

3축 가속도신호의 비선형 분석방법과 동적 베이지안 네트워크를 활용한 제스처 인식

황주원^o 민준기 조성배

연세대학교 컴퓨터과학과

juwon@sclab.yonsei.ac.kr, loomlike@sclab.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Gesture Recognition using Nonlinear Analysis Method of 3D Acceleration and Dynamic Bayesian Network

Ju-Won Hwang^o Jun-Ki Min Sung-Bae Cho

Dept. of Computer Science, Yonsei University

센서기술의 발전으로 인해 다양한 HCI(Human Computer Interaction) 기술이 개발되고 있다. 센서가 부착된 디바이스를 이용하여 사용자의 동작을 인식하는 기술이 이중 하나이다. 선형 가속도계는 정해진 방향의 가속도를 측정하는 센서로서, 디바이스에 장착하면 사용자의 동작에 따른 가속도 정보를 획득할 수 있으며 [1], 특히 3축 가속도 센서는 입체적인 움직임을 잘 인식할 수 있다.

본 논문에서는 스마트폰을 이용하여 수집한 가속도 신호를 슬라이딩 윈도우를 이용하여 세그먼테이션 한 뒤, 각 세그먼트에 해당하는 특징을 추출하고, 제안하는 비선형 분석법과 확률기반 세그먼테이션 네트워크를 이용하여 제스처와 non-motion을 분류한다. 또한 제스처 인식 네트워크를 이용하여 현재 어떠한 제스처인지 결정하는 것이 목적이다. 그림 1은 제스처 인식 방법의 순서이다.

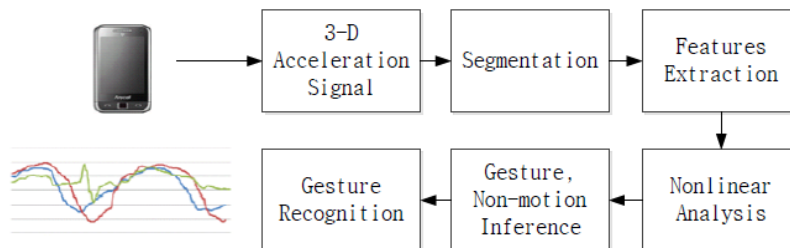


그림 1 제스처 인식 방법 순서

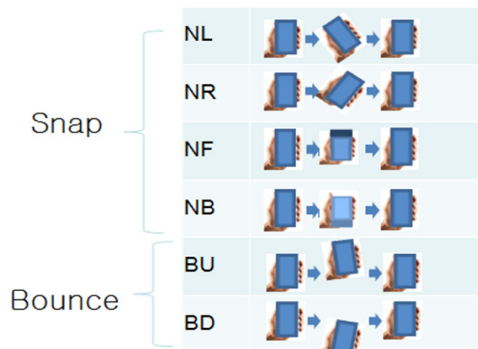
먼저, non-motion과 제스처 구간을 세그먼테이션하기 위해서 트레이닝 데이터를 이용하여 특징을 추출하였다. 특징 추출은 슬라이딩 윈도우를 이용하여 앞 시점에서의 데이터를 현재 시점에 적용시키기 위해서 현재 시점의 슬라이딩 윈도우를 앞 시점의 것과 일부분 중복(overlap)시킨 구간을 이용하였다. 특징은 평균, 절대값평균, 표준편차, 에너지, 공분산, 인접한 점간의 차이 값 누적, 제스처 시작점과 끝점에서의 중력방향 차이를 사용하여, 총 21(3축*7)개의 특징을 사용한다. 또한 가속도 값만을 고려하면 동작 중 가속도 값이 0을 통과하는 부분에서 올바른 값을 추출할 수 없으며, 변화량 값만을 고려하면 가속도 값이 최대 혹은 최소값을 가짐에도 불구하고 작은 값의 결과를 얻게 된다. 이 점에 착안하여 가속도 값을 구하고, 이에 대한 변화량이 큰 부분을 제스처 시작부분으로 인식하도록 하였다.

다음으로는, 제스처와 non-motion을 구분하고자 세그먼테이션 네트워크를 설계하였다. 하지만 특징 데이터를 추출하고 분석하는 과정에서 데이터에 불확실성이 포함될 수 있으며, 경우에 따라서는 non-motion구간에서 노이즈가 발생하여 제스처일 때와 유사한 특징을 보일 수 있다. 이와 같은 문제를 극복하고 각 특징과 non-motion 및 제스처 일 때의 관계들을 효율적이고 수학적으로 표현하기 위해서 동적 베이지안 네트워크를 이용하였다. 베이지안 네트워크는 노드와 아크(arc)로 표현되는 방향성 비순환 그래프(DAG: directed acyclic graph)형태로서, 노드마다 정의된 조건부 확률 테이블에 의해 적은 비용으로 노드간의 확률 관계를 효율적으로 표현하며[2], 인간의 인과적인 판단 및 추론 과정을 모델링하기에 유리한 도구이다[3]. 특히, 노드 사이의 외형적인 시간적 관계를 모델링할 수 있는 동적 베이지안 네트워크를 사용하여 가속도 데이터의 시계열 특성을 반영할 수 있고, 제스처의 시작점과 끝점의 불명확함으로 인해 많은 불확실성이 내재되어 있는 문제를 해결하고자 세그먼테이션 네트워크를 설계하였

다. 세그멘테이션 네트워크의 각 특징노드에 속하는 조건부확률테이블을 설계하기 위하여 제스처와 non-motion의 가속도 데이터를 특정한 값으로 명시할 수 있는 분석과정이 필요하다. 이를 위해서 non-motion과 제스처에 해당하는 가속도 데이터를 학습하여 각각의 특징 데이터에서 다섯 개(최소값, 상위25%, 상위50%, 상위 75%, 최대값)의 기준 값을 추출하였다. 추출한 10개의 값은 해당하는 각 클래스의 특정 패턴을 보이는 의미있는 값이기 때문에 분석 단계에서 오름차순으로 정렬된 값을 이용하였다. 믹스처 가우시안 모델을 사용하여 non-motion과 제스처 두 모델이 특징 정보 값이 겹치는 구간에서 기준값이 집중되어 있음을 확인하였다. 두 모델이 겹치는 구간이란 다른 범위의 데이터보다 non-motion과 제스처일 때 많이 발생할 수 있는 데이터 값을 의미한다. 본 논문에서의 비선형 분석방법을 통해 non-motion과 제스처에 해당하는 10개의 state값을 학습시키는 확률기반 세그멘테이션 네트워크를 설계하였다.

마지막으로 제스처 인식 단계에서는 앞서 획득한 세그멘테이션 결과를 이용하였을 때의 제스처 인식을 확인하기 위해 제스처 인식 네트워크를 설계하였다. 제스처 인식 네트워크 또한, 제스처 데이터에 대한 흐름과 데이터 사이에 있을지도 모르는 non-motion과 같은 노이즈가 있을 불확실성을 효과적으로 다루기 위해서 동적 베이지안 네트워크를 이용하여 설계하였다. 세그멘테이션 네트워크와는 다르게 각 제스처에 대한 데이터를 학습시켜 각각 다른 비선형 분석 과정이 필요하다. 제스처 인식 네트워크의 입력은 제스처일 때의 특징 값이며, 출력은 각 제스처 네트워크를 사용하여 얻은 추론 값이다. 이때, 행동별로 구성되어 있는 네트워크 중 가장 확률값이 높은 행동이 최종 행동으로 선택된다.

실험은 확률기반 세그멘테이션 네트워크의 정확도와 제스처 인식으로 구성하였다. 이를 위해 가속도 센서가 부착된 삼성전자의 T-omnia 스마트 폰을 이용하여 5명의 대학생이 이를 동안 수집하였으며, 수집한 데이터의 2/3는 학습 데이터로 사용하였고 1/3은 테스트에 사용하였다. 세그멘테이션 네트워크 성능 평가 결과 precision은 88.4%, recall은 88.23%, accuracy는 86.168%임을 확인하였다. 제스처 네트워크의 인식을 평가는 그림 2에 기술한 제스처를 이용하여 진행하였다. 실험결과 표 1에서 볼 수 있듯이, 모든 제스처에서 평균적으로 약 80%의 인식을 확인할 수 있었다.



	Precision (%)	Recall (%)	Accuracy (%)
BD	80.5	82.78	78.92
BU	82.22	75.04	75.24
NB	49.08	71.58	68.92
NF	82	84.47	75.2
NL	77.15	75	68.3
NR	80.41	77.95	72.27

그림 2 인식기 성능 평가에 사용된 제스처

표 1. 확률기반 제스처 인식 네트워크 성능

감사의 글

본 연구는 지식경제부 및 한국산업기술평가관리원의 산업원천기술개발사업의 일환으로 수행하였음 (10033807, 다중센서 및 협업을 위한 자율 학습 기반 상황인지 기술).

참고문헌

[1] S. K. Kim and S. J. Choi, "HMM-based motion recognition with 3-D acceleration signal," *Journal of KIISE: Computing Practices and Letters*, vol. 15, no. 3, pp. 216-220, 2009.

[2] K. B. Korb and A. E. Nicholson, *Bayesian Artificial Intelligence*, Chapman & Hall/CRC, 2003.

[3] M. L. Wong, W. Lam, and K. S. Leung, "Using evolutionary programming and minimum description length principle for data mining of Bayesian networks," *IEEE Trans. on Pattern Analysis and Machine intelligence*, vol. 21, no. 2, pp. 174-178, 1999.