

# 스마트 글라스 기반의 사회적 수용성을 고려한 사용자 인터페이스 기준 연구 분석

이민호\*, 허환\*, 김제우\*

\*전자부품연구원 지능형영상처리연구센터

e-mail : {amino7, gighks, jwkim}@keti.re.kr

## A Survey on User Interface Considering Social Acceptability based on Smart Glasses

Minho Lee\*, Hwan Heo\*, Jaewoo Kim\*

\*Intelligent Image Processing Research Center, Korea Electronics Technology Institute

### 요약

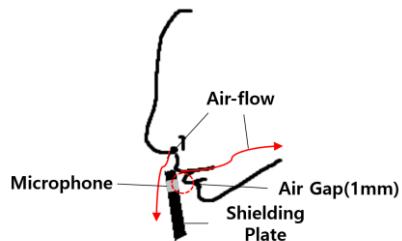
Augmented Reality(AR) 기반의 스마트 글라스가 제공하는 일반적인 사용자 인터페이스는 카페 혹은 도서관 등과 같은 공공장소에서 사용이 꺼려지는 경향이 있다. 콘텐츠 조작 시 주변 사람들에게 관심을 받거나 프라이버시를 지키기 어렵기 때문이다. 이런 문제는 사회적으로 수용 가능한 범위 내에서의 사용자 인터페이스를 수행함으로써 해결할 수 있다. 본 논문에서는 스마트 글라스에서 사회적 수용성을 고려한 사용자 인터페이스의 기준 연구들을 분석하고자 한다.

### 1. 서론

카페 혹은 도서관 등과 같은 공공장소에서 전용 컨트롤러, 터치 패드, 음성 인식 및 Mid-Air 제스처 등과 같은 스마트 글라스의 일반적인 사용자 인터페이스 사용 시, 사용자 혹은 주변 사람들 모두 부담감과 어색함을 느낄 수 있다. 사회적으로 수용하기 어렵다고 느끼기 때문이다. 이와 같은 문제점을 해결하고자 사회적 수용성을 고려한 사용자 인터페이스에 관련된 많은 연구가 진행되고 있다. 그 중 기준에 행하여지고 있는 연구로는, 숨을 들이쉬며 말하기를 통해 음성을 인식하는 사일런트보이스[1], 말하지 않고 입 움직임 인식을 통해 기기와 상호작용하는 Lip-Interact[2], 귀 혹은 턱과 뺨을 터치하는 Hand-to-Face[3] 등이 있다. 본 논문에서는 이러한 연구들을 통해 사회적 수용성이 고려된 사용자 인터페이스 기술 동향에 대해 알아보고, 해당 기술들이 갖고 있는 특징과 그 결과에 대해 분석 및 기술해 본다.

### 2. 사회적 수용성을 고려한 기준 연구 분석

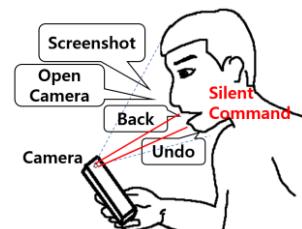
#### 2.1 사일런트보이스



(그림 1) 사일런트보이스의 Ingressive Speech

사일런트보이스는 새로운 음성 입력 인터페이스 장비이다. 일반적으로 아주 작은 음성을 입력하기 위해서는 마이크와 사용자의 입 사이의 거리를 매우 가깝게 하고 숨을 내뱉으며 말하는 “Egressive Speech” 방법을 이용한다. 이 경우에 팝-노이즈가 생기고 음성 인식의 정확도가 떨어지게 된다. 사일런트보이스는 음성을 입력할 때 그림 1과 같이 숨을 들이쉬며 말하는 “Ingressive Speech” 방법을 사용하여 음성 인식을 수행함으로써, 팝-노이즈 현상을 극복하면서 초소형(ultra-small) 음성 인식도 가능하며, 소란스러운 장소나 사무실, 또는 집에서 방해 없이 사용이 가능하다. 하지만 개인화된 시스템을 구축하기 위해서는 사일런트보이스의 특정 발음 사전을 만들면서 동시에 대규모 데이터 수집이 요구된다. 또한 사용 시 마이크를 입에 붙이고 사용해야하는 불편함도 존재한다.

#### 2.2 Lip-Interact



(그림 2) Lip-Interact 의 개요

그림 2 와 같이 Lip-Interact 는 사용자가 목소리를 내어 말하지 않고 입 움직임만으로 스마트폰에 명령을 실행할 수 있는 기술이다. 전면 카메라를 이용하여 사용자의 입 움직임을 취득하고, End-to-End 딥 러닝

모델을 이용하여 입력된 입 움직임을 인식한다. 이 시스템은 시스템-레벨 기능(앱 시작, 시스템 설정 변경 및 팝업창 설정) 및 응용 프로그램-레벨 기능(두 앱에 대한 통합 작업)에 액세스하기 위한 44 개의 명령을 지원한다.

Lip-Interact 는 스마트 워치, 헤드웨어 장치 등 다른 플랫폼에 적용될 수 있는 장점을 가지고 있다. 하지만 많은 노이즈가 존재하는 실제 사용 사례에 대한 실험이 진행되지 않았으며, 다양한 환경에서의 알고리즘 개선 및 테스트가 필요하다. 또한 Lip-Interact 는 스마트폰 카메라가 아닌 추가적인 카메라를 장착하여 구현되었다. 스마트폰에 추가로 장착하는 기기가 줄어들어야 기능을 수행하기 위한 응답 시간을 줄이고 사용자 경험을 더욱 향상시킬 수 있다. 언어적 한계 또한 단점으로 들 수 있다. 중국어로만 구현 됐기 때문에 추후 다른 언어에도 적용될 수 있도록 개발이 필요하다.

### 2.3 Hand-to-Face

스마트 글래스 사용 시 사용자의 얼굴은 일반적으로 제어에 방해를 받지 않기 때문에 쉽게 접근할 수 있다는 장점이 있다. 또한 얼굴을 만지는 행위는 인간의 일반적인 행동으로 사회적 수용성을 충분히 만족 한다. Hand-to-Face 는 사용자의 얼굴을 입력 장소로 활용하여 제어의 움직임을 방해하지 않으면서 추가 입력 방식으로 제안되었다.

제안된 두 가지 방식은 그림 3 과 같이 귀 접촉에 대한 카메라 추적과 엄지손톱으로 턱이나 뺨 접촉에 대한 민감도 활용이다. 이를 통해 소형화, 난독화, 가림, 위장, 용도 변경이라는 사회적으로 수용 가능한 5 가지 Hand-to-Face 설계 전략을 수행한다.



(그림 3) 귀 접촉에 대한 카메라 추적(좌), 엄지손톱으로 턱이나 뺨 접촉에 대한 민감도 활용(우)

사용자는 고개를 돌려 자신의 귀를 관찰자로부터 숨길 수 있다. 귀의 입력 공간이 작기 때문에 소형화가 가능하며, 굵기와 같은 일반적인 행동으로 위장할 수 있다. 따라서 난독화 입력을 수행할 수 있다. 사용성 평가 결과로 피실험자는 평소 귀를 만지는 것과 비슷하다는 응답을 나타냈다. 이는 위장과 난독화 설계 전략에 해당된다.

턱이나 뺨에 대한 엄지 손톱의 접촉에 대한 민감도 활용에서는 세 가지 전략을 구체화 했다. 엄지 손톱이 얼굴에 닿는 부분에 가려지기 때문에 가림을, 뺨을 굽거나 턱을 잡는 것은 일반적인 행동이기 때문에 용도 변경을, 그리고 입력 범위가 작기 때문에 소형화가 가능하다. 피실험자는 이러한 제스처가 관찰자

가 알아차리기에 상당히 작다는 응답과 다른 사람들이 알아차리지 못하게 엄지를 숨길 수 있다는 응답을 밝혔다. 이는 소형화와 가림 설계 전략에 해당된다. 오차율은 21.5%로 높은 편이며, 이는 단단한 턱과 부드러운 뺨의 차이 때문에 발생한 오차를 포함한다. 또한 손톱의 끝부분이 접근성이 더 좋았다는 결과를 나타냈다.

### 3. 결론

본 논문에서는 공공장소에서 스마트 글래스 사용 시 주변 사람들에게 관심을 덜 받고 프라이버시를 지킬 수 있는, 사회적 수용성을 고려한 사용자 인터페이스에 대한 기존 연구를 분석하였다.

스마트 글래스의 전방 인식 센서 범위 내에서 손동작을 인식하는 Mid-Air 제스처 방식보다는 평상 시 일반적으로 만지게 되는 턱, 볼, 또는 귀를 통해 인터페이스를 구현하는 방법은 상당히 자연스러운 방법이라고 할 수 있다. 숨을 들이쉬며 팝-노이즈 현상을 줄이면서 아주 작은 소리까지 정확하게 받아쓰기 기능을 구현한 사일런트보이스와 소리를 내지 않고 입 움직임 만으로 원하는 명령을 수행할 수 있는 Lip-Interact 는, 주변사람들에게 거슬릴 만한 소음을 생성하지 않는 점에서 사회적 수용성이 고려된 기술이라고 할 수 있다.

하지만 사일런트보이스는 마이크를 입에 붙이고 사용해야하고, Lip-Interact 는 스마트폰과 같은 추가적인 장치를 요구한다. 또한 이를 모두 언어적 한계와 개인화를 위한 많은 데이터 수집이 요구된다. 사람들은 모두 다른 발음과 입 모양을 가지고 있기 때문이다. Hand-to-Face 의 경우에도 귀 영상을 취득하는 카메라 혹은 엄지 손톱에 부착하는 추가 센서를 요구한다.

시간이 지나면 기존의 연구들이 더 발전될 것이라 기대하고, 향후에는 사용자의 시선과 눈에 띄지 않는 손 동작 기반으로 사회적 수용성을 고려한 사용자 인터페이스에 대해 연구할 계획이다.

### 사사의 글

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 '범부처 Giga KOREA 사업'의 지원을 받아 수행된 연구임 (No.GK19P0200, (4D 실감-총괄/1 세부) 4D 복원 및 동적 변형 거동 모델 기반의 초실감 서비스 기술 개발)

### 참고문헌

- [1] FUKUMOTO, Masaaki. SilentVoice: Unnoticeable Voice Input by Ingressive Speech. In: The 31st Annual ACM Symposium on User Interface Software and Technology. ACM, 2018. p. 237-246.
- [2] SUN, Ke, et al. Lip-Interact: Improving Mobile Device Interaction with Silent Speech Commands. In: The 31st Annual ACM Symposium on User Interface Software and Technology. ACM, 2018. p. 581-593.
- [3] LEE, DoYoung, et al. Designing Socially Acceptable Hand-to-Face Input. In: The 31st Annual ACM Symposium on User Interface Software and Technology. ACM, 2018. p. 711-723.