

# 표면 진동 기반의 허공 제스처 인식

김민혁, 이동훈, 김효수  
중앙대학교

{alsgur7859, ldh019, hskimhello}@cau.ac.kr

요약—허공 제스처 인식은 온라인 미팅 상에서 더욱 몰입감 있는 상호작용을 지원하기 위해 효과적으로 활용될 수 있다. 본 연구는 제스처로부터 발생한 진동을 활용하여 허공 제스처 인식을 지원하는 새로운 시스템을 제안한다. 사용자가 책상위에 두 팔꿈치를 놓고 손 제스처를 취하면 그 제스처로 인하여 진동이 발생하고, 그 진동은 책상을 타고간다. 이러한 관점에서 우리는 제스처로 유발된 진동들을 수집하고 그 진동들을 이용해 제스처를 인식한다. 실제 사용자들에 대한 우리의 평가는 우리가, 스마트폰과 스마트워치, 이 두 가지 사용기기만을 사용해서, 정밀하게 허공 제스처들을 인식할 수 있음을 보여준다. (예를 들어, 모든 유저들을 대상으로 96.02%의 평균 정확도가 나왔다.)

## I. 소개

코로나19가 발생하면서, 수많은 수업과 미팅이 비대면으로 전환되고 비대면 상황에서 발표, 자료 공유 등 다양한 일들이 이루어지고 있다. 상황이 지속되면서 사람들이 비대면 미팅에 자주 참여하게 됨에 따라, 비대면 미팅에서의 불편함 해소에 대한 필요성이 증가했다. 그 필요에 반응하여, 많은 연구들은 온라인 환경에서 자연스럽고 효과적인 상호작용을 지원하기 위해 손 제스처를 활용하는 시도를 했다 [5], [6]. 비대면 미팅에서 발표를 진행할 때, 사용자가 키보드나 마우스를 통해 조작하지 않고 손 제스처만을 사용하여 자연스럽게 프레젠테이션의 페이지를 넘나드는 것을 상상해보라. 그러나 기존의 연구들은 여러가지 문제가 있고, 특히 배포 가능성에서 문제가 있다. 이것은 기존의 연구들이 주로 적외선 카메라나 [7] 특수한 웨어러블 기기 [8] 같은 다른 하드웨어에 의존하고 있기 때문이다.

본 논문에서는 제스처로 유발된 진동을 활용해 제스처를 인식할 수 있는 잠재성을 탐구한다. Figure 1은 우리가 제안한 시스템의 사용 시나리오를 보여준다. 처음으로, 우리는 두 팔이 책상 위에 올려진 상황을 가정한다. 사용자가 박수와 같은 허공 제스처를 취하면 진동이 발생하고, 그 진동은 책상의 표면을 타고 이동한다. 우리는 그렇게 발생한 진동을 책상 위에 올려진 스마트폰과 왼손에 차고 있는 스마트워치를 이용해서 감지하고, 이를 이용해 제스처를 인식한다. 그 말은 즉, 오직 상용기기인 스마트폰과 스마트워치만을 이용해서 손 제스처를 인식한다는 것이다.

## II. 관련 연구

이전에도 탁자 같은 평평한 면에서의 진동을 이용하여 기기와의 상호작용을 시도해왔다. SurfaceSight는 [1] IoT장치에 LIDAR를 더하여 표면상에서 일어나는 동작, 표면 위의 객체, 그리고 표면에 사람이 존재하는 지를 인지한다. Vibwrite는 [2] 다양한 표면 상에서 일어나는 핀 번호 입력, 패턴 입력, 제스처 입력 시에 사용자마다 각각 고유한 진동패턴이 있음을 이용해 authentication 한다. SurfaceLink는 [3] multi-device가 놓인 표면상에서 일어나는 swipe 동작들을 acoustic sensor로 인식해 이를 multi device의 상호작용에 이용한다.



Fig. 1. 우리가 제안하는 시스템의 사용 시나리오. 사용자는 두 팔꿈치를 책상 위에 위치하고, 왼손에 스마트워치를 차고 오른쪽에 스마트폰을 비치한다.

Chen, M. et al.은 [4] 표면상에서 글자들을 쓰면서 생긴 소리를 acoustic sensor로 인식해 적은 글자를 복원한다. 요약하자면, 기존의 연구들은 표면 부근의 공간을 인풋 공간으로 활용해 표면상의 제스처들을 인식한다. 기존의 연구들과 다르게, 우리는 온라인 소통 환경에서 더욱 자연스러운 상호작용을 지원해줄 수 있는 허공 제스처 인식 시스템을 제안한다.

## III. 시스템 설계

이 부분에서는 허공 손 제스처를 인식하기 위한 새로운 방법을 소개한다 (Figure 2를 보라). 그것은 두 가지 단계, 트레이닝 단계와 인터랙션 단계를 담고 있다. 트레이닝 단계에서는 먼저 사용자에게 특정한 손 제스처를 여러 차례 (예를 들면 10번 혹은 20번) 수행하라고 요청한다. 이때 두 팔꿈치는 책상 위에 올려져 있어야 한다. 우리는 그 후에 책상 위에 놓인 스마트폰과 왼손에 찬 스마트워치로 진동 데이터를 수집하고, 모인 트레이닝 데이터를 가지고 분류 모델을 생성한다. 인터랙션 단계에서는 사용자는 트레이닝 제스처들을 책상 표면에 팔꿈치를 댄 채로 수행함으로써 온라인 상에 다른 이들과 상호작용한다. 우리는 그리고 인터랙션 중에 얻은 진동 데이터와 분류 모델을 가지고 제스처를 인식한다.

### A. 제스처로 유발된 진동 데이터 수집

실제로, 우리는 사용자가 언제 동작을 수행한다는 사전 정보가 없기 때문에, 제스처로 유발된 진동 데이터를 수집하는 데에 어려움이 있다. 우리는 이 챌린지를 CFAR(Constant False Alarm Rate) 알고리즘을 사용함으로 접근했다. CFAR

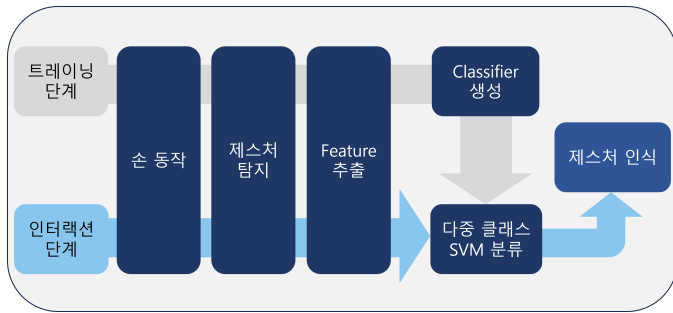


Fig. 2. 전체 흐름도

알고리즘은 인풋 시퀀스에서의 갑작스러운 변화를 탐지한다. 스마트폰과 스마트워치 모두에서 제스처 전후에 큰 변화를 보이므로, 우리는 이 알고리즘을 이용해 제스처를 정밀하게 탐지하고 모델 생성과 제스처 인식에 쓰일 진동 데이터를 수집할 수 있다.

### B. Feature 추출

제스처로 유발된 진동 데이터가 탐지되면, 그 데이터에서 우리가 필요한 feature를 뽑는다. 우리는 먼저 300Hz 이상의 고주파 요소를 필터링한다. 왜냐하면 손 제스처는 저주파 대역의 진동을 유발하기 때문이다. 그 후에, 우리는 필터링 된 데이터로부터 몇 가지 클래스의 feature를 얻는다. 예를 들어 우리는 integral absolute value, variance, root mean square, standard deviation, 그리고 mean absolute deviation 등의 통계적 feature들을 사용한다. 우리는 또한 Fast Fourier Transform과 Power Spectral Density를 사용한다. 선정된 feature들은 classifier에 사용되기 위해 하나의 벡터로 합쳐진다.

### C. 분류

우리는 분류에 multi-class Support Vector Machine (SVM)을 사용하는데, 그 이유는 SVM이 적은 양에 데이터에 대해서도 잘 동작하기 때문이다. 트레이닝 단계에서, 우리는 트레이닝 진동 데이터에서 추출한 feature를 이용해 SVM classifier를 build 한다. 그 후에 생성된 모델을 활용하여 제스처를 인식한다.

## IV. 실험

본 연구에서는 6명의 사용자에게서(4 males, 2 females) 손 동작 데이터를 수집했다. 사용자마다 10가지 손동작에 대해 각각 30번씩, 총 300개의 손동작 데이터를 모을 것을 요구했다. Google Pixel4 스마트폰을 탁자 위에 올려두고 양 팔꿈치는 일정한 간격을 유지한 채로 탁자의 고정된 위치에 두는 것을 default로 설정했다. 또한 Galaxy Watch 4 스마트워치를 왼손에 차도록 했다. 그리고 우리는 우리가 제안한 시스템의 각 사용자별 제스처 인식 정확도를 10-fold cross validation을 이용해 측정했다. 6명의 유저의 정확도는 table 1에서 볼 수 있는 것과 같다. 이들의 평균 정확도는 96.02% 한 동작을 30번 했을 때 1 2번 분류가 잘못되는 정도의 정확도이다. 이 정도의 정확도는 default 세팅 안에서 10가지 다른 손동작들이 각각 고유하고 서로 간에도 잘 구별됨을 보여준다. 그 중 정확도가 가장 높은 유저의 정확도는 99.33%로 (사용자1, 사용자2) 테스트한 300번의 동작 중 단 2번의 동작만 잘못 분류했다. 정확도가 가장 낮은 유저의 정확도는

TABLE I

6명의 사용자의 10-FOLD CROSS VALIDATION 손 제스처 분류 정확도

사용자	사용자1	사용자2	사용자3	사용자4	사용자5	사용자6	평균
정확도	99.33%	99.33%	95.67%	95.38%	95.05%	91.36%	96.02%

(사용자6) 91.36%로 300번의 동작 중 25개 정도가 잘못 분류되었는데, 이는 비교적 유사한 동작들을 아직 모든 유저에 대해서는 완전히 고유하게 구별해내지 못한 것으로 보인다. 종합적으로 보면, 우리의 연구 결과를 봤을 때 손 동작 별로 default 세팅에 대해서 고유하게 구별해낼 수 있다는 것을 확인했다. 이를 통해 보다 다양한 환경에서 유저의 손동작을 고유하게 구별해낼 수 있는 가능성을 보았다.

## V. 결론

이 연구에서 우리는 스마트폰과 스마트워치 같은 스마트 기기와 물체의 표면을 사용한 새로운 허공 손 제스처 인식 시스템을 소개한다. 특히, 우리는 제스처 인식을 위해 사용자의 제스처로부터 유발된 진동을 사용한다. 우리의 실험 결과는 제스처 인식에서 높은 정확도를 보이고, 그것은 우리가 자연스러운 인터랙션을 제공할 수 있도록 한다.

## 논문 사사

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2022R1C1C1012664).

## 참고 문헌

- [1] Laput, G. and Harrison, C. (2019) 'SurfaceSight', Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3290605.3300559>.
- [2] Liu, J. et al. (2017) 'VibWrite', Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. <https://doi.org/10.1145/3133956.3133964>.
- [3] Goel, M. et al. (2014) 'SurfaceLink', Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/2556288.2557120>.
- [4] Chen, M. et al. (2019) 'Your table can be an input panel', Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 3(1), pp. 1-21. <https://doi.org/10.1145/3314390>.
- [5] J. Wang, J. Chen, Y. Qiao, J. Zhou and Y. Wang, "Online Gesture Recognition Algorithm Applied to HUD Based Smart Driving System," 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Beijing, China, 2019, pp. 289-294, doi: 10.1109/ISMAR-Adjunct.2019.00-26.
- [6] Y. Fang, K. Wang, J. Cheng and H. Lu, "A Real-Time Hand Gesture Recognition Method," 2007 IEEE International Conference on Multimedia and Expo, Beijing, China, 2007, pp. 995-998, doi: 10.1109/ICME.2007.4284820.
- [7] Yi Li, "Hand gesture recognition using Kinect," 2012 IEEE International Conference on Computer Science and Automation Engineering, Beijing, China, 2012, pp. 196-199, doi: 10.1109/ICSESS.2012.6269439.
- [8] S. Kallio, J. Kela and J. Mantyjarvi, "Online gesture recognition system for mobile interaction," SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483), Washington, DC, USA, 2003, pp. 2070-2076 vol.3, doi: 10.1109/IC-SMC.2003.1244189.