Opportunities with Hand Gesture Technology in Mobile Gaming

Rasmus Klintsø Christensen
rakc@itu.dk
Student No.: 15829
BSc Software Development

Advisors: Fabricio Narcizo and Jørgen Staunstrup
Submitted: August 2021

IT UNIVERSITY OF COPENHAGEN

Abstract

During the past decade, there has been steady developments in the area of computer-vision based hand-gesture recognition (HGR) technologies, and expansion in the environments they are available for. Hand-gesture input, combined with head-mounted displays, has become the principal interaction method in virtual reality games. It also shows promise in other areas, such as sign-language recognition, interactive museum exhibitions, and interactive displays available in public spaces. This paper explores the possible introduction of HGR-based interaction in mobile games, based on the identification of key concepts in literature examining the aforementioned areas. The result is a proposition of four general heuristics guiding the design and development of mobile games based on HGR as the primary interaction method.

Contents

Chapter 1

Introduction

1.1 Context & Motivation

There is a lot of diversity in how humans interact with computer technology. Different devices and applications have varying degrees of freedom depending on the role and purpose of use. The first cell phones were designed with layout of buttons making up a basic numeric keypad in addition to a few basic actions such as *confirm* and *return*. This limited games for these devices to have their input actions mapped to the keypad buttons. This changed with the introduction and market success of smartphones, establishing the touchscreen as the most popular canvas for designing mobile games [1]. The touchscreen has dominated this last decade, and will probably continue to do so, since it provides developers with significant freedom in input design and visual artistic expression. Yet it begs the question of how mobile games might evolve to include technologies, which are based on newer developments in computer-vision and body-motion sensors, as seen with the recent success of virtual reality devices [2].

Mobile Games is the sector in the video game industry with the largest market share in 2021 [3]. It has grown naturally as a consequence of the global normalization of smartphone usage [1]. Since the requirement for playing these games is a device already in many people's possession, the games developed for mobile has the lowest barrier-of-entry, compared to other platforms.

Video games designed for mobile platforms are often characterized by convenience and short-duration game play loops. However, there exists a huge diversity in genres and the user segments they are targeted for. Furthermore, the prevailing trends and popularity of certain games are ar-

guably correlated to geographic and cultural differences. As an example, one of the largest user segments consists of Chinese commuters providing a demand for mobile entertainment as a consequence of their daily travel time [4].

One appealing aspect of implementing computer-vision hand gesturing in mobile gaming – as opposed e.g. VR-gaming – is that it would eliminate the requirement of any other equipment than what the mobile device itself has to offer.

1.2 Directions of the study

Going from an initial focus on reviewing specific frameworks and techniques on machine learning, computer-vision, and hand-gesture recognition technology, this study changed direction toward research in academic literature, with findings more related to the design and usability aspects of hand gesture recognition technology. Suggestions and insights from included studies, directed this study towards investigating and presenting concepts, which are deemed to reinforce the positives of existing utilization of hand gesture recognition based user interaction. With inspiration from especially the successful designs seen in VR-gaming, but also the additional three domains, the final contribution of this study became directed towards the resulting heuristics, intended to guide the purpose and design of future pioneering mobile games, so that they might find success in basing their interaction method on the exciting new opportunities of hand gesture recognition.

1.3 Problem Statement & Research Questions

During the past decade, there has been steady developments in the area of computer-vision based hand-gesture recognition technologies, and expansion in the environments they are available for. This study aims to explore usage of the technology, investigate why it seems to be under-utilized in commercially available applications – specifically in mobile games – and propose how related issues might be solved.

RQ1: By investigating previous and current utilization of hand gesture recognition technology, which key concepts can be identified that are applicable for developing future mobile games?

RQ2: Can the aforementioned key concepts be formulated into general heuristics – with the added purpose of motivating further research in hand gesture recognition utilization?

1.4 Objectives

- Introduce the usage of hand gesture recognition technology in selected areas – structured as distinct domains.
- Identify key concepts from these domains, which are relevant to, and can be applicable in developing future mobile games.
- Propose general heuristics for creating pioneering mobile games, based on hand gesture recognition as interaction method.
- Discuss the limited scope of this study, and suggest directions to continue the work in future research.

Chapter 2

Literature Review

2.1 Developments in Mobile Gaming

Design of mobile games has been influenced by the features offered by the average smartphone. More specifically, the sensors and input methods supported. The primary ones are the touch screen and the accelerometer [1].

An empirical study from 2011 involved users testing the same game with three different interfaces for interaction, followed by them providing comparable feedback for each of them[1]. In the study, Browne and Anand[1] described them as the novel mobile interfaces, touch screens and accelerometers. The study was motivated by the *"...phenomenal surge in market penetration."*[1] of smartphones happening at the time, and correctly predicted the following popularity of mobile games. The tests were based on what they describe as scroll-shooter game, with the three interfaces in question being *Touch gesture* (swiping on display), *Accelerometer* (using device orientation), and *Simulated button* (pressing buttons implemented visually in the game).

The result and main contribution of the study is five heuristics for designing mobile game interfaces. Browne and Anand[1] propose;
(H1) An accelerometer-based user interface should be available.
(H2) Multiple user interfaces should be available.
(H3) Touch gestures should be utilized when diagonal direction input is either required from or desired by the user.
(H4) Interface sensitivity should be configurable.
(H5) Physical properties of gestures should be directly translated into virtual properties. [1]

Of those, especially (H5) is interesting in the context of this study. It is important to note, that Browne and Anand's[1] notion of *gestures* covers the user's finger-touchscreen interaction. Elsewhere in this paper, the no-

tion of *hand gestures* as an interaction method is based on computer-vision techniques. Despite this difference, the general principles directing their research still applies to mobile games in 2021, and are therefore relevant to this study.

2.2 Domains of Hand Gesture Application

This study investigates four different domains of successful application of hand gesture recognition technology as an interaction method. Based on the concepts and characteristics found in these four domains, it will explore possibilities in mobile gaming, and from the knowledge attained, propose what key concepts must be applied in order to see the technology better utilized in this area - thus answering the research question.

2.2.1 Domain A: Virtual Reality Games

Virtual Reality technology has seen significant improvements during the last decade. It has developed from being a niche technology, only available in dedicated arcades or amusement parks, to something more comparable with (and supported by) the average home entertainment system. Companies compete on accessibility in terms of supported platforms and pricing, while continuously refining the technology and improving fidelity in the gaming experience offered. [5]

The domain of VR-gaming is relevant to this study, since it is characterized by having the player's entire body interacting with the virtual environment. This immersive experience is upheld by the inclusion of hand- and arm-gesturing as one of the primary methods of interacting with the games [5]. So instead of pressing buttons on a keyboard or controller, it is the player's body itself that acts as the primary interface. However these player interactions are being supported by a variety of advanced sensors, and thus requires a setup of external devices demanding a dedicated *play zone*, and increases the barrier-of-entry for potential customers.

Virtual Reality entering mainstream culture is exemplified by the 2018 movie from Steven Spielberg; Ready Player One. The movie helps fuel the introduction of the technology to a wider audience. A large part of the movie takes place in a virtual world, which the main characters interact with using advanced VR-technology. This presents an idea of what could be possible in the near-future. This blockbuster had an estimated budget of

$175,000,000 and grossed $582,893,671 globally [6], indicating its popularity and position in mainstream entertainment.

Beyond the general media exposure, exemplified by Spielberg's movie, the VR and VR-gaming market is currently seeing a stable and steadily growing user-base. In 2021 there are globally more than 171 million users of virtual reality [2]. Demand for standalone VR devices is expected to grow over 16 times between 2018 and 2022 [2]. 70% of VR headset-owning consumers have bought a game on it. [2]. Global VR-gaming revenues reached $ 22.9 billion in 2020 [2].

Since VR-technology is becoming more widespread, it has simultaneously created a larger market for software (games), which supports a growing supply and demand that further enables game studios to produce gaming experiences which are exclusively designed for virtual reality. The 2019 VR title *Beat Saber* [7] is a good example of such a game experience, winning several awards, including *The Game Awards*: Best VR/AR Game 2019 [7]. This is an important milestone that enables true innovation in interactive entertainment, and has arguably been the catalyst of a new paradigm in consumer electronics. Building on the initial wave of success, Valve Software released their long awaited next installment in the Half-Life franchise, *Half-Life: Alyx*, in March 2020 [8]. The game is exclusively available for PC with the VR devices supported by the *SteamVR* platform, including Valve's own *Valve Index* that released the previous year. The game is designed to take advantage of the hand gesture interface enabled by the Index' advanced controllers. Features include enabling the player to use the *gravity glove* to manipulate objects in the virtual environment. The glove tracks an accurate 1-to-1 match of the gesture and hand position the player makes with the controller. Thus it follows Heuristic 5 proposed by Browne and Anand [1]; "Physical properties of gestures should be directly translated into virtual properties.". The success of *Half-Life: Alyx* is confirmed by widespread critical acclaim and by following in footsteps of *Beat Saber* by winning *The Game Awards*: Best VR/AR Game 2020 [9].

It is interesting to find which key concepts are responsible for the VR success, and whether they can be applied to create comparable developments in how players interact with Mobile games. VR-games has proven, that hand gesturing can be an intuitive and user-friendly method of interacting with a virtual environment. Now the question is whether it can be down-scaled to work in e.g. Augmented Reality gaming experiences, based on mobile platforms – which would be more accessible to the wider public, and further

introduce more people to this potential new paradigm in human-computer interaction. However, there are clear differences to consider, among which is the dependence on external devices and sensors.

The real appeal in using computer-vision hand gesturing in mobile gaming, is that it would eliminate the requirement of any other equipment than what the mobile device itself has to offer.

2.2.2 Domain B: Health Technology

In a study from 2019, Yasen and Jusoh[10] concludes that the most popular application of hand gesture recognition techniques is for sign language.

Handicaps such as hearing-impairment has been a motivational factor for technological development for many years. Zafrulla Et al. [11] investigates the potential of the consumer oriented Microsoft Kinect, as the basis for developing sign language recognition in educational games for deaf children. The study revolved around the *CopyCat project*, which is described as an adventure game targeted for deaf children to help them acquire language skills - with the added purpose of advancing the research in computer-based sign language recognition. The main findings can be summarized with the Kinect being comparable to existing solutions in performance, while mitigating the need for the for specialized equipment.

In regards to computer-vision hand gesture recognition technology on mobile devices, some of the best examples found is related to sign-language recognition. Such applications demand high precision in the interpretation of gestures, since small differences can result in entirely different meanings. Jin Et al. concludes that visually distinctive gestures tend to have higher accuracy [12]. In their pioneering study from 2016, they were able to produce a framework recognizing 16 different American Sign Language gestures with an average accuracy measured to 97.13 % [12].

Successful applications with sign-language recognition, suggests that the technology is mature, and could be the basis of new innovations in game design for mobile devices. These developments in sign-language related technology is the primary inspiration for including this domain in the context of this study.

2.2.3 Domain C: Public Entertainment Systems

The domain of *Public Entertainment Systems* is characterized by cultural leisure or educational experiences, which are enhanced by gesture recognition technology in order to make them more interactive – and thus attractive to a wider audience. Some of the prime examples are seen in museum exhibitions [13]. Museums has been a driver of these innovations, since this is an area which has worked to enhance the historically passive nature of the experience offered. From this, a general theme has emerged; museums want to be fun and interactive, and thus more attractive to especially younger students [14].

Agate and Gaglio [15] present a framework for gesture recognition targeted for museums. The aim is to provide exhibition attendees with a *non-intrusive interface* for interacting with multimedia and digital versions of museum contents. The main idea is to add virtual editions of the museum's artifacts to the physical collection, which allows the visitors to interact with these through a motion sensor interface, using body movement and hand gestures.

Malerczyk [16] describes a mixed reality museum exhibit, which provides an interactive experience of the museum's artwork. These artworks are in the form of paintings. The sensitive nature of such artworks does not allow for physical contact. By including an interactive display in the exhibition, the museum allows visitors to select pieces from their collection and manipulate them on a screen using simple and intuitive hand gestures. As an example, valuable artworks often requires a significant distance from the observers, in order to protect them. However this exhibit allows them to magnify areas and examine specific details, which would be inaccessible otherwise. The study suggests that adding an element of interaction with virtual exhibits can directly increase the level of interest from the museum visitor. Furthermore, such increased interest will reinforce the quality of the educational aspect in an exhibition, through these hands-on experiences. The intended effect of such interactivity is reinforced by Jiwane and Khan [14]. The example from Malerczyk's study is an early example, and thus limited in practicality by the technical setup. However it exemplifies the value provided by adding interactivity to exhibits, even with the technological limitations present at the time.

The relation of this domain to that of mobile games, rests in the potential of including use of the attendees' smartphone as the platform for the implementation of the aforementioned beneficial interactivity.

2.2.4 Domain D: Public Service Displays

The domain of *Public Service Displays* is characterized by technology applications in public spaces, which are designed for providing self-service features hosted by devices such as interactive displays. These include publicly available screens in e.g. shopping malls, that allows visitors to interact with them in order to find specific stores or products. Self-service devices for shopping are also relevant; such as the devices customers use to order food, which have become a common sight at many fast-food restaurants. In general, it is regarding public or semi-public accessible self-service applications, which already feature, or could feature, hand gesturing as a method of user interaction. The purpose is to identify traits, which are general to these applications, and could be relevant to the issues on mobile gaming in public spaces.

Müller et al.[17] investigates the development and rising popularity of interactive displays in public spaces. They describe how the domain of multimedia has grown beyond use in personal electronics and has become a common sight in urban areas. However, the study is motivated by observations that the interactive features of such devices are underutilized and often ignored by the targeted users. They introduce the issue of lower-than-expected usage of deployed solutions, despite that such *interactive networked displays* are regarded as a promising method of deploying multimedia applications and content [17]. Furthermore, many displays seemingly fail to attract the expected attention from passers-by, and instead vanishing in all the visual noise and clutter of other impressions in the public space that compete for attention. Müller et al.[17] also found, that when they do manage to capture the attention of these passers-by, they seemingly fail to actually motivate them to interact with them. The study analyze the attention of potential users in public spaces, with explanations of the relevant mental models and a variety of interaction modalities. The main contribution of the study is a proposed taxonomy formulated as a comprehensive guide for the designers and developers of interactive multimedia on public displays.

In relation to the context of mobile gaming, some of Müller et al.'s[17] most interesting findings involve the behavior of potential users in public spaces. People show inhibitions resulting from the self-awareness prevalent when a potential user is being watched by others. Put very simply; people are afraid of looking silly in front of strangers. This common fear needs to be balanced out by a design that invokes enough interest and curiosity to motivate people to try something new [17]. This is relevant to mobile games, since these are often played in public spaces, which makes them affected by the same inhibitions. Mobile games based on hand gestures would attract

attention to the player, which might be unwanted, and could compel them to choose a game that allows them to be more discrete. This issue should therefore be considered going forward.

Chapter 3

Proposed Solution

Based on the insights and knowledge acquired from reviewing the four selected domains in Chapter 2, this chapter presents key concepts that should be considered for future developments in hand-gesture based mobile games. The derived result is a list of proposed heuristics summarizing the most useful findings. The list serves to guide the direction of pioneering work in creating what could become a new interaction paradigm on the mobile platform. For clarification, the heuristics following from this chapter are presented with similar notation style to the ones from Browne and Anand [1] included in Section 2.1. They should not be confused with each other.

3.1 Contribution: How to Adapt Hand Gestures for Mobile Games

This section presents the aforementioned key concepts from in the domains of this study. Some of the following paragraphs have headers denoted by **(H1)**, **(H2)**, etc., indicating that they contain the primary reasoning that support the corresponding heuristic, as they are presented in Section 3.2.

3.1.1 Key Concept from Domain A: Virtual Reality Games

The establishment of a consumer market in virtual reality gaming is based on two important aspects. The first and more obvious one revolves around the basis in the technology, which has reached the level of refinement and cost-efficiency required for the development of consumer-oriented electronics. Without having the electronic devices accessible to a large consumer-base, there is no purpose in developing the games or other software which runs on them. The second aspect revolves around the growth in *gamer culture* and its corresponding mass market, which is further supported by the

high level of media exposure that has been normalized during the 2010s.

(H1) Game Interface as a Feature

The first key point from Domain A, is that future hand gesture based mobile games should attempt to imitate the things that work well in the current generation of VR-gaming. The concepts that are transferable and viable with the mobile platform, mainly rests in how a game experience is designed specifically around how the user interacts with the technology. Beat Saber is a good example of a game experience that takes full advantage of the body engagement, which is enabled by a complete virtual reality setup [7]. The core game play loop is both short and simple, which makes it approachable for beginners. Building on an easy introduction, it is possible for experienced players to tune the difficulty to create a constant challenge, and therefore engaging game play with new goals to work towards. Thus Beat Saber creates a fun learning curve, which simultaneously makes the player more accustomed to using the VR technology. Arguably, the game acts as training ground for the player wanting to become an advanced user of VR-technology in general. What is learned from this, is that the pioneers designing hand gesture mobile games, should base the core game play elements on introducing the player – and training them in – how the interaction works. In short; turn the game interface into a core feature of the game itself.

(H2) Emphasize Accessibility

A full body immersive experience, as offered by virtual reality devices, is not suited for the mobile platform. This should not be the goal when attempting to create hand gesture mobile games. Instead, one must identify a limited scope of user interaction, and refine that into a clearly defined core, which the game interface can be based on. This could be 4-5 specific gestures, which are easy for the player to memorize, and, more importantly, allowing the recognition technology to consistently interpret them with a low degree of error. The point of limiting this interface scope, is that the level of immersive interaction will not be able to compete with what is offered by a dedicated platform. Rather the strength of hand gesture mobile games will be, that the barrier-of-entry to these will be negligible compared to the requirements for VR games, assuming smartphone ownership. The aim should be to allow every potential user with a modern smartphone to experience a new paradigm in game interaction. Thus, the second key point from Domain A, is that hand gesture mobile games should aim to offer the

player a lite-version of the intuitive body-engagement, that we have seen work well in VR. – A lite-version which is accessible to significantly larger amount of users.

In regards to the possible user-base, it is worth considering the global number of active mobile gamers in 2021, which is over 2.2 billion [3] according to Dobrilova in a Tech Jury blog from March in the same year. Furthermore, when looking into the 2020 figures for USA alone, mobile games revenue reached 10.73 billion USD [18], with the number of mobile game users reaching 147.13 million in 2020, according to a Statista report from June 2021 [18].

3.1.2 Key Concept from Domain B: Health Technology

Investigating the domain including sign language recognition in smartphones, is primarily interesting to this study based on the argument, that if such solutions can work, the supporting technology could potentially serve as the basis for HGR mobile games. Beyond the earlier described findings from Jin et al. [12], presenting a framework recognizing 16 ASL gestures, this project has not answered the question of how the underlying technology can be applied for mobile games. It remains subject for discussion and further research.

(H2+) Emphasize Accessibility

The findings from investigating the domain of health technology relates to the accessibility aspect of disabilities (e.g. hearing impairment), and the correlating issue of possessing technology to mitigate them. Even though technology assisting with disabilities is another subject than that of entertainment, the two are not mutually exclusive, as shown by Zafrulla et al.[11] – especially in the context of educating children. Zafrulla et al.[11] investigated the potential of the Microsoft Kinect platform, yet it would seemingly be even more accessible to base such disability mitigating educational games on the mobile platform, due to the greater prevalence of smartphones. Thus the aforementioned heuristic of emphasizing accessibility is further reinforced by the findings from the domain of health technology.

3.1.3 Key Concept from Domain C: Public Entertainment Systems

Museums has greatly improved the experience they offer their guests, by hosting technology that allows for interaction with an exhibition [13] [14]. There has been experimented with *mixed reality* and use of hand gesturing to manipulate digitized artifacts. Making an exhibition a more active participatory experience can directly increase the level of interest from the guests, which further improves the quality of the educational aspect [13] [14].

The lessons from this domain does not point directly towards concepts suited for deriving general heuristics, which match the structure of this study's contribution. Yet it provokes ideas in implementing interactivity in museum exhibitions and similar experiences. One might imagine exhibitions that enable mobile applications based on a combination of localization and augmented reality features. This could facilitate digital treasure hunts and other experiences of that nature. But in the context of this study, the connection to establishing hand-gesture recognition *mobile games* becomes too speculative, and investigating this direction further is subject for further research.

3.1.4 Key Concept from Domain D: Public Service Displays

This section present key concepts that are essential mobile games based on hand gestures or other body movement, if they are designed to be engaged with in public spaces. These include areas such as parks, city squares, or in the seat of train.

The important takeaways from Domain D revolves around human behavior in public spaces. More specifically, how we behave when interacting with technology, or if we even choose to do so. These are important to consider, since mobile games are often played when the user is mobile. A typical case is playing while using public transport. Furthermore, in the case of the popular *Pokémon Go* [19], it is required of the player to move around in the real world, with parks, monuments, and squares, etc., acting as important landmarks inside the game as well.

Müller et Al. [17] suggests that a public display can be perceived as a stage, which can make shy people avoid interacting with them. They further describe how such displays must capture the attention of the people who are potentially interested in their contents, while not annoying others. The most relevant points from their study relates to the self-awareness most

people express in public. Potential users of technology in public spaces are inhibited by cultural etiquette and the simple fact, that humans do not like to look silly in front of strangers. This means that comparable technology would benefit from a design that finds a balance between two key aspects:

(H3) Enable Composure in Public Spaces

A given mobile game – or other application using hand gestures and other body movements for interaction – should facilitate a method of interaction, that allows the user to maintain their composure in a way that enables them to avoid unwanted attention. This requirement could be reached by; **1)** including an alternative interaction method that does not involve significant body movements, or **2)** having the gesture-recognition technology be sufficiently sensitive and precise to interpret understated movements.

(H4) Invoke Curiosity

The features and game-play experience, offered by a gesture-based mobile game, should invoke curiosity that motivates the user to engage with a method of interaction, that may be outside their comfort-zone. An obvious entailment of this, is that it requires a high quality product. But this is arguably the case for any game or software that attempts to demand the attention of potential new users. Another justification of requiring hand-gestures from the user, can be to make the gesture-based interaction a game-play feature in itself (which is an idea also resulting from domain A). For gaming, mastering a new and challenging interface can be the entire purpose of a game. The entertainment can be based on the user trying to make the device interpret their input correctly (e.g. accelerometer used to control driving games, or games based on balancing objects with device orientation). Furthermore, the implementation of hand-gesture interaction can itself be perceived as a major feature, that creates interest and attention for a new game. In this case, some generic game-play design can be made more interesting by simply interacting with it in an innovative new way.

3.2 Heuristics

This section presents four proposed heuristics guiding future development of hand gesture mobile games. They are derived from the key concepts described in Chapter 3. Each heuristic is a compiled version of the respective findings and reasoning. They should be considered with support from the referred-to sections.

3.2.1 H1: Game Interface as a Feature

Pioneering games based on hand-gesture interaction will benefit from designing the interface itself as a core game play element, and thus serve as a training ground for new players to become accustomed with a new method of interaction [3.1.1 + 3.1.4].

3.2.2 H2: Emphasize Accessibility

The aim of creating hand-gesture based games for the mobile platform, should be to allow all potential users with a modern smartphone to experience this new paradigm in game interaction [3.1.1] [3.1.2].

3.2.3 H3: Enable Composure in Public Spaces

If a pioneering title of hand-gesture based mobile games is meant to be played in public spaces, it should facilitate a method of interaction, that allows the user to maintain their composure in order to avoid unwanted attention [3.1.4].

3.2.4 H4: Invoke Curiosity

Pioneering games based on hand-gesture interaction should promote features to invoke curiosity that motivates the user to engage with a method of interaction, that may be outside their comfort-zone [3.1.4].

Chapter 4

Discussion

4.1 Summary

The first research question of this paper [1.3, RQ1], directed the effort towards identification of key concepts deemed to be of value in future hand-gesture based mobile games. First of these was the lesson from VR-games, with Beat Saber as prime example, that a new interaction method is well introduced by having a game focusing on the interface as a core game play element. The concept was further reinforced by the reasoning from Domain D, suggesting that mastering of a new and challenging interface, can be the entire purpose of a game. The general idea is to base the first generation of hand-gesture games on designs, that will smooth out the learning curve for new players, and become the foundation of a new genre. Another key concept identified, was related to the issue of accessibility, and how the mobile platform has the potential to reach more people than what was previously possible. Previous instances of motion-sensing related interaction methods depends on devices dedicated to the purpose. Such limitations of mass-marketing potential presents opportunities for mobile-based alternatives. The findings from Domain C relates to the benefits of adding an element of interactivity to museum exhibitions. It did not directly result in heuristics for mobile games, however it does suggest the potential of utilizing the mobile platform as the basis for such interaction. Key concepts learned from Domain D first relates to human behavior in public spaces, and how it affects our use of devices in this arena. Natural self-awareness can result in inhibitions, which presents an essential challenge in interaction design involving body movement. The design should therefore allow users to maintain composure. From this issue follows the second key concept in Domain D, which suggests mitigating the first issue further by offering an experience which is sufficiently appealing for overcoming such inhibitions

and preference for remaining in our comfort zone.

4.1.1 Applying the result

As directed by the second research question [1.3, RQ2], the key concepts summarized above have been formulated into the four proposed heuristics presented in Section 3.2. They aim is to provide guidelines for shaping design visions and carve out a position in mobile gaming industry, that can support the initial introduction, and sustain the resulting genre on the longer term. The heuristics do not offer concrete solutions, since those would be characterized by a diversity resulting from many separate (and competing) creative processes towards distinctive game experiences. Furthermore, in order to direct the developments toward more specific solutions, any such heuristic would benefit from the support of more specific and dedicated research, and based on experience that supposedly does not exist yet. Instead the heuristics proposed in Section 3.2 are meant to inspire visions of pioneering games that might achieve success and define a genre of hand-gesture recognition based mobile games.

H1: Game Interface as a Feature does not need much further elaboration. The main point is to center game mechanics on the interface of hand-gesture interactions, with aforementioned benefits [3.1.1].
H2: Emphasize Accessibility is fulfilled by keeping the focus of design and development on the broad availability facilitated by the mobile platform. This should also be motivating the decision of basing developments on this platform. Exposing the largest possible amount of people to hand-gesture interaction help capture a stable user base for the genre, and ensure future interest and demand.
H3: Enable Composure in Public Spaces can become a greater challenge, but may also prove to be the least important, given that the others are achieved. As argued in Section 3.1.4, it may be fulfilled by implementing multi-modality, albeit with the risk of defeating the purpose of hand-gesture interaction all-together. Otherwise, it may be solved by the ability to recognize and interpret understated gestures, though it would require significantly refined techniques.
H4: Invoke Curiosity based on the reasoning in Section 3.1.4 provokes the statement: *make a good game*, which of course is no simple effort. In this context, it is a matter of creating game play mechanics that users will experience as naturally befitting use of hand-gestures. The designer might identify existing mechanics which are less intuitive with touchscreen, and use them to present this new interaction method. There may also be po-

tential in combining hand-gesture interactions with other computer-vision based features already present on the mobile platform, such as augmented reality. This combination suggests possibilities with 3D virtual environment and object manipulation, and might even inspire a plethora of use-cases beyond games. The challenge is to define a clear scope and core game play mechanics within this context of novelty, and translate them into an intuitive and entertaining experience.

4.2 Further Research

The beginning phases of this bachelor thesis was originally directed towards an empirical user-study and evaluation of a specific machine learning framework, namely MediaPipe. After deciding that this is not a suitable platform for the intended direction and limited scope of this bachelor thesis, the overall project was reformulated as a literature review. The purpose of this review was to investigate and evaluate the developments of the technology area as a whole, and propose whether the available technology has reached a level sufficient to become the foundation for a new generation of mobile games, based as hand-gesture recognition. However, attempts in reviewing the specific frameworks and techniques on machine learning, computer-vision, and hand-gesture recognition technology, which are documented in academia, did not turn out fruitful. This diminished the scope and potential contribution of this project, and begs the question of which directions that would motivate future work. This section provides such suggestions.

4.2.1 Specific Gestures

Further research is encouraged, to investigate which specific hand- and finger-gestures that are viable for creating new standards in interaction design. It may be beneficial to identify which gestures feel natural and intuitive in use, and have them be applicable across a variety of use-cases. One can imagine a future scenario with an established set of actions, being represented across a multitude of games, which reinforces an effect of such a title being easy to learn by applying the experience gained elsewhere. This is currently exemplified with the established controller-input layouts of current generation games for PC and console platforms. Very few PC-gamers need to be taught, that they must use the Space-key for jumping, or the WASD-keys for general movement. Likewise with console controllers' analog stick, and with the A- and X-button on the Xbox and PlayStation, respectively.

4.2.2 The issue of Domain B

The decision to include the subject of sign-language related technology, and use it to define *Domain B: Health Technology*, was based on the assumption that an obvious link to the subject of this study could be made. Sign-language is obviously related to hand-gestures as a general concept, which makes it an appealing entry-point of research and investigation. However, the initial intention was to present the specific technology that supports proven solutions in sign-language recognition, and use that knowledge to suggest a method of utilization in mobile games. This path of research produced lacking and mostly suggestive findings, and was thus inconclusive in the original intention. In retrospect, this intention turned out to be too hypothetical, and maybe even naive in the scope of this project. Yet it is arguably an interesting point of research, and may still invoke curiosity and speculation on how these technology areas might develop in parallel and mutual advancement. However, certainty in this regard calls for further research with a more specific and technical direction than this project.

4.2.3 Limitations on Mobile

The are some limiting factors inherent to the mobile platform related to the functionality and usability supported by it. These deserve revisiting in future work, since they are not sufficiently considered in the review of the domains in Chapter 2 or the resulting contribution of this paper. Examples of such limiting factors include the finite battery life of mobile device, the requirements for specific camera types relating to lighting conditions, and the implications of latency for internet-dependent games. Reviewing these possible issues will be beneficial in supporting the heuristics contributed by this project.

One Handed Interaction

One of the primary merits of introducing gesture-based interaction to the mobile platform, is the increased accessibility inherent to the widespread availability and use of smartphones [3.2.2]. In order to maintain this merit, introducing new gesture-based features should avoid dependence on external equipment. Consideration of this limitation is one of the things making *vision-based* gesture recognition appealing. Body-movement interactions supported only by the device camera as the only hardware requirement is a strong reinforcement of the availability aspect of the mobile platform. However, this means that the player will only have one hand free for the HGR-based interactions, which may create major limitations in the design of such games. Evaluating the significance of this issue calls for further re-

search and user-testing.

Gesture-based interfaces may not always be the most preferred or practical interaction method. Even though the novelty they provide causes interest and excitement, there can be drawbacks in regards to user fatigue, unintended input, inflexible use conditions, and lack of tactile feedback. The implications and potential impact of these issues is cause for further review, which will benefit the direction of this study.

Dependence on Augmented Reality

Are use-cases of hand-gesture interaction on the mobile platform limited to those also including augmented reality? Uncovering the exact relationship between the two areas provide cause for investigation. It may also prove interesting to review the AR topic specifically, and how it is supported with smartphone technology, since it also relies on techniques in computer-vision.

Vidal and Rodrigo [20] describes increased popularity of augmented reality use in applications, incentivized by the two standard application programming interfaces; Google's ARCore and Apple's ARKit. However, they are limited compared to HMDs (head-mounted displays), since they are only gathering information from a single RGB-camera and a digital gyroscope [20]. The study explored the feasibility of using a dual-back-camera (stereo camera) smartphone to develop AR applications featured hand gesture recognition(HGR) as a method of interaction. They used a HoloLens head-mounted system as a reference point for the framework developed in the study, and concluded that AR-HGR applications are possible on smartphones [20].

The next step building on Vidal and Rodrigo's conclusion, and in the context of this project, is to attempt implementation of the Section 3.2 heuristics in AR-HGR applications. The potential might be tested by building experimental demos of some distinct few game mechanics in AR-HGR based mobile games, and facilitate user-testing and evaluation of the interactive experience they provide.

Chapter 5

Conclusion

This study explored areas of existing utilization of hand-gesture recognition (HGR) and related methods of interaction, and from those identified some key concepts that shows promise as the basis for creating future mobile games based on that interaction method. The areas of this research was structured into four distinct domains of utilization. From Domain A on virtual reality (VR) gaming, the first contributed heuristic was derived, which propose that these mobile games have the HGR interface as basis for primary game features. This proposition was reinforced by the findings from Domain D on Public Service Displays. A second heuristic offered, propose that the creation of future HGR mobile games emphasize the great accessibility afforded by the platform, based on findings from Domain A, as well as Domain B on Health Technology. Findings from Domain C on Public Entertainment Systems did not contribute to any specific heuristics offered, yet it suggests use-cases for smartphones in museum exhibitions. Domain D offers a third heuristic proposing, that HGR mobile games should allow the user to maintain composure, if they are meant to be played in public spaces. Last the fourth heuristic propose that HGR games should invoke curiosity to a degree that motivates the seeking of experience outside the user's comfort zone, primarily based on the novelty offered.

The proposed heuristics are intended to guide developers wanting to create pioneering mobile games based on hand-gesture interaction, and help them decide the direction and overall targeting of such titles in the mobile gaming market. The goal is to help shaping design visions and carve out a position in mobile gaming industry, that can support the initial introduction, and reinforce creative diversity of mobile games.

Bibliography

[1] Kevin Browne and Christopher Anand. An empirical evaluation of user interfaces for a mobile video game. *Entertainment Computing*, 3(1):1–10, 2012.

[2] TechJury.net Christo Petrov. 43 virtual reality statistics that will rock the market in 2021, 2021.

[3] TechJury.net Teodora Dobrilova. 23+ mobile gaming statistics for 2021 – insights into a $76b games market, 2021.

[4] Pandaily.com Greg Grigorian. How mobile gaming conquered china, 2019.

[5] Johnathan Bown, Elisa White, and Akshya Boopalan. *Looking for the Ultimate Display: A Brief History of Virtual Reality*. Elsevier Inc., 2017.

[6] IMDb. Ready player one (2018).

[7] Beat Games. Beat saber (2019).

[8] Valve Corporation. Half-life: Alyx (2020).

[9] The Game Awards. 2020 winners.

[10] Mais Yasen and Shaidah Jusoh. A systematic review on hand gesture recognition techniques, challenges and applications. *PeerJ Computer Science*, 2019(9):e218, sep 2019.

[11] Zahoor Zafrulla, Helene Brashear, Thad Starner, Harley Hamilton, and Peter Presti. American sign language recognition with the kinect. *ICMI'11 - Proceedings of the 2011 ACM International Conference on Multimodal Interaction*, pages 279–286, 2011.

[12] Cheok Ming Jin, Zaid Omar, and Mohamed Hisham Jaward. A mobile application of American sign language translation via image processing

algorithms. In *Proceedings - 2016 IEEE Region 10 Symposium, TENSYMP 2016*, pages 104–109. Institute of Electrical and Electronics Engineers Inc., jul 2016.

[13] Rolf Steier, Palmyre Pierroux, and Ingeborg Krange. Embodied interpretation: Gesture, social interaction, and meaning making in a national art museum. *Learning, Culture and Social Interaction*, 7:28–42, 2015.

[14] A. Jiwane and A. F. Khan. Interactive museums: empowering visitors' engagement. *IET Conference Proceedings*, pages 308–313(5), January 2021.

[15] Vincenzo Agate and Salvatore Gaglio. A gesture recognition framework for exploring museum exhibitions. *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, pages 6–8, 2018.

[16] Cornelius Malerczyk. Interactive Museum Exhibit Using Pointing Gesture Recognition. Technical report, 2004.

[17] Jörg Müller, Florian Alt, Albrecht Schmidt, and Daniel Michelis. Requirements and design space for interactive public displays. *MM'10 - Proceedings of the ACM Multimedia 2010 International Conference*, (Figure 1):1285–1294, 2010.

[18] Statista.com J. Clement. Mobile gaming market in the u.s. - statistics & facts, 2021.

[19] Niantic Inc. Pokémon go (2016).

[20] Eric Cesar E. Vidal and Ma Mercedes T. Rodrigo. Hand gesture recognition for smartphone-based augmented reality applications. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 12190 LNCS, pages 346–366. Springer, jul 2020.