

손목 움직임 기반 휴식, 걷기, 달리기 분류에 관한 연구

하정호*, 김준호*, 최선탉*, 조위덕*

*아주대학교 전자공학과

e-mail:{hjh5470569, junhokim218, suntaag62, wdukecho}@gmail.com

A Study on Method of Classification by Walking Resting and Running Based on Motion of Wrist

Jeong-Ho Ha*, Jun-Ho Kim*, Sun-Taag Choe*, We-Duke-Cho*

*Dept of Electronic Engineering, Ajou University

요 약

본 논문은 손목에 부착된 단일 3축 가속도 센서를 이용하여 사용자 움직임 기반의 휴식, 걷기, 달리기(느린속도, 빠른속도)를 분류하는 방법에 관한 연구이다. 초당 32회 표본 값의 가속도 정보에서 특징 신호인 평균, 표준편차를 산출하고 사용자의 행동상태를 4가지 상태로 분류한다. 분류 기준이 모호한 상태전이 신호에 대해 6가지 상태로 분류하여 구해진 총 10개의 행동상태 정보를 2차원 평면에 사영하고 최종적으로 K-means 군집화 기법을 적용하여 사용자의 행동상태를 4가지 상태로 분류한다.

1. 서론

웨어러블 디바이스는 사용자의 일상생활에서 발생한 이동거리, 속도, 활동 강도, 소모한 칼로리와 같은 다양한 정보를 수치화 된 값으로 얻을 수 있다. 이를 활용한 피트니스 기능은 사용자의 운동량을 거시적으로 표현해주어 사용자가 본인의 활동량을 직접 확인할 수 있으며, 사용자에게 맞는 맞춤형 건강 정보와 운동 처방을 제공하고 있다. 그러나 대부분의 웨어러블 디바이스는 배터리 용량이 작기 때문에 장시간 사용하는데 문제가 있다.[1] 이러한 문제를 해결하기 위해 낮은 연산 강도를 이용한 사용자의 행동상태를 분류하는 방법이 필요하다. 본 논문에서 분류할 행동은 휴식, 걷기, 달리기(느린속도, 빠른속도) 4가지 상태이다. 신호의 특징을 확인하기 위해 각 행동 별 평균, 표준편차의 값을 사용하였고, 군집화 기법을 적용하여 행동상태를 분류한다.

본 연구는 (그림 1)처럼 X, Y, Z방향의 3축 정보를 수집하는 LifeSense 가속도센서를 이용하여 사용자의 움직임을 기반으로 하는 휴식, 걷기, 달리기(느린속도, 빠른속도)를 분류하는 방법에 관한 연구이다. 사용자의 행동상태를 분류하기 위해 (그림 2)와 같이 사용자의 오른쪽 손목에 가속도센서를 착용한 후 휴식, 걷기, 달리기 행동의 데이터를 수집한다. 휴식의 경우 사용자의 상태가 정적인 상태임을 고려하여 앉은 자세 혹은 가만히 서있는 자세를 모두 포함한다. 따라서 정적인 동작의 경우 손목의 움직임이 특정한 범위 내에서 발생하기 때문에 휴식 상태 가속도 정보는 그 값의 변화가 적다. 걷기 동작은 성인남녀의 경우 평균 4km/h이다.[2] 따라서 실험시 발생하는 이동속도의 오차를 고려하여 2km/h ~ 6km/h의 속도를 걷기 상태로 분류한다. 달리기 동작은 사용자의 이동속도의 증가에 따라 사용자의 두 발이 지면으로부터 떨어지는 순간으로 정하며, 인간의 달리기 최대 속도인 34.48km/h (100m 달리기 세계 신기록 보유자 우사인 볼트의 속도)의 속도를 넘지 않는 범위까지를 달리기 상태라고 분류한다. 수집된 정보로부터 에너지 신호를 추출하고, 추출된 신호는 평균, 분산, 표준편차의 값으로 변환 하는데 사용한다. 3가지로 분류된 특징 값 중 평균, 표준편차 2가지 신호를 이용해 2차원 평면에 사영하고 이로부터 군집을 얻는다. 실제 행동 데이터는 시간에 따라 선형적 특징을 갖기 때문에 행동상태를 단순히 4가지 특징으로 정확히 분류하는데 무리가 있다. 따라서 라벨이 주어진 4가지 특징 외의 행동상태 6개를 선정하여 총 10개의 특징 군집으로 분류하고 최종적으로 4가지 군집으로 분류한다.



(그림 1) LifeSense 가속도센서



(그림 2) LifeSense 가속도센서 부착

2. 본론

2.1 데이터 수집

본 연구에서 사용할 데이터의 수집은 단일 3축 가속도 센서인 LifeSense를 이용한다. 데이터는 초당 32개 표본이며 실험은 우레탄 트랙에서 실시한다. 데이터의 수집 시간은 휴식의 경우 정적인 동작에서 5분, 걷는 동작 2분, 달리는 동작 3분이다.

2.2 데이터 전처리

LifeSense로 부터 받아온 가속도 센서의 정보는 가속도 값 뿐 만 아니라 각 가속도에 대한 정보를 포함하고 있다. 따라서 식(1)을 이용하여 방향에 상관없는 가속도의 특징 신호를 추출한다.

$$A_E = \sqrt{A_x^2 + A_y^2 + A_z^2} \dots\dots\dots \text{식(1)}$$

2.3 평균, 분산, 표준편차 특징 값 산출

추출된 에너지 신호에서 사용자의 행동 상태 중 의미 있는 값을 추출하기 위해 특징이 뚜렷하게 나타나는 신호 표본 120개를 선정하고 식(2), 식(3), 식(4)을 이용하여 평균, 분산, 표준편차의 값을 구한다.

$$\bar{X} = \sum_{i=1}^n X_i / n \dots\dots\dots \text{식(2)}$$

$$s^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n-1) \dots\dots\dots \text{식(3)}$$

$$s = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)} = \sqrt{s^2} \dots\dots\dots \text{식(4)}$$

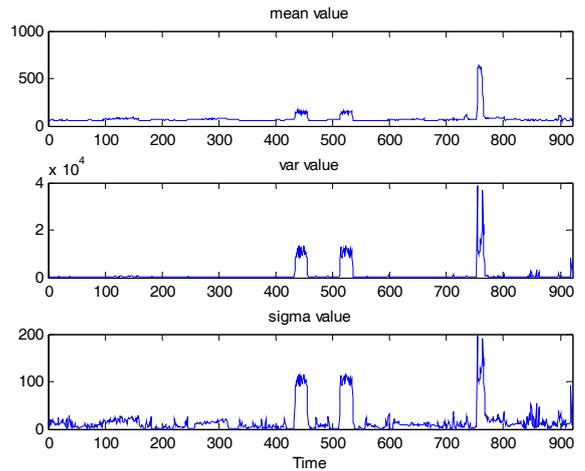
각 행동으로부터 얻은 특징 값은 <표 1>과 같다.

<표 1> 각 행동 별 평균, 분산, 표준편차 값

	평균	분산	표준편차	표본 수
휴식	59.23	1.77	1.33	120
걷기	73.72	367.29	19.16	120
느린 달리기	204.97	8336.21	91.30	120
빠른 달리기	615.55	11876.15	108.97	120

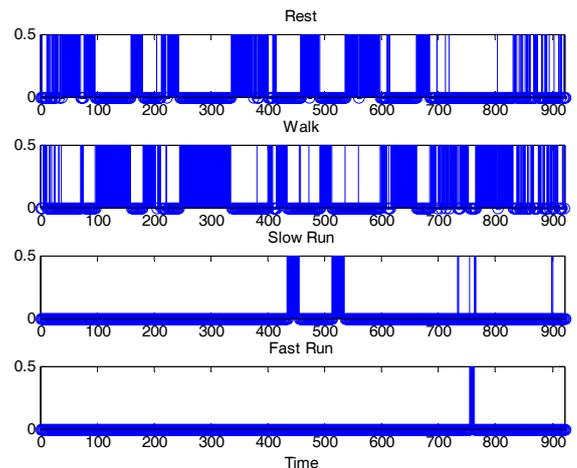
2.4 데이터 선별

선별된 표본에서 확인 가능한 특징의 분포를 (그림 3) 처럼 실제 실험 데이터에서 확인할 수 있으며 그 중 각 행동 특징 값의 분포의 범위가 매우 넓어 오차가 큰 분산 값을 제외한 평균, 표준편차를 이용하여 행동을 분류한다.



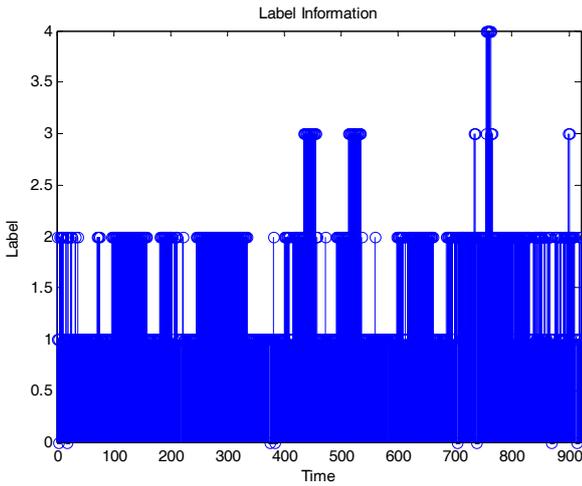
(그림 3) 각 행동상태의 평균, 분산, 표준편차 값

각 행동은 라벨을 이용하여 행동상태를 (그림 4)와 같이 분류하며 연속적인 표현은 (그림 5)와 같다. 분류된 신호는 각 행동상태에 따라 휴식 = 1, 걷기 = 2, 달리기 (느린 속도 = 3, 빠른 속도 = 4)이다.

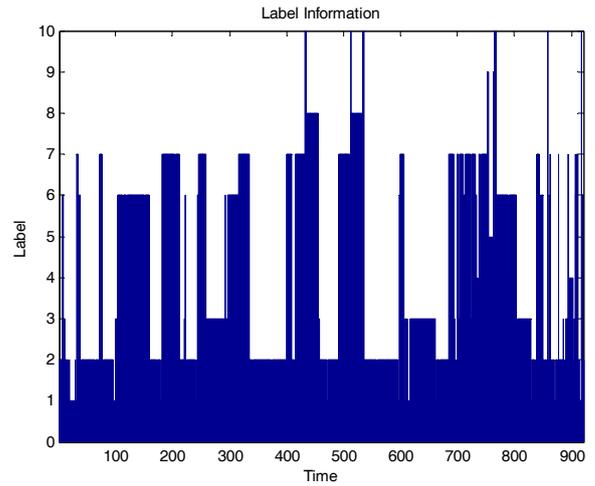


(그림 4) 각 행동상태 별 라벨링

행동 데이터는 시간에 대해 선형성을 나타내므로 연속적인 동작임을 확인 할 수 있다. (그림 4)와 같이 라벨을 이용한 분류는 그 값이 특정 행동상태로 유지되지 않고 수시로 변경되는 것을 확인 할 수 있다. 가령 (그림 4) x축의 값이 200 지점의 경우 실제 실험에서는 휴식인 상황이었지만 분류 결과 휴식과 걷기가 혼재된 상황으로 표현되고 있다. 이는 휴식 중 사용자의 손목 움직임이 걷기에 준하는 값으로 산출되기 때문이다. 따라서 분류기를 사용하지 않고 단순한 분석 수치만을 이용한 라벨에서는 정밀한 행동상태를 분류하기엔 그 기준을 나누기가 모호하다. 따라서 이러한 문제점을 보완하기 위해 분류기 거친 데이터의 분포를 정확하게 파악해야 할 필요가 있다.



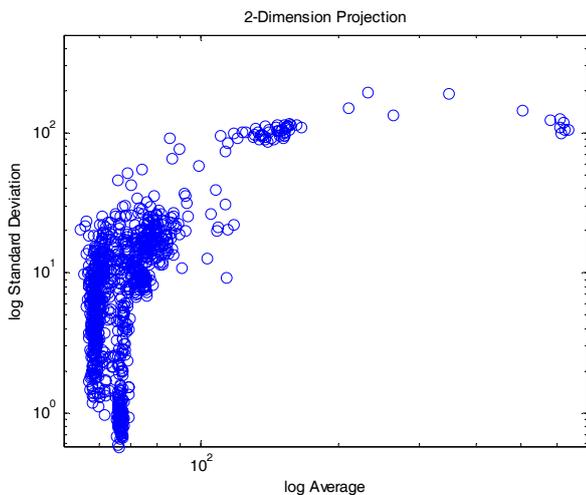
(그림 5) 라벨링을 이용한 행동상태 분류



(그림 7) 행동상태 전이를 고려한 고차원적 라벨링

2.5 군집화 적용

단순 라벨의 데이터는 분류 기준의 정보가 적기 때문에 행동상태를 정확하게 분류하기에는 무리가 있다. 따라서 군집화를 적용하기 위해 (그림 6)과 같이 평균 값을 x축, 표준편차 값을 y축으로 구성된 2차원 평면에 사영한다.



(그림 6) 2차원으로 사용된 행동상태 특징 값

축 범위는 데이터의 분포를 쉽게 확인하기 위해 log 공간으로 변경한다. x축의 범위는 평균 값 기준 50 ~ 700이며 y축의 범위는 표준편차 값 기준 0 ~ 500을 나타낸다.

평균과 표준편차로 구성된 2차원 평면에서는 단순히 라벨링된 (그림 5)와 비교 하였을 때 각 데이터가 어떤 행동상태를 나타내는지 확인하기 어려우며 그 분포에서 군집을 찾아내기 어렵다. (그림 6)과 같이 사영된 행동상태 특징 값은 시계열의 표현이 아니기 때문이다. 그러나 사영을 함으로써 데이터의 특징신호를 구분하는 세밀한 기준점을 확인할 수 있고 행동상태를 군집화 하는데 큰 이점이 있다.

고차원적인 행동상태를 분류하기 위해 행동상태의 전이를 고려하여 기존에 제시한 4가지 단계와 행동상태 전이가 고려된 군집의 개수를 선정한다. 기존 행동상태 분류에 더불어 휴식 ↔ 걷기, 걷기 ↔ 달리기(느린속도), 달리기(느린속도) ↔ 달리기(빠른속도) 세 가지 행동상태 전이의 역을 포함한 총 6가지를 추출하여 최종적으로 (그림 7)과 같은 10개 행동으로 분류한다.

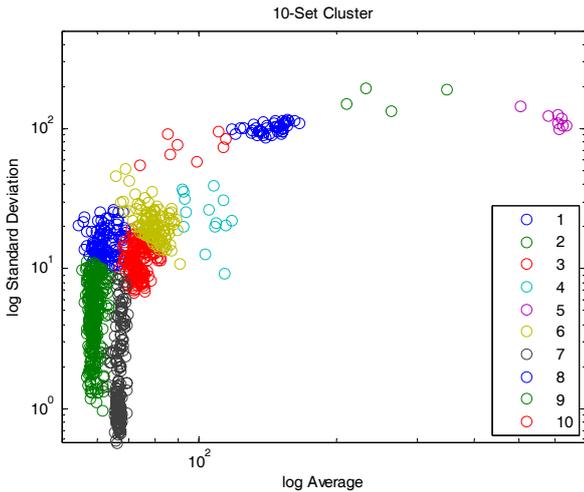
$$V = \sum_{i=1}^k \sum_{j \in S_i} |x_j - \mu_i|^2 \dots \dots \dots \text{식(5)}$$

$$S_i^{(t)} = x_p : |x_p - \mu_i^{(t)}|^2 \leq |x_p - \mu_j^{(t)}|^2 \forall j, 1 \leq j \leq k \dots \dots \dots \text{식(6)}$$

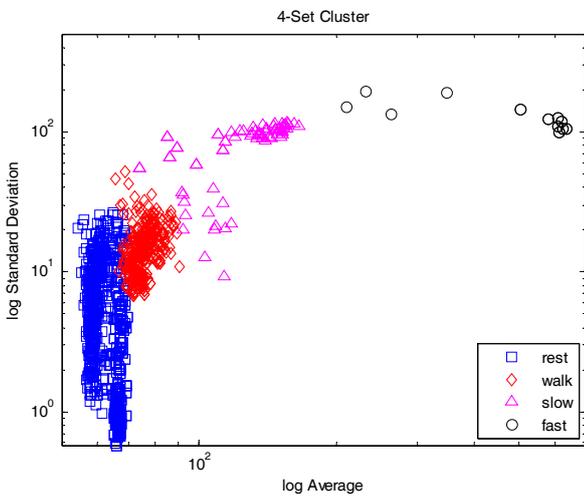
$$\mu_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \dots \dots \dots \text{식(7)}$$

K-means 알고리즘[3]은 입력된 데이터를 유클리디안 거리 기반으로 식(5)를 적용하여 분산을 최소화 하는 방식으로 동작한다. 식(6)에서처럼 순차적으로 데이터 간 평균을 구한 후, 거리를 계산하여 무게 중심을 선정한다. 선정된 무게중심을 기반으로 군집을 형성하게 되고 순차적으로 군집과의 포함 여부를 확인한다. 식(7)을 적용하여 최종적으로 각 군집에 있는 데이터들의 무게중심 값으로 재조정된다. 2차원으로 사용된 행동상태를 10개의 라벨 정보화 군집화 기법을 적용하여 (그림 8)과 같이 분류한다. 이 중 각 군집별 특성을 고려하여 (그림 9)와 같이 행동상태 의미의 연관성이 있는 데이터들 간 최적의 군집을 찾는다.

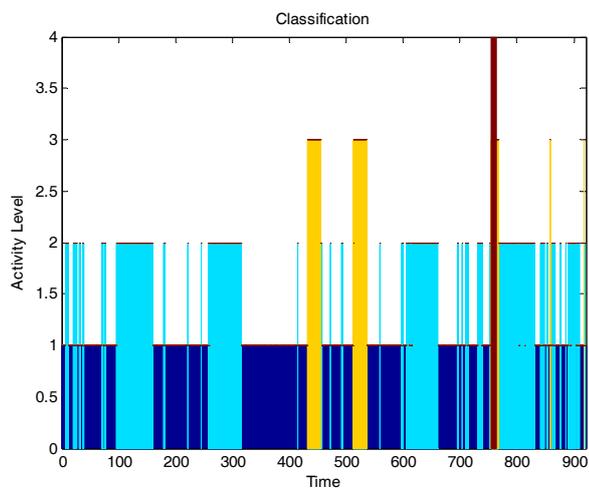
K-means 알고리즘은 라벨이 주어진 데이터들에 대해서만 군집을 분류하는 특징이 있다. 라벨이 주어지지 않은 데이터에 대해서 K-means 알고리즘은 가장 유사하다고 생각되는 군집으로 포함된다고 결론을 내리는 단점이 있다. 따라서 분류의 단계를 나누어 행동전이의 과정을 군집화 중간단계로 인식하고 유사한 행동으로 라벨링하는 단계를 거친다. 그러면 이전 결과보다 정확한 각 행동상태 최적의 군집을 찾을 수 있다.



(그림 8) 행동상태를 10가지 특징 상태로 군집화



(그림 9) 행동상태를 4가지 특징 상태로 군집화



(그림 10) 분류된 행동상태

2.6 결과

사용자의 단일 3축 가속도 센서 정보만을 이용하여 행동상태를 (그림 10)과 같이 휴식, 걷기, 달리기(느린속도, 빠른속도)의 4가지 단계로 분류한다. 분류 과정을 전부를 K-means 알고리즘에 의존하지 않고 행동상태가 모호한 특징 상태에 라벨링을 하여 사용자의 행동상태를 보다 정확하게 분류한다.

3. 결론

본 연구는 단일 3축 가속도 데이터를 이용하여 사용자의 손목의 움직임 값으로 행동상태를 분류하는 방법을 제안한다. 행동상태는 휴식, 걷기, 달리기(느린속도, 빠른속도) 4가지 단계로 분류하였다. 데이터의 분류과정을 단순히 K-means 알고리즘에 의존하지 않고 행동상태가 유사한 특징에 대해서도 분류가 가능하도록 설계한다. 행동상태의 분류 기준이 명확하지 않은 상태전이를 고려하여 행동상태를 전체 10가지 단계로 분류하고 의미상 특징이 알맞은 행동상태끼리 하나의 단계로 구분하여 최종적으로 휴식, 걷기, 달리기(느린속도, 빠른속도) 4가지 단계로 분류한다. 수치적인 데이터만의 사용으로는 명확하게 분류되지 않았던 기준에 대한 모호성을 줄이며 행동상태를 분류할 수 있었다. 향후 본 연구에서 제안한 알고리즘을 통해 사용자의 다양한 행동의 분류를 실시하며, 일상생활에서 빈번하게 발생하는 차량이동, 지하철 탑승, 자전거 탑승과 같은 이동 상태를 분류 할 수 있는 방법에 대해 연구할 것이며, 테니스와 배드민턴과 같이 동일한 스윙 동작이지만 다른 행동상태에서 나타나는 행동에 대해 분류하는 방법에 대해 연구하고자 한다.

Acknowledge

본 연구는 연구성과실용화진흥원의 "모바일 IoT 디바이스를 이용한 활동/수면/식사 자동 행위인식과 전주기 활동 패턴 모델링 기술 개발" 과제의 연구결과로 수행되었음.

참고문헌

- [1] Starner, Thad E. "Powerful change part 1: batteries and possible alternatives for the mobile market." IEEE Pervasive computing 2.4 (2003): 86-88.
- [2] 박세진, 이준수, 강덕희, 정은희, 전효정, 박성빈. (2007.11). 연령에 따른 보행속도 및 보폭에 대한 고찰. 대한인간공학회 학술대회논문집, 430-434.
- [3] Wagstaff, Kiri, et al. "Constrained k-means clustering with background knowledge." ICML. Vol. 1. 2001.