

동적 베이지안 네트워크를 이용한 멀티모달센서기반 사용자 행동인식

양성익^o 홍진혁 조성배

연세대학교 컴퓨터과학과

{unikys, hjinh}@sclab.yonsei.ac.kr, sbcho@yonsei.ac.kr

Activity Recognition based on Multi-modal Sensors using Dynamic Bayesian Networks

Sung-Ihk Yang^o Jin-Hyuk Hong Sung-Bae Cho

Dept. of Computer Science, Yonsei University

1. 서론

최근 상황인지를 활용하고 있는 다양한 서비스들이 제공되면서, 사용자의 현재 상황을 인식하는 방법이 활발히 연구 되고 있다[1]. 온도 등 사용자 주변 환경, 사용자의 위치, 사용자의 기분, 사용자의 행동 등이 대표적 컨텍스트이며, 이 중에서 사용자의 행동을 알아내는 것은 복잡하고 어려운 일이다. 사용자의 행동을 인식하는 연구로는 적외선 카메라로 찍은 비디오 영상을 기반으로 인식하는 연구[2]와 GPS를 기반으로 하는 연구들이 있는데[3], 비디오 영상을 기반으로 행동을 인식하는 경우에는 영상 처리 등과 같이 복잡한 전처리 과정이 있어 연산량이 많고, GPS를 기반으로 하는 연구는 위치와 이동 경로에 따른 행동 추론을 하기 때문에 상세한 신체 동작을 인식하는 것은 어렵다. 이에 따라 최근의 많은 연구들은 사용자의 행동을 직접적으로 표현해주는 센서를 기반으로 행동을 인식하고 있으며, 센서의 데이터를 수집 및 전처리를 일괄처리하고 있는 것에 초점이 맞춰져 있다. 사용자에게 서비스를 제공해주기 위해서는 시계열 분석을 통한 정확하고 적은 연산량이나 분산 처리 가능한 행동 인식 방법이 요구된다.

2. 본론

제안하는 방법의 전체적인 흐름은 다음과 같다. 먼저, 사용자가 행동을 하면, 센서로부터 센서 데이터를 수집하게 된다. 센서 데이터는 각 센서별로 수집 모듈을 통해서 입력받고, 멀티모달의 센서를 동일한 숫자형으로 변환 후 통합한다. 통합한 센서 데이터는 전처리 단계와 컨텍스트 추론 단계에서 사용되며, 전처리 단계의 데이터는 다시 컨텍스트 추론을 하는데 활용된다. 이 두 단계의 결과는 행동 인식기에서 활용하게 되며, 그 후 나온 행동 인식기의 결과는 공유 데이터 메모리에 저장하여 다른 서비스에서 활용할 수 있도록 제공하고 있다.

멀티모달 센서의 데이터 수집은 각 센서별 수집 모듈을 통하여 입력이 되고, 센서 통합 시스템에서 데이터를 통합한다. 통합된 센서 데이터를 바로 사용하게 되면 잡음이 많고 데이터가 연속적이기 때문에 잘못된 결과를 유도할 수 있으므로, 행동 인식에 바로 사용하기에 적합하지 않다. 따라서, 통합한 데이터의 잡음을 없애주고 연속적인 데이터를 이산화 시켜주는 전처리가 필요하다. 사용한 전처리 기법으로는 잡음 제거를 위한 데이터 평활화와 데이터 이산화를 위한 클러스터링을 사용한다. 데이터 평활화에서는 특정 크기의 데이터 윈도우가 이동하면서 각 지점별로 데이터의 평균값으로 잡음을 줄여주는 사각 평활화를 활용했고, 이산화를 위한 방법으로는 K 개의 군으로 데이터를 이산화 시켜주는 K -means 군집화 기법을 사용했다.

제안하는 방법은 다양한 행동 인식을 위한 확장성을 위해 다수의 확률 추론모델로 구성된 분산형 추론모델을 설계하여, 사용자의 행동을 인식한다. 분산형 추론 모델은 전역 추론모델과 지역 추론모델로 나뉜다. 이렇게 설계함으로써 연산량의 분산이 용이하며, 전역 추론모델을 통하여 행동을 선분류하고 지역 추론모델을 통해 후분류를 하게 된다. 전역 추론모델은 연산량이 적은 베이지안 네트워크로 추론하

게 되며, 입력되는 센서 데이터로 현재의 행동을 확률적으로 인식한다.

전역 추론모델은 저급의 센서 데이터를 논리적으로 해석한 고급 컨텍스트 정보를 활용하기 때문에, 그 결과를 빠르게 추론할 수 있다. 하지만 베이지안 네트워크를 통해서 행동을 인식하는 것은 한 순간의 센서 데이터에 의존적이기 때문에 연속적인 행동의 인식은 어렵다. 따라서, 전역 추론모델의 추론 결과에 따라 더 정확하게 인식할 지역 추론모델을 선택한다. 지역 추론모델은 각 행동별로 존재하고, 연속적인 행동의 인식에 용이한 동적 베이지안 네트워크를 활용하여 저급의 센서 데이터를 시계열 분석하게 된다.

행동을 추론하기 위한 전역 추론모델은 베이지안 네트워크이며, 각 부위별로 부착된 센서를 기반으로 특징 추출한 결과와 생리적 신호를 통하여 들어온 센서 데이터를 증거 값으로 사용한다. 각 행동은 동적 베이지안 네트워크를 OVR전략으로 학습하여 각각의 지역 추론모델이 존재한다. 지역 추론모델은 전역 추론모델과는 달리 센서 데이터를 고급의 컨텍스트로 해석하지 않고, 각 센서 데이터를 그대로 활용하여 행동을 인식하게 된다. OVR전략으로 각 행동별로 추론모델을 설계하게 될 경우, 동시에 다수의 행동에 대하여 동일한 추론 결과가 나올 경우 한 가지 행동을 선택하는 방법이 필요하다. 지역 추론모델들의 결과를 선택하는 방법은 전역 추론모델을 통한 추론결과의 행동들을 확률순으로 정렬하여 우선순위를 주고, 우선순위가 높은 행동부터 해당하는 지역 추론모델로 검증한다. 특정 행동의 추론모델에 대한 검증 결과가 특정 임계치 이상이면, 그 행동이 행동 인식의 결과가 된다. 지역 추론모델의 검증 결과가 임계치 이하로 나오면, 다음으로 우선순위가 높은 행동의 지역 추론모델로 넘어가 검증을 하게 된다. 이러한 단계를 거쳐서 임계치를 넘는 지역 추론모델이 나오게 되면, 나머지 우선순위가 낮은 행동들의 지역 추론모델에 대해서 검증을 하지 않고 결과를 구할 수 있다. 이는 연산량이 적은 전역 추론모델을 통해 현재 순간의 확률 높은 행동부터 지역 추론모델을 통해 검증하여 그 결과가 참이 되면, 이후 확률이 낮은 행동에 대하여 검증을 하지 않으므로 전체 OVR모델에서 전체 행동인식 모델에서 필요한 연산량을 줄여주고, 각 베이지안 네트워크는 분산 시스템 환경에서 분산이 가능하여 연산량의 균형 조절에도 유용하다.

실험에서 인식하는 행동으로는 사무실 환경에서 일어날 수 있는 식사, 낮잠, 휴식, 전화, 공부, 화장실, 운동, 회의 등의 8가지 행동을 정의하여 데이터를 수집했다. 실험에서 사용한 멀티모달 센서들은 생리적 신호 센서와 가속도 센서를 사용하였다. 생리적 신호 센서로 피부온도와 열유량 그리고 피부 전기 반응도를 수집하였고, 가속도 센서로 각 머리, 오른팔, 왼팔, 오른 손목과 왼 손목에 착용하여 상반신의 동작에 대한 데이터를 수집하였다. 실험의 결과는 전체 분류 확률이 97.4%로 나왔고, 모든 행동이 대체로 94%이상의 높은 정확률을 보여주고 있다. 각 행동은 틀리게 추론한 경우는 다른 행동을 맞다고 인식한 경우가 대부분이었으며, 행동을 인식하지 못하는 경우도 조금 있었다.

3. 결론

본 논문에서는 멀티모달 센서를 기반으로 계층적 동적 베이지안 네트워크를 활용하여 실시간으로 행동을 인식하는 방법을 제안하였다. 실시간으로 행동을 인식하기 위하여 멀티모달 센서를 통합하는 프레임워크를 활용하여, 데이터 수집과 데이터 전처리 과정이 실시간으로 이루어졌고, 이 데이터를 연산량이 적고 분산 시스템에 적용 가능한 계층적 동적 베이지안 네트워크에서 변수 값으로 사용하여 행동을 인식하였다. 실험 결과로 동적 베이지안 네트워크를 통한 행동 인식의 결과가 97.4%로 행동 인식을 위한 베이지안 네트워크의 성능을 입증하였다.

향후에는 계층적 동적 베이지안 네트워크와 일반 동적 베이지안 네트워크 및 다른 방법과의 성능 비교와 연산량 차이에 대한 분석을 수행하고, 개인별 행동의 크기와 세기의 차이를 극복하고 자연스럽게 모호한 행동을 인식하는 방법과, 실제로 분산 시스템에 적용할 경우 추론의 로드를 고려하는 방법에 대한 연구가 필요할 것이다.

참고문헌

- [1] J. Kolari, T. Laakko, T. Hiltunen, V. Ikonen, M. Kulju, R. Suihkonen, S. Toivonen and T. Virtanen, "Context-aware services for mobile users," *VTT Information Technology*, 2004.
- [2] J. Han and B. Bhanu, "Human activity recognition in thermal infrared imagery," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. 3, pp. 17-25, 2005.
- [3] L. Liao, D. Fox and H. Kautz, "Hierarchical conditional random fields for GPS-based activity recognition," *Robotics Research*, vol. 28, no. 3, pp. 487-506, 2007.