

임베디드 시스템에서 필획기반 다국어 입력 시스템

Stroke based Multilingual Input System for Embedded System

이진영*
Jin-Yeong Lee

홍성용**
Sung-Ryong Hong

이시진***
Si-Jin Lee

요약

최근의 정보기술 동향은 모바일 서비스를 중심으로 발전해 가고 있으며, 대부분의 휴대폰 이용자들이 통화뿐만 아니라 무선 인터넷을 통해 제공되는 다양한 서비스를 이용하고 있다. 다양한 응용 서비스를 가능하게 해주는 시스템 소프트웨어 및 미들웨어의 중요성이 더욱 확대되고 있으며, 그 중 하나가 문자 입출력 시스템이다. 본 논문에서 제안한 입력 시스템은 해당 언어의 생성원리를 기반으로 알파벳을 분해하여 글 쓰듯이 입력하는 필획(Stroke) 방식으로 구성하여, 언어를 알고 있는 사람이면 누구나 쉽게 입력할 수 있도록 개발하였다.

Abstract

Recently, development in information technology is mainly focused on mobile service, and most of mobile users are using various services based on wireless network. So the importance of system software or middleware, which enables such mobile services, is growing bigger and bigger, and one of those is character input/output system. This paper will introduce an alphabet input system, which decomposes a character to a series of strokes, by its formation principal. It is designed to make a person, who knows the character, to input characters in the way that he/she is actually writing down the character.

□ keywords: 입출력, 입력기, Input Method, 임베디드 시스템, 다국어, 언어 처리

1. 서론

인터넷의 보편화로 인해 최근의 정보기술 동향은 모바일 서비스를 중심으로 발전해 가고 있으며, 대부분의 휴대폰 이용자들이 통화뿐만 아니라 무선 인터넷을 통해 제공되는 다양한 서비스를 사용하고 있다. 현재까지 모바일 기기는 음성 서비스 위주의 CDMA 칩셋 등이 주요 핵심 기술이었던 반면에, 앞으로 다가올 데이터 서비스 위주의 시장에서는 다양한 응용 서비스를 가능하게 해주는 시스템 소프트웨어 및 미들웨어의 중요성이 더욱 확대되고 있다. 그 중 하나가 문자 입출

력 시스템이다. 휴대폰의 경우, 각 언어별로 10 ~ 12개의 버튼을 이용하여 문자를 입력하고 있다[1, 2]. 입력 시스템에는 T9, Zi 등이 있다. 그러나 이들 입력 시스템은 각 언어의 제자 원리와 나라별 문화 특성을 고려하지 않고 있다. 본 연구에서는 각 언어별 입력기를 해당 언어의 생성원리를 기반으로 제작해서, 언어를 알고 있는 사람이면 누구나 쉽게 입력할 수 있도록 개발하고자 한다. 예를 들어, 한글 입력기의 경우, 한글의 전신으로 알려져 있는 가림토를 기반으로 설계하고, 한글을 아는 사람이면 누구나 쉽게 인지할 수 있도록 개발했다. 또한, 기존 입력 시스템 보다 빠른 속도로 원하는 문자를 입력할 수 있도록 했다. 예를 들어, 기존 일본어 입력 시스템은 평균 3타 정도를 입력해야 한다. 하지만, 새롭게 제안한 일본어 입력 시스템은 히라가나 문자를 분해한 도형을 글쓰는 순서대로 조합하기 때문에 평균적으로 2

* 정 회 원 : 타임스페이스시스템(주) 기술연구소 부소장,
leejin@empal.com

** 정 회 원 : 타임스페이스시스템(주) 대표이사
hsr@timespace21.com

*** 종신회원 : 대전대학교 컴퓨터공학과 부교수
sjlee@daejin.ac.kr

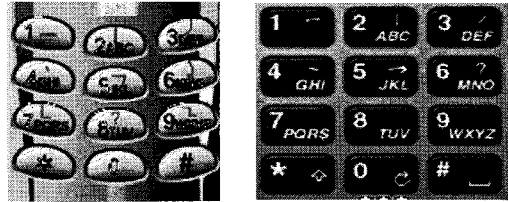
[2007/07/25 투고 - 2007/07/31 심사 - 2007/08/14 심사완료]

에서도 상대적으로 우위에 있다. 그러나 한 자판에 자음과 모음이 동시에 배치됨으로써 연속글자가 입력되지 않는다. 따라서 입력 도중 이동 키를 입력해야 하는 단점이 있다. 네오패드는 본 연구의 입력 방식과 유사한 원리를 바탕으로 하고 있다. 즉, 한글의 자음과 모음을 도형으로 구분하여 도형을 조합해 나가면서 입력하는 방식이다. 그러나 분해한 도형 자체가 이해하기가 난해하며, 전세계 언어를 거의 동일한 도형을 가지고 입력하기 때문에 각 나라의 언어 특성을 제대로 반영하지 못한다는 단점이 있다[7, 8].

위에서 살펴본 바와 같이 국내의 입력 시스템은 대부분 한글 입력 시스템만 지원하고 있는 실정이다. 또한, 전세계적으로도 유래를 찾아 보기 힘들 정도로 국내 시장만큼은 국내업체의 한글 시스템이 시장을 장악하고 있다. 세계적으로 다국어 문자 입력 시스템에 대한 전문 기업에는 Tegic Communication(T9)과 Zi Corporation(Zi)이 있다. 이들 두 업체는 전세계 문자 입력 시장의 대부분을 차지하고 있는 실정이다. 그러나 이 두 업체의 입력 시스템은 각 나라의 언어와 문화적 특성을 정확히 반영하고 있지 못하기 때문에 일부 언어의 경우, 인터페이스를 직관적으로 이해하기 힘들며 입력 타수 또한 많은 상태이다. 다음은 중국어 입력기를 기반으로 두 업체의 입력 시스템에 대해서 비교 평가해 보고자 한다.

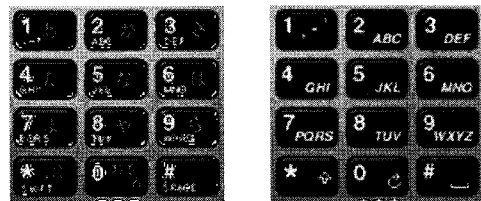
T9 중국어 입력 시스템은 필획 입력 방식을 제공하고 있다. 필획이란 한자를 구성하는 각 필획을 글쓰는 순서대로 쓰는 방식을 의미한다. 이때, 사용하는 필획은 그림 2에서 보는 바와 같이 총 5개를 이용하고 있다. 따라서, 한자를 구성하는 여러 개의 필획을 단 5개의 필획으로 함축하고 있기 때문에 필획을 선택하기가 쉽지 않으며, 하나의 한자를 입력하기 위한 총 타수가 많다는 단점이 있다[9].

Zi 중국어 입력 시스템 역시 T9과 같이 필획 방식을 제공하고 있다. Zi는 그림 2에서 보는 바와 같이 총 9개의 필획을 이용하여 한자를 입력



(그림 2) Zi(좌) T9(우) 중국어 입력 자판

하고 있으며, T9에 비해 한자의 필획을 조금 더 정확하게 표현하고 있다. 그러나 이 방식 역시 하나의 필획 버튼에 여러 필획을 대표하도록 설계되어 있기 때문에, 필획 선택을 하기 위해서는 각 패덴을 암기해야 하는 문제점이 있다. 예를 들어, ‘人’을 입력하기 위해서 ‘34’ 버튼을 입력한 후에 후보 군에서 한자를 선택해야 하며, 자판 모양을 자세히 보아 알 수 있듯이 일반적으로 쓰는 ‘人’과는 다른 모양의 필획을 조합해야 하는 문제를 안고 있다[10].



(그림 3) T9일본어, Zi영어 입력 시스템

두 회사는 중국어 외에도 전세계의 다국어 입력 시스템을 보유하고 있으며, 본 연구에서 개발한 입력 시스템도 한글, 영어, 중국어, 일본어에 대한 입력 시스템을 보유하고 있다. 그림 3은 T9과 Zi의 기타 입력 시스템에서의 자판 배열을 보여 주고 있다. 그림에서와 같이 중국어와 한글을 제외한 언어는 대부분이 한 자판에 해당 문자를 다수 개 배치한 방식임을 알 수 있다[11].

3. 필획기반 다국어 입력 시스템

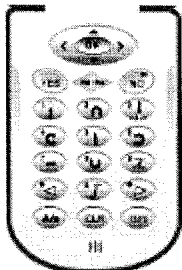
본 연구에서 개발한 다국어 문자 입력 시스템을 총칭하여 Effy 문자 입력 시스템이라 한다.

엔진부가 있으며, OS의 그래픽 API와 연동하여 출력을 담당하는 폰트 출력 엔진부가 있고, 키패드나 터치 스크린의 하드웨어와 연동되는 입력 하드웨어 제어부가 있다. 입력 엔진부, 출력 엔진부, 하드웨어 제어부를 객체로 모듈화 하여 시스템 문제 해결과 시스템 확장에 용이하도록 구성하였다(그림 6 참조).

3.2 휴대폰용 Effy 문자 입력 시스템

(1) 영어 입력 시스템

영어 입력 시스템은 알파벳의 전신으로 알려져 있는 ‘페니키안비트’를 참조하여 개발했으며, 영어 소문자 알파벳을 분해한 입력 방식이다. 그림 7은 영어 입력 시스템의 사용자 인터페이스를 보여 주고 있다.



(그림 7) 영어 입력 시스템 UI

'c'-'1' = a	'i'-'4' = k	'7'-'5' = t
'1'-'8' = b	'i'-'1' = l	'7'-'3' = u
'c'-'4' = c	'i'-'3' = m	'7'-'1' = v
'c'-'1' = d	'i'-'1' = n	'1'-'7' = w
'7'-'4' = e	'n'-'n' = o	'3'-'7' = x
'7'-'8' = f	'i'-'n' = p	'u'-'1' = y
'8'-'7' = g	'i'-'8' = q	'u'-'7' = z
'i'-'n' = h	'c'-'1' = r	
'2'-'1' = i	'i'-'8' = s	
'1'-'2' = j	'c'-'8' = t	
'8'-'1' = k	'c'-'1' = u	
	'u'-'1' = v	
	'u'-'7' = w	
	'2'-'2' = x	
	'2'-'1' = y	
	'2'-'2' = z	

(그림8) 영어 입력 오토마타

입력 방식은 입력하고자 하는 알파벳의 소문자를 연상하면서 두 개의 도형으로 하나의 알파벳을 조합하는 방식이다(그림8참조). 표 1에서 영어 입력 시스템을 이용하여 ‘time’을 입력하는 과정을 볼 수 있다. 영어 입력 시스템은 하프알파벳 기반 필획 모양을 조합하여 영어를 입력할 수 있도록 구성되어 있고, 빈도순이 정렬되어 있는 DB를 제공하여 영어 단어의 입력을 용이하게 만들었으며, 한번 입력한 단어는 히스토리를 버퍼에 저장하여 재입력을 용이하게 할 수 있도록 제작하였다.

(표 1) ‘time’ 입력 과정

예
“time”
입력 과정
1단계 : “7.5”를 입력해서 ‘t’를 입력합니다. “5.7” 등의 조합으로 지원하고 있습니다.
2단계 : “9.5”를 입력해서 ‘i’를 입력합니다.
3단계 : “2.2”를 입력해서 ‘m’을

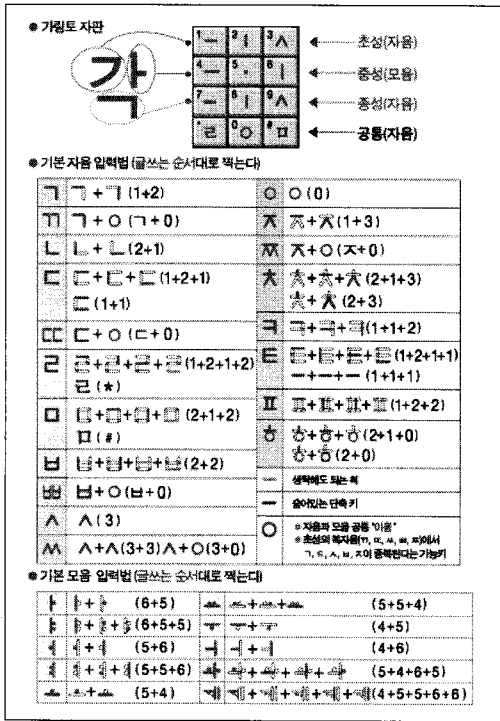
(표 2) 영어 입력기의 주요 클래스

클래스 명	주요 기능
TSWidget(English)	InputMethod 인터페이스
TSEngParser	가림토 필획 입력 파싱
TSGesture	제스처 처리
TSEngUI	버튼/UI 정보 처리

(2) 한글 입력 시스템

한글 입력 시스템은 한글의 전신인 가림토를 본떠서 개발했으며, 자음과 모음을 필획(Stroke)으로 분해한 것이 특징이다. 그림 9에서와 같이 자음은 ‘123’번 버튼에 있는 모양의 필획으로 조합하며, 모음은 ‘456’ 버튼에 있는 것처럼 ‘천지인’ 방식으로 조합한다. 이 때, 한글 입력은 3별식 체계를 따르고 있으며, ‘123’은 초성의 자음

입력 시에 사용하고, ‘456’은 중성의 모음, ‘789’는 중성의 자음 입력 시에 이용한다. 하단의 ‘ㄹㅇㅇ’은 초성과 중성의 자음 입력 시에 부수적으로 이용하여 입력할 수 있다. 또한, ‘ㅇ’은 기능 키로서의 의미를 가지며, ‘ㄱ, ㄷ’ 등의 쌍자음을 입력할 때 사용한다.



(그림 9) 한글 입력 오토마타

입력 방식은 입력하고자 하는 글자의 자음과 모음을 초/중/중성의 위치에 맞게 글쓰는 순서대로 필획을 조합한다. 표 3은 ‘한글’이라는 단어를 입력할 때의 과정을 보여 주고 있다. 또한, Effy 한글 입력 시스템은 미리 내장되어 있는 사전을 활용하여, 문장을 작성할 때 입력의 편의성을 높여 준다. 즉, 사용자가 ‘사’를 입력하면, 내장되어 있는 사전 내용인 ‘사랑해’를 출력함으로써 끝까지 입력하지 않고서도 후보 군을 활용할 수 있도록 한다.

(표 3) ‘한글’ 입력 과정

예제 1 ‘한글’ 입력 순	
입력 과정	
1단계 :	“2,1,0”을 입력해서 초성의 ‘ㅎ’를 입력합니다. 단축키로 ‘2,0’만 입력해도 됩니다.
2단계 :	“6,5”를 입력해서 중성의 ‘ㅏ’를 입력합니다.
3단계 :	“8,7”을 입력해서 중성의 ‘ㄴ’을 입력합니다.

한글 입력 시스템은 기림도 기반 필획 모양을 조합하여 한글을 입력할 수 있도록 구성되어 있고, 입력한 한글을 한자로 변환할 수 있는 기능, 빈도순이 높은 DB를 제공하여 한글 단어의 입력을 용이하게 만들었다. 또한 한번 입력한 단어는 히스토리를 버퍼에 저장하게 해서 재입력을 용이하게 할 수 있도록 개발했다.

(표 4) 한글 입력기 주요 클래스

클래스 명	주요 기능
TSWidget(Korean)	InputMethod 인터페이스
TSKorParser	기림도 필획 입력 파싱
TSKorHanja	한자 변환 클래스
TSGesture	제스처 처리
TSKorUI	버튼/UI 정보 처리

(3) 일본어 입력 시스템

일본어 입력 시스템은 히라가나를 분해하여 개발된 입력 시스템으로서, 각 히라가나 문자의 특성을 분석해서 12개의 도형으로 분해했다. 입력 방식은 입력하고자 하는 문자를 연상하면서 글쓰는 순서대로 2개의 도형을 조합하는 방식이다. 이때, 입력된 히라가나/가타가나는 내장된 한자 DB를 활용하여 한자로 변환해서 입력할 수 있다.



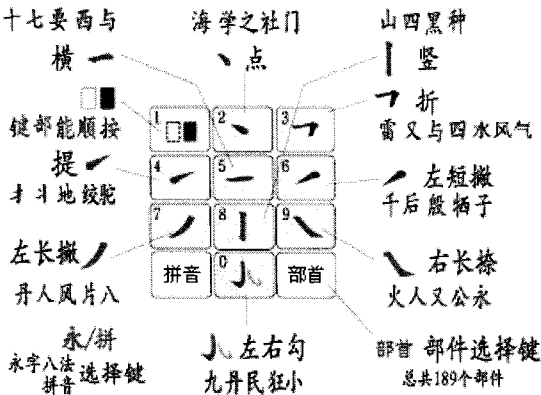
(그림 10) 일본어 입력 오토마타

(표 5) 일본어 입력기 주요 클래스

클래스 명	주요 기능
TSWidget(Japanese)	InputMethod 인터페이스
TSjpKanaParser	반자타법 입력 파싱
TSjpRomaParser	로마가나 입력 파싱
TSjptTrans	한자 변환
TSjpTransCompo	입력중인 문장 처리
TSjpTransCandi	한자 변환 리스트 출력
TSGesture	제스처 처리
TSjpUI	버튼/UI 정보 처리

(4) 중국어 입력 시스템

중국어 입력 시스템은 진나라 서예가인 왕희지의 “永 안에 한자의 모든 필획이 담겨져 있다.”라는 주장을 기반으로 해서 개발되었으며, ‘永’자의 필획을 분해했다.



(그림 11) 중국어 입력 오토마타

필획 입력 방식은 키보드상의 필획을 글쓰는 순서대로 조합하여 한자를 입력하는 방식이다. 미리 구축되어 있는 DB를 활용함으로써 필획 전체를 입력할 필요없이 원하는 한자를 빠르게 입력할 수 있다. 필획을 입력하는 도중에 1획이 추가될 때마다, 후보한자 출력 영역에는 해당 필획으로 시작하는 전체 한자중 빈도순이 높은 순서대로 후보한자를 출력해 준다. 후보 영역에서 입력하고자 하는 한자가 출력되면, 나머지 필획을 입력할 필요없이 바로 확정할 수 있다. 표 6은 필획 입력 방식을 활용하여 ‘你好’를 입력하는 과정을 보여준다.

(표 6) 필획 입력 방식 예제

예제 1 “你好” 입력 과정	
1단계 :	‘6’ 버튼을 선택합니다. 이 필획으로 시작하는 후보한자 출력됩니다.
2단계 :	‘8’, ‘6’ 버튼을 선택합니다. 출력된 후보한자에서 彡를 선택합니다.
3단계 :	彡와 관련되서 사용되는 한자 후보군에서 好를 선택합니다.

영자판법 기반 필획 모양을 조합하여 중국어를 입력할 수 있도록 구성되어 있으며, 빈도순으로 DB가 구성되어 있어서, 필획 입력도중에도 원하는 글자를 입력할 수 있다. 또한, 입력한 한자에 대해서 관련 글자를 제공하여, 보다 쉽고 빠르게 원하는 단어를 입력할 수 있도록 제작하였다.

(표 7) 중국어 입력기 주요 클래스

클래스 명	주요 기능
TSWidget(Chinese)	InputMethod 인터페이스
TSHanZiParser	영자판법 필획 입력 파싱
TSCNsCompo	입력중인 필획 처리
TSCNsCandi	후보군, 관련글자 출력
TSGesture	제스처 처리
TSCNsUI	버튼/UI 정보 처리

3.3. 입력 시스템의 성능 비교

본 입력 시스템은 한 버튼에 여러 개의 알파벳을 배치하는 방식이 아닌, 해당 언어의 생성 원리를 분석하여 도형을 분리함으로써 직관적이면서 보다 편리하고 빠르게 입력할 수 있는 방식임을 타수 비교 및 타수별 분포도를 통해서 알 수 있다. 표 8에서 대표적으로 Effy 중국어 입력 시스템과 기존 입력법과 비교했을 경우, 7.43% ~ 14.73% 정도 빠르게 입력할 수 있었음을 보여주고 있다.

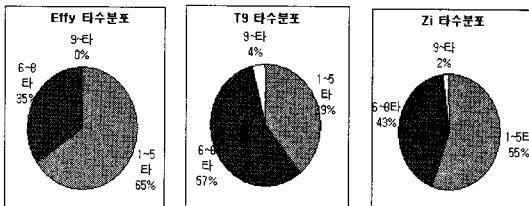
(표 8) 중국어 필획 입력 타수 비교

빈도순	한자	Effy	T9	Zi
1	的	1	1	2
2999	漣	7	8	8
3000	稔	6	8	6
합계		15236	17868	16459
평균타수		5.08	5.9%	5.49
타수감소율			-14.73%	-7.43%

또한 표9와 그림12의 분포도를 보면 다른 입력 시스템에 비해서 5타 이하로 문자를 입력할 수 있는 비율이 상대적으로 높다는 결과를 알 수 있다.

(표 9) 중국어 필획 입력 타수별 분포도

입력타수	Effy	T9	Zi
1~5타	1953	1160	1636
6~8타	1042	1715	1298
9~ 타	5	125	66



(그림 12) 중국어 필획 입력 타수별 분포도

4. 결론

휴대폰의 경우 차세대 서비스의 대부분은 기본적으로 문자 입력을 필요로 하고 있다. 그러나 기존 입력 시스템은 자판 배열을 포함해서 인터페이스의 인지도가 떨어지며, 입력 속도 또한 느린 형편이다. 본 논문에서 제시한 다국어 문자 입력 시스템은 언어의 생성원리를 기반으로 필획(Stroke) 방식을 채택함으로써 기존 입력 시스템에 비해 보다 사용하기 편리하고, 인지하기 쉬우며, 빠르게 원하는 문자를 입력할 수 있다. 따라서, 현재 문제가 제기되고 있는 핸드폰 단말기에서의 문자 입력에 대한 불편 해소 및 사용자 편의성 향상에 기여할 수 있다. 향후에는 본 연구의 결과물을 응용한 다양한 플랫폼을 기반으로 개발에 주력할 예정이다. 또한, 전세계 언어의 동시지원 문제도 향후 진행되어야 할 과제라고 할 수 있다.

Reference

- [1] <http://www.worldlanguage.com/>
- [2] <http://pitecan.com/bib/InputSystem.html>
- [3] <http://www.embeddedtechnology.com>
- [4] <http://www.embedded.com/>
- [5] <http://www.sec.co.kr/>
- [6] <http://www.lge.co.kr/>
- [7] <http://www.nacotech.com>
- [8] <http://www.neopad.com>
- [9] <http://www.t9.com>
- [10] <http://www.zicorp.com>
- [11] <http://www.mobilephone.co.jp/>

● 저 자 소 개 ●



이 진 영(Jin-Yeong Lee)

1999년 대전대학교 컴퓨터공학과 졸업(공학학사)
2001년 중앙대학교 대학원 컴퓨터공학과 졸업(공학석사)
2004년 중앙대학교 대학원 컴퓨터공학과 (공학박사 수료)
2001년~현재 타임스페이스시스템(주) 기술연구소 부소장, 중앙대학교 컴퓨터공학과 박사과정
관심분야 : 임베디드 시스템, 다국어처리, 문자 입력, 분산 시스템
E-mail : leejin0@empal.com



홍 성 용(Sung-Ryong. Hong)

1987년 홍익대학교 경영학 졸업(학사)
1986년~1997년 (주)진주 영업지점장
1997년~2000년 대상농장(주) 영업팀장
2001년~현재 : 타임스페이스시스템(주) 대표이사
<관심분야> 다국어처리, 문자입력, 다국어 콘텐츠, IT
E-mail : hsr@timespace21.com



이 시 진(Si-Jin Lee)

1990 중앙대학교 전자계산학과 (이학사)
1992 중앙대학교 대학원 컴퓨터공학과 (공학석사)
1997 중앙대학교 대학원 컴퓨터공학과 (공학박사)
1997 ~ 현재 대전대학교 컴퓨터공학과 부교수
관심분야 : 분산 시스템 및 시스템 소프트웨어, 리눅스 시스템, 유비쿼터스 컴퓨팅, 임베디드 시스템 등
E-mail : sjlee@daejin.ac.kr