

문제영역에 관한 지식을 이용한 컴퓨터 시각 시스템

崔 瑩 一

(正 會 員)

崇實大學校 工科大學 電子計算學科 助教授

I. 서 론

컴퓨터 시각 시스템(computer vision system)이란 물체 또는 장면을 보고 이해 할 수 있는 시스템을 말하며 공장에서의 부품검사,^[1] 문서판독,^[2] X선 가습사진 해석^[3] 등 다방면에 응용되고 있다. 모든 장면을 보고 이해할 수 있는 컴퓨터 시각 시스템의 구현은 바람직 하지만 실제로는 불가능한 일이다. 따라서 주어진 문제 영역에 관한 지식과 영상 해석에 관한 지식을 이용하여 영상으로부터 실세계에 관한 유용한 정보를 추출해 내는 컴퓨터 시각 시스템의 구현이 일반적인 접근 방법이다. 본 논문에서는 문제 영역에 관한 지식 표현 방법과 지식 이용 방법들을 조사하고 이를 컴퓨터 시각 시스템에 적용하는 사례를 살펴본다.

일반적으로 컴퓨터 시각 시스템에 관한 연구의 목표는 다음의 두 가지로 요약될 수 있다. 첫째는 인간의 시각 시스템에 관한 연구이고, 둘째는 이와같은 연구를 바탕으로 보고 이해할 수 있는 컴퓨터 시각 시스템의 구현이다. 인간이 실세계로부터 취득하는 지식의 80% 이상은 시각 시스템에 의한 것으로, 그림 1 과 같이 시각정보 처리가 이루어 진다.^[4]

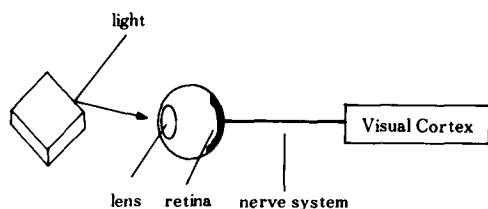


그림 1.

즉, 물체에 의하여 반사된 빛은 안구를 통하여 그 빛의 강도에 따라서 망막에 물체의 잔영을 형성하고 시신경을 거쳐 visual cortex에서 여러 종류의 특징을 추출하고 이들과 지식 베이스와의 비교 분석을 통하여 물체를 해석하게 된다. 이와같은 인간의 시각 시스템은 항상 목표 지향적인(goal-directed) 시스템이다. 다시 말해서 인간이 어떠한 장면에서 특정한 물체를 인식하는 과정은 그와 같은 물체가 존재하리라는 예상을 먼저하고 그 예상을 확인하는 과정을 통하여 이루어 진다. 예를 들어 사무실 문을 열고 들어 섰을때 파도가 출렁이는 바다가 나타난다면 우리는 깜짝 놀라게 될 것이다. 이는 책상, 의자, 서류함과 같은 사무실 장면을 미리 예상하고 있었는데 이와 부합되지 않는 장면이 나타났기 때문이다. 이와 같은 인간의 시각 시스템에 관한 이해를 통하여 주어진 문제 영역에 관한 사전 지식이 영상 해석을 위하여 중요한 역할을 하고 있다는 것을 알 수 있다.

컴퓨터 시각 시스템은 TV camera 또는 raster scanner와 같은 감지 장치(sensing device)를 이용하여 만들어진 실세계의 장면에 대한 디지털 농담 영상(digital gray-level image)을 입력으로 받아 영상 해석 작업을 통하여 주어진 문제 영역에 부합되는 정보를 기호형태(symbolic form)로 출력시키는 시스템을 말한다.^[5,6] 이를 위하여 컴퓨터 시각 시스템은 다음의 두 단계의 작업을 통하여 구현 되는데 첫째는 저수준 작업(low-level processing)이고, 둘째는 고수준 작업(high-level processing)이다. 저수준 작업에서는 주어진 농담 영상의 각 화소의 값을 인접 화소의 값과 비교하여 유사한 성질의 화소들의 집합으로 분할하는 작업(segmentation)으로서 분할된 각 영역(region)은 다음과 같은 여러 특징들로서 대표된

다. 질감(texture), 방향(orientation), 면적(area), 밀집도(compactness), 형태(shape), ... 고수준 작업은 분할된 각 영역을 그들의 특징 및 인접 영역과의 관계 그리고 주어진 문제 영역의 지식을 이용하여 실세계의 물체, 개념(concept)으로 해석하는 작업이다. 감지 장치에 의하여 얻어진 농담 영상은 잡음(noise)과 불확실한 정보(uncertain information)를 내포하게 되고, 이와같은 영상을 처리한 저수준 작업 결과 역시 모호성(ambiguity)을 내포하므로 고수준 작업 단계에서는 주어진 문제 영역의 지식을 이용한 가설 생성(hypothesis generation)과 검증(verification)을 통한 해석이 이루어져야 한다.

문제 영역에 관한 지식을 이용한 컴퓨터 시각 시스템을 구현하기 위하여는 다음과 같은 문제점이 대두된다.

(1) 어떠한 종류의 지식이 주어진 문제에 관련되며, 이들을 어떻게 표현할 것인가.

(2) 단계(1)에서 얻어진 지식들을 언제 어떻게 사용하여 영상 해석을 할 것인가.

본 논문의 II 장에서는 지식, 표현등과 같은 용어의 정의를 내리고 지식 표현 방법을 사례 연구를 통하여 살펴보고 III 장에서는 표현된 지식들을 어떻게 사용할 것인가를 결정짓는 시스템 제어 전략에 관하여 살펴본다.

II. 지식의 표현 (Representation of Knowledge)

지식의 표현은 인공 지능의 중요한 연구 분야의 하나로서 많은 표현 방법이 제안되어 왔다.^[7] 먼저 용어의 정의를 내린다. “지식”이란 다음 3 가지 형태로 나눌 수 있다.

(1) 실세계에 대한 관측의 묘사; 예를들어, 비행기는 동체와 왼쪽 날개 그리고 오른쪽 날개로 이루어진다.

(2) 주어진 문제를 이해하는 방법; 예를 들어, 주어진 영상에서 비행기의 존재를 확인하는 문제는 동체와 왼쪽 날개, 오른쪽 날개에 해당하는 부분을 찾아내고 이들 사이의 관계를 확인하는 과정이다.

(3) 문제해결 방법; 예를 들어, 비행기의 왼쪽 날개와 동체를 확인하기 위해서는 분할 영역 R1 이 분할 영역 R2와 인접해 있고 R2의 면적은 R1의 면적보다 상대적으로 크고 R1의 방향과 R2의 방향은 거의 직교하며 R1의 화소들의 centroid가 R2의 화소들의 centroid의 왼쪽에 있는 경우 R1은 왼쪽 날개에 해당하고 R2는 동체에 해당된다.

“표현”이란 어떠한 대상을 대표하는 부호(symbol)로서 그 대상의 구조적 특징을 보유하는 부호를 의미한다. 따라서 “지식의 표현”이란 자료 구조(data structure)와 해석 방법(interpretive procedure)의 복합체로서 정의될 수 있으며 적절한 제어 전략에 의하여 사용 되어지면 주어진 문제를 효율적으로 풀수 있어야 한다.

영상 해석을 위하여는 먼저 어떠한 종류의 지식이 필요한가를 결정하여야 한다. 일반적으로 다음과 같은 지식들이 고려된다. 물체들의 집합을 표현하는 지식, 예를 들어 일반적인 전기모터(cylinder와 축으로 이루어짐), 그 집합에 속하는 개개의 물체를 표현하는 지식, 예를 들어 밀반침이 있는 전기모터 또는 플랜지를 갖는 전기모터, 물체와 물체사이의 관계를 나타내는 지식, 예를 들어 전기모터와 선반은 서로 인접해 있고 전기모터가 선반의 윗쪽에 있다 실행 방법에 관한 지식, 예를 들어 영역 분할은 sobel edge operator를 사용하고 패턴매칭은 subgraph-isomorphism을 사용한다.

이처럼 어떠한 종류의 지식이 필요한가가 결정되면 적절한 표현 방법이 선택되어야 한다. 많은 표현 방법이 제안되어 왔으나 이들을 비교 분석하는 적절한 기준이 없으며 연구자의 취향에 따라 특정한 표현 방법을 선호하는 경향이다. 일반적으로 널리 사용되는 표현 방법은 의미망(semantic net), production rule, 프레임(frame) 그리고 서술 논리(predicate logic)이다.^[8]

1. 의미망 (Semantic Net)에 의한 지식표현

이것은 Quillan,^[9] Shapiro^[10]에 의하여 제안된 방법으로서 정보를 node와 이들을 연결하는 arc로 구성된 망(network) 구조로 나타낸다. 일반적으로 node는 개념 또는 물체를 나타내고 arc는 node와 node 사이의 관계를 나타내 준다. 의미망 표현 방법의 장점은 첫째, 이해가 쉽고 둘째, 한 정보를 여러 node가 공유할 수 있어 경제적인 표현 방법이며 셋째, 지식을 갱신하기가 용이하다. 반면에 단점으로는 세부적인 지식을 표현하기가 어렵고, 예외 사항을 다루기가 어렵다는 점이다.

의미망 구조는 Ballard^[11]에 의하여 부두가에 정박해 있는 배를 찾기 위하여, 그리고 X선 가슴 사진에서 종양을 확인하기 위하여 사용되었다. 예를 들어, “The bridge is at the intersection of road-57 and river-3 is near building-30”과 같은 지식은

그림 2의 의미망으로 나타내 진다.

그림 2에서 road, bldg, bridge 그리고 river는 물체들의 집합을 나타내고, intersec는 사건들의 집합을 나타내고, x와 xy는 예시화 된 물체 또는 사건을 나타내고, obj1과 obj2는 xy 라는 사건에 관련되어 있는 물체를 나타내며, there-exist-at 는 xy 사건이 발생한 위치를 나타낸다.

주어진 문제 영역의 물체 또는 장면에 관한 지식이 의미망 형태로 시스템의 지식베이스에 구현되면, 영상 해석 작업은 지식 베이스와 입력 영상과의 사상(mapping)을 통하여 이루어질 수 있다. 그림 3은 영상 해석 작업을 설명해 주고 있다.

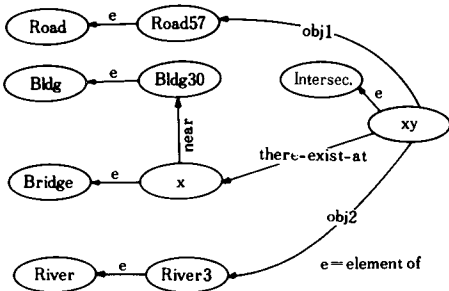


그림 2.

그림 3의 model에는 주어진 문제 영역에 관한 여러 형태의 지식이 의미망 형태로 구성되어 있다. 각 node는 관련된 물체, 개념 또는 상황이 만족하여야 할 특징들을 제약(constraints) 형태로 갖고 있으며 이와같은 제약의 강도를 나타내는 어떠한 척도를 이용하여 가격과 이윤(cost-and-benefit)을 계산하는 분석을 통하여 전체적인 매칭을 지휘 및 통제할 수 있다.^[12] 그림 3의 image data-structure는 입력 영상의 각 분할된 영역들을 여러 특징을 통하여 설명하고 있다. Image data-structure의 형성은 특정한 문제 영역에 무관하게 이루어 지거나 또는 주어진 문제 영역의 특성을 이용하여 이루어질 수 있다. Image data-structure와 model과의 사상(mapping)을 용이하게 하기 위하여 이들을 연결시켜주는 매개체로서 sketchmap을 사용한다. 즉, sketchmap의 node는 model의 node가 주어진 입력 영상에 맞도록 예시화 또는 구체화된 형태이다. 예를 들어, 그림 4는 SIGMA^[13]라는 시스템에서의 "house scene"의

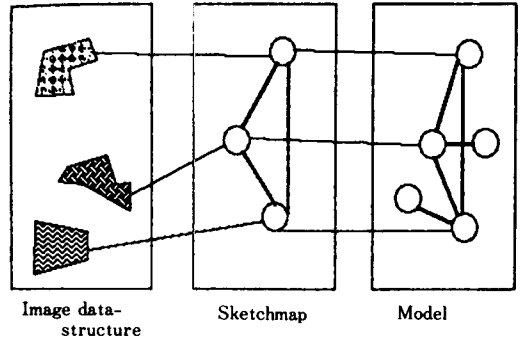


그림 3.

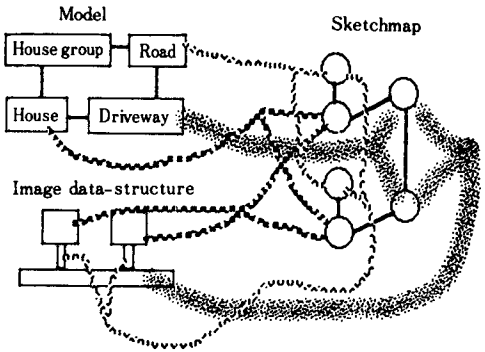


그림 4.

sketchmapping을 보여 주고 있다.

2. 프레임(Frame)에 의한 지식표현

이것은 Minsky^[14]에 의하여 제안된 방법으로서 주어진 문제영역에 관한 지식이 프레임이라 불리는 틀에 의하여 구성되어 있으며, 컴퓨터 시각 시스템 또는 자연어 처리와 같은 응용 분야에 널리 활용되고 있다. 프레임은 slot과 이에 해당되는 filler로서 형성되며, 일반적인 틀에 짜여진 상황 또는 물체를 나타내는데 사용되어질 수 있다. 프레임은 여러 형태의 지식을 포함할 수 있으며, 특히 문제를 푸는 절차적 지식(procedural knowledge)이 특정한 slot의 값으로 나타내질 수 있고 다른 프레임의 이름이 slot의 값으로 사용되어 계층 구조(hierarchical structure)의 지식 표현이 가능하다.

프레임 구조를 사용한 대표적인 컴퓨터 시각 시스템으로서 ACRONYM^[15]을 들 수 있다. ACRONYM에서의 지식 표현방법을 설명하기 위하여 먼저 일반 원통(generalized cylinder)에 의한 물체 묘사법을 살

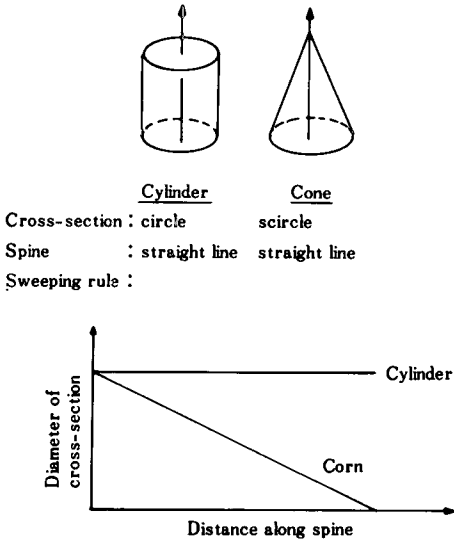


그림 5.

펴본다. 하나의 원통은 몇개의 일반 원통으로 나타내질 수 있고, 각 일반 원통은 매끄러운 곡선으로 이루어진 중심축(spine)과 단면(cross-section), 그리고 중심축을 따라서 단면의 연속적인 변화율(sweeping rule)로서 나타내질 수 있다. 그림5는 cylinder와 cone을 일반 원통으로 나타낸 예이다.

이와같은 일반 원통 묘사법을 이용하여 전기모터를 계층적 프레임 구조로 나타내면 다음과 같이 된다.

FRAME	SLOT	FILLER
Electric-motor	Class	Simple-cylinder
	Spine	Z0014
	Sweeping-rule	Constant
	Cross-section	Z0013
Z0014	Class	Spine
	Type	Straight
	Length	8.0
Z0013	Class	Cross-section
	Type	Circle
	Radius	2.5

영상 해석작업은 프레임에 부착된 정보를 이용하여 각 slot에 특정한 값을 할당하는 매칭 과정으로 다음과 같이 이루어진다.

(1) 주어진 문제 영역에 맞는 적절한 프레임을 선택한다.

(2) 만약 선택된 프레임이 상식치(dedault value)를 갖고 있지 않으면, 적절한 procedure를 사용하여 slot의 값을 정한다.

(3) 만약 slot의 값이 다른 프레임의 이름이면, 그에 해당하는 프레임으로 제어(control)가 옮겨진다.

선언적 구조(declarative structure)의 지식 표현은 표현이 명확하고 교정이 용이한 반면, 절차적 구조(procedural structure)에 의한 표현법은 제어(control)가 용이하고 직접적인 추론(direct inference)을 가능케 한다. Schema는 의미망 구조와 프레임 구조의 혼합형으로 선언적 지식과 절차적 지식을 모두 표현하며, 상속 추론(inherited inference)이 가능하다. Schema는 일반적으로 장면 또는 물체를 계층적 구조로 나타내며, 주어진 문제 영역에 관한 지식은 LTM(long-term memory)에 저장되어 있고 입력 영상에 관한 정보는 STM(short-term memory)에 저장되어 있다. 따라서 영상 해석작업은 LTM와 STM과의 사상(mapping) 과정으로 이해할 수 있다. 예를 들어 그림6은 VISION^[16] 시스템의 주택 장면을 나타내는 schema로서 영상 해석과정을 설명해 주고 있다.

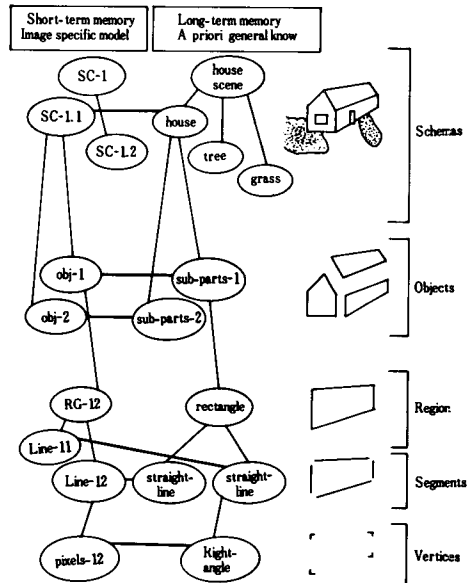


그림 6.

그림 6에서 region, segments, vertices 단계는 입력 영상을 묘사하는 1차적인 방법으로서 부호화된 출력(symbolic output)을 상위 단계인 object와 schema에 전달해 준다. 이와 같은 전달 과정은 여러 형태의 지식원(knowledge source)을 이용한 가설 생성(hypothesis generation)과 검증(verification)을 통하여 이루어진다. 예를 들어, camera와 물체와의 각도를 이용하여 물체의 특정한 표면(전면, 측면, ...)에 관한 정보를 유도할 수 있다.

3. 생성 규칙(Production Rule)에 의한 지식표현 이것은 Newell^[17]에 의해 인간의 인지를 모델링하기 위해 제안된 방법으로, Post의 기호대치 연산(symbolic replacement operation)을 위하여 생성 규칙이라는 문법을 사용한다. 생성규칙은 원인→결과(antecedent → consequence) 형태로 구성되어 LTM에 저장되어 있고, STM에 있는 데이터와 매치가 일어나면 적용되어(firing) STM의 데이터를 변화시킨다.

생성 규칙을 이용한 대표적인 컴퓨터 시각 시스템으로는 OPS5를 이용한 SPAM^[18]이라는 시스템을 들 수 있고, 여기서 사용되는 생성 규칙의 한 예는 다음과 같다.

```

; IF there is a task to interpret an unknown
; linear region
; and the region has ellipse length greater than
; the threshold for roads
; and there is no fragment interpretation for
; that region with a hypothesis of "road,"
; THEN make such a fragment.
(p interpret-as-road-1
 (context ↑task interpret-unknown ↑datum linear
  (rid))
 (constants ↑road-length (thresh))
 (region ↑region-id <rid> ↑ellipse-length >
  (thresh))
 - (fragment ↑region-id <rid> ↑hypothesis road)
 →
 (make fragment ↑hypothesis road...))
    
```

생성 시스템의 대표적인 특징은 LTM에 있는 생성 규칙들의 상호 연관성이 없어 보수 및 유지가 편리한 점이다. LTM에 있는 생성규칙들은 그림 7과 같은 세 가지 형태로 나누어질 수 있다. 첫째는 입력 영상의 1차적인 data인 region, line, area에 관

한 knowledge rules이다. 둘째는 knowledge rules의 적용 순서를 정하고 제어하기 위한 control rules이다. 셋째는 어떠한 control rule을 적용할 것인가를 결정하는 strategy rules이다.

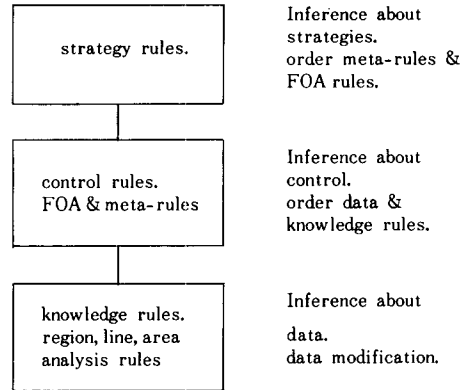


그림 7.

III. 제어 전략 (Control Strategy)

영상 해석은 영상분할(segmentation)과 분할된 영역해석(region interpretation)을 통하여 이루어 지는데, 제어 전략이란 영상해석을 위하여 주어진 문제 영역에 관한 지식들을 어떻게 사용할 것인가를 결정하는 방법이다. 일반적으로 domain-independent한 operator를 이용하여 초기의 영상분할이 이루어지며, 주어진 문제 영역안에서 잠정적인 해석을 내리고 이를 바탕으로 영상분할을 다시 구체화 시키는 주기적인 방법이 사용된다. 예를 들어 지식 베이스 안에 house와 road에 관한 정보가 저장되어 있고 이들 사이의 관계를 나타내는 정보 역시 저장되어 있다 하자. 만약 주어진 입력 영상의 특정한 영역이 house로 잠정적인 해석이 내려지면 house와 road 사이의 구조적인 관계를 나타내는 정보를 이용하여 road에 해당되는 영역을 추론할 수 있고, 이와 같은 추론을 바탕으로 초기의 영상 분할을 구체화 시킬 수 있다. 널리 통용되는 두 가지 형태의 접근 방법으로는 흑판 기법(blackboard model)을 이용한 제어 전략, 계획 수립(planning)에 의한 제어 전략 그리고 인간의 인식주기(human perceptual cycle)를 흉내낸 제어 전략을 들 수 있다.

먼저 data의 흐름에 따라 각 제어 전략을 다음과 같은 기준에 의하여 분류해 본다.

(1) 병렬 제어(parallel control) 대 순차적 제어(sequential control)

영상해석 과정은 원칙적으로 순차적이며 주기적이다. 즉, 가설을 생성하고 확인하는 과정을 통하여 해석이 이루어진다. 하지만, 여러개의 processor를 사용하여 협조적인 문제 해결 방법을 사용하면 어느 정도의 병렬성을 유지할 수 있다. 특히 저수준 처리 과정에서의 병렬 제어는 처리시간 단축을 위하여 효과적일 수 있다.

(2) 계층적(hierarchical) 제어 대 우발적(heterarchical) 제어

계층적 제어는 상향식 제어(bottom-up)와 하향식 제어(top-down) 방식으로 나누어 생각할 수 있다. 상향식 제어방식은 data-driven 제어방식이라 불리우며, 주어진 영상에서 특징을 추출하고 이를 부호로서 표현하여 주어진 모델과 비교를 통하여 결론에 도달하는 제어 방식이다. 이러한 제어 방식은 대상 영역이 단순하고 주어진 입력영상이 잡음(noise)을 내포하지 않는 선명한 영상인 경우에 효과적이며 산업용 컴퓨터 시각 시스템에 많이 사용된다. 하향식 제어 방식은 goal-directed 제어 방식이라 불리우며, 예측(prediction)에 의하여 예상되는 시험 모델을 제시하고 입력 영상의 각 표현 단계에서 대상의 모델과 일치하는 특성이 존재하는지를 검증하는 제어 방식이다. 이러한 제어 방법은 모호성을 내포하는 영상인 경우에 효과적이며 자연 장면해석(natural scene understanding)에 널리 사용되고 있다.

계층적 제어 방식은 연산 과정의 변화에 대응하는 유연성이 부족한 단점이 있다. 이와같은 단점을 고려하여 연산 순서를 미리 정해놓지 않는 우발적 제어 방식은, 가용한 여러 형태의 지식원 중에서 당면한 문제에 가장 적합한 지식원을 선택 적용시켜 상황을 변화시키며 변화된 상황에 대하여 같은 작업이 반복된다. 즉, 당면한 상황에 의존하는 situation-driven 제어 방식으로서, 어떠한 지식원을 선택할 것인가를 결정하는 기준에 따라 또는 주어진 상황을 해석하는 기준에 따라 여러 형태의 제어 방법을 생각할 수 있다.

1. 후판 기법에 의한 제어

문제 해결을 위한 효과적인 탐색 방법은 부분해를 먼저 구하고, 이들 부분해를 결합 확장시켜 최종적

인 해를 구하는 접근 방법이다. 예를 들어, HEAR-SAY 시스템^[19]은 speech를 이해하는 문제를 부분해와 완전해로 이루어진 문제 공간(problem space)을 탐색(search)하는 과정으로서 비유하여 해석하고 있다. 즉, 어떠한 메시지를 이해하기 위하여는 각각의 단어를 인식하는 과정을 통하여 전체적인 문맥과 문법을 확인하며, 반면에 문법과 문맥을 확인하는 과정은 각각의 단어를 인식하는 과정을 도와줄 수 있다.

후판기법^[20,21]을 이용하는 시스템의 주요 구성원으로는

(1) 지식원(knowledge sources)

주어진 문제영역에 관한 지식으로서 procedure, rule 또는 logical assertion등으로 표현될 수 있다. 각 지식원은 후판 또는 제어 자료구조(control data-structure)를 갱신시킬 수 있다. 각 지식원은 condition-part와 actionpart로 나누어지며, condition-part는 지식원의 action-part가 적용될 후판의 상황을 설정해 주는 가설들로 이루어지며 action-part는 후판의 상황이 condition-part와 일치할때 후판의 상황을 어떻게 변화시켜 줄 것인가를 결정해 준다.

(2) 후판

후판은 지식원이 필요로 하는 또는 지식원에 의하여 생성된 정보를 기록하고 있고, 다른 지식원과의 대화는 이 후판을 통하여 이루어진다. 입력된 영상의 각 물체에 대한 부분해석, 다른 인접 물체와의 상관관계 또는 subpart relation이 계층 구조로 나타나지게 된다. 즉, 하나의 후판은 여러개의 panel로 나누어지고 각각의 panel은 인접 panel과의 상관 관계를 나타내는 link로 연결 되어진다.

(3) 제어

현재의 후판의 상황을 분석하고 적절한 지식원을 선택 적용하여 후판의 상황을 변화시켜 나가는 과정으로서 scheduling queue, FOA(focus of attention)과 같은 여러 module을 이용하여 실행된다. FOA은 다음에 처리될 작업을 지정하는 module로서, 특정한 지식원 또는 후판위에 특정한 물체를 선택하거나 특정한 물체와 이에 적용될 지식원을 동시에 결정지을 수 있다. Scheduling queue에는 현재의 후판의 상황에 관련된 지식원들이 있다. Scheduler는 FOA이 지식원인 경우 현재의 후판의 상황에 적절한 지식원을 선택하는데 const-benefit analysis와 같은 기준을 사용할 수 있다. 만약 scheduler가 지식원의 condition part를 선택하게 되면 이에 상응하는 action part가 예시화 되어 scheduling queue

에 들어가게 되고, 만약 scheduler가 지식원의 action part를 선택하게 되면 후관의 상황이 적절히 갱신된다.

그림 8은 Levine⁽²²⁾의 컴퓨터 시각 시스템의 제어 방법을 설명해 주고 있다.

이 그림에서 low-level processor는 입력 영상을 분할시키는 작업을 하고, feature analyzer는 분할된 영상의 각 영역의 특징을 추출한다. Hypothesis initializer는 후관 위에 묘사된 각 영역의 특징들과 지식원들을 이용하여 가설을 형성하며, hypothesis verifier는 인접 영역과의 관계와 관련되는 지식원들을 이용하여 생성된 가설을 검증 확인한다. 즉, 가설을 생성하는 단계는 하향식 제어 방식으로 문제 해결을 위한 탐색 공간을 제한하는 역할을 하며, 가설을 검증하는 단계는 상향식 제어 방식으로 제한된 공간내에서의 탐색에 해당된다. 검증 및 확인된 가설은 그에 대한 신뢰도를 나타내는 confidence value를 갖게 되며, 이와같은 신뢰도는 인접 가설과의 상관 관계를 통하여 갱신된다. FOA는 현재의 상황을 분석하여 분할된 영상의 여러 영역 중에서 어떤 영역을 선택하여 처리할 것인가를 결정하여 준다. 즉, 각 영역의 잠정적인 해석에 대한 confidence value를 조사하여, 가장 신뢰도가 높은 영역을 기준으로 하여 다음에 조사할 영역이 결정되어진다. Scheduler는 현재의 상황을 분석하여 어떠한 processor를 적용할 것인가를 선정한다. 예를 들어, 분할된 특정한 영역에 대하여 인접 영역과의 관계를 고려하여 가설을 형성할 것인지 또는 그 영역의 특성을 조사하여 생성된 가설들의 신뢰도를 확인하는 검증할 것인지에 cost-benefit analysis를 통하여 결정하게 된다.

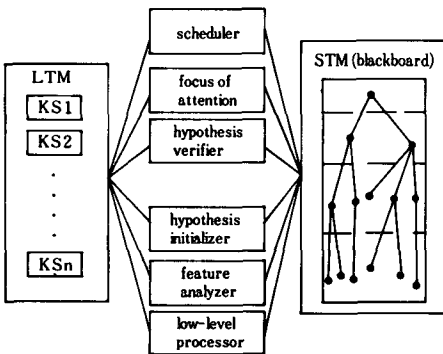


그림 8.

인접 영역과의 관계와 관련되는 지식원들을 이용하여 생성된 가설을 검증 확인하는 과정은 패턴 매칭(pattern matching) 과정으로 해석될 수 있으며, 일반적으로 graph isomorphism⁽²³⁾을 통하여 이루어진다. 즉, 생성된 가설들을 node (V)에 비유하고 인접 가설들 사이의 정보를 arc (E)에 비유하여 $G = (V, E)$ graph로 나타내고, 마찬가지로 입력 영상의 분할된 각 영역을 node (V1)에 비유하고 이들 사이의 관계를 arc (E1)에 비유하여 $G1 = (V1, E1)$ graph로 나타낸다. 따라서 G1 graph가 G graph와 isomorphic mapping이 이루어지는가를 확인하는 과정이 곧 가설 검증 과정에 해당된다. 하지만 잡음과 모호성을 내포하는 입력 영상에 대하여 정확한 isomorphism을 기대하기는 어려우며, 따라서 유사성의 정도를 나타내는 matching metric을 사용하게 된다. 일반적으로 사용되는 matching metric으로 다음과 같은 척도를 고려할 수 있다.⁽²⁴⁾

$$\begin{aligned}
 \text{COST} = & \sum_{e \in (V \cap V1)} d(e, F(e)) \\
 & + \sum_{e \in (E \cap E1)} d(e, F(e)) \\
 & + \sum_{e \in (G - G1)} \text{MissingCost}(e)
 \end{aligned}$$

여기서 F(.)는 G graph의 node(또는 arc)와 G1 graph의 node(또는 arc) 사이의 mapping을 나타내며, d(.,.)는 두 arguments 사이의 차이를 나타내는 척도이고, MissingCost(.)는 G graph와 G1 graph 사이의 구조적인 차이를 나타낸다.

2. 계획 수립(Planning)에 의한 제어 전략

Plan이란 문제 해결을 위하여 사용될 operator들의 행렬로서, 일반적으로 planner는 주어진 문제 영역의 coarse-to-fine analysis 또는 cost-benefit analysis를 통하여 수행될 operator들의 종류와 이들의 적용 순서를 결정한다. 즉, 가용한 operator들을 AND/OR graph 또는 사다리 모양의 트리(ladder-like decision tree) 형태로 구성하여 역방향 제어를 통하여 주어진 문제를 해결하려는 방법이다. 예를 들어, Ballard와 Sklansky⁽²⁵⁾에 의하여 제안된 X선 가슴 사진에서 종양을 찾는 방법을 살펴본다. 이들은 그림 9와 같이 coarse-to-fine 순서로 가용한 operator들을 사다리 형태의 decision tree로 구성하여 종양의 유무를 확인하려고 시도했다.

이 그림에서 1번 operator는 가슴 사진에서 lung에 해당되는 영역을 확인하고, 2번 operator는 lung 영역 안에서 종양이라고 여겨지는 작은 혹(nodule)

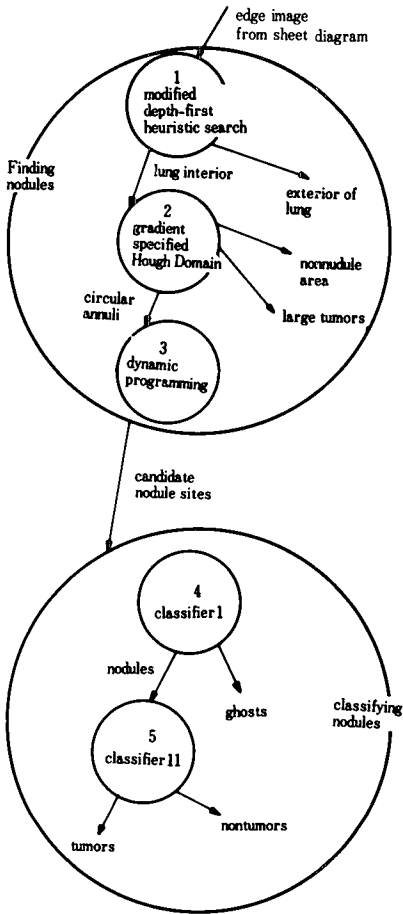


그림 9.

의 위치를 hough domain에서 확인하고, 3번 operator는 전 단계에서 확인된 예상되는 nodule들이 closed boundary를 형성하는지를 확인하여 closed boundary를 형성하지 못하는 경우는 제외 시키며, 4번 operator는 closed boundary를 형성하는 circular shape 중에서 nodule이 아니라고 여겨지는 것은 제외시키고, 5번 operator는 예상되는 nodule 중에서 종양(tumor)을 확인한다.

3. 인간의 인식 주기에 의한 제어

Neisser^[25]는 인간의 인식 시스템을 다음과 같이 설명한다. “인간의 시각 시스템은 먼저 예상되는 장면에 관한 schema를 형성시키고, 형성된 schema를 이용하여 관측의 범위 및 방향을 설정한 후 관측된 자료를 이용하여 초기의 schema를 변화시키는 과정을 통하여 이루어 진다.” 이와 같은 이론을 배경으로 한 제어 방법이 인식 주기에 의한 제어 방법으로 그림10에 의하여 설명 되어질 수 있다.

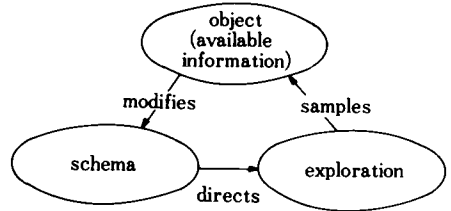


그림10.

먼저 지식 베이스에 있는 여러 schema 중에서 적절한 schema를 선정한 후, 선정된 schema에 의하여 입력 영상에 대한 탐색 방향이 결정되고 탐색을 통하여 얻어진 자료를 근거로 하여 지식 베이스의 schema를 변화시키고 변화된 schema에 의하여 다음에 적용될 schema를 결정하는 주기적인 방법을 통하여 영상 해석이 이루어 지게 된다. Neisser의 인식 주기 이론의 근간은 인간이 실세계를 이해하는 결정적인 요소는 동작 이해(motion understanding)라는 점이다. 즉, 실세계는 정적인 현상이 아니라 항상 변화하는 동적인 현상이라는 사실이다. 따라서 지식 베이스 내에 있는 schema는 연속적인 변화를 일으켜 실세계의 변화에 부합 되도록 갱신되어야 한다는 점이다. ☺

◆ 用語 解 說 ◆

PL/1

수치계산, 논리연산 및 사무 데이터 처리를 하기 위한 프로그래밍 언어의 하나
PL/1은 Programming Language/1의 略語