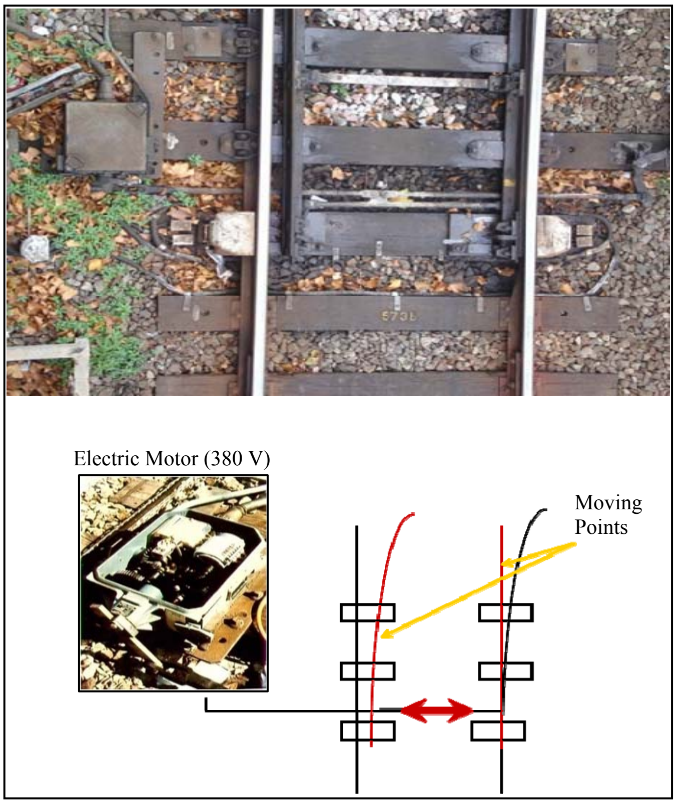

Figure 1: Switch mechanism driven by an electrical motor.

The proposed approach, which does not require a physical model of the system simply uses the signals of the electrical power consumed during the switch actuation period (see figure 2 and figure 3). It consists of two steps:

1. the feature extraction from signal,

2. the learning of the different class parameters (class without defect, class with minor defect and class with critical defect) using a labelized collection of signals.

The Mixture Discriminant Analysis (MDA) [6], which allows complex class distributions to be modeled precisely, is used to estimate the parameters of each class of signals.

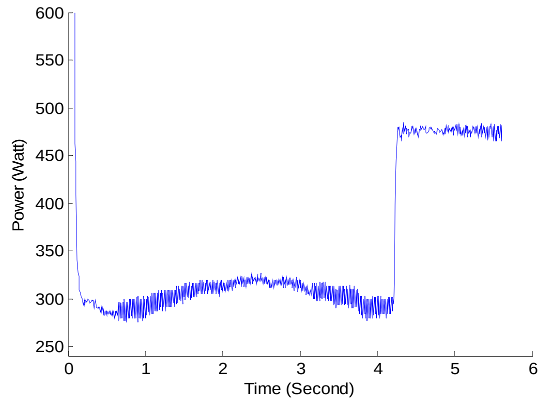

Figure 2: Example of signal corresponding to a normal switch operation

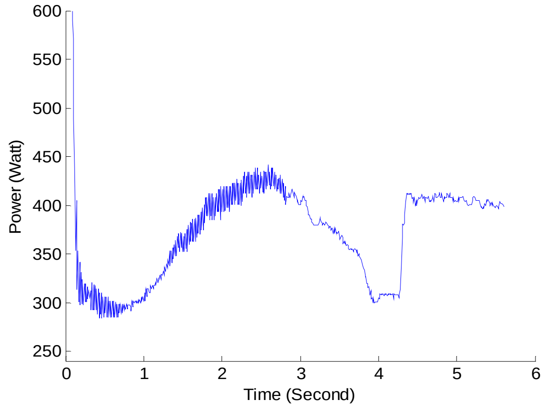

Figure 3: Example of signal corresponding to a switch operation with lake of lubrication

This computed aided decision method can be summarized according to the block diagram of figure 4.

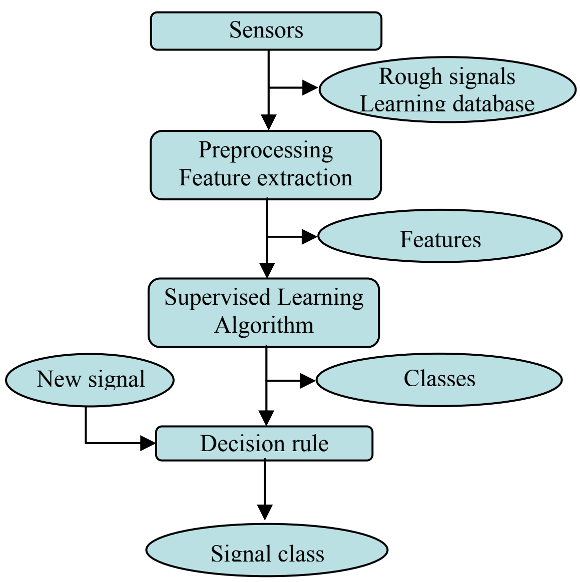

Figure 4: Block diagram of pattern recognition system.

The paper is organized as follows: in section two we describe the signal parameterization method. The section three presents the proposed mixture discriminant approach for signal classification and an experimental study is summarized in the final section.

2 Signal parameterization

As already mentioned, the available information for each switch actuation is given by the electrical power signal. Each signal consists of 550 points sampled at 100 hz. The dimension of the deduced representation space is too large for the pattern recognition task. Consequently, a pre-processing task is required to deliver a reduced input space that resumes information efficiently and allows avoiding the curse of dimensionality problem [1].

The switch actuation consists of successive mechanical motions of different parts of the mechanism, which are observed on the shape of the signal (see figure 5):

- starting phase,
- points unlocking,
- points translation,
- points locking,
- friction phase.

The duration of the switch operation is assumed to be uniform for all signals.

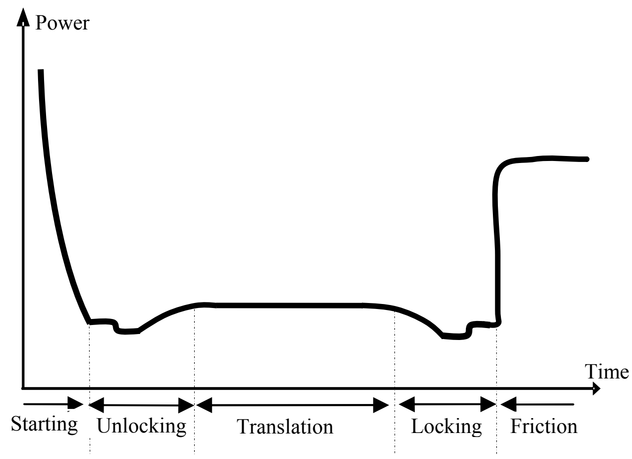

Figure 5: The curve of the electrical consumed power during a switch operation

So we choose to resume the information of each signal phase by the few following parameters:

- the maximum value of the signal,
- the minimum value,
- the mean value,
- the standard deviation,
- the coefficients of a second degree polynomial regression model; the used polynomial model is given by $h(t) = \lambda (t - \alpha)^2 + \mu$, where $h(t)$ is the signal

at time t, and (λ, α, μ) and are the polynomial coefficients used.

Consequently, each signal is parameterized by a parameter vector of dimension 21 (7 parameters for each of the 3 central phases that are sufficient for our recognition problem).

3 Learning of classes parameters

Three different classes of signal, corresponding to the different operating states of the switch mechanism, are considered:

- C_1: class without defect,
- C_2: class with minor defect,
- C_3: class with critical defect.

In this section, we denote by x_i the feature vector extracted from the signal of the i^{th} observation.

3.1 Mixture distribution modeling

Given a labelized collection of signals (signals with known classes), the parameters of each class are learned using the Mixture Discriminant Analysis (MDA) [6]. According to this approach, each class C_k density is modeled by a Gaussian Mixture distribution [4] [6], which can be defined as follows:

$$f(x_i \mid C_k) = \sum_{r=1}^{R_k} \pi_{rk} N(x_i \mid m_{rk}, \Sigma_{rk}), \quad (1)$$

where R_k is the number of components of the mixture, the π_{rk} ($r=1,\ldots,R_k$) are the proportions of the mixture verifying $\sum_{r=1}^{R_k} \pi_{rk} = 1$, and $N(x_i|m_{rk},\Sigma_{rk})$ is the Gaussian probability density function with mean m_{rk} and covariance matrix Σ_{rk}. This particular distribution allows complex classes shapes to be precisely modelized. Figure 6 shows an example of mono-dimensional four-component Gaussian mixture distribution.

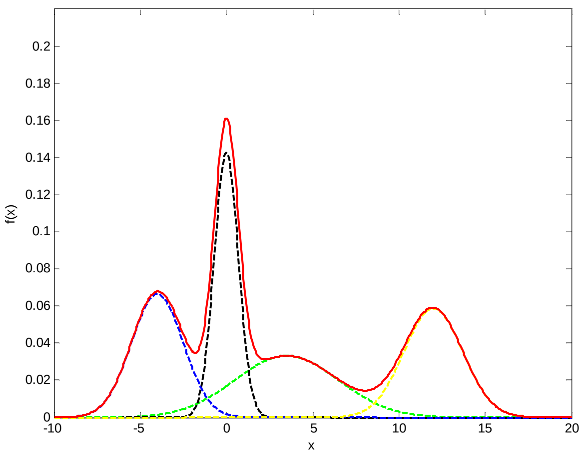

Figure 6: Example of a Gaussian mixture distribution (in solid red line) and its four Gaussian components (in dotted lines)

3.2 Parameter estimation

The mixture distribution parameters for each class C_k are estimated by the maximum likelihood method [5]. In our situation, the log-likelihood function can be written as:

$$\begin{aligned} L(\theta_k) &= \log \prod_{i=1}^{n} f(x_i \mid C_k) \\ &= \sum_{i=1}^{n} \log\left[\sum_{r=1}^{R_k} \pi_{rk}\phi\left(x_i \mid m_{rk}, \Sigma_{rk}\right)\right], \end{aligned} \quad (2)$$

where $\theta_k = (\pi_1,\ldots,\pi_{R_k}, m_1,\ldots,m_{R_k}, \Sigma_1,\ldots,\Sigma_{R_k})$ is the parameter vector of the class C_k and vectors $x_1,\ldots,x_n$ are the class C_k signals features.

The Expectation-Maximization (EM) algorithm [4] is used to perform the maximization of the log-likelihood (2) which cannot be solved directly. The optimal number of Gaussian distributions R_k for each class is computed by maximizing the Bayesian Information Criterion BIC [7].

$$BIC(R_k) = L(\theta_k) - \frac{v_{R_k}}{2}\log(n_k), \quad (3)$$

where θ_k is the estimate provided by the EM algorithm, v_{Rk} is the dimension of the parameter vector θ_k, and n_k is the cardinal number of the class C_k.

3.3 Signal classification

Given the parameters estimated by the EM algorithm for all classes, each new signal designed by the feature vector x_i is assigned to the class k^* which maximizes the posterior probability that x_i originates from the class C_k. This class is given by:

$$k^* = \arg\max_k P(C_k \mid x_i), \quad (4)$$

with

$$P(C_k \mid x) = \frac{f(x \mid C_k)P(C_k)}{\sum_{k=1}^{K} f(x|C_k)P(C_k)}, \quad (5)$$

where $P(C_k)$ is the prior probability of class C_k estimated by the proportion of signal belonging to class C_k into the learning phase. This classification rule is known as the *maximum a posteriori* (MAP) rule.

4 Experiments

This section is devoted to the evaluation of the proposed approach in terms of classification accuracy. A database of 119 real signals with known classes has been used for this experimentation. This database has been divided in two groups: a training base of 90 signals for the classes parameters learning and a test base of 29 signals for the evaluation of the classifier.

After the parameterization phase, the EM algorithm is run for each class. The number of mixture components estimated by the BIC criterion, for each class C_k, is given in table 1.

Class	C_1	C_2	C_3
Number of mixture components	4	2	2

Table 1: Number of selected mixture components

Each signal of the test base is classified according to the MAP rule given by equation (5). The results obtained by the proposed MDA approach are compared with those provided by:

- the Neural Network (NN) approach based on a multilayer perceptron [2][5] with a single hidden layer of 13 cells,
- the K Nearest Neighbours (KNN) approach [3][5],
- the Bayesian discrimination approach based on a single Gaussian distributions per class [5].

Table 1 shows the percentage of good classification obtained with the four concurrent approaches.

MDA	**95 %**
NN	90 %
KNN	88 %
Bayesian disc. with one Gaussian	75 %

Table 2: Results of the classification by the different approaches

These results show clearly that the proposed approach outperforms Neural Network, K-Nearest Neighbours and Bayesian discrimination approaches. A good classification rate of 95 % is obtained with MDA. This result confirms the fact that mixture modelling can improve decision boundaries.

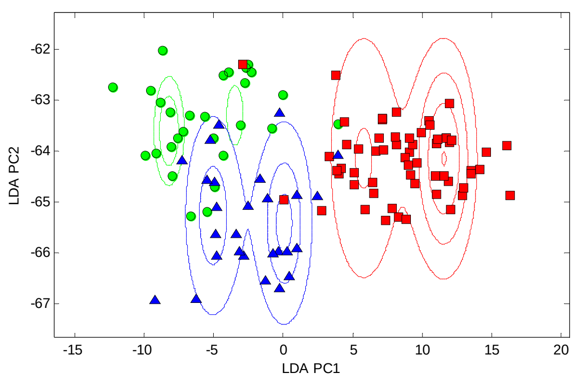

Figure 7: Projection of the dataset and the estimated mixture distributions (isolines) into the principal factor discriminant plane (class1=green circle, class2=blue triangle, class3=red square)

Figure 7 shows the projection of both the database and the estimated mixture distributions on the plane formed by the two axis of a factor discriminant analysis (FDA) [5]. Let's recall that the FDA aim to find the best representation subspace of the data where the inner-class variance is the less and the between-class variance is the larger. Despite of the dimension reduction, this graphic clearly highlights the convenience of mixture modelling. Moreover, it can be noticed that only two components of the class 1 among the four ones, are visible on this projection plane.

5 Conclusion

A new methodology for switch mechanism diagnosis driven by an electric motor was proposed in this paper. This approach does not require a physical model of the system but simply the electrical power consumption during the switch actuation period. It consists of two main steps: the feature extraction from the observed signals performed using polynomial fitting and elementary statistics, and the implementation of a probabilistic discriminant approach for the learning of classes (corresponding to the operating states of the switch mechanism) parameters. The used probabilistic approach models each class of signals by a mixture of Gaussian distributions allowing complex decision boundaries to be learned more precisely. The experimental study on real measured signals shows some good performances of the proposed approach, compared to alternative approaches.

References

[1] R. E. Bellman. *Adaptive Control Processes*. Princeton University Press, Princeton, NJ, (1961).

[2] C. M. Bishop. "Neural network for pattern recognition", Oxford University Press, NY (1995).

[3] B. V. Dasarathy (editors). "Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques", (1991).

[4] A. P. Dempster, N. M. Laird, and D. B. Rubin. "Maximum likelihood for incomplete data via the EM algorithm", The Journal of the Royal Statistical Society, Series B, volume 39(1), pp. 1-38, (1977).

[5] R. O. Duda, P. E. Hart and D. G. Stork. "Pattern classification (2nd ed.)", John Wiley and Sons, (2001).

[6] T. Hastie and R. Tibshirani. "Discriminant analysis by Gaussian mixtures", *Journal of the Royal Statistical Society, B*, Volume 58(1), pp. 155-176, (1996).

[7] G. Schwarz. "Estimating the dimension of a model", Annals of statistics, Volume 6 (2), pp. 461-464, (1978).