

글자 빈도수와 바이그램을 이용한 스마트폰 영어 글자판 설계*

(Design of Smart phone English Keypad using
Frequency Information of Characters and Bigram)

이 중 화*, 박 유 현*, 김 병 기**

(Jung-hwa Lee, Yoo-hyun Park, and Pyeoung-Kee Kim)

요 약 스마트폰에서는 문서작성, 채팅, 웹 검색 등의 다양한 기능들이 제공되고 있으며, 이에 따라 스마트폰에서 글자를 입력하는데 사용하는 글자판의 중요성이 더욱 부각되고 있다. 그러나 기존의 스마트폰 영어 글자판에서는 글자들을 단순히 글자 순서대로 배치하고 있기 때문에 입력의 효율성이 매우 떨어진다. 따라서 본 논문에서는 영어 글자의 빈도수 정보와 바이그램을 사용하여 글자를 배치한 개선된 영어 글자판을 설계하였다. 또한 성능평가를 통해 제안하는 자판의 성능을 기존의 자판과 비교하여 제안하는 자판의 성능이 더 우수함을 보였다.

핵심주제어 : 영어 키패드, 스마트폰, 모바일 장치, 글자 빈도수

Abstract Smart phones provide variety of functions like world processing, chatting, web browsing, etc. for users. Therefore, efficiency of the keypad that are used in text entry are becoming more important. But, in traditional english mobile keypad, text entry speed is not efficient because simple placement of characters in accordance with the order of characters. In this paper, we propose the improved english keypad for smart phone using the frequency information of english characters and bigram. And, we verified that the proposed keypad is more efficient than other keypads through the experiment.

Key Words : English Keypad, Smart phone, Mobile device, character Frequency

1. 서 론

최근 스마트폰 사용자가 증가하면서 스마트폰을 이용한 다양한 연구들이 진행되고 있다[1][2]. 스마트폰의 기반이 되는 기존의 모바일 폰에서는 음성통화와 단문메시지 서비스 등이 주요기능으로 제공되었다. 하지만 최근 스마트 폰에서는 문서작성, 채팅, 웹 검색

등의 다양한 기능들이 제공되고 있으며, 이에 따라 스마트폰에서 글자를 입력하는데 사용하는 글자판의 중요성이 더욱 부각되고 있다[3][4][5][6].

스마트폰을 포함한 모바일 폰에서 제공되는 글자판은 12개의 글쇠를 가지는 유선 전화기 자판을 기본으로 하고 있다. 그러나 영어뿐만 아니라 대부분의 언어들에서 필요한 글쇠의 수는 12개를 초과하기 때문에 한 글쇠에 여러 개의 문자를 입력할 수 있도록 하는 멀티 탭(multi-tab) 방식을 채택하고 있으며 글쇠에 글자를 배치할 때에도 단순히 글자의 순서대로 한 글

† 이 논문은 2011학년도 동의대학교 교내연구비에 의해 연구되었음(과제번호: 2011AA181)

* 동의대학교 컴퓨터소프트웨어공학과

** 신라대학교 IT학과 (junghwa@deu.ac.kr)

쇠에 두 개 또는 세 개의 글자를 배치하는 형태를 띠고 있다. 그러나 이러한 단순 글자 배치 방법은 글자의 빈도수나 사용 형태를 고려하지 않기 때문에 글자를 입력할 때 정확성과 효율성을 떨어뜨리게 된다. 기존의 유선 전화기나 모바일 폰에서는 전화기 자판을 이용하여 글자를 입력해야 하는 경우가 많지 않았기 때문에 기능이나 성능적인 측면에서 중요성이 부각되지 못했지만, 스마트폰이 등장으로 스마트폰에서 제공하는 다양한 응용들에 글자판이 사용됨에 따라 그 사용성이 더욱 증가되어 글자의 입력 효율성이 더욱 중요하게 다루어지고 있다[7][8][9].

따라서 본 연구에서는 스마트폰을 포함한 모바일 기기에서 사용할 수 있는 개선된 글자판에 대해 연구한다. 이를 위해 먼저 기존의 스마트폰에서 사용되고 있는 글자판을 분석하여 그 장단점을 파악하고, 특히 멀티 탭 방식의 글자판에서 발생하는 문제점들을 분석하여 이를 최소화 할 수 있도록 함으로써 입력 효율성을 증가시킨다. 글쇠에 글자를 배치할 때에도 글자의 사용 특성을 고려하여 배치함으로써 보다 빠르고 정확하게 글자를 입력할 수 있도록 글자판을 설계한다. 또한 성능평가를 통해 제안하는 자판의 성능을 기존의 자판과 비교하여 제안하는 자판의 성능을 검증한다.

2. 관련 연구

일반적인 모바일 폰 글자판에서는 여러 가지 방식에 의해 키를 누름으로써 글자를 입력한다. 지금 까지 모바일 폰 글자판에 대한 연구는 어떻게 글자를 입력할 것인가라는 글자를 입력하는 방식과 글쇠에 글자를 어떻게 배치할 것인가라는 글자를 글쇠에 배치하는 방법들에 대해 이루어 졌다. 본 장에서는 기존의 영어 모바일 폰 글자판에 대한 연구에 대해 살펴본다. 처음으로 살펴볼 글자판은 아래 <그림 1>에 있는 가장 전통적인 표준 영어 글자판이다[8].

1	2 abc	3 def
4 ghi	5 jkl	6 mno
7 pqrs	8 tuv	9 wxyz
*	0 space	#

<Fig 1> 표준 모바일 폰 글자판

위 방식은 기존의 전화기 영어 글자판을 그대로 옮겨온 것으로써, 각 글쇠에 영어 글자 3개 또는 4개씩을 배치한 형태이다. 이 글자판은 한 글쇠에 여러 개의 글자가 배치되어 있고 누름 동작만으로 글자를 입력하기 때문에 필요에 따라 글쇠를 여러 번 눌러야 하는 멀티 탭(multi-tab) 방식을 채택하고 있다. 예를 들어 글쇠의 가장 처음에 배치되어 있는 a, d, g 등은 해당 글쇠를 한 번만 누름으로써 입력할 수 있지만 s와 같은 글자는 7번 글쇠를 네 번 눌러야만 입력할 수 있다.

이 글자판은 글쇠에 글자를 배치할 때 단순히 글자의 순서대로 각 글쇠에 배치하였기 때문에 초보자도 그 위치를 쉽게 알 수 있으며 사용법이 단순해서 배우기 쉽다는 정점을 가진다. 그러나 영어 문장 상에서의 글자의 빈도수 등의 언어적 특성이나 글자간의 거리 등과 같은 글자의 입력의 효율성을 전혀 고려하지 않고 있기 때문에 글자판의 입력 효율성 측면에서는 바람직하지 않은 형태라고 할 수 있다. 그럼에도 불구하고 위의 자판은 매우 오랫동안 사실상 표준으로 자리 잡고 있기 때문에 현재에서도 기본 모바일 자판으로 스마트폰을 포함한 대부분의 모바일 기기에서 제공되고 있다.

위의 표준 영어 자판의 단점을 개선하고자 하는 연구들에 있어서 글쇠에 배치되어 있는 글자의 순서를 바꾸어 글자의 입력 효율성을 높이고자 하는 연구가 진행되었다[8]. 이 연구에 따르면, 글자의 입력 빈도를 예측하고 이에 따라 기존의 글자판에 배치되어 있는 글자의 위치를 <그림 2>와 같이 바꾸어 입력의 속도를 증가시키려고 하였다. 그러나 이 연구에서는 글쇠 그룹을 만들 때 각 글쇠 그룹에 속하는 글자들 간의 연속 입력 가능성에 대해 고려하고 있지 않기 때문에 멀티 탭 방식의 글자판에서 발생하는 모호성 문제의 발생 빈도를 줄이지 못하고 있다.

1	2 adm	3 egb
4 ily	5 thv	6 rcp
7 suk	8 ofxy	9 nwjq
*	0 space	#

<Fig 2> 개선된 모바일 폰 글자판

또 다른 방식으로 최대 네 번의 키 누름이 발생하는 기존의 멀티 탭 방식의 문제점을 개선하고자 하는 글자판 디자인에 대한 연구가 있다[9]. 이 연구에서는 제안하는 글자판을 일명 메시지-이지 글자판(MessageEasy keypad)라고 부르는데, 이 글자판에서는 모든 글자들을 두 번의 글쇠 눌러 입력하는 방법을 사용한다. 이 방법의 특징은 모든 글자를 동일한 횟수의 글쇠 누름으로 글자를 입력할 수 있어서 사용자가 글자를 입력할 때 동일한 리듬으로 글쇠를 누름으로써 오타를 줄일 수 있으며, 평균적인 키 입력 횟수를 줄임으로써 입력의 효율성을 높일 수 있다는 것이다. 그러나 모든 글자들을 두 번의 키 입력으로 입력해야 하기 때문에 단문의 입력에서는 상대적으로 효율성이 떨어질 수 있으며 무엇보다도 입력방식을 이해하기 어렵다는 단점이 있다.

최근에는 스마트폰에서 기존의 컴퓨터 글자판으로 사용되고 있는 쿼티 글자판을 그대로 사용할 수 있다. 그러나 자판에 배치되어 있는 글쇠의 수가 많고 스마트폰의 크기가 제한적이기 때문에 입력 시 많은 오타를 발생시킬 수 있는 문제점이 있다.

3. 빈도수 정보와 이동거리를 고려한 영어 글자판 설계

글자판을 설계하는데 있어서 가장 중요하게 고려되어야 하는 것은 사용자가 얼마나 정확하고 빠르게 원하는 글자를 입력할 수 있는가 하는 것과 사용자가 글자판에 배치되어 있는 글자들을 얼마나 빨리 습득할 수 있는가 하는 것이다[10].

오타율을 줄이고 글자의 입력 속도를 증가시키기 위해서는 자주 사용되는 글자들, 즉 빈도수가 높은 글

자들을 사용자가 누르기 편한 위치에 배치하고 글자판 내에서 글자간의 이동 거리를 최소화하는 것이 중요하다. 본 논문에서는 영어 말뭉치를 이용하여 영어 글자들의 빈도수와 글자간의 연관성을 조사하고 조사된 결과를 바탕으로 글자들을 글자판에 배치하여 글자 입력의 속도를 향상시키고자 한다.

본 논문에서 기존의 모바일 글자판과의 호환성을 위해서 3*4 형태의 12개의 글쇠를 가지는 기존 모바일 자판을 사용한다. 이는 3*4 자판이 현재까지 모바일 자판의 표준으로써 대부분의 기기에 기본적으로 제공되고 있으며 현재까지 많은 사용자들이 사용하고 있기 때문이다.

3.1 빈도수 정보를 이용한 글자 그룹 선정

12 글쇠 글자판에서는 글쇠의 수가 표현해야 할 글자의 수보다 적기 때문에 한 글쇠에 여러 개의 글자를 배치해야만 한다. 따라서 글자판은 설계할 때 가장 먼저 고려해야 할 사항은 어떠한 글자들을 같은 글쇠에 위치시킬 것인지, 또한 글자들의 배치 순서를 어떻게 할 것인지 하는 것이다.

기존의 모바일 폰 글자판의 경우를 살펴보면 숫자 2 글쇠의 경우 a, b, c의 세 글자가 배치되어 있는데, a는 글쇠 2를 한번 누르는 반면에 c를 입력하고자 하는 경우는 글쇠 2를 세 번 눌러야 하기 때문에 더 많은 시간이 필요하게 된다. 여기에서 만약 b가 상대적으로 a보다 자주 사용될 경우는 abc 순서보다는 bac 순서로 배치하는 것이 더 효율적일 것이다. 즉, 영어 문장 내에서 자주 사용되는 글자들을 각 글쇠의 첫 번째에 위치하는 1 수준(1'st level) 글자로 정하게 되면 글자를 입력하기 위해 글쇠를 여러 번 눌러야 하는 경우를 줄일 수 있기 때문에 자판의 성능을 향상시킬 수 있다. 따라서 본 논문에서는 영어 문장 내에서 나타나는 각 글자의 빈도수를 조사하여 이를 바탕으로 각 글쇠에 위치할 글자들의 수준(level)을 결정한다.

아래 <표 1>은 뉴욕타임즈 말뭉치(corpus)를 이용하여 영어 문장에서 각 글자들이 나타나는 빈도수를 조사한 것이다[11].

<Table 1> 영어 글자 빈도수

글자	빈도수	비율(%)	순위
E	7,741,842	12.55	1
T	5,507,692	8.93	2
A	5,263,779	8.53	3
O	4,729,266	7.67	4
N	4,535,545	7.35	5
I	4,527,332	7.34	6
S	4,186,210	6.79	7
:	:	:	:
V	653,370	1.06	21
K	460,788	0.75	22
X	123,577	0.20	23
Z	66,423	0.11	24
J	65,856	0.11	25
Q	54,221	0.09	26

위의 테이블을 살펴보면 각 글자들이 영어 문장 내에서 나타나는 빈도수가 최대 12.55%에서 최소 0.09%와 같이 많은 차이를 보이며 특히 상위 9개의 글자의 빈도수가 약 70.67%를 차지하고 있음을 알 수 있다. 만약 이 글자들을 각 글쇠의 첫 번째 수준에 위치시키면 같은 글쇠를 두 번 또는 세 번 누름 없이 한 번씩만 눌러 약 71%의 글자를 입력할 수 있게 된다는 뜻이다.

위의 결과를 바탕으로 본 논문에서는 e, t, a, o, n, i, s, r, h까지 빈도수 상위 9개의 글자를 각 글쇠에 위치할 1 수준 글자로 선정한다.

같은 방법으로 다음 빈도수 상위 9개의 글자를 선정하여 글쇠 두 번 누름으로 입력할 수 있는 2 수준 글자 그룹에, 나머지 글자들을 3 수준 그룹에 포함 시킨다.

선정된 각 수준에 속하는 글자들은 <표 2>와 같다.

<Table 2> 각 수준 별 글자

글자 수준	글자 종류
1 수준	e, t, a, o, n, i, s, r, h
2 수준	l, d, c, u, m, f, p, g, y
3 수준	w, b, v, k, x, z, j, q

앞에서 언급한 바와 같이 12 글쇠 글자판에서는 글자판 글자판의 수가 입력해야 할 글자의 수보다 적기

때문에 각 글자판에 3, 4개의 글자를 배치해야 한다.

따라서 두 번째 단계로 각 수준별 글자 그룹에서 글자들을 선정하여 같은 글쇠에 놓일 글자들의 그룹을 만든다.

글자 그룹을 만들 때에는 멀티 탭 입력 방식을 사용할 때 발생하는 모호성(ambiguity) 문제가 발생되는 것을 고려해야 한다. 예를 들어 기존의 12 글쇠 글자판에서 글쇠 2번에 위치한 adm의 경우를 살펴보면, 글쇠를 한번 누를 때는 a, 두 번 누를 때는 d, 세 번 누를 때는 m이 입력되게 되는데 ad를 연속적으로 입력해야 하는 경우는 같은 글쇠를 a입력을 위해 한 번, d 입력을 위해 두 번째서 총 세 번을 눌러야 한다. 하지만 이 경우는 d를 입력하기 위해 글쇠를 세 번 누르는 경우와 구별할 수 없게 된다.

이 문제는 일반적으로 두 가지 방법으로 해결하고 있는데, 첫 번째 방법은 탭 글쇠를 사용하는 방법이다. 위의 예에서 a를 입력할 때에는 글쇠 2번을 한 번 누르고, 이어서 m을 입력하는 경우는 중간에 탭 글쇠를 눌러 첫 번째 입력이 종료되었다는 것을 알려줌으로써 탭 글쇠 이후의 입력에 대해서는 첫 번째 누름으로 간주하여 m이 입력되도록 한다. 두 번째 방법은 탭 글쇠 입력하는 대신 글쇠를 처음 입력 이후에 일정시간을 지난 다음 입력하도록 하는 타임-킬(time-kill) 방식이다.

위의 두 가지 방법은 모두 멀티 탭 방식의 글자판에서 모호성 문제를 해결할 수 있지만 두 방법 모두 글자의 입력 효율성을 치명적으로 저하시킬 수밖에 없다. 탭 글쇠는 쓰는 방법은 글자판에 탭 글자를 위한 별도의 글쇠를 두어야 하는 부담이 있으며, 통상 탭 글쇠로 사용하는 글쇠는 글자판의 가장자리에 위치하고 있기 때문에 글쇠간의 이동거리가 길어 전체적인 입력 속도 저하를 가져온다. 두 번째 타임-킬 방법은 탭 글쇠를 별도로 둘 필요가 없다는 장점이 있지만, 타임-킬에 통상 1-1.5초 정도의 시간이 필요하기 때문에 첫 번째 방법과 마찬가지로 입력 속도의 저하를 초래하게 된다. 따라서 멀티 탭 방식으로 글자판을 설계할 때는 모호성 문제가 발생되는 횟수를 최소화 할 수 있도록 글자를 배치하는 것이 중요하다.

위에서 알 수 있듯이 모호성 문제는 한 글쇠에 같이 위치하는 글자들이 연속적으로 입력될 때 발생하게 된다. C1C2C3의 세 개의 문자가 배치되어 있는 글쇠의 경우 세 개의 문자로 조합할 수 있는 C1C2,

C1C3, C2C3로 이루어진 부분 문자열이나 역으로 C2C1, C3C1, C3C2로 이루어진 부분 문자열을 입력하는 경우 모호성이 발생한다. 따라서 본 논문에서는 말뭉치 속에 특정 글자들이 연속적으로 나타나는 빈도수를 조사하여 같은 글자 그룹에 배치 가능한 글자 중에서 연속적으로 나타날 확률이 최소인 글자들을 선정함으로써 모호성 문제를 최소화 한다.

여기에서는 두 개의 글자가 연속적으로 나타날 횟수를 2차원 행렬로 표현한 바이그램(bigram)을 사용하여 같은 그룹에 속할 글자를 선정한다. 본 논문에서 사용한 바이그램은 앞에서 영어 글자 빈도수 조사에서 사용한 말뭉치를 이용하여 가능한 두 개의 영어 글자의 쌍이 나타나는 횟수를 조사한 것이다.

<표 3>은 뉴욕타임즈 말뭉치에서 조사한 바이그램의 일부분으로써 빈도수를 확률로 다시 나타낸 것이다[11].

<Table 3> 바이그램 예

	a	b	y	z
a	0.03	10.60	3.48	21.44
b	2.04	0.96	1.36	0.62
:	:	:	:	:	:
x	0.32	0.00	0.03	0.00
y	3.36	10.60	0.01	2.42
z	0.21	0.00	0.26	5.78

위의 바이그램은 임의의 선행 글자 Ci에 대해 후행 글자 Cj가 나타날 확률 Bij을 계산한 것으로 바이그램을 이용하면 특정 글자 다음에 나타날 확률이 가장 적은 글자를 선정할 수 있다.

앞에서 언급한 바와 같이 CiCj가 연속적으로 나타날 경우 모호성이 발생한다면 CjCi와 같이 역으로 나타날 때에도 역시 모호성이 발생한다. 따라서 글자 그룹을 선정할 때에는 Bij와 Bji를 모두 고려하여야 한다. 따라서 본 논문에서는 Bij와 Bji의 합(Tij)이 최소인 글자 조합을 선정하여 같은 글자 그룹에 배치한다. 아래 <표 4>는 선행글자가 레벨 1의 e일 때 레벨 2의 각 글자가 가지는 바이그램 값이다.

<Table 4> 'e'에 대한 2 수준 글자의 바이그램 값

2 수준 글자	Bij	Bji	Tij
c	4.63	17.50	22.13
d	11.38	34.26	45.64
f	1.39	11.57	12.96
g	1.14	25.71	26.85
l	4.63	18.97	23.6
m	3.10	29.28	32.38
p	1.70	18.78	20.48
u	0.19	4.57	4.76
y	1.70	32.72	34.42

위의 표에 따르면 e와 d로 이루어진 부분 문자열의 Tij 값이 45.64로 가장 높고 반대로 e와 u로 이루어진 부분 문자열의 Tij 값이 4.76으로 가장 낮다. 즉 e와 d를 같은 그룹에 위치시키기 보다는 e와 u를 같은 그룹에 위치시킬 경우 두 개의 글자가 연속적으로 나타날 확률이 가장 낮기 때문에 멀티 탭 방식 때문에 발생하는 모호성이 나타날 확률이 낮다는 것을 알 수 있다.

같은 방법으로 3 수준의 글자들인 w, b, v, k, x, z, j, q 중에 위의 각 글쇠 그룹에 속할 글자들을 선정한다. CiCjCk의 세 개의 글자가 글쇠에 배치되는 경우는 세 개의 문자로 조합할 수 있는 CiCj, CiCk, CjCk로 이루어진 부분 문자열이나 역으로 CjCi, CkCi, CkCj로 이루어진 부분 문자열을 입력할 때 모호성이 발생한다. 따라서 이 경우에는 Tijjk를 모든 조합 가능한 부분 문자열의 Bij 값들의 합으로 정의하고 3 수준의 글자와의 조합에서 계산되는 Tijjk 값들 중에서 가장 작은 값을 가지는 글자를 선택한다.

<표 5>는 각 글쇠 그룹에 포함되는 1, 2, 3 수준 글자들을 최종적으로 선정한 결과이다.

<Table 5> 최종 글자 그룹

그룹	1 수준	2 수준	3 수준
1	e	u	w
2	t	d	q
3	a	y	k
4	o	g	x
5	n	p	j
6	i	c	b
7	s	f	v
8	r	l	z
9	h	m	-

3.2 이동거리를 고려한 글자 배치

지금까지 각 글쇠에 배치할 글자들의 빈도수 정보와 바이그램을 이용하여 모호성 문제가 발생하는 횟수를 최소화 할 수 있도록 글자 그룹들을 선정하였다. 본 절에서는 선정된 글자 그룹들을 글자판에 배치한다.

글쇠에 글자그룹을 배치할 때에는 사용자가 글자를 입력할 때 누르는 글쇠간의 거리를 최소화하면 전체의 입력시간을 줄일 수 있게 되는데, 이를 위해서는 글자판의 중앙을 기점으로 가장자리로 움직이는 것이 가장 효율적이다. 또한 각 글쇠 위치에 대해서 입력의 용이성이 다른데, 본 논문에서는 이를 모두 고려하여 글자들을 배치한다.

<그림 3>은 글자판에서 글자가 배치될 글쇠에 대해 사용자 선호도를 고려하여 배치 우선순위를 나타낸 것이다[12].

2	4	3
5	1	6
7	9	8
11	10	12

<Fig 3> 배치 우선 순위

다음으로 글자판에 글쇠그룹을 배치하기 위해 그룹들의 배치 순위를 결정한다. 글쇠그룹의 배치순위는 각 그룹에 속한 글자들의 빈도수 평균을 구해서 값이 큰 순서대로 선호 순위가 높은 글쇠에 배치하도록 한다.

<표 6>은 각 글쇠 그룹에 속한 글자들의 빈도수의 평균을 계산한 것이다.

<Table 6> 그룹별 빈도수

	평균빈도수	빈도수(%)	등수
euw	3456940	16.23269	1
tdq	2643911	12.41496	2
icb	2451300	11.51052	3
ayk	2262202	10.62258	4
rlz	2252508	10.57706	5
hm	2211617	10.38505	6
sfv	2045502	9.605021	7
ogx	2019863	9.484632	8
npj	1952327	9.167501	9

위의 결과를 이용하여 최종적으로 글자판에 글쇠그룹을 위치시킨다. hm이 속한 그룹은 다른 그룹과 달리 2개의 글자로 이루어져 있기 때문에 문장을 입력할 때 꼭 필요한 '.'을 이 그룹에 배치하도록 한다. <그림 4>는 본 논문에서 제안하는 최종적인 글자판 설계이다.

1	2	3
tdq	ayk	icb
4	5	6
rlz	euw	hm.
7	8	9
sfv	npj	ogx
*	0	#
	space	

<Fig 4> 제안된 글자판

4. 성능 평가

본 장에서는 본 논문에서 제안하는 글자판의 성능을 평가한다. 글자판의 성능평가에는 Fitts' law를 기반으로 하는 이동모델을 사용한다. Fitts' law는 이동 거리에 따르는 입력 시간을 예측할 수 있는 모델로써 글자판과 같은 물리적 장치의 성능을 측정하는데 사용된다[13].

Fitts' law에서는 하나의 키를 입력한 후 다음 키를 입력하는데 걸리는 시간을 나타내는 MT를 아래 식(1)과 같이 표현하고 있다.

$$MT = a + b \times ID \quad (1)$$

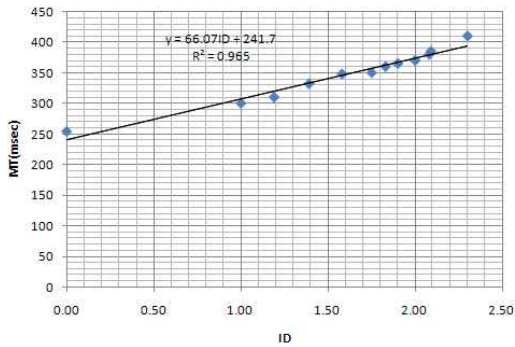
위의 식에서 a와 b는 실험에 의해 측정되는 측정치이며, ID는 입력의 난이도를 의미한다. ID는 글자판의 물리적인 크기와 이동거리에 의해 결정되는데, 식(2)와 같이 표현된다.

$$ID = \log\left(\frac{A}{W} + 1\right) \quad (2)$$

여기서 W는 자판의 가로크기이며 A는 글쇠 간의 이동 거리로써 글쇠의 크기에 따라 결정된다. 글쇠의 크기는 제품의 크기에 따라 서로 다른데, 본 논문에서는 글쇠의 가로크기를 10.2mm, 세로 크기를 8mm로

가정한다.

다음으로 위의 값들을 이용하여, 입력 시간 MT를 계산할 수 있도록 계수 a와 b를 결정하여 식(1)을 완성한다. 이를 위해 본 논문에서는 30명의 실험자를 선정하여 이동거리 별로 글쇠를 누를 때 걸리는 시간을 측정하여 다음 <그림 5>와 같은 선형 회귀 분석을 통해 각 계수 값을 결정한다.



<Fig 5> 선형 회귀 분석

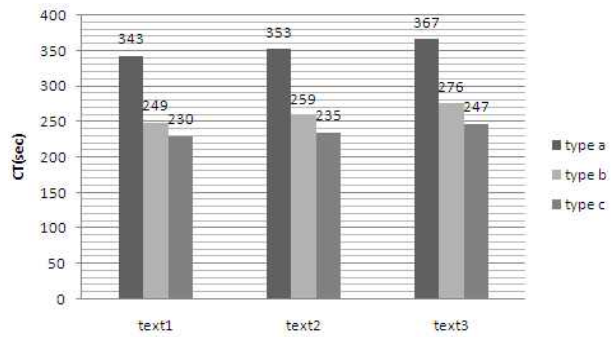
위의 결과에 의하면, a값은 241.7며 b 값은 66.07로 계산되었으며, 상관계수는 0.965로 나타났다. 따라서 MT는 다음과 같이 표시될 수 있다.

$$MT = 241.7 + 66.07 \times ID \quad (R^2 = 0.965) \quad (3)$$

이제 완성된 이동모델을 이용하여 성능평가를 수행한다. 성능평가는 앞의 <그림 1>의 기존의 표준 글자판(type a)과 <그림 2>의 개선된 글자판(type b), 그리고 <그림 4>의 본 논문에서 제안하는 글자판(type c)을 대상으로 이동모델을 사용하여 문장을 입력할 때 걸리는 시간을 측정한다. 실험에는 세 가지 형태의 문장들을 사용하였는데, 텍스트1은 SMS에서 사용될 수 있는 일상생활에서 사용되는 대화체 문장들로 구성하였으며, 텍스트 2는 트위터에서 사용된 문장들로 구성하였다. 또한 텍스트 3은 신문 기사에서 추출하였다. 각 텍스트들은 약 100 단어 정도의 길이를 가지도록 실험 데이터를 구성하였다.

<그림 6>은 세 가지 실험 데이터를 입력할 때 걸리는 시간을 비교한 것이다.

실험 결과에 따르면 본 논문에서 제안하는 자판이 type a에 비해 텍스트 입력에 소요되는 예상 시간이



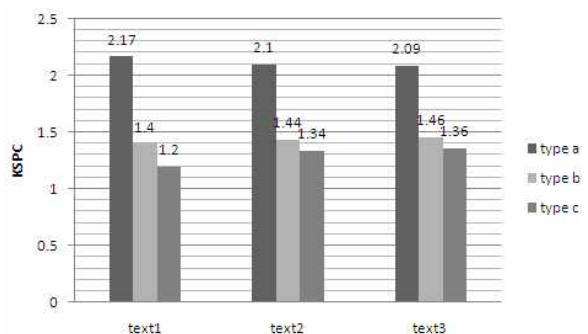
<Fig 6> 입력 시간 비교

각각 113(sec), 118(sec), 120(sec) 단축됨을 알 수 있으며 type b에 대해서도 세 개의 실험 데이터에 대해 각각 19(sec), 29(sec), 29(sec) 단축되어, 본 논문에서 제안하는 글자판(type c)이 다른 두 종류의 글자판에 비해 입력 속도가 향상되었음을 알 수 있다. 이러한 결과는 본 논문에서 제안하는 자판이 글쇠간의 이동거리가 짧아서 이동에 소요되는 시간이 작고, 또한 모호성 문제가 발생할 확률이 적어 땀 글쇠 누르거나 타임-킬(time-kill)에 소요되는 시간이 줄어들기 때문이다.

다음으로, 본 논문에서는 제안된 글자판과 기존의 글자판에서 문장을 입력할 때 필요한 타수를 측정하여 비교한다. 타수는 KSPC(key strokes per character)에 의해 측정될 수 있는데 KSPC는 식(4)와 같이 나타낼 수 있다[14].

$$KSPC = \frac{\text{총 글쇠 누름 횟수}}{\text{총 글자수}} \quad (4)$$

<그림 7>은 실험 데이터에 대한 각 글자판의 KSPC를 비교한 것이다.



<Fig 7> KSPC 비교

실험 결과에 따르면 본 논문에서 제안하는 type c 자판이 한 글자를 입력하는데 평균 1.3번 글쇠를 누르는 반면, type a는 2.12번, type c는 1.43번으로 나타났다. 위의 결과에서 살펴볼 수 있듯이 제안하는 자판은 빈도수가 높은 글자를 각 글쇠의 첫 번째 레벨에 배치하여 글자를 입력할 때 누르는 글쇠의 횟수를 최소화하였기 때문에 KSPC 측면에서도 기존의 자판과 비교하여 효율적임을 알 수 있다.

5. 결 론

본 논문에서는 스마트폰에서 사용할 수 있는 개선된 영어 글자판을 제안하였다. 기존의 글자판에서는 단순히 글자를 순서대로 배치하고 멀티 탭 방식으로 글자를 입력하기 때문에 모호성 문제를 발생시켜 문자를 입력할 때 많은 시간이 소요된다. 따라서 본 논문에서는 글자의 빈도수와 바이그램을 사용하여 모호성 문제의 발생을 최소화 할 수 있도록 하고, 운지거리를 최소화 할 수 있도록 글자를 배치하여 효율적인 글자 입력이 가능하도록 하였다. 또한 제안하는 자판의 성능을 검증하기 위해 Fitts' law에 기반한 이동모델과 KSPC를 이용하여 기존의 자판과의 성능을 비교하였다. 실험 결과 본 논문에서 제안하는 자판의 기존의 자판보다 더 효율적임을 알 수 있었다.

최근 스마트폰의 등장으로 모바일 폰 글자판의 중요성이 더욱 부각되고 있는 시점에서 본 연구와 같은 개선된 글자판에 대한 연구는 더욱 중요하다고 할 수 있다. 향후 본 연구에서 제안하는 글자판을 배포하여 실제 사용해 보고 그 효용성을 검증해 볼 것이며, 사용자 요구 사항을 반영하여 글자판의 성능을 더욱 개선할 계획이다.

참 고 문 헌

- [1] 송충원, 남광우, 이창우, "스마트폰 센서스트림을 이용한 운전 패턴 인식 시스템", 한국산업정보학회 논문지, v.17, no.3, pp.35-42, 2012년 6월
- [2] 안병익, 주영도, "스마트 폰 중심의 위치 기반 소셜 네트워크 서비스 모델 설계", 한국산업정보학회 논문지, v.16, no.5, pp.55-62, 2011년 12월
- [3] Nandar Pwint Oo and Ni Lar Thein, "iTextMM: Intelligent Text Input System for Myanmar Language on Android Smartphone", Lecture Notes in Electrical Engineering, Vol. 107, Part 8, pp. 661-670, 2011.
- [4] 김호식, 전재웅, 박태진, 최윤철, "터치스크린 기반 스마트폰을 위한 한글 입력 기법", 정보과학회 논문지:컴퓨팅의 실제, 제17권, 제10호, pp. 555-558, 2009.
- [5] Martin Klima, Vaclav Slovacek, "Vector Keyboard for Touch Screen Devices", HCI 2009, LNCS 5627, pp. 250-256, 2009.
- [6] Junghwa Lee, "Hangeul Keypad for Smart Phones Using Gesture", International Journal of KIICE, Vol. 9, No. 6, pp. 706-711, 2011.
- [7] Arpit Mittal, Arijit Sengupta, "Improvised layout of keypad entry system for mobile phones", Computer Standards & Interfaces, Vol. 31, Issue 4, pp. 693-698, June 2009.
- [8] Arpit Mittal, Arijit Sengupta, "Optimized Layout for Keypad Entry System", Conf. of Embedded Systems and Applications(ESA'07), pp.127-133, 2007.
- [9] Saied B, Nesbat, "A System for Fast, Tull-Text Entry for Small Electronic Devices", Proceedings of the ICMI 2003, pp.154-162, November, 2003
- [10] Barry O'Riordan, Kevin Curran, Derek Woods, "Investigating Text Input Methods for Mobile Phones", Journal of Telematics and Informatics, Vol. 23, Issue 1, pp. 1-21, February 2006.
- [11] Michael N. Jones, D.J.K.Mewhort, "Case-sensitive letter and bigram frequency counts from large-scale English corpora", Behavior Research Methods, Instruments & Computers, pp. 334-396, 2004
- [12] Gecynth Namu, "Improvised Layout of Mobile Keypad for Filipinos", International Journal of KIMICS, Vol. 8, No. 3, pp.328-332, 2010.
- [13] MacKenzie, I.S., "Fitts' law as a research and design tool in human-computer interaction", Human-Computer Interaction, Vol. 7, pp.91-139, 1992

[14] I. Scott MacKenzie, "KSPC(Keystrokes per Character) as a Characteristic of Text Entry Techniques", Mobile HCI 2002, LNCS 2411, pp.195-210, 2002.



이 중 화 (Jung-hwa Lee)

- 정회원
- 부산대학교 전자계산학과 이학사
- 부산대학교 전자계산학과 이학석사
- 부산대학교 전자계산학과 이학박사
- 동의대학교 공과대학 컴퓨터소프트웨어공학과 부교수
- 관심분야 : 데이터베이스, 한글정보처리, 멀티미디어, 모바일 시스템



박 유 현 (Yoo-hyun Park)

- 부산대학교 전자계산학과 이학사
- 부산대학교 전자계산학과 이학석사
- 부산대학교 전자계산학과 이학박사
- 동의대학교 공과대학 컴퓨터소프트웨어공학과 조교수
- 관심분야 : 인터넷 시스템, 클라우드 시스템, 융합IT 서비스



김 병 기 (Pyeong-kee Kim)

- 1988년 경북대학교 전자공학과(학사)
- 1990년 경북대학교 전자계산기공학과(석사)
- 1995년 경북대학교 컴퓨터공학과(박사)
- 1995년-현재 신라대학교 IT학과 교수
- 관심분야 : 패턴인식, 영상처리, 멀티미디어, 컴퓨터 게임

논문 접수일 : 2012년 09월 17일
 1차수정완료일 : 2012년 10월 02일
 게재확정일 : 2012년 10월 15일