

# 연속적인 제스처와 언어 모델을 이용한 한글 입력 인터페이스

김동호<sup>0</sup> 권영희 김진형  
한국과학기술원 전자전산학과  
(dkim<sup>0</sup>, kyhee, jkim)@ai.kaist.ac.kr

## A Hangul Input Interface Using Continuous Gestures and Language Models

Dongho Kim<sup>0</sup>, Younghee Kwon, Jin-Hyung Kim  
Dept. of Electrical Engineering & Computer Science, KAIST

### 요 약

휴대 전화와 PDA의 급속한 보급과 더불어, 휴대용 장치에서의 간단한 텍스트 입력 방식 연구의 중요성도 나날이 증대되고 있다. 현재 다양한 방식의 입력 방법들이 존재하나, 한글의 특성을 반영하지 않아 한글 입력에 적용하기 어려운 단점들이 존재한다. 본 연구에서는 2차원 평면상의 연속적인 제스처와 언어 모델을 이용하고, 한글의 특성을 고려한 새로운 한글 입력 방식을 제안한다.

### 1. 서론

키보드는 컴퓨터의 문자 입력 장치로 널리 이용되어 왔으나, 휴대성이 떨어지기 때문에 휴대전화나 PDA 등의 모바일 환경에서는 적합하지 않다. 이러한 휴대용 장치에서 키보드를 대체할 수 있는 보다 간편한 텍스트 입력 방식으로 여러 가지가 연구되고 있으며, 그 중 전자펜이나 제스처를 이용한 입력 방법으로, 일반적인 문자 인식 외에 고정된 레이아웃을 이용하는 방법[1,2], 언어 모델을 통한 동적인 레이아웃을 이용하는 방법[3,4] 등이 존재한다. 그러나 이러한 입력 방식들은 한글 구조의 특성이 반영되지 않았기 때문에 한글 입력에 적용하기에 맞지 않는 부분이 많았다.

### 2. Dasher[3]

Dasher는 연속적인 제스처와 언어 모델을 이용한 문자 입력 인터페이스이다. Dasher는 문자의 입력을 뜻하는 2차원 평면 상의 사각형들과 마우스 등으로 움직일 수 있는 포인터로 이루어진다. 이 때 사각형들은 알파벳 순서로 배열되어 있고(그림 1(a)), 사용자는 입력하고자 하는 문자의 사각형이 위치하는 방향으로 포인터를 움직여 문자를 입력할 수 있다.

포인터를 접근시키면 시점이 포인터 방향으로 접근하면서 사각형의 크기는 계속 확대되고, 사각형 내부에 표시된 다음 문자들의 사각형으로 다시 접근함으로써 다음 문자를 입력할 수 있다(그림 1(b)). 이 때 평면 상의 레이

아웃은 언어 모델을 통해서 결정되는데, 지금까지 입력된 문맥을 바탕으로 다음에 입력될 문자의 확률을 계산하고 그에 비례하게 사각형의 크기가 결정된다. 그러므로 확률이 높은 문자는 확률이 낮은 문자보다 쉽게 입력할 수 있다. 그림 1은 Dasher에서 영문 알파벳을 입력하는 모습을 보여준다.

### 3. 한글의 특성을 반영한 한글 입력

라틴문자에서는 26개의 알파벳을 좌우로만 나열하여 단어를 구성하지만 한글은 19가지의 초성, 21가지의 중성, 27가지의 종성이 결합되어 하나의 음절을 형성하고, 형성된 음절들이 좌우로 나열되어 어절과 문장을 구성한다. 조합 가능한 음절 수가 11172개에 달하기 때문에 라틴문자에서처럼 한번에 한 문자를 입력하기는 불가능하고, 음절 단위가 아닌 자모 단위로 입력을 해야 한다. 자모 단위로 한글 입력 인터페이스를 구현할 때, 자모의 조합이라는 한글의 고유한 특성으로 인해서 몇 가지 고려할 점이 존재한다.

#### 3.1 자모 입력의 제한

라틴문자에서는 한 문자 다음에 올 수 있는 문자의 제한이 없으므로 한 문자를 입력할 때 항상 26개의 알파벳 중 하나를 선택한다. 그러나 라틴 알파벳을 한글 자모로 치환하여 한글 입력을 한다면 한글 자모의 종류는 67개이기 때문에 매번 67개 자모 중 하나를 선택해야 하므로

인터페이스 상에서 복잡도가 크게 증가하게 된다. 그림 2는 Dasher에서 한글 음절 ‘한’을 입력하는 모습을 보여준다. 크기의 차이는 있지만 매번 자모를 선택할 때마다 초, 중, 종성의 모든 자모들이 표시되고 있음을 알 수 있다.

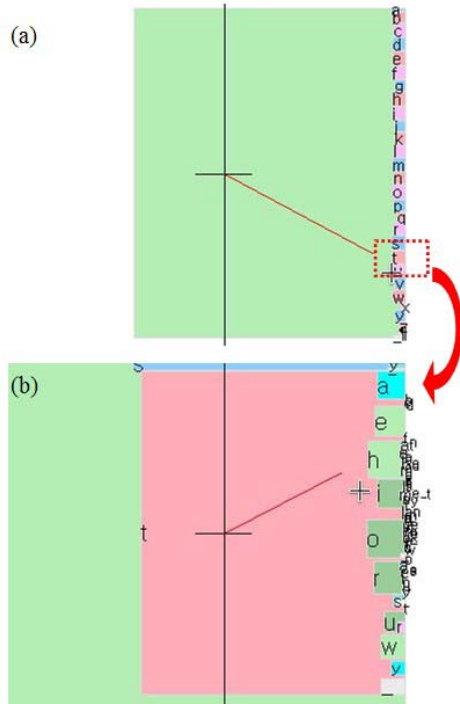


그림 1. Dasher에서의 영문 입력  
(a) 초기상태  
(b) ‘t’가 입력된 후의 상태

그러므로 초성 다음에는 중성, 중성 다음에는 초성 또는 중성, 중성 다음에는 초성이 오는 한글의 특성을 반영하여 선택할 수 있는 자모의 종류를 제한할 수 있다. 이 경우 한글 음절이 제대로 조합되지 않은 경우를 입력할 수 없지만 인터페이스를 보다 간결하게 만들 수 있다는 장점이 있다.

### 3.2 한글 조합 방법

초, 중, 종성 67개의 자모를 입력하는 방법 외에도 여러 방법으로 한글을 조합할 수 있다. 한글 음절에서 종성이 없는 경우도 종성의 한 종류로 분류하여 68개의 자모를 이용하면 한 음절은 항상 하나씩의 초, 중, 종성으로 조합된다. 또는 2벌식 키보드에서와 같이 전체 토큰의 수를 줄이기 위해서 몇몇 복합자음과 이중모음 등을 제외한 19개의 자음과 14개의 모음만을 이용하여 한 음절을 조합할 수 있다.

### 3.3 한글 표시 방법

한글 초성과 종성에는 동시에 쓰이는 16개의 자음(ㄱ,

ㅋ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅆ, ㅇ, ㅈ, ㅊ, ㅋ, ㅌ, ㅍ, ㅎ)이 존재한다. 그림 2의 (c)에서, 자음들이 배열된 위치를 고려하지 않으면 초성인지 중성인지를 구분하기 힘들기 때문에 입력 시 혼동을 가져올 수 있다. 그러므로 인터페이스 상에서 한글 자모를 표시할 때, 음절 단위로 조합해서 표시하는 방법을 고려할 수 있다. 그림 3은 음절 단위로 조합해서 표시한 예를 보여준다.

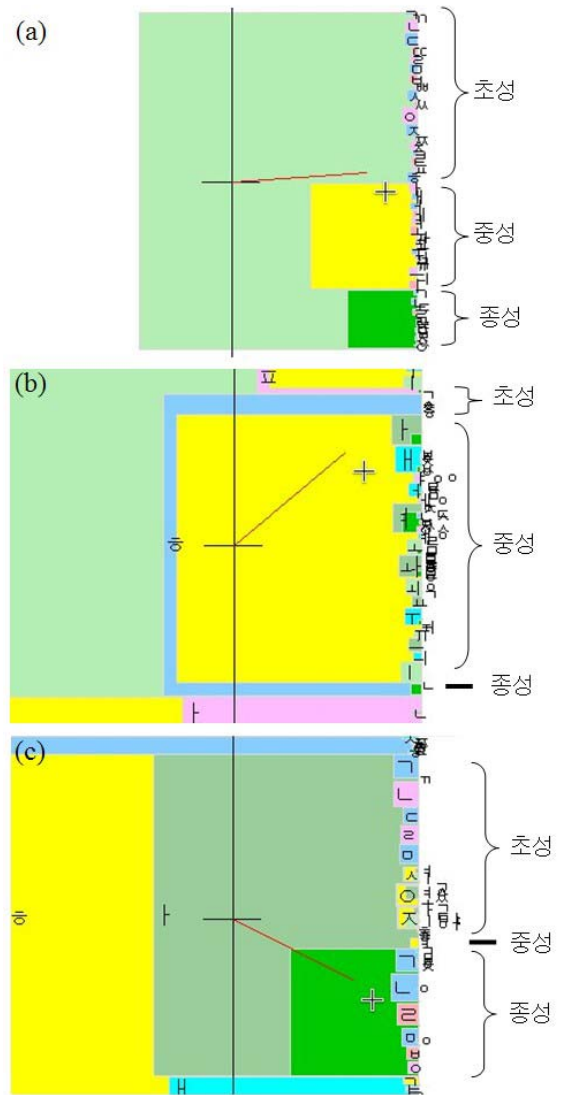


그림 2. Dasher에서의 한글 ‘한’ 입력

## 4. 실험

3.2에서 설명한 한글 조합 방법들에 대한 언어 모델의 성능을 평가하기 위해서 다음 3가지 방법에 대해 n-gram 언어 모델을 이용하여 cross-entropy를 계산하였다.

### 모든 한글 자모 이용 (그림 4. (a))

초성 19개 + 중성 21개 + 종성 27개 = 총 67개 토큰

모든 한글 자모 + <null> (중성이 없는 경우) (그림 4. (b))

67개 자모 + <null> = 총 68개 토큰

2벌식 키보드 (그림 4. (c))

자음 19개 + 모음 14개 = 총 33개 토큰

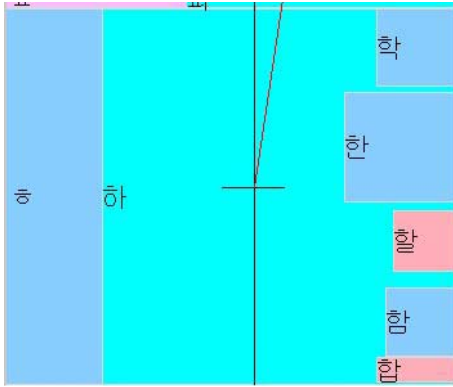


그림 3. 음절 단위로 조합한 예

실험에는 '97 KAIST raw text corpus 중 뉴스나 잡지에서 추출된 텍스트들을 이용하였고, 훈련과 테스트에 각각 공백과 특수문자, 숫자 등이 제외된 약 1300만자, 약 300만자의 텍스트가 이용되었다. 표 1은 테스트 데이터에 대한 cross-entropy 결과를 보여준다.

표 1. Cross-entropy 결과

n	모든 자모	모든 자모 + <null>	2벌식 키보드
3	3.04	2.73	3.05
4	2.73	2.49	2.70
5	2.45	2.27	2.42
6	2.28	2.11	2.24
7	2.20	1.99	2.14
8	2.17	1.93	2.11
9	2.18	1.92	2.11

5. 결론

본 연구에서는 Dasher를 바탕으로 연속적인 제스처와 언어 모델을 사용하고, 한글의 특성을 반영하는 한글 입력 인터페이스를 제안하였다. 3가지 한글 조합 방법을 실험 결과, 한글 음절 하나를 입력하기 위해서 초성, 중성, 종성 또는 <null> 토큰을 차례로 한번씩 입력하는 경우의 언어 모델이 가장 좋은 cross-entropy 결과를 나타냄을 확인하였다.

위와 같은 한글 조합 방법을 이용한 한글 입력 인터페이스를 숙련자가 마우스로 이용할 경우 분당 19개 문자의 속도로 입력 가능성이 실험을 통해 확인되었다. 그러나 이 실험에서는 문서로 인쇄된 문장을 보고 입력하는 방법을 택했기 때문에, 사용자의 시선이 인터페이스로부터 분산될 수 밖에 없었다. 시각적 피드백을 계속해서 받을 수 있는 환경에서는 더욱 높은 입력속도를 보일 것으로 예상된다.

참고 문헌

- [1] J. Mankoff and G. D. Abowd, "Cirrin: A Word-level Unistroke Keyboard for Pen Input", Proc. ACM UIST, 1998.
- [2] K. Perlin, "Quickwriting: Continuous Stylus-based Text Entry", Proc. ACM UIST, 1998.
- [3] David J. Ward, et al, "Dasher - a Data Entry Interface Using Continuous Gestures and Language Models", Proc. ACM UIST, 2000.
- [4] J. Williamson and R. Murray-Smith, "Dynamics and Probabilistic Text Entry", Univ. of Glasgow, Tech. Rep., 2002.

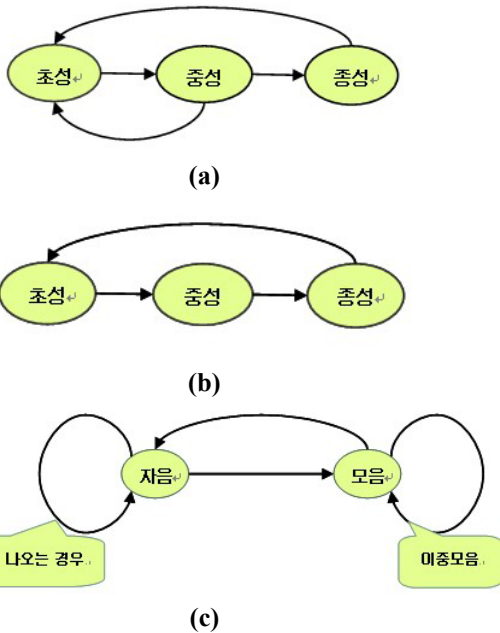


그림 4. 한글 조합 방법

실험 결과 언어 모델에서의 문맥의 길이에 관계 없이 모든 한글 자모와 <null> 토큰을 사용한 결과가 다른 방법보다 더 낮은 cross-entropy 값을 보였으며, 문맥의 길이가 8~9일 때 가장 낮은 cross-entropy 값을 보였다.