

데이터베이스에 근거한 자동 키보드의 입력 방법

이계숙[†], 옹환승^{**}

요 약

최근 들어 하드웨어 키보드가 제한적인 모바일 단말기에서 소프트웨어 키보드가 광범위하게 사용되고 있다. 본 연구에서는 데이터베이스 분석을 통해 최적의 키보드 공간을 생성하여 최소의 입력 단계로 원하는 입력이 가능한 새로운 소프트웨어 키보드 입력 방법을 제시하였다. 제안된 소프트웨어 키보드는 입력 가능한 모든 단어를 분석하여 각각의 입력 단계에 따라 동적으로 생성된다. 이때 최소한의 키 버튼만을 도시함으로써 키 버튼 선택 영역을 최소화하고 키 입력 실수를 방지한다. 또한 입력 가능한 단어들의 수가 제한적일 때 후입력 자동 완성 기능을 제공함으로써 키 버튼 선택 횟수를 최소화할 수 있다.

Research on the Automatic Software Keyboard Based on Database

Kye-Suk, Lee[†], Hwan-Seung, Yong^{**}

ABSTRACT

Recently software keyboard is widely used in mobile devices where restrictive hardware keyboard is available. In this paper, new software-driven keyboard input method is proposed, which use minimum number of keyboard input with small keyboard space generated after analyzing of database. In this software keyboard is generated dynamically at each input step by analyzing all possible input words. Software keyboard, only possible key buttons are displayed for minimizing keyboard space and preventing mistyping. And it also provide input word completion function when the number of the candidate words is within threshold scope.

Key words: Software Keyboard(소프트웨어 키보드), Automatic Keyboard(자동 키보드)

1. 서 론

컴퓨터 주변기기 중 성능과 기능적으로 큰 변화 없이 개선되지 않은 것을 꼽으라 하면 단연코 키보드 일 것이다. CPU, 그래픽 카드 등이 보다 빠른 속도와 성능 향상을 위해 꾸준히 개선되고 발전되어 온 반면 키보드의 경우에는 특별히 성능이나 기능이 개선되

거나 발전하지 않았다[2].

더구나 한글의 특성상 한글 자판의 구성에 관해서는 계속 해결해야 할 문제로 남아 있다. 영어를 포함한 라틴 계열의 문자는 하나의 자소도 호형으로 표현 될 때 일정한 크기로 쓰여지는 반면 한글은 2개 또는 3개, 경우에 따라서는 5개까지의 자소들로 구성된다. 즉 영어를 기계화 할 때는 입력, 출력, 처리의 단위가 모두 같아 컴퓨터에서 사용하는 코드는 자소 단위로 코드를 부여하면 된다. 그러나 한글은 자소 단위로 입력하고 음절 단위로 모아서 출력하도록 되어 있다. 결국 한글 코드는 자소 단위와 음절 단위의 의미를 동시에 갖고 있어야 입력력 및 처리가 효율적인 코드가 된다[1].

※ 교신저자(Corresponding Author): 옹환승, 주소: 서울시 서대문구 대현동(120-750), 전화: 02-3277-2592, FAX: 02-3277-2306, E-mail: hsyong@ewha.ac.kr

접수일: 2004년 2월 9일, 완료일: 2004년 6월 8일

[†] 용인송담대학 정보통신과 겸임교수

(E-mail: gyesuk71@hanmail.net)

^{**} 종신회원, 이화여자대학교 컴퓨터학과 부교수

최근의 키보드는 다양한 종류가 출시되고 있다. 그것은 무엇보다도 PDA와 휴대폰의 대중화 때문이다. PDA, 휴대폰과 같은 기기는 휴대하면서 사용하고 크기가 작기 때문에 기존에 일반 컴퓨터에서 사용하던 키보드를 그대로 사용할 수 없다. 그래서 이들 기기는 자체적으로 문자를 입력할 수 있는 시스템이 내장되어 있다. 그러나 기기의 크기가 작기 때문에 내장된 입력 시스템이 효율적이지 않은 것이 사실이다[2].

현대 정보화 사회에서 처리해야 할 정보의 양이 기하급수적으로 증가함에 따라서 필요한 정보를 효과적으로 입력 및 출력하고 분석 처리하는 능력이 요구되고 있다[4]. 더구나 정보 처리의 속도는 CPU 속도가 아닌 사용자의 입력 작업 속도에 달려 있다. 많은 양의 정보를 빠른 시간 내에 입력하기 위해서는 인체 공학적인 디자인 측면과 키보드를 두드릴 때의 매커니즘에 대한 개선이 이루어지는 하드웨어 키보드에 대한 연구뿐만 아니라 소형화 되어가는 정보 기기에 적합하게 최소의 키 버튼을 제공하고, 키 버튼 선택 횟수를 줄일 수 있는 지능적인 소프트웨어 키보드에 대한 연구가 필요하다.

본 논문은 하드웨어 키보드를 사용할 수 없는 정보 기기에서 한글을 입력할 때 기존의 PC키보드를 그대로 화면에 표시하는 소프트웨어 키보드가 가지는 키 버튼 선택 영역면과 키 선택 횟수면에서의 단점을 보완하기 위하여 첫째, 키 버튼 선택 영역을 고려하여 입력하고자 하는 단어의 집합이 주어졌을 때 주어진 단어 입력에 필요한 최소한의 키 버튼만을 표시하는 기법을 제안한다. 둘째, 단어를 입력함에 있어서의 신속성을 위해 키 버튼 선택 횟수를 최소화하는 기법을 제안한다. 주어진 단어 집합을 입력할 때 필요한 키 버튼 수의 감소로 인한 키 버튼 선택 영역의 축소와 키 버튼 선택 횟수의 감소로 인한 키 입력 시간 단축의 효과를 기대할 수 있다. 또한 본 논문에서 제안하는 입력 방식의 효율성을 검증하기 위하여 2000개의 샘플 단어 집합을 추출하고 단어 집합의 크기를 변경하면서 기존의 입력 방식과 본 논문에서 제안하는 데이터베이스에 근거한 자동 키보드 알고리즘의 성능을 평가하였다.

본 논문의 구성은 다음과 같다. 제2장은 한글 코드에 관한 기존의 연구와 여러 가지 정보 기기에서의 한글 입력 방식에 대해 기술한다. 제3장에서는 주어진 단어 집합을 효율적으로 입력할 수 있는 데이터베

이스에 근거한 자동 키보드 방식을 제안하고, 제4장에서는 제안한 입력 방식에 대한 성능 평가를 수행하며, 제5장에서 결론을 맺는다.

2. 모바일 기기에서의 한글 입력 방식과 구조적 키보드에 관한 연구

본 장에서는 모바일 정보 통신 기기에서의 한글 입력 방식에 대해 살펴보고, 일반 컴퓨터에서 사용하는 하드웨어 키보드 형태가 아닌 구조적 가상 키보드에 관한 연구에 대해 고찰한다.

2.1 모바일 기기에서의 한글 입력 방식

최근 정보 통신 기술과 무선 인터넷 기술의 발달로 휴대폰과 PDA에서의 한글 입력이 급속히 늘어나고 있다. 휴대폰과 PDA와 같은 모바일 정보 단말기에서 한글 입력 방식은 전화기 자판을 통한 입력 방식과 화면 키보드 방식, 그리고 펜을 통한 입력 방식이 대표적이다.

전화기 자판을 통한 입력 방식은 휴대폰에 있는 0~9까지의 10개의 숫자 버튼과 2개의 특수 버튼(*, #)을 포함하는 12개의 버튼을 이용해서 쉽고 빠르게 한글을 입력할 수 있는 방식이다. 이 12개의 버튼으로 67개의 한글 자모를 입력하기 위해서는 동일한 버튼에 여러 개의 자음과 모음을 배정해야 한다. 이때 쌍자음과 격자음과 같이 가족 관계에 있는 자모를 하나의 버튼에 배치하는 게 일반적이다. 예를 들면, 「ㄱ, ㄱ, ㅋ」과 같이 가족 관계에 있는 자모를 하나의 버튼을 통해 입력한다. 이와 같은 대표적인 방식은 삼성전자의 천지인 방식과 LG전자의 나랏글2000이 있다. 천지인 방식은 단자음을 세 번 눌러서 쌍자음을 입력하고, 격자음은 두 번 눌러서 입력한다. 중성은 「·, -, |」를 합성해서 입력하고 다음 글자를 입력하기 위해서는 2초동안 기다려야 한다. 나랏글2000의 경우는 쌍자음 변형을 위해서는 「#」 버튼을, 격자음 변형을 위해서는 「*」 버튼을 이용한다. 모음을 입력하는 방식은 「ㅏ ㅑ」 「ㅓ ㅕ」 「ㅡ ㅣ」를 각각 동일한 버튼에 배정하고 복모음은 「*」 버튼을 이용해 입력한다[8].

가족 관계에 있는 자모를 하나의 버튼에 배치하지 않고 하나의 버튼에 여러 개의 자음과 모음을 같이 배치하여 입력하는 방식으로는 노키아의 문자 예측

입력 기능과 엔엘피 솔루션의 소프트키패드 방식이 있는데 자모를 선택하기 위해 동일한 키를 여러 번 누르는 것 대신에 한 음절을 입력하고 나서 후보들 중 선택해야 하거나 입력 모드 변환을 위해 별도의 키 입력이 필요하므로 키 입력 횟수를 감소시킨다고 볼 수 없고, 12개의 기본 버튼외에 추가적인 키 버튼이 필요하다.

화면 키보드는 일반 PC에서 사용하는 키보드와 동일한 키보드를 화면에 나타내어 화면의 키보드를 치면 실제 키보드를 친 것과 같은 효과를 내주는 소프트웨어 키보드 방식이고, 펜을 통한 입력 방식은 LCD와 입력 장치인 태블릿(Tablet)을 결합한 디스플레이 장치가 있어서 태블릿에 펜으로 글씨를 쓰면 태블릿에서 펜의 위치를 인식한 다음 운영 체제가 펜의 움직인 자취를 화면에 보이고, 동시에 이 자취를 가지고 필기된 글씨를 인식하여 문자로 표시하는 방식이다[9]. 문자 인식 과정은 입력 과정에서 발생한 시간적, 공간적 정보와 획순, 획수, 필기 방향이나 속도 등의 부가적인 정보로 입력 데이터에서 특징을 추출하고, 미리 정해진 여러 모델 중의 한 모델로 분류하는 과정으로 구성된다.

2.2 구조적 가상 키보드에 관한 연구

일반적인 PC 키보드를 화면에 표현할 경우 한글 입력을 위한 키는 Shift 키를 눌러야 하는 쌍자음(ㄱ, ㅋ, ㆁ, …)과 복자음(ㄷ, ㅌ, …), 이중 모음을 포함해서 33개이다. 이 33개의 키를 모두 버튼으로 만들어 화면에 표시한다면 입력 창의 크기가 최소화되어야 할 경우 상대적으로 버튼의 크기가 작아질 수 밖에 없다. 이러한 단점을 해결하기 위해 제안한 소프트웨어 키보드가 구조적 가상 키보드이다[7].

구조적 가상 키보드는 입력 시간의 단축을 위해 초성, 중성, 종성용 키보드를 분리하는 구조형을 지향하고 초성, 중성, 종성의 입력 창이 자동 전환된다. 즉, 초성을 입력하면 중성 입력창이 활성화되고 중성을 입력하고 나면 종성 입력창이 활성화 된다. 따라서 전체 선택 영역을 축소시켜 자소 선택의 간편성과 신속성을 높여 준다[7].

이 기법에서는 초성과 종성용 키보드를 별개의 키보드 입력 창으로 분리시켰고, 쌍자음과 복자음, 이중 모음을 각각 다른 키 버튼으로 할당하였기 때문에 초성 키보드의 버튼 수는 19개, 중성 키보드는 21개,

중성 키보드는 21개이다. 따라서 일반 PC에서 사용하는 하드웨어 키보드의 한글 입력을 위한 키 버튼 수인 26개와 비교했을 때는 효율성이 그다지 높지 않으므로 키보드의 버튼 수를 줄일 수 있는 알고리즘이 필요하다.

3. 데이터베이스에 근거한 자동 키보드

본 장에서는 데이터베이스에 근거한 자동 키보드에 대해 정의하고, 기능을 소개하며, 이 방법을 통하여 주어진 단어들의 집합을 최소한의 키 집합으로 입력할 수 있는 방법과 키 선택 횟수를 최소화할 수 있는 방법을 제안한다.

3.1 데이터베이스에 근거한 자동 키보드 정의

데이터베이스에 근거한 자동 키보드란 모바일 기기와 같이 하드웨어 키보드나 키패드가 없는 단말기에서 데이터베이스나 파일 형태로 저장되어 있는 단어들의 집합을 입력할 때 필요한 키만을 보여 주는 지능형 소프트 키보드이다. 즉, 입력하고자 하는 단어의 초성에 해당하는 키를 선택하면 전체 단어들의 집합에서 선택된 초성 다음에 나올 수 있는 중성 키만을 보여 주게 된다. 그리고 중성 키를 선택하면 전체 단어들의 집합에서 지금까지 선택된 초성과 중성 다음에 올 수 있는 중성 키만을 보여 준다. 만약 중성이 없다면 다음 음절의 초성 키를 보여 준다.

만약 표 1과 같은 단어 집합이 주어진다면 초기에 보여지는 초성 키보드는 그림 3(a)와 같이 모든 단어의 첫 음절의 초성에 해당하는 자음을 중복 없이 도시한다. 이 상태에서 ‘ㅅ’ 키 버튼을 선택했을 때는 그림 3의 (b)와 같이 ‘ㅅ’을 초성으로 가진 단어의 중성에 해당하는 모음 키들이 도시 된다. 이때 ‘ㅓ’ 키를 선택하면 그림 3의 (c)와 같이 초성이 ‘ㅅ’이고 중성이 ‘ㅓ’인 음절의 종성에 해당하는 자음 키 버튼이 도시된다.

3.2 데이터베이스에 근거한 자동 키보드 설계

주어진 단어들의 집합을 효율적으로 입력하기 위한 자동 키보드의 설계 사양을 다음과 같이 정하였다.

- (1) 한글 입력에 있어서 기존의 하드웨어 키보드를 소프트웨어 키보드로 대체

표 1. 그림 1의 입력 단어 집합

삼도물산	ㅅ	ㅈ	ㅊ
경남모직우	ㄱ	ㅋ	ㅇ
인지컨트롤	ㅇ	ㅣ	ㄴ
삼성전자우	ㅅ	ㅈ	ㅊ
삼성증권우	ㅅ	ㅈ	ㅊ
경동보일러	ㄱ	ㅋ	ㅇ
엘지애드	ㅇ	ㅋ	ㄴ
센추리	ㅅ	ㅋ	ㄴ
성지건설	ㅅ	ㅋ	ㅇ
신풍제지	ㅅ	ㅣ	ㄴ

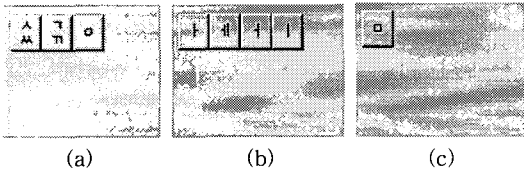


그림 1. 표 1과 같은 단어 집합이 주어졌을 때 키보드 변화 : (a) 초성 선택을 위한 키보드, (b) (a)상태에서 'ㅅ'을 선택했을 때 키보드, (c) (b)상태에서 'ㅈ'를 선택했을 때 키보드

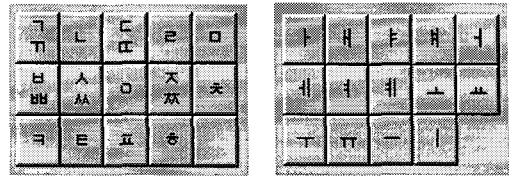
- (2) 초성과 중성 입력을 위한 자음 키보드와 중성 입력을 위한 모음 키보드로 분리
- (3) 주어진 단어들의 집합에 해당하는 자음과 모음 키 버튼만을 지능적으로 도시
- (4) 키 버튼 선택 횟수 최소화를 위한 후입력 자동 완성 기능

위와 같은 기능 구현을 위한 자음 키보드와 모음 키보드의 구성과 지능형 키 도시 기능, 후입력 자동 완성 기능은 다음과 같다.

3.2.1 자모의 구성

데이터베이스에 근거한 자동 키보드에서 초성과 중성 입력을 위해서는 그림 2의 (a)와 같은 자음 키보드를 사용한다. 일반 PC의 키보드와 개수와 비교할 때 빈칸 키가 추가되는데 이는 중성이 없는 단어와 중성이 있는 단어를 선택할 때 사용된다.

초성에는 이중 자음이 존재하지 않으므로 쌍자음을 입력할 경우는 더블 클릭과 같은 이벤트로 구현하여 단자음과 구분할 수 있다. 중성에는 쌍자음과 이중 자음이 모두 존재할 수 있으므로 쌍자음은 초성과 마찬가지로 방법으로 처리하고, 이중 자음의 경우에는



(a) 자음 키보드 (b) 모음 키보드

그림 2. 데이터베이스에 근거한 자동 키보드의 자모 키보드

하나의 자음 선택 후 다시 중성을 위한 키 버튼들이 도시될 때 연속적으로 선택한다.

중성 입력을 위해서는 그림 2의 (b)와 같은 키보드를 사용하는데 일반 PC 키보드와 비교해보면 'ㅈ'와 'ㅋ'키가 별개의 버튼으로 구분되었다. 이는 키 버튼 선택 횟수 최소화를 위한 방안으로 Shift 키 없이 선택이 가능하도록 하기 위함이다.

3.2.2 지능형 키 도시 기능

지능형 키 도시 기능은 사용자가 한 개의 자소를 입력할 때마다 초성/중성/중성의 분류에 따라 자음 키보드와 모음 키보드가 자동으로 전환되는 기능으로 다음과 같은 과정으로 처리된다.

- (1) 단어의 입력은 항상 초성 입력을 위한 자음 키보드부터 시작한다.
- (2) 자음 키보드로부터 초성을 선택하면 전체 단어 집합에서 선택된 초성만을 가진 단어들을 선택하여 중성에 해당하는 모음 키 버튼만을 도시한다.
- (3) 중성 입력을 위한 모음 키보드로부터 중성을 선택하면 이중 모음이 있는 단어들의 두 번째 모음 키 버튼과 단일 모음으로 구성된 단어들의 중성에 해당하는 자음 키 버튼이 동시에 도시 된다.
- (4) 중성까지 선택된 음절에 모든 단어들이 중성이 없다면 다음 음절의 초성에 해당하는 자음 키 버튼이 도시 되고 이 경우에는 자음 키보드의 빈칸 키가 필요 없다.
- (5) 중성이 없는 음절과 중성이 있는 음절이 같이 남아 있다면 자음 키보드 중 빈칸 키가 도시 되어 중성을 선택할 수 있도록 해준다.
- (6) 중성에 이중 자음이 있는 경우에는 첫번째 자음 키 버튼을 도시하고 단자음일 경우 해당 자음 키 버튼을 도시한다.
- (7) 중성 입력을 위해 자음 키 버튼을 선택했을 때, 단자음일 경우는 다음 음절의 초성에 해당하는

자음 키 버튼을, 이중 자음일 경우는 두번째 종성에 해당하는 자음 키 버튼을 도시한다.

3.2.3 후입력 자동 완성 기능

데이터베이스에 근거한 자동 키보드는 주어진 단어들의 집합을 효율적으로 입력하기 위한 소프트 키보드이므로 초성, 중성, 종성을 선택해 감에 따라 입력 가능한 단어들의 수는 감소하게 된다. 따라서 남아 있는 단어가 한 개일 경우는 자동으로 나머지 음절들을 채움으로써 키 선택 횟수를 줄일 수 있고, 화면에 공간적 제약이 없을 경우는 리스트 박스와 같은 컨트롤을 사용하여 입력 가능한 단어들을 나열하고 그 중 하나의 단어를 선택할 수 있다.

표 1을 예로 들어 설명하면 ‘ㅅ’과 ‘ㅈ’, ‘ㅊ’을 선택했을 때 입력 가능한 단어들은 ‘삼성전자우’와 ‘삼성증권우’ 두개로 한정된다. 이때 한번의 키 선택으로 두 단어 중 하나를 입력할 수 있고, 리스트 박스를 사용하여 입력 가능한 단어들이 한정적이 되었을 때 선택할 수 있다.

그림 3은 현재 등록된 1000개의 주식 거래소 이름을 샘플 집합으로 구성하고, 후입력 자동완성 리스트 박스의 크기가 10일때, ‘미래와 사람’이라는 주식 명을 입력하는 과정이다. ‘ㅊ’과 ‘ㅣ’를 선택하면 리스트 박스에 나머지 입력 가능한 단어들이 도시 되고, 여기서 선택하면 전체 3번의 선택으로 입력을 완성할 수 있다.

3.3 프로그램 순서도

데이터베이스에 근거한 자동 키보드의 처리 과정을 순서도를 통해 표현하면 그림 4와 같다

3.4 데이터베이스에 근거한 자동 키보드의 특징

본 논문에서 제안하는 주어진 단어 집합을 효율적

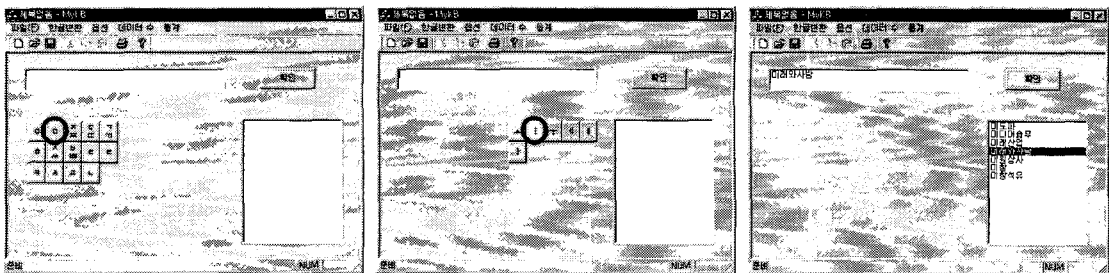
으로 입력하기 위한 데이터베이스에 근거한 자동 키보드의 특징은 다음과 같다.

(1) 기존의 하드웨어로 구현된 키보드나 현재 모바일 단말기에서 사용되고 있는 소프트 키보드는 한글 입력시 초성, 중성, 종성에 무관하게 항상 고정된 키 집합을 제공하는 반면 데이터베이스에 근거한 자동 키보드는 주어진 단어 집합을 분석하여 초성, 중성, 종성 입력 시 필요한 키 버튼만을 도시해 준다. 따라서 입력시 도시 되는 버튼 경우의 수를 크게 줄일 수 있다. 이는 모바일 단말기와 같이 화면이 제한적이고 하드 키보드가 없는 경우 유용하다.

(2) 주어진 단어 집합을 분석하여 입력할 수 있는 키 버튼만을 도시하므로 사용자의 입력 오류를 원천적으로 방지할 수 있다. 따라서 단어 집합을 사전으로 확장하면 단어의 맞춤법 오류를 입력 단계에서 원천적으로 예방할 수 있다.

(3) 데이터베이스에 근거한 자동 키보드는 입력하고자 하는 단어들의 집합이 주어지므로 소수의 음절만 입력하여도 입력 가능 상태인 단어들의 숫자가 현저히 줄어 든다. 따라서 화면의 제한이 있는 경우에는 입력 가능 상태인 단어가 한 개일 때 나머지 음절을 자동적으로 채워 주고, 화면의 제한이 없는 경우에는 입력 가능 상태인 단어가 충분히 줄었을 때 선택 리스트 박스를 통하여 선택하게 하면 단어 입력을 위한 키 선택 횟수를 최소화할 수 있다.

이와 같은 특징들을 바탕으로 구현된 데이터베이스에 근거한 자동 키보드가 이전 방식들과 비교했을 때 어느 정도의 성능 향상을 가져오는지 검증하기 위하여 데이터베이스에 근거한 자동 키보드 내에 성능 평가 모듈을 추가하였고, 다음 장에서는 여러 종류의 단어 집합에 대해 기존 입력 방식과의 비교를



(a) 초성 선택 화면

(b) (a)에서 ‘ㅊ’선택 후

(c) (b)에서 ‘ㅣ’선택 후

그림 3. 후입력 자동 완성 기능 예

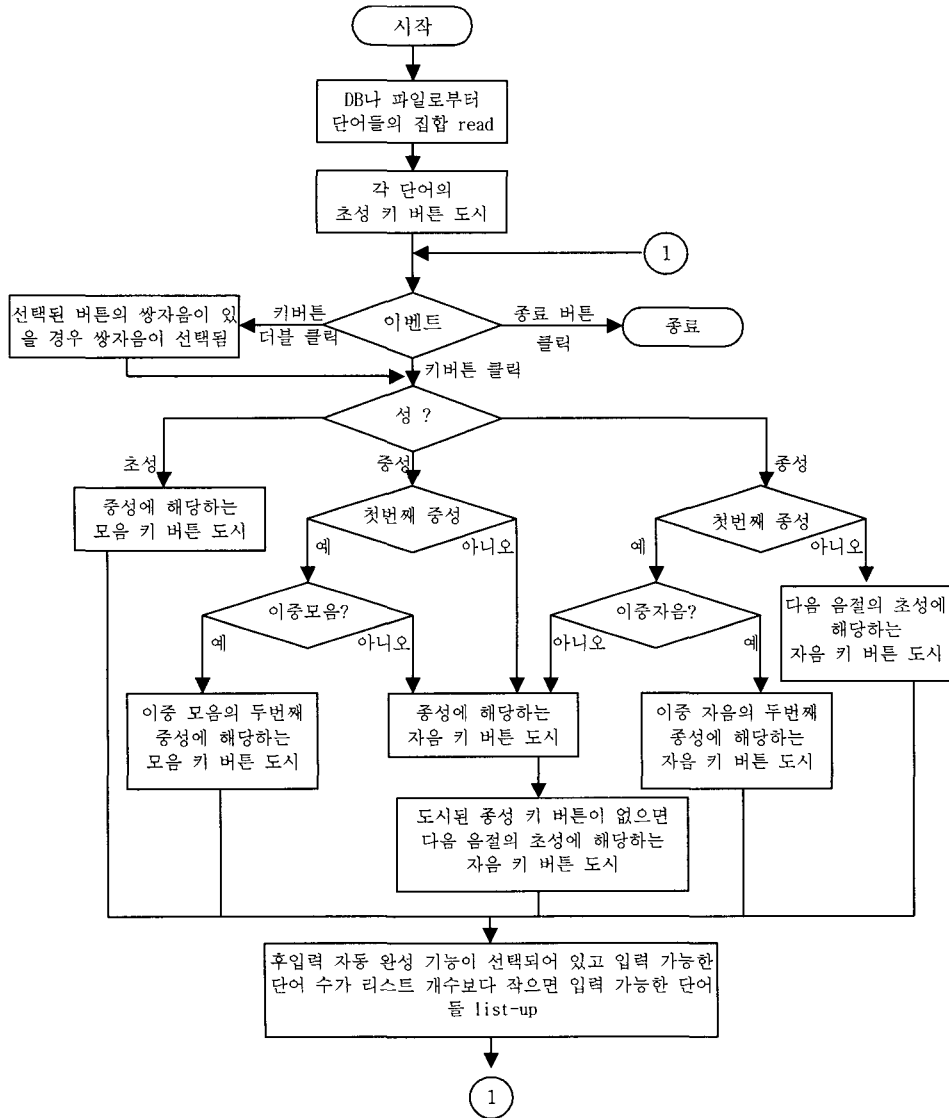


그림 4. 데이터베이스에 근거한 자동 키보드의 순서도

통해 본 논문에서 제안한 알고리즘의 성능을 평가해 본다.

4. 성능평가

본 장에서는 기존에 존재하는 일반 키보드 방식과 비교하여 다음과 같은 성능 평가를 수행한다. 첫째, 동일한 단어 집합에 대해 일반 키보드 방식과 데이터베이스에 근거한 자동 키보드 방식의 키 버튼 개수를 비교한다. 둘째, 동일한 단어 집합에 대하여 일반 키

보드 방식과 데이터베이스에 근거한 자동 키보드 방식의 키 선택 횟수를 비교한다. 이때 데이터베이스에 근거한 자동 키보드의 후입력 자동 완성 기능이 있을 때와 없을 때를 구분하여 살펴본다. 이를 통해 다양한 형태의 단어 집합이 주어질 때 어떤 방법이 가장 효율적인 입력 방법인지에 대한 기준을 제안한다.

4.1 샘플 단어 집합

성능 평가에 사용된 단어 집합은 최대 2000개이고, 1000개는 현 주식 시장의 거래소와 코스닥에 상

장된 주식 이름에서, 나머지 1000개는 야후 사전 (<http://kr.alldic.yahoo.com>)에 등록된 단어에서 발췌하였다. 성능 평가에 사용된 단어 집합의 크기는 100개, 500개, 1000개, 1500개, 2000개이고 단어의 최소 음절은 2음절, 최대 음절은 6 음절이다.

단어 집합의 크기를 2000개로 제한한 것은 본 논문에서 제시한 입력 방법이 제한된 크기의 단어 집합이 도메인으로 주어졌을 경우 유용한 입력 방법을 제안하는 것으로 상장 주식명이나 국가명, 부서명과 같이 계층적 구조를 가지지 않는 단어 집합의 경우 크기가 대부분 2000개 이내이기 때문이다.

4.2 일반 키보드 방식과 데이터베이스에 근거한 자동 키보드 방식의 키 버튼 개수 비교

PC에서 사용하는 일반 키보드 방식과 본 논문에서 제안하는 데이터베이스에 근거한 자동 키보드 방식의 키 버튼 개수면에서의 성능을 평가하기 위하여 각 단어 집합에 대해 입력에 필요한 키 버튼 수의 최대값과 최소값, 평균값으로 비교하였다. 그림 5는 단어 집합의 크기를 100개, 500개, 1000개, 1500개, 2000개의 다섯 가지 경우로 변경하면서 각 키보드 방식별 키 버튼 수의 최대값을 구한 결과이다. 이 때 후입력 자동 완성 리스트는 데이터베이스에 근거한 자동 키보드에 후입력 자동 완성 리스트 박스를 추가한 기능이고 리스트 박스의 크기는 10이다. 후입력 자동 완성은 데이터베이스에 근거한 자동 키보드에서 대응되는 단어가 한 개일 때 자동 완성시켜 주는 경우이다.

그림 5는 단어 입력을 위해 필요한 키 버튼 수의 최대값이 후입력 자동 완성 리스트 박스 기능을 가진 데이터베이스에 근거한 자동 키보드 방식이 가장 작게 나타나는 결과를 보여 준다. 또한 일반 키보드 방식은 단어 집합의 크기에 무관하게 항상 27개의 키 버튼이 필요하지만 데이터베이스에 근거한 자동 키보드 방식에서는 키 버튼의 최대값으로 비교해 봤을 때 500개 이하의 단어 집합에서 자동 키보드의 모든 방식이 효율적임을 나타내고 있다.

그림 6은 단어 집합 크기 변화에 따른 각 키보드 방식별 키 버튼 수의 최소값을 구한 결과이다.

그림 6에서 키 버튼 개수의 최소값의 경우 후입력 자동 완성과 후입력 자동 완성 리스트 기능이 있을 경우 첫 음절의 초성 선택 이후 별도의 키 버튼이

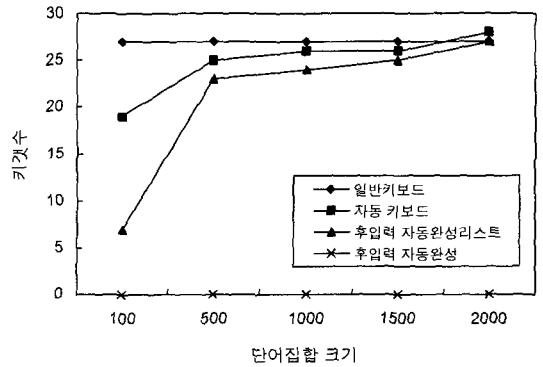


그림 5. 단어 집합 크기 변화에 따른 키 버튼 개수 최대값 비교

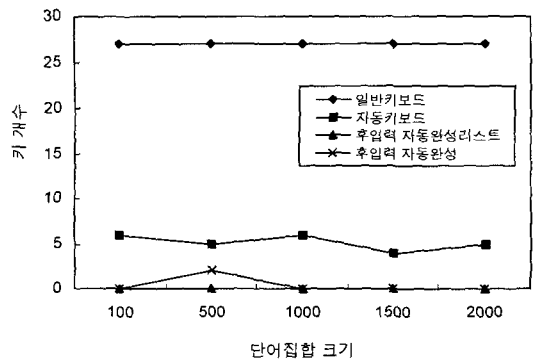


그림 6. 단어 집합 크기 변화에 따른 키 버튼 개수 최소값 비교

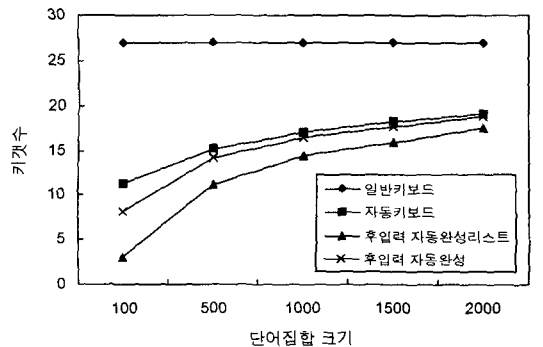


그림 7. 단어 집합 크기 변화에 따른 키 버튼 개수 평균값 비교

필요하지 않음을 보여 주고 있다.

그림 7은 단어 집합 크기 변화에 따른 각 키보드 방식별 키 버튼 수의 평균값을 구한 결과이다.

위의 그림 7은 단어 입력을 위해 필요한 키 버튼 수의 평균값이 단어 집합의 모든 크기에 대해서 일반 키보드보다 전반적으로 작게 나타나는 결과를 보여 준다. 일반 키보드 방식은 단어 집합의 크기에 무관하게 항상 27개의 키 버튼이 필요하지만 데이터베이스

스에 근거한 자동 키보드의 경우는 2000개의 단어 집합에 대해서도 20개 미만의 키 버튼만이 필요함을 나타낸다. 또한 자동 키보드의 모든 방식이 단어 집합의 크기가 증가하면서 키 버튼 수의 증가 비율이 점차 줄어드는 결과를 보인다.

4.3 일반 키보드 방식과 데이터베이스에 근거한 자동 키보드 방식의 키 선택 횟수 비교

이 절에서는 일반 키보드 방식과 데이터베이스에 근거한 자동 키보드 방식의 키 버튼 선택 횟수의 성능 평가를 위하여 각 단어의 입력에 필요한 키 선택 횟수의 최대값과 최소값, 평균값으로 비교하였다. 데이터베이스에 근거한 자동 키보드 방식 중 후입력 자동 완성 기능이 없는 경우는 일반 키보드와 키 선택 횟수에 있어서 동일하므로 리스트 박스의 크기가 10인 후입력 자동 완성 리스트 박스 방식과 후입력 자동 완성 방식으로만 성능 평가를 하였다. 그림 8은 단어 집합의 크기를 변경하면서 단어 입력 시 필요한 키 선택 횟수의 최대값을 구한 결과이다.

그림 8은 단어 입력을 위해 필요한 키 선택 횟수의 최대값이 후입력 자동 완성 리스트 기능을 가진 데이터베이스에 근거한 자동 키보드의 경우가 가장 작게 나타나는 결과를 보여 준다. 그리고 세 가지 입력 방식 모두 500개 이상의 단어 집합에 대해서는 키 선택 횟수의 최대값이 수렴됨을 나타내고 있다. 후입력 자동 완성 리스트 방식이나 후입력 자동 완성 방식에서 선택 횟수가 최대값으로 나오는 경우는 마지막 한 음절이 다른 단어가 존재할 때이다. 예를 들어 '가나다라'와 '가나다라마'와 같이 마지막 한 음절만 다른

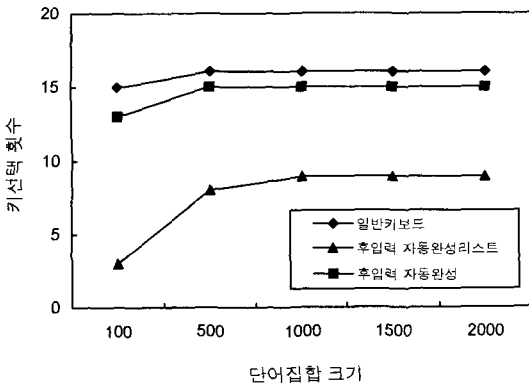


그림 8. 단어 집합 크기 변화에 따른 키 선택 횟수 최대값 비교

단어가 존재한다면 이 경우가 키 선택 횟수의 최대값을 가지는 경우이다.

그림 9는 단어 집합의 크기를 변경하면서 단어 입력 시 필요한 키 선택 횟수의 최소값을 구한 결과이다.

그림 9는 단어 입력 시 필요한 키 선택 횟수의 최소값이 샘플 단어의 음절이 최소 크기인 2음절이고 중성이 없는 경우 일반 키보드의 경우 4번의 선택이 필요하고 후입력 자동 완성 기능이 있을 경우는 한 번이나 두 번의 선택만으로 가능함을 보여주고 있다.

그림 10은 단어 입력 시 필요한 키 선택 횟수의 평균값으로 자동 키보드 입력 방식의 성능을 평가해 본 결과이다. 키 선택 횟수의 평균값으로 비교했을 때 후입력 자동 완성 리스트 기능이 키 선택 횟수가 가장 적는데 일반 키보드 입력 방식에 비해 50% 정도 선택 횟수가 감소하였다. 후입력 자동 완성 방식은 일반 키보드 방식에 비해 25% 정도 키 선택 횟수가 감소됨을 알 수 있다. 단어 집합의 크기가 500과 1000일 때 세 가지 방식 모두 키 선택 횟수가 가장 크게 나타나는데 이는 성능 평가에 사용된 단어 집합에서 단어 집합의 크기가 500이거나 1000일 때 단어

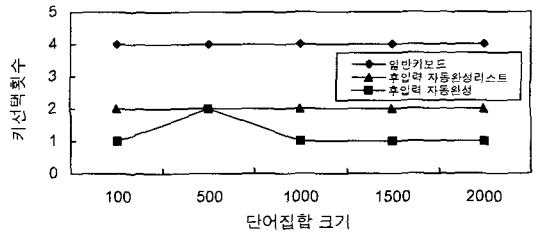


그림 9. 단어 집합 크기 변화에 따른 키 선택 횟수 최소값 비교

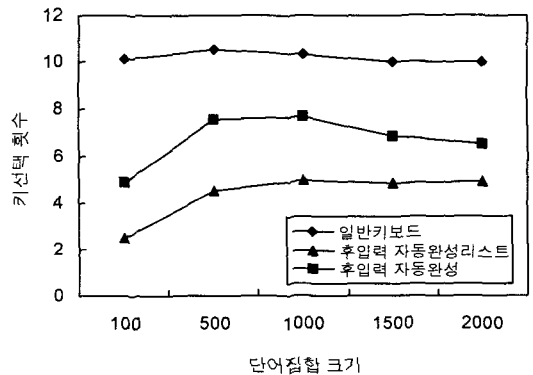


그림 10. 단어 집합 크기 변화에 따른 키 선택 횟수 평균값 비교

의 길이가 4음절이거나 5음절인 단어가 차지하는 비율이 상대적으로 높기 때문이다.

본 논문에서 수행한 성능 평가의 수행 결과, 데이터베이스에 근거한 자동 키보드 방식이 입력 시 필요한 키 버튼 개수면에서나 키 선택 횟수면에서 일반 PC 키보드 입력 방식보다 효율적임을 나타내었다. 키 버튼 개수면에서는 데이터베이스에 근거한 자동 키보드의 모든 방식이 일반 키보드 방식과 비교했을 때 높은 성능을 나타내었고, 키 선택 횟수면에서는 데이터베이스에 근거한 자동 키보드 방식 중 후입력 자동 완성 리스트 박스 방식이 가장 높은 성능을 나타내었다.

이러한 결과를 바탕으로 다양한 단어의 집합의 크기와 입력 창 크기의 제한 여부에 따라 지능형인 소프트웨어 키보드 입력 방식을 선택적으로 수행할 수 있다. 입력 창 크기에 제한이 클 경우는 대응되는 단어가 하나일 때만 자동 선택을 실행하는 후음절 완성 기능이 적합하고 그렇지 않을 경우는 후음절 완성 리스트 박스의 크기를 조절해 가면서 효율적인 입력 방식을 선택할 수 있다.

또한 단어 입력시 필요한 키 버튼 개수나 키 선택 횟수면에서 단어 집합의 크기가 500개 이하일 때 높은 성능을 나타내었다. 계층적으로 구성되지 않는 단어 집합의 경우 예를 들면 상장된 주식명이나 국가명, 학과명, 부서명, 주소의 동명등은 대부분 500개 이내의 크기를 가지므로 이와 같은 경우에 적용하면 실용적이다.

5. 결 론

최근 무선 인터넷과 정보 통신 기술의 발달로 정보 기기와 모바일 단말기의 크기가 점차 소형화 되어 감에 따라 부피가 크고 이동성에 문제가 있는 하드웨어 키보드를 그대로 사용할 수 없는 문제점이 있다. 본 논문에서는 모바일 기기와 같이 하드웨어 키보드가 없는 단말기에서 기존의 하드웨어 키보드 대신에 사용할 수 있는 소프트웨어 키보드 형태에 관해 살펴 보았고, 기존의 방법들이 단어 입력 창의 크기면에서나 입력 속도면에서 많은 단점을 가지고 있음을 지적 하였다.

본 논문에서는 소프트웨어 키보드를 구현할 때 주어진 단어 집합을 최소한의 키 버튼과 최소한의 키 선택 횟수를 통해 입력할 수 있는 데이터베이스에

근거한 자동 키보드를 제안하였다. 데이터베이스에 근거한 자동 키보드에서 키 버튼 개수를 줄이는 방법으로 초성과 중성 입력을 위한 자음 키보드와 중성 입력을 위한 모음 키보드를 가지는 구조적 키보드 형태를 제안하였는데 자음과 모음 키보드는 쌍자음과 이중 자음, 이중 모음을 위한 별개의 키 버튼이 존재하는 것이 아니라 단자음이나 단모음 버튼을 통해 입력하므로 키 버튼이 최대한 사용되더라도 일반 하드웨어 키보드의 키 버튼 수를 넘지 않는다. 또한, 초성을 선택하면 입력한 초성 다음에 나올 수 있는 중성을 선택하기 위한 모음 키보드로, 중성을 입력하면 중성이나 다음 음절의 초성을 선택할 수 있는 자음 키보드로 자동 전환된다. 데이터베이스에 근거한 자동 키보드에서 키 버튼 선택 횟수를 최소화하는 방법으로 후입력 자동 완성 리스트 박스 기능을 제안하였는데 단어의 초성, 중성, 종성 버튼 선택시마다 입력 가능한 단어들의 개수를 세어서 단어의 개수가 후입력 자동 완성 리스트 박스의 크기 이하가 될 때 리스트 박스에 단어들을 나열해 주고 한번의 클릭을 통하여 나머지 입력을 완성하는 기능이다. 후입력 자동 완성 리스트 박스의 크기는 입력 창의 크기 제한에 따라 조절하여 사용할 수 있다.

제안된 방법들과 일반 키보드 방식의 성능을 비교, 평가하기 위하여 2000개의 샘플 단어 집합을 구성하고, 단어 집합의 크기를 100개, 500개, 1000개, 1500개, 2000개로 변경해 가면서 세 가지 성능 평가를 수행한 결과 첫째, 단어 입력시 필요한 키 버튼 개수면에서 데이터베이스에 근거한 자동 키보드 방식이 모든 단어 집합 크기에서 평균적으로 높은 성능을 보였다. 또한 단어 집합의 크기가 증가할수록 키 버튼 수의 증가율은 감소함을 알 수 있었다. 둘째, 단어 입력시 필요한 키 선택 횟수면에서 데이터베이스에 근거한 자동 키보드의 후입력 자동 완성 리스트 박스 방식이 키 선택 횟수의 최대값이나 평균값에서 모두 효율적으로 나타났다. 후입력 자동 완성 리스트 박스의 크기가 1일 때와 10일 때를 비교해 보면 크기가 10일 때가 평균적으로 35%정도 더 선택 횟수를 줄일 수 있는 것으로 나타났다. 이러한 성능 평가 결과를 통하여 단어 집합의 크기와 입력 창의 크기 제한 여부에 따라 데이터베이스에 근거한 자동 키보드 방식을 선택적으로 수행할 수 있다.

본 논문에서는 제안한 방법의 성능 평가 시 단어 집합의 크기는 최대 2000개였다. 단어 집합을 사전으

로 확대하면 모든 단어에 대한 입력의 오류를 방지할 수 있는 자동 키보드가 가능할 것이다. 따라서 단어 집합을 사전으로 확대했을 때 프로그램의 성능 향상을 위한 알고리즘과 탐색 기법에 대한 연구가 필요하다. 또한 데이터베이스 기반 자동 키보드 방식은 키버튼의 종류와 위치가 매번 변화되므로 입력 시 혼란을 줄 수 있다. 입력시마다 도시 되는 키버튼이 가변적일 때 키버튼 인지와 선택에 좀더 효율적일 수 있도록 키버튼의 배열에 대한 고려가 필요하다.

참고 문헌

[1] 박동순, 이재현, 최은,강석, 컴퓨터 한글 코드의 사용과 표준화, 정보과학회지 제6권 제1호('88. 2), pages 69-74, 1988.

[2] 김지현, 휴대용 키보드, <http://www.i-biznet.com> [김지현의 디지털월드]

[3] 고희대, 컴퓨터에 의한 한글 입출력 처리에 관한 연구, 전남대 대학원 석사 학위 논문, pages 8-9, 1984.

[4] 양명섭, 구조적 형태분석에 의한 한글 자모분리기의 설계 및 구현, 전북대 대학원, pages 1-9, 1995.

[5] 이준, 고정어 풀어쓰기로 표현된 한글 처리에 관한 연구, 공군사관학교 논문집 41('98.2), pages 129-149, 1998.

[6] 한국어 정보처리 연구소, C로 구현한 한글 코드 시스템 프로그래밍 가이드, 한국어 정보처리 연구소, 2001.

[7] 이재혁, 한글입력을 위한 구조형 가상 키보드의 개발, 군산대학논집 9('84.12), pages 281-296, 1984.

[8] 구민모, 이만영, 전화기 자판의 한글 입력 효율

성 평가모형, 정보처리학회 논문지D 제8-D권 제 3호('02.6), pages 295-304, 2001.

[9] 이현주, 조문중, Pen Windows상에서의 On-line 필기인식 기술 현황, 정보과학회지 제11권 제5호('93.10), pages 104-114, 1993.

[10] 안창, 한글 인식 기술의 현황과 문제점, 한일장신대신학과사회 14, pages 419-430, 2000.



이계숙

1990년 이화여자대학교 전자계산학과 학사
 2002년 이화여자대학교 정보과학대학원 정보과학석사
 1994년~1999년 삼성전자 정보통신연구소 연구원
 1999년~2003년 용인송담대학

정보통신과 시간강사

2004년~현재 용인송담대학 정보통신과 겸임교수
 관심분야 : 데이터베이스 시스템, 네트워크 프로그래밍, 모바일 프로그래밍



용환승

1983년 서울대학교 컴퓨터공학과 학사
 1985년 서울대 대학원 컴퓨터공학과 공학석사
 1985년~1989년 한국전자통신연구소 연구원
 1994년 서울대 대학원 컴퓨터공

학과 공학박사

2002년 8월~2003년 2월 IBM T.J. Watson 연구소 객원 연구원
 1995년 이화여자대학교 컴퓨터학과 부교수
 관심분야 : 객체-관계 데이터베이스 시스템, 멀티미디어 데이터베이스, OLAP 및 데이터 마이닝, 바이오정보학, 유비쿼터스 컴퓨팅