



HAL
open science

An Exploring Study of Hidden Markov Model in Rolling Element Bearing Diagnostis

Ge Xin, Jerome Antoni, Nacer Hamzaoui

► **To cite this version:**

Ge Xin, Jerome Antoni, Nacer Hamzaoui. An Exploring Study of Hidden Markov Model in Rolling Element Bearing Diagnostis. *Surveillance* 8, Oct 2015, Roanne, France. hal-01285513

HAL Id: hal-01285513

<https://hal.science/hal-01285513>

Submitted on 9 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Exploring Study of Hidden Markov Model in Rolling Element Bearing Diagnostis

Ge Xin¹, Jerome Antoni¹, Nacer Hamzaoui¹

¹LVA- INSA de Lyon
25 bis, avenue Jean Capelle, 69621 Villeurbanne, France
{ge.xin}@insa-lyon.fr

Abstract

Rolling element bearing is a crucial component in rotating machinery, so the diagnosis of bearing fault has attracted a lot of attention in both scientific and application areas. The objective of fault diagnosis can be viewed as a separation of informative events concerning different types of fault in noisy measurements. Therefore, it is appealing to model the raw signal as a linear combination of few components with the prior knowledge and assumption. Hidden Markov model (HMM) is a probabilistic model of joint probability of a collection of random variables which represent the hidden states as state variables given the observation sequence. In this paper, an exploring study in rolling element bearing diagnostics based on HMM is investigated and a new fault separation scheme is proposed. We analyse the performance of the proposed scheme through numerical experiments and demonstrate its potentiality in industrial applications.

1 Introduction

As crucial components in rotating machinery, rolling element bearing have received constant research attention, in particular for their diagnostis [1]–[4]. Typically when a defect occurs during operation, it excites some resonances at a specific rate called ‘bearing characteristic frequencies’. The excitation, which is generally amplitude modulated, embodies all the information concerning different types of faults. In industrial applications, however, such transient signals are usually embedded in heavy background noise which jeopardizes their detection. For the sake of separating the fault signature from noisy measurements, a reasonable transformation should be devised to represent the fault signal and the noise, respectively.

A proper transformation is used to extract the fault property from the acquired raw signal, which enables the following step to achieve the diagnostic identification. In terms of transform domain, there exists several feature extraction techniques which can be divided into three main categories: time domain, frequency domain and time–frequency domain. Intuitively speaking, in terms of statistical signal processing, time domain and frequency domain analyses are both directly based on the stationary assumption, thus leading to scalar indicators (RMS value, crest factor, kurtosis factor, or some complex combination of them) or vectorial indicators (signal spectrum, cepstrum or envelope spectrum) [2, 5]. To the contrary, time–frequency domain reveals the instantaneous features jointly in time and frequency which is appealing for non–stationary signal analysis, e.g. Wigner–Ville distribution, wavelet transform, and high–order spectral analysis [6]–[8]. However, as it is proved by Antoni in [2, 4], rolling element bearing vibrations are random cyclostationary.

In our present work, a general algorithm is proposed to implement the decomposition of a vibration signal into two states: 1) a stationary component with a Gaussian probability density function and 2) a transient signal with a different Gaussian probability density function. The algorithm is inspired by the Gaussian mixture problem which has been well addressed in the Data Mining community: it uses the EM algorithm in each frequency band of a Short-Time Fourier Transform (STFT). The purpose is to capture the maximum-likelihood estimate of the parameters of an underlying distribution from a given data set and therefore better discriminate informative signals from noise. It provides a proper classification result for diagnostic identification. This algorithm is evaluated through simulations which validates the efficiency of the proposed classification scheme in detecting the transient signals without pre-selected parameters.

The rest of the paper is organized as follows. Section 2 introduces the mathematical description of the problem, the basic theory of HMM and its solving method, the EM algorithm. In section 3, a classification scheme based on HMM is proposed. Then, the performance of the proposed algorithm is evaluated through numerical experiments in section 4. The conclusion is made in section 5.

2 Mathematical model and HMM

2.1 Mathematical description

Our target is to extract the informative signal $x(t)$ embedded in noisy measurement $y(t)$. This simple model is described by the following equation:

$$y(t) = x(t) + n(t) \quad (1)$$

where $n(t)$ denotes an additive zero-mean stationary Gaussian noise with standard deviation σ .

As discussed before, time-frequency representation is apt to reveal the symptomatic property of the energy distribution to vary with pseudo-periodic intervals in the case of a fault (the mathematical model is demonstrated in Eq.5). The STFT is a common time-frequency representation that adds a time variable to the conventional Fourier spectrum in order to study the time-varying characteristics of signal. Therefore, the decomposition of vibration signal for damage detection is done by applying the STFT in order to capture the fault signature. Let $\{w[n]\}_{n=0}^{N_w-1}$ be a positive and smooth N_w -long data-window which shift R samples (from 1 to N) to truncate a segment of $x[n]$ at times $kR, \dots, kR+N_w-1$, and let f_b denote the frequency bin index. Then it is defined as:

$$STFT_x(k, f_b) = \sum_{n=kR}^{kR+N_w-1} w_k[n] \cdot x[n] \cdot e^{-j2\pi f_b n / N_w} \quad (2)$$

2.2 Hidden Markov model

A hidden Markov model is a probabilistic model of joint probability of a collection of random variables which represent the hidden states as state variables given the observation sequence. As a dual random process, a hidden Markov model has two kinds of stochastic variable, namely the hidden state and the observation vector. The state sequence is unobservable, but can be estimated by the observation sequence [9]–[11].

As mentioned before, vibration signals can be described as a linear combination of a random cyclostationary component and a Gaussian stationary noise. In the time-frequency domain, stationary noise is spread over the whole picture and the informative part is embodied in it. Hence Gaussian mixed model (GMM) can be employed to model the probability distribution of the observation. We assume that data Y is observed (assume it is i.i.d.) and X is latent random variables, meaning that it is hidden/unobserved. We assume that a complete data set exists $C = (X, Y)$ and also specify a joint density function:

$$p(c | \Theta) = p(x, y | \Theta) = p(x | y, \Theta) p(y | \Theta) \quad (3)$$

where Θ embodies the mean μ and covariance σ which belongs to the different states x (latent random variable). Suppose that we are given a training set $\{y^{(1)}, \dots, y^{(m)}\}$, $z^{(i)} \sim \text{Multinomial}(\phi)$ (where $\phi_j \geq 0$, $\sum_{j=1}^k \phi_j = 1$, and the parameter ϕ_j gives $p(z^{(i)} = j)$), and $x^{(i)} | z^{(i)} = j \sim N(\mu_j, \Sigma_j)$. Let k denote the number of values that the $z^{(i)}$'s can take on. The parameters of GMM are thus ϕ , μ and σ . To estimate them, the total likelihood of the measurement is evaluated:

$$\begin{aligned} l(\phi, \mu, \Sigma) &= \sum_{i=1}^m \log p(x^{(i)}; \phi, \mu, \Sigma) \\ &= \sum_{i=1}^m \log \sum_{z^{(i)}=1}^k p(x^{(i)} | z^{(i)}; \mu, \Sigma) p(z^{(i)}; \phi) \end{aligned} \quad (4)$$

2.3 EM algorithm

Any probability distribution could be described as a mixture of a few Gaussian distributions. To estimate these parameters of Gaussian densities, maximum-likelihood parameter estimation is employed. However, there exist latent variables which denote the probability of state. This is what will make the estimation problem difficult.

The EM algorithm is an efficient iterative procedure to estimate the maximum-likelihood parameter in the presence of missing or hidden data. It consists of two steps: the E-step and the M-step which are presented below [12]. In the E-step, the missing data are estimated given the observed data and current estimate of the model parameters. In the M-step, the likelihood function is maximized under the assumption that the missing data are known. Convergence is assured since the algorithm is guaranteed to increase the likelihood at each iteration. Here are the EM iteration steps:

Repeat until convergence:

(E-step) For each i, j , set

$$w_j^{(i)} := p(z^{(i)} = j | x^{(i)}; \phi, \mu, \Sigma)$$

(M-step) Update the parameters:

$$\begin{aligned} \phi_j &:= \frac{1}{m} \sum_{i=1}^m w_j^{(i)} \\ \mu_j &:= \frac{\sum_{i=1}^m w_j^{(i)} x^{(i)}}{\sum_{i=1}^m w_j^{(i)}} \\ \Sigma_j &:= \frac{\sum_{i=1}^m w_j^{(i)} (x^{(i)} - \mu_j)(x^{(i)} - \mu_j)^T}{\sum_{i=1}^m w_j^{(i)}} \end{aligned}$$

3 Bearing fault separation based on GMM

In this study, we propose an informative signal separation scheme based on GMM. It provides a proper classification result for diagnostic identification. The details of the separation process are as follows:

1. Transform the raw signal into time-frequency domain using STFT (see Eq.2), so that the stationary component is taken in each segment corresponding to frequency axis and the non-stationary component is carried by the time axis.
2. Assume the measurement is a mixture of Gaussian distributions (see Eq.3). For each frequency bin, it consists of two states: a stationary component with a Gaussian probability density function on the one hand and a transient signal with a different Gaussian probability density function on the other hand.
3. The EM algorithm is employed to estimate the power spectrum of the two states and the probability of being in any state given the data.
4. Based on the estimated parameters and probability of states, the underlying information can be extracted (as shown in following section). An indicator of state is calculated and therefore the state decision is made. Hereafter, a vibration signal is decomposed into two components. Finally, the informative signal is extracted.

4 Verification on Numerical experiments

The following discussion is based on numerical simulations. First of all, a synthetic signal is generated which simulates the vibrations (m/s^2) excited by an outer-race fault as shown in Figure 1. To see this more clearly, let us recall the more realistic model for the vibration signal produced by a faulty bearing [2]:

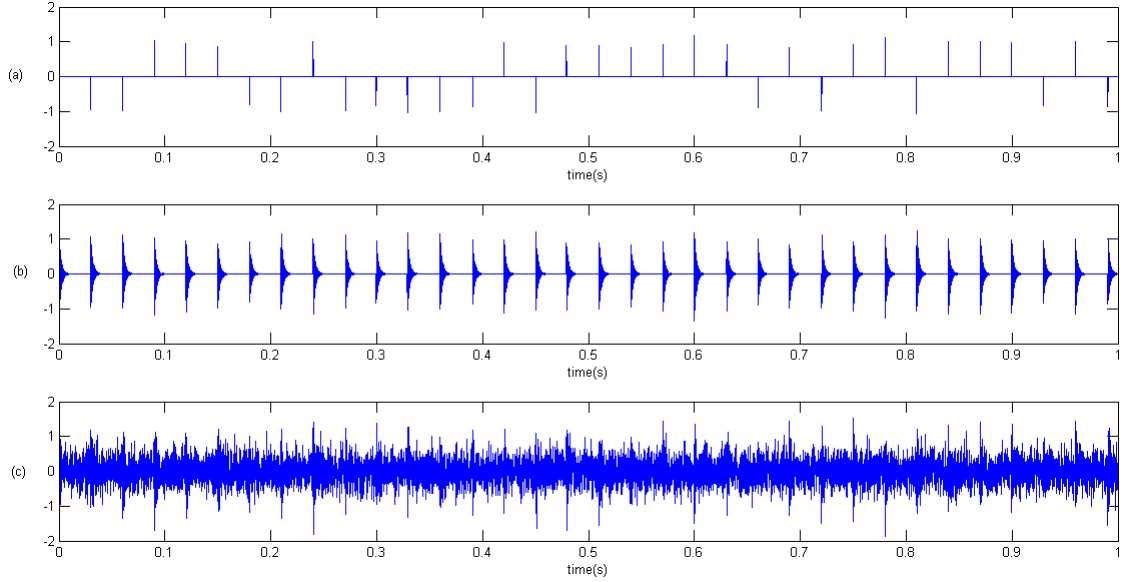


Figure 1: Synthesised vibration signal: (a) excitation, (b) response, (c) noisy signal with the NSR = 2.

$$y(t) = \sum_{i=-\infty}^{+\infty} h(t - iT - \tau_i)A_i + n(t) \quad (5)$$

where $\tau_i \sim (\mu_\tau = 0, \sigma_\tau)$ and $A_i \sim (\mu_A = 0, \sigma_A)$ account for the uncertainties on the arrival time (jitters) and on the magnitude of the i th transient, respectively. Therefore a noisy signal $y[n]$ (time duration is 1s, sampling frequency is 10^4 Hz) with length $N = 10^4$ is produced by Eq.5 as shown in Figure 1(c). The excitation with averaged interval ($T = 0.03$ s) and response of the system are presented in Figure 1(a)~(b), respectively. For the random part, $\sigma_\tau = 0.01T$ and $\sigma_A = 0.1$ are set as shown in Eq.5.

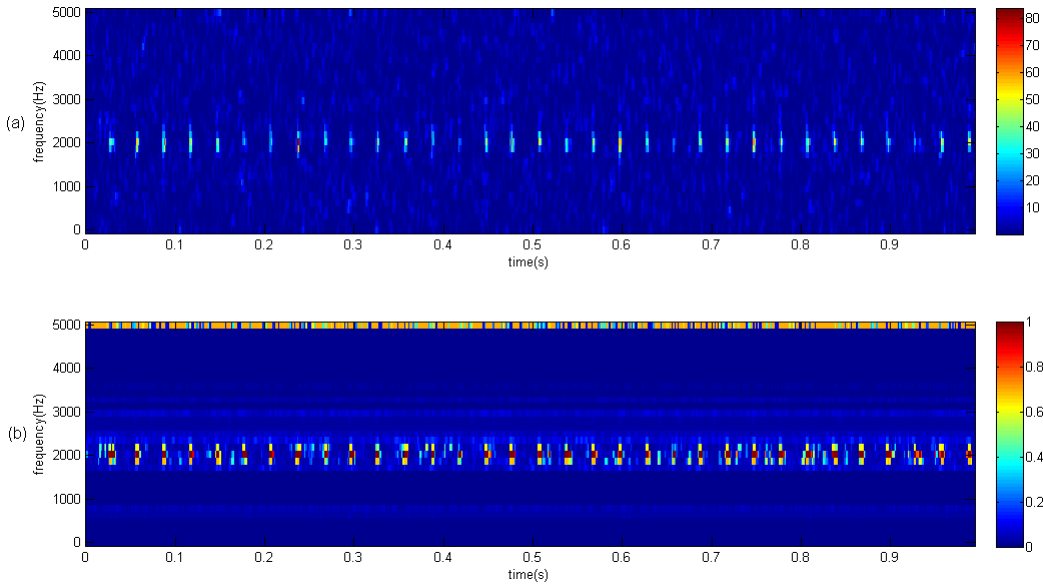


Figure 2: (a) Spectrum (b) $P(\text{state } 2|\text{data})$

Figure 2 shows the spectrum with $Nw = 2^6$ length hanning-window and estimated probability of being in state 2 given the data, respectively. $P(\text{state } 2|\text{data})$ is $w_2^{(i)} = p(z^{(i)} = 2 | x^{(i)}; \phi, \mu, \Sigma)$ normalized by differences between final likelihood at convergence of 1 and 2-component model along frequency-axis. As we can see in Figure 2, the probability of being in state 2 reveals the spectrum distribution of informative signal which

corresponds to the estimated spectrum of state 2 denoted in red dash line of Figure 3. The estimated spectrum is normalized between 0 and 1. There are two states which have a constant power spectrum corresponding to noise component and a non-flat power spectrum corresponding to transient signal.

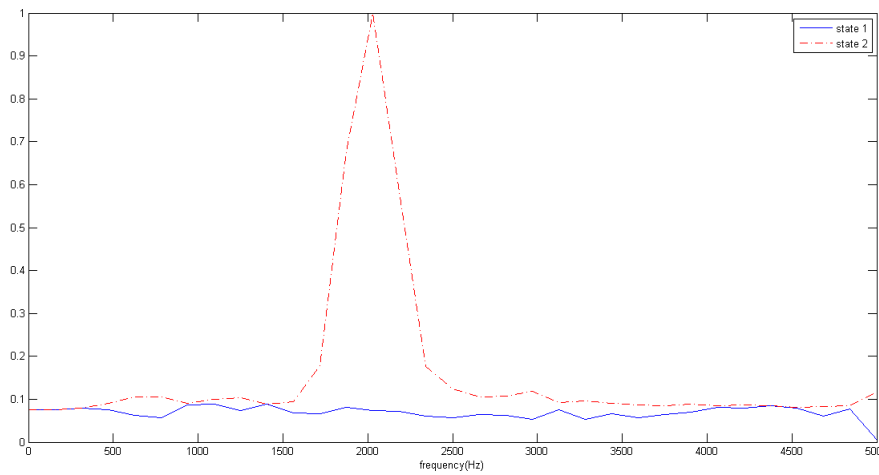


Figure 3: Estimated spectra in state 1 and 2

Therefore, based on the estimated parameters and probabilities of states, some underlying information can be extracted. As it is displayed in Figure 4(a), an indicator of state is calculated by the summation of estimated probabilities $P(\text{state 2}|\text{data})$ along the frequency-axis and a statistical threshold is set to select these segments carrying transient signals. The state decision, whether the indicator of state is above or below the threshold (corresponding to 1 or 0 respectively), is made as shown in Figure 4(b).

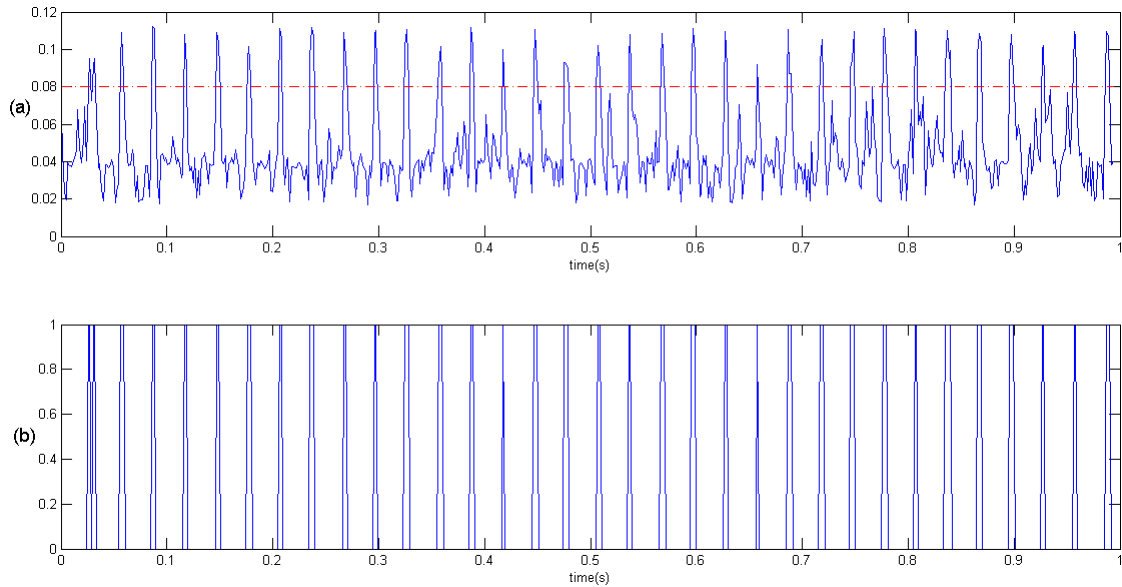


Figure 4: (a) State indicator with statistical threshold in red dash line and (b) State decision

The state decision consists of 0 and 1. So it implements the decomposition of a vibration signal into two states: a stationary component representing noise and a transient signal contaminated by noise. As we assume the measurement is recorded in a heavy noise, the influence of high NSR on the separated transient component should be taken into account. A simple way is to filter the spectrum with revised estimation of spectra based on the state decision result. Depending on the two separated signals, the re-estimated spectra of state 1 and 2 are displayed in Figure 5, which shows smoother spectra (especially in state 1 corresponding to noise component) compared with Figure 3.

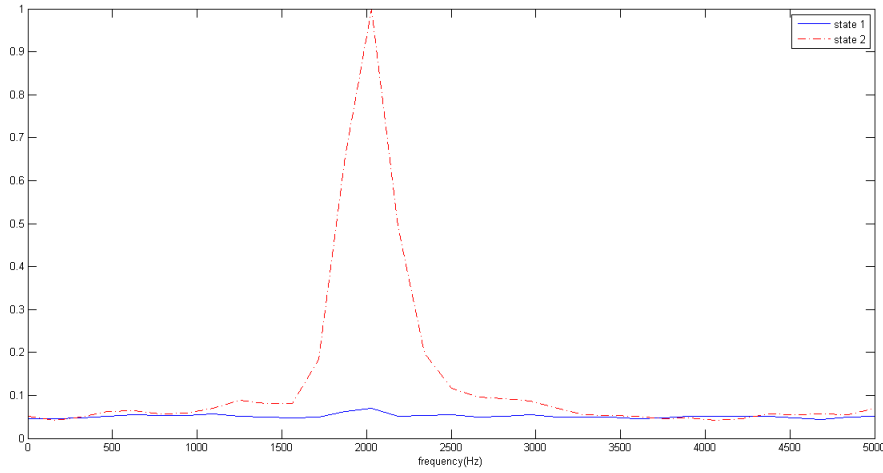


Figure 5: Re-estimated spectra state 1 and 2

According to the previous procedure, Figure 6 displays the separated signals, reference and noisy measurement in time domain which demonstrates the good performance of the proposed fault separation scheme. Meanwhile, it proves the validity of the estimated parameters and probability of states. It results in the significant extraction of underlying informations.

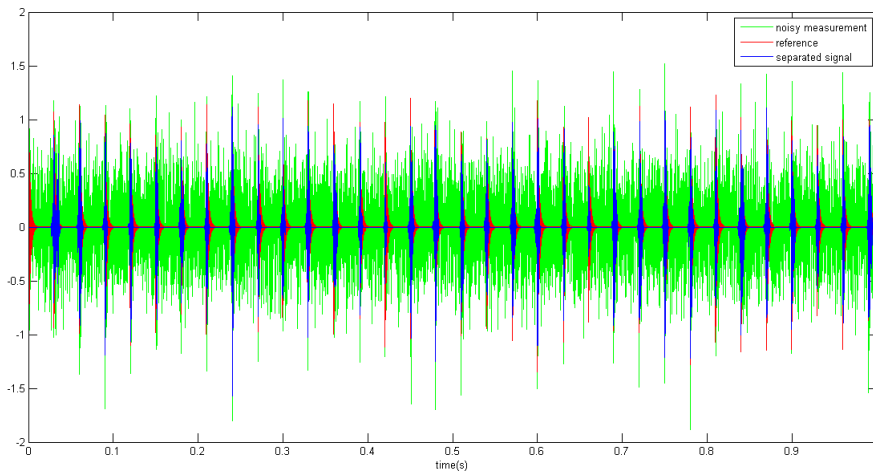


Figure 6: Noisy measurement, reference and separated signal

5 Conclusion

In this paper, an exploring study in rolling element bearing diagnosis based on HMM is investigated and a new fault separation scheme is proposed. Though in numerical experiments, the validity of the estimated parameters and probability of states is demonstrated. The indicator of state results in a state decision which decomposes the noisy measurement into two components based on different Gaussian distributions. Therefore the informative signal is discriminated from noise. It provides a proper classification result for diagnostic identification. What's more, it reveals the underlying information from a given data set without pre-selected parameter and has a potentiality in industrial applications.

References

- [1] R.B. Randall, J. Antoni, *Rolling element bearing diagnostics—a tutorial*, Mechanical Systems and Signal Processing, vol.25, 2011, pp.485-520.
- [2] J. Antoni, *Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions*, Journal of Sound and Vibration, vol.304, 2007, pp.497-529.
- [3] L. Yu, J. Antoni, Q. Leclere, *Combined Regularization Optimization For Separating Transient Signals From Strong Noise In Rolling Element Bearing Diagnosis*, Surveillance 7, Chartres, France, 2013 October, pp.1-13.
- [4] J. Antoni, F. Bonnardot, A. Raad, M. El Badaoui, *Cyclostationary modelling of rotating machine vibration signals*, Mechanical systems and signal processing, vol.18, 2004, pp.1285-1314.
- [5] H. Tang, J. Chen, G. Dong, *Sparse representation based latent components analysis for machinery weak fault detection*, Mechanical Systems and Signal Processing, vol.46, 2014, pp.373-388.
- [6] Q. Meng, L. Qu, *Rotating machinery fault diagnosis using Wigner distribution*, Mechanical Systems and Signal Processing, vol.5, 1991, pp.155-166.
- [7] W. Xiao, J. Chen, G. Dong, Y. Zhou, Z. Wang, *A multichannel fusion approach based on coupled hidden Markov models for rolling element bearing fault diagnosis*, Journal of Mechanical Engineering Science, vol.226, 2012, pp.202-216.
- [8] V. Rai, A. Mohanty, *Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform*, Mechanical Systems and Signal Processing, vol.21, 2007, pp.2607-2615.
- [9] L. Rabiner, *A tutorial on hidden Markov models and selected applications in speech recognition*, Proceedings of the IEEE, vol.77, 1989, pp.257-286.
- [10] T. Boutros, M. Liang, *Detection and diagnosis of bearing and cutting tool faults using hidden Markov models*, Mechanical Systems and Signal Processing, vol.25, 2011, pp.2102-2124.
- [11] Q. Miao, V. Makis, *Condition monitoring and classification of rotating machinery using wavelets and hidden Markov models*, Mechanical Systems and Signal Processing, vol.21, 2007, pp.840–855.
- [12] J.A. Bilmes, *A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models*, International Computer Science Institute, vol.4, 1998, pp.1-13.