The Thesis Committee for Haoran Xie
certifies that this is the approved version of the following thesis:

Identification of individual slit valves in semiconductor manufacturing fab based on their vibration signatures

APPROVED BY

SUPERVISING COMMITTEE:

Dragan Djurdjanovic, Supervisor

Michael Bryant

Identification of individual slit valves in semiconductor manufacturing fab based on their vibration signatures

by

Haoran Xie, B.E.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN ENGINEERING

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2014

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Dragan Djurdjanovic, who introduced me to this area. Without his guidance and support, this thesis would not be possible. I would also like to thank Dr. Michael Bryant for his review and comments on my draft.

I would also like to thank my colleagues in the LIMES Lab. In particular, Marcus Musselman who taught me a lot in how to apply theoretical knowledge in engineering practice and Merve Celen who spent her time helping revising my drafts.

Finally, I wish to thank my parents Jianguo Jia and Xiuyue Chen, as well as other members in my family, for all their support from afar.

Identification of individual slit valves in semiconductor manufacturing fab based on their vibration signatures

Haoran Xie, M.S.E
The University of Texas at Austin, 2014

Supervisor: Dragan Djurdjanovic

Slit valves play an important role in semiconductor manufacturing industry. They enable creation of a vacuum environment required for wafer processing. Due to the high volume of production in the modern semiconductor industry, slit valves experience severe degradation over their useful lifetime. If maintenance is not applied in due time, degraded valves may lead to defects in end products because of pressure loss and particle generation. In this thesis, we proposed methods for signal processing and feature extraction for analysis of slit valve vibration signatures. These methods would be used to demonstrate the ability of reliably, accurately and efficiently distinguish between each individual valve via a multi-class classification procedure. Such ability is a clear illustration of the feasibility of vibration based monitoring of the slit valve conditions.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Slit valves play an important role in semiconductor manufacturing industry. A slit valve is a gate that separates the process chamber and the transfer chamber of a semiconductor manufacturing tool, enabling creation of a vacuum environment required for wafer processing. Due to the high volume of production in the modern semiconductor industry, slit valves experience severe degradation over their useful lifetime. Once the valve is degraded, the vacuum condition inside the process chamber cannot be maintained, resulting in wafer defects. In addition, degraded slit valves may also lead to particle generation from degraded valve seals and guide ways, causing contamination in both the chamber and end products. Considering the very small margin of error in today's microelectronic manufacturing, there is an urgent need for establishing a monitoring and maintenance plan to set early alarms about degraded slit valves and prevent potential product flaws.

In majority of today's semiconductor fabrication plants (fabs), preventive maintenance of equipment is conducted following the reliability based maintenance (RBM) paradigm, i.e. based on the elapsed calendar time or usage and the statistical properties of the useful life distribution of the relevant

population of machines [1, 2]. Differences between individual machines in a population based on which RBM policies are postulated cause RBM to incur losses due to unnecessary maintenance of equipment that does not really need to be maintained, or due to unexpected failures of machines whose scheduled maintenance does not occur soon enough.

An alternative paradigm to address this drawback is the condition based maintenance (CBM), in which one builds and uses a connection between the condition of the individual pieces of equipment and sensor reading emitted by that machine. With such information, maintenance operations can be performed according to the actual working condition of the equipment, exactly when needed and exactly where needed.

In the past few years, CBM has drawn increasing attention in the semiconductor manufacturing industry. Advanced diagnostics and prognostics methods have been employed for various equipment and processes, such as etching equipment, chemical vapor deposition, chemical mechanical planarization and so on (more information will be covered in literature review chapter). Unfortunately, no CBM technology has ever been applied to slit valves. Therefore this study aims to explore the feasibility and aptness of application of CBM diagram on this important element in semiconductor manufacturing.

In the research presented in this thesis, we sensorized with accelerometers over 50 slit valves in a major domestic semiconductor fab and proposed methods for signal processing and feature extraction for analysis of their vibration signatures. These methods will be used to demonstrate the ability of

reliably, accurately and efficiently distinguish between each individual valve via a multi-class classification procedure. Such ability is a clear illustration of the feasibility of vibration based monitoring of the slit valve conditions, though such a study is outside of the scope of this thesis.

The remainder of this thesis is organized in the following manner. A literature review of CBM research and application in semiconductor manufacturing industry is given in Chapter 2. In Chapter 3, methods for processing of slit valve vibration signals and multi-segment classification of valves based on their signatures are described in detail. Chapter 4 gives an overview of the slit valve system operating in the major semiconductor fab as well as the data acquisition system, and shows the results of applying these methods to identification of 50 individual slit valves. Finally, conclusions and potential future work are presented in Chapter 5.

Chapter 2

Literature Review

In the past few years, condition based maintenance has drawn increasing attention in the semiconductor manufacturing industry. This chapter attempts to provide a brief review of the research and practices of CBM employed for various equipment and processes in semiconductor fabrication facilities.

(1) *Etching*: Shadmehr et al. [3] reported a technique combining principal component analysis (PCA) and neural network models to characterize the effect of process parameters on the optical emission and mass spectra of CHF_3/O_2 plasma. This technique was sensitive to changes in chamber contamination levels and proven to be a promising tool for real-time monitoring and control of reactive ion etching process. Kim et al. [4] built a neural time series model for reactive ion etching (RIE). The difference between predicted value of feed-forward neural network model and actual measured RIE response was an indication of potential equipment faults. He also employed Dempster-Shafer evidential reasoning and an inverse neural network model to infer the causes of faults and generate evidential belief. In [5], Hong et al. suggested use of modular neural networks as a new methodology for malfunction prognosis in reactive ion etching utilizing in-situ metrology data. He tested with

three types of malfunction scenarios for the RF power system and obtained satisfying results for predicting performance one run ahead. Spanos et al. [6] collected 5 real-time signals (position of the RF tune vane, position of the RF load coil, etc.) from a plasma-etch tool and fed them into individual time-series filters that produce multiple, cross-correlated identically, independently, normally distributed (IIND) residuals. These residuals were later combined by Hotelling's T^2 equations into a single real-time alarm signal that could be used in a statistical process control chart. Baek et al. [7] reported discovery of the correlation between the electron collision rate of plasma and chamber conditions after wet cleaning by using self-excited electron resonance spectroscopy in a dynamic random access memory gate etch process. This electron collision rate was able to identify small changes in chamber condition that could not be detected under conventional monitoring methods, suggesting a potent tool for chamber condition based maintenance and process control.

(2) *Chemical Vapor Deposition (CVD)*: In [8], a radio frequency (RF) impedance probe was integrated on a plasma-enhanced chemical vapor deposition chamber to explore the sensitivity of the reactor electrical characteristics on the events of input parameter variation. Such sensitivity information could be incorporated to chamber condition monitoring. Hopfe et al. [9] demonstrated a chemically sensitive CVD process control concept using non-invasive optical sensors based on diode laser spectroscopy in the near infra red range. This technology manifested its robustness and ability of enhancing CVD process performance predictability. Wu et al. [10] investigated the approach for

simultaneous fault detection and classification by using the principal component analysis method to detect designed faults of gas flow and RF parameters and classify the faults in PCA vector space on a plasma CVD tool. Yang [11] proposed to apply Bayesian Belief Network to investigate the relationship among process variables and wafer quality. The simulation on a CVD tool verified the feasibility of employing such network to perform wafer quality prognosis .

(3) *Chemical Mechanical Planarization (CMP)*: Tang et al. [12] conducted an experiment to investigate the correlation between the microscratches and the signal characteristics of acoustic emission (AE) generated during chemical mechanical planarization. The high sensitivity of AE signals to the CMP process state change indicated that this AE sensing technology may be used as a tool for in-situ microscratch detection and process monitoring in CMP. Lee [13] also introduced acoustic emission sensor to perform endpoint detection to insure desired polishing thickness of material during chemical mechanical planarization process. Chan et al. [14] presented the performance of a laser interferometry based In-situ Rate Monitor (ISRM) system on process control and monitoring in CMP . This non-contact, optical system relied on the ability to correlate the laser optical signal to the change of thickness of the polished film layer on wafers. It detected the endpoint and stopped the polishing process when the required amount of a polished film is removed. Wang et al. [15] studied timing correlation of multiple functional process variables (FPVs) such as coefficient of friction and polishing pad temperature for CMP process

condition monitoring and diagnosis. A nonlinear dynamics model was established to show the main and interaction effects among multiple FPVs through a design of experiment procedure. The extracted interaction patterns were proven helpful for detecting abnormal condition in CMP experiments with slurry contamination.

(4) *Lithography*: Facco et al. [16] demonstrated an image analysis based monitoring system for wafer quality evaluation after lithography. The semiconductor surface image was analyzed in different levels of resolutions via wavelet decompositions to extract commonly inaccessible features. Additionally, a two level nested PCA model was used for surface roughness monitoring, while a new strategy based on "spatial moving window" PCA was also proposed to analyze the shape of the patterned surface. Shen et al. [17] introduced fuzzy set theory to the analytic hierarchy process for the fault evaluations of lithography process. The new fuzzy analytic hierarchy process helped decision makers handle the fuzziness and ambiguity in real world environments and eased the process of root cause finding, hence improvement plan could be prioritized accordingly to allocate limited resources efficiently. Bao et al. [18] employed scatterometry to monitor the stepper focus and expose dose drift during lithography process. The measured spectrum from scatterometry was used to extract the current profile information which would be fed into a control parameter extractor to compute optimized control parameter values. Pampuri et al. [19] introduced a joint method of survival models theory with l_1 penalization techniques to predict the remaining lifetime of an equip-

ment with several meaningful process variables. This proposed methodology was validated with a set of observations of lithography steppers in the fabs, yielding promising preliminary results.

(5) *Handling and Cluster Tools*: In [20], the status of robot arm of cluster tools were captured by a CCD camera and processed by the image recognition technology. Then a series of arm location data was saved into a SPC chart and maintenance engineers were informed once the arm location violates the established limit lines. Guan et al. [21] proposed to use a new data driven CBM (DD-CBM) framework for test handlers in a domestic high-volume integrated circuit manufacturing environment. The DD-CBM incorporated condition monitoring variables into a control chart to estimate machine health and dynamically updated the control limit based on the equipment's operation conditions. Costuros [22] reported use of wavelet transform on motor torque signals sampled from a blade in transfer chamber. The wavelet transform coefficients were applied to construct the channel capacity of the signal (concept borrowed from Shannon's information theory) for fault diagnostics.

Though a number of breakthroughs about applications of CBM in the semiconductor industry are reported in literature, unfortunately no CBM technology has ever been applied to slit valves, in spite of their great importance in wafer processing. Some reasons explaining this may be lack of appropriate way in sensing vibrations from slit valves on typical tools, as well as complexity and non-stationarity of slit valve vibrations.

This thesis focuses on the vibration signatures during slit valve travels

(both opening and closing) and attempts to extract descriptive features that are able to characterize individual slit valves through a novel signal processing and feature extraction procedure. A multi-class study will be performed to narrow the feature set down to several most discriminative ones and these features will be regarded as candidates for condition monitoring of slit valves during semiconductor manufacturing.

Chapter 3

Background on Mathematical Methods

Due to the non-linearity and complexity of valve motions, the vibration signals emitted by slit valves are highly non-stationary. This means that frequency contents of those signals vary significantly over time [23], which invokes the need for non-stationary signal analysis tools, such as Cohen's class of time-frequency transform method [24]. Cohen's general class of time-frequency distribution (TFD) for the signal $x(t)$ can be described as

$$C_x(t,\omega) = \frac{1}{4\pi^2}\int\int\int \phi(\theta,\tau)x(u+\frac{\tau}{2})x^*(u-\frac{\tau}{2})e^{-j(\theta t+\tau\omega-\theta u)}d\theta d\tau du \quad (3.1)$$

where $x(t)$ and $x^*(t)$ denote the relevant signal and its complex conjugate respectively, while $\phi(\theta,\tau)$ is the so-called kernel function of the TFD. The kernel determines mathematical properties of the resulting TFD, such as realness of the resulting distribution $C_x(t,\omega)$, time and frequency support properties, upholding of the time and frequency marginals, cross term suppression, as well as group delay and instantaneous frequency properties [25][1]. Fulfillment of these properties enables interpretation of the function $C_x(t,\omega)$ defined by Transformation (3.1) as a joint 2-dimensional distribution of signal energy in time and

[1]Definitions of those properties as well as mathematical constraints on the kernels which are necessary to achieve them are summarized in [26].

frequency domains. In this study, we opted to use the binomial kernel, which can be considered to be the most advanced signal-independent reduced interference distribution (RID) kernel [27]. Its expression in the ambiguity domain is:

$$\phi(\theta, \tau) = cos^{|\tau|}(\theta) \tag{3.2}$$

The RID character of the binomial kernel and its consequent ability to suppress cross terms, which are inherently present in TFDs, are highly desirable, since it is well documented that cross terms can hamper signal interpretation and classification based on TFDs riddled with cross terms [24]. On the other hand, signal independent nature of the binomial kernel is also important because of the sheer volume of data considered in this study. Namely, signal dependent kernels, such as those in [28, 29], would be computationally infeasible in the realm of multiple vibration readings from a large number of valves collected at sampling rates in the kHz ranges, which is what we dealt with in this research.

After signal processing stage, numerous features were extracted from the resulting vibration TFDs and later used as inputs for the multi-class classifier built to recognize individual valves based on their vibration signatures. These features can be partitioned into three categories: timing based features, time domain based features and time-frequency distribution based features.

Timing based features consist of times required to complete various portions of the valve movement. In this thesis, we just recorded time intervals between valve travel start and end signals though a much more elaborate set of timing features describing the valve motion in finer detail could be

pursued, if adequate discrete control signals are available to reliably delineate these portions of the valve motions. Such timing based features are a good indication of machine working condition since the time required to accomplish various portions of the designated motion for a machine with moving parts will usually vary as the condition of that machine drifts.

Time-domain based features are calculated from the time domain waveforms of the signals and include signal entropy, mean signal energy, median energy of the signal, variance, skewness and kurtosis of the signal energy. All time-domain based features used in this study and the formulae according to which they are calculated are listed in Table 3.1. These time-domain based features are intuitive and extensively used as vibration based measurements of the working condition in many previous machine monitoring works [30, 31].

Feature	Formula
Signal energy	$SumE = \sum_{t=t_1}^{t_n} x(t)^2$
Signal entropy	$H = \sum_{t=t_1}^{t_n} -\frac{x(t)^2}{SumE} log \frac{x(t)^2}{SumE}$
Entropy of signal energy	$HE = \sum_{t=t_1}^{t_n} -\frac{x(t)^4}{\sum_{t=t_1}^{t_n} x(t)^4} log \frac{x(t)^4}{\sum_{t=t_1}^{t_n} x(t)^4}$
Maximal energy	$\max x(t)^2$
Time of maximal energy	$\arg\max_t x(t)^2$
Minimal energy	$\min x(t)^2$

Table 3.1 – *Continued on next page*

Feature	Formula
Time of minimal energy	$\arg\min_t x(t)^2$
Maximal amplitude	$\max x(t)$
Time of maximal amplitude	$\arg\max_t x(t)$
Minimal amplitude	$\min x(t)$
Time of minimal amplitude	$\arg\min_t x(t)$
Median energy	middle value of $x(t)^2$
Mean energy	$E[x^2] = \frac{1}{n}\sum_{t=t_1}^{t_n} x(t)^2$
Variance of energy	$Var[x^2] = \frac{1}{n}\sum_{t=t_1}^{t_n}(x(t)^2 - E[x^2])^2$
Skewness of energy	$Skewness = \frac{\frac{1}{n}\sum_{t=t_1}^{t_n}(x(t)^2 - E[x^2])^3}{(Var[x^2])^{\frac{3}{2}}}$
Kurtosis of energy	$Kurtosis = \frac{\frac{1}{n}\sum_{t=t_1}^{t_n}(x(t)^2 - E[x^2])^4}{(Var[x^2])^2}$

Table 3.1: Time-Domain Features

Time-frequency distribution based features used in this study consist of the so-called time-frequency distribution moments, entropy and several signal energy related features calculated from the binomial distributions of the slit

valve vibrations. Following [24], the moment terms $E[t^p\omega^q]$ can be calculated from

$$E[t^p\omega^q] = \sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m} t^p\omega^q C_x(t,\omega) \tag{3.3}$$

where $C_x(t,\omega)$ is the time-frequency distribution of the signal $x(t)$. According to [32] and [33], moments of low orders can be used to approximate the general characteristics of TFDs and successfully accomplish clarification based on those TFDs. Moments up to 3 were used in this study to describe the time-frequency patterns observed in the slit valve vibrations. Signal entropy based on its binomial TFD $C_x(t,\omega)$ can be calculated as

$$H = \sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m} -\frac{C_x(t,\omega)}{SumE} log \frac{C_x(t,\omega)}{SumE} \tag{3.4}$$

In addition, various signal energy related features can also be extracted from their binomial TFDs, including maximal energy, as well as the time instance and frequency at which that maximal energy appeared in the TFD. A complete list of time-frequency domain features used in this study is summarized in Table 3.2.

Feature	Formula
TFD moments of order up to 3	$E[t^p\omega^q]$ with $p+q \leq 3$
Entropy	$H = \sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m} -\frac{C_x(t,\omega)}{SumE} log \frac{C_x(t,\omega)}{SumE}$
Maximal energy	$\max C_x(t,\omega)$

Table 3.2 – *Continued on next page*

Feature	Formula
Time of maximal energy	$\arg\max_t C_x(t,\omega)$
Frequency of maximal energy	$\arg\max_\omega C_x(t,\omega)$
Minimal energy	$\min C_x(t,\omega)$
Time of minimal energy	$\arg\min_t C_x(t,\omega)$
Frequency of minimal energy	$\arg\min_\omega C_x(t,\omega)$
Median energy	middle value of $C_x(t,\omega)$
Mean energy	$E[C_x(t,\omega)] = \frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m} C_x(t,\omega)$
Variance of energy	$Var[C_x(t,\omega)] = \frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m}(C_x(t,\omega) - E[C_x(t,\omega)])^2$
Skewness of energy	$Skewness = \frac{\frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m}(C_x(t,\omega)-E[C_x(t,\omega)])^3}{(Var[C_x(t,\omega)])^{\frac{3}{2}}}$
Kurtosis of energy	$Kurtosis = \frac{\frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m}(C_x(t,\omega)-E[C_x(t,\omega)])^4}{(Var[C_x(t,\omega)])^2}$

Table 3.2: Time-Frequency Domain Features

Within this plethora of signal features, not all will provide useful information for valve recognition. To the contrary, the so-called curse of dimensionality plagues the performance of classifiers based on such highly dimensional

feature space and a feature reduction process is needed to improve the classification process by removing the redundant and irrelevant features [34]. The feature reduction procedure employed in this study is closely associated with the classification strategy, and therefore a brief description of the classification framework will be presented first.

Individual valve differentiation is a multi-class classification problem. One straightforward method dealing with this type of problems is to train a universal multi-class classifier that could take care of all of the classes simultaneously. In the recent literature, we see such approaches in [35, 36]. However, this strategy requires that the information needed to separate classes 1 and 2 also be suitable differentiating classes 2 and 3 and all other possible class-pairs. Unfortunately, this assumption turns out to be overly constraining, especially when the number of classes involved becomes very large, as it is in the case considered in this study (we are dealing with dozens of valves and hence must realize a multi-class classification problem involving dozens of classes).

In this study we adopted a classification approach introduced in [37] and [38], which is based on repeated pairwise classifications that successively distinguishes between all possible class-pairs in a multi-class classification problem. This method increases classification accuracy by enabling pairwise distinction between any given pair of classes, using a feature set specifically selected to optimize that particular classification problem. Thus, this approach uses a divide and conquer paradigm and a variable feature set for the multi-class classification problem, rather than utilizing one universal feature set to tackle

the entire multi-class classification.

More specifically, this approach decomposes an n-class problem into $\binom{n}{2}$ one-against-one pairwise classification subproblems, with a specific classifier being trained for each one of those subproblems using the most discerning features for that subproblem. When a query signal is to be classified, it is passed through all the pairwise classifiers, each providing a vote for one of the two classes involved. Eventually, all the outputs of these $\binom{n}{2}$ subclassifiers are aggregated and the query signal is assigned to the class receiving the most votes. Figure 3.1 better illustrates the complete process.

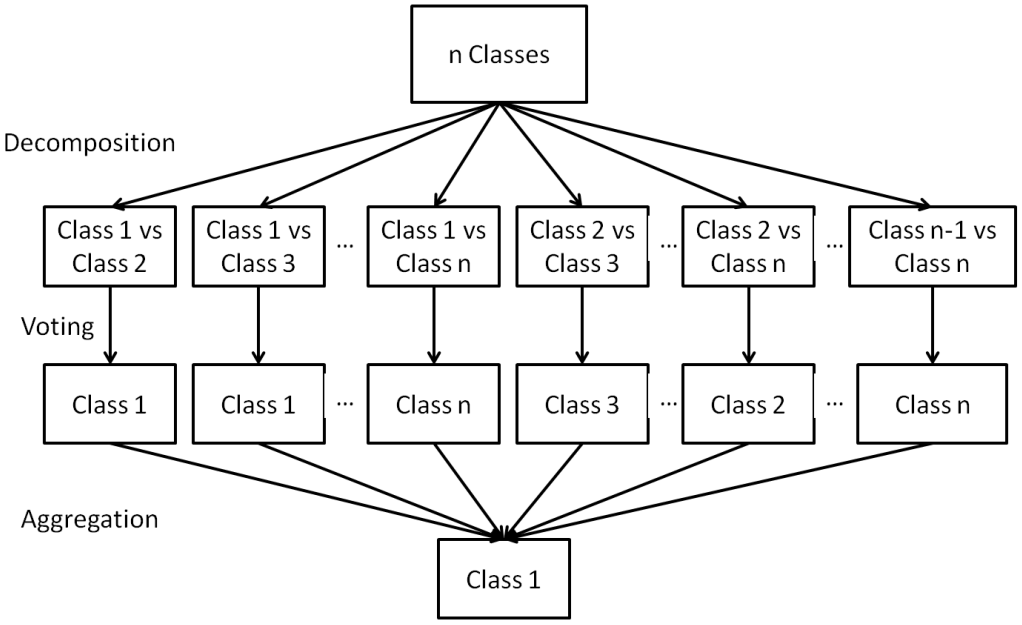

Figure 3.1: Pairwise Classification Process

For each of the pairwise classification problems, the features relevant

for that problem were selected from the exhaustive feature set described earlier in this chapter, using an inter/intra class distance ratio in the feature space (similar with the Dunn Index [39]). Namely, given an n-class problem and an exhaustive feature set consisting of m features, the selection process is conducted as follows:

- For each possible pair of classes (ω_i,ω_j), $i \in \{1, 2, ...n\}$, $j \in \{1, 2, ..., n\}$, $i \neq j$, and each feature l, $l \in \{1, 2, ...m\}$, calculate the maximum intra-class distance $d(\omega_i, \omega_j, l)$ between all points in class ω_i and the minimum interclass distance $D(\omega_i, \omega_j, l)$ between points in class ω_i and class ω_j. Then, the inter/intra class distance ratio is obtained as

$$r(\omega_i, \omega_j, l) = \frac{D(\omega_i, \omega_j, l)}{d(\omega_i, \omega_j, l)} \tag{3.5}$$

- For a specific pair of classes (ω_i,ω_j), $i \in \{1, 2, ...n\}$, $j \in \{1, 2, ..., n\}$, $i \neq j$, features with distance ratios greater than 1 are labeled, and among them, the one having the greatest number of distance ratios above 1 for all other subproblems is selected as the feature for this class-pair (ω_i,ω_j). In this way, the accuracy of each classifier is maintained and the number of features is effectively cut down by the choice of features that can be considered as generally versatile (applicable to other pairwise classification subproblems).

Another key point in a classification problem is the selection of the classification algorithm. In this study, the k-Nearest Neighbor (kNN) classifier

is chosen to discriminate between different valves. kNN is a non-parametric classification algorithm that determines the class memberships of an unknown testing point according to the k closest training points in the feature space [40]. Because of its simplicity, kNN is a highly suitable classification algorithm for a classification strategy based on numerous pairwise classification problems, such as the one encountered here.

Chapter 4

Results of Signal Processing and Identification of Slit Valves

In this chapter, the signal processing, feature extraction, feature selection and classification methods described in the previous chapter will be applied to accomplish recognition of individual slit valves in a major domestic semiconductor manufacturing facility. Before that, an overview of the slit valve and data acquisition system will be provided first.

4.1 Overview of Slit Valve and Data Acquisition System

All valves considered in this study were pneumatic valves of identical design, produced by the same manufacturer (i.e. nominally, they are supposed to be identical). When the valve closes, a pneumatic cylinder drives the valve plate down guide rails. Near the bottom of the valve travel, the valve head encounters a cam, which directs the valve head motion from downward to forward. At the end of the valve motion, the valve head makes contact with a base plate to create the seal. The motion is reversed when the valve opens and is schematically illustrated in Fig 4.1.

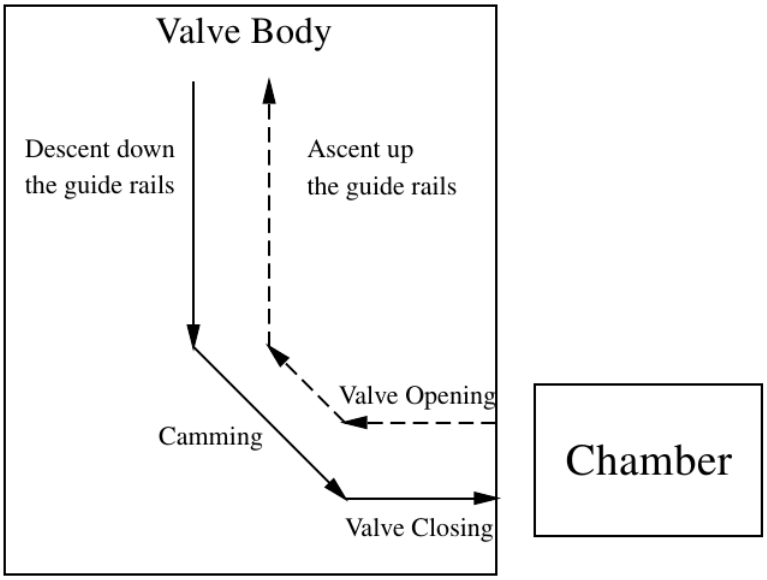

Figure 4.1: Illustration of Valve Operation and Motion During Opening (dashed line) and Closing (solid line)[2]

The data acquisition system was based on the sbRIO-9636 embedded control and acquisition device from National Instruments [41]. The valve vibrations were captured using 3-dimensional (3D) accelerometers ADXL327 [42] mounted on the valve housing and their readings were sampled at 5 kHz. The results in this thesis are based on the Root Mean Square (RMS) of the 3D vibration signals provided by this sensor, as well as the original 3D signals, since specific vibration directions can be used to characterize and monitor specific portions of the valve motion (descent down/ascent up the rail, camming, valve

[2]Due to the proprietary nature of the valve design, its detailed blueprint and characteristics could not be shown here.

closing/opening).

In addition, since control signals from the valve were not available, the timing information about valve motion was obtained using 2 photo-resistors placed over status lights on the valve housing. One of those lights indicated valve body passing by a fixed point near the top of the guide rail, while the other indicated the valve body passing a fixed point near the end of the valve movement, when the valve seal makes contact with the chamber wall. Signals from the photoresistors were used as automatic markers for valve motions and for normalization of the valve travel time, which reduced variability of valve signatures caused by variation in the valve travel times. Normalization was accomplished as follows. During valve closing motion, turning on of the light near the top of valve motion denoted normalized time 0, while turning off of the light near the bottom of the valve motion signified normalized time 1. Conversely, during the valve opening motion, normalized time 0 occurred when status light near the bottom of the valve motion turned on, while normalized time 1 occurred when the status light at the top of the valve motion path turned off. One should note that valve movement pre and post these light indicators were also collected, resulting in valve vibrations for normalized times -0.2 to 1.2. Each vibration signal was correspondingly divided into 3 stages in the following manner. During valve closing, the three segments were valve motion before the valve up signal (normalized time -0.2 to 0), valve motion between the valve up and valve down signals (normalized time 0 to 1) and valve motion after the valve down signal (normalized time 1 to 1.2). Conversely,

during valve opening, we observed valve motion before the valve down signal turned on (normalized time -0.2 to 0), valve motion between the valve down and valve up signals (normalized time 0 to 1) and valve motion after the valve up signal turned off (normalized time 1 to 1.2). Accordingly, features described in previous chapter were extracted using data points from these 3 stages separately or from the whole period (normalized times -0.2 to 1.2), yielding a feature library consisting of 1122 features[3].

From each of the 50 valves, vibration signals corresponding to 10 openings and 10 closings were collected, yielding the total of 1000 signals. From each of those signals, timing based, time domain and time-frequency domain characteristics described in chapter 3 were extracted, based on which this 50-class classification study was conducted [4]. The training set was constructed by randomly picking half of the recordings from each valve, while the remaining half was used for testing. Such selection of training and testing sets was repeated for 100 times to objectively evaluate the performance of the proposed classification method, independently of the choice of the training set.

[3]From each of the 4 segments, we got 16 time-domain based features and 19 time-frequency domain based features, for each of the 3 directions of vibrations, as well as the vibration RMS. Plus the movement time of both closing and opening motions, this yielded 1122 features.

[4]One should note that our classification method transformed this 50-class classification into 1225 pairwise classification problems, which were solved using the kNN classification algorithm.

4.2 Results of Slit Valve Identification

Classification accuracies based on different feature categories and movement directions of the valves are listed in Table 4.1. It is obvious that the use of time-domain features and time-frequency domain features greatly outperformed the accuracy of classification based on the timing features alone (performance of the timing based features is actually quite poor). In addition, it is visible that the elaborate feature extraction and classification methods introduced in this thesis enabled improved classification results via fusion of features from various domains and valve motion stages, yielding the best performance when features from all domains and all motion stages are included. In this case, perfect results were obtained in 7 out of 100 tests and an accuracy average of 98.74% was maintained. Furthermore, the consistency of this method in differentiating individual slit valves was evident in the low variance ($7.58 * 10^{-5}\%$) of accuracies during the 100 repetitions corresponding to different trainings. From these results, one can conclude that vibration patterns of slit valves are so distinctive that they identify individual valves, similarly how human's speech can be used to identify an individual. Advanced time-frequency analysis and sophisticated feature extraction methods introduced in this study were able to expose those discerning vibration patterns and enable almost perfect valve identification.

	Timing Features	Time Domain Features	TFD Features	Fusion of Timing, Time Domain, TFD Features
Opening	22.88%	96.88%	96.99%	98.18%
Closing	27.54%	96.64%	96.35%	97.93%
Opening & Closing	60.58%	97.15%	98.73%	98.74%

Table 4.1: Pairwise Classification Results

Delving deeper into how this remarkable performance was achieved, we can notice that many of the 1024 features generated from feature extraction stage were never selected or were only rarely selected for pairwise classification subproblems, while a few others, on the other hand, happened to be selected and used more frequently. For the case of jointly using opening & closing valve motions, and features from all domains (highest classification accuracy, as per in Table 4.1), top 10 most frequently used features are identified in the pareto chart shown in Fig 4.2, and are further explained in Table 4.2. Note that in the column describing portions of the signal in Table 4.2, "Pre" denotes portion of the signals from normalized times -0.2 to 0, "Between" denotes portion of the vibration signals between normalized times 0 to 1, "Post" denotes portion of the vibration signals from normalized times 1 to 1.2) and "Entire Signal" denotes the entire vibration signal obtained during a given valve motion (normalized times -0.2 to 1.2).

As can be seen in Fig 4.2, median energy in the time-frequency domain of the RMS of the entire vibration signal is the feature used most commonly. It was used in 800 out of 1225, or more than 68% of the pairwise classification subproblems. Mean energy in the time-frequency domain calculated for the RMS of the vibration signals during normalized times 0 to 1 ranks second in the chart and covers 100 out of 1225 pairwise subproblems, which is about 8.16% of subproblems. Besides these 2 features, none of the remaining ones was used in more than 3.1% classification subproblem, which implies that the aforementioned 2 features can be considered to be the most discerning features

in differentiating the valves from each other. Consequently, these features can be used as the top candidates for the use in the long-term valve degradation monitoring of slit valves in order to enable their condition based maintenance.

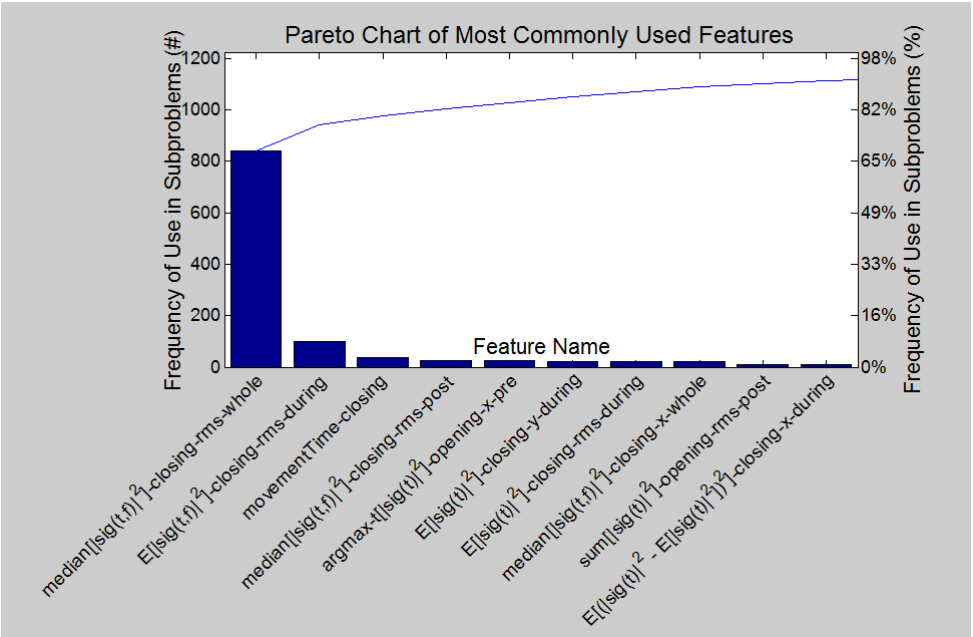

Figure 4.2: Pareto Chart of Most Commonly Used Features in the classification scheme that led to 98.74% accuracy (highest accuracy achieved in this study)

Feature Name	Detail				Usage
	Feature	Direction of valve motion	Channel of the vibration signal	Portion of the signal	
$median[\|sig(t,f)\|^2]$-closing-rms-whole	Median energy in the time-frequency domain	Closing	RMS	Entire Signal	68.65%
$E[\|sig(t,f)\|^2]$-closing-rms-during	Mean energy in the time-frequency domain	Closing	RMS	Between	8.16%
$movementTime$-closing	Movement time	Closing			3.02%
$median[\|sig(t,f)\|^2]$-closing-rms-post	Median energy in the time-frequency domain	Closing	RMS	Post	2.12%
$argmax - t[\|sig(t)\|^2]$-opening-x-pre	Position of maximum energy in the time domain	Opening	X-direction	Pre	2.04%
$E[\|sig(t)\|^2]$-closing-y-during	Mean energy in the time domain	Closing	Y-direction	Between	1.80%
$E[\|sig(t)\|^2]$-closing-rms-during	Mean energy in the time domain	Closing	RMS	Between	1.71 %
$median[\|sig(t,f)\|^2]$-closing-x-whole	Median energy in the time-frequency domain	Closing	X-direction	Entire Signal	1.63 %
$sum[\|sig(t)\|^2]$-opening-rms-post	Total energy in the time domain	Opening	RMS	Post	0.90 %
$E[(\|sig(t)\|^2 - E[\|sig(t)\|^2])^2]$-closing-x-during	Variance of energy in the time domain	Closing	X-direction	Between	0.82 %

Table 4.2: Most Commonly Used Features in the classification scheme that led to 98.74% accuracy

In addition, the 10 most commonly utilized features listed in Fig 4.2 and Table 4.2 happened to be selected for use in around 90% of all the pairwise classification subproblems, with 51 features covering all 1225 subproblem pairs. It's an indication that valve identification information could be stored in 51 or even fewer metrics instead of the whole vibration signature from 3 channels, suggesting that efficient valve monitoring could be accomplished without excessive data storage requirements.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

In this thesis, we conducted a classification study in which numerous slit valves used in a major semiconductor manufacturing fab were individually recognized using their vibration signatures. By applying an advanced time-frequency signal processing and feature extraction method, the slit valve vibration signals were transformed into a set of descriptive metrics that were used for characterization of each individual valve's vibration patterns. A kNN based multi-class classification approach was used to recognize the source of the unclassified vibration signals, leading to almost perfect recognition of 50 individual valves in the fab. The few misclassification occurred within the "quiet" valves, whose motion did not awaken a lot noticeable vibrations, and thus the valves ended up being confusing. Moreover, several features were found to be efficient in discriminating great majority of valves and these features are regarded as promising parameters for monitoring of the working condition and health degradation of slit valves considered in this manuscript. Such monitoring study is outside the scope of this thesis and will be considered in future publications.

5.2 Future Work

More efforts will be taken to acquire working condition information of slit valves in fabrication facilities based on their vibration signatures. Potential future work can further focus on the following three aspects:

(1) *Long-term Degradation Analysis*: Based on the results so far, several features are picked up to be efficient in showing difference between slit valves. These features can be further tested whether they are also able to be used for monitoring of the working condition and health degradation of slit valves in long-term cycles. The experiment can be conducted by intentionally changing the operation conditions of slit valves and determining if the variation trends of these features reflect the varying working status as well as how sensitive they are with those changes.

(2) *Fault Localization and Characterization*: As mentioned in previous chapter, valve motion initiation and ending were detected using 2 photoresistors in this study. If the valve motion control signals are available or more markers are utilized, the analysis procedures described in this study can be applied to different portions of the valve motion individually and help localize the fault occurrence precisely. At the same time, corresponding direction of accelerometer signals can be used for a specific portion of the valve motion to ensure more accurate results, since characterization information may be lost or blurred during the fusion of 3 direction signals or investigation of the whole traveling period.

(3) *Establishment of Condition Monitoring Frame*: Once the long-term analysis is finished, a set of features is determined as slit valve health index and can be fed into the SPC based condition monitoring framework. Feature values extracted from the earliest a few number of tests (e.g., first 20 cycles) can be used to establish the center line and control limits of the SPC charts. If the current input falls outside of the valid range, alarm is triggered to indicate working condition change of the slit valve under examination. Besides SPC charts, artificial intelligence techniques such as neural network can be employed to determine current working condition or even predict potential future failure and remaining useful lifetime based on the current feature values and usage history.

Bibliography

[1] S. Fulton and M. Kim, "Ismi consensus preventive and predictive maintenance vision ver. 1.1," *International SEMATECH Manufacturing Initiative*, 2007.

[2] M. Cholette, M. Celen, D. Djurdjanovic, and J. Rasberry, "Condition monitoring and operational decision making in semiconductor manufacturing," 2013.

[3] R. Shadmehr, D. Angell, P. B. Chou, G. S. Oehrlein, and R. S. Jaffe, "Principal component analysis of optical emission spectroscopy and mass spectrometry: Application to reactive ion etch process parameter estimation using neural networks," *Journal of the Electrochemical Society*, vol. 139, no. 3, pp. 907–914, 1992.

[4] B. Kim and G. S. May, "Real-time diagnosis of semiconductor manufacturing equipment using a hybrid neural network expert system," *Components, Packaging, and Manufacturing Technology, Part C, IEEE Transactions on*, vol. 20, no. 1, pp. 39–47, 1997.

[5] S. J. Hong, "Real-time malfunction diagnosis and prognosis of reactive ion etching using neural networks," 2003.

[6] C. J. Spanos, H.-F. Guo, A. Miller, and J. Levine-Parrill, "Real-time statistical process control using tool data [semiconductor manufacturing]," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 5, no. 4, pp. 308–318, 1992.

[7] K. H. Baek, Y. Jung, G. J. Min, C. Kang, H. K. Cho, and J. T. Moon, "Chamber maintenance and fault detection technique for a gate etch process via self-excited electron resonance spectroscopy," *Journal of Vacuum Science & Technology B*, vol. 23, no. 1, pp. 125–129, 2005.

[8] S. Raoux, K. Liu, X. Guo, and D. Silvetti, "In-situ rf diagnostic for pecvd process control," in *MRS Proceedings*, vol. 502, Cambridge Univ Press, 1997.

[9] V. Hopfe, D. Sheel, C. Spee, R. Tell, P. Martin, A. Beil, M. Pemble, R. Weiss, U. Vogt, and W. Graehlert, "In-situ monitoring for cvd processes," *Thin Solid Films*, vol. 442, no. 1, pp. 60–65, 2003.

[10] H. Wu, C. Chang, B. Chen, C. Lee, C. Chang, J. Ko, M. Zhou, and M. Liang, "Fault detection and classification of plasma cvd tool," in *Semiconductor Manufacturing, 2003 IEEE International Symposium on*, pp. 123–125, IEEE, 2003.

[11] L. Yang and J. Lee, "Bayesian belief network-based approach for diagnostics and prognostics of semiconductor manufacturing systems," *Robotics and Computer-Integrated Manufacturing*, vol. 28, no. 1, pp. 66–74, 2012.

[12] J. Tang, D. Dornfeld, S. K. Pangrle, and A. Dangca, "In-process detection of microscratching during cmp using acoustic emission sensing technology," *Journal of electronic materials*, vol. 27, no. 10, pp. 1099–1103, 1998.

[13] D.-E. Lee, I. Hwang, C. M. Valente, J. Oliveira, and D. A. Dornfeld, "Precision manufacturing process monitoring with acoustic emission," *International Journal of Machine Tools and Manufacture*, vol. 46, no. 2, pp. 176–188, 2006.

[14] D. A. Chan, B. Swedek, A. Wiswesser, and M. Birang, "Process control and monitoring with laser interferometry based endpoint detection in chemical mechanical planarization," in *Advanced Semiconductor Manufacturing Conference and Workshop, 1998. 1998 IEEE/SEMI*, pp. 377–384, IEEE, 1998.

[15] H. Wang, X. Zhang, A. Kumar, and Q. Huang, "Nonlinear dynamics modeling of correlated functional process variables for condition monitoring in chemical–mechanical planarization," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 22, no. 1, pp. 188–195, 2009.

[16] P. Facco, F. Bezzo, M. Barolo, R. Mukherjee, and J. A. Romagnoli, "Monitoring roughness and edge shape on semiconductors through multiresolution and multivariate image analysis," *AIChE journal*, vol. 55, no. 5, pp. 1147–1160, 2009.

[17] C.-w. Shen1/2, M. J. Cheng3/4, and C.-W. Chen, "A fuzzy ahp-based fault diagnosis for semiconductor lithography process," 2011.

[18] J. Bao and C. J. Spanos, "A simulation framework for lithography process monitoring and control using scatterometry," in *AEC/APC Symposium XIII*, 2001.

[19] S. Pampuri, A. Schirru, C. De Luca, and G. De Nicolao, "Proportional hazard model with 1 penalization applied to predictive maintenance in semiconductor manufacturing," in *Automation Science and Engineering (CASE), 2011 IEEE Conference on*, pp. 250–255, IEEE, 2011.

[20] W.-R. Jong and T.-W. Lin, "Statistical process control for e-diagnostic prediction of cluster-tool equipment," in *Industrial Electronics Society, 2007. IECON 2007. 33rd Annual Conference of the IEEE*, pp. 2916–2921, IEEE, 2007.

[21] T. Guan, Y. C. Kuang, M. Ooi, X. G. Cheah, Y. S. Tan, and S. Demidenko, "Data-driven condition-based maintenance of test handlers in semiconductor manufacturing," in *Electronic Design, Test and Application (DELTA), 2011 Sixth IEEE International Symposium on*, pp. 189–194, IEEE, 2011.

[22] T. V. Costuros, "Application of communication theory to health assessment, degradation quantification, and root cause diagnosis," 2013.

[23] L. Wang, M. G. Mehrabi, and E. Kannatey-Asibu, "Tool wear monitoring in machining processes through wavelet analysis," *TRANSACTIONS-NORTH AMERICAN MANUFACTURING RESEARCH INSTITUTION OF SME*, pp. 399–406, 2001.

[24] L. Cohen, *Time-frequency analysis*, vol. 778. Prentice Hall PTR New Jersey, 1995.

[25] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *Signal Processing, IEEE Transactions on*, vol. 40, no. 2, pp. 402–412, 1992.

[26] T. A. C. M. Classen and W. F. G. Mecklenbrauker, "The wigner distribution - a tool for time-frequency signal analysis- part i: Continuous-time signals," *Phillips Journal of Research*, vol. 35, no. 3, pp. 217–250, 1980.

[27] A. Papandreou-Suppappola, *Applications in time-frequency signal processing*, vol. 10. CRC press, 2002.

[28] D. L. Jones and R. G. Baraniuk, "An adaptive optimal-kernel time-frequency representation," *Signal Processing, IEEE Transactions on*, vol. 43, no. 10, pp. 2361–2371, 1995.

[29] M. Coates and W. Fitzgerald, "Regionally optimised time–frequency distributions using finite mixture models," *Signal processing*, vol. 77, no. 3, pp. 247–260, 1999.

[30] A. Saxena and A. Saad, "Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems," *Applied Soft Computing*, vol. 7, no. 1, pp. 441–454, 2007.

[31] R. Heng and M. Nor, "Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition," *Applied Acoustics*, vol. 53, no. 1, pp. 211–226, 1998.

[32] D. Djurdjanovic, S. E. Widmalm, W. J. Williams, C. K. Koh, and K. P. Yang, "Computerized classification of temporomandibular joint sounds," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 8, pp. 977–984, 2000.

[33] D. Djurdjanovic, J. Ni, and J. Lee, "Time-frequency based sensor fusion in the assessment and monitoring of machine performance degradation," in *ASME 2002 International Mechanical Engineering Congress and Exposition*, pp. 15–22, American Society of Mechanical Engineers, 2002.

[34] J. Kittler, "Mathematical methods of feature selection in pattern recognition," *International Journal of Man-Machine Studies*, vol. 7, no. 5, pp. 609–637, 1975.

[35] A. K. Mahamad and T. Hiyama, "Fault classification based artificial intelligent methods of induction motor bearing," *INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING INFORMATION AND CONTROL*, vol. 7, no. 9, pp. 5477–5494, 2011.

[36] L. Wuxing, P. W. Tse, Z. Guicai, and S. Tielin, "Classification of gear faults using cumulants and the radial basis function network," *Mechanical systems and signal processing*, vol. 18, no. 2, pp. 381–389, 2004.

[37] S. Pöyhönen, A. Arkkio, P. Jover, and H. Hyötyniemi, "Coupling pairwise support vector machines for fault classification," *Control Engineering Practice*, vol. 13, no. 6, pp. 759–769, 2005.

[38] U. H.-G. Kreßel, "Pairwise classification and support vector machines," in *Advances in kernel methods*, pp. 255–268, MIT Press, 1999.

[39] J. C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters," 1973.

[40] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification.* John Wiley & Sons, 2012.

[41] "sbRIO-9636 OEM Operating Instructions and Specifications." `http://www.ni.com/pdf/manuals/373378d.pdf`.

[42] "ADXL327 Data Sheet." `http://www.analog.com/static/imported-files/data_sheets/ADXL327.pdf`.