

은닉 마코프 모델을 이용한 행동 분류 연구

서정우*, 오현교**, 조승호*, 이호석***, 문봉희****

*강남대학교 컴퓨터미디어정보공학부

**팜웨어뱅크

***호서대학교 뉴미디어학과

****숙명여자대학교 컴퓨터과학과

e-mail:sjw0825@daum.net

A Study on Human Behavior Classification using a Hidden Markov Model

Jeong_U Seo*, Hyeon_kyo Oh**, Seung_ho Cho*, Ho_Seok Lee***, Bong_hee Moon****

*Dept of Computer Media Engineering, Kang-Nam University

**Firmware Bank

***Dept of New Media, Ho-Seo University

****Dept of Computer Science, Sook-Myung Women's University

요약

최근 다양한 센서들이 일상생활에 활용되어, 일정한 환경에서 사람의 행동을 분류하고 인식하기 위한 연구들이 활발하게 진행되고 있다. 본 연구에서는 2개의 진동센서 값과 1개의 적외선 센서 값을 은닉 마코프 모델에 적용하여 침대 위에 있는 사람의 3가지 행동유형-눕기, 뒤척임, 일어나기-을 분류하고자 한다. 3개 센서 값의 특징들을 기초로 은닉 마코프 모델에 학습시키고, 특징집합과 학습 데이터량을 변화시키면서 사람의 행동유형에 대한 인식 실험을 수행하였다. 특징 개수 혼합에 따른 인식률의 차이는 거의 없는 것으로 나타났으나, 학습 데이터량을 증가시켜 가면서 수행한 실험에서는 인식률이 평균 78.127%로 향상되는 성과를 거두었다.

1. 서론

최근 다양한 센서들이 일상생활에 활용되면서 일정한 환경에서 사람이 행동 패턴을 파악하기 위한 연구들이 활발하게 진행되고 있다. 특히 집 안에서 주로 거주하고 있는 노인, 환자나 어린아이들의 경우 이들의 일상생활에 불편함을 초래하지 않는 비 침습적인 방식으로 이들의 일상 행동 패턴을 관찰하고자 하는 요구들이 증가하고 있다.

수면은 일생의 3분의 1을 차지하며, 올바른 수면은 일상생활을 영위하고 질병을 예방하는데 큰 도움을 준다. 따라서 수면 모니터링에 대한 연구가 활발히 진행되고 있으며, 최근에는 스마트폰에 내장된 가속도 센서, 조도센서 등 다양한 센서를 이용하여 수면 모니터링을 하는 어플리케이션들이 사용자들로부터 좋은 평가를 받고 있다. 대표적인 어플리케이션인 Azumio사의 Sleep Time[1]은 가속도 센서로 수면 중 뒤척임을 인식하여 가장 적합한 기상시간에 알람을 울려주며, Squall Line Software사의 Sleepmeter [2]는 수면시간 및 시간대별 수면의 깊이를 측정하여 사용자의 수면패턴을 분석해주는 스마트폰 어플리케이션들이다.

지난 수년간 기존 연구에서는 무선 센서 네트워크를 이용하여 비침습적 방식으로 사람의 움직임을 분류하고자 하는 연구들이 진행되어 왔다[3,4]. 그중에서도 강남대학교의 연구에서는 이벤트 발생시, 입력 센서값을 이용하여 행

동유형의 분류를 시도하였다. 이러한 방법은 입력 센서값이 기준에 설정된 기준 범위에 미달하거나 초과하는 경우, 인식률이 저하되는 한계가 있었다. 본 연구에서는 은닉 마코프 모델의 적용을 통해 이러한 한계를 극복한 새로운 행동유형 분류 모델을 제시하고자 한다.

본 논문의 2장에서는 관련된 연구 배경에 대하여 기술하고, 3장에서는 본 연구에 은닉 마코프 모델을 어떻게 적용하였는지 설명하고, 4장에서는 특징 집합 또는 학습 데이터 량에 따른 인식 결과의 영향에 대하여 기술하고, 5장에서 결론을 맺는다.

2. 연구배경

그동안 강남대학교 유비쿼터스 컴퓨팅 연구실에서는 무선 센서 네트워크를 이용하여 비침습적 방식으로 침대 위 사람의 움직임을 3가지로 분류하는 연구를 진행해 왔다 [4]. 이 연구를 통해 구축된 시스템은 무선 센서노드들, 베이스 스테이션, 행동 분류기, 데이터 베이스 등으로 구성된다. Huins사의 UBee430 무선 통신 모듈을 사용하였고, 진동센서 2개와 적외선 감지 센서 1개를 통해 사람의 움직임 신호를 수집하였다[5].

베이스 스테이션은 무선 센서 노드들이 감지한 움직임 신호들을 ZigBee 무선 통신으로 전송받아 PC 모니터링 프로그램으로 전달한다. 모니터링 프로그램은 행동 분류

알고리즘을 통해 수신된 데이터로부터 현재 어떤 행동이 일어났는지 분류하였다. 이 연구에서는 침대에 눕기, 뒤척이기, 침대에서 일어나기 등 3가지 행동유형에 대한 분류를 시도하였으며, 행동유형 분류 시 행동유형의 특징벡터로 사용된 특징은 각종 센서의 일정 기준값 또는 최대값을 기반으로 한 6가지 특징이었다. 이러한 특징벡터를 기준으로 통계 기반의 행동분류 알고리즘을 적용한 바 있다.

그런데, 이 연구에서는 학습 데이터로 선택된 피실험자의 데이터 특성이 인식 결과에 커다란 영향을 끼치는 한계가 있었다. 본 연구에서는 이러한 데이터들에 대해 은닉 마코프 모델의 적용을 통해 한계를 해결하고자 시도하였다.

3. 은닉 마코프 모델

3.1 센서 데이터 전처리

센서 노드로부터 수집되는 데이터는 시간에 연속적이므로, 행동분류 모델의 입력으로 제공하기 위해서는 데이터를 적절한 구간으로 분할하여 추출할 필요가 있다. 행동유형 데이터의 단위는 사람이 움직이기 시작하여 움직임이 멈추는 상황까지를 구간으로 설정하였으며, 이는 적외선 센서 값이 급격하게 올라가는 시점부터 일정 값 300 이하로 내려갈 때까지로 설정하였다. 전처리를 수행한 결과 연속적인 데이터 중에서 움직임이 발생한 구간의 데이터만 추출하였다[4].

1,	8,	RI_인체감지,	2010-01-13 04:22:09:328
2,	504,	RI_베개 진동감지,	2010-01-13 04:22:09:328
3,	784,	RI_침대 옆 진동감지,	2010-01-13 04:22:09:328
4,	23,	RI_문도,	2010-01-13 04:22:09:328
4,	21,	RI_습도,	2010-01-13 04:22:09:328
4,	2712,	RI_조도,	2010-01-13 04:22:09:328
1,	18976,	RI_인체감지,	2010-01-13 04:22:09:375
2,	504,	RI_베개 진동감지,	2010-01-13 04:22:09:375
3,	784,	RI_침대 옆 진동감지,	2010-01-13 04:22:09:375
4,	23,	RI_문도,	2010-01-13 04:22:09:375
4,	21,	RI_습도,	2010-01-13 04:22:09:375
4,	2712,	RI_조도,	2010-01-13 04:22:09:375
1,	18976,	RI_인체감지,	2010-01-13 04:22:09:453
2,	480,	RI_베개 진동감지,	2010-01-13 04:22:09:453
3,	784,	RI_침대 옆 진동감지,	2010-01-13 04:22:09:453
4,	23,	RI_문도,	2010-01-13 04:22:09:453
4,	21,	RI_습도,	2010-01-13 04:22:09:453

[그림 1] 전처리 전 원 데이터

8	504	784
18976	504	784
18976	480	784
18976	480	728
7560	480	728
7560	496	728
7560	496	816
3096	496	816
3096	496	816
3096	496	696
32712	496	696
32712	512	696
32712	512	672
88888	512	672

[그림 2] 전처리 후 센서값

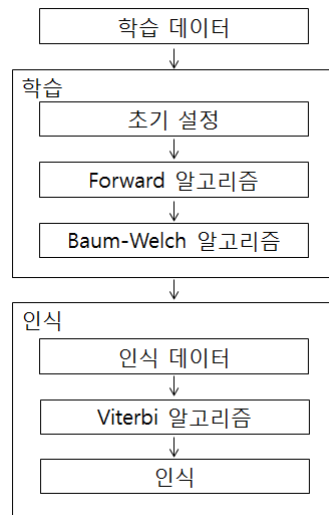
[그림 1]은 침대에 눕는 동작의 데이터 집합을 보여준

다. 인체감지가 처음 발생하고, 그 다음 1초 후 침대 옆 진동이 발생했으며, 다시 1초 후 마지막으로 베개 진동이 발생하였으므로, 침대에 접근하여 침대위 베개에 눕는 동작임을 유추할 수 있다.

[그림 2]는 은닉 마코프 모델에 입력되는 침대 위 사람의 움직임에 대한 데이터를 보여준다. 각 행은 6/10초 시간에 따른 데이터 집합이며, 각 열은 왼쪽부터 순서대로 적외선 센서, 베개 진동센서, 침대 옆 진동 센서의 ADC값으로 시간 t에 대한 특징 3개이다. 각 센서 값의 범위는 0~32,765이며 이 값이 클수록 많은 움직임이 일어났음을 뜻한다. 우리는 시간 흐름에 종속적인 행동 데이터들을 은닉 마코프 모델에 학습시킨 후, 다양한 행동 패턴들을 인식하고자 한다.

3.2 은닉 마코프 모델 적용

은닉 마코프 모델은 마코프 모델을 구성하는 내부 상태(state)를 알지 못한다고 생각하고, 상태의 개수, 상태 사이의 이동, 상태의 결과 출력 등을 확률로 정의한 모델이다. 은닉 마코프 모델을 활용하기 위해서 Forward 알고리즘, Viterbi 알고리즘, Baum-Welch 알고리즘 등을 사용한다. Forward 알고리즘은 출력을 발생시킨 은닉 마코프 모델을 찾는 알고리즘이며, Viterbi 알고리즘은 출력을 발생시킨 내부 상태를 파악하기 위한 알고리즘이고, Baum-Welch 알고리즘은 은닉 마코프 모델을 학습시키기 위한 알고리즘이다[6,7,8].



[그림 3] HMM 동작 순서도

본 연구에서는 진동센서 2종의 센서값과 적외선 센서의 센서값이 동일한 시간에 입력된 3개의 값이므로, 이들을 시간에 대한 특징벡터로 설정하여, 은닉 마코프 모델에 적용하였다. 즉, [그림 2]의 데이터는 3개의 센서로부터 시간의 흐름에 따라 검출한 값으로서, 상태 사이의 이동(변화)을 나타낸다. 따라서 센서 값들을 입력하고 forward 알고리즘을 사용하여 은닉 마코프 모델을 훈련한 후, 센서 값

들의 변화를 인식하면, 시스템 상태 사이의 이동을 알 수 있다. 본 연구에서는 3가지의 상태 변화를 인식한다. 따라서 3 종류의 상태 변화를 은닉 마코프 모델에 학습시켜서 모델을 준비하고, 이후에 센서에서 검출된 데이터를 입력하고 forward 알고리즘을 사용하여 상태 변화를 인식하는 시스템을 구성하였다. 다음 4절에서 설명하는데, 본 연구에서 사용하는 상태 변화는 눕기(S→L), 뒤척임(L→L), 일어나기(L→S) 3종류이다. [그림 2]는 눕기(S→L) 상태 변화를 나타낸 것이다.

4. 실험 및 결과

본 연구에서는 침대에 눕기, 뒤척임, 일어나기 등 3가지 행동유형에 대해 분류를 시도하였다. 침대에 눕기 행동유형은 침대에 앉아 있다가 눕는 행동(S→L 상태)이고, 뒤척임은 침대에 누워 뒤척이는 행동유형(L→L 상태)이며, 일어나기는 침대에 누워 있다가 일어나는 행동유형(L→S 상태)을 의미한다.



[그림 4] 실험 광경

본 실험에서는 피실험자 8명이 3가지 행동유형에 대해 5회씩 반복하여 총 120회 행동유형 실험을 수행하였다. 5회 수행한 실험 중에서 2회 실험 데이터는 은닉 마코프 모델의 학습을 위해 사용하였고, 나머지 3회 실험 데이터는 인식 실험을 위해 사용하였다. 이 실험에서 인식 테스트로 선택된 데이터도 이미 피실험자가 어떠한 행동한 것인지 알고 있다. 실험은 크게 2가지로 수행되었다. 첫 번째 실험에서는 특징집합을 변화시켜 가면서 인식에 끼치는 영향을 분석하여 보았고, 두 번째는 학습 데이터량을 변화시켜 가면서 인식 결과를 분석하였다. 본 실험에서 특징집합은 다음과 같이 구성하였다.

- 1) 전 처리 후 얻어진 3가지 센서값을 특징 3개로 구성
- 2) 위의 특징1)을 복사하여 특징을 6개로 확대
- 3) 행동유형의 분별성이 양호한 베개 진동센서 데이터만 추가하여 특징 수를 4개로 확대
- 4) 베개 진동센서 데이터만 제공하여 특징으로 반영하고 특징 수는 그대로 3개로 구성

다음 [표 1]은 실험 결과를 제시한다. 이 실험에서는 특징집합의 변화에 따라 인식률에서 차이가 나타나지 않아,

인식률에서 경향성을 관찰하기는 어려웠다. 따라서, 기본 특징을 근간으로 특징 수만 변화시키는 것은 인식률에 별 영향을 주지 않는 것으로 분석된다. 다른 한편, “일어나기” 동작유형이 다른 “눕기”와 “뒤척임”의 인식률에 비해 잘 분류되는 결과를 보였다. 그 이유로는 베개 진동 센서 값을 거의 발생시키지 않는 “일어나기”는 베개 진동 센서 값을 뚜렷하게 발생시키는 “눕기”와 “뒤척임”에 비하여 인식이 더 잘 이루어진 것으로 분석된다.

[표 1] 특징집합의 변화에 따른 인식률 (단위: %)

특징 3개	눕기	뒤척임	일어나기
눕기	67.92	15.83	16.25
뒤척임	21.46	66.04	12.50
일어나기	5.42	5.42	89.17

특징 6개	눕기	뒤척임	일어나기
눕기	68.54	18.75	12.71
뒤척임	20.00	66.88	13.13
일어나기	8.13	8.54	83.33

특징 4개	눕기	뒤척임	일어나기
눕기	69.79	17.29	12.92
뒤척임	25.42	59.79	14.79
일어나기	3.75	8.96	87.29

2번째 특징 제공값	눕기	뒤척임	일어나기
눕기	66.25	17.71	16.04
뒤척임	23.54	62.71	13.75
일어나기	4.38	3.96	91.67

다음 [표 2]는 학습 데이터량에 따른 인식률을 측정 한 실험 결과를 제시한다. 수집된 5회 데이터 중에서 각각 2회, 3회, 4회 데이터에 대하여 학습시키고, 나머지 3회, 2회, 1회 데이터에 대하여 인식 실험을 수행하였다. 학습 데이터량이 늘어날수록 인식률이 향상되는 경향을 관찰할 수 있었다. 특히, 4회 데이터량을 학습 단계에서 사용한 실험에서는 상당한 인식률의 향상을 나타내었다. 앞으로 실험 데이터를 더 많이 수집하게 되면 인식률이 높아질 것으로 기대된다.

[표 2] 학습 데이터량에 따른 인식(특징 3개, 단위: %)

2회 학습	눕기	뒤척임	일어나기
눕기	67.92	15.83	16.25
뒤척임	21.46	66.04	12.50
일어나기	5.42	5.42	89.17

3회 학습	눅기	뒤척임	일어나기
눅기	65.94	16.56	17.50
뒤척임	15.63	67.19	17.19
일어나기	3.75	6.88	89.38

4회 학습	눅기	뒤척임	일어나기
눅기	69.38	11.88	18.75
뒤척임	16.88	72.50	10.63
일어나기	3.75	3.75	92.50

참고문헌

- [1] Azumio Inc., <http://www.azumio.com>
- [2] Squall Line Software, <http://www.squalllinesoftware.com>
- [3] Marko Borazio, Kristof Van Laerhoven. "Combining Wearable and Environmental Sensing into an Unobtrusive Tool for Long-Term Sleep Studies", IHI '12 Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, 2012.
- [4] 조 승호, 손 선동, 김 진태, 문 봉희, "개인별 행동 특징을 중심으로 한 수면과 연관된 일상행동 판단", 한국정보과학회 논문지 vol.16, no.12, pp.2104-1219, 한국, 2010.12.
- [5] Huins co., *Design and Application of Ubiquitous Sensor Network based on TinyOS 2.x*, 2008(in Korean).
- [6] 한 학용, *패턴인식 개론*, 한빛미디어, 2009.
- [7] Lawrence Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. of IEEE, Vol. 77 No. 2, 1989.2.
- [8] Daniel Jurafsky and James Martin, *Speech and Language Processing*, Prentice-Hall, 2009.

다음 [표 3]은 적외선 센서의 데이터는 행동유형에 공통되는 것으로 보고, 진동센서 2개만을 특징집합으로 사용하여 실험한 결과이다. 표 1과 같이 실험 데이터 5회중 2회는 학습 단계에서, 3회는 인식 단계에서 사용하여 실험하였다. 특징 3개를 사용한 실험에 비해서 눅기 동작유형의 인식률은 10% 향상되었으나, 일어나기 동작유형의 인식률은 10%가 낮아졌다. 이는 적외선 센서 특징이 행동유형 2가지에는 영향을 미치고 있으나, 뒤척임 행동유형 인식에는 영향이 없다고 분석된다.

[표 3] 특징 2개에 따른 인식률 (단위: %)

특징 2개	눅기	뒤척임	일어나기
눅기	77.50	14.38	8.13
뒤척임	24.79	60.83	14.38
일어나기	14.38	4.58	81.04

5. 결론

기존 연구에서는 이벤트 발생시, 센서값들의 평균과 표준편차를 구하여 행동 분류를 시도하였다. 이러한 방법은 입력 센서값이 기존에 설정된 기준 범위에 미달하거나 초과하는 경우, 인식 오류가 발생하였다. 본 연구에서는 센서값들의 평균과 표준편차를 계산하는 대신, 시간에 따른 센서값들의 변화를 은닉 마코프 모델로 인식하여 행동을 분류하는 알고리즘을 제시하였다. 또한, 본 논문에서는 센서값의 정도와 상관없이 피실험자 전체의 행동 데이터를 대상으로 분류하는 알고리즘을 제안하였다.

진동센서 2종과 적외선 센서 1종의 데이터를 특징벡터로 사용하여 은닉 마코프 모델에 적용하였고, 특징집합을 변화시켜 가면서 인식률의 변화를 관찰하였다. 특징의 개수와 인식률간 상관관계는 없는 것으로 분석되었다. 또한, 행동유형에 공통되는 특징인 적외선 센서는 행동유형 분류에 영향을 주고 있는 것으로 분석되었다. 그리고 학습 데이터 량이 증가할 수록 인식률이 향상되어 평균 78.127%로 향상되는 성과가 있었다. 향후 더 많은 데이터로 학습해 보는 것이 성능 향상에 긍정적인 것으로 기대된다.