

화상이해를 위한 지식의 이용

南宮在贊
(光云大學 教授)

■ 차례 ■

- 1. 화상이해란?
- 2. 화상이해의 일반적인 방법
- 3. 화상데이터의 구조화
- 4. 화상이해에 있어서의 지식표현
 - 4.1 3차원 물체의 형상표현
 - 4.2 지식의 구조적 표현
 - 4.3 화상처리 기술에 관한 지식의 이용
- 5. 인식 과정의 제어
 - 5.1 가설검정 (Hypothesis-and Test) 방식
 - 5.2 이완법
 - 5.3 신뢰성의 계산
- 6. 결론
참고문헌

1 화상 이해란?

화상이해 (Image Understanding)란 낱말이 이제 와서는 널리 알려졌고 그에 관한 연구도 활발히 진행되고 있다.

일반적으로 화상 이해란 단순히 패턴을 분류 하는 것으로 그치지 않고 화상으로서 주어진 데이터를 해석해서 대상 세계에 관한 지식을 기초로 추론을 행함으로써 그 화상이 나타내는 현실 세계의 정경에 관한 기술을 작성하는 것이다.

정경에 대한 기술이 이루어지면 그것을 이용해서 정경에 관한 질문에 답한다든지 대상물에 물질적인 힘을 가해 움직임이 가능하게 한다든지 하는 의미로 계산기가 그 화상을 이해하는데에는 지식이 필요하다. 지식의 양이 많고 그것들을 묻는 추론이 유연할 수록 이해의 정도는 깊고 외계와의 상호작용도 풍부해진다.

통상 인간은 개, 고양이, 집, 비행기라는 개념을 요소로 해서 지식을 구조화 해서 쌓아놓고 있다고 생각되어진다. 화상이해의 경우 이러한

개념과 관측된 2차원 화상간에는 정보의 질적 차이가 크다.

화상 이해의 목적은 이 정보 차이를 연산 추론에 의해 메꾸어서 개념과 화상데이터간의 대응을 행하는 일이라 할 수 있다.

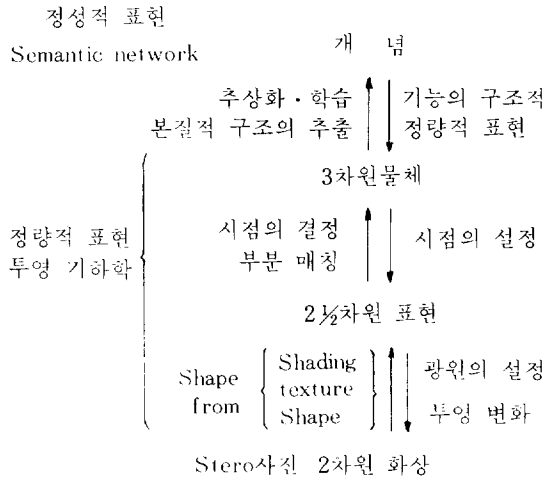


그림 1. 화상이해에서의 정보레벨

특히 최근에 와서는 그림 1에 나타낸 것처럼 양자간에 3 차원 물체, $2\frac{1}{2}$ 차원이란 중간단계를 만들어 각 인접 level간에 정보변화, 처리 방식을 명확히 하기 위한 연구가 활발하게 진행되고 있다. 여기서는 화상이해를 개념 차원에서 이해를 포함한 문제를 고려해서 개개의 화상해석 기술 뿐 아니라 시스템적인 관점에서부터 화상이해 연구의 현황 및 문제점을 고찰한다.

2 화상이해의 일반적 방법

인간 혹은 환경에의 움직임에 포함한 화상이해 시스템의 구성은 일반적으로 그림 2와 같다. 화상이해의 과정은 크게 나눠 두개의 과정으로 되어있다.

(a) Bottom up해석 : 입력화상으로 부터 edge 및 영역이라는 화상특징을 추출해서 화상구조기술의 요소로 한다. 그 가운데 화상특징의 속성 및 그들간의 기하학적 관계를 구하는 구조화를 행한다. 그 결과 입력 화상이 갖는 정보가 symbolic한 구조기술로 표현돼 model과의 조화가 얻어진다.

(b) Top down해석 : bottom up해석에 의해 인식 대상의 구체적 모델이 예측할수있는 경우 그 모델의 기술에 따른 특징추출, 구조화, 조회의 process를 제어해서 아직 해석이 되어있지 않는 부분의 해석을 한다. 그 결과 새로운 화상특징의 검출이 있는 화상의 구조기술 수정, 조회 범위의 제한, 새로운 대상물의 인식등이 가능하게 된다.

화상 이해에 있어서 중요한 문제로서는

(1) 관측한 화상 데이터를 다각적으로 해석해서 충분한 정보를 갖는 구조기술을 작성하는 일

(2) 3 차원 물체 및 운동 물체등 시점 및 시간에 의해 보는 방법이 변화하는 대상의 모델화 또한 개념과의 상호관계, 유사성등에 관한 지식을 잘 기술하는 일.

(3) 두개의 기술간의 유연한 matching 및 상호작용(해석의 제어)방식을 고려하는 일이다.

또한 최근 주목을 받고있는 지식공학 관점으로 부터는

(4) 단순히 해석 결과를 나타낼 뿐 아니라 왜 이런 결과에 다다르게 되었는가를 설명할 수 있

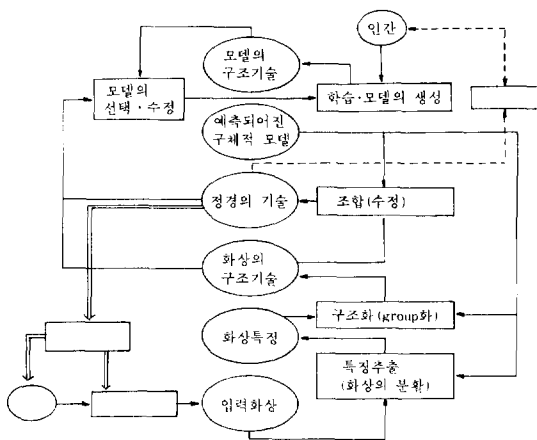


그림 2. 화상이해계의 블록도

는 것등이 필요하다.

해석에 이용된 추론 과정의 이력을 명시적인 형태로 시스템 내에 비축해 두는 일은 인간에 대한 설명기능을 위한 것 뿐만아니라 시스템 자신이 backtracking에 의해 자기 판단 및 수정을 행할 수 있는 것도 중요한 문제이다.

3 화상데이터의 구조화

화상의 구조 기술의 요소로는 일반적으로 edge 및 영역이 사용되고 그것들을 화상으로부터 추출하기 위해 화상처리 수법이 이때까지 많이 시도되었다. 그 결과 복잡한 화상으로부터 의미 있는 edge 및 영역을 완전하게 추출 하는데에는 그런 용이한 문제가 아니라는 것이 명확하게 됐다. 그러므로 일시적으로 화상 특징의 추출단계의 대상 세계에 관한 지식을 적극적으로 도입하려고 하는 움직임이 활발해졌다.

그러나 이러한 방법은 어떤 한정된 대상의 전용 해석 시스템을 작성하는 데에는 좋지만 화상이해의 일반적 도구로서는 충분하다고는 말할 수 없다.^{1),2)}

현재로서는 해석의 제일단계로 화상데이터의 구조화에 있다. 특정한 대상에 관한 지식에 의존하지 않는 범용적 수법으로 화상을 해석해서 화상이 갖는 정보를 될 수 있는 한 충실하게 표

현하려는 것이 일반적이다. 이 생각 방법은 Marr의 primal sketch³⁾에 있어 명확하게 시사되고 있다. 특히 MIT에서는 명도 경계 추출법으로서 zero crossing⁴⁾이 대표되는 것처럼 인간의 시각 기능과의 대응관계에 기초된 화상 특징의 추출, 구조화에 관한 연구가 활발하게 행해지고 있다. 이에 대해 Haralick⁵⁾은 국소영역에 있어서 명도변환의 pattern을 좌표치(x, y)의 다항식 함수를 이용해서 근사화를 행하는 것에 의해 각종의 화상 특징(각, edge, 꼬리, 계곡, 등곡선점)이 정도 높게 안정하게 추출될 수 있음을 보였고 함수 근사에 기초한 명도 패턴 해석의 유효성을 주장했다.

화상해석의 제 일단계에서 얻어진 edge영역은 인식대상의 모델의 조회를 취한 기술 level이 너무 낮은 경우가 많다. 이 때문에 몇개의 edge 영역은 그룹화 해서 화상의 구조기술 단위를 고도화 할 필요가 있다.

group화의 방법으로서 소위 Gestalt법칙을 기초로한 각종의 알고리즘이 계산된 것이 이제는 하나로 되어 있지 않다. 이에 대해 실용이라는 관점으로는 인식 대상의 구조를 한정하는 것에 의해 그것에 맞는 화상특징 group화를 행하는 방법이 자주 사용된다.

이의 예로서는 ACRONYM⁶⁾에 있어서 리본검출, 기계 부품의 인식에 있어서 타원 평행직선의 추출 혹은 거리 해석에 있어 평면 및 곡선의 인식등이다.⁷⁾ Ballard⁸⁾는 이런 고차의 화상특징 추출을 위해 일반적 방법으로서 일반화 Hough 변환을 제안했다.

4] 화상이해에 있어서 지식 표현

앞에서도 기술했던 것처럼 대상물을 이해하는데 지식이 필요하고 지식의 양에 따라 이해level이 다르다.

화상 이해의 경우 대상의 형상에 관한 지식이 특히 중요한 문제로 되고 그 표현법에 관해서 많은 방법이 제안되어 있다.⁹⁾

공업 부품 인식에서처럼 인식 대상의 형상이 미리 고정되어 있는 경우는 개개 부품의 3차

원 형상을 어떻게 기술 표현을 하는가가 문제로 되고 CAD에 있어 물체 형상표현법과 공통되는 부분이 많다. 그러나 화상이해의 목적은 2차원 물체를 이용해서 3차원 물체를 인식하는 일도 있고 시점의 변화 감취짐에 의해 화상평면 상에서 대상의 형상 변환에 어떻게 대처해서 model과 data의 조회를 효율 좋게 행하는 등이 형상 표현법을 생각하는 것 외에 큰 문제이다.

한편 이해의 level로 부터 말하면 이러한 물체 검출적인 해석은 더욱 level이 낮은 것이라고 말할 수 있다.

예를 들면 B747 여객기의 검출과 비행기라는 개념의 인식간에는 큰 차가 있다. 전자의 경우는 그의 기하학적 형상이 명확하게 정의 돼 있고 그 정보를 이용하면 된다. 그러나 후자의 경우에는 각각각색인 비행기 형상의 간단한 표현법, 혹은 다른 개념(예를 들면 비행장, 활주로등)과의 관계 기술이 필요하게 되고 대량의 서로 다른 지식의 구조화, 유연한 추론이 불가결하다.

4.1 3차원 물체의 형상표현

화상이해에 있어 3차원 물체의 표현법에서는 일반원통(원추)가 종래부터 잘 사용되어 왔다. 일반 원통은 (1)축이 되는 3차원 공간에서의 곡선(spine) (2)축상의 각점에 있어 단면용 축으로 이어서 이동 시킬때의 형상변화를 정하는 규칙에 의해 정의된다.

복잡한 3차원 물체의 경우는 몇개의 일반원통을 조합함으로써 표현한다. 일반 원통은 그 정의로부터 체적 지향형의 표현법으로 되고 화상으로부터 얻어진 물체표면이 형상 데이터(2 1/2차원표현)과의 조회를 직접하는 것은 어렵다.

이때문에 ACRONYM에서는 일반원통의 윤곽 형상을 이용해서 화상특징으로서 Ribbon과의 대응을 갖는것을 행하고 있다. 또한 Shafer은 종래 불명확하던 일반원통의 정의 및 그의 기하학적 성질에 관해서 상세한 검토를 하고있다. 물체 표면 형상에 중점을 둔 표현법으로서는 Horn, Smith에 의한 Extended Gaussian Image 및 Oshima 등에 의한 평면, 2차 곡면 batch에 의한 것들이다. 전자의 방법은 그림 3에 나타낸 것처럼 물

체 표면의 각점에 있어서 법선 vector를 그 시점이 원점에 일치하도록 이동해서 동일 방향의 vector는 그들을 바르게 맞추는 것이다.

이 방법은 2½차원 표현과 대응이 좋고 간결하지만 복잡한 물체의 경우 표현의 일의성 조작성에 문제가 생긴다.

화상 이해에서는 2차원 화상과 3차원 물체와의 조화를 하는 일이 큰 목적이다. 따라서 그 조화 과정이 행하기 쉽게 하기 위하여 3차원 물체의 형상을 우선 “compile”해서 시점의 이동에 의한 2차원 형상의 변화로서 3차원 형상을 표현

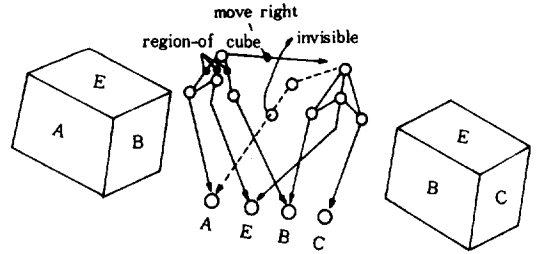
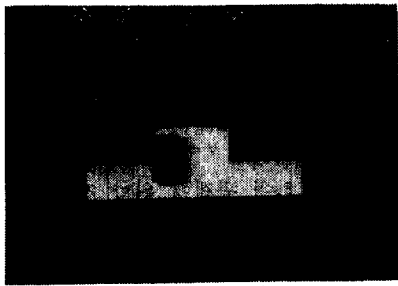


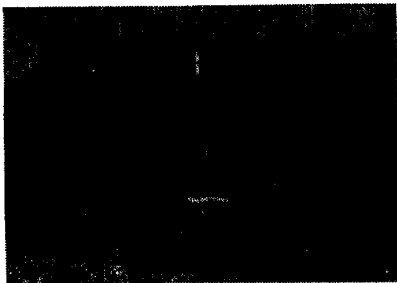
그림 4. 시점의 이동에 의한 이차원 형상의 변화



(a) 3차원 물체



(b) 접평면의 방향을 나타내는 단위벡터



(c) Extended Gaussian Image

그림3 . Extended Gaussian Image

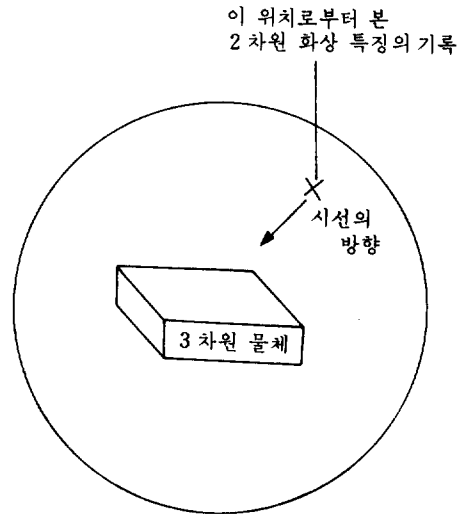


그림 5. Property Spheres에 의한 3차원 물체의 형상표현

하려고 하는 방법이다 (그림 4). 이 생각은 최초 Minsky¹⁰⁾ 의해 제안돼 ARGOS¹¹⁾에서 저고도 항공사진의 이해에 이용됐다. 최근에는 3차원 물체의 기하학적 형상분석에 기초한 다면체 뿐만 아니라 곡면 물체의 형상 표현에도 이러한 생각이 적용하려고 하는 aspect graph¹²⁾가 제안되고 있다. 또한 그림 5에 나타난 것처럼 물체를 구의 중심에 두고 구면상의 각 점으로부터 물체를 볼때의 2차원의 형상특징 (moment-용)을그림의 값으로 하는 property sphere¹³⁾도 제안되어 있다.

이러한 aspect graph 및 property sphere를 이용한 시점의 위치를 결정하는 문제는 앞으로의 과제이다.

4.2 지식의 구조적 표현

일반적으로 복잡한 개념간의 관계를 표현하는 경우 그 기본구조로서 상위-하위관계 (specialization), 부분-전체관계 (part-whole)에 기초로한 계층구조가 잘 이용된다. 이들의 관계의 의미에 관해서는 인공지능의 분야에서 상세한 검토가 되어있다. 이를 테면 B747은 비행기라는 개념의 하위개념에서 비행기에 관한 속성 및 제약조건은 그대로 B747에도 적용된다는 것이다.

이 속성의 후계기능에 의해 지식의 중복이 대폭으로 절감되어 간결한 표현이 실현될 수 있다. 또한 물체의 형상이 큰 의미를 갖는 화상이해에서는 부분-전체 관계에 의한 복잡한 물체의 구조표현은 지식표현이 표준적 수법으로 되어있다.

계산기 내에 있어서 이러한 지식의 표현법으로서 frame^{(9),(14),(15),(16)}이 잘 이용된다. frame은 지식을 정리한 단위로 표현하는 것으로 소위 production법칙에의 지식표현에 비해 전망이 좋은 표현법으로 되어있다. 그림 6은 ACRONYM에 있어 전기 motor의 형상표현을 frame을 이용해서 기술한 것이다. 각 frame(그림 중에서는 NODE)에는 다수의 slot가 포함돼 그 frame의 속성 및 다른 frame과의 관계 혹은 attached procedure라 불리는 수속을 쓰는 것이 허락되어 있다. 통상

```

Node :      GENRIC_ELECTRIC_MOTOR_CONE
CLASS :      SIMPLE_CONE
SPINE :      Z0014
SWEEPING_RULE : CONSTANT_SWEEPING_RULE
CROSS_SECTION : Z0013

Node : Z0014
CLASS :      SPINE
TYPE :      STRAIGHT
LENGTH :     MOTOR_LENGTH

Node :      CONSTANT_SWEEPING_RULE
CLASS :      SWEEPING_RULE
TYPE :      CONSTANT

Node : Z0013
CLASS :      CROSS_SECTION
TYPE :      CIRCLE
RADIUS :     MOTOR_RADIUS
    
```

그림 6. Frame에 의한 모니터 형상 표현

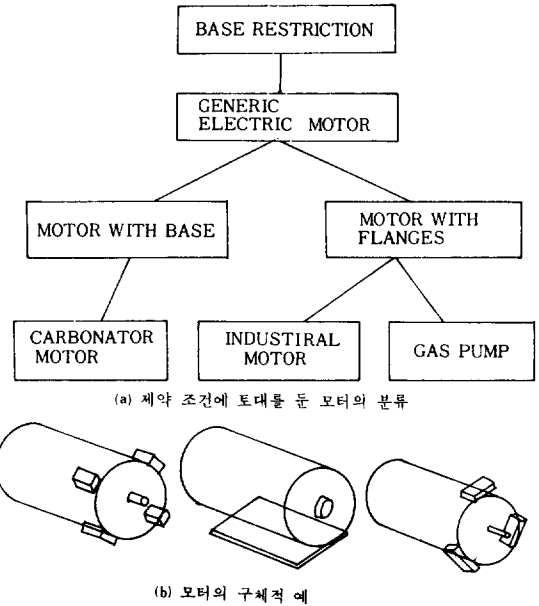


그림 7. 상위 하위 관계에 의한 지식의 계층구조

frame 표현에는 slot 값으로서 수치와 이름등의 상수가 대입되는데 ACRONYM에서는 MOTER-RADIUS에서 본 모양으로 변수가 허용되고 이러한 변수에 대한 제약조건 (예를 들면 $9 \leq \text{MOTOR-RADIUS} \leq 20$)에 의해서 개념간의 계층구조와 속성간의 관계가 정의된다.

그림 7은 이러한 제약조건에 대한 모타의 분류(상위-하위 개념)을 나타낸 계층구조로 각 node에는 그 class를 규정하기 위한 제약조건이 기술되고 상위의 class에 대한 제약조건은 그대로 하위 class에서 계승된다.

화상이해에서 이용되는 계층구조로서는 위에 기술한 외에도 해상도 변화에 의한 물체, 정경 기술의 변화를 표현하는 것등이 있다. 이대표적인 예로서는 Nishihara, Marr에 의한 원통을 이용한 3차원 물체의 계층표현(그림 8), Rise man 등에 의한 VISIONS시스템,¹⁷⁾ Rosenthal등에 의한 항공사진해석 시스템이 있다. 이러한 계층표현은 화상처리에서 피라미드 구조와¹⁸⁾ 기본적으로는 동일하게 생각하는 방법으로 후자의 2개 시스템에서는 계층모델과 피라미드 구조와의 대응 관계에 바탕을 둔 해석을 행하고 있다.

대국적 분류로 부터 상세한 해석이 라는 고

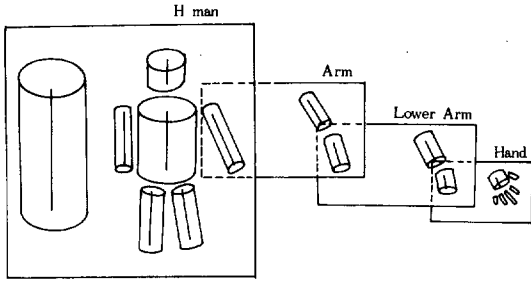


그림 8. 일반 원통에 의한 물체의 계층적 표현

러 방법은 인식과정에서 주의의 집중(focus of attention)을 실현 하는 방법으로서 다른 시스템에서도 이용되고 있다.¹⁹⁾ 해상도 축에 바탕을 둔 계층적 표현법에서의 문제는, 해상도 변화를 모델의 구조 변화로서 표현하기 위한 구체적 방법이 확립되어 있지 않은 점으로 해상도가 다른 모델을 어떤 모양으로 조합시켜 해석에 이용하는 가는 명확하지 않다.

일반적으로 개념간의 관계를 나타내는 것으로는 위에 기술한 것같은 계층구조에다 다른 개념간의 횡방향 관계기술이 필요로 한다. 통상 어떤 기하학적, 의미적 관계가 필요한가, 대상세계, 기술 레벨의 깊이에 의해서 달라지는가 등이 인식, 이해라는 관점으로 부터 중요한 일반적 관계로서 유사-차이 관계(similarity)¹⁰⁾가 있다.(그림 9). 일반적으로 2개의 다른 대상간에는 공통하는 속성, 구조 또는 양자를 구별하기 위한 특징이 존재한다.

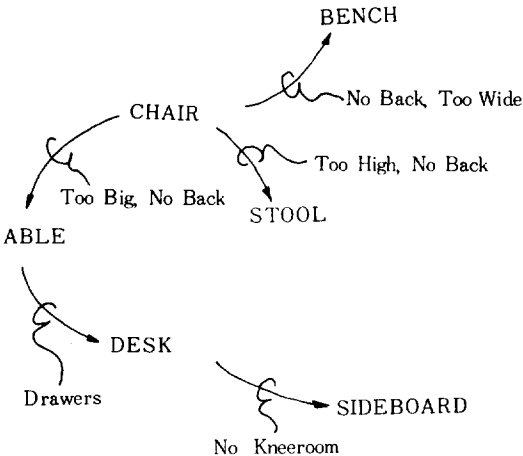


그림 9. 유사-차이관계에 의한 개념간의 관계표현

이러한 유사성 및 차이에 관한 기술을 이용하는 것에 의해 지식표현이 간결하게 되고 많은대상을 효율 좋게 인식하는 것이 가능하게 된다. Tsotsos¹⁵⁾는 심장의 좌심실벽의 운동 이해를 행하는 시스템에서 이 유사-차이 관계를 이용하는 것에 의해 운동을 나타내는 개념간의 식별이 효율 좋게 행해지는 것을 나타내었다.

4.3 화상처리 기술에 관한 지식의 이용

화상 이해에서 항상 문제로 되는 것으로 화상의 세그멘테이션이 있다. 3에서도 서술한 것처럼 화상특징의 추출에 관해서는 지금까지 많은 알고리즘이 개발되었는데 복잡한 화상으로 부터 의미가 있는 화상특징을 완전히 추출하는 것은 대단히 어렵다.

이 문제에 대한 최근의 새로운 움직임으로써 화상처리 기술 자체에 관한 지식을 화상해석에 이용하는 것이 있다. 즉 지금까지의 화상처리에 관한 연구결과 새로운 화상처리 알고리즘의 개발과 동시에 그러한 특성(효율, 유효하게 동작하기 위한 조건등)에 관한 비교점도도 상당히 되어졌다. 여기서 그러한 알고리즘을 화상해석을 위한 도구로 생각해 그러한 특성에 관한 지식에 바탕해서 대상데이터, 목적에 맞게 도구를 잘 사용한다는 고려 방법이다. Selfridge²⁰⁾는 이러한 생각을 항공사진의 해석에 적용하고 그림 10에 나타낸 것처럼 시스템을 개발했다. 이 시스템에서는 Appearance Model Expert로부터 주어진 인식대상에 관한 기술에 따라 (1)Operator Expert는 화상특징 추출을 위해 적절한 화상처리

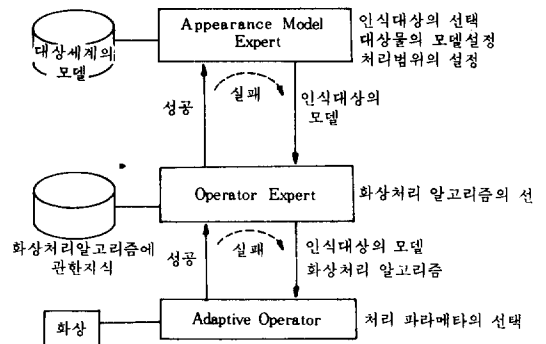


그림 10. 화상처리 알고리즘에 관한 지식의 이용.

알고리즘을 선택하고(2) Adaptive Operator는 대상화상의 성질에 맞는 처리 파라미터를 선택한다. 이 시스템에서는 그만큼 고도지식의 구조화와 깊은 추론이 행해지고 있다는 의미는 아니지만 인식대상에 관한 지식에 화상처리 알고리즘 자체에 관한 지식을 이용하는 것에 의해 유연한 해석이 실현되는 것을 나타낸점은 아주 흥미있다.

5 인식 과정의 제어

화상이해에서 항상 문제가 되는 것으로 데이터 지식의 애매성, 국소특징과 대국적 특징과의 관계, 국소 특징간의 정합성이란 문제가 있다. 인식과정에서는 이러한 애매성을 제거하고 국소적 특징간의 정합성을 바탕으로 국소적인 화상 특징으로부터 대국적인 정경의 이해를 행하는 것이 필요하다. 인식과정의 제어에서는 인식대상 자체에 관한 지식쪽에서 어떤 모양으로 해석을 진행해 나가는가란 제어 추론 방식에 관한 지식도 필요하다.

일반적으로 인식과정의 제어방식으로서는 가설검정(hypothesis-and-test) 방식과 이완법이 있다. 전자는 축차처리적인 것에 대해 후자는 병렬 처리적이다.

5.1 가설검정(Hypothesis-and-test) 방식

이 방법에서는 먼저 bottom-up 해석에서 얻어진 가장 신뢰성이 높은 화상특징을 이용해서 모델의 선택 시점의 결정을 행하고 구체적 모델을 예측한다. 다음에 예측된 모델에 관한 지식을 이용해서 top down 해석을 행하고 그 모델의 타당성 검증, 다른 대상의 인식을 행한다. 이러한 bottom up 해석과 top down 해석의 반복에 의해 정경의 기술이 축차 작성되어 간다. 또한 새로운 해석결과 모순이 생겨 모두 작성된 정경의 기술이 수정되어지는 back tracking이 행해지는 것도 있다.

이 방식에서는 어떤 대상의 인식 예측을 행하기에 충분한 국소적 특징이란 어디에서 어느정도 넓이의 범위를 조사하면 충분히 신뢰할 수 있

는가란 것이 큰 문제로 된다.

이 문제는 학습문제와도 밀접하게 관계되어 있고 앞으로의 큰 연구과제이다. 또 back tracking을 행하는 것으로는 인식과정의 이력 기억법 및 그 변경법에 관한 검토도 필요하다.

5.2 이완법

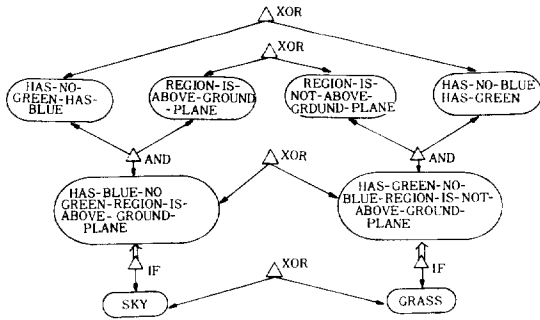
이 방법에서는 우선 각 화상특징에 대해 그것과 대응할 가능성이 있는 모든 정경특징을 선택하고 각각의 대응관계에 대해서 그 신뢰도를 나타내는 확률을 준다. 다음에 인식대상에 관한 지식(주로 기하학적 구조)에 바탕하여 각 화상특징과 그 주위 것과 해석의 정합성으로 부터 정경 특징과의 특징을 나타내는 확률을 각각 갱신한다.

이 확률 갱신 process를 반복하는 것이 의해서 각 화상 특징에 대한 다중해석이 차차 제거되고 최종적으로는 unique한 해석이 얻어진다. Waltz의 constraint filtering²¹⁾은 이 대표적인 예이다.

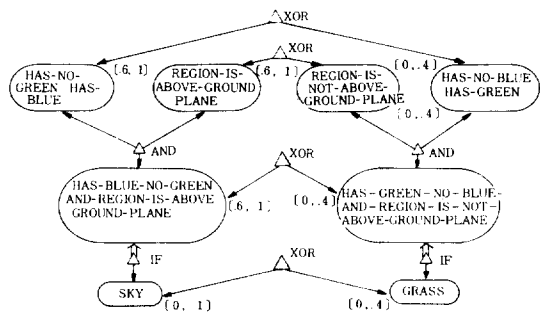
이 방법은 화상특징의 group화 정경 인식에서 매칭이란 저레벨, 고레벨 어느처리에도 적용 가능하고 잡음에도 강하기 때문에 최근에는 여러 가지 응용에 시험되고 있다.^{14),15),22)} 그러나 해외 수속성, 국소 특징간의 정합성을 나타내는 중계수 설정법, 또는 해석결과에 대한 설명 기능의 결여 모든 국소특징이 똑같은 모양으로 취급되는 처리의 유연성이 부족한 결점이 있다. Price²²⁾는 이완법에 부분-전체 관계에 의한 계층구조의 도움을 시도하고 있는데 일반적으로는 화면내의 어떤부분에 주목하고 그 부분에서는 다른 것은 틀린 해석을 해야한다는 주의의 집중(focus of attention)같은 mechanism을 어떤 모양으로 짜 넣는가 가 앞으로의 과제이다.

5.3 신뢰성의 계산

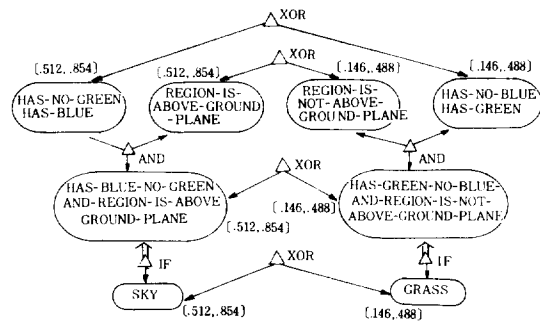
화상이해에서 제한 받지 않는 추론·인식을 행하는 경우 항상 큰 문제가 되는 것으로 데이터 지식의 신뢰성(애매성)이 있다. 특히 지식공학처럼 전문가가 지닌 지식자체가 확실한 것이 아니고 많은 단편적인 지식을 바탕으로 무엇인가 의미있는 결론을 얻는 것으로 하는 경우는 신뢰



(a) Dependency Graph에 의한 지식의 표현



(b) 어떤 영역이 확률 0.6에서 above ground, blue, not green이라는 정보가



(c) 똑같은 영역이 확률 0.3에서 not blue, green, not above ground이라는 정보가 (b)에 추가된 결과

그림11. Dependency graph기초로 한 신뢰성의 계산

성의 표현 계산이 본질적 문제로 된다. 종래 이러한 신뢰성을 표현하는 방법으로는 Bayes 확률을 이용한 것이 많다.

그러나 이 방법에서는 (1) 사전확률, 조건부 확률이 미리 알려진 것 (2) 신뢰의 제거와 불신용이 구별되지 않는다는 제한이 있다. 패턴 인식에서는 많은 샘플이 주어지고 그러한 특징량의 분포

에 바탕을 두고 행하면 좋고, 이러한 제한은 특히 문제로 되지 않는다. 이것에 대해 화상이해에서는 화상특징의 애매성이다, 개념레벨에서 지식의 애매성이란 지식공학에서의 문제와 똑같은 문제가 생기고 인식과정의 제어를 고려하는데에서 신뢰성의 표현 계산이 중요한 의미를 갖는다.

최근에는 Bayes 확률의 제한으로부터 벗어나고 불확실한 정보를 유연하게 표현하기 위한 기초 개념으로써 Dempster & Shafer 이론을 이용하는 것이 지식공학 분야에서 활발히 시도되고 있다. 화상이해에서도 이 영향을 받아 Dependency Graph에 의한 지식의 표현, 다중류의 센서로부터 얻어진 정보의 종합화, Dempster & Shafer 이론의 이완법예의 응용등이 시도되고 있다. 이 이론에서는 어떤사상의 신뢰성이 $[0, 1]$ 의 단위구간에 포함되어 있는 부분구간 $[SUPT, PL]$ ($0 \leq SUPT \leq PL \leq 1$)로 표현된다.

SUPT, PL은 그 사상이 올바른 것의 하한, 상한 확률을 나타내고 있고, $[1, 1]$ 의 완전한 신뢰 $[0, 0]$ 는 완전한 부정 $[0, 1]$ 은 전부를 판단하는 정보의 완전한 제거를 나타낸다. 일반적으로 그 사상을 지지하는 사실이 일어나지면 SUPT가 증가하고, 부정하는 사실이 일어나지면 PL이 감소한다. 이러한 사실(관측 데이터의 획득)에 의한 $[SUPT, PL]$ 의 갱신은 Dempster의 결합법칙을 이용해서 계산된다. (상세한 것은 34참조) 그림11은 Dependency Graph를 이용한 신뢰성의 계산과정을 나타낸 것으로 Dempster의 결합법칙에 바탕해서 독립적인 관측사실의 조합으로부터 어떤 화상특징의 인식결과가 확률로서 표현된다. 그림11로부터 명확히 된 것처럼 이 방법에서는 먼저 식별하는 대상의 집합(SKY, GRASS)를 정하고 어떤 화상특징이 그것내의 어느쪽인가를 식별한다는 패턴 분류에 가까운 기초개념이 기본으로 되고 있다. 일반적인 화상이해 시스템에서 이러한 방법을 적용하는 것으로는 먼저 서술한 상위-하위 관계, 부분-전체관계 또는 유사-차이 관계등에 의한 지식표현과 신뢰성의 계산 과정을 어떤 모양으로 조합시키는가 가 문제로 된다.

6 결 론

본 원고에서는 화상이해 시스템이라는 관점에 서의 문제점을 갖고 연구현황에 대해서 설명했다.

문헌 1)에도 서술되어 있는 것처럼 1970 년대 후반에 개발된 화상이해 시스템 대다수는 어떤 제한된 대상세계의 특성에 크게 의존하는 것으로 되어 있는데 일반적인 화상이해 시스템을 만드는데는 아직 많은 문제가 남아있다.

그 중에서도 3 차원 물체의 표현법, 인식법은 현재의 화상이해 연구의 중심과제로 되어 있고 활발한 연구가 진행되고 있다. 또 시스템적 관점으로 부터는 (1) 다차원의 데이터 대상에 관한 지식을 적절하게 나타내기 위한 어휘의 설정¹⁵⁾ (2) 주의 집중, 예측, 추론 등에 관해서 명확한 방식의 개발 (3) 애매한 데이터 지식의 표현추론 방식²³⁾의 확립등이 앞으로의 과제로 남아 있다. 유연하고 효율적인 화상이해 시스템을 만들기 위해서는 이러한 문제에 대해서 자세히 검토하고 각각의 문제에 대해서 일반적인 방식을 확립할 필요가 있다.

참 고 문 헌

- 1) Binford, T.O., Survey of Model-Based Image Analysis, The International Journal of Research, Vol.1, No.1, pp.18-64(1982).
- 2) Kanade, T., Region Segmentation : Signal vs Semantics, Computer Graphics and Image Processing, Vol.13, pp.279-297(1980).
- 3) Marr, D., Early Processing of Visual Information, AI Memo, No. 340, MIT(1975).
- 4) Hildreth, E.C., The Detection of Intensity Changes by Computer and Biological Vision System, Computer Vision, Graphics and Image Processing, Vol.22, No.1, pp.1-27(1983).
- 5) Haralick, R. et al., The Topographic Prima Sketch, Proc. of Trends & Application, pp. 198-209(1983).
- 6) Brooks, R.A., Goal-Oriented Edge Rinking and Ribbon Finding, Proc. of Image Understanding Workshop, pp. 72-78(Aprij)(1979).
- 7) Oshima, M. and Shirai, Y., Object Recognition Using Three-Dimensional Information, Proc. of 7th IJCAL, pp. 601-606(1981).
- 8) Ballarb, D.H., Generliziog the Hough Transform to Detect Arbitrary Shaes, Pattern Recognifion, Vol.13, No.2, pp.111-122(1981).
- 9) Bajcsy, R., Three-Dimensional scene Analysis, proc of 5th ICPR, pp.1064-1074(1980).
- 10) Minsky, M., A Framework for Representing Knowledge, The Psychology of Computer Vision(P.Winston ed.), McGraw-Hill(1975).
- 11) Rubin, S.M., The ARGOS Image Understanding System, Proc. of Image Understanding Workshop, pp. 159-163(Nov.1978)
- 12) Koenderink, J.J. and Van Doorn, A.J., The Internal Representation of Solid Shape with Respect to Vision, Biol. Cybernetics, Vol. 32, pp. 211-216(1979).
- 13) Fekete, G., Technical Report, Univ. of Maryland(in preparation).
- 14) Brooks, R.A., Symbolic Reasoning among 3-D Models and 2-D Images, Artificial Intelligence, Vol. 17, pp. 285-348(1981).
- 15) Tsotsos, J.K., Temporal Event Recognition : An Application to Left Venticular Performanees, Proc. of 7th IJCAL, pp. 900-907(1981).
- 16) Bobrow, D.G. and Winograd, T., An Overview of KRL, a Knowledge Representation Language, Cognitive Science, Vol. 1, pp.3-46(1977)-
- 17) Hanson, A. and Riseman, E., VISIONS : A Computer System for Interpreting Scenes, Computer Vision System(Hanson and Riseman eds.), Academic Press(1978).
- 18) Rosenfeld, A., Quadrees and Pyramide for Pattern Recognition and Image Procecing, Proc. of 5th ICPR, pp. 802-807(1980).
- 19) Nagao, M. and Matsuyama, T, A Structural Anaiysis of Complex Aerial Photographs, Plenum(1980).
- 20) Selfridge, P.G., Reasoning about Sucess and Failure in Aerial Image Understanding, Ph. D. Thesis, TR 103 Univ. of Rochester, N. Y. (1982).
- 21) Walta, D.L., Understanding Line Drawings of Scenes with Shadows, The Psychology of

- Computer Vision(P. Winston ed.), McGraw-Hill(1976).
- 22) Price, K.E., Symbolic Matching of Images and Scene Models, Proc. of Workshop on Computer Vision, pp. 105-112(Aug.1982).
- 23) Davis, L. and Rosefeld, A., Cooperation Processes for Low-Level Vision : A Survey, Processes for Low-Level Vision : A Survey, Artificial Intelligence, Vol. 17, pp. 245-263(1981).