

# FFT를 활용한 제조데이터 전처리 및 제품분류

김한솔 · 진교홍\*

창원대학교

## Manufacturing Data Preprocessing Method and Product Classification Method using FFT

Han-sol Kim · Kyo-hong Jin\*

Changwon National University

E-mail : kimsol2344@changwon.ac.kr / khjin@changwon.ac.kr

### 요 약

스마트 공장 구축사업을 통해 생산 설비로부터 전력, 진동, 압력, 온도 등의 센서 데이터가 수집되고 있으며 데이터 분석을 통해 예지보전, 불량예측, 이상탐지 등의 서비스 개발이 진행되고 있다. 일반적으로 제조데이터의 경우 정상과 비정상 데이터의 불균형이 극심하여 이상탐지 서비스가 선호되고 있다.

본 논문에서는 이상탐지 서비스 개발의 전단계로 제조데이터의 특징 데이터 추출을 위해 FFT 방법을 사용하였으며, 이를 통해 생산되는 제품을 분류해보고 그 결과를 확인하였다. 즉, 제품별 대표 패턴을 FFT 변환 후 상관계수를 계산하여 제품분류가 가능한지 확인하였다.

### ABSTRACT

Through the smart factory construction project, sensor data such as power, vibration, pressure, and temperature are collected from production facilities, and services such as predictive maintenance, defect prediction, and abnormality detection are developed through data analysis. In general, in the case of manufacturing data, because the imbalance between normal and abnormal data is extreme, an anomaly detection service is preferred. In this paper, FFT method is used to extract feature data of manufacturing data as a pre-stage of the anomaly detection service development. Using this method, we classified the produced products and confirmed results. In other words, after FFT of the representative pattern for each product, we verified whether product classification was possible or not, by calculating correlation coefficient.

### 키워드

Manufacturing Data, Fast Fourier Transform, Feature Engineering, Pearson Correlation, Classification

### 1. 서 론

산업 현장에는 장비로부터 추출되는 생산 데이터, 장비에 부착된 센서로부터 수집되는 센서 데이터, 생산 환경 데이터 등 다양한 데이터가 존재한다[1]. 데이터를 수집하고 분석해서 의미 있는 정보를 추출한 결과를 통해 제조업에서는 공정 최적화와 손실을 최소화할 수 있다[5].

수집된 센서 데이터 중 진동 및 전력 데이터는 유사한 형태의 패턴들이 반복적으로 나타나는 구

조를 가지며, 탐색적 데이터 분석(Exploratory Data Analysis, EDA)을 통해 이들 데이터의 패턴들이 제품을 생산할 때마다 나타난다는 것을 확인할 수 있었다. 주목할 점은 이들 패턴의 추세는 유사하지만, 센서값의 크기가 같은 경우는 극히 드물다는 것이다. 이런 특징은 같은 제품임에도 다른 제품으로 잘못 판단하게 하는 원인이 된다.

본 논문에서는 고속 푸리에 변환(Fast Fourier Transform, FFT)을 사용하여 이 문제를 해결하고자 한다[2]. 센서값만 보면 차이가 심해 다른 제품으로 판단되는 경우에 FFT가 적용되면 주파수 특성이

\* corresponding author

유사하여 같은 제품으로 판단할 수 있을 것으로 예상된다. 제안하는 방법은 제품별 진동 데이터를 통해 대표 패턴을 생성한다[3]. 그리고 대표 패턴을 FFT를 수행하여 주파수 영역으로 변환한 패턴 간 상관계수를 비교한다.

이때 사용되는 상관계수는 두 변수가 모두 연속형 자료이므로 피어슨 상관계수(Pearson correlation coefficient)를 이용한다[4].

이 방법으로 2가지 제품을 생산하는 설비에서 수집된 진동 데이터를 이용해 주파수 특성에 대한 분석을 시행하고 이를 통해 제품 간 분류가 가능한지 확인하였다.

본 논문의 구성은 다음과 같다. 2장은 수집된 데이터와 데이터 분석 과정에 관해서 설명하였으며 3장은 상관계수를 활용하여 임계값을 설정하고 제품분류 성능을 평가하였다. 마지막으로 4장은 결론 및 향후 연구 방향을 제시하였다.

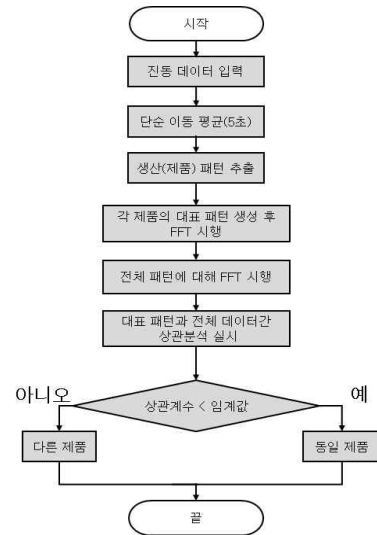


그림 1. 데이터 분석 흐름도

## II. FFT를 활용한 제조 데이터 분석

### (1) 제조데이터

MCT 설비에서 전력, 진동, 온도, 유압, 습도 등의 센서로 1초 단위의 제조데이터를 수집하였다. 이들 데이터를 대상으로 EDA를 진행하였으며, 그 중 진동 데이터와 전력 데이터가 제품별 특성을 보여주었다. 한편, 진동 데이터는 공정주기와 비가동 정보를 담고 있으며, 공정 패턴을 추측하는데 전력 데이터보다 용이하였다.

### (2) 데이터 분석 과정

데이터분석과정은 그림 1과 같다. 먼저, 입력된 진동 데이터에 윈도우 크기를 5로 설정하고 단순이동평균(Simple Moving Average)을 적용하여 누락된 데이터를 보간하고 불규칙한 시간 간격을 일정하게 맞추었다.

그 후 전처리된 진동 데이터를 생산 패턴 추출기에 넣어 A 제품을 생산할 때의 패턴과 B 제품을 생산할 때의 패턴을 추출하였다.

그림 2 (a)는 A 제품을 생산하면서 발생한 진동 데이터를 모두 취합하여 평균값으로 생성된 대표 패턴이다. 그리고 그림2 (b)는 A 제품 대표 패턴에 대한 FFT를 실시간 결과이다.

그림 2의 (c)와 (d)는 B 제품에 대한 대표 패턴과 FFT를 한 결과이다. 그리고 전체 데이터에서 제품별 패턴을 추출하고 각 패턴에 FFT를 적용하였다.

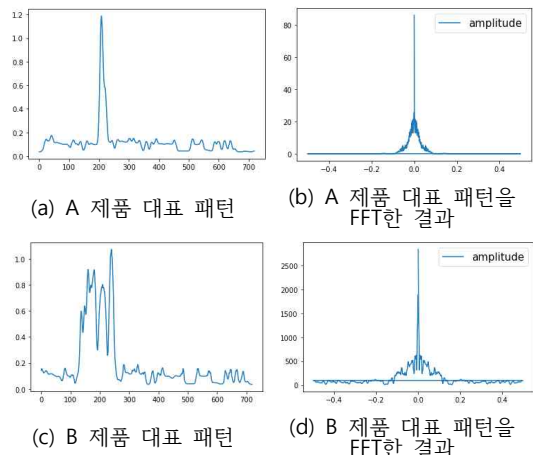


그림 2. 각 제품의 대표 패턴과 FFT 결과

마지막으로 제품별 FFT결과와 대표 패턴에 대한 FFT 결과에 대한 피어슨 상관계수를 계산하였다. 이를 통해 제품별 상관계수에 대한 분포를 만들어 제품분류가 가능한지 확인하였다.

## III. 성능 평가

성능 평가는 설비에서 두 달 동안 수집된 진동 데이터를 사용했다. 전체 데이터에 단순이동평균을 5초 간격으로 실시하고 FFT를 시행하였다. 그림 3은 FFT가 완료된 전체 패턴과 A 제품의 대표 패턴의 상관관계 분석하여 이를 히스토그램으로 나타내었다.

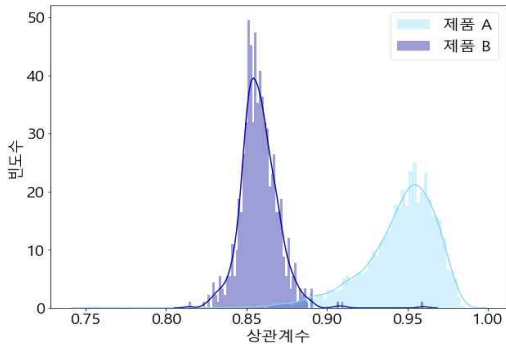


그림 3. 대표 패턴 A와 전체 데이터에 대한 상관계수 분포

그림 4는 B 제품의 대표 패턴과 전체 패턴간의 상관관계를 분석한 결과이다.

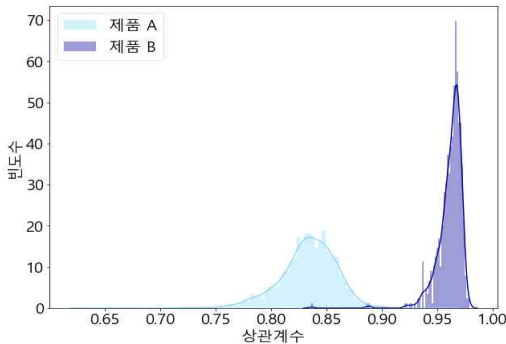


그림 4. 대표 패턴 B와 전체 데이터에 대한 상관계수 분포

그림 3, 4를 통해 각 제품의 상관계수는 고유의 평균과 분산을 가지는 정규분포의 형태를 가진다는 것을 확인 할 수 있다. 또, 같은 제품 간 상관관계를 비교하면 다른 제품을 비교할 때 보다 높은 상관계수를 가진다.

이러한 특징으로 인해 두 정규분포가 겹치는 부분이 적어 두 분포 사이를 구분 짓는 임계값의 설정이 쉽다. 제품 A의 대표 패턴으로 두 분포를 나눌 수 있는 임계값은 0.9, 제품 B의 대표 패턴으로 분포를 구분하는 임계값은 0.92로 설정하였다.

표 1의 내용은 실제 분류 결과와 예측한 분류 결과가 얼마나 차이가 나는지 오차 행렬(Confusion matrix)을 통해 도출한 결과이다.

표 1. 대표 패턴으로 전체 데이터를 분류한 결과

대표패턴	accuracy	recall	precision	f1 score
A 제품	0.954798	0.945761	0.998861	0.971586
B 제품	0.998532	0.995184	0.996784	0.995983

성능 평가 지표로 accuracy, recall, precision, f1-score를 사용하였다. 실험 결과 A 제품을 대표 패턴으로 설정하였을 때, f1-score는 97%였다. 하지

만 B 제품을 대표 패턴으로 설정하였을 때는 앞의 결과보다 2% 더 높은 99%를 보였다.

#### IV. 결론 및 향후 연구

본 논문에서는 서로 다른 두 종류의 제품을 구분하기 위해 제품 생산 과정에서 발생하는 진동 데이터의 대표 패턴을 추출하고 FFT를 하여 상관 분석을 하였다. 그 결과 두 제품의 상관계수 분포가 뚜렷하게 구분되는 것을 확인하였다.

앞으로는 본 연구결과를 머신러닝과 딥러닝 모델의 입력 데이터로 사용하여 설비 예지 보전, 이상탐지 등의 모델 개발에 활용할 것이다.

#### Acknowledgement

본 논문은 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 사회맞춤형 산학협력 선도 대학(LINC+) 육성사업의 연구결과입니다.

#### References

- [1] K. S Noh and S. H Park, "An Exploratory Study on Application Plan of Big Data to Manufacturing Execution System," Journal of Digital Convergence, vol. 12, No. 1, pp. 305 - 311, Jan. 2014.
- [2] G. D. Bergland, "A guided tour of the fast Fourier transform," in IEEE Spectrum, vol. 6, No. 7, pp. 41-52, July 1969.
- [3] Vit Niennattrakul, Dararat Srisai and Chotirat Ann Ratanamahatana, "Shape-based template matching for time series data", Knowledge-Based Systems, Vol. 26, Pages 1-8, 2012.
- [4] S. K. Yadav, K. Tyagi, B. Shah and P. K. Kalra, "Audio Signature-Based Condition Monitoring of Internal Combustion Engine Using FFT and Correlation Approach," in IEEE Transactions on Instrumentation and Measurement, vol. 60, No. 4, pp. 1217-1226, April. 2011.
- [5] D. Y. Lee and J. S. Yoon, S. G. Lee. "제조데이터의 가치와 활용." Journal of the KSME Vol. 57, No. 8, pp. 49-53, Oct. 2017.