은닉 조건부 랜덤 필드를 이용한 인간 행위 인식 시스템의 설계

김혜숙, 한유미, 김인철
경기대학교 컴퓨터과학과
e-mail : {chia, hym6936, kic}@kyonggi.ac.kr

Design of a Human Activity Recognition System using Hidden Conditional Random Fields

Hye-Suk Kim, Yu-Mi Han, In-Cheol Kim
Dept of Computer Science, Kyonggi University

요 약

본 논문에서는 키넥트 센서 데이터에 은닉 조건부 랜덤 필드 모델을 적용하여 인간의 일상 행위를 인식하는 시스템을 제안한다. 많은 고수준의 일상 행위들은 다수의 부속 행위들이 순차적 혹은 반복적으로 수행되어 나타나는 하나의 계층구조로 볼 수 있다. 따라서 제안하는 시스템에서는 이러한 고수준의 일상 행위들을 순차성과 계층성을 잘 표현할 수 있는 확률 그래프 모델의 하나인 은닉 조건부 랜덤 필드 모델로 모델링함으로써, 행위 인식률을 높이려고 시도하였다. 또한 제안하는 시스템에서는 효과적인 행위 모델의 학습과 적용을 위해, 모션 특징, 구조 특징, 손 위치 특징과 같은 다양한 종류의 특징들을 키넥트 센서 데이터로부터 추출하여 이들을 이용하였다. 그리고 12 가지 일상 행위들에 관한 코넬 대학의 CAD-60 데이터 집합을 이용한 다양한 실험을 통해, 제안하는 시스템의 우수한 인식 성능을 확인할 수 있었다.

1. 서론

카메라 영상을 이용해 인간의 일상 행위를 인식하는 기술은 지능형 로봇(intelligence robot), 스마트 헬스케어(smart healthcare) 그리고 스마트 홈(smart home)등에 폭넓게 이용될 수 있다[1-3]. 카메라 영상을 이용한 인간 행위 인식 기술은 신체에 별도의 센서를 부착해야 하는 불편함이 없기 때문에, 보다 자유로운 상태에서 인간의 일상 속 행위들을 인식해낼 수 있다는 장점이 있다. 그러나 카메라 영상은 빛의 반사나 조명 여건에 매우 큰 영향을 받으며, 인식 대상들이 카메라가 포착할 수 있는 공간 범위 안에 있어야 한다는 제약과 시야가 장애물에 가려지는 경우 행위 인식이 어려워진다는 문제점이 있다. 최근에 소개된 마이크로소프트(Microsoft)사의 키넥트(Kinect)는 RGB영상에 깊이(depth) 정보를 추가로 제공함으로써, 기존의 카메라 영상이 가진 단점을 완화시켜 주었다.

키넥트 센서를 이용한 기존의 행위 인식 연구들에서는 SVM(Support Vector Machine), HMM(Hidden Markov Model), MEMM(Maximum Entropy Markov Model) 등의 다양한 학습 모델을 적용한 행위 인식 시스템을 제안하였다. 하지만, SVM 모델[4]의 경우는 행위와 부속 행위(sub-activity)들 간의 계층 구조도, 시간에 따른 부속 행위들 간의 순차 구조도 표현하지 못한다는 한계점이 있다. HMM 모델[5]은 생성 그래프 모델(generative graphical model)의 하나로서 행위의 계층성과 순차성을 잘 표현할 수 있으나, 부속 행위를 나타내는 은닉 상태와 관측들 간의 결합 확률(joint probability)들을 학습해야 하는 부담을 가지고 있다. 한편, 판별 그래프 모델(discriminative graphical model)의 하나인 MEMM 모델[3]의 경우는 이러한 HMM 모델의 한계성을 벗어나 행위의 계층성과 순차성을 잘 표현할 수 있으나, 대신 레이블 편향 문제(label bias problem)을 가지고 있는 것으로 잘 알려져 있다.

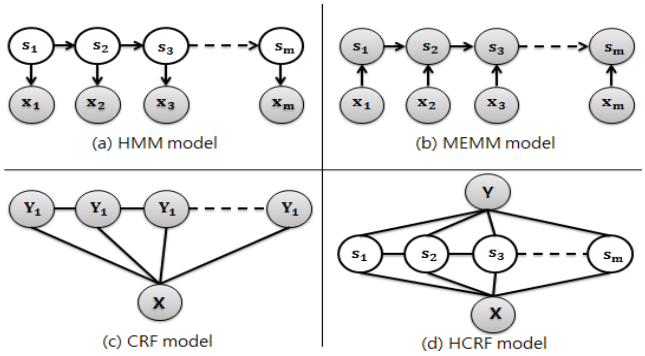
본 논문에서는 키넥트 센서 데이터에 은닉 조건부 랜덤 필드(Hidden Conditional Random Field, HCRF)[6]모델을 적용함으로써 인간의 일상 행위를 효과적으로 인식하는 시스템을 제안한다. 많은 고수준의 일상 행위들은 다수의 부속 행위들이 순차적 혹은 반복적으로 수행되어 나타나는 하나의 계층구조로 볼 수 있다. 따라서 제안하는 시스템에서는 이러한 고수준의 일상 행위들을 순차성과 계층성을 잘 표현할 수 있는 확률 그래프 모델의 하나인 은닉 조건부 랜덤 필드 모델로 모델링함으로써, 행위 인식률을 높이려고 시도하였다. 또한 제안하는 시스템에서는 효과적인 행위 모델의 학습과 적용을 위해, 모션 특징(motion feature), 구조 특징(structure feature), 손 위치 특징(hand position feature)과 같은 다양한 특징들을 키넥트 센서 데이터로부터 추출하고 이들을 이용한다. 본 논문에서 제안하는 시스템의 성능을 분석하기 위해, 코넬 대학의 CAD-60 데이터 집합을 이용한 다양한 실험을 수행하고 그 결과를 소개한다.

2. HCRF 모델

카메라 영상을 이용한 인간의 행위 인식에 많이 사용되는 확률 그래프 모델에는 은닉 마코브 모델(Hidden Markov Model, HMM), 최대 엔트로피 마코브 모델(Maximum Entropy Markov Model, MEMM), 조건부 랜덤 필드(Conditional Random Field, CRF) 그리고 은닉 조건부 랜덤 필드(Hidden Conditional Random Field, HCRF) 등이 있다. (그림 1)의 (a)는 HMM 모델, (b)는 MEMM모델, (c)는 CRF 모델 그리고 (d)는 HCRF 모델을 각각 나타낸다. 그림에서 변수 X는 관측치(observation), 변수 Y는 클래스 레이블(class label), 변수 s는 내부 상태(internal

※ 본 연구는 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음

state)를 나타내며, 각 그림에서 빈 원으로 표시된 노드는 은닉 변수(hidden variable), 속이 채워진 노드는 관측 변수(observation variable)을 나타낸다.



(그림 1) 확률 그래프 모델

(그림 1) (a)의 HMM 모델은 확률을 이용한 유향 그래프 모델(directional graphical model)로써, 모델링하려는 대상 시스템이 은닉 상태를 가진 마코브 프로세스(Markov Process)라고 가정한다. 그리고 은닉 상태를 통해 관측 값을 생성할 수 있다. HMM은 주어진 관측 값을 통해 결합 확률이 최대가 되는 방향으로 초기 확률(prior probability), 상태 전이 확률(state transition probability), 관측 확률(observation probability)을 학습한다. 하지만 예측 단계에서는 결합 확률이 아닌 조건부 확률을 사용하므로 학습 모델과 예측 모델이 불일치한다. 또한 관측 값이 현재 상태에 종속되므로 장거리 종속성을 표현할 수 없다. 따라서 관측에 현재 상태만 영향을 줄 수 있으며, 관측 값이 중복되는 경우 모델의 학습이 어렵거나 불가능하다는 문제점이 있다[4].

(그림 1) (b)의 MEMM 모델은 HMM 모델과 최대 엔트로피(Maximum Entropy)모델의 특징을 결합한 확률 유향 그래프 모델이다. MEMM 모델은 조건부 확률을 이용해 모델을 학습하고 예측한다. 또한 MEMM 모델은 상태가 관측에 종속적이므로 과거와 미래의 관측이 현재 상태에 영향을 줄 수 있다고 가정한다. 그러므로 HMM 모델과 다르게 MEMM 모델은 장거리 종속성을 표현할 수 있다. 그러나 MEMM 모델은 다음 상태를 결정하기 위한 조건부 확률을 계산할 때 상태 별로 지수 모델(exponential model)을 적용하므로 레이블 편향 문제가 발생한다.

(그림 1) (c)의 CRF 모델은 조건부 확률을 이용하여 최적의 클래스 레이블을 계산하는 확률 무향 그래프 모델이다. 상태와 관측 사이에 방향이 없으므로 상태별로 확률을 계산하지 않고 잠재성 함수(potential function) 또는 특징 함수(feature function)를 이용하여 모델을 학습한다. 또한 상태 별로 지수 분포를 적용하여 조건부 확률을 계산하는 MEMM 모델과 다르게 주어진 전체 관측에 한 개의 지수 분포를 적용하여 전체 클래스 레이블을 결정하므로 레이블 편향 문제가 발생하지 않는다. 그러나 CRF는 모델을 학습할 때 전체 시계열 데이터(sequence data)의 관측 값마다 레이블을 지정해야한다. 그러므로 관측 값의 레이블을 미리 알고 있지 않은 경우 관측 값 마다 레이블을 지정하기 위해 추가적인 학습을 필요로 한다.

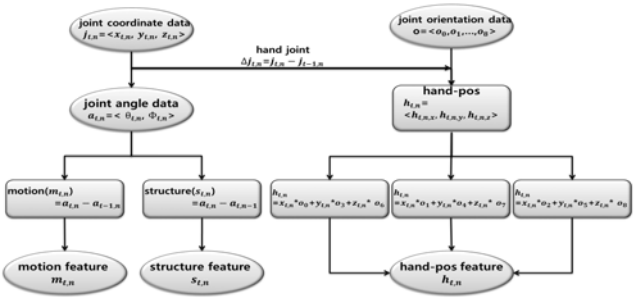
(그림 1) (d)의 HCRF 모델은 CRF 모델에 은닉 상태 층을 추가한 확률 무향 그래프 모델이다. HCRF는 CRF에 은닉 상태 층을 추가함으로써, 고수준의 복잡한 행위를 인식하기 위해 행위 내 부속 행위의 순서를 지정할 필요가 없다. 또한 HCRF 모델은 CRF 모델에서 발전되었으므로 HMM과 MEMM 모델이 가지는 문제점이 발생하지 않는다. HCRF는 일대다(1-vs-all) 구조와 다중 클래스(multi-class) 구조를 가진다. 일대다 구조는 행위 별로 한 개의 HCRF 모델이 생성된다. 그리고 생성된 행위 별 모델에 관측 값을 주어 행위의 레이블마다 조건부 확률을 계산하고, 가장 큰 확률 값을 가지는 행위의 모델을 관측 값의 행위로 결정한다. 다중 클래스 구조는 여러 행위의 관측 값을 이용해 한 개의 HCRF 모델을 학습한다. 그리고 테스트 시계열 데이터와 학습된 모델을 이용해 가장 큰 확률 값을 가지는 행위를 선택한다.

3. 행위 인식 시스템 설계

3.1 시스템 개요

본 논문에서는 키넥트 센서 데이터로부터 입 헹구기(rinsing mouth), 양치하기(brushing teeth), 물 마시기(drinking water), 컴퓨터 작업하기(working on computer), 젓기(stirring), 썰기(chopping), 의자에 앉아서 휴식하기(relaxing on couch), 약병 열기(opening pill container), 렌즈 끼기(wearing contact lens), 전화하기(talking on phone), 화이트보드에 적기(writing on whiteboard), 의자에 앉아 대화하기(talking on couch) 등 총 12가지 서로 다른 인간의 일상 행위를 인식하는 시스템을 제안한다. 제안하는 행위 인식 시스템은 크게 특징 추출 단계, HCRF 행위 모델 학습 단계 그리고 행위 인식 단계로 구성된다. 특징 추출 단계에서는 관절 좌표 데이터를 바탕으로 시간에 따른 관절의 변화, 관절 사이의 변화 그리고 양손의 위치에 대한 특징 데이터를 추출한다. 그리고 특징에 이산화 과정을 적용하여 훈련 및 인식 시계열 데이터를 생성한다. 모델 학습 단계에서는 훈련 시계열 데이터와 클래스 레이블을 바탕으로 HCRF 모델을 학습한다. 마지막으로 인식 단계에서는 새로운 테스트 시계열 데이터에 인식 시계열 데이터 및 학습된 HCRF 모델을 적용하여 행위를 인식한다.

3.2 특징 추출



(그림 2) 특징 추출 세부 단계

특징 추출 단계에서는 사람의 신체 관절들 중 머리, 목, 몸통 등을 포함한 15개 관절의 좌표 데이터와 머리와 몸통에 해당하는 2개의 관절의 방향 데이터를 바탕으로 특징을 추출한다. (그림 2)는 제안하는 행위 인식 시스템의 특징을 추출하는 세부 과정을 나타낸다.

$$\theta = \cos^{-1}\left(\frac{z}{r}\right) \qquad (1)$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \qquad (2)$$

관절 좌표 데이터는 카메라와 대상과의 방향 및 거리에 민감하게 변하므로 관절별 좌표 데이터를 극각(polar angle)과 방위각(azimuthal angle)으로 이루어진 관절별 각도 데이터로 변환한다. 식 (1)과 (2)에서 r은 원점으로부터 관절 좌표까지의 거리를 나타내며, θ는 벡터의 극각, ϕ는 벡터의 방위각에 해당한다.

$$\text{motion}(m_{t,n}) = a_{t,n} - a_{t-1,n} \qquad (3)$$

$$\text{structure}(s_{t,n}) = a_{t,n} - a_{t,k} \qquad (4)$$

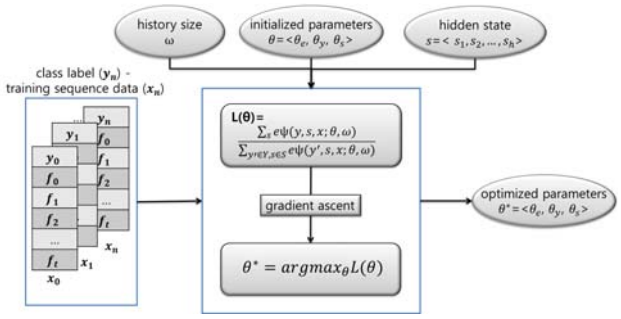
$$\text{hand-pos}(h_{t,n}) = x_{t,n}{}^{*}o_{l} + y_{t,n}{}^{*}o_{l+3} + z_{t,n}{}^{*}o_{l+6} \quad (5)$$

특징 집합은 모션 특징, 구조 특징 그리고 손 위치 특징으로 구성된다. 식 (3)은 (그림 2)의 모션 특징, 식 (4)는 (그림 2)의 구조 특징, 식 (5)는 (그림 2)의 손 위치 특징에 해당한다. 식(3), (4), (5)에서 변수 t는 프레임의 인덱스(index)에 해당하며 변수 n과 k는 관절의 번호, 변수 l은 관절 방향 데이터의 인덱스에 해당한다. 모션 특징은 관절의 시간적 변화, 구조 특징은 관절의 공간적 변화를 나타내기 위한 특징이다. 또한 인식할 행위가 손의 위치에 영향을 받는다고 가정하여 손의 위치를 나타내기 위한 손 위치 특징을 선정하였다.

일반적으로 특징 벡터의 차원이 증가하면 모델 학습과 인식을 위한 계산량이 증가한다. 본 논문에서는 사용하는 특징은 총 252차원이므로 모델 학습과 인식의 효율성을 높이기 위해, 고차원 특징벡터에 k-평균 군집화(k-means clustering)을 적용해 차원 축소(dimension reduction)와 이산화(discretization)를 수행한다. 그 결과 1차원 정수 형태의 시계열 데이터가 생성된다. 이 때, 시계열 데이터의 길이는 행위별 수행 시간에 따라 결정되므로 행위별로 다른 길이로 생성된다. 마지막으로 이산화 된 시계열 데이터는 각각 학습 시계열 데이터(training sequence data)와 테스트 시계열 데이터(test sequence data)로 나누어진다.

3.3 모델 학습

HCRF 모델을 적용하면 복잡한 일상의 행위를 여러 개의 부속 행위로 나눌 수 있다. HCRF 모델은 은닉 상태를 이용하여 부속 행위 별 클래스 레이블을 표시하기 때문에 클래스 레이블을 결정하기 위한 추가적인 학습이 필요하지 않는다. 행위의 클래스 레이블 데이터는 시계열 데이터 당 한 개의 클래스 레이블 값을 가지므로 학습 시계열 데이터의 개수와 동일한 길이로 생성된다. 본 논문에서는 HCRF 모델의 두 가지 구조 중 한 개의 학습된 모델로 여러 행위를 인식할 수 있는 다중 클래스 구조를 통해 HCRF 모델을 학습한다.



(그림 3) HCRF 모델 학습

(그림 3)은 HCRF 모델 학습과정을 나타낸다. 그림에서 Θ는 HCRF 모델의 매개변수 벡터(parameter vector)를 나타내며, 구성요소 중 θ_e는 은닉 상태들에 관련된 매개변수를, θ_y는 클래스 레이블 y와 은닉 상태들에 관련된 매개변수를, θ_e는 클래스 레이블 y와 연결된 두 개의 은닉 상태 쌍에 관련된 매개변수를 각각 나타낸다. 또한, Θ와 ω를 매개변수로 갖는 잠재성 함수 Ψ(y, s, x; Θ, ω)는 클래스 레이블 y와 은닉 상태들 s, 그리고 관측 시계열 데이터들 x, 이들 간의 상호 적합도(compatibility)를 나타낸다. HCRF 모델 학습을 위한 각 훈련 데이터는 관측 시계열 데이터 x와 이에 대응하는 행위 레이블 y의 쌍으로 구성되며, 관측 시계열 데이터 x는 특징 벡터 f_k들의 시퀀스로 구성된다. 모델 학습과정에서는 초기화된 HCRF 모델 매개변수들 Θ과 훈련 데이터 집합 {(x_i, y_i)| i=1, .., n}을 이용해, 목적 함수 L(Θ)가 최대가 될 수 있는 최적의 매개변수 값 Θ*를 경사 상승법(gradient ascent)으로 찾아낸다.

HCRF 모델을 학습하기 위해서는 은닉 상태의 수, 히스토리 사이즈(history size)를 미리 결정해야 한다. 우선 훈련 시계열 데이터를 바탕으로 은닉 행위 별로 은닉 상태의 개수를 결정하였다. 그리고 행위 별 은닉 상태의 개수 중 최대값인 7을 HCRF 모델을 학습하기 위한 은닉 상태로 결정하였다. 히스토리 사이즈는 현재 상태에 직전 상태가 영향을 줄 수 있다고 가정하고 1로 설정하였다. 또한 HCRF 모델의 특징의 가중치를 학습하기 위한 최적화 함수는 LBFGS(Limited-memory Broyden Fletcher Goldfarb Shanno)를 사용하였고, 마지막으로 HCRF 모델을 학습하기 위해 학습 시계열 데이터 집합에 300번의 반복 횟수를 설정하였다. 그 결과 12가지 행위를 인식 할 수 있는 한 개의 HCRF 모델이 학습된다.

3.4 행위 인식

행위 인식 단계에서는 학습된 1개의 HCRF 모델과 테스트 시계열 데이터를 이용하여 12가지 행위에 대한 조건부 확률을 계산한다. 식 (6)은 학습된 모델 θ^*과 테스트 시계열 데이터 x 그리고 히스토리 사이즈 ω를 이용해 행위에 대한 클래스 레이블 y를 결정하는 식이다. 세 개의 매개변수 x, ω, θ^*가 주어졌을 때 y가 나타날 조건부 확률을 계산한다. 그리고 조건부 확률이 가장 큰 y가 인식된 행위에 해당한다. 행위 인식 단계의 결과는 각 테스트 시퀀스에 해당하는 레이블이 의미하는 행위이다.

$$\text{argmax}_{y \in Y} P(y|x, \omega, \theta^*) \quad (6)$$

4. 실험 및 평가

본 논문에서 제안하는 행위 인식 방법의 성능을 평가하기 위한 실험을 전개하였다. 실험에 사용된 행위 인식 시스템은 Window 7에서 C++과 매트랩(Matlab)을 사용하여 구현하였다. 그리고 실험에 사용된 데이터 집합은 키넥트 센서를 사용하여 수집된 코넬 대학의 CAD-60 데이터 집합[3]이다.

(그림 4) 인식 대상 행위들

코넬 대학의 CAD-60 데이터 집합은 사무실, 부엌, 침실, 화장실, 거실의 서로 다른 다섯 가지 환경에서 초당 30프레임으로 약 45초 동안 수집되었다. 수집된 데이터는 일상생활에서 나타나는 12가지 종류에 대한 관절 좌표 및 방향 데이터, RGB 이미지, 깊이 영상으로 구성되어 있다. (그림 4)는 인식하고자 하는 12가지 행위에 대한 장면이다. (그림 4)의 순서대로 양치하기, 젓기, 화이트보드에 적기, 컴퓨터 작업하기, 전화하기, 렌즈 끼기, 의자에 앉아서 휴식하기, 약병 열기, 물마시기, 썰기, 의자에 앉아 대화하기, 입 헹구기에 해당한다.

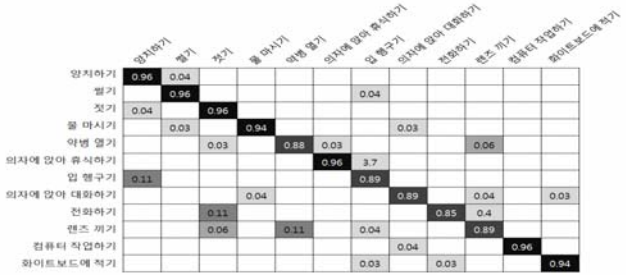
본 논문에서는 코넬 대학의 CAD-60 데이터 집합을 이용하여 행위 인식 시스템의 성능을 분석하기 위한 세 가지 실험을 수행하였다. 첫 번째 실험은 HCRF 모델의 두 가지 구조의 인식 성능을 비교하기 위한 목적으로 수행하였다. HCRF 모델의 두 가지

구조에는 일대다 구조와 다중클래스 구조가 있다. 일대다 구조는 훈련 데이터에 인식하고자 하는 행위의 데이터만 포함되어 있으며 다중클래스는 모든 행위 데이터를 이용하여 훈련 데이터를 생성하였다. 또한 두 가지 구조에 장거리 종속성이 미치는 영향을 확인하기 위해 히스토리 사이즈를 0과 1로 설정하여 실험을 진행하였다. <표 1>은 HCRF 모델의 두 가지 구조로 행위 인식 성능을 실험한 결과이다. <표 1>에서 볼 수 있듯이 일대다 구조보다 다중클래스 구조인 HCRF 모델이 평균 인식률이 더 높음을 볼 수 있다. 또한 히스토리 사이즈를 0으로 했을 때보다 1로 했을 때의 평균인식률이 더 높음을 확인할 수 있다.

<표 1> HCRF 모델의 구조 별 행위 인식률

모델 구조	평균 인식률(%)
HCRF(일대다) ω = 0	86.11
HCRF(일대다) ω = 1	88.27
HCRF(다중클래스) ω = 0	90.23
HCRF(다중클래스) ω = 1	92.18

두 번째 실험은 HCRF 학습 모델의 행위 별 인식률을 분석하기 위한 목적으로 수행하였다. 실험을 수행하기 위해 HCRF 모델의 다중클래스 구조를 사용하였으며 히스토리 사이즈는 1로 설정하였다. 다중 클래스 구조를 사용하였으므로 훈련 데이터에 12가지 행위의 데이터가 모두 포함되어 있다. (그림 5)는 코넬 대학의 CAD-60 데이터 집합에 대한 행위 별 오차 행렬(confusion matrix)을 나타낸다. (그림 5)에서 약병 열기, 입 헹구기, 의자에 앉아 대화하기, 전화하기, 렌즈 끼기 행위의 경우 다른 행위에 비해 인식 결과가 낮다는 것을 확인할 수 있다. 예를 들어 입 헹구기의 경우 양치하기로 오분류 되는 경우가 많았다. 이는 양치하는 행위 도중에 입을 헹구는 행위가 포함되어 있기 때문에 발생한 결과로 추정된다. 그리고 약병 열기나 렌즈 끼기의 경우 다른 행위에 비해 행위를 수행하는데 걸리는 시간이 짧기 때문에 정보가 충분하지 않아 발생한 결과로 추정된다.



(그림 5) 행위 별 오차 행렬

세 번째 실험은 HCRF 모델과 다른 확률 그래프 모델 중 HMM 모델과 CRF 모델의 평균 인식률을 비교하기 위한 목적으로 수행하였다. HMM은 방향 그래프 모델이며, CRF와 HCRF는 무향 그래프 모델이므로 실험을 통해 유향 그래프 모델과 무향 그래프 모델의 행위 인식 성능을 비교할 수 있다. 또한 CRF 모델은 계층을 표현하지 못하므로 은닉 상태를 이용한 계층화가 행위 인식 성능에 미치는 영향을 확인할 수 있다. 실험을 위해 HMM 모델은 행위 별로 은닉 상태를 설정하여 주었다. 그리고 CRF 모델은 HCRF 모델과 다르게 프레임 별로 클래스 레이블을 지정하여 레이블 데이터를 생성하였다. 또한 HMM 모델은 장거리 종속성을 표현할 수 없으므로 히스토리 사이즈가 0이며, CRF와 HCRF 모델의 경우 히스토리 사이즈가 0과 1인 경우에 대하여 실험을 진행하였다.

<표 2>는 HMM 모델, CRF 모델, HCRF 모델의 행위 인식 결과를 나타낸다. <표 2>에서 CRF가 히스토리 사이즈 0일 때, 평균 86.88%로 가장 낮은 인식률을 보였다. 이를 통해 행위를 계층화 시켜 학습한 모델이 더 좋다는 것을 확인할 수 있었다. 그리고 HMM 모델과 다중클래스 구조의 HCRF의 인식률을 비교한 결과 히스토리 사이즈를 1로 설정했을 때 평균 인식률이 높았다. 이는 HMM 모델이 장거리 종속성을 표현할 수 없는 반면 HCRF 모델은 장거리 종속성을 이용하여 모델을 학습할 수 있기 때문이라고 추정된다.

<표 2> 학습 알고리즘 별 행위 인식률

학습 모델	평균 인식률(%)
HMM ω = 0	90.84
CRF ω = 0	86,88
CRF ω = 1	87.96
HCRF(다중클래스) ω = 0	90.23
HCRF(다중클래스) ω = 1	92,18

5. 결론

본 논문에서는 키넥트 센서 데이터에 다중클래스 HCRF 모델을 적용한 일상 행위 인식 시스템을 제안하였다. 제안하는 시스템에서는 고수준의 일상 행위들이 가지는 계층 구조와 순차 구조를 다중클래스 HCRF 모델로 모델링함으로써, 행위 인식률을 높이려고 시도하였다. 또한 제안하는 시스템에서는 효과적인 행위 모델의 학습과 적용을 위해, 모션 특징, 구조 특징, 손 위치 특징과 같은 다양한 종류의 특징들을 키넥트 센서 데이터로부터 추출하여 이들을 이용하였다. 그리고 12 가지 일상 행위들에 관한 코넬 대학의 CAD-60 데이터 집합을 이용한 실험을 통해, 제안하는 시스템의 우수한 인식 성능을 확인할 수 있었다.

참고문헌

[1] A. Jalal, Md. Z. Uddin and T. S. Kim, “Depth Video-based Human Activity Recognition System Using Translation and Scaling Invariant Features for Life Logging at Smart Home”, IEEE Transaction on Consumer Electronics, vol. 58, no. 3, pp. 863-871, 2012

[2] H. S. Koppula, R. Gupta and A. Saxena, “Learning Human Activities and Object Affordances from RGB-D Videos”, International Journal of Robotics Research, vol. 32, no. 8, pp. 951-970, 2013

[3] J. Sung, C. Ponce, B. Selman and A. Saxena, “Human Activity Detection from RGBD Images”, Proc. of AAAI Workshop on Pattern, Activity and Intent Recognition, 2011

[4] C. Zhang and Y. Tian, “RGB-D Camera-based Daily Living Activity Recognition”, Journal of Computer Vision and Image Processing, vol. 2, no. 4, 2012

[5] Md. Z. Uddin, N. D. Thang, J. T. Kim and T. S. Kim, “Human Activity Recognition Using Body Joint-Angle Features and Hidden Markov Model”, Journal of ETRI, vol. 33, no. 4, pp. 569-579, 2011

[6] S. B. Wang, A. Quattoni, L.P. Morency, D. Demirdjian, T. Darrell, “Hidden Conditional Random Field for Gesture Recognition”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1521-1527, 2006