

論文 95-32B-9-4

# Hough 변환을 이용한 캐드 기반 삼차원 물체 인식

## (CAD-Based 3-D Object Recognition Using Hough Transform)

具 滋 性 \*, 李 商 郁 \*

(Ja Seong Ku and Sang Uk Lee)

### 요 약

본 논문에서는 삼차원 물체의 표현에 삼차원 Hough 변환 공간을 이용하여 삼차원 물체 인식 문제를 고찰하였다. 물체의 모델링 과정에서는 인식하려는 물체의 캐드 형식으로부터 인식에 필요한 특징들을 추출해 내었다. 캐드 모델의 이용을 통해 정밀도와 확장성을 향상시켰다. 정합시에는 입력 영상과 왜곡(회전 이동과 평행 이동)을 가정한 기준 물체를 삼차원 Hough 변환 공간 상에서 비교하였다. 이 때 효율적인 인식을 위해 왜곡을 정의하는 고차원 파라미터 공간(6차원)을 저차원으로 나누었다. 먼저 왜곡을 정의하는 인자를 회전 이동 인자와 평행 이동 인자로 나누고, 회전 이동 인자는 다시 물체를 바라보는 시점에 관한 인자와 회전 각도에 관한 인자로 나누었다. 입력 영상의 삼차원 Hough 변환 공간을 직접 정합에 이용하기 때문에 잡음에 대한 민감성과 계산의 복잡성이 크게 감소되었다. 실험 결과 실제 거리 영상에서 만족할 만한 성능을 나타내었다.

### Abstract

In this paper, we present a 3-D object recognition system in which the 3-D Hough transform domain is employed to represent the 3-D objects. In object modeling step, the features for recognition are extracted from the CAD models of objects to be recognized. Since the approach is based on the CAD models, the accuracy and flexibility are greatly improved. In matching stage, the sensed image is compared with the stored model, which is assumed to yield a distortion (location and orientation) in the 3-D Hough transform domain. The high dimensional (6-D) parameter space, which defines the distortion, is decomposed into the low dimensional space for an efficient recognition. At first we decompose the distortion parameter into the rotation parameter and the translation parameter, and the rotation parameter is further decomposed into the viewing direction and the rotational angle. Since we use the 3-D Hough transform domain of the input images directly, the sensitivity to the noise and the high computational complexity could be significantly alleviated. The results show that the proposed 3-D object recognition system provides a satisfactory performance on the real range images.

\* 正會員, 서울大學校 制御計測工學科

(Dept. of Control and Instrumentation Eng., Seoul Nat'l Univ)

接受日字: 1995年3月10日, 수정완료일: 1995年9月11日

### I. 서 론

인공적인 시각 시스템은 주위 환경에 대한 시각 정보로부터 물체의 기하학적, 물리적인 성질을 유추해내는 것을 목적으로 한다. 카메라를 통하여 받아들여진

이차원 영상 정보를 통해 삼차원의 실세계를 분석하기 위하여 많은 연구가 진행되어 왔다. 이와는 달리 실세계로부터 삼차원 정보를 직접 얻는 방법에 대해서도 연구가 이루어지기 시작하였고 발전해왔다. 삼차원 정보를 직접 얻을 경우 실세계를 이차원으로 투영하여 얻는 것에 비해 정보의 손실이 적기 때문에 실세계를 분석하는 데에 있어서 더욱더 효과적이라고 할 수 있다. 주어진 모델이 존재할 때, 삼차원 영상을 분석하여 이 모델의 존재 여부 및 방향, 위치 등을 찾아내는 것이 모델 기반 삼차원 물체 인식(model-based 3-D object recognition)이다. [11]

삼차원 물체 인식은 크게 두 가지 부분으로 나뉘어진다. 한가지는 물체의 모델링(표현)이고, 다른 한가지는 이미 저장되어 있는 모델링 된 물체와 새로 입력된 삼차원 영상과의 정합이다.

물체의 모델링이란 인식하려는 대상 물체의 여러 가지 성질 중에서 인식에 필요한 특징들만을 모아서 효과적으로 표현하고자 하는 작업이다. 지금까지 주로 수동적인 작업에 의존하거나 기준 영상을 통해 훈련시키는 방법에 의존해 왔는데 두 방법 모두 물체의 수가 많을 때에는 상당히 번거로운 작업일 뿐만 아니라 정밀도를 저하시킨다. 이를 극복할 수 있는 체계적인 물체의 모델링 방법이 CAD(Computer aided design) 모델을 이용하는 것이다. [12] 물체의 CAD 모델은 물체가 설계될 때부터 존재하는 경우가 많기 때문에 물체의 수가 많아도 모두 체계적으로 데이터베이스에 저장할 수 있다. 또한 CAD 모델의 분석을 통해 가장 효과적인 정합 기법을 도출해 낼 수 있다. CAD 모델의 분석은 대상 물체가 정해지는 시점에서 가능한 작업이기 때문에 실제 물체 인식시에는 따로 시간을 필요로 하지 않아 정합 시간을 줄일 수 있다. [13-5]

정합 단계에서는 삼차원 물체의 표현 방법이 먼저 고려되어야 한다. 많은 경우 삼차원 영상을 정합에 이용하기까지는 물체의 면들을 분류해 내기 위해 영상을 영역화하고 영역화 된 부분을 표현할 수 있는 인자(parameter)를 추출해 내는 등의 전처리 단계가 필요하다. 그런데 이 과정은 잡음에 민감할 뿐만 아니라 계산 시간 면에서 전체 삼차원 물체 인식 시스템의 가장 큰 병목 구간으로 작용한다. 삼차원 Hough 변환 공간을 이용하여 삼차원 물체를 표현할 경우 이런 전처리 단계 없이 바로 삼차원 Hough 변환 공간상에서의 정합이 가능하다. [16] Hough 변환을 물체 인식에 이용

하려는 연구는 많이 소개되어 왔는데 이차원 Hough 변환뿐만 아니라 직선형 이차원 Hough 변환을 확장시킨 삼차원 Hough 변환을 삼차원 데이터에 적용시키려는 연구도 소개되어 왔다. [7]

본 논문에서는 삼차원 물체의 표현에 삼차원 Hough 변환 공간을 이용하여 삼차원 물체 인식 문제를 고찰한다. 물체의 모델링 과정에서는 물체의 CAD 모델로부터 기준 물체의 삼차원 Hough 변환 공간을 분석적으로 구해 낸다. 정합 단계에서는 삼차원 Hough 변환 공간상에서 왜곡(회전 이동과 평행 이동)을 가정한 기준 물체와 입력 영상을 비교한다. 이 때 왜곡을 가정하는 고차원 파라미터 공간을 저차원 공간으로 분할하여 인식의 효율성을 높인다. 그림 1 에 본 논문에서 제안하는 삼차원 물체 인식 시스템의 전체 구성을 제시하였다.

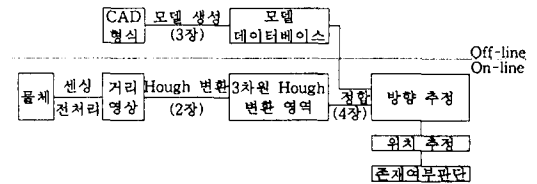


그림 1. 삼차원 물체 인식 시스템의 전체 구성  
Fig. 1. Block diagram of the 3-D object recognition system.

## II. 삼차원 Hough 변환

삼차원 Hough 변환은 이차원에서의 직선 Hough 변환의 확장이다. 이차원 직선 Hough 변환은 X-Y 공간의 한 벡터를 크기와 방향의 공간으로 변환하여 준다. 삼차원 Hough 변환도 마찬가지로 X-Y-Z 공간 상의 한 벡터를 크기와 방향의 공간으로 변환시켜 준다. 삼차원 Hough 변환이 이차원의 경우와 크게 다른 점은 방향을 나타내기 위해 두 개의 인자가 필요하다는 것이다. 삼차원 공간상에서 방향을 표현하기 위해 단위 구면에 균일하게 분포하는 단위 벡터들을 설정하고 이를  $u_n$  이라 하자. 이 때 삼차원 Hough 변환을 다음과 같이 주어진 데이터 벡터  $X$  와 단위 벡터 사이의 내적으로 정의한다.

$$p = X \cdot u_n \quad (1)$$

이 값을 양자화 하여 이차원 공간  $H(n, p)$  에 누적시킨다. 모든 단위 벡터  $u_n$  과 모든 데이터 벡터  $X$

에 대해 반복하면 결과적으로 새로운 삼차원 Hough 변환 공간  $H(n, p)$  가 생성된다. 이 삼차원 Hough 변환 공간의 의미를 그림 2 에 나타내었다. 단위 벡터  $u_n$  과 데이터 벡터  $X$  의 내적 값은 그림에서 길이  $p$  로 나타내 진다. 결국 이로 인해 누적된  $H(n, p)$  값은 단위 벡터  $u_n$  의 방향으로  $p$  만큼 떨어진 곳에 두께  $\Delta p$ (양자화 수준) 인 얇은 판에 포함된 데이터 벡터의 개수가 된다.

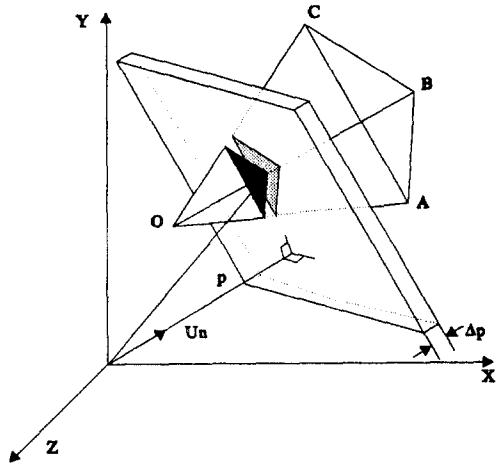


그림 2. 삼차원 Hough 변환 공간의 의미  
Fig. 2. Meaning of the 3-D Hough transform domain.

만약 모든 데이터들이 한 평면을 구성하고 있다면 그 평면의 법선 벡터에 해당되는 단위 벡터  $n$  에 대해서는 모든  $p$  값이 같은 값을 가지게 될 것이다. 따라서 삼차원 Hough 변환 공간  $H(n, p)$  의 값은 평면을 이루는 데이터의 개수가 된다. 그림 2 를 예로 들면 그림에 나타난 단위 벡터에 대해  $H(n, p)$  의 값은  $p$  가 사면체의 꼭지점 O 에 이르렀을 때부터 조금씩 증가하다가 면 ABC 에 이르면 면에 포함되는 데이터 벡터들을 모두 포함하게 되므로 갑자기 증가할 것이다. 이를 그림 3(a) 에 나타내었다. 만약 주어진 단위 벡터 방향으로 면이 존재하지 않는다면 그림 3(b) 와 같이 큰 극대치가 나타나지 않을 것이다. 이를 통해 쉽게 면을 검출할 수가 있다.

한편 이 성질을 관찰해 보면 EGI(Extended gaussian image) 와의 연관성을 알 수 있다. EGI 는 면의 방향에 대해 그 면의 면적을 대응시켜 물체를 표현하는 방법이다. 이는 삼차원 Hough 변환 공간이 포함하고 있는 정보의 일부분이라 할 수 있다. 삼차원

Hough 변환 공간은 앞에서 보았듯이 극대치가 나타나는 곳에서의  $H(n, p)$  값이 면의 면적을 나타내고 있을 뿐만 아니라 극대치가 나타나는  $p$  의 위치 값을 통해 면의 위치에 관한 정보도 나타내고 있는 것이다. 이를 통해서 물체의 평행 이동량에 관한 정보도 구할 수 있다. 이는 이동량 불변(Translation-invariant)의 성질을 가지는 EGI 와는 대조적인 것이다. 또한 EGI 의 경우 오목한 물체에 대해서 모호성이 생기지만 삼차원 Hough 변환 공간은 그렇지 않다.

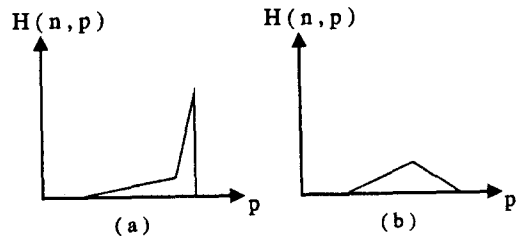


그림 3. 특정  $n$ 에 대한  $p$  축으로서의 사영  
(a) 평면이 있는 경우 (b) 평면이 없는 경우  
Fig. 3. Projection onto  $p$  axis of a specified  $n$   
(a) With plane (b) Without plane

이제 삼차원 Hough 변환 공간이 물체의 크기 변화, 평행 이동, 회전 이동에 대해 어떤 성질을 가지는지를 살펴보면 다음과 같다. [7]

- ① 물체가  $s$  배 되었을 때 :  $p$  축 상에서만  $s$  배

$$H_s(n, sp) = H(n, p) \quad (2)$$

- ② 물체가 회전 변환 행렬을  $A$  에 의해 회전되었을 때 :  $n$  축 상에서의 재 배열

$$H_r(m, p) = H(n, p), \text{ 여기서 } u_m = Au_n \quad (3)$$

- ③ 물체가 벡터  $t$  만큼 평행 이동되었을 때 :  $p$  축 상에서의 이동

$$H_t(n, p + t \cdot u_n) = H(n, p) \quad (4)$$

### III. 물체의 모델링

- 1. IGES(Initial graphics exchange specification) 캐드 형식

본 논문에서는 물체의 모델링에 IGES 캐드 형식을 이용한다. IGES 캐드 형식은 캐드 시스템간의 데이터 교환을 목적으로 제정된 캐드데이터의 한 형식이다. [8]

여기에는 이차원, 삼차원 물체의 표현뿐만 아니라 회로도나 그 밖의 도면들을 위한 여러 가지 객체들이 정의되어 있다. 이 중 본 논문에서 이용될 몇 가지 객체들을 살펴보면 다음과 같다.

- ① 선 : Line (type 110), Circular arc (type 100), Conic arc (type 104)
- ② 면 : Plane (type 108)
- ③ 입체 : Block (type 150), Right angular wedge (type 152), Right circular cylinder (type 154), Right circular cone frustum (type 156), Sphere (type 158)

대부분의 캐드 소프트웨어들이 IGES 형식으로서의 변환을 지원하므로 이를 이용할 수가 있다. 한가지 단점으로는 IGES 형식이 물체의 표현을 목적으로 만들어진 것이 아니고 데이터 교환을 목적으로 만들어 졌기 때문에 물체의 표현 방법이 그리 체계적이지 못하다. 따라서 IGES 캐드 형식을 분석하여 모델링에 필요한 특징들을 유추해 내는 과정이 적절한 방법으로 이루어져야 한다.

2. 모델 생성 과정

IGES 파일을 입력으로 하여 정합에 이용될 특징들을 모델 데이터베이스로 저장해 놓는 것이 모델 생성 단계이다. 정합에 이용될 특징은 삼차원 Hough 변환 공간이므로 IGES 파일을 분석하여 해당 물체의 삼차원 Hough 변환 공간을 구해 놓아야 한다.

물체의 표면에 관한 인자들을 알 때 이로부터 삼차원 Hough 변환 공간을 구하는 방법은 두 가지로 고려해 볼 수 있다. 첫 번째 방법은 주어진 인자를 가지고 표면 상의 점들의 좌표를 합성해 내어 전장에서 살펴본 삼차원 Hough 변환을 직접 계산하여  $H(n, p)$  에 누적시켜 삼차원 Hough 변환 공간을 구하는 것이다. 두 번째 방법은 표면의 방정식을 통해 특정 방향  $n$  에 대해 원점에서의 거리  $p$  를 구하고 그 표면의 면적값을 직접  $H(n, p)$  에 대입하는 것이다.

본 논문에서는 두 가지를 모두 이용한다. 이차 곡면 이하의 단순한 면에 대해서는 후자의 분석의 방법을, 복잡한 면에 대해서는 전자의 합성의 방법을 이용한다. 분석적으로 구할 수 있는 몇 가지를 살펴보면 다음과 같다.

① 평면

평면의 경우 IGES 캐드 형식으로부터 쉽게 평면 방

정식에서의 계수  $a, b, c, d$  를 얻을 수 있다. 먼저 삼차원 Hough 변환에 이용하는 단위 벡터의 집합 중 벡터  $(a, b, c)$  와 가장 가까운 벡터  $u_n$  을 선택한다. 이 때 선택된  $n$  에 대한  $p$  는 바로 평면의 원점으로부터의 거리  $d$  가 된다. 이렇게 평면 방정식의 계수  $a, b, c, d$  로부터  $n, p$  를 쉽게 얻을 수 있고 이 위치에서의 삼차원 Hough 변환 공간의 값  $H(n, p)$  에 이 평면의 면적을 넣는다. 이 때 평면의 법선 벡터  $(a, b, c)$  와 선택된  $u_n$  이 완전히 일치하지 않으므로 어느 정도 오차를 발생시키게 된다. 이를 보완하기 위해 가장 가까운  $u_n$  하나를 선택하는 대신  $(a, b, c)$  를 둘러싸는  $u_n$  을 세 개 택해서 각각의 거리에 따라 가중치를 주어 세 곳의  $H(n, p)$  에 대입하면 어느 정도 잡음에 대해 강한 성질을 지니게 된다.

② 원통면

원통면의 경우 IGES 캐드 형식으로부터 축 방향 단위 벡터, 높이, 반지름, 한 면의 중심점 등을 얻을 수 있다. 여기서 얻은 원통면을 원통면의 중심축이  $Z$  축에 오도록 회전 변환과 이동 변환을 한다. 만약 이렇게 이동한 원통면의 삼차원 Hough 변환 공간을 구한다면 삼차원 Hough 변환 공간의 회전, 이동 변환에 관한 성질을 이용해 쉽게 원래 원통면의 삼차원 Hough 변환 공간도 구할 수 있을 것이다.

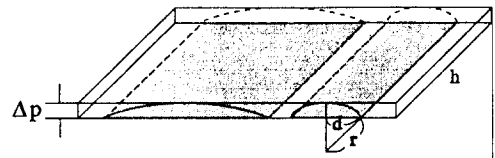


그림 4. 원통면의  $H(n, p)$   
Fig. 4.  $H(n, p)$  of cylindrical surface.

이제 중심축이  $Z$  축에 있는 원통면의 삼차원 Hough 변환 공간을 구해 보자. 먼저 방향  $n$  은  $Z$  축이 0 인 모든 단위 벡터가 선택되어야 할 것이다. 중심축이  $Z$  축 상에 있으므로 모든 방향에 대해 면까지 떨어진 거리  $p$  는 같고 이 값은 원통면의 반지름  $r$  이 된다. 이렇게  $n, p$  가 선택되었고 이제 여기에 대입해 줄 값  $H(n, p)$  를 결정해야 한다. 평면의 경우 면적이 되었지만 곡면의 경우 면적이 없으므로 근사값을 구한다. 그림 4 에서 보듯이  $H(n, p)$  의 값은, 삼차원 Hough 변환 공간의 의미 대로라면, 두께가  $\Delta p$  인 판

에 포함된 원통면 상에 있는 데이터의 수가 된다. 이 데이터들을 센서의 방향에서 바라보고 취득하게 되기 때문에 빗금친 부분으로 사영을 내린 면적이 실제 획득할 수 있는 데이터의 개수가 된다. 이를 구해 보면 다음과 같다( $d$  는 단면의 직사각형 폭의 반,  $h$  는 원통면의 높이).

$$\begin{aligned} H(n, p) &= 2 \times d \times h \\ &= 2 \times \sqrt{r^2 - (r - \Delta p)^2} \times h \\ &= 2 \times \sqrt{2r\Delta p - \Delta p^2} \times h \end{aligned} \quad (5)$$

만약 양자화 수준  $\Delta p$ 가 물체의 크기에 비해 아주 작다면 다음과 같이 된다.

$$H(n, p) \approx 2h\sqrt{2r\Delta p} \quad (6)$$

③ 구면

구면의 경우 IGES 캐드 형식으로부터 반지름, 중심점 등을 얻을 수 있다. 역시 원뿔면의 경우와 같이 구의 중심점을 좌표계의 중심점으로 옮겨 놓고 삼차원 Hough 변환 공간을 구해 보자. 구면의 경우 모든 방향의 단위 벡터에서 삼차원 Hough 변환 공간 값이 존재한다. 이 때  $p$  는 모든 방향에 대해서 구의 반지름  $r$  이 될 것이다. 그림 5 를 통해  $H(n, p)$  의 값을 구해 보면 다음과 같다( $d$  는 단면의 원의 반지름,  $r$  은 구의 반지름).

$$\begin{aligned} H(n, p) &= \pi \times d^2 \\ &= \pi \times (\sqrt{r^2 - (r - \Delta p)^2})^2 \\ &= \pi \times (2r\Delta p - \Delta p^2) \\ &\approx 2\pi r\Delta p \end{aligned} \quad (7)$$

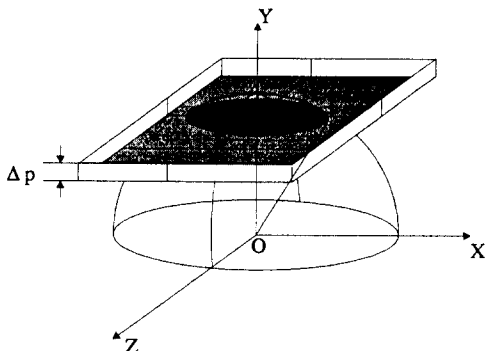


그림 5. 구면의  $H(n, p)$   
Fig. 5.  $H(n, p)$  of spherical surface.

IV. 정 합

1. 개요

IGES 캐드 형식으로부터 얻어진 기준 물체의 삼차

원 Hough 변환 공간  $H(n, p)$  와 입력 데이터로부터 얻어진 삼차원 Hough 변환 공간  $H'(n, p)$  를 비교하여 물체를 찾아내는 것이 물체 인식의 마지막 단계이다. 또한 물체의 회전 이동량과 평행 이동량을 추정해 내는 것이 이 단계이다. 본 논문에서는 기준 물체의 회전, 평행 이동을 가정하여 이것과 입력 영상과의 정합을 통해 물체의 존재 여부를 확인한다. 즉, 가정한 회전, 평행 이동량에 대해 일정 정도 이하의 오차를 나타내면 물체가 존재한다고 판단하는 것이다.

2 장에서 설명한 삼차원 Hough 변환 공간  $H(n, p)$  의 회전, 평행 이동에 대한 성질을 이용하여 기준 물체의  $H(n, p)$  를 변환시켜 가며 입력 영상의  $H'(n, p)$  와 비교하면 물체의 존재 여부와 함께 회전, 평행 이동량을 구할 수 있다. 하지만 삼차원 공간상에서 고려할 수 있는 회전, 평행 이동량의 수가 많음을 고려할 때 이들을 직접 비교하는 것은 너무 많은 계산량을 필요로 한다. 따라서 고차원 파라미터 공간을 저차원 파라미터 공간 몇 개로 분할하는 과정이 반드시 필요하다. 지금 찾고자 하는 파라미터는 회전 이동에 대한 파라미터 세 개, 평행 이동에 대한 파라미터 세 개를 합하여 모두 6 개이다. 이 6 차원 파라미터 공간을 먼저 회전 이동과 평행 이동에 대한 삼차원 공간 두 개로 나누어 접근한다. 이는 삼차원 Hough 변환 공간  $H(n, p)$  는  $n$  축이 방향을  $p$  축이 거리를 나타내고 있기 때문에 가장 쉬운 접근 방법이라고 할 수 있다. 이 중 회전 이동에 대한 파라미터는 다시 이차원과 일차원으로 나누어진다. 이는 다음 절에 설명될 것이다.

2. 삼차원 Hough 변환 공간에서의 방향 특성

고차원 파라미터 공간의 분할에 따라 먼저 삼차원 파라미터 공간인 회전 이동량에 대한 산출을 행하게 된다. 삼차원 Hough 변환 공간  $H(n, p)$  에서  $n$  축이 방향을 나타내므로 기준 영상과 입력 영상의 삼차원 Hough 변환 공간에서  $n$  축 상에서의 비교가 요구된다. 즉,  $H(n, p)$  중  $p$  는 무시하고  $n$  축으로 사영을 내려 수열  $H(n)$  을 구한 후에 후에 이들을 일차원 상에서 비교하면 된다. 이  $n$  에 관한 수열  $H(n)$  은 삼차원 Hough 변환 공간 중 방향에 관한 특성을 나타내는 수열이 된다. 먼저  $H(n)$  을 구하는 방법을 고려해 보자.

여기서  $n$  에 관한 수열을 구한다는 것은 각 단위 벡터가 가리키는 방향에 대한 정보를 구한다는 것이다.

정해진 방향으로 데이터들이 어떤 특성을 가지는지(예를 들면 면이 존재하는지)에 관한 정보를 구하는 것이다. 2장에서 설명되었듯이 삼차원 Hough 변환 공간에서는 쉽게 면을 검출할 수 있다. 즉, 특정 방향  $n$  이 정해져 있을 때  $p$  에 대한  $H(n, p)$  의 그래프는 면이 존재할 때는 뾰족한 극대치를 나타낸다 (그림 3 참조).  $H(n, p)$  의 값이 정해진 문턱값 이상이고  $p$  방향의 이웃에 대해 아주 크다면  $n$  의 방향으로 면이 존재한다고 판단할 수 있다. 이 때 극대치를 나타내는  $p$  는 면의 위치를,  $H(n, p)$  의 값은 면적(실제는 면이 포함하고 있는 데이터의 개수)을 나타내게 된다. 이 면적값을  $H(n)$  에 저장해 이를 회전 이동량 산출을 위한 정합에 이용하게 된다.

그런데 단위 벡터는 양자화 되어 있는 것이기 때문에 실제 데이터의 면의 방향은 이 한정된 단위 벡터와 일치하지 않을 때가 많다. 이 경우  $H(n, p)$  의 값은 상대적으로 작아지고 대신 주위의  $H(n, p+1)$  과  $H(n, p-1)$  의 값이 커지게 된다. 즉, 두께가 양자화 수준  $\Delta p$  인 판이 면에 포함된 데이터를 한꺼번에 모두 포함하지 못하고 주변의 판들이 일부 데이터들을 포함하게 되는 것이다. 따라서 보다 정확한 면적값을 추출하기 위해서 극대치 주변의 몇 개를 검사해 적정 선까지 합하여 이 값을  $H(n)$  으로 이용한다. 그림 6 이 이를 나타내고 있다.

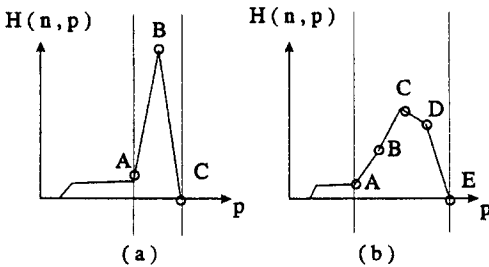


그림 6. 면의 검출과 면적 산출  
 (a)  $S = B$  (b)  $S = B + C + D$   
 Fig. 6. Detection of the plane and evaluation of the area.

3. 회전 이동량 추정

앞에서 설명한 대로 구한 수열  $H(n)$  을 가지고 기준 영상과 입력 영상을 비교하여야 한다. 이 때는 가능한 모든 회전 이동에 대해 고려를 해야 하는데 삼차원에서의 회전 이동은 이차원과 비교하여 상당히 많은 가지 수가 존재한다. 예를 들어 회전 이동량에 대한 해

상도를  $6^\circ$  가량으로 한다고 가정하자. 단위구 상에 존재할 수 있는 회전축의 수는 1,000 개 가량이고 이에 대해 60 가지의 회전이 가능하므로 모두 60,000 가지나 됨을 알 수 있다. 이는 60 가지밖에 안되는 이차원의 경우와 비교하면 엄청나게 많은 수이다.

이렇게 60,000 가지의 가능한 회전 변환에 대한 기준 영상의  $H(n)$  을 미리 저장해 놓으려면 대략 모델 하나에 대해 60 Mbyte 가 필요해 모델이 많아질 경우 저장에 어려워진다. 한편 60,000 가지나 되는 회전 변환을 매번 정합시에 계산하는 것도 무리가 따른다. 본 논문에서는 이 두 가지의 절충안으로 회전 이동에 관한 삼차원 파라미터를 다시 이차원과 일차원으로 분할해 이 중 이차원 파라미터에 해당하는 부분을 미리 계산하여 저장해 놓고 일차원 파라미터에 해당하는 부분을 정합시에 계산하도록 하였다.

모델 생성시에 미리 계산되는 부분은 시점에 관한 정보이다. 이는 이차원 파라미터이다. 먼저 단위 구면에 고르게 분포하는 단위 벡터들을 시점으로 설정한다. 이 단위 벡터의 방향에서 바라보았을 때 보이는 면들을 결정하고 보인다면 시점을 기준으로 어떤 위치(삼차원 Hough 변환시 설정된 단위 벡터 중에 선택)에서 보이는지, 또 얼마만큼의 면적으로 보이는지를 계산해 각 시점에 대해  $H(n)$  을 구해 놓는다. 이는 불룩한 물체의 경우 단순히 면의 방향과 시점간의 내적을 이용하면 쉽게 계산이 가능하다. 오목한 물체의 경우 물체의 가리어짐을 고려해야 하기 때문에 다소 복잡해진다. 이는 aspect graph 의 생성과 유사한 과정을 통해서 가능하다. Aspect graph 는 물체를 보는 시점에 따라 모양이 어떻게 변하는가에 관한 정보를 담고 있다. 한편 물체의 대칭성을 이용한다면 단위구면 상의 모든 시점을 고려할 필요가 없다. 예를 들어 물체가 한 면에 대해 대칭이라면 반구면 상의 단위 벡터로 충분하고 회전체라면 회전축을 지름으로 하는 반원 상의 호에 있는 단위 벡터로 충분하다. 이를 통해 정합 대상을 크게 줄일 수 있어 시간 상의 이득을 볼 수 있다.

정합시에 계산되는 부분은 시점이 고정되었을 때 그 시점을 회전축으로 한 회전이다. 본 논문에서 삼차원 Hough 변환에 사용된 단위 벡터들은 관찰자가 보는 시점을 극점으로 하는 구 좌표계를 이용해 생성이 되었다. 예를 들어 반구면 상에 단위 벡터 256 개를 잡는다면 천정각을  $7.5^\circ$  단위로 쪼개 반구 상에 12 개의 위선을 긋고 그 위선 상에 단위 벡터를 분포시키되 극

점에서는 1 개로 시작해 적도로 갈수록 많은 단위 벡터를 분포시켜 전체적으로 균일하도록 한다. 그림 7에 그 결과가 있다. 그림에서 극점으로부터의 위선에 분포하는 단위 벡터의 수는 1, 5, 9, 14, 18, 22, 25, 28, 31, 33, 35, 35 로 모두 합이 256 개가된다. 이렇게 단위 벡터들을 설정함으로 인해서 시점을 회전축으로 하는 회전 변환 결과의  $H(n)$  은 아주 쉽게 얻을 수 있게 된다. 즉, 기준 영상의  $H(n)$  을 같은 위선 상에 있는 것끼리 순환 이동시키는 것으로 결과를 얻을 수 있다. 이로 인해 정합시에는 실시간의 빠른 계산이 가능해 진다.

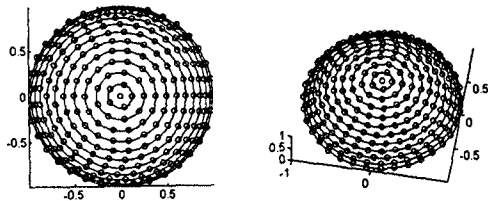


그림 7. 구 좌표계를 이용한 단위 벡터  
Fig. 7. Unit reference vector using the spherical coordinate system.

정합 점수는 회전을 가정한 모델의  $H(n)$  과 입력의  $H(n)$  사이의 내적으로 정해진다. 또한 모델의  $H(n)$  에서 국소최대치를 나타내는  $n$  에서는 반드시 입력의  $H(n)$  에서도 국소최대치를 나타내어야 한다. 즉, 모델에서 그 방향으로 면이 존재한다고 알고 있으면 반드시 입력 데이터에 이것이 검출되어야 한다는 것이다.

#### 4. 평행 이동량 추정

위의 정합 결과 정합 점수가 높은 몇 개를 후보로 선정해 평행 이동량을 계산한다. 기준 물체를 구해진 회전량만큼 회전 이동 시킨 결과를 먼저 구한다. 이 결과 기준 물체의 여러 면들 중에서 관찰자의 시점에서 보이는 면들을 골라 낼 수 있고 이 면들의 방향에서의  $p$  값을 입력 영상과 비교함으로써 평행 이동량을 추정하게 된다. 여기에서, 관찰자에게 보이는 면들의 수가 적어도 세 개 이상이 되어야만 평행 이동량을 최소자승법의 해로서 구해 낼 수 있다. 만약 물체가 아주 단순한 모양이어서 관찰자의 시점에서 두 면 또는 한 면만이 보인 영상이 입력된다면 평행 이동량을 완전히 구할 수 없다. 예를 들면 정육면체의 한 면만이 보인다면 면은 한 방향에서만 검출될 것이고 이 방향에 대한  $p$  의 비교 결과  $Z$  방향의 평행 이동량은 알 수 있지

만  $X, Y$  방향의 평행 이동량은 알 수 없다. 이 경우는 입력 영상에서 경계에 관한 정보를 별도로 추출해 이용해야만  $X, Y$  방향의 평행 이동량을 추정할 수 있을 것이다.

최소자승법의 해를 통해 평행 이동량을 구한 후에 다시 이 평행 이동량에 의거한  $p$  값을 구해서 입력 영상과의 오차를 구할 수 있다. 이 오차를 통해 최종적으로 물체의 존재 여부를 판단하게 된다.

### V. 실험 결과

입력 영상의 획득을 위해 캐나다 Servo-Robot 사의 레인지 파인더를 사용하였다. 이 레인지 파인더는 auto-synchronized scanning 방식을 사용한다. 실험에 사용한 레인지 파인더는 근거리에서  $Y$  축 방향으로 108 mm 를 해상도 ( $Z$  축 방향 깊이) 0.03 mm 로 스캔할 수 있고, 원거리에서 579 mm 를 해상도 1.4 mm 로 스캔할 수 있다. 한 라인을 스캔할 때  $Y$  축 방향으로 256 (또는 512) 개의 데이터를 얻을 수 있다.

삼차원 Hough 변환에 사용하는 단위 벡터는 전장에서 설명하였듯이 구 좌표계를 이용하여 256 개를 설정하였다. 이 때 평균적인 각도의 해상도는  $8^\circ$  정도가 된다. 한편  $p$  축의 양자화 수준  $\Delta p$  는 1 mm 로 하였다. 입력 영상은 단일 물체 영상의 경우 한 스캔당 256 개의 데이터를 0.5 mm 간격으로 256 스캔을 얻어  $256 \times 256$  크기의 깊이 영상을 사용하였다. 대상 물체로는 최대 길이 100 mm 정도의 크기를 가지는 다면체 10 개를 선정하였다.

알고리즘은 IBM PC Pentium 66 상에서 C++ 언어를 이용하여 구현하였다. 정합에 걸리는 시간은 입력 영상에 따라 다르지만 대체로 한 물체와 비교하는 데 2 초 안팎이 소요되었다. 기준 물체 10 개를 대상으로 실험하였기 때문에 30 초 이내로 정합이 가능하였다. 정합에 걸리는 시간에 비해 삼차원 Hough 공간을 구하는데 더 많은 시간(2 분 정도)을 필요로 한다.

그림 8 은 정삼각뿔 모양의 단순한 물체에 대한 결과이다. 그림 (a) 는 이 물체의 기준 위치를 나타내고 있다. 그림 (b) 는 입력 거리 영상이고 그림 (c) 는 이로부터 얻은 의사 밝기 영상이다. 그림 (d) 는 삼차원 Hough 변환 공간  $H(n, p)$  로부터 구해진  $H(n)$  이다. 입력 영상에 나타나는 면이 세 개뿐임에도 불구하고

하고 세 개 이상의 방향에서 큰 국소최대치를 나타내는 것을 볼 수 있다. 이는 입력 영상에 나타난 면의 방향과 단위 벡터가 정확히 일치하지 않고 오차가 있기 때문에 주위의 몇 개의 방향에서 동시에 국소최대치가 나타나는 것이다. 그림 (e) 는 미리 생성된 모델 중 45 번째 열을 선택해서 나타낸 것이다. 그림 (f) 는 이것을 172.5 도 회전시켰을 때 나타나는  $H(n)$  이다. 이를 그림 (d) 와 비교해 보면 모델에서 나타나는 국소최대치가 입력 영상에서 모두 큰 값으로 나타나는 것을 확인할 수 있다. 결과적으로 45 번째 시점을 중심으로 172.5 도 회전시킨 상태를 후보로 선정하게 되고 결국 이것을 토대로 평행 이동량을 구해 검증해 보면 오차가 작아 물체가 존재한다고 판단하게 된다.

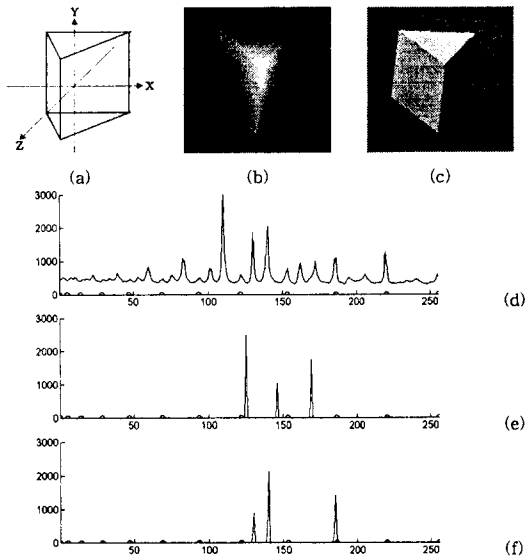


그림 8. 단순한 물체에 대한 분석  
Fig. 8. Analysis of the simple object.

다음의 표는 정삼각뿔 모양의 물체에 대한 기존의 방법과의 비교 결과이다. 기존의 방법으로는 주로 Hoffman<sup>[9]</sup>의 방법을 이용하였다. 먼저 영상을 영역 분할해서 면을 구분해 내고 평면으로 근사 시켜 평면의 방정식을 구한다. 정합 단계에서는 입력 영상에서의 어떤 면이 기준 모델의 어떤 면에 해당되는가를 육안으로 판단하여 지정해 주었다.

	회전이동( $\alpha, \beta, \gamma$ ) 도	평행이동(x,y,z) mm
제안 방법	( -26.6, -14.5, 175.9 )	( -4.3, 4.2, 31.2 )
기존 방법	( -29.1, -8.5, 177.9 )	( -2.7, 2.5, 35.0 )
차이	( -2.5, 6.0, 2.0 )	( 1.6, -1.7, 3.8 )

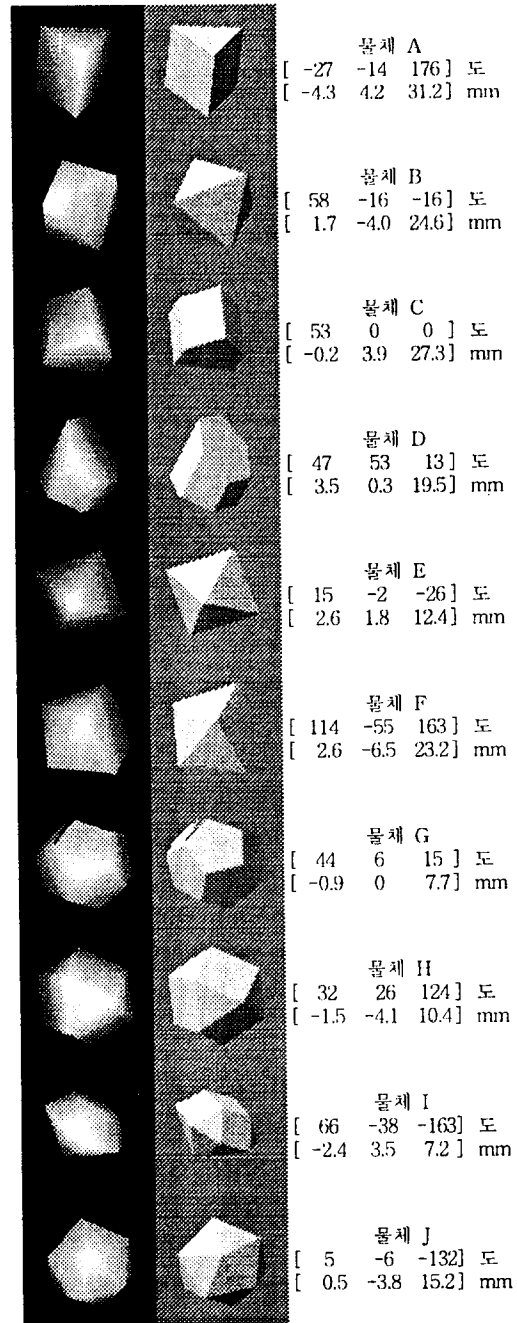


그림 9. 단일 물체 영상에 대한 실험 결과  
Fig. 9. Results on the single object images.

위에서 보이는 차이는 본 논문에서 설정한 256 개의 단위 벡터들의 해상도가 대략 8 도임을 고려할 때 만족할 만한 수준이다. 기존의 방법에도 오차가 있기 때문에 두 방법의 차이가 본 논문에서 제시한 결과에 대



한 오차라고는 말할 수 없다. 하지만 두 방법이 아주 비슷한 값을 나타내는 것을 통해 본 논문이 제안하는 기법의 타당성을 검증할 수 있다.

그림 9 는 단일 물체 영상에 대한 결과를 나타내고 있다. 결과의 그림들은 각각 거리 영상과 의사 사탕기 영상이다. 각 입력 영상에 대해 인식 결과로써 회전량과 이동량을 구할 수 있다. 10 매의 입력 영상 모두 정확히 인식이 되는 것을 확인할 수 있다.

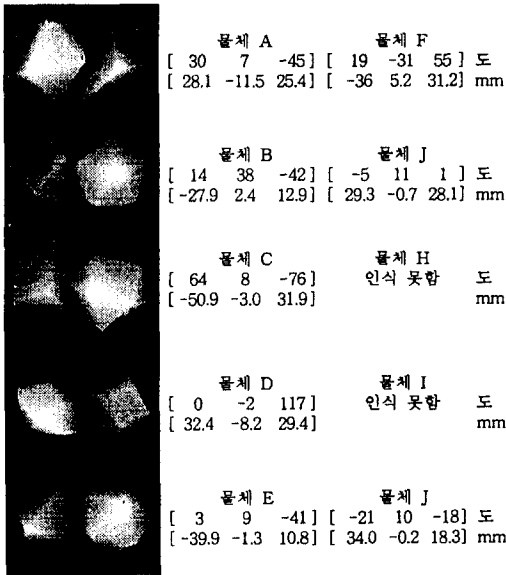


그림 10. 다중 물체 영상에 대한 실험 결과  
Fig. 10. Results on the multiple object images.

그림 10 은 두 개의 물체가 포함되어 있는 영상에 대한 결과이다. 다중 물체 영상은 스캔 수를 늘려서  $340 \times 256$  크기의 거리 영상을 사용하였다. 다중 물체 영상의 경우 단일 물체 영상에 비해 어려운 점은 정해진 단위 벡터의 방향에 대해 물체가 여러 개 있기 때문에 여러 개의 국소최대치를 가질 수 있다는 것이다. 여러 개의 국소최대치를 가질 경우 정합의 첫 단계부터 약간의 혼선이 생긴다. 입력 영상의  $H(n)$  과 특정 모델의 회전을 가정한  $H(n)$  을 비교할 때 입력 영상의  $H(n)$  은 두 개 이상의 물체를 복합해 놓은 결과이므로 잘못된 정합 결과를 낼 수 있다. 물론 평행 이동량을 이용한 검증 단계에서 잘못된 결과는 탈락은 되겠지만 이들로 인해 정확한 결과가 상대적으로 정합 점수가 낮게 간주되어 검증 대상 후보에 들지 못할 경우 인식을 못하게 된다. 이로 인해 인식률이 떨어진다.

그림 10 에서 물체 H 와 물체 I 가 그 예이다. 검증 대상 후보를 많이 잡으면 어느 정도 해결할 수 있었지만 이 경우 계산량이 늘어나게 된다. 한편 평행 이동량을 이용한 검증 단계에서 국소최대치가 나타나는 위치를 근거로 하여 평행 이동량을 구하게 되는데 이 위치가 두 곳 이상에서 나타나기 때문에 모두 검증 대상으로 삼을 경우 계산량이 늘어나게 된다.

## VI. 결론

본 논문에서는 모델 기반 삼차원 물체 인식 문제를 고찰하였다. 삼차원 물체의 표현 방법으로서 삼차원 Hough 변환을 사용하여 이에 적합한 물체의 모델링 과정과 정합 과정을 연구하였다.

물체의 모델링 과정에서는 정합에 필요한 특징들을 캐드 모델로부터 자동으로 추출해 내었다. 모델의 삼차원 Hough 변환 공간을 구하는 과정은 분석적인 방법으로 이루어지기 때문에 모델 자체의 오차 요인을 극소화시킬 수 있다. 또한 모델링 과정에서 물체의 시점에 따라 달라지는 면의 특성들을 미리 저장해 놓음으로써 실제 인식 시간을 단축시킬 수 있다. 캐드 모델을 사용함으로써 대상 물체의 개수가 많아져도 쉽게 물체를 모델링 할 수 있어 데이터베이스의 확장이 용이하다.

정합시에는 입력 영상과 왜곡(회전 이동과 평행 이동)을 가정한 기준 물체를 삼차원 Hough 변환 공간 상에서 비교하였다. 특히 왜곡을 정의하는 고차원 파라미터 공간을 저차원으로 나누어 인식의 효율성을 높였다. 먼저 왜곡을 정의하는 인자를 회전 이동 인자와 평행 이동 인자로 나누고, 회전 이동 인자는 다시 물체를 바라보는 시점에 관한 인자와 회전 각도에 관한 인자로 나누었다. 이 중 시점에 관련된 인자는 캐드 모델을 이용한 모델의 생성 과정에서 미리 계산되어 저장되어 있기 때문에 실제 인식에 필요한 시간이 단축된다. 또한 회전 각도에 관한 인자도 삼차원 Hough 변환을 위한 단위 벡터들을 적절히 설정해 빠른 계산이 가능하게 하였다. 입력 영상의 삼차원 Hough 변환 공간을 직접 정합에 이용하기 때문에 잡음에 대한 민감성과 계산의 복잡성이 크게 감소되었다.

여러 대상 물체를 선정해 실험한 결과 본 논문에서 제안하는 기법의 타당성이 검증되었다. 한 장면에 하나의 물체가 있는 경우에 정확하게 인식을 수행하였음은

물론이고 두 개의 물체가 존재하여 부분적인 가리어짐이 있는 경우에도 좋은 결과를 얻을 수 있었다.

이 시스템은 공장 자동화를 위한 시각 인식 분야에 응용되어질 수 있다. 특히 다양한 모델을 생산해 내는 FMS(Flexible Manufacturing System) 에서 모델의 변화에 쉽게 적응하며 빠른 시간 안에 물체를 인식할 수 있다.

앞으로의 연구 방향으로는 대상 물체의 확장을 들 수 있다. 특히 곡면이 포함된 물체에 대한 실험적인 검증이 필요하다. 한편 삼차원 Hough 변환 공간 상에서 면의 방향 정보를 가지고 정합을 하므로 단순한 물체의 경우 보이는 면의 수가 적으면 이동량을 추정해 내지 못하는 경우가 생긴다. 이를 해결하기 위해 거리 영상의 경계 정보를 부분적으로 이용하는 것이 앞으로 고려되어야 할 것이다. 또한 인식에 필요한 시간을 단축시키기 위해서는 삼차원 Hough 변환의 고속 알고리즘 개발이 필수적일 것이라 생각된다.

#### 참 고 문 헌

- [1] F. Arman and J. K. Aggarwal, "Model-based object recognition in dense-range images - a review," *ACM Computing Surveys*, Vol. 25, No. 1, pp. 5-43, Mar. 1993.
- [2] B. Bhanu, "Guest editor's introduction: CAD-based robot vision," *Computer*, Vol. 20, No. 8, pp. 13-16, Aug. 1987.
- [3] F. Arman and J. K. Aggarwal, "CAD-based object recognition in range images using pre-compiled strategy trees," in *Three Dimensional Object Recognition Systems*, A. K. Jain and P. J. Flynn, Ed., Elsevier, Amsterdam, pp. 115-134, 1993.
- [4] P. J. Flynn and A. K. Jain, "CAD-based computer vision: From CAD models to relational graphs," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 2, pp. 114-132, Feb. 1991.
- [5] P. J. Flynn and A. K. Jain, "BONSAI: 3D object recognition using constrained search," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 10, pp. 1066-1074, Oct. 1991.
- [6] R. Krishnapuram and D. Casasent, "Determination of three-dimensional object location and orientation from range images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, No. 11, pp. 1158-1167, Nov. 1989.
- [7] D. C. W. Pao and H. F. Li, "Shapes recognition using the straight line hough transform : theory and generalization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No. 11, pp. 1076-1089, Nov. 1992.
- [8] K. Reed, D. Harrod, Jr., and W. Conroy, *The Initial Graphics Exchange Specification (IGES) Version 5.0*, U.S. Department of Commerce, Sep. 1990.
- [9] R. Hoffman and A. K. Jain, "Segmentation and classification of range images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 9, No. 5, pp. 608-620, Sep. 1987.

#### 저 자 소 개



具 滋 性(正會員)

1970年 4月 21日生. 1993年 2月 서울대학교 공과대학 제어계측공학과 졸업. 1995年 2月 서울대학교 대학원 제어계측공학과 졸업(공학석사). 1995年 현재 동 대학원 박사과정 재학중. 주

관심분야는 컴퓨터비전.

李 商 郁(正會員) 濟 29卷 B編 第 12號 參照.

현재 서울대학교 공과대학 제어계측공학과 교수