

M1-0104-00-0165

국가지정연구실 사업

차세대 PC 환경을 위한 Multi-Modal 입력 시스템 개발

Development of Multi-Modal Input System for Post PC

한국과학기술원 전산학전공
인공지능패턴인식 연구실

과 학 기 술 부

제 출 문

과학기술부 장관 귀하

본 보고서를 “차세대 PC 환경을 위한 Multi-Modal 입력 시스템 개발”과제의 보고서로 제출합니다.

2003 . 9 . 10

주관연구기관명 : 한국과학기술원

주관연구책임자 : 김진형

연 구 원 : 김민수

” : 김자환

” : 조성정

” : 류성호

” : 최현일

” : 이택헌

” : 권영희

” : 장만대

” : 조규태

” : 김연경

” : 김정훈

” : 최승락

” : 이기용

” : 조대현

” : 서승원

” : 현동준

” : 김은경

” : 최준기

보고서 초록

과제관리번호	M1-0104-00-0165	해당단계 연구기간		단계 구분	(해당단계) / (총단계)
연구사업명	중 사업명	예시) 중점국가연구개발사업, 선도기술개발사업 등			
	세부사업명	예시) 기계설비요소기술개발사업, 고온초전도기술개발사업 등			
연구과제명	중 과제명	중과제가 있을 경우에는 기재 (단위과제일 경우에는 아래 기재)			
	세부(단위)과제명	차세대 PC 환경을 위한 Multi-Modal 입력 시스템 개발 (Development of multi-modal input system for post PC)			
연구책임자	김진형	해당단계 참여연구원수	총 : 21 명 내부 : 21 명 외부 : 명	해당단계 연구비	정부: 천원 기업: 천원 계: 천원
연구기관명 및 소속부서명	한국과학기술원 전산학과		참여기업명		
국제공동연구	상대국명 :	상대국연구기관명 :			
위탁연구	연구기관명 :	연구책임자 :			
요약(연구결과를 중심으로 개조식 500자이내)				보고서 면수	
<ul style="list-style-type: none"> - Multi-Modal 통합 입력시스템의 핵심 기술 개발 - 문서, 일반 사진, 펜 기반 인터페이스, 음성을 바탕으로 유기적 통합 입력을 가능하게 함 - 문서 영상 인식 기술 <ul style="list-style-type: none"> - 영상내의 연결 요소의 유사성에 기반한 문서 구조 분석 수행 - 인쇄체 문자 인식(영,숫자: 99.5%, 한글: 97.5%) - 계층적 랜덤 그래프의 확률적 의존 관계에 기반한 필기체 한글 인식(87.5%) - 개방 어휘 환경하의 문맥 정보를 이용하여 어절단위 한글 인식기 성능을 20% 향상 - 일반 사진 영상에서의 문자 영역 추출 및 인식 <ul style="list-style-type: none"> - 정형화된 구조 영상에서의 문자 영역 추출 및 인식(차량 번호판 대상 99.7%) - SVM 와 CAMShift 알고리즘을 이용한 일반 사진 영상에서의 문자 영역 추출(91.3% 탐색률) - 펜 기반 인터페이스를 위한 요소 기술 <ul style="list-style-type: none"> - 신경망 검증을 통한 은닉 마르코프 모델 기반 인식기의 성능 개선(59.9%의 오류 감소) - Bayesian Network 기반 인식 방법론의 개발 <ul style="list-style-type: none"> - 문자 구성 요소간의 상관관계를 계층적 확률 모델로써 표현 - 상용 인식기와 자체 비교 평가 결과 95.7% 로 가장 높은 성능을 보임 - 요소기술 응용 시스템 설계 및 구축 <ul style="list-style-type: none"> - 문서 입력/교정 시스템 개발 - 온라인 문자 인식 기반 문자 입력기 개발 - 키보드 입력 방식 개선에 대한 연구 수행 - 문자 데이터베이스 구축 					
색인어 (각 5개 이상)	한글	문서인식, 필기인식, Pen-based Interface, 입력장치, Tablet PC			
	영어	Document Recognition, Character Recognition, Pen-based Interface, Input Device, Tablet PC			

요 약 문

본 보고서는 Multi-Modal 통합 입력시스템의 핵심 기술 개발 및 성능의 고도화에 대해 서술한다. 이는 문서, 일반 사진, 펜 기반 인터페이스, 음성을 바탕으로 유기적 통합 입력을 가능하게 하는 시스템을 말한다. 첫번째 핵심 기술은 문서 영상 인식 기술이다. 영상내의 연결 요소의 유사성에 기반한 문서 구조 분석, 신경망을 이용한 인쇄체 문자 인식을 수행하였다. 필기체 한글인식을 위해 문자 구성 요소의 확률적 의존 관계를 계층적 랜덤 그래프로 표현하였다. 또한 한국어 문맥 정보를 이용하여 한글 인식기의 성능을 높였다. 두번째는 일반 사진 영상에서의 문자 영역 추출 및 인식이다. 영상에서 문자 영역과 주위 영역의 상관관계를 신경망으로 학습하여 탐색하고 결과에 대해 영상 처리 및 군집화를 통해 최종 문자 영역을 추출한다. 자동차 번호판과 같이 어느정도 정형화된 문자 영역 추출 및 인식은 99.7%, 비정형 문자 영역 추출을 91.3% 를 보였다. 세번째로는 펜 기반 인터페이스를 위한 요소 기술이다. 기존의 은닉 마르코프 모델 기반 인식기의 성능을 개선하기 위해 문자의 구조적 지식을 신경망으로 검증하는 방법을 개발하여 약 60%의 오류를 감소시켰다. 주목할만한 결과로써 문자 구성 요소간의 계층적 상관관계를 Bayesian Network 을 이용하여 모델링 하였다. 이는 결합 확률 분포를 통해 문자의 지역적/전역적 특성을 모델링 하기 때문에 정교한 인식을 수행할 수 있다. 자체 비교 실험 결과 상용인식기보다 높은 95.7%의 성능을 보였다. 마지막으로는 핵심 기술을 바탕으로한 응용 시스템 설계 및 구축이다. 문서 구조 분석 및 문자 인식기를 통한 문서 입력, 펜 기반 인터페이스를 응용한 교정 시스템의 설계 및 구축을 수행하였다. 또한 수기 입력이 가능하도록 온라인 문자 인식 기반 문자 입력기 개발하였다. 기존 키보드 입력 방식을 개선하기 위해 동시입력 및 한글 창제 원리에 기반한 새로운 입력방식에 대한 연구를 수행하였다. 또한 국내외 학자들의 원활한 연구 수행을 위해 다양한 문자 데이터베이스를 구축하여 운영하고 있다.

S U M M A R Y

In this report, core techniques for multi-modal input system are described. The main components of the system are document, scene image, pen-based interface, and speech. All the components are integrated systematically in terms of input method. The first component is document recognition. A structure of document is analyzed by finding geometrical relations of connected components. Then SVM-based OCR is used to recognize printed characters. For the handwritten Hangul, we used hierarchical random graph with probabilistic dependencies. Hangul context information is also used to achieve better performance. Secondly, scene text detection and recognition. By learning correlation of a pixel and its neighborhood with SVM, pixels that belong to text region are successfully detected. To get more accurate text region, image processing and clustering are performed. In case of rigid text regions, such as vehicle license plate, 99.7% of recognition rate was achieved. For non-rigid text regions, detection rate was 91.3%. The third component is technologies for pen-based interface. Since HMM-based on-line character recognition only models local information, incorporation of verification with global information by neural network has been done. About 60% of errors are reduced. We also developed Bayesian network based on-line character recognition system. It models hierarchical relationships between character components with help of probabilistic framework so as to do coarse-to-fine modeling. Since the network has ability to approximate the full joint probability, it is possible to model both local and global features of characters, accurately. In our experiment, the performance was 95.7% which is the highest rate among other system. Final component is applications of core techniques as follows; i) document input system using document structure analysis method and OCR ii) pen-based document correction system iii) handwriting input method iv) a new keyboard layout which supports parallel key press so as to enable fast typing. Additionally, we have been maintaining character/document/image databases for the researchers who are interested in related fields.

C O N T E N T S

Chapter 1. Introduction	9
1.1 Purpose of the research	9
1.2 Necessity of the research	12
1. Technical significance	12
2. Economical significance	13
3. Cultural significance	14
Chapter 2. State of Arts: domestic and foreign countries	15
2.1. State of arts of foreign countries	15
2.2. Domestic state of arts	16
2.3. Prior achievement of this lab.	16
2.4. Weakness of current state	17
2.5. Visions about related field	18
Chapter 3. Contents and results	19
3.1. Research focus	19
3.2. Research results	21
3.2.1. Development of optical document recognition technique	22
3.2.2. Text retrieval and recognition of natural scene images	26
3.2.3. Development of basic technologies for pen interface	27
3.2.4. Design and implementation of applications	30
3.2.5. Construction of text and image database	33
Chapter 4. Achievement rate and contributions	34
4.1. Achievement rate of the research	34
4.2. Major achievements	36
1. Journal papers	36
2. Conference papers	37
4.3. Major successful examples	38
4.4. Contributions	39

Chapter 5. Plans to application of the research results	41
Chapter 6. References	42

목 차

제 1 장 연구개발과제의 개요	9
1절 연구개발의 목적	9
2절 연구개발의 필요성	12
1. 기술적 중요성	12
2. 경제, 산업적 중요성	13
3. 사회, 문화적 중요성	14
제 2 장 국내외 기술개발 현황	15
1절 국외 연구 실적	15
2절 국내 연구 실적	16
3절 본 연구실의 연구실적	16
4절 현 기술상태의 취약성	17
5절 앞으로의 전망	18
제 3 장 연구개발수행 내용 및 결과	19
1절 연구 추진 전략	19
2절 연구 수행 내용 및 결과	21
1. 문서 영상 인식 기술 개발	22
가. 문서 구조 분석	22
나. 인쇄체 문자 인식기 개발	23
다. 필기체 문자 인식기 개발	23
라. 문맥 정보를 이용한 한글 인식기 성능 향상	25
2. 일반 사진 영상에서의 문자 영역 추출 및 인식	26
가. 정형화된 구조 영상에서의 문자 영역 추출 및 인식	26
나. 일반 사진 영상에서의 문자 영역 추출	27
3. 펜 기반 인터페이스를 위한 요소 기술	27
가. 은닉 마르코프 모델 기반 인식기의 성능 개선	28
나. Bayesian Network 기반 인식 방법론의 개발	28
4. 요소 기술 응용 시스템 설계 및 구축	30
가. 문서 교정 기본 시스템 설계 및 구축	30
나. 문서 입력 기본 시스템 개발	31
다. 자동차 번호판 인식 시스템	31

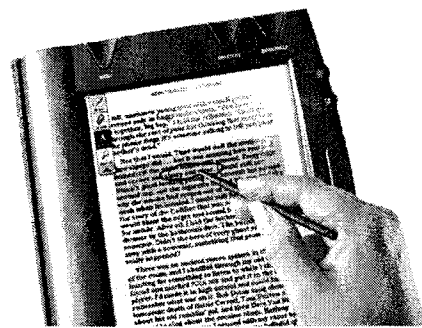
라. 온라인 문자 인식 기반 문자 입력기 개발	31
마. 음성 인식 기술 동향 조사	32
바. 기존 키보드 입력 방식 개선에 대한 연구	32
5. 문자 데이터베이스의 구축	33
제 4 장 목표달성도 및 관련분야에의 기여도	34
1절 연구 개발 목표 달성도	34
2절 연구개발성과	36
1. 논문게재	36
2. 학술발표	37
3절 대표적 성공사례	38
1. 확률적 관계 모델링을 이용한 계층적 한글 문자 인식	38
2. 개방 어휘 환경 하에서의 문자 인식을 위한 문맥 정보 모델 개발	38
3. 인공 신경망과 연속 탐색 알고리즘 기반 차량 번호판 인식	38
4. 베이지안네트워크(Bayesian Network) 기반 온라인 필기 인식 기술	38
5. 문자 영상 DB 구축	39
4절 대외기여도	39
1. 홈페이지 운영 및 활용 현황	39
2. 산·학·연 협력거점 활동현황	39
3. 기술지도 및 성과활용 등	39
4. 첨단기술정보의 제공, 연구기기 공동 활용	40
5. 국내외 연구기관과의 협력	40
6. Workshop, seminar 및 단기강좌를 통한 기술협의, 기술교류 등 실질적 자문 활동 등	40
제 5 장 연구개발결과의 활용계획	41
제 6 장 참고문헌	42

제 1 장 연구개발과제의 개요

1절 연구개발의 목적

본 과제에서는 차세대 개인용 컴퓨터를 위한 Multi-Modal 통합 입력시스템을 개발한다. Multi-Modal 통합 입력시스템이란 스캐너나 카메라에 의한 영상, 펜을 이용한 필기, 마이크로폰을 통한 음성, 그리고 전통적인 Keyboard 입력을 통합하는 입력 시스템이다. 획득한 문서영상과 음성을 인식하여 문자정보화하고, 필요하면 펜을 이용하여 인식내용을 교정하는 시스템을 Tablet PC와 같은 휴대용 펜 컴퓨터 상에서 개발하고자 한다.

패턴인식 기술은 패턴정보를 자동으로 추출하여 응용하는 기술로서 인간의 시청각을 통한 인지능력을 컴퓨터 상에 모방하는 것을 목표로 한다. 문자인식 기술은 문자패턴정보를 인식하는 기술로서 컴퓨터의 자동 입력 수단으로 활용되고 펜 인식 시스템은 키보드를 대체할 입력 방법으로도 사용된다.



[그림 1] Tablet PC에서 펜을 이용한 교정

문자정보의 입력방식은 문자를 쓰는 때와 인식할 때와의 차이에 따라 오프라인 인식과 온라인 인식으로 나눌 수 있다. 오프라인 인식은 종이문서 위에 쓰여 있는 문자정보, 즉 문서의 내용을 인식하는 기술을 일컬으며, 우편봉투 자동분류, 금융권의 전장표 자동화, 전자도서관 구축, 인구센서스 자료처리, 세금신고서의 자동처리에 활용되고 있다. 온라인 인식은 전자펜 등으로 문자를 쓰는 순간 곧바로 인식하여 입력하는 방식을 말하며, PDA, 이동전화 등 휴대형 단말기 및 전자책에서 활용되고 있다.

문서/문자인식에 대한 연구는 특정분야에서 제한적으로 결실을 보기도 했다. 그러나 인식기술의 측면에서 살펴보면, 현재의 발표된 인식시스템은 내재된 기술적 취약성 때문에 성능과 응용범위의 확대에 제약을 받고 있다. 오프라인인식의 경우, 언어의 종류에 따라서 인식율이 현격히 차이가 나며, 또한 문자가 기록된 매체, 조명, 카메라의 성능 등에 따라 인식성능이 영향을 받는다. 온라인 인식기술은 이동 중 필기로 인한 불균일성, 필기 습관을 무시한 필기 방식의 제한, 그리고 이와 같은 것들로부터 기인한 낮은 인식율과 불편함은 앞으로 극복해야 할 기술적 과제로 지적되고 있다.

본 과제에서는 이러한 기술적 취약성을 체계적으로 극복하고 각 요소 기술을 통합하여 카메라로부터 획득한 문서/문자 영상을 인식하고 필요하면 펜을 이용하여 인식내용을 교정하여 활용하는 시스템을 개발한다. 이 과정에서 개발되는 요소 기술은 추후 다양한 응용시스템에도 사용될 수 있도록 한다. 이를 위하여 개발하는 기술은 그림 2와 같다.

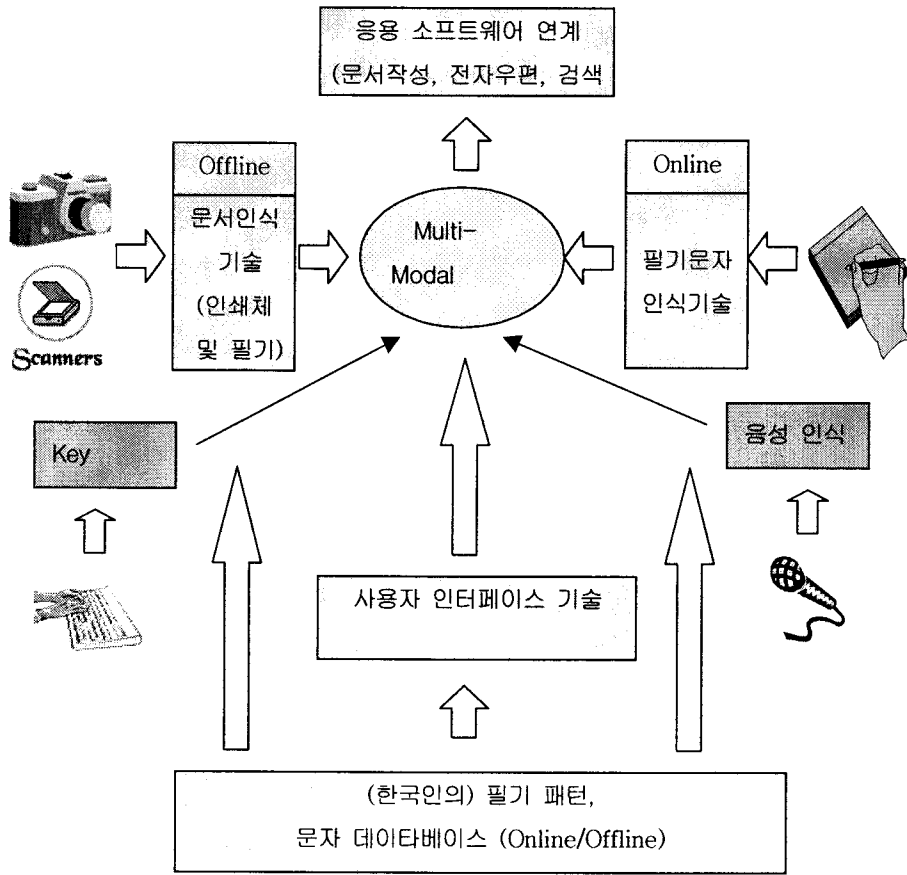
첫째, 본 연구실에서 보유하고 있는 문서 자동인식 핵심기술을 발전시켜서 안정화, 고도화 한다. 본 연구실에서는 문서인식 전반에 걸친 다양한 기본기술을 확보하고 응용분야에 따라서 적합한 요소기술의 심화 및 안정화를 추구하고 있다. 이미 개발되어 있는 시스템도 인식성능을 향상하고 안정성을 제공하면 그 응용범위를 확대할 수 있다. 본 과제에서는 입력 대상 문서별 특성에 따라, 그리고 입력 장치별 특성에 따라 핵심기술의 안정성 확보와 주변기술(전처리, 문맥정보처리 등의 후처리)의 고도화를 위한 연구개발에 초점을 맞춘다.

둘째, 보유한 온라인 필기문자인식 기술을 고도화하고, 도형, 교정부호 등을 인식할 수 있는 기술로 확장한다. Tablet PC로 대표되는 펜 컴퓨터의 보급 및 응용분야는 더욱 확대될 것으로 예상된다. 필수적인 요소기술인 한글, 영문자, 한자의 필기문자 및 도형, 교정부호의 인식 기술을 개발한다.

셋째, 인식기술을 Multi-Modal 입력시스템에 응용하기 위한 사용자 인터페이스에 관한 연구를 수행한다. 단순한 인식기능만으로는 사용자가 만족하지 못한다. 본 과제를 통해 문서인식 기술과 펜 인식 기술을 결합할 때 요구되는 사용자의 요구를 파악하고 편의성 증진을 도모한다.

넷째, 문서/문자인식분야의 연구를 위하여 필기문자 및 문서영상 데이터베이스를 구축한다. 관련 정보 및 지식을 정리하여 필기패턴인식, 특히 동양언어 문서인식의 세계적 연구센터로서의 역할을 자임한다.

본 연구실에서는 지금까지 과학재단지정 ERC인 인공지능연구센터 체제하에서 관련된 연구를 장기간 수행하였다. 이 과정에서 특허, 개발 환경 및 관련 기술을 많이 확보하여, 기 확보된 기술을 기반으로 국내 관련분야 학자들의 모임인 문자인식연구회, 우정자동화사업을 책임지고 있는 ETRI 등 국내 기관들을 선도하여 연구 협력하고자 한다. 또 미국, 캐나다, 영국의 해외 대학 및 연구기관과의 기술교류를 통해 공동 연구 및 인프라를 공유할 계획이다.



[그림 2] Multi-Modal 통합입력 시스템의 구성

2절 연구개발의 필요성

1. 기술적 중요성

자동화와 정보화에 있어서 핵심성을 인정받고 있으면서도 현행 문자인식기술의 취약으로 인한 처리율의 제한이 응용분야에서 문제점으로 지적되고 있다. 예를 들어, 우편물 자동처리 시스템의 경우 필기로 쓰여진 주소를 인식하여 처리하는 비율은 미국이 55%정도이다[USP98]. 그러나 이 정도의 처리율이라도 수작업에 비해 비용 감축효과가 크기 때문에 지속적인 투자가 이루어지고 있다. 기술이 조금씩 향상될 때마다 다양한 분야에서 보다 적극적으로 인식기술을 도입하여 자동화, 정보화를 촉진하게 될 것이다.

인식은 100%를 향한 끝없는 행진이다. 아무리 잘해도 100% 인식이란 불가능하다. 따라서 인식 결과를 쉽게 검증하고 수정할 수 있는 시스템적인 보완이 필수적이다. 지금까지 이점에 소홀하여 인식 시스템들이 시장에서 실패하였다. 본 과제에서는 인식된 결과를 쉽게 수정할 수 있는 방법으로 펜으로 교정하는 시스템을 제안한다. 펜으로 다량의 문자를 입력하는 것은 속도 및 편의성에서 문제가 있다. 그러나 적은 숫자의 입력 및 교정은 Keyboard보다 펜이 편하다. 특히 위치가 중요한 의미를 갖는 경우, 즉 수식을 입력할 때나 Spread Sheet 등에서의 입력은 펜이 자연스럽게 편리하다.

이동 중에 정보를 획득하는 경우가 빈번하게 됨에 따라서 카메라, 스캐너, 펜 등의 휴대용 입력 장치는 지속적으로 작아지고 있으며 무선 전송이 보편화 되고있다. 극 소형 카메라 및 스캐너 장치가 이제는 PDA부품으로 또 만년필의 머리에도 장착된다. Micro Electro-Mechanical System의 발전에 힘입어 만년필형 스캐너 및 가속도를 감지할 수 있는 펜이 개발되고 있다. 손에 들고있는 카메라나 손으로 미는 형태의 스캐너로 영상을 획득하면 손 떨림 등의 영향을 감쇄하여야 하는 특별한 기술이 필요하다. Multi-Modal 통합입력 시스템이 개발되면 즉석에서 인식하고 교정하여 전송하는 편의성을 제공할 것이다.

한글 중심의 국내 문서 및 필기 인식 기술은 영어를 중심으로 개발된 외국 기술과 크게 차별화 된다. 우리나라에서 사용되는 문서에는 한글이외에도 영어, 한자, 숫자 등의 문자가 혼용되는 특징을 가지고 있다. 따라서 이들 각각의 문자에 대한 인식기술이 필요하고 또 이들을 같이 사용하였을 때에도 인식할 수 있는 능력이 요구된다. 한글의 구조는 영문자와 판이하게 달라 외국 기술을 수입하여 사용할 수 없다. 우리는 한글을 인식했던 기술력을 바탕으로 더 나아가 동양권의 다른 문자 인식기술도 개발할 수 있다고 믿는다.

문자인식 기술은 인공지능의 핵심으로서 지능시스템 연구의 Testbed로 활용된다. 글씨를 인식할 수 있는 수준의 지능시스템을 만들 수 있는 기술을 갖추면 유사한 성능의 인공지능시스템 개발도 쉽게 가능하다고 믿어진다. 이 분야에 대한 타 분야의 관심도 높아서 이 분야의 학술 논문지인 IEEE Transactions on Pattern Analysis and Machine Intelligence의 impact factor가 2.085로서 전산학 관련 주요 학술지 평균치인 0.842보다 높다. 그리고 지금도 다수의 학술대회가 세계 곳곳에서 성황리에 개최되고 있다는 사실에서도 이 기술의 중요성을 유추할 수 있다.

2. 경제, 산업적 중요성

연구개발의 경제, 산업적 중요성은 인식 관련 시장이 크고 응용분야가 다양하다는 데 있다. 온라인 필기인식 기술을 바탕으로 한 PDA의 세계 시장이 2003년 2100만 대(약 73억달러)에 달할 것으로 예측하고 있다. 또한 2000년 가을 COMDEX에서 Bill Gates가 들고 나와서 선풍적인 인기를 모았던 Tablet PC류는 2004년까지 890만대가 보급될 것이라고 예측했다[DT21].

요소 기술의 하나인 문서 인식 기술은 자동화 및 정보화 사업에서의 중요한 기반 기술이다. 1999년부터 금융결제원이 주관하여 시행되고 있는 금융기관 수납장표 정보화 사업은 수표, 지로, 각종 전표 등의 문자정보를 인식하여 자동 처리하는 시스템을 각 금융기관에 설치하려는 사업으로 이 수납장표 자동화 사업의 규모는 약 4,700억원에 이를 것으로 추정하고 있다[KB98]. 우편봉투 주소인식에 의한 우편물 자동분류에서 문자 및 필기인식기술이 핵심역할을 한다. 우정본부는 2003년까지 22개의 우편집중국의 우편물 자동분류시스템을 설치에 2003년까지 8,350억원을 투자할 예정이다.

문자인식 기술의 발전에 따라 펜 스캐너 시장이 빠른 속도로 성장하고 있다. Desktop스캐너를 중심이던 시장이 소형 스캐너 중심으로 변하고 있다. 한 예로 1999년부터 2004년까지 펜 스캐너 시장이 연109% 씩 성장하여 2004년에는 연 1,340,000 대가 팔릴 것으로 예측하고 있다[GAT99].

향후 10여년간 대규모로 투자되는 우편물 자동분류, 디지털라이브러리 구축 등과 기업에서 추진중인 각종 자동화, 정보화 사업에 국내 인식기술의 참여는 관련기술을 성숙시키는 계기가 될 것이다. 또 분야마다 독특한 특성이 있기 때문에 다양한 틈새 시장도 창출될 것으로 예상된다. 특히 한글의 경험은 쉽게 한자에 대한 인식 기술로 이어지고 중국 및 일본 등 동양권 시장 진출 가능성이 있다. 이미 국내에 인식기술을 기반으로 한 기술 창업회사가 적지 않으나 기술발전에 따라 더욱 많은 기술 창업회사가 탄생하고 성공사례를 접하게 될 것이다.

3. 사회, 문화적 중요성

필기된 문서가 컴퓨터에 입력된다거나 펜을 사용한 입력은 어린이나 컴퓨터에 익숙하지 않은 사람들에게도 편리하고 친숙한 입력 수단을 제공한다. 특히 장애인을 위한 시스템을 통하여 사회복지 증진에도 기여한다. 예를 들어, 펜을 이용한 입력 장치는 말을 못하는 환자를 위한 도구로 이용될 수 있으며, 영상을 음성으로 변환해 주는 Image-to-Text-to-Speech 시스템은 맹인이 문서의 내용을 파악하는데 도움을 준다.

언어문화적 관점에서 볼 때 우리 민족 고유의 문자 보유와 한글의 과학성에 대한 자긍심이 컴퓨터 시대에서도 이어지기 위해서는 한글 인식은 반드시 국내 기술로 완성되어야 한다. 더구나 필기 습관 등 필기 문화의 차이로 인해 외국 기술의 직접 도입에 의한 사용은 한계가 있어 국내 개발 필요에 대한 당위성을 갖고 있다.

제 2 장 국내외 기술개발 현황

1절 국외 연구 실적

문서/문자인식에 대한 연구는 미국, 일본, 독일 등 대부분 선진국에서 오랫동안 활발하게 진행되어 오고 있다. 미국에서는 대부분의 대학에서 다양한 연구를 수행해 왔다. 특히 State University of New York at Buffalo등에서는 U.S. Postal Service의 지원을 받아 영상 처리, 숫자 인식, 영어단어 인식 및 주소사전 참조 기법 등에 대해 집중적으로 연구해 온 결과 우편봉투의 필기주소 인식시스템 기술을 개발하여 1996년 관련 업계에 이관하였다. 이런 노력의 결실로 현재 미국 총 우편물의 55%를 자동처리하고 있으며 이를 60%이상으로 올리기 위한 연구가 계속되고 있다[USP98].

캐나다에서는 Concordia대학의 CENPARMI가 연구거점 역할을 수행해 오고 있다. 영어, 숫자, 한자 인식을 포함하여 형식문서, 수표인식에 이르기까지 다양한 연구 결과로 정부 및 업계를 지원하고 있다. 뿐만 아니라 세계 여러 나라의 문자인식을 연구하는 학자, 기술자에게 방문연구를 지원하여 인적, 기술적 교류도 활발히 하고 있다.

일본은 1960년대 후반기부터 정부의 적극적 지원하에 우편물 자동분류기 개발을 시작하면서 오늘날 문서/문자인식 분야에서 세계적으로 우수한 기업을 육성했다. 우정성 산하의 IPTP에서는 1992년과 1993년 두 차례에 걸쳐 인식성능 경진대회를 개최하는 등 자국산 기술의 경쟁력 강화를 위한 노력을 하고있다.

산업계에서는 IBM, Kodak, AT&T Bell Lab., Motorola, HP, Xerox, Siemens, NEC, Hitachi, Toshiba, Rico, Olivetti 등 세계 유수의 기업과 많은 소프트웨어 회사들이 문자인식 분야의 사업에 참여하고 있다. 영문자 OCR시스템은 Caere Corp. 과 Xerox 등에서 상용시스템을 판매하고 있다. 펜 인식 시스템은 Paragraph 등이 두각을 나타내고 있다.

각국의 공공연구소에서는 문서/문자인식 연구를 위한 데이터베이스를 구축하여 민간에게 공급하고 있다. CEDAR, NIST, University of Washington(이상 미국), CENPARMI(캐나다), IPTP, ETL(이상 일본), 4MSL(중국), ITRI(대만) 등의 DB는 연구지원 및 객관적 평가를 위하여 구축하였으며 이들은 현재 세계 여러 곳에서 연구 개발에 사용되고 있다.

2절 국내 연구 실적

국내에 10여 개의 문자인식 연구팀이 활동을 하고 있으며 문서/문자인식을 전문으로 하는 벤처기업도 서너개 존재한다. 이 분야 학자들을 중심으로 해마다 문자인식 워크숍을 개최하는데 대학원 학생 포함하여 약 40명이 모인다. 이들이 국내의 문자인식 연구원의 전부라고 할 수 있다. 일부 대학에서는 필기인식 기술에 대한 연구가 진행되고 있으나 연구 지원이 간헐적이고 그 규모도 크지 않았기 때문에 문자인식을 전문적으로 연구하는 팀이 적다. 열악한 상황에서도 한불 문서인식 공동 워크숍을 상호방문하며 개최하고 국제학술대회를 유치하는 등 분전하고 있다.

업계에서는 OCR 인쇄문자 인식시스템은 몇몇 소프트웨어 업체에 의해 상용화 제품도 출시되고 있다. 최근 PDA시장에서는 온라인 필기한글 인식기술이 탑재된 제품이 등장하기도 했다. 그러나 국내 시장이 협소하고 우리의 문서 인식에 요구되는 기술 수준이 영어권에 비하여 난이도가 높아 성공적인 제품은 보기 힘들다. 기업에서도 지속적으로 연구팀을 유지하고 있는 곳이 없어서 미국, 일본에 비하여 그 수준이 낮은 편이다.

3절 본 연구실의 연구실적

본 연구실은 1985년 이후 17년간 패턴인식/인공지능, 특히 한글인식의 연구를 세계적인 수준으로 진행하고 있다. 특히 필기한글의 인식을 도전해 불만할 패턴인식의 문제로 국제사회에 소개하고 있다.

1987년 Silnoon project를 통해 국내 최초로 인쇄체 한글 문서인식 시스템 시제품을 발표하여 인식기술의 상용화의 견인차 역할을 하였고, 1990년부터 1995년까지는 펜 컴퓨터를 위한 온라인 필기 인식 연구를 중점적으로 수행하였다. Hidden Markov Model을 이용하여 개발한 온라인 필기 인식 기술은 필기에 제약을 두지 않고 인식하므로 흘려 쓴 글씨까지도 인식할 수 있는 성능을 가지고 있으며 차세대 펜컴퓨터의 인식기로서 기업에 기술이전 되었다. 1996년 이후에는 한글 오프라인인식, 형식문서 인식에 대해 연구를 집중하고 있다. 다년간 기술을 축적하여 필기 한글인식, 고속 숫자인식 등 기본적인 기술은 이미 확보했다.

특히 1991년부터 4년간 수행된 ‘노트패드 컴퓨터 개발 컨소시엄’ 과 1995년부터 3년간 수행된 ‘문자인식 기반기술 연구 컨소시엄’ 을 통해 다수의 기업과 대학이 참여한 공동 연구를 수행하여 기본적인 필기한글 인식기술을 확보하는 등 국내 문자인식 분야 연구발전 및 인력양성에 큰 역할을 하였다. 이 두 번의 컨소시엄 사업은 다수의 기업이 공동 투자하고 여러 대학에서 공동연구를 수행함으로써 연구자

원의 효율성을 제고하고 개발된 기술을 공유하는 체제로 산학협력의 모범이 되었다.

김진형 교수는 1997년 International Conference on Document Analysis and Recognition 학회에 초청되어 Plenary session에서 “Problems and Approaches for Oriental Document Analysis” 라는 제목으로 기조연설을 하였고, International Association for Pattern Recognition (IAPR)의 Fellow로 추대되는 등 지금까지의 연구 성과를 국제적으로 인정 받고 있다. 1998년에는 이 분야의 저명한 국제학술대회인 제6회 International Workshop on Frontiers in Handwriting Recognition를 대전에 유치하여 성공적으로 개최하였다.

본 연구실에서는 그 동안 59명의 석사와 24명의 박사를 배출하였으며 SCI급 국제 학술지에 55편의 논문을 게재하였고 국제학술회의에서 63편의 논문을 발표하는 등 활발한 연구활동으로 국내 패턴인식 연구계를 선도하며 세계의 우수 연구그룹과 어깨를 나란히 해오고 있다. 특히 1995년에는 필기인식 기술에 관한 미국 특허와 2편의 국내 특허를 획득함으로써 기술력을 국내외적으로 인정 받았다. 본 연구실 출신들이 핸드소프트, 레코그램 등 다수의 기술 중심 벤처회사를 창업하여 활동하고 있다.

4절 현 기술상태의 취약성

현행 문서/문자 인식기술의 취약성은 다음과 같은 문제에서 기인한다. 오프라인 인식의 경우, 깨끗한 종이에 쓰고 촬영이나 왜곡이 없는 고해상도의 영상일 때에만 안정적인 결과를 얻는 것이 오늘의 기술 수준이다. 문자가 기록된 매체의 상태, 입력장치의 성능 등에 따라 문자 영상의 질은 현저히 달라진다. 입력 장치의 경우 주로 스캐너를 통한 입력이 중심을 이루고 있으며, 아직 카메라 입력에 대한 연구는 취약한 상태다. 스캐너로 입력한 경우도 입력영상의 해상도에 따라 결과가 달라지며 국내에선 한글의 복잡도로 인해 대개 300DPI이상의 고해상도를 요구한다. 문자를 인쇄한 프린터의 유형(dot matrix, line printer, laser printer, 활자 인쇄 등)에 따라서 처리할 수 있는 한계가 달라진다. 레이저 프린터급으로 출력된 문서에 한해서는 안정된 기술이 확보된 상태다. 특히 FAX 문서의 경우 외국에서는 인식시장이 형성되어 있음에도 불구하고 국내에서는 낮은 해상도와 촬영 때문에 거의 연구가 이루어 지지 않고 있다.

필기문자 인식기술은 필기자에 따라 변형이 다양하고 여기에 흘림의 문제도 겹쳐 있어서 연구가 더 필요하다. 온라인 인식기술은 최근 시장에 등장한 PDA 등의 단말기에 탑재되어 관심의 대상이 되었다. 그러나 이동중 작성으로 인한 불균일한 입

력, 사람의 자연스런 필기 습관을 무시한 필기 방식의 제한, 저조한 인식율, 그리고 이와 같은 것들로부터 기인한 불편함 등이 기술적 문제로 지적되고 있다. 필기자에 따른 문자변형에 대한 적응성, 흘림체 처리, 한글/영어/숫자/한자의 혼용 사용시의 인식 문제 등은 필히 해결되어야 할 문제이다.

우리나라에서는 한글/영어/숫자/한자 등이 함께 사용되고 있다. 다중언어의 사용은 우리나라 언어문화의 한 특징이다. 문자인식 기술측면에서 보면, 여러 문자가 혼용된 경우의 인식은 하나의 문자만 사용된 경우에 비해 기술적 난이도가 매우 높다. 다중문자 인식에 대한 연구는 인식기술이 받아들여지기 위하여 많은 연구가 요구되고 있다. 필기한글은 기술적 난이도가 높아 이에 대한 연구를 수행하고 있는 국내 기관이 극히 적어 저변 확대와 동시에 공동연구 등 효과적인 연구 추진체계에 대한 검토가 필요하다.

또한 국내의 문서/문자인식 연구는 핵심 인식기술에 치중되어 진행되어옴으로써 주변기술이라고 할 수 있는 전처리 및 후처리에 대한 연구결과가 미약하다. 현실의 문제를 해결하기 위해서는 이 분야에 대한 깊이 있는 연구가 필요하다. 예로, 우리나라 문서에서 한글이름, 한자이름, 영문이름을 함께 기록하는 것을 흔히 접할 수 있다. 세 가지 이름을 각각 인식하는 것보다 상호 참조하여 인식하면 더 좋은 인식 결과를 얻을 수 있다. 비슷한 예로 우편번호와 주소, 이름과 주민등록번호 등이 있다. 이와 같이 주변기술에 대한 연구가 미약한 것도 현기술의 취약성의 한 원인이 되고 있다.

5절 앞으로의 전망

문서/문자 인식기술에 대한 수요는 다음의 두 가지 이유에서 앞으로 크게 증가할 것으로 전망된다. 첫째는 인건비 상승으로 인하여 자동화의 요구가 더욱 강해질 것이며 문서/문자 인식은 자동화, 정보화의 핵심이기 때문이고 둘째는 관련 시장의 규모와 새로운 응용분야가 급속히 확대되고 있기 때문이다. 최근에 급속하게 성장하는 PDA시장에서 문자인식 기술의 응용을 확인할 수 있다. Tablet PC, Web Pad 등의 이름으로 불리우는 펜 컴퓨터는 Post PC의 총아로 부상되고 있다.

인식기술은 인간의 인지능력을 모방하는 기술이다. 문자는 인간의 정보전달의 주요 수단 가운데 하나로써 인간과 인간 사이의 정보전달 방식으로 문자가 사용되는 한, 인간과 기계사이의 정보 전달수단으로써 이에 관한 연구는 필연적으로 지속될 것이다.

제 3 장 연구개발수행 내용 및 결과

1절 연구 추진 전략

1단계 연구 기간에는 Multi-Modal 입력 시스템의 기능에 대해 정의하고, 각 입력 시스템을 구현하기 위한 요소 기술들을 개발하며, 그 성능을 고도화 및 안정화 시킴으로써 상용화 가능성을 확보하는 것을 주안점으로 삼았다.

현재의 GUI 패러다임은 키보드와 마우스를 주 입력수단으로 하는 데스크탑 환경에 최적화되어있어, PDA나 스마트폰 등과 같이 이동성이 중시되는 차세대 컴퓨팅 환경에서는 기기 자체의 크기 제약 등으로 인한 다양한 문제점이 발생한다. 따라서 본 과제에서는 이같은 문제점들을 해결할 수 있도록 펜이나 영상, 음성 등 다양한 매체를 통하여 입력을 수행할 수 있는 Multi-Modal 입력 시스템을 설계, 구현하고자 한다.

본 연구에서 상정하는 Multi-Modal 입력 시스템의 대략적인 개요는 다음과 같다.

시스템의 기본 입력 기능은 펜으로 이루어진다. 그래픽 사용자 인터페이스와 액정 태블릿 기술이 보편화됨에 따라 마우스로만 가능했던 시스템 제어 등과 같은 많은 기능을 펜으로 대체할 수 있다. 또한, 펜을 이용하여 종이에 쓰듯이 텍스트를 입력할 수 있고, 기존의 문서 및 여러 입력 수단을 통해 입력받은 내용에 대한 교정이 가능하다.

시스템에는 휴대용 스캐너가 장착되어 대용량 문서의 빠른 입력을 가능하게 한다. 시스템은 입력받은 문서의 구조 및 형식을 분석하고, 문자 인식을 통해 문서의 텍스트 정보를 얻는다. 또한, 시스템에 장착되어진 카메라를 이용하여 일반 영상을 촬영하고, 입력된 영상으로부터 건물외 간판이나 자동차 번호판과 같은 텍스트 정보를 검출한다.

버스나 기차와 같이 펜의 사용이 불편한 환경을 위하여 시스템에 장착된 마이크를 이용하여 음성을 입력받는다. 입력된 음성을 인식하여 시스템을 제어하거나 텍스트 입력이 가능하다. 또한, 기존 키보드의 제약을 여러 다양한 형태의 키보드 기술을 활용하여 개선한다.

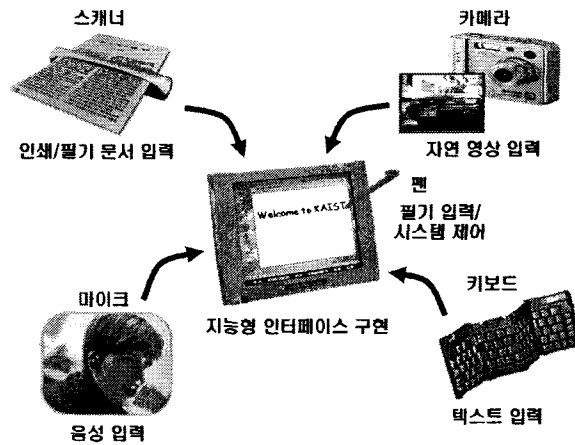


그림 1 Multi-Modal 입력 시스템의 구성 요소

이와 같이 정의된 Multi-Modal 입력 시스템을 위해, 본 연구에서는 각 입력 매체에 필요한 요소 기술의 개발을 목표로 하였다. 그 중, 1단계로서는 입력 비중이 높은 문서 입력, 영상 입력, 펜 입력 기술 개발 및 키보드 입력 방식의 개선에 중점을 두어 연구를 진행하였다.

각 요소 기술 개발을 위한 추진 전략은 서로 유사하여, 다음과 같은 추진 절차를 거쳤다. 우선, 연구실이 보유하고 있는 자체 기술과 국내외 연구기관과의 협력 및 정보 수집을 통해 효과적인 성능 개선 방법을 기획하였다. 성능 개선에 필수 요소인 충분한 양의 문서 및 필기 데이터베이스를 수집, 확보하여 안정성 있는 통계적 자료를 추출함으로써 기획한 성능 개선 방법의 타당성을 객관적으로 증명하였다. 또한, 개선된 연구 결과를 실용화 및 라이브러리화하여 연구 성과를 가시화하고, 향후의 연구를 위한 인프라를 구축하며, 2단계 연구에 진행될 시스템 통합에 대비하였다. 본 연구를 통해 개발된 기술은 외부의 기술과의 성능 비교 또는 타 연구기관으로의 기술 전수를 통해 그 성능을 평가하고, 여러 가지 문제점 및 요구 사항을 피드백 받아 다시 기술 향상을 위한 연구 자료로 활용하였다.

2절 연구 수행 내용 및 결과

본 연구에서는 Multi-Modal 입력 시스템의 개발에 필요한 기반 기술의 성능 개선을 위해 각 입력 요소 기술별로 연구를 추진하였다. 또한, Multi-Modal 입력 시스템의 기본 요소로서 문서 교정 기본 시스템을 설계 및 구현하였으며, 기반 기술 연구를 뒷받침할 수 있도록 문서 및 필기 데이터베이스를 구축하였다. 각 연구별 세부 추진 내용은 다음과 같다.

첫째, 문서 영상 인식 관련 기술에서는 문서의 구조 분석 기술, 인쇄체/필기체 문자 인식 기술, 한국어 문맥정보 처리 기술 등을 개발하였다. 특히, 한글, 한자, 영어, 숫자 등 다양한 문자들에 대하여 인쇄체/필기체 인식기를 구현하여 문서 인식 시스템의 구현을 위한 기반을 마련하였다. 이 중 필기체 한글 인식기는 현재 한글 11,172자 모두를 인식할 수 있는 유일한 인식기이다. 또한, 개방 어휘 환경 하에서의 문자 인식을 위한 한국어 모델을 개발, 인식에 적용하여 높은 성능 향상을 얻었다.

둘째, 자연 영상을 대상으로 하는 인식 관련 기술에서는 자동차 번호판, 간판, 표지판 등이 촬영된 영상을 대상으로 문자 정보가 존재하는 영역을 판별하고, 해당 영역으로부터 문자 정보를 인식하는 기술을 개발하였다.

셋째, 펜 기반 인터페이스를 위한 요소 기술 개발로서 은닉마르코프모델 및 베이지안네트웍 등 다양한 방법론을 이용하여 필기체 문자 인식기 및 제스처 인식기를 구현하였다. 베이지안네트웍 인식기는 기존의 은닉마르코프 모델 기반 인식기의 구조적 한계를 해결할 수 있는 차세대 인식 방법론으로써 세계 연구진의 주목을 받았다. 독자적인 비교 실험 결과, 본 연구실에서 제안한 베이지안네트웍 인식기는 현재 시장에 출시되어있는 상용 인식기보다 우수한 성능을 보임으로써 그 가능성을 입증하였다.

넷째, 각 요소 기술의 응용 시스템으로서 문서 교정 기본 시스템, 문서 입력 기본 시스템, 자동차 번호판 인식 시스템, 온라인 문자 인식 기반 문자 입력기 등을 설계 및 구현하였으며 음성 인식 기술 동향에 대한 조사 및 키보드 입력 방식 개선에 대한 연구를 병행하였다.

다섯째, 문자·영상 및 펜 기반 인식에 관련된 다양한 데이터베이스를 국내외로부터 수집하여 연구 개발을 위한 인프라를 구축하여 인터넷을 통해 이를 공개하였다.

1. 문서 영상 인식 기술 개발

문서 인식 기술은 디지털카메라나 스캐너 등을 통하여 입력된 문서의 영상으로부터 문자 정보를 추출하는 기술이다. 문서 영상의 인식은 기존의 입력 매체들과 비교할 때 대량의 문자 정보를 빠르게 입력할 수 있는 장점을 지닌다.

문서의 인식은 그림 2와 같이 입력 받은 문서에 대하여 잡영 제거, 기울기 보정 등의 전처리 작업을 수행한 뒤, 문서의 구조를 분석하여 문자 영역을 추출하고, 해당 영역에 대하여 인식을 수행하는 순으로 진행된다. 본 과제에서는 각 단계별로 요구되는 각종 기반 기술들을 구현하여 이후 문서 영상 인식 시스템의 개발을 위한 기반을 마련하였으며, 이중 필기체 한글 및 한국어 인식 기법의 독자적 개발 및 성능 향상을 위하여 주력하였다.

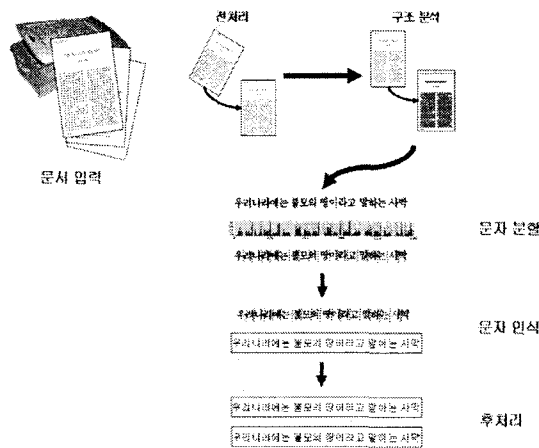


그림 2 문서 영상의 인식 과정

가. 문서 구조 분석

문서 구조 분석은 문서의 형태를 파악하고, 문자의 위치 정보를 추론하는 작업이다. 문서 구조 분석은 문서 영상의 인식을 위하여 필수적으로 선행되어야 하는 주요 작업이다. 본 연구에서는 기존에 발표된 문서 구조 분석 기법들을 한국어 문서들에 대해 실험함으로써 우리 환경에 적합한 문서 구조 기법을 구현하는 방안을 모색하였다.

실험에는 Run length 분석 알고리즘, Recursive X-Y cut 알고리즘, Docstrum 알고리즘의 세 가지 방법이 사용되었다. 첫째, Run length 분석 알고리즘은 문서 영상에서 일정 크기 이상의 수평, 수직 런을 추출하고 이들을 통합해나가며 문서 내의 문장 및 단락을 판별한다. 이 방법은 구현이 간단하고 수행속도가 빠르지만, 휴리스틱 파라미터에 의존도가 높아 일반화하기 어렵다. 둘째, Recursive X-Y cut 알고리즘은 문서 전체를 분석 단위로 간주하고 X 축, Y 축 방향 프로젝션 프로파

일 분석을 통하여 서로 구분이 뚜렷한 다수의 부영역으로 분할한다. 이후 각각의 부 영역에 대하여 재귀적으로 동일한 분할 과정을 더 이상 영역 분할이 발생하지 않을 때까지 반복한다. 이 방법은 수행속도가 매우 빠르고 문서의 계층적인 표현이 가능하다는 장점이 있다. 셋째, Docstrum 알고리즘은 연결 요소 분석을 기반으로 서로 유사한 성질을 가지는 구성 요소들을 차례로 합쳐 나감으로써 문장 및 단락을 판별하는 방식이며, 잡영 및 기울어짐의 영향을 상대적으로 덜 받는 장점이 있다.

실제 문서를 대상으로 각각의 알고리즘을 비교 평가한 결과, Docstrum 알고리즘이 상대적으로 가장 좋은 성능을 보였다. 그러나 Docstrum 알고리즘 역시 한글의 2차원적 구조를 표현하기 위해서 별도의 수정 및 개선 작업이 필요하였으며, 전반적인 분석 성능 향상을 위하여 최적 이진화 기법, 기울기 보정 기법, 영상 평활화 기법 등의 추가 기능들을 구현하였다.

나. 인쇄체 문자 인식기 개발

인쇄체 문자들은 규칙적인 형태를 가지며 변형이 적어 높은 성능의 인식기를 구현하기가 상대적으로 용이하다. 본 연구에서는 한글, 영어, 숫자 등 다양한 종류의 문자를 대상으로 하는 인쇄체 기반 인식기를 구현하였다.

영어 및 숫자 인식기는 한글에 비하여 상대적으로 문자의 수가 적고 형태 구분이 쉬워 인식기 구현이 용이하다. 본 연구에서는 윤곽선 방향 특징을 입력 받고 가우시안 커널을 사용하는 Support Vector Machine 인식기를 이용하여 구현하여, 99.53%의 인식률을 보였다.

반면, 한글은 자모의 2차원 배치를 통하여 날자가 형성되는 독특한 구조를 가지고 있기 때문에, 이를 고려하여 인식기의 디자인이 이루어져야 한다. 본 연구에서 개발한 인쇄체 한글 인식기는 크게 문자 추출기, 형식 분류기, 날자 인식기의 세 단계 구조로 이루어져있다. 문자 추출기는 문자간 경계를 신경망을 통하여 판별해 주는 인식기로서, 문자열을 입력받아 날자단위로 분할된 영상을 출력해 준다. 형식 분류기는 그물눈 특징을 입력받는 인공신경망을 사용하여 입력된 날자를 한글의 6형식에 따라 분류한다. 이후, 형식 분류 결과에 따라 초성/중성/종성의 영역을 설정한 뒤, 해당 영역 별로 전용 신경망 인식기를 사용하여 날자를 인식한다. 실험 결과 97.5%의 문자 추출률 및 97.5%의 문자 인식률을 보였다.

다. 필기체 문자 인식기 개발

필기체 문자는 인쇄체 보다 훨씬 다양한 변이를 가지고 있기 때문에, 모든 종류의 문자들에 대하여 좋은 성능을 보이는 보편적인 인식 방법론은 존재하지 않는다. 따라서 각각의 인식 대상에 따라 특화된 인식 방법론의 개발이 요구된다.

한글은 간단한 모양을 가지는 자모들의 조합으로 이루어지기 때문에, 단순한 획의 차이만으로 구분되는 유사한 모양의 문자들이 많이 존재한다. 따라서 한글의 경우 글자간의 미묘한 차이를 효율적으로 구분해 낼 수 있으면서, 동일한 글자에서 발생하는 변이를 효율적으로 흡수할 수 있는 정교한 인식 방법론의 개발이 필요하다.

본 연구에서는 필기체 날자 한글 영상의 인식을 위하여 랜덤 그래프를 이용한 인식기를 제안하였다. 제안된 인식기에서는 랜덤 그래프의 위상(topology)을 통하여 한글의 자모 기반 구조를 표현하고, 그래프의 각 노드에 획, 자모 모델, 날자 모델 등을 확률 변수의 형태로 대응시킴으로써 필기 문자의 형태 변형을 확률적으로 모델링하였다. 따라서 다양한 모양의 변이를 확률적 프레임워크 하에서 효율적으로 흡수할 수 있으면서도, 인식 결과를 획이나 자모와 같은 개념적 단위를 사용하여 해석할 수 있는 장점을 지닌다.

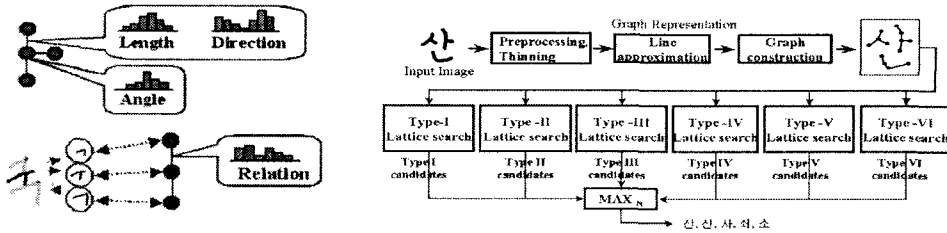


그림 3 필기체 한글 영상 인식기의 구조

입력된 문자 영상은 전처리, 획 추출, 그래프 생성을 거쳐 한글 6 형식에 대응되는 6개의 독자적인 인식기에 전달된다. 이 인식기들로부터 생성된 인식 후보 중 최적의 정합 확률을 가지는 글자가 인식 결과로 제공된다. 그러나 기본적으로 랜덤 그래프를 이용한 한글 인식 방법론은 입력 영상과 각 모델들 간에 조합 가능한 모든 사상을 조사해 보아야 하므로 높은 시간 복잡도 문제를 내포하고 있다. 이를 해결하기 위하여 각 자모 후보들을 격자 형태로 배열하고, DP matching 및 best first search 를 기반으로 하는 동적 탐색 기법을 적용하여 인식 속도를 향상시켰다. 또한, 인간의 인지과정에서 발견되는 특징인 선택적 주의집중 현상을 모델링함으로써 탐색 과정에서 발생하는 불필요한 자모 후보의 생성을 최소화하였다.

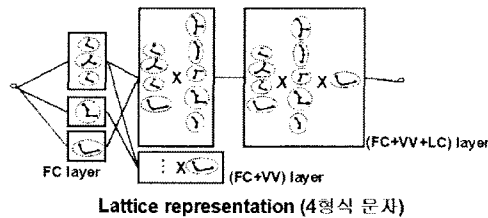


그림 4 자모 후보들의 격자 형태 표현 예

또한, 인식을 향상을 위해 개별 특징들 및 모델들 각각의 상호 연관 관계를 다차원 결합 확률 분포함수를 통하여 모델링하였다. 이들 다차원 확률 분포함수는 다변항 정규분포를 바탕으로 하는 선형회귀모델을 사용함으로써, 계산 시간 복잡도를 희생하지 않고도 보다 정교한 모델링 효과를 얻을 수 있었다. KU-1 필기체 한글 영상 데이터베이스에 대한 실험결과, 제안된 방법은 87.5%의 낱자별 인식을 보였으며, 기반 시스템에 비해 약 40.5% 가량의 실행 속도 향상 효과를 얻을 수 있었다.

한자는 한글에 비하여 상대적으로 복잡한 모양과 구조를 가지고 있어, 자모와 같은 세부단위로 모델로 표현하기 어렵다. 따라서 한자의 경우 각 낱자마다 개별적으로 모델을 작성하는 편이 바람직하다. 본 연구에서는 낱자를 이루는 획들과 그들의 상호 위치 관계를 확률적인 모델로서 표현하는 한자 인식 기법을 제안하였다. 제안된 모델에서는 문자 내부의 획들 간의 상호 의존관계를 정규 분포를 가정한 결합 확률 분포 함수를 이용하여 표현함으로써 모델의 변형을 흡수하였다. 다차원 의존관계의 표현으로 인한 모델의 복잡도 증가를 막기 위하여, 제안된 방법에서는 획들 간의 상호 정보량을 기준으로 각 획마다 가장 큰 영향을 미치는 획들에 대해서만 결합 확률 분포를 모델링하였다. 또한, 신경망 기반의 전용 구분기법을 적용함으로써 유사한 모양의 한자들을 구분할 수 있게 하였다. 실험 결과 제안된 방법은 99.46%의 낱자 인식을 보였다. 또한, 고문헌의 분석을 위하여 윤곽선방향특징을 이용하는 통계적 한자 인식기를 구현하였으며, 별도의 고문헌 자료를 대상으로 실험한 결과 93%의 인식을 얻었다.

숫자는 인식 대상의 수가 10개로 제한되어 있고 풍부한 양의 필기체 데이터베이스가 존재하기 때문에 명시적인 문자 구조의 모델링 없이도 높은 성능의 인식기를 구현할 수 있다. 본 과제에서는 누적 윤곽선 방향 특징을 기반으로 3차 다항식 커널을 사용하는 SVM을 사용하여 필기체 숫자 인식기를 구현하였으며 평균 98.7%의 인식을 얻었다.

라. 문맥 정보를 이용한 한글 인식기 성능 향상

한글 문서들을 조사해 보면, 방대한 한글의 조합 가능성에 비하여 실제로 사용되는 표현들의 수는 제약되어 있음을 발견할 수 있다. 이러한 언어적 특성으로부터 야기되는 문맥 정보를 활용할 경우 어절 및 문장과 같은 상위 단위에서의 인식을 전반적으로 향상시키는 것이 가능하다. 이와 같이 언어적 특성을 표현하는 모델을 언어 모델(Language model)이라고 지칭한다. 언어 모델은 일반적으로 언어의 생성을 연속된 단어들의 마르코프 과정으로 표현함으로써 임의의 문장의 사용빈도를 확률의 형태로 추정해 주는 확률적 언어 모델이 널리 사용된다. 그러나 한글의 어절은 여러 단어들의 조합으로 이루어지기 때문에 영어의 단어에 비하여 훨씬 다양한 가짓수를 갖는다. 따라서 효율적인 언어모델을 작성하기 위해서는 어절보다

세부적인 단위가 사용되어야 한다. 그러나 각 어절 내부에는 별도의 구분 기호가 없기 때문에, 형태소 분석과 같은 별도의 어절 분석 과정이 언어 모델 내에 포함되어야 한다.

본 연구에서는 개방 어휘 환경 하에서의 문자 인식을 위한 문맥 정보 모델을 개발하였다. 제안된 모델은 한국어에서 빈번히 사용되는 임의 길이의 문자열들을 기본 단위로 사용하는 동적 베이지안네트웍 구조를 지닌다. 제안된 방법은 언어 모델에서 사용할 어휘를 훈련 텍스트로부터 자체적으로 학습하기 때문에 별도의 어휘 지정 작업이 불필요하다. 또한, 문장 단위에서 각 어휘들 간의 종합적인 상관관계를 고려하기 때문에, 보다 넓은 문맥상에서의 상관관계를 보다 정확히 표현하는 것이 가능하다. 제안된 모델의 학습에는 EM 알고리즘을 기반으로 하는 반복 학습 알고리즘이 사용된다. 제안된 모델을 실제 인식기와 결합하여 필기체 한글의 어절 및 문장 단위 인식을 수행한 결과, 온라인 및 오프라인 인식기 모두 약 20% 가량의 어절단위 인식을 성능 향상을 얻을 수 있었다.

2. 일반 사진 영상에서의 문자 영역 추출 및 인식

일반 사진 영상으로부터 자동차 번호판, 간판 및 표지판 등과 같이 문자 정보를 포함한 개체를 추출하고, 해당 문자 영역의 내용을 인식하여 사용자에게 필요한 정보를 제공한다. 본 기술을 이용하여 차량 정보 또는 간판의 내용 정보 등을 활용할 수 있다. 그림 5는 본 기술의 진행 과정을 소개한다.



그림 5 일반 영상으로부터의 문자 영역 추출 및 인식

자연 영상은 다양한 색채를 가지고 있으며, 동일한 물체를 촬영할 경우에도 조명, 촬영각 등에 따라 다양한 변화가 나타날 수 있다. 따라서 자연 영상으로부터 문자 정보를 추출하기 위해서는 이와 같은 변형들을 흡수할 수 있는 안정적인 인식 방법론이 별도로 개발되어야 한다. 본 과제에서는 우선 정형화된 구조를 가지는 대상을 이용하여 자연 영상에서 발생 가능한 변이를 흡수할 수 있는 방법을 개발하고, 이를 점차 일반적인 대상으로 확장해 나가는 방식을 택하였다.

가. 정형화된 구조 영상에서의 문자 영역 추출 및 인식

자동차 번호판과 같이 문자 영역의 구조가 정형화되어 있는 영상의 경우, 일반적인 대상에 비해 문자 영역의 검출이 용이하다. 자동차 번호판의 경우, 색상 대비가 높은 단일 색상들로 이루어져 있으며, 정형화된 구조를 가지고 있기 때문에 번호판들 사이의 변이가 적다. 따라서 본 연구에서는 자연 영상에서의 변이를 안정적으로 흡수하는 영역 판별 및 문자 인식 기법의 개발을 위하여 자동차 번호판을 우선적인 연구 대상으로 삼았다.

영상으로부터 번호판이 존재하는 영역을 판별하기 위하여 색상 및 텍스처 정보를 입력받아 사용하는 SVM기반 인식기를 개발하였다. 해당 인식기는 CAMShift 알고리즘을 사용하여 선택된 일부 영역에 대해서만 판별 작업을 수행함으로써 전반적인 인식 속도를 향상시킬 수 있었다. 추출된 문자 영역은 색과 형태의 변이를 흡수하기 위하여 우선 회색조 영상으로 변환한 뒤, 외곽선 방향 특징을 기반으로 하는 인식기를 사용하여 그 내용을 인식하였다. 실험 결과, 수집된 자동차 데이터베이스에 대하여 제안된 알고리즘은 영상 판별율 99%, 해당 영역 내 문자 인식을 99.7%라는 높은 성능을 보였다.

나. 일반 사진 영상에서의 문자 영역 추출

간판 및 표지판 등 보다 일반적인 대상을 촬영한 영상은 대상 문자의 글꼴, 크기, 배치 등이 자유롭기 때문에, 조명 및 원근감등 주위 환경에 의하여 발생하는 변이 뿐 아니라 인식 대상 자체에 내재된 변이도 함께 모델링 되어야 한다. 따라서 자동차 번호판 인식에서 사용된 색상 및 텍스처 기반 인식기는 이와 같은 변이를 흡수하기에는 부적합하며, 일반 사진 영상에 대한 문자 영역 추출에 대한 연구를 별도로 추진하였다.

본 연구에서는 문자 영역 내에서 단위 면적당 경계선의 분포가 높게 나타나는 사실에 기인하여, 영상에 존재하는 경계 부분의 분포를 바탕으로 문자 영역을 판별하였다. 기본적으로 경계 검출 필터를 사용하여 입력된 컬러 영상을 경계 영상으로 변환한 뒤, 지정된 크기의 윈도우 내의 경계선 분포를 바탕으로 문자 영역 여부를 판단해 주는 인공 신경망 기반 인식기를 사용하여 문자 영역의 후보를 판별한다. 이들 연결요소 분석 및 모폴로지컬 연산 등을 통하여 후보들을 군집화한 후, 후보 영역의 크기, 위치, 모양 등의 분석을 통하여 문자 영역을 추출한다. 일상적인 대상을 촬영한 다양한 영상 데이터를 이용하여 제안된 방법을 실험한 결과 약 91.3%의 성공적인 검출 성능을 얻었다.

3. 펜 기반 인터페이스를 위한 요소 기술

펜 기반 인터페이스는 기존 GUI 환경의 마우스를 대체할 수 있는, 차세대 PC의 전체 인터페이스의 근간을 이루는 구성 요소이다. 펜 기반 인터페이스의 주요 기능

은 크게 필기 문자 인식과 명령 실행을 위한 제스처 인식으로 나눌 수 있다. 기능은 다르나, 문자 또는 제스처 모두 정형화된 기호를 인식한다는 점에서 이에 필요한 요소 기술의 원리는 동일하다. 본 연구에서는 필기 문자 입력으로서의 펜 기반 인터페이스에 초점을 맞추어 연구를 추진하였다. 제스처 인식의 경우, 1단계에서는 문서 교정 시스템에 특화하여 설계 및 구현하였고 본 결과를 통해 2단계에서 본격적인 활용 방안을 모색한다.

가. 은닉 마르코프 모델 기반 인식기의 성능 개선

은닉 마르코프 모델은 통계적인 학습 알고리즘을 가지고 있으며 문자의 시간적/공간적 변이를 효율적으로 흡수할 수 있는 특성을 지니고 있어 온라인 필기 문자 인식기의 개발에 널리 사용되고 있다. 그러나 은닉마르코프모델은 관측 기호들 간의 지역적인 의존 관계만을 모델링하기 때문에, 복잡한 구조를 가지는 문자의 경우 여러 획들 사이의 전역적인 상관관계를 제대로 표현하지 못하는 단점이 있다. 본 과제에서는 이와 같은 문제점을 해결하기 위하여 문자의 전역적인 구조 정보를 표현하는 인공 신경망 인식기를 은닉 마르코프 모델 인식기와 결합하는 방법을 제안하였다.

문자의 구조적 정보의 표현을 위하여 기본 획이라는 별도의 단위가 사용되었다. 이 기본 획들은 주어진 입력에 대한 은닉 마르코프 모델 상의 최적 상태 시퀀스를 기반으로 생성함으로써 문자의 변형에 의한 영향을 최소화하였다. 문자의 구조적 정보 모델은 각 자모별로 다층 신경망 인식기를 통하여 구현되었다. 이 인식기는 은닉 마르코프 모델에 의하여 제공된 인식 결과를 검증하여 후보들을 재정렬하는 역할을 수행한다.

실험 결과, 제안된 방법은 필기체 한글 인식기의 경우 59.9%의 오류를 감소시키는 효과를 가져왔다.

나. Bayesian Network 기반 인식 방법론의 개발

은닉 마르코프 모델이 가지는 구조적 한계를 근본적으로 해결하기 위하여, 본 연구에서는 베이지안네트웍을 이용한 새로운 온라인 필기 문자 인식 방법론을 제안하였다.

베이지안네트웍은 표현하고자 하는 대상을 그래프의 형태로 표현하는 확률적 모델이다. 베이지안네트웍 상에서 확률 변수는 그래프의 노드로, 이들 사이의 조건부 의존관계는 아크로 표현된다. 은닉 마르코프 모델이 이웃한 관측 기호 간의 1차 의존 관계만을 표현 가능한 것과 달리, 베이지안네트웍에서는 그래프의 구조에 따라 임의의 의존관계를 표현하는 것이 가능하다. 본 과제에서는 각 확률 변수들이 다변량 정규분포를 가진다고 가정하고, 이들 간의 의존관계를 선형 회귀 모델을 사용하

여 표현함으로써 모델 내부의 변수 증가에 따른 복잡도 증가를 최소화하였다.

제안하는 인식 방법론에서는 베이지안네트워크를 이용하여 점, 획, 자모 등 다양한 문자 구성요소들 간의 상관관계를 계층적 확률 모델로서 표현한다. 베이지안네트워크 모델의 그래프 구조는 각 구성요소들 간의 추상적인 관계를 바탕으로 조건부 독립 특성을 파악하여 설계한다. 일례로, 그림 6의 모델에서는 자모는 획의 조합으로 이루어지고, 획은 점의 조합으로 이루어진다는 지식에 따라 자모와 점 사이에는 획에 대한 조건부 독립 가정이 적용되고 있다.

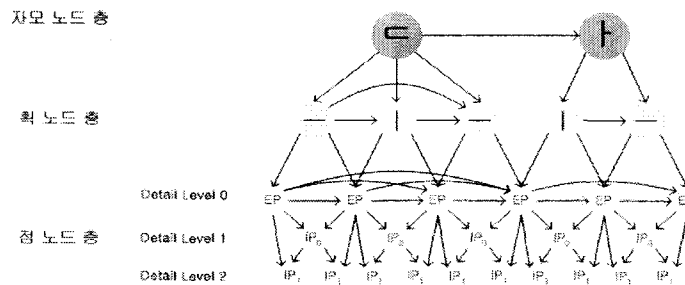


그림 6 한글 베이지안네트워크 인식기 구조 ('다'의 모델)

베이지안네트워크 인식기의 가장 큰 특징은 문자 전반에 대한 결합 확률 분포를 표현하는 모델이라는 점이다. 따라서 문자의 입력이 일부분만 이루어지더라도 그에 연관된 조건부 확률 분포에 따라 모델의 파라미터들이 입력에 맞추어 동적으로 변화되는 특징을 가지고 있다. 한 예로, 그림 7과 같이 숫자 '4'의 처음 세 획의 위치가 변화함에 따라, 동적으로 변화하는 마지막 획의 형태에 대한 예측이 가능하다. 이와 같이 베이지안네트워크 인식기에서는 문자의 전반적인 형태에서부터 각 획의 세부적인 모양에 이르기까지 다단계에 걸친 변이들을 체계적으로 흡수하는 것이 가능하다.



그림 7 미완성된 입력의 영향에 의한 베이지안네트워크 모델의 예측 지점 변화

실험 결과 제안된 방법은 95.72%의 낱자 단위 인식률을 보여 기존의 은닉마르코프 모델에 비하여 약 3% 가량의 인식 성능 향상을 보였다. 또한, 현재 출시된 상용 인식 시스템과 비교 실험을 수행한 결과 가장 높은 성능을 보여 제안된 방법의 실용적 우수성을 증명하였다.

필기 한글 인식기	인식률
KAIST Bayesian Network 인식기	95.72
KAIST HMM 인식기	92.90
MS Windows XP Tablet PC Edition 내장 인식기	86.31
상용 인식기 (국내 D사)	93.24

4. 요소 기술 응용 시스템 설계 및 구축

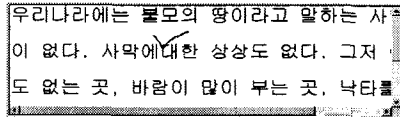
위에서 설명한 각 요소 기술을 바탕으로 본 연구는 Multi-Modal 입력 시스템의 일부 기능을 구현하였다. 1단계 연구로서, 인식된 문서를 펜을 이용하여 교정할 수 있는 교정 시스템, 문서 영상의 형식 분석 및 인쇄체 문자 인식이 가능한 문서 입력 시스템, 사진 영상으로부터 자동차 번호판의 내용을 입력받는 자동차 번호판 인식 시스템, PC 또는 PDA 상에서 필기 입력이 가능한 문자 입력기를 개발하고 기존 키보드 입력 방식의 개선 방안을 연구하였다.

한편, 본 연구에서는 2단계 연구에서 음성 인식 기술을 아웃소싱하여 입력 시스템으로서 구현하려 하므로, 음성 인식의 기술 동향에 대해 조사하고 그 활용 방안에 대해 제시하였다.

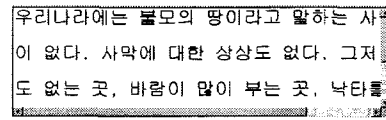
가. 문서 교정 기본 시스템 설계 및 구축

문서 인식 기술을 통해 인식된 결과는 사용자의 추가, 삭제, 변경 등의 자유로운 편집이나 교정 연산을 수행했을 때 더욱 만족스러운 결과를 가져오게 된다. 이러한 교정을 위해서는 펜 기반 문자 입력뿐만 아니라 여러 교정 부호를 인식할 수 있는 제스처 인식기의 구축이 필요하다. 제스처 인식기를 개발하기 위해서는 문서 교정 시 사용하기 쉽고 직관적인 제스처의 형태를 정의할 필요가 있다. 당해년도에는 호환성 및 데이터 수집 등을 고려하여 MS Windows XP Tablet PC Edition의 제스처 집합과 호환되는 8종의 편집용 제스처를 정의하여 사용하였다.

설계된 제스처 인식기 및 고안된 제스처를 이용하여 펜 기반 문서 교정 시스템의 프로토타입을 구현하였다. 교정 시스템은 문서 입력 시스템의 인식 결과를 입력으로 받아 텍스트 형태로 화면에 보여준 뒤 제스처 및 온라인 필기 인식을 통해 이전의 잘못된 인식 결과를 수정할 수 있도록 한다. 또한, 펜 기반 교정이 미숙한 사람도 사용하기 쉽게끔 키보드 및 마우스를 이용한 교정도 가능하도록 구현하였다. 교정 시스템의 실행 화면은 그림 8과 같다.



(a) 교정 부호 사용 (공백 입력)



(b) 교정 후 화면

그림 8 문서 교정 기본 시스템의 실행 화면

나. 문서 입력 기본 시스템 개발

본 연구에서는 문서 형식 인식 및 인쇄체 문자 인식 기술을 이용하여 한글 문서 영상을 입력받아 문서 형식을 분류하고, 분류된 각 영역 내의 문자를 인식하는 시스템을 구현하였다. 본 시스템을 이용하여 한글 문서의 내용을 빠른 시간 내에 PC로 입력이 가능하다. 그림 9는 본 문서 입력 기본 시스템의 실행 화면을 나타낸다.

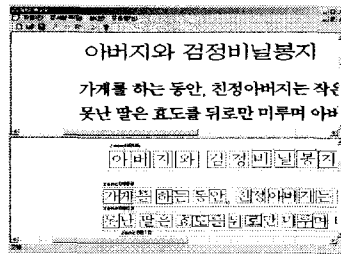


그림 9 문서 입력 기본 시스템의 실행 화면

다. 자동차 번호판 인식 시스템

본 연구에서는 일반 사진 영상에서의 문자 영역 추출 및 인식 기술을 바탕으로 자동차 번호판 인식 시스템을 구현하였다. 자동차의 전방 또는 후미가 촬영된 영상을 입력받아 번호판의 영역을 추출하고 해당 영역의 번호판 내용을 인식하여 출력한다. 그림 10은 본 자동차 번호판 인식 시스템의 인식 결과 화면을 나타낸다.

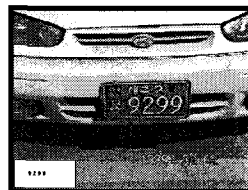
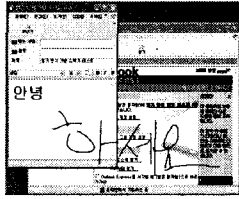


그림 10 자동차 번호판 인식 시스템의 실행 화면

라. 온라인 문자 인식 기반 문자 입력기 개발

본 연구에서는 온라인 문자 인식기를 문서 교정 시스템에서의 교정 인터페이스로 사용함과 동시에 응용 프로그램을 위한 문자 인식 서비스를 제공하기 위해 문자 인식 기반 입력기를 구현하였다. 개선된 은닉 마르코프 모델 기반 방식 인식기를 사용하여 한글과 영어, 숫자에 대한 필기 인식 기능을 제공하였다. 일반 데스크

탐 환경뿐만 아니라 Windows CE 기반의 PDA에서도 동작하도록 설계되었다.



(a) Windows 환경



(b) PDA 환경

그림 11 온라인 문자 인식 기반 문자 입력기

또한, 본 연구에서는 은닉 마르코프 모델 기반 한자 인식기와 베이지안네트웍 기반 기호 인식기를 독립적으로 구현하였다. 본 인식기들은 2단계 연구 시 기존의 인터페이스에 추가될 예정이다.

마. 음성 인식 기술 동향 조사

기술 동향 조사는 KAIST 전산학과 음성 언어 연구실의 자문을 얻어 이루어졌다. 현재 국내외적으로 많은 연구 기관 및 기업에서 음성 인식 기술에 대한 연구가 활발히 진행 중에 있으며, 음성 인식 기술을 세부 요소별로 나누어 정리하면 다음과 같다.

음성 인식 기술은 크게 연속음성 인식, 고립단어 인식, 화자 정보 인식, 음성 합성 등의 기술과 이를 뒷받침하는 신호 처리 기술, 잡음처리 연구 등으로 이루어져 있다. 고립단어 인식기술은 전화선을 통한 정보 검색, 이동전화 및 PC 명령어 인식 등에 활용되고 있으며, 화자 정보 인식 기술 또한 1-5%의 낮은 오류율을 보이며 상용화 수준에 도달해 있다. 전화망 또는 자동차 실내 등의 환경에서의 음성 인식을 위한 잡음 처리 연구도 활발하게 진행 중에 있다. 연속음성인식의 경우 국내의 화자독립/화자적응 인식을 지원함과 함께 96.7%의 인식률을 나타냄으로써 이미 실용화 단계에 근접해 있다.

본 연구에서는 이후 Multi-Modal 입력 시스템 구축 시, 고립 단어 인식을 통한 명령어 입력 시스템, 화자 인식을 통한 사용자 인증, 음성 합성을 이용한 Text-to-Speech / Image-to-Speech 등의 기능 구현을 계획 중이다.

바. 기존 키보드 입력 방식 개선에 대한 연구

본 연구에서는 기존 키보드의 입력 효율성을 향상시키기 위한 연구의 일환으로, 미국 성공회 소속 안마태 신부와 함께 개선된 컴퓨터 자판 설계에 관한 연구를 수행하였다. 제안된 자판은 자음과 모음의 동시 입력을 가능하게 하여 입력속도를 향상시켰으며, 한글 창제 원리를 고려하여 자모를 배열함으로써 학습을 용이하게 하

였다.

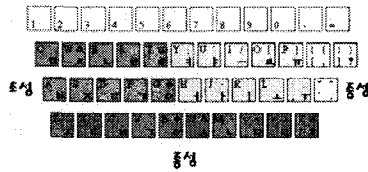


그림 12 개선된 한글 자판의 구성

5. 문자 데이터베이스의 구축

방대한 양의 문자 및 영상 데이터베이스는 관련 인식 기술 개발 및 훈련, 검증, 평가 등에 있어서 필수 불가결한 요소이다. 각종 변이에 강인한 인식기를 구현하기 위해서는 가능한 다량의 데이터를 사용하여 훈련시키는 것이 바람직하며, 테스트에 사용된 데이터의 규모가 클수록 실제 환경 하에서의 인식 성능의 편차가 줄어든다. 또한, 상이한 방법론을 이용하여 개발된 인식기들의 공정한 비교 평가가 이루어지기 위해서는 훈련 및 평가 과정에서 동일한 데이터베이스가 사용되어야 한다.

본 과제에서는 다양한 대상에 대한 요소 인식 기술을 개발하기 위하여, 그동안 국내외 연구진에 의하여 공개된 문자 및 영상관련 데이터베이스들을 수집하고, 한글 및 한국어 인식관련 자료를 독자적으로 구축하였다. 수집된 데이터베이스의 종류 및 규모에 관한 자세한 사항은 별도의 자료로 정리하여 첨부하였다.

수집된 자료들은 기본적으로 인터넷을 통하여 일반에 자유로이 공개하며, 별도의 라이선스 취득이 필요한 자료의 경우 해당 기관으로부터 라이선스를 인증 받았음이 확인된 경우에 한하여 배포하고 있다.

제 4 장 목표달성도 및 관련분야에의 기여도

1절 연구 개발 목표 달성도

번호	세부연구개발목표	달성내용	달성도 (%)
1	문서 영상의 전처리 및 문자 구조 분석 기법의 구현	<ul style="list-style-type: none"> ○ 문서 영상용 영상 처리 기법들의 구현 <ul style="list-style-type: none"> - 적응적 이진화 기능 (Otsu's method /Niblack's method) - 잡영 제거 및 각종 필터링 기능 - 기울기 보정 기능 ○ 한국어 문서 구조 분석 기법 개발 <ul style="list-style-type: none"> - Recursive X-Y cut / Docstrum 기법 구현 및 개선 	100
	오프라인(영상기반) 문자 인식기법 개발 및 인식기 구현	<ul style="list-style-type: none"> ○ 계층적 랜덤 그래프 기반 필기체 한글인식기 개발 (인식률:87.5%) <ul style="list-style-type: none"> - 인지 과학적 특성에 기반한 효율적 탐색 기법 개발 적용 - 다차원 의존관계 모델링을 통한 인식 성능 향상 - 기존에 비해 40%의 인식 속도 개선, 5% 인식률 향상 ○ 인쇄체 문자 인식기 구현 <ul style="list-style-type: none"> - 인공 신경망 기반 한글 인식기 구현 (인식률: 97.5%) - SVM 기반 숫자 및 영문자 인식기 구현 (인식률: 99.53%) ○ 필기체 문자 인식기 구현 <ul style="list-style-type: none"> - SVM 기반 숫자 인식기 구현 (인식률: 98.7%) - 한자 인식기 구현 (인식률: 99.46 %) 	
	문맥 정보를 이용한 인식 결과의 후처리 기술 개발	<ul style="list-style-type: none"> ○ 개방어휘 환경 하에서의 문자인식을 위한 문맥정보 모델 개발 <ul style="list-style-type: none"> - 필기체 한글 어절 단위 인식률의 20% 성능 향상 	
2	변형이 심한 환경 하에서의 문자 인식 기술 개발	<ul style="list-style-type: none"> ○ 정형화된 구조를 가지는 대상의 영상 인식 <ul style="list-style-type: none"> - 자동차 번호판 인식 시스템 구현 (추출율: 99%, 인식률:97%) > 색상/텍스처 기반 SVM 인식기 및 CAM Shift 알고리즘을 이용한 문자 영역 탐색 > 인공 신경망 기반 번호판 문자 인식기 구현 ○ 비정형 대상으로부터의 문자 영역 검출 (검출율: 91.3%) <ul style="list-style-type: none"> - 경계 영상 기반 인공신경망 인식기를 이용한 문자영역 후보판별 - 군집화 기법 및 검증을 통한 문자 영역 추출 	90

3	온라인 필기 인식 기술 개발 및 성능 개선		<ul style="list-style-type: none"> ○ 기 보유중인 HMM 기반 인식기의 오류 유형 분석 <ul style="list-style-type: none"> - HMM의 구조적 한계로 인한 오류 발생 유형 파악 (자모 모델링 오류 / 자모 위치관계 오류 / 혼동되는 글자 / 잘못 쓰인 글자) ○ 은닉 마르코프 모델 기반 인식기의 성능 향상 <ul style="list-style-type: none"> - 기본획 단위 자모 인식 신경망 검증기를 이용하여 문자의 구조적/전역적 정보를 모델링함 - 필기 한글 인식기 오류의 59.9%감소 ○ 베이지안네트워크 기반 인식 방법론 개발 <ul style="list-style-type: none"> - 문자의 구조적 지식과 변이 형태를 일관적인 확률적 구조 내에서 표현 - 다차원 의존관계 모델링을 통한 문자모양의 정교한 표현 - 상용 제품보다 우수한 인식률을 보임 ○ 베이지안네트워크 기반 필기 인식기 구현 (인식률 - 한글: 95.72%, 숫자: 99.7%, 기호: 90.68%) ○ HMM 기반 필기 인식기 구현 (인식률 - 한글: 92.9%, 영어: 93.2%, 한자: 92.5%) 	100
4	문서 교정 기본 시스템 설계 및 구축	<p>인식 결과 검증/교정 인터페이스 설계 및 구현</p> <p>음성인식 기술 동향 조사</p>	<ul style="list-style-type: none"> ○ 편집용 제스처의 고안 <ul style="list-style-type: none"> - Windows XP Tablet PC Edition의 표준 제스처와 호환 ○ 베이지안네트워크 기반 제스처 인식기 개발 ○ 인식결과 검증/교정 인터페이스 프로토타입 구현 <p>○ 상용화 수준의 음성 인식 기술</p> <p>○ 교정 시스템 등 폭 넓은 분야에 적용 가능성을 확인</p>	100
5	문자 DB 구축		<ul style="list-style-type: none"> ○ 한글/한자/숫자/영어/기호 등 다양한 대상에 대한 필기체/ 인쇄체 데이터베이스 수집 <ul style="list-style-type: none"> - 필기체: ETRI, KU-1, PE, KAIST 등 - 인쇄체: ETRI - 온라인: KAIST, Turkey, UNIPEN 등 ○ 자동차 번호판 / 간판 / 표지판 등 자연 영상내의 문자 정보 데이터베이스 구축 ○ 전용 서버 운영을 통한 수집 자료의 공개 	100

2절 연구개발성과

구분	기술이전		기술료 수입			특허출원			특허등록			논문게재			학술발표		
	건수	건수	금액 (억원)	국내	국외	계	국내	국외	계	국내	국외	계	국내	국외	계		
1차년도 (’01)													1	1	3	2	5
2차년도 (’02)													4(2+)	4(2+)	4	2(4*)	6(4*)
총계													5(2+)	5(2+)	7	4(4*)	11(4*)

+ 심사중 * 발표예정

1. 논문게재

논문제목	저자	학술지명	제출일	발행년도 권호	발행기관 (국명)	SCI 여부
Pair-wise Discrimination Based on a Stroke Importance Measure	김인중 김진형	Pattern Recognition	2001년 8월 6일	2002 Vol.35 No.10	Elsevier Science (네덜란드)	O
Verification of Graphemes Using Neural Networks for HMM-based On-line Handwritten Hangul Recognition	조성정 김진형	International Journal of Computer Processing of Oriental Languages	2002년 3월	2002 Vol.15 No.2	World Scientific (싱가폴)	X
Model-based Stroke Extraction and Matching for Handwritten Chinese Character Recognition	C. L. Liu 김인중 김진형	Pattern Recognition	2000년 10월	2001 Vol.34 No.12	Elsevier Science (네덜란드)	O
Texture-based Approach for Text Detection in Images Using Support Vector Machines and Continuously Adaptive Mean Shift Algorithm	김광인 정기철 김진형	IEEE Transactions on Pattern Analysis and Machine Intelligence	2001년 10월25일	채택	IEEE (미국)	O
Statistical Character Structure Modeling and Its Application to Handwritten Chinese Character Recognition	김인중 김진형	IEEE Transactions on Pattern Analysis and Machine Intelligence	2002년 3월 28일	채택	IEEE (미국)	O
Utilization of Stochastic Relationship Modeling in Hierarchical Character Recognition	강경원 김진형	IEEE Transactions on Pattern Analysis and Machine Intelligence	2002년 12월2일	심사중	IEEE (미국)	O
Bayesian Network Modeling of Strokes and Their Relationships for On-line Handwriting Recognition	조성정 김진형	Pattern Recognition	2002년 2월 15일	심사중	Elsevier Science (네덜란드)	O

2. 학술발표

학술발표제목	발표자	발표장소 (국명)	일시
Color Texture-based Object Detection: An Application to License Plate Localization	김광인 정기철 김진형	International Workshop on Pattern Recognition with Support Vector Machines(미국)	2002년 8월
Automatic Completion of Korean Words for Open Vocabulary Pen Interface	류성호 김진형	5th International Conference on Document Analysis Systems (미국)	2002년 8월
Bayesian Network Modeling of Strokes and Their Relationships for On-line Handwriting Recognition	조성정 김진형	6th International Conference on Document Analysis and Recognition(미국)	2001년 9월
On-Line Signature Verification Using Model-Guided Segmentation and Discriminative Feature Selection for Skilled Forgeries	이택헌 조성정 김진형	6th International Conference on Document Analysis and Recognition (미국)	2001년 9월
문자의 구조적 제약과 동적 격자 탐색을 이용한 필기 한글 문자 인식	강경원 김진형	제13회 한글 및 한국어 정보처리 학술대회 (한국)	2001년 10월
유사 문자쌍을 구별하기 위한 한글 인식의 후처리	장승익 김진형	제13회 한글 및 한국어 정보처리 학술대회 (한국)	2001년 10월
온라인 필기 문자 인식을 위한 베이지안 망에서의 확과 확관계의 표현	조성정 김진형	컴퓨터비전 및 패턴인식 연구회 추계 워크샵 (한국)	2001년 11월
자모 단위 유사쌍 구분을 이용한 필기체 한글 인식 후처리	장승익 강경원 김진형	제14회 영상처리 및 이해에 관한 워크샵 (한국)	2002년 1월
확률적 언어 모델을 위한 자료 기반 어휘 구축	류성호 김진형	제14회 한글 및 한국어 정보처리 학술대회 (한국)	2002년 11월
An Efficient Digitalizing Scheme of Handwritten Documents in Oriental Languages	곽희규 류성호 김진형	컴퓨터 비전 및 패턴인식 연구회 추계 워크샵 (한국)	2002년 11월
분절 비교 기반의 온라인 서명 검증	이택헌 김진형	컴퓨터비전 및 패턴인식 연구회 추계 워크샵 (한국)	2002년 11월
Learning the Lexicon from Raw Texts for Open Vocabulary Korean Word Recognition	류성호 김진형	7th International Conference on Document Analysis and Recognition (영국)	2003년 8월 발표예정
Bayesian Network Modeling of Hangeul Characters for On-line Handwriting Recognition	조성정 김진형	7th International Conference on Document Analysis and Recognition (영국)	2003년 8월 발표예정
Handwritten Hangeul Character Recognition with Hierarchical Stochastic Character Representation	강경원 김진형	7th International Conference on Document Analysis and Recognition (영국)	2003년 8월 발표예정

Generation of Handwritten Characters with Bayesian network based On-line Handwriting Recognizers	최현일 조성정 김진형	7th International Conference on Document Analysis and Recognition (영국)	2003년 8월 발표예정
--	-------------------	--	------------------

3절 대표적 성공사례

1. 확률적 관계 모델링을 이용한 계층적 필기체 한글 문자 인식

본 연구실에서는 확률적 관계 모델링을 이용한 계층적 필기체 한글 인식 기술을 개발하였다. 랜덤 그래프의 위상으로 한글의 구조를 표현하고 확률 변수를 이용하여 필기 문자의 형태 변형을 확률적으로 모델링함으로써 변형이 심한 글자에 대해 안정적인 성능을 나타내어, 기존 필기 한글 인식기에 비해 30%의 오류 감소율을 보였다. 본 연구를 통해 관계 모델링의 중요성을 보였으며, 관계 모델링은 향후 문자 인식의 중요 기술로 쓰일 것으로 전망된다.

2. 개방 어휘 환경 하에서의 문자 인식을 위한 문맥 정보 모델 개발

본 연구실에서는 개방 어휘 환경에서의 문자 인식을 위한 한국어 언어 모델을 생성하는 방법을 제안하였다. 제안된 방법은 언어 모델에 사용되는 어휘를 일반 텍스트로부터 자체적으로 학습하며, 문장 단위와 같은 넓은 문맥상에서의 어휘들 간의 상관관계를 보다 정확히 표현하는 것이 가능하다. 제안된 언어 모델을 적용한 필기 한글 인식에 사용한 결과, 약 20% 가량의 어절단위 인식률 향상 효과를 얻을 수 있었다.

3. 인공 신경망과 연속 탐색 알고리즘 기반 차량 번호판 인식

영상 처리 기법과 패턴 인식 방식의 장점을 혼용하여 차량 번호판의 탐색 및 이의 인식을 수행하여 99.0%의 탐색율과 99.7%의 높은 인식률을 보였다. 자연 영상에서 문자 영역 탐색 시 패턴 인식적 접근 방식을 도입하여 관련 분야의 국내외 학자들에게 본 알고리즘의 응용 및 개선에 대한 연구 기반을 제공하였다. 앞으로 높은 탐색율과 정확도를 기반으로 실용화 수준의 응용 시스템의 개발이 예상된다.

4. 베이저안네트웍(Bayesian Network) 기반 온라인 필기 인식 기술

본 연구실에서는 기존의 은닉 마르코프 모델의 구조적 한계를 뛰어넘는 새로운 베이저안네트웍 기반 온라인 필기 인식 기술을 개발하였다. 본 기술은 문자 구성 요소간의 의존 관계를 결합 확률 분포를 이용하여 조건부 확률로 표현하므로 글자

형태의 정확한 모델링이 가능하다. 본 기술은 문자 인식 기술 개발의 새로운 패러다임을 창출했으며, 현존하는 타 인식기와의 성능 비교 결과 가장 우수한 성능을 보여, 앞으로 이에 대한 활발한 연구가 예상된다.

5. 문자 영상 DB 구축

본 연구실에서는 국내 문자 인식 기술 개발을 돕기 위해 다양한 종류의 문자 영상 DB를 구축하였다. 필기체 한글을 중심으로 하여, 필기체 영·숫자, 인쇄체 한글 및 영·숫자, 온라인 필기 데이터, 간판 및 번호판 촬영 영상 등을 총 12종, 66.7G의 데이터를 수집하였다. 현재 이 자료들은 국내외 연구자들이 자유로이 이용할 수 있도록 전용 서버를 통해 인터넷으로 공개하고 있다. (<ftp://aidb.kaist.ac.kr/>)

4절 대외기여도

1. 홈페이지 운영 및 활용 현황

본 연구실에서는 연구실의 연구 현황 제공과 국내외의 기술 동향 안내라는 두 가지 역할을 중심으로 하여 홈페이지(<http://ai.kaist.ac.kr>)를 운영, 활용하고 있다. 연구실의 현황에는 현재 진행 중인 연구 분야에 대하여 문제 및 접근 방법, 실험 결과, 연구실에서 발표한 국내외 학술지, 학회 논문, 석·박사 학위 논문의 목록 및 일부 내용이 나타나 있다. 국내외 기술 동향 정보 제공을 위해 관련 분야의 새 소식, 학회 및 해외 탐방기, 국내외 관련 기술 동향 등을 제공한다.

2. 산·학·연 협력거점 활동현황

본 연구실은 문자인식 관련 분야를 이끌어 가는 연구실로서 문자 인식과 관련된 산·학·연의 연구 동향과 연구 주제 등에 대해 토의하는 문자인식 워크샵에 2002년과 2003년 양해에 본 연구실의 연구 성과 및 실제 산업 현장에서 요구되는 문자 인식 기술의 이슈들 및 해결 방안에 관한 논의에 주도적으로 참여 하였다. 또한 연구 관련 메일링 리스트를 운영하여 학회 및 워크샵 정보, 국내외 기술 동향, 연구실 내 연구 성과 등의 소식을 적극적으로 학계/산업계에 전달하고 있다.

3. 기술지도 및 성과활용 등

본 연구실에서는 본 과제를 통해 개발된 온라인 문자인식 알고리즘을 응용하여 삼성종합기술원이 주관하는 지능형 마이크로시스템 개발 사업에 위탁 과제로서 마

이코로 무선입력장치의 문자인식 알고리즘을 개발, 필기체 한글인식 알고리즘을 응용하여 전자통신연구원이 주관하는 우편 배달 순로 시스템 개발 사업에 위탁 과제로서 문자인식 시스템 신뢰도 향상에 관한 연구를 수행하였다. 또한, (주)동방 S&C 에서 개발 중인 고문서 디지털화 시스템에 본 연구실에서 개발한 필기체 한자 인식 기술을 전수하여 현재 활발하게 사업이 진행 중에 있다.

4. 첨단기술정보의 제공, 연구기기 공동 활용

본 연구실에서는 본 과제를 통해 개발된 내용을 학술지 및 학회 발표, 또는 기술이전 등을 통해 제공하고 있을 뿐만 아니라, 본 과제 또는 문자 인식과 관련한 문서 및 문자 단위 영상 데이터베이스와 온라인 필기 데이터베이스 등 각종 데이터베이스를 수집하여 FTP 서비스(<ftp://aidb.kaist.ac.kr>)를 통해 제공한다. 또한, 본 연구실에서 영상 처리와 관련되어 제작한 KLib 라이브러리(<http://ai.kaist.ac.kr/~klib>)를 배포하였다. 특히 (주)동방 S&C 등의 외부 기관의 모듈 및 소프트웨어 구현에 큰 역할을 하고 있다.

5. 국내외 연구기관과의 협력

국내의 경우 숙명여대와는 영상으로부터의 문자 추출에 대한 연구를, 삼성종합기술원과는 온라인 필기 문자인식에 대한 연구를, 전자통신연구원과는 필기체 한글 영상 문자인식 및 문맥정보 활용에 관한 연구에 관해 긴밀히 협의하였다. 국외의 경우, 중국의 清華大學과는 2002년 3월 북경에서 “제 1회 KAIST-Tsinghua 문서 인식에 관한 협력 워크샵”을 공동 개최하여 18 편의 논문을 발표하였으며, 2001년 7-9월에 영국의 University of Essex 에 연구원을 파견하여 영국 국립 자연사박물관 카드 인덱스의 디지털화에 따른 문서 인식에 관한 연구를 협조하는 등 해외 연구기관과도 긴밀한 협조를 유지하고 있다.

6. Workshop, seminar 및 단기강좌를 통한 기술협의, 기술교류 등 실질적 자문활동 등

본 연구실에서는 1회의 국제 공동 워크샵(제 1회 KAIST-Tsinghua 문서 인식에 관한 협력 워크샵, 2002.3.28-29, 북경)을 개최하였으며, 2회의 한국 문자인식연구회 문자인식워크샵 (2002.2.22-23, 무주; 2003.2.14-15, 무주)에 참여하였다. 또한, 2년간 총 6회에 걸쳐 국내외 강사들을 초빙하여 세미나를 개최, 국내외 연구진들의 연구동향을 파악하는 시간을 가졌다.

제 5 장 연구개발결과의 활용계획

1단계 기간 동안 개발된 다양한 인식 관련 요소 기술들과 인식기들은 이후 구현될 Multi-Modal 입력 시스템의 기본 구성요소로서 활용될 예정이다. 개발될 입력 시스템은 스마트폰, PDA, Tablet PC 등과 같은 차세대 이동형 컴퓨터 장비 상에서 펜, 음성, 영상 등 다양한 매체를 입력 수단으로 사용할 수 있는 지능형 환경을 제공할 것이다.

또한, 각각의 요소 기술들은 관련 분야의 연구진 및 기업들과 공동 작업을 통해 다양한 제품 개발에 활용할 예정이다. 일례로, 본 연구실은 온라인 필기 인식 기술을 삼성 종합기술원에서 개발 중인 3차원 전자펜에서 사용가능하도록 제공하였으며, ETRI 우정 자동화 사업팀과의 공동 연구를 통해 우편물 자동 구분기의 국산화를 위한 한글 주소 인식기 개발에 참여하고 있다. 또한, 규장각 소장자료 전산화 작업을 수행중인 (주) 동방미디어에 필기 한자 인식 관련 기술을 제공하여 해당 기업이 관련 분야에서 독보적인 기술적 우위를 선점하고 세계 시장으로의 진출을 위한 기반을 마련하는데 기여하였다. 이후 연구 기간에도 개발된 기술을 활발히 공유함으로써 관련 분야의 활성화에 기여할 예정이다.

제 6 장 참고문헌

- [USP98] Comprehensive Statement for FY 1998, U.S Postal Service
(<http://new.usps.com/pdf/comp98.pdf>)
- [DT01] 디지털타임즈 2001년 3월 29일
(http://www.dt.co.kr/dt_srcview.html?gisaid=2001032902100352540001)
- [KB98] 금융결제원과 한국은행의 금융기관 수납장표의 정보화 추진에 관한 보고서, 1998
- [GAT99] The State of the Document Technologies Industries: 1999-2003, Gartner Group, AIIM Internation, 1999
- [GAT02] Gartner Report on Tablet PC sales 2002, Gartner Group,
http://www4.gartner.com/5_about/press_releases/2002_11/pr20021106b.jsp, 2002/11/6
- [APP02] Mac OS X - Feature - Inkwell, Apple,
<http://www.apple.com/macosx/jaguar/inkwell.html>
- [NF03] Are Tablet PCs All That, News factor,
<http://www.newsfactor.com/perl/story/21418.html>, 2003/5/2
- [GES03] Design Considerations for Gestures, MSDN,
<http://msdn.microsoft.com/library/default.asp?url=/library/en-us/tpcsdk10/html/whitepapers/designguide/tbconuxdggesturedesignconsiderations.asp>
- [PKP03] Handheld sales down by 10 percent, pocketpcBlast! News,
<http://pocketpc.pdablast.com/articles/2003/4/2003429-Handheld-Sales-Down-11.html>, 2003/4/29
- [CNET02] Tablet PC rivalry sets in, CNET News,
<http://news.com.com/2100-1001-963131.html?tag=nl>, 2002/10/24
- [REG03] Mobile handset sales up 6 per cent in 2002, The Register,
<http://www.theregister.co.uk/content/68/29693.html>, 2003/11/03
- [WIN03] Sluggish Economy Hits PDA Sales, Wireless News,
<http://www.wirelessnewsfactor.com/perl/story/21415.html>, 2003/5/1
- [ETR02A] MH 시스템 진단 및 마스터 플랜 수립, 한국전자통신연구원
우정기술연구센터, 2002
- [ETR02B] "2002년 영상정보처리산업 육성 및 활용사례", 한국전자통신연구원