

論文93-30B-7-7

# 면 법선 영상 기반형 3차원 물체인식에서의 새로운 매칭 기법 (A New Matching Strategy for SNI-based 3-D Object Recognition)

朴鍾勛\*, 崔宗秀\*\*

(Jong Hoon Park and Jong Soo Choi)

## 要約

본 논문은 면 법선 영상 기반형 인식을 위한 좀 더 효율적이며, 신뢰성있는 새로운 매칭 기법을 제안한다. 기존에 면 법선 영상 기반형 인식을 위해 제안된 바 있는 유사도 결정함수를 이용하는 매칭 기법은 모델 베이스에 속한 모든 모델 물체들의 특징들을 장면 물체의 특징들과 비교해 보아야 하며, 특징들의 대표 속성들에 부여된 가중치가 경험적 방법에 의존할 수 밖에 없다는 한계를 가지고 있다. 본 논문에서 제안하는 매칭 기법은 장면 물체로부터 얻어진 회전 입력 영상들에서 추출된 특징들을 모두 만족하는 모델 물체를 모델 베이스에서 탐색해내는 방법을 사용함으로써, 유사도 결정함수를 사용하는 기존의 방법에서 나타난 효율성과 신뢰성 저하의 문제를 해결할 수 있다.

본 방법의 구현 가능성을 확인하기 위해서, 26개의 모델 물체로 모델 베이스를 구축하고, 이들을 포함하는 거리 영상들을 대상으로 실험을 행 하였다. 실험 결과, 본 방법이 면 법선 영상 기반형 인식에 효율적으로 적용될 수 있다는 가능성을 확인하였다.

## Abstract

In this paper, a new matching strategy for 3-D object recognition, based on the Surface Normal Images(SNIs), is proposed. The matching strategy using the similarity decision function(9,10) lost the efficiency and the reliability of matching, because all features of models within model base must be compared with the scene object features, and the weights of the attributes of features is given by heuristic manner. However, the proposed matching strategy can solve these problems by using a new approach. In the approach, by searching the model base, a model object whose features are fully matched with the features of scene object is selected.

In this paper, the model base is constructed for the total 26 objects, and synthetic and real range images are used in the test of the system operation. Experimental result is performed to show the possibility that this strategy can be effectively used for the SNI based recognition.

\*正會員, 韓國電子通信研究所 미디어 應用研究室  
(Media Application Section, ETRI.)  
接受日字: 1992年 9月 29日

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

I. 서론

3차원 물체 인식은 컴퓨터 시각의 궁극적 목표이며, 많은 연구자들에 의해 연구되어 온 분야이다. 여러가지 형태의 인식 방법들 중에서도 인식하고자 하는 물체의 모델을 특징(feature)의 형태로 기술하고, 입력되는 영상으로부터 추출된 특징들과 비교하여 영상내에 존재하는 물체의 종류와 위치 및 자세를 인식해내는 모델 기반형 인식기법이 가장 널리 사용되고 있다. [1,2,3] 모델 기반형 3차원 물체 인식은 모델의 특징(feature)들과 장면 물체의 특징들을 비교함에 의해 이루어진다. 그러나 불행하게도 3차원 공간 상에 놓인 물체는 관찰자의 위치에 따라 그 형상이 심하게 달라지게 된다. 그러한 이유로 모델과 장면 물체의 비교를 위해서는 양자를 같은 형상을 가지도록 하는 변환(transformation) 작업이 필요하다. 결국, 이 변환 과정은 비교를 위해 모델의 특징들과 장면 물체의 특징들 사이의 정합(correspondence)를 찾는 과정이 되며, 따라서 모델과 장면 물체간의 정합을 위한 정확한 변환을 찾는 문제는 인식에 있어 가장 중요한 연구 과제중의 하나이다. 일반적으로 변환은 3축 방향으로의 병진운동(translation)과 3축을 중심으로 하는 회전(rotation)을 뜻하며, 실제로 이를 찾는 문제는 그리 쉽게 풀리지 않는 어려운 문제이다. 그림 1은 3차원 물체 인식에서의 정합문제(correspondence problem)를 설명해 준다.

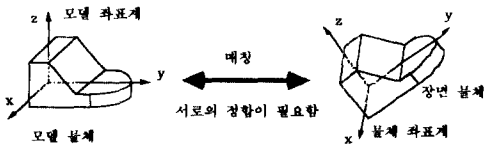


그림 1. 3차원 물체 인식에서의 모델과 장면 물체 간의 정합 문제

Fig. 1. Correspondence problem between a model and a scene object in 3-D object recognition.

그림 1에서 볼 수 있듯이, 모델은 모델 좌표계에 의해 기술되어 있고, 장면 물체는 관찰자 시선 방향에서의 물체 좌표계에 의해 나타나므로, 그 사이의 정합이 필요하게 된다. 입력 영상이 2차원 영상인 경우에는 이 문제가 더욱 심각해져서 2차원에서 3차원, 혹은 3차원에서 2차원으로서 투영(projection)이 매우 어렵다. 입력 영상이 3차원 영상, 즉 거리 영상(range image)인 경우라도 양 좌표간의 정확한 변

환(transformation)을 찾기란 쉬운 일이 아니다. [4,5] 결국, 모델물체나 장면물체중 어느 하나를 변환하여 같은 형상을 갖도록 해 주어야 하겠으나, 그를 해결하기 위해서는 장면 물체의 자세(pose)에 대한 가설 설정(hypothesis generation)과 가설 검증(hypothesis verification)이 행해져야 하므로, 그래프탐색(graph search) [7] 이나, 반복적 혹은 비반복적인 최소화(minimization) [5,6,8] 과정이 필요하게 된다. 그러므로 정확한 투영 혹은 변환을 찾기 위해 너무 많은 시간과 노력이 요구된다. 만약, 입력 영상에서 물체들의 중첩(occluding)에 의해 형상의 왜곡이 일어 난다면, 그 물체들의 인식은 더욱 더 힘들어진다.

이와 같은 정합의 어려움을 줄여보고자 한 노력의 하나로 면 법선 영상 기반형 물체 표현방법이 제안된 바 있다. [9,10] 면 법선 영상 기반형 인식 방법은 각면의 법선 방향을 정합을 찾기 위한 기준으로 설정하고, 모델과 장면 물체를 그 기준에 의해 강제로 회전하여 비교함에 의해 정합의 어려움을 줄일 수 있었다. 박 중훈등 [9,10] 에서 이러한 물체 표현 방식을 3차원 물체 인식에 적용한 예로서 모델 구축과 매칭에 대한 한가지 방향이 제시되었으나, 이는 인식과정에서 모델과 장면 물체의 각 특징들 간의 비교과정에 유사도 결정함수를 사용하여 모델의 모든 면 법선 영상들에 대한 비교를 함으로써, 매칭의 효율성을 저하시켰다. 또한 유사도를 검증하기 위해 특징들의 각 속성(Attribute)들에 부여된 가중치들의 결정이 경험적인 방법에 의존할 수 밖에 없었다. 그러나 이들 가중치들의 변화에 의해 인식의 결과가 많은 영향을 받을 수 있기 때문에 인식의 신뢰도를 저하시키는 결과를 낳았다.

본 논문에서는 기존에 제안되었던 면 법선 영상 기반형 인식 기법의 단점들을 보완하기 위한 새로운 매칭 전략을 제안한다. 본 논문에서 제안된 매칭 기법은 인식 과정에서 유사도 평가 방법을 사용하지 않고, 장면 물체로부터 추출되는 특징들을 모두 만족하는 모델을 모델 베이스에서 찾아내는 매칭 방법을 사용함으로써, 기존의 방법 [9,10] 에 비해 구현이 다소 복잡하다는 단점이 있으나, 매칭의 신뢰성과 효율성을 높여 줄 수 있다는 장점을 가진다. 본 논문의 II장에서는 면 법선 영상의 기본 개념과 기존의 면 법선 영상 기반형 물체 인식 방법을 간단히 설명하고, III장에서는 본 논문에서 제안하는 매칭 기법을 소개한다. IV장에서는 제안된 기법을 3차원 물체 인식에 적용한 예로 실험한 결과를 보이고, 그에 대한 검토를 한다. 마지막으로 V장에서 결론을 맺기로 한다.

## II. 면 법선 영상 기반형 3차원 물체 인식

### 1. 면 법선 영상(Surface Normal Image: SNI)

면 법선 영상 기반형 물체 표현 기법은 기존의 3차원 물체 표현 방법들에서 나타나는 정합의 문제, 즉 모델과 장면 물체의 각 특징들 간에서 매칭을 위한 정합을 얻기 어렵다는 문제를 줄여보고자 제안된 물체 표현 방법이다.<sup>[9,10]</sup> 이 기법은 물체의 표면(Surface Boundary)을 기본 요소로 하여 각 면의 정면 방향(면 법선 관찰 방향)에서 보여진 물체의 형상(면 법선 영상)들을 이용하여 물체를 표현한다. 그림 2에 한 물체에 대한 면 법선 영상들의 예를 보인다.

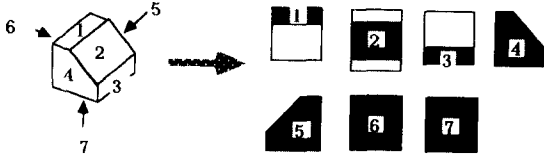


그림 2. 다면체의 면 법선 영상(■ 표시는 각 면 법선 영상의 기준면을 나타낸다)

Fig. 2. Surface normal images of a polyhedra. (The shaded surfaces are the base surfaces of the surface normal images).

그림 2에서 볼 수 있듯이 하나의 물체에서 그 물체의 면 수 만큼의 면 법선 영상이 얻어질 수 있다.

이러한 면 법선 영상의 개념은 한 면의 고유 형상은 그 면을 정면에서 보았을 때 가장 잘 나타난다는 생각으로부터 출발한 것이다. 한 물체는 그 물체의 면 수 만큼의 면 법선 영상들로서 표현이 가능하며, 입력되는 장면 물체에 나타난 각 면의 법선 벡터를 관찰자 방향과 일치하도록 회전한다면 같은 면을 기준으로 하는 면 법선 영상과 회전된 장면물체(Rotated Input Image: RII)는 똑 같은 형상을 가지게 된다. 따라서 모델로 구축되어 있는 면 법선 영상들과 장면 물체를 회전한 영상인 회전 입력 영상(RII)들 간의 비교에 의해 그 물체의 종류와 자세를 인식해 낼 수 있다. 그러므로, 면 법선 영상 기반형 인식에서는 모델과 장면 물체 간의 정확한 변환(Transformation)을 위한 기준(여기서는 면 법선 관찰 방향)이 제공될 수 있고, 그로 인하여 매칭을 위한 정합의 어려움은 상당히 줄어들게 된다. 그림 3에 면 법선 영상 기반형 인식의 기본적 흐름을 보인다.

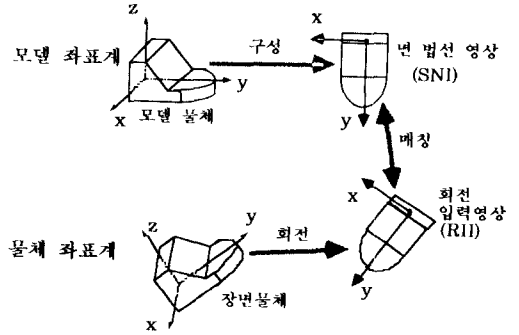


그림 3. 같은 기준면을 가지는 SNI와 RII간의 매칭  
Fig. 3. Matching between an SNI and an RII which have the same surface.

그림 3에서 볼 수 있듯이 모델과 장면 물체간의 3차원 매칭 문제는 면 법선 영상과 회전 입력 영상간의 2차원 매칭 문제로 단순화 될 수 있다.

### 2. 기존의 모델 구축 방법과 매칭 기법

박 중훈등<sup>[9,10]</sup>은 그 물체에서 얻어질 수 있는 면 법선 영상들을 기반으로 층 구조의 모델 베이스를 구축하였다. 하나의 물체는 소수의 면 법선 영상들로 구성되며, 하나의 면 법선 영상은 그 면 법선 영상의 기준 면과 그 인접면들로 구성된다. 그 기본 구조를 그림 4에 보인다. 그림 4에서 볼 수 있듯이, 한 물체에 대한 모델의 기본 구조는 물체 층(object layer), SNI 층(SNI layer), 면 층(surface layer) 등, 3개의 계층을 가지는 나무구조(3-layered tree structure)이다.

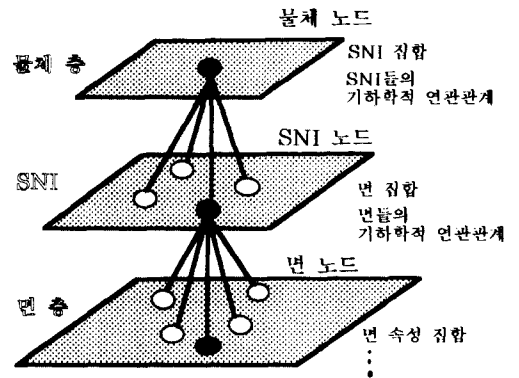


그림 4. 모델 베이스의 기본 구조  
Fig. 4. Basic structure of SNI-based model base.

꼭대기 층(top layer)의 물체 노드(object node)는 그 물체의 면 법선 영상(SNI)을 표현하는 부노드(subnode)들의 집합으로 구성되며, 각 SNI 노드(SNI node)는 그 SNI에 포함된 기준면과 인접면들을 표현하는 부노드들의 집합으로 구성된다. 이러한 모델의 구조는 물체 노드와 면 노드사이의 중간적 계층인 SNI노드가 존재함으로써, 좀 더 확고한 계층 구조를 이루게 된다는 점에서, 기존의 표면표현법(surface boundary representation)과 중요한 차이점을 가지게 된다.

한편, 박 중훈등<sup>[9,10]</sup>은 물체의 인식을 위한 매칭과정에서 각 특징들에 대한 유사도 결정함수를 사용하는 Osima 와 Shirai<sup>[11]</sup>의 접근 방식과 유사한 접근 방식을 사용하였다. 그를 위해 각 특징들의 속성(Attribute)들에는 그 중요도에 따라 가중치가 부여되었으며, 유사성 함수 값과 가중치의 곱의 합(weighted sum) 형태로 매칭 여부를 판단하는 유사도 결정함수가 사용되었다. 후보 모델 물체의 모든 면 법선 영상(SNI)들은 장면 물체의 모든 회전 입력 영상(RII)들과 비교되어 결정함수 값이 결정되며, 그 값이 가장 큰 모델 물체가 매칭된 것으로 선택된다. 이때, 모델 베이스내에 존재하는 SNI들과 입력 영상으로부터 만들어지는 RII들간에 매칭을 위한 비교 횟수  $P_s$ 는

$$P_s = \sum_{i=1}^{N_o} n_i \cdot m \quad (1)$$

이 된다. 여기서,  $N_o$ 는 모델 베이스내의 모델 물체 갯수이며,  $n_i$ 는 각 모델 물체들을 구성하는 SNI의 수,  $m$ 은 RII의 수를 나타낸다. 결국, 기존의 방법에서는 모델 물체에 포함된 모든 면 법선 영상들의 모든 특징들이 비교과정을 거쳐야 하는 비효율성을 감소해야 한다. 또한 각 속성치들의 가중치들은 인식 결과에 지대한 영향을 미치는데도 불구하고, 가중치의 결정이 경험적인 방법에 의존해야 하기 때문에 인식 결과의 신뢰도를 저하시키는 결과를 낳았다. 이와 같은 이유로 면 법선 영상 기반형 인식에 보다 적합한 매칭 기법의 개발이 필히 요구된다.

### III. 모델 베이스 탐색에 의한 매칭 기법

유사도 결정 함수에 의한 기존의 매칭 방법이 모델에 속한 모든 면 법선 영상들을 비교함으로써 매칭의 효율을 저하시킬 뿐만 아니라, 각 특징들의 속성들에 주어지는 가중치의 변화에 따라 엉뚱한 인식 결과를 초래 할 수 있으므로, 본 절에서는 모델 베이스를 탐

색하여 장면 물체에서 얻은 특징들을 모두 보유한 모델을 찾아내는 새로운 매칭 기법을 제안한다. 모델 베이스 탐색에 의한 매칭은 결정함수에 의한 방법에 비해, 신뢰도를 높여 줄 수 있고, 모델 베이스의 모든 면 법선 영상과의 비교가 필요 없으므로, 매칭 시간의 관점에서 볼때 보다 효율적이라는 장점이 있다.

면 법선 영상 기반형 인식 기법에서는 하나의 장면 물체로부터 그 면 수 만큼의 RII가 얻어질 수 있다. 이때, SNI와 RII들은 물체의 각 면을 기준면으로 하여 똑같은 방향(모델 좌표계의 Z 축 방향)으로 회전되어 있으므로, 동일한 기준면을 가지는 SNI와 RII는 똑같은 형상을 가지게 된다. 또한, 장면 물체의 면수는 모델의 면수보다 작으므로, 만약 한 모델 물체가 장면 물체와 성공적으로 매칭되려면 장면 물체의 RII집합이 후보 모델 물체(candidate model object)에 존재하는 SNI 집합의 부분 집합이 되어야 한다. 장면 물체와 모델 물체의 매칭 여부를 결정하기 위해서 장면 물체로부터 만들어 지는 RII들은 정해진 순서에 따라 차례대로 모델의 SNI들과 비교된다. 인식 과정이 시작될 때에는 모델 베이스 내에 존재하는 모든 모델 물체들이 장면 물체와 매칭될 후보가 되며, 인식이 진행되면서 후보에서 탈락하게 되어 유일한 후보가 남을 때까지 인식 과정이 계속된다. 따라서, 제안된 매칭 기법에서 SNI들과 RII들간에 매칭을 위한 비교 횟수  $P_m$ 은

$$P_m = \sum_{i=1}^{N_r} \sum_{j=1}^{N_{ci}} n_j \quad (2)$$

이 된다. 여기서,  $N_r$ 은 매칭이 끝날때까지 사용된 RII의 수이며,  $N_{ci}$ 는  $i$ 번째 RII 매칭시 남아있는 후보 모델 물체들의 수이며,  $n_j$ 는  $j$ 번째 후보 모델 물체의 SNI수이다. 위의 식 2를 식 1과 비교해볼 때,  $N_r$ 은 만들어진 RII의 수  $m$  보다는 작으며, 특히  $N_{ci}$ 는 모델 베이스내의 모델 물체 총수  $N_o$ 에 비해 매우 작게 되므로, 전체 비교 횟수  $P_m$ 은  $P_s$ 에 비해 현저히 줄어들게 된다. 제안된 방법에서 매칭의 효율을 좀 더 높여주기 위해서는 하나의 장면 물체로부터 얻어지는 복수개의 RII들을 어떠한 순서로 모델 베이스의 SNI들과 매칭해 나가야 하는가 하는 것이 매우 중요하다. 매칭 순서는 각 RII들에 속한 기준면들의 표면 유형(surface type), 상대 넓이(relative area), 그리고 면의 상황(situation, 즉 면의 기울어짐 정도, 중첩 여부등)에 따라 정해지는 것이 바람직하다. 그 기준은 다음과 같다.

- i) 평면이 곡면에 우선한다.
- ii) 면적이 큰 것이 우선한다.

iii) 중첩되지 않은 것이 우선한다.

이러한 기준은 곡면의 경우는 SNI가 근사(approximation)에 의해 만들어지기 때문에 평면에 비해 큰 오차를 낳을 가능성이 많으며, 면적이 넓은 면 그리고 중첩이 일어나지 않은 면이 오차에 의한 영향을 적게 받게 된다는 사실로부터 만들어진 것이다. 그림 7에 모델 베이스 탐색에 의한 매칭 과정의 개략도를 보인다.

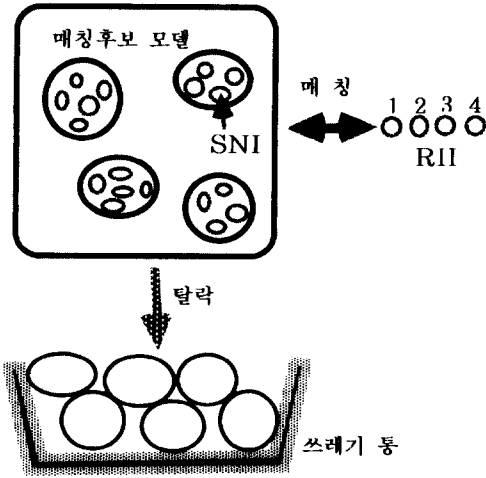


그림 7. 모델 베이스 탐색에 의한 매칭 과정

Fig. 7. Matching procedure by model base search.

장면 물체와 모델 물체간의 매칭 과정중 각 RII와 SNI의 매칭은 각 면 특성들 간의 매칭과 면과 면 사이의 연관관계 속성들(면간 접합각도, 접합유형 등)의 비교에 의해 이루어진다. 그 과정 중 가장 먼저, RII와 SNI에 포함된 면의 수가 비교된다. 서로 매칭이 되기 위해서는 RII의 면 수가 SNI의 면 수 보다 작거나 같아야 하므로, 다음의 관계를 만족하지 않으면 그 SNI는 매칭에 실패한 것으로 간주한다.

$$\text{number\_of\_surfaces(RII)} \leq \text{number\_of\_surfaces(SNI}_{i(j)}) \quad (3)$$

여기서 SNI<sub>i(j)</sub>는 i번째 모델 물체의 j번째 SNI를 의미한다. 위의 관계를 만족하는 SNI는 매칭 여부의 판정을 위해 RII와 SNI의 기준면들에 대한 매칭이 이루어지며, 그 후 기준면의 매칭이 성공적으로 이루어진 SNI들에 대해서는 그 인접면들의 매칭이 행해

진다. 이때, 인접면과 기준면과의 접합 각도가 참조되며, 모든 면 대 면 매칭이 성공적으로 이루어진 경우, 면 특성의 연관관계 속성들의 최종 점검을 거친 후에 매칭의 성공 여부가 판정된다.

RII와 SNI 간의 매칭 과정에서의 각 면 특징들 간의 매칭은 RII의 각 면 특징들로부터 구해진 속성치(attribute value)들과 모델 베이스의 SNI에 미리 저장된 속성치들과의 비교과정에 의해 이루어진다. RII의 한 면 특징의 모든 속성치들이 SNI의 한 면 특징의 속성치들과 일치하면, 그 SNI의 면이 RII의 면에 매칭된 것으로 한다. 그러나 속성치들 간의 매칭 과정에서 RII의 일부가 타 물체 혹은 자신의 일부에 의해 가려진 경우, 각 속성치들은 SNI의 속성치들과 정확히 일치할 수 없으므로 중첩에 의한 영향이 고려되어야 한다. 그래서, 본 논문에서는 물체의 각 속성들을 중첩에 의해 변화하는(occlusion-variant: OV) 속성과 중첩에 불변인(occlusion-invariant: OI) 속성의 두가지 부류로 나누고, 장면 물체의 일부가 가려진 경우, 속성치들의 비교 과정에서 중첩의 영향을 고려하도록 한다. OI 속성들의 경우, 예를 들면 표면 유형(surface type), 면 방향(surface orientation), 면간 접합각도(conjunction angle between surfaces), 등등의 경우, RII와 SNI의 각 속성치들은 서로 일치하게 된다. 즉,

$$OI_k(SR_i) = OI_k(SS_j) \quad (4)$$

의 관계를 만족해야만 한다. 여기서 SR<sub>i</sub>는 RII의 i번째 면 특징, SS<sub>j</sub>는 SNI의 j번째 면 특징이다. 또한, OI<sub>k</sub>(SR<sub>i</sub>)는 SR<sub>i</sub>의 k번째 속성치, OI<sub>k</sub>(SS<sub>j</sub>)는 SS<sub>j</sub>의 k번째 속성치를 뜻한다. 예를 들어, SR<sub>i</sub>와 SS<sub>j</sub>는 반드시

$$\text{surface\_type}(SR_i) = \text{surface\_type}(SS_j) \quad (5)$$

를 만족해야 한다. 그러나 OV속성들의 경우, 중첩에 의해 면의 일부가 가려져 있다면, SNI와 RII의 각 속성치들이 일치하지 않는다. 따라서, OV 속성들은 가려진 면(occluded surface)의 매칭 과정에 사용될 수 없다. 그러나, 중첩의 영향으로 발생하는 속성치들의 변화 양상을 찾아낼 수 있는 일부 속성들은 매칭 과정에서 유용하게 사용될 수 있다. 예를 들어, 인접면수, 혹은 구멍의 수 등은 중첩이 일어나 왜곡된 RII의 면에 대한 속성치들이 고유 형상을 지닌 SNI의 면에 대한 속성치 보다 항상 작다. 즉,

$$OV_k(SR) \leq OV_k(SS) \quad (6)$$

의 관계가 성립하므로, 매칭을 위한 정보로 사용될 수 있다. 그러나, 그외의 속성들은 중첩에 의해 그 값이 변화하므로 이를 매칭에 적용할 수 없고, 따라서 매칭 과정에서 제외시킨다. RII의 한 면이 중첩되었는가 여부는 그면이 거리 불연속(depth-discontinuity)이 존재하는 윤곽선(edge)의 뒷 부분에 존재하는가, 혹은 그 면이 곡면인가 아닌가 하는 기준에 의해 쉽게 결정될 수 있다.

지금까지 설명한 면 법선 영상 기반형 인식을 위한 새로운 매칭 방법은 매칭된 후보 모델을 찾아내면, 그 즉시 매칭 과정을 끝낸다. 어느 한 장면 물체의 매칭을 완료하기 위해서 몇개의 RII가 필요한가 하는 것은 모델 베이스의 크기와 인식하고자 하는 물체의 복잡도에 크게 좌우된다. 즉, 모델 베이스 내에 매우 비슷한 모델 물체들이 다수 존재하게 된다면, 그만큼 매칭에 걸리는 시간은 많아지게 된다. 그러나, 모델 물체들 사이의 유사도가 비교적 작다면, 소수의 RII들만을 가지고도 충분히 인식이 가능하다.

입력 영상내의 장면 물체에 대한 종류 확인(object identification)을 위한 매칭 과정이 모두 끝나면, 모델 베이스의 물체 노드에 존재하는 기준 좌표계와 장면 물체 및 모델 물체의 위치/방향 속성치들을 이용하여 장면 물체의 자세를 쉽게 계산해낼 수 있다.

#### IV. 실험 및 검토

본 장에서는 제안된 새로운 매칭 기법을 3차원 물체 인식에 적용하는 것이 가능한지, 그리고, 그 때 일어나는 제반 문제점들은 어떤 것들인지를 알아보기 위해 시험용 3차원 물체 인식 시스템을 구성하고, 간단한 블럭과 기계 부품 및 장난감들로 이루어진 거리 영상들을 대상으로 인식 실험을 행하였다. 실험은 256\*256\*256의 데이터를 사용하였고, SUN/4 Workstation상에서 C언어에 의해 수행하였다.

##### 1. 모델 베이스 구축과 실험용 거리 영상

본 논문의 인식 시스템에서는 orthogonal projection의 가정하에서 26개의 물체에 대한 모델 베이스를 구축하고, 그 모델 물체들을 포함하는 다수의 실제 거리 영상(real range image)과 인공 거리 영상(synthetic range image)을 대상으로 인식 실험을 행하였다. 모델로 사용된 물체들은 주로 평면과 원통면, 그리고 구면으로 이루어져 있다. 모델 베이스는 각 물체들의 표면을 3차원적으로 모델링하여,

그 3차원 정보를 이용하여 SNI를 얻고, 그에 포함된 각 특징들의 속성치들을 구하였다. 각 속성들중 크기와 관계된 속성들, 예를 들면 면의 면적, 선의 길이 등은 크기의 영향을 받지 않도록 하기 위하여, 그 물체에 대한 상대적 값으로 하였으며, 한 모델에 포함되는 SNI들은 기준면의 크기 순, 그리고 평면 우선 순으로 배열하여, 매칭시 효율을 높이도록 설계하였다. 표 1은 26개의 모델 물체에 대하여 그 물체의 면수와 SNI수를 보여준다. 모델을 구성하는 SNI들은 인식의 복잡도를 줄이기 위해 비교적 면적이 큰 면들을 기준면으로 하는 SNI들만을 모델 베이스에 저장하였으며, 이때 물체의 방향성과 대칭성을 함께 고려하였다. 따라서 표 1에서 볼 수 있듯이 모델 물체의 전체 면 수에 비해 모델의 SNI 수가 작게 되었다. 그러나 곡면이 포함된 모델 물체의 경우에는 한 곡면에 대해 여러 개의 SNI가 존재하므로<sup>(9,10)</sup>, SNI 수가 다소 많아지게 되었다. 결과적으로, 같은 방향을 가지는 면들을 많이 보유하고 있거나, 대칭성을 가지는 물체는 모델 구축에 있어 상대적으로 작은 규모를 가지게된다. 그림 8은 모델 구축에 사용된 26개의 모델 물체들이다.

표 1. 모델 물체들의 형상 정보  
Table 1. Shape of the model objects.

(a) 블럭 모델

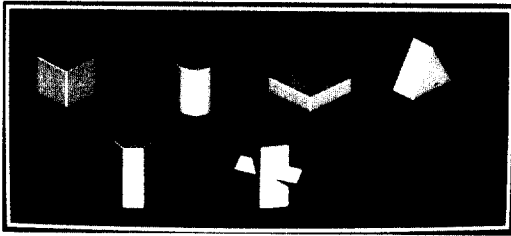
	면 수	SNI수	평면 수	원통면 수	구면 수
Block1	6	1	6	0	0
Block2	3	2	2	1	0
Block3	6	3	6	0	0
Block4	5	2	5	0	0
Block5	8	2	8	0	0
Block6	16	8	16	0	0

(b) 기계 부품 모델

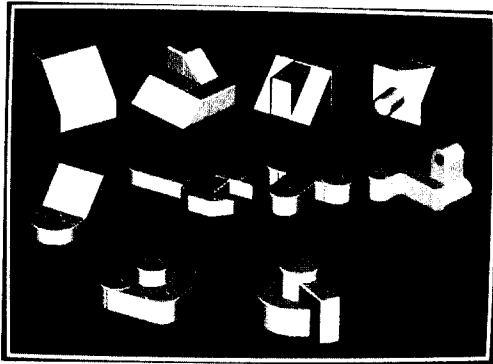
	면 수	SNI수	평면 수	원통면 수	구면 수
Workpiece1	7	7	7	0	0
Workpiece2	8	7	8	0	0
Workpiece3	10	8	10	0	0
Workpiece4	11	5	9	2	0
Workpiece5	8	9	7	1	0
Workpiece6	14	11	13	1	0
Workpiece7	10	12	7	3	0
Workpiece8	13	11	10	3	0
Workpiece9	12	10	6	6	0
WorkpieceA	9	10	7	2	0

(c) 장난감 모델

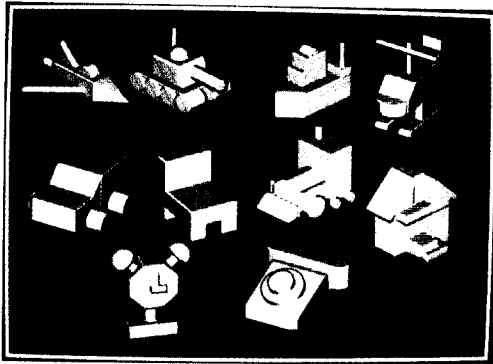
	면 수	SNI수	평면 수	원통면 수	구면 수
Car	16	10	12	4	0
Airplane	25	11	23	2	0
Tank	22	12	15	6	1
Ship	21	7	19	2	0
Telephone	20	11	14	6	0
Watch	29	14	24	3	2
Train	30	10	22	8	0
Helicopter	44	8	42	2	0
Chair	20	6	20	0	0
House	28	7	28	0	0



(a)



(b)



(c)

그림 8. 모델 베이스 구축에 사용된 모델 물체  
(a) 블럭 모델 (b) 기계부품 모델  
(c) 장난감 모델

Fig. 8. Model objects used in model base construction. (a) Block model, (b) Workpiece model, (c) Toy model.

한편, 모델 베이스 구성과 매칭시 사용된 속성들은

면 특징과 선 특징에 관해 여러가지가 있는데, 다음의 표 2에 요약하여 나타내었다.

표 2. 인식에 사용된 면과 선 특징의 속성들  
Table 2. The attributes of surface and line, used in recognition.

	속 성	속성의 내용	
면	면 방향	각 면의 방향 성분	
	상대 면적	각 면의 상대적 면적	
	면 유형	평면, 곡면, 구면	
	이심률	면의 길쭉한 정도	
	원형도	면의 둥그란 정도	
	인접면 수	그면과 인접한 면의 개수	
	침합선 길이	면과 면사이의 침합선 길이	
	선	침합 유형	면과 면의 침합유형(오목, 볼록)
		침합 각도	두면의 침합 각도
		윤곽선 개수	면의 윤곽을 이루는 선분의 개수
선 유형		선분의 유형(직선, 곡선)	
면	상대 길이	선분의 상대 길이	
선	침합 각도	두 선분 사이의 침합 각도	

실제 인식 실험에 사용된 입력 거리 영상들은 256\*256의 해상도(resolution)를 가지며, 실제 거리 영상들의 경우는 블럭들이, 그리고 인공 거리 영상의 경우는 기계 부품과 장난감들이 각각 2 - 4 개씩 포함되어 있다. 중첩의 영향을 알아보기 위해, 일부 입력 영상은 여러개의 물체가 중첩되어 있도록 하였으며, 좀더 열악한 환경에서의 실험 과정을 고려해 보기 위하여, 임의의 잡음이 첨가된 영상들을 만들어 실험에 사용하였다. 먼저 거리 해상도에 따라 나타나는 양자화 잡음(quantization noise)을 고려하였다. x와 y축 해상도는 변화시키지 않은 상황에서 거리 해상도 범위를 256에서 128과 64로 줄여가며 만들어진 영상을 사용하여 실험하였다. 다음으로 가우스 랜덤 잡음(Gaussian random noise)이 첨가된 영상에 대한 실험을 하였다. 보통 거리영상에서는 물체 내부에 지붕형 윤곽선이 많이 나타나므로, 물체 내부에 섞인 랜덤 잡음은 명암 영상에 비해 훨씬 치명적인 오류 요인이 된다. 따라서 랜덤 잡음이 섞인 거리 영상은 매우 조심스러운 전처리 과정이 필히 요구된다. 본 실험에서는 표준 편차  $\sigma$  가 0.5, 1.0, 2.0인 가우스 랜덤 잡음이 첨가된 영상들이 인식 실험에 사용되었다.

그림 9는 실험에 사용한 입력 거리 영상들의 일부 예이다. 여기서, (a)와 (b)는 인공 거리 영상이며, (c)는 실제 거리 영상(Real range image)이다.

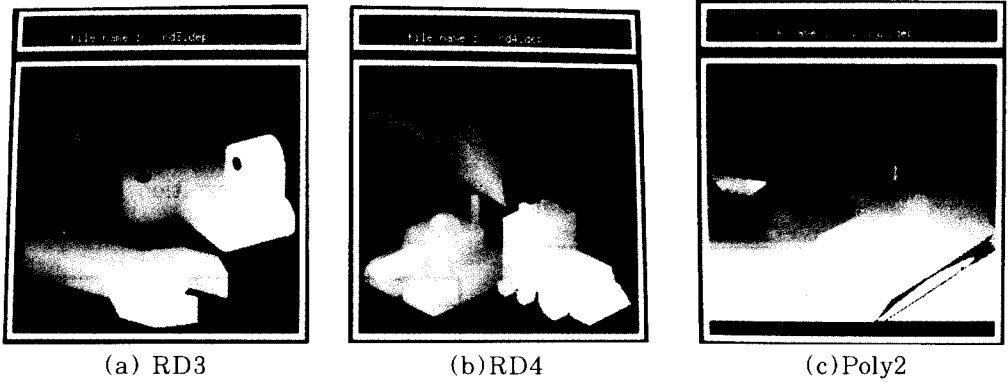


그림 9. 입력 거리 영상  
Fig. 9. Input range images.

### 3. 거리 영상 분할 실험

입력된 거리 영상은 전처리 과정을 거쳐 영상 분할이 이루어진다. 전처리 과정은 주로 잡음을 제거하기 위한 평활화 과정이며, 본 논문의 실험에서는 가우스 평활화를 사용하였다. 평활화된 영상으로부터 거리 불연속 점들을 찾아내고, 물체 내부의 방향 불연속 점들을 찾는다.<sup>[12]</sup> 구해진 불연속 점들은 거리 영상에서 평활하지 않은 부분이므로, 그 점들을 제외한 평활한 영역들에 대해 각 면소의 가우스 곡률과 평균 곡률 값들을 구한다.<sup>[13,14]</sup> 다음에는 완전한 영역 분할과 특징 추출을 하기 위하여, 추출된 불연속 점들의 정보를 이용하여 인접한 불연속 점들을 이어주는 간단한 작업을 통해, 물체 내의 각 면들의 끊어지지 않은 윤곽선을 구한다. 마지막으로 구해진 윤곽선들을 세선화하고 잡음성의 소영역들을 인접한 대영역에 병합시키는 과정을 행한다. 소영역의 병합 과정은 구

해진 각 영역의 면 유형과 방향 정보를 이용한다. 입력 영상들에 대한 영상 분할 결과는 대체적으로 만족스러웠으나, 일부 영상에서는 물체 내부의 윤곽선들이 정확하게 찾아지지 않고, 끊어져 두면이 하나의 면으로 잘못 분할되거나, 하나의 면이 여러 개의 작은 면들로 잘못 나누어진 경우를 볼 수 있었다. 이러한 경우는 주로, 평면과 곡면이 서로 평탄하게 만나고 있는 부분에서, 그리고 매우 작은 면들이 밀집되어 있어서 불연속 점들이 많이 존재하는 곳에서 발생한다. 본 논문에서는 하나의 면이 작은 면들로 잘못 나누어진 경우 인식 과정에서 여러 속성들을 고려하여 병합 처리하도록 하였다. 그러나, 두면이 하나로 잘못 병합된 경우는 사후 보정이 불가능하였으며, 인식 결과에서 오류를 낳았다. 다음의 그림 10에서 입력 거리 영상들에 영상 분할 결과를 보인다.

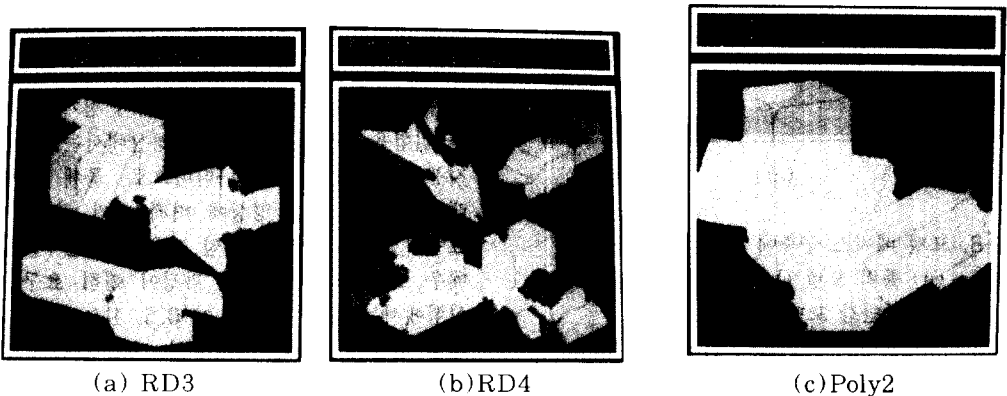


그림 10. 영상 분할 결과  
Fig. 10. The results of segmentation.



## 4. 특징 추출과 매칭 과정

입력영상에 대한 영상 분할이 이루어지면, 영상내에 존재하는 모든 면들에 대한 번호붙이기(labeling)를 행하고, 각 면들에 대해 1차 속성 즉, 면 유형, 면 방향, 면간 접합 각도가 구해진다. 1차 속성들은 모두 회전에 전혀 영향을 받지 않는 속성들이며, 그 중 면 방향은 매칭을 위한 RII들을 얻기 위해 사용된다. 각 면들에 대한 면 방향은 잡음을 많이 포함하는 윤곽선 부근의 면소들을 제외한 모든 면소들의 법선 벡터의 평균치로서 구해진다.

다음은, 각 면의 면 방향을 Z축에 일치하도록 장면 물체를 회전하여 각 면에 대한 회전 입력 영상(RII)을 얻는다. 실험에서는 문턱치를 두어 어느 정도 이상의 면적을 가지는 면들에 대해서만 RII를 얻도록 하였다. 면적이 좁은 면은 그 만큼 방향 오차를 많이 포함하므로 RII가 구해진다 하더라도 오히려 인식에

악 영향을 미치게되기 때문이다. 보통의 경우 면적이 큰 몇개의 RII만으로도 인식은 가능하므로 작은 면들의 무시는 타당하다.

RII들이 얻어지면 그 영상들에서 모델 베이스의 면 법선영상과 매칭하기 위한 2차 속성들을 추출한다. 2차 속성이란 물체의 회전에 의해 영향을 받는 속성들을 의미하며, 면의 면적, 이심률(Eccentricity), 원형도(Compactness), 인접면과의 관계등이 이에 속한다. RII들은 각 면들의 분리와 각 선분들의 분리 과정을 거쳐, 면과 선 특징들의 속성들에 의해 매칭이 가능한 형태로 변형된다. 각 면들은 이미 번호가 붙여져 있으므로, 간단히 분리되나, 각 선분의 분리는 윤곽선을 추적하여 꼭지점을 구한 뒤 선분들을 분리하여야 한다.

회전된 입력영상에서 추출된 특징들은 모델베이스의 모델들과 비교되며, 매칭의 기본 단위는 면 법선

---

```

SCENE # : RD3
SCENE NAME : WORKPIECE 3

NUMBER OF RIIS OF JOB.1 PREPARED FOR MATCHING : 6
MATCHING ORDER OF JOB.1 : RII.2 RII.3 RII.1 RII.4 RII.6 RII.5
RII.2 IS MATCHED TO MOB.1 SNI.1
RII.2 IS MATCHED TO MOB.3 SNI.1
RII.2 IS MATCHED TO MOB.5 SNI.1
RII.2 IS MATCHED TO MOB.14 SNI.6
RII.2 IS MATCHED TO MOB.15 SNI.2
MATCHING CANDIDATE IS MOB.1 MOB.3 MOB.5 MOB.14 MOB.15
RII.3 IS MATCHED TO MOB.1 SNI.2
RII.3 IS MATCHED TO MOB.3 SNI.5
RII.3 IS MATCHED TO MOB.15 SNI.2
MATCHING CANDIDATE IS MOB.1 MOB.3 MOB.15
RII.1 IS MATCHED TO MOB.3 SNI.7

MATCHING IS COMPLETED !!!
JOB.1 IS MATCHED TO MOB.3
NUMBER OF RIIS OF JOB.1 USED FOR MATCHING : 3

NUMBER OF RIIS OF JOB.2 PREPARED FOR MATCHING : 8
MATCHING ORDER OF JOB.2 : RII.5 RII.3 RII.4 RII.6 RII.2 RII.1 RII.7 RII.8
RII.5 IS MATCHED TO MOB.8 SNI.9
RII.5 IS MATCHED TO MOB.11 SNI.3
RII.5 IS MATCHED TO MOB.13 SNI.2
MATCHING CANDIDATE IS MOB.8 MOB.11 MOB.13
RII.3 IS MATCHED TO MOB.8 SNI.5
RII.3 IS MATCHED TO MOB.11 SNI.2
RII.3 IS MATCHED TO MOB.13 SNI.1
MATCHING CANDIDATE IS MOB.8 MOB.11 MOB.13
RII.4 IS MATCHED TO MOB.8 SNI.6

MATCHING IS COMPLETED !!!
JOB.2 IS MATCHED TO MOB.8
NUMBER OF RIIS OF JOB.2 USED FOR MATCHING : 3

NUMBER OF RIIS OF JOB.3 PREPARED FOR MATCHING : 6
MATCHING ORDER OF JOB.3 : RII.1 RII.3 RII.2 RII.5 RII.4 RII.6
RII.1 IS MATCHED TO MOB.6 SNI.1
RII.1 IS MATCHED TO MOB.7 SNI.1
MATCHING CANDIDATE IS MOB.6 MOB.7
RII.3 IS MATCHED TO MOB.6 SNI.4

MATCHING IS COMPLETED !!!
JOB.3 IS MATCHED TO MOB.6
NUMBER OF RIIS OF JOB.3 USED FOR MATCHING : 2

```

---

그림 11. 매칭 과정의 한가지 예

Fig. 11. An example of matching procedure.

영상이다. 그림 11은 하나의 입력 영상에 나타난 3개의 장면 물체들에 대한 모델 베이스 탐색에 의한 매칭 과정을 보여 주고 있다.

실험에 사용된 거리 영상들에 대하여, 인식의 결과는 대체적으로 만족스러운 것이었다. 거의 모든 장면 물체들이 정확한 인식 결과를 보여주었다. 그러나, 일부 장면 물체들의 경우, 매칭 결과에서 그에 매칭되는 모델 물체를 찾아내지 못하였다. 이는 주로 영상 분할 과정에서의 오류가 인식 결과에 까지 영향을 미친 경우나, 일부 물체가 너무 심하게 중첩되어 가려져 있는 경우로 사료된다. 한편, 잡음 영상의 인식 실험은 양자화 잡음에 대해서는 비교적 강건한 결과를 보였으나, 가우스 잡음이 섞인 영상에서는 좋은 결과를 얻지 못했다. 특히 가우스 잡음이 심하게( $\sigma = 2.0$  이상) 첨가된 영상들의 경우에는 영상 분할 및 매칭 결과에서 전혀 인식을 해내지 못하는 결과를 보았으며,  $\sigma = 0.5$ 나  $\sigma = 1.0$ 인 경우에도 잘못 매칭되는 경우를 자주 볼 수 있었다. 이러한 결과는 주로 영상 분할의 결과가 불량하기 때문에 나타나는 것으로, 본 논문에서 사용한 방법이 비교적 잡음에 민감하였던 탓으로 보인다. 따라서 매우 열악한 환경에서는 영상 분할 과정에 세심한 배려가 있어야 할 것이다.

다음의 표 3은 본 논문에서 보여주고 있는 3개의 입력 영상들에 대한 매칭 결과를 나타낸 것이다.

표 3. 입력 영상 #RD3, RD4, Poly2의 매칭 결과

Table 3. The matching result of the input image #RD3, RD4, Poly2.

영상 번호 장면 번호	RD3			RD4				Poly2	
	IOB.1	IOB.2	IOB.3	IOB.1	IOB.2	IOB.3	IOB.4	IOB.1	IOB.2
영상분할 후 면의 갯수	9	10	8	18	23	19	18	13	5
영역 면의 갯수	6	10	6	11	21	12	15	7	3
매칭을 위한 RII수	6	8	6	7	10	8	10	6	3
매칭에 사용된 RII수	3	3	2	3	4	3	2	2	3
매칭된 모델 물체	MOB #3	MOB #5	MOB #6	MOB #11	MOB #14	MOB #12	MOB #17	MOB #26	MOB #25

표 3에서 볼 수 있듯이, 실험 예에서 보여준 입력 영상 3개에 대해서는 정확한 인식 결과를 얻었다. 결과에서 입력 영상 RD3의 장면 물체 IOB.1의 예를 살펴보면 매칭에 사용된 RII수는 3개이다. 이때, 그림 11의 매칭 과정을 가지고 SNI와 RII간의 전체 비교횟수를 계산해보면 194(매칭 시작시 SNI수) + 42(첫번째 RII 비교 후 남은 후보 모델의 전체 SNI수) + 26(두번째 RII 비교 후 남은 후보 모델의 전체

SNI수) + 7(세번째 RII 비교 후 남은 후보 모델의 전체 SNI수) = 269회가 된다. 반면에 기존의 유사도 결정 함수를 사용하는 경우에는 전체 비교 횟수가 194(전체 SNI수) \* 6(전체 RII 수) = 5044 회가 필요하므로, 매칭의 시간면에서 제안된 매칭 기법이 효율적임을 단적으로 알 수 있다.

### V. 결론

본 논문은 기존의 면 법선 영상 기반형 인식 기법의 단점을 보완하기 위한 새로운 매칭 기법을 제안하는데 그 목적이 있다. 제안된 매칭 기법은 장면 물체로부터 추출된 특징들을 모두 만족하는 모델 물체를 모델 베이스에서 탐색해내는 방법으로, 장면 물체에서 얻어진 회전 입력 영상들을 모델 베이스의 모든 면 법선 영상들과 비교해 볼 필요가 없으므로, 그 만큼 매칭의 효율성을 더해 줄 수 있다. 또한 기존의 방법에서 문제가 되었던 각 특징의 속성들에 주어질 가중치의 경험적 결정 문제가 해결될 수 있으므로, 인식 결과의 신뢰성을 증가시키는 장점을 가진다. 그러나 제안된 매칭 기법은 장면 물체와 가장 유사한 모델 물체를 선택하는 기존의 방법과는 다르게, 장면 물체의 모든 특징들을 만족하는 모델 물체를 탐색하는 개념이므로, 영상 분할의 결과, 혹은 중첩의 영향에 따라 매칭되는 모델을 찾아내지 못하는 경우가 발생할 수 있다. 이는 기존의 유사도 결정함수에 의한 매칭에 비해 융통성이 적어진다는 점에서 단점으로 생각될 수 있으나, 그 만큼 인식의 신뢰도를 높인다는 장점이 될 수도 있다.

본 논문에서는 제안된 3차원 물체 인식의 새로운 접근 방식의 실현 가능성을 알아보기 위해 제안된 모델 구축과 매칭 방법을 적용하여 시험용 3차원 물체 인식 시스템을 구성하고, 26개의 모델 물체에 대한 인식 실험을 행하였다. 실험에서는 256\*256의 해상도를 가지며 거리 해상도가 256인 거리 영상들을 사용하였다. 여러가지 경우, 즉 중첩과 잡음의 영향을 고려한 여러 경우의 입력 영상들에 대한 실험 결과, 대체적으로 믿을 만한 인식 결과를 얻을 수 있었다. 그러나, 잡음이 심하게 포함된 영상들과 중첩이 심하게 일어난 장면 물체들의 경우, 그 장면 물체들과 매칭되는 모델을 찾아내지 못하는 경우가 자주 발생되었는데, 이는 주로 영상 분할 과정의 오류에 의한 것으로 생각된다.

본 논문에서 제안된 면 법선 영상 기반형 인식의 새로운 모델 매칭 기법은 그 구현상 다소 복잡해진다 는 단점이 있으나, 단지 방향 제시를 위해 소개되었

던 기존의 방법에 비해 좀 더 효율적이며, 신뢰성있는 인식 결과를 가져올 수 있다.

#### 감사의 말

본 연구 수행시, 끊임없는 검토와 토론에 시간을 할애해 주신 중앙대 제어계측공학과 부교수 장태규 박사에게 심심한 사의의 뜻을 표합니다.

#### 參考文獻

- [1] P. J. Besl and R. C. Jain, "Three-dimensional object recognition," *ACM Computing Surveys*, vol. 17, no. 1, pp. 75-145, Mar. 1985.
- [2] R. T. Chin and C. R. Dyer, "Model-based recognition in robot vision," *ACM Computing Surveys*, vol. 18, no. 1, pp. 68-108, Mar. 1986.
- [3] T. O. Binford, "Survey of model-based image analysis systems," *Int. Journal of Robotics Research*, vol. 1, no. 1, pp. 18-63, Spring 1982.
- [4] D. G. Lowe, "Three-dimensional object recognition from single two-dimensional images," *Artificial Intelligence*, vol. 31, pp. 355-395, 1987.
- [5] C. H. Chen and A. C. Kak, "A robot vision system for recognizing 3-D objects in low-order polynomial time," *IEEE Trans. System, Man, and Cybernetics*, vol. 19, no. 6, pp. 1535-1563, Nov./Dec. 1989.
- [6] G. Stockman, "Object recognition and localization via pose clustering", *Computer Vision, Graphics, and Image Processing*, vol. 40, pp. 361-387, 1987.
- [7] K. Ikeuchi, "Generating an interpretation tree from a CAD model for 3D-object recognition in bin-picking tasks," *Int. Journal of Computer Vision*, vol. 1, no. 2, pp. 145-165, 1987.
- [8] O. D. Faugeras and M. Hebert, "The representation, recognition, and locating of 3-D objects," *Int. Journal of Robotics Research*, vol. 5, no. 3, pp. 27-52, Fall 1986.
- [9] 박종훈, 장태규, 최중수, "면 법선 영상을 이용한 3차원 물체 인식", 대한전자공학회 논문지, 제 28권, B편, 제 9호, pp. 36-47, 9, 1991.
- [10] J.H.Park, T.G.Chang and J.S.Choi, "3-D object recognition and localization using surface normal images", *Proc. of 1st Korea-Japan Joint Conference on Computer Vision*, pp. 209-216, 1991.
- [11] M. Oshima and Y. Shirai, "Object recognition using three-dimensional information," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 5, no. 4, pp. 353-361, July 1983.
- [12] T. J. Fan, G. Medioni, and R. Nevatia, "Segmented descriptions of 3-D Surfaces," *IEEE Trans. Robotics and Automation*, vol. 3, no. 6, pp. 527-538, Dec. 1987.
- [13] P. J. Besl and R. C. Jain, "Invariant surface characteristics for 3D object recognition in range images," *Computer Vision, Graphics, and Image Processing*, vol. 33, no. 1, pp. 33-80, 1986.
- [14] 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, Sept. 1987.

#### 著者紹介

崔宗秀(正會員) 第28卷第9號 參照  
현재 중앙대학교 전자공학과 교수

朴鍾勳(正會員) 第28卷第9號 參照  
현재 한국전자통신연구소 미디어 응용연구실 선임연구원