2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 978-1-6654-4057-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/VRW52623.2021.00034

Distance Estimation with Mobile Augmented Reality in Action Space: Effects of Animated Cues

Soumyajit Chakraborty*
Vanderbilt University, USA

Jeanine K. Stefanucci†
University of Utah, USA

Sarah Creem-Regehr‡
University of Utah, USA

Bobby Bodenheimer§
Vanderbilt University, USA

ABSTRACT

Augmented reality is standard on many modern smartphone platforms. The distance of virtual objects seen through the smartphone display should be perceived accurately for easy interaction with virtual objects with high fidelity. We investigate whether distance perception through mobile augmented reality devices is affected by an animated avatar. The avatar walks and is positioned from near to far action space. We conduct a distributed experiment "in the wild" to investigate these effects.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Augmented reality; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Collaborative interaction

1 INTRODUCTION

Augmented Reality (AR) is quickly being adopted by many applications like those for navigation (such as Google Maps), and those for large-scale tasks such as room design (e.g., to explore how placement of furniture in a room might look). Because of these new applications, it is becoming important to understand how users perceive distances to virtual targets projected onto the real world. If a difference exists between the perceived distance to a virtual object and the perceived distance to a real object at the same distance, then a user could develop a distorted spatial representation of augmented reality, and the perceptual registration of the virtual object will be biased. Some evidence exists [2, 12, 17] that distance judgments in AR are distorted in head-mounted displays.

However, today many mobile devices such as tablets and phones are capable of running AR applications, using the camera on the device as a video see-through display. Due to the prevalence of smartphones and tablets, these apps are the most common form of AR experienced by ordinary users. There has been little work done on the perception of distance through these devices, though, and what exists has been mixed. Swan et al. [20] found distance underestimation in tablet-based AR, whereas a recent study by Liu et al. [11] found accurate distance perception through a smartphone.

The present paper presents work aimed at clarifying the accuracy of distance estimation through a smartphone as well as exploring a factor that may improve it if bias in estimation is observed. Part of the motivation of this work was to see if the results of Liu et al. [11] were replicable. A second consideration was that we were interested in the addition of a motion cue to the target stimulus. In the head-mounted display literature, e.g., Rosales et al. [17], subjects had difficulty discriminating whether virtual objects were in contact with the ground or not. If ground contact is not correctly perceived, then it is straightforward to understand how distances could be misperceived. We hypothesized that a motion stimulus such as walking would help the user perceive contact between virtual object and ground better. Finally, because this work was done during the COVID-19 pandemic and conducted on smartphones, we conducted it "in the wild". Users downloaded the distance estimation app onto their smartphone and did the experiment in their own surroundings. Challenges associated with that practice are discussed.

2 RELATED WORK

There has been a significant amount of work done on distance estimation in AR [2–10, 12, 14, 15, 17–21]. However, there is still no consensus on the accuracy of perceived distance through various platforms for viewing AR. There could be various reasons for the lack of a consensus, including (1) environmental contexts and available cues, (2) methodologies and designs employed, (3) distances considered, and 4) the AR devices themselves. Although previous results mostly suggest that people underestimate distance in Augmented Reality, there is work that shows that people overestimate distances if the distance is farther than a certain distance or if the experiment has been conducted in an outdoor space rather than in an indoor space. For example, distances estimated in *personal* to *action* spaces using various optical see-through displays (such as the Microsoft HoloLens) are consistently underestimated [6, 17, 19]. However, Swan et al. [18, 19] found that people underestimated distances up to 23 meters but overestimated beyond it when they were asked to estimate the egocentric distances of some virtual objects placed at different locations ranging from 5 to 45 meters. To test whether indoor or outdoor environment plays a role in the decision-making process of the observer, Livingston et al. [12] ran a similar experiment in an indoor environment as well as an outdoor environment. They found that people underestimated the distances in the indoor environment but overestimated the distances in the outdoor environment.

With the advent of various mobile AR devices (such as various Android or iOS devices), researchers have started exploring accuracy of perceived distance in these devices compared to the optical see-through head-mounted displays used previously. Even though there is less work on mobile devices in this area, the findings are interesting. The typical wide-angle camera used in most tablets means that there is likely a difference between where the viewer will look at the displayed scene and where the correct viewing position is based on the center of projection of the camera. This difference in viewer- and camera-based point of projection is discussed in the picture perception literature and findings suggest that it can lead to perceptual distortions in the geometrically predicted direction of expansion or elongation of space [16, 20]. Dey and colleagues investigated the influence of different AR visualization methods and sizes of tablet displays on egocentric distance perception [2–5]. Even though they found that there was underestimation in egocentric distance estimation by participants, the underestimation of distance was reduced with smaller devices (such as iPhones) and increased with larger devices (such as iPads).

Given the prior work on distance estimation using mobile AR devices, to our knowledge the effect of other depth cues (such as motion cues) to a virtual object (such as a life-size human avatar) on an observer's ability to estimate the absolute egocentric distance

---

*soumyajit.chakraborty@vanderbilt.edu
†jeanine.stefanucci@psych.utah.edu
‡sarah.creem@psych.utah.edu
§e-mail: bobby.bodenheimer@vanderbilt.edu

978-1-6654-4057-8/21/$31.00 ©2021 IEEE
DOI 10.1109/VRW52623.2021.00034

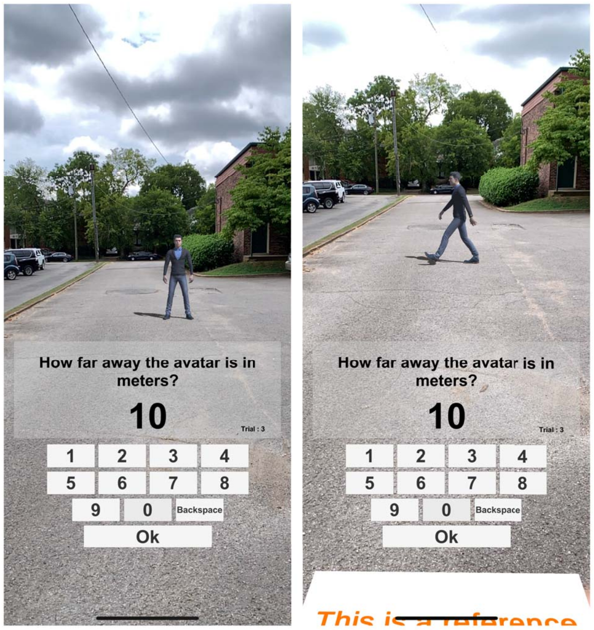
easier has not been assessed. In our current work, we implement this phenomenon and present our findings.

3 EXPERIMENT

The goal of the current experiment was to assess whether adding a walking animation to a life-size human virtual avatar would help people to estimate the absolute egocentric distance of the avatar accurately. At the time of this research, in-person experimentation with a head-mounted device was not possible at our institution, so we decided to follow the method of Liu et al. [11] and implement this experiment on a smartphone. Liu and colleagues did their work in a controlled experimental setting; however, such a setting was not available to us, so we coded our application for both Android and iOS. The app could be downloaded and installed on a user's phone. Informed consent was obtained through the app, and the user's data was emailed to us where it was anonymized and analyzed.

We hypothesized that that distance would be underestimated, since most prior work in AR [4, 5, 12, 17–19] finds distance underestimation, although Liu et al. [11] is an exception. Likewise, distance estimation in the real world in action space as measured through verbal reports typically shows underestimation [1, 13]. We also predicted that a motion cue would improve distance estimates. As there is not much evidence that environmental context influences distance estimation in AR [7], we gave users the freedom to choose whether they did the experiment indoors or outdoors. Regardless, we informed them that they would need about about 30 m of clear space to do the experiment, so we expected that most participants would do the experiment outdoors.

3.1 Participants


How far away the avatar is in meters?

10

Trial : 3

How far away the avatar is in meters?

10

Trial : 3

Figure 1: Left: A view of the *Standing Avatar* around 10 meters. Right: A view of the *Walking Avatar* around 10 meters. The user has estimated the absolute egocentric distances as *10 meters* for both the avatars using the GUI.

We recruited twenty-five participants from several places throughout the globe, which included countries from the USA and India. We were able to do this as we published the app in the App Stores for Android and iOS devices. Although we published the Android App in the Google Play Store publicly, we were unable to do so in the case of the Apple App Store. Apple did not allow us to release this app publicly in their App Store, due to security reasons. However, we were able to send the beta app to our participants via email, and thus the iOS users were also able to run the app. Among the participants, there were 13 female and 12 male participants whose ages ranged between 19 years to 69 years. An online consent form was included with the app so that when the participant ran the app, they could go through it and give their consent.

3.2 Materials and Design

We created an app that could be run in any ARCore supported Android devices which included various Android smartphones and tablets, and ARKit supported iOS devices which included various iPhones and iPads. The main app was built using the Unity game engine. We used Unity 2019.4 for developing this app. Because no experimenter could be physically with the participants while they ran the app to answer any questions they may have had, we developed a very easy to understand, thorough, and step-by-step graphical user interface (GUI) to guide the participants on what to do throughout the experiment. Because of the remote location of the participants, we incorporated a built-in, auto-email-sending facility to our app so that when the participants completed the experiment, the app sent the data to us via email as a CSV file along with other relevant information as simple text in the body of the email. This facility requires a continuous and stable internet connection. So, upon starting the app, the app checked whether there was a stable internet connection or not. If it could find one, it allowed the participant to run the experiment. Otherwise, it prompted the participant to connect to a stable internet connection first.

Because we could not monitor participants while they were running the app, we did not know whether a participant was sending multiple data files from the same device or not. Further, we could not know if multiple participants used the same device to send their data to us. To eliminate this problem, we simply created a text file and stored it in the device's memory after the user sent the data to us successfully. If they were to open the app again, it would find the text file in the device's memory and would not allow the user to go further with the experiment.

We used one Male character (Josh), and one Female character (Kate) from Mixamo for our experiment. The participants were given a gender-matched avatar. Note that we did not have an a priori hypothesis that gender would affect the distance estimation results, but we thought a gender-matched avatar was best, as is standard practice in many such experiments. For the walking animation, we chose the default walking animation provided by Mixamo for both avatars. The participants were shown a standing avatar or a walking avatar randomly throughout the experiment, which was instantiated at different positions from the mobile phone in each trial. For the walking avatar, it walked side-to-side up to 1 meter on each side from where it was instantiated by the app in each trial. We implemented a side-to-side walking animation in order to provide motion as a potential cue for depth while keeping the actual distance to the avatar constant. We added shadows to the avatar using a Unity Asset called *Simple AR-VR Invisible Light and Shadow Receivers*, which could be purchased from the Unity Asset store, so that the participants could easily understand that the avatar was standing on the ground. We did not directly ask participants whether they felt that the avatar was on the ground or not during our experiment. However, our anecdotal experience in pilot testing was that the shadows were helpful for estimating the distance to the avatar. We used Unity's default Skybox lighting in our app. We used the amount of real light detected by the phone camera to gauge intensity and to change the direction of the shadow of the avatar during the run time of the app. Due to Unity's default Skybox lighting settings, the avatar should be easily visible in most real-world indoor or outdoor lighting conditions. Also, if participants ran the app indoors, ambient and positional lighting was detected in the room, and the shadow of the avatar was projected accordingly. For publishing the app, we used the Google Play Console for Android devices and the Apple App Store Connect from the Apple Developer portal for the iOS devices. Once approved by Google and Apple, we were able to share the apps among participants.

3.3 Procedure

After obtaining consent from the participants, the app would gather demographic information including the name, age, and gender of the participants. It then asked the participants to find an indoor or outdoor space that extended to 30 meters without obstructions. When the participants identified the space, the app would ask the participants whether they would like to estimate distance in feet, yards, or meters. After choosing their preferred units, the app would ask the participants to point the mobile phone down so that it could detect the floor. When the floor was detected, the app asked the participants to move the phone upwards slowly to point towards the farthest distance in front of the person and then tap anywhere on the detected floor to instantiate the gender-matched avatar at 3 meters away from the device. The app asked whether the avatar was standing in the right place or not. If the position of the avatar did not seem good to the participants, they could redo the previous step again to place the avatar in the right place. After placing the avatar at a suitable place, the app showed participants a white square plane on the ground that participants were told could serve as a reference to help them to estimate distances. On this reference square, the length of its sides were shown to participants in either meters, feet, or yards to correspond to the units chosen by the participant. The reference was exactly 1 meter square. After finishing all of these setup phases, the participants were told not to move from where they were standing, and the experiment started thereafter.

The experiment was within-participants and there were twelve total trials, where each of the avatars (standing and walking) randomly instantiated around 10 meters, 18 meters, and 25 meters away from the device. At each place, the avatar placement was jittered between ±1 meter from the device as well as either 0 meters or 1 meter to the left or right side of its instantiation point randomly, so that the participant could not use consistent references in the real environment to estimate the distance to the avatars. During each trial, the participants were given a GUI where they were able to enter the value of the estimated egocentric distance of the avatar and hit the ”Ok” button to go to the next trial. After the end of the twelve trials, the participants were asked to rate whether the walking animated avatar helped them to estimate the egocentric distance of the avatar more easily than the standing avatar. All of the recorded data was then sent by the app to the experimenter via email and the app showed the participants a confirmation.

3.4 Analyses and results

We ran a 2 (avatar animation: walking vs. standing) x 3 (distance) repeated measures analysis of variance (ANOVA) on the distance estimations. There was a significant main effect of distance, $F(2,48) = 13.88, p < .001, \eta_p^2 = .36$, as expected, showing that estimations increased with actual avatar distance. Notably, estimations were relatively accurate with some underestimation at 25 meters ($Ms = 9.27$m, 17.21m, and 21.78m for 10m, 18m, and 25m actual distances, respectively). There was no effect of avatar animation, $F(1,24) = 1.44, p = .24$. The results are shown in Figure 2.

After analyzing the answer to the survey question asked at the end of the experiment about whether participants felt the walking avatar helped them estimate distance, we found that people choose the ”Neither agree nor disagree” option most frequently, on average. This finding may have contributed to the lack of effect of the walking avatar on distance estimates (compared to the standing avatar) if people did not find the walking avatar more helpful. Three participants reported running the experiment in an indoor environment and the rest reported running the experiment outdoors. There was no obvious difference in the results given this small sample size.

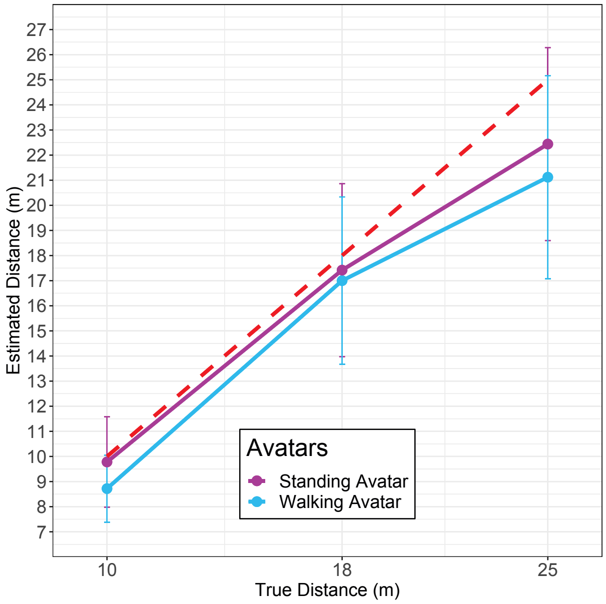


Figure 2: Mean distance estimates as a function of distance for the standing (magenta) and walking (cyan) avatar conditions, collapsed across environment. The red dashed line indicates veridical estimation. Error bars indicate one standard error of the mean.

4 GENERAL DISCUSSION

The results from our experiment replicated the findings of Liu et al. [11], who found that participants accurately estimated the egocentric distance of a life-size human avatar placed at various distances using a mobile AR display compared to distance estimates made to a real person standing in the same locations viewed through the same mobile AR display. We found that as the distance to the avatar increased, people tended to underestimate the egocentric distance of both the standing and the walking avatars, which supports the results of Liu et al. [11], as well as other studies in AR using different devices [4, 5, 12, 17–19]. Contrary to our hypothesis, the added animation of the avatar did not help people more accurately estimate the absolute egocentric distance of the walking avatar. However, in retrospect this lack of a finding is likely due to the fact that people are relatively accurate at distance estimation, so an added motion cue provided little benefit to their already accurate estimates. There was a slight but consistent underestimation of distance for both the standing and walking avatar as distance increased, but this underestimation was present for both avatar conditions. Given the large variability in distance estimates that was found across participants (as evidenced by the large error bars in the Figure), any effect of motion of the avatar may have been hard to detect. Future work could test for the influence of motion cues on distance estimation in mobile AR in a more controlled laboratory setting to try to reduce some of this between-subject variability.

In line with the findings of Liu et al. [11], we showed fairly accurate distance perception in mobile AR. This general accuracy in estimation of distance is in contrast to several studies for which distances were underestimated when presented through head-mounted, optical see-through AR devices that showed underestimation. It is interesting that distance perception in mobile AR has generally been more accurate. As noted by Swan et al. [20], it is possible that there are perceptual distortions that result from a viewpoint in the real world that was offset from the viewpoint of the device. This conflict in the center of projection would predict distortions that would expand the perceived environment, possibly counteracting the distance underestimation that is typically seen in the real world and in optical see-through devices such as the Microsoft HoloLens. Future work should consider directly comparing distance estimates for static and moving avatars at farther distances in both devices.

An important contribution of the current work is that it shows that running mobile AR experiments across distributed users who are outside of a laboratory setting is feasible. Of course, there are technical and procedural challenges to running such studies. Technically, the experimental apps must be significantly more robust than typical laboratory applications. Obtaining consent and properly gathering data are technical challenges that can be overcome, as demonstrated here. Understanding what users are doing during the experiment and why they are doing it becomes more problematic, as well as understanding if users are properly following the directions. A larger sample size is likely needed to reduce the variance of environmental factors as well as variability across participants. However, the current work provides a foundation for future work to build on with regard to running a distributed AR experiment with mobile devices.

5 CONCLUSION

Our study investigated egocentric distance estimation to a standing, life-size human avatar and a walking, life-size human avatar that were viewed through mobile AR displays. We hypothesized that distances to the avatar would be underestimated, but that motion cues produced from the walking avatar might aid in distance estimation. Contrary to our hypothesis, we found that people were generally very accurate in estimating the distances to the avatars regardless of motion condition. Future work could consider replicating this experiment with more participants in order to decrease between-participant variability, which may have contributed to the lack of a motion effect. Also, due to the remote nature of the experiment, we could not control the environments in which the avatars were viewed or participants' behaviors while viewing. Overall, our initial findings suggest that added animation does not help participants to estimate egocentric distance to a life-size, virtual avatar in *medium-field spaces*. Future work could address questions left open by this initial study, including running a follow-up in a more controlled laboratory setting.

ACKNOWLEDGMENTS

The authors thank Julien Pettré for helpful suggestions in the early stages of this work, and the reviewers for constructive comments. This material is based upon work supported by the National Science Foundation under grant 1526448, and by the Office of Naval Research grant N00014-18-1-2964.

REFERENCES

[1] J. Andre and S. Rogers. Using verbal and blind-walking distance estimates to investigate the two visual systems hypothesis. *Perception & Psychophysics*, 68(3):353–361, 2006.

[2] A. Dey, M. Billinghurst, R. W. Lindeman, and J. E. Swan. A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. *Frontiers in Robotics and AI*, 5:37, 2018. doi: 10.3389/frobt.2018.00037

[3] A. Dey, A. Cunningham, and C. Sandor. Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, pp. 211–218. ACM, 2010.

[4] A. Dey, G. Jarvis, C. Sandor, and G. Reitmayr. Tablet versus phone: Depth perception in handheld augmented reality. In *2012 IEEE international symposium on mixed and augmented reality (ISMAR)*, pp. 187–196. IEEE, 2012.

[5] A. Dey and C. Sandor. Lessons learned: Evaluating visualizations for occluded objects in handheld augmented reality. *International Journal of Human-Computer Studies*, 72(10-11):704–716, 2014.

[6] C. Diaz, M. Walker, D. A. Szafir, and D. Szafir. Designing for depth perceptions in augmented reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 111–122. IEEE, 2017.

[7] H. C. Gagnon, L. Buck, T. N. Smith, G. Narasimham, J. Stefanucci, S. H. Creem-Regehr, and B. Bodenheimer. Far distance estimation in mixed reality. In *ACM Symposium on Applied Perception 2020*, pp. 1–8, 2020.

[8] J. A. Jones, J. E. Swan, II, G. Singh, E. Kolstad, and S. R. Ellis. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV '08, pp. 9–14. ACM, New York, NY, USA, 2008. doi: 10.1145/1394281.1394283

[9] L. Kuparinen, J. II, S. Rapson, and C. Sandor. Depth perception in tablet-based augmented reality at medium-and far-field distances. In *SAP'13 Proceedings of the ACM Symposium on Applied Perception (pp. 121). New York: Association for Computing Machinery. ISBN 978-1-4503-2262-1*. Association for Computing Machinery, 2013.

[10] W. Lages, Y. Li, L. Lisle, T. Höllerer, and D. Bowman. Enhanced geometric techniques for point marking in model-free augmented reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 301–309. IEEE, 2019.

[11] J. M. Liu, G. Narasimham, J. Stefanucci, S. Creem-Regehr, and B. Bodenheimer. Distance perception in modern mobile augmented reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 196–200, March 2020.

[12] M. A. Livingston, Z. Ai, J. E. Swan, and H. S. Smallman. Indoor vs. outdoor depth perception for mobile augmented reality. In *2009 IEEE Virtual Reality Conference*, pp. 55–62. IEEE, 2009.

[13] J. M. Loomis. Three theories for reconciling the linearity of egocentric distance perception with distortion of shape on the ground plane. *Psychology & Neuroscience*, 7(3):245, 2014.

[14] G. Pointon, C. Thompson, S. Creem-Regehr, J. Stefanucci, and B. Bodenheimer. Affordances as a measure of perceptual fidelity in augmented reality. In *2018 IEEE VR 2018 Workshop on Perceptual and Cognitive Issues in AR (PERCAR)*, pp. 1–6, March 2018.

[15] G. Pointon, C. Thompson, S. Creem-Regehr, J. Stefanucci, M. Joshi, R. Paris, and B. Bodenheimer. Judging action capabilities in augmented reality. In *Proceedings of the 15th ACM Symposium on Applied Perception*, pp. 1–8, 2018.

[16] S. Rogers. Perceiving pictorial space. In *Perception of space and motion*, pp. 119–163. Elsevier, 1995.

[17] C. S. Rosales, G. Pointon, H. Adams, J. Stefanucci, S. Creem-Regehr, W. B. Thompson, and B. Bodenheimer. Distance judgments to on-and off-ground objects in augmented reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 237–243. IEEE, 2019.

[18] J. Swan, M. A. Livingston, H. S. Smallman, D. Brown, Y. Baillot, J. L. Gabbard, and D. Hix. A perceptual matching technique for depth judgments in optical, see-through augmented reality. *Proceedings of the IEEE Virtual Reality Conference, 2006*, pp. 19–26, 2006.

[19] J. E. Swan, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE transactions on visualization and computer graphics*, 13(3):429–442, 2007.

[20] J. E. Swan, L. Kuparinen, S. Rapson, and C. Sandor. Visually perceived distance judgments: Tablet-based augmented reality versus the real world. *International Journal of Human–Computer Interaction*, 33(7):576–591, 2017.

[21] J. Wither and T. Höllerer. Pictorial depth cues for outdoor augmented reality. In *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, pp. 92–99. IEEE, 2005.