

실내 환경에서 보이지 않는 목표 물체를 탐색하기 위한 활동기반 계층적 베이저안 네트워크

송윤석⁰, 조성배

연세대학교 컴퓨터과학과

corlary@sclab.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Hierarchical Bayesian Networks based on Activity for Localizing Hidden Target Objects in Indoor Environment

Youn-Suk Song⁰, Sung-Bae Cho

Dept. of Computer Science, Yonsei University

요 약

서비스 로봇 분야에서 물체를 인식하고 장면을 이해하는 것은 매우 중요하다. 전통적인 방법들은 기하학적 모델을 기반으로 물체를 인식하였으나 불확실하고 동적인 환경에서 이러한 방법은 한계를 갖는다. 이에 최근 지식 기반의 접근 방법을 통해 이러한 부분을 보완하는 연구가 이루어지고 있다. 본 논문에서는 효과적인 물체 탐색을 위해 베이저안 네트워크를 사용하여 대상 물체의 존재 여부를 추론하는 방법을 제안한다. 이를 위해 트리구조의 계층적 베이저안 네트워크를 사용하였고 물체들의 관계를 활동을 기준으로 모델링 하였다. 6가지 장소를 기반으로 한 실험 결과, 86.5%의 정확도를 보여주었다.

1. 서론

최근 서비스 로봇에 대한 연구가 많이 이루어지고 있다[1]. 특히 선진 국가에서는 고령화 사회에서 노인들의 생활을 보조하기 위한 연구 등, 실내 환경에서 기능하는 로봇에 대한 연구가 증가하고 있다[2]. 그 중 로봇이 서비스를 제공하기 위해 물체를 인식하고 상황을 판단하는 연구는 서비스 성능의 향상을 위해 중요하다. 기존의 전통적인 연구들은 영상 안에 있는 정보만을 통해서 물체의 존재 여부와 위치 등을 결정했다. 이러한 경우 불확실성과 다양한 변화를 포함하는 가정이나 사무실 같은 실내 환경에서 물체가 가려져 있는 경우 물체를 인식하는데 실패할 수 있다[3].

본 논문에서는 실내 서비스 환경에서 로봇이 보다 효과적으로 물체를 찾기 위해 발견된 물체를 증거로 대상 물체의 존재 가능성을 추론하는 방법을 제안한다. 이를 위해서 활동을 기준으로 물체들 간의 존재 관계를 만들어 트리 구조의 계층적 베이저안 네트워크 모델로 표현하였다. 계층적 베이저안 네트워크는 물체들의 존재 관계를 모델링 하는데 있어 표현력과 계산량 측면에서 다른 구조보다 효과적이다[4]. 또 효율적인 설계를 위해서 '공통-원인' 구조를 베이저안 네트워크의 빌딩 블록으로 사용하여 계층적으로 결합하였는데 이는 재사용 측면에서 유용하였다.

2. 배경

2.1 관련 연구

영상으로부터 물체를 탐지하기 위한 연구는 오래 전부터 있어왔다. 전통적인 연구들은 두 가지 가정에 기반하여 수행되었다. 첫 번째 가정은 대상이 되는 모든 물체는 상대적으로 적은 수의 형태 모델들(shape models)을 통해서 정의될 수 있어야 한다는 것이고 두 번째 가정은 지역적으로 측정될 수 있는 특징을 가지고 있어야 한다는 것이다[5]. 이런 가정 하에 전통적 시스템들은 주로 산업 현장에서 기하학적인 모델을 사용하여 물체의 위치나 방향을 알아내는데 사용되었다[5]. 하지만 다양한 물체가 존재하고 환경이 고정되어 있지 않은 실내 환경에서는 이와 같은 기하학적인 모델만으로는 물체를 인식하고 판단하는데 있어 한계가 있다. 이를 위해 지식 기반의 접근 방법이 제안되었는데

Marengoni 등은 항공 사진 인식 시스템인 Ascender I 에 계층적 베이저안 네트워크와 유틸리티 값을 적용하여 효과적으로 시각 합수를 선택하도록 함으로써 계산량을 줄이고 보다 정확한 결과를 얻고자 하였다[6]. Torralba 등은 Hidden Markov 모델을 영상에서 모아진 특징 벡터들로부터 장소를 인식하기 위해 사용하였고 그 결과 얻어진 정보를 통해 물체 인식의 우선 순위를 결정하였다[7]. 본 논문에서는 효과적인 물체 탐색을 위해서 찾고자 하는 물체의 존재 가능성을 추론하기 위한 베이저안 네트워크 모델을 제안한다.

2.2 베이저안 네트워크

베이저안 네트워크는 베이즈 규칙을 기반으로 변수들 사이의 의존 관계를 사용하여 각 변수들의 신뢰 값을 평가하는 DAG(Directed Acyclic Graph) 모델이다. 각 노드는 랜덤 변수를 나타내며 여러 가지 상태 값을 가질 수 있는데 이들의 합은 1 이다. 링크는 그들 사이의 관계를 표현한다. 이 때 원인이 되는 쪽이 부모 노드가 되고 결과가 자식 노드가 된다. 두 노드 사이의 링크는 결합 확률 분포를 만드는데 부모 노드는 초기 확률 값 $P(p)$ 을 가지며 자식 노드는 조건부 확률 값 $P(c|p)$ 을 갖는다. 조건부 독립 조건을 사용하면 노드들 사이의 결합 확률 분포 $P(x_1, x_2, x_3, \dots, x_n)$ 는 다음과 같이 표현될 수 있다.

$$P(x_1, x_2, \dots, x_n) = \prod_i P(x_i | Parents(X_i))$$

3. 제안하는 베이저안 네트워크

3.1 서비스 로봇

제안하는 방법은 15가지 장소와 29가지 물체가 존재하는 실내 환경을 고려하여 적용되었다. 관계된 장소와 물체들이 그림 1과 표 1에 있다.



장 의실

회의실

휴게실

그림 1. 관계된 장소들

표 1. 서비스 로봇 환경

| 분류 | 내용 |
|----|---|
| 장소 | 강의실, 회의실, 세미나실, 컴퓨터실, 교수실, 행정실, 수위실, 연구실, 복도, 계단, 홀, 엘리베이터, 화장실, 화장실전실, 휴게실연구실 |
| 물체 | 책상, 긴 의자, 팔 걸이 없는 의자, 강의대, 찬장, 책장, 쓰레기통, 세면대, 변기, 벽시계, 에어컨, 전화기, 데스크 탑, 노트북, 마우스, LCD, 키보드, 빔 프로젝터, 스크린, 오디오, 마이크, 벽 칠판, 이동칠판, 파티션, 커튼, 물통, 문, 창문 |

서비스 로봇은 물체를 찾기 위해 실내 환경을 탐색할 때, PCA를 통해 분석된 결과와 장소를 기반으로 설계된 장소-물체 베이지안 네트워크를 통해 전체적인 장면에서 물체의 존재 여부를 판단하게 된다[8]. 이 결과에 따라 현재 장소에서 물체 탐색 여부를 결정하는데 이러한 방법은 이미지 속에서 물체를 검사하는데 드는 계산 비용을 줄일 수 있게 한다. 활동-물체 베이지안 네트워크는 실제 물체 인식 과정 시 탐색된 결과로부터 대상 물체의 존재 가능성을 추론하고 현재 장소에서 물체 검사를 계속 할지 여부를 결정하기 위해서 사용된다. 활동에 따라 물체들의 존재 관계를 표현하는 것은 보통 물체들은 특정 활동과 관계되어 존재하고 한 장소에서 여러 가지 활동이 존재 할 수 있기 때문에 물체의 존재를 추론하는데 있어 보다 정교하고 효과적이다.

3.2 활동-물체 베이지안 네트워크의 구조

활동-물체 베이지안 네트워크는 트리 구조이고 노드는 크게 활동 노드와 클래스 노드, 프리미티브 노드로 나누어진다. 전체적인 구조는 부분-트리라고 볼 수 있는 공통-원인 구조를 빌딩 블록 단위로 하여 이를 계층적으로 결합하여 만들어진다. 공통-원인 구조는 3개의 노드 관계에서 2개의 노드가 다른 하나의 노드를 부모로 갖는 형태를 의미한다. 이러한 구조는 자식 노드에 할당된 파라미터를 통해서 자식들 간의 관계를 표현해주는데 유용하다. 이와 관계된 개념은 Reichenbach가 1956년에 발표한 '공통 원인의 원칙'에서 살펴볼 수 있다[9].

공통 원인의 원칙: "만약 두 개의 변수가 확률적인 의존 관계에 있다면, 하나의 변수가 직접 혹은 간접적으로 다른 변수에 원인이 되거나 두 노드는 공통된 조상을 갖는다"

공통-원인 구조를 빌딩 블록 단위로 결합해가며 만드는 것은 설계 관점에서 매우 유용하다. 이는 베이지안 네트워크의 설계와 확장, 재사용을 용이하게 한다. 베이지안 네트워크를 구성하는 각 노드들과 관계된 구체적인 내용이 다음에 있다.

- 활동 노드: 루트 노드, 물체들 사이의 관계를 만드는 상위 분류 기준, 결과 출력으로만 사용.
- 클래스 노드: 부분-트리의 루트 노드, 물체들간의 관계를 만드는 기준, 클래스 노드의 파라미터 값은 계층적 관계를 표현, 결과 출력으로만 사용.
- 프리미티브 노드: 물체를 표현, 파라미터 값은 같은 빌딩 블록 안에 있는 물체들과의 상관 관계를 표현. 물체가 발견되는 경우 증거 값으로 사용, 물체가 존재할 확률을 출력 값으로 가짐.

활동 베이지안 네트워크의 모든 노드들은 이진 속성을 가지고 있고 증거 값의 입력은 프리미티브 노드를 통해서만 들어올 수 있게 하는데 이는 추론과 관련되어 다음 절에서 논의 된다. 다음은 전체적인 구조와 관계된 몇 가지 설계 원칙이다.

- 활동 노드는 공용 클래스 노드와 private 클래스 노드만을 자식 노드로 갖는다.
- public 클래스 노드는 다른 활동-물체 베이지안 네트워크에서 많이 사용되는 물체들의 빌딩 블록 단위로 되어 있고 private 클래스 노드는 활동 노드와 밀접하게

관계된 물체들로 구성되어 있다.

- 클래스 노드의 파라미터는 계층적 관계를 의미하고 프리미티브 노드들의 파라미터 값은 같은 자녀 노드들 사이의 관계를 의미한다.

활동-물체 베이지안 네트워크의 기본적인 구조가 그림 2에 있다.

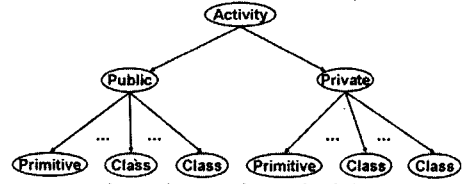


그림 2. 활동-물체 BN의 기본 구조

이러한 계층적인 구조는 naive 베이지안 구조보다 표현력이 좋고 다중-연결, 완전-연결 구조보다 훨씬 적은 계산량으로 추론이 가능하다. 각 구조들의 계산량에 대한 비교가 표2에 있다.

표 2. 여러 구조의 스펙

| | Naive | Fully-C. | Hierar.S- |
|-------|------------------------|----------------------|------------------------|
| 물체수 | n | n | n |
| 확률값수 | 2 ²⁽ⁿ⁻¹⁾ +2 | 2(2 ⁿ -1) | 2 ²⁽ⁿ⁺¹⁾ +2 |
| 계산복잡도 | O(n) | O(2 ⁿ) | O(n) |

3.3 확률 테이블과 d-분리

물체의 존재를 추론하기 위해서는 증거가 없을 때의 각 노드들의 확률 값이 (0.5, 0.5)로 유지되는 것이 중요하다. 이를 위해 자식 노드 C에서 P(C=yes|P=yes) 일 때의 확률 값을 a라고 할 때, 1-a를 P(C=yes|P=no) 일 때의 확률

$$\begin{aligned}
 P(C_{yes} | P) &= P(C_{yes} | P_{yes})P(P_{yes}) + P(C_{yes} | P_{no})P(P_{no}) \\
 &= \alpha \times 0.5 + (1-\alpha) \times 0.5 \\
 &= 0.5
 \end{aligned}$$

값으로 할당한다. 루트 노드의 초기 확률 값이 (0.5, 0.5)이면 다음의 식에 의해서 연쇄적으로 모든 노드들은 증거가 없을 때의 확률 값으로 (0.5, 0.5)를 갖게 된다.

확률 값 설정의 예가 그림 3의 왼쪽에 있다.

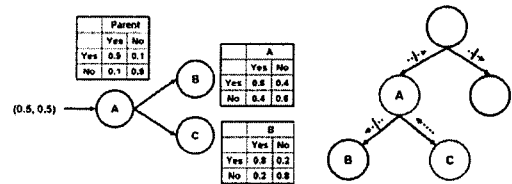


그림 3. 왼쪽: 확률 테이블의 예. 오른쪽: d-분리

추론은 네트워크의 "정보 흐름"에 의해 발생한다. 이 때 방향은 크게 두 가지로 부모에서 자식으로 가는 예측 형태와 자식에서 부모로 가는 진단 형태로 나누어진다. d-분리는 두 노드 집합 사이에서 특정 노드들에 의해 확률 값의 전파가 막히는 경우를 의미한다. 공통-원인 구조에서는 부모 노드에 증거 값이 설정되는 경우 자식 노드들 사이와 자식과 조상 노드들 사이에서 이러한 현상이 발생한다. 이는 물체들간의 존재 관계를 추론할 때 모든 노드로 확률 값이 전파되는 것을 막는다. 이를 위해 본 논문에서는 부모 노드는 클래스 노드만 가능하다고 하여 d-분리가 발생하지 않도록 하였다.

4. 실험 및 결과

4.1 실험 환경

실험은 6개의 장소(컴퓨터실, 연구실, 휴게실, 회의실, 세미나

실, 수위실)에서 프리젠테이션 활동-물체 베이지안 네트워크를 통해 수행되었다. 프리젠테이션 활동-물체 베이지안 네트워크의 전체적인 구조가 테이블 3과 그림 4에 있다.

표 3. 프리젠테이션 활동-물체 베이지안 네트워크의 스펙

| 노드종류 | 사용된 노드 이름 | 개수 |
|-------|---|----|
| 활동 | 프리젠테이션 | 1 |
| 클래스 | public{가구(기본1), 컴퓨터관련(액세서리), 오디오관련}, private{프리젠테이션 도구} | 8 |
| 프리미티브 | 의자, 책상, 마우스, 키보드, 모니터, 컴퓨터, 스피커, 오디오, 강의대, 마이크 x 2, 칠판, 스크린, 빔 프로젝터 | 14 |

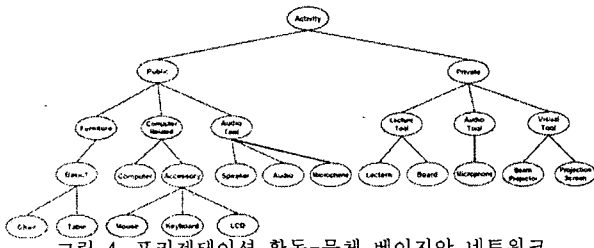


그림 4. 프리젠테이션 활동-물체 베이지안 네트워크

실험은 전문가에 의해서 설계된 활동-물체 베이지안 네트워크의 성능을 평가하기 위해서 수행되었으며 서비스 로봇은 여러 장소들을 다니면서 발견되는 물체들을 바탕으로 대상 물체의 존재를 추론한다고 하였다. 실험 결과는 각 장소에서 발견되는 물체에 따라 찾고자 하는 대상인 빔 프로젝터의 존재 확률 값 변화와 성공률을 기록하였다.

4.2 실험 결과

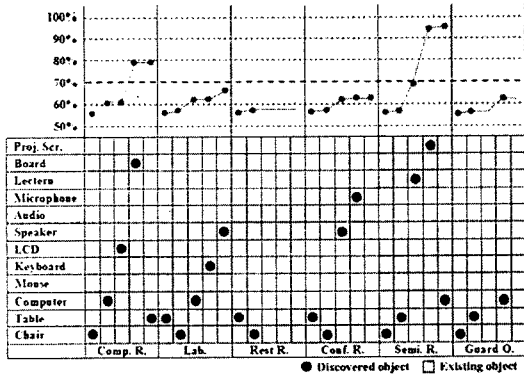


그림 5. 각 장소에서 빔 프로젝터가 존재할 확률 값의 변화

로봇이 5개의 물체를 발견하는 동안 기준 값 70%에서 각 장소에 대해 빔 프로젝터의 존재 확률 값이 그림 5에 있다. 예측 값은 한 경우를 제외하고는 베이지안 네트워크가 신뢰할 만하다는 것을 보여준다. 빔 프로젝터가 존재하지 않는 컴퓨터실 같은 경우 로봇은 물체가 존재할 것이라고 예상하였다. 이는 빔 프로젝터가 존재할 수 있는 환경과 유사한 경우 증거 값들이 쌓임에 따라서 '틀린-긍정' 오류가 실제로 발생할 수 있음을 보여준다. 실제 환경에서 경계 값과 물체를 찾는 횟수가 물체의 존재를 추론하는데 중요한 영향을 미칠 수 있다. 전체적인 성공률이 그림 6에 있다. 결과는 각 장소에서 10회의 실험을 하고 이 때의 성공률을 기록하고 같은 과정을 총 10회 반복한 결과로서 앞의 실험과 같은 사실을 보여준다. 전체적인 성공률은 86.5%로 제안하는 방법이 물체를 추론하는데 있어 신뢰할 만하다는 것을

보여주었다.

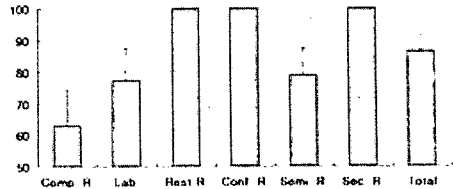


그림 6. 프리젠테이션 활동-물체 베이지안 네트워크

5 결론

본 논문에서는 불확실한 실내 환경에서 로봇이 찾고자 하는 물체의 존재 확률을 추론하여 여러 장소에서 효과적으로 물체를 찾기 위한 트리 구조의 베이지안 네트워크 모델을 제안하였다. 효과적인 설계를 위해 공통-원인 구조를 빌딩 블록 단위로 하여 계층적으로 결합하는 방법을 사용하였고 이를 위해 초기 확률 값과 d-분리와 같은 개념들을 설명하였다. 또한, 실험을 통해서 이러한 방법이 물체의 존재를 추론하는데 있어 신뢰할 만함을 보였다.

후후에는 부정 노드를 추가함으로써 발견된 물체에 따라서 확률 값을 떨어뜨리는 방법 등을 함께 적용해 보고 여러 종류의 활동-물체 베이지안 네트워크를 실제 로봇에서 함께 사용해 보고자 한다.

감사의 글

본 논문은 프린터 과제 지원에 도움을 받았다.

참고문헌

- [1] P. Dario, et al., "Robot assistants: Applications and evolution," *Robotics and Autonomous Systems*, vol. 18, pp. 225-234, 1996.
- [2] F. Mizoguchi, "Smart office robot collaboration based on multi-agent programming", *Artificial Intelligence*, vol. 114, pp. 57-94, 1999.
- [3] K. Murphy, et al., "Using the forest to see the trees: A graphical model relating features, objects, and scenes," *Proc. Neural Info. Proc. System*, vol. 16, pp. 1499-1506, 2003.
- [4] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.
- [5] T. M. Strat and M. A. Fischler, "Context-based vision: Recognizing objects using information from both 2-D and 3-D imagery," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 10, pp. 1050-1065, 1991.
- [6] M. Marengoni, et al., "Decision making and uncertainty management in a 3D reconstruction system," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 852-858, 2003.
- [7] A. Torralba, et al., "Context-based vision system for place and object recognition," *Proc. Intl. Conf. on Computer Vision*, pp. 273-280, 2003.
- [8] K. S. Hwang, et al., "Bayesian network design for high-level context reasoning in uncertain Indoor environment," *Soft Computing Lab. Tech. Report*, 2005.
- [9] H. Reichenbach, *The Direction of Time*, University of California Press, 1956.