

지능형 홈네트워크 서비스를 위한 PDA 기반 음성처리 및 정보검색 기술

KT 김학균 · 정영준 · 구명완
솔트룩스 정용일 · 이경일

1. 서 론

현재 정부에서는 IT839 전략을 통하여 지식정보화 사회를 고도화하고 지능화하여 지능기반사회인 u-Korea의 진입을 서두르고 있다. 홈네트워크는 8대 신규서비스에 포함되어 있으며, 현재 활발히 연구가 진행되고 있다. 미국에서는 2003년부터 유비쿼터스 컴퓨터기반 소프트웨어를 개발하여 망연결 통신기술, 센서 네트워크 기술 및 신뢰성과 안정성 제고 기술을 개발하고 있으며, 유럽에서는 Home Environment 라는 과제를 수립하여 2006년까지 36억 유로를 투자하고 미래가정을 위한 테스트베드 운영 및 관련 기술을 개발하고 있다. 그리고, 일본에서는 2007년까지 모든 가정에 정보가전을 보급하여 e-life 생활양식을 확산시키기 위해 e-Japan IT전략사업을 통해 홈네트워크 정책을 제시하고 있다[1].

홈네트워크 서비스에서 제공하게 될 다양한 서비스는 궁극적으로 사용자의 상황(context) 및 요구에 최적으로 대응하는 것이다. 현재 미래의 홈네트워크 서비스에서 고려되고 있는 상황은 집안에서의 사용자의 위치, 개인의 정서상태, 그리고 개인의 선호 정보 등이 있다. 또한 사용자와 홈네트워크 서비스와의 인터페이스로는 사람과 사람 사이의 자연스러운 인터페이스 방식인 멀티모달 인터페이스가 최적의 해법으로 고려되고 있다.

본 고에서는 지능형 홈네트워크 서비스에서의 음성 기반 멀티모달 인터페이스 및 맞춤형 정보검색 기술에 관하여 기술한다. 본 고에서 정의하고 있는 지능형 홈네트워크 서비스는 사용자의 위치정보, 선호정보를 기반으로 하여 사용자의 요구에 최적의 결과를 가전제품 제어 등으로 제공하고 있으며, 인터넷 검색 서비스도 제공할 수 있다. 사용자의 요구를 위한 인터페이스 단말기로는 PDA를 사용하였으며 홈네트워크 서비스와 무선랜 환경에서 연결된다고 가정하였다.

PDA 단말기의 리소스를 최소로 사용하기 위하여

무선랜이라는 모바일 환경하에서 패킷 기반의 음성정보 처리를 수행하는 분산음성인식(DSR: Distributed Speech Recognition) 기술이 사용되었으며, 사용자의 인터페이스를 최적화할 수 있는 멀티모달 인터페이스 기술도 개발되었다. 이러한 인터페이스가 구현되고 있는 홈네트워크 환경은 정보기기, A/V 시스템, 가전기기 등 다양한 성능과 형태의 기기를 대내의 몇몇 구성원이 동일하게 이용하는 서비스 환경이며, 멀티모달 인터페이스 기술은 사용자에 따른 이용형태, 이용환경의 변화에 무관하게 통일된 방법으로 서비스를 제공할 수 있는 수단이 된다. DSR기술은 무선랜에서의 패킷 손실에 의한 음성인식 성능 저하를 막기 위해 단말기에서는 잡음제거와 음성의 특징만을 추출하여 전송을 하고, 서버에서 이 특징을 기반으로 음성인식 처리를 하는 방식으로 무선환경에서의 음성처리에 필수적인 기술이다.

1장의 서론에 이어 2장에서는 음성인식 기반 멀티모달 인터페이스에 대해 기술하고, 3장에서는 차세대 웹 기술인 시맨틱 웹에 기반한 맞춤형 정보 검색 및 텍스트 마이닝 기술에 대해 소개한다. 그리고 4장에서는 상기 소개된 기술로 구현된 무선랜 기반의 지능형 홈네트워크 시스템에 대한 구성과 시험 서비스에 대해 간단히 언급하고 마지막으로 5장에서 결론을 맺기로 한다.

2. 음성 기반 멀티모달 인터페이스 기술

최근 음성처리 기술이 발전하고, 성능이 개선된 지능형 초소형 단말기가 출현하면서 새로운 인터페이스 기술에 대한 연구가 본격적으로 진행되고 있다. 대표적인 연구가 멀티모달 인터페이스이다. 멀티모달 인터페이스란 음성(speech), 시각(visual), 촉각(tactile) 등과 같은 단일 모달리티(Modality)를 복합적으로 이용하여(Multi Modality) 보다 편리하고 쉽게 컴퓨터와 대화하는 것을 말한다. 멀티모달 인터페이스는 음성

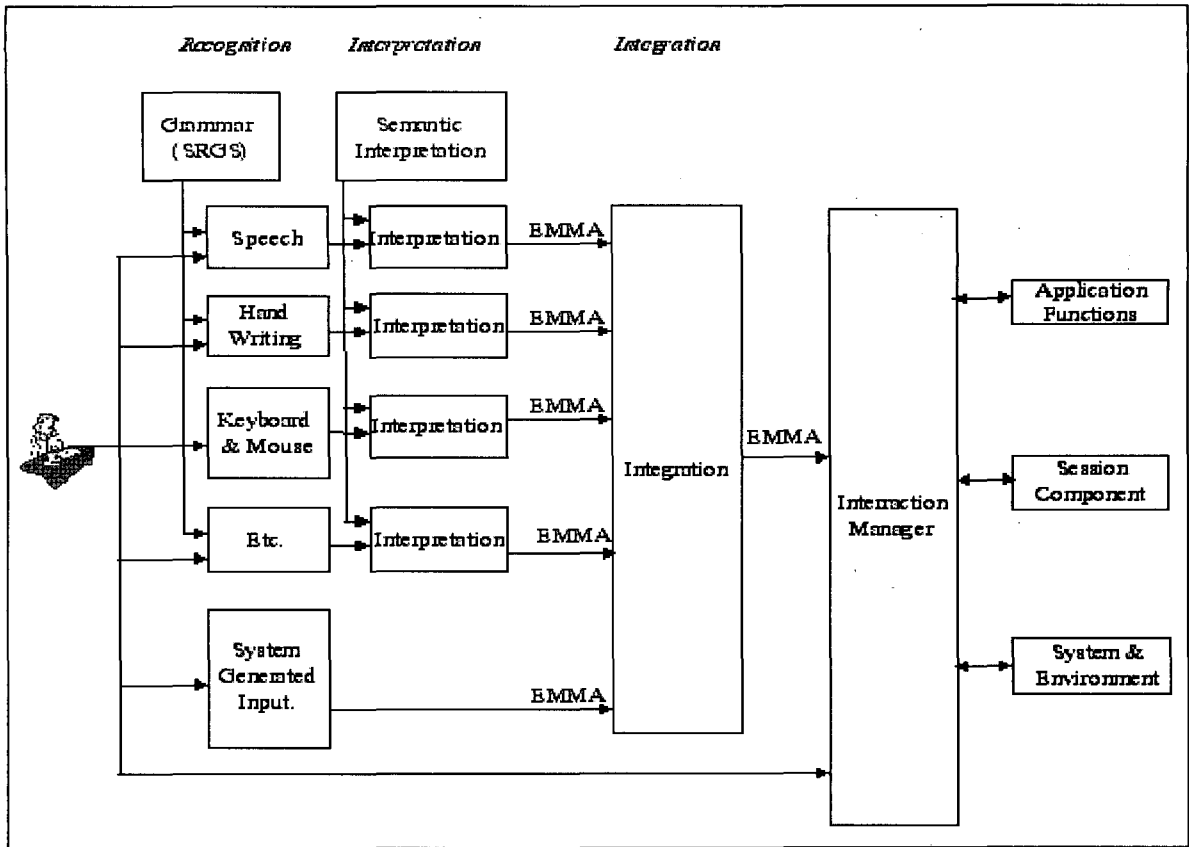


그림 1 입력 요소

입, 출력을 처리하는 음성 인터페이스, 펜 필기체를 인식하는 잉크 인터페이스 및 보편적인 그래픽 인터페이스를 모두 포함하고 있으며, 앞으로 새로이 발전하는 인터페이스 기술들을 추가할 수 있다. 이와 같은 다양한 모달리티 인터페이스간의 상호관계(Interaction)에 따라 다음과 같은 Multimodal Interaction Framework을 정의한다[2].

2.1 Multimodal Interaction Framework

Multimodal Interaction Framework 은 멀티모달 인터페이스를 이용한 응용서비스를 위한 구성 요소를 말하며, 입/출력 요소(Modality Components), 인터랙션 매니저(Interaction Manager), 세션(session) 요소, 시스템 환경 요소 및 응용 서비스 실행 요소로 나누어진다.

2.1.1 입력 요소(Input Components)

입력 요소는 그림 1과 같이 음성, 필기체, 키보드 등을 통해 사용자의 입력을 받고, 이를 인지하는 인식 모듈(Recognition), 인식된 결과의 의미(semantic) 정보를 판단하는 해석모듈(Interpretation)을 거쳐서 사용자의 명령 및 모달리티의 정보를 포함하는 데이터(EMMA: Extensible Multi-Modal Annotation)를

생성한다. 펜으로 원을 그리면서 “이 근처 식당 추천해주세요”와 같이 하나 이상의 입력 요소의 결과가 복합적으로 사용되어야 사용자의 명령이 완성되는 경우는 통합모듈(Integration)을 거쳐서 하나의 데이터로 통합되어 인터랙션 매니저로 전달된다. 음성/필기체 등이 입력되면 인식 모듈은 음향, 언어 및 문법 정보, 필기체 모형 등을 이용하여 텍스트로 변환한다. 그리고 포인팅 디바이스가 입력이 되면 버튼을 누르는 순간의 x-y 좌표 정보를 추적하게 된다. 해석 모듈에서는 인식된 결과를 의미적으로 해석해서 인식결과에 의미 태그를 붙인다.

2.1.2 멀티모달 기술 언어(EMMA)

EMMA는 입력 요소의 결과를 표현하는 표준 언어이다. 음성, 펜, 키보드와 같은 입력 모달리티를 통해 입력 받은 사용자 명령을 해석한 데이터와 입력 모달리티의 특성, 환경 등에 관련된 정보가 포함된다. 따라서 EMMA는 멀티모달 플랫폼의 컴포넌트들간의 표준화된 데이터의 교환 포맷으로 사용된다. EMMA 문서는 다음과 같이 세 가지 구성 요소를 가진다.

- **의미 데이터(Instance data)** : 입력요소의 해석 모듈(Interpretation)을 통해서 해석된 결과로써 XML 문서로 기술한다.

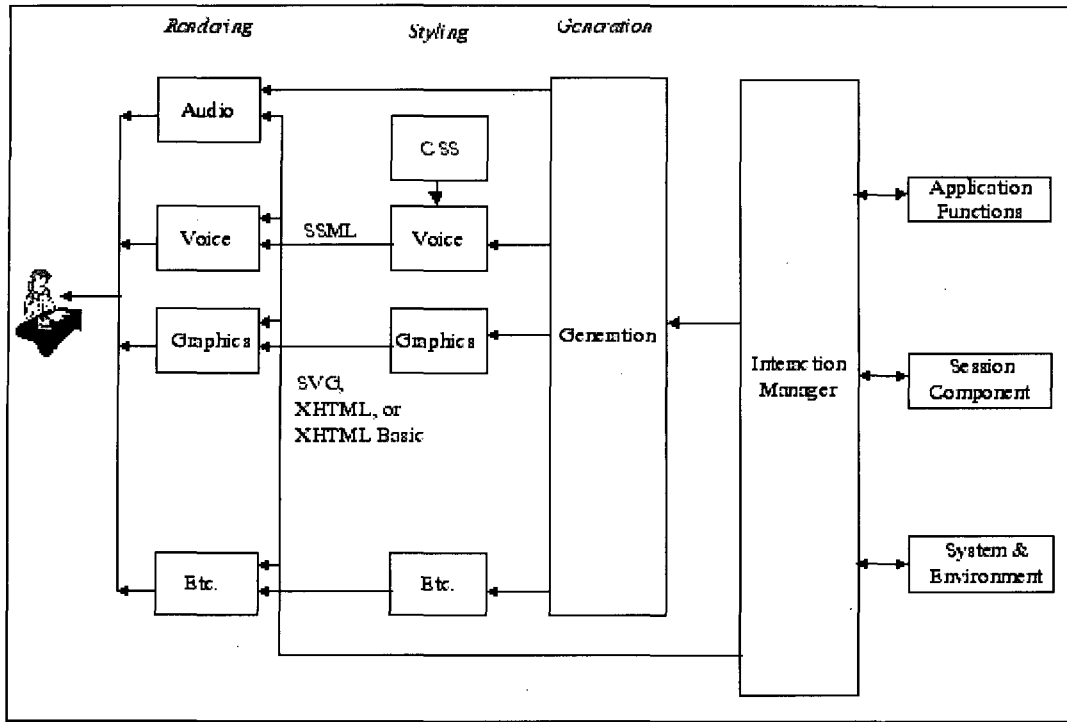


그림 2 출력 요소

- **데이터 모델(Data model)** : 의미 데이터의 구조에 대한 제약을 정의한다. 의미 데이터의 DTD(Document Type Definition) 이다.
- **메타데이터(Metadata)** : 입력 모달리티의 특성을 반영한 속성이다. 입력 모달리티의 종류, 입력 시간, 입력 종료 시간, 인식 결과의 신뢰값(confidence) 등이 포함된다.

2.1.3 출력 요소(Output Components)

출력 요소는 그림 2와 같이 생성 모듈(Generation), 스타일 모듈(Styling) 및 렌더링 모듈(Rendering)로 나누어진다. 생성 모듈은 인터랙션 매니저가 사용자에게 제공할 정보를 어떤 모달리티를 활용하여 출력할 것인지를 결정한다. 이 때, 출력 모달리티는 음성, 그래픽 등이 될 수 있다. 스타일 모듈에서는 모달리티의 특성에 맞게 포매팅 할 때 스타일 정보들을 기술한다. 렌더링 모듈에서는 스타일 모듈에서 생성된 정보를 해석하여, 모달리티의 특성에 맞게 사용자에게 음성, 그래픽 등을 출력한다. 따라서, 렌더링 모듈은 사용자에게 가장 가까운 브라우저를 통해 구현된다.

2.1.4 인터랙션 매니저(Interaction Manager)

인터랙션 매니저는 입력 요소로부터 얻은 EMMA 데이터를 이용하여 실제 응용 서비스를 실행 한 후, 그 결과를 출력 요소에 제공한다. 또한, 음성 I/O 인터랙션, 키보드/마우스 입력과 그래픽 출력을 하는 인터랙션 등과 같이 다양한 모달리티의 인터랙션을 통합하여,

모달리티 간의 동기화(Synchronization)를 이룬다. 즉, 사용자가 보고 듣고 만지는 모든 모달리티가 표현하는 정보가 모두 동일한 내용을 가지기 위함이다. 동기화가 이루어지지 않으면, 모달리티간에 서로 다른 정보를 제공하는 모호한 상태(ambiguity)가 발생한다. 인터랙션 매니저는 다양한 모달리티로부터 들어오는 이벤트에 따라서 상태 변화(state transition)를 일으키는 상태 천이 기계(state machine)으로 구성되며, 현재 상태 및 이벤트 정보를 바탕으로 응용 서비스 실행 및 사용자와의 상호작용(Interaction)을 관리한다. 예를 들어, 사용자가 그래픽 화면의 입력 폼의 포커스를 변경한 경우 포커스의 변경 이벤트가 발생되며, 인터랙션 매니저는 그 이벤트를 받아 상태 변화를 일으킨다. 이 때, 동기화를 위해서 출력 요소를 이용하여 다른 모달리티에 포커스의 변경을 알림으로써, 모달리티 간의 동기화를 이룬다. 인터랙션 매니저를 위한 상태 천이 기계를 표현하는 언어로써, SCXML(State-Chart XML)이 사용된다. 인터랙션 매니저와 연결된 세션 요소에서는 다양한 디바이스가 사용되더라도 세션이 끊어지지 않도록(seamless) 멀티모달 응용서비스와의 지속적인 연결 및 다양한 단말기 출력을 위한 싱크 기능이 있다. 시스템 환경 요소는 DCI(Delivery Context Interface) 라 불리며, 디바이스 선정 및 사용자 선호에 따라 자동으로 변화하도록 시스템 환경을 표현해 준다. 즉, 출력 모드를 주변 환경 및 상황에 따

라 모바일 디바이스, 자동차 용 디바이스 및 데스크 탑 등으로 사용자가 쉽게 바꿀 수 있도록 환경을 구성해 준다.

3. 시맨틱 웹 기반 정보 검색 및 텍스트 마이닝 기술

기존 웹 환경에서 포괄적으로 사용되고 있는 문자열 일치 방식의 전문 검색(full text retrieval) 기술은 유비쿼터스 컴퓨팅 환경에서 사용자 상황과 관심에 적합한 맞춤형 정보검색 서비스를 구현하는데 많은 한계를 보이고 있다. 이러한 한계를 극복하기 위한 대안으로 온톨로지와 의미 메타데이터에 기반한 정보 처리 기술인 시맨틱 웹(Semantic Web)[3]에 대한 연구가 활발히 연구되고 있다. 시맨틱 웹은 W3C에서 산업 표준으로 제안되었으며, 본 연구에서는 웹 온톨로지 언어인 OWL[4]을 사용해 정보 검색을 위한 온톨로지를 표현하고, WWW 상에 이미 존재하는 텍스트 콘텐츠를 의미 메타데이터인 RDF로 자동 변환, 표현하여 이를 사용자 맞춤형으로 검색하는 지능형 정보 검색 서비스 시스템을 구현하였다. 온톨로지 기반의 정보 처리 기술은 유비쿼터스 환경에 존재하는 다양한 디바이스와 서비스 에이전트 간의 명확한 의사소통을 가능케 하며 이를 통한 상호 운영성을 보장한다[5].

본 장에서는 사용자의 관심정보를 수집하기 위한 focused web crawling 기술과 수집된 정보를 자동 분류하고 의미 정보를 추출(information extraction), 각 사용자 단말 환경에 적합하게 자동 요약(automatic summarization)하는 기능을 수행하는 텍스트

마이닝 기술, 의미 메타데이터를 저장, 관리, 검색하기 위한 context managing 기술과 정보 검색용으로 설계된 온톨로지에 대해 설명을 한다.

3.1 Focused Web Crawling과 Text Mining 기술

Focused web crawling이란 웹에 존재하는 수 많은 정보 site에서 사용자의 관심 정보만을 선별, 수집하는 기술로 현재의 web crawling 기술을 발전시킨 것이다. 수집된 정보 중 사용자 관심 정보를 변별하기 위해 자동 분류 및 필터링 기술이 사용될 수 있는데, 본 연구에서는 SVM이 사용되었다. 그림 3과 같이, 수집된 정보는 온톨로지 기반의 RDF로 변환이 되며, 이 때 기존 콘텐츠를 단순 RDF로 변환하는 것이 아니라, 키워드 추출, 개체명(named entity)과 같은 중요 정보 추출(information extraction)을 수행하며, 각 유비쿼터스 단말기에 표현되기 적절한 수준의 길이와 복잡도로 콘텐츠를 자동 요약(summarization) 하는 등의 텍스트 마이닝 기술이 동시 적용되었다. 이렇게 자동 수집, 필터링 및 자동 분류, 정보 추출 및 자동 요약된 RDF 형태의 정보는 리스닝 과정을 거쳐서 각 사용자에게 맞춤형 정보 제공 서비스를 제공하게 된다.

아래의 OWL 문서는 기업 정보 수집을 위한 온톨로지의 예를 보여주며, RDF문서는 수집된 웹 사이트 정보에서 정보를 추출하여 이를 RDF로 변환한 결과를 보여준다. 예에서 보는 바와 같이 <http://www.saltlux.com> 이란 webpage는 Company라는 의미를 갖고 있다는 것을 알 수 있고, 이 회사의 이름은 Saltlux이고 제품은 [IN2]TMS라는 추가적인 정보도 알 수 있다.

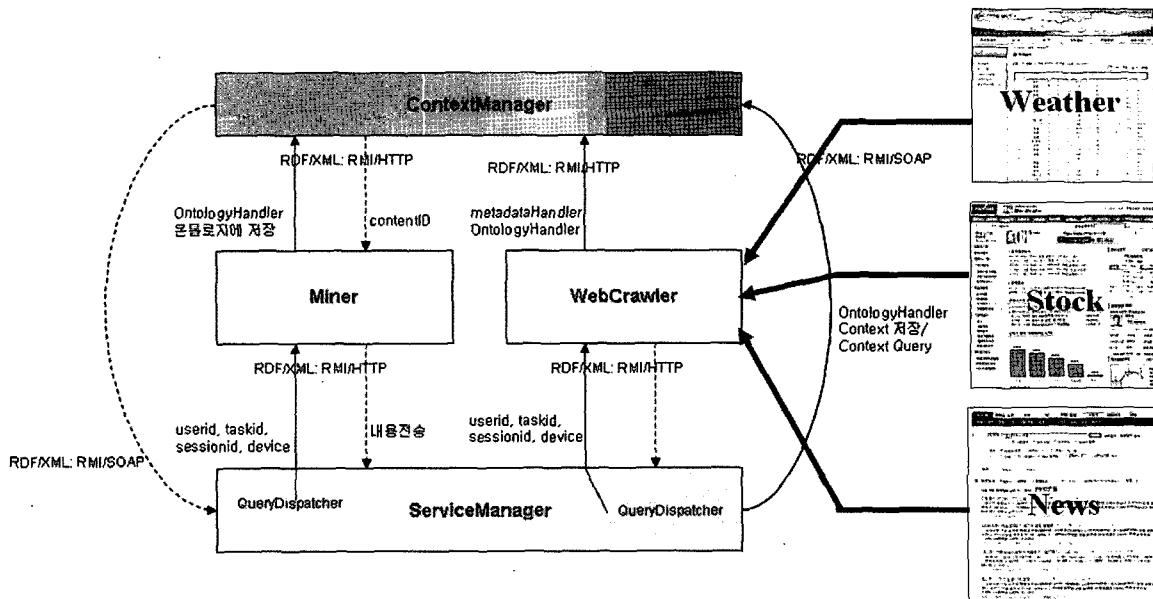


그림 3 Focused Web Crawling과 시스템 구성

```

<owl:Class rdf:ID="Enterprise">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:cardinality rdf:datatype="
        http://www.w3.org/2001/XMLSchema#int"
        >1</owl:cardinality>
      <owl:onProperty>
        <owl:DatatypeProperty rdf:ID="hasName"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:DatatypeProperty rdf:ID="hasProduct"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
  ...
</owl:Class>

<rdf:Description about=http://www.saltlux.com>
  <rdf:type resource="#Company">
    <s:hasName>Saltlux</s:hasName>
    <s:hasProduct>[IN2]TMS</s:hasProduct>
    ...
</rdf:Description>

```

3.2 Context managing 기술

Context managing 기술이란 현재의 웹이나 유비쿼터스 환경에서 발생하는 다양한 콘텍스트를 효과적으로 수집, 관리, 저장, 검색하기 위한 기술이다. 그림

4는 본 연구에서 개발된 지능형 홈네트워크 서비스를 위한 미들웨어에서 사용자 정보 요청에 따른 각 manager간의 콘텍스트 흐름을 보여주고 있다. 그림 4에서 context manager는 OWL 및 RDF로 표현된 콘텍스트를 저장, 관리하고, 지식 저장소(Knowledge Repository)의 각종 메타데이터 관리와 RDBMS와의 연결, OWL문서에 대한 검색과 다양한 파일 개체에 대한 저장, 관리하는 기능을 수행한다. Service manager는 상황인지 서비스를 위해 각 사용자 및 사용자 그룹을 위한 서비스 세션과 태스크 및 서브태스크를 생성, 관리하고, 각 사용자의 privacy 정책을 관리한다. 사용자가 PDA를 사용해 멀티모달 형식으로 정보를 요청하면 service manager에서 태스크가 생성되고, 리즈닝 과정을 통해 사용자 질의 분석 및 서브태스크가 생성되며, 이를 통해 focused crawling, semantic annotation 및 텍스트 마이닝 등의 일련의 과정이 진행되게 된다.

3.3 정보 검색을 위한 domain ontology

특정 도메인에 대한 지능형 정보검색 서비스를 위해서는 도메인에 대한 온톨로지의 구축이 필수적이다. 현재 홈네트워크 서비스를 위해 인물, 지역, 날씨, 주식, 영화, 회사, 교통 등이 온톨로지가 설계되어 있다. 그림 5는 홈 네트워크를 위해 설계된 온톨로지의 개념도를 보여주고 있으며, 그림 6는 특정 지역(평창)에 대해 사용자가 날씨 정보 검색을 요청하였을 때의 검색 결과를 보여주고 있다. 검색 결과를 보면, 사용자 관

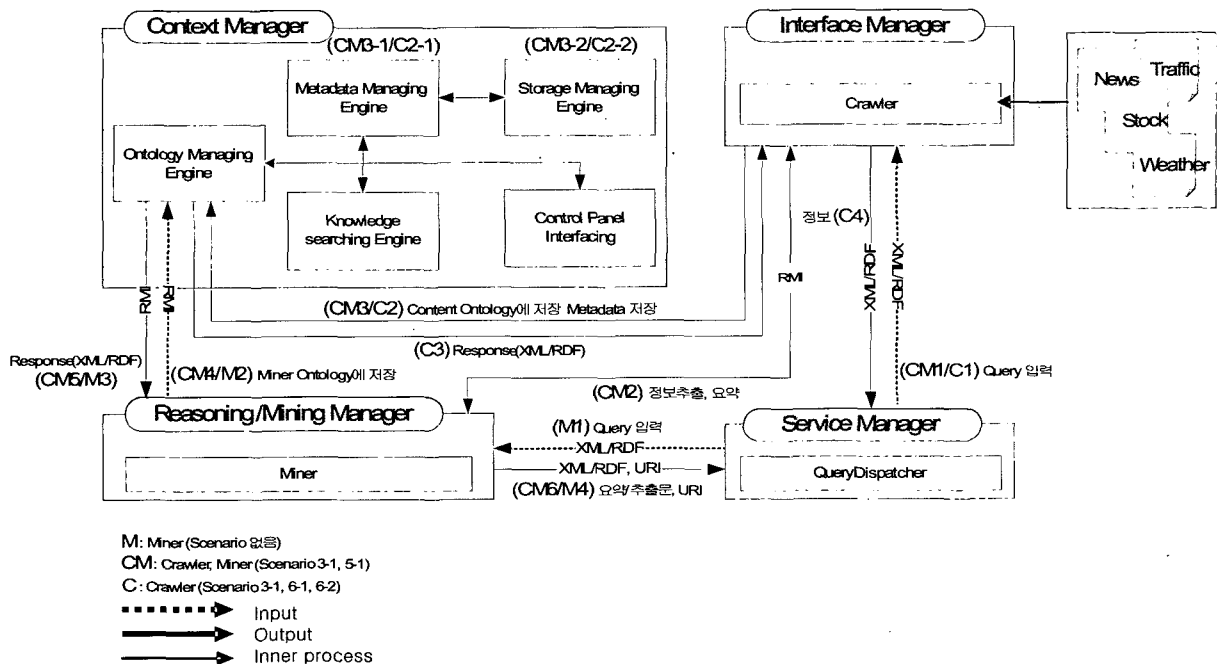


그림 4 manager간의 context flow

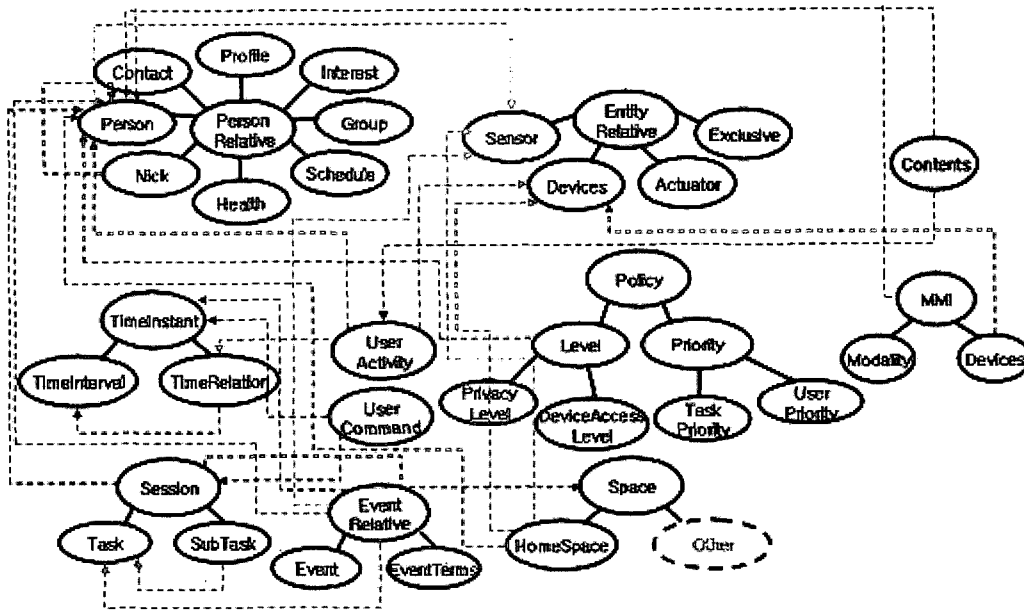


그림 5 홈 네트워크에서의 컨텍스트 온톨로지 개념도

```

<terminated> QueryDispatcher [Java Application] C:\Program Files\Java\j2re1.4.2_08\bin\javaw.exe (2005. 12. 14. 오후 4:08)
[SERVICE MANAGER] QueryDispatcher
[*]CMIconnection : successfully connected to: //localhost:3399/CMF_FACTORY
11687 ms): + query [0] = <?xml version='1.0' encoding='euc-kr'?><rdf:RDF xmlns:rdf='ht
WeatherMiner_200512141608431134544123734]
<rdf:RDF
  xmlns:j.0="http://221.148.161.165:3390/20050315/miner.owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:j.1="http://221.148.161.165:3390/20050315/ir.owl#">
  <j.1:Weather rdf:about="http://221.148.161.165:3390/20050315/miner.owl#WeatherMiner_2005
  <j.0:hasTaskName>IR</j.0:hasTaskName>
  <j.0:hasKeywords>평창|날씨정보|날씨</j.0:hasKeywords>
  <j.1:hasOutingIndex>50</j.1:hasOutingIndex>
  <j.1:hasSportsIndex>50</j.1:hasSportsIndex>
  <j.1:hasHeatingIndex>90</j.1:hasHeatingIndex>
  <j.0:hasUri>http://www.kma.go.kr/kor/weather/forecast/forecast_03_01.jsp?region3=11D1
  <j.1:hasColdIndex>90</j.1:hasColdIndex>
  <j.0:hasDeviceInfo>PDA</j.0:hasDeviceInfo>
  <j.1:hasLowTemp>-</j.1:hasLowTemp>
  <j.0:hasSummarizeForPDA>강원도 평창 지역의 날씨정보입니다. 오늘의 날씨는 맑음, 기온은 최고 -3도 입니
  <j.1:hasCarWashingIndex>40</j.1:hasCarWashingIndex>
  <j.1:hasHighTemp>-3</j.1:hasHighTemp>
  <j.0:hasPersonID>Person_0003</j.0:hasPersonID>
  <j.1:hasRainAMPercent>-</j.1:hasRainAMPercent>
  <j.0:hasTitle>평창의 날씨정보입니다.</j.0:hasTitle>
  <j.1:hasRegion>강원도 평창</j.1:hasRegion>
  <j.1:hasRainPMPercent>-</j.1:hasRainPMPercent>
  <j.1:hasWeatherDescription>맑음</j.1:hasWeatherDescription>
  <j.0:hasContentURI>Content_200512141608431134544123734</j.0:hasContentURI>
  <j.1:hasClothesWashingIndex>50</j.1:hasClothesWashingIndex>
  <j.0:hasTaskID>Task_00002</j.0:hasTaskID>
  </j.1:Weather>
</rdf:RDF>
  
```

그림 6 날씨정보를 검색한 결과

심 지역의 날씨 정보가 PDA에서 음성으로 출력되기 위해 자연언어로 생성, 자동 요약된 것을 확인할 수 있다.

4. 지능형 홈네트워크 시스템 및 서비스

4.1 멀티모달 단말기

멀티모달 단말기는 스타일러스 펜과 음성을 통한 멀티

티모달 입력을 받아들이는 역할을 수행한다. 즉, 사용자와 그래픽 및 펜 인터페이스를 수행하는 HTML 브라우저가 내장되어 있다. 그리고, 음성 입출력을 처리하기 위한 Voice I/O 프로세서가 있다. 음성 인식 환경은 DSR(Distributed Speech Recognizer) 표준을 따른다. ETSI DSR표준에는 그림 7과 같이 분산 음성 인식을 위한 MFCC 계산, 양자화, 에러 처리 등

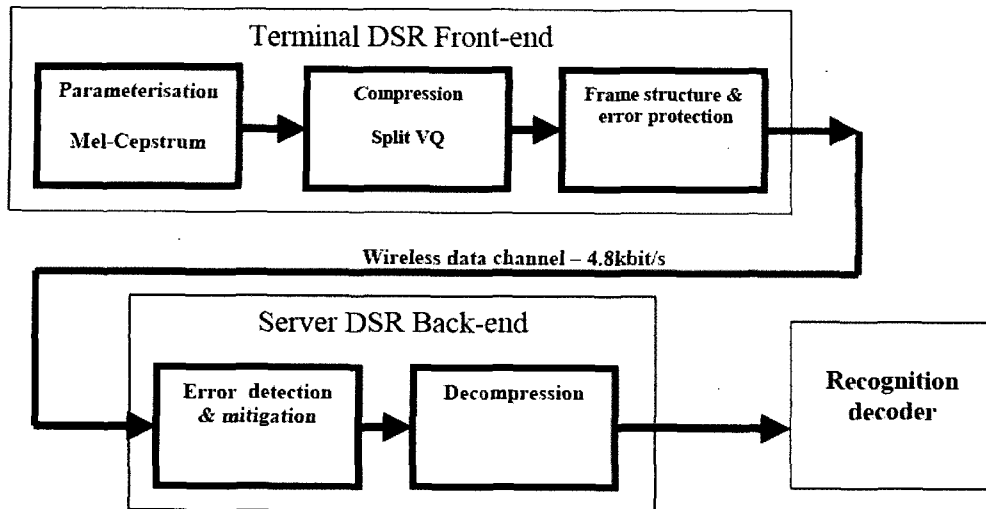
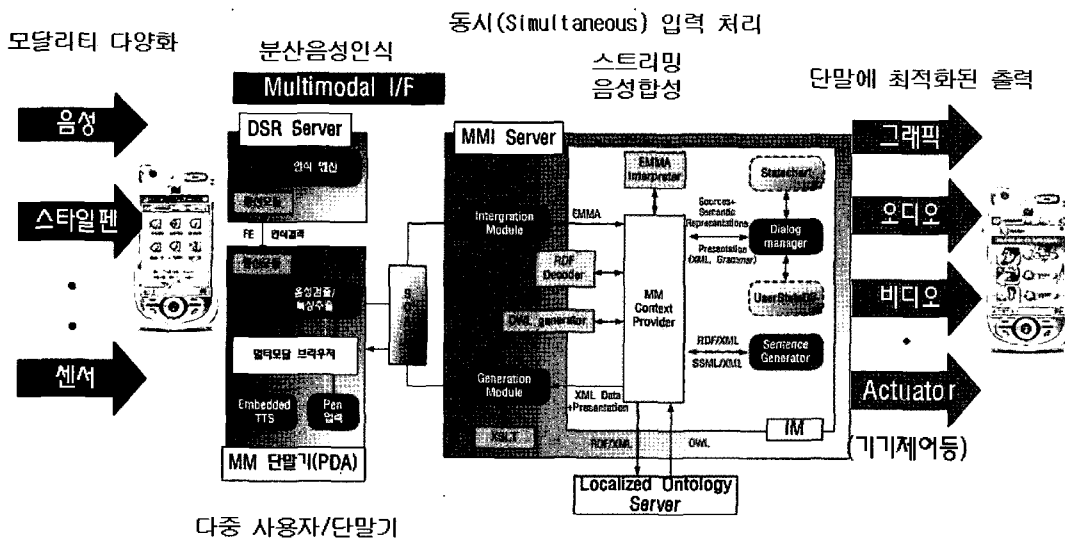


그림 7 DSR 블록 다이어그램



다중 사용자/단말기

그림 8 멀티모달 인터랙션 아키텍처

에 대한 틀이 제공된다[6]. 따라서, Voice I/O 프로세서는 음성입력을 받아 표준화에 따라 특징을 추출한 뒤, 무선랜 구간을 통해 DSR 서버에 전달되어 입력된 특징 정보에 따라 인식을 수행한다[7].

4.2 멀티모달 플랫폼

멀티모달 플랫폼은 앞서 제시된 MultiModal Interaction Framework에 따라서 아키텍처를 구성한다. 그림 8과 같이 주요 구성요소로서, LOS(Localized Ontology Server), MMCP(MultiModal Context Provider), IM(Interaction Manager), Integration Module 및 Generation Module이 있다. LOS는 데이터의 마이닝 및 정보 요약, 추론 등이 수행되는 지식 서비스의 핵심 요소이다. MultiModal Context Provider는 LOS와 단말기로부터 오는 메시지를 관리 배분하고, IM은 멀티모달의 인터랙션을 통제하는

역할을 수행하며, SCXML 문서를 해석하고 천이 기계(state machine)를 구성하는 Dialog Manager가 핵심요소로 존재한다.

4.3 서비스 시나리오

PDA 기반 음성 기반의 멀티모달 인터페이스 및 언어 검색 기술 연구를 활용한 지능형 홈네트워크 서비스는 집안의 사용자가 PDA를 가지고 무선랜 환경에서 음성, 스타일러스 펜, 필기체 인식 등의 멀티모달 방식으로 입력을 하여 집안의 다양한 기기를 제어하거나 인터넷 검색 기술에 의해 요청한 정보를 찾아내고, 사용자의 선호도에 맞게 결과를 요약하며 그 결과를 가진 기기 혹은 PDA와 관련된 디스플레이를 통하여 출력한다. 다음은 구현될 시나리오의 전개를 위한 환경을 정리하였다.

- 가족 구성원들은 핸드폰 기능을 포함하는 PDA를

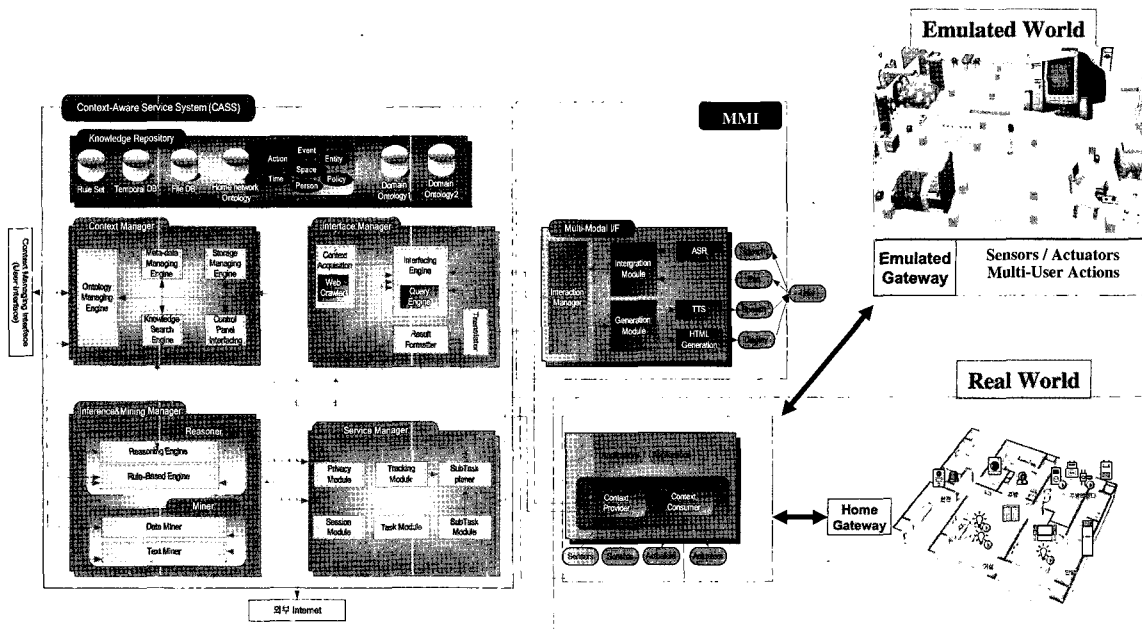


그림 9 홈네트워크 서비스의 실제 공간과 가상 공간

보유

- 맥내에 설치된 홈게이트웨이는 가정 내의 모든 센서 및 가전 제품들과 양 방향 커뮤니케이션을 수행
- PDA는 맥내 혹은 맥외에 설치된 멀티모달 지능형 홈서버와 통신하여, 개인화된 정보 제공
- 거실, 개인의 방과 주방 등에는 PC 기능의 단말 디스플레이가 설치되고, 이 단말기는 홈게이트웨이에 연결됨
- 현관에는 소형의 디스플레이가 설치되어 가족들이 외출과 귀가를 할 때 필요한 정보 등을 제공
- 멀티모달 지능형 홈서버는 자동으로 수행하는 일과 사람의 지시(수동 조작, 음성 명령 등)에 의하여 수행할 일에 대한 프로그램이 디폴트로 조정되어 있으며, 이 프로그램은 학습 기능에 의하여 가족 구성원의 특성과 생활 습관에 따라 능동적으로 변화됨.

이러한 환경하에서 다음과 같은 시나리오에 따라서 멀티모달 인터페이스 및 언어 검색 기술을 활용한 지능형 홈네트워크 서비스를 구성하였다. 서비스는 KTF 차세대연구소의 홈네트워크 데모를 활용하여 구현이 되었으며, 또한 그림 9와 같은 가상 공간을 활용하여 서비스의 현재 상태를 보여준다.

아래와 같이 크게 세 개의 꼭지로 서비스를 제시하고, 여기에는 포함되지 않았지만 지능형 멀티모달 환경으로 구성할 수 있는 다양한 서비스들이 있다.

4.3.1 유비쿼터스 환경지원 Context 정보 관리

- 가족확인 서비스 - 그림 10과 같이 집 외부에서 인

증된 가족 구성원이 음성 혹은 버튼 클릭으로 집안에 가족 구성원과 그들의 현재 상태를 확인할 수 있다.

- 홈네트워크 고장 진단 서비스 - 맥내외부에서 고장 진단을 통해 고장난 기기들을 판단할 수 있다.
- 영화보기 서비스 - 개별적인 기기제어를 넘어선 복합적인 기기제어를 의미한다. 영화를 보기 위해서는 커튼을 닫고, TV를 켜고, 전등은 끄는 복합적인 기기제어가 필요하다. 따라서, 사용자가 영화보기라는 명령을 내리면, 현재 사용자의 선호 환경 및 위치를 기준으로 주변의 기기들이 제어된다.

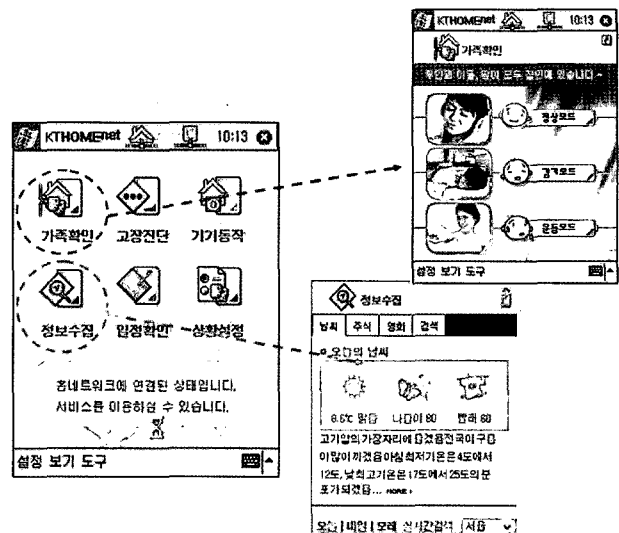


그림 10 가족확인 서비스

4.3.2 지식서비스 지원 정보 마이닝 기술

- **사용자 단말기 기반의 정보 검색 및 요약 서비스** - 사용자가 "정보수집" 메뉴를 통해 날씨, 주식, 영화 추천, 검색 서비스를 받는다. 일반적인 검색 서비스와 달리, 사용자 단말기 및 사용자의 위치에 기반하여 가장 가까운 홈 디스플레이에 그 결과가 표시된다. 또한, 검색 결과는 단계별 정보 요약 기술이 적용되어, 단말기의 특성에 따라 적합한 형태의 요약된 정보로 사용자에게 표시된다.
- **아침인사 서비스** - 가족 구성원이 아침에 일어나면, 센서를 통해 구성원의 움직임을 확인하고, 구성원이 선호하는 콘텐츠, 조건 뉴스 혹은 오늘의 일정을 홈 디스플레이와 단말기에 표시한다.

4.3.3 정보서비스 개인화 기술

- **사용자 사생활 정보보호 서비스** - 다른 가족 구성원의 일정 등을 조회하고자 할 경우, 구성원의 권한과 다른 가족 구성원의 권한에 따라 일정 등이 적절하게 공개, 비공개 되는 서비스이다.

5. 결 론

본 고에서는 지능형 홈네트워크 서비스를 위한 PDA 기반 음성 및 언어 처리 기술 연구를 소개하였다. 지능형 홈네트워크 서비스는 사용자의 자연스러운 인터페이스인 멀티모달 인터페이스가 보장되어야 하며 사용자의 상황에 최적의 결과를 정보 가전기기를 통하여 제공하여야 한다. 이를 위해서는 무선랜과 PDA라는 제약된 환경에서 사용자에게 다양한 서비스를 최적의 형태로 제공할 수 있는 기술들로서 분산음성인식 기술, 멀티모달 인터페이스 기술, 그리고 정보 검색/요약 기술이 필요하며, 이러한 기술들을 활용할 수 있는 지능형 홈네트워크 시스템이 필요하다.

본 고에서는 멀티모달 인터페이스 및 시맨틱 웹 기반의 정보 검색 기술을 지능형 홈네트워크 서비스 시스템에 적용함으로써 효과적인 상황인지 서비스 구현이 가능함을 보였다. 현재 이러한 인터페이스 및 상황인지 기술은 국제적으로 표준화가 활발히 진행되고 있으며 향후 국내에서도 지속적인 활동과 기술개발이 필요한 분야이다.

무선환경에서의 음성처리 기술과 이에 기반한 멀티모달 인터페이스 기술, 그리고 사용자 선호도나 단말사양에 따라 다양한 정보를 최적화된 형태로 요약하여 제공할 수 있는 차세대 정보 검색 기술은 향후 홈네트워크 서비스가 더욱 지능화되고 텔레매틱스나 로봇틱스 등의 다양한 서비스들과 연계되는 유비쿼터스 환경으로 진화하는데 있어서 기반이 되는 기술로 자리잡을 것이다.

참고문헌

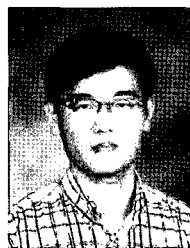
- [1] IT839 전략 기술개발 Master Plan, 정보통신부, 2005년 6월
- [2] W3C MultiModal Interaction Framework, <http://www.w3.org/TR/mmi-framework/>, W3C Note, May, 2003.
- [3] W3C, <http://www.w3.org/2001/sw/>.
- [4] Heiner Stuckenschmidt and Frank van Harmelen, "Ontology Based Metadata Generation from Semi-Structured Information," Proceedings of the First Conference on Knowledge Capture(K-CAP'01), Victoria, Canada, Oct. 2001.
- [5] Tim Berners-Lee, "Semantic Web Road map", "<http://www.w3.org/DesignIssues/Semantic.html>", 1998.
- [6] ETSI standard document, "Speech Processing, Transmission and Quality aspects(STQ); Distributed speech recognition; Extended advanced front-end feature extraction algorithm; Compression algorithm; Back-end speech reconstruction algorithm," ETSI ES 202 212 v1.1.1, Nov. 2003
- [7] 박성준, 한기수, 구명완, "음성인식 WiFi 전화시스템 개발", 제22회 음성통신 및 신호처리 학술대회 논문집, pp. 89-92, 2005.

김 학 군



1997 연세대학교 컴퓨터학과(이학사)
2000 연세대학교 컴퓨터학과(공학석사)
2000~현재 KT 미래기술연구소 음성언어 연구담당 음성응용연구부 선임연구원
관심분야: 음성정보처리 시스템 개발 및 응용, 멀티모달
E-mail : playbug@kt.co.kr

정 영 준



1986 연세대학교 전자공학과(공학사)
1989 한국과학기술원 전기 및 전자공학과(공학석사)
1989~현재 KT 미래기술연구소 음성언어 연구담당 멀티모달연구부장
관심분야: 음성정보처리 시스템 개발 및 응용, 음성인식, 음성합성, 멀티모달
E-mail : yjjung@kt.co.kr

구 명 완



1982 연세대학교 전자공학과(공학사)
1985 한국과학기술원 전기 및 전자공학과
(공학석사)
1991 한국과학기술원 전기 및 전자공학과
(공학박사)
1996~1997년 미국 벨연구소 방문연구원
2003 서울대학교 KT-MBA 과정 수료
1985~현재 KT 미래기술연구소 음성언어
연구담당장(상무)
관심분야: 음성정보처리 시스템 개발 및
응용, 음성인식, 음성합성, 멀티모
달, 음성통역
E-mail : mwkoo@kt.co.kr

정 용 일



2004 동국대학교 전자상거래학과(석사)
2004 동국대학교 전자상거래연구소 연구원
2004~현재 주식회사 솔트룩스 HLT연구
소 전임연구원
관심분야: Semantic Web, Ontology,
Knowledge Retrieval
E-mail : yijeong@saltlux.com

이 경 일



1996 인하대학교 전자재료공학과(공학석사)
1996~1999 LG 중앙연구소(연구원)
1999~2000 현대전자 연구소(연구원)
2000~현재 주식회사 솔트룩스(부사장)
2004~현재 언어음성산업협회 부회장
2004~현재 인하대학교 정보통신공학부
겸임교수
관심분야: 자연언어처리, 정보검색, 텍스
트 마이닝, 시맨틱 웹
E-mail : tony@saltlux.com
