

Efficient Mobile Writing System with Korean Input Interface Based on Face Recognition

Jong-Hyun Kim*

*Professor, Dept. of Software Application, Kangnam University, Yongin, Korea

[Abstract]

The virtual Korean keyboard system is a method of inputting characters by touching a fixed position. This system is very inconvenient for people who have difficulty moving their fingers. To alleviate this problem, this paper proposes an efficient framework that enables keyboard input and handwriting through video and user motion obtained through the RGB camera of the mobile device. To develop this system, we use face recognition to calculate control coordinates from the input video, and develop an interface that can input and combine Hangeul using this coordinate value. The control position calculated based on face recognition acts as a pointer to select and transfer the letters on the keyboard, and finally combines the transmitted letters to integrate them to perform the Hangeul keyboard function. The result of this paper is an efficient writing system that utilizes face recognition technology, and using this system is expected to improve the communication and special education environment for people with physical disabilities as well as the general public.

▶ **Key words:** RGB camera, Mobile device, Face recognition, Korean input interface, Korean keyboard system, Special education environment

[요 약]

가상 키보드방식의 문자 입력 시스템은 고정된 버튼 위치를 터치하여 입력하는 방식으로 손가락의 움직임이 불편한 사람들이나 노인들은 입력하기가 매우 불편하다. 이러한 문제를 완화시키기 위해, 본 논문에서는 모바일 디바이스의 RGB 카메라를 통해 얻은 영상과 사용자의 움직임을 통해 자판 입력과 필기 기능이 가능한 효율적인 프레임워크를 제안한다. 이 시스템을 개발하기 위해 얼굴인식을 활용하여 입력 영상으로부터 컨트롤 좌표를 계산하고, 이 좌표 값을 이용하여 한글을 입력하고 조합할 수 있는 인터페이스 개발한다. 얼굴인식을 기반으로 계산된 컨트롤 위치는 자판의 글자를 선택하고 전달하는 포인터 역할을 하며, 마지막으로 전달된 글자들을 조합하여 한글 자판 기능을 수행할 수 있도록 통합한다. 본 연구의 결과는 얼굴인식 기술을 활용한 효율적인 필기 시스템이며, 이 시스템을 사용하면 일반인 뿐 만 아니라 신체가 불편한 지체장애인의 의사소통 및 특수교육 환경도 개선시킬 수 있을 거라 기대한다.

▶ **주제어:** RGB카메라, 모바일 디바이스, 얼굴인식, 한글 입력 인터페이스, 한글 키보드 시스템, 특수 교육환경

-
- First Author: Jong-Hyun Kim, Corresponding Author: Jong-Hyun Kim
 - *Jong-Hyun Kim (jonghyunkim@kangnam.ac.kr), Dept. of Software Application, Kangnam University
 - Received: 2020. 04. 20, Revised: 2020. 05. 25, Accepted: 2020. 06. 01.

I. Introduction

손가락을 이용한 터치 기반 한글 입력 디스플레이는 대부분의 입력 시스템에서 자주 활용되는 방식이다 [1,2,3]. 이 방법은 버튼의 위치를 손가락으로 터치하면 해당 문자가 입력되는 방식이며, 이 과정을 반복해서 좀 더 길고 복잡한 문장을 입력한다. 하지만 노인이나 장애인들은 이처럼 일반적인 한글 입력 방식을 사용할 때 몇 가지 불편한 점이 있다. 모바일 디바이스에서 활용되는 문자 입력 시스템은 입력 버튼의 위치가 고정되어 있으며 기본적으로 버튼의 위치를 인식하는 정보를 기반으로 문자를 입력하고 있기 때문에 장애인들이 사용하기에 매우 불편한 구조이다 : 1) 입력 버튼의 위치를 시각적으로 확인하기 어렵고, 이 과정에서 손가락으로 원하는 문자를 입력한다는 것은 더욱더 어려운 일이다. 특히, 2) 작은 디스플레이 안에 오밀조밀하게 모여져 있는 자판 배열을 손가락 하나만으로 조절한다는 것은 노인이나 장애인들에게는 힘든 일이다.

지체장애를 갖고 있는 사람은 일반인과는 달리 일상에서 무언가를 쓰거나 메시지를 보내는 것이 매우 불편하다. 이런 불편함은 일상적인 메시지의 전달 뿐만 아니라 지체장애를 앓는 특수교육 환경에서도 영향을 끼친다. 이러한 문제를 보상해주는 지원에는 1) 타인이 대화를 듣고 적어주는 제도나, 2) 네트워크를 통한 문자통역사 시스템이 있지만, 이것은 일방적으로 대화를 보여줄 뿐이며 아직까지 지체장애인을 위한 완벽한 타자 입력 시스템은 없다.



Fig. 1. Share-typing (text interpretation app for the hearing impaired)

최근에 쉐어타이핑(Share-typing)이라는 문자 공유 시스템이 SK주식회사, 마이크로소프트, 현대 자동차 등에서 활용되고 있지만 그 기능이 청각장애인들을 위한 실시간 자

막 서비스로만 한정되어 입력/필기 시스템으로는 충분하지 않다 [4] (Figure 1 참조). 이 시스템은 말소리를 모바일 디바이스에서 보여주는 기능으로, 세미나, 포럼, 교회 등 다양한 장소에서 문자통역사, 자원봉사자들이 타이핑하는 내용을 자막으로 제공하여 청각장애인들의 소통을 지원하는 프레임워크이다. 이러한 문제를 해소하기 위한 대안으로 음성인식 기반 접근법들이 제안되었지만 여전히 나랏글, 천지인과 같은 기존의 한글 입력 방식을 선호하고 있다. 뿐만 아니라, 한글 입력 인터페이스는 제조사마다 다르기 때문에 입력 인터페이스 방식을 모두 학습하고 교육받기에는 무리가 있다. 통합된 환경이 없는 입력 시스템의 부재는 노인이나 장애인들에게는 기본욕구의 좌절 등에서 초래되는 부정적인 심리를 감소시키지 못하기 때문에 사회적 문제로 야기될 수도 있다. 본 논문에서는 이 문제를 최소화하기 위해 지체장애인들도 쉽게 한글 자판 입력 시스템을 활용할 수 있는 새로운 인터페이스를 제안한다. 이러한 인터페이스는 사람간의 대화나 관계 뿐만 아니라 장애인들을 위한 특수 교육 환경에서의 자판 입력으로 활용될 수 있다.

II. Preliminaries

1. Related works

문자 입력 방식 중 천지인 방식은 천(·), 지(-), 인(1)이라는 3가지 조합을 이용하여 모든 한글의 모음을 입력할 수 있는 방법이다 [5]. 하지만 키보드와 다르게 같은 문자임에도 불구하고 천지인 방식은 모음을 입력하는데 필요한 입력 동작이 많기 때문에 몸이 불편한 노인이나 장애인들은 사용이 어렵다. 뿐만 아니라 동작 수의 증가는 오타를 입력 할 가능성 또한 높아지기 때문에 장애인들은 선호하지 않는 방식이다.

시각장애인들을 위한 대표적인 방식이 점자 입력 방식이다. 이 방식은 디스플레이 영역을 여섯 곳으로 구분하여 시각적으로 버튼의 위치를 인지하기 어려운 장애인들을 위한 문자 입력 대체기술로 제안되었다 [3,6]. 하지만, 점자 인식만으로 문자를 해석하는데 한계가 있어서 그 사용은 파급적이지 못했다 [7]. 손가락의 복잡한 움직임은 단순화하고자 적외선 센서를 이용한 하드웨어 키보드 기법이 제안되었지만, 무게가 큰 하드웨어를 직접 들고 다녀야하기 때문에 이동성에 제약을 받는다 [8,9].

쉐어타이핑과 유사한 접근법으로 디바이스가 입력된 음성을 문자로 변환해주는 음성인식 방법이 있다 [10,11]. 하지만, 정확한 음성인식은 소형 단말기에서는 처리하기

힘든 만큼의 처리성을 요구하기 때문에 외부 서버를 이용하는 경우가 많다. 이러한 단점을 해결하고자 사용자가 직접 음성을 문자로 변환해주는 수동 서비스가 있지만 이 역시 매우 제한적이다. 또 다른 문자 입력 방식으로는 스캐치 기반 문자 입력 기법이 있다. 그 중에서 Swype는 현재 안드로이드 플랫폼 기반 모바일 디바이스에서 사용되는 방식으로 쿼티(QWERTY) 자판 위를 스캐치하듯 드래그하여 문자를 입력하는 방식이다 [12]. 이 외에도 좀 더 직관적인 스캐치패턴을 위해 원형의 키보드 인터페이스 방식을 이용한 방법들도 있다 [13]. Costagliola 등은 서버 메뉴 방식을 이용하여 빈번하게 사용하는 메뉴를 추가적으로 등록시키는 접근법이다 [14]. 이 인터페이스는 기존의 쿼티 자판에서 하나의 키를 눌렀을 경우, 선택된 키를 중심으로 서브메뉴가 나타나는 방식이다.

최근에는 기계학습을 이용하여 음성인식을 처리하는 기법들은 RNN(Recurrent natural networks)과 GAN(Generative adversarial network) 기반으로 개발되고 있다 [15,16,20,22]. 이 접근법은 은닉층이 재귀적인 구조를 갖는 인공신경망이며, 기존 신경망은 시간이 길어질수록 경사도 소실 문제(Vanishing gradient problem)가 발생한다. 이 문제를 완화시키고자 장기 의존성(Long-term dependency)을 학습시킬 수 있는 LSTM(Long-short term memory) 등이 제안되었으며, 이 기법도 음성 인식 및 자연어 처리에 널리 활용되고 있다 [17,18,21]. 이러한 다양한 네트워크 방법을 통해 문자 입력 및 인식 분야에서는 정확도 및 활용성이 획기적으로 개선되고 확장되었지만, 사용자 인터페이스 측면에서는 눈에 띄게 개선된 사항이 없으며, 유니버설(Universal)하게 다양한 사용자들을 수용하기에는 인터페이스의 개선이 필요한 상황이다.

III. The Proposed Scheme

제안하는 시스템 개요는 모바일 디바이스를 통해 사용자의 얼굴영상을 입력 받아 아래와 같은 순서로 수행된다.

1. 모바일 디바이스의 전면 RGB 카메라를 통해 얻어진 사용자의 얼굴 영상을 이용하여 얼굴 좌표 계산 : 얼굴 인식 과정을 통해 얼굴의 위치변화를 추적하여 화면상에 포인터로 사용
2. 포인터를 이용한 문자 입력 : 모바일 디바이스의 화면상에 가상 키보드를 만들어 얼굴 추적으로 계산된

포인터의 위치와 인접한 문자를 선택

- 3 문자들을 이용한 한글 조합 : 입력된 자음과 모음을 조합하여 한글 문장으로 표현

1. Calculation of facial coordinates with RGB camera of mobile device

본 연구에서는 모바일 디바이스의 전면 RGB 카메라를 이용하여 사용자가 모바일 디바이스에서 타자를 입력할 수 있도록 얼굴의 좌표를 계산한다. 제안하는 기법은 RGB 카메라로부터 얻어진 영상을 이용한다는 부분에서 모바일 디바이스 뿐 만 아니라 RGB 카메라가 내장되어 있는 노트북에서도 활용이 가능하다. 입력으로 들어온 영상 기반으로 얼굴 좌표를 계산하기 위해 먼저 얼굴인식 과정을 거친다. 일반적으로 입력영상은 주변 조명과 잡음이 포함되어 있기 때문에 이를 최소화하기 위한 2가지 전처리 과정을 거친다. 영상이나 사진에서는 같은 장소임에도 불구하고 불규칙한 자연광이나 조명이 다르기 때문에 영상 내 포함되어 있는 빛의 양을 조절해야 잡음을 줄일 수 있다. 우리는 회색계 알고리즘(Grey world algorithm)[19]을 이용하여 영상 내 빛 보정을 수행한다 (수식 1 참조).

$$\hat{I}_r(x,y) = \alpha I_r(x,y), \hat{I}_b(x,y) = \beta I_b(x,y) \quad (1)$$

$$\alpha = \frac{g_{avg}}{r_{avg}}, \beta = \frac{g_{avg}}{b_{avg}} \quad (2)$$

수식 2에서 $r_{avg}, g_{avg}, b_{avg}$ 는 영상의 각 채널 평균이며, 여기서 g_{avg} 는 이미지의 밝기 채널로 사용된다. 역시 밝기를 가정하고 있는 이유는 r 과 b 의 보정 계수를 이용하여 g 의 밝기를 최종적으로 보정하기 위함이다. 결과적으로, 전체 영상의 빛의 평균값을 이용하여 빛의 차이를 줄이는 과정이다. 우리는 추가적으로 모폴로지(Morphology)를 사용하여 영상의 잡음을 한 번 더 제거한다. 이러한 과정을 통해 잡음이 제거된 영상으로부터 사용자의 눈, 코, 입을 추출한다. 이렇게 추출된 눈, 코, 입으로부터 경계상자(Axis aligned bounding box)를 생성하고, 이 상자의 중심 위치를 모바일 디바이스 상에서의 포인터 위치로 설정한다.

Figure 2는 제안된 기법을 이용하여 얼굴을 인식한 결과이며, 인식과정에서 얼굴의 위치를 움직였을 때도 안정적으로 얼굴의 경계상자를 계산하였다. Figure 2a에 삽입된 작은 이미지는 눈, 코, 입을 중심으로 인식한 결과이며, 이 방식을 이용하여 포인터 위치를 결정해도 된다. 하지만, 이 방식은 눈, 코 등을 개별적으로 인식하기 때문에 포

인터의 역할을 하기에는 다소 불안정한 부분이 존재한다. 약간만 움직여도 떨리는 부분이 있기 때문이며 이 문제를 완화시키기 위해 경계상자를 이용하였다. 실제로 경계상자는 눈, 코, 입 전체를 아우르는 상자으로써 내부의 작은 움직임보다는 전체 경계의 크기가 변하느냐가 더 중요하기 때문에 상대적으로 눈, 코, 입을 개별적으로 계산하는 것 보다 안정적이다 (Figure 2 참조). 본 논문에서는 문자를 입력하기 위해 얼굴인식으로부터 계산된 경계상자의 중심 위치를 포인터 위치로 이용한다.

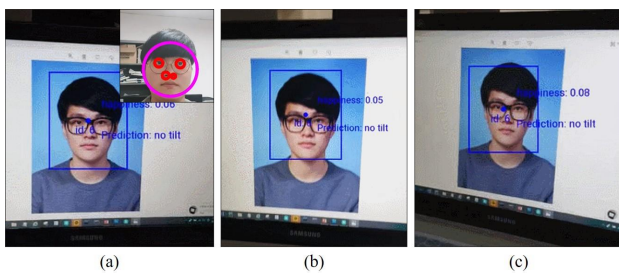


Fig. 2. Facial recognition with our method

2. Character input using pointers based on face coordinates

얼굴인식을 활용하여 계산한 포인터의 위치를 이용하여 모바일 디바이스에서 문자를 입력하기 위해 이번 장에서는 포인터 위치에 반응하는 인터페이스를 제작한다. 포인터의 위치가 모바일 디바이스 화면에서 오른쪽으로 이동하면 가상 키보드가 자동으로 가시화되도록 만들었으며, 키보드 상에서 포인터 위치에 따라 가상 키보드를 쉽게 입력 할 수 있도록 확대/축소하는 방식을 이용하였다.

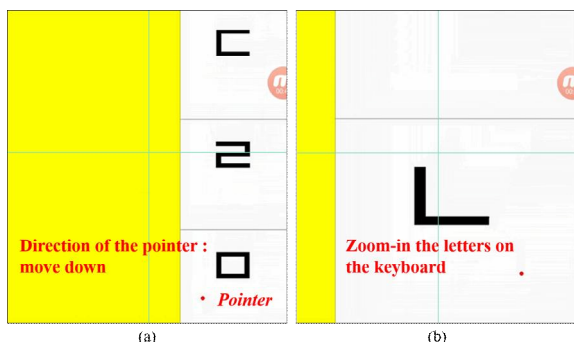


Fig. 3. Move and zoom-in the keyboard depending on the position of the pointer : (a) move, (b) zoom-in, red dot : pointer

실제로 쿼티 방식이나 자판 버튼 방식들은 작은 화면에 조밀하게 배치된 문자 버튼이기 때문에 노인들이나 장애

인들이 사용하는 것이 쉽지 않았다. 하지만 제안하는 인터페이스는 간단한 포인터 방식으로 인해 화면 확대/축소가 가능하여 시력이 나쁜 사람들에게도 쉽게 문자 입력이 가능하도록 했다.

Figure 3에서 보듯이 사용자의 포인터가 모바일 디바이스 화면의 오른쪽 구역의 일정 범위를 넘으면 가상 키보드가 나타난다. 이 때 포인터의 위치를 위/아래쪽 방향으로 움직이기 되면 실제로 자판이 위/아래로 넘어가면서 사용자들이 쉽게 문자열을 선택할 수 있도록 한다. 디바이스의 화면에서 위/아래의 선택 기준은 디스플레이 중심 위치를 기준으로 위쪽이라면 위쪽 방향으로 스크롤링하고, 아래쪽 방향도 마찬가지로 기법으로 적용한다. 이때 중요한 것이 스크롤의 속도이며, 본 논문에서는 디스플레이의 중심 위치를 기준으로 속도를 서서히 감소하도록 했다 (수식 3 참조).

$$scroll_{speed} = \tanh(p_y)\sigma \quad (3)$$

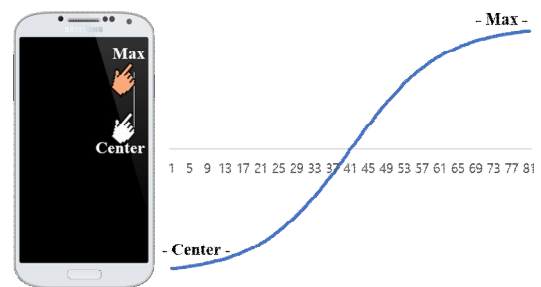


Fig. 4. Controlling scroll speed

수식 3은 시그모이드(Sigmoid) 함수와 유사한 형태로 디스플레이의 중심 위치에서 화면 끝 쪽으로 갈수록 속도를 서서히 감소하여 사용자가 문자를 선택하는데 있어서 좀 더 수월하도록 모델링했다 (Figure 4 참조). 여기서 σ 는 스크롤의 속도를 제어하는 변수로, 우리는 2.0으로 설정했다.

3. Hangeul combination using characters

한글은 연속된 문자들을 일정 규칙에 맞게 조합되며, 이러한 특징을 기반으로 본 논문에서는 다음과 같은 규칙을 사용한다 : 일반적으로 초성과 종성의 경우나 연속된 입력의 조합이 한글체계에서 유효하면 문자를 조합하고, 그렇지 않으면 조합하지 않는다 (Figure 5 참조).

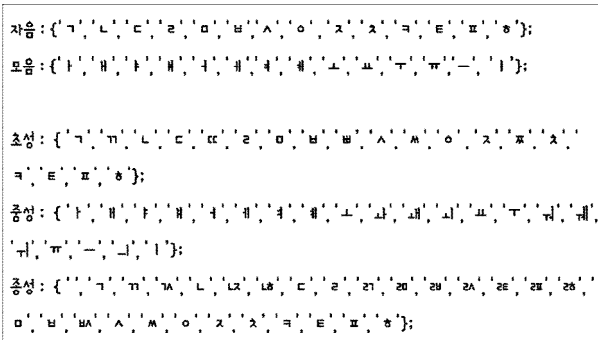


Fig. 5. Pattern for Hangeul combination

이 인터페이스는 된소리를 자음체계에 포함시킨 것이 아니고, 같은 자음이 연속된 입력을 통해 된소리를 생성하였기 때문에 자음이 연속적으로 입력되었을 때 조합 가능한 유무를 계산한다. 연속된 문자들의 입력을 각각 자음과 모음으로 분류하여, 이전에 입력했던 문자가 모음인지 자음인지 여부에 따라 조합 조건을 계산한다. 일반적으로 한글은 유니코드로 표현되기 때문에 본 논문에서는 유니코드를 조합하여 화면에 한글을 표시하였다.

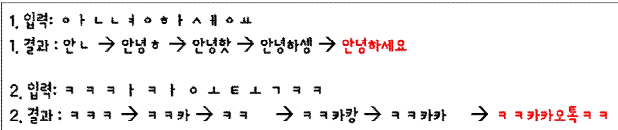


Fig. 6. Hangeul sentences made up of characters combination

Figure 6은 본 논문에서 제안한 방법을 이용하여 입력된 문자들로부터 한글을 조합한 결과이다. 한글은 초성, 중성, 종성의 결합으로 이루어지며, 초성은 자음과 자음의 조합, 중성은 모음과 모음의 조합, 종성은 자음과 자음의 결합으로 계산한다.

IV. Results

본 연구의 결과들을 만들기 위해 실험한 환경은 Intel Core i7-7700K CPU, 32GB RAM, Geforce GTX 1080Ti GPU가 탑재된 컴퓨터를 이용하였다. 본 논문에서는 모바일 디바이스 환경에서 얼굴인식을 통해 계산한 포인터 위치를 이용하여 한글을 입력할 수 있는 새로운 한글 자판 인터페이스를 제안했다.

Figure 7은 본 논문에서 제안하는 방법을 실제로 사용하는 장면으로써 모바일 디바이스의 전면 RGB 카메라를 이용하여 사용자가 쉽게 한글 자판 입력이 가능하도록 했다.

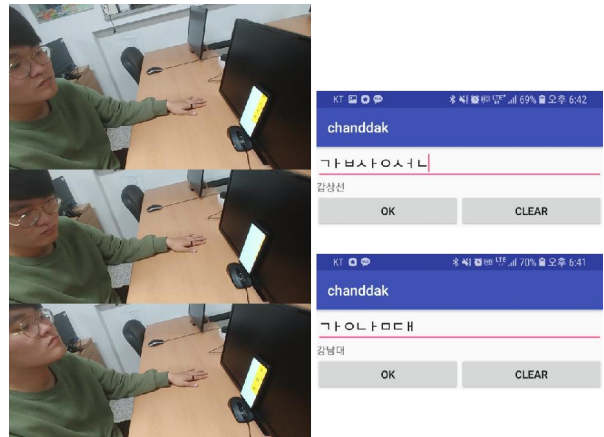


Fig. 7. Sentence result created through face recognition-based Hangeul keyboard

우리는 제안한 방법이 실제 자판 입력을 이용하는 콘텐츠에 어떠한 영향을 미치는지를 이해하기 위해 다수의 사용자 실험을 진행했다. 사용자 실험을 진행하기 위해 20명의 인원들에 대해서 본 논문에서 제안하는 방법을 이용하게 했다 (일반인 10명, 장애인 10명). 제안하는 방법을 이용하여 한글 자판을 입력할 수 있는 조작법에 대해 짧게 설명한 후, 참가자들은 본 논문에서 제안하는 모바일 디바이스에서 수행되는 사용자 인터페이스를 이용하여 문장을 입력한다. 각 문장은 한글로만 설정하였으며 짧고/긴 문장들을 자유롭게 입력하도록 했다.

Table 1. Q&A form (Q1,Q4 : satisfaction(5 points), regular(3 points), dissatisfaction(0 points), Q2 : easy(5 points), regular(3 points), difficulty(0 points), Q3 : none(5 points), less than 3(3 points), 5 or more(0 points))

Q1	Was sentence input using the face recognition-based Hangeul keyboard interface more interesting than using the finger interface?
Q2	Was the face recognition based interface difficult to operate??
Q3	Was there a lot of typos in typing the sentence?
Q4	(Only students with disabilities) Do you think the face recognition-based interface helped your disability?

우리는 짧은 문장 10개, 긴 문장 10개를 주어주고 똑같이 입력하도록 했으며, 이 과정에서 시간제한은 두지 않았다. 향후, 제안된 기법을 입력된 문장의 정확성과 편리성 측면에서 분석을 했기 때문에 시간제한은 두지 않았으며, 손가락이 아닌 얼굴인식 기반이기에 시간제한은 큰 의미가 없다고 판단하였다.

참가자들에게는 20개의 문장을 입력 할 때 1) 손가락을 통해 입력하는 일반적인 핑거 방식과 2) 본 논문에서 제안한 인터페이스를 이용하는 방식을 모두 사용한 뒤, 아래와 같은 설문지를 작성하도록 요청했다 (Table 1 참고).

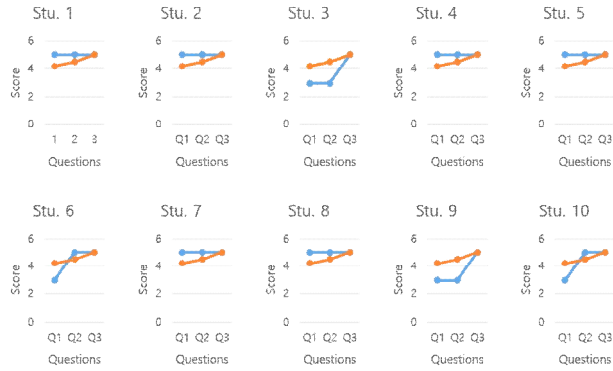


Fig. 8. Survey of normal students (blue line : score for the question, orange line : avg. score for the question)

얼굴인식 기반 한글 자판 입력 시스템이 문자를 입력하는데 있어서 어떠한 영향을 미치는지를 조사했으며, 단순한 문자 입력 측면이 아니라, 신체가 불편한 장애인들도 제안하는 인터페이스를 어떻게 평가하는지를 알기 위해 일반인과 장애인을 분류해서 설문조사를 진행했다.

Figure 8는 일반인들을 상대로 조사한 결과이다. 모든 질문에 대해서 5점 만점에 4.5가 나왔으며, 대부분 제안하는 자판 입력 인터페이스에 만족하였다. 손가락만 이용하던 기존의 작동방식과 비교해 신선하고 새로웠다는 의견이 있었으며, 작동방식의 편리성 측면에서는 다소 불편했다는 의견이 있었다. 일반 학생들은 손가락을 이용한 터치 기반에 익숙해져 있기에 오히려 얼굴인식 기반 문자 입력 방식이 좀 더 불편해 보였다. 하지만, 익숙함에 대한 문제이지, 작동방식이 어려웠다는 의견은 거의 없었으며, 긴 문장을 지속적으로 입력할 경우 목이 무리가 갈 수 있다는 의견이 있었다.

Figure 9는 장애인들은 대상으로 조사한 결과이며, 일반인과 유사하게 대부분의 결과에서 긍정적인 평가를 받았다. Q4는 장애인들에게만 주어진 질문으로써 제안하는 자판 입력 인터페이스가 본인의 장애에 얼마큼이나 긍정적으로 영향을 끼쳤는지를 조사하기 위함이며, 대부분의 장애인들에 대해서 좋은 점수를 받았다.



Fig. 9. Survey of handicapped students (blue line : score for the question, orange line : avg. score for the question)

이 질문에 대해서는 개인차가 있었으며, 제안하는 인터페이스가 자기 자신이 앓고 있는 장애에 도움이 되지 않는다는 생각하는 경우 다소 낮은 점수로 평가되는 경우가 종종 있었다 (Figure 9에서의 Stu. 2와 Stu. 8 참조). 그럼에도 불구하고 대부분의 문자 입력 방식에 대해서는 일반인과 유사하게 긍정적인 평가를 받았다.

V. Conclusions

본 논문에서는 얼굴인식을 이용하여 컨트롤 위치를 계산하고 이로부터 한글 자판을 입력할 수 있는 새로운 인터페이스 프레임워크를 제안했다. 입력 영상은 RGB 카메라로부터 얻었으며, 조명과 잡음으로부터 표현되는 문제를 줄이기 위해 회색계 알고리즘과 모폴로지 필터를 활용하였다. 결과적으로 RGB 카메라로부터 얼굴을 인식하고 작은 움직임에도 떨어지는 컨트롤 위치가 아닌 좀 더 안정적으로 위치를 계산하기 위해 눈, 코, 입에 해당하는 경계상자의 중심위치를 포인터 위치를 설정하였다. 우리는 노트북과 모바일 디바이스에 내장된 RGB 카메라에서 실험했으며 안정적으로 문자를 입력하는데 성공했다.

제안한 기법은 일반인들보다는 몸이 불편한 장애인들을 위한 기능이다. 일반인들은 실제로 손가락으로 자유롭게 터치가 가능하기 때문에 굳이 얼굴인식 기반 입력 시스템이 필요 없지만, 몸이 불편한 사람이나 노인들은 손가락의 디테일한 움직임을 통해 입력하는 것이 어렵기 때문에 제안하는 기법은 ICT복지기술에서 의미가 크다고 생각된다. 뿐만 아니라, 입력하고 싶은 문자의 크기가 포인터 위치가 따라 확대/축소가 되기 때문에 시력이 낮은 노인에게도 큰 도움이 될 것이라고 생각된다.

하지만, 여전히 몇 가지 단점이 있다. 앞서서도 언급했듯이 긴 문장을 지속적으로 입력할 경우 손가락 보다 근육에 무리가 될 수 있다. 실제로 우리는 일반 학생들에게 아주 긴 문장을 입력하게 했으며 목에 무리가 된다는 의견을 받았다. 향후, 얼굴인식으로만 의존하는 입력 시스템이 아니라, 다양한 움직임을 추가적으로 접목하여 좀 더 효율적인 인터페이스로 확장할 예정이다. 또 하나의 단점은 오타가 생긴 경우 자동으로 교정해주는 방법이 현재 프레임워크에는 포함되어 있지 않다. 손가락을 이용한 터치기반 인터페이스는 오타가 생긴 경우 해당 단어를 지우고 다시 작성하면 되지만, 얼굴인식 기반인 경우에는 오타가 난 부분까지 포인터 위치를 이동해서 지우고, 문장을 재입력하는 것은 매우 부담스러운 작업이다.



Fig. 10. Typing error

실제로 Figure 10은 얼굴인식 기반 자판 입력 인터페이스를 이용하던 중 나타난 오타이다. 사용자는 “까마귀”를 입력하고 싶었지만 첫 글자에 “ㄱ”이 세 번 이상 입력되기 때문에 한글 조합이 되지 않는 결과이다. 이 오타를 제거하기 위해서 얼굴인식을 통해 포인터 위치를 다시 처음으로 이동시키는 것은 매우 번거로운 과정이며, 현재 시스템에서는 이 부분을 자동으로 보정해주는 과정은 없다. 마지막으로, 제안하는 방법은 문자 입력의 정확도가 결과적으로 얼굴인식에 의존하는 시스템이다. 물론, 눈이나 코와 같은 작은 특징을 이용하기 보다는 특징들은 경계상자로 표현하기 때문에 정확도 측면에서 민감하지는 않지만 향후 정확도를 개선시킬 수 있는 방법으로 개선이 필요할 것으로 생각된다. 향후, 이 문제를 자동화하여 얼굴인식 기반 자판 입력 인터페이스에서 목 근육에 무리가 가는 움직임을 최소화하는 방법으로 최적화를 할 예정이다.

REFERENCES

- [1] S. E. Park, S. Y. Moon, H. S. Yoon, D. C. Shin, and H. S. Jeong, "An Input Method for Hangul Character and Implementation for Personal Digital Assistant," Proc. of Fall Conference of KIISE, vol. 21, no. 2, pp. 657-660, 1994.
- [2] W. H. Choi, S. C. Kim, and K. R. Song, "Apparatus and Method for Inputting the Korean Alphabet Based on the Character-Generative Principles of HUNMINJUNGEUM," Patent no. 1999-052648, 1999.
- [3] J. W. Kim, "Apparatus for Inputting Key and Method for Inputting Character by Direction Key," Patent no. 2001-0006208, 2001.
- [4] Share-typing, <https://www.sharetyping.com/>.
- [5] Kang, Seung-Shik, Hahn, Kwangsoo, "Hangul Vowel Input System for Hand-held Devices ", KIPS Transactions on Software and Data Engineering, pp. 507-512, 2005. DOI : 10.3745/KIPSTB.2005.12B.4.507.
- [6] T. J. Eom, J. B. Lee, and B. K. Kim, "Design of Smart Phone-Based Braille Keyboard System for Visually Impaired People," Journal of Information Security, vol. 12, no. 1, pp. 63-70, 2012.
- [7] Y. C. Byun, "2011 Survey on Impaired People," Research Report 2011-82, 2012.
- [8] H. Y. Kim, Gesture Recognition Interface by Keyboard Insertion IR Module, Master Thesis, KAIST, 2011.
- [9] D. H. Kim, Y. H. Kwon, and J. H. Kim, "A Hangul Input Interface Using Continuous Gestures and Language Models," Proc. of KCC2005, pp. 586-588, 2005.
- [10] Won Il Lee, "Natural Language Processing for Korean Speech Recognition," Proc. of KCC2005, pp. 586-588, 2005.
- [11] Sang Hun Jeong, "The Study on Phonetical Information for Speech Recognition", The Contents and Language Phenomena of Seokbosangjeol, vol. 49, pp.135-160, 2007.
- [12] Keyboard for Android : <http://swypeinc.com/>.
- [13] Jared Cechanowicz, Steven Dawson, Matt Victor, Sriram Subramanian, "Stylus based text input using expanding CIRRI," AVI '06, May. 2006. DOI : 10.1145/1133265.1133299.
- [14] Gennaro Costagliola, Vittorio Fuccella, Michele DiCapua, Giovanni Guardi, "Performances of Multiple-Selection Enabled Menus in Soft Keyboards", DMS'09 - The 15th International Conference on Distributed Multimedia Systems., pp. 359-364, 2009.
- [15] Amberkar, Aditya, Parikshit Awasarmol, Gaurav Deshmukh, and Piyush Dave. "Speech Recognition using Recurrent Neural Networks." In International Conference on Current Trends towards Converging Technologies, pp. 1-4, 2018. DOI : 10.1109/ICCTCT.2018.8551185.
- [16] Lim, Chee Peng, Siew Chan Woo, Aun Sim Loh, and Rohaizan Osman. "Speech recognition using artificial neural networks." In Proceedings of the First International Conference on Web Information Systems Engineering, vol. 1, pp. 419-423, 2000. DOI : 10.1109/ICCSP.2018.8524333.
- [17] Graves, Alex, Navdeep Jaitly, and Abdel-rahman Mohamed. "Hybrid speech recognition with deep bidirectional LSTM." In IEEE workshop on automatic speech recognition and

- understanding, pp. 273-278, 2013. DOI : 10.1109/ASRU.2013.6707742.
- [18] Sundermeyer, Martin, Ralf Schlüter, and Hermann Ney. "LSTM neural networks for language modeling." In Thirteenth annual conference of the international speech communication association. 2012.
- [19] Wang-Jun Kyung, Dae-Chul Kim, Yeong-Ho Ha, "Enhancement of Faded Images Using Integrated Compensation Coefficients Based on Multi-Scale Gray World Algorithm", vol. 39, pp. 459-466, The Journal of Korea Information and Communications Society, 2012. DOI : 10.7840/kics.2014.39A.8.459
- [20] Schneider, Steffen and Baevski, Alexei and Collobert, Ronan and Auli, Michael, "wav2vec: Unsupervised pre-training for speech recognition", arXiv preprint arXiv:1904.05862, 2019.
- [20] Schneider, Steffen and Baevski, Alexei and Collobert, Ronan and Auli, Michael, "wav2vec: Unsupervised pre-training for speech recognition", arXiv preprint arXiv:1904.05862, 2019.
- [21] Tian, Xu and Zhang, Jun and Ma, Zejun and He, Yi and Wei, Juan and Wu, Peihao and Situ, Wenchang and Li, Shuai and Zhang, Yang, "Deep LSTM for large vocabulary continuous speech recognition, arXiv preprint arXiv:1703.07090, 2017.
- [22] Sriram, Anuroop and Jun, Heewoo and Gaur, Yashesh and Satheesh, Sanjeev, "Robust speech recognition using generative adversarial networks, 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5639-5643. 2017. DOI : 10.1109/ICASSP.2018.8462456

Authors



Jong-Hyun Kim received the B.A. degree in the department of digital contents at Sejong University in 2008. He received M.S. and Ph.D. degrees in the department of computer science and engineering at Korea University,

in 2010 and 2016. Prof. Kim is an assistant professor in the department of software application in Kangnam University. His current research interests include fluid animation and virtual reality.