Non-stationary Condition Monitoring Through Event Alignment

Niels Henrik Pontoppidan and Jan Larsen
Intelligent Signal Processing, Informatics and Mathematical Modelling
Technical University of Denmark, DK-2800 Lyngby, Denmark.
{nhp,jl}@imm.dtu.dk, www.imm.dtu.dk/∼{nhp,jl}

Introduction

We present an event alignment framework which enables change detection in non-stationary signals. Classical condition monitoring frameworks have been restrained to laboratory settings with stationary operating conditions, which are not resembling real world operation. A traditional engine is controlled by the cam-shaft position; nowadays electronic control of large diesel engines, is emerging allowing for small timing changes of certain events such as injection, valve openings and closings etc. From a condition monitoring perspective this leads to changes that systems have to be invariant to.

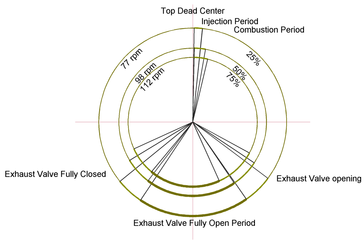
In this paper we apply the technique for non-stationary condition monitoring of large diesel engines based on acoustical emission sensor signals. The main machine learning aspect in this application overfitting, e.g., when the invariance wrt. the known changes prevents the system of correctly detecting faults.

The performance of the event alignment is analyzed in an unsupervised probabilistic detection framework based on outlier detection with either Principal Component Analysis or Gaussian Processes modeling. We are especially interested in the true performance of the condition monitoring performance with mixed aligned and unaligned data, e.g. detection of fault condition of unaligned examples versus false alarms of aligned normal condition data. Further, we expect that the non-stationary model can be used for wear trending due to longer and continuous monitoring across operating condition changes.

Event Alignment

Event alignment should provide a warp between two known conditions, so that warped data can be classified using a condition model trained on the reference data. In other words we train a condition model using normal say 25% load data (see data description below), we also take normal 50% load data and produce the warp. With the warp we should be able to detect faulty conditions under both 25% load and 50% load.

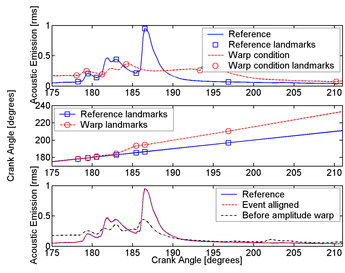
In order to build the warp we have gathered a list of important events. Each mode is described by a set of angular positions defining say start, peak and end of these events.



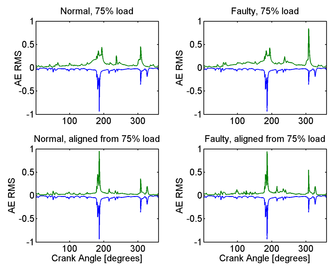
Timing of some important events at different load settings. The length of each circle is proportional to the rotation period, so both time and angular lengths of events at different loads can be compared. The figure clearly shows that merely scaling or translation is not enough to align the events, so that higher order methods as splines is necessary. Further the figure back our idea of interpolating between a few known warps in order to scope the whole range of possible load and speeds.

$$x_A = (f(x_W, L_R, L_W) - \mu_W) .* g + \mu_R, \quad (1)$$
$$\mu_W = \langle f(x_W, L_R, L_W) \rangle \quad (2)$$
$$g_i = \begin{cases} 1 & , \sigma_{iR} > \sigma_{Wi} \\ \sigma_{Ri}/\sigma_{Wi} & , \sigma_{Ri} > \sigma_{Wi} \end{cases}, \; i = 1, 2, \ldots, d \quad (3)$$

The above equations apply the necessary time stretching and amplitude modification to invert the result of the timing changes. The figure below display to AE RMS signals zoomed in on the injection period (180° out of phase with above figure), showing the timing differences and landmarks.

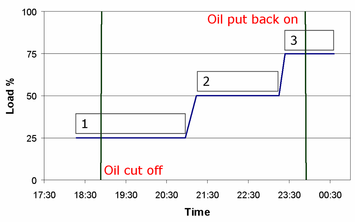


By alignment of the landmarks, i.e. bringing the red line to the blue (in middle) we see how the un-aligned peaks in the upper plot becomes aligned, however still with amplitude mismatch.s



Event alignment of normal and faulty data. All figures display the reference mode mirrored down. A warp that aligns normal 75% load data with normal 25% load data was also applied to faulty 75% load data. The two lower plots show how the faulty example is still deviating after the event alignment.

Data description



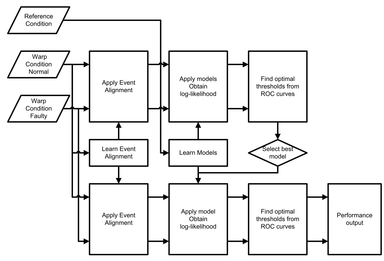
Acoustic Emission signals was conditioned using Root Mean Square and sampled with 2048 points per revolution. The data set consist of real signals acquired under piecewise increasing load (25, 50 and 75% load). Data was acquired 10 seconds every minute for 6 hours, giving 2227 full cycle examples.

Changing the load is a normal operating change that the system needs to be invariant to. The changes that we want to detect is the: a) shut down of lubricating oil during 25% load, b) the temporary fault during 50% load (believed to originate from the water brake) and c) the re-enabling of lubricating oil during the 75% load.

Acoustic Emission

is ultrasonic stress waves appearing on the surface of the materials. The increased damping (in the material), as opposed to vibration signals, improves on the signal to noise ratio, e.g., making cylinder crosstalk minimal.

Detection models

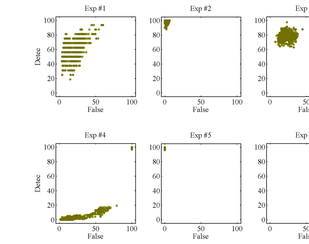


Principal Component Analysis Model

$$X_{\mathcal{T}} = [\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_N] \qquad d \times N \quad (4)$$
$$X_{\mathcal{T}} = U \cdot \Lambda V^{\top} = UPC \qquad d \times N(k) \cdot N(k) \times N \quad (5)$$
$$s_k = U_k^{\top} \tilde{x} \quad (6)$$
$$\tilde{x} = U_k s_k + \varepsilon \quad (7)$$
$$\theta = \{\Lambda_k, \sigma_\varepsilon^2\} \quad (8)$$
$$p(\tilde{x}|\theta, k) = p(s_k|\theta)\, p(\varepsilon|\theta) \quad (9)$$
$$\mathcal{L} = \log p(s_k|\theta) + \log p(\varepsilon|\theta) \quad (10)$$

From a set of N centered normal examples $\tilde{x}$ we build the training matrix X. From this training matrix we estimate a principal component matrix PC, and a projection matrix U through the Singular Value Decomposition (SVD). The number of principal components k is controlled by using the first k columns of U and (k) rows of PC. When applying the PCA model to new examples we multiply with the first k transposed columns of U from the left and obtain s_k plus the noise ε, as the remaining $d-k$ components that span a Gaussian noise space.

It follows directly from the properties of the SVD that the principal components of Normal Condition (NC) examples follow a multivariate zero mean Gaussian with covariance Λ_k (using the first k columns and rows of Λ).

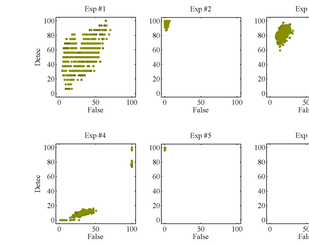


Gaussian Process Model

As an alternative to the PCA subspace model we can perform modeling directly in the observed domain. Through Gaussian Process (GP) modeling we obtain a measure of how much an example deviates from the reference condition. From Gibbs and MacKay we have (with interchanged t and x relative to Gibbs MacKay notation).

$$Q(t_i, t_j) = \theta_2^2 \exp\left(-\frac{(t_i - t_j)^2}{2\theta_1^2}\right) + \theta_3^2 \delta(t_i, t_j) \quad (11)$$
$$p(\tilde{x}|\mathbf{Q}, \mathbf{t}) = \frac{1}{Z} \exp\left\{-\frac{1}{2}\tilde{x}^{\top}\mathbf{Q}^{-1}\tilde{x}\right\} \quad (12)$$
$$\mathcal{L} = \frac{1}{2}\log|\mathbf{Q}| + \frac{1}{2}\tilde{x}^{\top}\mathbf{Q}^{-1}\tilde{x} \quad (13)$$
$$\widehat{\theta}_{n_\theta} = \arg\min_{\theta}\left\{\frac{1}{2}\log|\mathbf{Q}| + \frac{1}{2}\tilde{x}_{n_\theta}^{\top}\mathbf{Q}^{-1}\tilde{x}_{n_\theta}\right\} \quad (14)$$
$$\widehat{\theta} = \log\left(\frac{1}{card(N_\theta)}\sum_{n_\theta \in N_\theta}\exp(\widehat{\theta}_{n_\theta})\right) \quad (15)$$

where $\mathbf{t}$ is the vector of crank positions, and $\tilde{x}$ is the corresponding observed centered values. The covariance matrix Q is a function of the index vector t and the parameters θ, and the last term is the noise part. For each training examples $\tilde{x}_{n_\theta}, n_\theta \in N_\theta$ we train an independent GP with parameters $\{\theta_{n_\theta}\}$ through minimization of $\mathcal{L}$ using `minimize.m`. Finally we perform an average over the parameters obtained from different training examples to obtain the final model parameters. The exp and log in the averaging is due to the reparameterization that the minimizer uses to enforce non-negative constraints. The original implementation of the Gaussian Processes was due to Carl Rasmussen, and we have customized the input/output structure to fulfill our needs, e.g., allowing the training and use of the Q matrix.



Experiments and results

One model was trained under the reference condition: normal 25% load. This condition model is used on the warped examples. The experiment and numbers are *not* identical with the previous figures. However similar results can be seen.

Exp	Normal data	Faulty data	PCA	Detec/	False	# Comp	GP	Detec/	False
1	25 % load, lube oil on	25 % load, lube oil off		95/	5	2		95/	5
2	75 % load, lube oil on	75 % load, lube oil off		95/	5	2		95/	5
3	50 % load, lube oil "on"	50% load, unstable speed		60-65/	20-30	3		80-90/	15-20
4	75 % load, lube oil on	25 % load, lube oil off		95/	5	2		95/	5
5	50 % load, lube oil "on"	25 % load, lube oil on		50-60/	65-80	36		80-95/	0-15
6	Mixed loads, lube oil on	Mixed loads, lube oil off		95/	5	2		95/	5

List of experiments. During experiment 3 and 5 examples acquired without lube oil was warped into the reference condition with lube oil. Condition Monitoring Performance. Detec/False denotes detection versus false alarms rate in percentage.

In experiment 1 no warping is applied, thus this is the stationary performance that we are targeting.

Conclusion and outlook

The experiments show that non-stationary condition monitoring is indeed possible. It is important to notice that the event alignment does not decrease the overall condition monitoring performance as the results obtained in experiment 2 and 4 are equal to the stationary results in experiment 1. Furthermore, the performance obtained using both mixed aligned and original data in experiment 6 is the same as in the individual experiments, indicating that the optimal rejection thresholds are fairly constant even with several warp conditions each having its own set of event alignment parameters. The conclusion is that non-stationary Condition Monitoring (CM) indeed can be obtained by extending a stationary Condition Monitoring System (CMS) with event alignment.

Future work will concentrate on refining the method to handle a larger range operation conditions, automatic detection of landmarks, and further investigations related to overfitting. In addition, we will evaluate whether the framework will allow for wear trending, which of course calls for new experiments involving much larger time scales. We will also pursue fast Gaussian Processes in order to avoid signal downsampling.

IEEE Workshop on Machine Learning for Signal Processing 2004
São Luís, Brasil, September, 2004

The work is supported by EU Competitive and Sustainable Growth Programme GRD2-2001-50014: AE-WATT
Data acquired and provided by MAN B&W Diesel A/S, Copenhagen
Discussions with Heriot-Watt University was greatly appreciated