Hand Gesturing Human-Computer Interaction

Jinxiang Liu

Abstract

In a long time, human-computer interactions (HCI) were restricted to the using of a mouse, a keyboard and a graphic display. However, because of the development of techniques including projective displays, wireless sensor, and the recognition of sound and vision, it allows more natural, multimodal and creative interactions between machine and human. As a new advancement, the application of hand gestures presents an attractive way for communicating with computers. Through hand gesturing interfaces, users can achieve tasks more efficiently and dynamically. This paper provides a overview of hand gesturing HCI. The advantages and disadvantages of hand gesturing HCI also be analyzed, which provide the direction for future development and research.

1. INTRODUCTION

As computers have become more and more popular in the modern life, it is impossible for us to live without such appliances. Therefore, having good interactions between human and machines is important. Since the majority of human-computer interactions (HCI) mode is based on the using of keyboards, graphic displays and mice, which are the devices that restrict the naturalness and speed of interactions between people and computers, the usage of available flow of information might be limited by the existing techniques of HCI, although the technologies about communication, computing and display have been made great progress. Such limitations are very obvious for new technologies, like virtual reality.

To address this problem, many novel techniques and devices have been introduced into HCI, including applications of hand gestures. Hand gestures are the movements of human's hands, which can be used as a way for interactions without using any words. It involves not only some simple actions by using hands, like moving objects or pointing at something, but also the complicated works that convey information and represent the feelings. In this sense, hand gestures are the necessary part of the language, and thus, the using of hand as a device helps the communication between human and computer become more intuitive.

Recently, the hand gesturing HCI has been wildly applied into our real life. According to the work from Rautaray, Siddharth S., et al. (2015), applications of hand gesturing HCI generally are categorized into two types, core applications and advanced application. For the core applications, they are mainly used for the processing of information or the visualization of system, including the applications in the area of sign language translators, robotics and virtual reality. Figure 1 shows the details of this category. For the advanced applications, it includes the applications from other related areas, like medicine environments, games, and augmented reality. Figure 2 presents each part of this category.

This paper offers a review of hand gesturing HCI, including its working process and applications. Moreover, it discusses the advantages and disadvantages of hand gesturing HCI, and corresponding future work. The rest of this paper is organized as following. Section 2 is about how hand gesturing HCI works. It focuses on three main stages, which are detection, tracking and recognition of hand gestures. Section 3 introduces the applications of hand gesturing HCI, and shows some corresponding examples. In the section 4, the superiorities and limitations of hand gesturing HCI has been analyzed, which also imply the future work of hand gesturing HCI. Finally, Section 5 makes the conclusion of this paper.

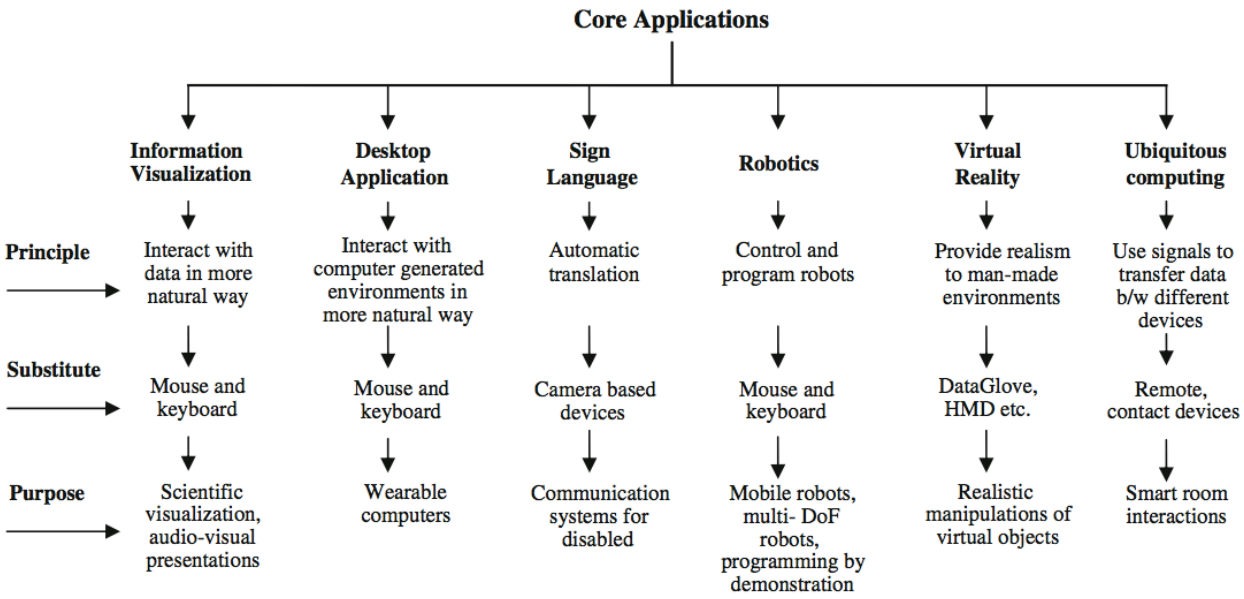

Figure 1[2] Core applications of hand gesturing HCI

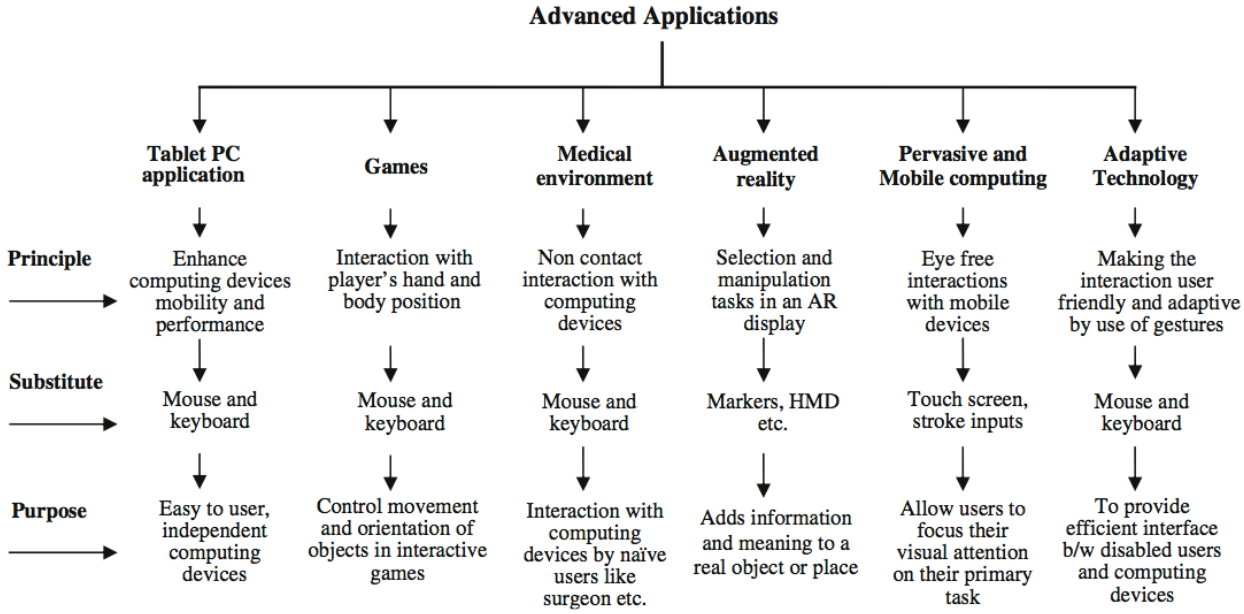

Figure 2[2] Advanced applications of hand gesturing HCI

2. MODELING AND ANALYSIS OF HAND GESTURES

For the hand gesturing HCI, hand movements are used to control the processes of system, therefore, it is important to identify the kinds of hand gestures. Generally the systems with hand gesturing interfaces include three main parts, detection, tracking and recognition of hand gestures. Figure 3 shows the diagram of such systems. Cameras are required to capture the images of hand gestures. Once the system receives such images, it starts to detect and track the hand gestures by estimating the parameters of hand gesture model. The stages of detection and tracking happen in the analysis part shown in Figure 3. Based on the estimated parameters and the knowledge with high level, the system is in the stage of hand gesture recognition.

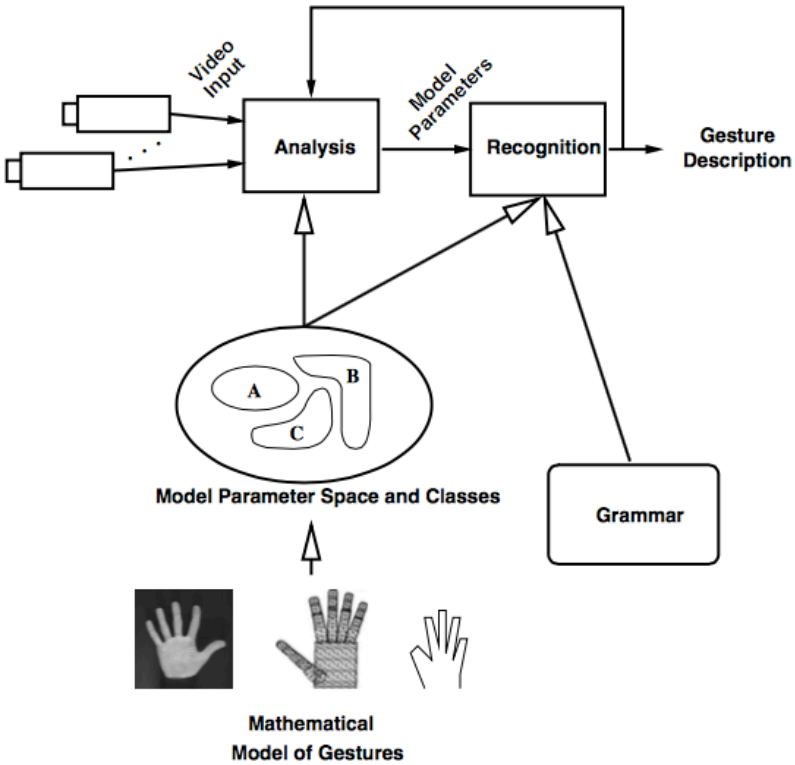

Figure 3[1] Diagram of hand gesture interpretation system

2.1 DETECTION

The detection of hand gesture is most important stage for the systems using hand gesturing interfaces. In this stage, it locatizes and extracts the hand gestures, thus the data of hand movements will be separated from the background of images. The extracted data will be passed to the stage of tracking and recognition. Many techniques have been proposed to detect the hand gestures, and generally they are based on some visual features or its combinations. These visual features can be the color of human skin, the shape and motion of hands.

Because human skin has special color footprint, color of human skin can be used as a feature to detect hand gestures. However, due to the human races or individual characteristics, the human skin color varies greately. To solve such problem, some other variabilities might be applied. The

contours of human hands in the image also contain a lot of information. Moreover, the shape of hands is independent on the color of human skin. Generally in order to extract the contour of hand, it needs to detect the edges of hands, and these edges should be irrelevant to the background of image. From this aspect, combining the hand shape extraction with color of skin and the subtraction of background helps the detection of hand gestures. The using of hand motion as a feature needs to assume that the movements of hands are the only motion in the image, which is a disadvantage of the techniques using motion cue. But there are many method have been proposed to overcome this limitation.

2.2 TRACKING

The stage of tracking is similar with the stage of detection, but it requires that the processing of images should be fast enough. The tracking of hand gestures usually is the most difficult part because of the moving and changing speed of object in the image. This stage is about the correspondence of frames. To find out the hand movements in a set of images, the segmentation of visual features of hand should be coincident from one frame to another.

Having a strong tracking stage is very important. First, the stage of tracking offers the connections of hand appearances for a set of frames, which imply the trajectories of hand gesture features over the time. Such trajectories carry crucial information about hand gestures, and this information might be used directly or after the analysis. Moreover, the tracking of hand gestures provides a method for the maintaining of estimated parameters, features and variables of the model, and these estimates might not be straightly observable at some moment.

2.3 RECOGNITION

The purpose of recognition stage is interpreting the information from extracted hand gestures. This stage includes two categories, static or dynamic hand gestures recognition. For static hand gestures, like the postures of human hands, a matcher of template or a general classifier is usually used for their recognition. These matchers or classifiers can be further divided into two groups, linear group and non-linear group, according to the type of learning algorithm. The linear techniques can only be employed on the case that data is linearly separable, otherwise, it needs to use non-linear methods. For dynamic hand gestures, because hand movements might change with the time, their recognition requires to deal with temporal dimension. Therefore, if the temporal aspect of hand gestures are not modeled, the techniques including Hidden Markov Models are usable.

There are many learning algorithms to choose for the recognition of hand gestures. These algorithms could be supervised or unsupervised learning, reinforcement learning and semi-supervised learning. However, the representation of hand gestures should be the significant factor for the learning algorithm selection. Recently, the techniques based on the automata are very popular, for example, hidden markov model (HMM), finite-state machine (FSM), past-now-future network (PNF). These automata have a set of states and corresponding transitions. The states are about static hand gestures, and the tranisitions indicate the possible changes that are allowed by the contraints. Therefore, a static hand gesture is considered as a state, while a dynamic hand gesture is represented by a path from an starting state to an ending state. The related work includes the approach proposed by Lu and Little (2006), which is about the use of a

PCA-HOG global descriptor, and the HMM model presented by Lee and Kim (1999). Although such automata-based techniques solve the hand gesture recognition problem very well, its computational complexity is huge as it proportionally increases with the number of recognized hand gestures. In addition, it needs to update such hand gestures recognition when a new hand gesture adds.

3. APPLICATION OF HAND GESTURING HCI

The application of hand gesturing HCI become more and more popular in this modern world. By introducing hand gestures into HCI, the classical devices are not essential, for example, computer mouse, joysticks and control button. Therefore the interactions between human and machines are improved into a more natural and easy way. The general information about the applications of hand gesturing HCI has been discussed in Section 1. Here, some of applications are reviewed, and Figure 4 exhibits these applications.

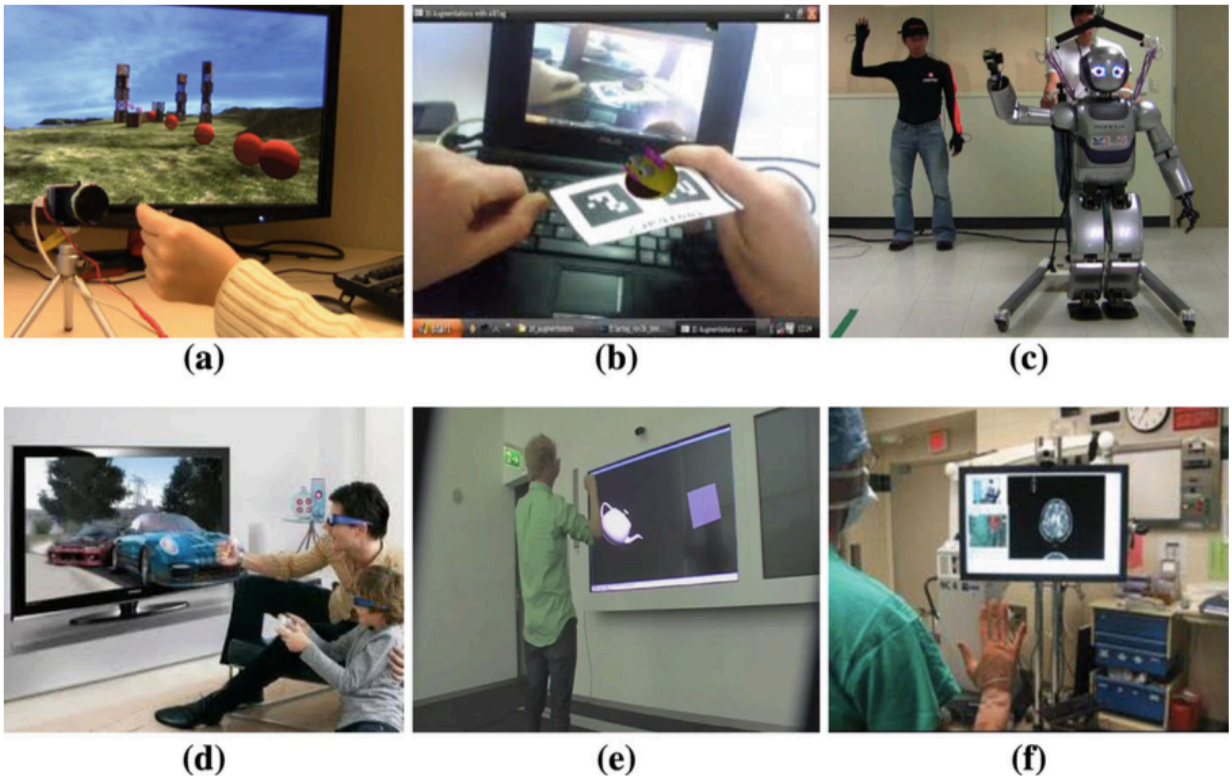

Figure 4[2] Some of examples of hand gesturing HCI

Figure 4a is about the desktop applications of hand gesturing HCI. In such applications, hand gestures substitute the interaction devices of keyboard and mouse. Through hand movements, users can control the movements of graphic object, edit and annotate the documents, browse websites, etc. Hand gestures also can be used to input information in a non-direct way, like responding to the boxes of dialogue through the gesture of nodding, however, in the desktop applications, most of them are applied as the direct input sources. Similarly, tablet computer applications can employ hand gestures for the interactions.

Figure 4b shows the hand-gesturing-HCI applications of augmented reality. Such applications usually use markers for the presenting of virtual items in the augmented reality displays. Such markers are the visual cues, consisting of designed markings on physical objects, which can be tracked more easily through computer vision. In the work from Buchmann et al. (2004), markers have been used with the bare hand gestures to finish the tasks of manipulation and selection in an augmented reality display.

In virtual reality, hand gestures have been applied for the interactions that manipulate the virtual items realistically. Such applications can track the hand movements of users to manage the orientation and the motion of the visual objects, or to navigate the environment in a virtual world, or to control the activities of avatars in a game, etc. Generally, the applications of hand gesturing HCI in virtual reality have three types, which are non-immersive interactions, semi-immersive interactions, and fully-immersive interactions. For the non-immersive interactions, within a virtual world, the users are not indicated. For the semi-immersive interactions, users are indicated as avatars within a virtual world. While for the fully-immersive interactions, users can interact like he/she is inside the virtual world. Figure 4d exhibits an application for 3D non-immersive interactions, where hand gestures are for a computer game to control the car, and the representations of users are not necessary.

The hand gesturing interactions for robots controlling are similar with the fully-immersive virtual reality interactions. Instead of in a virtual world, such interactions usually happen in the real world. It has cameras installed on the robots. Through feeding the video from these cameras to the system, the user can use hand gestures to control the activities of robots. Figure 4c is about such kind of application for robotic learning.

For the applications about computer supported collaborative work, hand gesturing interactions allow several users to interact through a shared display. Devices including desktop computer, large screen displays are needed to run such applications. Figure 4e presents such a case about applying hand gestures in HCI. Moreover, some applications have the functionality to share the annotations and notes through video streams within groups for the remote interactions.

In the medical area, numerous techniques of computer information have been used. Since some serious mistakes can result in mortal incidents, it is crucial to prevent these mistakes by using the technology in a safe way. However, as the most common devices in HCI, mice, keyboards and touch screen displays that often used by nurses or doctors are found to be the medium to spread infection in the hospital. Therefore, hand gestures provide a good way for the interactions with computer. Through hand gesturing HCI, people can control the surgeons without touching devices, and simplify the complicated surgeons. Figure 4f shows in the medical environment, a doctor performs the surgeon by using hand gestures to interact with computer devices.

As the applications using hand gestures for HCI become more and more practical, hand gesturing HCI has widespread use in our life. The hand gesture interpretation would allow the natural interactions with computers. The applications mentioned in this section are only small part of the area using hand gestures in HCI. Although a great advancement has been made, the improvements and further development are still necessary.

4. DISCUSSION OF HAND GESTURING HCI

The interactions between people and computers have great influnce on the utilization of information. In order to efficiently use the available information from computers, computer applications require to interact with human more frquently. Although the classical computer devices like mouse and keyboard provide a good way to communicate with machines, it still has some of situations that such devices are not compatible for human-computer interactions. A particular example is about the 3D interactions. It is hard for the mouse with 2 degrees of freedom to imitate a space with 3 dimensions. However, the application of hand gestures in HCI supplies an natural and attractive method. In this way, human hands gestures are the replacements of these classical but clumsy devices for HCI, and then people can communicate with computers more intuitively. The movements of human hands are important to convey the information when people interact with each other. Similarly, the use of hand gestures for HCI helps to improve the user experience and the communication with computers. Besides, hand gesturing HCI allows users to express their feelings or thoughts through hand gestures.

Although hand gesturing HCI has its irreplaceable advantages, it still has challenges in the applications. Accoring to the stages that the system works, the corresponding disadvantages will be discussed in the following.

First, in the stage of hand gestures detection, color of human skin are used as a cue to detect the hand gestures, however, it is easy to identify these color-based segmentations if the background objects have similar color distribution with human hands. When the shape of human hands is the cue for detection of hand gestures, the results highly depend on the selected classifier, and a weak classifier can lead to faulty outcomes. For the hand gestures detection based on hand motion, this technique is not very practical because of its undertaken assumption.

Second, the biggest challenge for the stage of hand gestures tracking is the ability of operation on the acquisition with high frame rate. If this problem is solved and the sytem become efficient enough, the hand gestures detection can improve to the phrase of tracking. But the limitations in the stage of detection still need to be addressed. Besides, other drawsbacks in this stage vary from the tracking methods that have been used. For example, the smoothness of contours is the issue that contour-based tracking approaches suffer, the effectiveness needs to be improved for the tracking methods based on co-related features.

Similarly, for the stage of hand gestures recognition, it also has a set of shortcomings. Generally the techniques used at each stage for the implemntation decide the overall performance of the system using hand gesturing HCI. Therefore, finding optimum combinations of techniques at each stage is essiential for the development of systems.

Developing a robust system using hand gesturing HCI is a necessary but complex task, especially, it does not have standard baseline to define the robustness specifically. However, depending on the typical issues faced, to define the robustness of system using hand gesturing HCI, it should consider from three aspects, which are user, gesture, and application. In addition, the developed system should at least have wide acceptability, good compatibility, and provide friendly user experience.

5. CONCLUSION

Recently, the utilization of hand gestures for human-computer interacitons is a thriving domain of applications and research. Because of the scientific challenges and the increasing demands of possible applications using hand gesturing HCI, more and more effort has been devoted into this area. This paper provides a brief overview of hand gestures in the HCI. It reviews the working process of hand gestures recognition, and its applications in HCI. Some of common limitations for such technogolies are still waiting to address, which also provide the direction for future development and research. This paper discusses the advantages and disadvantages of hand gesturing HCI. The effort will be needed to overcome the drawbacks and thus to develop an efficient robust system using hand gesturing HCI.

REFERENCE

[1] Pavlovic, Vladimir I., Rajeev Sharma, and Thomas S. Huang. "Visual interpretation of hand gestures for human-computer interaction: A review." IEEE Transactions on pattern analysis and machine intelligence 19, no. 7 (1997): 677-695.
[2] Rautaray, Siddharth S., and Anupam Agrawal. "Vision based hand gesture recognition for human computer interaction: a survey." Artificial Intelligence Review 43, no. 1 (2015): 1-54.
[3] Lu, Wei-Lwun, and James J. Little. "Simultaneous tracking and action recognition using the pca-hog descriptor." In Computer and Robot Vision, 2006. The 3rd Canadian Conference on, pp. 6-6. IEEE, 2006.
[4] Lee, Hyeon-Kyu, and Jin-Hyung Kim. "An HMM-based threshold model approach for gesture recognition." IEEE Transactions on pattern analysis and machine intelligence 21, no. 10 (1999): 961-973.
[5] Buchmann, Volkert, Stephen Violich, Mark Billinghurst, and Andy Cockburn. "FingARtips: gesture based direct manipulation in Augmented Reality." In Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia, pp. 212-221. ACM, 2004.
[6] Du, Hongwang, Wei Xiong, and Zuwen Wang. "Modeling and interaction of virtual hand based on virtools." In Multimedia Technology (ICMT), 2011 International Conference on, pp. 416-419. IEEE, 2011.
[7] Bellarbi, Abdelkader, Samir Benbelkacem, Nadia Zenati-Henda, and Mahmoud Belhocine. "Hand gesture interaction using color-based method for tabletop interfaces." In Intelligent Signal Processing (WISP), 2011 IEEE 7th International Symposium on, pp. 1-6. IEEE, 2011.
[8] Chen, Yen-Ting, and Kuo-Tsung Tseng. "Developing a multiple-angle hand gesture recognition system for human machine interactions." In Industrial Electronics Society, 2007. IECON 2007. 33rd Annual Conference of the IEEE, pp. 489-492. IEEE, 2007.
[9] Von Hardenberg, Christian, and François Bérard. "Bare-hand human-computer interaction." In Proceedings of the 2001 workshop on Perceptive user interfaces, pp. 1-8. ACM, 2001.
[10] Rogers, Yvonne, Helen Sharp, and Jenny Preece. Interaction design: beyond human-computer interaction. John Wiley & Sons, 2011.