

K-Means 군집화 알고리즘을 이용한 주기 변동 랜덤 신호의 정렬 및 이벤트 추출

변성훈
선박해양플랜트연구소
byunsh@kriso.re.kr

Alignment and event extraction of period-varying random signal using K-Means clustering algorithm

Sung-Hoon Byun
Korea Research Institute of Ships and Ocean Engineering (KRISO)

요 약

본 연구는 주기적으로 반복되지만 시간에 따라 주기가 임의적으로 변화하는 주기 변동 신호에 대해 K-Means 알고리즘을 사용해 군집 신호를 추출하고, 추출된 군집 신호를 이용해 실제 주기를 추정하는 신호 정렬 및 추출 알고리즘에 대한 연구 결과를 소개한다. 대상 신호는 왕복 또는 회전 사이클 운동에 의해 주기성을 가지나, 기계적인 특성으로 사이클 길이(주기)가 변화하는 기계 진동 신호가 대표적이다. 제안된 알고리즘은 신호가 일정한 임계값 이하로 작아지는 경우 신호의 값을 0 으로 치환하게 되면, 남은 신호가 군집화되는 원리에 착안하여 K-Means 알고리즘을 이용해 군집화를 진행하였다. 이 때, 각 군집의 파형은 각 사이클마다 랜덤하지만 평균 전력은 유사하게 유지되는 점을 이용하여 기준이 되는 군집을 선정하고, 기준 군집 사이의 간격을 측정하여 주기를 추정하였다. 제안된 알고리즘을 실제 선박 발전기의 진동 신호에 적용하여 사이클 길이를 추정하였으며, 이를 발전기 사양의 표시 회전수에 의한 사이클 길이와 비교하여 알고리즘 수행 결과의 유효성을 확인하였다.

I. 서 론

최근 다양한 기계 시스템의 진동 신호로부터 시스템 건전성을 모니터링하는 기술에 대한 연구가 활발하다. 건전성 모니터링을 위해 측정되는 진동 신호들은 많은 경우 주기적인 실린더 왕복 또는 프로펠러 회전으로 인하여 주기성을 가지지만, 기계의 작동 특성으로 주기가 일정하게 유지되지 않고 시간에 따라 조금씩 임의적으로 변화하는 특성을 보여준다. 따라서 타코미터를 이용해 왕복(회전)에 동기를 맞추고 신호를 수집한 경우가 아니면, 수집된 신호의 주기를 정확하게 파악하기 어렵고 다양한 이벤트 신호들의 시간에 따른 변화를 확인하는데 어려움이 따른다. 본 연구는 이와 같이 주기적으로 반복되지만 시간에 따라 주기가 임의적으로 조금씩 변화하는 주기 변동 신호의 주기를 추정하고, 한 주기 내의 이벤트 신호를 군집화하여 추출하는 방법에 대한 연구 결과를 소개하고자 한다.

II. 문제 설명

본 연구에서 대상으로 하는 주기 변동 신호는 다음과 같은 특징을 가진다.

- 신호의 측정 시작 시간은 임의적이며, 측정 주기가 왕복(회전) 운동에 동기화되지 않음
- 신호는 주기적인 사이클이 반복되나, 각 사이클 길이, 즉 주기는 변화함
- 한 사이클은 기관의 행정 단계 특성을 반영하여 파악 가능한 개수의 군집 신호로 구성됨

- 각 군집 신호의 파형은 임의적이거나, 각 군집 신호의 평균 전력은 유사하게 유지됨

Fig. 1은 선박에서 사용되는 발전기의 진동 신호[1]를 표시 회전수(nominal RPM) 정보를 이용해 사이클 단위로 분리하여 나타낸 것이다. 표시 회전수에 의한 사이클 길이는 실제 사이클 길이와 차이가 존재하므로, 각 신호 군집의 위치가 일정하지 않고 변화하는 것을 확인할 수 있다. 본 연구의 목적은 위와 같은 특징을 가지는 시계열 신호를 같은 기준 위치를 가지는 사이클 단위로 분리하고, 각 사이클의 주요 이벤트 신호를 추출하는 방법을 고안하는 것이다.

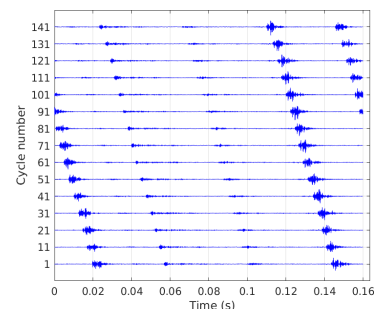


Fig. 1 Example of period-varying signal aligned using the nominal cycle length.

III. K-Means 알고리즘을 이용한 신호 정렬 및 이벤트

추출

본 연구에서는 신호가 일정한 임계값 이하로 작아지는 경우 신호의 값을 0 으로 치환하게 되면, 남은 신호가 군집화되는 원리에 착안하여 K-Means 알고리즘[2]을 이용해 군집화를 진행하였다. 이 때, 각 군집의 파형은 시간에 따라 달라지지만 평균 전력은 유사하게 유지되는 점을 이용하여 기준이 되는 군집을 선정하고, 기준 군집 사이의 간격을 측정하여 한 사이클의 길이, 즉, 주기로 판단하였다.

Fig. 2는 본 연구에서 제안한 알고리즘을 정리하여 나타낸 것이다. 처음 초기화 단계 (Initialization)에서는 군집 중심 위치 정보가 없으므로 다수의 초기값을 입력하여 결과를 계산하고, 그 중에서 가장 작은 점-중심 거리합을 가지는 결과를 최종 결과로 사용한다. 이 때, 각 초기값은 K-Means++ 방법을 사용하여 생성한다[3]. 이후 반복 연산 단계(Iteration)에서는 이전 사이클 처리 과정에서 측정된 중심을 초기값으로 입력하여 K-Means 알고리즘을 수행하도록 하였다.

Input: signal sequence $\underline{x} = [x_0, x_1, \dots, x_{L-1}]$, threshold v , number of clusters N
Output: cycle starting index i_0 , estimated cycle lengths $\underline{M}$, index sets of clusters $\underline{I}$

Initialization

1. Set initial starting index $i_0 (= 0)$ and initial cycle length $M_0 (= \text{nominal cycle length})$
2. Set input signal sequence for clustering, $\underline{x} = [x_{i_0}, x_{i_0+1}, \dots, x_{i_0+M_0-1}]$
3. Thresholding $\underline{x}$: make x_i zero for $i \in I_{\text{thd}}$ where $x(i) < v$ and set $I_{\text{thd}} = I_{\text{thd}}^C$.
4. Cluster I_{thd} using K-Means algorithm with random initial centroid: get N clusters of index $\underline{I}_0 = \{I_0^1, I_0^2, \dots, I_0^N\}$ and centroid locations $\underline{c}_1$.
5. Select target cluster I_0^i and update i_0 under the criterion:
 $i_0 = \underset{i \in I'}{\operatorname{argmin}} |x_i| \geq 0.5 * \text{maximum of target cluster signal level}$
6. Update the initial centroid $\underline{c}_0 = \underline{c}_1$.

Iteration

7. for ($k=1$:MAX_ITER) do
 - a. If $i_0 + M_0 > L$ do
 $K=k-1$
 break
 - end if
 - b. Set input signal sequence for clustering, $\underline{x} = [x_{i_0}, x_{i_0+1}, \dots, x_{i_0+M_0-1}]$
 - c. Thresholding $\underline{x}$: (same as 1-c)
 - d. Cluster I_{thd} using K-Means algorithm with the initial centroid $\underline{c}_0$: get N clusters of index $\underline{I}_k = \{I_k^1, I_k^2, \dots, I_k^N\}$ and centroid locations $\underline{c}_1$.
 - e. Select the target cluster I^i and set i_1 under the criterion:
 $i_1 = \underset{i \in I'}{\operatorname{argmin}} |x_i| \geq 0.5 * \text{maximum of target cluster signal level}$
 - f. Update the cycle length $M_0 = M_0 + i_1 - i_0$ and save $M_k = M_0$.
 - g. Update the initial centroid $\underline{c}_0 = \underline{c}_1$ and the starting index $i_0 = i_1$.
 - h. $k = k+1$
8. end for
9. return $i_0, \underline{M} = [M_0, M_1, \dots, M_K], \underline{I} = [I_0, I_1, \dots, I_K]$

Fig. 2 Algorithm for cycle length estimation

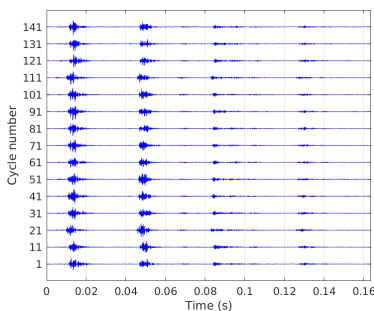


Fig. 3 Signal (in Fig. 1) aligned using the estimated cycle lengths.

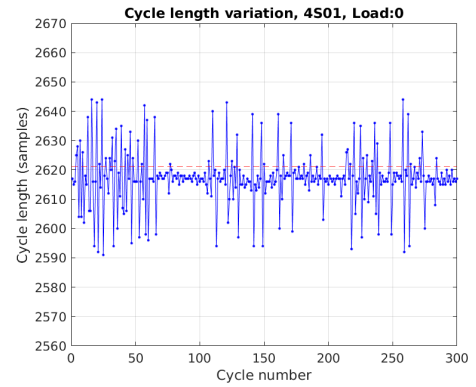


Fig. 4 Estimated cycle length (solid line with dots) and nominal cycle length (dashed) in samples.

Fig. 3은 제안한 알고리즘을 이용하여 Fig. 1의 신호를 사이클 길이 단위로 정렬한 결과이다. 유사한 전력 크기를 가지는 군집 신호들이 같은 시간 위치를 가지도록 정렬된 것을 확인할 수 있다. Fig. 4는 알고리즘으로부터 추정된 사이클 길이를 표시 회전수와 비교한 것으로, 시간에 따라 사이클 길이가 변화하며 평균적인 사이클 길이는 표시 회전수에 의한 사이클 길이 값과 유사하지만 약간 작게 나타났다.

IV. 결론

본 연구는 주기적으로 반복되지만 시간에 따라 주기가 임의적으로 조금씩 변화하는 주기 변동 신호에 대해 K-Means 알고리즘을 사용해 사이클 내의 군집 신호를 추출하고, 추출된 군집 신호를 이용하여 실제 주기를 추정하는 알고리즘을 개발하였다. 제안된 알고리즘을 실제 선박 발전기의 진동 신호에 적용하여 사이클 길이를 추정하였으며, 이를 발전기 사양의 표시 회전수에 의한 사이클 길이와 비교하여 유사하게 사이클 길이가 추정되는 것을 확인하였다.

ACKNOWLEDGMENT

본 논문은 산업통상자원부 국가연구개발사업 “자율운항선박 핵심 기관시스템 성능 모니터링 및 고장예측 진단 기술 개발”으로 수행된 연구결과입니다.

참고 문헌

- [1] 변성훈 외, “자율운항선박 기관시스템 상태 진단을 위한 진동 신호 분석 기법 고찰”, 대한조선학회 2021년 춘계학술대회 논문집, 2021.
- [2] 오일석, 패턴 인식, 교보문고, 2008, pp.349-354.
- [3] Arthur, David, and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. Stanford, 2006.