멀티미디어 시스템을 위한 입력 멀티모달 제공방법

A Method with Input Multi-modal for Multi-media System

이광로* · 정경석** · 김인홍**

1. 서 론

인터넷을 기반으로 한 멀티미디어 시스템의 보급이 최근 급속히 확산되고 있다. 이러한 변화에 따라 컴퓨터의 사용환경도 고정 사용환경에서 이동 사용환경으로, 기계 중심적인 사용환경에서 인간 중심적인 환경으로 변화하고 있다. 따라서 멀티미디어 시스템의 효율적인 사용을 위해서는 멀티미디어 데이터의 처리 뿐만 아니라 멀티미디어 시스템의 사용환경에 대한 연구가 시급한 상황이다. 즉, 원하는 정보를 쉽게 접할 수 있는 미래 정보화 사회에 적합한 컴퓨터 사용환경을 위해서는 언제, 어디서든지 컴퓨터의 사용이 가능한 이동성(mobility) 지원과 누구든지 사용 가능한 친숙성을 지원할 수 있는 사용자 인터페이스의 개발이 요구된다[1,2]. 이를 위해 사람과 사람의 의사 전달에 있어서 여러 가지의 통신 채널을 사용하듯이 컴퓨터와 사용자 간의 의사 전달에 있어서도 여러 가지의 통신채널을 제공함과 동시에 사용자가 쉽게 들고 다닐 수 있는 크기로 제공되어야 할 것이다. 최근에 들어 음성이나 펜을 이용한 단일 입력 모달리티(uni-modality)가 부분적으로 도입되고 있으나 현재의 입력수단을 대치하기에는 여러 가지 문제점을 가지고 있다. 이러한 문제점은 단순히 인식성능이 개선된다고 해결되는 문제는 아니며, 멀티 입력 모달리티(multi-modality)와 이동성 지원을 위한 근본적인 방법이 요구된다.

본 논문에서는 입력 수단으로 친숙성과 사용성이 좋은 펜과 음성을 이용하여 명령어를 처리할 수 있는 멀티입력 처리방법을 제안하고, 또한 휴대형 단말기의 이동성 지원을 위해 휴대형 단말기로부터 인식엔진이나 서비스 제공 모듈이 동적으로 구성 가능한 사용자 인터페이스 구조를 제안한다. 논문의 구성은 1장에서 서론, 2장에서 단일모달 입력방법, 3장에서 멀티모달의 입력 방법, 4장에서 휴대형 멀티모달 인터페이스 구조, 5장에서 지도 시스템의 구현, 6장에서 평가, 7장에서 결론 및 연구방향으로 구성되어 있다.

2. 단일모달 입력방법

현재 컴퓨터의 입력 수단으로는 키보드와 마우스가 일반적으로 많이 사용되고 있다. 그러나 사용자에 좀더 편리한 사용환경과 휴대성을 지원하기 위해서는 휴대성이 편리한 펜과 마이크의 사용이 가능한 멀티모달이 지원되어야 한다. 또한 이러한 디바이스를 통해 제스처, 명령어 수준의 쓰기와 말하기, 자연어 수준의 쓰기와 말하기를 지원할 수 있어야 한다. 현재 많이 사용되고 있는 단일모달 입력방법의 특성을 살펴보면 다음과 같다.

*정회원, 하이콤 리서치
**준회원, 하이콤 리서치
***종신회원, 하이콤 리서치

키보드 입력방식은 입력 속도가 비교적 빠르고, 입력의 정확도가 아주 높고, 입력 비용이 거의 필요 없기 때문에 컴퓨터에 정보를 입력하는 가장 보편적인 입력 수단으로 사용되고 있다. 그러나 초심자에게 매우 불편하고 휴대성에 있어서 매우 좋지않다.

사람과 사람 사이의 정보전달에 가장 유용한 입력 수단으로 사용되고 있는 음성 입력방식은 입력 속도가 아주 빠르고, 음색이나 발성자의 특징 등 부수적인 정보를 얻을 수 있고, 사용에 있어서 매우 편리하다. 특히 다른 입력방식과 달리 손을 사용하지 않고도 입력이 가능하다. 그러나 사용자나 주변환경의 영향, 인식 적용영역 등에 따라 인식율이 크게 좌우되고, 인식을 위한 많은 컴퓨팅 파워를 요구한다.

펜 입력방식은 시각적 명령 지시 및 대상(object) 선택, 문서의 오류 수정, 수식의 입력 등의 작업에 적합하며, 자료입력 방법이 친숙하고, 휴대성이 좋고, 주변환경에 영향을 받지않는다. 특히 회의 중과 같이 소리를 내서는 안 되는 상황이나 길안내와 같이 시각적 정보를 필요로 하는 상황에서 유용하다. 그러나 자료를 입력할 때 음성에 비하여 제약조건이 많고 인식을 위한 컴퓨팅 파워를 요구한다.

3. 멀티모달 입력방법

멀티모달리티를 사용한 입력이 단일모달리티를 사용할 때보다 사용자에게 더욱 자연스럽고 편리하다. 또한 멀티모달리티에 사용되는 단일모달리티는 서로 보완적이고 잉여정보를 가지고 있어 좀더 정확한 정보전달을 할 수 있다[3,4,5]. 멀티모달을 사용할 때 단일모달의 조합방식과 조합해석 방법은 매우 다양하다. 일반적으로 멀티모달리티는 배타적 멀티모달리티(exclusive multi-modality), 선택적 멀티모달리티(alternate multi-modality), 협동적 멀티모달리티(synergistic multi-modality) 등 세가지로 나누어진다[2]. 현재, 사용자 인터페이스를 개발할 때 직접조작 기술은 조작이 간단하고 가시화 되어 있어 현재 가장 많이 사용되고 있다. 그러나 직접조작 방법은 기능이나 가시화에 있어서 단순한 것 만이 가능하다. 여러 개의 기능을 조합한 작업이나 가시화 하기 어렵거나 할 수 없는 작업을 수행할 수 없다. 이에 반해 자연어 처리(NLP : Natural Language Processing) 인터페이스는 여러 개의 기능을 조합한 작업이나 가시화 하기 어렵거나 할 수 없는 작업을 수행할 수 있다. 그러나 모호성과 처리 기술의 한계 때문에 적용영역이 극히 제한적 있다. <표1>에 자연어 처리 기술과 직접조작 기술의 특성을 비교하여 나타낸다. <표1>에서 알 수 있듯이 이들은 서로 상호 보완적이며, 친숙성과 조작성이 뛰어나다. 따라서 이 두 기술의 상호 보완적인 면을 이용해 휴대 단말기에서 협동적 멀티모달리티 지원을 위한 멀티모달 제공방법을 제안한다.

표 1. 자연어 처리 기술과 직접조작 기술의 특성 비교

	자연어 처리 기술	직접조작 기술
사용법의 친숙성/조작성	○	○
Subset 제약	○	△
작업지시의 모호성	×	○
개발비용/수행비용	×	○
복잡한 작업/병렬작업	○	×
가시화 곤란한 작업	○	×

○: 우수, △: 양호, ×: 불량

시스템의 블록구조는 (그림1)과 같이 데이터 입력부분과 변환부분으로 구성된다. 데이터 입력부분은 펜과 음성을 통해 입력을 받는 음성입력

모듈, 문자입력모듈, 제스처입력모듈과 입력된 데이터를 인식하여 아스키로 변환하는 각각의 인식엔진, 펜이나 음성 입력 후 일정 시간이 지나면 특정 이벤트를 발생시켜 입력종료 시점을 알려주는 시계모듈로 구성되어 있다. 변환부분은 데이터 입력부분으로부터 받은 아스키를 수행 가능한 명령어로 변환하는 결합기, 응용영역에서 사용되는 ICL(Interagent Communication Language) [6]의 타입을 결정하기 위한 필터들을 관리하는 필터 테이블, 외부의 다른 에이전트와의 통신을 위해 ICL의 해석과 포장을 하는 통신모듈로 구성되어 있다.

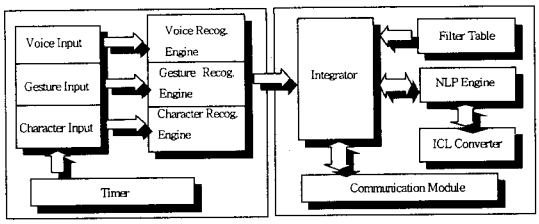


그림 1. 다중 입력 모달리티의 블록구조

기본 동작원리는 먼저 사용자가 펜이나 마이크를 사용하여 데이터를 입력하면 이 데이터는 해당 인식기에 전달되어 인식된다. 인식결과는 아스키 코드로 변환하여 결합기의 버퍼에 저장된다. 결합기는 입력된 데이터를 자연어 처리기를 이용하여 문장을 분석하고, 분석된 결과는 ICL 변환기에 의해 ICL로 변환된다. ICL로 변환된 데이터로부터 프리디킷(predicate)을 추출한다. 추출된 프리디킷을 사용하여 필터 테이블에 등록된 목록에서 적합한 필터를 선택한다. 선택한 필터와 ICL로 변환된 데이터를 이용하여 입력된 정보의 인수를 확인하여 정보의 누락 여부를 확인한다. 이때 누락된 정보가 있으면 사용자에게 정보가 해당 인수의 요구를 위한 메시지를 표시하여 추가정보를 받는다. 입력된 데이터가 완전한 ICL이나 수행 명령어로 변환되면 이것을 조정 에이전트에게 보내 해당 작업을 수행시킨다.

결합기에서의 멀티모달 입력에 대한 데이터의 통합방법은 다음과 같다.

$$R = R(t1) \cup R(t2) \text{ or } R(t1) \quad \cdots\cdots (1)$$

$$R(t1) = S(t1) \cap C(t1) \cap G(t1), \quad R(t2) = S(t2) \cap C(t2) \cap G(t2)$$

여기서 R은 멀티모달의 입력결과, t1은 사용자의 입력이 끝난 뒤 2초경과 후의 시간, t2는 R(t1)의 내용 중 누락된 정보가 있을 경우 재입력이 이루어진 후 2초 경과 후의 시간, S는 음성입력 데이터, C는 문자입력 데이터, G는 제스처입력 데이터, ∪는 단순 통합(merge), ∩는 다음과 같은 연산자이다.

$A \cap B = B$ ($A \neq \phi, B \neq \phi$), $A \cap B = B$ ($A=\phi, B \neq \phi$),

$A \cap B = A$ ($A \neq \phi, B = \phi$), $A \cap B = (A = \phi, B = \phi)$ ..... (2)

여기서 는 데이터가 null임을 나타낸다.

필터의 구조는 프리디킷과 아규먼트로 구성된다. 즉 프리디킷(아규먼트 #1, 아규먼트 #2, . . , 아규먼트 #N)의 형식을 갖는다. 다음은 프로토타입으로 구현한 지도 시스템에서 사용한 필터의 예를 나타낸다.

- 대상정의 : site(Map, Id, Label, X, Y, Icon, AttrList)
- 미디어 데이터 정의 : media_data(Map, Icon-Type, Id, MediaList)
- 미디어 타입 정의 : media_type(Map, Media-Type, ViewType)
- 지도상의 아이콘그룹 정의 : map_icon_group (Map, Group, Icon, SingularType, Msg, Has-Constraints, Btn)

- 정보의 출력 : info_display(Map, Group, InfoToDisplay)
- 대상의 정의 : object_def(Map, Group, ObjectId)

4. 휴대형 멀티모달 인터페이스 구조

휴대성 기기의 이동성과 멀티모달 지원을 위해 요구되는 특성은 크게 6가지로 분류할 수 있다. 첫째로 이동성 측면에서 클라이언트는 휴대성을 고려하여 하드웨어나 소프트웨어의 크기가 경량이어야 한다. 또한 멀티모달을 지원하기 위해서는 많은 컴퓨팅 파워가 필요함으로 이를 지원할 수 있어야 한다. 둘째로 확장성 측면에서 새로운 서비스 지원 기능을 쉽게 확장할 수 있고, 필요한 부분만을 확장할 수 있어야 한다. 또한 시스템의 사용 중에도 새로운 동적인 기능확장이 가능해야 한다. 셋째로 재사용성 측면에서 기존에 이미 만들어진 소프트웨어를 가능하면 그대로 사용한다. 또는 기존에 만들어진 소프트웨어의 기능을 사용할 수 있어야 한다. 넷째로 유연성 측면에서 지원하는 서비스 중 일부가 사용할 수 없게 되더라도 나머지를 통해 사용할 수 있어야 한다. 또한 수행환경이나 사용환경에 영향을 가능하면 받지 않아야 한다. 다섯째로 편리성 측면에서 사용자는 원하는 작업을 음성과 펜을 사용하여 자연어로 지시하면, 사용자의 편의에 따라 임의로 선택된 모달리티를 지원해야 한다. 여섯째로 분산성 측면에서 사용자의 작업지시를 위탁 받아서 분산적으로 실행해 주며, 사용자가 원하는 기능을 만들어 기존에 이미 만들어진 다른 기능들과 상호 협력하며 사용자의 요구를 수행해 준다. 또한 분산되어 있는 자료와 컴퓨팅 파워를 사용할 수 있어야 한다.

본 연구에서는 이러한 특성을 수용할 수 있는 멀티 에이전트 구조인 OAA(Open Agent Architecture) 기반구조[6]를 기반으로 하고 있다.

OAA는 하나의 조정 에이전트와 둘 이상의 응용 에이전트, 그리고 에이전트 통신언어 (ICL: Interagent Communication Language)로 이루어진 분산환경과 플러그 앤 플레이를 바탕으로 한 동적인 구조를 갖는다. 에이전트는 그 역할에 따라 분류가 되는데 조정 에이전트는 응용 에이전트들 간의 교류를 위한 장소 및 조정기능을 제공한다. 이와 대조적으로 응용 에이전트는 특정한 도메인의 작업을 전문적으로 처리하는 역할을 맡는다. OAA에서는 에이전트의 효율적인 제어를 위하여 조정 에이전트를 통해서만 응용 에이전트들이 다른 에이전트들과 교류할 수 있다. 응용 에이전트들은 자신들의 능력을 조정 에이전트에게 미리 선언해 놓음으로써, 자신이 잘 처리할 수 있는 작업의 수행을 요구 받을 수 있도록 한다.

ICL은 에이전트들 상호간에 정보교류를 위한 메시지 형태의 통신언어로 content 층, protocol 층, wrapper 층으로 구성된 3개의 층으로 이루어진다. Content 층은 에이전트 간에 교류 되는 메시지 내용을 나타내고, Protocol 층은 에이전트들 간의 약속된 규칙을 의미한다. Wrapper 층은 에이전트가 그 내용을 보고 그 메시지가 ICL로 해석되어 질 수 있는 메시지인지 아닌지 판단하며, 어느 에이전트가 보낸 것인가를 알아낼 수 있도록 한다. 에이전트들은 이 ICL을 이용하여 다른 에이전트에게 작업을 요구하거나 수행결과를 돌려주며 자신이 제공할 수 있는 서비스를 ICL로 표현한다.

OAA에서는 클라이언트 개념에 해당하는 응용 에이전트들이 조정 에이전트가 갖고있는 블랙보드에 해결할 작업을 기록하는 식으로 진행된다.

조정 에이전트나 응용 에이전트들은 모두 독립적인 하나의 프로세스이다. 에이전트들은 미리 정의된 통신 프로토콜에 의해 메시지를 주고 받을 수 있다. 따라서 각 에이전트는 네트워크가 연결된 컴퓨터라면 어디에나 놓여질 수 있고, 네트워크를 매체로 원거리에 있는 에이전트와 상호 협력하면서 작업을 처리할 수 있다.

5. 지도 시스템의 구현

본 연구에서 제안한 방법의 유용성을 검증하기 위해 OAA(Open Agent Architecture) 기반구조를 바탕으로 펜과 음성을 조합한 프로토타입인 지도 시스템을 개발했다. 현재 구현된 시스템은 서버로 Sun Sparc. 10을 사용했으며 에이전트는 프로로그와 C로 구현되었다. 클라이언트는 노트북 PC 상에 델파이(Delphi)로 작성되었다.

지도의 표면은 사용자에 의해 문자와 제스처의 입력이 가능하다. 또한 자연어 문장 수준의 입력이 가능하다. 사용자는 멀티모달 인터페이스를 통해 인터넷이나 로컬네트워크에 연결되어 있는 데이터베이스에서 사용자가 원하는 대상(집, 호텔, 관광지,)을 찾거나 필요한 정보를 필터링 해서 얻을 수 있다.

멀티모달 지원기능은 지도제어, 특정 대상선택, 거리계산, 정보검색 등을 할 때 직접조작법과 자연어처리를 복합 시킨 명령의 입력이다. (그림 2)에 지도시스템의 구조를 나타낸다. 지도 시스템을 구성하는 에이전트들과 이들의 역할은 다음과 같다.

- 조정 에이전트(Facilitator) : 에이전트 간의 통신을 중개하고 각 에이전트에 대한 작업 할당 및 관리를 한다. 따라서 모든 에이전트 조정 에이전트에 연결 함으로서 다른 에이전트의 기능을 제공 받을 수 있다.
- 기능 에이전트(Functional Agent) : 특정 영역에 대한 작업을 수행하는 에이전트로 웹 에이전트, 전화 에이전트, 데이터베이스 에이전트, 지도에이전트, 전자우편 에이전트 등 아주 다양하다. 기능 에이전트는 사용자의 요청에 의해 개발자가 새롭게 만들어 자유롭게 사용할 수 있다.
- 사용자 인터페이스 에이전트(User Interface Agent) : 사용자로부터 펜, 마이크, 키보드, 마우스 등을 사용한 멀티모달 입력을 받거나 에이전트로부터 정보를 사용자에게 보여준다. 사용자는 휴대형 단말 상에 인터페이스 에이전트를 탑재하여 다니면서 전화선이나 네트워크가 있는 곳이면 어디서든지 에이전트 시스템과 연결하여 서비스를 받을 수 있다.

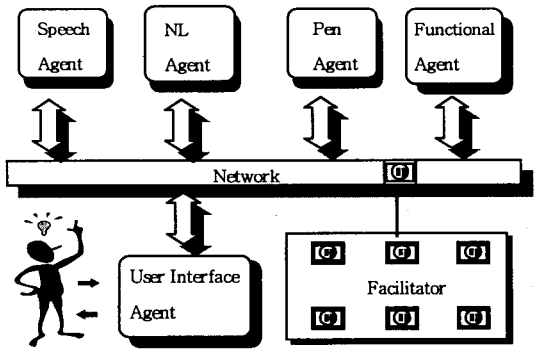


그림 2. 지도 시스템의 구조

- 자연어 처리 에이전트(NL Agent) : 사용자가 입력한 자연어나 음성인식 에이전트로부터 도착한 자연어 텍스트를 처리하여 에이전트가 이해할 수 있는 형태로 변환한다. 자연어 처리 엔진은 DCG-NL과 Gemini[7] 시스템을 사용한다.
- 음성인식 에이전트(Speech Agent) : 사용자가 입력한 음성을 인식하고 그 결과를 텍스트 형

태로 돌려준다. 음성인식 엔진은 SRI에서 개발한 Corona[8] 시스템을 사용한다.

ㅇ 펜 에이전트(Pen Agent) : 사용자가 펜을 이용하여 입력한 데이터를 인식하고 그 결과를 텍스트 형태로 돌려준다. 문자인식 엔진은 CIC에서 개발한 handwriting recognizer와 MS PenWindows API를 사용하고, 제스처 엔진은 TAPAGE[2]에서 사용한 gesture recognizer를 사용한다. 지도 응용 시스템에서 지원하는 제스처는 <표 2>에 나타낸 것과 같이 지우기, 선택, 이동, 거리, 확대 등 5가지이다.

표 2. 지도 응용 시스템에서 지원하는 제스처

	지우기	선택	이동	거리	확대
제스처					

단일 모달리티를 이용한 경우와 멀티 모달리티를 이용한 작업지시 방법을 "특정 대상간의 거리계산"이라는 것을 예로 들어 나타내면 다음과 같다.

- 제스처 : 특정대상 간의 사이를 선으로 연결한다. 즉 계산하고자 하는 거리 만큼 선을 긋는다.
- 음성 : 음성으로 "What is the distance from City hall to Hilton hotel?" 이라 입력한다.
- 문자(펜) : 펜으로 "Distance from City hall to Hilton hotel?" 이라 입력한다.
- 음성과 제스처 : 음성을 통해 "What is the distance from here to this hotel?" 이라 입력하고 동시에 펜 제스처를 이용하여 특정대상을 선택한다.
- 펜과 제스처 : 펜 제스처를 이용하여 특정대상을 선택한 후, 펜을 통해 "What is the distance?" 이라 입력한다.

(그림3)은 지도 시스템에서 음성과 제스처를 이용하여 특정위치를 중심으로 반경 1마일 내에 있는 호텔들을 찾으라는 요구에 대해 출력한 결과를 나타낸다. 이때 사용자는 음성으로 "find hotels within 1 mile from here"라 입력하고, 동시에 특정위치를 포인팅 하거나 원을 그린다.

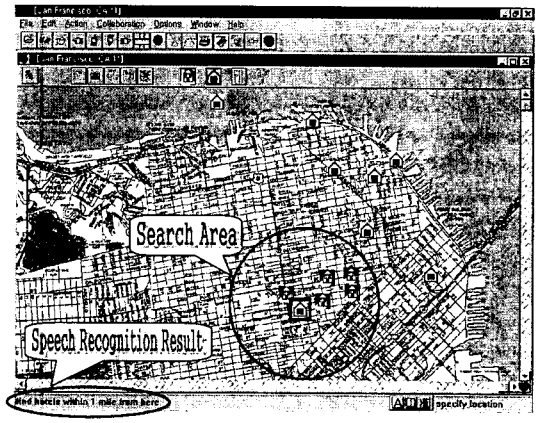


그림 3. 음성과 제스처 명령에 의한 수행결과 화면

이 지도 시스템에서 제공하는 명령어 집합을 작업종류에 따라 분류하면 거리계산, 위치검색, 정보 필터링, 정보검색, 지도 정보갱신 등 5가지로 분류할 수 있다. 또한 제공하는 서비스는 지도정보, 호텔정보, 식당정보, 관광지 정보, 거리정보 등이며 이러한 정보는 인터네트에 구축된 정보를 직접 액세스 하여 제공한다.

6. 평가

본 연구에서 제안한 멀티모달리티 시스템의 평가를 구조적인 측면과 멀티모달리티 측면으로 살펴보면 다음과 같다.

구조적인 측면:

- 확장성 : 새로운 기능을 추가할 때 다른 에이

전트의 수정 없이 독립적으로 만들어 조정자(facilitator)에 연결하여 사용할 수 있는 plug and play기능을 지원한다.

- 모듈성 : 인터페이스 에이전트만을 PDA나 휴대형 단말에 탑재 시켜 사용할 수 있어 계산량이 많거나 특정 서버를 사용해야 하는 응용에 쉽게 적용할 수 있다.
- 재사용성 : 다양한 플랫폼에서 개발된 기존의 응용프로그램을 쉽게 에이전트화 하여 사용할 수 있다.
- 협력성 : 같은 에이전트 사회에 참여한 다른 에이전트의 기능을 쉽게 사용할 수 있다.
- 분산성 : 네트워크에 연결된 다양한 플랫폼을 지원함으로 시스템의 이형성과 특정 시스템의 부하를 경감 시킬 수 있다.

멀티모달리티 측면 :

- 음성, 문자, 제스처에 대한 개별 모드는 물론 이들을 서로 조합한 멀티모달을 지원 함으로써 사용자의 입력 방법에 대한 선택의 폭을 넓게 한다.
- 단일 모달을 상호 보완적으로 적절히 조합 함으로써 개별 인식 기술의 한계를 극복할 수 있고, 사용자의 편리성을 제공한다.
- 상황에 따라 가장 적절한 모달을 사용자가 선택할 수 있다.
- 멀티모달을 통해 입력되는 데이터의 중복성을 이용하여 좀더 정확한 인식이 가능하다.
- 사용방법이나 조작이 쉽고, 단말의 크기를 작게 할 수 있다.

7. 결론 및 연구방향

본 논문에서는 입력 수단으로 친숙성과 사용성이 좋은 펜과 음성을 이용하여 명령어를 처리할 수 있는 멀티입력 처리방법을 제안했고, 또한 휴대형 단말기의 이동성 지원을 위해 휴대형 단말기로부터 인식엔진이나 서비스 제공 모듈이 동적으로 구성 가능한 사용자 인터페이스 구조를 제안했다. 이 방법은 직접조작 방법과 자연어 처리 기술을 사용하여 음성-제스처, 문자-음성, 문자-제스처 등의 모달을 입력순서에 관계없이 사용할 수 있는 협동적 멀티모달리티(synergistic multi-modality)를 지원한다. 따라서 멀티미디어 시스템의 사용자 환경에서 사용자와 컴퓨터 사이에 좀더 자연스럽고 오류가 적은 사용자 인터페이스 환경 지원이 가능하고, 휴대형 기기로부터 인식엔진이 분리되어 있기 때문에 휴대형 기기의 경량화에 관계없이 컴퓨팅 파워를 향상 시킬 수 있고, 추후 인식엔진의 업그레이드 및 유지 보수가 용이하다. 특히 사용자와 컴퓨터 간의 통신에 있어서 사람에게 친숙한 환경제공에 필수적인 음성, 문자, 제스처 등 각종 인식 관련 기술에 대한 한계를 상당 부분 극복할 수 있다.

앞으로의 연구방향은 멀티모달 입력의 통합 시에 사용자의 행동이나 과거의 경험을 바탕으로 사용자의 특성에 맞게 학습할 수 있는 학습모듈의 추가하고 사용자 인터페이스를 자바(Java)로 구현함으로써 분산환경 지원성을 높인다. 또한 현재 구현한 지도 시스템은 멀티모달을 사용하는 하나의 응용에 불과하다 따라서 실세계의 여러 응용에 적용하여 제안한 멀티모달 시스템의 유용성을 입증할 것이다.

참 고 문 헌

[1] Tadao Kobayasi, "Mobile Computing with Multi-media," 일본전자정보통신학회, MVE97-13, pp. 29-34, April 1997.

[2] Faure, C. and Julia, L. "An Agent-Based Architecture for a Multi-modal Interface," In

Proc. AAAI'94-IM4S(Stanford), pp.82-86, 1994.

[3] Kazuhiro Miyahara, Toshio Okamoto, "Repr-esentation of Personal Information Interests and its Acquisition Method for Collaborative Filtering," 일본전자정보통신학회, AI96-32, pp. 15-22, January 1997.

[4] Minh Tue Vo and Cindy Wood, "Building an Application Framework for Speech and Pen Input Integration in Multi-modal Learning Interfaces," http://www.cs.cmu.edu/afs/cs/user/tue/WWW/papers/icassp96/paper.html, pp. 1-13, May 1996.

[5] Vo, M.T., Houghton, R., Yang, J., Bub, U., Meier, U., Waibel, A., and Duchnowski, P., "Multi-modal Learning Interfaces, " Proc. ARPA SLT Workshop 95 (Austin, Texas), 1995.

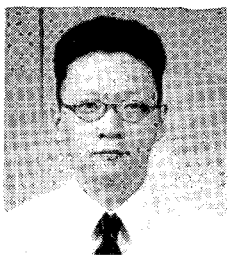
[6] Cohen, A. Cheyer, M. Wang, and S.C. Baeg, "An Open Agent Architecture," In O. Etzioni, editor, *Proceedings of the AAAI Spring Symposium Series on Software Agents*, pp.1-8, Stanford California, March 1994. American Association for Artificial Intelligence.

[7] Dowding, J.M. Gawron, D. Appelt, J. Bear, L. Cherny, R.Moore and D. Moran, "Gemini: A natural language system for spoken-language understanding," In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, pp. 54-61, Columbus, Ohio, June 1993.

[8] Michael Cojen, Hy Murveit, Jared Bernstein, Patti Price, and Mitchel Weintraub, "The DECIPHER speech recognition system," In *IEEE ICASSP*, pp. 77-80, 1990.

이 광 로

- 1986년 일본 Fukuoka 공업대학교 전자기계과 학사
- 1988년 일본 Ritsumeikan 대학교 전기공학과 석사
- 1988년~1998년 한국전자통신연구원 인공지능연구실 선임연구원
- 1994년~1995년 미국 SRI International(International Fellow)
- 1998년~현재 하이콤 리서치 부장
- 관심분야 : 에이전트 시스템, HCI, 패턴인식, 인공지능



정 경 석

- 1998년 국립 경상대학교 컴퓨터과학과 학사
- 1997년~현재 드림인테크 대표
- 1998년~현재 하이콤 리서치 연구원
- 관심분야 : 에이전트 시스템, 멀티미디어 영상처리

김 인 홍

- 1975년~1983년 경북대학교 전자공학과 학사
- 1983년~1985년 한국 과학 기술원 전산학 석사
- 1983~1988년 한국전자통신연구원 연구원
- 1988~1990년 삼성종합기술원 선임연구원
- 1990~1996년 Maspar Computer Co. 선임연구원
- 1996~1997년 하이콤 테크날러지 이사
- 1997년~현재 하이콤 리서치 대표
- 관심분야 : Realtime System, 멀티미디어 통신, 시스템 소프트웨어