

시각 정보에 의한 의도 추론 기법

박진희, 이조셉, 김호준*
한동대학교 정보통신공학과
e-mail : *hjkim@handong.edu

A Method for Intention Inference from Visual Information

Jin-Hui Park, J. S. Lee, Ho-Joon Kim
Dept. of Information Technology, Handong Global University

요 약

본 논문에서는 유비쿼터스 환경에서 인간의 행동패턴을 인식하고 이 결과를 활용하여 사용자의 의도를 추론하는 방법론에 관해 기술한다. 인간행동의 예측에 관한 지식표현으로부터의 추론기능과 예제패턴 기반의 학습기능을 동시에 지원하는 모델을 제시하고 이론의 타당성과 유용성을 고찰한다. 의도 추론 문제에서 지식기반 기법이 갖는 불완전성을 극복하기 위하여 예제기반 학습능력의 필요성을 도출하는 한편, 다양한 변이가 존재하는 응용에서 학습데이터 선정의 어려움을 보완하기 위한 방법론을 제시한다. 세부적으로 인간행동에 관한 특징표현과 행동패턴 클래스를 정의하고 이들간의 관계를 고유한 지식표현 규칙으로 정형화 한다. 또한 제안된 지식표현을 수용하는 추론 메커니즘을 제시하며, 제시한 모델의 부수적 특징으로 학습과정을 통한 지식 정련기능의 유용성을 고찰한다.

1. 서론

인간의 행동패턴을 관찰하여 행위에 대한 의도를 예측하거나 추론하는 연구는 인간과 컴퓨터간의 진보된 인터페이스를 위한 연구의 측면에서 다양한 방법으로 시도되어 왔다[1-3]. 하지만, 시각센서 정보를 기본 데이터로 사용하는 행동패턴 인식기법은 그 기술적인 한계로 인하여 전형적으로 인식결과의 불확실성을 내재하며, 대상자와 환경의 변화에 의하여 일관성이 결핍될 수 있다는 문제점을 갖는다.

실용적 시스템 구축을 위한 어려움으로 다음과 같이 몇 가지 요소를 들 수 있다. 우선, 지식의 표현과 획득의 어려움이다. 매우 다양한 형태로 표현되는 인간의 행동패턴을 제한된 양의 규칙으로 완전히 정의하기가 어려울 뿐만 아니라, 규칙에 사용되는 특징정보의 표현도 명확하게 규정되기 어렵다. 두 번째로는 예제기반 학습모델 구축의 어려움이다. 신경망기법 등 관찰된 예제로부터 스스로 지식을 획득하는 방법론이 알려져 있으나, 인간 행동패턴의 경우 매우 다양한 변이가 존재하여 학습을 위한 표준패턴을 정의하기가 어렵다. 예컨대 동일한 행위분류에 대한 동일인의 행동패턴도 환경이나 여건의 변화에 따라 일관된 형태를 보이지 않을 수 있으며, 다양한 대상자를 고려하면 무수히 많은 변이를 고려하여야 한다. 그러므로 지식기반의 추론기법과 예제기반의 학습기법의 장단점을 상호 보완하여 결합하는 방법론의 제시가 필요하다.

이에 본 연구에서는 지식표현을 수용하는 추론 모델을 정의하고 이 모델이 예제데이터를 통한 학습과정을 통하여 지식을 정련할 수 있음을 보인다. 지식

표현은 특징정보와 행위패턴간의 6 가지 종류의 관계로서 정의되며, 이들을 수용하는 추론 메커니즘을 제시한다. 실험은 실제 사무실 환경에서 정의되는 6 가지 의도 행위에 대하여 예측하는 문제를 고려하였으며, 20 개의 기본 행동의 추출 결과를 특징정보로서 사용하였다.

2. 의도 행위 추론을 위한 지식 표현

시각정보를 기반으로 하여 사용자의 의도행위를 정의하는 일은 매우 복잡하고 어려운 작업이다. 기존의 연구에서 행위패턴에 대한 몇 가지 지식표현과 추론 방법이 제시된 바 있으나[4-5], 이들 대부분의 연구에서 지식표현에 사용한 특징정보는 명확하게 추출할 수 있는 경우임을 전제로 한다. 반면, 컴퓨터 비전 기술에 근거한 행동패턴의 정의는 특징정보 표현의 명확성과 지식범위의 완전성을 갖춘 지식표현이 불가능하다. 이에 본 연구에서는 시각정보기반의 의도추론을 위하여, 고유한 지식표현 기법을 제시하고 지식의 불완전성을 학습과정을 통하여 정련(refinement)하는 방법론을 제시하고자 한다.

본 연구에서는 다음과 같은 형태의 지식표현을 제안하고 이를 수용하는 추론 모델을 제시한다. 편의상 행동패턴에 대한 특징의 종류를 단위동작(unit action), 행동패턴(action) 및 의도행위(behavior)로 구분하고, 이들간의 관계를 다음 6 가지 형태로 정의하였다.

- positive relation: 하위 특징의 활성화 정도가 상위 클래스에 주는 영향을 반영
- negative relation: 하위특징의 비활성화 정도가 상위

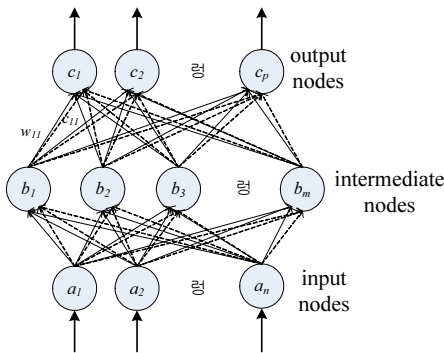
클래스에 주는 영향을 반영

- **excitatory relation:** 하위특징의 값이 상위특징의 활성화에 기여함을 반영
- **inhibitory relation:** 하위특징의 값이 상위특징의 활성화를 억제함을 반영
- **independent relation:** 해당특징과 클래스가 상호 연관성이 없음을 반영
- **unknown relation:** 해당 특징과 특정 클래스의 관계가 알려져 있지 않음을 의미

이들 각각은 0 과 1 사이의 값으로 표현되는 확신도 요소(certainty factor)와 결합되어 표현되며, 이는 다음절에 기술한 추론 네트워크에서 특징들간의 연결강도로써 반영된다.

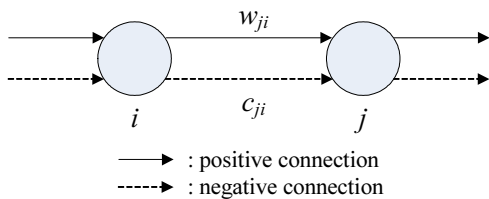
3. 적응형 추론 네트워크

저자의 이전 연구에서 지식기반 추론 능력과 예제기반 학습능력을 동시에 지원하는 패턴 분류 신경망 모델을 제안한 바 있다[6]. 본 연구에서는 행동패턴 인식 및 의도추론의 문제를 위하여 이 모델을 확장 적용한다. 이는 다층구조를 갖는 일종의 신경망 모델로 이른바 ‘적응형 추론 네트워크(Adaptive Inference Network)’이며, 고유의 연결구조 및 동작특성을 갖는다. 본 절에서는 네트워크의 구조, 동작특성, 학습 알고리즘 및 유용성에 관해 기술한다.



(그림 1) 추론 네트워크의 구조

(그림 2)에 보인 바와 같이 노드들간의 연결형태는 이중 연결구조를 가지며 각 연결은 별도의 가중치가 적용된다.



(그림 2) 노드들간 이중 연결구조

네트워크의 활성화 함수는 다음과 같다.

$$f(Net_j) = \begin{cases} 1 & \text{if } Net_j > 1 \\ \frac{1}{2}(Net_j + 1) & \text{if } -1 \leq Net_j \leq 1 \\ 0 & \text{if } Net_j < -1 \end{cases} \quad (1)$$

$$Net_j = \sum_{i=1}^{m_j} w_{ji} x_i + \sum_{i=1}^{m_j} c_{ji} (1 - x_i) \quad (2)$$

네트워크를 구성하는 노드간에 두 개의 연결구조는 각각 하위특징의 활성화 정도와 비활성화 정도에 특정가중치를 결합하여 반영한다. 식(2)에서 보인 바와 같이 활성화 정도에 대하여 양성가중치(positive weight) w_{ji} 가 반영되며, 비활성화 정도에 따라 음성가중치 (negative weight) c_{ji} 가 반영된다.

이러한 형태의 동작특성은 기존의 다층 퍼셉트론 (multi-layer perceptron) 신경망에서 바이어스(bias) 노드의 역할을 각 하위노드에 분산 시킴으로써 각 노드가 외부적인 특징표현으로 반영될 수 있는 좀더 실질적인 의미를 갖도록 한다. 위의 활성화 함수는 학습과정에서 델타 규칙을 적용시키기 위하여 식(1)에 관하여 구간별로 미분을 취하거나, 함수 자체를 다음과 같은 시그모이드 함수로 근사화 할 수 있다.

$$f(Net_j) = \frac{1}{1 + e^{-Net_j/T}}$$

이러한 추론 네트워크의 구조 및 동작특성은 기존의 전형적인 신경망 모델과는 달리, 각 노드에 외부적인 의미 부여가 자연스럽게 가능하고 규칙기반으로 표현되는 지식표현을 수용하여 추론 기능을 지원할 수 있게 한다. 이에 관해서는 다음절에 기술한다.

4. 지식기반 추론 기능

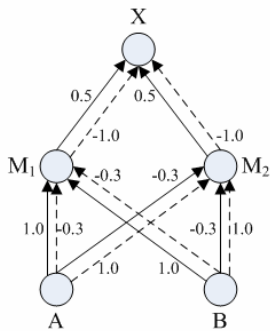
앞 장에서 기술한 네트워크는 2 장에 기술한 지식표현을 수용하는 추론기능을 지원한다. 다시 말해서 확신도(cf: certainty factor) 값과 결합된 6 가지 관계 표현으로부터 노드들간의 가중치를 다음과 같은 규칙으로 초기화 한다.

- **positive excitatory relation:** 해당 노드간의 가중치 w_{ji} 를 cf 값으로 설정
- **negative inhibitory relation:** 해당 노드간의 가중치 w_{ji} 를 $-cf$ 값으로 설정
- **negative excitatory relation:** 해당 노드간의 가중치 c_{ji} 를 cf 값으로 설정
- **negative inhibitory relation:** 해당 노드간의 가중치 c_{ji} 를 $-cf$ 값으로 설정
- **independent relation:** 해당 노드간의 가중치 w_{ji} , c_{ji} 를 모두 0 으로 설정
- **unknown relation:** 해당 노드간의 가중치 w_{ji} 및 c_{ji} 를 작은 수치의 무작위 값으로 초기화

예로서 논리연산 XOR 의 규칙을 <표 1>과 같은 relation 으로 표현할 수 있고 이를 반영한 추론 네트워크는 (그림 3)와 같다.

<표 1> XOR 연산을 위한 규칙 표현

feature variables	relation type	cf
A → M ₁	pos. excitatory	1.0
A → M ₂	pos. inhibitory	0.3
A → M ₁	neg. inhibitory	0.3
A → M ₂	neg. excitatory	1.0
B → M ₁	pos. excitatory	1.0
B → M ₂	pos. inhibitory	0.3
B → M ₁	neg. inhibitory	0.3
B → M ₂	neg. excitatory	1.0
M ₁ → X	pos. excitatory	0.5
M ₁ → X	neg. inhibitory	1.0
M ₂ → X	pos. excitatory	0.5
M ₂ → X	neg. inhibitory	1.0



(그림 3) XOR 연산을 위한 추론 네트워크

다시 말해서, 어떤 문제에 관하여 전형적인 규칙표현인 <IF-THEN> 규칙으로부터 특징정보와 목표패턴 클래스간의 관계를 도출할 수 있고 각각에 대하여 전문가의 직관 등에 의하여 확신도를 정의할 수 있다면 이를 수용하는 추론 네트워크의 구현이 가능하다.

5. 학습 알고리즘

앞절에서 기술한 바와 같이 제안된 추론 네트워크 모델은 주어진 지식으로부터 노드들간의 연결가중치를 초기화 하고, 예제기반의 학습기능을 통하여 지식을 정련할 수 있다. 학습 과정은 다음과 같다.

학습을 위하여 출력 노드에서의 에러를 식(3)과 같이 정의 하는데, 식에서 t_j^p 와 z_j^p 는 각각 학습패턴 p 에 대하여 해당 노드의 기대 출력값(desired output)과 실제 출력값(actual output)을 의미한다.

$$E^p = \frac{1}{2} \sum_j (t_j^p - z_j^p)^2 \quad (3)$$

학습패턴 p 에 의하여 각 노드들간에 두 종류의 연결가중치 w_{ji} 와 c_{ji} 에 가해지는 변화분은 다음 식과 같이 산출된다.

$$\Delta^p w_{ji} = \eta_1 \frac{\partial E^p}{\partial z_j^p} f'(Net_j^p) x_i^p$$

$$\Delta^p c_{ji} = \eta_2 \frac{\partial E^p}{\partial z_j^p} f'(Net_j^p) (1 - x_i^p)$$

즉, 역전파된 에러의 크기와 활성화 함수의 미분값 및 하위노드의 활성화 값에 비례하여 변화의 크기와 방향을 결정하게 된다. 이 때 적용되는 에러값은 상위노드가 출력노드인 경우와 중간노드의 경우를 구분하여 다음과 같이 결정된다.

$$\frac{\partial E^p}{\partial z_j^p} = \begin{cases} -(t_j^p - z_j^p) & \text{: output node} \\ \sum_k \frac{\partial E^p}{\partial Net_k^p} (w_{kj} - c_{ji}) & \text{: intermediate node} \end{cases}$$

6. 응용 및 실험

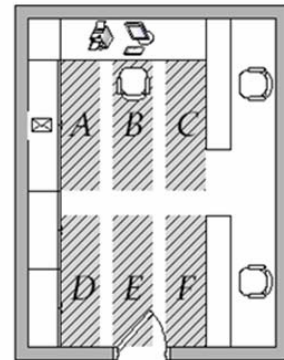
6.1 환경설정 및 특징표현

제안된 내용의 유용성을 고찰하기 위하여 실제 실험에 적용하였다. 응용 대상은 사무실 환경에서 6 가지 의도행동에 대한 예측문제를 고려하였으며, 기본 특징으로서 20 종류의 단위 특징을 정의하였다. 이를 <표 2>에 나타내었다.

<표 2> 기본 특징과 행위 패턴 클래스 정의

기본 특징	f_1 : 최종위치 A	f_2 : 최종위치 B
	f_3 : 최종위치 C	f_4 : 최종위치 D
	f_5 : 최종위치 E	f_6 : 최종위치 F
	f_7 : 의자에 앉기	f_8 : 서기
	f_9 : 문을 열기	f_{10} : 상체 숙이기
	f_{11} : 쪼그리고 앉기	f_{12} : 두리번 거리기
	f_{13} : 머리 숙이기	f_{14} : 뒤돌아 서기
	f_{15} : 물건 들기	f_{16} : 물건 놓기
	f_{17} : 팔 내리기	f_{18} : 팔 뻗기
	f_{19} : 팔 들기	f_{20} : 다리 들기
패턴 클래스	C_1 : 우편물 찾기	C_2 : 컴퓨터사용
	C_3 : 민원 문의	C_4 : 물품 대여
	C_5 : 물품 반납	C_6 : 비정상적 물품대여

대상공간의 공간적 정보를 표현하기 위하여 (그림 4)와 같은 평면도의 사무실을 대상으로 하여 기본 행동과 행위패턴을 고려하였다.



(그림 4) 실험대상 사무실 공간 평면도

6.2 의도 추론을 위한 지식 표현

앞 절에서 제시한 특징 표현 및 패턴클래스에 대하여 의도추론을 위한 지식을 정의하였다. <표 3>은 대상응용에 대하여 6 가지 관계로 표현되는 지식표현의

예이다. 표에 나타난 바와 같이 각 특징정보와 목표 클래스 간의 관계를 확신도 요소와 결합하는 형태로 기술 하였다.

<표 3> 특징 정보와 패턴 클래스와의 관계 표현

하위 특징	상위 클래스	relation	cf
f_{11}	C_1	positive excitatory	0.10
f_7	C_2	positive excitatory	0.90
f_7	C_5	positive inhibitory	0.80
f_{17}	C_1	positive inhibitory	0.30
f_{11}	C_3	negative excitatory	0.70
f_{10}	C_3	negative excitatory	0.70
f_8	C_3	negative inhibitory	0.75
f_9	C_1	negative inhibitory	0.60
f_{12}	C_3	unknown	-
f_{20}	C_5	independent	0.00

6.3 실험결과

실제환경에서 촬영한 일련의 동영상으로부터 행위 패턴을 분류하기 위하여 기본 특징을 추출한다. 이러한 특징값들이 추론 네트워크에 적용되어 의도 행위를 분류하게 된다. 실험에는 15 개의 패턴에 대하여 주어진 지식만을 적용한 추론 결과를 관찰하였으며, 10 개의 예제패턴으로 학습시킨 후 5 개의 패턴에 대하여 인식한 결과를 고찰 하였다.

(그림 5)는 실험에 사용한 동영상의 한 장면에 대한 예와 해당 행위에 대하여 추출한 특징 값이다.



(그림 5) 입력 영상에 대한 특징 추출 결과

<표 4> 의도 행위 분류 결과

지식 표현만을 적용한 추론 결과	학습과정 이후의 추론 결과
$C_1: C_4$	$C_1: C_1$
$C_2: C_2$	$C_2: C_2$
$C_3: C_3$	$C_3: C_3$
$C_4: C_5$	$C_4: C_4$
$C_5: C_4$	$C_5: C_4$

이에 대해 적용한 두 종류의 실험 결과를 <표 4>에

보였다. 실험결과에 보인 바와 같이 학습과정이 가해진 추론결과는 지식표현만으로 추론한 결과에 비해서로 다른 활성화 특성을 보인다. 이는 초기의 지식 표현이 현재의 환경과 대상자에 대해서는 다소의 왜곡이 있음을 의미한다. 다시 말해서, 알려진 지식은 실제 상황에 관찰되는 데이터에 의해 정련될 수 있음을 보인다.

<표 5> 정련된 지식표현의 예

하위특징	상위클래스	relation	cf
f_1	C_4	pos. excitatory	0.60→0.70
f_{15}	C_3	pos. inhibitory	0.50→0.80
f_{16}	C_4	neg. excitatory	0.40→0.36
f_9	C_1	neg. inhibitory	0.60→0.70

7. 결론

본 연구에서는 제시한 추론 네트워크는 예제기반 학습 능력과 규칙기반의 추론능력을 동시에 지원한다. 또한 이중연결구조의 특성으로부터 특징정보와 패턴 클래스간의 세분화된 관계를 보다 잘 반영하며, 학습 과정은 전문가로부터 획득한 지식의 미비점을 보완할 수 있게 한다. 컴퓨터 비전 기술에 근거한 행동패턴의 추출과정은 환경적 요인과 대상자의 변화에 따라 그 정확성과 일관성이 결여될 수 있다는 점에서 지식기반과 학습기능을 결합하는 추론 기법은 상호의 장단점을 조화시킬 수 있게 한다. 또한 제시된 모델의 특성을 활용하여 지식표현에 대한 검증기법과 학습패턴의 유효성을 판정하는 방법으로도 활용될 수 있을 것이다.

*본 연구는 21 세기 프론티어 연구개발사업의 일환으로 추진되고 있는 지식경제부의 유비쿼터스컴퓨팅및네트워크원천기반기술개발사업의 08B3-O4-11T 과제로 지원된 것임.

참고문헌

- [1] Henry A. Kautz: A formal theory of plan recognition and its implementation, Reasoning about plans, Morgan Kaufmann Publishers Inc., San Francisco, CA, 1991.
- [2] Sandra Carberry: Techniques for Plan Recognition, User Modeling and User-Adapted Interaction, Vol. 11. 31-48, 2001.
- [3] Neal Lesh, Charles Rich, and Candace L. Sidner: Using plan recognition in human-computer collaboration: Proceedings of the Seventh International Conference on User Modeling, 23-32, 1999.
- [4] Weiming Hu, Tieniu Tan, Liang Wan, Steve Maybank: A Survey on Visual Surveillance of Object Motion and Behaviors, IEEE Transaction on Systems, Man and Cybernetics, Part C, Vol.34, No.3, 334-352, 2004.
- [5] Alper Yilmaz, Mubarak Shah: Actions Sketch: A Novel Action Representation. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1. 984-989, 2005.
- [6] Ho-Joon Kim, Hyun S Yang: A Neural Network Capable of Learning and Inferencing for Visual Pattern Recognition, Pattern Recognition, Vol. 27, No. 10, 1291-1302, 1994.