Hand Gesture-based Smartphone Interaction

VENISSA CAROL QUADROS, McGill University, Canada

Smartphones stand to benefit from alternate input techniques like gesture-based input. Since most modern-day smartphones come equipped with inertial measurement unit (IMU) sensors, gesture recognition-based interaction systems can be implemented on these devices without any additional hardware. However, given that gesture-based interaction with smartphones is relatively uncommon, it is necessary to evaluate the acceptability of such a system in social settings. A simple hand gesture-based interaction system was implemented and a week-long in-the-wild study was used to glean insight on social acceptability, gesture preferences and user sentiment. Despite variations among participants, a general positive trend in social acceptability was observed over the seven-day period in some cases and some rationales for gesture preferences and overall user views were identified.

CCS Concepts: • **Human-centered computing → Gestural input**; *Smartphones*; *Field studies*; Accessibility technologies.

Additional Key Words and Phrases: gestural input, smartphone interaction, social acceptability

ACM Reference Format:
Venissa Carol Quadros. 2025. Hand Gesture-based Smartphone Interaction. 1, 1 (May 2025), 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Input techniques in touchscreen smartphones are primarily touch-based [13] with physical buttons and voice commands being largely limited to special functions like volume control and voice assistants respectively. The small screen size in smartphones makes these interactions heavily reliant on visual attention and precision. Despite this, touch-based interfaces are pervasive not just in the mainstream but also in technology geared toward smartphone accessibility for individuals with visual impairments. This includes built-in accessibility software like “TalkBack” for the Android platform and “VoiceOver” for the Apple iOS which overlay a screen reader and some variations in the touch gestures.

The need for alternate input interfaces is, however, not limited to the accessibility space. The motivation to develop and use eyes-free interaction can rather be influenced by various factors including lighting conditions, social settings, and lower perceived effort [15]. This has spurred research to explore alternative input modes like hand gesture detection.

Most smartphones are equipped with IMU sensors which find use in applications like location tracking and navigation services. The research focused on gesture recognition using the data from these sensors has been quite successful in the recognition of continuous gestures based on accelerometer data [12], and with both accelerometer and gyroscope data for continuous [6] and discontinuous gestures [14, 16] in resource-constrained settings in real-time. Other works have

Author’s address: Venissa Carol Quadros, venissa.quadros@mail.mcgill.ca, McGill University, 845 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 0G4.

XXXX-XXXX/2025/5-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

combined IMU-based gesture recognition with eye-tracking [8] and spatial audio [4] in their user interface design.

However, the use of gestures for smartphone interactions is relatively uncommon and, despite phones being widely accepted in social settings, the use of gesture-based interaction is currently limited to some publicly deployed apps and games. Therefore, there is a need to design gestures that are subtle, similar to existing interactions and/or natural movements and also enjoyable in order to increase their acceptability [11].

Based on this need, this project focuses on building a simple hand gesture-based smartphone interaction system and evaluating its social acceptability. In this work, a basic application was developed to recognize six simple hand gestures using the IMU data in an Android smartphone. The gesture-based interaction system worked as an overlay on the existing touch-based user interface by allowing the user quick access to six applications- one mapped to each gesture- by bringing them to the foreground screen when the occurrence of the corresponding gesture was detected.

This application was installed on the users' personal smartphones and the mapping of gestures to applications was based on user preference. A week-long in-the-wild user study was conducted to evaluate the social acceptability of the hand-gesture-based interaction system.

This preliminary work focused on understanding the feasibility and desirability of smartphone interactions using hand gestures and also to potentially direct the focus of future development. The system prototype and preliminary study detailed in this work do not seek to achieve the robustness of a commercially ready product or even past works focused primarily on the gesture recognition algorithm. Instead, this project seeks to evaluate user comfort and acceptance of the new interaction mode, identify user preferences among gestures, and understand the effect of recognition reliability on system favorability and user sentiment.

2 SYSTEM DESIGN

The system consists of three major components: the base Android application that runs perpetually in the background, the gesture recognition logic which uses the IMU data to identify gestures and the mapping strategy which defines how the user utilizes gestures to interact with their device.

2.1 Android Application

The first step in implementation was developing an Android application which was capable of continuously running the gesture recognition service in the background and also calling an application to the foreground when the corresponding gesture was detected. Running a service in the background is possible using features like Android Foreground services which can run even when the user isn't directly interacting with the application. However, as of Android 10 (API level 29) restrictions have been placed on starting applications from the background in order to enhance the user's control of what's displayed on the screen and displaying notifications has been recommended when feasible [1]. Therefore, although the gesture recognition could run in the background of an ordinary application the restrictions necessitate that the user opens the application for it to be able to launch the application mapped to the detected gesture. This wasn't favourable for the outcomes of the study since without the benefit of quick eyes-free access offered by the new interaction scheme there would be little incentive or benefit to a user from the general population to use it.

In order to overcome this limitation and given the system's alignment and potential benefit to accessibility, the base application was developed as an AccessibilityService. Given that an AccessibilityService is one among the few cases that might require launching a UI, the restrictions mentioned previously do not apply.

I benefitted greatly from the publicly available code base and tutorial for GlobalActionBarService by Google Codelabs for getting started with building an AccessibilityService. I made some minor

(a) Twist (b) Fold (c) Tilt

Fig. 1. Three primitive gestures for gesture-based interaction system

modifications to the code provided to adapt it to my use case and then moved on to building the Gesture Recognition service.

2.2 Gesture Recognition

Three primitive gestures namely twist, fold and tilt were defined as shown in Figure 1 based on gyroscope data for angular speed around each axis. These along with three more gestures- double twist, double fold and double tilt - derived by the repetitive execution of the primitive gestures formed the six gestures used in this work.

To limit the implementation complexity, a simple threshold-based gesture recognition logic was developed. It identified the primitive three gestures when the gyroscope measure for the corresponding axis exceeded a set threshold value. If a primitive gesture occurred for the second time within one second of the initial detection of the gesture, it was classified as a 'double' version of the primitive gesture. A flowchart of the logic is shown in Figure 2.

Initially, constant values were selected as the threshold value based on angular velocity values observed using the Physics Toolbox Suite Android application when performing the gestures. The directionality of the threshold velocity was set based on the convenience of performing gestures. However, as the participant preferences varied and the system reliability rating was intended to be self-reported, the velocity values as well as the gesture directionality were adjusted as per their individual preferences and the consequent observations were recorded.

2.3 User Customization and Mapping Strategies

Despite the benefits of having a logical gesture to application mapping owing to the limited gesture-based vocabulary, the gesture mapping was customized based on user preferences in order to incentivize the user to use the system more frequently. Each of the six gestures worked as a shortcut to launch an application of the user's choice. This customization was possible only during the initial setup and could not be modified further by the participant. This helped to avoid variations in

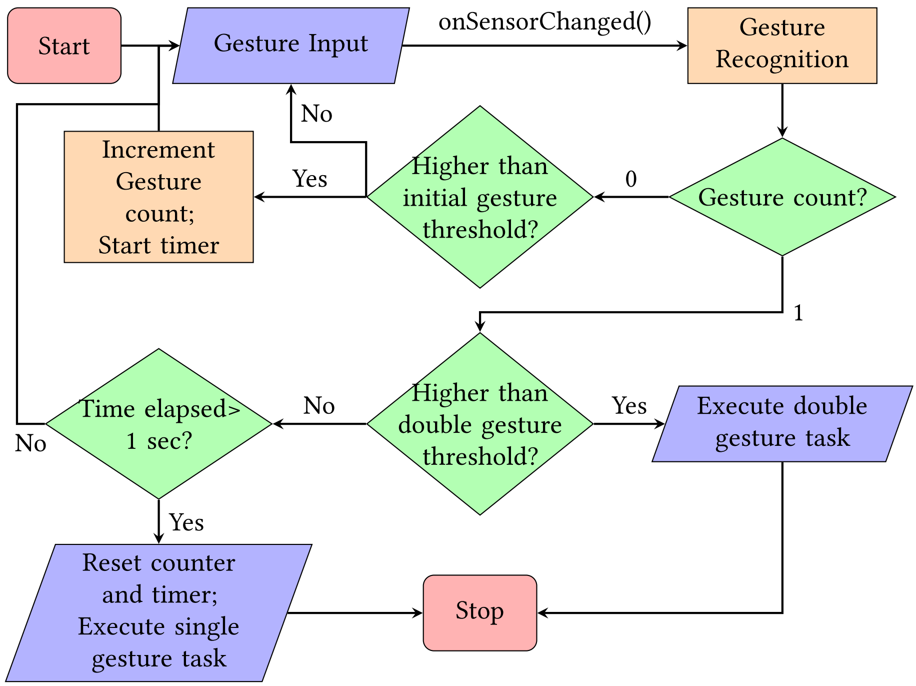


Fig. 2. Flowchart of Gesture Recognition Algorithm

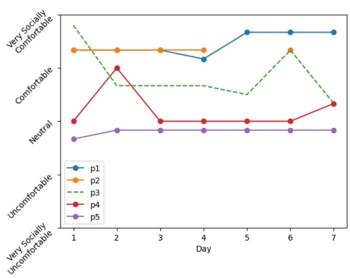
remembering the gesture mapping owing to repeated changes by the participant during the study period.

3 PILOT STUDY

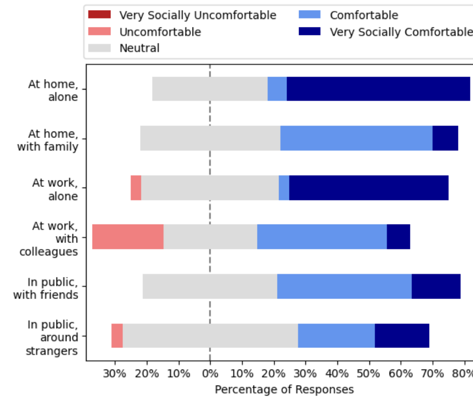
3.1 Experimental Design and Procedure

Conducting social acceptability studies in naturalistic settings is generally encouraged [7, 11]. However, owing to the lack of standardized questionnaires for evaluating social acceptability [7], designing an in-the-wild pilot study proved trickier than initially anticipated. Although not a direct measure, the use of locations and audiences introduced by Rico and Brewster [2009, 2010a] is one popular method to quantify social acceptability [7]. In this method, the participants' ratings of willingness to perform gestures at six locations or settings alone or in the company of five audience types were used to assess social acceptability. A 5-point Likert scale adaptation of this method used by Alallah et al. [2018a, 2018b] was adopted in the daily questionnaire. Further, a modified version of the questionnaire employed by Freeman et al. [2014] that consists of only six settings, namely home- alone, home- family, work- alone, work- colleagues, public- friends and public- strangers, was used to reduce the number of questions the participants were required to answer from 36 to six. Likert scale questions on system reliability, the difficulty of recalling the gesture mapping, and overall preference for the gesture-based system over touch-based interactions were also included. The survey also allowed for optional text-based entries for the participants to record any notable observations when the gesture-based interaction system did not work.

The study methodology was as follows: The participants were asked to select a gesture to application mapping for six smartphone applications of their preference. The gesture-based interaction system was set up on the participants' smartphones and configured to work with their customized



(a) Variation in ratings over time



(b) Variation in ratings in different settings

Fig. 3. Variation in social acceptability ratings

mapping during the initial session. The experimenter demonstrated the gestures on their personal smartphone and the participants were asked to try the same on their own devices. The participants were encouraged to use the gesture-based interaction whenever they used their smartphones but advised to disable it using the Accessibility button whenever necessary. Each participant was asked to complete the daily questionnaire described above in the evenings/ nights for a period of one week. After one week, the participants met with the experimenter for about half an hour for a final semi-structured interview to collect feedback and gain more insight into the participants' experiences.

4 RESULTS AND DISCUSSION

Five participants (p1- p5) participated in and completed the study. Four participants (p1- p4) were physically present for the initial and final sessions whereas the entire process was conducted remotely for p5. Although this study did not yield statistically significant results given the limited number of participants, the data gaps and the limitations of the participant demographic as detailed in Section 5 some general trends and observations were made.

4.1 Social Acceptability in different social settings and over time

The daily survey response distribution was analyzed over time to check whether the social acceptability varied over the seven-day period. The variation of the average of the ratings given by each participant is shown in Figure 3a. Although p1, p4 and p5 reported varying degrees of overall social acceptability of the system, the ratings had some variations but remained generally constant and displayed a small increase over the period of the study. The increase in acceptability over repeated use is consistent with the findings of a study by Rico and Brewster [2010b] which compared the user comfort reported over two study sessions. While p2 reported constant social acceptability on the survey, owing to the missing data on days 5 and 7, it seems presumptuous to draw conclusions about the trend exhibited by p2.

Unlike the other participants, large variations were observed in the case of p3. On further enquiry, this turned out to be a result of an oversight on my part during the setup stage rather than the participant's perception of social acceptability. The participant reported that they selected the "Neutral" option on the occasions a certain setting did not apply to them, such as days they did

Setting	Acceptability	Setting	Acceptability
Home, alone	81.43%	Home, family	73.33%
Work, alone	80.54%	Work, colleagues	66.43%
Public, friends	76.19%	Public, strangers	70.00%

Table 1. Acceptability in different social settings

not go in to work or have their family at home, rather than opting out of answering the question as was expected in such cases. It was unclear whether the participant also selected the "Neutral" option to express the social acceptability of certain situations. Therefore, the survey data from p3 was removed from further analysis and only their feedback during the final interview has been included.

The overall response distribution was also aggregated based on the question as shown in Figure 3b. Table 1 shows the mean acceptability score for each setting. These percentages were obtained by averaging the ratings provided by each participant over the seven days to obtain a single rating per participant in each setting and then computing the mean of the ratings across the participants. A value of 100% in Table 1 would indicate that all the participants considered the system "Very Socially Comfortable" in the corresponding setting. The ratings obtained indicate a perception of social acceptability of gesture interactions in private or familiar settings over public settings. However, despite the limitations noted in Section 5, the participants reported being more comfortable performing gestures among strangers in public than among colleagues unlike the observations made by Freeman et al. [2014].

The responses for perceived reliability and difficulty in recalling the gesture to application mapping were also analyzed both individually and in combination with social acceptability ratings for correlation with the reported preference for the gesture-based interaction system over touch. Among these comparisons, social acceptability was found to have the strongest correlation with system preference with a value of 0.56. The system reliability and gesture mapping were, however, found to impact gesture preferences and user sentiment as discussed in the subsequent subsections.

4.2 Rationale behind gesture preferences and perception of Social Acceptability

While the use of locations and audiences simplifies the process of obtaining quantifiable feedback on perceived social acceptability, the questionnaire does not provide insights about what factors contribute to an interaction being more or less preferable or acceptable [7]. Given that all the gestures tested with this system were very similar in that they involved subtle hand movements while holding a smartphone the user preferences were more specific to hand gestures rather than the general gesture recommendations made by Rico and Brewster [2010b]. The most frequently discussed rationales for gesture preferences were as follows:
1. Convenience with natural grip: The convenience of performing a gesture with the participants' natural grip of their phones was the most frequently recurring reason for preferring one gesture over another. However, since the participants' natural grip styles varied so did their gesture preference.

In three out of five cases, tilt was the most preferred gesture followed by twist and fold. The full list of ranking of gesture preferences has been presented in Table 2. The participants showed less preference for gestures which made the movement feel unnatural. In particular, participants rated gestures that required adjusting their grip as their least favourite. This was the rationale for fold being ranked as the least favourite gesture by all participants except p2 and p4.

Participant	Twist	Fold	Tilt
p1	1	3	2
p2	3	2	1
p3	2	3	1
p4	2	1	3
p5	2	3	1

Table 2. Ranking of Participants' Gesture Preferences

Further, it was noted that in two out of four cases when convenience of performing gestures with natural grip was noted as the major rationale for the gesture ranking, the participants stated that they had formed their gesture preferences shortly after they were demonstrated and made their mapping selections accordingly.

2. Similarity to common actions: Owing to the limitations in system reliability and consequent increase in false triggers when the gestures matched common movements the participant made with their smartphones, the similarity to common gestures negatively influenced a gesture's likability, contrary to the recommendations made by Rico and Brewster [2010b]. This was especially the case for p2 who described encountering false positives for the twist gesture and occasionally the fold gesture as the system mistook their regular interactions with their smartphone as a gesture. Consequently, their rationale for gesture preferences was almost entirely based on this factor. While other participants did not report facing much trouble with false positives, the situation was perhaps exacerbated in the case of p2 as unlike the other participants they opted to have the twist gesture threshold configured to trigger when the device was 'twisted' inwards rather than outwards.

3. Effort: Although this has some similarities to the first rationale, another noted factor for ranking the gesture preferences was perceived effort. The threshold for triggering was the same across the three gestures, however, the participants expressed a higher preference for gestures they perceived as requiring less effort to trigger. Also, participants generally liked single over double gestures. This sentiment is consistent with one of the identified motivations for developing eyes-free interaction systems [15].

Additionally, p4 also described feeling the need to exercise caution when executing gestures like the tilt. While describing their reasoning for preferring the fold gesture over tilt, they said "...it doesn't have the chance of me dropping this (it)" and also thought that they were more likely to secure their phone with the other hand if it slipped out of grip for the fold rather than the tilt.

4.3 Additional observations for further improvement

The participants also made some additional general observations which are likely to be relevant for the further development of the system. Three out of five participants explicitly stated that they did not use all of the six gestures on a regular basis as they used fewer than six apps regularly and that they would have preferred a system with fewer options. Additionally, some participants stated that including more customization on enabling or disabling individual gestures would have been valuable. Some participants also expressed a preference for gestures that would have worked more conveniently with their natural grip e.g. shaking the phone, but their views on what would be ideal additions varied.

Further, only two out of five participants perfectly recalled the gesture to application mapping. Both these participants said that they grouped the apps e.g. mapping messaging apps to tilt gestures or transit-related apps to fold gestures and so on.

4.4 System Adoption

Overall, despite reports of some inconveniences, all participants opted against uninstalling the gesture-based interaction application and chose to retain the application on their personal devices after the conclusion of the week-long study.

One participant opted for a modification to limit the number of gestures detected and change the mapping after the conclusion of the study in order to better suit their needs for continued use.

5 LIMITATIONS

A week-long in the wild seems like a long and short time all rolled into one. Some evenings/ nights the participants forgot to fill in the survey and missed reminders and on other occasions, they left certain questions unanswered because they didn't apply to them. To avoid too many data gaps due to the former scenario, I asked the participants to fill in the survey the following morning, hoping their perspectives would remain largely unchanged. This occurred once for p1, p2, p4 and p5 and twice for p3. Additionally, p2 forgot and later opted out of filling out the daily survey on days 5 and 7.

One more significant factor that likely impacted the study results was the participant demographic. All five participants were males with an engineering background. Four of the five participants also work in the same lab. This likely impacted their views, acceptance of the system, and social settings.

Further, the study evaluated the overall acceptability of the system rather than the individual gestures in order to limit the survey length. The user-customized gesture mapping could have also introduced an additional bias.

6 CONCLUSION AND FUTURE WORK

Through this work, I intended to evaluate a gesture-based smartphone interaction system's social acceptability and understand the effect of various factors on user adoption. While the study affirmed the findings from the literature and also resulted in some new findings, it was rife with challenges and limitations.

This study helped gauge the relative social acceptability of the gesture-based interaction system in various settings, provided some valuable insights into factors that influenced gesture preferences and helped me identify some mistakes to avoid in the future. The general positive trend in social acceptability ratings among three out of five participants after a week and the inclination of the participants to continue using the gesture-based interaction system provide some motivation to carry this work forward.

Improvements like implementing other gesture recognition methods to improve system reliability and limit its impact on gesture preference, recording system usage to better understand the social acceptability change over time and trying other gesture variations and customizations could potentially add to the value of the system and yield richer insights.

7 ACKNOWLEDGEMENTS

I thank Cyan Kuo for her valuable insight and pointers on where to look for designing the user study. I also thank the study participants for their time and patience in testing a far-from-perfect prototype.

REFERENCES

[1] [n. d.]. *"Developer guides | Android Developers"*. Retrieved July 23, 2023 from https://developer.android.com/guide/index.html

[2] Fouad Alallah, Ali Neshati, Yumiko Sakamoto, Khalad Hasan, Edward Lank, Andrea Bunt, and Pourang Irani. 2018. Performer vs. Observer: Whose Comfort Level Should We Consider When Examining the Social Acceptability of Input Modalities for Head-Worn Display?. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 10, 9 pages. https://doi.org/10.1145/3281505.3281541

[3] Fouad Alallah, Ali Neshati, Nima Sheibani, Yumiko Sakamoto, Andrea Bunt, Pourang Irani, and Khalad Hasan. 2018. Crowdsourcing vs Laboratory-Style Social Acceptability Studies? Examining the Social Acceptability of Spatial User Interactions for Head-Worn Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3173574.3173884

[4] Christina Dicke, Katrin Wolf, and Yaroslav Tal. 2010. Foogue: Eyes-Free Interaction for Smartphones. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (Lisbon, Portugal) *(MobileHCI '10)*. Association for Computing Machinery, New York, NY, USA, 455–458. https://doi.org/10.1145/1851600.1851705

[5] Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2014. Towards Usable and Acceptable Above-Device Interactions. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services* (Toronto, ON, Canada) *(MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 459–464. https://doi.org/10.1145/2628363.2634215

[6] Hari Prabhat Gupta, Haresh S. Chudgar, Siddhartha Mukherjee, Tanima Dutta, and Kulwant Sharma. 2016. A Continuous Hand Gestures Recognition Technique for Human-Machine Interaction Using Accelerometer and Gyroscope Sensors. *IEEE Sensors Journal* 16, 16 (2016), 6425–6432. https://doi.org/10.1109/JSEN.2016.2581023

[7] Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. 2020. Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3313831.3376162

[8] Andy Kong, Karan Ahuja, Mayank Goel, and Chris Harrison. 2021. EyeMU Interactions: Gaze + IMU Gestures on Mobile Devices. In *Proceedings of the 2021 International Conference on Multimodal Interaction* (Montréal, QC, Canada) *(ICMI '21)*. Association for Computing Machinery, New York, NY, USA, 577–585. https://doi.org/10.1145/3462244.3479938

[9] Julie Rico and Stephen Brewster. 2009. Gestures All around Us: User Differences in Social Acceptability Perceptions of Gesture Based Interfaces. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Bonn, Germany) *(MobileHCI '09)*. Association for Computing Machinery, New York, NY, USA, Article 64, 2 pages. https://doi.org/10.1145/1613858.1613936

[10] Julie Rico and Stephen Brewster. 2010. Gesture and Voice Prototyping for Early Evaluations of Social Acceptability in Multimodal Interfaces. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction* (Beijing, China) *(ICMI-MLMI '10)*. Association for Computing Machinery, New York, NY, USA, Article 16, 9 pages. https://doi.org/10.1145/1891903.1891925

[11] Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 887–896. https://doi.org/10.1145/1753326.1753458

[12] Farhan Sufyan, Subhash Sagar, Zubair Ashraf, Shoaib Nayel, Mohd Sameen Chishti, and Amit Banerjee. 2023. A Novel and Lightweight Real-Time Continuous Motion Gesture Recognition Algorithm for Smartphones. *IEEE Access* 11 (2023), 42725–42737. https://doi.org/10.1109/ACCESS.2023.3255402

[13] S. Shyam Sundar, Eugene Cho, and Jinping Wang. 2018. *Interacting with Mobile Media*. John Wiley Sons, Ltd, Chapter 27, 615–639. https://doi.org/10.1002/9781118976005.ch27 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118976005.ch27

[14] Xian Wang, Paula Tarrío, Eduardo Metola, Ana M. Bernardos, and José R. Casar. 2012. Gesture Recognition Using Mobile Phone's Inertial Sensors. In *Distributed Computing and Artificial Intelligence - 9th International Conference, DCAI 2012, Salamanca, Spain, 28-30th March, 2012 (Advances in Intelligent and Soft Computing, Vol. 151)*, Sigeru Omatu, Juan F. De Paz Santana, Sara Rodríguez-González, José M. Molina, Ana M. Bernardos, and Juan M. Corchado Rodríguez (Eds.). Springer, 173–184. https://doi.org/10.1007/978-3-642-28765-7_21

[15] Bo Yi, Xiang Cao, Morten Fjeld, and Shengdong Zhao. 2012. Exploring User Motivations for Eyes-Free Interaction on Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2789–2792. https://doi.org/10.1145/2207676.2208678

[16] Shenglin Zhao, Haoyuan Cai, Wenkuan Li, Yaqian Liu, and Chunxiu Liu. 2021. Hand Gesture Recognition on a Resource-Limited Interactive Wristband. *Sensors* 21, 17 (Aug 2021), 5713. https://doi.org/10.3390/s21175713