

# Ontological 지식 기반 영상이해시스템의 구조

## Framework for Ontological Knowledge-based Image Understanding Systems

손 세 호, 이 인 근, 권 순 학

영남대학교 전자정보 공학부

Seo H. Son, In K. Lee, Soon H. Kwon

School of Electrical Engineering and Computer Science

Yeungnam University

E-mail : kazma@yumail.ac.kr

### ABSTRACT

In this paper, we propose a framework for ontological knowledge-based image understanding systems. Ontology composed of concepts can be used as a guide for describing objects from a specific domain of interest and describing relations between objects from different domains. The proposed framework consists of four main subparts: i) ontological knowledge bases, ii) primitive feature detectors, iii) concept inference engine, and iv) semantic inference engine. Using ontological knowledge bases on various domains and features extracted from the detectors, concept inference engine infers concepts on regions of interest in an image and semantic inference engine reasons semantic situations between concepts from different domains. We present a outline for ontological knowledge-based image understanding systems and application examples within specific domains such as text recognition and human recognition in order to show the validity of the proposed system.

Keywords : image understanding system, ontology, ontological knowledge bases,

### 1. 서론

지난 20여년동안 SIGMA, SCHEMA, SYGAL 그리고 PROGAL과 같은 다양한 지식 기반 영상이해시스템들이 제안되었다[1-3]. 대부분의 지식 기반 시스템들은 적용분야에 대해 특별히 정의된 지식베이스를 사용하여 영상을 해석한다. Yanai는 각 object에 대해 정의된 object-oriented agent를 사용하여 object를 추출한 후 objects들의 상대적 관계를 이용해서

영상을 이해하는 multi-agent 구조를 가지는 영상이해시스템을 구축하였다[4]. 하지만 이러한 지식베이스의 구축은 적용분야와 사용되는 영상 처리 방법들에 따라 다른 구조를 가지고 있어 범용적인 영상이해시스템의 구축에 한계를 가진다. 이러한 문제를 해결하기 위해 최근에는 온톨로지를 사용한 영상이해에 대한 연구가 이루어지고 있다. 온톨로지(ontology)란 특정 영역에 정의되는 개념들(concepts), 개념들의 속성(properties), 그리고 개념들 간의 의미

관계들(semantic relations)로 표현되는 지식이다. Maillot이 제안한 온톨로지 기반 인지 비전 시스템은 색 개념 (color concept), 질감 개념 (texture concept) 그리고 공간적 시간적인 개념(spatio-temporal concept)의 3가지 시각적 개념(visual concept)을 사용하여 domain-independent한 지식베이스를 구축하여 영상을 해석하였다[5,6]. Maillot은 의미학적인 분류(taxonomy)에 기반하여 개념의 온톨로지를 정의하였다. 하지만 개념의 시각적인 표현은 개념들의 의미학적인 분류에 의해서가 아니라 구성학적 분류(partonomy)에 의해서 이루어진다.

본 논문은 온톨로지의 표현방법인 의미학적 분류와 구성학적 분류를 이용하여 특정 영역에 대한 개념들에 대해 domain-independent한 지식 베이스를 구축하여 영상을 해석한다. 본 논문에서는 구성학적 표현으로 구축된 개념들의 지식베이스를 이용하여 개념을 인식한 후 개념들 간의 관계로부터 관심물체를 추론한다. 그리고 추론된 관심 물체들의 의미 관계를 의미 관계 지식베이스를 사용하여 추론하여 영상을 이해하고 해석한다.

## 2. 본론

본 논문은 ontological 지식기반 영상이해시스템의 개괄적인 구조를 제안하고자 한다. 제안된 시스템의 구조는 그림 1과 같다.

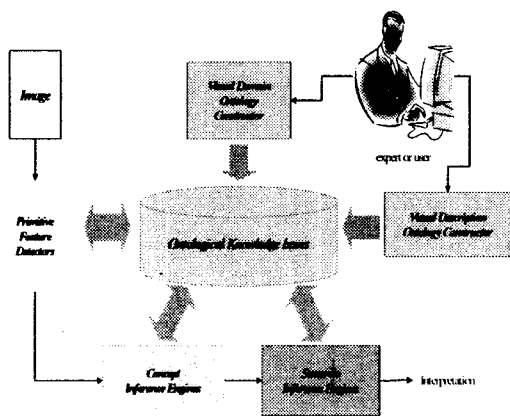


그림 1. 제안된 시스템의 구성도

제안된 영상이해시스템은 크게 개념과 개념

들 간의 구성학적, 의미학적 관계에 대한 지식을 가지고 있는 ontological 지식베이스, 개념의 형태학적 특징을 추출하는 영상 처리부, 추출된 특징들로부터 개념을 추론하는 개념 추론부, 그리고 개념들 간의 관계로부터 의미를 추론하는 의미 추론부로 구성된다.

### 2.1 Ontological 지식 베이스

Ontological 지식베이스는 크게 개념의 속성을 형태학적인 정보로 표현하는 visual description의 온톨로지에 관한 지식베이스, 관심 영역의 개념들의 구성학적인 분류에 관한 온톨로지를 표현하는 지식베이스, 개념들의 의미학적 분류에 관한 온톨로지를 표현하는 지식 베이스, 그리고 다른 분야의 개념들 간의 의미 관계에 대한 온톨로지를 표현하는 지식베이스로 구성된다.

먼저, 개념은 다음의 두 가지 visual description으로 시각적인 정보를 표현한다. 그림 2에서처럼 개념은 형태학적인 정보를 형태학적 표현정보들(Morphological Descriptions, MDs)과 개념들 간의 형태학적 구성정보들(Organizational Descriptions, ODs)로 구분하여 정의한다. 또한 개념들은 형태학적 표현정보들, 형태학적 구성관계정보들뿐만 아니라 하위 레벨의 개념들의 관계에 의해서도 정의되어진다.



그림 2. 개념의 정의

개념의 형태학적인 정보를 표현하는 형태학적 표현정보는 색, 모양, 질감 등등과 같은 기본적인 시각 정보들로 이루어진다. 개념들 간의 형태학적인 구성 관계를 나타내는 형태학적 구성정보는 같은 레벨에서의 개념들 간의 공간적 관계, 상위 개념들 간의 공간적 관계 등등과 같은 개념들 간의 공간적인 구성 정보를 표현한다. visual description 영역에 대한 온톨로지는 그림 3과 같다.

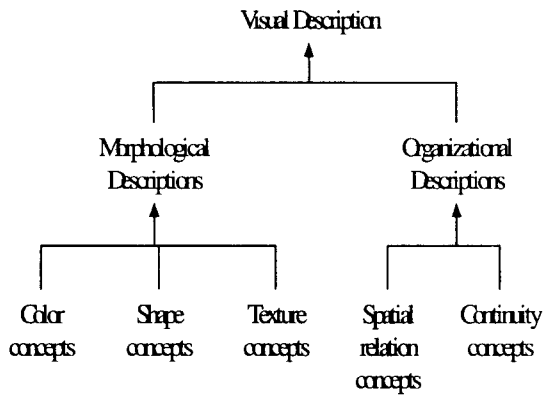


그림 3. visual description 온톨로지

visual description에 관한 온톨로지와 개념들 간의 구성학적인 정보를 사용하여 관심 물체의 개념들의 온톨로지에 대한 지식베이스를 구축한다. 개념들의 구성학적인 분류에 대한 일반적인 구조는 그림 4와 같다.

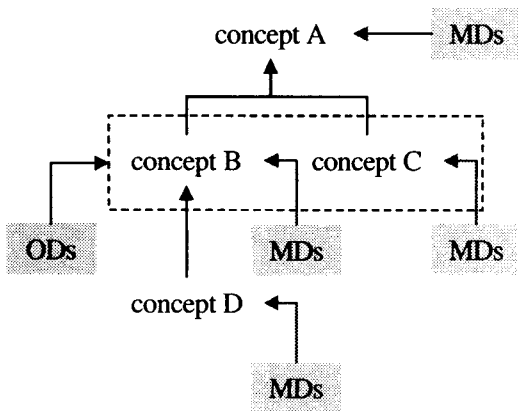


그림 4. 개념의 구성학적 분류

그림 4에서 개념 A는 개념 B와 C의 부모 노드이고 B와 C는 형제노드이다. 개념 B는 하위 레벨의 D와 자신의 형태학적 표현정보에 의해 정의되며 상위 개념 A는 자신의 형태학적 표현정보, 같은 레벨의 개념 B, C들 간의 관계, 그리고 B, C와 상위 레벨의 개념 A와의 형태학적 구성정보에 의해 정의된다. 그림 5는 인간의 구성학적 분류의 예이다.

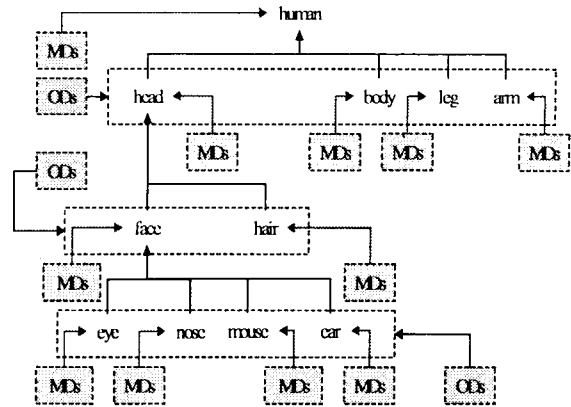


그림 5. 인간의 구성학적 분류

개념들의 의미학적 정보를 사용하여 의미학적인 분류에 관한 온톨로지를 표현하는 지식베이스를 구축한다. 그림 6은 의미학적인 정보를 가지는 개념들의 예를 나타낸다.

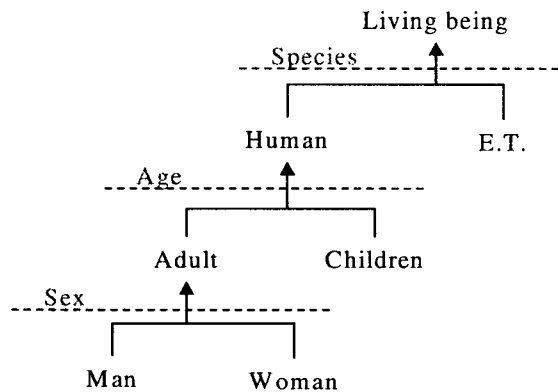


그림 6. 인간의 의미학적 분류

의미 관계 지식베이스는 다른 분야들의 개념들 간의 의미 관계에 대한 온톨로지를 표현한다. 이 지식베이스는 다음의 정의들을 사용하여 서로 다른 분야의 개념들의 관계에 대해 표현한다.

정의 1)

$\mu_{CONCEPT}$ (영역  $\Omega_i$ , 개념 A)는 영역  $\Omega_i$ 가 개념 A를 나타낼 가능성이다. 영역  $\Omega_i$ 에 대해 가장 큰 가능성을 가지는 개념  $A^*$ 을  $CONCEPT(\text{영역 } \Omega_i, \text{개념 } A^*)$ 로 나타낸다. 여기서  $\Omega_i$ 는 개념  $A^*$ 의 영역이다.

정의 2)

$\mu_{CONCEPT}(\bigcup_j A_j^*, \text{개념 } B)$ 는 개념  $\bigcup_j A_j^*$ 의 상위 개념이  $B$ 일 가능성이다. 개념  $\bigcup_j A_j^*$ 에 대해 가장 큰 가능성을 가지는 개념  $B^*$ 를  $CONCEPT(\bigcup_j \Omega_j, \text{개념 } B^*)$ 로 나타낸다. 여기서  $\Omega_j$ 는 개념  $A_j^*$ 의 영역이다.

정의 3)

$\mu_{SR}(\text{개념 } A^*, \dots, \text{개념 } B^*, \text{관계 } C)$ 는 개념들이 관계  $C^*$ 를 가질 가능성이다. 단, 의미 관계 추론에 사용되는 개념들은 2차원 공간에 대한 개념을 3차원 공간에 대해 확장한 개념들이다. 개념  $A^*, B^*$ 에 대해 가장 큰 가능성을 가지는 관계  $C^*$ 를  $SR(\text{개념 } A^*, \dots, \text{개념 } B^*, \text{관계 } C^*)$ 로 나타낸다.

예를 들어, 의미 관계 지식베이스의 개념들의 관계 정보로부터  $CONCEPT(\Omega_1, \text{사람})$ 과  $CONCEPT(\Omega_2, \text{차})$ 으로부터 두 개념은  $SR(\text{사람, 차, 안에 있다})$ ,  $SR(\text{사람, 차, 가리다})$  등등의 여러 가지 의미 관계들로 표현 가능하다.

### 2.2 특징 검출부

특징 검출부에서는 ontological 지식베이스를 이용하여 영상을 색, 모양, 질감 등등의 형태학적 정보들로 분석한다. 형태학적 정보들의 분석에 앞서 영상을 유사한 정보를 가지는 영역들을 합병하여  $I = \bigcup_{i=1}^n \Omega_i$ 으로 재구성한다.

ontological 지식 베이스 내의 각 영역에 대한 개념들의 형태학적 표현정보를 이용하여 영역  $\Omega_i$ 의 개념에 대한 가능성을 다음과 같이 표현한다.

$$\begin{aligned} \mu_{CONCEPT.MDs}(\Omega_i, A) &= \mu_{CONCEPT.Color}(\Omega_i, A) \\ &\quad \wedge \mu_{CONCEPT.Shape}(\Omega_i, A) \\ &\quad \wedge \mu_{CONCEPT.Texture}(\Omega_i, A) \end{aligned}$$

여기서, 영역  $\Omega_i$ 에 대한 개념  $A$ 의 색, 모양, 질감에 대한 각각의 가능성은  $\mu_{CONCEPT.Color}(\Omega_i, A)$ ,  $\mu_{CONCEPT.Shape}(\Omega_i, A)$ ,  $\mu_{CONCEPT.Texture}(\Omega_i, A)$ 이다.

### 2.3 개념 추론부

특징 검출부에서 얻은 개념에 대한 가능성을 바탕으로 영역  $\Omega_i$ 의 개념을 추론한다. 영역  $\Omega_i$ 의 개념  $A^*$ 은 가장 큰 가능성  $\mu_{CONCEPT.MDs}(\Omega_i, A^*)$ 을 가지는 개념이다. 이 때 영역  $\Omega_i$ 를  $CONCEPT(\Omega_i, A^*)$ 와 같이 나타낸다.

개념들의 위치관계 정보인 형태학적 구성 정보들을 이용하여 상위 개념을 추론한다. 즉, 개념  $\bigcup_j A_j^*$ 의 상위 개념은 개념들 간의 형태학적

구성관계인  $\mu_{CONCEPT.ODs}(\bigcup_j A_j^*, B)$ 을 사용하여 추론한다. 추론된 상위 개념의 영역은  $CONCEPT(\Omega_k, B^*)$ 와 같이 표현 가능하다. 여기서  $\Omega_k$ 는  $\bigcup_j \Omega_j$ 이다.

### 2.4 의미 추론부

두 개 이상의 다른 분야에 대해 정의되어진 지식베이스를 사용하여 개념들을 추론할 경우 동일한 영역에 복수개의 개념을 추론하는 모순이 발생한다. 이러한 모순은 물체의 3차원 입체 정보를 2차원의 평면 정보로 변환하여 영상으로 표현하기 때문에 발생한다. 즉, 3차원 공간에서는 서로 다른 위치에 존재하는 개념들이 2차원 영상으로 투영되어질 때 각 영역들이 중복되어 나타난다. 이를 해결하기 위해 입력 영상을 보는 가상 관측자를 가정하여 2차원 공간상의 개념을 3차원 공간상에서의 개념으로 확장한다. 2차원 영상의 임의의 두 개념 A, B에서

개념 A의 영역이 개념 B에 의해 분리되거나 개념 B의 영역을 포함하고 있을 때 가상 관측자의 관점에서 개념 B는 개념 A보다 가상 관측자에게 가까이 있다. 이 사실로부터 동일한 영역의 복수개의 개념에 가상 관측자 기준에서 본 개념들의 3차원 정보를 추가하여 개념들 간의 관계를 추론에 이용한다. 또한 분류학적 개념 지식베이스와 의미 관계 지식베이스를 바탕으로 다른 분야들의 개념들 간의 의미 관계를 추론한다. 여기에 사용되어 지는 분류학적 개념 지식베이스와 의미 관계 지식베이스는 WordNet, Open Directory Web 등등의 다양한 지식베이스들을 사용하여 구축되어 진다.

### 3. 모의 실험

이 절에서는 본 논문의 타당성을 증명하기 위해 문자 인식과 사람 인식, 두 분야에 대하여 ontological knowledge bases를 구축하여 제안된 영상이해시스템을 사용하여 모의 실험한다. 그림 7은 문자 인식과 사람 인식에 사용한 구성학적 분류에 따른 각각의 온톨로지에 대한 표현이다.

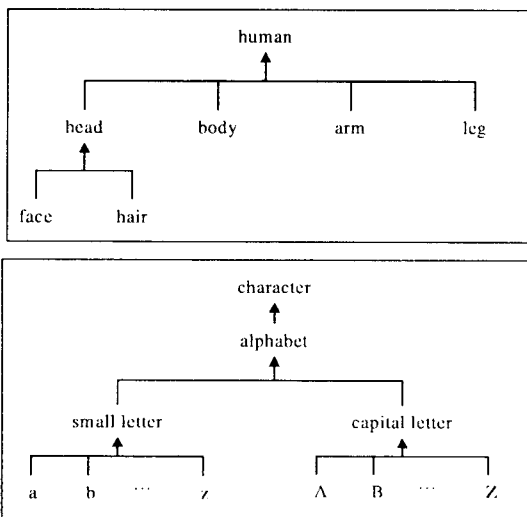


그림 7. 문자와 사람의 구성학적 온톨로지

그림 8은 문자와 사람으로 구성된 입력 영상과 사람과 문자에 대한 ontological 지식 베이스를 가진 시스템에 의해 분석된 결과 영상이다. 제안된 시스템은 그림 7의 구성학적 분류에 대한

온톨로지 뿐만 아니라 문자 인식의 문자 모양에 관한 모양 온톨로지를, 사람 인식의 피부색에 관한 색상 온톨로지와 사람 모양에 관한 모양 온톨로지를 사용하였다. 입력 영상의 문자 인식 결과로 "KFC STOP TORTURING CHICKS"라는 문장이 구성된다. 입력 영상이 가지고 있는 문장을 해석하는 문제는 자연언어처리의 형태소 해석, 구문해석, 그리고 의미 해석 문제로 변환하여 해결한다.

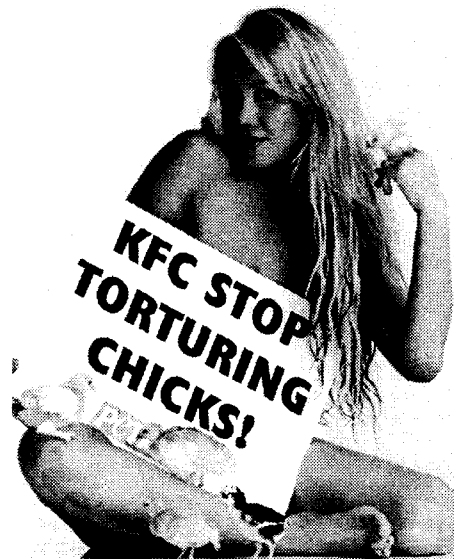


그림 8. 입력 영상과 결과 영상

복수개의 인식영역에 의해 동일한 영역에 두 개 이상의 개념이 추론되어 개념의 모순성이 발생한다. 하지만 이 모순성은 의미추론 과정

에서 해결된다. 문자는 2차원 공간의 표현이기 때문에 3차원 공간에서는 문자를 표현하는 영역으로 나타내어야만 한다. 이 사실과 개념의 모순성에 의해 의미추론부로부터 두 영역의 합집합은 문자를 표현하는 것이라는 개념이 추론된다. 마찬가지로 문자를 표현하는 영역과 사람 영역은 동일한 영역에 두 개 이상의 개념이 존재하므로 모순성이 발생한다. 이 문제 역시 다음과 같이 해결된다. 비록 2차원 영상 공간에서는 두 개념이 같은 영역에 존재하지만 3차원 공간에서는 문자를 표현하는 영역이 사람 영역을 분리한다. 이러한 3차원 상의 관계에 의해  $SR(\text{개념 문자를 표현하는 것, 개념 사람, 가리다})$ 이 추론되어 진다. 즉, 문자를 표현하는 영역이 사람 영역을 가리고 있다.

#### 4. 결론

본 논문에서는 물체들을 구성하는 개념들의 온톨로지를 이용하여 ontological 지식 기반 영상이해시스템에 대한 개략적인 구조와 개요를 제안하였고 문자인식과 사람인식에 대한 모의 실험함으로써 제안된 시스템의 타당성을 보였다. 제안된 시스템은 물체의 구성학적 분류와 의미학적 분류에 대한 온톨로지를 사용하여 물체들 간의 관계를 추론하여 영상을 인식하고 이해하였다.

추후 연구과제로는 보다 많은 분야의 개념들의 온톨로지 구축과 개념들 간의 의미 관계에 대한 지식베이스의 구축 및 확장이다. 또한 각 특징들의 일반적인 영상처리 기법에 대한 연구가 이루어져야 하겠다.

#### 5. 참고 문헌

[1] T. Matsuyama, and V. Hwang, SIGMA: A Knowledge-Based Aerial Image Understanding System, Plenum, New York, 1990.

[2] D. Crevier and R. Lepage, "Knowledge-Based Image Understanding Systems: A Survey," Computer Vision and Image Understanding, Vol. 67, No. 2,

pp. 161-185, 1997.

[3] B. A. Draper, A. R. Hanson, and E. M. Riseman, "Knowledge-directed vision: control, learning, and integration," Proceedings of the IEEE, Vol. 84, No. 11, pp. 1625-1637, 1996.

[4] K. Yanai, "An image understanding system for various images based on multi-agent architecture," In Proceedings of The Third International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'99), pp. 186-190, 1999.

[5] N. Maillot, M. Thonnat, and A. Boucher, "Towards Ontology Based Cognitive Vision," In Proceedings of The Third International Conference On Computer Vision Systems (ICVS 2003), LNCS 2626, pp.44-53, Austria, 2003.

[6] N. Maillot, M. Thonnat and A. Boucher, "A Visual Concept Ontology for Automatic Image Recognition," In Proceedings of The Second International Semantic Web Conference (ISWC 2003), Florida, 2003.