

다면체 인식을 위한 탐색 공간 감소 기법

A Reduction Method of Search Space for Polyhedral Object

이상용

경남대학교 정보통신공학부

Sang-Yong Rhee

Division of Information and Communication Engineering

Kyungnam University

E-mail : syrhee@kyungnam.ac.kr

요 약

본 논문에서는 다면체의 인식을 위하여 사용되는 여러-방향-보기 방법 (multiple-view approach)에서 모델베이스의 탐색공간 크기를 줄이기 위한 방법을 제시한다. 이 방법에서 모델베이스는 물체를 둘러싸고 있는 보기 구 (view sphere)의 미리 정해진 보는 곳(viewpoint)에서 보여지는 2차원 투영체들로 구성된다. 물체를 인식하기 위하여 개략-상세 전략을 채택하되, 본 논문에서는 모델베이스의 개략 탐색을 위하여 ART-1을 이용하는 방법을 제시한다.

1. 서론

3차원 물체인식은 센서를 통하여 포착한 물체가 무엇인가를 알기 위하여 혹은 그렇게 인식된 물체의 위치와 자세를 알기 위한 과정이다. 물체의 인식을 위한 방법은 모델베이스에 있는 모델들과 비교해 일치되는 것을 찾아가는 방법(model-based method)[1-3]과 모델의 구축의 어려움과 모델의 자체의 한계 등에 의해서 생김새에 기초한 방법(appearance-based method)으로 나누어진다[4-6].

모델베이스를 이용한 물체인식문제는 어떻게 물체를 표현할 것인가는 표현방법과 얼마나 빨리 그리고 효과적으로 모델베이스의 물체와 비교하여 일치되는 것을 찾아내느냐 하는 인식방법, 이 두 가지 문제로 나눌 수 있다. 이 두 가지 문제는 서로 밀접한 관계를 갖는데 물체인식을 위한 방법은 물체 표현 방법에 의해서 직접적인 영향을 받으며, 반대로 물체 표현 방법은 물체인식 방법에 의해서 영향을 받게 된다.

생김새에 기초한 방법은 정규화된 크기와 밝기의 영상을 이용하여 고유공간(eigen space) 데이터베이스에 구축하고, 이 고유공간에서 거리가 최소인 물체를 찾는 방법을 사용한다[5,6].

모델베이스 물체인식의 근본적인 문제는 어떤 방법을 이용하여 모든 모델들을 탐색하지 않고 몇 개의 후보 모델로 탐색범위를 좁힐 것인가이다. 최근에는 인덱싱 혹은 해싱 기법도 제안되고 있다[7].

물체를 표현하기 방법은 크게 두 가지로 나눌 수 있다. 한가지 방법은 한 물체를 표현하기 위하여 하나의 표현방법을 채택하는 단일 3차원 모델(single 3D-model) 표현법[2,8]과 다른 방법은 하나의 물체를 여러 각도에서 바라본 형태 즉, 여러 개의 2차원 투영체로 저장하는 여러-방향-바라보기 방법 (multiple-view approach) 이 있다[3,4]. 본 연구에서는 물체의 표현을 위하여 여러-방향-바라보기 방법을 채택한다.

3차원 물체 표현을 위하여 여러-방향-바라보기 방법을 사용했을 때의 장점은 물체 인식 단계에서 모델베이스의 한 원소와 센서에 의해서 포착된 영상이 직접적으로 대응된다는 장점을 지니고 있다. 다시 말하면, 단일 3차원 모델을 사용할 경우에 부분 비교를 수행하여 일치되는 물체를 찾거나 혹은 3차원 변환을 통하여 비교일치되는 물체를 찾아야 하는 수고를 겪지 않아도 된다는 것이다. 물론 단점도 있는데 그것은 모델베이스

가 굉장히 크다는 것이다. 즉 각 방향에 따라 보여지는 다양한 형태의 2차원 투영체를 저장해야 하므로 큰 기억장소가 요구되며, 큰 모델베이스를 탐색하여 비교일치되는 원소를 찾기 위한 시간도 오래 걸린다는 문제점을 지니고 있다.

여러-방향-바라보기 방법을 사용하여 물체를 인식한다는 것은 센서에 의해서 포착된 영상과 모델베이스의 원소들 중에서 비교일치정도가 가장 큰 것을 찾아내는 것이라고 할 수 있는데, 이러한 경우에는 모델베이스를 탐색하는 방법이 인식과정의 효율성에 지대한 영향을 끼친다.

일반적인 접근방법으로는 먼저 개략 탐색을 거쳐 후보 물체들을 선별하고 난 후에, 상세 비교하는 방법을 사용한다. 이러한 방법이 처음부터 상세 비교를 하는 것보다 탐색 시간이 짧고, 탐색 공간이 적은 것은 자명하다고 할 수 있다. 따라서, 본 연구는 물체 인식을 위한 기본방향으로 개략-상세 방법 (course and fine search method)으로 설정하되, 본 논문은 개략 탐색에서 모델베이스의 탐색 공간을 줄이기 위한 방법에 대한 것이다.

본 논문은 다음과 같이 구성된다. 제 2장에서는 물체 표현 방법과 신경망에 학습을 시키기 위하여 이진 행렬로 변환하는 방법을 설명한 후, 3장에서는 실험결과를 제시하고, 결론을 맺는다.

II. 물체 표현 방법

2.1 여러-방향-바라보기 방법

사람들이 여러 방향에서 본 물체의 외형들을 기억하는 것과 같이, 여러-방향-바라보기 방법은 보기-구체(viewing sphere) 표면의 미리 결정된 방향들에서 물체를 바라본 2차원 투영체의 집합으로 물체의 모델베이스가 구성된다. 본 논문에서 대상 물체를 바라보는 방향을 다음과 같이 결정한다. 대상 물체를 감싸는 보기 구(viewing sphere)를 각 면이 삼각형으로 이루어진 300면체로 근사화 하고, 각 면의 중심에서 대상 물체에 대한 2차원 투영체들을 얻는다.

구를 300개의 면으로 근사화 하였기 때문에 300개의 투영체를 얻게 되지만, 300개의 투영체 전부를 그대로 저장한다는 것은 비합리적이므로, 투영체에서 특징 몇 가지만을 추출하여 저장하도록 한다. 우리가 고려하는 특징들은 다음과 같다.

- (1) 탐지된 요소들의 폐곡선의 갯수 : m
- (2) (n_1, n_2, \dots, n_m) : 여기에서 n_i 는 2차원 투영체에서 보이는 폐곡선을 구성하는 경계선의 갯수들로 편의상 n_1 에서 n_m 은 크기순

으로 나열한다.

- (3) 두개의 크기 순으로 나열된 리스트 $((L_{11}, L_{21}, \dots, L_{k1}), (L_{12}, L_{22}, \dots, L_{k2}))$:

여기에서 L_{i1} 은 임의의 폐곡선에 대하여 그 폐곡선이 주위에 있는 다른 폐곡선과 접하는 에지가 있을 때, 그 에지가 전체적으로 닿아 있는 것의 갯수를 나타내고, L_{i2} 는 임의의 폐곡선에 대하여 그 폐곡선이 주위에 있는 다른 폐곡선과 접하는 에지가 있을 때, 그 에지가 부분적으로 닿아있는 것의 갯수를 나타낸다. 그리고, 각 리스트는 편의상 크기순으로 정렬한다. 따라서 $L_{11} \geq L_{21} \geq \dots \geq L_{k1}$ 과 $L_{12} \geq L_{22} \geq \dots \geq L_{k2}$ 의 관계가 성립한다.

- (4) (p, q) : 여기에서 p 는 n_1 개의 경계선을 갖는 폐곡선의 내각이 둔각인 개수, q 는 둔각의 수가 2개 이상인 경우, 둔각 사이의 최소 꼭지점 수

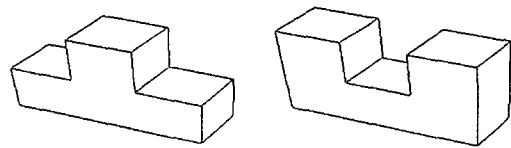


그림 1. 두 다면체

위에서 언급한 경계선들은 선분 혹은 곡선들로 구성되어 있다. 위의 특징들 중에서 (1)에서 (3)까지는 특징은 물체를 구성하는 각 면이 모두 볼록한 경우에 이웃한 면들 사이의 위상적인 변화를 탐지할 수 있는 최소 집합임이 증명된 바가 있다[10]. 그러나, 그림 1와 같이 물체를 구성하고 있는 면에 내각에 둔각이 존재하는 경우, (1)에서 (3)까지의 특징만으로는 동일하게 표현되어, 두 물체의 차이를 구별해낼 수가 없다.

물론 신경망을 이용하여 물체를 인식하기 위한 기본적인 접근은 개략 탐색을 하여 해당 카테고리만을 구별해내기 위한 것이기는 하지만 그림 1의 두 물체 정도는 구별할 수 있어야 한다. 본 연구에서는 (4)의 특징을 추가함으로써 이러한 문제를 해결하였다. 보기-구체의 한 면에서 본 투영체를 위의 특징들은 그림 2(b)와 (c)처럼 표현할 수 있다.

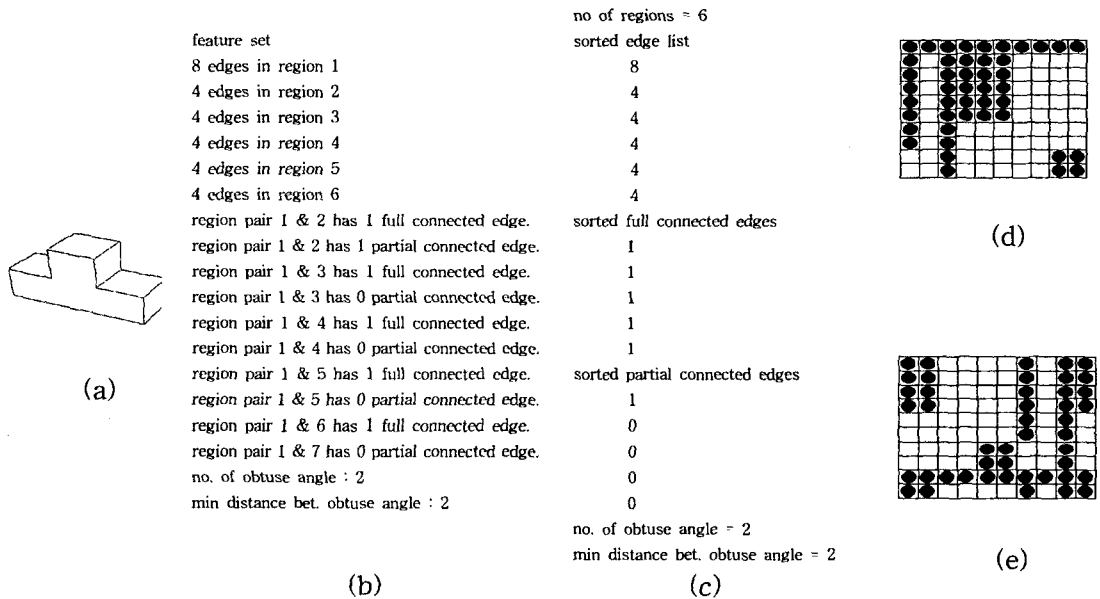


그림 2. 2차원 투영체의 특징 표현 예제 (a) 2차원 투영체. (b) 추출된 특징 집합. (c) 요약된 특징 집합 (d) 원 특징 행렬. (e) 이동된 특징 행렬

2.2 특징 벡터로의 변환

ART-1은 이진 패턴만을 사용하므로 위의 특징집합들을 이진 특징 벡터로 변환해야 한다. 각 특징 집합의 각 요소들을 이진 문자열로 코딩하는 것은 특징 벡터들의 각 값들은 정수형이기 때문에 어려운 일은 아니지만 이진 문자열의 단순한 나열보다는 각 특징들의 중요도가 고려되도록 표현해야 한다. 본 연구에서는 하나의 정수를 동일한 수의 이진 자리로 표현한다. 즉, 3이라면 3개의 1로 표현한다. 이진법을 이용하여 각 정수 값을 표현할 수도 있지만, 자리수도 늘어나기 때문에 바람직하다고 할 수 없다.

전체 특징 집합을 나타내는 이진 자리의 문자열들은 10×10 행렬로 만들기 위해 정렬한다. 본 논문에서는 특징 집합의 정수값을 10이 넘지 않는다고 가정하였는데, m이 10이상인 물체도 있을 수 있지만, 이러한 경우 10으로 간주하여 처리하거나, 필요하면 위의 행렬을 확장하여 사용한다. 앞 절에서 설명한 특징 표현 기법에 따라, 특징 집합의 요소들을 $m, (n_1, n_2, \dots, n_m)$,

$(L_{11}, L_{21}, \dots, L_{m1}), (L_{12}, L_{22}, \dots, L_{m2}), (p, q)$ 순서로 정렬한다. m은 가장 중요한 특징으로 m개의 이진 자리로, $n_i, 1 \leq i \leq m$ 와 p, q도 각각 이진 자리로 표현한다.

$(L_{11}, L_{21}, \dots, L_{m1}), (L_{12}, L_{22}, \dots, L_{m2})$ 는 상대적으로 덜 중요한 특징값이다. 따라서 개략 탐색 단계에서는 $S_1 = \sum_{i=1}^K L_{i1}, S_2 = \sum_{i=1}^K L_{i2}$ 과 같이 두 개의 특징값들의 합을 구한 후, 이진 자리로 표현하여 사용한다.

이제 위의 값을 행렬로 표현한다. 먼저 S_1 과 S_2 는 첫 번째와 두 번째 열에 넣는고, $n_i, 1 \leq i \leq m$ 들은 세 번째 열을 기준으로 행으로 차례대로 넣는다. 이렇게 하면 3번째 열이 m값이 된다.

p와 q는 각각 9번째와 10번째 열에 넣되 10번째 행에서 거꾸로 넣는다. 기본적인 행렬을 만들고 난 후 해밍 거리를 크게 하기 위하여 두 개의 중요한 특징인 m과 n_1 을 이용하여 열으로 m만큼, 행으로 n_1 만큼 이동시킨다. 이렇게 생성된 행렬을 ART-1의 입력으로 사용한다. 그림 2(d)와 3(e)가 기본 행렬과 이동된 행렬을 나타낸 것이다.

2.3 ART-1 신경망으로 표현

입력 이진 벡터는 10 곱하기 10의 100개의 원소를 갖으므로, F1에서 각 노드는 하나의 원소를 대응되도록 100개로 한다. F2에서 사용가능한 노드의 수는 학습순서와 경계 변수(vigilance

parameter)의 어떻게 정하느냐에 따라 변하므로 고정하지 않는다. 그러나 모델링을 하는 보기-공의 총면수인 300면을 넘지 않도록 한다. 따라서 편의상 F2는 중복되기는 하지만 300개의 노드를 갖도록 한다. 결과는 F2에 있는 노드들을 탐지하여 얻도록 한다[9].

III. 실험결과

우리는 제안된 방법의 실험을 위하여 보기 구체를 근사화하여 사용하였다. 먼저 구체의 모든 시점으로부터의 영상에서 추출된 특징집합으로 표현된 이진 벡터 집합을 사용하여 ART-1이 안정상태에 도달할 때까지 학습시켰다. 여기에서 안정상태란 새로운 이진 입력 벡터에 의해서 새로운 그룹이 형성이 되지 않는 것을 의미한다. 경계 변수 값은 0.8로 하여 학습을 시행하였다.

테스트 샘플이 300개의 시점(viewpoint)로부터 추출된 것 중에 하나이고, 면과 꼭지점이 정확하게 추출되었다면, 성공률은 100퍼센트였다. 그러나 300개의시점과 시점 사이의 영상을 이용한 경우 인식율은 90퍼센트로 줄었는데, 이 이유는 학습시킨 시점들은 이산적인 점들이므로 다면체의 모든 투영체가 표현된다고 볼 수 없기 때문이다. 따라서 어떤 특정한 각도에서 보여진 투영체에서 추출된 특징 집합은 해당되는 카테고리를 찾을 수 없는 경우도 발생한다. 또한 면과 꼭지점들이 정확하게 추출되지 못한 경우도 올바른 결과를 구할 수 없었는데, 이는 본 연구의 의도와는 거리가 있는 것이므로 고려하지 않는다.

IV. 결론

본 논문에서는 여러 방향 보기법의 문제점인, 모델베이스의 큰 탐색공간을 줄이기 위하여 ART-1을 이용하는 방법을 제시하였다. 3차원 물체 인식을 위한 기본 접근 방법으로 개략 탐색 후 상세 탐색을 가정한다면, 본 논문에서 제시된 방법은 개략 탐색 단계에 해당된다.

본 연구는 물체의 일반적인 특징인 외곽선과 그들의 관계만을 특징으로 추출하여 사용하였기 때문에 해당물체의 형태에 크게 종속되지 않는다. 또한 ART-1의 특성상 임의의 학습환경에서도 안정되게 카테고리가 이루어진다. 이것은 필요한 경우에 시점의 수를 늘릴 수 있기 때문에 여러-방향-보기 접근법에 아주 적절한 성질이라고 할 수 있다.

추후로 연구해야할 과제는 다음과 같다. 본 논문에서는 대상 물체를 다면체, 특히 각 면이

다각형인 물체로 한정하였으나, 각 면이 다각형이 아닌 물체로 확장이 필요하며, 추가적으로 자유곡면으로 이루어진 물체의 개략탐색을 할 수 있도록 연구를 진행할 계획이다.

참고문헌

- [1] P. Besl and R. Jain, "Three-Dimensional Object Recognition," ACM Computing Surveys, Vol. 17, No. 1, pp. 75-145, 1985.
- [2] R. T. Chin and C. R. Dyer, "Model-Based Recognition in Robot Vision," ACM Computing Surveys, Vol. 18, No. 1, pp. 67-108, 1986.
- [3] M. Korn and C. Dyer, "3-D Multiple-view Object Representations for Model-Based Object Recognition," Pattern Recognition, Vol. 20, No. 1, pp. 91-103, 1987.
- [4] Andera Selinger and Randal C. Nelson, "Appearance-Based Object Recognition Using Multiple Views," Proc. IEEE Conf. on CVPR, Vol. 1, pp. 905-911. 2001.
- [5] H. Borotschinig, L. Paletta, M. Prantl, and A. Pinz, "Appearance-based active object recognition," Image and Vision Computing, Vol.18, pp. 715-727, 2000.
- [6] H. Murase and K. Nayar, "Visual learning and recognition of 3-d objects from appearance," Int. J. of Computer Vision, Vol. 14, No. 1, pp. 5-24, 1995.
- [7] June Ho Yi and David M. Chelberg, "Model-Based 3D Object Recognition Using Bayesian Indexing," Computer Vision and Image Understanding, Vol. 69, No. 1, pp. 87-105, 1998.
- [8] R. Horaud, "New Methods for Matching 3-D Objects with Single Perspective Views," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-9, No. 3, pp. 401-412, 1987.
- [9] G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine," Computer Vision, Graphics, and Image Processing, 37, pp. 52-115. 1987.
- [10] F. Y. Liao, "Multiple-View Approach to 3-D Object Recognition," Ph. D. Dissertation, 1990.