

음성 및 손동작 결합 인터페이스를 통한 가상객체의 생성

김창섭 남현길 박종일[†]

한양대학교

{cskim, hgnam}@mr.hanyang.ac.kr, jipark@hanyang.ac.kr

Virtual Object Generation Technique Using Multimodal Interface With Speech and Hand Gesture

Kim, Changseob Nam, Hyeongil Park, Jong-Il

Hanyang University

요약

가상현실 기술의 발전으로 보다 많은 사람이 가상현실 콘텐츠를 즐길 수 있게 되었다. PC나 스마트폰과 같은 이전의 콘텐츠 플랫폼과 달리 가상현실에서는 3차원 정보를 전달할 수 있는 인터페이스가 요구된다. 2차원에서 3차원에서의 변화는 보다 높은 자유도를 가지는 반면, 사용자는 새로운 인터페이스에 적응해야 하는 불편함 또한 존재한다. 이러한 불편함을 해소하기 위하여 본 논문에서는 가상현실상에서 음성과 손동작을 결합한 인터페이스를 제안한다. 제안하는 인터페이스는 음성과 손동작은 현실 세계에서의 의사소통을 모방하여 구현하였다. 현실 세계의 의사소통을 모방하였기 때문에 사용자는 추가적인 학습 없이 가상현실 플랫폼에 보다 쉽게 적용할 수 있다. 또한, 본 논문에서는 가상객체를 생성하는 예제를 통하여 기존의 3차원 입력장치를 대신할 수 있음을 보인다.

1. 서론

이전에 차세대 콘텐츠 플랫폼으로 여겨지던 가상현실(Virtual Reality, VR)이 많은 연구 및 콘텐츠 다양화를 통하여 현재는 일반 사용자들도 쉽게 접할 수 있게 되었다. 현재 대부분의 가상현실 기기는 사용자의 시야를 가상의 현실의 시야로 변경해주는 디스플레이 장치를 기본으로, 사용자의 입력을 받기 위한 별도의 입력 장치로 이루어져 있다. 이러한 입력장치는 사용자의 의도를 분석하기 위하여 3차원 정보를 해석할 수 있어야 한다. 가상현실 이전의 플랫폼의 경우 일반적인 컴퓨터의 마우스나, 스마트폰의 터치 인터페이스를 통하여 사용자의 의도를 해석하였다. 새로운 입력 장치는 사용자로 하여금 입력 장치에 대한 학습을 필요로 한다.

이러한 추가적인 학습을 줄이기 위하여 NUI (Natural User Interface)에 관하여 많은 연구가 진행되었다[1, 2]. NUI의 일환으로 본 논문에서는 인간의 의사소통을 모방한 인터페이스를 제안한다. 제안하는 인터페이스는 음성과 손동작이 결합된 형태이며, 가상객체를 생성하는 예제를 통하여 인터페이스의 효율성에 관한 연구를 진행하였다. 가상객체 생성에 있어 사용자의 손동작을 분석할 필요가 있다. 또한 사용자가 의도한 시점에서만의 분석이 필요하고, 음성을 이용하여 이러한 분석 시점을 결정하였다.

2장에서는 음성과 손동작을 결합한 인터페이스에 대하여 설명하고, 3장에서는 음성과 손동작의 합성에 대하여 설명한다. 4장에서는 손동작 가상객체 생성에 대하여 설명하고, 마지막으로 5장에서는 본 논문의 결론과 추후의 연구 방향에 대하여 기술한다.



그림 1. 음성과 손동작을 결합한 인터페이스의 사용 예시

2. 음성과 손동작을 결합한 인터페이스

그림 1은 음성과 손동작을 결합한 인터페이스를 실제 사용하는 장면이다. 그림 1에서 사용자는 생성하고자 하는 가상객체의 형태를 손으로 정해주고 있다. 이러한 예제뿐만 아니라 다양한 예제에서도, 연속적인 손동작 중 사용자가 의도하는 동작만을 인식할 필요가 있다. 이전의 손동작 연구에서는 손의 움직임 중 유의미한 부분만을 추출하기 위하여 다양한 방법들을 적용했다. 반복되는 손동작의 패턴을 분석하거나, 수화와 같은 손의 모양으로 해석 시점을 결정하는 방법들을 사용하였다[3, 4]. 본 논문에서는 손의 움직임 중 유의미한 움직임을 결정하기 위하여 음성을 사용하였다. 또한 앞서 언급한 바와 같이, 인간의 의사소통 방식을 모방하였기 때문에 음성 명령은 일상적인 대화에서 사용하는 단어들이 *like*와 *this*로 구성하였다.

[†]교신저자

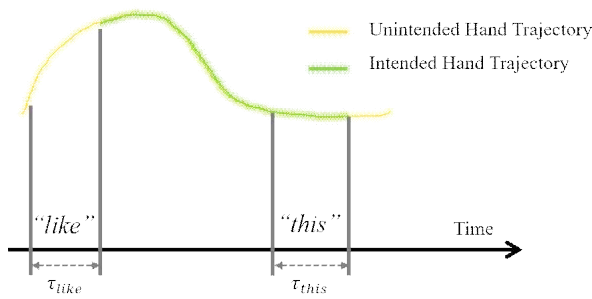


그림 2. 음성과 손동작의 합성을 통한 손동작 해석 시점 결정 방법

3. 음성과 손동작의 합성

그림 2에서 노란색 및 초록색 곡선은 손의 3차원 움직임을 표시하고 있다. 초록색 곡선은 음성에 의해 유의미한 움직임으로 결정된 움직임이며, 이는 음성에 대한 시간 축에서 'like'와 'this'가 각각 발음이 끝나는 사이를 의미한다. 음성의 시간 축에 있어서 τ_{like} 와 τ_{this} 는 각각 'like'를 'this'를 발음하는 데 걸리는 시간이다.

4. 손동작으로부터 가상객체 생성 방법

본 논문에서는 손동작 인식을 위하여 LeapMotion을 사용하였으며 음성 인식의 경우 단어를 등록하여 사용하였다. 다시 말해, 사용자의 음성 중에 목표 단어(keyword)가 있을 경우 명령이 수행된다. 본 논문에서 사용하는 예제는 회전체(body of revolution)의 생성이나 제안하는 인터페이스는 회전체뿐만 아니라 박스나 원통형 등 다양한 기초적인 3차원 도형 생성에 적용 될 수 있다. 그림 2는 이러한 가상객체 생성의 과정을 나타낸다.

4.1 손동작의 2차원 투영

센서로부터 출력되는 손동작 인식 결과는 손의 모양 손의 3차원 위치로 이루어져 있다. 본 논문의 방법에서는 손의 3차원 위치를 사용하여 회전체의 형태를 정한다. 회전체 계산의 간편함 때문에 3차원 경로를 2차원 평면에 투영시켜 평면을 회전하여 회전체를 생성한다. 그림 3의 (b)가 이를 나타내고 있다. 사용자의 시야에서는 변하지 않으나 다른 시야에서는 특정 평면에 투영된 것을 확인할 수 있다.

4.2 손동작 최적화

본 논문에서는 센서로부터 출력되는 손의 3차원 위치를 단위시간(약 0.5초)마다 샘플링하여 사용한다. 이는 출력값의 떨림(jittering)에 보다 강인해지는 반면 손의 경로가 부자연스러워지는 단점이 있다. 이를 해결하기 위하여 Bézier 커브를 적용하였다. 그림 3의 (c)는 (b)에 Bézier 커브를 적용한 결과를 나타낸다.

4.3 양손동작의 합성

회전체는 한 곡선을 회전축을 중심으로 회전하여 생성되는 도형을 뜻하고, 이를 위하여 양손의 경로를 하나로 합칠 필요가 있다. 또한, 앞서 구한 Bézier 커브는 실제로 점들의 집합이며 이 점들을 이용하여 합쳐진 경로를 생성한다. 그림 3의 (d)는 합쳐진 경로를 나타낸다. 이는 한 손의 경로와 대칭 시킨 반대 손의 경로의 점들을 번갈아 연결한 결과이다. 그 외, 양손 경로의 비대칭성에 대응하기 위하여 회전축은 직선으로 이루어져 있지 않고 같은 시점의 샘플들의 중심들로 이루어져 있다.

5. 결론

본 논문에서는 가상객체의 생성을 통하여 음성과 손동작이 합쳐진 인터페이스를 제안하고 효용성에 대하여 논의하였다. 이는 인간의 의사소통방식을 모방하였으며, 이를 통하여 사용자가 새로운 플랫폼에 보다 용이하게 적응할 수 있을 것으로 기대된다. 또한 본 논문에서는 가상객체의 생성에 그쳤으나 보다 자유로운 상호작용이 가능할 것으로 예상된다. 추후에 음성인식에 NLP(Natural Language Processing)를 합친다면 의사소통 외에도 보다 많은 분야에 적용될 수 있을 것이다.

감사의 글

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No.2017-0-01849, 실내외 임의공간 실시간 영상 합성을 위한 핵심 원천기술 및 개발툴킷 개발)

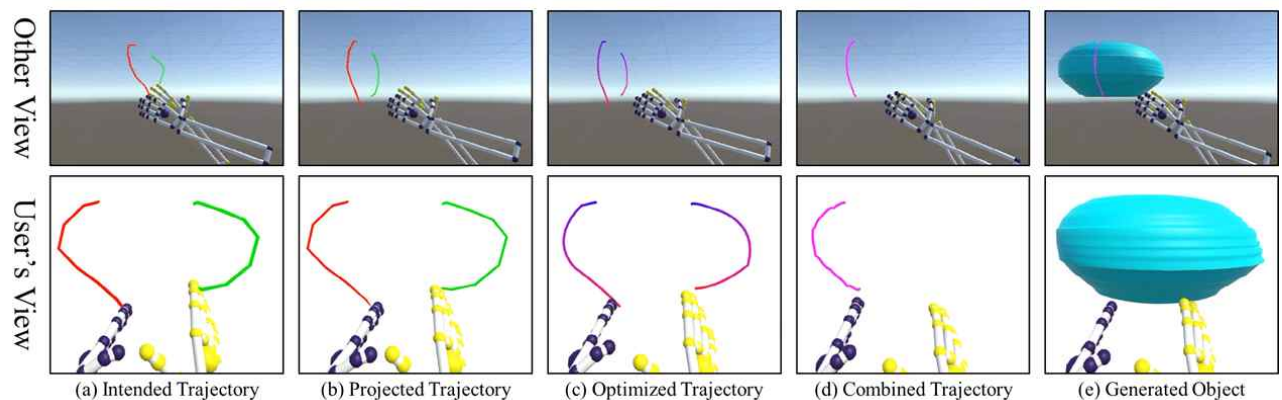


그림 3. 가상객체의 생성 과정

참고문헌

- [1] Jain, Jhilmil, Arnold Lund, and Dennis Wixon. "The future of natural user interfaces." CHI'11 Extended Abstracts on Human Factors in Computing Systems. ACM, 2011.
- [2] Cheng, Hong, Lu Yang, and Zicheng Liu. "Survey on 3D hand gesture recognition." IEEE transactions on circuits and systems for video technology 26.9 (2015): 1659-1673.
- [3] 김지예, 박종일. "다문화 손동작 인식을 위한 HOG-HOD 알고리즘." 멀티미디어학회논문지 20.8 (2017): 1187-1199.
- [4] 김지예, 박종일. "주피수 분석을 이용한 반복적인 손동작 인식." 한국방송미디어공학회 학술발표대회 논문집 (2017): 166-167.