

Hidden Information State 대화 관리자를 이용한 멀티모달 대화시스템

김 경 덕, 이 근 배

포항공과대학교 컴퓨터공학과 지능소프트웨어 연구실

Multimodal Dialog System Using Hidden Information State Dialog Manager

Kyungduk Kim, Gary Geunbae Lee

Intelligent Software Laboratory, Dept. of CSE, POSTECH

E-mail : {getta, gblee}@postech.ac.kr

Abstract

This paper describes a multimodal dialog system that uses Hidden Information State (HIS) method to manage the human-machine dialog. HIS dialog manager is a variation of classic partially observable Markov decision process (POMDP), which provides one of the stochastic dialog modeling frameworks. Because dialog modeling using conventional POMDP requires very large size of state space, it has been hard to apply POMDP to the real domain of dialog system. In HIS dialog manager, system groups the belief states to reduce the size of state space, so that HIS dialog manager can be used in real world domain of dialog system. We adapted this HIS method to Smart-home domain multimodal dialog system.

I. 서론

멀티모달 대화 시스템은 음성 대화를 기본으로 하여 펜 입력, 몸짓, 비전, 눈동자의 움직임 등을 입출력 방법으로 가지는 대화 시스템을 말한다. 시스템은 음성만을 입출력 수단으로 가지는 대화 시스템에 비해 자연스럽고 효과적인 의사소통을 가능하게 하기 때문에 최근 새로운 인터페이스의 하나로 각광을 받고 있으며 무인안내정보시스템(Kiosk)이나 개인용 정보 단말기에 실제로 적용되고 있다 [1].

대화시스템에서 사용자의 입력이 주어졌을 때, 어떤 시스템 출력을 낼 것인가를 결정하는 것이 대화 관리자의 역할이다. 유한 상태 네트워크(finite state network)를 이용한 방식이나 정보 상태 갱신(information state update)을 이용한 방식은 시스템 설

계자가 미리 정해놓은 대화흐름으로 대화가 이어질 경우 좋은 성능을 보일 수 있으나, 인식 에러에 약하며 다양한 대화 흐름을 수용하지 못한다는 단점이 있다. 마르코프 의사결정 (MDP, Markov Decision Processing)에 기반한 대화 관리자[2]의 경우, 앞서 언급한 대화 관리 방식과는 달리 통계방식을 이용하여 유연하면서도 최적화된 시스템 출력을 결정할 수 있게 해 준다. 하지만 MDP 기반 대화 관리자는 사용자 입력의 인식 에러에 취약하다는 단점을 가지기 때문에 POMDP (Partially Observable MDP)를 이용한 대화 관리자가 큰 관심을 모으고 있다.

하지만 POMDP를 이용한 대화 관리자도 단점을 가지고 있다. MDP와 마찬가지로 현재 대화가 어느 상태 인가를 나타내는 대화 상태 공간(state space)이 지나치게 커져서 최적해를 계산하기가 현실적으로 힘들다는 점이 그것이다. 이를 해결하기 위하여 대화 진행 중 같은 속성을 지니는 상태들을 묶어서 파티션(partition)로 처리하는 Hidden Information Approach (HIS)가 제시 되었다 [3]. 본 논문에서는 HIS에 대해서 간략히 소개하며 HIS가 적용된 스마트 홈 도메인의 멀티모달 대화 시스템에 대해서 설명한다.

II. Hidden Information State 대화 관리자

2.1 음성대화 시스템을 위한 POMDP

POMDP는 MDP와는 다르게 시스템이 사용자로부터 인식하는 정보가 부분적이라는 가정을 기본으로 한다. MDP에서는 현재의 상태가 무엇인지 분명하게 정해지

고, 그 상태에서 어떤 행동이 누적 보상을 가장 크게 하는지를 강화 학습 (reinforcement learning)을 통해 학습한다. 하지만 POMDP에서는 현재의 상태를 특정한 하나의 상태로 단정하지 않고, 각각의 상태에 대해서 그 상태일 확률을 가지고 있게 된다. 이를 신뢰 상태(belief state)라고 하며, 강화 학습을 통해서 주어진 신뢰 상태의 분포에 따라 최적화된 행동을 선택하게 된다.

POMDP는 $(S, A, T, R, O, Z, \lambda, b_0)$ 로 정의되며 각 기호는 다음과 같이 정의 된다.

- S : 상태(state)들의 집합
- A : 행동(action)들의 집합
- T : 상태 전의 확률. s에서 새로운 상태 s'으로의 전이확률 = $P(s'|s, a)$
- R : 보상(reward)의 기대값 = $r(s, a)$
- O : 관측값(observation)들의 집합
- Z : 관측 확률 = $P(o'|s', a)$
- λ : 할인 상수 (discount factor)
- b_0 : 최초 신뢰 상태 (initial belief state)

POMDP는 각 상태들에 관한 신뢰 분포 $b(s)$ 를 유지하며 각각의 상태에서 특정 행동을 취하였을 경우 보상값이 어떻게 되는지를 계산하여 그 기대값이 가장 큰 행동을 시스템 출력으로 정한다. 현재의 시간을 $t=i$ 라 할 때 $t=i+1$ 의 신뢰 상태는 다음과 같이 Bayes rule에 의해 구할 수 있다.

$$\begin{aligned} b'(s') &= \frac{P(s'|o', a, b)}{P(o'|a, b)} & (1) \\ &= \frac{P(o'|s', a, b)P(s'|a, b)}{P(o'|a, b)} \\ &= \frac{P(o'|s', a, b) \sum_{s \in S} P(s'|a, b, s)P(s|a, b)}{P(o'|a, b)} \\ &= \frac{P(o'|s', a, b) \sum_{s \in S} P(s'|a, b, s)b(s)}{P(o'|a, b)} \\ &= k \cdot P(o'|s', a) \sum_{s \in S} P(s'|a, s)b(s) \end{aligned}$$

그리고 보상값의 기댓값은 다음과 같이 구할 수 있다.

$$\begin{aligned} R &= \sum_{t=0}^{\infty} \lambda^t R(b_t, a_t) & (2) \\ &= \sum_{t=0}^{\infty} \lambda^t \sum_{s \in S} b_t(s) r(s, a_t) \end{aligned}$$

POMDP를 대화시스템에 적용하기 위해서는 대화의 상태를 POMDP에서의 상태로 어떻게 표현할 것이냐 하는 문제가 생기게 된다. 이를 위해 분할된 POMDP (factored POMDP)를 대화 시스템에 적용한

다 [4]. 즉 시스템의 상태 s_m 를 (3)과 같이 세 가지 구성요소로 나누어 생각한다. s_u 는 사용자 목적 상태 (user goal state)로서 사용자의 최종 목적이 무엇인지를 나타내는 상태이며, a_u 는 사용자의 입력을 말하며, s_d 는 대화 흐름 상태 (dialog history state)로서 현재까지의 대화 내력에 대한 상태를 나타낸다.

$$s_m = (s_u, a_u, s_d) \quad (3)$$

식 3을 식 1에 대입하여 식 4를 구할 수 있다[4].

$$b'(s'_m) = k \cdot P(o'|s'_m, a'_u) \sum_{s_m \in S} P(s'_m|a_m, s_m)b(s_m) \quad (4)$$

여기서 a_m 은 시스템이 사용자에게 출력한 행동을 의미한다.

2.2 Hidden Information State Approach

POMDP는 MDP와는 다르게 사용자 입력의 인식오류를 고려한다는 장점을 가지고 있으나, MDP에서처럼 상태수가 너무 많다는 문제점 또한 여전히 가지고 있다. HIS는 이를 해결하기 위해서 제시된 방법으로 사용자 목적 상태를 나타내는 s_u 의 상태 공간 S_u 에 대해서 동등한 (equivalent) 상태들을 서로 묶어서 하나의 파티션(partition)으로 생각하여 각 파티션에 대한 신뢰 상태를 구하는 방법이다 [3]. 이는 사용자에 의해 구체화 되지 않은 정보를 가지는 사용자 목적 상태들은 POMDP에서 동등하게 취급된다는 점에 착안한 방법으로 POMDP의 상태공간을 크게 줄일 수 있다.

대화가 시작되면 모든 상태 $s_u \in S_u$ 는 하나의 partition p_0 에 속하게 된다. 그리고 대화가 진행됨에 따라 파티션은 $P(p'|p)$ 의 확률로 작은 파티션으로 나뉘게 되며 각 파티션이 가지는 신뢰 상태 또한 다음과 같이 값이 나뉘게 된다 [3].

$$b(p') = P(p'|p)b(p) \quad (5)$$

$$b(p-p') = (1 - P(p'|p))b(p)$$

HIS에서는 식 (1)에서 파티션을 사용자 목적 상태 s_u 에 대입한 값을 신뢰 상태 분포로 사용한다 [3].

$$\begin{aligned} b'(p', a'_u, s'_d) &= k \cdot P(o'|a'_u)P(a'_u|p', a_m) & (6) \\ &\sum_{s_w, a_w, s_d} P(s'_d|p', a'_u, s_d, a_m)P(p'|p)b(p, s_d) \end{aligned}$$

시스템 설계자는 도메인에 알맞은 온톨로지 규칙을 제작하여 그에 맞도록 파티션들이 분할되도록 할 수 있다. 온톨로지 규칙들은 CFG와 유사한 형태로 되어 있으며 최초의 파티션 p_0 에 온톨로지 규칙들을 차례로 적용하게 되면 파티션은 트리 형태를 취하게 된다. 그리고 각 규칙은 사전 지식(prior knowledge)으로서 해당 규칙이 적용될 확률을 가지고 있다.

표 1 스마트 홈 도메인 온톨로지 규칙 예

task -> status(entity)	0.7
task -> control(entity)	0.3
entity -> device(type, name)	1.0
type -> light(loc, on_off)	0.3
type -> tv(loc, on_off, schedule)	0.3
...	

파티션은 사용자의 입력이 시스템에 들어오거나 시스템의 행동이 사용자에게 출력될 때마다 체크를 하여 필요에 의해 분할을 하게 된다.

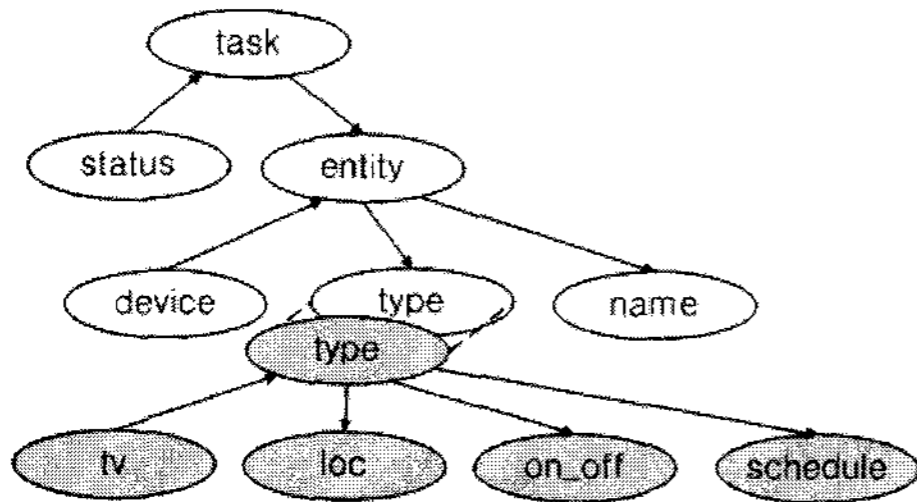


그림 1 파티션 분할의 예

그림 1은 파티션 분할의 예로 회색 부분의 트리가 새롭게 생성된 파티션이다. 새로운 파티션은 식 5에 의해 신뢰 상태의 값이 정해진다. 만일 흰색의 파티션의 상황에서 사용자가 TV에 대한 정보를 요구한다면 (e.g. $a_u = \text{request}(tv.on_off=on)$) 시스템은 온톨로지 규칙 $type \rightarrow tv(loc, on_off, schedule)$ 을 적용하여 회색 부분의 파티션을 생성하게 된다.

시스템은 여러 파티션들을 관리하며 각각의 사용자 입력에 대해서 적절하다고 판단되는 시스템 행동 후보를 생성하게 된다. 이때에 계산 복잡도를 줄이기 위해서 SPBVI (Summary Point Based Value Iteration) [5]이나 CSPBVI (Composite SPBVI) [6] 방법이 사용된다. 이 방법은 대화 상태와 시스템 행동을 간략히 나타낼 수 있게 요약 (summarizing) 하여 공간이 크게 줄어든 상태와 행동에 대해서 강화 학습을 적용하는 방법이다. 현재 본 논문에서는 알고리즘이 복잡한 SPBVI나 CSPBVI를 적용하지 않고, 휴리스틱으로 시스템 행동 후보를 생성하게 하여 미리 정의된 보상함수를 적용하여 가장 큰 보상값을 갖는 시스템 행동을 최종값으로 사용자에게 출력하였다.

III. 스마트 홈 도메인의 멀티모달 대화 시스템

본 논문에서는 펜 입력과 음성 입력이 동시에 가능한 스마트 홈 도메인의 멀티모달 대화 시스템에 대해 HIS를 적용하였다. HIS를 사용하였기 때문에 사용자

입력 오류에 강인하며, 멀티모달 입력을 지원하기 때문에 사용자에게 효과적이고 자연스러운 인터페이스를 제공한다는 것을 특징으로 삼을 수 있다.

스마트 홈 도메인의 멀티모달 대화 시스템의 전체 구조도는 다음과 같다.

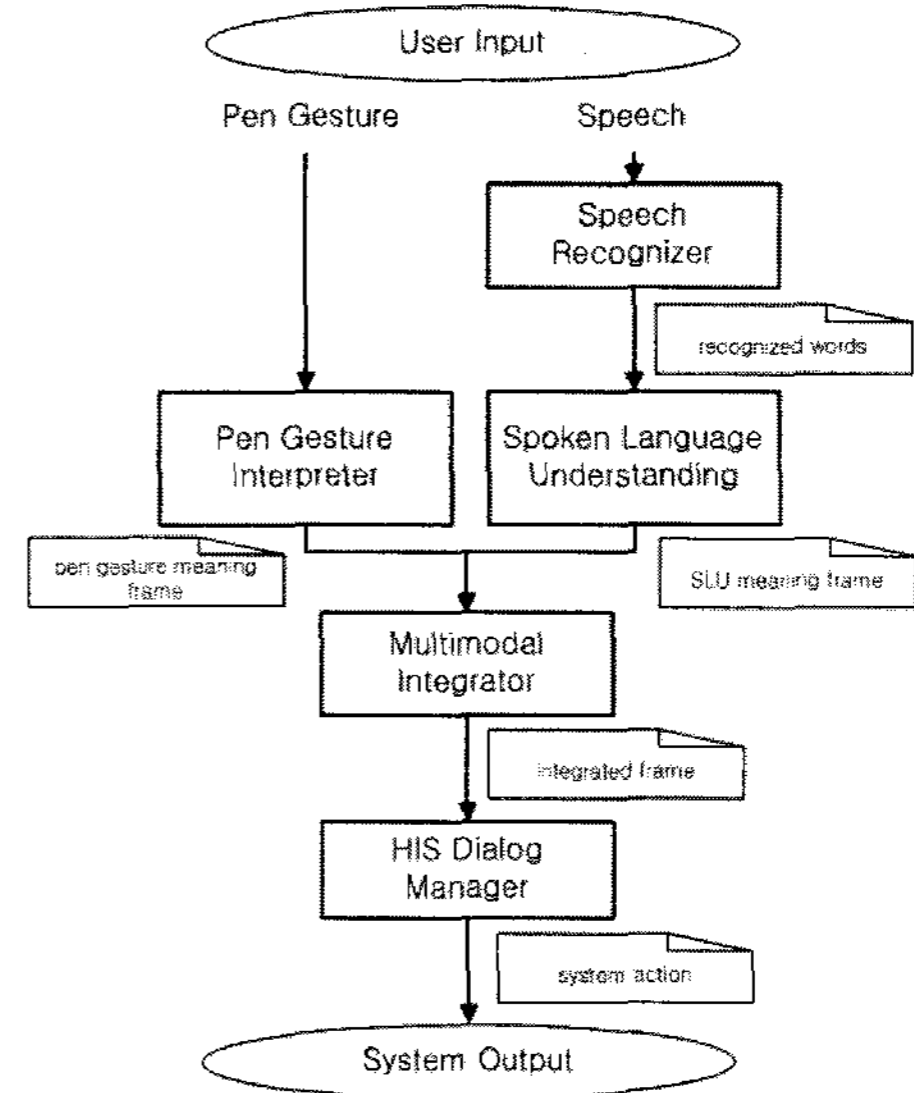


그림 2 스마트 홈 도메인 멀티모달 대화 시스템의 전체 구조도

음성인식기로는 HTK를 사용하였으며[7] 음성언어의 이해를 위해 개념 잡기식의 언어이해(concept spotting SLU) 모듈[8]을 사용하였다. 각 유니모달 입력인 펜 입력과 음성입력 이해 결과를 하나로 통합하기 위해서 Greedy 알고리즘 방식의 멀티모달 통합기 [9]를 적용하였다. HIS 대화 관리자는 통합된 의미 프레임을 받아들이고 그것을 사용자의 입력(a_u)으로 한 파티션의 신뢰 상태를 관리하며 최적화된 정책에 의해 가장 적합한 시스템 행동을 결과로 사용자들에게 출력해 준다.

사용자는 화면을 통해서 현재 집안의 상태나 가전기기의 상태를 직접 확인 할 수 있으며, 시스템에게 음성이나 펜 입력을 통하여 상태를 물어 볼 수 있다. 또한 시스템에게 명령하여 조명을 조절한다거나 전자기기의 동작을 제어할 수 있다. 실제 스마트 홈에 멀티모달 대화 시스템을 적용하기에는 비용과 시간이 많이 들기 때문에 시스템의 유효성을 검토하기 위해서 GUI로 된 시뮬레이터를 구현하였다. 시뮬레이터에서 사용자는 음성과 펜 입력을 통하여 시스템과 의사소통을 할 수 있으며 GUI를 통해 현재 집 내부의 상황과 대화의 흐름 및 상태를 확인할 수 있게 하였다.

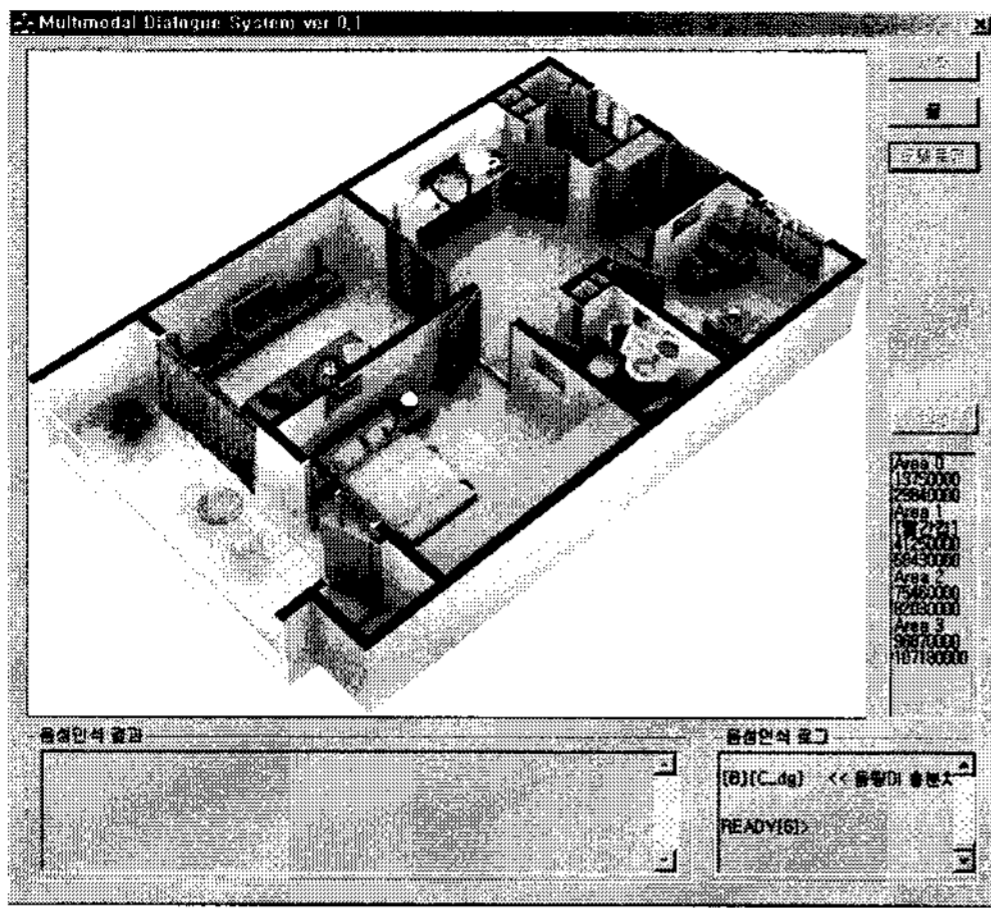


그림 3 GUI 시뮬레이터 스크린샷

IV. 결론 및 향후 연구 계획

HIS는 POMDP를 이용한 대화관리자의 문제점으로 지적되었던 상태공간을 혁신적으로 줄일 수 있는 수단으로 제시된 기술이다. 본 논문에서는 HIS 대화관리자를 스마트 홈 도메인의 멀티모달 대화시스템에 적용하여 보았다. 실제 스마트 홈에 적용하기에는 비용과 시간이 많이 들기 때문에 GUI를 통한 시뮬레이터를 구현하였다.

HIS를 사용하게 되면 사용자 입력의 인식 결과를 완전히 신뢰하지 않는다는 가정을 하기 때문에 사용자 입력 오류에 대해서 강인한 특성을 가지게 된다. HIS는 대화 상태 공간을 줄일 수 있는 장점을 가지고 있으나 시스템 행동의 종류를 줄일 수는 없기 때문에 그대로 강화 학습을 적용하여 최적화된 정책을 구하기에는 무리가 있다. 이를 해결하기 위한 방법인 SPBVI, CSPBVI는 대화 상태와 시스템 행동의 공간을 요약할 때에 적절한 휴리스틱을 적용해야 하며, 정책에 의해 선택된 요약 시스템 행동을 실제 시스템 행동으로 매핑하기 위해서는 휴리스틱 적용이 필요하다. 이렇게 휴리스틱 적용과정에서 근사 에러가 발생할 수 있기 때문에 현재 이 방법을 더욱 개량하여 사용할 수 있는 방법에 대해서 연구 중에 있다.

향후 연구는 HIS의 단점을 보완하기 위한 방향으로 진행될 것이다. MDP/POMDP기반의 대화관리자는 대화의 흐름이 보상함수에 너무 민감하여 적절한 보상함수를 찾아내기 힘들 수 있다는 문제점이 있다. 그리고 인간-기계간의 대화 자체를 최적화의 문제로 볼 수 없다는 견해도 있다. 뿐만 아니라 MDP/POMDP 기반의 대화관리자에서는 최적화된 정책을 찾기 위한 강화학습의 복잡도도 크다. 이런 문제를 해결 할 수 있는 대화 모델링 기술 개발을 위해 MDP/POMDP기반 대화관리자에 비해 계산 복잡도가 작으며 현재 사용자 입

력에 대해 가장 적절한 시스템 출력을 대화예제에서 검색하는 예제 기반 대화 모델링[10] 기술을 HIS 대화 모델링에 적용하여 두 방법의 장점을 고루 가지는 대화 모델링 기술 개발에 초점을 맞출 것이다.

감사의 글

본 논문은 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었다. (IITA-2005-C1090-0501-0018)

참고문헌

- [1] M. Johnston, and S. Bangalore. "MATCHKiosk: A Multimodal Interactive City Guide". *ACL-2004*, 2004
- [2] 최준기, 은지현, 장두성, 김현정, 구명완, "마르코프 의사결정 과정에 기반한 대화 관리자 설계", *대한 음성학회 가을 학술대회 발표 논문집*, 2006
- [3] S. Young, J. Schatzmann, K. Weilhammer and H. Ye. "The Hidden Information State Approach to Dialog Management." *ICASSP 2007*, 2007.
- [4] J. Williams and S. Young. "Partially Observable Markov Decision Processes for Spoken Dialog Systems." *Computer Speech and Language* 21(2): 231-422.
- [5] J. Williams and S. Young. "Scaling up POMDPs for Dialogue Management: the Summary POMDP Method." *ASRU2005*, 2005
- [6] J. Williams and S. Young. (2007). "Scaling POMDPs for Spoken Dialog Management." *IEEE Audio, Speech and Language Processing*, accepted for publication, 2007
- [7] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, G. Moore, J. O. D. Ollason, D. Povey, V. Valtchev, P. Woodland, *The HTK book: version 3.3.*, UK. <http://htk.eng.cam.ac.uk>, 2005
- [8] J. Eun, M. Jeong and G. G. Lee, A multiple classifier-based concept-spotting approach for robust spoken language understanding. *Interspeech 2005-Eurospeech*, 2005.
- [9] J. Chai, Z. Prasov, J. Blaim, and R. Jin, Linguistic Theories in Efficient Multimodal Reference Resolution: an Empirical Investigation. *IUI-05*, 2005.
- [10] 이청재, 정상근, 이근배, "대화 예제를 이용한 상황 기반 대화 관리 시스템", *대한음성학회지: 말소리*, *MALSORI*, 1226-1173, 제56호, pp.185-194, 2005