

# 에이전트 행동에 기반한 의도 인식 컴퓨팅

## Agent's Activities based Intention Recognition Computing

김진옥\*  
Kim Jin Ok

### 요약

에이전트의 의도를 인식하는 것은 사물지능형 컴퓨팅에서 인간컴퓨터 상호작용의 주요 부분이다. 컴퓨팅 시스템에서 인식 대상의 의도를 정확하게 유추하면 다수의 에이전트간의 협력 상황 이해와 특정 행동이 취해질 때의 상황 파악이 쉽기 때문이다. 본 연구는 다른 이의 행동을 해석하고 행동의 근거가 되는 의도와 목적을 추론하는 인간의 기제를 바탕으로, 컴퓨팅 시스템이 행동을 인식하여 습득한 사전 경험 데이터를 이용, 대상의 의도를 빠르게 인식하는 방법을 제안한다. 의도 인식을 수행하기 위해 제안 방법은 에이전트의 목적에 따른 행동 변화를 검출하고 시스템이 사전에 학습한 행동 정보를 모델링하기 위해 특정 형태의 행동 은닉마코프 형식을 이용한다. 에이전트의 의도를 추론하는 데 관점을 다양하게 취함으로써 시스템이 에이전트의 행동이 끝나기 전에 미리 의도를 추론하도록 한다. 의도 인식의 정확도, 조기 검출률과 정확 지속률에 대한 실험으로 여러 가지 행동을 취하는 에이전트의 의도 검출 결과를 정량적으로 제시함으로써 제안 연구가 효과적인 의도 인식 시스템 구현에 기여함을 보여준다.

### ABSTRACT

Understanding agent's intent is an essential component of the human-computer interaction of ubiquitous computing. Because correct inference of subject's intention in ubiquitous computing system helps particularly to understand situations that involve collaboration among multiple agents or detection of situations that can pose a particular activity. This paper, inspired by people have a mechanism for interpreting one another's actions and for inferring the intentions and goals that underlie action, proposes an approach that allows a computing system to quickly recognize the intent of agents based on experience data acquired through prior capabilities of activities recognition. To proceed intention recognition, proposed method uses formulations of Hidden Markov Models (HMM) to model a system's prior experience and agents' action change, then makes for system infer intents in advance before the agent's actions are finalized while taking the perspective of the agent whose intent should be recognized. Quantitative validation of experimental results, while presenting an accurate rate, an early detection rate and a correct duration rate with detecting the intent of several people performing various activities, shows that proposed research contributes to implement effective intent recognition system.

☞ keyword : Intent Inference(의도 추론), Activity Recognition(행동 인식), Human Computer Interaction(인간컴퓨터 상호작용), Hidden Markov Model(은닉마코프모델)

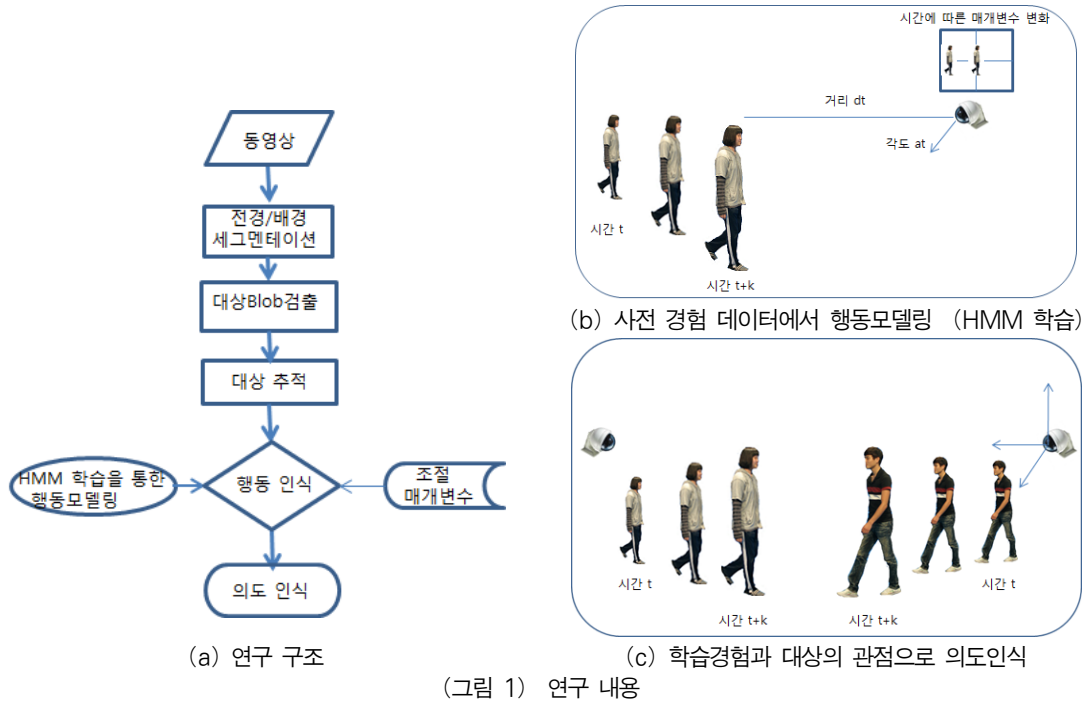
## 1. 서론

다른 사람의 의도를 인식하는 것은 휴먼 인지의 중요한 부분으로 특히 대상의 행동에서 다른 사람과 협업하려는지 또는 경쟁하려는 지에 대한 내재된 의도를 파악하는 것은 인지 과정에서 중요한 추론 문제로 여겨지고 있다[1,2]. 컴퓨팅 발전 단계에서 컴퓨팅 시스템은 사람의 인지 기능을 모델링하여 사람과 의사소통하는 방법으로 구현하고 있으며 그 결과는 사람과 컴퓨터간의 원활한 상호작용을 연구하는 HCI, 융합인지과학 분야에서 다

양하게 접목되고 있다. 특히 사물지능형 컴퓨팅에서 시스템이 인간 중심의 환경과 효과적인 협력의 일환으로 컴퓨터가 사람과 커뮤니케이션하고 사람의 의도를 추론하는 기술을 지속적으로 연구하면서 사용자에게 자연스러운 환경서비스를 제공하고자 한다.

컴퓨팅의 의도 추론 및 인식 방식은 사람의 인지과정을 모방한다. 사람은 특정 행동을 취할 때 끊임없이 외부 정보를 감지하여 행동과정에 그 내용을 반영하며, 이 경험 바탕으로 관찰 대상으로 삼은 사람의 시각에서 그 사람이 외부와 상호작용하는 것을 관측하여 의도를 추론한다. 특히 사람은 감지한 관측치가 과거 경험과 맞아 떨어질 때 해당 관측치를 통해 같은 상황에서 무슨 의도가 있었는지를 적절하게 파악하는 회상기억작용을 이용한다. 이러한 과정은 무감독 학습 중심의 의사결정 모델로 정

\* 종신회원 : 대구한의대학교 국제문화정보대학  
모바일콘텐츠학부 교수 bit@du.ac.kr  
[2011/10/25 투고 - 2011/11/04 심사 - 2012/01/13 심사완료]



리되어 기계학습에 새로운 학습 이론과 모델이 시도되면서 은닉마코프 모델(HMM, Hidden Markov Model)과 같은 확률그래프 모델로 대표되는 생성적 학습모델이 대두되고 있다[3].

본 연구는 생성적 학습 모델을 바탕으로 지능형 컴퓨팅 시스템에 필요한 대상의 의도 인식을 대상의 행동으로 파악하는 시스템을 제안한다. 연구 수행을 위해 사람이 다른 사람의 행동을 지각하고 예측하며 해석하는 인지 과학적 근거[4]와 더불어 사람은 상대의 시각으로 그 사람이 특정 행동을 취하는 목적과 의도를 추론할 수 있다는 점[5,6]을 반영하여 의도 인식과정을 설정한다. 기계학습을 위해서는 생성적 학습 모델의 대표적 방법인 HMM 형식을 이용하여 행동 학습을 수행하고 학습경험을 통해 행동에 반영된 의도 변화과정을 처리한다.

제안 연구 시스템은 (그림 1)과 같이 기계학습을 위한 행동 모델링 단계와 학습경험과 대상의 관점으로 상대의 의도를 인식하는 의도 인식 과정 두 단계로 구성한다. 첫 번째 단계에서는 특정 행동을 수행하는 경험 데이터를 통해 시스템이 차후에 인식해야 하는 해당 행동에 대한 HMM 형식을 학습한다. 시스템은 학습을 위해 ‘만나기’와 같은 상황에서 두 에이전트가 가까워지면 한 사람이

가는 방향의 거리와 머리간의 각도와 같은 파라미터가 변하는 상태를 관찰한다. 두 번째 단계에서는 관찰자가 된 시스템이 학습한 HMM을 통해, 관측한 에이전트의 시점에서 동일 목적의 파라미터 변화를 측정하여 다른 에이전트의 행동 변화를 모니터한다.

본 논문은 2장 관련 연구, 3장 제안 의도 인식 컴퓨팅 시스템 구조, 4장 대상 검출과 추적, 5장 실험 그리고 6장 결론으로 구성되어 있다.

## 2. 관련 연구

의도는 대상이 품고 있는 사고와 행동의 목적을 의미하며 목적을 달성하기 위해 의도를 연속적이고 상세한 행동 수준으로 확대하면 의도의 개념은 ‘계획’과 동의어가 된다. 의도 추론에 대한 기존 연구[7,8]에서는 자연어와 같은 명시적 정보를 이용한 의도인식 방법을 제시하여 왔으나 사물지능형 시스템은 전적으로 감지한 데이터의 암묵적 단서에 의존하기 때문에 이와 다른 기제를 필요로 한다.

암묵적 단서인 사용자 행동을 통한 의도 인식 방법은 로봇 분야에서 대표적으로 제시되어 왔다. 대상에 대한

다양한 데이터를 수집하기 위해 물리적 센서를 이용하는 로봇은 수집한 센서 데이터에서 사용자의 콘텐츠 정보를 추출하기 위해 다양한 기계학습 방법을 사용한다[9]. 의도 인식을 위해 카메라와 마이크를 이용한 멀티모달 센서를 통합하여 주변 상황으로부터 컨텍스트 정보를 수집한 후 통계학적 확률밀도함수 등의 학습 방법을 통해 사용자 상황데이터에서 사용자의 상태를 통계적으로 결정하기도 한다[10]. 의도 그래프를 이용하여 사용자의 의도를 데이터 구조적으로 설명하는 방법도 제시되고 있다[11].

최근 연구에서는 사용자 행동을 파악하는 여러 가지 센서를 이용하여 물리적 측정 결과로부터 심리적 상태를 추론하려는 인지 과학적 노력이 계속되고 있다[12]. 사용자의 행동을 인식하고 그 결과로부터 특정 의도를 포착하려는 연구에서 가장 다양하게 적용되는 절차 모델링 도구 및 기계학습 방법이 HMM이다. 시간적 시퀀스와 관련한 절차를 모델링하는 강력한 도구인 HMM은 의도 인식과 행동 인식 연구에 다양하게 적용하여 행동 모델링과 의도 추론 과정을 처리한다[13-15].

하지만 행동 인식에 관한 기존 연구에서 제안된 시스템들은 행동을 인식하기 전에 개별 동작을 전체 시퀀스로 설정 완료해야 함을 전제[16,17]로 하는데 이는 행동의 사전 설정은 여러 상황에 따라 행동이 다양하게 변할 수 있다는 점을 간과하고 있다. 또한 잠재적 위험 상황을 검출하거나 긴밀하게 협력해야 하는 사물지능형 시스템 시나리오에서 실제로 행동이 완료되기 전에 미리 그 의도를 검출하는 것이 중요함에도 사람의 의도를 학습하는데 HMM을 적용한 이전의 연구에서는 대부분 목적 발생을 관측한 이후에만 사용자의 의도가 무엇인지 파악하는 한계성을 보이고 있다[18,19]. 따라서 행동이 완료되기 전에 미리 대상의 의도를 파악하여 빠르게 대응하는 연구가 수행되어야 하며 시간에 따라 변하는 긴 태스크의 행동 시퀀스를 모델링하는 것보다는 행동 간의 전이 전에 은닉 상태에서 행동의 목적이 의도에 따라 어떻게 변하는지를 제시하는 방법론이 필요하다.

제안 연구와 유사한 행동 인식 연구[20]와 HMM 응용 행동 인식 연구[21]는 템플릿을 설정하거나 비정상적 행동을 감지하는 방법을 주로 다뤘으나 이 방법들은 관측한 행동과 미리 존재하는 행동 모델 집합간의 불일치성을 검출함으로써 기대한 행동 패턴에서 예외적인 것을 검출하는 데는 유용하지만 관측한 행동의 의도에 관한 정보를 제공하지 않는다는 취약점이 있다.

이와 같은 기존 연구에서 도출된 문제를 해결하기 위해 새로운 형태의 행동 HMM 형식을 설정하고 행동을

학습하여 경험 데이터를 축적한 다음 정확히 에이전트의 의도를 파악하고 에이전트가 행동을 마치기 전에 조기에 의도를 인식하는 방법이 필요하다. 본 연구는 낮은 수준의 목적을 가진 행동별 의도를 추론하는 학습단계를 통해 경험 데이터를 축적하고 이 데이터를 의도 인식에 반영하여 사람이 행동을 마치기 전에 의도 유추 목적을 달성하도록 한다. 또한 특정 훈련 환경에서만 적용하는 HMM의 전형적 접근방식과는 달리 제안 행동 모델은 여러 상황에 일반적이며 또한 다른 도메인으로 전달되어 다른 환경에서 훈련했더라도 원래 도메인에 관계없이 필요한 도메인에서 행동 특징을 설정하도록 한다.

### 3. 의도 인식 컴퓨팅 시스템

#### 3.1 행동 분석을 위한 HMM 형식

HMM은 일시성이 내재되어 있으며 시간이 확대되는 변화 과정을 표현하기 위한 문제에서 가장 많이 사용되는 기계 학습 방법이다. 행동을 통한 HMM 기반 의도 인식 시스템은 이산 상태  $\{s_i\}$  집합  $N$ 으로 표현한다. 각 시간 단계에서 시스템은 이산 상태 중 한 상태에 놓이며  $P(s_j(t+1)|s_i(t)) = a_{ij}$  와 같은 확률에 따라 다른 상태로 전이된다. 시스템이  $t$ 시간에 상태  $s_i$ 에 놓였다고 가정할 때  $a_{ij}$ 는  $t+1$ 시간에 상태  $s_j$ 에 있을 확률을 의미한다. 하지만  $t$ 시간에 시스템의 상태는 직접 관측할 수 없고 대신 은닉상태에 종속적인 변수(상태)집합  $\{v_i\}$ 만이 관측 가능하며 각 상태  $s_j$ 에 대해서는 특정하게 보이는 상태  $\{v_k\}$ 를 관측할 확률  $P(v_k(t)|s_j(t)) = b_{jk}$ 이 주어진다. 이와 같이 전통적 HMM 학습 방법은 모델의 구조가 주어지는 방식으로, 관측 가능한 훈련데이터 집합과 함께 은닉상태 수, 드러나는 상태 수, 상태간의 전이 토폴로지가 주어지며 이 구조를 통해  $a_{ij}$ 와  $b_{jk}$ 와 같은 전이 확률을 계산한다.

여기서 제시하는 HMM 형식은 에이전트가 행동을 취하면서 세계와 상호작용하는 과정과 태스크의 목적이 변하는 것을 증가, 감소, 지속상수 등으로 인코딩한 파라미터를 통해 모델링한 것이다. 제안 형식은 단순히 정적 상태간의 전이를 모델링하는 전통적 HMM방법과 달리 태스크 목적 파라미터 상에서의 변화를 드러나는 상태로 인코딩하고 수행 행동의 숨은 기저 의도를 은닉상태로 나타낸다. 즉 제안 모델에서 대상의 드러나는 행동 변화는 추적가능하고 관측 가능한 상태이며 에이전트의 의도 목적변화는 은닉상태가 된다. 같은 형태의 행동으로 시

작했다 하더라도 결과가 다르게 나타내는 행동 변화의 목적을 추상화된 환경상태 즉 목적 대상까지의 거리, 목적 대상과의 각도 등의 파라미터로 나타낸다. HMM으로 모니터링하는 파라미터로 행동을 선택한 이유는 행동의 목적이 의도를 전달하고 또한 행동 변화를 추적하는 것이 에이전트의 의도를 추론하고 이해하는데 필수적이기 때문이다.

### 3.2. 행동 모델링 단계

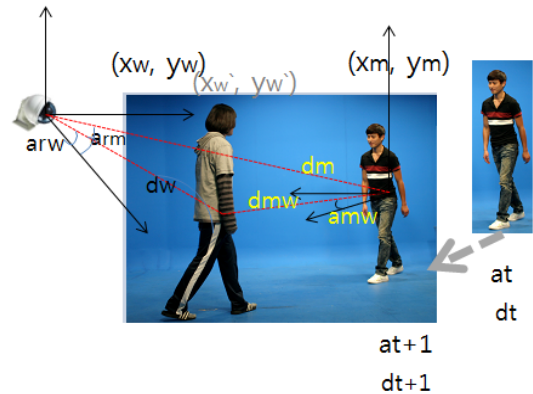
행동 모델링 단계는 인식 대상 추적에 필수적인 단계로 시스템 카메라가 각 에이전트를 관측하는 사전 처리 단계에서 별도로 구축하여, 움직이는 대상의 외형모델을 행동 학습에 이용하도록 한다. 이 과정에서 대상은 전경과 배경 분리를 통해 검출한다. 에이전트 모델이 한번 학습되면 시스템은 다른 동작과 의도 파라미터를 학습하기 위해 미리 설계된 시나리오를 수행한다. 시스템이 움직일 때 배경 추출 처리 단계는 중단되고 시스템은 개선된 평균이동 추적 방법을 이용하여 전경 대상을 추적한다. 그리고 행동을 취하는 자체 경험에서 차후 인식해야 할 각 행동에 해당하는 HMM 형식을 학습한다.

예를 들어 ‘만나기’와 같은 행동에서는 사람의 머리와 몸이 향하는 방향간의 거리와 각도는 두 에이전트가 점점 가까워지면 값이 줄어든다. 따라서 에이전트에서 볼 때 ‘만나기’의 대상이 되는 다른 에이전트에 대한 각도와 거리는 목적에 해당하는 파라미터 {각도=0, 거리=1m}가 된다. 시스템의 관측 가능한 상징은 증가, 감소, 상수, 미지수와 같이 일어날 수 있는 모든 변화의 조합을 모델링하여  $v_k = \{거리: -, 각도: ++\}$ 와 같은 드러나는 값으로 표현한다. 행동 모델에서 드러나는 상징은 추적 모듈의 결과로 만들어진 파라미터의 변화에 해당하고 행동 수행의 중요 의도는 HMM의 은닉상태로 인코딩한다.

주어진 행동의 반복된 실행은 Baum-Welch 알고리즘 [22]을 이용하여 모델 전이 확률  $a_{ij}$ ,  $b_{jk}$ 을 예측하기 위한 데이터를 제공한다. 학습결과 시스템은 각 행동에 대해 한 개의 HMM 집합을 구성한다. 학습에서 관측된, 나타난 행동 상태는 관찰자 시점으로 계산한다.

### 3.3. 의도 인식

의도 인식 문제는 관측된 에이전트에 대해, 이전에 훈련한 HMM 형식을 통해 수행하려는 행동의 의도를 추론하는 과정으로 이루어진다. 의도는 행동의 목적이 은



(그림 2) 의도 인식 과정

된 상태를 말한다. 시스템은 위치별 관심 에이전트의 행동을 모니터링하고 적용 가능한 HMM에 대해 관측 가능한 상징을 평가하며 의도 인식 과정에서는 행동 모델링단계의 학습과는 다른 방법으로 나타난 상징인 파라미터를 계산한다. 즉 관찰자가 장면밖에 있기 때문에 관찰자의 시점보다는 관찰의 대상인 에이전트의 시점으로 행동 특징을 계산하도록 한다. 관측 가능한 상징을 관측 대상이 특정 행동을 취하는 목적과 의도를 추론하여 관측 대상의 시점에서 취하는 것이다. 이 관측 과정은 HMM 학습에 사용된 동일한 파라미터인 상대와의 거리와 각도를 관찰하여 처리한다.

예를 들어 (그림 2)에서 시스템은 관찰하는 에이전트의 시점을 취한다.  $x_{\{m,w,w\}}$ 과  $y_{\{m,w,w\}}$ 간의 거리  $d_{\{m,w,mw\}}$ 는 2차원 좌표이고  $\alpha_{\{r,m,rw,mw\}}$ 은 시스템과 남자간의 각도를 나타낸다. 시스템이 남자의 의도를 검출하기 위해 i)먼저 등을 보인 여자의 위치 좌표를 구한다. ii)그 다음 모니터 대상인 남자에게 좌표시스템을 전달한다. iii)그리고 남자의 시점에서 여자의 위치를 계산하고 iv)남자의 좌표시스템에서 관측 가능한 상징을 계산한다. 남자의 머리는 이전 위치를 통합하여 계산한 다음 ii)단계에서 좌표 시스템의 방향을 결정하도록 한다.

시스템은 전방향 알고리즘[23]을 이용하여 모든 HMM과 각 대상 에이전트에 대해 각 모델이 만든 관측치의 유사도를 계산한다. 에이전트의 의도를 나타내는 가장 가능성 높은 상태를 검출하기 위해 최고 확률의 모델이 나타내는 의도 상태를 선택한다. 모델에 대해서는 Viterbi 알고리즘[24]을 적용하여 은닉상태에서 가장 확률이 높은 시퀀스를 검출한다.

HMM을 이용하는 인식 연구는 관측한 시퀀스와 인식 과정간의 정확한 동기화와 관측한 행동 시퀀스를 명확하게 나눌 때 정확도가 결정된다. 제안 연구에서는 에이전트가 중요하게 생각하는 동작을 모르기 때문에 시퀀스를 나누는 과정이 아무 때나 이루어질 수 있음을 고려해야 한다. 시스템의 관측치가 고정시퀀스가 아닌 연속스트림으로 오는 상황에서 시퀀스 길이가 길어지면 특정 모델의 확률이 0으로 줄어들 수 있기에 이 문제를 해결하기 위해 관측 시퀀스를 가장 최근의  $k$ 개 관측 데이터로 나눈다. 관측 데이터 크기가  $k=30$ 일 때 더 나은 결과를 보이는 것은 휴리스틱하게 찾아낸 것이다.

## 4. 대상 검출과 추적

에이전트가 취하는 행동을 인식하고 모델링하는 기능을 장착한 시스템을 구축하기 위해 비전 기반의 지각 기능이 필요하다. 일반적으로 에이전트들의 외형은 미리 알려지지 않기에 에이전트를 감지하고 추적하는 데 사용할 수 있는 유일한 시각적 단서는 이미지 동작이다. 전역적 움직임을 담은 이미지 시퀀스에서 분리를 수행할 수 있지만 전형적으로 광류 예측에 기반한 접근방식은 시간 소모가 크므로 제안 연구에서는 실시간 감시 방법에서 차용한 대상 모델링과 전경/배경 분리 기술을 이용한다 [25,26].

행동 모델링 단계에서는 동작 학습을 위해 각 에이전트를 관찰하는 기능을 적용하는데 이 시각적 인지 기능은 전경과 배경 분리 기술을 이용하여 대상 에이전트를 검출한다. 에이전트 행동 모델을 한번 학습한 이후, 시스템은 다른 동작과 의도 파라미터를 학습하기 위해 설정된 시나리오를 수행한다. 이 때 배경 추출과정이 이루어지며 이후 시스템은 평균 이동 추적 방법을 통해 전경 대상을 추적한다. 의도 인식 단계에서는 시스템이 다른 에이전트가 취한 행동을 관측하면서 외형모델을 구축하고 대상 추적기의 속도와 강건함을 개선하기 위해 전경/배경 분리 기술을 이용한다.

### 4.1. 전경과 배경 분리

행동 모델링과 의도 인식 단계에서 적용되는 대상 검출과 추적 방식은 히스토그램 기반의 색상 분포로 각 검출 영역을 표현하는 외형 모델을 이용한 커널 기반의 방식을 이용한다[27,28].

각 장면 배경에 나타난 것을 새로운 이미지 프레임과

비교하여 검출을 처리하는데 장면 배경의 통계적 표현을 통해 실시간으로 전경 대상을 검출하기 위해 픽셀의 색상 강도를 이용한다. 정적 배경의 이미지 시퀀스에서 색상 값 또는 픽셀 강도의 분포는 가우시안 분포함수로 모델링하도록 한다. 각 픽셀의 표준 강도를 계산하여 이 평균값과 다른 각 프레임의 픽셀 강도를 찾는 것이 가장 간단한 전경 검출 방법이다[25, 26].

### 4.2. 추적

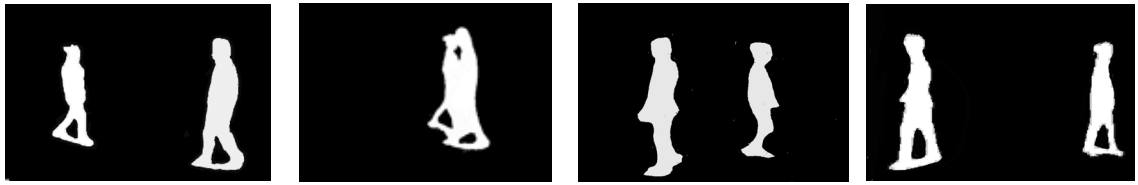
움직이는 대상을 추적하는 시스템은 개별 대상에 대한 외형모델을 설정하여 새로운 프레임에서 검출한 대상자들 가운데 해당 인물을 찾는다. 각 목표 대상을 새로운 프레임에서 찾으면 이 대상에게 단독 ID가 부여된다. 모델 갱신 단계에서는 보이는 대상에 대한 새로운 위치와 기하학적, 픽셀 강도 정보를 갱신한다. 이 정보는 시간상 나중 프레임에서 대상을 인식하고 새로운 위치를 복원하는데 이용한다.

추적을 처리하기 위해 먼저 추적 대상의 신체 상하부 위를 각각 히스토그램 집합으로 표현하고 관련 목록을 유지한다. 관심 대상의 목록은 각 개별 대상의 단일 ID와 공간적 색상 정보로 구성한다. 추적과정에서 시스템은 관측된 대상 리스트로써 가공되지 않은 성분 리스트를 이용하고, 정렬된 대상 리스트 유지와 각 대상의 개별 추적을 위해 통계적 연관 매칭을 이용한다.

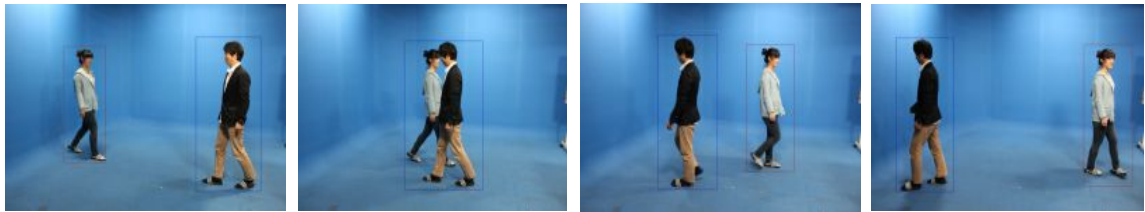
추적 모듈은 외형 모델링, 연관 매칭, 모델 갱신의 세 가지 과정으로 구성한다[29]. 외형 모델링은 가공하지 않은 연관된 성분리스트에서 각 대상에 대해 대상자 밀집 영역의 중심, 대상의 높이와 넓이, 상하 영역의 전경 마스크, 상하 영역 픽셀의 다중변동 가우시안 분포 값을 포함하여 모델을 생성한다. 연관 매칭은 가공하지 않은 각 대상의 상하 영역의 픽셀을 추적한 대상의 정렬 리스트에서 각 모델에 대해 적용한다. 최적 값의 모델 ID로 대상을 표시한다. 추적이 한번 이루어지면 모델을 갱신한다. 성분 리스트에서 보이지 않은, 새롭게 나타난 대상에게는 새로운 ID를 부여하고 적절하게 모델을 갱신하여 다수의 객체를 추적하도록 한다.

### 4.3. 충돌 처리

대상간의 충돌처리를 하여 두 사람이 아주 가까이 있어 겹쳐 보이는 것을 해결하기 위해 검출된 대상의 거리를 이용하여 충돌을 검출한다. 충돌을 검출한 후에는 각 개별대상을 대표모델과 정합하여 최소 정합 점수의 대상



(a) 추적 과정



(b) 비디오 프레임

(그림 3) 비디오 시퀀스 프레임에서의 대상 추적 과정



(a) 따라가기

(b) 만나기

(c) 지나치기

(그림 4) 행동 모델링

은 겹쳐진 상태로 간주한다. 겹쳐진 대상의 모델은 갱신하지 않고 새로운 위치만을 칼만 필터로 예측한다. 겹쳐진 대상의 위치를 갱신하고 평균 이동 알고리즘으로 추적한다.

#### 4.4. 추적 결과

(그림 3)은 에이전트의 검출과 추적 결과를 보여준다. 사람들이 장면에서 나타나고 서로 지나치며 공간을 떠난 후 다시 장면에서 나타나는 내용으로 이루어져 있다. 두 번째 프레임에서 시스템이 겹침을 해결하고 세 번째 프레임에서 한번 겹쳐진 대상이 다시 나타나는 것을 아주 빠르게 복원하는 것을 볼 수 있다.

### 5. 실험

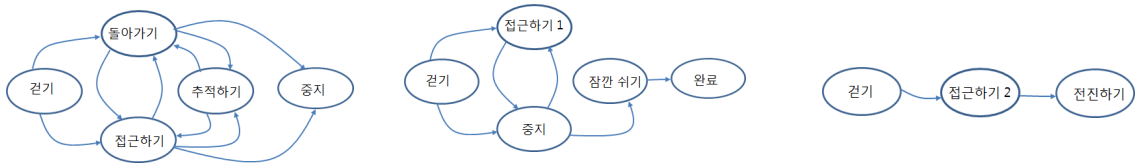
실험은 행동 모델링과 의도 인식 단계에서 각각 수행

한다. 적용한 비디오 시퀀스 프레임 율은 초당 15프레임이다. 시스템이 의도인식을 처리하기 위해 대표적 행동을 모델링하는데 행동 모델링 단계에서는 ‘따라가기’, ‘만나기’, ‘지나치기’의 세 가지 행동을 수행하는 사람의 동작을 학습한다. 본 연구가 대상으로 삼는 행동은 행동 완료 전에 다른 상태로 바뀔 수 있는 행동의 변화와 행동 변화의 파라미터로 삼는 상정을 고려하여 설정하였다.

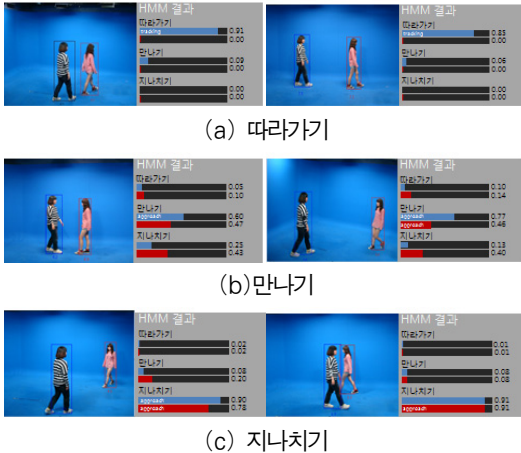
설정된 행동 간의 차이점은 다른 사람에 대한 각도와 거리와 같은 목적 파라미터의 변화로 모델링한다. 만약 행동의 목적이 다른 사람을 만나는 것이라면 다른 사람과의 거리와 각도는 줄어든다. 하지만 다른 사람을 단순히 지나치는 것이라면 거리는 줄어들지만 관측 시점에서 다른 사람과의 각도는 커진다. 이런 관측치를 HMM에 대한 관측 가능한 상정으로 모델링하여 어떻게 에이전트의 인지적 정보가 행동을 수행하는 동안 변하는지를 인코딩한다.

(그림 4)는 행동 모델링 단계에서 행동 추적 결과와 더불어 시스템의 시점으로 본 샘플 프레임이다. 이 과정





(그림 5) 세 가지 행동 HMM 형식 설정

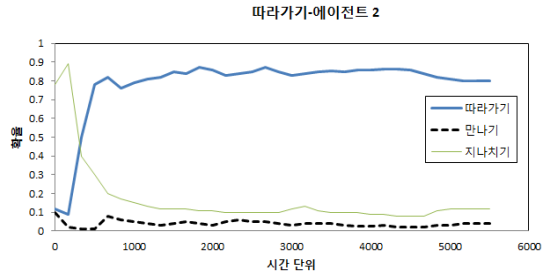
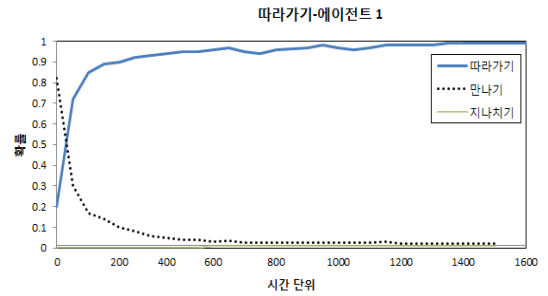


(그림 6) 여러 행동의 의도 인식 결과

에서 얻은 관측치는 (그림 5)와 같이 설정한 행동 HMM 형식을 학습하는데 사용한다. 관측 가능한 상정을 계산하기 위해 모니터한 목적 파라미터는 시스템의 카메라 시점에서 에이전트까지의 거리와 각도이다.

이와 같이 세 가지 HMM 형식을 설정하고 행동을 학습한 후 의도 인식 단계에서 낮은 수준의 행동 목적을 통합하여 높은 수준의 목적인 에이전트의 의도를 인식한다. 이를 위해 전체 행동 HMM 형식과 각 대상 에이전트에 대해 각 모델이 만든 관측치의 유사도를 계산한다. 에이전트의 의도를 나타내는 가장 가능성 높은 상태를 검출하기 위해 최고 확률 의도를 선택한다.

의도 인식 단계에서 사물지능형 컴퓨팅 시스템은 ‘따라가기’, ‘만나기’, ‘지나치기’ 행동을 반복적으로 번갈아 수행하는 시나리오 상의 행동 관찰자가 된다. 이 주요 행동 시나리오 외에 추가한 두 개의 복잡한 시나리오에서는 두 에이전트가 세 가지 행동을 반복하면서 번갈아 수행한다. 첫 번째 세 개의 시나리오의 각각을 두 번 수행하여 시스템이 다른 시점에서 행동을 볼 수 있도록 하고 다양한 환경 조건에서 의도 인식 과정의 강건함을 유지하도록 한다. 추가한 두 가지 복잡한 시나리오의 목적은

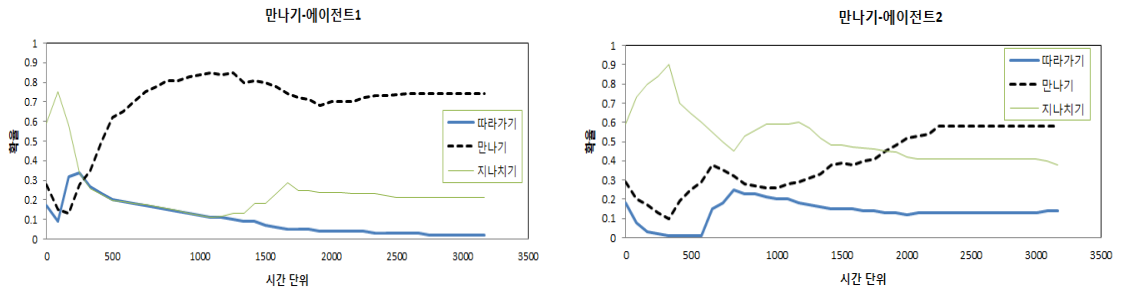


(그림 7) ‘따라가기’ 시나리오에서의 의도인식 확률

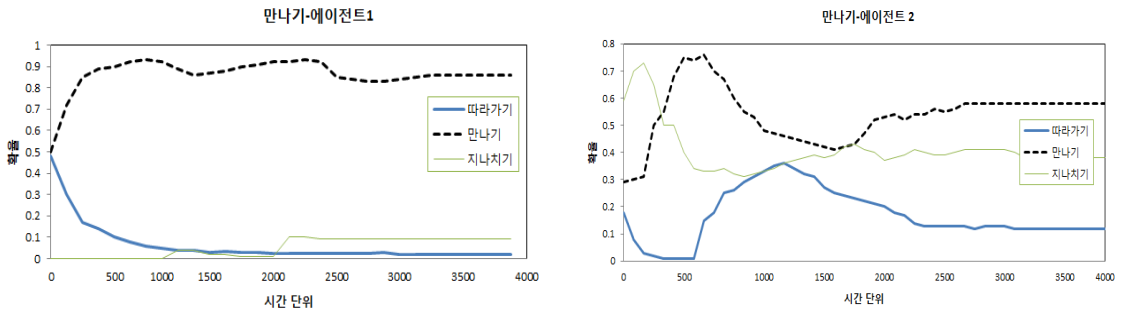
에이전트가 한 행동을 다른 행동으로 바꿨을 때 시스템이 의도의 변화를 추론하도록 하기 위함이다.

세 가지 주요 행동 시나리오에서 세 가지 HMM 형식 각각에 대해 모델이 관측치를 통해 나타내는 확률을 기록한다. (그림 6)은 각 시나리오가 수행되는 동안 다른 시점에서 두 사람을 검출하고 의도 인식하는 과정을 보여준다. 붉은색과 푸른색 바의 길이는 모델의 추적 유사도를, 바 안의 텍스트는 최고 유사모델의 의도적 은닉상태를 나타낸다.

(그림 7-9)는 한 개 비디오 시퀀스 전체 시간대에서 각 모델의 유사도를 나타낸다. 그래프에서 시간 단위는 시퀀스의 한 프레임에 해당하고 확률은 예측한 행동이 실제 행동과 일치하는 유사 확률을 의미한다. 각 그래프에서 정확한 모델의 확률은 낮은 유사도를 보이는 다른 모델의 확률을 빠르게 초과함을 알 수 있다. (그림 8)의 ‘만나기’와 ‘지나치기’ 시점 1, 2는 두 에이전트 각각의 시점

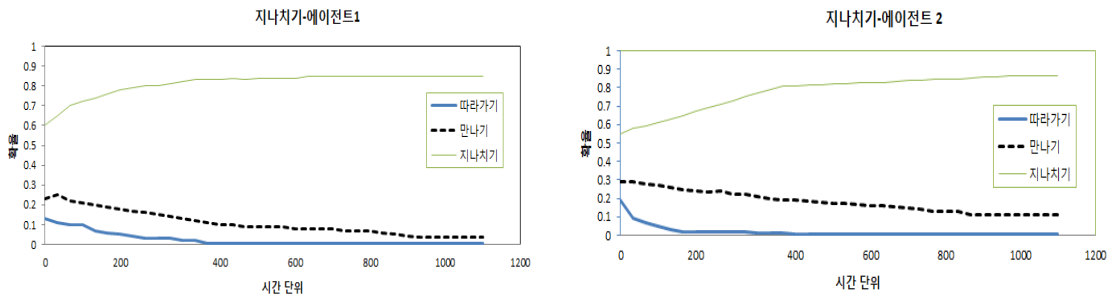


(a) 시점 1

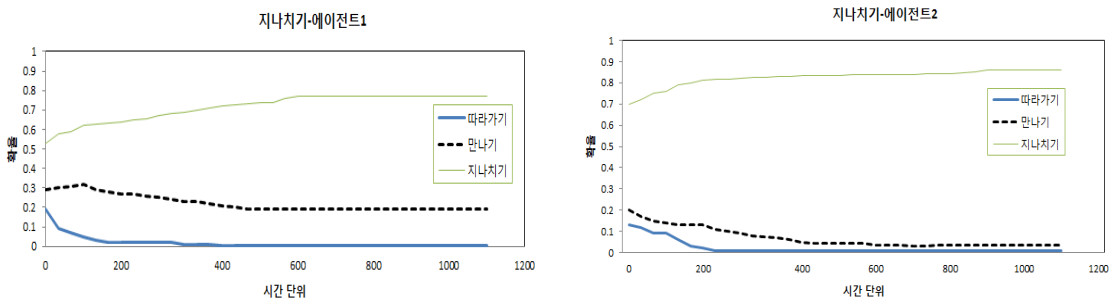


(b) 시점 2

(그림 8) '만나기' 시나리오에서의 의도 인식 확률



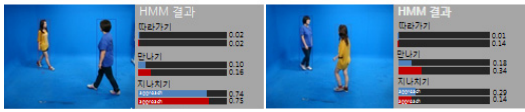
(a) 시점 1



(b) 시점 2

(그림 9) '지나치기' 시나리오에서의 의도 인식 확률





(a) 지나치기



(b) 따라가기

(그림 10) 지나치기-지나치기-변갈아 따라가기 시나리오 결과

에서 행동을 모델링한 것이다.

‘따라가기’ 시나리오에 대해서는 따라가는 사람만의 의도를 보여주고 다른 시나리오에서는 행동을 취하는 두 사람의 의도를 모두 보여주어 시스템이 ‘만나기’ 또는 ‘지나치기’ 의도를 가진 모든 사람들을 검출하도록 한다.

지나치기-지나치기-변갈아 따라가기 행동을 복잡하게 수행하는 시나리오에서는 시스템이 사람의 행동변화에 빠르게 적응하는 지와 그 사람의 정확한 의도 상태를 검출하는 지를 실험한다. 에이전트들이 다른 행동을 연속적으로 취하는 동안 시스템에는 언제 행동이 바뀌는지 시작하는지에 대한 정보를 제공하지 않는 상태에서 (그림 10)에서와 같이 그래프 바가 ‘tracking’ 또는 ‘approach’와 같은 은닉 상태를 표시하는 행동이 가장 높은 현재 확률의 행동 모델이며 기대한 행동 모델을 잘 인식함을 보여주고 있다.

제안 방법의 정량적 평가를 위해 HMM 성능 평가 시 주로 사용하는 세 가지 측정치인 정확도, 조기 검출 정도, 정확한 지속 시간 정도를 실험 데이터를 통해 추정한다. 정확도는 관측 시퀀스 수의 확률로 전체 테스트 시퀀스 수 대비 관측한 의도가 실제 의도와 같을 확률이다. 조기 검출률은  $t^*/T$ 로 관측 시퀀스의 지속 시간  $T$  대비  $t^* = \min\{t | \text{가장 강한 의도적 행동이 } t \text{부터 } T \text{시간동안 최고치}\}$ 로 관측 시퀀스 지속 시간을 100%를 보았을 때 어느 시점부터 의도를 정확하게 유추했는지 그 정도를 의미한다. 정확 지속률은  $C/T$ 로 관측 시퀀스의 지속 시간  $T$  대비 유추한 의도  $C$ 가 시퀀스 지속 시간 전체에서 차지하는 정도이다.

우수한 인식 결과는 높은 정확도와 낮은 조기 검출 정도, 긴 지속시간을 보여야 한다. 제안 시스템은 (그림 6-9)까지의 실험을 통해 ‘따라가기’ 2개, ‘만나기’ 4개, ‘지나치기’ 4개 시퀀스에 대해 정확하게 인식해 의도인

(표 1) 의도 조기 검출률과 정확 지속률 평가 결과

시나리오	제안 방법	
	조기 검출정도(%)	정확 지속률(%)
따라가기 1	1.23	98.8
따라가기 2	3.70	96.3
만나기 1-에이전트 1	0	100
만나기 1-에이전트 2	47.24	86.1
만나기 2-에이전트 1	8.24	97.8
만나기 2-에이전트 2	52.45	47.6
지나치기 1- 에이전트 1	0	100
지나치기 1- 에이전트 2	0	100
지나치기 2- 에이전트 1	0	100
지나치기 2- 에이전트 2	0	100

식 시나리오에 대해 100%의 인식 정확도를 보였다.

(표 1)은 실험에 대한 조기 검출과 정확한 지속시간을 보여준다. 두 가지 경우를 제외하고, 시스템은 행동의 10%가 수행되기 훨씬 전에 정확하게 의도를 유추하고 다섯 가지 경우는 시작 시점부터 의도를 빠르게 검출하여 0%의 조기 검출 정도를 보여 시퀀스 시작부터 정확하게 의도를 유추함을 알 수 있다. 의도를 전체적으로 정확하게 유추한 지속 시간 정도는 아주 높아서 대부분 90%가 넘었다. ‘만나기’의 두 가지 경우만 에이전트 2의 의도를 추론할 때 조기 검출 정도에서 오류를 보였는데 첫 번째 47.42%의 만나기 1-에이전트 2의 경우 시스템이 초기에는 정확한 의도를 추론했으나 중간 지점에서 ‘지나치기’처럼 이해해서 행동이 취해진 지 47.24%의 시간이 지난 다음에야 제대로 의도를 검출했다. 두 번째 오류의 경우 두 번째 에이전트의 시점에서 ‘지나치기’를 ‘만나기’ 행동으로 잘못 이해해서 52.45%의 행동이 취해진 시간이 지난 다음 정확하게 의도를 유추했다. ‘만나기’에 대한 대부분의 시나리오에서는 정확한 지속 정도가 평균 86.09%에 달해 관측 시퀀스 전체적으로 에이전트의 의도를 비교적 잘 유추함을 알 수 있다. 다만 에이전트 2의 시점에서 관측 가능한 상정을 계산할 때 ‘만나기’와 ‘지나치기’ 행동 간의 높은 유사성 때문에 측정값이 다소 낮아졌다.

(표 1)과 같이 동일한 행동시나리오를 이용하여 제안 방법을 템플릿+SVM 기반의 행동 인식 방법[20]과 비교한 결과 드러난 결과 행동만으로는 은닉된 대상의 행동 의도를 포착하기 어렵다. 즉 행동을 인식하기 전 행동의 개별 동작 모델을 미리 설정한 다음 관측하는 행동이 모두 완료된 이후에 관측행동과 개별 동작모델을 비교하는

과정으로 행동의 결과를 인식하기 때문에 행동의 목적이 의도하는 바를 미리 행동 과정 상태에서 파악하기 어려울 수 있다.

그에 반해 제안 연구는 행동 간의 전이 전에 은닉 상태에서 행동 목적이 의도에 따라 어떻게 변하는지를 제시하고 사물지능형 시스템 시나리오에서 실제로 행동이 완료되기 전에 미리 그 의도를 유추하여 미리 대상의 의도를 파악함으로써 에이전트의 의도에 빠르게 대응할 수 있음을 알 수 있다. 이 점이 높은 의도 인식 결과와 더불어 기존 연구가 간과하고 있는 은닉된 목적을 포착하는 과정을 개선한 것이다.

## 6. 결 론

본 연구에서는 사물지능형 컴퓨팅에 적용할 수 있는, 행동을 통해 대상의 의도를 인식하는 방법을 제안했다. 구현 방법은 대부분의 행동 인식 연구에서 취하고 있는 행동이 이루어진 후에 행동의 의도를 검출하는 방법과 달리 다양한 동작을 수행하는 여러 사람의 의도를 행동 과정 중에 빨리 검출하여 시스템에 적용시키는 구조로써 사물지능형 컴퓨팅 환경이 에이전트의 의도를 재빨리 파악해 대응하는 적응적 환경 구현에 기여한다.

이를 위해 행동 HMM 형태를 설정하고 사전에 쌓은 시스템의 경험 데이터를 이용하여 에이전트가 다양한 행동을 수행할 때 이를 학습한 HMM 형태와 비교하여 행동의 의도를 인식하고 사용자와 컴퓨팅 시스템이 원활하게 상호작용하도록 한다. 또한 에이전트의 의도를 추론할 때 여러 에이전트의 시점을 이용하는 방법을 통해 에이전트의 행동이 끝나기 전에 행동의 의도를 미리 유추하도록 한다.

실험 결과를 통해 제안 시스템이 에이전트의 의도를 추론하는 과정에서 사물지능형 환경과 에이전트간의 상호작용을 모델링하여 단순히 주변을 지나치거나 사람을 만나는 것과 같은 행동에 따른 목적 의도를 주어진 비디오 시퀀스에서 빨리 구별함을 알 수 있다. 추론된 의도는 컴퓨팅 시스템의 의사결정과 제어에 다양하게 사용할 수 있다.

제안 연구에서 제시한 지나치기, 만나기, 따라가기와 같은 행동 모델 외에도 관측 상태를 위한 여러 가지 의도 상태를 추가하여 더 정확한 표현으로 정교하게 행동을 모델링할 수 있다. 예를 들어 만나기 전에 두 에이전트의 ‘천천히 다가서기’를 인코딩하는 은닉상태를 추가

하여 거리와 각도, 속도에 대한 다양한 변화 정도를 고려할 수 있다. 또한 필요한 의도인식 과정에 대해 필요한 특징을 설정하고 자동적으로 검출하도록 시스템을 다각도로 설계할 수 있다.

향후에는 제안한 내용을 바탕으로 사물지능형 컴퓨팅 시스템의 다양한 HCI 협력 시나리오를 더 정교하게 구성해 의도를 추론하도록 한다. 누군가 물건을 훔치려는 의도뿐만 아니라 물건을 제공하려는 것인지 제공받으려는 것인지 등에 대한 의도를 유추하여 목적물을 다루는 다양한 행동을 대상으로 실험을 계속할 예정이다. 또한 사물지능형 컴퓨팅에서 숨기거나 가로채기와 같은 더 복잡한 탐지 시나리오에 대해 시스템 대응 범위를 확대하며 다수의 대상을 동시에 추적하도록 하여 의도 인식 과정의 강건성을 높이도록 한다.

## 참 고 문 헌

- [1] M. Iacobini, I. Molnar-Szakacs, V. Gallese, G. Buccino, J. Mazziotta, G. Rizzolatti, Grasping the Intentions of Others with One's Own Mirror Neuron System, PLoS Biol, Vol.3, No.3. pp.529-535, 2005.
- [2] I. Kaliouby, P. Robinson, Real-time inference of complex mental states from facial expressions and head gestures, Proceedings of the 2004 Conf. Computer Vision and Pattern Recognition Workshop, Vol.10, pp.154, 2004.
- [3] D. Koller, N. Friedmann, Probabilistic Graphical Models : Principles and Techniques, MIT Press, 2009.
- [4] D. Premack, G. Woodruff, Does the chimpanzee have a theory of mind?, Behav.Brain Sci. Vol.1, No. 4, pp.515-.526, 1978.
- [5] A. Gopnick, A. Moore, Changing your views: How understanding visual perception can lead to a new theory of mind, Children's Early Understanding of Mind, pp.157-181, 1994.
- [6] A. Woodward, J. Sommerville, J. Guajardo, How infants make sense of intentional action, Intention and Intentionality, MIT Press, pp.149-169, 2001.
- [7] B. J. Grosz, C. L. Sidner, Plans for discourse, Intentions in Communication, Chap. 20, pp. 417-444, 1990.

- [8] D. Avraami-Zilberbrand, G. A. Kaminka, H. Zarosim, Fast and Complete Plan Recognition : Allowing for Duration, Interleaved Execution and Lossy Observations, Proceedings of IJCAI Workshop on Modeling Others from Observations, 2005.
- [9] D. AarNo, D. Kragic, Motion Intention Recognition in Robot Assisted Applications, Robotics and AutoNomous Systems, Vol.56, pp.692-705, 2008.
- [10] P. Krauthausen, U. D. Hanebeck, Intention Recognition for Partial-Order Plans Using Dynamic Bayesian Networks, Proceedings of the 12<sup>th</sup> International Conference on Information Fusion, pp. 444-451 2009.
- [11] S. J. Youn, K. W. Oh, Intention Recognition using a Graph Representation, World Academy of Science, Engineering and TechNology, Vol.25, pp. 13-18, 2007.
- [12] Aaron Beach, Michael Gartrell, Xinyu Xing, Richard Han, Qin Lv, Shivakant Mishra, Karim Seada, SocialFusion: Context-Aware Inference and Recommendation By Fusing Mobile, Sensor and Social Data, Technical Report, CU-CS-1059-09, University of Colorado, 2009.
- [13] K. Ogawara, J. Takamtsu, H. Kimura, K. Ikeuchi, Modeling manipulation interactions by Hidden Markov Models, Proceedings of Int. Conf. Intelligent Robots and Systems, pp.1096-1101, 2002.
- [14] 김진옥, 표정 HMM과 사후 확률을 이용한 얼굴 표정 인식 프레임워크, 한국정보과학회 논문집 : 컴퓨팅의 실제, 제 11호 3권, pp.284-291, 2005.
- [15] N. Nguyen, D. Phung, S. Venkatesh, H. Bui, Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model, Proceedings of IEEE Int. Conf. Computer Vision and Pattern Recognition, pp.955-960, 2005.
- [16] U. Maurer, A. Smailagic, D. Siewiorek, M. Deisher, Activity recognition and monitoring using multiple sensors on different body positions, Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks, pp.99-102. 2006.
- [17] P. Thompson, Weak models for insider threat detection, Proceedings of SPIE, Vol.5403, pp.40-48, 2004.
- [18] T. Gu, Z. Wu, X. Tao, H. K. Pung, J. Lu, epSICAR: An Emerging Patterns based approach to sequential, interleaved and Concurrent Activity Recognition, Proceedings of the IEEE International Conference on Pervasive Computing and Communications, pp.1-9, 2009.
- [19] T. Hu'ynh, M. Fritz, B. Schiele, Discovery of Activity Patterns using Topic Models, Proceedings of the Tenth International Conference on Ubiquitous Computing, pp.10-19, 2008.
- [20] M. Harville, D. Li, Fast, Integrated Person Tracking and Activity Recognition with Plan-View Templates from a single Stereo Camera, IEEE Conf. on Computer Vision and Pattern Recognition, pp. 398-405, 2004.
- [21] T. Duong, H. Bui, D. Phung, S. Venkatesh, Activity recognition and abnormality detection with the switching hidden semi-Markov model, Proceedings of IEEE Int. Conf. Computer Vision and Pattern Recognition, Vol.1, pp.838-845, 2005.
- [22] L. E. Baum, T. Peterie, G. Souled, N. Weiss, A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, Ann. Math. Stat, Vol.41, No.1, pp.164 - 171, 1970.
- [23] L. R. Rabiner, A tutorial on hidden-Markov models and selected applications in speech recognition, Morgan Kaufmann Publishers, 1990.
- [24] G. D. Forney Jr., The Viterbi algorithm, Proceedings of IEEE, Vol.61, No.3, pp.268-278, 1973.
- [25] 김진옥, 색상 조합 모델과 LM알고리즘을 이용한 얼굴 영역 검출, 한국정보처리학회논문지 B, 제 14-B권, pp.255-262, 2007.
- [26] 김진옥, 복잡한 배경의 비디오에서 에이전트 행동 추적 방법, 기초과학, 제 15권 2호, pp.47-54, 2011.
- [27] N. Dalal, B. Triggs, Histogram of Oriented Gradients for Human Detection. Proceedings of International Conference on Pattern Recognition, pp.886-893, 2005.
- [28] D. Comaniciu, V. Ramesh, P. Meer, Kernel-based object tracking, IEEE Trans. Pattern Anal. Mach. Intell. Vol.25, No.5, pp.564-77, 2003.
- [29] D. Ramanan, D. Forsyth, A. Zisserman, Tracking

People by Learning Their Appearances, IEEE Trans.  
Pattern Anal. Mach. Intell, Vol.29, No.1, pp.65-81,  
2007.

● 저 자 소 개 ●



**김 진 옥 (Kim Jin Ok)**

1989년 성균관대학교 졸업(학사)

1998년 성균관대학교 대학원 정보통신공학과 졸업(석사)

2002년 성균관대학교 대학원 전기전자 및 컴퓨터공학과 졸업(박사)

2004년~현재 대구한의대학교 국제문화정보대학 모바일콘텐츠학부 교수

관심분야 : 멀티미디어공학, 패턴 인식, 영상처리, 유비쿼터스 컴퓨팅

E-mail : bit@dhu.ac.kr