Abnormality Detection in Hard Disk Drive Assembly Process Using Support Vector Machine

Masayuti Simongyi
Computer Science Program
Department of Computer Engineering
Chulalongkorn University,Bangkok, Thailand
Masayuti.S@student.chula.ac.th

Prabhas Chongstitvatana
Department of Computer Engineering
Chulalongkorn University, Bangkok, Thailand
Prabhas.C@chula.ac.th

***Abstract*— The research proposes a method to detect abnormality in assembly of a hard disk drive. Machine learning techniques are employed to recognise the behavior of good and bad components. Both good and bad components of disk drive assembly are collected from the assembly line. To test the components, the current in voice coil motor is measured and collected for using as training data set. Since the amount of abnormal drive in hard disk drive assembly process is very small, this paper also set the experiment of varying amount of training data set that can satisfy practical hard disk drive assembly process. Support Vector machine is chosen as it is very good in binary classification. It is found to be suitable for this task.**

Keywords— Hard Disk Drive, Voice Coild Motor, Detection, Support Vector Machine

I. INTRODUCTION

In the process of manufacturing and assembling the hard disk drive. It involves the assembly of various hard disk drive components together. Hard disk drive consists of the following main components: voice coil motor (VCM), pivot, ramp and actuator arm, etc. In addition to the main components that make the hard disk drive work properly. The hard disk drive also contains components used to prevent damage to the read/write head and the disk. For example, a latch has the function of preventing the read/write head from get off garage when the shock occurs. A breather filter acts to prevent leaks of helium gas while writing a servo signal. The breather filter is required to be installed and align with top cover properly otherwise it may be hit by a sliding actuator arm, which is frequently moved in and out, because of the breather filter is mounted on top cover nearby position of read/write head. The collision between the actuator arm and the breather filter not only cause the damage to the head but also produces the particle that circulate inside the hard disk drive. This particle has been proven to be a major cause of head and disk media collision while writing servo signal.

From above, it was found that the problems encountered in assembling the hard disk can be classified into two categories. The first is that the parts are not completely assembled. Most of them are small pieces that are not noticeable. For example, a latch which event though it has visual mechanic inspection with human eyes before closing hard disk drive top cover but in some case there still have a problematic hard disk drive that can pass this check and continue to the next process. The second type of problem is that parts are completely assembled but some parts are not properly installed, for example, installing a breather filter that mentioned above.

Machine learning method is a technique used for detecting this fault. The method is applied to the process of writing the servo signals which is done after the hard disk components are completely assembled from the clean room. Support vector machine (SVM) is a popular technique in Machine learning that can be used to classify two-classes data (good/bad). The voice coil motor current is used as input data to train SVM. The SVM classifies the hard disk drive into two classes which is good drive and bad drive. The good drive is the drive that completely assembled and properly installed. The bad drive is the hard disk drive that has abnormality in assembly process as mentioned earlier.

Due to difficulty of finding abnormality drive from assembly process because the failure rate is very small, on average the abnormality assembled drive have less than 100 parts per million. So, this work uses small number of bad drives to train SVM when compare to the good drive. In this paper, we also experiment with varying the proportion of good and bad drives that use to train SVM to find out the good proportion for training the machine.

II. RELATED WORKS

Support Vector Machine is a pattern classification algorithm develop by C. Cortes and V. Vapnik at AT&T Bell Labs [1]. The algorithm is designed for two-group classification problems by non-linearly mapped input vectors to a very high dimensional feature space where a linear decision surface is constructed [2]. Training a SVM is equivalent to solving a linearly constrained quadratic programming (QP) problem which is challenging when the number of data points exceeds few thousands. E. Osuna R. Freund F. Girosi proposed a decomposition algorithm [3] that is guaranteed to solve the QP problem with larger number of support vectors. Kernel selection is one of the major problems in the study of SVM, T. R. Baitharu, S. K. Pani, and S. K. Dhal did a comparison of kernel selection for SVM by using Diabetes Dataset [4] and found that linear kernel is the best choice.

There are works that use SVM as based model for Model Predictive Controller (MPC). D. Liang, F. Xu, H. Chen and S. Yu use SVM as a based model for MPC to manage the engine speed to the set-point of idle speed [5]. Saludes S, Fuente M.

J. use SVM as model of the plant in nonlinear MPC [6]. Support Vector Machine is one of popular technique that use to identify defect. A. Kumar and R. Kumar propose a method to detect defect from vibration signal of centrifugal pump using time-frequency analysis and support vector machine [7].

M. Ebrahimi, M. Khoshtaghaza, S. Minaei, and B. Jamshidi uses SVM technique to detect pest in strawberry greenhouses [8]. SVM had been proved to be good detection method of detecting hard disk drive failure [9]. S. Shakya and S. Sigdel develop a hybrid algorithm base on SVM and Naïve Bayes for anomaly detection [10]. Hodge, Victoria J., and Jim Austin did a survey of outlier detection using various techniques, including support vector machine [11].

III. Hard Disk Drive Assembly Process

The hard disk drive assembly process must be done in a clean room. After the components are cleaned, the first step is to install the disk media and then the disk clamp to hold the disk media together. After that, it will be verified that if the hard disk drive media has been properly installed before installing the ram and other devices, as shown in Fig. 1. It can be seen that in the hard disk drive assembly process it is not only involve the installation of each component together but it is also include of the verification of the component that has been installed. For example, the verification of disk media balance and the leak test but there are also a number of components that could not be verified. For example, a ram, voice coil motor, actuator arm, read/write head and a latch. This is because there is no proper way to verify these components better than visual mechanical inspection.

After all the hard disk drive components are assembled, they are closed with the top cover and passed to the next process with printed circuit board assembly (PCBA). It's ready for supply the voltage and the reader is pushed into the disk media to write a servo signal

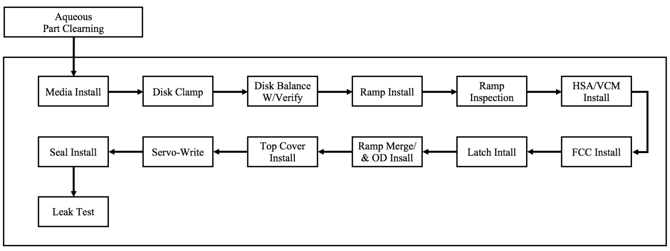


Fig. 1. Process of assembly hard disk drive in clean room.

In order to read and write servo signal on disk media, the first step is to spin up spindle motor to make it spin up at constant velocity which typically is around 7200 RPM and then loading the heads onto the disk. If the head is loaded while the spindle motor is not spinning at our target speed it has the potential to cause the read/write head crash with the disk media. Generally, in the firmware of the hard disk drive, it is always necessary to have this verification step before loading head onto the disk.

When the spindle motor is spinning at speed, the current is supplied to voice coil motor so that the actuator arm is pushed onto the disk. This process requires a different level of current because the force that need to push actuator out of garage until get actuator onto the disk is varying and because of the voice coil motor itself is a like coil which generate the back EMF when supply the current through it. So, it is important to have close loop system built in as shown in Fig. 2 to accurately control the speed of the actuator.

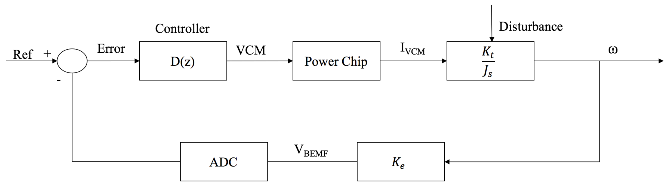
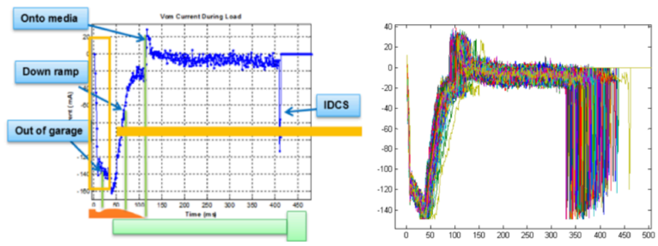


Fig. 2. Closed-loop control system of actuator arm.

The current that supply to voice coil motor since head parking on the garage until head successfully load onto disk media is plot and shown in Fig. 3(A). At the beginning, enough amount of current need to be supplied to make actuator out of garage and latch and then the current needed is gradually reduce until close to zero when head is on disk media and detect the crash stop.

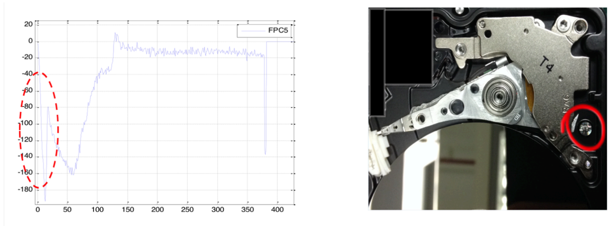
When we collect voice coil motor current during loading head operation over 1000 times, the voice coil motor current has a consistent shape as shown in Fig.3 (B) this is because the current that supply to the coil has feedback control as described above.



(A) (B)

Fig. 3. (A) Single VCM current. (B) 1000x overlay of VCM current.

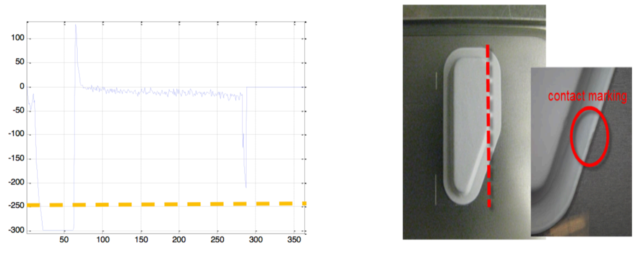
The current that supply to the voice coil motor is continuous value but the current that collect for experiment is sampling from continuous value in every one millisecond. This sampling time is equal to the interrupt time for close loop feedback control system. In this experiment, we allocate memory of 1024 values for each load operation. This will include interval time of 1024 milliseconds which enough for our validation.



(A) (B)

Fig. 4. (A) VCM current of missing latch drive. (B) missing latch position.

From the experiment with the problematic drive it is found that the form of voice coil motor current is change when experiment with incomplete assembly drive. For example, when latch is not installed as shown in Fig. 4(A) the shape of the voice coil current is different from Fig.3(A), Fig. 4(B) show the position where latch is not installed. Figure 6 show the abnormality of voice coil motor shape when breather filter is install improperly.

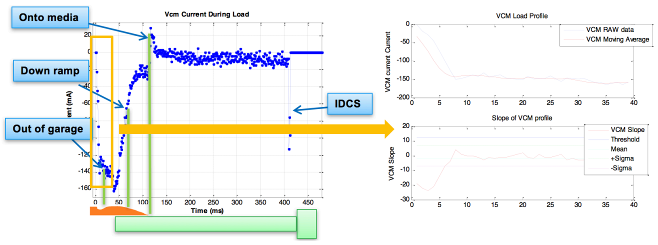


(A) (B)

Fig. 5. (A) VCM current of improperly install breather filter. (B) breather filter.

IV. ABNORMALITY DETECTION USING TRAPPING METHOD

Currently there is a method to detecting abnormality in hard disk drive assembly at servo writing process by developing a program to trap specific abnormality. The development process in this method requires to collect voice coil motor current from specific abnormality drive, investigate the abnormality pattern and develop a program to trap it. This method requires a number of hard disk drive to ensure the correctness of the program and that it does not over-kill the good drive, for example, in case of missing latch. The algorithm of trapping program is to finds the moving average of the voice coil current and then calculates the slope of it to check if it is in the appropriate range as shown in Fig. 6.



(A) (B)

Fig. 6. Demonstration of missing latch detection algorithms.

V. ABNORMALITY DETECTION USING MACHINE LEARNING METHOD

Machine learning technique can be used to capture the characteristics of good drive. Support vector machine is chosen as it is a very good method for classifying two-classes data (good/bad). In this paper, the voice coil motor current data have been collected from 20 abnormality drives and 7519 normal drives, totally 7539 drives. The numbers of abnormality drive in assembly process is very small when compare to the number of good assembly drive. So, the experiment has been divided into three groups by the number of hard drives used in training set and verification set as shown in Table 1. in which Group 1, 2 and 3 have the proportion of abnormality drives in 3%, 10% and 10% accordingly.

Table 1: Data set separated by group of experiment.

Experiment # ➔	#1		#2		#3	
Class	Good	Bad	Good	Bad	Good	Bad
Training Set	485	15	90	10	45	5
Validation Set	7034	5	7429	10	7474	15
Total	7519	20	7519	20	7519	20

A. *Data collection process*

In the process of collecting data of voice coil current to use for training support vector machine. The data is stored in pair format between voice coil motor current data and the class of assembly process. This class indicates that which hard disk drive is assembly correctly. The Good class is representing the hard disk drive that assembly correctly and the Bad class represent the hard disk drive that have abnormality in hard disk drive assembly process. The voice coil motor current data is sampling every one millisecond starting when load command is invoked for 1024 milliseconds. This time interval is long enough to cover event from loading head until head loading onto media and found the crash stop. On average this process takes around 300 to 600 milliseconds as depicted in Fig. 3(B).

B. *Training and Validation*

Support vector machine is a supervised learning method. the data is separated into two set, one is training data set and another one is validation data set as depicted in Table 1. After voice coil current data are collected then it is used as training data set for training support vector machine with Matlab. The function that is used to train support vector machine in Malab is svmtrain() function. From the training algorithm, the model is created. This model can be used to classify validation data set with svmclassify() function in Matlab. The result from validation is compared with actual class of hard disk drive assembly process as shown in Fig.7.

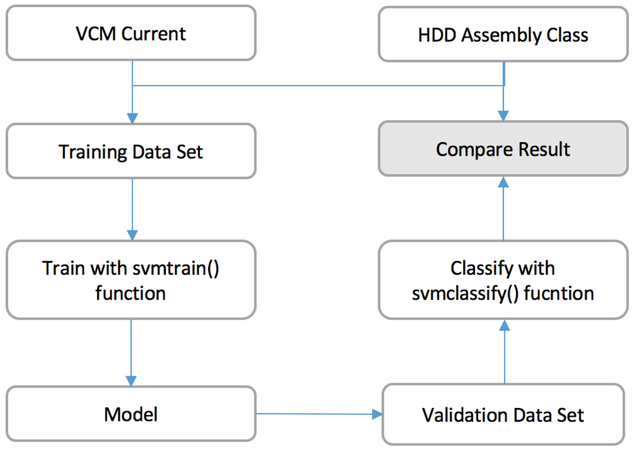


Fig. 7. Testing and validation work flow.

VI. Experiment Result

From the result of classification with support vector machine with different proportion of Good and Bad class and different number of training data set, described in section 6A, it is found that the experiment #1 which uses 500 drives as training data set can classify 100% of good and bad drive.

Table 2: Experiment result

Experiment	HDD Class	SVM Classification		Error Part Per Million	
		Bad	Good	Bad	Good
#1 (500)	Bad	5	0	0	0
	Good	0	7034	0	0
#2 (100)	Bad	10	0	0	0
	Good	5	7424	673	0
#3 (50)	Bad	15	0	0	0
	Good	13	7461	1739	0

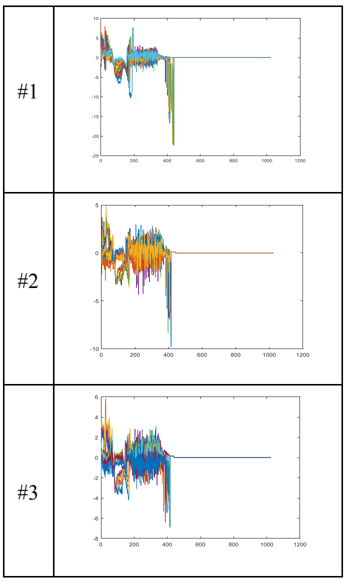


Fig. 8. Support vector of each experiment Group.

VII. Conclusion

From the experiment, we can conclude that with training data set of 500 drives in which the proportion of bad drive is 3% or 15 drive, support vector machine can classify good and bad drive with 100% accuracy. In practice, it is possible to find the training set with this sample size and proportion. However, when the training sample set is decrease to 100 drives, it is found that support vector machine can classify bad drive correctly but have error in classify good drive in the rate of 673 part per million. This error rate is consider to be very high because the hard disk drive manufacturing process produces more than 10 million hard disk drives for each quarter. When compare machine learning detection method with the method of developing the program to trap the abnormality of hard disk drive assembly, it is found that when enough sample size has been chosen to train the support vector machine, the support vector machine can classify correctly for both good and bad drive. SVM also has the advantage over the trapping method because it is easier to develop and it does not require the knowledge of specific abnormality to develop a correct program.

References

[1] C. Cortes and V. Vapnik. “Support vector networks. Machine Learning,” 2O:l-25, 1995.

[2] Han, Jiawei, Jian Pei, and Micheline Kamber. Data mining: concepts and techniques. Elsevier, 2011.

[3] E. Osuna R. Freund F. Girosi "An Improved Training Algorithm for Support Vector Machines" Proc. IEEE Workshop on Neural Networks and Signal Processing 1997.

[4] T. R. Baitharu, S. K. Pani, and S. k. Dhal, “Comparison of kernel selection for support vector machines using diabetes dataset,” Journal of Computer Sciences and Applications, vol. 3, no. 6, pp. 181–184, 2015.

[5] D. Liang, F. Xu, H. Chen and S. Yu, "Support vector machine based model predictive control for engine idle speed control," 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 2772-2777.

[6] Saludes S, Fuente M J, “Fault tolerant SVM based nonlinear Model Predictive Control,” Control Conference (ECC), 2007 European. IEEE, 2007: 4124-4129.

[7] A. Kumar and R. Kumar, “Time-frequency analysis and support vector machine in automatic detection of defect from vibration signal of centrifugal pump,” Measurement, vol. 108, no. Supplement C, pp. 119 – 133, 2017.

[8] M. Ebrahimi, M. Khoshtaghaza, S. Minaei, and B. Jamshidi, “Vision-based pest detection based on svm classification method,” Computers and Electronics in Agriculture, vol. 137, no. Supplement C, pp. 52 – 58, 2017.

[9] Joseph F. Murray, Gordon F. Hughes, Kenneth Kreutz-Delgado, “Machine Learning Methods forPredicting Failures in Hard Drives: A Multiple-Instance Application” Journal of Machine Learning Research, vol. 6, pp 783–816, 2005.

[10] S. Shakya and S. Sigdel, "An approach to develop a hybrid algorithm based on support vector machine and Naive Bayes for anomaly detection," 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, 2017, pp. 323-327.

[11] Hodge, Victoria J., and Jim Austin. "A survey of outlier detection methodologies." Artificial intelligence review 22.2 (2004): 85-126.