

# 데이터 마이닝 기법 및 경험적 모드 분해법을 이용한 회전체 이상 진단 알고리즘 개발에 관한 연구

윤상환\*, 박병희\*\*, 이창우\*\*.#

\*창원대학교 컨소시엄사업단, \*\*창원대학교 기계설계공학과

## A Study on Fault Diagnosis Algorithm for Rotary Machine using Data Mining Method and Empirical Mode Decomposition

Sang-hwan Yun\*, Byeong-hui Park\*\*, Changwoo Lee\*\*.#

\*Consortium organization, Changwon UNIV.,

\*\*School of Mechanical Engineering, Changwon UNIV

(Received 8 April 2016; received in revised form 26 May 2016; accepted 31 May 2016)

### ABSTRACT

Rotary machine is major equipment in industry. The rotary machine is applied for a machine tool, ship, vehicle, power plant, and so on. But a spindle fault increase product's expense and decrease quality of a workpiece in machine tool. A turbine in power plant is directly connected to human safety. National crisis could be happened by stopping of rotary machine in nuclear plant. Therefore, it is very important to know rotary machine condition in industry field. This study mentioned fault diagnosis algorithm with statistical parameter and empirical mode decomposition. Vibration locations can be found by analyze kurtosis of data from triaxial axis. Support vector of data determine threshold using hyperplane with fault location. Empirical mode decomposition is used to find fault caused by intrinsic mode. This paper suggested algorithm to find direction and causes from generated fault.

**Key Words** : Support Vector Machine(서포트 벡터 머신), Empirical Mode Decomposition(경험적 모드 분해), Kurtosis(첨도), Skewness(왜도), Fault Diagnosis(이상 진단)

### 1. 서 론

회전체는 자동차 내부의 모터, 공작기계 스피들, 플랜트에서 펌프 혹은 발전기등 산업 전반에 걸쳐 사용되고 있다. 플랜트에서 펌프가 멈추게 된다면 천문학적 손실이 발생한다. 하루 생산해

야하는 석유에너지등 1차 손실에서 2차 손실까지 야기하게 된다. 풍력발전기 및 수력발전기등 발전기 내부에 사용되는 회전체에 고장이 발생하게 되면 전기를 생산할 수 없다. 개발도상국의 고립된 지역의 발전기의 경우 지역 주민들의 전기를 사용할 수 없는 상태가 된다. 따라서 회전체의 모니터링은 필요한 분야가 다양하다는 것을 알 수 있다. 특히, 공작기계 분야에서 스피들의 진동데이터를 수집하여 가공품질 혹은 스피들의 수명상태를 예측하는 분야에 많이 사용되고 있다<sup>[1-2,13-14]</sup>.

# Corresponding Author : leecw@changwon.ac.kr

Tel: +82-55-213-3618, Fax: +82-55-263-5221

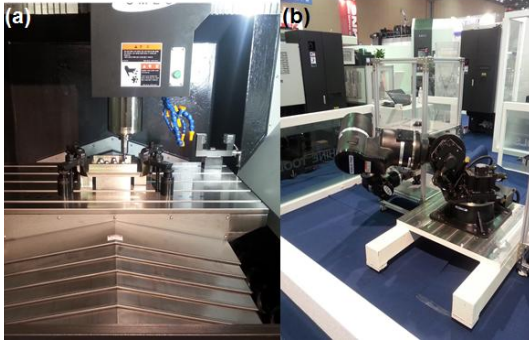


Fig. 1 Rotary example (a) spindle of machine tool (b) robot arm

회전체의 이상 상태 및 수명에 따라 생산 공정의 휴지 시간, 작업 및 주변 환경이 결정되는데 회전체의 주요부품 베어링은 기본적인 수명에도 불구하고 공정조건에 따라 회전체의 수명은 예측 수명과 달라서 발생하게 된다. 이상 진단은 회전체뿐만 아니라 태양광 패널, 기어박스등 다양한 분야에 사용되고 있다. 태양광 발전 시스템의 태양광 패널(photovoltaic system)의 모듈의 전압, 전류 그리고 온도를 Artificial neural network toolbox를 통해 기계학습이론을 적용하여 이상을 진단한 경우도 있다<sup>[3-4]</sup>. 풍력 터빈의 비정적인 상태에서 발생하는 비주기적이며 복잡한 진동신호를 시간-변화 특성 주파수를 Vold-Kalman filter과 고차의 미분에너지 함수분리법을 이용한 풍력 터빈의 이상진단을 연구한 사례도 있다<sup>[5]</sup>. 진동신호와 음향신호를 측정할 후 데이터마이닝 기법중 하나인 랜덤 포레스트 기법을 심화한 Two deep boltzmann machine 방법을 사용하여 기어 박스의 pitting, 97.68% 진단을 한 연구도 있다<sup>[6]</sup>. 베어링에 이상이 발생할 경우 진동데이터를 측정하여 베어링 이상 원인을 구분한 연구도 있다<sup>[7-8]</sup>. 인공지능회로망기법을 사용하여 수학적 모델에서 얻은 데이터들을 회로망에 학습시킨 후 실험에서 측정된 값들을 입력하여 베어링 이상을 진단한다. 하지만 수학적 모델은 오직 베어링에만 적용할 수 있으므로 복합시스템에서 발생하는 이상 원인을 진단하는

데는 어려움이 있다. 베어링 롤러에서 측정되는 진동신호를 EMD(Empirical mode decomposition) 기법을 이용하여 원신호에서 내재모드를 추출하여 이상 신호를 추출한 후 이상 원인을 구분한 연구도 있다. 하지만 EMD기법은 내재모드함수 만족조건 극댓값 및 영통과점 개수를 구하는데 많은 연산시간이 필요한 단점이 있다<sup>[9]</sup>. 이를 극복하는 것이 중요하다.

본 연구에서는 산업현장 회전체에서 발생하는 진동데이터를 통계적 파라미터를 이용하여 진동이 발생하는 위치를 추정하고 SVM 기법을 적용하여 임계값보다 높은 데이터들을 수집함으로써 이상원인을 추정하는 알고리즘에 관한 연구를 진행하였다.

## 2. 본론

### 2.1 통계적 파라미터

첨도(kurtosis), 왜도(Skewness), 평균값 및 표준편차의 통계적 파라미터를 이용하여 데이터를 분석하였다. 첨도는 상대적으로 데이터가 0에 얼마나 수렴하는지 나타내며 높을수록 0에 데이터가 집중되어 표시되며 작으면 작을수록 데이터가 분산된 것을 보여준다. 데이터 대부분이 0보다 클 경우 왜도는 0보다 크며 0보다 작을 경우 왜도는 0보다 작다. 표준정규분포의 첨도는 3이고 왜도는 0이다. 측정된 진동데이터들을 표준정규분포의 첨도와 왜도를 비교한다. 각 축에서 피크 값이 발생하게 되면 그 값들의 분포는 큰 영역에서 발생하게 될 것이고 첨도가 낮다는 것을 의미한다. 이를 통해 진동 발생방향을 추정할 수 있다. 왜도가 0보다 큰 것은 진동이 축에서 양의 방향에 있다는 것을 의미하고 왜도가 0보다 작다는 것은 진동이 음의 방향에 있다는 것을 의미한다. 어느 축에서 진동이 발생하는지 나타내며 X, Y, Z축 별로 왜도를 구할 수 있다. 평균값은 RMS(Root-Mean-Square)를 이용하여 구하였다. 표준편차는 평균값 대비 데이터들이 얼마나 멀리 위치해 있는가를 나타낸다. X, Y, Z축의 통계적 데이터는 Matlab R2012b 이용하여 분석한다. 첨도 및 왜도 공식은 식(1)과 (2)와 같다.

$$k = \frac{E(x-\mu)^4}{\sigma^4}$$

$$k_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^2} \quad (1)$$

$$s = \frac{E(x-\mu)^3}{\sigma}$$

$$s_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^{3/2}} \quad (2)$$

Table 1은 임의의 X, Y 및 Z축의 데이터에서 첨도와 왜도를 구한 결과이다. X축의 첨도가 가장 낮으며 왜도는 0보다 작다. Y축은 첨도가 가장 높으며 왜도는 0보다 크고 0을 기준으로 가장 작은 변화량을 나타낸다. Z축은 첨도가 2번째로 크며 왜도는 가장 크다. 첨도가 가장 낮은 X축은 데이터가 분산되어 있으며 0보다 작은 데이터들이 많으므로 가장 큰 진동은 X축 역방향으로 나타나는 것을 추정할 수 있다. Y축이 첨도 및 왜도가 가장 낮으므로 진동이 가장 작게 그리고 Y축 정방향으로 발생하는 것을 추정할 수 있다. X축, Y축 및 Z축 데이터의 첨도 및 왜도를 이용함으로써 기계에서 발생하는 진동 발생 방향을 추정할 수 있다.

## 2.2 Support vector machine

SVM(Support vector machine)은 입력 데이터의 분포를 고려하여 확률분포로 평균을 구한 일반화 오차를 최소화할 수 있는 방향으로 학습을 수행하는 선형분류기이며 인공신경회로망의 복잡한 회로망의 가중치 부여를 간략화한 기법이다. 학습표본에 두 개의 범주로 분류된 N개의 객체로 이루어진 함수가 있다면 하나의 선형함수로 두 범주를 완벽하게 분리할 수 있다. 이 때 p차원 공간에서의 선형함수를 하이퍼플레인(hyperplane) 혹은 선형결정경계라고 한다. 두 범주를 구분하는 하이퍼플레인은 무수히 많을 수 있으나 가장 적절한 것을 결정하고자 한다. 기계학습 기법 중 하나이며 SVM은 초평면 상에서 선형결정경계를 식(1)을 통해 구하여 데이터를 분류한다.

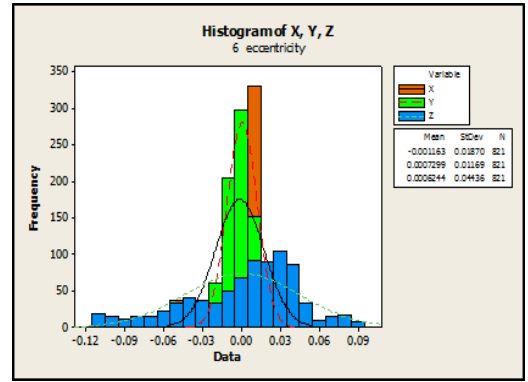


Fig. 2 Example for histogram with kurtosis

Table 1 Example of kurtosis

	X	Y	Z
Data	0.8884	-0.1022	-0.8637
	-1.1471	-0.2414	0.0774
	-1.0689	0.3192	-1.2141
	-0.8095	0.3129	-1.1135
	-2.9443	-0.8549	-0.0068
	1.4384	-0.0301	1.5326
	0.3252	-0.1649	-0.7697
	-0.7549	0.6277	0.3714
	1.3703	1.0933	-0.2256
	-1.7115	1.1093	1.1174
Kurtosis	2.0891	2.2841	2.1133
Skewness	-0.1231	0.0512	0.4882

$$A(x) = w \cdot x + w_0$$

$$= w^T x + w_0 = \sum_{i=1}^n w_i x_i + w_0 \quad (3)$$

$$f(x) = \text{sign}(A(x))$$

$$w^T x_i + b \geq 1 \text{ for } y_i = 1 \quad (i=1, \dots, N) \quad (4)$$

$$w^T x_i + b \leq -1 \text{ for } y_i = -1 \quad (i=1, \dots, N) \quad (5)$$

$$\text{Max} : \frac{2}{w^T w} \quad (6)$$

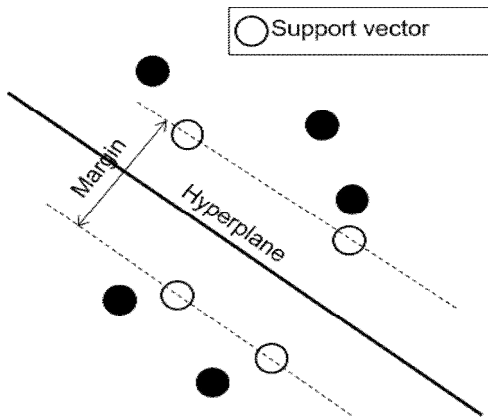


Fig. 3 Support vector machine

함수  $f(x)$ 가 1의 값을 가지는 경우 입력 데이터  $x$ 는 클래스  $C_1$ 으로 할당되며, 반대로 -1의 값을 가지는 경우는 클래스  $C_2$ 로 할당된다. 학습을 통하여 주어진 학습데이터들이 올바른 클래스에 할당될 수 있도록 함수  $A(x)$ 를 찾아야 한다. 즉, 함수  $A(x)$ 는 서포트 벡터까지 거리를 최대화 하는 선형 결정경계를 찾는 문제로 된다. 학습표본이  $N$ 개의 객체이고,  $I$ 번째 객체를  $p$ 개의 변수로 이루어진 벡터  $x_i$ 로 표기하며 이에 대응하는  $i$  분류된 범주를  $y_i$ 로 표기하면 분리 선형결정경계를 찾는 문제는 다음과 같이 된다. Fig. 3과 같이 선형결정경계를 기준으로 특정한 길이(margin)만큼 떨어진 데이터를 찾는 것이며 이를 서포트 벡터 머신 기법이라 할 수 있다. 일정 지점에 우치한 점들을 서포트 벡터라고 한다. 진동데이터에서 각 축의 평균 진동량을 RMS(Root-Mean-Sqaure)을 구한 후에 이 값을 기준으로 Matlab을 이용하여 서포트 벡터를 구해보았다. Fig. 4과 Fig. 5와 같이 데이터에서 서포트벡터를 구해보았다. 서포트 벡터의 마진은 0.01로 정하였으며 Fig. 4는 선형결정경계를 0.01보다 큰 값을 설정하였고 Fig. 5는 선형결정경계를 0.01과 0.001 사이 값을 설정하였다. Fig. 4에서 서포트 벡터는 (0.01, 0.01)과 (0.02, 0.02)로 결정되었다.

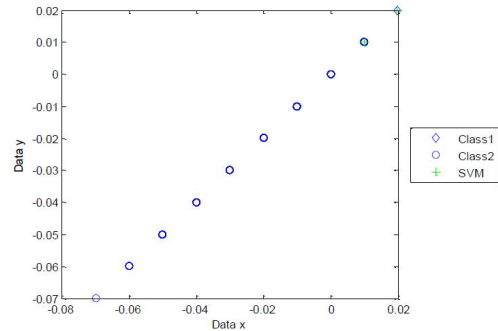


Fig. 4 Support vector machine at data is higher than 0.01

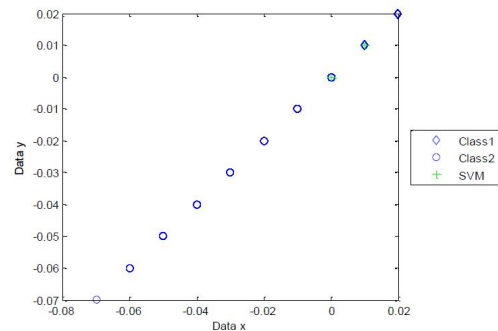


Fig. 5 Support vector machine at data is higher than 0.001

이는 결정경계에서 마진 0.01로 설정하였기 때문에 서포트 벡터는 Fig. 4와 같이 정해졌다. Fig. 5의 서포트 벡터는 0.001보다 큰 값이며 마진이 0.01이기 때문에 0.001에서 마진 0.01을 만족하는 0과 0.01이 서포트 벡터로 결정되는 것을 알 수 있다. 선형결정경계에 대한 특정한 길이를 만족하기 때문이다. 서포트 벡터머신 기법을 적용함으로써 진동데이터내에서 선형결정경계를 찾고 서포트 벡터를 찾음으로써 진동데이터의 임계값을 구할 수 있다. 본 연구에서는 서포터 벡터 머신을 이용하여 임계값(threshold)을 넘어서는 진동데이터들을 구별한다.

### 2.3 Empirical mode decomposition

경험적 모드 분해(EMD: Empirical Mode Decomposition)이라고 하며 신호에서 내재모드 함수를 추출하는 과정이다. Hilbert-transform의 일종이며 신호에서 내재모드함수를 추출하기 위해 모든 국소 극값을 반복적으로 계산한다. 복잡한 노이즈가 섞인 신호에서 국소적인 자료 패턴을 추출할 수 있다<sup>[10]</sup>. 복잡한 신호를 추출하기 위해 경험적 모드분해법은 신호에서 각각의 주파수 성분 신호를 추출하는 기법이며 다음의 조건을 만족해야 한다. 먼저 반복되는 형태의 함수의 임의의 지점 국소 극대값과 그 다음 국소 극대값 사이의 구간에서 0을 지나는 점들의 개수를 구해야하고, 국소 극대 값들과 국소 극대 값들 각각에 대해 보간법으로 산출한 함수의 상위막과 하위막이 있으며 이는 상위막과 하위막 사이에 존재하는 평균막은 대칭적인 구조를 이루고 있다. 내재 모드함수는 경험적 모드 분해 각각의 성분으로 정의할 수 있으며, Huang은 다음의 2가지 성질을 만족하는 내재모드 함수를 정의하여 신호를 추출하였다<sup>[11-12]</sup>. 내재 모두 함수는 국소 극값의 개수와 영통과점의 개수 차이는 한 개 이하이고 평균막의 국소 평균은 0이어야 한다. 그리고 평균막과의 차이를 구한후 신호의 합은 0을 만족해야 한다. 주어진 시그널을 가장 빠른 주파수를 가지는 내재모드 함수와 잔여신호로 분해한다. 선정된 시그널의 국

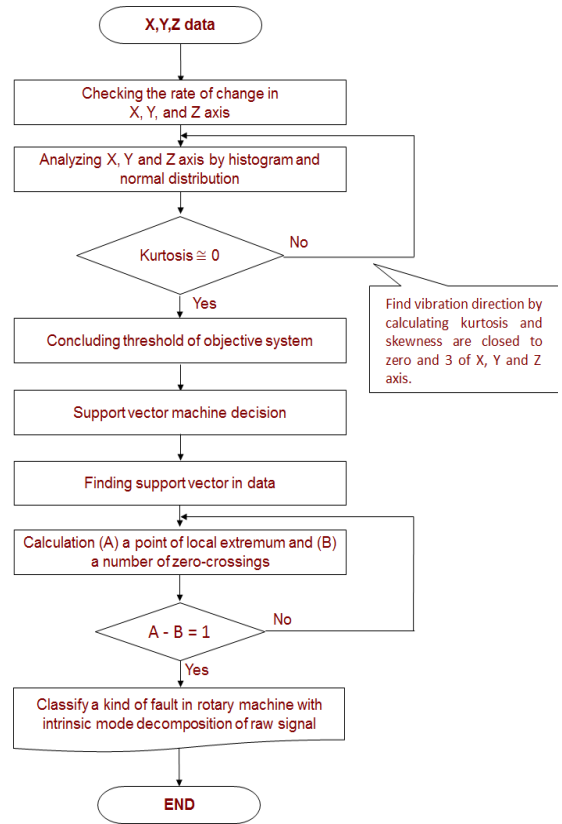


Fig. 7 Flow chart of Fault classified with support vector and empirical mode decomposition

소 극대값과 국소 극소값을 모드 식별한다. 국소 극대값에 대해 보간법(interpolation)으로 상위막을 생성한다. 동일한 과정으로 하위막도 생성한다. 상위막과 하위막의 평균인 평균막을 계산한다. 신호 첫 번째 내재모드함수와 평균막의 차이로 고주파 성질을 가지는 신호성분을 생성한 후 내재모드 함수를 성질을 만족하면 내재모드함수 J1로 정의한다. 신호 s를 내재모드함수 J1로 분해하여 최종적으로 임의의 내재모드함수 J1, Jk와 잔여 신호를 산출한다. Fig. 6은 경험적 모드 분해법 예시를 나타낸 그림으로서 복잡한 신호는 노이즈와 특정 주파수 성분의 신호가 함께 섞여 있다. 각각의 신호를 분리함으로써 복잡한 기계시스템에서 발생하는 여러 진동 데이터들을 진동원인 별로 구분할 수 있다.

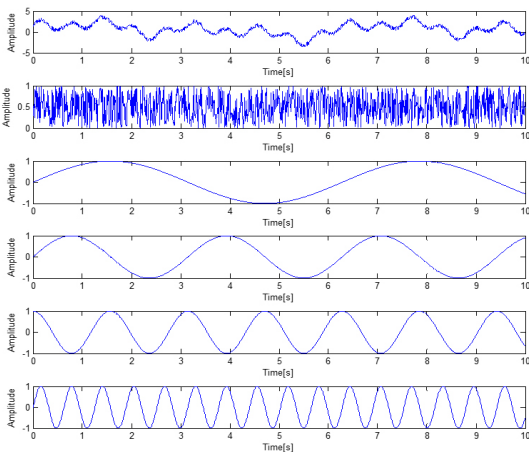


Fig. 6 Example of empirical mode decomposition

## 4. 결론

Fig. 7과 같이 회전체 이상진단 알고리즘 연구를 진행하였다. 통계적 파라미터 표준정규분포, 첨도 및 왜도를 분석하였다. 각 축의 첨도치를 비교함으로써 진동이 발생하는 축 방향을 추정하였다. 진동 발생을 추정한 방향에서 평균값을 구한 후 서포트 벡터 머신기법을 통해서 진동발생방향의 임계값을 설정할 수 있었다. 임계값 선정이후 진동원인을 구별하기 위해 경험적 모드 분해법을 사용하여 복잡한 신호에서 내재모드를 찾음으로써 이상신호를 추출할 수 있다. 기존의 모니터링 진단 분야는 통계적 파라미터를 이용하여 진동발생 방향을 추정한 연구사례가 없다. 본 연구에서는 진동데이터를 통계적 파라미터를 이용하여 진동 발생방향을 추정하고 서포트 벡터 머신을 적용하여 임계값과 정상 데이터를 구분할 수 있다. 최종적으로 진동이 발생하여 이상유무를 확인할뿐만 아니라 EMD기법을 적용하여 이상원인을 진단할 수 있다.

## 후 기

이 논문은 2015년도 산업통상자원부(10051632)의 지원을 받아 수행된 연구임.

## REFERENCES

1. Lee, K. J., Lee, T. M., and Yang, M. Y., "Tool wear monitoring system for CNC end milling using a hybrid approach to cutting force regulation." *The International Journal of Advanced Manufacturing Technology*, Vol. 32, No. 1-2, pp. 8-17, 2007.
2. Choi, H. J., Park, C. W., Bae, J. S., Ahn, J. H., and Choi, S. D., "Design of High Speed Spindles Active Monitoring and Control Algorithm" *The Korean Society of Manufacturing Process Engineers*, Vol. 10, No. 5, pp. 13-19, 2011.
3. Rai, A. and Upadhyay, S. H., "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings," *Tribology International*, Vol. 96, pp. 289-306, 2016.
4. Chine, W., Mellit, A., Lughì, V., Malek, A., Sulligoi, G., and Pavan, A. M., "A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks," *Renewable Energy*, Vol. 90, pp. 501-512, 2016.
5. Feng, Z., Qin, S., and Liang, M., "Time-frequency analysis based on Vold-Kalman filter and Higher order energy separation for fault diagnosis of wind turbine planetary gearbox under nonstationary condition," *Renewable Energy*, Vol. 85, pp. 45-56, 2016.
6. Dolenc, B., Boskoski, P., and Juricic, D., "Distributed bearing fault diagnosis based on vibration analysis" *Mechanical Systems and Signal Processing*, Vol. 66, pp. 521-532, 2016.
7. Dolenc, B., Boskoski, P., Pfajfar, J., and Juricic, D., "Vibration Based Diagnosis of Distributed Bearing Faults" *Vibration Engineering and Technology of Machinery*, Vol. 23, pp. 651-661, 2016.
8. Yu, D., Cheng, J., and Yang, Y., "Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings," *Mechanical Systems and Signal Processing*, Vol. 19, No. 2, pp. 259-270, 2005.
9. Park, M. S., Kim, D., and Oh, H. S., "Empirical mode decomposition using the second derivative" *Korea Journal of Applied Statistics*, Vol.26, No.2, pp. 335-347, 2013.
10. Kim, N. H., Lee, E. S., Lee, D. W., and Kim, N. K., "A Study on the Monitoring of the Micro Grooving using the AE Technology" *The International Journal of Advanced Manufacturing Technology*, Vol. 25, No. 7-8, pp. 663-667, 2005.
11. Gu, D. S., Lee, J. H., Yang, B. S., and Choi, B. K., "Application of Envelop Analysis and Wavelet Transform for Detection of Gear Failure," *Transactions of the Korean Society of Mechanical Engineers A*, Vol. 32, No. 11, pp. 905-910, 2008.

12. Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., and Liu, H. H., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and Non-stationary time series analysis" Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, Vol. 454, pp. 903-995, 1998.
13. Li, C. P., Kim, M. Y., Park, J. K., & Ko, T. J., "A study on the development of rotary ultrasonic machining spindle" Journal of The Korean Society of Manufacturing Process Engineers, Vol. 14, No. 4, pp. 160-166, 2015.
14. Kim, J. S., Kim, B. H., Lee, C. S., Kim, Y. J., and Park, Y. H., "Study on The Status of Welded Parts According to The Types of Shielding Gas in TIG welding" Journal of The Korean Society of Manufacturing Process Engineers, Vol. 14, pp. 38-43, 2015.