Manufacturing process monitoring using two-step feature selection and classifier fusion

by

Juil Yum

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in The University of Michigan
2012

Doctoral Committee:

Professor Elijah Kannatey-Asibu Jr., Chair
Professor Judy Jin
Professor Jwo Pan
Professor Albert Shih
Research Investigator Tae Hyung Kim

To my wife and parents

Acknowledgements

I express my sincere gratitude and respect to my supervisor, professor Elijah Kannatey-Asibu Jr. for his great inspiration and gentle guidance during my doctoral study in the University of Michigan, Ann Arbor.

I wish to extend my gratitude to all my friends and officemates for their friendship and encouragement during my stay at the University of Michigan. It was also nice to attend a Friday Tennis Club every week to manage my physical strength as well as my mental healing.

Finally, I would like to give my special thanks to my lovely wife, Maeran Uhm, for her support to keep me motivated despite of her difficult situation. And the pray from parents (Taewoon Yum and Soonyi Jung) and parents-in-law (Dongseob Uhm and Jinsoon Park) kept me awaken during my doctoral program with their endless support. I owe much to express in words.

Table of Contents

List of Figures

Figure

List of Tables

Table

List of Symbols

M_i	= number of patterns in class S_i
S_i	= state of nature for class i
N	= number of classes
x_{ji}	= $j^{th} pattern in class S_i$
$\mu_{\mathbf{i}}$	= mean of class i
$\mathbf{\Phi}_{xi}$, $\mathbf{\Phi}_{xs}$	= scatter within each class S_i and scatter between classes, respectively
$\mathbf{\Phi}_x$	= overall system covariance matrix
Q	= feature selection criterion
$\mathbf{y_i^t}$	= frequency feature set consisting of n amplitudes at frequency i, t = 1,2,3, ..., n
$\mathbf{Y}$	= feature set ranked by Class Mean Scatter (CMS) criterion
$C(i,j)$	= correlation coefficient between i and j
$cov(x,y)$	= covariance between x and y
$\mathbf{G}$	= feature subset
$\mathbf{F}$	= new feature set from a modified redundancy and relevancy analysis
$\mathbf{A}$	= state transition matrix
$\mathbf{B}$	= observation symbol probability distribution
$\mathbf{x}$, $\mathbf{x_t}$	= observation sequence and observation symbol at time t, respectively
π	= initial state probability distribution
$p(\mathbf{x} \mid S_i)$	= likelihood of $\mathbf{x}$ given S_i
$P(S_i)$	= priori probability of class i
$N(\mu_i, \Sigma_i)$	= multivariate normal distribution
Σ_i	= covariance matrix of class i

L = likelihood function

$d_{i,j}$ = the class decision of observation j for classifier i

$\mathbf{D}$ = decision matrix

$\mathbf{W}$ = weighting vector

w_i = performance vector of classifier i

w_{f_i} = performance vector of classifier i with penalty

y_j = true class of observation j

$\delta(k, l)$ = 1 if k=l, 0 otherwise

$\mathbf{V}$ = voting matrix

I = identity matrix

i, j, m, n = index numbers $(0, 1, 2, 3 \dots)$ for analytical series solutions

c_i = damping coefficient of component i

k_i = stiffness coefficient of component i

m_i = mass of i

$\dot{x}$ = velocity

$\ddot{x}$ = acceleration

SA_i = regression model of estimated spectrum amplitude at frequency i

Abstract

Two-step feature selection and classifier fusion methods have been investigated for coroning and ultrasonic welding processes.

First, an acoustic emission (AE) based feature selection is studied for monitoring the coroning process. This involves significant data reduction since the AE signal requires a high sampling rate in order to capture useful information. Features, which are spectral components in the frequency domain, are processed using the class mean scatter (CMS) criterion. The features that are selected in the first step are then combined to reduce dimensions without averaging in the second step. With these features, classification is then performed and the results compared with those obtained using a conventional feature selection method.

Next, a classifier fusion method is developed to enhance the reliability and robustness in decision making. This is based on state performance weighting, which incorporates information on performance of classifiers for each state. A penalty voting concept is also investigated to further enhance classifier performance. Using equal weighting for each of the classifiers investigated, the overall classification rate achieved is 87.7%, while with state performance weighting, the classification rate improves to 98.5%. Using penalty voting further enhances the performance to 99.7%.

The signal processing techniques developed is then further used to investigate the feasibility of real time monitoring of ultrasonic weld quality using audible sound. A two degree-of-freedom model of the system is developed to help provide better understanding of the process characteristics. A series of experiments is also conducted to define the robust weld quality range for metal welding using the T-peel test. The relationship between weld quality and sound signals generated is then analyzed, and a strong correlation is obtained between the stiffness variation and spectral compo-

nents of the audible sound. The results show that a good weld is dominated by one frequency component in the audible sound range near 10 kHz, which is half the vibration frequency of 20 kHz, for the condition used, while that for a cold weld is characterized by two frequency components at 9 and 11 kHz. Over weld conditions do not generate unique frequency components.

Chapter I

Introduction

Modern manufacturing systems are highly automated and as such, require real time systems for monitoring the machine condition to prevent possible damage to parts and tool, and also ensure quality products. Depending on the type of manufacturing process, sensors need to be chosen carefully to obtain useful data about the process, and the resulting signals, which are usually combined with background noise, properly analyzed to extract useful information. The key is to extract the useful data from the overwhelming signals and determine the machine health. Two processes are the focus of this research, namely coroning and ultrasonic welding processes.

Coroning is a final gear finishing process before the gears are assembled into a transmission system, and can result in gear profile error which affects gear quality. It is believed to be the most accurate gear finishing process, resulting in a smooth surface and small transmission error (TE). TE has been found to be a major source of noise in automotive transmission system. Abbes et al. (2007) and Fuentes et al. (2006) reported that transmission error causes vibration in the transmission housing and eventually manifests itself as noise. Goto (2001) reported that dynamic stiffness of gear teeth has strong correlation with transmission error, while Mitchell and Daws (1982) and Krishnaswami et al. (2001) indicated that improvements in gear profile can reduce gear noise by 5-10 dB. The coroning process can significantly influence the TE, so it is important to monitor it, especially since tool wear associated with the coroning process is related to the product quality. Coroning is a multiple direction process (Schenk et al., 2003) and tool wear is a complex process. The signals of interest often appear in the high frequency ranges. Thus an accelerometer is not

adequate for tool wear monitoring. The major advantage of using acoustic emission (AE) to monitor tool condition is that its frequency range is much higher than that of machine vibrations and ambient noise (Li, 2002). Ravindra et al. (1997) also used AE signals as effective features for tool condition monitoring in cutting. Iwata and Moriwaki (1977) and Kim et al. (1999) reported that RMS AE voltages increase with tool wear. Jayakumar et al. (2005) were able to detect die wear, cracking, friction properties, etc. during grinding and forming processes such as drawing, punch stretching, blanking, and forging. Thus AE has the potential for being used to monitor the coroning process. Wear of the coroning tool was investigated and a monitoring system developed using an AE sensor for a high volume production line.

Ultrasonic welding has been increasingly investigated for joining battery components due to the clean environmental attractiveness and advantages of solid state welding. Resistance spot welding (RSW) is used widely in the automotive body industry (Tang et al., 2002; Cho et al., 2003a), but in general, highly conductive materials require higher current, shorter welding time, and more precise control of the process (Cho et al., 2003a), which requires a more massive RSW system. And it is extremely difficult to make a fusion weld between aluminum and copper, which are the principal battery tab materials. Since it is a solid state process, ultrasonic welding makes a good bond between dissimilar materials. Laser welding is also attractive since it is a non-contact process, but it is difficult to use on reflective materials. Ultrasonic welding is widely used for wire bonding (Tsujino et al., 1994; Shah et al., 2009; Lv and Han, 2008; Krzanowski, 1990) in the electronic chip packaging industry, metal sheets (Ueoka and Tsujino, 2002), thin foils (White, 2003), plastics (Grewell, 1999), glass (Wagner et al., 2003), and ceramics (Matsuoka, 1998). The mechanism of ultrasonic welding is not fully understood, but the first key mechanism is that the oxide layer breaks up by vibration (Kong et al., 2005; Siddiq and Ghassemieh, 2008; 2009) for enhancing the metal to metal bond. Thouless et al. (2006) and Cullen et al. (2008) used T-peel strength for ultrasonic welding. There are three main parameters that affect the weld strength, viz. amplitude, pressure, and welding time (Kong et al., 2003; 2005; Thouless et al., 2006; Hetrick et al., 2009) and they are related to

the weld strength by measuring the T-peel and lap shear loads. These parameters have been related to the optimal weldability zone of ultrasonic welding as shown in Fig. 1.1, (Joshi, 1971). In this research, a series of experiments was conducted to find the optimal welding zone for the materials used. The T-peel strength and failure modes were used as a weld quality reference. Weld quality monitoring was then applied using a microphone signal, focusing on the audible sound range since sound signals are generated during the process.

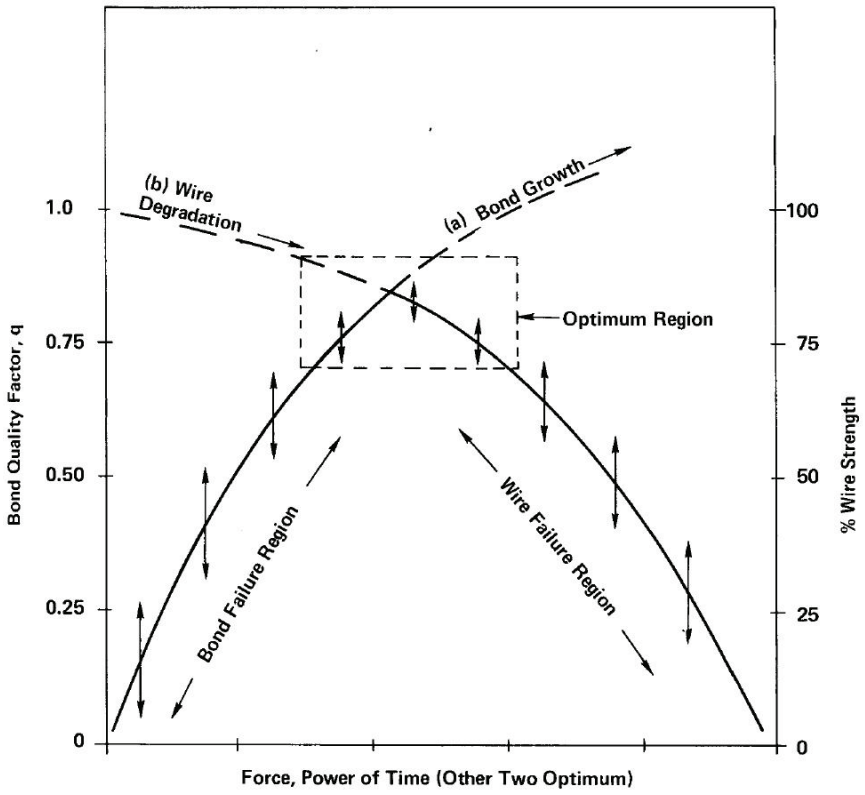


Fig. 1.1. Effect of principal process parameters on weld quality (Joshi, 1971).

Appropriate sensor selection leads to good feature candidates and provides an effective means to a successful manufacturing process monitoring. Two major approaches in feature selection are individual evaluation and subject evaluation. Individual evaluation, also known as feature weighting/ranking, assesses individual features and assigns them weights according to their degrees of relevance. A subject evaluation is often selected from the top of a ranking list, which approximates the set of relevant features. This approach is efficient for high-dimensional data (Guyon and Elisseeff,

2003; Blum and Langley, 1997; Hastie et al., 2001; Yu and Liu, 2004; Yu et al., 2008). The class mean scatter (CMS) criterion, an individual evaluation method, chooses the features which maximize the ratio of the scatter between classes to the scatter within classes. Sun et al. (2002) used this separability criterion (CMS) in laser welding and obtained a 100% classification. Wang (2008) also used this criterion for classification, but pointed out the challenges of feature selection for non-linearly separable data sets and noisy features. Along with feature selection, the number of features is also important. Fewer features result in faster computational time. By a ranking approach, the features could be correlated among themselves (Ding and Peng, 2003). Excessive features slow down the learning process and irrelevant or redundant features may confuse the learning algorithm (Yu et al., 2008). Highly correlated features can be grouped as one feature to compensate for the limitation of the averaging method to reduce the sensitivity to frequency changes. The two-step feature selection process will combine the features, giving it the same advantage as the averaging method.

In order to increase reliability and robustness, a fusion method that involves classifiers is also investigated. Sensor fusion has been used in various fields such as tool wear estimation (Bhattacharyya and Sengupta, 2009), underwater vehicle navigation (Nicosevici et al., 2004), vehicle fault diagnostics (Muldoon et al., 2002), and in optical implementation (Volfson, 2006). Luo and Kay (1990) introduced four different levels of fusion as signal-level, pixel-level, feature-level, and symbol-level fusion. Signal-level fusion means that the original signals are combined to create new information. Pixel-level is used in increasing image information by multiple images. Feature-level fusion can be used to increase the likelihood of selected features using multiple features and symbol-level fusion is the highest level of fusion that combines multi sensory information and feature-level fusion. The main purpose of fusion is to overcome the inaccuracy of single sensor reading and reduce sensitivity to noise. It has been adopted in many fields on classification and identification as a strong support method such as bagging and boosting (Bauer and Kohavi, 1999; Breiman, 1996). If this method is extended to classifier fusion, then decision will be more robust by giving more weight to the most reliable classifier. State performance weighted

classifier fusion and penalty voting are also studied and evaluated on coroning and ultrasonic welding processes.

In short, this research investigates on-line condition monitoring of coroning and ultrasonic welding processes, and involves the development of appropriate sensor selection along with feature selection and classification methods that incorporate fusion algorithms. Specifically, it focuses on acoustic emission (AE) and microphone signals, and the outcomes are described in Chapters 2, 3, and 4.

Chapter 2 covers the two-step feature selection method, which enhances classification performance via increasing the separability of the selected feature. This chapter has been accepted for publication in the ASME Journal of Manufacturing Science and Engineering, and presented at the proceedings of the ASME 2012 International Manufacturing Science and Engineering Conference, Notre Dame, Indiana, MSEC2012-7380.

Chapter 3 deals with classifier fusion, specifically state weighted classifier fusion and penalty voting. The proposed method is to increase reliability and robustness on decision making. This chapter has been accepted for publication in the International Journal of Advanced Manufacturing Technology.

Chapter 4 describes the feasibility of using audible sound features for real time weld quality monitoring during ultrasonic welding. It develops a two degree-of-freedom mechanical model that correlates audible sound generation with variation in the system stiffness. This chapter is under preparation for submission to the IEEE Sensors Journal.

Conclusions and future work are presented in Chapter 5.

Chapter II

A two-step feature selection method for monitoring tool wear, and its application to the coroning process

2.1 Introduction

A gear box in an automotive transmission system plays an important role in transmitting engine power to the driving wheels. The transmission gear box consists of a complex profile of gears that provide smooth driving and fuel efficiency. Gear fabrication normally involves hobbing, shaping, shaving, and honing. Among these processes, honing is the final gear finishing process and thus determines the final gear quality. Coroning was also developed as the final finishing process in gear fabrication. It is similar to, but more accurate than gear honing (Davis, 2005) and a coroning tool coated with diamond abrasive particles removes small amounts of material. Coroning is a complex multi dimensional metal removal process in gear manufacturing. It has a ring shape with teeth inside, which are coated with diamond that grinds hardened gear. The tool is engaged with a gear and then rotates under pressure. The pressure applied at the contact surface enables material to be removed and the gear geometry corrected. The high accuracy of the coroning process results in a smooth surface and small transmission error. Thus, it has been applied in volume production in gear manufacturing.

A coroning tool is shown in Fig.2.1.

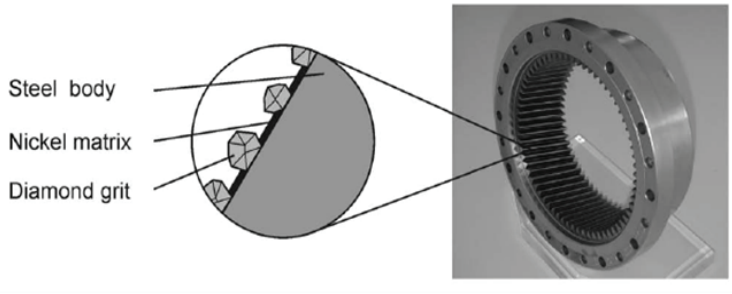


Fig. 2.1. Coroning tool (Schenk et al., 2003)

However, tool wear still exists since it involves material removal. The wear mechanism of a coroning tool is complex due to multi-directional motions and metal removal of gear teeth. The tool wear can result in gear profile error which would be a major source of noise in automotive transmission systems (Abbes et al., 2007; Fuentes et al., 2006), causing vibration in the transmission housing and eventually manifesting itself as noise. Thus, it is essential to monitor the condition of the coroning tool during production.

In this paper, an acoustic emission sensor (AE) is used to capture information on tool wear during gear finishing by the coroning process. AE has been widely used for tool condition monitoring in machining processes. Researchers have applied statistical features such as kurtosis, skew, and dominant frequencies from the AE signal to analyze the signals. Kannatey-Asibu Jr and Dornfeld (1982) and Dornfeld and Kannatey-Asibu (1980) obtained relationships between the cutting tool wear and root mean square (RMS), skew, and kurtosis of the AE signal from machining. Ravindra et al. (1997) also used AE signals as effective features for tool condition monitoring in cutting. Iwata and Moriwaki (1977) and Kim et al. (1999) reported that the RMS AE voltage increases with tool wear. However, those features may result in loss of important information, especially since for tool wear, there is significant useful information in the high frequency ranges. Since wear associated with the coroning process is not extensive, those features may not show any significant changes as wear progresses, unlike the observations of Kim et al. (1999). De Oliveira and Dornfeld (2001) reported that there are several factors which might influence the AE RMS level in a non predictable way in real production such as coolant flow, machine noise, and

electrical noise, unlike laboratory environments. Webster et al. (1994) and Aguiar et al. (2006) also indicated that the AE RMS might be insensitive due to the inherent averaging operation. However, the major advantage of using AE to monitor tool condition is that its frequency range is much higher than that of machine vibrations and ambient noise (Li, 2002). Hence, it is necessary to extract features that contain high frequencies to estimate the tool condition in the coroning process. In this paper, all frequency components obtained through AE were considered as initial features and then a feature selection algorithm was developed to remove redundant features and find the best feature set.

Various feature selection methods have been investigated for condition monitoring that involve mean and variance (Zhou et al., 2005; Jin, 2004; Zhou and Jin, 2005). Among the feature selection methods is one that is based on separability enhancement (Fukunaga, 1972). Using the separability criterion, the data is used to calculate the scatter within each class and the scatter between the classes, and the ratio of the two used to rank the features. Sun et al. (2002) used this separability criterion (class mean scatter (CMS)) in laser welding and obtained 100% classification. Wang (2008) also used this criterion for classification, but pointed out the challenges of feature selection for non-linearly separable data sets and noisy features. Along with feature selection, the number of features is also important. Fewer features result in faster computational time. However, increasing the number of features often improves classification results. Guyon and Elisseeff (2003) introduced single and multiple features for classification and reported that multiple features can improve the classification results when features are selected using the variable ranking method. Among selected features, there exist relevant and redundant features. More features can improve the classification performance, but excessive features may increase computation time and confuse the learning algorithm, the so called 'curse of dimensionality' (Hastie et al., 2001). Yu and Liu (2004) proposed relevance and redundancy analysis to eliminate redundant features. However, eliminating features may result in loss of important information. Rather than elimination, combinations of these features have the potential to overcome the averaging effect. Therefore a two-step feature selection using

CMS and modified relevance and redundant analysis is proposed for coroning tool wear monitoring in this paper. The proposed two-step feature selection method enables the best feature set to be selected by combining features, resulting in feature dimension reduction without averaging, and improved system performance.

The rest of this paper is organized as follows. The second section explains experimental setup which describes the sensor used, its attachment, and data collection. The third section provides a description of the proposed feature selection method for the coroning process. The fourth section gives descriptions of classifiers used. The fifth section discusses the results of applying the proposed method with classifiers described in the preceding section. The final section presents the conclusions.

2.2 Experiments

An AE sensor (Kistler AE 8152B22sp3) was used to monitor the coroning process and data was collected at a 2 MHz sampling rate. The raw data was then converted to the frequency domain for further analysis. Tool wear features were selected using the CMS criterion and used for classification.

Fig. 2.2 shows a schematic of the overall experimental setup. The signal from the AE sensor, which was attached to the coroning tool as shown in Fig. 2.3, was passed through a band pass filter which was set from 100 kHz to 1 MHz to filter out signals due to mechanical vibration. The signal was then further amplified with a gain of 10 (equivalent to 20 dB), and sampled at a rate of 2 MHz using the digitizer. A sample AE signal in the time domain is shown in Fig. 2.4.

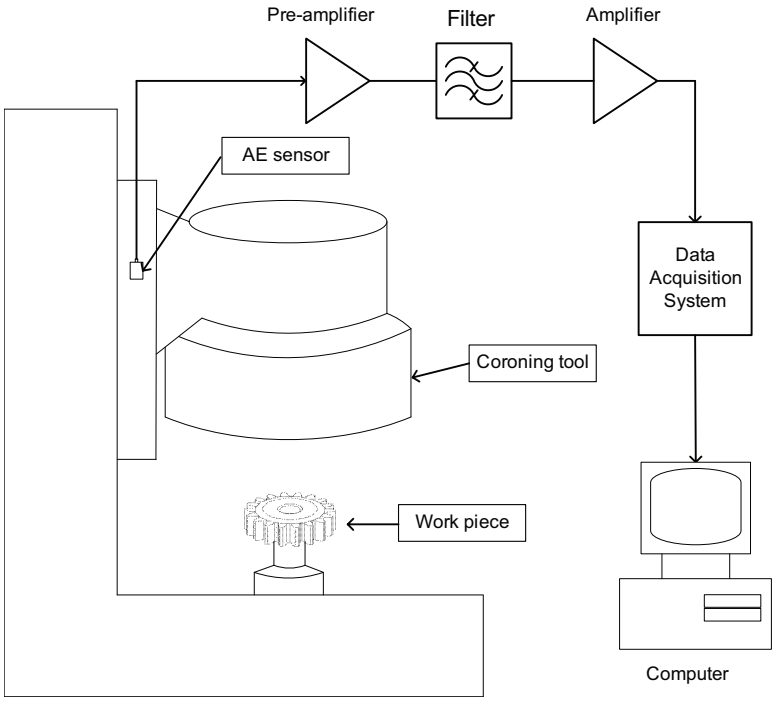


Fig. 2.2. Schematic of experimental setup

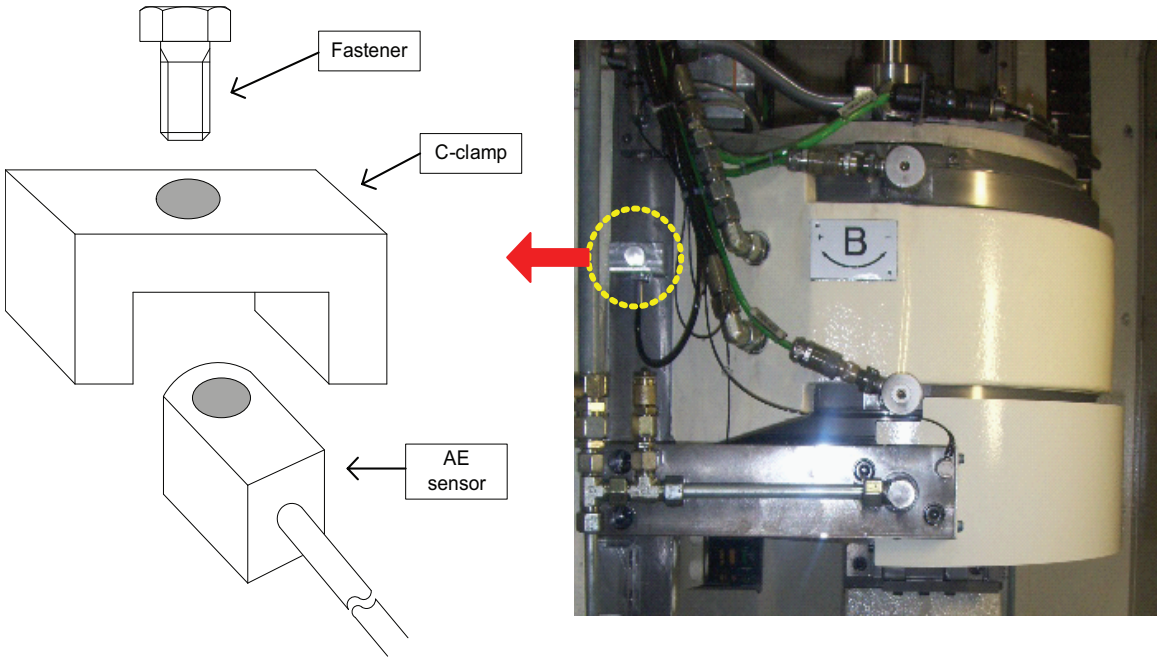


Fig. 2.3. Coroning machine and sensor attachment

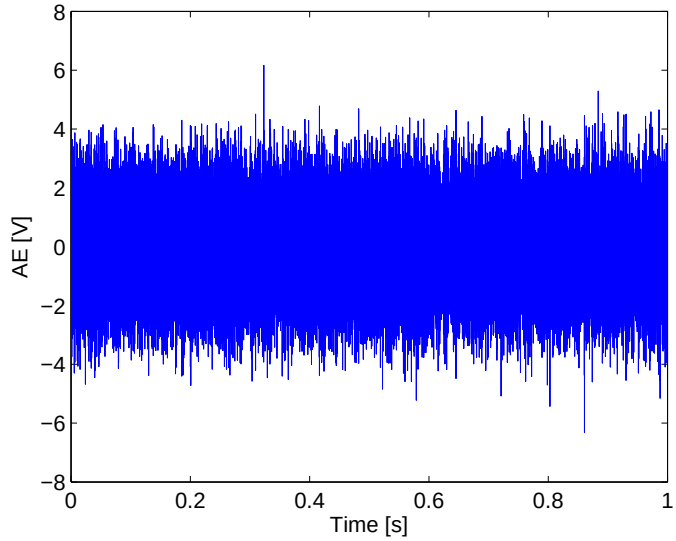


Fig. 2.4. AE signal in time domain

To identify the relationship between a feature and profile error, indirect measurement of the fabricated gear dimensions was also made for every 200th gear sample fabricated during production.

2.3 Two-step feature selection

The proposed two-step feature selection method is described in this section, and is based on CMS and modified relevance and redundancy analysis. The goal is to find the best possible features for monitoring the process. The CMS method chooses the features which maximize the ratio of the scatter between classes to the scatter within classes. A brief background on this method is presented next (Kannatey-Asibu Jr., 2009). Let the transformed data set be represented by $\mathbf{x}$ as shown in Fig. 2.5. The feature mean for each class is then given by

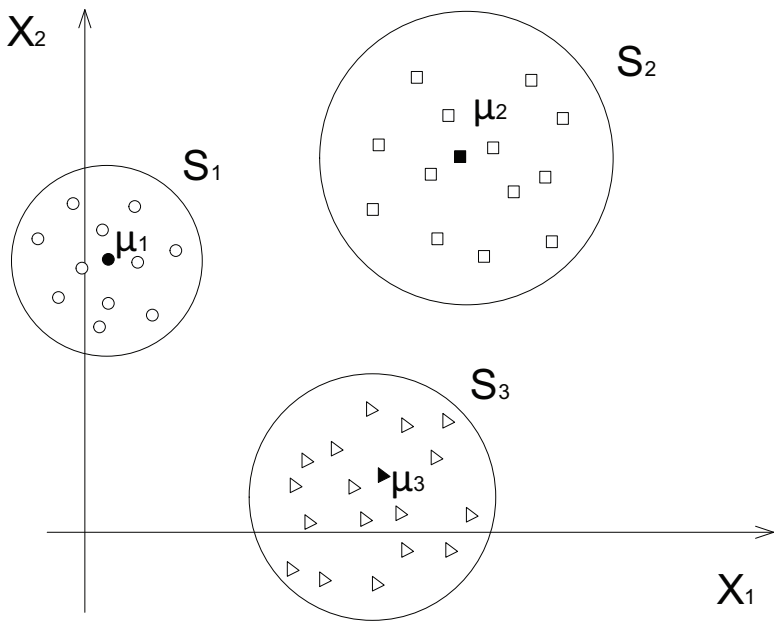


Fig. 2.5. Three classes in a two-feature space (Kannatey-Asibu Jr., 2009).

$$\mu_i = \frac{1}{M_i} \sum_{j=1}^{M_i} \mathbf{x}_{ji} \tag{2.1}$$

where M_i = the number of patterns in class S_i

$S_i = 1, 2, \ldots, \mathrm{N}$

$\mathbf{x}_{ji}$ is the j^{th} pattern in class S_i

N = number of classes.

The overall system mean is

$$\bar{\mu} = \sum_{i=1}^{N} p_i \mu_i \tag{2.2}$$

where p_i = a priori probability of class S_i.

The scatter within each class is obtained by calculating the covariance matrix of

$$\mathbf{\Phi}_{xi} = \frac{1}{M_i} \sum_{j=1}^{M_i} (\mathbf{x}_{ji} - \mu_i)(\mathbf{x}_{ji} - \mu_i)' \tag{2.3}$$

which leads to an overall system covariance matrix as

$$\mathbf{\Phi}_x = \sum_{i=1}^{N} p_i \mathbf{\Phi}_{xi} \tag{2.4}$$

The scatter between classes is

$$\mathbf{\Phi}_{xs} = \sum_{i=1}^{N} p_i (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})' \tag{2.5}$$

Finally, the feature selection criterion is defined as

$$\mathbf{Q} = \frac{\mathbf{\Phi}_{xs}(i,i)}{\mathbf{\Phi}_x(i,i)} \tag{2.6}$$

where $\mathbf{\Phi}_{xs}(i,i)$ and $\mathbf{\Phi}_x(i,i)$ are the i^{th} diagonal elements of the covariance matrices $\mathbf{\Phi}_{xs}$ and $\mathbf{\Phi}_x$, respectively. Thus, the higher values of $\mathbf{Q}$ indicate the stronger candidate features.

Among selected dominant frequency features from the CMS criterion, usually in the high frequency range for wear (Byington et al., 2003), there still exist redundant features. Using a ranking approach, the features that are correlated can be identified (Ding and Peng, 2003). A large number of features slows down the learning process and irrelevant or redundant features may confuse the learning algorithm (Yu et al., 2008).

Two major approaches in feature selection are individual evaluation and subject evaluation. Individual evaluation, also known as feature weighting/ranking, assesses individual features and assigns them weights according to their degrees of relevance. A subject evaluation is often selected from the top of a ranking list, which approximates the set of relevant features. This approach is efficient for high-dimensional data (Blum and Langley, 1997). The CMS method is basically an individual evaluation. In this paper, the subject evaluation is also applied by choosing top ranked CMS features before going through the next step.

AE signal levels for a faulty process are generally higher than for a normal process (Saravanan et al., 2006), which means that the energy level increases as the tool wears. In monitoring the coroning process using AE, it is important to consider the energy accumulation from different frequency ranges. Highly correlated features can

be grouped as one feature to compensate for the limitation of the averaging method to reduce the sensitivity to frequency changes (Yu et al., 2008). Unlike Yu and Liu (2004) who eliminated redundant features, the two-step feature selection process combines them, giving it the same advantage as the averaging method. Correlation coefficients of ranked features from the CMS method are calculated. Features are then grouped using the modified relevance and redundancy analysis as illustrated in Fig. 2.6 and described below. Thus, the two-step feature selection method uses not only subject evaluation but also modified relevance and redundancy feature selection to improve classification.

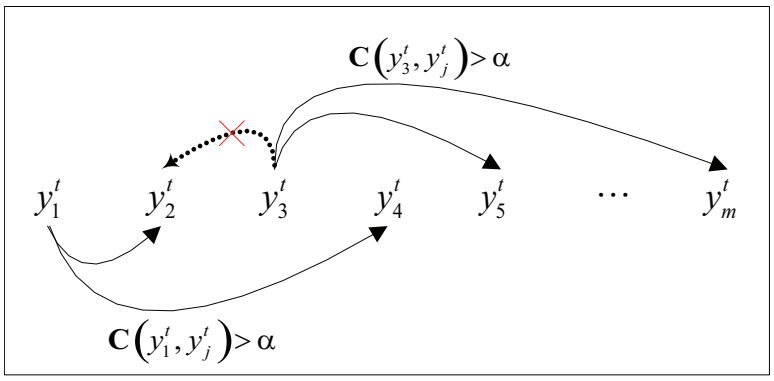


Fig. 2.6. Schematic illustration of new relevance and redundancy analysis feature selection process

To understand the proposed method, consider $\mathbf{y}_i^t = \{y(1) \quad y(2) \quad \ldots y(t)\}$ as a frequency feature set consisting of amplitudes at select frequencies, i = 1,2,. . . ,m, for n observations, $t = 1, 2, \ldots, n$ (Pandit and Wu, 1983) where m is the total number of frequency components. Using the CMS criterion, the features are ranked as $\mathbf{Y} = \{\mathbf{y}_1^t \quad \mathbf{y}_2^t \quad \ldots \mathbf{y}_m^t\}$. By calculating the correlation coefficient, $C(\mathbf{y}_i^t, \mathbf{y}_j^t)$, between CMS ranked features $\mathbf{y}_i^t$ and $\mathbf{y}_j^t$ with standard deviation σ, subsets of ranked CMS features are formed. $C(\mathbf{y}_i^t, \mathbf{y}_j^t)$ is given as:

$$\mathbf{C}(\mathbf{y_i^t}, \mathbf{y_j^t}) = \frac{cov(\mathbf{y}_i^t, \mathbf{y}_j^t)}{\sigma_{\mathbf{y}_i^t} \sigma_{\mathbf{y}_j^t}}, \tag{2.7}$$

where $cov(\mathbf{y}_i^t, \mathbf{y}_j^t)$ is the covariance between $\mathbf{y}_i^t$ and $\mathbf{y}_j^t$

i = 1, 2, . . . , m, j = 1, 2, . . . , m.

Define a subset as:

$$\mathbf{G}_1 = \{\mathbf{y}_1^t \quad \mathbf{y}_2^t \quad \mathbf{y}_4^t\}, \ldots, \mathbf{G}_M = \{\mathbf{y}_3^t \quad \mathbf{y}_5^t \quad \mathbf{y}_m^t\} \tag{2.8}$$

where $\mathbf{G}_k \cap \mathbf{G}_l = \emptyset \ \forall k, l = 1, 2, \ldots, M$ and $M < m$

$C(\mathbf{y}_i^t, \mathbf{y}_j^t) > \alpha$

α is a threshold value

$\mathbf{y}_1^t, \mathbf{y}_2^t, \mathbf{y}_4^t$ are correlated, and

$\mathbf{y}_3^t, \mathbf{y}_5^t, \mathbf{y}_m^t$ are also correlated.

The threshold value of α for the correlation coefficient, for this research, was set at a high of 0.95. The final features are obtained as the sum of all the elements in a subset:

$$\mathbf{F}_k = \sum_{\mathbf{y}_i^t \in \mathbf{G}_k} \mathbf{y}_i^t \tag{2.9}$$

This process continues until all ranked CMS features satisfying the threshold are accounted for. The two-step feature selection framework is illustrated in Fig. 2.7.

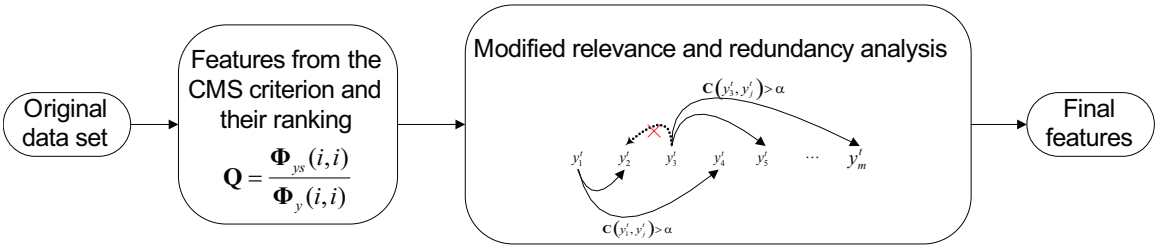


Fig. 2.7. Two-step feature selection method

To illustrate, consider six features that are randomly selected by ranking from the CMS criterion,
$\mathbf{Y} = \{\mathbf{y}_1^t \quad \mathbf{y}_2^t \quad \mathbf{y}_3^t \quad \mathbf{y}_4^t \quad \mathbf{y}_5^t \quad \mathbf{y}_6^t\}$. By calculating the correlation coefficients, $C(\mathbf{y}_i^t, \mathbf{y}_j^t)$, the resulting correlation matrix is as shown in Table 2.1.

Table 2.1. Correlation matrix obtained by calcualtion the correlation coefficients

	$\mathbf{y}_1^t$	$\mathbf{y}_2^t$	$\mathbf{y}_3^t$	$\mathbf{y}_4^t$	$\mathbf{y}_5^t$	$\mathbf{y}_6^t$
$\mathbf{y}_1^t$	1	0.97	0.89	0.98	0.77	0.76
$\mathbf{y}_2^t$	0.97	1	0.86	0.96	0.88	0.94
$\mathbf{y}_3^t$	0.89	0.86	1	0.96	0.99	0.96
$\mathbf{y}_4^t$	0.98	0.96	0.96	1	0.41	0.46
$\mathbf{y}_5^t$	0.77	0.88	0.99	0.41	1	0.87
$\mathbf{y}_6^t$	0.76	0.94	0.96	0.46	0.87	1

Setting the threshold that defines redundancy as $\alpha > 0.95$ correlation, the subsets become

$\mathbf{G}_1 = \{\mathbf{y}_1^t, \mathbf{y}_2^t, \mathbf{y}_4^t\}$ and $\mathbf{G}_2 = \{\mathbf{y}_3^t, \mathbf{y}_5^t, \mathbf{y}_6^t\}$ satisfying the criterion $\mathbf{G}_k \cap \mathbf{G}_l = \emptyset \quad \forall k, l.$ Then the final features become $\mathbf{F}_1 = \mathbf{y}_1^t + \mathbf{y}_2^t + \mathbf{y}_4^t$ and $\mathbf{F}_2 = \mathbf{y}_3^t + \mathbf{y}_5^t + \mathbf{y}_6^t$.

2.4 Classifiers

In order to verify the proposed feature selection algorithm, four classifiers, Hidden Markov model, minimum error rate Bayesian, Gaussian mixture model, and K-mean classifiers, are selected. In this section, these four classifiers are described and their monitoring performances for coroning tool wear discussed in the next section.

2.4.1 Hidden Markov model

We represent the compact form of the complete set of the model as $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$ with the following considerations:

- a finite set of N states, $\mathbf{S} = \{S_1,...,S_N\}$ and the state at time t as q_t
- a state transition probability matrix

  $\mathbf{A} = \{\mathbf{a}_{ij}\}_{N\times N}, \quad 1 \le i, j \le N$

  where $\mathbf{a}_{ij} = \mathrm{p}(q_{t+1} = S_j \mid q_t = S_i)$

  $1 \le i, j \le N, \quad 1 \le t \le n-1$
- an observation symbol probability distribution,

  $\mathbf{B} = \{b_j(\mathbf{x}_t)\}$

where $\mathbf{x}$ is an observation sequence,

$\mathbf{x} = \mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_t}$

$\mathbf{x}_t$ is an observation symbol at t,

$b_j(\mathbf{x}_t) = \mathrm{p}(\mathbf{x}_t \mid q_t = S_j), \quad 1 \leq j \leq N, \quad 1 \leq t \leq n.$

- an initial state probability distribution, $\pi = \{\pi_i\}$

  where $\pi_i = \mathrm{P}(q_1 = S_i), \quad 1 \leq i \leq N$

Rabiner (1989) gave examples of how to find the maximum likelihood state sequence, maximizing the probability of the observation sequence using the Baum-Welch method or Expectation-Maximization algorithm.

2.4.2 Minimum error rate Bayesian classification

Using the minimum-error-rate Bayesian rule (Duda et al., 2006), classification is based on the following condition:

$$\text{Decide } S_i \text{ if } \quad p(\mathbf{x} \mid S_i)P(S_i) > p(\mathbf{x} \mid S_j)P(S_j) \tag{2.10}$$

where S_i the state of nature for class i

$p(\mathbf{x} \mid S_i)$ = likelihood of $\mathbf{x}$ given S_i

$P(S_i)$ = a priori probability of class i.

Assuming a multivariate normal distribution,
$N(\mu_\mathbf{i}, \mathbf{\Phi_{xi}})$, for the class conditional probability density function $p(\mathbf{x} \mid S_i)P(S_i)$, and taking its logarithm, we obtain the discriminant function:

$$\begin{aligned} g_i(\mathbf{x}) &= -\frac{1}{2}(\mathbf{x} - \mu_\mathbf{i})^T \mathbf{\Phi}_{xi}^{-1}(\mathbf{x} - \mu_\mathbf{i}) - \frac{d}{2}\ln 2\pi \\ &\quad -\frac{1}{2}\ln|\mathbf{\Phi}_{xi}| + \ln P(S_i) \end{aligned} \tag{2.11}$$

where d = is the dimension of vector $\mathbf{x}$

$\mu_\mathbf{i}$ is the mean of class i

$\mathbf{\Phi}_{xi}$ is the covariance matrix of class i.

2.4.3 Gaussian mixture model

The Gaussian mixture model determines the Gaussian mixtures that maximize a likelihood of n samples of N different classes ($N < n$), $S = S_1, S_2, \ldots, S_N$ with initial guess of mean $K = \{\mu_1, \mu_2, \ldots, \mu_N\}$, assuming a Gaussian distribution (Duda et al., 2006). In a set of observations $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t)$, $\mathbf{x}_i$ is a d-dimensional real vector. By calculating the likelihood:

$$p(\mathbf{x} \mid S_i) = \sum_i p(\mathbf{x} \mid S_i, \mu_1, \mu_2, \ldots, \mu_N) P(S_i) \tag{2.12}$$

then, the likelihood function, L, becomes:

$$\begin{aligned} L &= p(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_t \mid S_i) \\ &= \prod_{i=1}^{n} \sum_i p(\mathbf{x} \mid S_i, \mu_2, \ldots, \mu_N) P(S_i) \end{aligned} \tag{2.13}$$

Now we maximize the likelihood function such that $\frac{\partial L}{\partial \mu_\mathbf{i}} = 0$ using the Expectation-Maximization (EM) algorithm.

2.4.4 K-means

K-means (Duda et al., 2006) is a clustering method which involves assigning a set of n samples to the closest mean vectors, μ_i, for a finite set of N classes. The initial number of N states has to be chosen properly. In a set of observations $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t)$, K-means minimizes the sum of squares from points to the class mean:

$$\arg\min_S \sum_{i=1}^{N} \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \mu_i\|^2 \tag{2.14}$$

where μ_i is the mean of each class, S_i

2.5 Results and discussion

AE signals obtained during the coroning process were analyzed in the frequency domain and the CMS method used to identify candidate features, which were then ranked, Fig. 2.8. As an example, the variation of the 976 kHz spectral component amplitude with time is shown in Fig. 2.9.

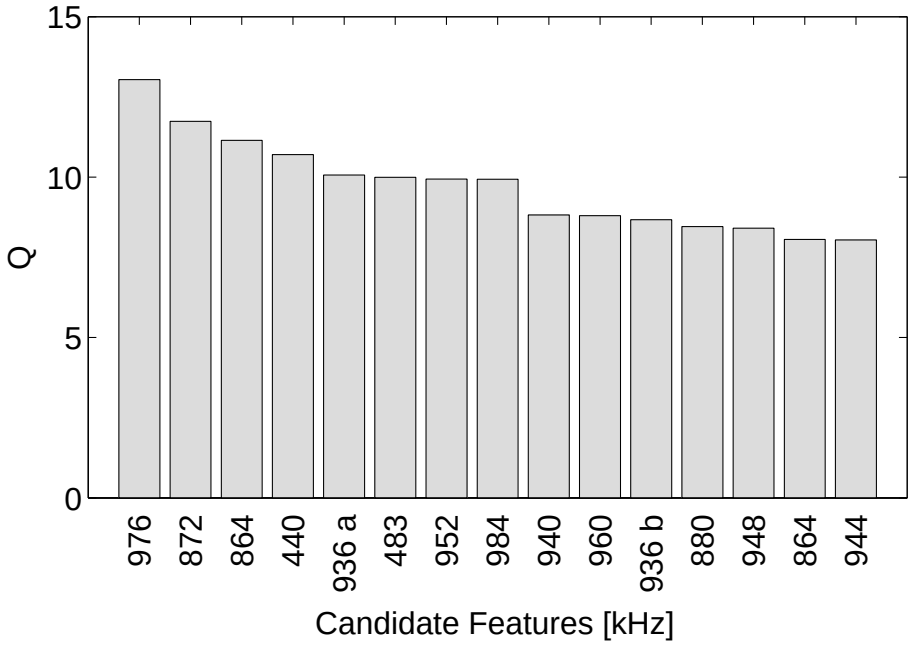


Fig. 2.8. Top 15 ranked CMS candidate frequency features. 936a and 936b kHz are distinguished using higher significant figures as 936015 and 936016 Hz respectively.

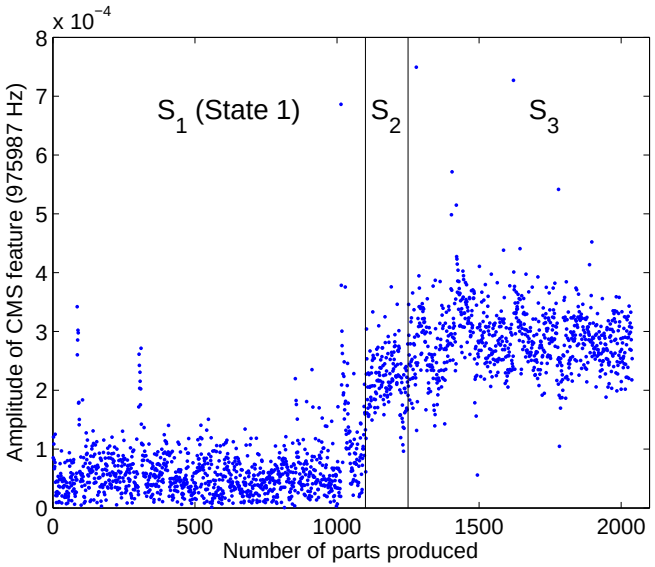


Fig. 2.9. Variation of the CMS feature at 976 kHz with tool wear (number of parts produced)

The time domain features such as RMS, skew, and kurtosis were calculated as shown in Fig. 2.10. Unlike previous research (Kim et al., 1999; De Oliveira and Dornfeld, 2001), the time domain features of the AE signal measured during the coroning process were observed to be stable during the cycle time. However, as shown in 3.2, two frequencies at 936 kHz and 483 kHz gave an indication of tool wear from indirect measure of part profile error.

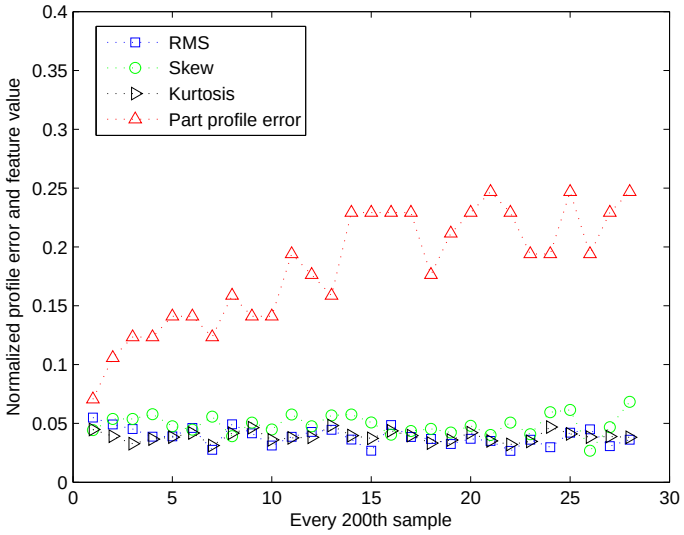


Fig. 2.10. Variation of measured profile error and statistical features (RMS, skew, kurtosis) with tool wear (number of gears produced)

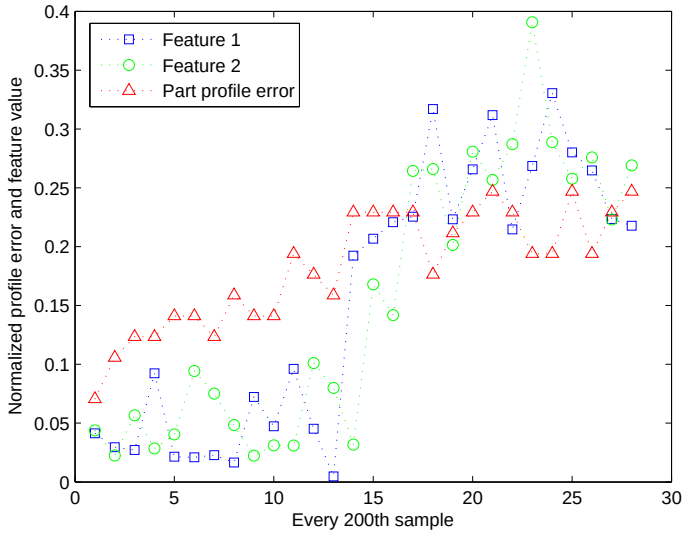


Fig. 2.11. Variation of measured profile error and the CMS features (normalized spectrum amplitude at 936 and 483 kHz) with tool wear (number of gears produced)

The ideal monitoring system would track the continuous progress of tool wear. Since wear of the cutting tool is a gradual process, doing so by classification would ideally require an infinite number of classes. However, as an initial step, we group the progress of tool wear into three states. State 1 represents the sharp tool, state 2 for a slightly worn tool, and state 3 for an extensively worn tool. Since state 2 shows rapid wear, this state is narrow in duration. Using fifteen features selected by the CMS criterion, the classification rates for the three states are summarized in Table 2.2 for different classifiers. The performance of the HMM and Bayesian classifiers for state 2 were comparably lower than the other classifiers, while the performance for states 1 and 3 were much higher than the GMM and K-mean classifiers. The overall performance among all classifiers was highest for the HMM and Bayesian methods.

Table 2.2. Classifier performance (%) using CMS-selected frequency features

	HMM	Bayesian	GMM	K-means
State 1	100.0	99.2	90.8	97.1
State 2	29.3	38.7	98.7	92.0
State 3	99.7	99.0	71.6	21.6
Overall	94.1	94.1	84.0	67.5

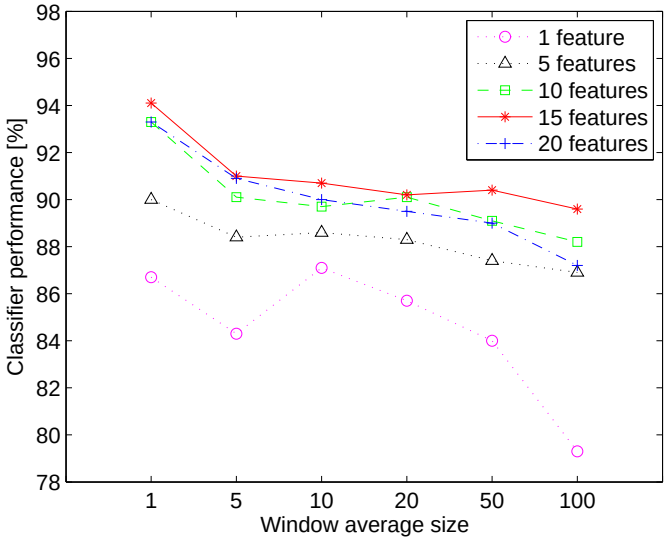


Fig. 2.12. Tool wear classification rates for different window sizes using the minimum error rate Bayesian method

As Table 2.2 shows, classification rates obtained using specific frequency features selected by the CMS criterion were generally less than 95%. HMM and Bayesian classifiers showed excellent capability of classifying states 1 and 3, but state 2 classification rate was low. Even though the overall classification rate reached 94.1%, the single frequency features may have a risk of lower classification rate if the feature frequency changes. To reduce the risk associated with using specific frequencies as features, band averaging was investigated. This involved averaging a contiguous set of frequency components. However, increasing the average window size resulted in a decreased classification rate as shown in Fig. 2.12.

Thus, conventional band averaging does not necessarily improve performance. Frequency band averaging also runs the risk of losing feature information. On the other hand, two-step feature selection does not result in information loss and it also has the potential of reducing frequency change effect on classification. Table 2.3 shows the classifier performances based on CMS features after removing redundant features. Other than the HMM classifier, the overall performances did not improve compared to the results obtained before removing the redundant features. This implies that the features removed may contain important tool wear information.

Table 2.3. Classifier performance (%) after removing redundant features

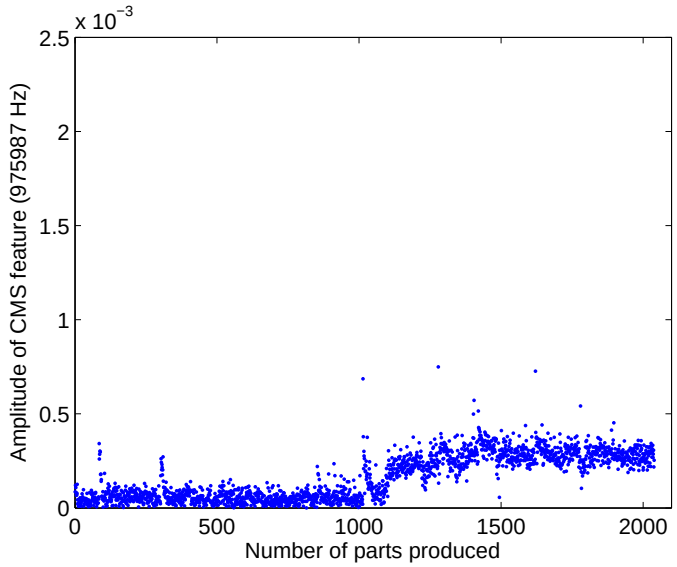
	HMM	Bayesian	GMM	K-means
State 1	97.7	97.7	91.6	95.7
State 2	72.7	45.3	93.3	90.7
State 3	95.3	94.9	66.2	21.6
Overall	94.9	92.8	82.3	66.7

Using the two-step feature selection with modified relevance and redundancy analysis method resulted in a significant improvement in the overall performance, Table 2.4. Also, the state performances that were low when CMS features were used increased with the two- step feature selection method as can be seen by comparing Tables 2.2-2.4.

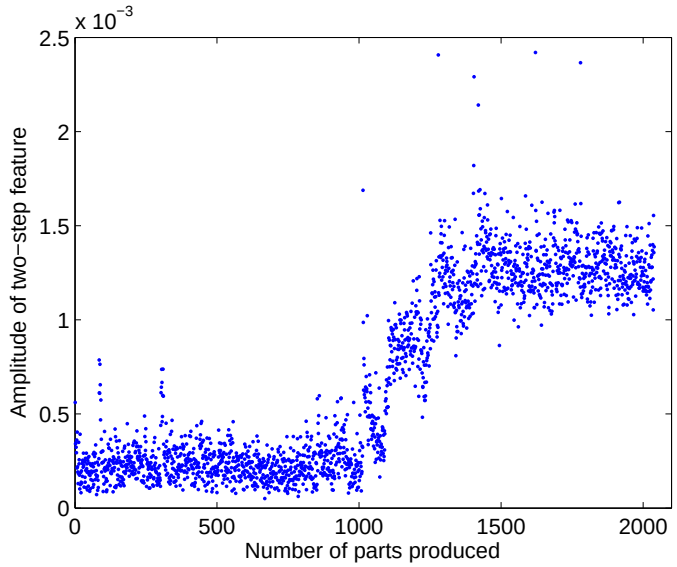
Table 2.4. Classifier performance (%) using two-step feature selection with modified relevance and redundancy analysis method

	HMM	Bayesian	GMM	K-means
State 1	98.2	98.5	92.6	92.5
State 2	91.3	68.7	97.3	88.0
State 3	99.7	98.1	73.4	95.9
Overall	98.3	96.2	85.5	93.5

Variation in features selected by the CMS and two-step feature selection methods shows that two-step feature selection creates a more separable data distribution than CMS, Fig. 2.13.



(a) CMS



(b) two-step

Fig. 2.13. Scatter plot of principal features from (a) CMS (976 kHz) (b) two-step feature selection

The classifier performances are summarized in Fig. 2.14, and the proposed features show significant improvement when compared with the other feature selection techniques, for each of the four classification methods. Two-step features resulted in the best performance for all the classifiers with a maximum classification rate of 98.3%.

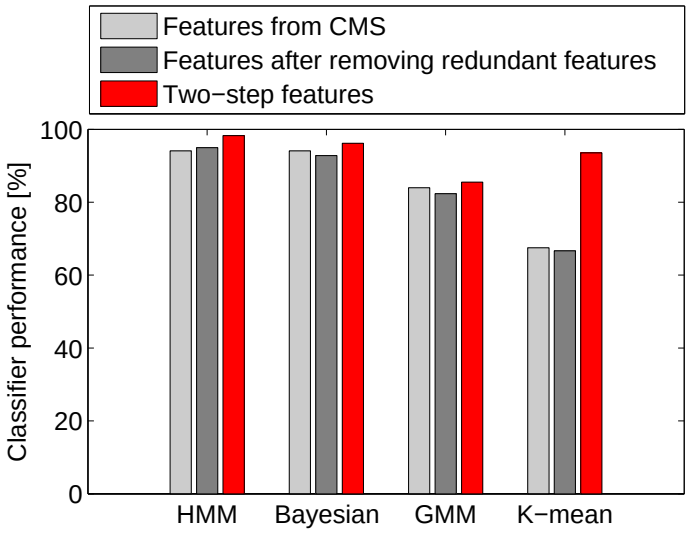


Fig. 2.14. Classification rate comparison between existing and proposed feature selection methods evaluated with four classification methods

2.6 Conclusions

A new feature selection method developed under this study resulted in 98.3% classification rate for monitoring tool wear during the coroning process.

Using conventional window averaging resulted in a decrease in the classification rate. The two-step feature selection method incorporates the advantage of window averaging, thereby reducing the sensitivity of the classification results to frequency changes. It improved classification significantly for the hidden Markov, minimum error rate quadratic classification based on the Bayesian rule, the Gaussian mixure, and K-means models. While the proposed feature selection method improves the performances for all the classifiers investigated, it is interesting to note that different classifiers have different performances.

Chapter III

Classifier fusion for acoustic emission based tool wear monitoring

3.1 Introduction

Coroning is a complex multi-dimensional metal removal process that is used for gear fabrication. Gears produced by polishing and finishing improve functional flank topology and reduce gear noise (Schenk et al., 2003). A coroning tool and system are shown in Fig.3.1. It has a ring shape with teeth inside, which are coated with diamond. The tool is engaged with a gear and then rotates under pressure. In addition to the tool rotation, there is also simultaneous grinding action parallel to the rotation axis. Thus, the coroning process ensures final gear quality before its assembly in a transmission box. It has been used for transmission manufacturing, especially in volume production (Schenk et al., 2003; Yum et al., 2009). Such a mass production process requires a real-time monitoring system to ensure quality and productivity. However, tool condition monitoring (TCM) for the coroning process has not been reported in the literature.

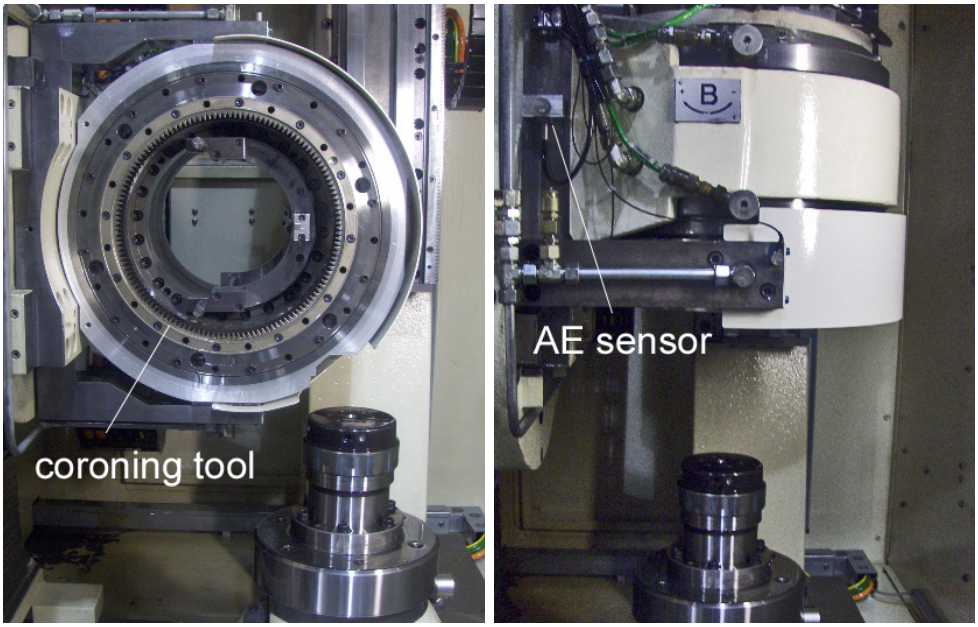


Fig. 3.1. Coroning machine and AE sensor attachment.

In TCM, the continuously varying nature of the process makes it difficult to estimate the tool condition. Thus, various classification methods have been investigated for wear monitoring (Leem et al., 1995; Kanthababu et al., 2008), including maximum likelihood, support vector machine (Donat et al., 2008; 2007), k-nearest neighbor (Li et al., 2010), hidden Markov model (Rabiner, 1989), and artificial neural network (Althoefer et al., 2005). However, no single classifier has been found to be suitable for all applications. To address this problem, a number of papers have proposed the use of multiple classifiers for improving monitoring performance (Chen et al., 2009a; Kuncheva, 2002; Tolba et al., 2010; Ruta and Gabrys, 2000; Sewell, 2011). There are selection and fusion methods for multiple classifiers. Similar to the class mean scatter (CMS) criterion (Kannatey-Asibu Jr., 2009), which ranks candidate features that describe the process, Roli et al. (2001) proposed a classifier selection method for choosing the best among a number of classifiers by evaluating the individual performances based on the so-called "overproduce and choose" concept. Ertunc and Loparo (2001) developed a decision fusion center algorithm for TCM in drilling. Three individual methods were designed and their outputs were used as weighting factors to calculate the global decision. Binsaeid et al. (2008) presented machine ensemble tech-

niques such as majority vote and generalized stacking ensembles during end milling. Yu (2011) proposed an artificial neural network (ANN) based tool wear prediction model in drilling operations. Several ANNs are generated by using modified Bagging method and then the best ANN selected. Fusion methods are applied not only in engineering but also in other fields such as medicine. Lederman et al. (2011) reported that the fusion method is more robust and consistent for identifying the risk of breast cancer. Those decision algorithms are normally based on a voting system, such as minimum, maximum, average, or majority voting (Kittler et al., 1998; Kittler and Alkoot, 2003). If the classifiers agree on class decision, as Petrakos et al. (2000) reported, then the fusion of classifiers will not achieve any better classification than using a single classifier. If they do not agree, then voting plays an important role. A voting method, however, still has challenges associated with 'tie votes'. A weighted voting is proposed in pattern recognition (Kahn Jr. et al., 2011; Hullermeier and Vanderlooy, 2010) in order to increase classifier performance and reduce 'tie votes' at the same time. But the performance of each classifier is not the same throughout the process due to the tool degradation and each classifier can perform differently at each state.

In this paper, the concept of decision fusion is extended to classifiers with the expectation that this will enhance the monitoring system performance by reducing the error associated with individual classifiers. Thus, the multi-classification fusion method is proposed here based on decision fusion. This is extended to incorporate state-performance weighting factors and penalty voting. As a first step, we facilitate classification by categorizing the entire tool wear regime into three states: sharp, slightly worn, and worn tools. An acoustic emission (AE) signal is measured and transformed into the frequency domain. Then, the CMS criterion (Emel and Kannatey-Asibu Jr., 1988) is applied to select the best features among all frequency components. The selected features and the tool condition are used for training and testing. The performances of individual classifiers and the performances at each state are used as weighting factors in the classifier fusion algorithm, which is then evaluated.

3.2 Experiments

The AE signal was used to monitor the coroning process and data was collected at a 2 MHz sampling rate. A total of 2039 samples were collected, and approximately half of them were used as a training set for the classifiers. The raw data was then converted to the frequency domain and frequency components were extracted as features. To identify a relationship between the tool condition and extracted features, the profile error of the fabricated gear was obtained by measuring profiles of gears reference to its tolerance using a coordinate measuring machine (CMM). Fig.3.2 shows variation of two select features, 976 kHz and 483 kHz, and the profile error with number of parts produced (every 200th gear sample produced by the coroning machine). The two features are among the CMS ranked features, and they were selected out of the lot to illustrate how the frequency components correlate with the profile error.

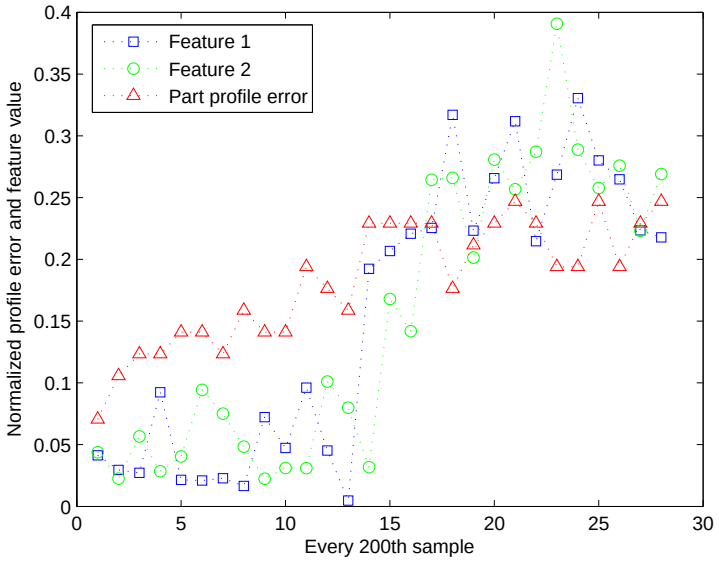


Fig. 3.2. Profile error measured and its comparison to select features (normalized spectrum amplitude at 976 and 483 kHz).

The tool condition was then categorized into three states using the profile error and features as shown in Fig. 3.3. State 1 represents the sharp tool, state 2 for a slightly worn tool, and state 3 for an extensively worn tool. Fig. 3.4 shows a two

dimensional feature space of these three states for two frequency features. However, finding the best features to improve the classifier performance is difficult. In this study, the CMS criterion proposed in (Emel and Kannatey-Asibu Jr., 1988) was used to select the best features. Among one million possible feature candidates, 15 features were selected using the CMS criterion.

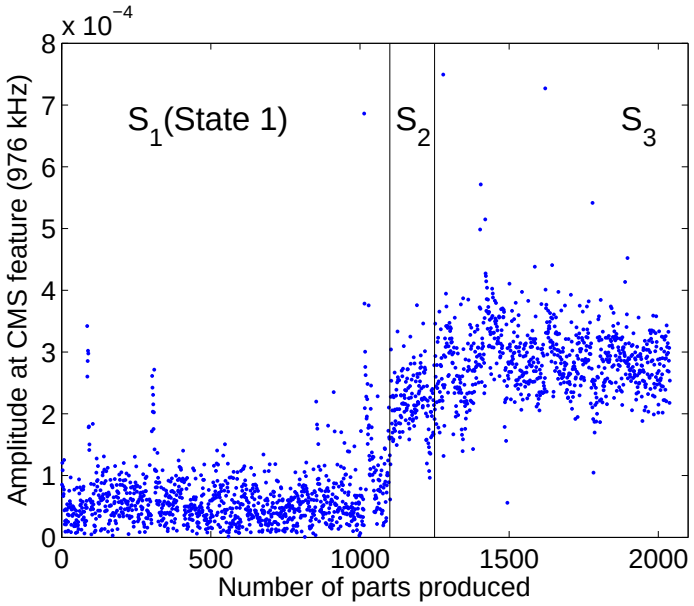


Fig. 3.3. Variation of the amplitude spectrum of AE signal at 976 kHz with tool wear (number of parts produced).

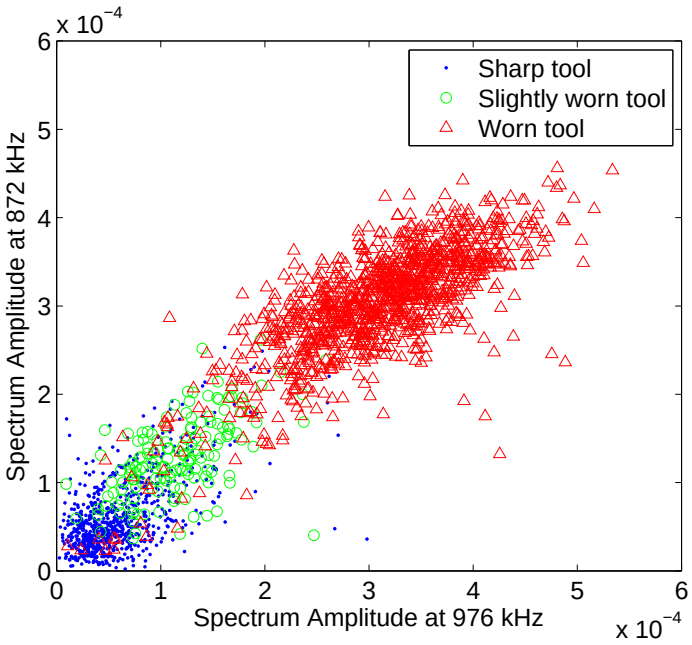


Fig. 3.4. The state of the cutting tool as represented by two frequency components in the feature space.

3.3 Classifiers

The proposed classifier fusion algorithm was based on four classifiers, Hidden Markov model (HMM), minimum error rate Bayesian (Bayesian), Gaussian mixture model (GMM), and K-means classifiers (K-means). In this section, these four classifiers are briefly described and their individual performances for coroning tool wear are presented.

3.3.1 Hidden Markov model

We represent the compact form of the model as $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$ with the following considerations:

(i) a finite set of N states, $\mathbf{S} = \{S_1,...,S_N\}$ and the state at time t as q_t

(ii) a state transition probability matrix

$\mathbf{A} = \{\mathbf{a}_{ij}\}_{N\times N}, \quad 1 \leq i, j \leq N$

where $\mathbf{a}_{ij} = \mathrm{p}(q_{t+1} = S_j \mid q_t = S_i)$

$1 \leq i, j \leq N, \quad 1 \leq t \leq n-1$

(iii) an observation symbol probability distribution,

$\mathbf{B} = \{b_j(\mathbf{x}_t)\}$

where $\mathbf{x}$ is an observation sequence,

$\mathbf{x} = \mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_t}$

$\mathbf{x}_t$ is an observation symbol at t,

$b_j(\mathbf{x}_t) = \mathrm{p}(\mathbf{x}_t \mid q_t = S_j), \quad 1 \leq j \leq N, \quad 1 \leq t \leq n.$

(iv) an initial state probability distribution, $\pi = \{\pi_i\}$

where $\pi_i = \mathrm{P}(q_1 = S_i), \quad 1 \leq i \leq N$

Rabiner (1989) gave examples of how to find the maximum likelihood state sequence, maximizing the probability of the observation sequence using the Baum-Welch method or Expectation-Maximization algorithm.

3.3.2 Minimum error rate Bayesian model

Using the minimum-error-rate Bayesian rule (Duda et al., 2006), classification is based on the following condition:

$$\text{Decide } S_i \text{ if } \quad p(\mathbf{x} \mid S_i)P(S_i) > p(\mathbf{x} \mid S_j)P(S_j) \tag{3.1}$$

where S_i is the state of nature for class i

$p(\mathbf{x} \mid S_i)$ is the likelihood of $\mathbf{x}$ given S_i

$P(S_i)$ is the priori probability of class i.

Assuming a multivariate normal distribution, $N(\mu_i, \Sigma_i)$, for the class conditional probability density function $p(\mathbf{x} \mid S_i)P(S_i)$, and taking its logarithm, we obtain the discriminant function:

$$\begin{aligned} g_i(\mathbf{x}) &= -\frac{1}{2}(\mathbf{x} - \mu_\mathbf{i})^T \Sigma_i^{-1}(\mathbf{x} - \mu_\mathbf{i}) - \frac{d}{2}\ln 2\pi \\ &\quad -\frac{1}{2}\ln|\Sigma_i| + \ln P(S_i) \end{aligned} \tag{3.2}$$

where d is the dimension of vector $\mathbf{x}$

$\mu_\mathbf{i}$ is the mean of class i

Σ_i is the covariance matrix of class i.

3.3.3 Gaussian mixture model

The Gaussian mixture model determines the Gaussian mixtures that maximize a likelihood of t samples of N different classes ($N < t$), $S = S_1, S_2, \ldots, S_N$ with initial guess of mean $K = \{\mu_1, \mu_2, \ldots, \mu_N\}$, assuming a Gaussian distribution (Duda et al., 2006). In a set of observations $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t)$, $\mathbf{x}_t$ is a d-dimensional real vector. By calculating the likelihood:

$$p(\mathbf{x_t} \mid S_i) = \sum_i p(\mathbf{x_t} \mid S_i, \mu_1, \mu_2, \ldots, \mu_N) P(S_i) \tag{3.3}$$

then, the likelihood function, L, becomes:

$$\begin{aligned} L &= p(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_t \mid S_i) \\ &= \prod_{j=1}^{t} \sum_i p(\mathbf{x_j} \mid S_i, \mu_2, \ldots, \mu_N) P(S_i) \end{aligned} \tag{3.4}$$

Now we maximize the likelihood function such that $\partial L / \partial \mu_\mathbf{i} = 0$ using the Expectation-Maximization (EM) algorithm.

3.3.4 K-means model

K-means (Duda et al., 2006) is a clustering method which involves assigning a set of t samples to the closest mean vectors, μ_i, for a finite set of N classes, which has N different means. The initial number of N states has to be chosen properly. In a set of observations $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t)$, K-means minimizes the sum of squares from points to the class mean:

$$\arg\min_S \sum_{i=1}^{N} \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \mu_i\|^2 \tag{3.5}$$

where μ_i is the mean of each class, S_i

3.3.5 Individual classifier performance

The overall performance of individual classifiers for monitoring continuous tool wear during the coroning process are shown in Fig. 3.5. The performances of the four classifiers (HMM, Bayesian, GMM, and K-means), are 94.1%, 94.1%, 84.0%, and 67.5%, respectively, Table 3.1. Most of the error for the HMM classifier came from state 2 as shown in Fig. 3.6.

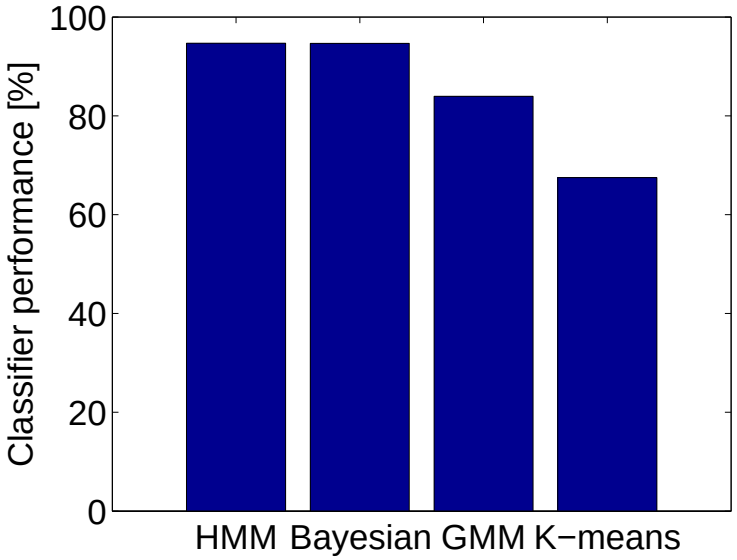


Fig. 3.5. Overall classification rates of individual classifiers for monitoring tool wear of the coroning process

Table 3.1. Single classifier performance (%) at each state.

	HMM	Bayesian	GMM	K-means
State 1	100.0	99.2	90.8	97.1
State 2	29.3	38.7	98.7	92.0
State 3	99.7	99.0	71.6	21.6
Overall	94.1	94.1	84.0	67.5

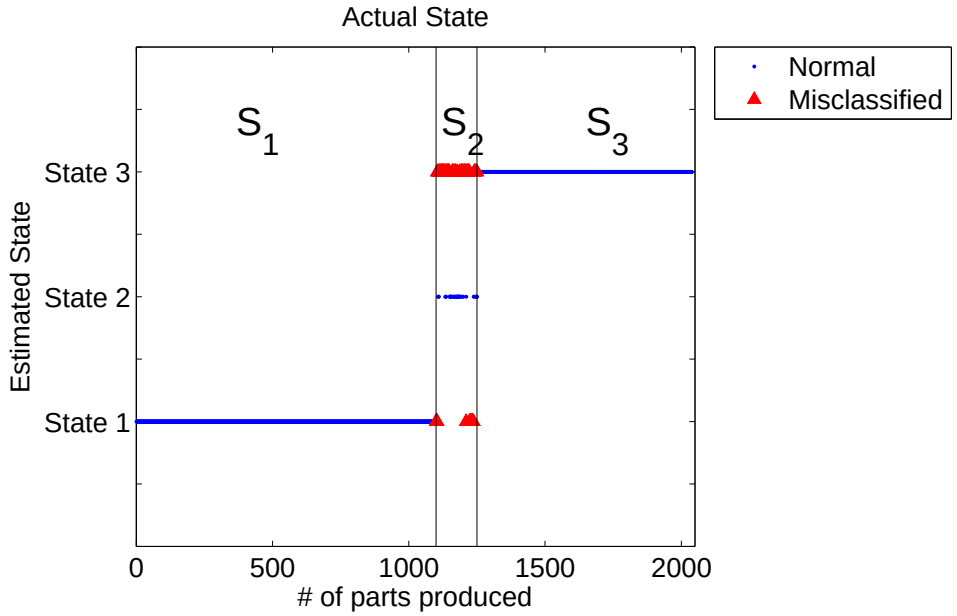


Fig. 3.6. HMM state estimation

3.4 Classifier Fusion

The classifier fusion procedure is shown schematically in Fig. 3.7, using the HMM, Bayesian, GMM, and K-means models.

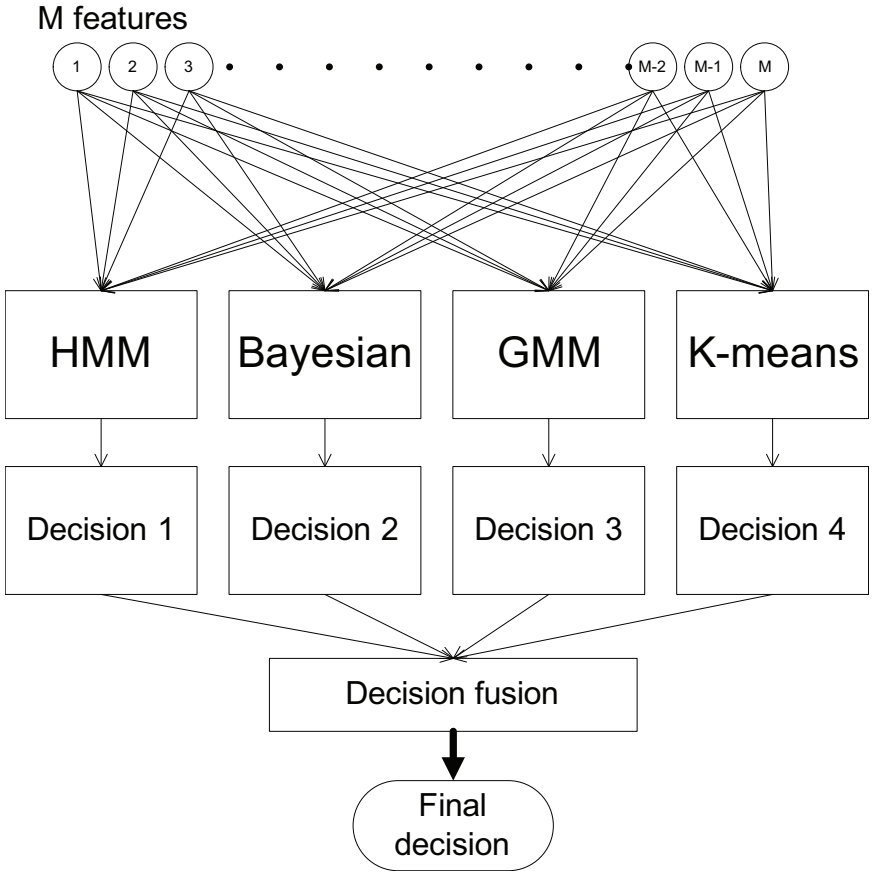


Fig. 3.7. Multi-classifier fusion

We first define a matrix **D** consisting of the class pattern (decision) associated with individual classifiers:

$$\mathbf{D} = \begin{bmatrix} d_{1,1} & d_{1,2} & d_{1,3} & \dots & d_{1,m} \\ d_{2,1} & d_{2,2} & d_{2,3} & \dots & d_{2,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n,1} & d_{n,2} & d_{n,3} & \dots & d_{n,m} \end{bmatrix}, \quad d_{i,j} \in \{1, 2, 3, \dots, N\} \tag{3.6}$$

where $d_{i,j}$: the class decision of observation j for classifier i

m: the number of observations throughout the life of the tool

n: the number of classifiers

N: the number of classes or states.

Now consider a weighting vector $\mathbf{W}$ that consists of individual classifier performances, w_i:

$$\mathbf{W} = [w_1 \quad w_2 \quad w_3 \quad \dots \quad w_n] \tag{3.7}$$

where w_i is the performance vector of classifier i.

w_i is calculated for m observations as follows:

$$w_i = \frac{\sum_{j=1}^{m} \delta(d_{i,j}, y_j)}{m}, \quad i = 1, 2, \dots, n \tag{3.8}$$

where y_j is the true class of observation j,

$j = 1, 2, 3, \dots, m$,

$$\delta(k, l) = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{otherwise} \end{cases}$$

Before the final decision takes place, let us define the voting matrix $\mathbf{V}$

$$\mathbf{V} = \begin{bmatrix} v(d_{1,1}) & v(d_{1,2}) & v(d_{1,3}) & \dots & v(d_{1,m}) \\ v(d_{2,1}) & v(d_{2,2}) & v(d_{2,3}) & \dots & v(d_{2,m}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v(d_{n,1}) & v(d_{n,2}) & v(d_{n,3}) & \dots & v(d_{n,m}) \end{bmatrix} \tag{3.9}$$

Where $v(i)$ is the i^{th} row of the identity matrix, $\mathbf{I}$, whose rank is determined by the number of classes of interest. For example, for a 3-class system (N=3), $d_{i,j}$ is then an element of $\{1, 2, 3\}$, and we have a 3x3 matrix $\mathbf{I}$. Elements of $\mathbf{D}$, $d_{i,j}$, are used to form the matrix $\mathbf{V}$. For example, $v(d_{i,j})$ becomes [1 0 0] from the first row of identity matrix, $\mathbf{I}$, when $d_{1,1} = 1$. Likewise $v(d_{i,j})$ becomes [0 1 0] when $d_{1,1} = 2$ and [0 0 1] when $d_{1,1} = 3$. Thus the voting matrix $\mathbf{V}$ is formed from the classifier, $d_{i,j}$. The final decision equation is then the matrix multiplication of the weighting vector $\mathbf{W}$ and

voting matrix $\mathbf{V}$:

$$\textit{Final decision} = \mathbf{FD}_j = \left[\arg \max_{column} \left[\mathbf{WV}(j)\right]\right] \tag{3.10}$$

where $V(j)$ is the j^{th} column of the voting matrix, $\mathbf{V}$

We now take an example to illustrate the influence of the weighting factor by comparing non-weighted and weighted decisions. Let us assume we have 4 classifiers, 5 observations, and 3 different states, which correspond to n=4, m=5, and $d_{i,j} \in \{1, 2, 3\} \quad \forall i, j$, with unity weight values and random class patterns. $\mathbf{W}$ and $\mathbf{D}$ then become:

$$\mathbf{W} = [w_1 \quad w_2 \quad w_3 \quad w_4] = [1 \quad 1 \quad 1 \quad 1]$$

$$\mathbf{D} = \begin{bmatrix} d_{1,1} & d_{1,2} & d_{1,3} & d_{1,4} & d_{1,5} \\ d_{2,1} & d_{2,2} & d_{2,3} & d_{2,4} & d_{2,5} \\ d_{3,1} & d_{3,2} & d_{3,3} & d_{3,4} & d_{3,5} \\ d_{4,1} & d_{4,2} & d_{4,3} & d_{4,4} & d_{4,5} \end{bmatrix} = \begin{bmatrix} 2 & 1 & 2 & 3 & 3 \\ 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 3 & 2 \\ 1 & 2 & 2 & 3 & 2 \end{bmatrix}$$

That results in the following voting matrix, $\mathbf{V}$:

$$\mathbf{V} = \left[\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}\right]$$

From Eqn. (3.10), $\mathbf{WV}(j)$ then becomes:

$$\mathbf{WV}(j) = \left[\begin{bmatrix} 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 2 \end{bmatrix}\right]$$

Finally, the decision becomes:

$$\textit{Final decision} = [1 \quad 1 \quad 2 \quad 3 \quad 2\text{or}3]$$

In this particular example, there is a tie, as the final decision indicates. The weighting factor plays an important role in eliminating such situations by placing more emphasis on classifiers with higher weights. For example, the overall performances of individual classifiers for continuous tool wear using recently obtained data, are 94.1%, 94.1%, 84.0%, and 67.5% for HMM, Bayesian, GMM, and K-means models, respectively. These overall performances are now used as weighting factors as follows:

$$\mathbf{W} = [w_1 \quad w_2 \quad w_3 \quad w_4] = [0.941 \quad 0.941 \quad 0.840 \quad 0.675]$$

Then from Eqn. (3.10), $\mathbf{WV}(j)$ becomes:

$$\mathbf{WV}(j) = \left[\left[\begin{bmatrix} 2.456 & 0.941 & 0.000 \end{bmatrix} \begin{bmatrix} 2.722 & 0.675 & 0.000 \end{bmatrix} \begin{bmatrix} 0.840 & 2.557 & 0.000 \end{bmatrix} \begin{bmatrix} 0.000 & 0.941 & 2.456 \end{bmatrix} \begin{bmatrix} 0.000 & 1.882 & 1.515 \end{bmatrix} \right] \right]^T$$

resulting in the following final decision:

$$\textit{Final decision} = [1 \quad 1 \quad 2 \quad 3 \quad 2]$$

Thus there are no more tie votes. Since the individual classifier performances vary with the states, the classifier fusion concept was further enhanced by updating the weighting factors using the respective classifier performances for each state. Finally, penalty accommodated weighting was used, giving a penalty on voting for classifier fusion in order to increase reliability.

3.5 Results and discussion

Table 3.1 summarizes the results for individual classifier performances. The HMM and Bayesian classifiers show the best overall performance. However, in state 2, the classification rate was only 29.3% and 38.7% for the HMM and Bayesian classifiers, respectively. Table 3.2 shows the results from classifier fusion with different weighting factors. No weights were used for the results in the first column. In the second column, the overall performance of each classifier was used as its weight. Finally, in the third column, the performance of each state for a classifier was used as its weight.

Table 3.2. Classifier fusion performance (%) at each state with different weighting factors

	No weight performance	Overall performance	State performance
State 1	99.2	99.2	99.2
State 2	88.0	51.3	91.3
State 3	71.6	99.0	99.0
Overall	87.7	95.6	98.5

By applying the classifier fusion algorithm, the state 2 performance increased to 88% with no weight factor, but the overall performance decreased to 87.7% compared to 94.1% without classifier fusion. Using the weighted classifier fusion, with overall performance of individual classifiers as weighting factors, interestingly, the overall performance increased to 95.6% while the state 2 performance decreased to 51.3%. This is because the classifier performance is poor for the HMM and Bayesian classifiers in state 2, as Fig. 3.8 shows, but the weight was drawn from the overall performance, Fig. 3.5.

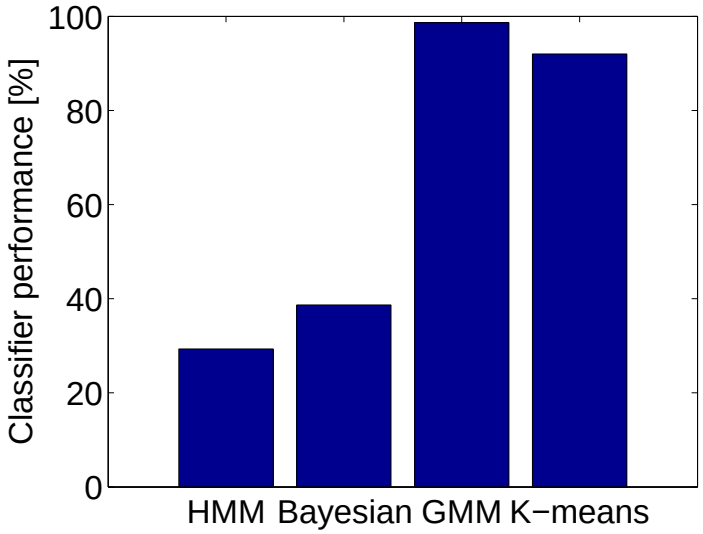


Fig. 3.8. Classification rates of individual classifiers for monitoring the coroning process in state 2 (slightly worn tool)

Since the HMM and Bayesian classifiers had better overall performance, they were overemphasized, leading to a wrong decision for state 2. Thus it is necessary to use the performance in each state as the weighting factors. The results were significantly improved with the state weighted factors, Table 3.2. The average classification rate increased from 95.6% to 98.5%, which is also higher than the single classifier performance of 94.1%. It is obvious from Fig. 3.9 and Table 3.3 that classifier fusion increased the overall performance. However, the non-weighted classifier fusion decreased performance when compared to HMM and Bayesian single classifiers.

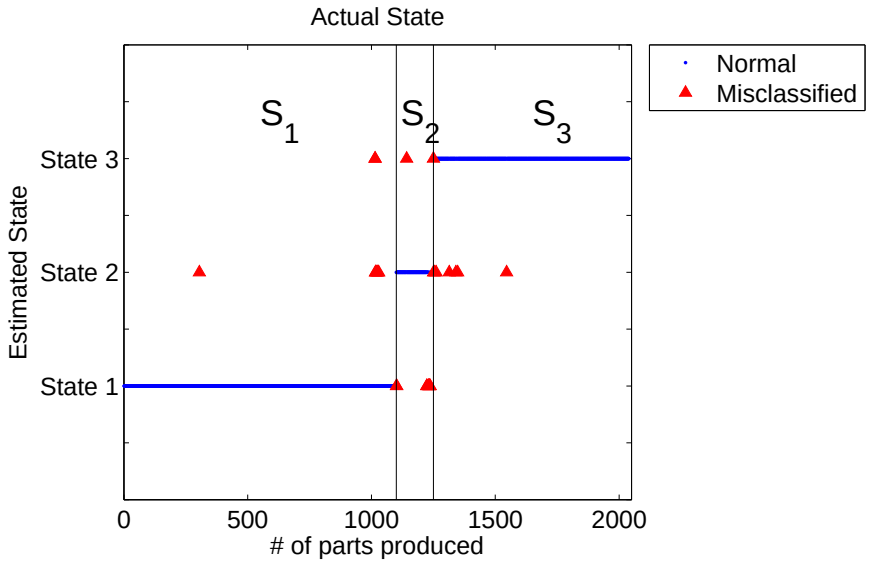


Fig. 3.9. Classification results based on classifier fusion (state weighted majority vote)

Table 3.3. Single classifier vs. classifier fusion performance (%) with weighted majority vote

	HMM	Bayesian	GMM	K-means
Single classifier	94.1	94.1	84.0	67.5
No weight performance			87.7	
Overall performance			95.6	
State performance			98.5	

State weighted classifier fusion increased overall performance, but it did not always improve it in terms of state performances. Classifier fusion with weighting factors enables more reliable decisions as a result of the high performance classifiers' voting. However, if three out of four classifiers have low performance, their combined effort could outweigh that of the high performing classifier, resulting in possible wrong decision. To minimize this possibility, a classifier fusion with weighted voting is considered. A vote penalty is incorporated into the decision, resulting in the following weighting factor:

$$w_{f_i} = \begin{cases} w_i & \text{if } w_i \geq 95\% \\ 0 & \text{otherwise} \end{cases} \tag{3.11}$$

Substituting w_{f_i} for w_i gives the results shown in Fig. 3.10. The penalty threshold can be determined based on the manufacturer's requirements. Table 3.4 shows the improvement in classification using classifier fusion with 95% vote penalty based on the classifier reliability.

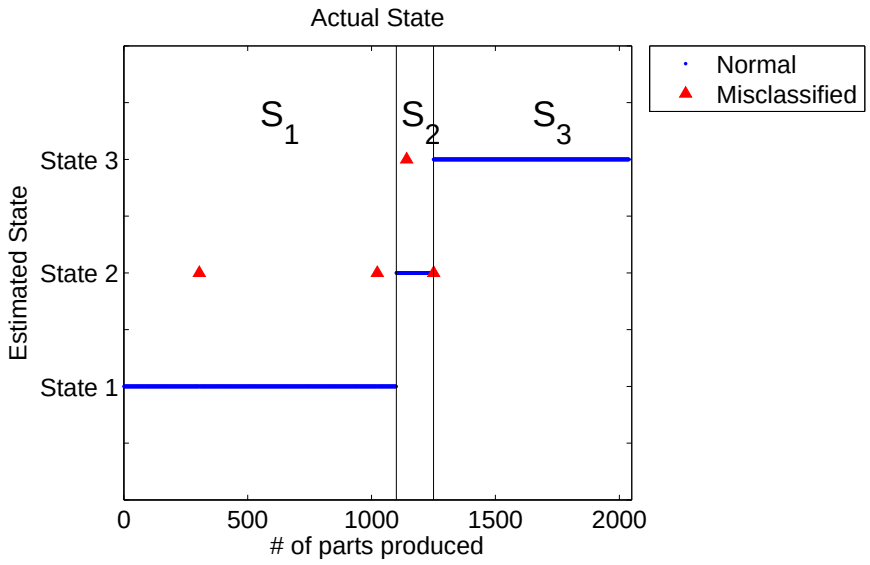


Fig. 3.10. Classifier fusion with state weighted vote penalty

Table 3.4. Classifier fusion compared to the state weighted classifier fusion

	Classifier fusion (State weight)	Classifier fusion (State weight vote penalty)
State 1	99.2	99.8
State 2	91.3	98.7
State 3	99.0	99.7
Overall	98.5	99.7

3.6 Conclusions

A classifier fusion algorithm adapted from decision fusion significantly increased performance compared to the non-fused algorithm in monitoring tool wear during coroning. Even though the overall performance improved with the new technique, the performance for some states improved at the expense of others. This trend was

eliminated by applying a classifier fusion, where a vote penalty was introduced, preventing multiple wrong votes, and this improved classification rate from 98.5% to 99.7%.

Even though the technique was demonstrated for acoustic emission monitoring of the coroning process, it should also be applicable to any process of interest, using any suitable sensors.

Chapter IV

Analysis of sound signals generated during ultrasonic welding of battery cells

4.1 Introduction

Ultrasonic welding (UW) is a joining technology for similar and dissimilar materials, where high frequency vibration induces friction to form a joint between materials. Traditionally, laser (Shackleton and Rischall, 1964) and arc welding processes have been widely applied to metal bonding. However, such fusion welding methods are difficult for joining dissimilar metals due to metallurgical incompatibility (Sun and Karppi, 1996; Montanarini and Steffen, 1976), and differences in melting temperature (Phanikumar et al., 2011) and physical properties. Dissimilar metal joining has a broad range of applications, and thus researchers continue to investigate advanced technologies. Govekar et al. (2009) introduced droplet joining in laser welding of dissimilar materials, and Mai and Spowage (2004) reported a technique that controls the melting ratio of dissimilar metals in laser welding. High reflectivity and thermal conductivity metals, such as aluminum and copper, are especially difficult to weld using a laser. UW is an alternative method that can overcome these challenges with rapid process time, process automation, and high throughput. It has the main advantage of solid-state bonding (Joshi, 1971), and has been used in electronic packaging (Tsujino et al., 2001; Cho et al., 2003b), medical devices (Devine, 1998), and plastic welding for several decades (Tsujino et al., 2002; 1998; Khmelev et al., 2008a;b). In particular, Electric Vehicle (EV) battery packs require the assembly of

several hundreds of battery cells in order to provide the high capacity needed, where the anode and cathode materials are commonly copper and aluminum respectively. UW is an efficient method for joining dissimilar and multi-layer battery tabs in automotive battery applications. The numerous battery cells must be connected to have reliable mechanical strength and electrical conductivity. If any single joint fails, the entire battery pack may lose the desired performance. Thus it is also important to ensure weld quality during the process. Monitoring the product quality in real-time during rapid manufacturing is essential to minimize cost and to control the process.

In order to facilitate an efficient real time monitoring system, sensors have to be chosen properly. Appropriate sensor selection leads to good feature candidates and provides an effective means to a successful manufacturing process monitoring. An UW process has the following characteristics:

•Ultrasonic vibration: 20 kHz transverse movement introduced by the sonotrode
•Ultrasound radiation: sound energy transferred to the air
•Material deformation: property change of the work pieces.

In order to identify the welder and weld status, the laser vibrometer has been introduced in UW (Gaul et al., 2010), but it has the disadvantage of high cost. Norgia et al. (2010) reported an optical instrument for ultrasonic welder inspection and control using laser vibrometer and optical triangulator. Xudong and Xiaochun (2007) investigated heat generation in ultrasonic metal welding using micro sensor arrays. It was reported that heat is necessary but not sufficient condition for a weld to form, and there is no direct correlation between heat and weld quality. Due to the nature of the ultrasonic welding process, radiated audible sound may contain useful information for the state of welding or the change of states. Sound signal has been widely used in monitoring manufacturing systems (Lu and Kannatey-Asibu, 2002; Chen et al.,

2009b). Wang et al. (2009) reported that welding parameters have great influence on the welding sound. And the radiated sound is an important quantity for detecting arc stability as well as weld quality (Cudina et al., 2008; Manz, 1981; Matteson et al., 1993; Pal and Pal, 2011; Sun et al., 1999). Masuzawa and Ohdaira (2000) reported that the amplitude of radiated ultrasound is related to the change of welding state during ultrasonic plastic welding. In many cases, sound signal is analyzed in the frequency domain and features are then extracted (Bi et al., 2010; Pan et al., 2009). However, there are very few papers that have been reported for the feasibility of the sound signal to monitor weld quality for ultrasonic metal welding. Therefore, this paper focuses on the radiated sound signals for characterizing and monitoring the ultrasonic welding process for dissimilar metal joining. Time-frequency analysis is conducted in order to identify the frequency change during the UW process. The rationale for the frequency variation is investigated through a two degree of freedom mechanical model, which involves stiffness coupling. Audible sound features are then selected to assess the feasibility for a real time monitoring system.

The paper is organized as follows: (1) background of the ultrasonic metal welding process and measurement system, (2) two degree-of-freedom modeling of the sound signal; (3) results and discussion for audible sound signal analysis and its comparison to weld quality; (4) classification performance; and (5) conclusions.

4.2 Background

4.2.1 Ultrasonic metal welding

Ultrasonic metal welding is a solid-state bonding process with high frequency frictional vibration to form a bond between work-pieces clamped under pressure. Compared to traditional fusion welding, it works for dissimilar, conductive materials such as copper and aluminum. A typical ultrasonic metal welding system consists of four subsystems: 1) a controller, 2) an ultrasonic transducer (or converter), 3) a booster,

4) a horn (or sonotrode) and anvil as shown in Fig. 4.1.

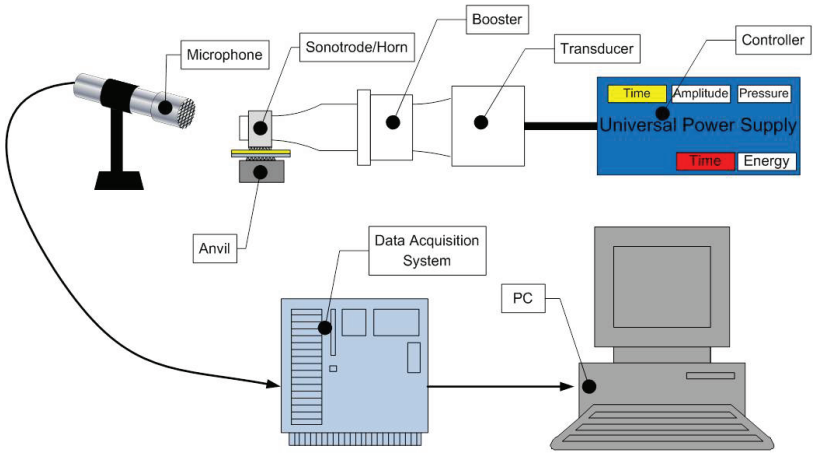


Fig. 4.1. Ultrasonic set up with monitoring system.

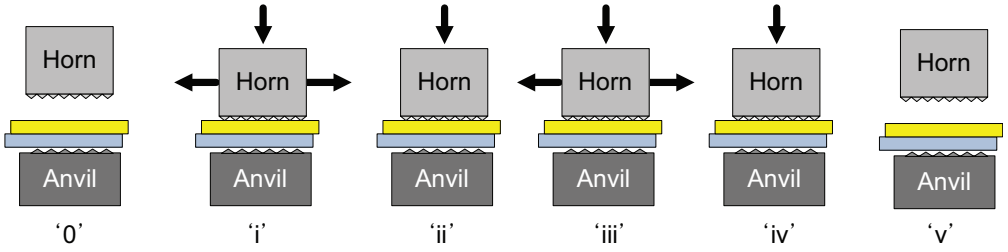
An ultrasonic welding process is initiated by applying a static force to the work pieces. A 20 kHz frequency electrical energy is then converted into mechanical energy by the transducer to vibrate the horn. The lateral frictional movement first removes contaminants on the surface, and then generates a weld at an elevated temperature (Lee et al., 2010). The major input parameters to the system are welding time, sonotrode amplitude and pressure. Different inputs will result in different weld quality. Not delivering enough energy to the interface between the workpieces may result in a cold weld such that the bonding between metals is weak.

4.2.2 Signal and weld formation

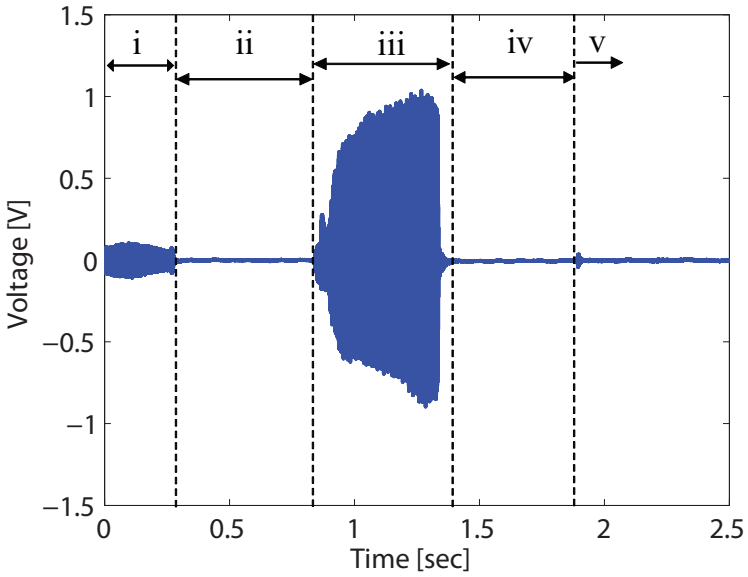
During ultrasonic metal welding, vibration energy is transferred to the work-pieces (thin metal layers) to create a joint, while partial energy is radiated as ambient noise. Vibration energy induces friction, which sometimes generates sharp noise. In brake systems in automotive wheels, this noise is known as squeal noise, and has been studied to control system instability due to vibrations from squeal noise (Currier

et al., 2011; Papinniemi et al., 2002; Shin et al., 2002; Spurr, 1961). With regard to squeal noise, it involves a stick-slip phenomenon (Ruiten, 1988; Vatta, 1979), which is strongly related to frictional force (Lehr et al., 2011). A similar phenomenon is observed during the ultrasonic metal welding process, which radiates sound signals during welding. Masuzawa and Ohdaira (2000) reported that radiated ultrasound contains important information that tells the welding status of joining thin polymer films. The radiated ultrasound can also be applied to ultrasonic metal welding. This paper will further investigate sound energy and how it relates to the weld qualities described in the following sections.

Fig. 4.2 illustrates five steps associated with the ultrasonic welding process. In the first step, the workpieces are clamped under pressure, and a 20 kHz pre-burst then applied to ensure a firm clamp. The pressure is held for a few milliseconds without any vibration in the next stage. Step iii is where the horn is vibrated at 20 kHz to form a bond between the workpieces, and this is followed by another hold step, iv. Finally, the workpieces are released with an after-burst, step v.



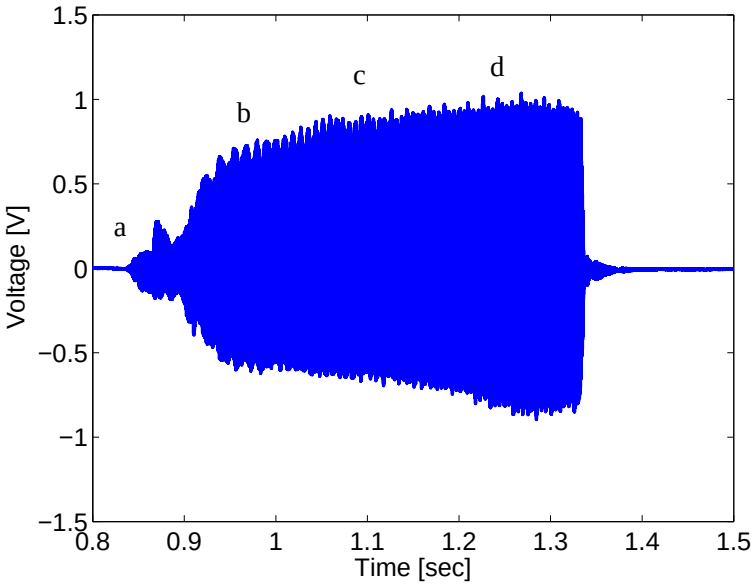
(a) Ultrasonic welding process



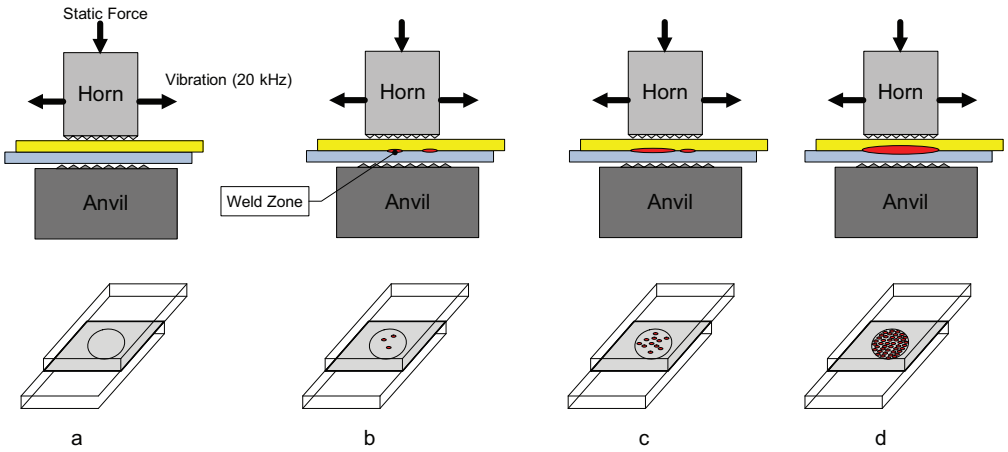
(b) Corresponding sound signal

Fig. 4.2. Ultrasonic welding process and corresponding sound signal.

Fig. 4.3(a) is an enlargement of stage iii, and illustrates a typical sound signal generated during the weld growth and formation step. This signal was obtained for an ultrasonic weld between pure copper and nickel coated copper sheets. The high frequency frictional motion creates a bond between the workpieces. In Fig. 4.3(b), stages a to d illustrate the nugget growth in time and how that corresponds to the sound signal that was measured during the process.



(a) Time domain sound signal



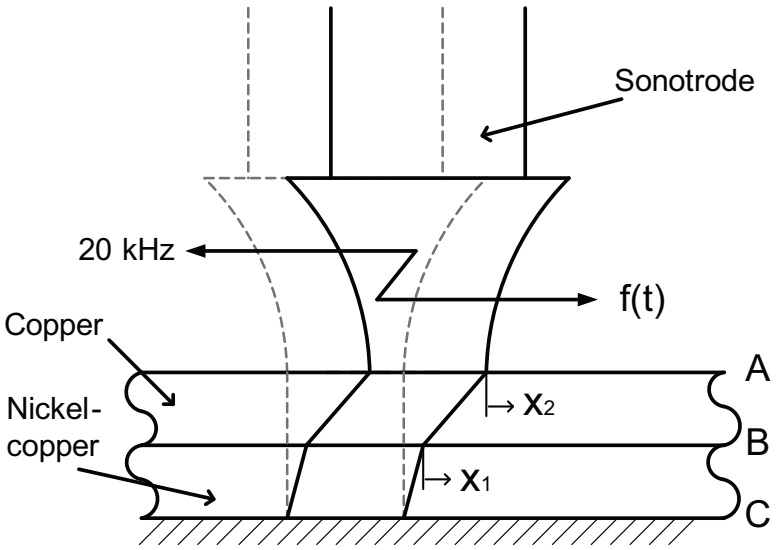
(b) Weld growth in relation to time

Fig. 4.3. Raw sound signal with corresponding weld nugget growth.

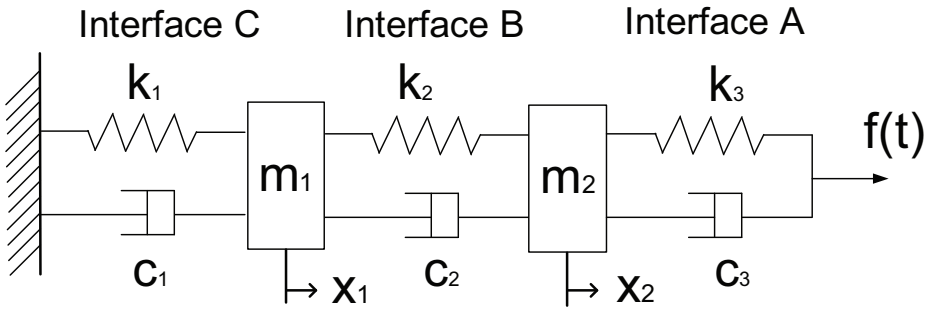
Weld formation is directly related to the weld quality since the bonded area between the workpieces determines the strength. Thus it should be possible to monitor the weld quality by measuring and analyzing the sound signals generated during the process. Progress in the welding process results in stiffness variation. Thus to understand the characteristics of the sound signal generation, the process is modeled using a two degree-of-freedom mechanical system in the next section.

4.3 Two degree-of-freedom modeling of sound signal from the process

This section studies the relationship between the audible sound signal and weld quality during the process using state space modeling of a mass-spring-damper system. The equation of motion is based on the mass-spring-damper system shown in Fig. 4.4, where m_1 and m_2 are masses for nickel coated copper and pure copper respectively. k_1, k_2 and k_3 represent the stiffness coefficients and c_1, c_2 and c_3 are damping coefficients for pure copper, interface bonding, and nickel coated copper respectively.



(a) Actual process



(b) Corresponding model

Fig. 4.4. Two degree of freedom mass-spring-damper system.

The equation of motion for the system shown can be written as:

$$m_1\ddot{x}_1 + (c_3 - 2c_2 + c_1)\dot{x}_1 + (c_2 - c_3)\dot{x}_2 + (k_3 - 2k_2 + k_1)x_1 + (k_2 - k_3)x_2 = 0 \tag{4.1}$$

$$m_2\ddot{x}_2 + (2c_3 - c_2)\dot{x}_2 + (c_2 - c_3)\dot{x}_1 + (2k_3 - k_2)x_2 + (k_2 - k_3)x_1 = f(t) \tag{4.2}$$

where f(t) is the frictional force that results from the static load applied to the work piece, which vibrates at 20 kHz cycle. This can be expressed in state space form as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{2k_2-k_3-k_1}{m_1} & \frac{k_3-k_2}{m_1} & \frac{2c_2-c1-c_3}{m_1} & \frac{c_3-c_2}{m_1} \\ \frac{k_3-k_2}{m_2} & \frac{k_2-2k_3}{m_2} & \frac{c_3-c_2}{m_2} & \frac{c_2-2c_3}{m_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{m_2} \end{bmatrix} f(t) \tag{4.3}$$

$$y = \begin{bmatrix} \frac{(k_3-k_2)}{m_2} & \frac{k_2-2k_3}{m_2} & \frac{c_3-c_2}{m_2} & \frac{c_2-2c_3}{m_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \frac{f(t)}{m_2} \tag{4.4}$$

where $x_1 = x_1$, $x_2 = x_2$, $x_3 = \dot{x}_1$, and $x_4 = \dot{x}_2$.

Since m_2 moves with the sonotrode, the output was based on its acceleration, assuming m_1 is fixed with the anvil. The acceleration is the output, equation 4.4. In the next section, the experiments that were performed to evaluate the analysis are discussed.

4.4 Experimental

A (Bruel & Kjaer 4190-L-001) microphone was used to measure signals from the welding process to enhance our understanding of the process and monitor weld quality. As Fig. 4.1 shows, the radiated sound signal was recorded during the entire process at 100 kHz sampling rate. The audible sound signal was then analyzed using the Fourier Transform (frequency analysis) and Short Time Fourier Transform (time-frequency decomposition) to assess the feasibility of monitoring and to help understand the mechanisms of the process. The relationship between the input parameters to the welding system and output signals from the sensors are compared. Time domain signals are analyzed and transformed to the frequency domain for further analysis. The workpieces used in this experiment were pure copper and nickel coated copper with dimensions 45x25x0.2 mm.

Table 4.1 summarizes the welding conditions. 25 different conditions were used, and each one was repeated ten times.

Table 4.1. Factors and levels of experimental system inputs.

Factor	Factor name	Levels
P	Welding pressure (psi)	10, 25, 40, 55, 70
	[kPa]	69, 172, 276, 379, 483
T	Welding time (sec)	0.2, 0.35, 0.5, 0.65, 0.8

A T-peel test was used to assess the weld quality. A T-peel test, Fig. 4.5, measures how much load is applied to break the welded workpieces while pulling the clamped ends of the workpieces. The workpieces are T-shaped using a bending fixture to pro-

vide a consistent clamping position. This determines the weld quality based on the load-displacement curve and the failure mode after the T-peel test.

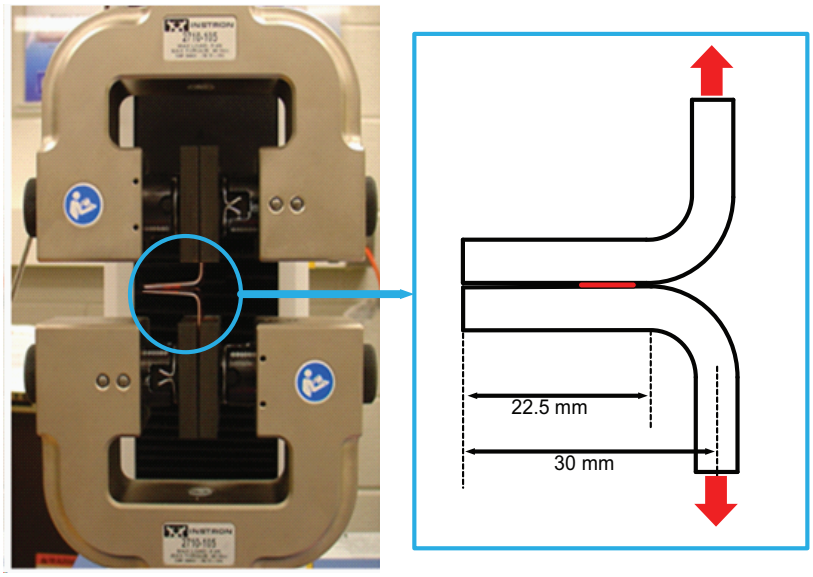


Fig. 4.5. T-peel machine and testing procedure (Kim et al., 2011).

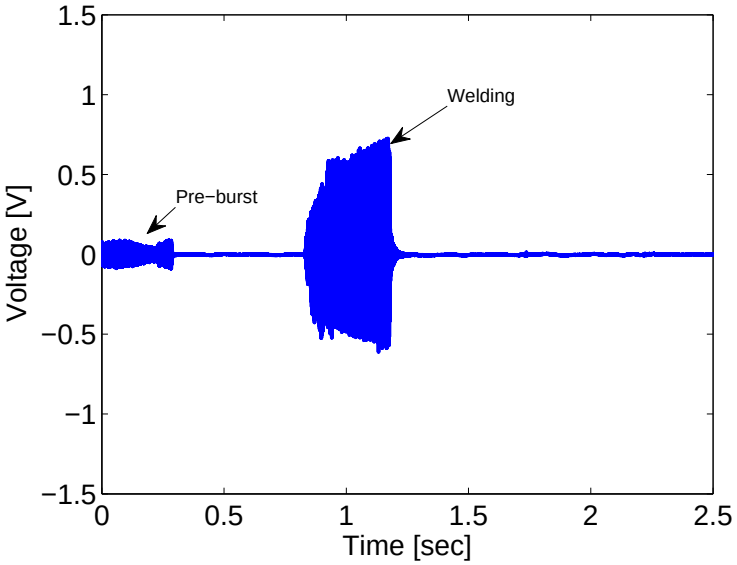
4.5 Results and discussion

In this section, weld quality results from a mechanical test, mechanical model, and sound signals are analyzed. This is followed by a discussion on the spectral characteristics of the sound signal, in order to extract features. The audible sound features selected are then used for classification.

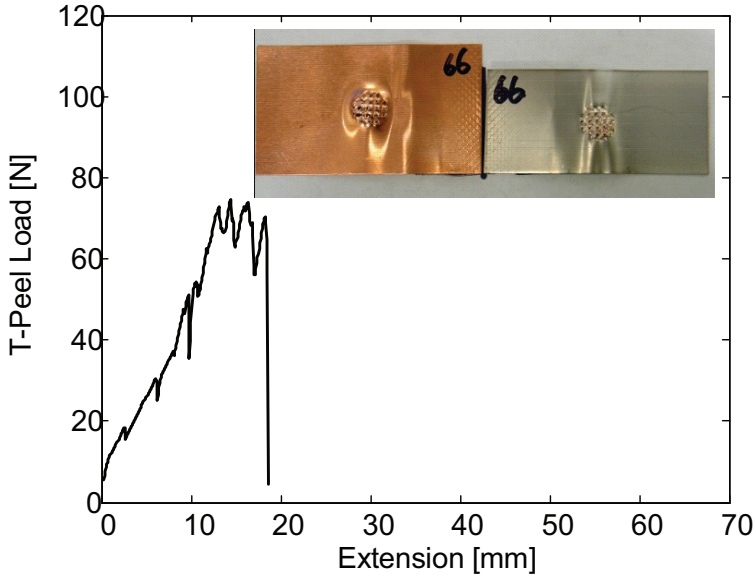
4.5.1 Weld quality from T-peel test and failure mode

Three states of the weld were identified based on the experiments: cold weld, good weld, and over weld. Figs. 4.6(a), 4.7(a), and 4.8(a) show signals associated with the cold weld condition (25 psi (172 kPa) for 0.35 sec), good weld condition (55 psi (379 kPa) for 0.5 sec), and over weld condition (70 psi (483 kPa) for 0.8 sec) respectively. Each weld involves five steps: pre-burst, hold, weld, hold, and after-burst. T-peel test results for weld samples from those conditions are shown in Figs. 4.6(b), 4.7(b),

and 4.8(b).

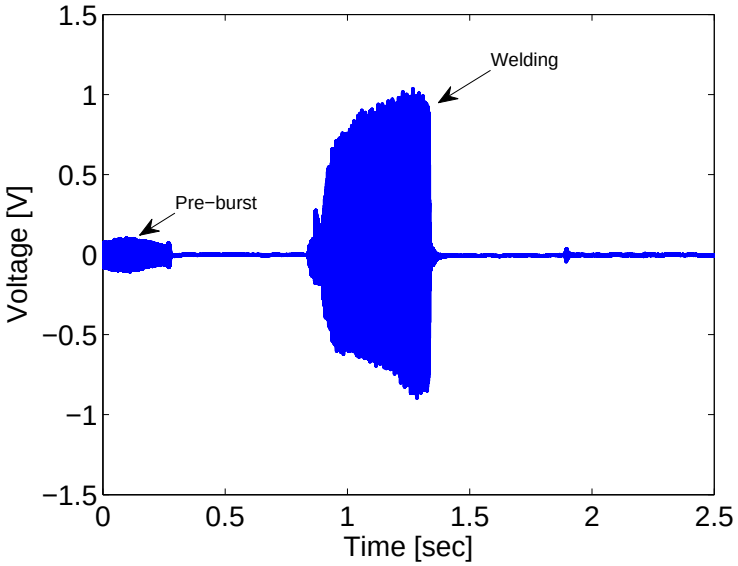


(a) Microphone signal from a cold weld condition

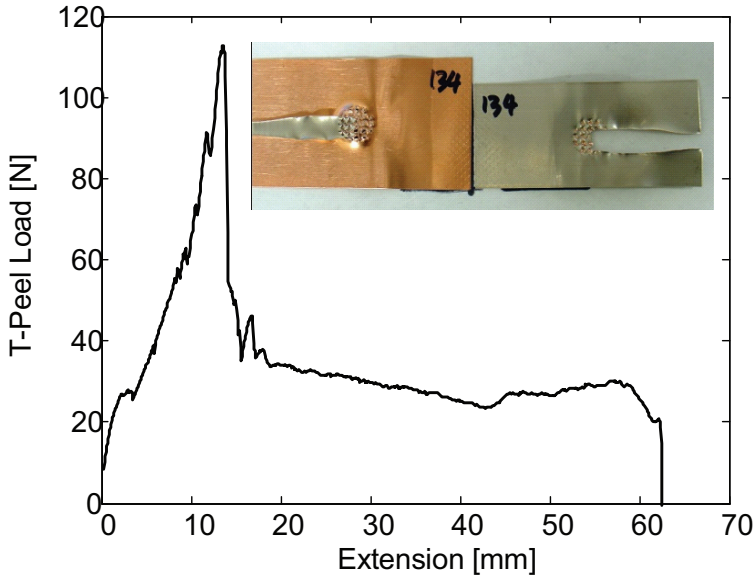


(b) Load-displacement curve from T-peel test and tested specimen

Fig. 4.6. Sound signal generated during a cold weld condition (25 psi (172 kPa), 0.35 sec) and associated load-displacement curve from a T-peel test with the failure mode.

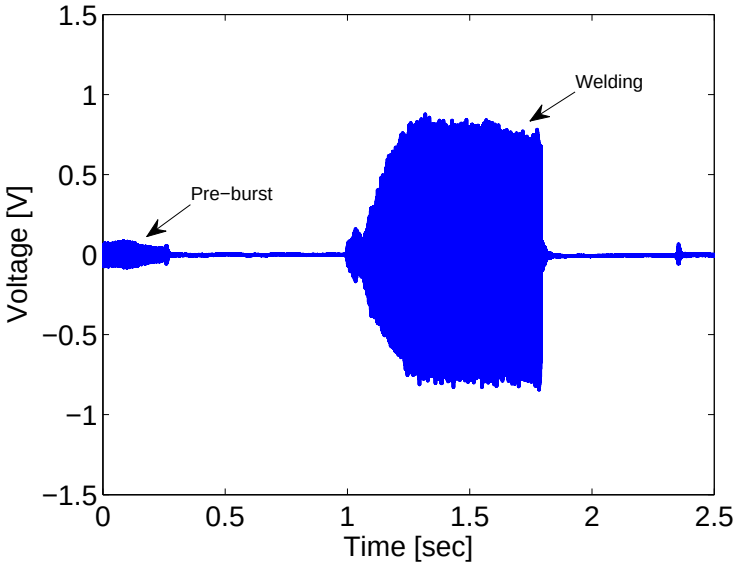


(a) Microphone signal from a good weld condition

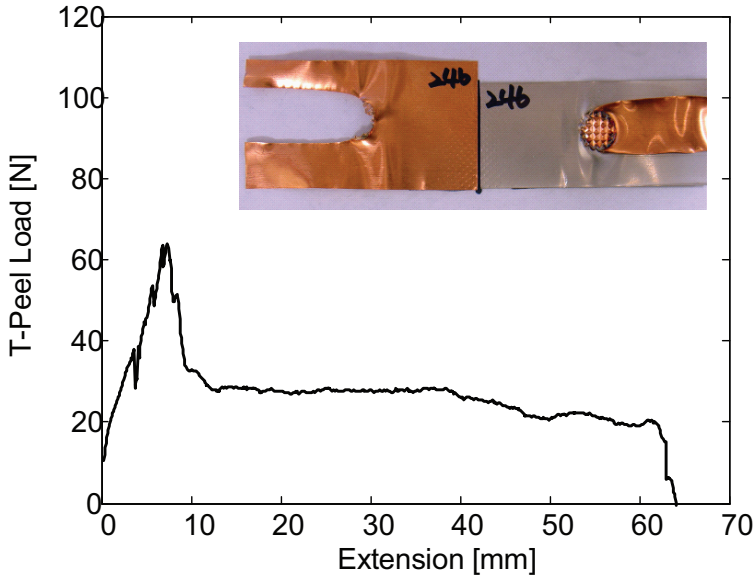


(b) Load-displacement curve from T-peel test and tested specimen

Fig. 4.7. Sound signal generated during a good weld condition (55 psi (379 kPa), 0.5 sec) and associated load-displacement curve from a T-peel test with the failure mode.



(a) Microphone signal from an over weld condition



(b) Load-displacement curve from T-peel test and tested specimen

Fig. 4.8. Sound signal generated during an over weld condition (70 psi (483 kPa), 0.8 sec) and associated load-displacement curve from a T-peel test with the failure mode.

From the time domain signal, one can observe that the duration of the sound is directly related to the weld time while the amplitude is related to the pressure. However, these by themselves cannot tell the status of the weld since some of the welds show different strengths by a T-peel test for the same welding conditions. Thus,

the frequency characteristics of $\ddot{x}_2$ as the system output, and time-frequency analyses are further studied in the following sections to investigate possible feature candidates that can describe the weld status during the welding process.

4.5.2 Effect of stiffness variation

Let us define k_1, k_2, and k_3 as $\alpha_1 K, \alpha_2 K$, and $\alpha_3 K$ respectively, where α is a constant of each stiffness coefficient. By setting the stiffness ratio as $\alpha_1 : \alpha_2 : \alpha_3 = 1 : 2.86 \times 10^{-5} : 1$, and using a value of 2 grams for each of the masses m_1 and m_2, the resulting acceleration was calculated and then transformed into the frequency domain. Fig. 4.9 shows the variation of spectral components with change in stiffness coefficients. The interfacial stiffness is assumed to be much smaller than the other stiffness coefficients since it is initially sliding and starts to form a joint as process evolves while the others are clamped by the sonotrode and anvil.

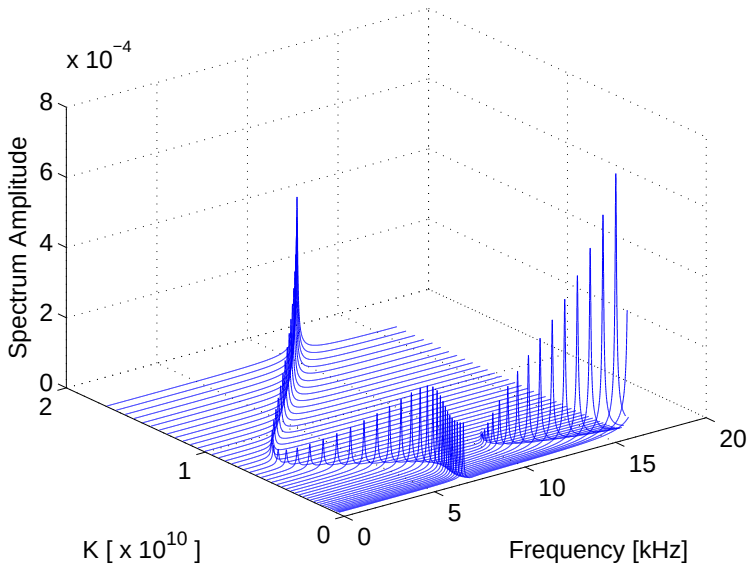


Fig. 4.9. System frequency variation due to stiffness change.

At relatively low stiffness values, there exist two frequency spikes in the audible sound range (0-20 kHz). However, the spectrum becomes dominated by one frequency

component as the stiffness increases. This corresponds to growth of the weld, which results in the joint becoming more rigid. Thus, one can conjecture that weld growth affects the system dynamics due to stiffness change, which has strong correlation with the weld quality.

4.5.3 Effect of mass variation

In this section, the variation of mass is studied. m_1 and m_2 are varied over a range of five times the initial mass, in other words, a workpiece thickness variation from 0.2 mm to 1 mm. As the anvil side mass, m_1, increases, the frequency components gradually shifts to the lower frequency range, Fig. 4.10 as expected. However, interestingly enough, a second frequency peak also appears in the high frequency range.

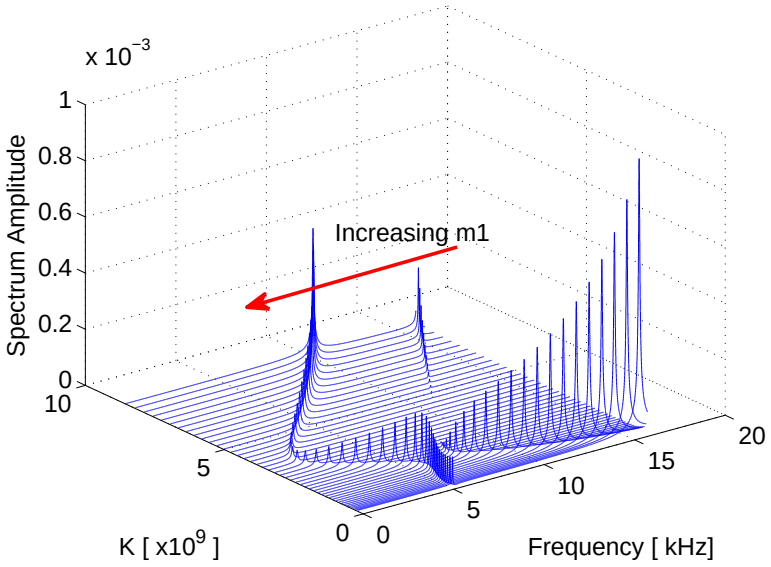


Fig. 4.10. System frequency variation due to mass change (m_1).

This indicates that there may be other coupling effects at play. From the root locus plots, Figs. 4.11- 4.13, it is realized that the single frequency peak results from pole-zero cancelation in a two degree-of-freedom system.

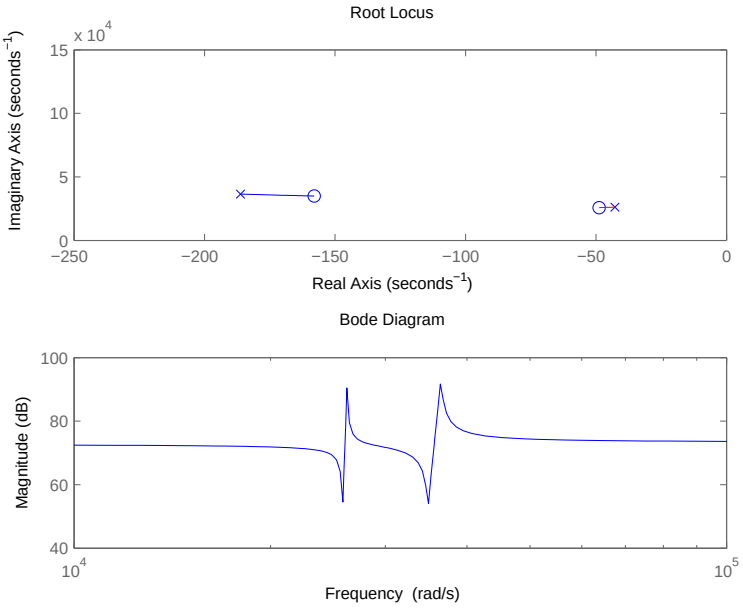


Fig. 4.11. Root locus and bode diagram for K = 7.07 x 10^6.

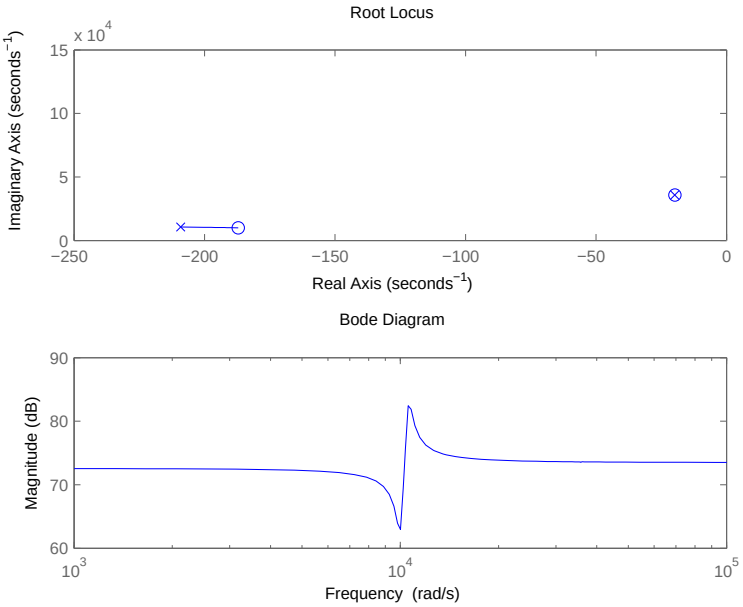


Fig. 4.12. Root locus and bode diagram for K = 12.73 x 10^6.

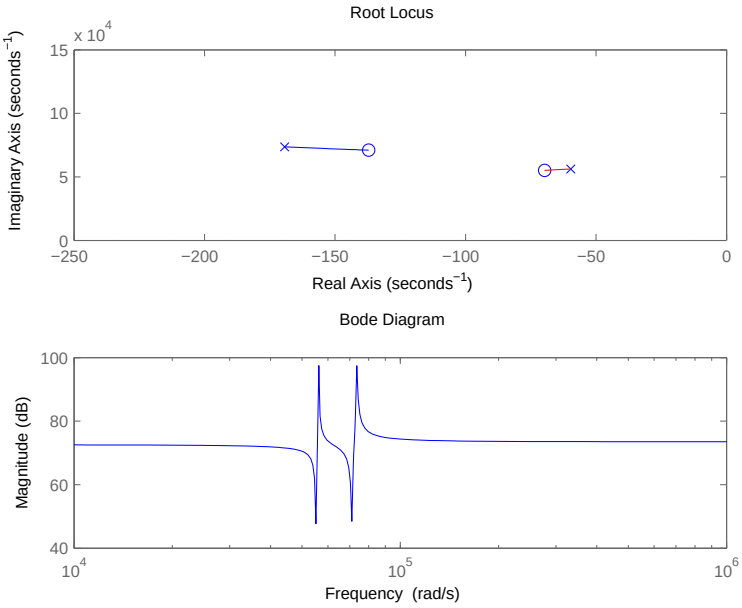


Fig. 4.13. Root locus and bode diagram for K = 34.13 x 10^6.

However, as m_2 increases, Fig. 4.14, the single peak at higher stiffness values gradually shifts to the lower frequency range.

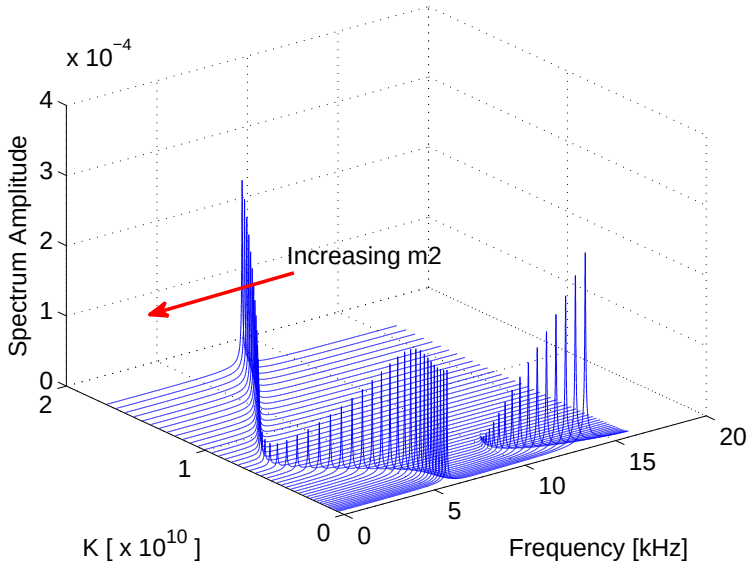
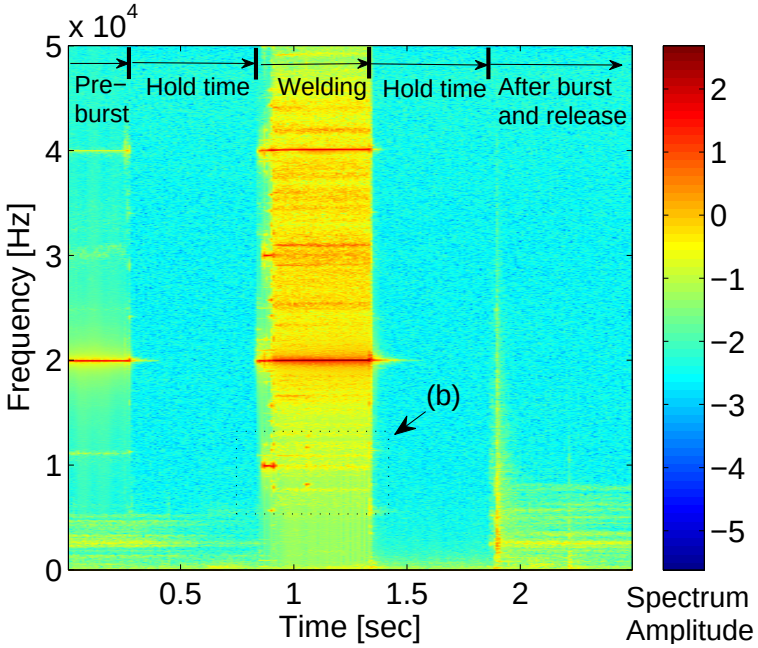


Fig. 4.14. System frequency variation due to mass change (m_2).

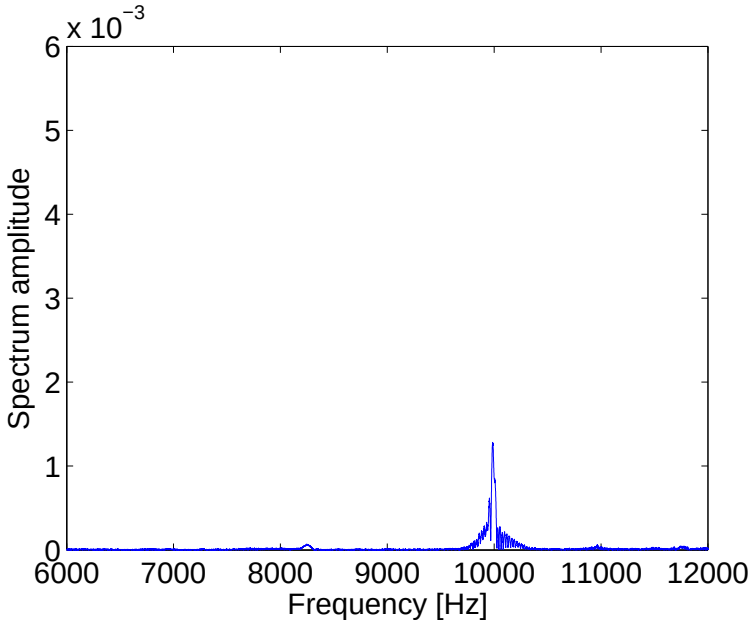
4.5.4 Time-frequency analysis

In this section, the sound signals from the microphone were converted to the frequency domain for further analysis using the fast Fourier transform (FFT). In addition, time-frequency analysis was also conducted to determine how the frequency components change as the weld progresses. Chu et al. (2004) studied the relationship between the short-circuiting frequency and process stability using time-frequency analysis in short-circuiting gas metal arc welding. This paper shows a strong correlation between weld quality and radiated ultrasound.

Fig. 4.15 shows the time-frequency spectrum for a good weld. As Fig. 4.15(a) clearly shows, the main welding frequency of 20 kHz dominates the spectrum. A zoom-in section, Fig. 4.15(b), shows only the 10 kHz component as the dominant frequency, with others being almost negligible. Fig. 4.16 shows similar plots for a cold weld and here, the frequency components are spread throughout the entire spectrum, unlike the good welding condition. Compared to Fig. 4.15(b), two other spectral components are more dominant than the 10 kHz component, and these are near 9 and 11 kHz as shown in Fig. 4.16(b). The over weld case is also shown in Fig. 4.17, and there, no major spikes are observed in the spectrum. The cold, good, and over weld states are determined by a combination of pressure, time, and vibration frequency and amplitude. Thus they have different stiffness characteristics and sound signatures.

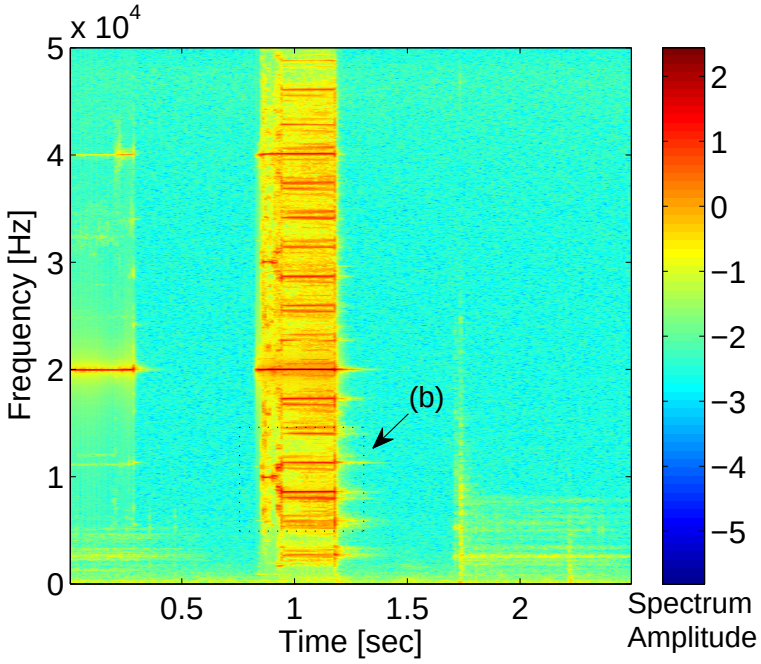


(a) Time-frequency of sound signal

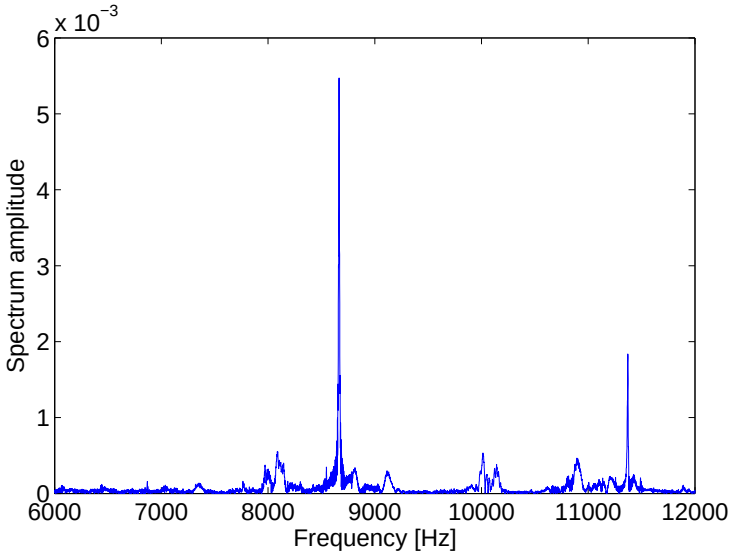


(b) Zoom-in FFT

Fig. 4.15. Time-frequency sound signals for a good weld condition and zoom-in FFT.

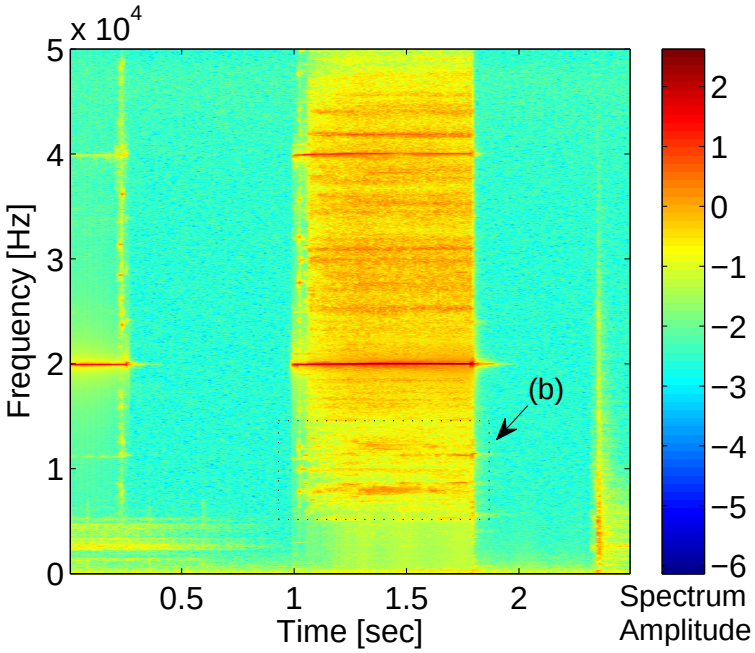


(a) Time-frequency of sound signal

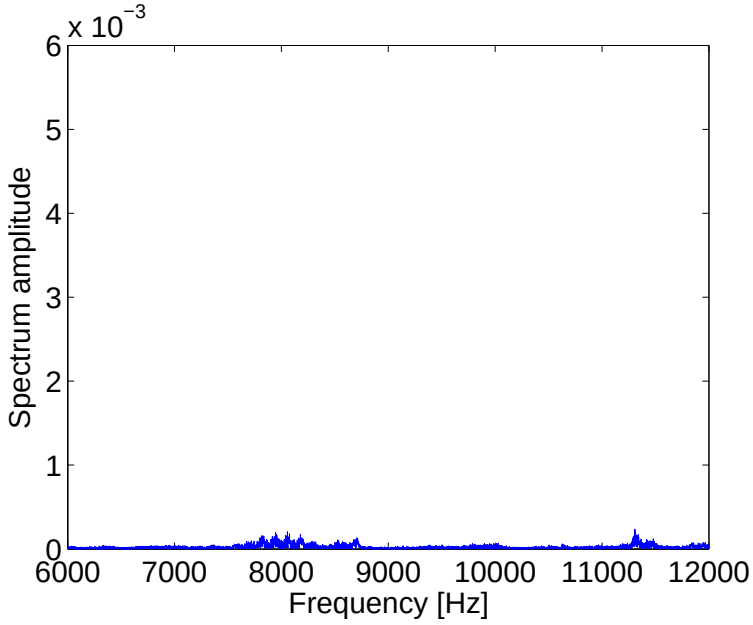


(b) Zoom-in FFT

Fig. 4.16. Time-frequency sound signals for a cold weld condition and zoom-in FFT.



(a) Time-frequency of sound signal



(b) Zoom-in FFT

Fig. 4.17. Time-frequency sound signals for an over weld condition and zoom-in FFT.

There is a clear distinction between good and cold welds in the frequency domain as observed in the frequency spikes during the welding process. A spectral peak exists for a good welding condition near 10 kHz, which is half the vibration frequency of 20 kHz. However, peaks near 9 kHz and 11 kHz are more significant in the cold weld

condition. These results correlate with the frequency distribution results associated with the stiffness coefficients variation from the sound signal modeling. For the over weld condition, none of the frequency components are significant in the audible sound range. This indicates the potential for using sound signals for detecting the weld quality during ultrasonic welding. The sound signal is capable of identifying the three states of ultrasonic welding, namely cold, good, and over weld conditions.

Frequency peaks near 9 kHz and 11 kHz were further correlated with weld quality using a third-order polynomial regression model that relates the spectral amplitude at that frequency to the welding pressure and time. Since data was only obtained for specific conditions, a regression model was developed to capture the continuous effect of time and pressure. The corresponding equations (4.5 and 4.6) are as follows, where SA, P and T are spectrum amplitude, pressure and time respectively:

$$\begin{aligned} SA_{9kHz} &= -1.055 \times 10^{-4} + 1.297 \times 10^{-5} P \\ &\quad +2.278 \times 10^{-4} T - 1.074 \times 10^{-5} P \times T \\ &\quad -3.089 \times 10^{-7} P^2 + 3.088 \times 10^{-4} T^2 \\ &\quad +2.419 \times 10^{-9} P^3 + 2.140 \times 10^{-8} P^2 \times T \\ &\quad -7.274 \times 10^{-6} P \times T^2 - 4.918 \times 10^{-4} T^3 \end{aligned} \tag{4.5}$$

$$\begin{aligned} SA_{11kHz} &= -6.127 \times 10^{-5} + 7.770 \times 10^{-6} P \\ &\quad +7.536 \times 10^{-5} T - 2.129 \times 10^{-6} P \times T \\ &\quad -1.963 \times 10^{-7} P^2 - 2.027 \times 10^{-5} T^2 \\ &\quad +1.489 \times 10^{-9} P^3 + 2.928 \times 10^{-8} P^2 \times T \\ &\quad -2.138 \times 10^{-6} P \times T^2 + 1.215 \times 10^{-4} T^3 \end{aligned} \tag{4.6}$$

Figs. 4.18 and 4.19 show the response surface of the frequency spectrum near 9 kHz and 11 kHz respectively, and indicate that lower pressures and longer times for the process tend to produce sharp sound signals.

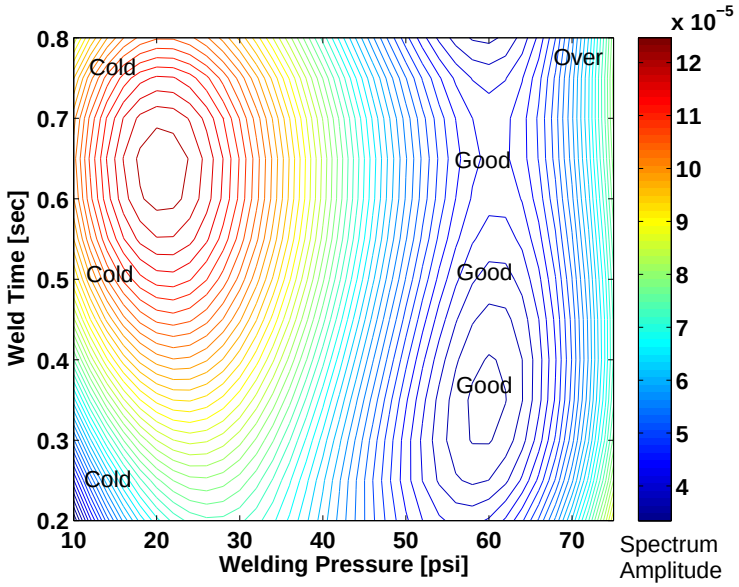


Fig. 4.18. 9 kHz response surface.

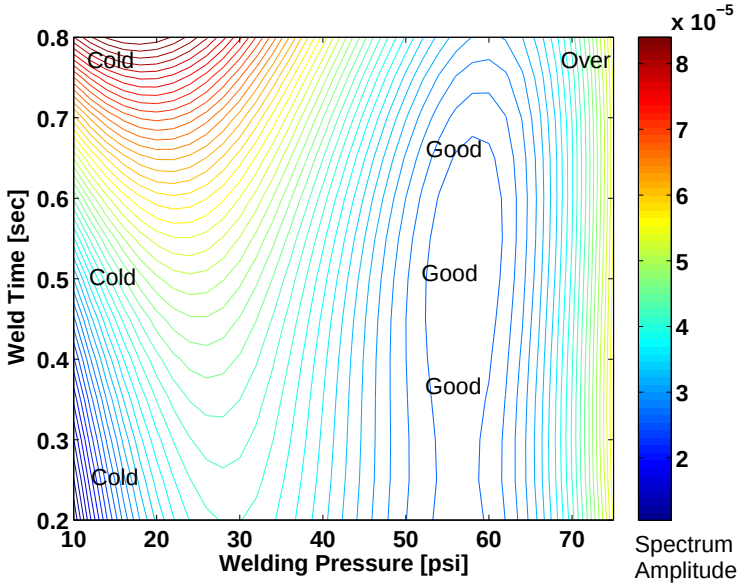


Fig. 4.19. 11 kHz response surface.

Fig. 4.20 shows the weld quality from a T-peel test. The higher intensity 9 and 11 kHz spectral components have similar trends with the mechanical T-peel tests. These

indicate that the presence of sharp noise correlates with a poor quality weld joint.

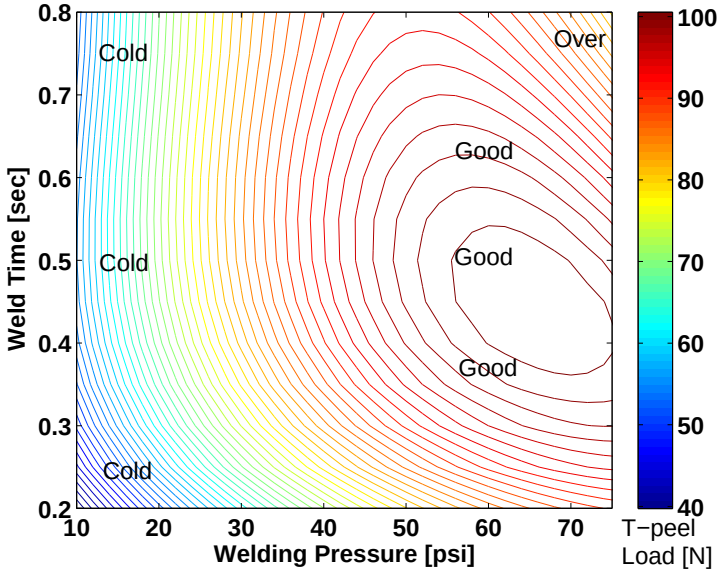


Fig. 4.20. Weld quality from T-peel test (Kim et al., 2011)

4.5.5 Classification

From the initial results (both analytical and experimental), four audible sound features, which are based on spectral components, are used for classification of the cold, good, and over weld conditions. These consist of the average frequency spectrum between: (1) 8 and 9 kHz, (2) 9.5 and 10.5 kHz, (3) 11 and 12 kHz, (4) 19.5 and 20.5 kHz. Four classifiers (Yum et al., 2012), namely Hidden Markov Model (HMM), minimum error rate Bayesian (Bayesian), Gaussian Mixture Model (GMM), and K-means, were evaluated to determine their ability to distinguish between the three weld states. A training data size of 50 samples was used, and a total of 72 separate samples were used for classification. Fig. 4.21 shows the classification rate of each classifier, 88.9, 86.1, 83.3, and 56.9 %, and these were for the HMM, Bayesian, GMM, and K-means respectively.

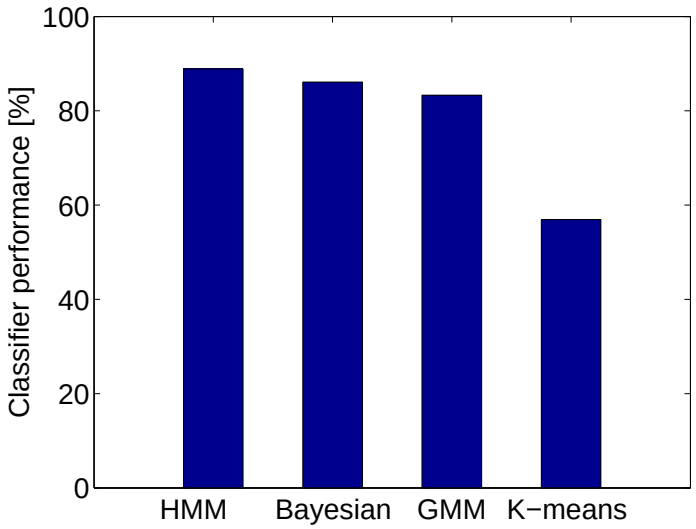


Fig. 4.21. Weld quality classification using audible sound features for cold, good, and over weld conditions.

The HMM results in the best performance compared to the other classifiers. In order to enhance the robustness and reliability of classification performance, as reported by Yum et al. (2012), the classifier fusion method was implemented, Fig. 4.22. Here, State 1, State 2, and State 3 represent cold, good, and over weld conditions, respectively.

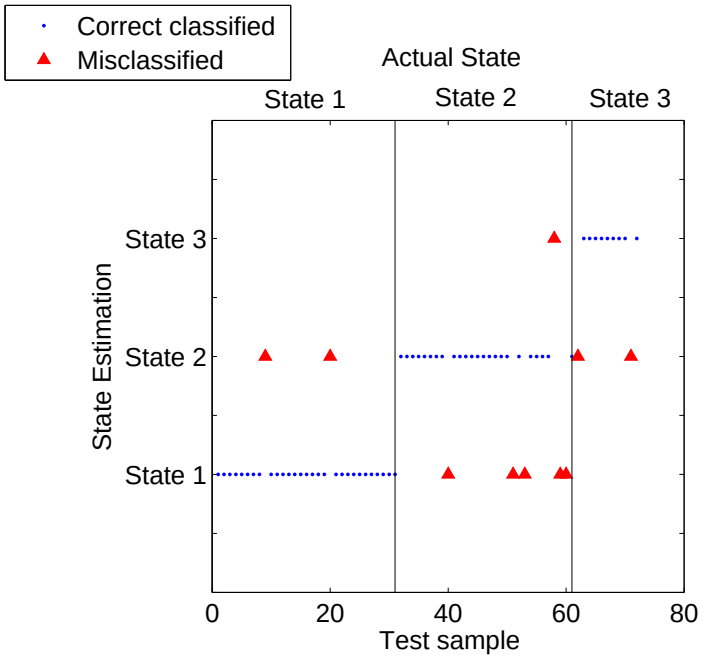


Fig. 4.22. Results for classifier fusion using 4 classifiers and audible sound features.

the classifier fusion method resulted in an 86 % classification rate, which implies that the audible sound signal has the potential of being used to detect the weld condition.

The data in the frequency domain was further analyzed by band-averaging them to investigate the effect of window averaging on classification. Fig. 4.23 shows the classification performance when the frequency amplitudes were averaged to generate a new feature.

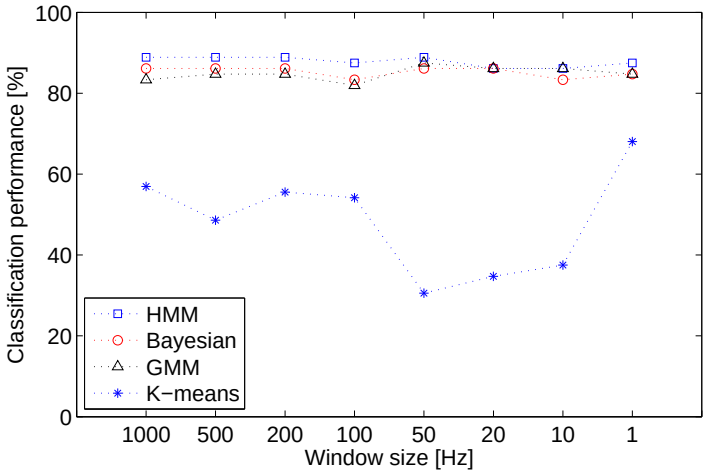


Fig. 4.23. Effect of window averaging on classification.

Averaging has the advantage of reducing computation time, and as the results show, it did not have a significant effect on the final classification rate for three of the classifiers. For the K-means classifier, the performance deteriorated with averaging.

4.6 Conclusions

The results of experiments done on the ultrasonic metal welding process show that:

1) The spectrum for a good weld is dominated by one frequency component in the audible sound range near 10 kHz, which is half the vibration frequency of 20 kHz, for the conditions used, while that for a cold weld is characterized by two frequency components at 9 and 11 kHz. Over weld conditions did not generate unique frequency components.

2) Analysis of the process, based on a two degree of freedom mass-spring-damper system, shows that stiffness variation that results from growth of the nugget during the process correlates with variation in the audible sound frequency components.

3) With the characteristic spectral components as features, audible sound was suc-

cessfully used to identify weld quality, distinguishing between cold, good, and over weld conditions.

Chapter V

Conclusions and Future work

Two-step feature selection and classifier fusion methods are proposed and evaluated for the coroning and ultrasonic welding processes. The ultrasonic welding process is also modeled to relate the sound signal features to the process characteristics.

In Chapter 2, a two-step feature selection method was developed and evaluated for the coroning process. During the first step, significant data reduction was done using the CMS method. Then the modified redundancy analysis further reduced the feature dimensions without loss of information, but rather increasing the separability for each state. The two-step feature selection resulted in 98.3% correct estimation of tool degradation compared to 94.9% for the CMS method only. The two-step feature selection method increased the classification performance while conventional window averaging decreased the classification rate, Fig. 2.12.

In Chapter 3, a classifier fusion algorithm was implemented using state performance weighting factor and a penalty voting concept. It is often difficult for a single classifier to achieve perfect classification during process monitoring. The classifier fusion significantly increased performance compared to the non-fused algorithm in coroning tool wear monitoring. Overall performance weighting resulted in 95.6% performance, and the state performance weighting showed 98.5 % classification while equal weighting for each classifier achieved 87.7%. The introduction of a vote penalty, meant to prevent multiple wrong votes, further improved the classification rate to 99.7%.

In Chapter 4, the feasibility of using audible sound for monitoring the quality of ultrasonic metal welding is presented. Weld quality is strongly related to the spectral components of audible sound. The results of experiments done on the ultrasonic metal welding process show that:

1) The spectrum for a good weld is dominated by one frequency component in the audible sound range near 10 kHz, which is half the vibration frequency of 20 kHz, for the conditions used, while that for a cold weld is characterized by two frequency components at 9 and 11 kHz. Over weld conditions did not generate unique frequency components.
2) Analysis of the process, based on a two degree of freedom mass-spring-damper system, shows that stiffness variation that results from growth of the nugget during the process correlates with variation in the audible sound frequency components.
3) With the characteristic spectral components as features, audible sound was successfully used to identify weld quality, distinguishing between cold, good, and over weld conditions.

The results of this study indicate that the two-step feature selection and classifier fusion methods improved classification in monitoring tool wear of the coroning process. However, in reality, manufacturing processes are frequently faced with uncertainties or disturbances. In other words, the process characteristics can change, and thus the monitoring algorithm has to be capable of adapting to such changes. This will ensure consistency of the algorithm and make the monitoring system more robust.

In audible sound analysis, mechanical modeling showed the relationship between spectral component of audible sound and coefficients of stiffness during nugget growth in ultrasonic welding. However, the spectral components obtained experimentally

were not identical with those predicted by the model. Thus a multi-degree of freedom model will need to be investigated for more complete understanding of the mechanisms of the ultrasonic welding process.

Bibliography

Bibliography

Abbes, M. S., T. Fakhfakh, M. Haddar, and A. Maalej (2007), Effect of transmission error on the dynamic behaviour of gearbox housing, *International Journal of Advanced Manufacturing Technology*, *34*(3-4), 211–218.

Aguiar, P. R., P. J. A. Serni, F. R. L. Dotto, and E. C. Bianchi (2006), In-process grinding monitoring through acoustic emission, *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, *28*(1), 118–124.

Althoefer, K., B. Lara, and L. D. Seneviratne (2005), Monitoring of self-tapping screw fastenings using artificial neural networks, *ASME Journal of Manufacturing Science and Engineering*, *127*(1), 236 – 245.

Bauer, E., and R. Kohavi (1999), An empirical comparison of voting classification algorithms: Bagging, boosting, and variants, *Machine Learning*, *36*(1-2), 105 – 139.

Bhattacharyya, P., and D. Sengupta (2009), Estimation of tool wear based on adaptive sensor fusion of force and power in face milling, *International Journal of Production Research*, *47*(3), 817 – 833.

Bi, S., H. Lan, H. Zheng, and L. Liu (2010), Characteristic analysis and pattern recognition of arc sound under typical penetration status in mig welding, *2010 IEEE International Conference on Information and Automation*, pp. 537–542.

Binsaeid, S., S. Asfour, S. Cho, and A. Onar (2008), Machine ensemble approach for simultaneous detection of transient and gradual abnormalities in end milling using multisensor fusion, *Journal of Materials Processing Technology*, *209*(10), 4728–4738.

Blum, A. L., and P. Langley (1997), Selection of relevant features and examples in machine learning, *Artificial Intelligence*, *97*(1-2), 245–71.

Breiman, L. (1996), Bagging predictors, *Machine Learning*, *24*(2), 123 – 140.

Byington, C., M. Watson, D. Edwards, and B. Dunkin (2003), In-line health monitoring system for hydraulic pumps and motors, *Proceedings of IEEE Aerospace Conference 2003*, *7*, 3279–3287.

Chen, L., et al. (2009a), Multiple classifier integration for the prediction of protein structural classes, *Journal of Computational Chemistry*, *30*(14), 2248–2254.

Chen, T.-H., M.-C. Lu, S.-J. Chiou, C.-Y. Lin, and M.-H. Lee (2009b), Study of sound signal for tool wear monitoring system in micro-milling processes, pp. 57 – 65.

Cho, Y., S. Hu, and W. Li (2003a), Resistance spot welding of aluminium and steel: A comparative experimental study, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *217*(10), 1355 – 1363.

Cho, Y., S. Hu, and W. Li (2003b), Resistance spot welding of aluminium and steel: A comparative experimental study, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *217*(10), 1355 – 1363.

Chu, Y. X., S. J. Hu, W. K. Hou, P. C. Wang, and S. P. Marin (2004), Signature analysis for quality monitoring in short-circuit gmaw, *Welding Journal (Miami, Fla)*, *83*(12), 366–S – 343–S.

Cudina, M., J. Prezelj, and I. Polajnar (2008), Use of audible sound for on-line monitoring of gas metal arc welding process, *Metalurgija*, *47*(2), 81–85.

Cullen, J., N. Athi, M. Al-Jader, P. Johnson, A. Al-Shamma'a, A. Shaw, and A. El-Rasheed (2008), Multisensor fusion for on line monitoring of the quality of spot welding in automotive industry, *Measurement: Journal of the International Measurement Confederation*, *41*(4), 412 – 423.

Currier, J. H., D. E. Anderson, and D. W. V. citters (2011), A proposed mechanism for squeaking of ceramic-on-ceramic hips, *Wear*, *269*, 782–789.

Davis, J. R. (2005), *Gear materials, properties, and manufacture*, ASM International, Materials Park, Ohio.

De Oliveira, J. F. G., and D. A. Dornfeld (2001), Application of ae contact sensing in reliable grinding monitoring, *CIRP Annals - Manufacturing Technology*, *50*(1), 217–220.

Devine, J. (1998), Ultrasonic rx for medical device assembly, *Assembly*, *41*(11), 5pp–5pp.

Ding, C., and H. Peng (2003), Minimum redundancy feature selection from microarray gene expression data, *Proceedings of the 2003 IEEE Bioinformatics Conference*, pp. 523–528.

Donat, W., K. Choi, W. An, S. Singh, and K. Pattipati (2007), Data visualization, data reduction and classifier fusion for intelligent fault detection and diagnosis in gas turbine engines, *Proceedings of the ASME Turbo Expo*, *1*, 883 – 892.

Donat, W., K. Choi, W. An, S. Singh, and K. Pattipati (2008), Data visualization, data reduction and classifier fusion for intelligent fault diagnosis in gas turbine engines, *ASME Journal of Engineering for Gas Turbines and Power*, *130*(4), p 041,602–1–8.

Dornfeld, D. A., and E. Kannatey-Asibu (1980), Acoustic emission during orthogonal metal cutting, *International Journal of Mechanical Sciences*, *22*(5), 285–296.

Duda, R. O., P. E. Hart, and D. G. Stork (2006), *Pattern classification*, John Wiley & Sons, New York.

Emel, E., and E. Kannatey-Asibu Jr. (1988), Tool failure monitoring in turning by pattern recognition analysis of AE signals, *ASME Journal of Engineering for Industry*, *110*(2), 137–145.

Ertunc, H. M., and K. Loparo (2001), A decision fusion algorithm for tool wear condition monitoring in drilling, *International Journal of Machine Tools and Manufacture*, *41*(9), 1347–1362.

Fuentes, A., F. L. Litvin, D. Vecchiato, K. Yukishima, I. Gonzalez-Perez, and K. Hayasaka (2006), Reduction of noise of loaded and unloaded misaligned gear drives, *Computer Methods in Applied Mechanics and Engineering*, *195*(41-43), 5523–5536.

Fukunaga, K. (1972), *Introduction to statistical pattern recognition*, Academic Press, London.

Gaul, H., A. Shah, M. Mayer, Y. Zhou, M. Schneider-Ramelow, and H. Reichl (2010), The ultrasonic wedge/wedge bonding process investigated using in situ real-time amplitudes from laser vibrometer and integrated force sensor, *Microelectronic Engineering*, *87*(4), 537–542.

Goto, Y. (2001), Gear noise reduction of a transmission, *Toyota CRDL*, *36*, 1–2.

Govekar, E., A. Jeric, M. Weigl, and M. Schmidt (2009), Laser droplet generation: Application to droplet joining, *CIRP Annals - Manufacturing Technology*, *58*(1), 205–208.

Grewell, D. A. (1999), A prototype "expert" system for ultrasonic welding of plastics, *Plastics Engineering*, *55*(2), 33 – 37.

Guyon, I., and A. Elisseeff (2003), An introduction to variable and feature selection, *Journal of Machine Learning Research*, *3*(7-8), 1157–1182.

Hastie, T., R. Tibshirani, and J. Friedman (2001), *The elements of statistical learning: data mining, Inference, and prediction*, Springer, New York.

Hetrick, E., J. Baer, W. Zhu, L. Reatherford, A. Grima, D. Scholl, D. Wilkosz, S. Fatima, and S. Ward (2009), Ultrasonic metal welding process robustness in aluminum automotive body construction applications, *Welding Journal*, *88*(7), 149–158.

Hullermeier, E., and S. Vanderlooy (2010), Combining predictions in pairwise classification: An optimal adaptive voting strategy and its relation to weighted voting, *Pattern Recognition*, *43*(1), 128 – 142.

Iwata, K., and T. Moriwaki (1977), An application of acoustic emission measurement to in-process sensing of tool wear, *CIRP Annals*, *26*(1), 21 – 26.

Jayakumar, T., C. K. Mukhopadhyay, S. Venugopal, S. L. Mannan, and B. Raj (2005), A review of the application of acoustic emission techniques for monitoring forming and grinding processes, *Journal of Materials Processing Technology*, *159*, 48–61.

Jin, J. (2004), Individual station monitoring using press tonnage sensors for multiple operation stamping processes, *Transactions of the ASME. Journal of Manufacturing Science and Engineering*, *126*(1), 83–90.

Joshi, K. C. (1971), Formation of ultrasonic bonds between metals, *Welding Journal*, *50*(12), 840–848.

Kahn Jr., C. E., J. Kalpathy-Cramer, C. A. Lam, and C. E. Eldredge (2011), Accurate determination of imaging modality using an ensemble of text- and image-based classifiers, *Journal of Digital Imaging*, pp. 1 – 6.

Kannatey-Asibu Jr., E. (2009), *Principles of laser materials processing*, John Wiley & Sons, New York.

Kannatey-Asibu Jr, E., and D. A. Dornfeld (1982), A study of tool wear using statistical analysis of metal-cutting acoustic emission, *Wear*, *76*(2), 247–261.

Kanthababu, M., M. S. Shunmugam, and M. Singaperumal (2008), Tool condition monitoring in honing process using acoustic emission signals, *International Journal of Automation and Control*, *2*(1), 99–112.

Khmelev, V. N., A. N. Slivin, and A. D. Abramov (2008a), Research of parameter influence of ultrasonic welding process on conjuncture of polymeric thermoplastic materials, *2008 9th International Workshop and Tutorials on Electron Devices and Materials*, pp. 272–8.

Khmelev, V. N., A. N. Slivin, A. D. Abramov, and S. V. Levin (2008b), Weld strength test of thermoplastics obtained by ultrasonic welding, *2008 9th International Workshop and Tutorials on Electron Devices and Materials*, pp. 227–30.

Kim, J.-S., M.-C. Kang, B.-J. Ryu, and Y.-K. Ji (1999), Development of an online tool-life monitoring system using acoustic emission signals in gear shaping, *Inter-*

national Journal of Machine Tools and Manufacture, *39*(11), 1761–1777.

Kim, T. H., J. Yum, S. J. Hu, J. P. Spicer, and J. A. Abell (2011), Process robustness of single lap ultrasonic welding of thin, dissimilar materials, *CIRP Annals - Manufacturing Technology*, *60*, 17–20.

Kittler, J., and F. M. Alkoot (2003), Sum versus vote fusion in multiple classifier systems, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(1), 110–115.

Kittler, J., M. Hatef, R. P. W. Duin, and J. Matas (1998), On combining classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(3), 226–239.

Kong, C., R. Soar, and P. Dickens (2003), Characterisation of aluminium alloy 6061 for the ultrasonic consolidation process, *Materials Science and Engineering A*, *363*(1-2), 99 – 106.

Kong, C. Y., R. C. Soar, and P. M. Dickens (2005), A model for weld strength in ultrasonically consolidated components, *Proceedings of the Institution of Mechanical Engineers, Part C (Journal of Mechanical Engineering Science)*, *219*(C1), 83–91.

Krishnaswami, S. D., S. D. Kaatz, D. H. Hilderbrand, J. P. Hiatt, and P. E. Phelan (2001), Gear whine reduction for a new automatic transmission, *SAE Technical Paper*, p. 1506.

Krzanowski, J. (1990), A transmission electron microscopy study of ultrasonic wire bonding, *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, *13*(1), 176 – 181.

Kuncheva, L. I. (2002), A theoretical study on six classifier fusion strategies, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(2), 281–286.

Lederman, D., B. Zheng, X. Wang, X. Wang, and D. Gur (2011), Improving breast cancer risk stratification using resonance-frequency electrical impedance spectroscopy through fusion of multiple classifiers, *Annals of Biomedical Engineering*, *39*(3), 931 – 945.

Lee, S. S., T. H. Kim, S. J. Hu, W. W. Cai, and J. A. Abell (2010), Joining technologies for automotive lithium-ion battery manufacturing - a review, *Proceedings of the ASME 2010 International Manufacturing Science and Engineering Conference*.

Leem, C. S., D. Dornfeld, and S. Dreyfus (1995), Customized neural network for sensor fusion in on-line monitoring of cutting tool wear, *ASME Journal of Engineering for Industry*, *117*(2), 152 – 159.

Lehr, A. V., V. N. Khmelev, V. A. Nesterov, and A. N. Silvin (2011), Influence of surface friction on continuous ultrasonic welding of thin polymer films, *2011 12th*

International Conference and seminar of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM2011), pp. 281–284.

Li, X. (2002), A brief review: Acoustic emission method for tool wear monitoring during turning, *International Journal of Machine Tools and Manufacture*, *42*(2), 157–165.

Li, Z., X. Yan, C. Yuan, and L. Li (2010), Gear multi-faults diagnosis of a rotating machinery based on independent component analysis and fuzzy k-nearest neighbor, *Advanced Materials Research*, *108-111*, 1033 – 1038.

Lu, M.-C., and E. Kannatey-Asibu (2002), Analysis of sound signal generation due to flank wear in turning, *Transactions of the ASME. Journal of Manufacturing Science and Engineering*, *124*(4), 799 – 808.

Luo, R. C., and M. G. Kay (1990), A tutorial on multisensor integration and fusion, *IECON Proceedings (Industrial Electronics Conference)*, *1*, 707 – 722.

Lv, L., and L. Han (2008), Effects of bonding pressure on nonlinear dynamic characteristic of the ultrasonic wire bonding system, *2008 International Conference on Electronic Packaging Technology & High Density Packaging (ICEPT-HDP)*, pp. 1–5.

Mai, T. A., and A. C. Spowage (2004), Characterisation of dissimilar joints in laser welding of steel-kovar, copper-steel and copper-aluminium, *Materials Science and Engineering A (Structural Materials: Properties, Microstructure and Processing)*, *A374*(1-2), 224–33.

Manz, A. F. (1981), Welding arc sounds, *Welding Journal*, *60*(5), 23–27.

Masuzawa, N., and E. Ohdaira (2000), Information contained in the radiating ultrasound during ultrasonic welding, *Ultrasonics*, *38*(1-8), 609–613.

Matsuoka, S.-i. (1998), Ultrasonic welding of ceramics/metals using inserts, *Journal of Materials Processing Technology*, *75*(1-3), 259 – 265.

Matteson, A., R. Morris, and R. Tate (1993), Real-time gmaw quality classification using an artificial neural network with airborne acoustic signals as inputs, *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE*, *3*, 213–278.

Mitchell, L. D., and J. W. Daws (1982), Basic approach to gearbox noise prediction, *SAE Special Publications*, pp. 1–14.

Montanarini, M., and J. Steffen (1976), Investigations on laser welding, *IEEE Journal of Quantum Electronics*, *QE-12*(2), 126 – 129.

Muldoon, S. E., M. Kowalczyk, and J. Shen (2002), Vehicle fault diagnostics using a sensor fusion approach, *Proceedings of IEEE Sensors 2002. First IEEE International Conference on Sensors*, *2*, 1591–1596.

Nicosevici, T., R. Garcia, M. Carreras, and M. Villanueva (2004), A review of sensor fusion techniques for underwater vehicle navigation, *Oceans '04 MTS/IEEE Techno-Ocean '04*, *3*, 1600–1605.

Norgia, M., M. Annoni, A. Pesatori, and C. Svelto (2010), Dedicated optical instruments for ultrasonic welder inspection and control, *Measurement*, *43*(1), 39–45.

Pal, K., and S. K. Pal (2011), Monitoring of weld penetration using arc acoustics, *Materials and Manufacturing Processes*, *26*(5), 684–693.

Pan, C., P. Zhao, S. Du, and J. Wang (2009), Quality assessment of aluminum alloy resistance spot welding based on wavelet and statistic analysis, *2009 IEEE International Conference on Information and Automation*, pp. 1438–1442.

Pandit, S. M., and S.-M. Wu (1983), *Time series and system analysis with applications*, New York: John Wiley & Sons.

Papinniemi, A., J. C. S. Lai, Z. Jiye, and L. Loader (2002), Brake squeal: a literature review, *Applied Acoustics*, *63*, 391–400.

Petrakos, M., I. Kannelopoulos, J. A. Benediktsson, and M. Pesaresi (2000), The effect of correlation on the accuracy of the combined classifier in decision level fusion, *IEEE 2000 International Geoscience and Remote Sensing Symposium*, *6*, 2623–2625.

Phanikumar, G., K. Chattopadhyay, and P. Dutta (2011), Joining of dissimilar metals: Issues and modelling techniques, *Science and Technology of Welding and Joining*, *16*(4), 313–317.

Rabiner, L. R. (1989), A tutorial on hidden markov models and selected applications in speech recognition, *Proceedings of the IEEE*, *77*(2), 257–286.

Ravindra, H. V., Y. G. Srinivasa, and R. Krishnamurthy (1997), Acoustic emission for tool condition monitoring in metal cutting, *Wear*, *212*(1), 78–84.

Roli, F., G. Giacinto, and G. Vernazza (2001), Methods for designing multiple classifier systems, *Multiple Classifier Systems. Second International Workshop, Proceedings of MCS 2001*, pp. 78–87.

Ruiten, C. J. M. v. (1988), Mechanisms of squeal noise generated by trams, *Journal of Sound and Vibration*, *120*, 245–253.

Ruta, D., and B. Gabrys (2000), An overview of classifier fusion methods, *Computing*

and Information Systems, *7*(1), 1 – 10.

Saravanan, S., G. S. Yadava, and P. V. Rao (2006), Condition monitoring studies on spindle bearing of a lathe, *International Journal of Advanced Manufacturing Technology*, *28*(9-10), 993–1005.

Schenk, T., U. Uebel, and F. Wolfel (2003), Coroning: High performance gear honing without dressing, *Proceedings of the ASME Design Engineering Technical Conference*, *4 B*, 751–759.

Sewell, M. (2011), Ensemble learning, `http://www-typo3.cs.ucl.ac.uk/fileadmin/UCL-CS/research/Research_Notes/RN_11_02.pdf`.

Shackleton, J., and H. Rischall (1964), Laser welding for microelectronic interconnections, *IEEE Electronic Components Conference Proceedings*, pp. 145 – 151.

Shah, A., H. Gaul, M. Schneider-Ramelow, H. Reichl, M. Mayer, and Y. Zhou (2009), Ultrasonic friction power during al wire wedge-wedge bonding, *Journal of Applied Physics*, *106*(1), 013,503 (8 pp.).

Shin, K., M. J. Brennan, J. E. Oh, and C. J. Harris (2002), Analysis of disk brake noise using a two-degree-of-freedom model, *Journal of Sound and Vibration*, *254*, 837–848.

Siddiq, A., and E. Ghassemieh (2008), Thermomechanical analyses of ultrasonic welding process using thermal and acoustic softening effects, *Mechanics of Materials*, *40*(12), 982 – 1000.

Siddiq, A., and E. Ghassemieh (2009), Theoretical and fe analysis of ultrasonic welding of aluminum alloy 3003, *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, *131*(4), 0410,071 – 04100,711.

Spurr, R. T. (1961), Theory of brake squeal, *Institution of Mechanical Engineers – Proceedings*, pp. 33–52.

Sun, A., J. Kannatey-Asibu, E., and M. Gartner (1999), Sensor systems for real-time monitoring of laser weld quality, *Journal of Laser Applications*, *11*(4), 153 – 68.

Sun, A., E. Kannatey-Asibu Jr, and M. Gartner (2002), Monitoring of laser weld penetration using sensor fusion, *Journal of Laser Applications*, *14*(2), 114–121.

Sun, Z., and R. Karppi (1996), Application of electron beam welding for the joining of dissimilar metals: an overview, *Journal of Materials Processing Technology*, *59*(3), 257–267.

Tang, H., W. Hou, and S. Hu (2002), Forging force in resistance spot welding, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering*

Manufacture, *216*(7), 957 – 968.

Thouless, M. D., Z. Bin, and S. M. Ward (2006), Predicting the failure of ultrasonic spot welds by pull-out from sheet metal, *International Journal of Solids and Structures*, *43*(25-26), 7482–500.

Tolba, A. S., H. A. Khan, A. M. Mutawa, and S. M. Alsaleem (2010), Decision fusion for visual inspection of textiles, *Textile Research Journal*, *80*(19), 2094–2106.

Tsujino, J., T. Mori, and K. Hasegawa (1994), Welding characteristics of ultrasonic wire bonding using high-frequency vibration systems, *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*, *33*(5 B), 3048 – 3053.

Tsujino, J., T. Uchida, K. Yamano, N. Iwamoto, and T. Ueoka (1998), Welding characteristics of ultrasonic plastic welding using two-vibration-system of 90 khz and 27 or 20 khz and complex vibration systems, *Ultrasonics*, *36*(1-5), 67–74.

Tsujino, J., Y. Harada, and T. Sano (2001), Ultrasonic complex vibration welding systems of 100 khz to 200 khz with large welding tip area for packaging in microelectronics, *2001 IEEE Ultrasonics Symposium. Proceedings. An International Symposium (Cat. No.01CH37263)*, *vol.1*, 665–8.

Tsujino, J., M. Hongoh, and T. Ueoka (2002), Welding characteristics of 40 khz ultrasonic plastic welding system using fundamental and higher resonance frequency vibrations, *2002 IEEE Ultrasonics Symposium. Proceedings (Cat. No.02CH37388)*, *1*, 699–702.

Ueoka, T., and J. Tsujino (2002), Welding characteristics of aluminum and copper plate specimens welded by a 19 khz complex vibration ultrasonic seam welding system, *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*, *41*(5 B), 3237 – 3242.

Vatta, F. (1979), On the stick-slip phenomenon, *Mechanics Research Communications*, *6*, 203–208.

Volfson, L. (2006), Visible, night vision and IR sensor fusion, *2006 9th International Conference on Information Fusion*, pp. 1–4.

Wagner, G., F. Walther, T. Nebel, and D. Eifler (2003), Glass/glass joints by ultrasonic welding, *Glass Technology*, *44*(4), 152 – 155.

Wang, J. F., B. Chen, H. B. Chen, and S. B. Chen (2009), Analysis of arc sound characteristics for gas tungsten argon welding, *Sensor Review*, *29*(3), 240–9.

Wang, L. (2008), Feature selection with kernel class separability, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(9), 1534–1546.

Webster, J., I. Marinescu, R. Bennett, and R. Lindsay (1994), Acoustic emission for process control and monitoring of surface integrity during grinding, *CIRP Annals - Manufacturing Technology*, *43*(1), 299–304.

White, D. R. (2003), Ultrasonic consolidation of aluminum tooling, *Advanced Materials and Processes*, *161*(1), 64 – 65.

Xudong, C., and L. Xiaochun (2007), Investigation of heat generation in ultrasonic metal welding using micro sensor arrays, *Journal of Micromechanics and Microengineering*, *17*(2), 273–82.

Yu, J. (2011), Online tool wear prediction in drilling operations using selective artificial neural network ensemble model, *Neural Computing and Applications*, *20*(4), 473–485.

Yu, L., and H. Liu (2004), Efficient feature selection via analysis of relevance and redundancy, *Journal of Machine Learning Research*, *5*, 1205–1224, 1044700.

Yu, L., C. Ding, and S. Loscalzo (2008), Stable feature selection via dense feature groups, *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 803–811.

Yum, J., A. Kamouneh, W. Wang, and E. Kannatey-Asibu Jr. (2009), Acoustic emission condition monitoring of the coroning process, *Proceedings of the ASME International Manufacturing Science and Engineering Conference*, *2*, 157–163.

Yum, J., T. H. Kim, and E. Kannatey-Asibu Jr. (2012), Classifier fusion for acoustic emission based tool wear monitoring, *Proceedings of ASME 2012 International Manufacturing Science and Engineering Conference, MSEC2012-7380*.

Zhou, S., and J. Jin (2005), Automatic feature selection for unsupervised clustering of cycle-based signals in manufacturing processes, *IIE Transactions*, *37*(6), 569–584.

Zhou, S., N. Jin, and J. Jin (2005), Cycle-based signal monitoring using a directionally variant multivariate control chart system, *IIE Transactions*, *37*(11), 971–982.