# A 3D-Touch Interface by Using EMG

JinHyuck Park
Sungkyunkwan University
2066, Seobu-ro, Jangan-gu Suwon-si, South Korea
vkqxkr@gmail.com

YoungHoon Seo
Sungkyunkwan University
2066, Seobu-ro, Jangan-gu Suwon-si, South Korea
y8seo@skku.edu

DongRyeol Shin
Sungkyunkwan University
2066, Seobu-ro, Jangan-gu Suwon-si, South Korea
drshin@skku.edu

ChoonSung Nam
Sungkyunkwan University
2066, Seobu-ro, Jangan-gu Suwon-si, South Korea
namgun99@gmail.com

## ABSTRACT

The 2D-Touch interaction for users to provide on-screen position values as user input provides various input methods such as Touch, Long-Touch, Drag and so on. However, it does not provide a way for the user to provide input through control of the force. To achieve this, 3D-Touch interface came out. This is a new way of interacting that can be used by adding force values as input in 2D-Touch. 3D-Touch adds depth of force to add the strength z of force to the coordinates of the existing position input x, y. Therefore, it is possible to diversify the input by force input in an environment where 2D-Touch is not possible, such as constrained space. However, 3D-Touch has the disadvantage that it can be measured only by a screen device capable of measuring force. For this reason, the 3D-Touch interface is not popularly used and is used only in limited products. Another way to measure force is through electromyogram, EMG. EMG (surface electromyography signal) is a biological signal used to sense the degree of activation and mobilization patterns of the nerve roots that are regulated by the nervous system during muscle. The EMG signal can cause a change in the signal to distinguish the 3D-Touch from the 2D-Touch by the applied force. Therefore, if users use a device that can measure EMG signals, they can use 3D-Touch interaction on devices that do not provide a 3D-Touch screen. In this paper, we test whether the three input methods of Touch, Peek, and Pop can be classified into EMG signals. CNN (Convolutional Neural Networks) is used to distinguish the EMG signals of each input.

## CCS Concepts

**• Human-centered computing → Pointing; Ubiquitous and mobile computing; Mobile devices; Haptic devices**

## Keywords

EMG; 3D-Touch; UI; Deep learning; CNN


*IECC ' 19*, July 7–9, 2019, Okinawa, Japan

ACM ISBN 978-1-4503-7177-3/19/07…$15.00
DOI: https://doi.org/10.1145/3343147.3343155

## 1. INTRODUCTION

The appearance of the 3D-Touch interface enables various inputs by adding the value of the user's force to the way that the previous 2D-Touch interface uses only position information on the screen[1]. In other words, the input was varied by dividing the force by Touch, Peek, Pop according to the value of force[2]. This type of interface has been proposed as a method that can be widely used for devices requiring various inputs in a limited space like mobile screen. However, 3D-Touch requires a device equipped with a screen that can measure force. If there is no such device, 3D-Touch interaction is not available.

The equipment that can measure force for 3D-Touch is surface electromyography signal, EMG[3]. EMG is a biological signal used to sense the degree of activation and mobilization patterns of the nerve roots that are regulated by the nervous system during muscle contraction. Myo[4], a commercially available device capable of measuring EMG, allows users to easily measure EMG signals.

The relationship between EMG signal and force may be measured by some muscles exhibiting a nonlinear relationship or by overestimating EMG signals over actual forces[5]. However, it is evident that the user's EMG signal changes due to the strength of the force. For the 3D-Touch, the force-generating muscles can be measured in the forearm through the pressure of the finger. Therefore, using Myo, it is possible to measure the force of the user for the 3D-Touch from the forearm. Deep learning methods are emerging to measure the input of 3D-Touch with EMG signal and

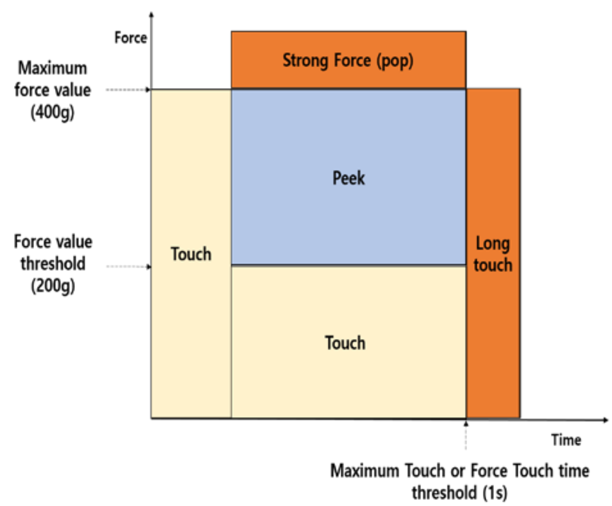


**Figure 1. 3D-Touch Area.**

classify the EMG signal according to 3D-Touch input[6]. Especially, CNN (Convolution Neural Network) method is used. EMG signal is converted into Image and its characteristics are extracted. For the CNN model learning, the extraction time of the EMG data is determined based on the reaction time in 3D-Touch [7]. We also set window size and overlap size to classify data and image each data. Therefore, in this paper, we propose a method to provide a 3D-Touch interface by using EMG signal which can be used universally.

# 2. RELATED WORKS

## 2.1 3D-Touch Interface

Unlike the usual 2D-Touch interface, the 3D-Touch interface divides the input step by dividing the threshold according to the force value. In the case of the iPhone, there are two kinds of touch according to the amount of force that the user presses the screen. It is a 2D-Touch that uses only position values and a 3D-Touch that uses position and force values. 'Peak' is a weak force in the strength of power, and 'Pop' is a strong force[8]. For example, a user can select 'Peek' to preview a specific photo in the photo search, and 'Pop' to view the photo as a whole. This 3D-Touch interface is also used to reduce the typing error and the convenience of reaching the user's page of a specific app. As shown in Fig. 1, there are three areas of the force, that is, the Touch area between 0g and 200g, the Peek area between 200g and 400g, and the Pop area over 400g. This area is set as the Long Touch area if it cannot reach the Peek area after 1s according to input time. And since the time domain of the touch takes time to reach the force, it excludes the region generated within 200ms[9].

## 2.2 EMG Classification Using CNN

The reason for using the CNN (Convolutional Neural Network) when using the EMG signal is that unlabeled data is available. Thus, it can learn potentially better features than hand-crafted features. Also, this method is suitable for users to know different EMG patterns[10]. The architecture of CNN is different from the existing ANN (artificial neural network). CNN includes three main types of layers. The three layers are a convolutional layer, pooling layer, and fully connected layer. Each layer represents a block that contains the number of layers. Convolution and pooling operations are applied to input data using filters to provide an optimized feature map. This feature map is used as the final output of the convolution layer. The classification works on the last fully connected layer[11]. A flattened layer is placed between the part that extracts the features of the image and the part that classifies the image, which makes the data of the image type array. This feature has the following differences compared to a fully connected layer: First, the shape of the input/output data of each layer is maintained. Second, it effectively recognizes the characteristics of the adjacent image while maintaining the image space information. Third, image feature extraction and learning are performed with a plurality of filters. Fourth, it has a pooling layer that strengthens the features of the extracted image. Fifth, since the filter is used as a shared parameter, the learning parameter is very small compared to a general ANN.

**Figure 2. Experiment for collecting EMG data for 3D-Touch**

# 3. CNN experiment for using EMG-3D-Touch

## 3.1 3D-Touch EMG Data Collection

In order to measure the EMG signal when using 3D-Touch, three inputs of 'Touch', 'Peek', and 'Pop' were collected from EMG signal for 5 seconds from 4 subjects. Two of them (sub1, sub2) were subjects who had used 3D-touch and the other two (sub3, sub4) were not. In figure 2, each subject was equipped with Myo on his right arm and collected data five times in total. We used iPhone 7 Plus as a device to collect EMG signals in 3D-Touch, and the app for the collection was developed as shown in Fig 3. The collected data generated about 1,000 EMG signals for 5 seconds. 40 EMG signals corresponding to 200ms were set as the window size, and the overlap was set to 10, which was used as data for learning the CNN model. In the previous study, the maximum time latency in real-time control was set to 300ms[12]. Recent studies suggest that the time latency should be maintained at 100-125ms[13,14]. In the case of Myo, however, realtime is possible only within 300ms because data is transmitted through the low frequency. The collected EMG signals were converted into images and 97 images were generated for 5 seconds to generate 485 images during a total of 5 acquisitions. Of these, 291 were used for validation, 97 for a test, and 97 were used for a train. We use the VGGNet model as the CNN learning model[15].

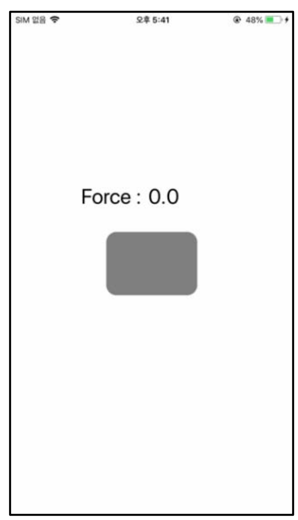


**Figure 3. App for Force-EMG signal collection.**

Myo extracts 8 EMG signal data from EMG sensor 1 to 8 at the same time. Since 8 sensor data are represented by 8 bits, EMG signal data is expressed as -128 ~ 127. In other words, in order to find one operation, it is necessary to recognize the operation by extracting and associating the characteristics of the eight EMG signal data generated in the operation. Eight sensor data must be imaged for CNN analysis. Figure 5 shows eight EMG sensor signals generated by the user in the 'Touch' gesture. This signal is expressed as a change over time. From the blue EMG sensor 1 to the neon EMG sensor 8 In Figure 5, we can see the characteristic

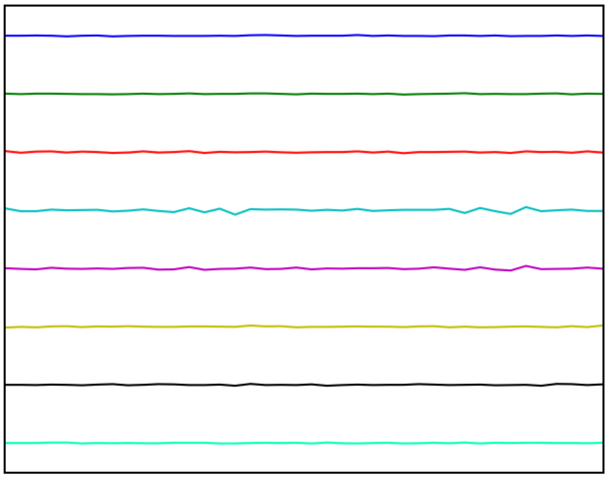

**Figure 5. Images for data representation of each sensor**

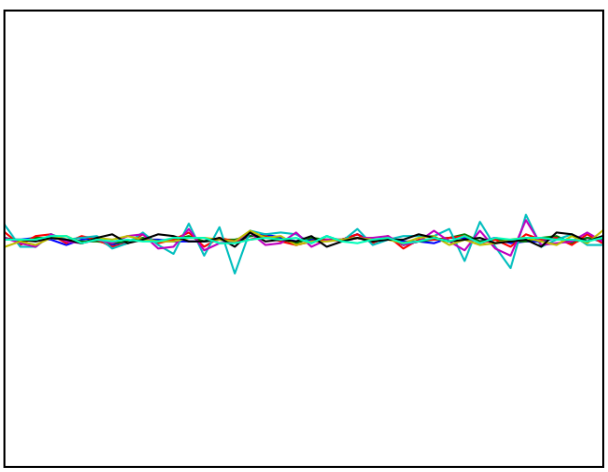

**Figure 6. Images for data representation of each sensor**

by changing the EMG value at 'Touch'. However, since the change of the EMG signal in 'Touch' is very small, using this image has less than 50% accuracy to distinguish Touch, Peek, and Pop. Therefore, if each sensor is represented on the same graph as shown in Fig. 6, it is possible to more easily grasp the change of the EMG sensor used in 'Touch'. In other words, EMG sensors 4, 5 and 7 show larger values than the other EMG sensors. In this paper, we use this image to distinguish 'Touch', 'Peek' and 'Pop' through CNN model.

## 3.2 Classification of Touch, Peek and Pop

The results of classifying each subject's 3D-Touch input as CNN is shown in Fig 7. In Fig. 7, the accuracy of classification of Touch, Peek, and Pop in 3D-touch through EMG is less than 90%. In particular, the 3D-Touch accuracy of subject 3 is very low, about

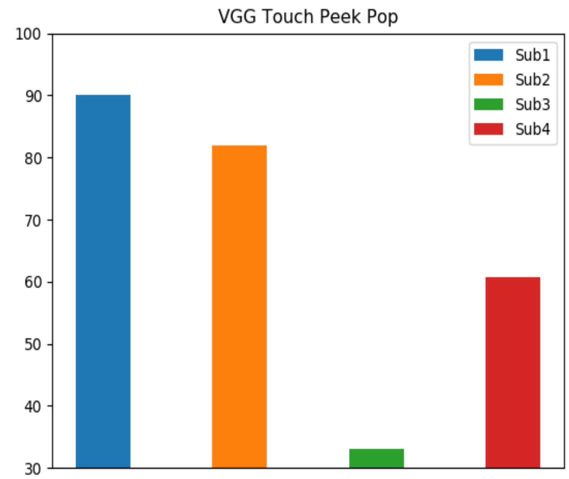


**Figure 7. Results through VGGNet model in Touch, Peek and Pop**

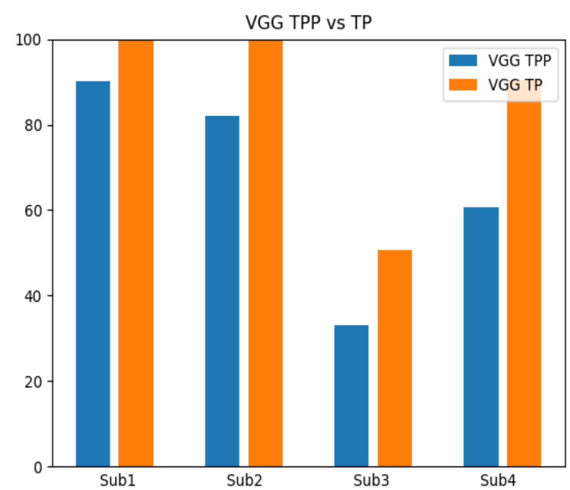


**Figure 8. Comparison of TPP and TP by each subject**

33%. This is because the magnitude of the EMG value for the force generated by each step of the user does not change much. For example, if the subject enjoys a Touch range with a force of 180g and presses Peek's range with a force of 210g, this difference may not be noticeable in the EMG data at only 30g. Likewise, in the areas of touch and pop, the subject can make less of a difference in force. Therefore, in order to divide a small amount of force into EMG data, the difference between the two forces must be clearly set. Based on these results, we try again to classify the difference between Touch and Pop with CNN at least 200g difference.

## 3.3 Classification of Touch and Pop

The touch area is less than 200g and Pop area is more than 400g so the difference between them can show a difference of force of at least 200g. Therefore, Touch and Pop can have a difference in magnitude above a certain force than the difference in force between the previous Touch and Peek, Peek and Pop. These two are classified through VGGNet as shown in Fig 8. As shown in Figure 8, it is more accurate to classify only Touch and Pop than the previous Touch, Peek, and Pop. In other words, by using EMG, the user can use the 3D-Touch interface partly by using Pop, which

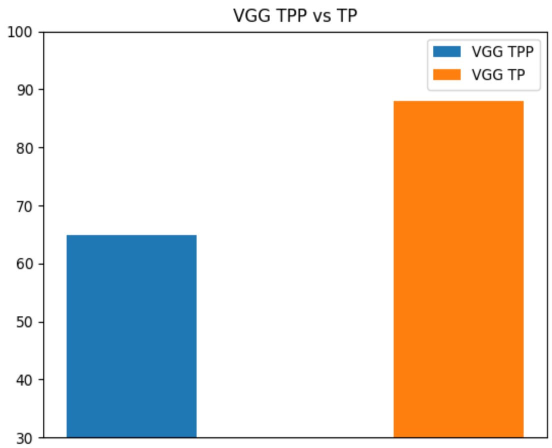


**Figure 9. Comparison of TPP and TP by all subjects**

is a force input that shows a difference of more than a predetermined force from the touch. In addition, this method can measure more than 400g of force, which is limited by the existing force-touch screen, so that more input is possible. As shown in Figure 4, in the case of subjects 3 and 4, clearly distinguishing between Touch and Pop is much better. It can be seen that the subject 3 has about 50% accuracy and subject 4 has an accuracy of 90%. However, due to low accuracy the 3D-Touch interface is still difficult to use.

Figure 9 shows the accuracy of learning TPP(Touch, Peek, Pop) and TP(Touch and Pop) based on collected EMG data of all subjects. It can be seen that the case of distinguishing only TP has higher accuracy than TPP. Accuracy is about 88% for TP and about 66% for TPP. However, since both of these results are similar to the average values of TPP and TP for each individual, it is difficult to obtain good accuracy even if the data to be learned increases. In other words, it means that the characteristics of TPP and TP cannot be properly extracted by the VGGNet model.

## 4. DISCUSS

In order to distinguish each input through EMG generated by a small force, VGGNet Models should be modified to change the model to match the EMG characteristics so that fewer force differences can be identified. VGGNet model was not designed to extract features of EMG data. Therefore, a model that can reflect more EMG data characteristics is needed. Second, image conversion is required through preprocessing rather than raw data. Previous studies have shown good results in capturing Raw Data, but pre-processing of Raw Data is required to extract more accurate features. Third, the accuracy of the characteristics of each subject is also important, but an experiment that satisfies all the characteristics of all the subjects is necessary. In other words, if the subjects 1 to 4 are combined, the model should be improved so that the accuracy can be increased.

## 5. CONCLUSION

We proposed a method to provide a 3D-Touch interface according to the magnitude of the user's pressing force without a screen providing a limited 3D-Touch interface. Because the proposed method changes the EMG immediately after the change of force, 3D-Touch can be realized by EMG measurement alone. However, we have confirmed through experiments that the existing models such as VGGNet are not accurate enough to distinguish EMG data. Another feature of this experiment is that 3D-Touch users using iPhone are able to distinguish well because their power control is higher than other device users. However, it still does not understand the characteristics of all users. Therefore, we will improve accuracy by developing a new model for EMG and preprocessing process that reflects characteristics of data.

## 6. ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2017R1D1A1B03032855).

## 7. REFERENCES


[1] L. Du, C. C. Liu, A. Tang, Y. Zhang, Y. Li, K. Cheung and M. C F. Chang. 2016. A novel single layer 3D touch sensing system for human/mobile devices interactions. In *53nd ACM/EDAC/IEEE Design Automation Conference* (Austin, USA, June 05-09, 2016). 14, 4 (April. 2018). DOI= https://doi.org/10.1145/2897937.2901902 .

[2] *Human Interface Guidelines*, https://developer.apple.com/design/human-interface-guidelines/ios/user-interaction/3d-touch/.

[3] M. Pang, S. Guo and S. Zhang. 2015. Prediction of Interaction Force suing EMG of Characteristic Evaluation of Touch and Push Motions. In *International Conference on Intelligent Robots and Systems* (Hamburg, Germany, September 28 - 02, 2015). IEEE, 2099-2104. DOI= https://doi.org/10.1109/IRO S.2015.7353656 .

[4] Seema Rawat, Somya Vats and Praveen Kumar. 2016. Evaluating and exploring the MTO ARMBAND. In *Proceedings of the 5th International Conference on System Modeling & Advancement in Research Trends* (November 25 – 27, 2016). IEEE, DOI= https://doi.org/10.1109/SYSMART. 2016.7894501 .

[5] A. F. T. Ibrahim, V. R. Gannapathy, L. W. Chong and I. S. M. Isa. 2015. Analysis of Electromyography (EMG) signal for Arm Muscle: A Review. In *Advanced Computer and Communication Engineering* Technology. 362, (December. 2015), 567-575. DOI= https://doi.org/10.1007/978-3-319-245 84-3_49 .

[6] N. Duan, L. Z. Liu, X. J. Yu and Q. Li. 2018. Classification of multichannel surface-electromyography signals based on convolutional neural networks. In *Journal of Industrial Information Integration*. (September. 2018). DOI= https://doi.org/10.1016/j.jii.2018.09.001.

[7] Z. Z. Ding, C. Yang and F. Jiang. 2018. sEMG-Based Gesture Recognition with Convolution Neural Networks. In *journal of Sustainability* (June, 2018). DOI= https://doi.org/10.3390/su1 0061865.

[8] *3D-Touch*, https://developer.apple.com/ios/3d-touch/.

[9] C. S. Nam and D. R. Shin. 2018. Force-touch measurement methodology based on user experience. In International

Journal of Distributed Sensor Networks. 14, 4 (April. 2018). DOI= https://doi.org/10.1177/15501477187677 94.

[10] A. Waris, I. K. Niazi, M. Jamil, K. Englehart, W. Jensen and E. N. Kamavuako. 2018. Multiday Evaluation of Techniques for EMG Based Classification of Hand Motions. In IEEE Journal of Biomedical and Health Informatics (August, 2018). DOI= https://doi.org/10.1109/JBHI.2018.2864335

[11] M. Liu, J. Shi, Z. Li, C. Li, J. Z. Zhu and S. Liu. 2017. Towards Better Analysis of Deep Convolutional Neural Networks. In *IEEE Transactions on Visualization and Computer Graphics*. 23, 1 (Jan, 2017). DOI= https://doi.org/10.1109/TVCG.2016. 25988 31

[12] B. Hudgins, P. Parker, and R. N. Scott. 1993. A new strategy for multifunction myoelectric control. In *IEEE Transactions on Biomedical Engineering* (Jan, 1993). DOI= https://doi.org /10.1109/10.204774.

[13] T. R. Farrell and R. F. Weir. 2007. The optimal controller delay for myoelectric prostheses. *In IEEE Transactions on neural systems and rehabilitation engineering* (March, 2007). DOI= https://doi.org/10.1109/TNSRE.2007.891391 .

[14] B. Peerdeman, D. Boere, H. Witteveen, R. H. Veld, H. Hermens, S. Stramigioli, H. Rietman, P. veltink and S. misra. 2011. Myoelectric forearm prostheses: State of the art from a user-centered perspective. In *Journal of Rehabilitation Research & Development* (August, 2011). DOI= https://www.ncbi.nlm.nih.gov/pubmed/21938658

[15] K. Simonyan and A. Zisserman. 2015. Very Deep Convolution Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations 2015* (San Diego, California, May 07 - 09), DOI= https://arxiv.org/ abs/1409.1556.