

論文97-34C-10-9

효율적인 인덱싱 기법을 이용한 3차원 물체 인식: Part I-Bayesian 인덱싱

(Three-Dimensional Object Recognition Using Efficient Indexing: Part I- Bayesian Indexing)

李俊昊 *

(June Ho Yi)

요 약

본 연구에서는 효율적인 인덱싱 기법을 사용한 3차원 물체 인식 시스템을 제안하였다. 본 논문에서는 입력 영상에서 검출되는 물체의 특징(feature)들에 가장 잘 부합하는 후보 모델의 인덱싱을 위해 물체의 특징의 모델 물체에 대한 분별력(discriminatory power)이라는 척도를 정의하여 사용하였다. 이 분별력은 Bayesian 이론을 사용하여 확률적으로 도출되었고 계산된 분별력 값은 인덱싱 테이블에 물체의 특징과 모델 물체들의 대응관계에 대한 정보에 부가하여 저장하였다. 물체 인식시에 off-line으로 계산된 인덱싱 테이블이 입력되고 입력영상에서 검출된 물체의 특징들에 대해서 생성된 물체의 가설들을 검증할 때 검출된 물체의 특징들의 물체들에 대한 분별력의 순서대로 검증을 행함으로써 인식속도를 향상 하였다. 또한, 어떤 물체의 특징이 얼마나 잘 추출될 수 있는지의 정도를 물체의 특징 그 자체, 시점(viewpoint), 센서의 특성, 물체의 특징 검출 알고리즘의 함수로 정의하여 위에 언급한 물체의 특징의 모델 물체에 대한 분별력의 계산에 포함시켰다.

Abstract

A design for a system to perform rapid recognition of three dimensional objects is presented, focusing on efficient indexing. In order to retrieve the best matched models without exploring all possible object matches, we have employed a Bayesian framework. A decision-theoretic measure of the discriminatory power of a feature for a model object is defined in terms of posterior probability. Detectability of a feature defined as a function of the feature itself, viewpoint, sensor characteristics, and the feature detection algorithm(s) is also considered in the computation of discriminatory power. In order to speed up the indexing or selection of correct objects, we generate and verify the object hypotheses for features detected in a scene in the order of the discriminatory power of these features for model objects.

I. 서 론

모델을 이용한 물체인식(model-based object re-

cognition)에 있어서 가장 중요한 문제중의 하나는 모델 데이터베이스에 있는 모든 모델들을 다 조사하지 않고도 인식하고자 하는 입력 영상의 물체에 가장 잘 부합되는 몇개의 후보 모델들만을 골라내는 것이다. 컴퓨터 비전 분야에서는 이 문제를 인덱싱(indexing)이라 정의한다. 이 인덱싱 기술은 최근 연구가 활발히 진행되고 있는 멀티미디어의 디지털 라이브러리의 구현에도 필수적인 핵심기술로 인식되어 있다. 인덱싱을

* 正會員, 成均館大學校 電氣電子 및 컴퓨터工學部
(School of Electrical and Computer Engineering,
Sung Kyun Kwan University)

接受日字:1996年9月13日, 수정완료일:1997年9月20日

위해 인덱싱 테이블을 off-line으로 미리 만들어 실제 인식시에 불러들이게되는데 인덱싱 테이블(indexing table)에는 인식을 위해 사용되는 물체의 특징들이 나열되어 있고 각 특징을 갖는 물체 모델들이 해당 특징에 연결되어 저장되어 있다. 이 인덱싱 테이블은 입력 영상의 물체 인식을 위해 행하여야 할 모델 데이터베이스의 검색을 테이블 룩업(table look-up)으로 대체함으로써 물체 인식에 소요되는 시간을 줄여준다.

인덱싱 기법을 사용한 몇가지 대표적인 물체 인식 시스템을 표 1에 요약하였다. Wheeler 와 Ikeuchi^[1]는 각 물체를 컴퓨터 그래픽스의 ray-tracing 기법을 이용하여 렌더링(rendering)한 영상들로부터 물체 인식을 위해 사용하는 물체의 특징에 대한 정보를 구하고 또 그 물체의 특징이 실제 영상에 나타날 때 렌더링 영상에서의 특징과 다르게 나타나는 사실도 함께 고려하였다. 그들은 가장 신뢰할만한 물체에 대한 가설(object hypothesis)들을 고르기 위해 각 가설의 likelihood와 가설들간의 상호 관계를 MRF(Markov Random Field)를 이용하여 표현하였다. 하지만 off-line으로 인덱싱 테이블을 계산하지 않고 인식시에 영상에 나타난 평면형의 면들과 모델 물체의 평면형 면들의 모든 쌍을 고려하여 MRF를 계산함으로써 계산량이 매우 많은 단점이 있다. 또한, 그들의 시스템은 단지 평면만으로 이루어진 다면체 모델만을 사용하였고 입력 영상에 단지 한 면만이 나타나는 물체에는 적용할 수 없다. Beis와 Lowe^[2]도 확률적인 방식을 사용하였는데 인덱싱을 위해 사용한 물체의 특징은 4-straight-line-segment chain이다. 하나의 4-straight-line-segment chain은 연속되는 네 개의 직선이 이루는 세개의 각(angle)과 네 직선들의 길이 비율로 나타내었다. 그들은 인덱싱 함수라는것을 만들어 다양한 시점(viewpoint)에서 합성한 영상들로부터 인덱싱 함수를 훈련(train)시켰고 훈련된 인덱싱 함수로부터 각 가설이 올바른 가설(즉, 올바른 물체 인식 결과를 주는 가설)일 확률을 구하였다. 하지만 그들이 인덱싱을 위해 사용한 물체의 특징은 직선으로만 이루어졌기 때문에 곡선 에지를 갖는 물체에 적용할 수 없는 단점이 있다. Calfano와 Mohan^[3]은 두 단계의 테이블 룩업 방식을 사용했으며 290개의 2D 모델들로 이루어진 데이터베이스에 대하여 실험하였다. 그러나 입력 데이터로서 실제의 2D영상이 아니라 임의로 그린 2D 윤곽선(drawing)을 사용하였다. 위에 간략하게

기술한 시스템들을 포함한 물체인식 시스템들은 대개가 다른 가정들에 기저하여 구현되었을 뿐 아니라 물체 인식을 위해 선택된 물체의 특징도 다르며 적용 환경도 다르기 때문에 그 성능들을 직접 비교하기는 어렵다.

본 논문에서는 신속하고 정확한 모델 물체의 인덱싱을 위하여 Bayesian이론을 사용한 확률적 해법을 제시하였다. 모델 물체에 대한 물체의 특징의 분별력(discriminatory power)이라는 척도를 정의하여 인덱싱 테이블에 물체의 특징과 모델 물체들의 대응 관계에 대한 정보 뿐 아니라 계산된 척도를 함께 제공함으로써 물체 인식시 여러 물체들에 산재해 있는 평범한 특징이 아닌 보다 뚜렷한 특징들이 제공하는 물체 가설들을 먼저 고려하여 검증하였다. 또한, 어떤 특징이 얼마나 잘 추출될 수 있는지의 정도를 물체의 특징 그 자체, 시점, 센서의 특성, 물체의 특징 검출 알고리즘의 함수로 정의하여 위에 언급한 물체의 특징의 모델 물체에 대한 분별력의 계산에 포함시켰다. 입력 영상에서 검출된 물체의 특징들에 대해서 생성된 물체의 가설들을 검증할 때 검출된 물체의 특징들의 물체들에 대한 분별력의 순서대로 검증을 행함으로써 처음에 검증하는 몇개의 물체 가설 중에 올바른 인식 결과를 포함시킴으로써 인식 속도를 빠르게 하였다.

위에 기술한 시스템 설계 원칙에 따라 3차원 물체 인식 시스템을 구현하였으며 본 시스템은 곡면을 가진 물체들뿐 아니라 인식하고자 하는 물체가 한 면만이 보이는 경우도 적용할 수 있다. 인공의(man-made) 물체들의 대부분은 2차 곡면(quadric surface)들을 사용하여 모델링 할 수 있기 때문에 2차 곡면들로 20개의 CAD 모델을 만들어서 구축한 데이터베이스를 이용하여 실험하였다 (그림 1참조). 모델 물체는 속성 상관(attribute-relational) 그래프로 표현하였다. 그 한 예가 그림 2에 나타나 있다. 속성 상관 그래프에서는 각 노드(node)가 물체의 표면(surface patch)을 나타내고 arc는 면들 사이의 에지를 나타낸다. 현재의 시스템은 LSG(Local Surface Group)라는 물체의 특징을 사용하여 구현되었으나 다른 종류의 특징을 사용하여도 전체의 시스템 구조는 그대로 적용된다. 물체 가설을 검증하는 방법으로서, 후보가 되는 물체가 입력 영상에서의 자세(pose)로 보이게하는 시점을 가설로부터 계산하고, 구해진 시점에 대해 후보 물체를 컴퓨터 그 시스템은 크게 인식 시간 이전에 인덱싱

표 1. 기하학적 인덱싱 테크닉들의 요약

Table 1. A summary of survey on geometric indexing/ hashing techniques.

year	system	acquisition /recognition	input data	indexing key	complexity	comments
95	Wheeler and Ikeuchi [1]	3D/3D	range image	group of consistent planar surface matches between image and model	not reported	Markov random field to represent to likelihood of hypotheses
94	Califano and Mohan [2]	2D/2D	2D drawing	seven dimensional global invariants	recognition= $N_T \cdot (k + cQ)$ where N_T is the number of feature triplets in the image and Q is the number of models in the database.	290 models
93	Rigoutsos and Hummel [3]	2D/2D	intensity image	coordinates of scene points computed in the coordinate system formed by an ordered pair of scene points	building indexing table= $O(M^n)$, recognition= $O(s^3 M^n)$ where M is the number of models, n is the max number of feature points per model, and s is the number of interest points found in the scene.	point-set matching, modeling of positional uncertainty in the hash space
93	Beis and Lowe [4]	2D/3D	range image	three angles and ratio of the interior edge lengths from four straight-line segment chain	not reported.	learning likelihood of each hypothesis
92	Stein and Medioni [5]	2D/2D	intensity image	super segments with several different cardinalities for edges	$O(n) < \text{recognition} < O(n^2 m^3)$ where n is the number of features in the scene and is the number of models in the database	super segment to represent edges
92	Stein and Medioni [6]	3D/3D	range image	3D super segments with several different cardinalities for edges and splashes	$O(n) < \text{recognition} < O(n^2 m^3)$ where m is the number of features in the scene and is the number of models in the database	Surface patches and edges are represented by splashes and 3D curves, respectively.
92	Flynn and Jain [7]	3D/3D	range image	two invariant feature values computed from a triple of scene surface patches that are simultaneously visible	building indexing table = $O(mn^3)$, recognition = $O(m^3 s^3)$ where m is the largest number of surfaces in any of object models and is the number of scene surfaces.	interpretation tables (ten 2D tables)
88	Lamdan et al. [8]	2D/2D	intensity image	coordinates of scene points in the affine-transformed coordinate system formed by an ordered triplet of three scene points	not reported	point-set matching

래픽 기술을 이용하여 렌더링하였다. 렌더링 영상에서 보이는 다른 물체의 면들(즉, 가설에 포함되지 않았던 면들)에 대응하는 면들을 입력 영상에서 찾아냄으로써 가설을 검증하였다.

본 논문에서는 물체의 특징의 물체에 대한 분별력 척도 계산의 근간이 되는 Bayesian 이론을 기술하고 전체 시스템의 구조와 함께 LSG라는 물체의 특징을 사용하여 인덱싱 테이블을 생성하는 부분에 대해 설명한다. 이 논문의 후속 논문에서는 인식시에 행하여지

는 모듈들, 즉, 물체의 특징 검출, 가설의 생성과 검증 부분에 대해 기술하고 실제 레인지(range) 영상과 합성(synthetic) 레인지 영상을 사용하여 시스템의 성능을 평가하였다.

II. 시스템 개관

이 절에서는 전체 시스템을 개략적으로 살펴보기로 하겠다. 그림 3의 시스템의 블록도가 보여주듯이 전체

테이블 구축을 위해 모델에 대한 정보를 계산하는 부분(off-line part)과 인식 시간에 계산되는 부분(on-line part)으로 대분된다. off-line 계산 부분은 다시 세부분으로 구성되는데 첫 구성요소는 CAD 모델 데이터로부터 물체 인식에 편리한 표현으로 자동 변환시켜주는 연산을 한다. 두번째 구성 요소는 컴퓨터 그래픽 기술을 이용한 레인지 영상 시뮬레이터이다. 이 중 한 모듈은 물체 특징의 검출 가능성(detectability)를 구하기 위해 센싱과정을 시뮬레이션하고 다른 모듈은 여러 다양한 시점들에 대해 모델 물체의 레인지 영상을 렌더링한다. 렌더링 영상들로부터 인식 대상이 되는 모든 물체들(즉, 데이터베이스에 있는 CAD 모델들)에 대해 인식을 위해 필요한 정보가 구해진 후 세째 구성 부분이 각 특징의 모델 물체들에 대한 분별력을 계산하게 된다. 구해진 분별력은 입력 영상으로부터 어떤 특징이 검출되었을 때 모델 물체가 과연 입력 영상에 존재할 확률이 얼마인가를 나타내게 된다. off-line 계산 부분의 결과로서 인덱싱 테이블이 만들어 지는데 이 인덱싱 테이블에는 각 특징들과 해당 특징을 갖는 모델 물체들이 분별력 값과 함께 저장되어 있다. 이렇게 구해진 인덱싱 테이블은 인식시에 입력된다.

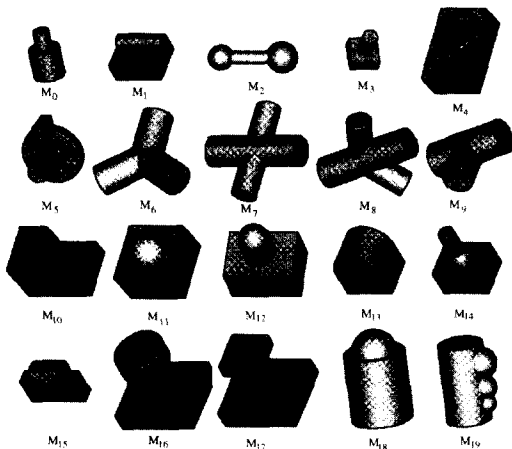


그림 1. 모델 데이터베이스
Fig. 1. Model database.

on-line 계산 부분은 물체의 특징 검출 모듈, 검출된 물체의 특징과 동일 혹은 비슷한 특징의 위치를 인덱싱 테이블에서 찾아 물체에 대한 가설(object hypothesis)들을 만들어 내는 모듈, 그리고 가설을 검증하는 모듈로 구성되어 있다. 물체의 특징 검출 모듈

의 입력은 하나의 레인지 영상이다. 특징 검출 모듈이 검출한 특징들과 동일 혹은 비슷한 entry들을 인덱싱 테이블에서 찾아 생성하게 되는 여러 개의 가설들은 각각의 가설에 함께 저장되어 있는 분별력이 큰 순서대로 나열된다. 여기에 유효한 가설이 만족해야 할 기하학적 제한 조건을 몇가지 적용하여 최종적으로 유효한 가설들의 리스트를 얻게되고 이 리스트에 배열된 순서대로 가설의 검증이 이루어지게 된다.

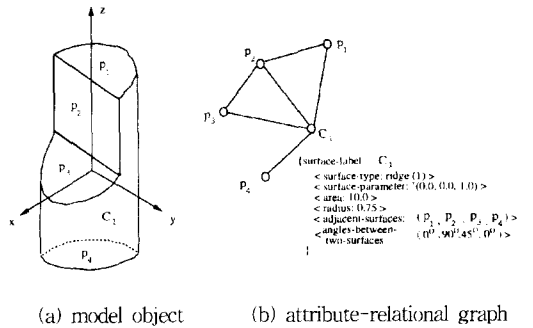


그림 2. 속성 상관 그래프 표현 방법 (a) 모델 물체 (b) 왼쪽 물체의 속성 그래프와 C_1 노드가 갖는 속성들의 집합
Fig. 2. The attribute-relational graph representation: (a) a model object and (b) its attributerelational graph and the attribute set for the node C_1 .

III. 물체 특징의 물체에 대한 분별력 (discriminatory power)

앞에서 언급한 바와 같이 본 연구에서는 입력 영상에서 검출되는 물체의 특징들에 가장 잘 부합되는 모델의 인덱싱을 위해 물체의 특징의 물체에 대한 분별력을 이용한다. 이 분별력을 Bayesian 이론의 a posteriori 확률로 도출하기 위하여 joint probability, $P(m_k, M_i, viewpoint_j)$, $k = 1, \dots, f$, $i = 1, \dots, N$, and $j = 1, \dots, v$, 로부터 출발한다. m_k, M_i 그리고 $viewpoint_j$ 는 각각 인덱싱을 위해 사용하는 물체의 특징, i 번째 모델 물체, 그리고 선택된 시점들의 집합중 j 번째 시점을 나타낸다. f, N, v 는 각각 물체의 특징, 모델 물체, 시점들의 갯수를 표시한다. 이 joint probability에는 모델 물체들에 대한 각 특징이 갖는 정보가 함축되어 있다. 사용하고자 하는 물체의 특징이 시점을 달리하여도 변하지 않는 특징(즉, viewpoint

independent feature)라도 단지 *viewpoint*,만을 생략 하면 앞으로 전개되는 수식을 공히 적용할 수 있다.

1. 사용되는 기호들의 정의

$P(m_k, M_i, viewpoint_j)$: joint probability of m_k (인덱싱을 위해 사용하는 물체의 특징), M_i (i 번째 모델 물체), $viewpoint_j$ (시점들의 집합중 j 번째 시점). $k = 1, \dots, f, i = 1, \dots, N, j = 1, \dots, v$.

$P(M_i)$: 주어진 물체가 M_i 일 확률을 나타내는 *a priori* 확률. $\sum_{i=1}^N P(M_i) = 1.0$.

$P(m_k / M_i)$: likelihood를 나타내는 조건부 확률로 $[P(m_k / M_i) > P(m_k / M_j)]$ 는 물체의 특징, m_k 가 모델 물체 M_i 의 특징일 확률이 모델 물체 M_j 의 특징일 확률보다 큼을 나타낸다.

$P(M_i / m_k)$: 입력 영상에서 물체의 특징, m_k 가 검출될 때 모델 물체 M_i 가 입력 영상에 존재할 확률을 나타내는 *a posteriori* 확률.

D_{m_k} : 물체의 특징, m_k 의 검출가능도 아래의 설명을 참조

Definition

$P(M_i / m_k)$ 를 물체의 특징, m_k 의 모델 물체 M_i 에 대한 분별력으로 정의한다.

서론에서 언급한 바와 같이 물체의 특징, m_k 의 검출가능도, D_{m_k} 를 $P(M_i / m_k)$ (즉, 물체의 특징, m_k 의 모델 물체 M_i 에 대한 분별력)의 계산에 포함시키는데 D_{m_k} 는 아래와 같이 물체의 특징 그 자체, 시점, 센서의 특성, 물체의 특징 검출 알고리즘의 함수로 정의된다.

$$0 \leq D_{m_k} = f(m_k, viewpoint, sensor, feature\ detection\ algorithm) \leq 1 \quad (1)$$

우선, D_{m_k} 는 물체의 특징 그 자체(즉, feature class)의 함수이다. 예를 들어, 물체의 면을 물체의 특징으로 이용하는 경우가 물체의 꼭지점(vertex)을 특징으로 이용하는 경우 보다 영상으로부터의 검출이 용이하다. D_{m_k} 는 또한 같은 종류의 특징이라도 어느 시점에서 그 특징을 검출하느냐에 따라 검출의 용이도가 달라진다. 평면(planar surface)을 여러 다른 시점에서 같은 알고리즘을 사용하여 검출한다고 했을 때 몹시 경사진

면으로 보이게 하는 시점에서보다 수평한 모습으로 보여지는 시점에서 그 면의 검출이 쉽다. 물론, 센서의 성능과 물체의 특징을 검출하기 위해 사용하는 알고리즘 또한 물체 특징의 검출가능도에 영향을 미치는 요소들이다.

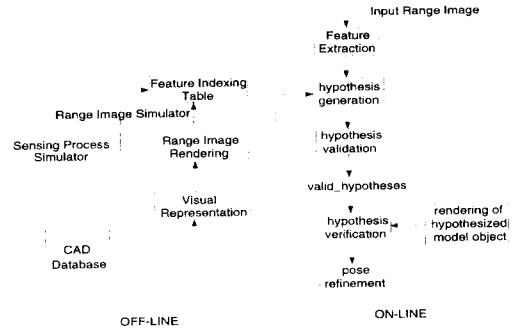


그림 3. 시스템 개관
Fig. 3. System Overview.

2. 분별력(discriminatory power)의 계산

이 절에서는 물체 특징의 모델 물체에 대한 분별력(discriminatory power)으로 정의한 $P(M_i / m_k)$ 를 계산하는 방법에 대해 기술한다. 다음에서는 위에서 정의된 것들의 계산치를 기호의 위에 hat()을 붙여 나타내기로 한다. $\hat{P}(M_i)$ 는 입력 영상에 물체 M_i 가 나타나는 빈도를 관찰하여 그 빈도를 모든 물체들의 관찰 빈도수로 나눔으로써 구할 수 있다. 예를 들어 공장 환경에서 컨베이어 벨트에 어떤 완성 제품을 이루는 여러가지 부품들이 흘러 가고 있을 때 경험적으로 어떤 부품이 다른 부품에 비해 얼마나 자주 나타나는지 작업자의 관찰에 의해 알 수 있다. 일단 모델 물체의 인덱싱을 위해 사용할 물체의 특징을 결정하면 $\hat{P}(m_k, viewpoint_j / M_i)$ 는 시점 $viewpoint_j$ 에서 M_i 를 볼 때 과연 물체의 특징, m_k 가 몇번이나 관찰되는지의 빈도수로 계산할 수 있다. 하지만, 물체의 특징, m_k 가 입력 영상에 존재한다고 하더라도 실제로 검출된다는 보장은 없으므로, 물체의 특징의 입력 영상에서의 검출 가능 정도를 분별력(discriminatory power) 계산에 포함시켰다. 즉, 물체의 특징의 검출 가능 정도를 1.0(perfect detectability)이라 가정하고 모델 관련 정보인 $\hat{P}(m_k, viewpoint_j / M_i)$ 를 계산하고

1 즉, $D_{m_k} \neq 1.0$

센싱 관련 정보인 \hat{D}_{m_i} 를 곱해줌으로써 센싱 관련 정보도 함께 포함되는 식 (2)를 얻는다.

$$\hat{P}(m_k, \text{viewpoint}_j/M_i) \cdot \hat{D}_{m_k} = \left(\frac{\# \text{ occurrences of the feature } m_k \text{ in } M_i \text{ for } \text{viewpoint}_j}{\# \text{ occurrences of all features } m_l, l=1, \dots, f_m \text{ in } M_i \text{ for all viewpoints}} \cdot \hat{D}_{m_k} \right) \quad (2)$$

그러므로 CAD데이터를 이용하여 모델의 정보를 계산할 때는 이러한 방식으로 물체 특징의 검출 가능 정도를 분별력(discrimatory power) 계산에 포함시킬 수 있다. 식 (2)로부터 likelihood, $\hat{P}(m_k/M_i)$ 는 식 (3)과 같이 계산되고 우리가 구하고자 하는 $\hat{P}(M_i/m_k)$ 는 Bayesian 법칙을 이용하여 식 (4)와 같이 구한다.

$$\hat{P}(m_k/M_i) = \sum_{j=1}^N \hat{P}(m_k, \text{viewpoint}_j/M_i) \cdot \hat{D}_{m_k} \quad (3)$$

$$\hat{P}(M_i/m_k) = \frac{\hat{P}(M_i) \sum_{j=1}^N \hat{P}(m_k, \text{viewpoint}_j/M_i) \cdot \hat{D}_{m_k}}{\sum_{j=1}^N \hat{P}(M_i) \sum_{k=1}^M \hat{P}(m_k, \text{viewpoint}_j/M_i) \cdot \hat{D}_{m_k}} \quad (4)$$

식 (3)과 (4)로부터 물체 M_i 에 물체의 특징, m_k 가 존재하지 않는다면 $\hat{P}(m_k/M_i) = 0.0$ 임을 알 수 있다. 또한 위에서 언급하였듯이 다양한 시점들에 대해 똑같이 보여지는(즉, viewpoint independent) 물체의 특징을 인덱싱을 위해 사용할 때도 식 (3)과 (4)에서 viewpoint_j 만을 생략한 같은 식을 적용할 수 있다.

어떤 특정한 물체의 특징과 시점이 주어졌을 때 D_{m_i} 를 계산하는 것은 다른 종류의 특징 검출 알고리즘들의 다른 센서 특징하에서의 성능(예를 들어 signal/noise ratio)을 결정하는 것과 같다. 만일 에지를 물체의 특징으로 사용할 경우, 에지 검출 알고리즘으로서 Sobel 연산자가 잡음(noise)이 있는 영상에서 Robert's cross 연산자보다 성능이 더 낫다는 것은 잘 알려져 있다. 실제로 어떤 signal/noise ratio값을 갖는 잡음이 있는 영상으로부터 위의 두 알고리즘을 사용하여 에지를 검출할 확률을 각각에 대해 구할 수 있다.

IV. 인덱싱 테이블의 구축

이 절에서는 3차원 물체 인식을 위해 사용할 물체의

특징으로서 LSG(Local Surface Group)라는 특징을 새로 제안하고 LSG를 이용하여 인덱싱 테이블을 만드는 과정을 설명하기로 한다.

1. LSG (Local Surface Group)

본 논문에서는 LSG를 물체 인식을 위한 물체의 특징으로 사용하여 물체 인식 시스템을 구현하였으나 분별력(discrimatory power) 즉, $P(M_i/m_k)$ 의 계산이 가능한 한, 모든 종류의 특징들이 사용될 수 있다. LSG는 단순한 특징이라기보다는 몇가지 속성들을 포함하는 구조를 갖는 특징이다. LSG는 하나의 면('seed 면'이라 명명함)과 주어진 시점에서 함께 보이며 이웃하고 있는 면들의 집합이다. 그림 4에는 면 C_1 을 seed 면으로 하고 과 이웃하면서 주어진 시점에서 함께 보이는 면들인 P_1 과 P_2 로 이루어진 LSG의 한 예를 보여주고있다. 일단 주어진 시점에서 함께 보이면서 이웃한 면들을 구하면 모델의 속성 상관 그래프의 노드에 저장되어 있는 속성들로부터 LSG의 각 슬롯의 값들을 채워 넣을 수 있다. LSG의 네번째 슬롯인 <simultaneously-visible-adjacent-surfaces : <list of surface>>에는 seed 면과 함께 보이며 이웃한 면들의 이름(혹은, 번호)과 surface type, 면적, 반지름(surface type이 실린더나 구형일 경우), 그리고 seed 면과 이루는 각도등이 함께 나타난다. seed 면과 이웃한 면이 이루는 각도는 두 면의 surface type이 평면형(planar)이거나 실린더형(cylindrical)일때만 적용되며 각 면의 방향의 차이로부터 구한다. 면의 방향은 평면형(planar)일 경우 면의 법선 벡터 방향으로, 실린더형(cylindrical)일 경우 실린더 축의 방향으로 정의한다. 다른 종류의 면들 사이의 각도는 정의되지 않으므로 NIL로 표시한다. 또한 LSG에서는 seed 면과 이웃한 면들 사이에는 C^0 에지(즉, jump 에지)는 허용하지 않는다. 왜냐하면 C^0 에지는 self-occlusion 아니면 물체의 boundary에 의해 생기게 되는데 위의 조건을 줌으로써 seed 면과 C^0 에지를 공유하는 이웃 면이 seed 면이 속하는 물체의 면인지 아닌지를 결정하는 매우 어려운 문제를 유보할 수 있기 때문이다. 어떤 시점에서 한 물체를 볼때 LSG의 수는 이론적으로 주어진 시점에서 보여지는 면들의 수임을 알 수 있다. LSG는 color나 texture와 같은 다른 속성들도 포함하기 위해 슬롯의 갯수를 확장할 수 있는 이점이 있다. 다시 한번 상기하면 LSG

는 시점에 의존하는 물체의 특징이며 그 안에 속성을 에 대한 슬롯들을 갖는 구조를 갖는 특징이다.

인덱싱을 위해 사용하는 물체의 특징이 복잡한 특징이라면 (예를 들어 물체 전체를 인덱싱을 위한 물체의 특징으로 사용하는 경우), 인덱싱은 효율적이지 않지만 단지 몇 개의 후보 물체만이 인덱싱의 결과로 나올것이다. 한편, 단순한 물체의 특징을 인덱싱을 위해 사용하면 (예를 들어 하나의 면을 인덱싱을 위한 물체의 특징으로 사용하는 경우), 인덱싱은 용이하지만 인덱싱의 결과로 나오는 많은 후보 물체들을 검증하여야 할 것이다. 본 시스템에의 경우 위에 소개한 LSG 전체를 인덱싱을 위한 특징으로 사용하는 경우 인덱싱이 용이하지 않고 계산양도 많이지므로 위의 LSG를 간편화한 것을 인덱싱을 위한 특징으로 사용하였다. 물체 인식 성능의 견지에서 최적의 복잡성을 갖는 인덱싱을 위한 물체의 특징을 결정하는 것은 앞으로 많은 연구를 필요로 하는 문제이다. 본 시스템에 사용된 인덱싱을 위한 특징의 한 예가 그림 4에 나타나 있다. 이 인덱싱을 위한 물체의 특징을 'Indexing_LSG'로 칭한다. Indexing 이웃하고 있는 면들의 각각을 구별하지 않고 그들의 surface type 정보와 seed 면과 이루는 각도들의 합을 사용하였다. 이 Indexing_LSG 여러 다른 instance(즉, 속성 슬롯의 값이 다른 것)들이 discriminatory-power 에서 기술된 m_k 들이다.

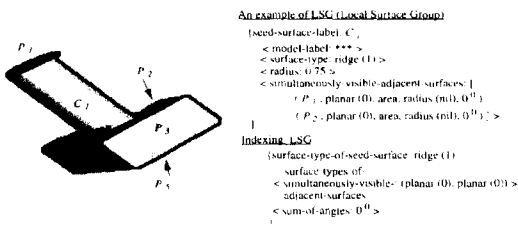


그림 4. LSG의 한 예
Fig. 4. An example of LSG.

2. $\hat{P}(M_i/m_k)$ 계산을 위한 LSG의 자동 추출

한 모델 물체로부터 모든 시점에 대해 LSG를 추출하는 방법은 다음과 같다. 모델 물체를 데이터베이스로부터 가져다가 각 시점에 대해 z-buffer 알고리즘^[9]을 이용하여 면의 라벨(즉, 번호)와 함께 렌더링하고 렌더링된 영상을 스캔(scan)하면서 LSG를 추출하게 된다. 모델 물체의 렌더링을 위해 구 표면상에 대체로 균일하게 분포된 162개의 시점들을 사용했는데 이 시

점들은 정20면체를 구에 투영한 다음 더 세분해서 얻어지는 geodesic 다면체의 dual을 이용하여 구하였다 [10] [11]. 단위 구(unit sphere)상의 한 시점이 그림 5에 표시되어 있다. LSG를 계산하는 개략적인 전체 과정은 다음과 같이 요약된다.

```

for i = 1, 2, ..., N
  for j = 1, 2, ..., v
    Render the range image of the object,  $M_i$ , for
      viewpoint  $j$ , along with surface labels.
    Scan the range image and collect LSGs.
  end for j
end for i
Compute  $\hat{P}(M_i/m_k)$ 's and return the indexing table.
    
```

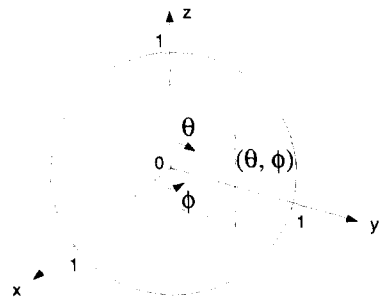


그림 5. 단위 구상의 하나의 시점, (θ, ψ)
Fig. 5. A viewpoint, (θ, ψ) , on a unit sphere.

모델 데이터베이스의 각 물체를 162개의 시점에 대해 렌더링을 하는 부분이 전체 시간을 지배하며 time complexity는 $O(n^2)$ 이다. 여기서 n^2 은 가로, 세로의 길이가 각 n 픽셀개를 갖는 z-buffer의 크기이다. 그림 1의 20개 물체들로 이루어진 데이터베이스에 대해 계산한 $\hat{P}(M_i/m_k)$ 의 분포를 그림 6에 가시화하였다. 그림 6은 다음과 같은 중요한 정보를 보여준다. 물체의 특징, m_k 의 모델 물체 M_i 에 대한 분별력이 1.0이면 (즉, $\hat{P}(M_i/m_k) = 1.0$ 이면), 물체의 특징 m_k 는 모델 물체 M_i 만이 갖고 있는 특징이다. 다시 말하면 입력 영상에서 물체의 특징, m_k 가 검출되면 모델 물체 M_i 가 입력 영상에 확실히 존재함을 알려준다. 한편, 입력 영상에서 물체의 특징, m_2 가 검출되면, $[\hat{P}(M_3/m_2) = 0.425] > [\hat{P}(M_0/m_2) = 0.401] > [\hat{P}(M_5/m_2) = 0.174]$ 의 결과로부터 검출된 물체의 특징, m_2 에 대해 입력 영상에 M_3, M_0 , 그리고 M_5 의 순서로 물체가 존재할 확률이 높음을 보여준다. 그

러므로 검출된 물체의 특징, m_2 에 대해 물체 M_3 의 가설을 먼저 고려한 다음 가설의 검증이 실패할 경우 M_0 , M_5 에 대한 가설의 순서로 검증하게 된다.

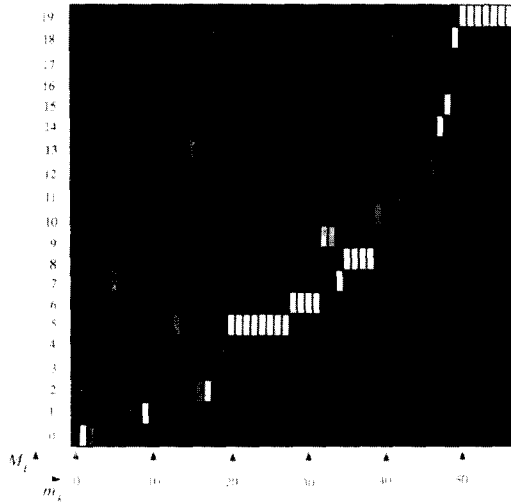


그림 6. 그림 1에서 보여준 20개 물체들로 이루어진 모델 데이터베이스에 대한 $P(M_i/m_k)$ 의 분포: $P(M_i/m_k) = 0.0$ (black), $P(M_i/m_k) = 1.0$ (white)

Fig. 6. Distribution of $P(M_i/m_k)$'s for 20 object model database shown in Fig. 1.

V. 결론

본 논문에서는 효율적인 3차원 물체 인식을 위한 인덱싱 기법으로서 Bayesian 이론을 이용한 확률적 해법을 전체 물체 인식 시스템의 구조와 함께 제안하였다. 또한 물체 인식을 위한 물체의 특징으로서 LSG(Local Surface Group)라는, 구조를 갖는 물체의 특징을 제안하고 LSG의 모델 물체에 대한 분별력을 계산하여 분별력이 실제 시스템에서 후보 물체의 인덱싱을 위해 어떻게 사용되는 지를 기술하였다.

참고 문헌

[1] M. Wheeler and K. Ikeuchi. Sensor modeling, probabilistic hypothesis generation, and robust localization for object recognition. *IEEE Trans. Patt. Anal. Machine Intell.*, 17(3):252-265, March 1995.

[2] Jeffrey Beis and David Lowe. Learning indexing functions for 3-D model-based object recognition. In *AAAI Workshop*, April 1993.

[3] Andrea Califano and Rakesh Mohan. Multidimensional indexing for recognizing visual shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, April 1994.

[4] I. Rigoutsos and R. Hummel. Distributed Bayesian object recognition. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, June 1993.

[5] F. Stein and G. Medioni. Structural indexing: Efficient 2-D object recognition. *IEEE Trans. Patt. Anal. Machine Intell.*, 14(12):1198-1204, December 1992.

[6] F. Stein and G. Medioni. Structural indexing: Efficient 3-D object recognition. *IEEE Trans. Patt. Anal. Machine Intell.*, 14(2):125-145, February 1992.

[7] P. Flynn and A. K. Jain. Object recognition using invariant feature indexing of interpretation tables. *CVGIP: Image Understanding*, March 1992.

[8] Y. Lamdan, J. Schwartz, and H. Wolfson. Object recognition by affine invariant matching. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, June 1988.

[9] J. Foley, A. van Dam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice*. Addison Wesley, 1992.

[10] C. H. Chen and A. C. Kak. A robot vision system for recognizing 3-D objects in low-order polynomial time. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1535-1563, November/December 1989.

[11] Philippe Brou. Using the Gaussian image to find the orientation of objects. *The International Journal of Robotics Research*, 3(4):89-125, 1984.

 저 자 소 개



李 俊 昊(正會員)

1985년 2월 서울대학교 전기공학과(공학사). 1987년 8월 The Pennsylvania State University, Department of Electrical Engineering(공학석사). 1994년 8월 Purdue University, School of Electrical Engineering(공학박사). 1989년 삼성종합기술원 연구원. 1994년 9월 ~ 1995년 10월 University of California, Riverside 연구원. 1995년 11월 ~ 1997년 2월 한국과학기술연구원 선임연구원. 1997년 3월 ~ 현재 성균관대학교 전기전자 및 컴퓨터 공학부 조교수. 주 관심분야는 컴퓨터 비전, vision-based man-machine 인터페이스 등임