Text Entry in Augmented and Virtual Environments

Steven Ding

✦

1 ABSTRACT

Modern computers and mobile devices offer sophisticated touch-based interfaces. However, if we are operating in an augmented or virtual environment, such interactions can be difficult. The ways of interaction have always been a critical issue along with the development of AR/VR technology. Text entry, being an extreme case of interaction, has not yet been solved. In this paper, we analyze issues and concerns for various text entry methods in augmented and virtual environments. Then we propose a design idea of a potential way of solving the problem. Finally, we discuss flaws and seek possible improvements to this idea.

2 INTRODUCTION AND BACKGROUND

With years of development of virtual reality (VR) and augmented reality (AR), these technologies now provide us with new affordable platforms that replace traditional monitors. Also benefiting from vastly improved computing power currently available for devices, the application field expands more than ever. However, even for recent commodity level products of VR and AR, the interaction methods are limited. Virtual reality devices, such as Oculus, and HTC Vive heavily depend on joysticks, which is a fine way of interaction for games but it also limits the use in other scenarios because of the low accuracy of joystick tracking and limited numbers of buttons that users can interact with. This makes text input an extreme case of interaction for VR devices which requires both high accuracy, and more buttons can be fixed on a pair of joysticks. As for augmented reality hardware, such as Google Glass and Hololens from Microsoft, they typically work with voice recognition. Voice recognition is a great way of giving simple instructions or making decisions, but it can be less accurate while in noisy environments and have issues in user experience and security for long text entry.

This paper proposes an idea of interaction with an easy setup that aims at text entry. It suits daily indoor use of VR/AR devices and provides both accurate entry and natural experience. This paper also reviews some concepts and ideas from the works of others.

2.1 VR/AR Concepts and Definitions

As a matter of fact, many of the theoretical and practical foundations of VR and AR studies can date back to the 1980s, plenty of attempts have been made to find appropriate definitions to describe different types of reality.

For virtual reality, a simple definition found in the dictionary [4] states: an artificial environment that is experienced through sensory stimuli (as sights and sounds) provided by a computer and in which one's actions partially determine what happens in the environment. A more recent definition of VR was proposed by LaValle [13]: "Inducing targeted behavior in an organism by using artificial sensory stimulation, while the organism has little or no awareness of the interference." Comparing to the previous definition, this one

---

• *Steven Ding, Michigan Technological University. E-mail: fding@mtu.edu.*

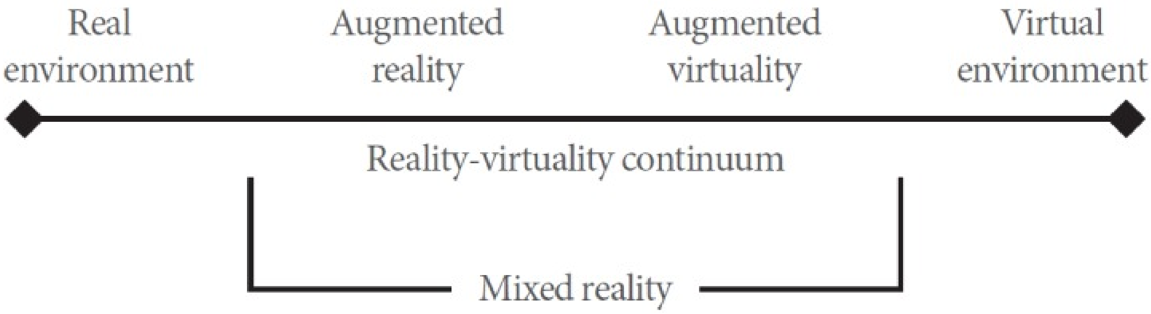

Fig. 1. Reality continuum [9]

broadens much to the field of application of virtual reality, so that the application of VR can be seen as a wide simulation environment, not only applicable to humans but generally to everything else. Moreover, it insists on the fact that the experience is deliberately created by an author with the intention to fool the senses of the user or organism. This concept removes the difficult term reality which can be subject to various interpretations, both philosophical as well as behavioral. It draws VR closer to the definition of simulation, namely that it is designed for a particular purpose. In this sense, it rejoins a concept that has been present in medicine for a long time and is a proven concept for both teaching and training.

AR is generally referred to as a system in which the user has a direct view of their environment and where a specially constructed device allows additional information or graphical elements to be blended with the real environment in the form of an overlay. The idea of augmented reality is actually from head-up displays(HUDs). The head-up displays have been used in military aircraft by pilots for easy access to information such as flight speed and altitude by creating an information overlay that users can see through. And this has been applied since the 1950s. Interestingly, there is also a concept that is opposite to AR, diminished reality. As opposed to AR, which describes a visual enhancement with artificially added information, the diminished reality refers to the processed environment from which insignificant or unwanted parts are omitted. Objects can be removed from an image by using image processing, for example, when they are occluding the view of a more important or significant [11].

There are also other concepts between and beyond VR and AR, One commonly quoted definition is the model introduced by Milgram and Kishino describing a continuum and a gradual transition as shown in the figure 1 depicts the continuum from the real world to the virtual world, leaving space in between for AR as well as for augmented virtuality and considering everything between these two worlds the mixed reality (MR) [16].

2.2 Input Methods

Input devices are always game changers for VR/AR devices. It provides users a sense of immersion and determines the way a user communicates with the device. The ultimate purpose of it is to help users to navigate and interact within a virtual or augmented environment to make it intuitive and natural as possible. But unfortunately, even after decades of development, the current state of technology is not advanced enough to support this yet.

Sutherland, with his student Bob Sproull, created the first virtual reality HMD, named *The Sword of Damocles*. This head-mount connected to a computer rather than a camera and was quite primitive as it could only show simple virtual wire-frame shapes. These 3D models changed perspective when the user moved their head due to the tracking system. It was never developed beyond a lab project because it was too heavy for users to comfortably wear; they had to be strapped in because it was suspended from the ceiling.

Fig. 2. Picture of a Sayre glove [7]

This system's only input is the move of the user's head.

The first gesture-based controller for VR ever known is the Sayre Glove [22]. Back in 1977, Thomas DeFanti and Daniel Sandin created a computer-based input device using flexible tubes wired over a physical glove shown in figure 2. Sensors measure how much light traveled through the tubes and, based on how much the tube was bent, could tell how much the finger had moved. They named the glove "The Sayre Glove" after a colleague at the laboratory, Richard Sayre, who had come up with the idea that inspired the actual glove. The Sayre Glove provided an effective method for multidimensional control, such as mimicking a set of sliders. This may have been the beginning of gesture recognition. Thomas Furness then created a working model of a virtual flight simulator, for the military, called the Visually Coupled Airborne Systems Simulator (VCASS), which is the first use of the glove-based controller in VR.

2012 is the year of the beginning of commodity-level HMDs, Luckey launched a Kickstarter campaign for the Oculus Rift which raised $2.4 million. Two years later, in 2014, Facebook bought the Oculus VR company for $2 billion. This was a defining moment in VR's history because VR gained momentum rapidly after this.

Started from 2016, many companies are developing their own VR headsets, including HTC, Google, Apple, Amazon, Microsoft Sony, Samsung, etc. Sony may be developing a similar location-tracking tech to HTC's VIVE for the PlayStation 4. Haptic interfaces were the general solution of interaction to modern VR/AR devices. Haptic interfaces are systems that allow humans to interact with a computer using their touch and movements, in other words, button-operated controllers. They are commonly used by recent commodity models such as Oculus and Vive, shown in figure 3. There is also a commodity-level product for AR named LITHO [15], which is A small wearable controller that sits between the index and middle fingers and connects via Bluetooth. Litho allows for intuitive and precise input with its touch surface on the underside, it resembles basic mouse interactions. There are also custom haptic feedback systems and an array of motion-tracking sensors.

There are also other ways of interaction used today, the most commonly used input devices are joysticks, force Balls/Tracking balls, controller wands, data gloves, trackpads, On-device control buttons, motion trackers, bodysuits, treadmills and motion platforms (virtual Omni).

We can see that with the development of the input methods, the application of the VR/AR system has greatly expanded. The freedom and ways of information inputs define the use of these systems. The input methods are of vital importance for the system's evolution process.

Fig. 3. Joysticks for HTC Vive(Left) and Oculus(Right)

2.3 Problems and Challenges

The major problem with these input methods is it lacks, or at least poorly suits text entry. The best possible way of text entry utilizing a haptic interface is a mouse-like point and click on a virtual keyboard, which is relatively slow. Another possible way of text entry is by voice recognition. But the problems with voice recognition are also not neglectable [12]. First of all, it cost more time and effort not only on entering the text, but also correcting them. Besides, it isn't necessarily true of voice recognition systems can speed things up, one may have to invest more time than expected into the process. Other problems including environmental sounds, accents, physical discomfort, and vocal problems need also take into account.

In recent times, the most natural way of text entry for a computer system under most circumstances is keyboard typing. There are also problems with using a keyboard along with the VR/AR system. The most intuitive problem with the VR system is that the user, under normal using situation, cannot see the physical keyboard. Typing with no sight to the physical keyboard is reported to be of less efficient [18]. However, the major problem with a physical keyboard is it doesn't fit daily AR/VR using scenarios. AR/VR systems are mobile systems. Most specifically, this means that the hardware required to implement an application is something that you take with you, or available wherever you go. And for text entry, the most applicable is home or indoor areas.

Therefore, we need a keyboard-like input system, which is applicable to indoor surfaces.

3 RELATED WORKS

There is a large number of attempts and work done on input systems that work with a mobile AR/VR system. A variety of sensor types are applied for different purposes.

3.1 Acoustic Systems

Microphones are a common choice for interact sensors. There are many reasons that people choose microphones. First of all, they are normally low-cost, it is not required to have high sample rates to detect different acoustics events. Besides, they are easy to install due to the size of a microphone is relatively small, which makes it suitable attachments to most indoor scenarios.

RapTapBath [23] is acoustics based interface system that converts the edge of a bathtub into a controller that recognizes hand tap gestures. A projector is installed in the ceiling to display the button on the edge of the bathtub makes it an augmented reality system, and a speaker system provides feedback on user interaction. Different types of tapping gestures: Fist, Knuckle, Pad (finger pads), Tip (tip of the finger), Fingers (multiple fingers) have a different enough waveform pattern that enables identification. This work is a great example of how acoustics sensors can differentiate ways of interactions. The problem with this approach is that there can only be a limited number of buttons on the area, suggesting the low accuracy of locating the interactions, which is critical for keyboard interactions.

There is another similar work utilized microphones. *Whoosh*[21] is a low-cost acoustics system that captures non-voice acoustic commands(ex: Blowing air, shooshing, etc.). The purpose of this work is mainly to solve situations where speech recognition may be tedious, such as some common interactions like scrolling and swiping. This work is inspired by the design of flutes. By manipulating the frequency, the system inbuilt microphone can differentiate events including short blow, double blow, long blow, swipe-up, swipe down, clockwise blow, shoosh, open exhale and sip-and-puff with accuracy over 90%.

We can see that the problem with the acoustics system is low accuracy, ill-placed microphones can be hard for location identification, and the error is inevitable during distinguishing gestures or events.

3.2 Optics Systems

Apart from acoustics systems, another choice is optics sensors. In 2014, Niikura et al. [19] attempted by fusion vision and audio recognition systems to turn flat surfaces into an input surface area. The system consists of a small high-speed camera, and a contact microphone was used to recognize several kinds of interaction between the fingers of the user and the surface. The infrared camera attached to the wrist is set-up such that it can see the tips of the fingers. The visual information from the camera is fused with the audio received by the mic. It is noteworthy that this fused data can not only recognizes the events like tap, left and right swipes, but also the finger that was used for the act. This can have a good application in text input. Since this system monitors which finger is being used. A smart keyboard system can dedicate specific letters to each finger and can develop an any-surface keyboard.

Optics systems are usually accurate and able to serves multiple purposes. The downsides are the data stream is usually heavy and we need more processing power to analyze vision data.

3.3 On-Body and Wearable Systems

On-body and wearable sensors is another major approach of interface for AR/VR systems. Chen et al. [6] developed a novel system called *iKey*. The system enables users to use the back of their palm(Opisthenar) as a virtual keyboard. The virtual keyboard system recognize keystrokes based on a location-based training model via body vibration. Mechanomyogram(MMG) is a mechanical signal that can be observed at the surface of the muscle. Which is essentially sound waves within the muscles. The human muscle MMG range is 1Hz to 150Hz which can be picked up by vibration sensors. There is a small piezoelectric ceramic vibration sensor which is embedded into a smart wristband. The sensor has a sample rate of about 600Hz. Ensembling technique is used to improve the classification accuracy. In this work, a 10-key input interface was created. The system estimates the location and tap based on the input received by the sensors and associates it to one of the 10 inputs. The average accuracy was estimated to be 92.4%.

Another work by Booth et al. [2] also used the piezoelectric contact sensors in the form of a wrist band to identify the events of finger tapping using the electric charge patterns from these sensors generated by the muscle movements around the wrist during these events. The system includes an array of 6 piezoelectric sensors, which is disk-shaped, sized 10 mm in diameter each. In order to process the data piezoelectric contact sensors picked up, signal amplification was done using an amplification circuit comprised of two resistors and a capacitor. The system only detects alternating pressure(force), static pressure on the sensor yields no signal. For this system, five-finger-tap events were correctly classified with about 96% accuracy. It also supports Vector Machine technique was used for classification of events and over 200 features was developed from the data stream coming from the 6 sensors.

Skinput [10] also resolves the location of finger taps on the arm and hand by analyzing mechanical vibrations that propagate through the body by using an array of cantilever-style piezoelectric disks attached to a forearm band. The band sensed the vibrations produced and classified taps to one of ten different locations on the user's forearm.

All of these works utilized piezoelectric vibration sensors. It is an application of the piezoelectric effect, which is the effect that certain materials have the property generating electric potential difference when subjected to physical forces(shear/stress). Quartz crystal is an example, when the crystal is either compressed or stressed, it creates an electric field across the structure. This property operates both ways: A quartz crystal can pick up even very high-frequency sound waves and transform it into alternating pulses of electric current. Or, when supplied with alternating current pulses, Quartz can produce corresponding sound frequency.

The reason why piezoelectric vibration sensors are popular is its low delay. Humans perceive real-time as an event occurring less than 300 milliseconds after the action is initiated. High-frequency systems that use piezoelectric vibration sensors can easily achieve signal window smaller than 300ms.

Others also have investigated devices worn on the fingers, wrists, or arms with different types of sensors. *Magic Finger* [24] featured a small laser emitter worn on the finger that tracked the finger motion relative to any surface. In addition, their model also recognized texture through a tiny camera. The changes in the intensity of the reflected radiation helped identify the touch events, and the closeup images taken were used to recognize a surface's texture, e.g. recognizing the difference between the table surface and a cloth. The texture detection accuracy was 98%. Funato et al. [8] estimated the contact force applied by fingers based on the speed of vibrations through the fingers. A vibration actuator and a contact microphone were attached to the index finger separated by a small distance. Low-frequency vibrations were generated using the actuator. When the index finger applied a force on a contact surface, this force changed the wave propagation patterns in the finger and the microphone received signals with different delays based on the amount of pressure the index finger had applied to the surface. A machine learning model estimated the exerted pressure based on the data from the actuator and the microphone. *The Sound of Touch*[17] is a work that transformed the forearm into an input surface by applying a wave propagation approach. Unlike others that just detect taping events on the skin, this paper presents an approach that detects continuous touch as well as arm grabbing events. An ultrasound transmitter and a receiver were wrapped around both ends of the forearm. Low-frequency ultrasound waves were transmitted onto the skin by the transmitter. The wave profiles observed at the receivers were used to infer the possible touch locations. Sound waves attenuate quickly through the skin, but when pressure is applied on the skin, the rate of attenuation decreases quickly and high amplitude signals are received. This idea was used to develop methods to classify touch events at two to six targets between a user's wrist and elbow. While classification accuracy deteriorated with more locations, the average classification accuracy was 82%.

A work by Nirjon et al. [20] is an approach to type on an imagining QWERTY keyboard. The product, *TypingRing*, is a wearable ring that determines the keys pressed and transmits the data to the device. The keyboard was divided into zones of sets of three keys to enable the user uses only three fingers to navigate on the keyboard. At every moment a specific set of three keys is highlighted on the on-screen keyboard. The user can move the palm with the ring around to move between different zones on the keyboard. The ring has an accelerometer, and proximity sensor to detect the movements of the fingers on either side. A displacement sensor is placed underneath the ring to enable locating on the flat surface. However, this approach is noticeably slower than an ordinary keyboard because there is no touch feedback and the palm movement is not as accurate.

A major benefit of on-body and wearable systems is they are extremely mobile. It is well combined with VR/AR systems that are also of mobility. The downside of these systems is the limitation of interface sizes. A wearable system cannot be too over-sized to wear, therefore the limitation of the interactive area makes it hard for performing natural keyboard typing.

3.4 Gesture systems

WiGest [1] describes itself as a ubiquitous WiFi-based gesture recognition system, the WiGest system uses the changes in the strength of the WiFi signal to sense in-air hand gestures around the mobile device. The idea was, when hands are moved in the proximity of receiver, the received signal strength indicator(RSSI) values tend to show specific patterns that can be used for gesture recognition. In this work only three states of the hand position relative to the receiver were used, UP (far from the receiver), DOWN(close to the receiver), DOWN-UP, UP-DOWN, and some other combinations of these three states. These gestures were then tied to a media player and were used to perform basic actions like volume up, volume down, play/pause, next-song, previous-song, and shuffle.

A similar idea was utilizing millimeter-wave radar, Lien et al. [14] developed a very accurate hand gesture sensing technology that is based on millimeter-wave radar. This system can track gestures with sub-millimeter accuracy, running at over 10,000 frames per second. The idea was when a radar wave emitted from a sensing device towards a moving or static object, it scatters most of it, but some get reflected back towards the source device. The time delay, phase or frequency shift, loss in amplitude, etc give a lot of information about the properties of the target object (distance, size, velocity, shape, surface smoothness, material, and orientation, etc.) The radar design illuminates the hand with a broad 150-degree radar beam at a high frequency (1 - 10kHz). The reflected signal is a superposition of reflections from multiple dynamic scattering centers that represent dynamic hand configurations. Then these superpositions can be analyzed to used to uniquely identify various hand configurations. Several static and moving hand gestures were identified with 90% accuracy.

The gesture systems are widely used in entertainment applications. The reason for that is the sensors are usually not placed on the user's body, or at least it doesn't affect user's movement. The problem is obvious too, gestures cannot carry much information you want to transmit in a fixed amount of time.

4 DESIGN AND METHODS

After analyzing the different method of interacting, if we want to have a reliable keyboard-like text entry system for VR/AR system, we want the following attributes:

- High information capacity
- Low latency
- Ability to provide real-time feedback
- High accuracy

A high information capacity can help lower the difficulty of conveying information, the gesture, therefore, is a bad choice, a normal gesture pattern recognition are limited to fewer units of information than ordinary keyboard input. The latency is determined by the complexity of processing the input. This makes optics less optimal. Usually, it takes more processing power to analyze a picture because of the higher dimension of a picture. These attributes make it optimal for using acoustics systems. The acoustics system is capable of transmitting a high capacity of information with low latency. We can use multiple acoustic sensors to instrument indoor flat surfaces into text input QWERTY keyboards. Prior work has focused on classifying taps into a small number of location targets and in some cases differentiating between different types of taps (e.g. tapping with a finger versus with a knuckle). However, no prior work has explored the idea of a more continuous two-dimensional tap localization using acoustic sensors.

When a user taps a surface instrumented with microphones, vibrations move outward from the tap at the material's speed of sound. Although the system cannot directly determine when the tap occurred, microphones can measure when the vibrations arrive at their location. Specifically, for every pair of microphones,

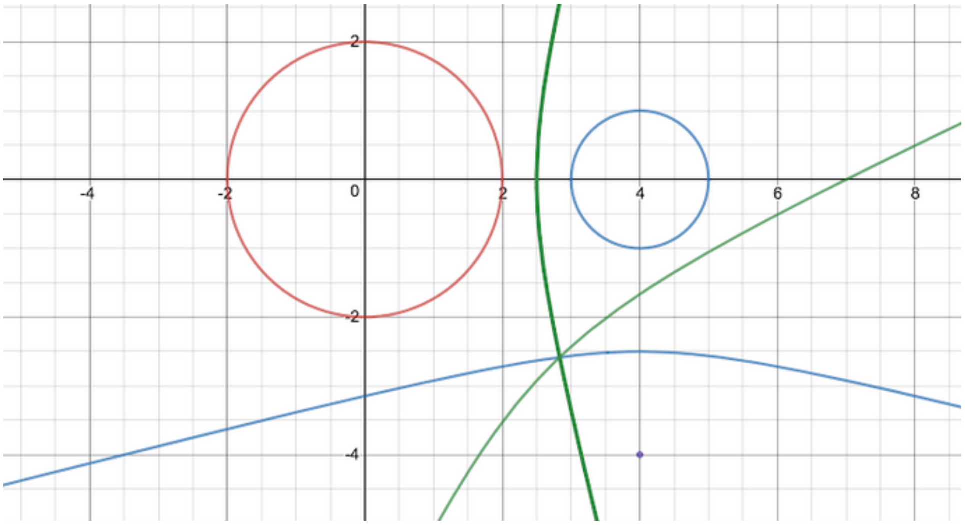

Fig. 4. Intersection of hyperbolas

we measure the number of samples (i.e., time) between the arrival times. When more than two mics are used, a single tap will produce multiple time arrival difference values, one for each pair of mics. This can be solved with a Time Difference Of Arrival(TDOA) algorithm [5], which is a well-established technique for the geolocation of RF emitters. Using three or more receivers, TDOA algorithms locate a signal source from the different arrival times at the receivers. The algorithm is basically a calculation of intersecting points of pairs of hyperbola, shown in figure 4. In this case, one hyperbola resembles the possible tap locations calculated from a time difference we get from a pair of microphones. This helps to solve the tap location, if more microphones are used, it is possible that we can locate the tapped point accurately.

As for real-time feedback, a solid surface is proven to be more beneficial and efficient with on surface targets, such as a keyboard, comparing to mid-air touching and manipulating [3]. Therefore by its nature, this method provides more feedback than other systems. Besides this, we can give acoustic feedback on detection of a user input. This will help accelerate user text input rates.

The accuracy of the system heavily depends on the locating algorithm and the sensitivity of the hardware, which needs further experiments to reinforce the idea.

5 LIMITATIONS

One apparent limitation is that this system is not mobile. Even if we can turn daily accessible flat surfaces into input areas, there can still be situations where we can't find a suitable surface. Calibration can also be a problem, the locating algorithm won't work unless we know the exact location of the microphones. This can be solved by creating a calibration step which guides the user to adjust the position of the microphones to the optimal locations. Another drawback for this method is that it requires VR devices to render a virtual keyboard to allow user to locate the keys. This can involves locating the surface's orientation and position.

REFERENCES

[1] H. Abdelnasser, M. Youssef, and K. A. Harras. Wigest: A ubiquitous wifi-based gesture recognition system. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 1472–1480. IEEE, 2015.

[2] R. Booth and P. Goldsmith. Validation of a piezoelectric sensor array for a wrist-worn muscle-computer interface. *CMBES Proceedings*, 39, 2016.

[3] G. Bruder, F. Steinicke, and W. Sturzlinger. To touch or not to touch? comparing 2d touch and 3d mid-air interaction on stereoscopic tabletop surfaces. In *Proceedings of the 1st Symposium on Spatial User Interaction*, SUI '13, page 9–16, New York, NY, USA, 2013. Association for Computing Machinery.

[4] M.-W. c2016. Virtual reality. simple definition of virtual reality [internet] springfield (ma). `http://www.merriam-webster.com/dictionary/virtualreality`.

[5] Y.-T. Chan and K. Ho. A simple and efficient estimator for hyperbolic location. *IEEE Transactions on signal processing*, 42(8):1905–1915, 1994.

[6] W. Chen, Y. Lian, L. Wang, R. Ruby, W. Hu, and K. Wu. Virtual keyboard for wearable wristbands. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, pages 1–2, 2017.

[7] E. V. L. (EVL). Sayre glove (first wired data glove). `https://www.evl.uic.edu/entry.php?id=2162`.

[8] N. Funato and K. Takemura. Estimating three-axis contact force for fingertip by emitting vibration actively. In *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 406–411. IEEE, 2017.

[9] A. Hamacher, S. J. Kim, S. T. Cho, S. Pardeshi, S. H. Lee, S.-J. Eun, and T. K. Whangbo. Application of virtual, augmented, and mixed reality to urology. *International neurourology journal*, 20(3):172, 2016.

[10] C. Harrison, D. Tan, and D. Morris. Skinput: Appropriating the body as an input surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, page 453–462, New York, NY, USA, 2010. Association for Computing Machinery.

[11] J. Herling and W. Broll. Pixmix: A real-time approach to high-quality diminished reality. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 141–150. IEEE, 2012.

[12] D. Kambeyanda, L. Singer, and S. Cronk. Potential problems associated with use of speech recognition products. *Assistive Technology*, 9(2):95–101, 1997.

[13] S. LaValle. *Virtual reality*. National Programme on Technology Enhanced Learning (NPTEL), 2016.

[14] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev. Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics (TOG)*, 35(4):1–19, 2016.

[15] LITHO. Litho. a controller for the real world. `https://www.litho.cc/`.

[16] P. Milgram and H. Colquhoun. Ismrpaper, 06 2014.

[17] A. Mujibiya, X. Cao, D. S. Tan, D. Morris, S. N. Patel, and J. Rekimoto. The sound of touch: on-body touch and gesture sensing based on transdermal ultrasound propagation. In *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces*, pages 189–198, 2013.

[18] H. Nicolau, K. Montague, T. Guerreiro, A. Rodrigues, and V. L. Hanson. Typing performance of blind users: An analysis of touch behaviors, learning effect, and in-situ usage. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*, ASSETS '15, page 273–280, New York, NY, USA, 2015. Association for Computing Machinery.

[19] T. Niikura, Y. Watanabe, and M. Ishikawa. Anywhere surface touch: utilizing any surface as an input area. In *Proceedings of the 5th Augmented Human International Conference*, pages 1–8, 2014.

[20] S. Nirjon, J. Gummeson, D. Gelb, and K.-H. Kim. Typingring: A wearable ring platform for text input. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, pages 227–239, 2015.

[21] G. Reyes, D. Zhang, S. Ghosh, P. Shah, J. Wu, A. Parnami, B. Bercik, T. Starner, G. D. Abowd, and W. K. Edwards. Whoosh: non-voice acoustics for low-cost, hands-free, and rapid input on smartwatches. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, pages 120–127, 2016.

[22] D. J. Sturman and D. Zeltzer. A survey of glove-based input. *IEEE Computer Graphics and Applications*, 14(1):30–39, Jan 1994.

[23] T. Sumida, S. Hirai, D. Ito, and R. Kawakatsu. Raptapbath: User interface system by tapping on a bathtub edge utilizing embedded acoustic sensors. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, pages 181–190, 2017.

[24] X.-D. Yang, T. Grossman, D. Wigdor, and G. Fitzmaurice. Magic finger: always-available input through finger instrumentation. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 147–156, 2012.