

모델 근거 물체 인식을 위한 데이터 베이스 구성

(Data Base Construction for Model-based Objects Recognition)

金宗培*, 崔然成**, 崔宗秀***

(Jong Bae Kim, Yeon Sung Choi and Jong Soo Choi)

要 約

모델 근거 인식법의 근본적인 과제인 정확성과 융통성의 결여를 해결하기 위하여 데이터 베이스를 이용한 새로운 물체인식 시스템을 본 논문에서 제안하였다. 먼저 정합에 사용될 여러 특징들을 추출하게 되는데, 이 특징들을 효율적으로 관리하고 정합에 필요한 정보를 체계적으로 제공하는 계층 데이터 베이스가 설계되었다. 정합의 결과는 다시 데이터 베이스로 되돌려져 저장된다. 계층데이터 베이스는 처리내용 뿐만 아니라 정합 결과까지도 저장하므로, 인식이란 데이터 베이스의 공간을 채우는 것이라 할 수 있다. 다면체의 각목 세계를 대상으로 실험한 결과, 인식의 오차는 없었으나 처리시간이 긴 단점이 있었다.

Abstract

A new system for three dimensional object recognition is proposed in an attempt to overcome the lack of accuracy and flexibility, the essential problem at which model-based recognition technique faces. Hierarchical data base was designed to manage a number of features efficiently which are extracted for the purpose of matching by the system and to provide information necessary to match systematically. The results of matching, in turn, is fed back to the data base, and stored in. The recognition is done by saving the results on higher level of hierachical data base.

I. 서 론

지금까지 이루어진 컴퓨터 시각(computer vision)

*正會員, 金星産電中央研究所
(GoldStar Industrial Elec. System Co., Ltd.)

**正會員, 濟州大學校 情報工學科
(Dept. of Information Eng., Cheju Nat'l Univ.)

***正會員, 中央大學校 電子工學科
(Dept. of Elec. Eng., Chungang Univ.)

接受日字: 1987年 8月 27日

(※本 研究는 韓國科學財團 1987年度 一般研究의 一部로 이루어진 것임.)

에 관한 연구를 방법론의 입장에서 보면 두가지로 대대별할 수 있다. 첫째, 제한된 대상과 공간에서 성립하는 기법과 이론에 대한 범위를 점차 확대시키면서 수정하여 나중에 일반화 시키는 방법, 둘째, 범용성이 있고 원론적인 것을 응용의 단계에서 적용의 범위를 좁여가면서 그것에 알맞게 수정하는 방법이다. 전자의 가장 대표적인 예가 각목세계(block world)이다. 단순한 다면체의 집합인 각목세계의 인식으로부터 많은 컴퓨터 시각의 알고리즘들이 탄생했음은 주지의 사실이다. 즉, 그 동안의 각목세계는 여러 이론의 시험장 역할을 충실히 해왔다.^[1]

본 고에서는 이 각목의 세계를 대상으로 새로운 모델근거 인식법(model-based recognition)을 제안한다. 물체 인식의 성능 평가를 위한 변수는 속도(speed)와 정확성(accuracy)과 융통성(flexibility)의 셋이 흔히 지적된다.^[2] 이 요소들을 향상시키기 위한 도구로서 모델 근거 인식법이 많이 연구되어 왔고, 또한 실용화가 되기도 하였다.^[3]

모델근거 인식법은 다음의 세가지 물음에서 시작된다.^[2](1) 장면에서 물체의 물리적 특성과 그것의 공간적 관계를 기술하기 위해 어떤 특징(features)을 추출해야 할까? (2) 물체에 의미를 부여하기 위해 특징(features)과 그들끼리의 관계는 어떻게 표현되어야 할까? (3) 모델과 영상의 특징들 사이의 정합은 어떻게 이룰까? 따라서 모델에 바탕을 둔 물체인식 시스템은 항상 이상 언급한 특징추출, 물체모델, 그리고 정합의 세가지 기본 요소로 이루어진다.

2차원으로 투영된 물체는 관측 방향에 따라 달라질 뿐만 아니라 이웃한 물체끼리의 부분적인 겹침(partial occlusion), 노이즈 등으로 인해 완전한 인식의 실마리를 제공하지는 못한다. 얻어진 특징들은 확실하지 않은, 단지 그럴듯한 정보(plausible information)에 불과하다. 그래서 본 논문에서는 가능한 한 많은 특징들을 추출하여 이들 사이의 구조적 관계로써 물체 모델을 표현하였고, 또, 특징들을 효율적으로 관리하고 저장하기 위하여 계층 데이터 베이스(hierarchical data base : HDB)를 구축하였다. 계층데이터 베이스는 입력 영상에 관한 처리 결과를 저장하고 모델 베이스와 정합되어 그 결과를 또한 저장한다. 제안된 방식은 상황에 따라 변화가 심한 3차원 물체의 인식에 적합하다. 유용성을 입증하기 위해 많은 다면체를 대상으로 실험하였다.

II. 시스템의 개요

시스템은 각 처리과정에서 부여된 임무를 수행하는 여러 구성요소(components)들로 이루어지며, 이들은 처리과정에 필요한 정보를 제공하거나, 처리된 결과를 저장하고 관리하기 위해 서로 정보를 교환한다. 그리고 구성요소들은 입력된 영상으로부터 의미 있는 특징을 추출하며, 이들과 모델과의 정합을 행한다. 그림 1은 시스템의 구성요소들과 그들의 내부 연결을 보인다.

영상입력 장치는 CCTV 카메라와 영상디지타이저(image digitizer)로 구성된다. 전처리기(preprocessor)는 장면의 고유한 특징을 보존하면서 화소중심의 영상을 영역중심의 영상으로 변환시켜 다음과정

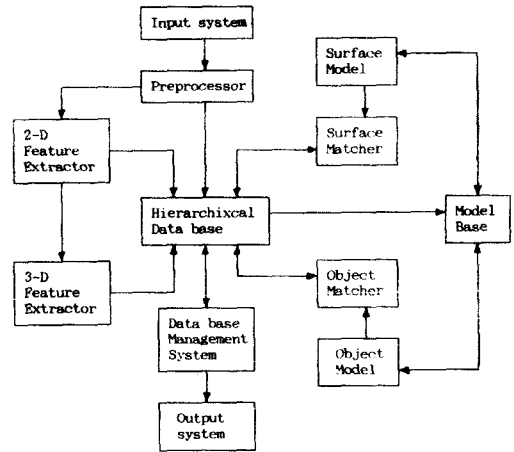


그림 1. 시스템의 계통도

Fig. 1. Block diagram of recognition system.

에서 처리할 정보량을 줄이게 된다. 또 이 처리에서는 영상을 분할하고 분할된 영역을 분류한다.

특징 추출기(feature extractor)에서는 전처리기에서 분류된 영역이 갖고 있는 특징들을 구하는 2차원 특징 추출기와, 이 특징들을 조합하고 분석하여 물체의 구조적, 위상학적 특징과 영역간의 관계성을 유도하는 3차원 특징 추출기가 있다.

모델베이스(model base)에는 여러 개의 다각형과 곡면들의 수치적 불변량(invariants)으로 구성되는 면 모델(surface model)과 물체의 구조적 특징으로 구성되는 물체모델(object model)이 있다.

계층 데이터 베이스는 필요에 따라 정보를 추가하거나, 삭제 및 호출할 수 있는 기능을 갖춘 여러 계층으로 구성되는데, 각각 장면 준위(scene level), 물체준위(object level), 면 준위(surface level)로 구성된다.

물체의 표현과 모델을 비교하여 상호의 일치성 여부에 따라 최종적인 인식을 행하는 정합기에는 영역의 형상에 대해 가설을 설정하는 면정합기(surface matcher)와 물체인식을 위한 물체정합기(object matcher)가 있다.

III. 물체인식 시스템

물체인식 시스템은 앞절에서 약술한대로 크게, 처리 부분과 데이터 베이스 부분으로 나눌 수 있다. 처리에는 전처리, 특징추출, 정합이 있다.

1. 전처리

전처리는 추출하고자 하는 정보 혹은 특징이 무엇

이든 상관없이 입력된 영상을 다음 처리를 위해 보다 만족스러운 상태로 변환시키는데, 독립적인 처리단계이다. 본 시스템에서는 영상을 기본영역(primitive region)으로 분류하게 된다. 여기서 기본 영역이란 밝기의 변화가 매끄러운(smoothing) 화소들의 집합을 말한다. 우선 서로 유사한 계조도(gray level)를 갖는 화소들을 한 영역을 병합하기 위해 멀티윈도우 히스토그램(multi-window histogram)법을 사용하여 영상을 분할하였다. 멀티윈도우 히스토그램법은 전영상의 계조도의 히스토그램에는 여러개의 골이 생기게 되는데, 이때, 골과 골사이에 하나의 윈도우를 씌우고 이 윈도우 안에서 가장 빈도가 높은 계조도로 그 윈도우 안의 모든 계조도를 대치하는 다준위 문턱치 처리(multilevel thresholding) 방식을 일컫는다.

영역 분할에 의해 다수의 영역, 즉, 물체의 면이 존재하게 된다. 이들이 각각 어느 물체에 속하느냐, 다시 말해서, 각 물체는 어느 면들로 구성되느냐를 판단하는 작업이 영역 분할이다. 이는 영역의 경계선 추적(boundary tracking)으로 가능하다.

정의: 한 화소와 그 4 근방이 모두 같은 계조도이면 그 화소는 내부점(inner point)이고, 그렇지 않으면 경계점이다.

내부점과 경계점의 구분은 수축후 XOR법(shrinking and exclusive-OR)을 사용한다. 원 영상과 그것을 수축시킨 영상을 XOR 하면, 내부는 변화가 없으므로 1을, 경계는 수축에 의해 손실되었으므로 0을 갖게 된다. 실제 영상은 2 치가 아니지만, 내부는 값이 균일하므로 1로, 그외는 0으로 대치해도 영역 단위의 처리에서는 상관이 없으나, 단, 다소 부수적인 처리가 요구된다. 그림 2에 처리과정을 보인다.

멀티윈도우 히스토그램을 선행하지 않으면 검출된 경계선이 절단되고 고립 혹은 顆粒點(isolated or granular points)이 발생하여 다음의 처리를 곤란하게 만든다. 추적은 8 방향 체인 코드를 수정한 모드 설정 체인코드법을 사용한다.^[4] 이는 원래 선도형 정보의 압축을 위해 제안되었지만 모서리 검출 능력이 뛰어나므로 선화解釋(line drawings interpretation)에도 유용하다. 물체의 분할은 두 단계로 나누어지는데, 경계선 코딩과 경계선 상호 관계설정이 그것이다.

첫째의 경계선 코딩에서는 장면에 존재하는 모든 영역의 경계선이 라스터주사(raster scanning)와 모드설정 체인 코드법에 의해 코딩되고 각각 고유한 영역 레이블을 부여받는다.

관계의 입장에서 보면 장면내의 물체는 단독 물체

0	0	0	0	1
0	0	0	1	1
0	0	1	1	1
0	1	1	1	1
1	1	1	1	1

(a)

0	0	0	0	0
0	0	0	0	1
0	0	0	1	1
0	0	1	1	1
0	1	1	1	1

(b)

0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	1	0	0	0
1	0	0	0	0

(c)

그림 2. 경계점과 내부점의 구분원리

- (a) 원 영상
- (b) 수축된 영상
- (c) 수축 후 XOR의 결과

Fig. 2. Identificatin of boundary and inner points.

- (a) original image.
- (b) shrinking image.
- (c) results of XOR operation after shrinking.

와 겹친 물체(occluded object)의 두 종류가 있다. 둘째의 관계 설정 단계에서는 T자형 경계선에 의해 겹친 물체를 분리하고, 경계선이 인접한 영역은 동일한 물체의 영역이므로 고유한 물체 레이블을 부여받는다. 물론 단독 물체에는 T자형 경계선이 존재하지 않는다. 이상과 같은 레이블링(labeling)과 물체의 분리 기술은 컴퓨터 시각연구의 초창기에 다면체 이해(polyhedral understanding)에서 흔히 이용되고, 완성되었다.^[5] 이상의 과정으로 전처리의 목표인 영역과 물체의 분할이 이루어진다.

2. 특징 추출

특징 추출은 두 단계로 나누어진다. 하위 단계인 2차원 추출 과정에서는 각 영역의 형상을 결정하는데 필요한 여러 특징들을 구한다. 이들에겐 형상의 기본적인 특징인 꼭지점의 수, 함축도(compactness), 면적, 둘레, 영역의 주축과 수평축이 이루는 각도인 방향(orientation), 영역의 최장현과 이에 수직하는 최장현의 비율인 편심율(eccentricity), 영역의 위치

를 대표하는 무게중심 등이 있다. 데이터 베이스의 입장에서 보면 각 영역의 레이블은 레코드의 기본열쇠(primary key)가 되므로 레코드를 튜플(tuple)로, 필드를 세포(cell)로 부르는 것이 일반적이고 또 이해하기도 쉬워진다. 이 관계 데이터 베이스(relational data base : RDB)는 면 정합 과정에서 사용된다.

영역의 꼭지점은 앞서 언급한 모드 설정 체인코드에 의해 검출된다. 둘레는 각 기준 모드에서 방향코드가 1 일때는 $\sqrt{2}$, 그 외에는 1로 한, 영역의 경계선 길이이다.

면적은 영역의 크기를 나타내는 기본량으로 영역의 내부점과 경계점의 총 갯수가 된다.

함축도는 종종 원형도(circularity)라고도 하며, 형상이 얼마나 굴절이 심하고 복잡한가를 측정할 수 있는 형상 고유치이다. 면적을 A, 둘레를 L 이라 하면, 함축도 C 는¹¹⁾

$$C = \frac{4\pi \times A}{L^2} \quad (1)$$

이다.

영상에서 한 영역의 상대적인 위치를 의미하는 무게 중심, 신장도와 회전은 모두 모멘트를 이용해 측정할 수 있다.¹⁰⁾

주어진 도형 R의 pq차 모멘트는

$$m_{pq} = \sum_x \sum_{y \in R} x^p y^q \quad (2)$$

으로 정의되며, 여기서 (x,y)는 경계 혹은 내부점이다. 영차(zero order) 모멘트는 단순히 도형의 화소 수, 즉, 면적이 된다.

영역의 무게 중심을 (x₀, y₀)라 하면

$$x_0 = \frac{m_{10}}{m_{00}} \quad (3)$$

$$y = \frac{m_{01}}{m_{00}} \quad (4)$$

이다. 식(2)에 정의된 모멘트는 원점을 식(3)~(4)에서 구한 중심으로 이동시킴으로서 위치에 무관하도록 다시 정의할 수 있다. 즉, 위치 불변 모멘트(position invariant moments)는

$$\mu_{pq} = \sum_x \sum_{y \in R} (x-x_0)^p (y-y_0)^q \quad (5)$$

이며, $\mu_{01} = \mu_{10} = 0$ 이다. 그리고 μ_{11} , μ_{02} , μ_{20} 은 통상 관성 모멘트(moments of inertia)라 불린다.

방향은

$$\theta = \frac{1}{2} \tan^{-1} \left[\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right] \quad (6)$$

로 정의되며, 도형을 각 θ 만큼 회전시키면 회전에 무관(rotation invariant)하게 된다.

근사적 편심율 e는 다음과 같이 주어진다.

$$e = \frac{(\mu_{02} - \mu_{20})^2 + 4\mu_{11}}{A} \quad (7)$$

시스템이 대상으로 하는 3 차원 물체는 2 차원 영상으로 투영되었을 때, 관찰 방향에 따라 형상이 크게 달라진다. 그러므로 지금까지 논의한 2 차원 특징은 그대로 3 차원 물체인식의 파라미터가 될 수 없고, 앞에서도 언급한 바 있지만 함축도와 편심율은, 영역의 2 차원 형상 결정에 꼭지점, 방향, 무게중심 등은 관계기술에 주로 이용되며 RDB에 보관되어 있으므로, 인식에 부수적인 역할을 담당할 수도 있다. 아주 간단한 예로 한 장면에 크기가 다르지만 형상이 같은 물체가 존재할 때, 로봇에게 작은 물체를 집어라는 명령을 내리면, 로봇은 RDB의 면적 세 포들을 검색해야 할 것이다. 또 컨베이어 벨트에 의해 운반되는 제품 중에서 넘어진 것을 바로 세우는 작업을 담당하는 로봇은 방향세포들을 검색할 것이다. 이런 부류의 예들은 실상황에서 흔히 등장하게 되고, 이때 2 차원 특징들은 인식에 직접 이용될 수 있다.

3 차원 특징 추출기는 표현기라고도 말할 수 있는데, 왜냐하면 물체의 위상학적 특징(topological feature)을 구할 뿐만 아니라 물체 상호간의 관계를 기술하기 때문이다. 위상학적 특징에는 모서리(edge)와 접속점(juntions)의 형태 그리고 그것의 수 등이 있다.

선화 해석에 사용되고 있는 모서리와 접속점의 레이블링 기법¹²⁾ 위상학적 특징을 추출하기 위해 사용되었지만, 새삼 이 기법을 소개할 필요는 없으리라 여겨진다. 다만, 전처리에서 구한 영역의 경계선이 두 영역의 접경에서도 서로 공유되지 않고 인접하거나 평행, 또는 거의 평행하게 되므로 이들을 하나의 선으로 통합하는 과정이 선행되었다. 실험에서는 55크기의 윈도우가 경계선을 일주하며 접경의 선들을 통합하고, 꼭지점이 존재하는 곳에서는 꼭지점을 윈도우의 중앙에 위치시킨 후, 모서리의 수와 모서리의 상호가 이루는 각도에 의해 접속점의 형태를 결정했다. 그러나 모서리에 + (convex edge) 또는 - (cocave edge)의 레이블을 모순없이 부여하기는 쉽지가 않으며, 완화법을 이용하여 반복 연산을 하지 않으면 안된다. 완화법에 의한 레이블링(relaxation labeling)은¹⁰⁾ 기존의 연구가 풍부하여 구현은 어렵지 않으나, 테스트 한 결과 장면의 복잡도(complexity)에

따라 레이블링에 소요된 시간이 전체 처리시간(processing time unit)의 30~40%를 점유하여 사용함에 어려움이 있다. 특히, 본 연구의 목적이 선화의 레이블링에 의한 다면체 인식이 아니라 물체와 장면의 데이터 베이스 구성에 있고, 모서리의 종류는 데이터 베이스의 한 속성일 뿐이며 인식에 결정적인 역할을 하지 않기 때문에, 약간의 오차를 감수하고 Huffman-clowes법을 채택했다. 실험에서 사용한 각목 세계가 단순했기 때문이겠지만, 다행히 오차는 없었다.

하나의 장면은 그래프로 표시될 수 있는데, 이때 마디(nodes)는 영역에, 호(arcs)는 영역과 영역사이의 관계에 대응한다. 관계는 각 영역의 중심의 위치를 비교하여 얻어진 영역간의 연결 상태이다. 연결 상태에는 왼쪽-위(left-above), 오른쪽-위(right-above), 위(above), 옆(side) 등이 있다. 이 3차원 특징추출 과정에서 얻어진 결과들은 의미기호(semantic symbol)로 변환되어, HDB로 보내져 저장된다.

IV. 계층 데이터 베이스

효율적인 정합을 위해 앞절에서 추출한 많은 특징들을 저장하고 관리할 수 있는 데이터 베이스가 필요하다. 이 데이터 베이스의 기능은 물체의 인식에 소요되는 정보를 제공하고, 인식된 결과를 저장한다. 그림 1을 좀 더 자세히 살펴보자. 전처리기에서 나온 분할된 영상은 2차원 특징 추출기로 제공되고, 부여받은 레이블은 HDB로 보내진다. 2차원 특징추출동안 데이터 베이스 관리시스템(data base management system: DBMS)은 레이블을 토대로 HDB의 규모와 형태를 결정한다. 2차원 특징추출기를 지난 영상은 3차원 특징추출기로 가고 결과는 HDB로 보내져 저장된다. 3차원 특징이 구해지는 동시에 HDB는 면 모델과 정합을 하게 된다. 정합의 결과와 3차원특징 추출의 결과가 HDB에서 물체 모델과

정합을 시작한다. 이 일이 종료되면 DBMS는 HDB를 토대로 장면을 해석하게 된다. 물론, 개별 물체의 인식이 작업의 목표라면 최종단계는 생략된다.

그림 1은 데이터의 흐름만을 도시했으나, 실제로는 입출력 장치의 제어와 DBMS와 교신하는 운영체제가 추가된다.

인식이란 결국 HDB의 빈 튜플과 세포들을 채우는 것이다. 이는 완전히 채워질 수도 있고 일부만 채워질 수도 있다. 전체 처리는 병행처리를 하며, 데이터 베이스 관리자(data base administrator: DBA)는 새로운 계층과 속성을 추가시키거나 혹은 삭제, 수정을 한다. 또한 HDB의 내용을 조회하여 새로운 학습 지침을 세운다.

학습시에는 그림 1의 회색 처리기들만 관계하고, 이때 HDB는 단지 정보전달만을 위한 모조처리기(dummy processor)가 된다. 그림 3은 기존의 모델근거 인식 시스템이다.¹⁾ 적당한 입력장치를 통해 얻어진 영상은 갖가지 특징이 추출되어지고, 이 특징들은 사전 지식에 의해 모델베이스에 저장된다. 이는 학습 과정이다. 인식 과정에서는 완전히 동일한 방법으로 추출된 영상의 특징을 모델 베이스의 모델들과 정합하게 된다. 그림 1과 기존의 모델근거 인식 시스템을 비교해 보면, 본 시스템의 특징을 뚜렷이 알 수 있다. HDB는 정합에 필요한 정보를 효율적으로 제공할 뿐만 아니라, 특징과 정합결과를 계속 보관하므로 일련의 작업에서 얻어지는 연속된 영상 프레임을 처리하는데 아주 유용하다. 그러나 기존의 시스템은 특징의 양이 한정되어 있고, 특징 상호간의 관계 기술능력이 미약하며, 정합 과정이 구조적이지 못해 융통성이 부족하다.

HDB는 장면 준위, 물체준위, 면준위 등 3개의 계층으로 구성되어 있다. 장면 준위는 가장 상위로서 일종의 장면기술(scene description)로 물체인식의 결과를 저장한다. 또한, 물체의 위치와 물체들 간의 관

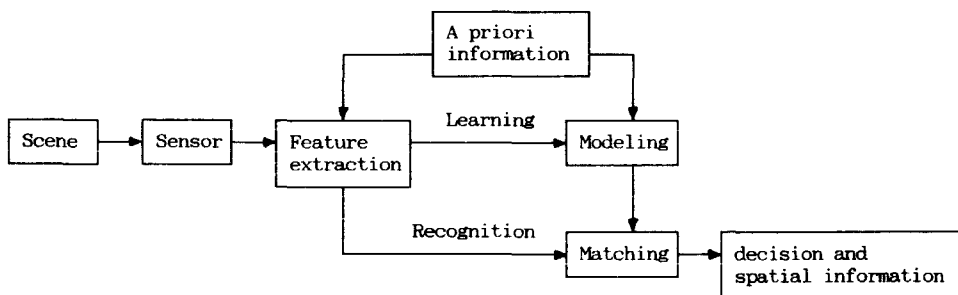


그림 3. 모델근거 인식시스템의 구성도

Fig. 3. Components of a model-based recognition system.

계를 나타내며, 물체 관계와 물체 연결 관계로 구성된다. 물체 관계에서는 영상에 존재하는 물체들과 이들의 형태와 위치를 나타내고, 연결 관계에서는 물체들을 둘씩 쌍으로 엮어 이들 사이의 위치 관계와 겹침여부를 나타낸다. 그림 4 에는 장면 준위의 DB가 있다. 예로, 물체 관계는 SCENE, OBJECT, TYPE OF OBJECT, POSITION의 네가지 속성으로 구성되어 있다.

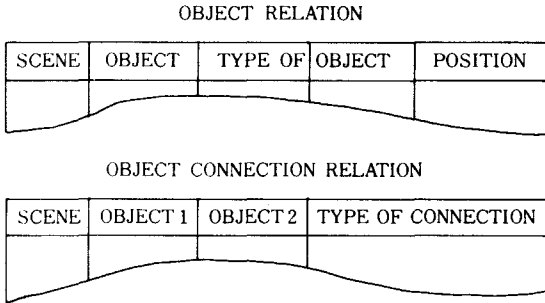


그림 4. HDB의 장면 준위 구조
 Fig. 4. Structure of scene level of HDB.

물체 준위는 면관계, 면 연결 관계, 모서리 관계, 접속점 관계로 구성된다. 면 관계에서는 면이 속해 있는 물체의 레이블, 면의 형태 및 면의 위치를 나타낸다. 면 연결관계에서는 면들을 둘씩 짝을 지워 이들의 연결형태 및 소속물체의 레이블을 나타내고, 모서리 관계에서는 모서리의 레이블과 모서리의 양 끝점이 되는 꼭지점의 쌍을, 접속점 관계에서는 접속점의 레이블과 이 접속점에서 분기되는 모서리의 레이블과 접속점의 형태와 위치로 구성된다. 그림5에 물체 준위의 DB가 있다.

면 준위는 영상내의 모든 면에 대한 기하학적 특징들로 구축되며, DB의 구조를 그림 6 에 제시하였다. HDB의 각 계층은 상호 관련된 둘 이상의 레이블로 구성되고, 이 레이블 자체는 III - 2 절에서 언급한 대로 일종의 RDB이다.

V. 모델 베이스 구성과 표현

모델은 정합에 활용될 지식을 갖추고 있어야 하며, 이들 지식은 정확하고 일관성 있게 표현되어야 한다. 인공지능에서 사용되는 지식표현법은 여러가지가 있으나 물체 인식을 위한 모델 베이스를 구성하는데 사용하기 위해 개체들 사이의 관계를 구조적으로 표현할 수 있어야 한다.

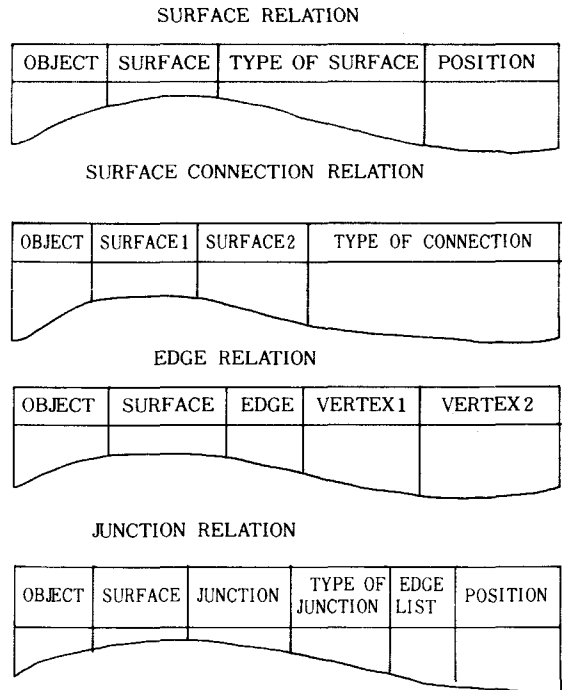


그림 5. HDB의 물체 준위 구조
 Fig. 5. Structure of object level of HDB.

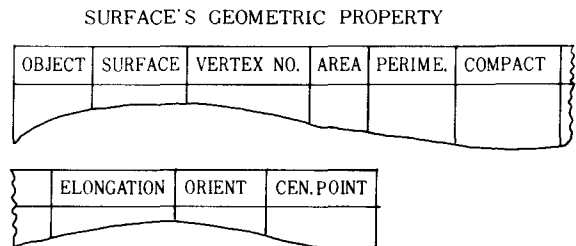


그림 6. HDB의 면 준위 구조
 Fig. 6. Structure of surface level of HDB.

본 시스템에서는 3 차원 물체는 면들로 이루어지고, 또 이 면들의 관계 및 위상학적 구속(constraints)으로 표현될 수 있다고 보아, 가시적인 면을 중심으로 특징에 바탕을 둔 표현(feature based representation)을 하였다. 영상은 여러 특징을 갖는 영역으로 분할되고, 각 영역은 물체의 면에 대응되므로, 영역을 기본으로 물체를 기술하는 것이 정합을 용이하게 한다.

모델베이스는 면 모델과 물체 모델로 구성된다. 면

모델은 면의 형상을 추정하는데 사용되는 면의 기하학적 기본량(geometric primitives)들로 구성되며 면 정합시 지식을 정합한다. 기본량들은 관찰방향과 카메라로부터의 거리를 변수로 갖는 함수 형식이다. 즉, 관찰각과 거리에 따라 면의 특징량은 변화된다.

물체 모델은 그래프로 표현된다. 즉, 마디는 물체의 면을, 호는 이들간의 연결관계를 표현한다. 정합시 처리를 간단히 하기 위해, 관계는 의미기호로 표현 되어 있다. 그림 7은 육면체에 대한 지식표현을 도해한 것이다. Consist-of와 must-be는 구조적 구속관계를, right-of, left-above와 side는 연결관계를 나타낸다. 명제표현(propositional representation)으 바 바꾸면 'S1은 S2의 오른쪽 위에 있고, S1, S2, S3는 모여서 육면체를 이룬다'가 된다. 이런 표현법은 의미망의 유사표현(analogical representation)에 해당되며, 그래프나 네트워크의 자료구조로 쉽게 구현된다.

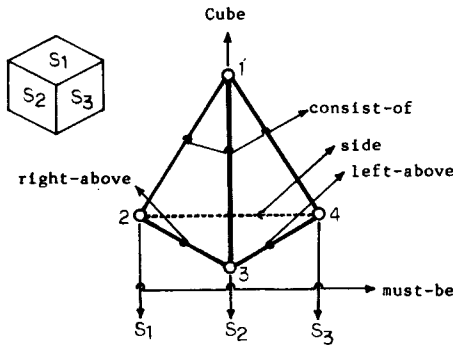


그림 7. 육면체 모델의 구조적 표현
Fig. 7. The schematic representation of model for cube.

VI. 정 합

인식은 학습과정을 통해 만들어진 모델과 영상으로부터 구한 특징들과의 정합을 통해 이루어진다. 본 시스템은 면의 형태를 추정하는 면 정합과 개별 물체의 인식을 위한 물체 정합의 두 단계가 있다.

면 인식은 HDB의 면 순위 정보와 면 모델간의 일치성을 계산하기 위해 다음과 같은 연산자를 적용한다.

$$D_i(p_i(S_k), p_i(M_u)) = w \frac{|p_i(S_k) - p_i(M_u)|}{p_i(M_u)} \quad (\ell = 1, 2, \dots, 5) \quad (8)$$

여기서 D_1 은 꼭지점수, D_2 는 함축도, D_3 는 면적, D_4

는 둘레, D_5 는 편심율에 대한 연산자이다. S_k 는 영상의 k 번째 영역이며, M_u 는 i 번째 모델의 j 번째 데이터이다. 함수는 S_k 는 M_u 의 특성을 비교한다. p_i 는 영역의 ℓ 번째 특성을 나타내고, D_i 은 ℓ 번째 특성의 차를 계산하는 연산자이다. 또, w 는 ℓ 번째 특징에 대한 가중치이다.

위 연산자를 예상되는 모든 면 모델에 대해 적용하여 구한 일치성의 값으로부터 면 형상을 추정하기 위해 다음 평가식을 세운다.

$$E_i(S_k, M_u) = \sum_j D_i \quad (9)$$

$$\text{Type of surface} = \text{Min}[E_i(S_k, M_u)] \quad (10)$$

$E_i(S_k, M_u)$ 가 가장 적은 값을 갖는 j 번째 모델을 영역의 형상으로 추정하고, 인접한 영역으로 확대하여 위의 과정을 반복한다. 실제로는 탐색시간을 줄이기 위해 꼭지점 수에 대한 연산자 D_1 이 작은 모델부터 평가식을 적용했다. 추정된 형상은 의미 기호로 변환되어 HDB의 물체 준위에 저장된다. 면 정합은 모델 구동처리(model-driven process)로 이루어진다.

물체 정합은 HDB에서 물체 순위의 정보와 물체 모델과의 면 형태 및 연결상태를 검사하는 평가함수를 적용하여 이루어진다. 이 평가함수는

$$E_2(M_k) = \frac{\sum a_i + E_k \beta_j}{F_k + L_k} \times 100 (\%) \quad (11)$$

이때,

$$a_i = \begin{cases} 1 & \text{if same surface} \\ \sigma & \text{if otherwise} \end{cases}$$

$$\beta_j = \begin{cases} 1 & \text{if same connection state} \\ 0.5 & \text{if otherwise} \end{cases}$$

이다. 여기서 $E_2(M_k)$ 는 k 번째 모델에 대한 함수값, F_k 는 k 번째 모델의 면의 수, L_k 는 호의 수이다. 또, a_i 는 모델의 i 번째 면에 대한 유사도로, HDB의 면과 모델의 면의 형태가 일치하면 1을, 그렇지 않으면 Kano의 유사도를 취한다.

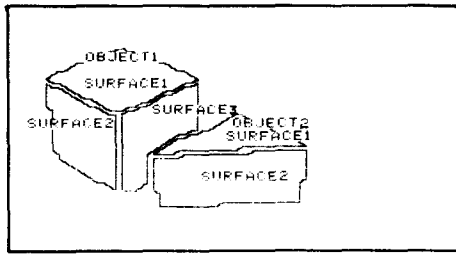
σ 는 Kano의 유사도이다.^[10] β_j 는 모델의 j 번째 호에 대한 유사도로서, HDB의 연결상태와 모델의 연결상태가 일치하면 1로, 다르면 0.5를 갖는다. $E_2(M_k)$ 의 값이 85% 이상이면 대상 물체를 모델의 k 번째 물체로 결정하고, 85%가 여럿이면 가장 큰 값을 갖는 것으로, 반대로 하나도 없으면 미확인 물체로 판단한다. 이 경우에는 E 가 가장 큰 물체가 가능성이 높다고 보아, 훼손된 면, 또는 겹친 면의 여부를 경험적 지식(heuristic knowlegde)으로 결정한다.^[11]

손실된 면은 복구되어 다시 평가치가 계산된다. 그래도 평가치가 개선되지 않으면 unknown으로 레이블링하고 다음 문제로 넘어간다. 판단 기준치 85%는 실험치로서 시스템의 상태와 알고리즘의 세부적 차이에 의해 달라질 수 있다.

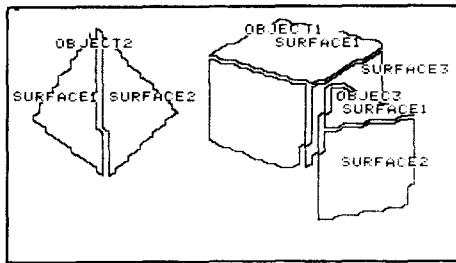
인식의 결과는 HDB로 이송되고, 즉시 DBMS는 슈퍼바이저 호출에 의해 제어 시스템을 넘긴다.

Ⅶ. 실험 및 고찰

실험의 대상은 석고로 만든 다면체들이며, 입력 영상의 해상도는 64×64×64이다. 그림 8은 레이블링된 장면이다.



(a)



(b)

그림 8. 레이블링된 각목 세계

Fig. 8. Examples of the labeled block world.

표 1은 그림 8(a)에 대한 HDB이다. 여기서 위치(position)는 화면의 좌측 상단을 원점으로 한다.

표 2는 그림 8(c)에 대한 면 정합을 위한 평가식 적용 결과이다. 또, 표 3은 물체 정합을 위한 평가식 적용 결과이다. 면 정합결과, 물체 1은 사각형+사각형+육각형으로 구성되었음을 HDB는 보여주고 있다. 물체 정합의 결과, 물체 1은 모델 1로, 물체 2는 모델 4로 각각 인식되었다.

실험결과 가장 큰 문제점은 처리시간이다. 즉, 정합이 계층적이기 때문에 정확하기는 하나 시간이 많

표 1. Fig. 8(a)에 대한 HDB

- (a) 장면 준위의 내용
- (b) 물체 준위의 내용
- (c) 면 준위의 내용

Table 1. Contents of HDB for Fig. 8(a).

- (a) contents of scene level HDB.
- (b) contents of object level HDB.
- (c) contents of surface level HDB.

* Object Relation *

Scene	Object	Type of object	Position
Scene4	OB 1	Cubic	22, 30
Scene4	OB 2	Prism	34, 42

* Object Connection Relation *

Scene	Object 1	Object 2	Type of connection
Scene 4	OB 1	OB 2	Occluding, side

(a)

* Surface Relation *

Object	Surface	Shape of surface	Mass point
OB 1	SU 1	Square	17, 21
	SU 2	Square	11, 35
	SU 3	Hexagon	22, 35
OB 2	SU 1	triangle	34, 37
	SU 2	Pentagon	34, 46

* Surface Connection Relation *

Object	Surface 1	Surface 2	Type of connection
OB 1	SU 1	SU 2	right-above
	SU 1	SU 3	left-above
	SU 2	SU 3	side
OB 2	SU 1	SU 2	above

* Edge Relation *

Object	Edge	Junct. 1	Junct. 2	Type of edge
OB 1	E1	J1	J2	convex
	E2	J2	J3	convex
	E3	J3	J4	convex
	E4	J4	J1	convex
	E5	J2	J5	convex
	E6	J5	J3	convex
	E7	J5	J6	convex
	E8	J6	J7	convex
	E9	J7	J8	convex
	E10	J8	J9	convex
	E11	J9	J4	convex
OB 2	E1	J1	J2	convex
	E2	J2	J3	convex
	E3	J3	J1	convex
	E4	J2	J4	convex
	E5	J4	J5	convex
	E6	J5	J6	convex
	E7	J6	J3	convex

* Junction Relation

Object	Junction	Type of junction	Position
OB 1	J 1	ELL	17, 13
	J 2	ARROW	5, 22
	J 3	FORK	15, 30
	J 4	ARROW	29, 22
	J 5	ELL	7, 40
	J 6	ARROW	17, 49
	J 7	ELL	21, 47
	J 8	ELL	22, 39
	J 9	ELL	29, 33
OB 2	J 1	ELL	35, 30
	J 2	ARROW	22, 41
	J 3	ARROW	45, 38
	J 4	ELL	23, 53
	J 5	ELL	45, 51
	J 6	ELL	46, 46

(b)

Surface Primitive	Object 1			Object 2	
	SU 1	SU 2	SU 3	SU 1	SU 2
Vertex No	4	4	9	3	5
Compact.	.7545	.6335	.5107	.5236	.779
Area	234	216	198	132	331
Perimeter	62.43	65.46	69.79	56.28	73.07
Elongat. 1	1.433	.666	.4229	1.244	.7504
Elongat. 2	6539.8	28425	38498	27025	31965
Mass x	17	11	22	34	34
center. y	21	35	35	37	46

(c)

표 2. 그림 8(a)에 대한 면 인식을 위한 평가 함수의 적용결과

Table 2. Results of evaluation function for surface recognition of Fig.8(a).

surface function	SURFACE 1	SURFACE 2	SURFACE 3
$E_1(i, 1)$	10.1309	13.5312	39.3369
$E_1(i, 2)$	0.45839	0.88931	26.8794
$E_1(i, 3)$	6.5176	7.02321	22.5278
$E_1(i, 4)$	12.3984	14.8577	21.3503
$E_1(i, 5)$	67.4005	64.7123	46.9357

Where $E_1(i, 1)$: results for triangle
 $E_1(i, 2)$: results for square
 $E_1(i, 3)$: results for pentagon
 $E_1(i, 4)$: results for hexagon
 $E_1(i, 5)$: results for circle

이 결된다. 지금의 상향식(bottom-up)이 아닌 하향식(top-down)의 정합이 가미되어야 하겠다. 또한 IV절에서 각 구성요소는 병행처리를 한다고 했으나, 현재 전용 운용체제를 설계하지 못했기 때문에 직접 구현

표 3. 그림 8(a)에 대한 물체인식을 위한 평가 함수의 적용결과

Table 3. Result of evaluation function for object recognition of Fig. 8(a).

parameter object	M_k	α	β	E_2
OBJECT 1	1	2.5	3	91.66%
	1	1.51	0.5	67%
OBJECT 2	2	0.99	0.5	49.66%
	3	0.9	1	66.33%
	4	1.6	1	86.66%

은 못했으나, 충분히 가능하다고 판단된다.

두번째 문제는 현상태로는 시스템이 범용적이지 못하다는 것이다. 왜냐하면 3차원 물체인식을 위해 2차원의 기하학적 정보를 사용했기 때문이다. 3차원 실세계는 2차원으로 투영되면서 엄청난 정보를 손실한다. 따라서, 입력정보 자체가 3차원인 입체시(stereopsis), 거리측정기(range finder)법을 쓴다든가 또는, 3차원의 형상정보를 제공하는 shape-from-X나 조도차 입체시(photometric stereo) 등을 사용해야 일반화 될 수 있겠다.¹⁷⁾

그리고, 다수의 장면을 인식하고자 할때는 HDB가 총체적인 데이터베이스로 바뀌어야 하기 때문에 스키마(schema)와 서브스키마(subschema)의 정의 등 전체적인 복잡도가 상당히 증가할 것으로 예상된다. 본고에서 다룬 내용은 DB의 규모도 작고 대상도 각목세게로 제한된, 따라서 인식에도 오차가 없는 시론적인 시스템에 관한 것이다. 그러나 위에서 지적한 입력정보의 문제만 해결된다면 정확성과 융통성에 있어서 기존의 방식보다 우월한 것으로 기대된다.

Ⅷ. 결 론

최근에 로봇 공학에 대한 관심이 높아짐에 따라 제조과정의 자동화를 위한 컴퓨터 시각이 상당한 주목을 끌고 있다. 로봇과 산업 자동화를 위한 몇몇 상업화된 시각 시스템이 있긴 하지만, 그것들의 능력은 아직 매우 초보적이다. 기계시각 연구의 역사에 비해 실용화로의 진전은 무척 느리다. 이유는 많은 제조업이 상세한 시각적 해석 능력을 원하며, 시스템은 저렴, 정확, 고속이어야 되고, 또한 융통성이 있어야 하기 때문이다.

본 고에서는 이와 같은 문제를 극복할 수 있는 새로운 시스템을 제안하였다. 시스템은 추출된 특징들

을 효율적으로 관리하고, 정합에 필요한 정보를 체계적으로 제공하는 계층 데이터 베이스를 갖는다. 그리고, 정합의 결과 자체도 DB의 가장 상위 준위에 저장된다. 즉, HDB를 구성하는 과정이 바로 인식과정이다. 정합을 위해 두가지 평가함수를 정의했으며, 모델 베이스는 테이블 형태의 면 모델과 그래프 형태의 물체 모델이 준비되었다. 제안된 시스템은 각 목 세계라는 제한된 상황에서 동작하도록 설계되었으므로 V절에서 제기한 여러가지 문제점을 가지고 있다. 앞으로는 곡면체나 무늬(texture)를 포함하는 물체 그리고 복잡한 비정형의 물체도 인식할 수 있도록 다양한 특징 추출법이 본 연구를 기초로 연구되어야 하겠다. 그리고 HDB를 프레임(frame)으로 변경하여 궁극적으로는 전문가 시각시스템(expert vision system)으로 발전시켜 보는 것도 재미있는 문제가 될 것이다.

参 考 文 献

[1] M. Oshima and Y. Shirai, "Object recognition using three-dimensional information," *IEEE Trans. PAMI*, vol. 5, no. 4, pp. 353-361, July, 1983.
 [2] R.T. Chin and C.R. Dyer, "Model-based recognition in robot vision," *ACM Computing Surveys*, vol. 18, no. 1, 1986.

[3] T.O. Binford, "Survey of model based image analysis system," *Int. J. Robotics Research*, vol. 1, no. 1, pp. 18-16, 1982.
 [4] 장기철, 최연성, 최종수, "선도형의 부호화를 위한 모드설정 체인코드," 한국통신학회 논문지, 13권 1호, pp. 41-53, 1988. 2.
 [5] M.B. Clowes, "On seeing things," *AI*, vol. 2, no. 1, pp. 79-116, spring 1981.
 [6] D. Waltz, "Understanding line drawings of scenes with shadow," in *The Psychology of Computer Vision*, Ed. by P.H. Winston, McGraw-Hill, pp. 19-92, 1975.
 [7] D.H. Ballard and C.M. Brown, *Computer Vision*, Prentice-Hall, pp. 254-260, 1982.
 [8] M.K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Information Theory*, pp. 179-187, Feb. 1962.
 [9] A. Rosenfeld, R.H. Hummel and S.W. Zucker, "Scene labelling by relaxation operators," *IEEE Trans. SMC*, vol. 6, no. 6, pp. 420-453, June, 1976.
 [10] 高野英彦, 形状パターンの認識技術, 情報調査會, pp. 135-170, 1984.
 [11] Y. Shirai, 'Analyzing intensity arrays using knowledge about scene,' in *The Psychology of Computer Vision*, Ed. by P.H. Winston, McGraw-Hill, pp. 93-114, 1975. *

著 者 紹 介



金 宗 培(正會員)
 1963年 1月 13日生. 1984年 2月
 중앙대학교 전자공학과 졸업. 1986
 年 8月 중앙대학교 전자공학과
 대학원 석사학위 취득. 1988年 1
 月~현재 금성산전(주) 중앙연구
 소 연구원. 주관심분야는 인공 지
 능, 컴퓨터시각, 컴퓨터시스템 등임.

崔 然 成 (正會員) 第25卷 第12號 參照
 현재 제주대학교 정보공학과
 전임강사

崔 宗 秀 (正會員) 第25卷 第12號 參照
 현재 중앙대학교 전자공학과
 부교수