An Adaptive Approach to Exergames with Support for Multimodal Interfaces

Juan Manuel Silva Salmerón

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
for the PhD degree in Computer Science

School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

Abstract

Technology such as television, computers, and video games are often in the line for reasons of why people lack physical activity and tend to gain weight and become obese.

In the case of video games, with the advent of the so called “serious games initiative”, a new breed of video games have come into place. Such games are called “exergames” and they are intended to motivate the user to do physical activity. Although there is some evidence that some types of Exergames are more physically demanding than traditional sedentary games, there is also evidence that suggests that such games are not really providing the intensity of exert that is at the recommended levels for a daily exercise. Currently, most exergames have a passive approach. There is no real tracking of the players progress, there is no assessment of his/her level of exert, no contextual information, and there is no adaptability on the game itself to change the conditions of the game and prompt the desired physiological response on the player.

In this thesis we present research work done towards the design and development of an architecture and related systems that support a shift in the exertion game paradigm. The contributions of this work are enablers in the design and development of exertion games with a strict serious game approach. Such games should have “exercising” as the primary goal, and a game engine that has been developed under this scheme should be aware of the exertion context of the player. The game should be aware of the level of exertion of the player and adapt the gaming context (in-game variables and exertion interface settings) so that the player can reach a predefined exertion rate as desired.

To support such degree of adaptability in a multimedia, multimodal system, we have proposed a system architecture that lays down the general guidelines for the design and development of such systems.

Acknowledgements

I want to express my deepest gratitude to Professor Abdulmotaleb El Saddik for his advice and guidance during my research at the multimedia communications research laboratory. Thank you professor for all your time and for going beyond your responsibilities as a supervisor and become a friend who always cared for the well being of me and my family.

Next I would like to thank my colleagues at the research lab with whom I held interesting conversations and who were always prompt to spark in me new ideas and directions. In particular I would like to thank my good friend Mauricio Orozco. Thank you Mauricio for your help in, and outside the lab, it was always a pleasure to have those lengthy conversations about research and life. Special thanks also to my other co-authors Jongeun Cha and Nasser Dardas. Collaboration with the three of you was very fruitful and I very much appreciate your work and dedication.

This thesis is dedicated to the greatest loves of my life. To my wife Dulce, who has been my support for the greatest part of my life. Thank you for constantly pushing me forward and giving me the confidence to fulfill my dream of finishing my PhD program. To my two daughters Sofia and Laura who are my two gems and the source of inspiration for all the things that I embark upon. To my parents Juan and Rosalba who taught me the value of pursuing higher education and whose legacy I will pass on to my daughters. To my brother Moises and my sister Nayeli, for the countless hours of fun and laughter that we have shared in our lives.

Table of Contents

List of Figures

List of Tables

1 Introduction

According to the World Health Organization (WHO), overweight and obesity are risk factor in serious medical conditions such as diabetes, cardiovascular diseases and even cancer. WHO estimates that by 2015 2.3 billion adults will be overweight and more than 700 million will be obese across the globe [WHO_10]. It has been reported that the key causes are increased consumption of energy-dense foods and reduced physical activity. In Canada, the 2007 Canadian Community Health Survey showed a self reported obesity rate of 17% in adult population (ages 18 and over), while the Public Health Agency of Canada estimates that probably the actual rate is at 25% [PHAC_10].

A lot has been said about the sedentary life that the new generations of kids and young adults are having. Particularly, technologies such as television, computers, and video games are often in the line for reasons for people's lack of physical activity and tendency to gain weight and become obese [Marshall et al_04] [Salmon et al_05].

Computer scientist have taken their part is the effort to alleviate this problem by encouraging people to engage in physical activity through those same technologies that are so addictive and which are causing the problem itself. In the case of video games, with the advent of the so called "serious games initiative" and particularly the games for health movement, a new breed of video games have come into place [GamesHealth_11] [SeriousGames_11]. Such games are called exertion games or simply "exergames". They are intended to motivate the user to do physical activity beyond barely moving the fingers on the game pad to play. As defined by Vossen [Vossen_04], in general, an exergame is where:

"*Physical activity must actually influence the game outcome by either omission or commission*" – Vossen D. 2004

Exergames come in many modalities. Usually a common feature of exergames is a novel user interface which may incorporate one or more sensors such as accelerometers, gyroscopes, laser pointers, video cameras, and even global positioning systems. Such array of sensors and devices has mostly been used to track and assess the movement of the game players as they perform a physical activity, so that the game can respond accordingly [Exergames_12]. Other games incorporate cardiovascular fitness machines (such as stationary bikes), but there is a lack of formal studies that provide evidence as to how these machines contribute to increase the amount of exercise the subject is performing.

Exertion videogame titles, consoles, interfaces and arcades have been developed in both, the commercial industry, and in the academic research. Commercial products have focused in enhancing the “fun” side of videogames by incorporating novel modes of interaction. Examples of such types of games include the arcade game Dance Dance Revolution, Wii Sports, EyeToy Kinetic series and Microsoft’s Kinect series for the Xbox platform. On the other hand, scientific research has treated exergames in the context of serious games, focusing in the motivational aspects that can drive a person to perform physical activity while still having fun in the process.

All such initiatives have the same goal in common: to motivate players to engage in active physical activities by playing a game. However, their limitation lies on the lack of feedback for the application, so that it can better respond and adapt to the different contexts (i.e. level of physical fitness) of the players.

1.1 Motivation and application scenario

Although there is some evidence that certain types of Exergames are more physically demanding than traditional sedentary games [Warren_06] [Graves et al_07], there is also some evidence that such exergames are not really providing the intensity of exertion that is at the recommended levels for a daily exercise. Some studies show that usage of modern video game consoles aimed at promoting physical activity, like Wii or EyeToy, do not necessarily provide a meaningful exercise session [Graves et al_07] [Luke_05].

Currently, most exergames have a passive approach. For the most part, the games are focused on tracking the movements of the players and map those movements to the typical input that would normally be received from a hand held game controller device. The overall objective of this strategy is to aim at increasing user engagement by being "more fun" to play, while encouraging the user to perform some physical activity. Nevertheless, in this approach there is no real tracking of the players exercising progress, there is no assessment of his/her level of exertion, there is also no contextual information involved to make in-game decisions and verify that the level of exertion is appropriate for the particular conditions of the player, and finally there is no adaptation approach to change the conditions of the game that would prompt the desired physiological response on the player to increase or decrease the amount of exercise required to play the game.

A game with such characteristics could be a valuable tool for improving the overall health of a person. In a "Serious Games" context, exergames should have "exercising" as the primary goal, and yet they should provide "Fun", through a media rich experience. We propose a shift on the way that exergames are conceived and developed, where the actual enactment of physical activity would be firstly addressed, and the game part would just be the means to motivate the user. Based on that premise, we promote the term "Game based exercising" as a differentiation from "exergames"; while the primary goal on the latter is to have fun (and hopefully get some exercise), the former is meant for doing meaningful exercise, while having fun in the process.

To better illustrate the motivation for our research consider the following scenario:

A teenager boy with a problem of obesity has been put into a diet and exercise regime by his family doctor. He is supposed to do 20 minutes of active exercise every day in order to achieve his exercise requirements. To give the teenager a motivation to stick to the exercise program the physician recommends that he plays an exertion game every day for at least 20 minutes.

The exertion game consists of a typical target shooting game. However the game input is directed by an interface that requires physical exertion to play the game: A stationary bicycle requires the boy to pedal in the bike in order to aim at the targets that he wants to shoot, and a camera enabled with a motion sensor can track his body on top of the bike and detects "throwing/shooting" gestures that the boy makes with his arms in the air above his head.

Additionally the boy's age and the level of exertion required are input in the game before the session. While the boy plays, he is wearing a heart rate monitor that wirelessly feeds its data to the game engine. The game knows how much the player should exercise and knows how much he is exercising at all times. The engine actively adapts to the player's context and modifies the game dynamics so the heart rate of the boy is driven to the target zone so that he can reach the desired level of exertion, remaining there for the duration of the game session.

In order to realize a system like the one depicted in the scenario, several requirements must be fulfilled.

1) **Context Awareness**: The system (in this case the game) must have context awareness of the player interacting with the game. Context in this case is a two-fold concept: One is the current level of effort applied to play the game, which is the heart rate of the player, and second is the level of exertion desired for the session, which is derived from contextual variables such as the demographics of the player (i.e. sex, age) and the physical state of the player (i.e. weight and level of fitness).
2) **Adaptability**: This awareness of the context of the player must be used by an adaptation strategy that defines how the variables of the game should be changed in order to fulfill two main goals: Keep the game enjoyable and keep the game as an exertion tool.
3) **Natural Interface**: Given the nature of exertion interfaces, players cannot use traditional input methods such as joysticks, gamepads, mice or keyboard. Hence, the system must support the use of multimodal natural interfaces.
4) **Backward compatibility**: Furthermore, because the majority of computer games currently available were not designed with exertion natural interfaces in mind, there is a need for middleware that enables backwards compatibility of new exertion interfaces into pre-existing computer games.
5) **Exertion output control**: The game should not only adapt to the conditions of the player, but it should adapt with objectivity in a measurable way. This is, the game should act as a control algorithm of the variable(s) used as an indicator of the level of exertion. In the scenario described above, the heart rate of the player is such variable.

> The system should drive the player’s heart rate up and down to maintain it within the boundaries of a predefined target exertion zone.

Design and development of such system involves meeting requirements from different subsystems, including multimodal natural interfaces, adaptation engine and context awareness. Because of those requirements, there is a need to maintain the quality of experience of the multimedia system by keeping a synchronized flow of all the media channels converging in the game. Such media channels include audio, video and possibly others such as haptics. Because of this, before embarking on the design and development of such system, we must first understand the relationship between media channels and the synchronization issues that might arise. Particularly there is a need to know the threshold of amount of skew between media channels that humans can tolerate before being aware that a synchronization problem exists in the multimedia system. Then, that threshold should guide the design and implementation of the related multimedia system.

In order to achieve a scenario like the one presented before and fulfilled all of its requirements, several research questions must be addressed. In this work we examine those research questions and next present them in terms of objectives and the methodology to follow to reach them.

1.2 Scope of research

1.2.1 Research statement

Given the motivations described in the previous section, our aim with this research is to proof the feasibility of an exertion gaming system that can adapt to the different levels of physical fitness of the players. The system should dynamically change the properties that define the game interface itself, to better suit the physical capabilities of the player. For this purpose, we propose an system architecture that can be used by others in the design and implementation of games with such characteristics. Such architecture should support the design and development of exertion games enabled by interfaces with multiple modes of interaction.

1.2.2 Objectives

In order to fulfill the requirements as described before, we have defined the following objectives.

Objective 1: Design an adaptive exergame system architecture:

Identify the functional requirements for exertion video games that comply with the precepts described by our motivation in terms of context awareness, adaptability and enhanced modes of interaction. Describe the system components that would fulfill the requirements in an architecture that can be used as a reference for the realization of such exertion video games.

Objective 2: Design and develop exertion estimation, and adaptation algorithms for an exertion game

Design and development of algorithms to estimate the level of exertion of an individual playing the game, and adapt to change the game conditions demanding more/less exertion from the player, accordingly.

Objective 3: Design and develop an adaptive exertion game implementation that supports multiple modes of interaction.

By following the precepts described in the proposed architecture, implement a game that is capable of adapting and changing its dynamics, based on contextual information about the level of exertion of the player. Integrate the game with multimodal interfaces, including body tracking and gesture recognition, using smart-phones and sensors as input devices, and supporting haptic interfaces.

Objective 4: Evaluate the adaptive game implementation

Assess the effectiveness of the overall adaptation strategy to increase the demand of exertion from the player, when compared to a non-adaptive version of the game. Perform usability testing of the multimodal interfaces integrated in the game.

1.3 Contributions

The major contributions of this thesis are:

- I. Design and development of an exertion game architecture for development of adaptive, multimodal, exertion games that primarily support an increased level of exercising during game sessions, while still providing a rich multimedia entertaining experience.
- II. Design and development of adaptation algorithms to control the heart rate of an exertion-game player, by modifying the rules and properties of videogames in order to prompt the desired physiological repose.
- III. Design and development of an architectural model for rapid integration of natural interaction exertion interfaces into pre-existing videogames, without the need to modify the videogame's source code.
- IV. Design and development of smartphone applications used as input interfaces to drive actions in an exertion game.
- V. Design and development of gestural exergame input interfaces to control a videogame using hand and body gestures.
- VI. Design and development of a videogame hardware device implemented of two haptic displays that enable spatial perception of game objects in order to improve accessibility of computer games for the visually impaired.

1.4 Scholarly output

1. Juan M. Silva, Mauricio Orozco, Jongeun Cha, Emil m. Petriu and Abdulmotaleb El Saddik. Human Perception of Haptic to Video and Haptic to Audio Skew in multimedia applications. ACM Transactions on Multimedia Computing Communications and Applications.
2. Juan M. Silva, Abdulmotaleb El Saddik. Enhancing the Exergame Experience by Using Smartphones as Multimodal Game Controllers. Springer Multimedia Systems Journal.
3. Nasser H. Dardas, Juan M. Silva, Abdulmotaleb El Saddik. A Target-Shooting Exergame with Hand Gesture Control. Springer Multimedia Tools and Applications Journal. P

Papers in refereed Conferences

1. Juan M. Silva, Abdulmotaleb El Saddik. 2011. Evaluation of an Adaptive Target-Shooting Exergame. In *Proceedings of 2nd International Conference on Serious Gaming and 7th Science meets Business Congress (GAMEDAYS 2011), Darmstadt, Germany, September 12-13, 2011.*
2. Juan M. Silva, Abdulmotaleb El Saddik. 2011. An adaptive game-based exercising framework. In *Proceedings of 2011 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems (VECIMS 2011), Ottawa, ON, Canada, September 19-21, 2011.*
3. Juan M. Silva, Alain Mouttham, and Abdulmotaleb El Saddik. 2009. UbiMeds: a mobile application to improve accessibility and support medication adherence. In *Proceedings of the 1st ACM SIGMM international workshop on Media studies and implementations that help improving access to disabled users* (MSIADU '09).
4. Juan M. Silva, Abu Saleh Md. Mahfujur Rahman, and Abdulmotaleb El Saddik. 2008. Web 3.0: a vision for bridging the gap between real and virtual. In *Proceedings of the 1st ACM international workshop on Communicability design and evaluation in cultural and ecological multimedia system* (CommunicabilityMS '08).

Book chapters

1. Mauricio Orozco, Juan Silva, Abdulmotaleb El Saddik and Emil Petriu. The Role of Haptics in Games. In Haptics Rendering and Applications. INTECH Open Access. ISBN 979-953-307-420-6

1.5 Thesis structure

The thesis is organized into 7 chapters. Chapter 1 presents an introduction to the research work, outlining the motivation, objectives, research tasks and the contributions and publications derived from this work.

Chapter 2 presents first an overview of the background field of serious gaming, providing the context for the current research, and then it presents the related work in exergaming. This includes an overview of the commercially available products and the latest research efforts in this area.

Chapter 3 is a preliminary study that investigates the threshold of human tolerance to inter-stream skew between Haptics-Audio/Video in a multimedia application. The findings of this study are the basis for the requirement of quality of experience that guided the rest of the work.

Chapter 4 presents the design principles and description of the proposed architecture components and algorithms.

Chapter 5 gives the implementation details of the game and interfaces that were developed as a proof of concept of a system derived from the proposed architecture.

Chapter 6 consists of the user studies carried out in relation to the applications and subsystems that were implemented, as a validation of the architecture concepts. It explains the methodology that was followed in each case and presents the results and discussion related to each one of those evaluations.

Finally Chapter 7 is a summary of the research work. It presents the conclusions reached and suggests possible directions for future work.

2 Background and related work

2.1 Serious Games

Playing games has always been an integral part of human development. Since childhood games are played to enforce social behaviour, to practice skills that will become vital in the future and to gain knowledge about the world around us [Roskos_00] [Bergen_98]. More structured games have been designed to impart knowledge about academic matters like math and science, and with the advent of computer and the digital age the use of computer games for learning intensified [Mitchell and Savill-Smith_04].

Almost in any given scenario anyone would agree that games should be fun and entertaining. However games may also serve another purpose than just entertainment, such as those games children play to learn and practice skills. Modern computer based games are designed to teach and train general population and specialized professionals about some subject matters, while being entertained. Such games are called “serious games”.

This does not mean that purely entertainment games are not serious business. In fact the computer gaming industry is one of the fastest growing and profitable of recent years. In 2009, Bobby Kotick, the CEO of Activision (one of the largest videogame development companies) estimated that the industry will grow from $39 billion in 2008 to $55 billion in 2012 [Kotick_09]. Serious game industry is also growing. As an example, in 2009 the governments of Korea and France independently announced funding for $63 and $50 million dollars respectively [Gamepolitics.com_10][01Net.com_10].

As the industry grows, new technology is developed and it becomes evident that such a comprehensive technological infrastructure could and should be leveraged to develop games that could contribute more than just entertaining. However, the proper way to realize games that can train or educate, while keeping the entertainment factor on the loop, still requires some research.

In 2002, Ben Sawyer and others started the Serious Games Initiative with the aim to enforce and promote the development of games with other purposes than entertainment [Serious Game Initiative_10]. This is pretty much the time that the term “serious game” came into wide use. Other notable groups have formed in universities and research centers around the world in several countries such as The United States, Canada, United Kingdom, Norway, Denmark, Finland, Germany, and Sweden [Sussi *et al*_07].

In this overview of serious games we first address the issue of the formal definition of the term, which has been a point of discussion among experts [Narayanasamy *et al*_06] [Johnston and Whitehead_09]. We present a classification of games per domain, and discuss the issue of evaluation of effectiveness of using serious games as it has been debated in the academic research [Raybourn and Bos_05]. We also present some studies around the psychology and social behaviour of gamers [Clarke and Duimering_06] [Björk_08] [Schuurman_08]. Finally we analyze the overall research challenges and the future trends of this field and present our conclusions.

2.1.1 Definition and delimitation

It is nowadays generally accepted that the term “Serious Games” refers to computer video games that are used with a purpose beyond or other than entertainment. However some discussion has been going on for some time about the formal definition and some other details about what may or may not constitute a serious game.

According to a definition stated by the people on the Serious Games initiative, started in 2002 by Ben Sawyer and others [Serious Game Initiative_10], a serious game can be defined as follows:

> *“applications of interactive technology that extend far beyond the traditional videogame market, including: training, policy exploration, analytics, visualization, simulation, education, health, and therapy.”*

We think that in fact, the current videogame market is already including serious games. It should be noted that the domains listed in this definition are not exhaustive.

For Michael Zayda, serious games are an extension of common video games originally intended for entertainment; adding a pedagogical component, which involves activities within the game context aimed at educating or training the player, imparting knowledge or skills [Zayda_05].

As Zayda emphasizes the importance of the pedagogical component, it is important to note that serious games should still be entertaining. This is particularly emphasized by Clark C. Abt who argues that two key elements of the definition of serious games are: the main purpose of the game is other than entertainment, and this does not mean that the game will not be amusing for its user. In fact, given that one of the intrinsic characteristics of playing a game in general is that it is a voluntary act [Huizinga_38], it is expected that the game should deliver some degree of entertainment in order to keep the user interested enough to achieve the ulterior motive of education.

Johnson and Whitehead [Johnston and Whitehead_09], came up with a method to tell apart games, serious games and simulations based on the basis of intent. They have basically moved the source of the definition away from the original purpose and content of the game, and put it on the user's intent. In their method for classification, they initially make a distinction of Games and Serious Games by asking what is the intention of use from the user. If the user's intent is entertainments then we are looking at a general game. Else, if the primary goal of the user is education (in any of its forms), then it can be considered a serious game. Furthermore, they argue that if the serious game is closely related to the user's reality (like his profession), then the serious game becomes a training simulation. Figure 2.1 shows their hierarchical classification of games and simulations. We think this is a good way of making the distinction; particularly for games and serious games. Just because we call it a serious game it does not remove it from its game classification.

Figure 2.1 Classification and interrelation of Games, Serious Games and Simulation [Johnston and Whitehead_09]

2.1.2 Characterization of Serious Gamers

The term "digital native" is used to refer to those individuals who were born and have grown on the new digital era where the Internet, mobile communications and multimedia systems are ubiquitous, readily available and fully integrated into their everyday lives. According to a report, students in ages 6 to 17 who go online make use of the Internet as a tool for homework and research on a daily basis and as much as thirty percent of them play online games at least once a week [Connected to the Future_03].

However, playing games is not an activity solely related to young digital natives. The United States Entertainment Software Association reports that 50 percent of the American population and as much as 68 percent of American heads of households play computer and video games [Top Ten Industry Facts_10]. They also report that as much as 43 percent of players are between 18 and 49 years old [Facts & Research_10].

According to a study [Jukes and Dosaj_06], those digital learners have some preferences about how they want to learn. For them it is important:

- Receiving information quickly from multiple multimedia sources
- Parallel processing and multitasking

- Processing pictures, sounds, and video before text
- Random access to hyperlinked multimedia information content
- Interacting/networking simultaneously with many other individuals
- Move seamlessly between real and virtual spaces instantaneously
- Learning with a “just-in-time” approach
- Instant access to friends, services, responses, gratifications and rewards
- Acquire knowledge by a process that is relevant, instantly useful, and fun

There is a clear parallelism between the above listed requirements and the features available in digital video games, hence the potential for developing serious games that can target those digital oriented individuals.

There is still a lack of studies that take into consideration the training and learning aspects of game playing motivation, and draw some conclusions as to what players of serious games would expect and desire while playing games for those purposes.

2.1.3 Classification of serious games

Serious games have been developed for many markets, including education, government, military, corporate, healthcare, politics, religion and arts. Design, development and means of distribution for each of these markets should be considered individually, since each market has special requirements [Michael and Chen_05].

For this classification, we present the domains on which serious games have mostly been developed or researched. It is important to note, that his is not an exhaustive list of domains and terminology. As serious games gain popularity, new domains may be added and new terms may surge in this industry.

Education

Also called “edutainment”, educational games refer to the games meant to teach some academic subject, probably in a classroom or in support to some type of formal education.

Edutainment is the main field that is commonly related to Serious Games in education. However, it is important to point out that even when Edutainment is a subfield of serious games, it is does not comprehend the entire scope of serious games for education. Edutainment has been around for quite some time, and it has had a series of design flaws that have generally prevented its full success [Egenfeldt-Nielsen_07]. Mostly, Edutainment has failed to do a proper integration of the educational component into the game-play experience. New approaches of serious games for education aim at solving the issues of edutainment, keeping the game-story as the main component and subordinating the pedagogical component to it, introducing a subtle learning experience. [Zayda_05].

Mitchell and Savill-Smith performed a comprehensive study of previous literature work on the use of video games for learning. They analyzed: (1) The impact of computer games on young people, (2) why should computer games be used for learning, (3) how have computer games been used for learning, (4) what are the experiences of young people with such technology and (5) the recommendations and conclusions that can be drawn from those experiences [Mitchell and Savill-Smith_04].

Healthcare

Research and industrial developments in this field could be further subcategorized based on the purpose of the game. On one side we have games that can educate the general population about health related concepts and good habits. On the other hand there are games that actively prompt the user to engage in physical activity, so to improve their overall wellness. The later are commonly known as "exergames".

Suhonen and others did a case study with adolescents, analyzing their technology usage, game habits and gaming motivations in order to draw a guideline for the development of health related computer games[Suhonen *et al. 2008*]. As part of their study, they have identified a series of games that exemplify both of the subcategories discussed here. *Bronkie the Bronchisaurus* and *Packy the Marlon* are examples of games for disease self management (asthma and diabetes respectively) [Lieberman_01]. Some health games take it a bit further and act as motivators for staying in a diet regime or keep up with physical activity, without directly demanding it from the user. *My health couch* for

Nintendo DS is a mobile game that can do some tracking about the level of physical activity (by means of a pedometer) and user food intake (by direct user feedback) and do recommendation for the next day [My Health Couch_10].

Corporate

Kevin Corti, has identified some potential benefits of computer based learning to the corporate arena: Performance Improvement, increased awareness of employee roles, competency testing and assessment, recruitment, customer education, promotional tool, induction and motivation of new hires, and implementation of best practices [Corti_06].

Military

Serious games have been used by the military with two main purposes: recruitment and training. In the context of recruitment, America's Army is a game that evolved beyond its original recruitment purpose, and has been used as a platform for other military training simulations [Testa_08]. Other uses include strategy and tactics, skill and team building, joint task force operation training among others [Macedonia_01][Prensky_01]. Paul Roman and Dough Brown, made a comparison study between constructive simulation and serious games approach for training purposes in a military context. [Romand and Dough_07].

Government

The application of serious games at the government level include games intended to outreach to general population for various topics, employee training in the different governmental departments, and policy and decision making [RILEY_07] [Mayer_09]. As an example of the increased interest from governments for serious gaming, Korea and France independently announced funding for $63 and $50 million dollars respectively in 2009 for serious games initiatives [Gamepolitics.com_10][01Net.com_10].

Politics

Political games are commonly focused in some form of light simulation or role-playing around part or the whole political system of a country. One example of a political related game is the elections related game Campaign Rush [Bogost_10]. Another example is the

Tropico, which is a game that places the player as the leader of a fictitious Caribbean island. The game plots some political scenario, and the player is left to make his/her own decisions as to how to rule the country [Tropico_10].

Social

Social Games, are those related to increasing public awareness of some of the major social problems like poverty, human rights, climate change, global conflict, and other topics alike [Games for Change_10]. One interesting example of a game aimed at enforcing some desired social and/or civic behaviour is *The Booze Cruise* [Parker *et al.*_09]. This is a serious game aimed at educating the general public about the difficulties of driving while impaired. The game is based on a car driving scenario, where the user is supposed to drive home while impaired. The game simulates some of the effects of alcohol ingestion, like delayed response and blurry visual perception.

Religion

Video games with religion content have been a rather controversial topic. Most major game development firms have generally avoided developing games with heavy inclusion of such content. Michael Thompson makes a pretty comprehensive analysis of cases where video games and religion have come together, sometime with good and other times with bad results [Thompson_09].

Examples of titles in this field include *Bible Adventures*, and *Super 3D Noah's Ark* from publisher Wisdom Tree [Wisdom tree_10].

2.1.5 Evaluation of effectiveness

Serious games are a reality and are being used in different scenarios in both, industry and academia. But, since the purpose of serious games is other than pure entertainment, how can we evaluate the degree to which such technology actually allows the player to reach the goals of education, training, health improvement etc. for which the game was developed?. In this section we present work done to address this question.

Richard Blunt presents the results of three studies related to the evaluation of effectiveness of serious games to improve the learning performance and academic achievement on university students. The comparative study is based on performance test of students who used a video game as a learning tool against those who did not used the tool for the same course. Their results show that the academic performance of students who did use the video game was statistically significant and superior compared to the other students [Blunt_09].

Wong and others [Wong *et al*_07] did a study to evaluate the effectiveness of a particular serious game vs a more traditional learning media. In their study they assess the learning effects of interactivity available in games vs non interactive media formats. In their study they exposed groups of users to learning sessions about psychology concepts, using *Metalloman*, a video game developed by them [Marsh et al_05] and traditional learning media such as hypertext and a physical text book. They found that interactivity does not seem as much of a crucial factor for learning as one might initially assume. They rather found that media richness seems to be more important in supporting knowledge gain.

Tashiro [Tashiro_09] contributes to the analysis of effectiveness of serious games in the context of healthcare. They developed a typology of instructional serious games for healthcare, and then evaluated some of the games for their strengths and weaknesses for improving clinical judgment among healthcare staff. They evaluated the games in relation to the standard for instructional materials as defined by the US National Research Council of the federation of American scientists [FAS_10]. They made a critique on several games used at the time within nursing, medical assistant, and paramedic education, pointing out that they did not complied with the above mentioned standards.

Although it is generally accepted that learn and skill gain through play is plausible, it is evident that the consequences of just assuming that the game will train and educate without formal evaluation can be dramatic. The issue of evaluating a game is not a trivial task. As it can be inferred from the work presented here, the evaluation of effectiveness of a game will greatly depend on the domain and the purpose of the game. However a challenge in this area will be to realize a set of guidelines that can allow game developers to do formal testing of the games in terms of its effectiveness.

2.2 Serious Exertion Gaming

Exertion games are generally understood to be a subset of the broader field of serious games. As such, they are defined as videogames that in addition to entertaining they support engagement in physical activity that is translated in health benefits for the player. In order to fulfill this ulterior goal, players are required to do some degree of physical exertion to achieve the primary goal of the game. As defined by Vossen [Vossen_04], in general, an exergame is where:

"*Physical activity must actually influence the game outcome by either omission or commission*" – Vossen D. 2004

For videogames to provide a meaningful exercising session, they must require from the user a significant physical effort, and this might be in direct conflict with the player having fun. This is true especially when players do not have in mind doing exercise as the primary objective. This conflict of interests is the cause for commercially available games not providing the levels of exercise as they would be recommended by a physician [Graves et al_07] [Luke_05].

Achieving this balance between meaningful exercise sessions, fun and enjoyable game sessions is critical for the success of serious exertion games. In this section we present the state of the art exertion games and exertion interfaces. The related work presented here has been selected because it's relevance to address the problem of bringing exertion games to a serious gaming context.

The topics in this section are organized as follows: We first present some initiatives that propose design patters in the form of frameworks or guidelines for the development of exertion games. In this case, researchers have drawn such guidelines based on their own experiences in developing, and evaluating such systems.

Next, we give an overview of common exertion interfaces. In this regard, we present some commercially available consoles and technologies, and then we examine some research work that generally address the issue of encouraging the player to engage in physical activity. The remaining of this section presents examples of different types of exertion interfaces including games driven by hand gestures, the use of smart phones and their sensors as input devices, and an overview of the role of haptic technology in games.

Finally we discuss the use of heart rate as the key piece of information used as a reference to estimate the level of exertion of a player, which can then be used to drive adaptation changes on the game dynamics.

2.2.1 Frameworks and guidelines

Some researchers have focused on defining some sort of guideline or framework for the development of exergames. Mueller and others propose a theoretical framework for exertion games whose purpose is to *"describe how designers can use technology to create more engaging exertion experiences mediated by technology"* [Mueller *et al*_11]. They focus on the analysis of how the human body moves, senses, responds and relates to other people and to stimulus in the context of physical exercising, drawing conclusions that might be of use for exergames developers.

Play Mate! is an interface and design methodology for traditional computer games. It motivates users to perform physical activities that lead to gaining in-game commodities, which allow the player's character to perform better and advance in the game [Berkovsky et al_09].

Another relevant work is that of Jeff Sinclair and others. They have identified some success factors from their experience in the development of exergames and propose a model called "the dual flow" [Sinclair *et al*_07]. Their model is based on the "the flow" theory by Mihaly Csikszentmihalyi from 1975, which defines a state of total engagement in an activity [Csikszentmihalyi_75].

In the dual flow, the main precept is that exergames should sustain a balance between perceived skills and perceived challenge. In other words the game should not exceed by far or fall short of the skills and physical fitness of the player. They furthermore propose a generalized structure for exergames that are based on the dual flow model and implemented a game as a proof of concept [Sinclair *et al*_09]. This game is discussed more in detail in section 3.4 of this chapter.

2.2.3 Exertion Interfaces

In this section we present an overview of the current trends and technologies around exertion game interfaces. We start by presenting commercial games and research initiatives that attempt to drive players to be more active while playing videogames.

On the commercial side, games like Dance Dance Revolution have been of interest not only to the general public but also for academic research [Hoysniemi_06]. In this game, players must follow the beat of music being played by tapping with their feet on a pressure sensitive pad, stepping on the appropriate pad quadrant according to visual feedback provided on a screen. Wii Sports [Wii_11] is another example of an exergame that offers a number of options in terms of type of input controllers. Using the accelerometers and optical sensors inside the controllers, the game can track speed, orientation and acceleration, which allows for all sorts of physical games such as tennis, baseball, bowling, golf, and boxing. The design of the remote control used by Wii system has triggered research to evaluate its usage in non gaming contexts [Schreiber et al_09]. With the advent of new technologies such as the EyeToy camera, which allows for tracking of human movement without the need of other input devices, other exergames have also come into the market, such as the EyeToy Kinetic series [EyeToy_11]. More recently, Microsoft's Kinect for the Xbox platform offers a movement tracking solution which allows the player's body to act as the input controllers [Kinect_11]. This last technology is also being used outside the gaming context; for example, for hands free retrieval of medical images while doctors are operating on a patient [KinectSurgeons_12].

In the research arena there has been some interest in the idea of motivating physical exercising by game-play in order to reduce the effects of being overweight and obese, which has been translated into research projects. Just a few examples of such initiatives include Exertion Interfaces [Mueller and Agamanolis_08], which is a project that aims at increasing the physical activity of computer users and states that persuasive physical activity intervention will contribute to overall wellbeing. MacLellan and others perform measurements of physical activity and the location of users to promote the utilization of walking as a form of transportation as a primary mechanism to encourage people to become more active [MacLellan et al_09]. Similarly, PiNiZoRo is a GPS-based exercise game that promotes walking by enabling a mobile device to trigger a fighting game mini session, based on the location of the user along a predefined map [Stanley et al_10]. Virku is a technological setup that uses an exercise bike to navigate a Virtual Environment while optionally playing a game [Mokka et al_03]. Mueller has done some initial work towards a taxonomy of these types of games [Mueller et al_08].

In most of this work, a lot of thought and research has been focused on increasing the quality of the exertion interface and its relevance to the game play to supporting social awareness and connection.

2.2.3.1 Gestural interfaces

Vision-based hand gesture interface gained a lot of interest in recent years since they can be used to control other applications or videogames. In [Sparacino_08], a natural interface to navigate inside a 3D Internet city was presented using hand gestures. The user stands in front of the screen and uses hand gestures to navigate through the Internet 3D city. All gestures begin from a rest position given by the two hands on the table in front of the body. Gesture recognition is achieved by Hidden Markov Model (HMM) modelling of the navigating gestures. The feature vector contains velocity and position of hands and head, and blobs' shape and rotation.

In [Chen *et al*_11], a real-time hand tracking and hand posture recognition technique was utilized for the Jing-Hang Grand Canal Serious Heritage Game. This method permitted the players to interact with their customized avatar by natural hand gestures, observe specific models and navigate in the virtual constructions. The hand was detected using MCT-based (modified census transform) method [Just *et al*_06]. Then, a multi-cue hand tracking algorithm [Pan *et al*_10] was utilized to track the hand. In the third step, the hand was segmented using a Bayesian skin-colour model [Weng *et al*_10]. Finally hand posture was recognized by the feature based on density distribution.

In [Hurst and Wezel_12], games and gesture-based recognition were augmented into mobile phone interfaces. Finger tracking was experimented in an augmented reality (AR) board game on mobile phones and showed that it sustains an increased level of engagement and entertainment. Using markers attached to the fingers, canonical interactions were evaluated such as translating, scaling, and rotating virtual objects in a mobile AR setting.

In [Takahashi *et al*_11], a human gesture recognition technique was presented that uses 4-D spatiotemporal features. The technique used a time-of-flight camera for input so that depth information can be obtained. Besides, a man–machine interface was developed that senses human gestures and postures for TV viewing that allows intuitive operation through a device-free interface. In [Song *et al*_11], a localized, continuous and probabilistic video representation was explored for human action recognition. The proposed representation makes use of the probabilistic distribution to encode the visual-motion information of an ensemble of local spatial temporal features in a continuous and localized manner. In [Ntalianis *et al*_10], an integrated framework for analyzing human actions in video streams was proposed. The proposed approach introduces the implicit user in-the-loop concept for dynamically mining semantics and annotating video streams.

In terms of videogame implementation that leverage vision-based gesture recognition, we can find work such as that of Kostomaj and Boh [Kostomaj and Boh_09]. They developed an Ambient Interactive Storybook framework for children with videogames that promote physical activity detecting motion of the body to trigger actions in the game. Their approach is to use a webcam to track gross motion changes of the centre of mass of the body to know if the player is moving left, right, down or up, and detect the transition between any of these positions and recognize them as postures. The location of the player is handed over to the game engine which is responsible to respond and manage all game variables.

In a similar work, Varona and others focus on detecting user motion to derive body gestures. In their approach, they use non-parametric techniques to recognize the body gestures, which then they apply in the control of a videogame in real time. In this case, the gesture recognition is not limited to the centre of the body mass, but it makes a more fine detection of different joints of parts of the body including feet, thorax, shoulders and elbows. Their application of body gestures is used to control a Tetris videogame [Varona *et al*_09].

In terms of hand gesture recognition applied to games, Li and others make use of a rear projector and a traditional web cam placed on a table to implement a game using ordinary hand gesture primitives for manipulating the game scenario. Their technique leverages the fact that the rear projector under the table provides a backlit image of the hand making gestures over the table. They make the segmentation by specifying an appropriate light intensity threshold. They support the tracking of several hands interacting over the surface of the table [Li *et al*_11].

2.2.3.2 Smartphones and sensors as input controllers

As mobile phones become smarter and include a wider and more powerful array of sensory components, the opportunity to leverage those capabilities in contexts other than telephony grows. Those sensory capabilities can be used as key components for modern user interfaces that can detect movement, actions and intentions to enrich human-computer interaction in a natural way. In this section, we present some research using smartphones as input controllers in the context of exertion videogames

We have classified the related work into two areas. First we discuss similar work that uses phones as a general tool for interaction with public displays, as this kind of interaction can be extended to the gaming context. Then we discuss more specific research that deals with interaction with videogames in general and exergames in particular.

Researchers have done work to explore the use of smart phones for interaction with public displays. They have investigated the different types of motion detection and their effectiveness to interact with diverse applications on public screens.

In [Boring *et al*_09], Boring and others examine the advantages and drawbacks of three types of input interaction (tilt, move and scroll) to control a pointer displayed on a large public display. From their research, they conclude that move and tilt are the fastest, but the less accurate in terms of positioning the pointer. Similarly, Kray and others explore the use of gestures using a smartphone to interact with public displays, tabletops or other phones. They present their assessment of which sensors embedded in the phones might be more suitable to detect specific gestures to support different types of interaction [Kray *et al*_10].

Scheible and others explored gesture recognition for immediate file transfer of a media file from a smartphone to a public display, so they can manipulate the media file with annotations and then transfer it back to the device [Scheible *et al*_08]. Henrysson explored the capability of smartphones to provide an interface with 6 degrees of freedom isomorphic interaction for the manipulation of 3D objects [Henrysson *et al*_05].

Although the principles of controlling a public display are in part applicable to controlling a videogame, there are some notable differences. Most of the efforts, when talking about control of public displays with smartphones, are focused on controlling the pointer or controlling some productivity application (word processing, presentation slides etc). However, videogames have a higher level of requirements for response time, and accuracy, so depending of the nature of the videogame, different types of interactions should be used.

In a smartphone-to-videogame scenario, Gilbertson and his collaborators investigate the ‘tilt’ movement of a smartphone to control 3D videogames without the need to use buttons [Gilbertson *et al*_08]. In particular, they contrast the use of a typical joypad vs. the tilt interaction in a car driving videogame. They found the tilt interaction as being more “fun” and attractive to the subjects of their study. In a similar study, Vajk and others make use of a mobile phone to implement a controller that behaves similar to the popular Wii game controller, in order to play games on a large public display. They showcased and tested their implementation during a conference with 35 participants and captured their impressions. In their conclusions, they emphasize the increased level of engagement that the novel interface brought to the game experience, according to their observations and data collection [Vajk *et al*_08].

To the best of our knowledge the only research work that uses mobile phones to address interactions with exergames is that of Kiili and Merilampi [Kiili and Merilampi_10]. They make use of smartphones as input controllers using simple motion detection for four different basic exergames developed by them. By detecting sudden movements of the phone, they can detect exercising movements such as jumping or squatting. Their objective is to provide a more engaging experience for children playing the cited exergames. The specific game implementations are straightforward, and their focus is on simple motion detection using the phone's accelerometer. They use a client-server architecture for sending data from the phone (client) to the videogame host computer (server). They use a multiplayer approach, but only a single interaction modality (low-level motion detection). Their approach to evaluating the level of engagement was based on focus groups. They conducted four focus groups where children participated playing the exertion games that they designed. Their data collection included questionnaires and observation of group interaction. The level of motivation was based on their observation of players and factors such as their willingness to participate, the level of excitement while playing and others such as their interest in acquiring these games to use them in their daily lives. Through this subjective assessment, they conclude that the exertion interfaces provided a high motivation for children to get them really immersed in the exergame activity.

2.2.3.3 The role of haptics in games

In the real world, people receive and disseminate information in three-dimensional space. Computers, thru graphical user interfaces, allow users to perceive an imitated three-dimensional world that exists in the real world. Such a virtual world can be enhanced in a more complete imitation of the real space by the introduction of an artificial support technology called haptics. A haptic interface is a device that allows a user to interact with a computer by receiving tactile and force feedback. The interaction can embrace the entire body or only the tip of a finger, giving the user information about the nature of the objects inside the world. Applications of this technology have been increasingly incorporating devices that allow the user to interact with sophisticated graphical user interfaces (GUI's), games, multimedia publishing, scientific discovery and visualization, arts and creation, editing sound and images, the vehicle industry, engineering, manufacturing, Tele-robotics and Tele-operations, education and training, as well as medical simulation and rehabilitation. In a gaming context, the user experience is driven by four aspects: physical, mental, social, and emotional (El Saddik_07). It is on the physical aspects that, force feedback technology (haptics) enhances the game experience by creating a more realistic physical feeling of playing a game. This translates into a rich environment which provides to players (subjects) a higher sense of immersion as well as new and interesting ways to interact with the game environment. In addition this simulated world can be used to do research on applications such as physical rehabilitation, driving training simulation and other serious gaming scenarios.

In this section we present an overview of how haptic technology can and has been leveraged to enhance the user experience in games.

Enhancing the game experience with haptics

Traditional haptic interfaces in games usually consist of vibrotactile feedback incorporated into Joysticks, Wheels and Gamepads. In the case of Joysticks and Wheels, sometimes kinaesthetic force feedback is incorporated as well. Although this enhancements help to have a more immersive experience for traditional videogames, this interfaces remain unsuitable in an exertion gamin scenario, as they usually require the player to be holding the device with their hands and it makes it difficult to control the game while exercising.

A more suitable haptic interface for exergaming is provided by the haptic vests and jackets. Originally designed for medical investigation, Physician Mark Ombrellaro developed a haptic vest that enabled the wearer to feel the impact of bullets, explosions and/or even hand taps in the trunk of the player's body. The 3RD Space Vest is commercialized by TN Games and is advertised as the only gaming peripheral device that allows you to feel what your game character feels. It works with compressed air to provide the player pressure and impact forces that can emulate a wide scope of direction and magnitude [TN Games_11].

Others have worked on similar systems, and although they have not focused on the application in games specifically, it is clear how such technologies could also be applied in a game scenario. In the University of Ottawa in Canada, Jongeun Cha and others have worked in a device to enhance teleconferencing. They developed a jacket that enhances communication with the physical and emotional connection by allowing participants to tickle or tap on the shoulders of each other remotely [Cha *et al*_09].

Similarly, researchers at the National University of Singapore developed the Huggy Pajama wearable system, where the remote communication between parent and child is enabled thru virtual hugs. This is possible by using a doll with embedded pressure sensors as input device and a haptic jacket as the actuator for the hug at the other end. The hug is reproduced by air pockets and the experience is further enhanced by adding a heating element to the device to mimic the warmth of a hug [Teh *et al*_08].

Haptics and serious gaming

In the general scope of serious gaming, haptic interfaces have played an important role in two main fields: healthcare and education. Next, we review some haptic interfaces that, although have not been employed strictly in an exertion gaming scenario, they are of interest, as the technologies and contributions of such works present an opportunity to illustrate the capabilities and potential of haptic devices in serious gaming.

In the context of healthcare, there has been great interest in research how computer haptics technology can assist in the rehabilitation process of people who has had some damage to their motor skills, while playing a game. Alamri and others proposed a framework for the rehabilitation of stroke patients using a virtual environment and a haptic device. They designed five virtual exercises for the diagnosis and rehabilitation of patients with hand impairments. These include moving a cup, arranging blocks, navigating a maze, training with a dumbbell, and grasping a rubber ball. They used a CyberGrasp device to track the subjects hand movements and to provide kinaesthetic force feedback regarding the virtual environment and the objects related to each exercise. With their evaluations, they proved the validity of the proposed framework as a tool to quantitatively measure patient’s progress in the rehabilitation process [Alamri *et al*_08].

The combination of the properties inherent to computer haptics (force and tactile feedback) with the appealing and motivating factor offered by Virtual Environments have provided a framework for the development of various rehabilitation systems that often involve the patient, and the therapists.

Haptic Hand writing sessions and Ten Pin Bowling are two examples of haptic assisted game applications developed by Xu and others at the University of Shanghai for Science and Technology in China. Their Ten Pin Bowling is intended for training of the motor function of post-stroke patients, based on Virtual Reality. They also demonstrate that haptic based hand writing is an efficient way to improve motor skills, postural stability, control and hand to eye coordination [Xu et al_10].

Jack and others worked on serious games rehabilitation using a haptic glove. In their work, the interaction with a virtual environment is enabled by using a CyberGlove and a Rutgers Master II-ND (RMII) force feedback glove. The Virtual Environment was designed to promote the training rehabilitation of a specific parameter of hand movement: range, speed, fractionation or strength. Their system would adapt to the level of rehabilitation achieved by the patient [Jack et al_01].

Similarly Huber and collaborators developed a home based tele-rehabilitation system for children with Hemiplegia. They made modifications to a play station game console to support the use of a haptic glove and their custom made virtual environment game. The game consisted in making specific hand movements to scare away butterflies that would appear on the virtual environment [Huber et al_08].

Broeren et al. Studied the effects of virtual reality and haptics in stroke rehabilitation by using a VR station loaded with a library of games and a hand held haptic stylus device. User interacted with the virtual objects of the games, while the workstation collected data about the 3D hand movements of the patients. They found that the enhanced rehabilitation experience was highly motivational to the patients [Broeren et al_08].

In the education field, most of the work has been towards providing a more immersive and hands-on approach to learning class content. In particular chemistry is one of the subjects often targeted thru haptic assisted learning tools [Fjeld and Voegtli_02].

As a good example, Sato and others designed a haptic grip and an interactive system with haptic interaction as a teaching aid. They focused on the interaction between two water molecules by constructing an environment to feel Van der Waals force as well as electrostatic force with haptic interaction [Sato et al_08]. In a similar way, Comai and collaborators propose a framework for the implementation of haptic bases systems for chemistry education [Comai et al 2010].

Basori and others [Basori *et al.*_08] introduced the concept of human haptic emotion to enhance interactivity and immersion in a virtual reality game. The idea is to convey the emotional state between virtual characters and the player by using the sense of touch, using a haptic device. They argue that the use of haptics as an expression of emotions (i.e. vibrations) can improve the level of immersion and keep the attention of the player on the game. This can be particularly useful for serious games where the users must stay focused on the task to achieve the ultimate serious goal of the game.

A second point of interest for haptics in education is that of learning hand writing skills. Eid and others seveloped a multimedia system for learning hand writing in different languages. The system is based on a haptic stylus controlled by software in such a way that it can guide the movement of the learner in an analogous way that a teacher would hold his/her hand. The system supports various languages including Arabic and Japanese. The amount of strict guidance of the system can be adjusted so that overtime the user makes the writing by him/her self [Eid et al_07].

In a similar work, kindergarten children were subject to hand writing training using a Visuo-Haptic device to increase the fluency of handwriting production of cursive letters. Forty two children participated in the experiment that showed that the fluency of handwriting production for all letters was higher after the training with the tool than those children which were not subject to it [Palluel-Germain et al._07].

Improving Accessibility with haptics

Although computer generated haptics has the potential to be a valuable technology to help gamers in having a more immersive and realistic experience, it is also a very well suited type of interaction to address the problem of accessibility to games for the visually impaired. There are currently different technologies that have been proven to improve the accessibility of electronic media to the blind people, such as automated reading software, voice synthesis, tactile displays with Braille and speech recognition. Yet, those are mostly focused in input/output of text and there is still some type media that is difficult to represent in a non-visual modality, particularly graphics (2D or 3D), which form the basis for any computer game.

Computer haptics seems to be a suitable solution for this problem of game accessibility, but there are still plenty of problems and research to be done on the field. Having an accurate haptic representation of an object in a virtual environment so that a blind person can recognize it is a critical task for accessibility of games for the visually impaired. Furthermore, games involve multiple objects that interact with one another and trigger events that need to be perceived by the player. Representing all such objects and events must be done in a way that the cognitive load for the player is kept in balance. In fact, in most instances the haptic interface must be accompanied by some audio feedback as a complement, as to no overload the user with haptic signals [Yuan_09].

A survey of strategies for making games accessible to the blind shows that audio is the most common used channel to complement or substitute the visual channel [Yuan *et al*_10]. AudioGames.net is an online community portal that has a repository of games whose interface is solely based on audio and the natural haptic interface of keyboard and mouse buttons [Audiogames_12]. In terms of pure haptic games, there are examples such as Haptic Sudoku for the blind, developed by Gutschmidt and others [Gutschmidt *et al*_10]. In this technological solution they emulate visual perception completely by haptic perception. Players can feel the Sudoku board and scan the numbers through the sense of touch using a haptic display, while audio cues alert the user of the outcome of their actions. Another example is Finger Dance, a sound based game for blind people [Miller *et al*_07]. In FingerDance, users try to match musical rhythm patters to keystrokes in the keyboard. The game has no visual feedback at all.

The second approach is sensory substitution. In this case the cues that would normally come from the visual channel are replaced by haptic stimuli. This allows for the modification of currently existing games that were originally designed for non disabled individuals and adapted to be played by the blind.

Blind Hero is a typical example of a game where sensory substitution has been applied to a very well known video game (Guitar Hero from Red Octane). Guitar Hero is a rhythm action game that is played by using a guitar shaped input device with coloured buttons that must be pressed following the corresponding visual cues that appear on the screen at the peace of some rock tune. In the modified version, Yuan and Folmer replaced the visual cues with haptic stimuli coming from a haptic glove that has small pager motors that stimulate the tip of each finger [Yuan and Folmer_08].

In terms of exertion interfaces, VI-Tennis and VI-Bowling are other examples for sensory substitution based on their analogous version of the Wii Console. In this case the haptic interface is based on a motion sensing controller enhanced with vibrotactile and audio cues that allows the players detect the key events in game play. In the case of VI-Tennis the controller would provide vibrotactile feedback to reflect the event of the ball bouncing and the timeframe at which the ball should be hit in return, while audio cues were left in a similar way as on the original Wii Tennis game [Morelli *et al*_10]. For the VI-Bowling they implemented a technique called Tactile Dowsing where the Wii Remote is used to sense the direction to where the bowling ball should be thrown. With the Tactile Dowsing the player moves the remote left and right in a horizontal pattern. The receiver will detect the motion of the remote and the software prompts a vibrotactile feedback on the remote in a pulsing pattern. The delay between vibrotactile pulses is regulated so that the closer the player points to the optimal direction of the throw, the lower the delay and vice versa. This way the player can sense the direction the same way by trying to position the remote in a direction that has a near continuous vibrotactile feedback [Morelli *et al*_10b].

2.2.4 Heart rate as game input

One of the drawbacks of traditional exergame approaches is that the game engine is not aware of the performance of the player in terms of physical exertion. The game engine is only aware of how the player is performing in relation to the game objective (i.e. make points, move to the next level, recover collectable items etc.), but ignores how much physical effort the player is employing to achieve those tasks.

When there is a shift of the primary goal of the game from entertaining to health improvement through physical exercise, the information about the level of exertion becomes crucial, and it can be used to change things in the game to make it more or less difficult for the player. This can in turn have an impact on the performance of the player in terms of advancing and obtaining rewards in the game.

In scenarios involving pure physical activity, such as sports training or exercising according to instruction by a physician, a common method to evaluate the level of effort made during the exercise sessions is to monitor the heart rate. The heart rate is an objective way of measuring the level of effort of an individual and its usage is normally defined by target heart rate zones. Each zone delimits to boundaries at which the heart rate reading of an individual should remain to perform at the level specified by the zone.

Provided that the heart rate can be used as an objective measurement of the level of effort of an exergame player, the nature of the change that this reading would prompt in the game can vary depending of the kind of exertion interface and the general game engine rules. In general, regardless of what changes happen in the game context, the desired physiological response from the player would be to *exert more*, *exert less* or *stay equal*, depending on whether the current heart rate reading is below, above or precisely at the target heart rate zone respectively.

In this regard, some researchers have worked on an approach where heart rate monitoring from the player is used as input to the game to prompt some sort of change in the game context.

Health Defender is an example of a game where the heart rate directly affects game play. In this game the player must make active physical movement (at his discretion) to elevate their heart rate in order to gain bonuses and ammunition, which would allow them to perform better in the game [Garcia and Coulton_08]. In this particular example, the heart rate was considered as an input that had some effect on the game. However, heart rate monitoring was loosely tracked only to recognize significant changes of the heart rate at time intervals with no intention to drive the level of effort to some specific target zone. There was no change either on how challenging the game was for the player at any time [Garcia and Coulton_09].

Göbel and others propose the use of a tool for personalization of exergames [Göbel *et al*_10]. In their approach, exergames can be customized in the form of storytelling, where an number of contextual variables such as training plans, and vital sign readings can be used to make decision on the direction that the story of the game will go. Those story paths may pose different level of challenge, both physically and in terms of skill. Although they do address the change of game variables and properties based on vital signs reading, there is no indication that they are actively driving the level of exertion of the user to some predefined target zone.

Jeff Sinclair and others identified some success factors from their experience in the development of exergames and propose a model called "the dual flow" [Sinclair *et al*_07]. In the dual flow, the main precept is that exergames should sustain a balance between perceived skills and perceived challenge. In other words the game should not exceed by far or fall short of the skills and physical fitness of the player. Furthermore they propose a generalized structure for exergames that are based on the dual flow model and implemented a game as a proof of concept [Sinclair *et al*_09]. To the best of our knowledge Sinclair's is the only work that actively implements an adaptive engine that continuously updates the game dynamics based on the player's physical and skill performance. However, their management of the physiological input from the user's condition does not affect the level of the challenge of the game; instead, it is kept as a separate variable that simply affects the resistance of the exercise bike that they use as an input controller for the game. Furthermore, their framework relies on a predefined exercise plan to be input into the system and there is a tight link between the game in question and the adaptive engine, where the latter needs detailed knowledge of the game to estimate the skill level.

2.2.5 Summary

Exergaming is a broad field and there is a thin line between games that entertain and happen to have an exertion interface and games that aim to have the user do exercise and provide an entertaining experience. In both cases applications rely on rich media interfaces where interaction must be as natural as possible and the relationship between the interface and game control must be intuitive.

The trends for those interfaces are based on several technologies. Technologies such as hand and body tracking by employing computer vision techniques, or by leveraging on sensors that players might be carrying with them, such as those found on modern smartphones.

Haptic technology although relatively new, is having a great impact on how humans interact with videogames. Early implementations of haptic interfaces are meant to improve the quality of the multimedia experience, but those same techniques and advances can easily be extended and applied specifically for exertion games. One good example of this is the use of haptics to improve accessibility of exertion games for the visually impaired.

Finally, as the focus of exertion games moves more into a serious gaming context, the need to monitor heart rate to drive the game dynamics becomes more important. There are some initiatives already taking steps in that direction and some specific implementations and general guidelines have been put in place, albeit their lack of focus on actively driving the level of effort of the player to a predefined target zone.

It is the combination of these sophisticated multimedia interfaces with the intelligence of adaptive game engines what will drive the next generation of exertion games that truly support physical activity as the primary serious goal.

In table 2.1 we have made summary of related work that addresses to some degree the requirements identified in Chapter 1.

From this table it can be seen that none of the revised works fully addresses all 5 requirements. As it will be presented in the next chapters, the proposed architecture addresses all five.

Scope	Type	Product / Project	Natural interface	Context Awareness	Adaptability	HR Control	Backward compatibility
Commercial Product	Console	Wii Sports	Yes	No	No	No	No
	Console	EyeToy Kinetic series	Yes	No	No	No	No
	Console	Microsoft Kinect Series	Yes	No	No	No	No
	Arcade	DDR	Yes	No	No	No	No
Research Project	Framework	[Mueller et al_11]	Yes	Yes	Yes	No	No
		Dual flow [Sinclair et al_09]	Yes	Yes	Yes	Yes	No
		PlayMate! [Berkovsky et al_09]	Yes	Yes	No	No	No
		Personalization of exergames [Göbel et al_10]	Yes	Yes	Yes	No	No
	Mobile game	PiNiZoRo [Stanley et al_10]	Yes	Yes	No	No	No
		Health Defender [Garcia and Coulton_09]	Yes	Yes	No	No	No
	Hand gesture Interface	[Kostomaj and Boh_09]	Yes	No	No	No	No
		[Varona et al_09]	Yes	No	No	No	No
		[Li et al_11]	Yes	No	No	No	No
	Haptic Interface	3RD Space Vest [TN Games_11]	Yes	No	No	No	No
		VI-Tennis [Morelli et al_10]	Yes	No	No	No	No
		VI-Bowling [Morelli et al_10b]	Yes	No	No	No	No
	Smartphone Interface	Simple motion detection [Kiili and Merilampi_10]	Yes	No	No	No	No
	Bodily interface	Virku [Mokka et al_03]	Yes	No	No	No	No

Table 2.1 Summary of related work and requirements.

3 Threshold of human tolerance to inter-stream synchronization skew in a multimedia application

3.1 Introduction

When considering the design and development of a multimedia system where multimodal natural interfaces intersect with complex game engines that include algorithms for adaptation and context awareness, there is a requirement to maintain the quality of experience of the multimedia system by keeping a synchronized flow of all the media channels converging in such system or game.

Typical media channels found in videogames can include audio, video and possibly others such as haptics. Because of this, before embarking on the design and development of such system, we must first understand the relationship between media channels and the synchronization issues that might arise, particularly in scenarios where the game could be deployed in a networking environment. This is particularly true for haptic applications where the haptic media units must be acquired and transmitted independently from the other media streams and are more sensitive to network delays, decreasing the quality of experience on the application.

Hence, it is necessary to understand the threshold of inter-stream temporal asynchrony (against haptics) that humans can withstand before having a perception of a delay in one of the streams. Then, that threshold should guide the design and implementation of the related multimedia system.

Previous work can be classified in two main areas. Some have investigated the effects that skew, due to network delays, can have on the quality of experience of the resulting multimedia application [Tatematsu *et al*_10] [Fujimoto *et al*_08]. Others have investigated the threshold of perception by using physical and digital means. Among those works, the majority have focused on Video/Audio skew alone [Dixon and Spitz_80][McGrath and Summerfield_85][Allan and Kristofferson_74][Steinmetz_96]. Some others analyze the perception of Haptic/Video synchronization skew [Miyasato et al._96] [Vogels_04] [Kameyama and Ishibashi_06][Jay et al._07]. And a few have studied Haptic/Audio skew [Levitin et al._00] [Adelstain et al._03].

However in this work we empirically intend to define the threshold of delay at which a user will perceive asynchrony of audio and video with respect to haptic in a multimedia application, using machine haptics devices. By studying the related work in this field, we conclude that understanding the perceived delay between the different media streams including haptics is an important issue. To the best of our knowledge, there is no direct study that tries to quantify the maximum threshold of perception of asynchrony between both, audio-haptic and video-haptic, using the same experimental setup and a software controlled haptic device such as Novint Falcon or the PHANToM.

Although there are certainly quite some studies that concern with the asynchrony of media streams involving haptics and another stream (audio/video), we argue that those works are focused on a single pair of streams at a time, or when addressing both pairs (haptic-audio, haptic-video) the focus is on the quality of experience or how the network delay affects collaborative work, often for a particular type of multimedia application.

Our work is different to in that we focus on finding a threshold of perception of asynchrony, independent of the application and using machine haptics and considering both types of modalities (haptic-audio, haptic-video) within the same experimental setup and conditions. We argue that this can give consistency to the results and our findings can be used for the design and implementation of synchronization algorithms where the maximum error to which the algorithm should perform can be based on a reference value similar to what we found to be the perception threshold.

3.2 Experimental Setting

For the purpose of evaluating the perception of delay between different media streams, we designed a simple multimedia presentation setup for the subject to perceive audio, video and haptic channels simultaneously. In this setup, the user is only required to passively perceive the sensory information, and is not required to actively interact with the computer. Our experimental setting addresses three components: visual, audio and haptics.

For the visual component we have a screen in front of the user, where they can see an animation of a ball bouncing up and down between two horizontal rackets each represented by a rectangle. The ball travels at a constant speed starting from the lower racket. Then, when the ball reaches the upper racket, it bounces back in the same straight vertical trajectory to the lower racket. The whole animation resembles the typical ping-pong game, except that the rackets are static in the center and the ball autonomously bounces between them. We purposely conceived this simple scenario in order to keep the participant in a “passive” mode where they are focused solely on the perception and not in any particular task, while still having all the typical variables involved in haptic enabled multimedia applications.

For the audio component we asked the users to wear earphones, which allow them to hear the playback of an audio sample that applies to the collision event between the ball and the rackets. The audio sample consists of a typical console beep sound with duration of 300 milliseconds and without any fading effects. For the haptics component, we have set up a haptic device next to the subject. They hold the device grip as if holding the ping-pong racket, while the animation plays on the screen. For each event where the ball collides with the lower racket, the device applies a soft descendant force feedback into the hand of the user, giving them the touch perception of that impact.

3.3 Technical Setup

The experiment was conducted using two haptics devices: The Falcon Haptic Device [Novint_12] and the Phantom device [Phantom_12] at the Multimedia Research Lab (MRCLab) of the University of Ottawa. The game was developed with C++ programming language and the CHAI 3D [Conti *et al.*_05] libraries to create the haptic stimuli. Figure 3.1 depicts the technical setup.

The experiments run on a PC with Intel Core 2 Quad processor, 3 GB in RAM and a graphics card Nvidia GeForce 8800 GTX. The Falcon is an off the shelf device and is used mostly for video games. However the sensitivity and accuracy in terms of resolution is not as high as the Phantom device which is almost eight times more expensive and mostly used in research laboratories where the requirements for reliability of the device are higher (Phantom device reports >450dpi in nominal position resolution whereas Falcon device ~ 400 dpi). We decided to use the two different types of haptic devices as a way to validate our findings and to investigate if the differences in hardware can make an influence on the perception of asynchrony for any of the modalities being tested.

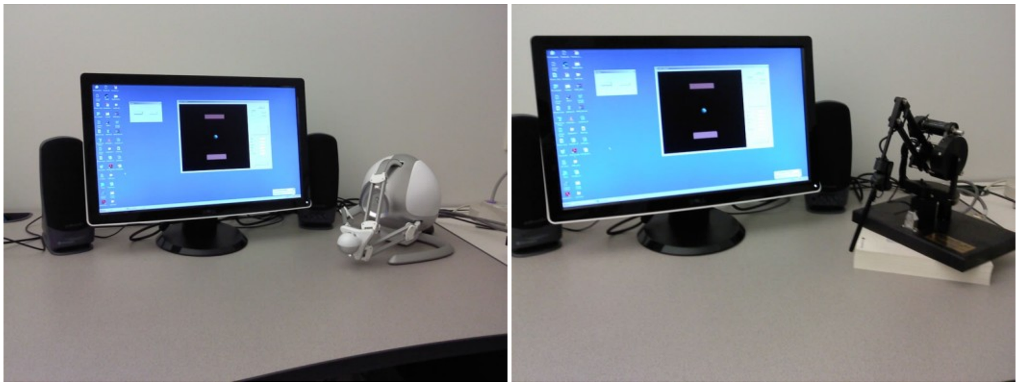


Figure 3.1 Technical setup with Falcon (left) and Phantom (right) devices

3.5 Methodology

The main assessment of the experiment was to determine subjects' perception of asynchrony for collision events between the ball and the lower racket and to see how the two technically different haptic devices affect the perception.

3.5.1 Characterization of a phenomenon driven by sensitivity

The analysis of the data obtained from a subjective phenomenon, such as a person's perception for asynchrony between two media streams for a single event, deals with continuous variables that cannot be measured in practice and that relay on proper placement of observations using parameters of estimation.

Our particular experimental investigation deals with negative or positive delays on audio or video streams with respect to a haptic stimulus, which users may or may not perceive according to some ambiguous factors such as the physiological and psychological conditions of the subject at the time of the experiment. The approach undertaken fundamentally focuses on haptic to video and haptic to audio skew in order to identify the threshold of tolerance to real inter-stream asynchrony before a "perceived" asynchrony is detected by a user. Consequently we assume there is a critical delay (threshold) associated with each subject, and that a subject will perceive an asynchrony when the value of a skew between two streams is greater than the threshold, either positive or negative. Equally the user will not perceive such asynchrony when the positive or negative delay is lesser than this critical delay (threshold). The population of users is thus characterized by a continuous variable - the critical delay (threshold) - which we intend to measure in a subjective empirical way.

Because of the characteristics of the data to be analyzed, we required a methodology for proper parameter estimation and efficient placing of observations. An adaptive procedure commonly called "up and down method" is one in which the stimulus level on any one trail of tests is determined by the preceding stimuli and responses (Levitt 1970]. Up-down testing methods are a subset of a broader class of testing procedures generally known as sequential experiments, and their variations have received much attention in many research fields such as psychophysics [Levitt_70, Dixon and Mood_48].

3.5.2 The "Up and down" methodology

The conceptual idea behind the Up And Down method is to choose some initial step, in our case, this initial step is a reference point localized at zero delay (both media streams are synchronized), represented by d_0 and then a sequence of fixed steps of +20ms (increased delay) represented by d_1, d_2, d_3,... above d_0 in the case of positive delay step, and another sequence of steps of -20ms represented by d_{-1},d_{-2},d_{-3},... below d_0 in the case of a negative delay representation. The first step is tested at d_1 (by increasing the delay +20ms from d_0). If the first delay was detected by the subject, the second step is tested again at d_0, otherwise the second step will increase by 20ms to define at d_2 (meaning the 20ms from d_1 step plus another 20ms). In overall, any delay step is tested at the level immediately below or immediately above the level of the previous test according to whether there was or not a perceived stimulus delay on the previous test.

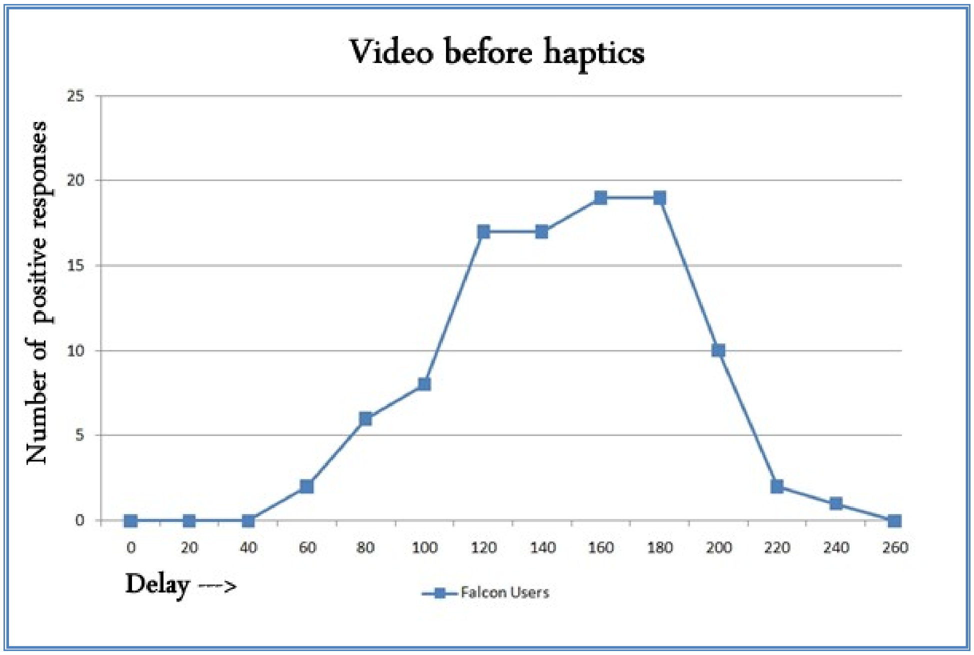


Figure 3.2 Record of sample data of video before haptic event test

Counting the total number of positive (perceived delayed) responses for each delay step applied in the experiments, we would end up with one table of totals for each pair of streams being evaluated in a given delay mode. As per the methodology applied, the analysis of the collected data is done in the context of either positive responses or negative responses (but not both), depending on the focus of the study [Dixon and Mood_48]. In our case we focused in finding the threshold of perceived asynchrony and hence it is the `positive` responses from the users that we consider. To illustrate this, Figure 3.2 depicts a graphic derived from such a table of totals. In this sample, we see results from responses given in the experiment when judging asynchrony between video and haptics, having the video stream ahead of haptics, and using the Falcon device. The horizontal axis shows the step (amount of delay introduced) interval at which responses were recorded, and the vertical axis shows the number of positive responses (perceived asynchrony) reported at each step interval. In this graph we can observe the number of positive responses that the subjects gave for each delay step applied.

For this particular sample, it can be observed that subjects did not generally perceived skew between streams when the introduced delay was between 0 and 40 milliseconds (no one gave a positive answer). Then in the area between 60 ms and 220 ms the number of positive responses varies across that range, which indicates that some subjects perceived and others did not perceived the skew at the various step intervals. Such uncertainty would be expected, as perception is different from one person to another and is dependent on the sensitivity of the auditory apparatus of the subject. Finally it can be seen that after 240 ms the number of positive responses again falls to 0 which means that all subjects at this time had already reported perceiving delay. Please note that because the method is based on a staircase algorithm, as soon as a user reports perceived asynchrony the delay step starts decreasing, which is why there appears not to be perceived delay beyond 240, when in reality what it means is that all users reported perception of delay before this point.

The up and down method has the primary advantage of automatically placing the tests near the mean, which when combined with the right statistical analysis, it increases the accuracy when estimating the mean. This means that for a given accuracy, the method will require fewer tests. The only requirement here is that each specimen (in this case an experiment participant) be tested separately, which is not a problem in our case, since the experiment was designed in this way.

There are several methods for dealing with the analysis of this type of sensitivity information as suggested by Levitt [Levitt_70]. One method is fitting the psychometric function using conventional techniques such as probit model, however Dixon and Mood suggest a method that is specifically designed for data obtained when employing an "up and down" methodology.

If interested, we encourage the reader to refer the details and the theory underlying the method proposed by Dixon & Mood, published by the journal of the American Statistical Association, and which is not included in this text. In the context of this work, it should suffice to know that the purpose of the method is to obtain a reliable mean and standard deviation from data obtained using the "up and "down" method for measuring sensitivity levels, such as the perception of individuals to an applied stimuli [Dixon and Mood_48].

In this experiment the values of the means for each modality being tested (haptic ahead of video, haptic behind video, haptic ahead of audio and haptic behind audio) were subject to a T-Test to assess if the means from the group using the Falcon device had a significant difference when compared to those of the group using the Phantom device.

3.6 Procedure

By using the experimental setup and the methodology described above, we evaluated the asynchrony of two media channels at a time. We had four different variations of delayed synchronization. Two in the Haptic-Audio stream with audio ahead of haptics, and audio behind haptics, and two in haptic-video stream with video ahead of haptics and video behind haptics. In each case the third channel that was not being evaluated was turned off. Our experiment was based on the up-down procedure, using a staircase algorithm [Levitt_70].

With this approach, a delay is introduced and increased at fixed step intervals on one of the two media channels under evaluation. For each step, the subject is to judge if there is a perceived delay between the two media streams being examined, by saying out loud “sync” or “not sync”, depending on whether they perceived synchrony or asynchrony among the media channels for a collision event. When there is a positive response (delay perceived), the delay will change direction and start to decrease at the same rate. When there is a negative response (delayed not perceived), the delay will change direction again and start to increase, and so forth. Each step evaluated by the subject is considered a trial, and a series of consecutive steps in the same direction is considered a run. The point at which there is a change in direction of the steps is the reversal point. The staircase loop was designed to do a set of 8 runs.

We performed a pre-experiment session to try out the entire setup. Some volunteers went through the entire process of the experiment as a general test run. We wanted to have an initial test of our methodology and to establish how conformable users would be with the multimedia presentation setup (ping pong game and haptics devices). It also helped to determine the step size (time increment in milliseconds) of the delay change for the experiment. Choosing the step size was critical and required some empirical trial and error. We did not want to use a too small size, since the stair case algorithm would render the experimental session too long and tiresome for the user. Also we did not want to use a large step size because that would make it difficult to pinpoint the threshold of perception with a good degree of accuracy. In this pre-experiment session we initially tested with a step size of 40ms in order to establish the general range of human sensitivity to a change in the skew and refined the number with some trial and error. As a result of the pre-experiment, we decided to use a step size of 20ms when conducted the experiments as described below.

For the experiment, users were asked to go through the multimedia presentation, and alert the experimenter whenever they felt comfortable with the haptic feedback and the overall multimedia experience in order to begin the test. Once they began the test, they were left to pay attention to the presentation, while the time delay steps were introduced according to the staircase algorithm. Subjects were instructed to say out loud “Sync” or “Not sync”, depending on weather they perceived the collision event as synchronized or as not-synchronized, judging the synchrony considering the 2 media streams evaluated at the time (either haptic-audio, or haptic-video). The experimenter noted the response into the computer, using a control panel that was displayed on a second screen monitor only visible to the experimenter.

The used up-down method in our experiment is best suited for estimating the middle point of the psychometric function. For this, it is required that observations are placed close to the focus of interest (the threshold) and hence there is a need to have a step size that is not too large (observations would be too far from the threshold) and is not too small (many observations are wasted in converging to the focus of interest). Our initial assessment was done for this purpose. We tested with a relatively large step size (as compared to previous work 60 msec), and then started reducing the step size by 10 msec down, i.;e. 50 msec, 40 msec, 30msec, 20 msec, 10 msec. As expected when the step size was very large (60, 50, 40 msec, etc.) the eight reversal points were reached very quickly and the resulting data was not very good at indicating where the threshold of perception was located. In the case of the small step size such as 10 milliseconds, the experiments were extremely long and tiresome for the users. After this heuristic approach we concluded to use the 20 milliseconds value that provided a good balance between the duration of the experiment and the accuracy on finding the threshold.

Every subject went through four sets of 8 runs in the staircase, combining both media types (audio and video) and both initial delay modes (one media before the other and vice versa). The size of 8 runs (or counting 8 times the reversal points) is based on evidence found by Levitt and Wetherhill (Levitt, 1970, Wetherhill and Levitt, 1965). They argue that the up-down method can outperform other methods such as the method of limits (which terminates after the first reversal point) and they recommend pursuing the experiment and continuing the procedure until either 6 or 8 reversal points are obtained. We decided to go for the 8 reversal points.

Subjects were unaware of which media was coming before or after for each set of the 8 runs because the order was randomized. In every case, the length of each set of runs was only limited by the user's responses in relation to their perception of an asynchrony between the two media streams. In total, 25 subjects participated in the experiment. 10 participants used the Falcon device, and another group of 10 used the Phantom device. Five additional participants performed the experiment using both devices (one after the other). In this fashion, a total of 30 experiments were conducted, 15 on each type of device.

3.6.1 Control of the Media Stream Skew

For the purpose of evaluating the perception of asynchrony between the media streams, we have introduced delay to present both, video and audio, before and after the haptic channel respectively. The media synchronization skew introduced into the audio and visual cues was handled within the presentation code itself. The presentation would start with no skew at all. That is, the streams are synchronized on the collision events. Then, an increment or decrement of 20 milliseconds is introduced every time the subject gives an answer, regarding his perception of synchronization with the current delay. The same applies for the visual to haptic skew. The presentation randomly selects which media will come first at the beginning of each run.

3.6.2 Subjects

All 25 participants were graduate students in the School of Information Technology & Engineering at the University of Ottawa, with an age ranging between 25 and 35. 19 subjects were male while 6 were female participants. They all reported having normal hearing and no neuro-motor impairment. They all were also unaware of the details of the experiment before their participation.

3.7 Discussion & Results

We first briefly present and discuss the results of applying the methodology presented above to the data for each of the four different types of tests, haptic ahead of video, haptic behind video, haptic ahead of audio and haptic behind audio. For each type of test we present the estimated values of μ and σ according to the type of used haptic device. We also present the p value to show if there is a statistically significant difference when comparing those values from one device to another. The value of p was calculated by performing a t-test in which the mean of a normally distributed population has a value specified in a null hypothesis (general or default position). We reiterate the usage of two different types of haptic devices, the phantom and the falcon. The former is characterized by its high sensitivity and is mainly used in research laboratories, while the second is consumer product mainly used for games. We depict and demonstrate the regions of annoyance according to the data recorded from both devices.

The right side of Figure 3 shows an overview of the results of the perceived Audio-Haptic synchronization with a positive delay (Audio after haptics). The horizontal axis shows the delay steps (20 ms per step), and the vertical axis shows the total number of positive (asynchrony is detected) responses reported for that step, which are the ones that directly reflect the detection rates, and on which we applied the statistical analysis method.

In this modality the application of the method shows a mean of 123 msec. and a standard deviation of 95 msec. for the Falcon device. Meanwhile for the Phantom device the results were an average response of 97 +/- 26 msec, which were not significantly different (p-value = 0.08) from the results presented on the Falcon device.

As it can be seen on the left side of Figure 3.3, when audio cue was presented before haptic, there are also interesting observations that can be discussed. For this modality the values of the mean delay and the standard deviation were 70ms and 143ms for the Falcon device, and 115 msec. and 80 msec. for the Phantom device. In this case there seems to be some meaningful difference, and the regions of perception are displaced if the two devices are considered, however the difference is not statistically significant (p-value = 0.105). Still just for sake of discussion we would like to point out that Phantom device users seem to be keener to immediately perceive some asynchrony as early as -20 millisecond delay, while Falcon users did not perceived any until after 80 milliseconds. This could possibly be attributed to the precision with which the haptic stimulus can be delivered from the device. The phantom device tends to give a more sharp feeling, while the Falcon device has been found to have a higher damping effect which may decrease the proper discrimination of applied forces [Lode et al._10]. This effect could be related to the apparent early detection on the Phantom.

Please note that the Y axis label reads “Positive Responses”. This is not the number of users that participated on the experiment, but an addition of all the positive responses reported, across all users for each step interval. This addition value can, and actually varies from one modality to another. This is because each user can potentially give different responses for the same step interval in any of the 8 runs. We have focused on the addition of the positive responses because those are the one representing the detection of asynchrony and it helps us to better illustrate on the graphics the boundaries for the area where the participant is not annoyed. That is why Figures 3.5 and 3.6 below show greater numbers on the Y axis, since it is adding up the bulk of positive responses for both devices.

Figure 3.4 (right) shows the results for video before haptics. We notice that in this case the introduced delay is positive, as opposed to the “audio before haptics” modality. The reason for this is that for practical purposes, the algorithm would actually introduce the delay into the haptic stream. So in this fashion a positive delay on the haptic stream gives an effect of having “video before haptics”.

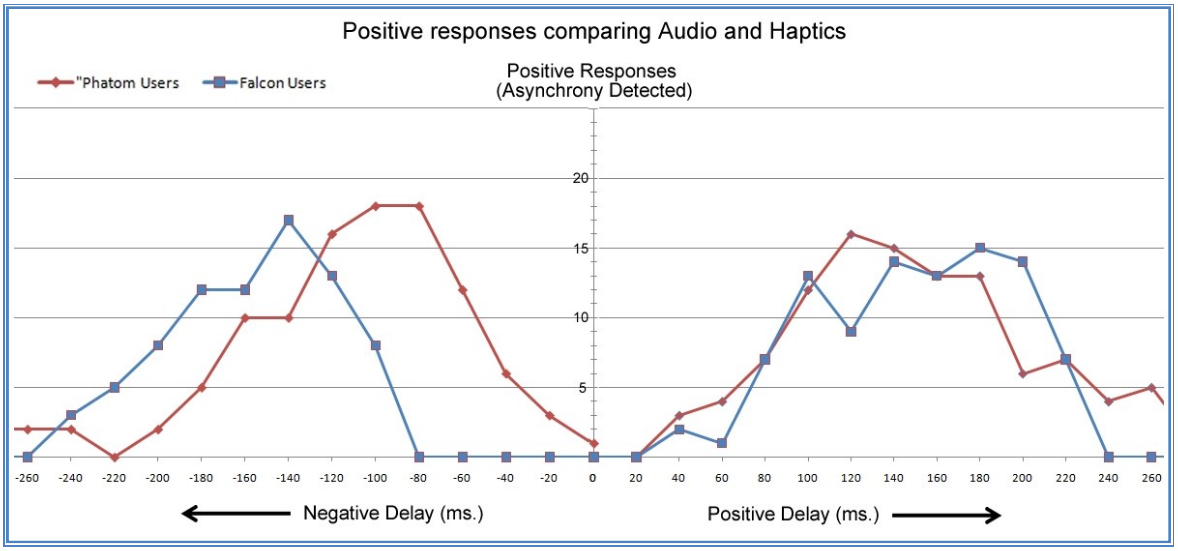


Figure 3.3 Perceivable delay for audio before (negative delay) and after (positive delay) haptics.

In this case it can be seen that the total of responses given for each step size are quite similar for both devices, and do not have a significant difference (p-value = 0.624). In both instances detection rates begin at around 60 msec. delay, and the perceivable annoyance interval is between 120 and 190 msec. The mean and standard deviation in these instances were calculated to 147 msec. and 57 msec. for Falcon device and 102 msec. and 66 msec. for the Phantom device, respectively.

Figure 4 (left) shows the case where video comes after haptics. Here, this effect is achieved by introducing a negative delay in the haptic stream. In this case, the average of noticeable annoyance for the asynchrony between a haptic event with reference of visual effect was -79 msec. and 32 msec. for users using the Falcon Device meanwhile the average response bias was of -95 +/-77 ms which was not significantly different from Falcon users (p-value = 0.088).

Figure 3.4 shows a continuous representation of the perceivable delay when comparing Video to Haptic. This shows a continuous flow of the frequency of positive responses for both types of delay, positive and negative. In the left it can be seen the case where the haptic event happened before the video, and in the right the case where haptics comes after the video.

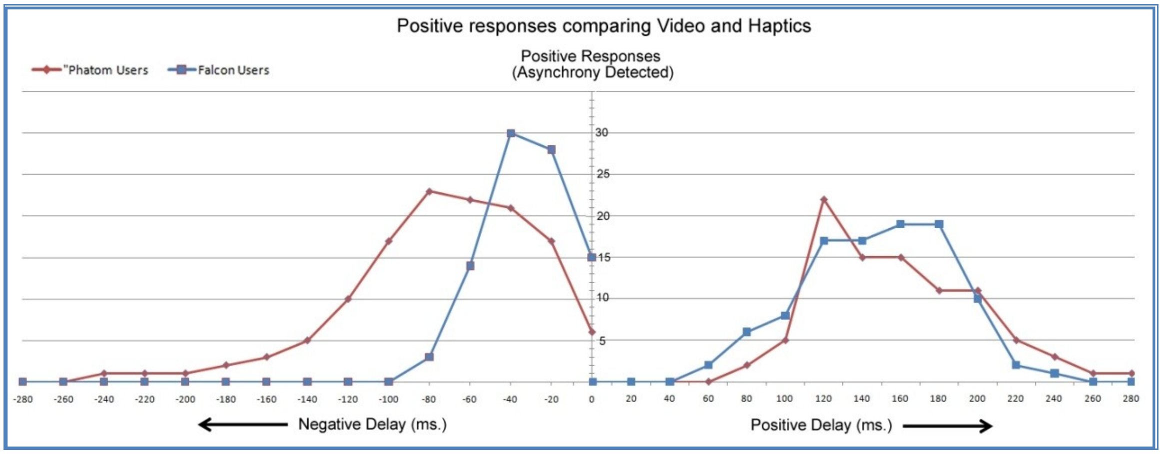


Figure 3.4 Perceivable delay for video before (positive)
and after (negative) haptic

Although the data presented before can be interesting when comparing one device against the other, we have further combined the results of both devices in order to provide a more general representation of the overall shape of perception responses. Again, it should be noted that the lack of positive responses for large values of delay means that the users perceived the delay before that large delay value. Hence there is a point where there will be no count of positive responses given that the stair case algorithm will not ask any further in the same direction after the user has reported perceived delay.

Figures 3.5 and 3.6 depict the result of combining data from both types of device and both types of delay (positive and negative) for the Haptic-Audio and Haptic-Video experimental modalities respectively. The shaded area of the graphic represents the noticeable asynchrony, meaning areas where there were some positive responses (perceived delay). This shaded area is presented for both sides, the negative and the positive delay (audio before and after haptics respectively). The horizontal axis shows the amount of delay introduced into the audio stream, while the vertical axis shows the total number of positive responses. The valley in the middle represents the range at which little or no noticeable delay was reported by participants.

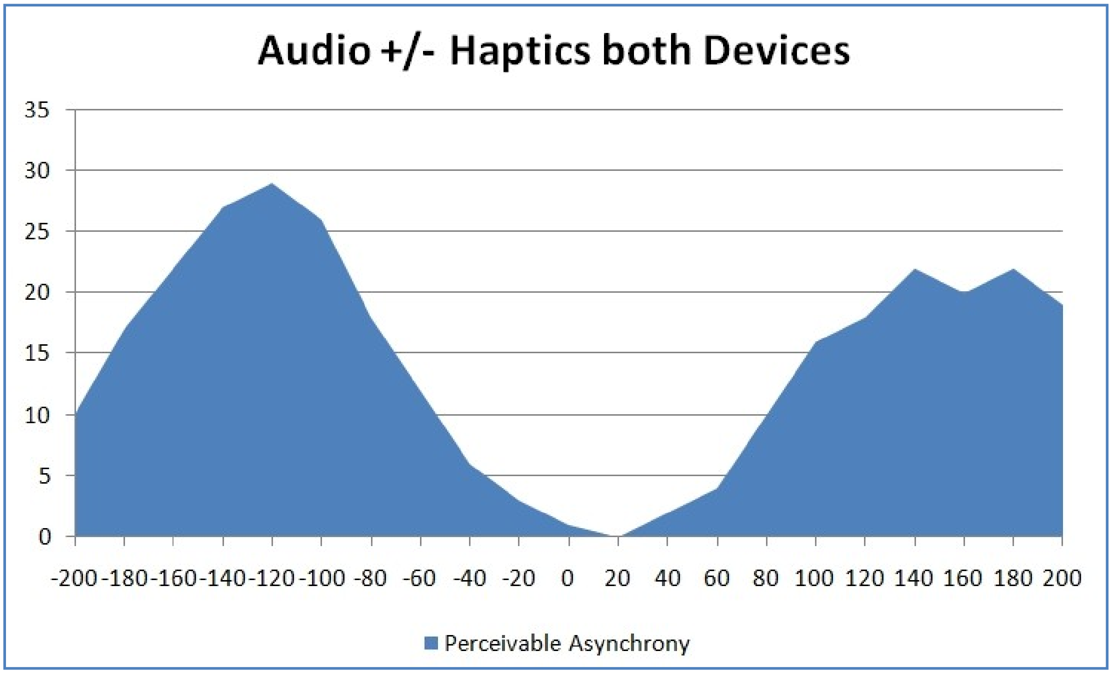


Figure 3.5. Areas of noticeable asynchrony between haptic and audio streams

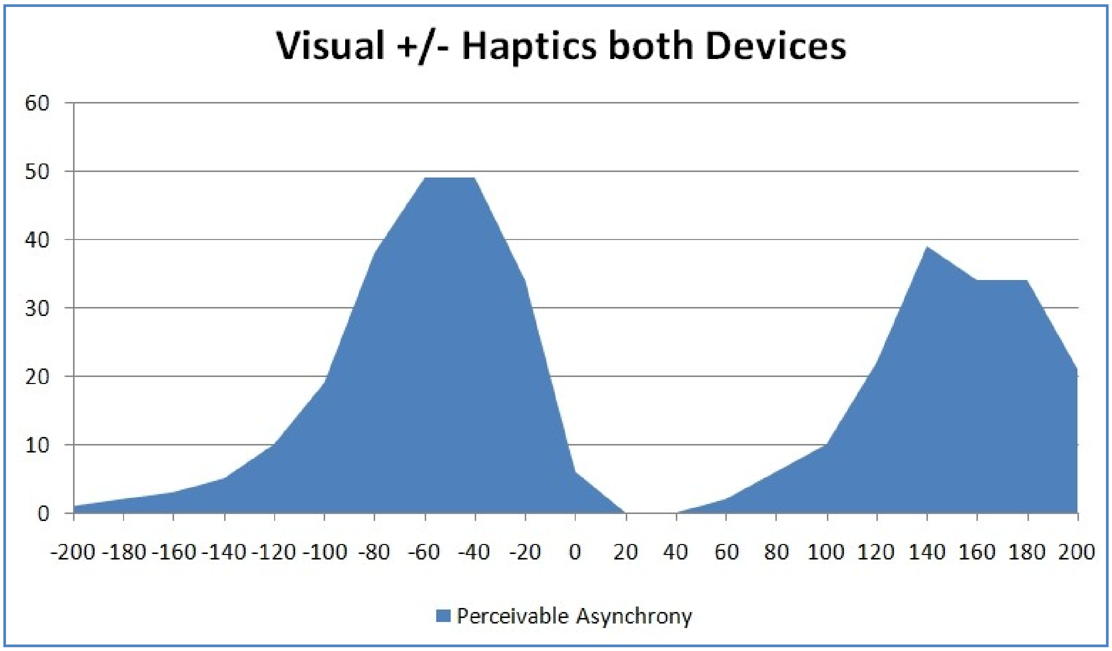


Figure 3.6 Areas of noticeable asynchrony between haptic and video streams

From both graphics, it can be seen that users are generally more sensitive to having a negative delay on the secondary stream, when compared to the reference stream. This is shown by the steep sides of the shaded area, close to zero delay shown on the negative delay side of the graphics. When taking a reference point in the vertical axis, such as 10 responses of perceived delay, it can be seen that for audio before haptics, those 10 responses correspond to a delay of around -60 msec., while the value is about 80 msec. when audio is after haptics. A similar trend is observed for the haptic-video modality with a range of about -20 to 100 msec.

One thing to note is that on Figure 3.6, there are some responses reporting perceived delay at delay 0. Because of the nature of the up-down staircase algorithm and the sensibility to perception of delay even at low intervals, some user would go further down on a run when giving responses, before realizing that there is in fact no delay anymore and correct the course (go up again). However given our current methodology analysis, these responses should not have an effect on the overall normalized delay (μ) values reported.

3.8 Summary

We have evaluated the human perception sensitivity to asynchrony between two media channels within the context of a multimedia environment. Our approach has been to keep the haptic channel on the middle and assess the synchrony of visual and audio cues with respect to the central haptic reference. Haptic to Audio and Haptic to Video assessments have been performed with 25 subjects using two haptic devices. For each assessment, we have evaluated the positive and negative delay (before/after haptics) of one of the streams (audio or video) with respect to the reference stream (haptic). To the best of our knowledge there are no previous studies that include all modalities (haptic-audio, and haptic-video) in a purely asynchrony perception investigative approach, in the context of the emerging field in multimedia applications with machine haptics.

Results show a high degree of sensitivity to perceiving asynchrony between the media channels. On average, when users are assessing audio to haptics, they report detection of asynchrony at 110 ms when audio is presented after haptics event and at 92 when audio is presented before haptics. In the assessment of video to haptics, detection threshold is on average found at 125 ms when haptics come after video, and at 87 when haptics come before video.

We have found that detection points are not symmetric between positive and negative delay. Users are usually keener to detect asynchrony when the event on the affected stream comes before the event of the reference stream. It is interesting to note on this regard that, although for the purpose of our experiment the haptic stream was the reference stream, in the case of haptic to video evaluation, some users (around 5) reported to intuitively consider the visual cue the reference point to evaluate the synchrony of the channels.

When using one device or the other, although a slight shift on the detection points (possibly due to the sensitivity inherent to the device) occurs, our experiments show that the detections happen in the same pattern for both of the devices. Furthermore, our study shows that even a low-end, of-the-shelf haptic device like the Falcon has the same general performance in terms of how delay in audio or video affects the perception of asynchrony of the user, when compared to high-end devices such as the Phantom device.

It has also been found that few users may detect delay at a time where no delay was actually introduced on any of the channels. We have speculated that this may be caused by the inherent properties of the up-down staircase algorithm used, where users may continue one or two steps (20 ms each) in the same direction (in this case down), before realizing that there is no delay between the streams. Hence, a few of the responses report delay when there was none.

Given those average values, we can see that perception rates are generally around 100 msec (in fact 92-110, 87-125 msec) regardless of modality and type of device . This means that participants were relatively sensitive to temporal delay between the media channels. These results can be use as guidelines to meet a target quality of service for any type of multimedia application where synchrony is vital for the performance of the task, including games that handle multiple media streams.

The architecture proposed in this thesis, and the derived adaptive game implementation, support multimodal interfaces that involve all three media channels subject to this study: Video, Audio and Haptics. Hence, our findings were used as a guideline in the design and development of the components described in the following chapters.

4 An Adaptive Approach to Exergames with Support for Multimodal Interfaces

4.1 Introduction

In this chapter we present the design principles and architecture for the development of an adaptive game engine for exergames. In our approach to exergames, games are based on an adaptive, context-aware engine that changes game dynamics to demand the appropriate amount of exertion from the user. We propose an architecture that considers the player's contextual and physiological sensory variables and how they can be used to affect the game dynamics, which in turn should prompt a physiological response from the player. These adaptive properties are complemented by the ability to addition multimodal natural interfaces that contribute to the entertainment experience and also contribute to fulfill the exertion requirements of a player.

4.2 Functional requirements

The functional requirements for a game derived from our architecture have been defined in direct relationship with the requirements identified from the scenario presented in Chapter 1. Next we present each one of these functional requirements.

User context management: Since we are working with adaptive exergames, the context of the user is of vital importance. The age, gender, overall fitness and any other relevant demographics and physical environment properties of the context of the user should be accounted for, stored and organized for reference and retrieval.

Estimation of level of exert: Given the goal of adequate exertion for the user, we need a mechanism to estimate the physical effort that the user is performing at all times. For this purpose there is a need of physiological sensors to be attached to the user which provide appropriate feedback about parameters such as heart rate, respiration rate, body temperature, blood pressure, level of oxygen and any other parameters that might be relevant to specific game implementation. Furthermore there must be an intelligent algorithm that receives all such inputs and computes a value which reflects the level of exertion.

Knowledge base management: In the process of making decisions across the entire system, there is a need to have a reference knowledge base that dictates the limits and boundaries for all the parameters being adjusted as the game runs. In such knowledge base there must be information that maps the user context to specific value ranges for each one of the parameters being measured and managed in the system. An example of this knowledge based could be a chart of heart rate values and target zones for aerobic exercise, given by the age and physical condition of the users.

Adaptation strategy: There must be an adaptation strategy. This is an algorithm that dictates the changes that ought to be made to the game parameters, based on the other system estimations in order to prompt the desired physiological response from the user. Such decisions will be guided by the desired level of exertion. The adaptation strategy acts as a control entity guiding the users to reach their optimal exertion level and keeping them there for the duration of the exercise/game session.

Retrofit of natural interfaces: A computer videogame is designed to respond to typical input devices such as mouse and keyboard. This kind of input devices is the one less suitable for exertion interfaces in videogames, given the fact that exertion game players are not sitting in front of a desk. In our design, the mouse and keyboard are replaced by one or more natural interfaces based on sensors and devices that track the player's movements and intentions to operate the game with an ubiquitous approach.

For this, there is a need of a custom application that addresses interaction with the specific sensors being used for the natural interface, and a host application that can translate these interactions into actions that would normally be triggered by mouse, keyboard or joystick for the game in question.

For this to work there is a need for a communication channel to support the transmission of data between the natural interface application and the host application. Given the objective of having the devices act as ubiquitous interfaces, a wireless communication is chosen.

Sensor management: The estimation of the level of exert can be given by one or more physiological variables that ought to be sensed and interpreted in order to make an appropriate estimation. Such sensors can have different programming interfaces. Our architecture defines an abstraction layer to interact with such low level sensors such as a heart rate monitor, a perspiration sensor or a respiration rate vest.

4.3 Adaptive exergame architecture

Figure 4.1 depicts the different components of the proposed architecture. Following the picture we describe each component and how it relates and addresses the problems stated by the functional requirements identified above.

In the design of our architecture we have kept in the middle the key idea of separation of concerns. Each one of the components deals with a very specific problem related to the system implementation. This approach is aimed at facilitating technical implementations where the realized system would have loosely coupled components which would favour code maintainability and allow for the incorporation of new modules easily.

4.3.1 Prerequisites

The components that are at the core of the architecture are meant to help in the development and extension of videogame into serious games that adapt to user context and support natural interfaces. For such components to work there is the assumption that the components of a traditional videogame are already in place, or are developed with traditional videogame approaches.

At the top we have the User Interface, which is the part of the system the user interacts with. This will typically be a computer screen and related input controllers. The controllers can be typical video game controllers such as joysticks, driving wheels, or other types more related to exergames such as stationary bikes, foot pads etc.

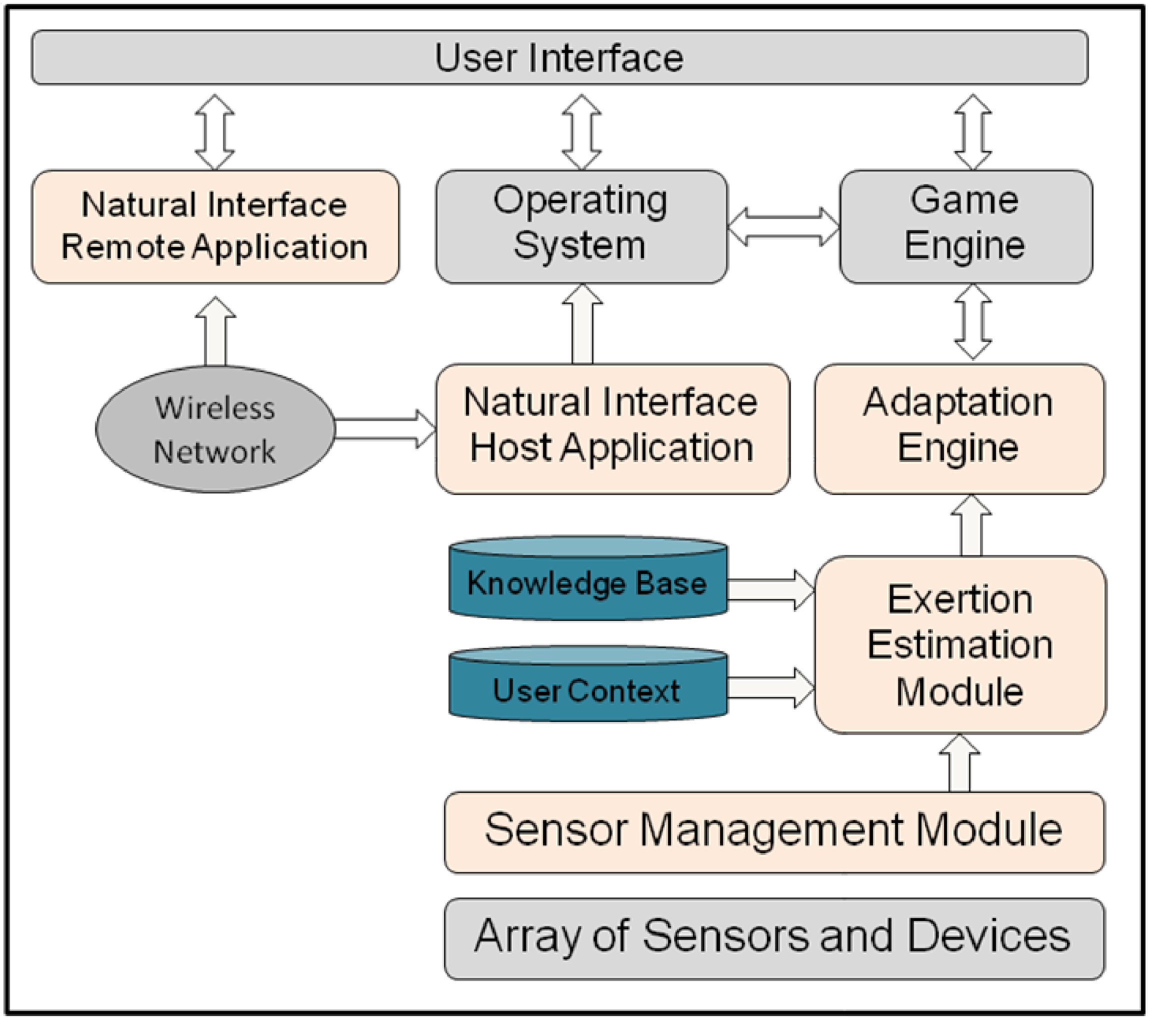


Figure 4.1 Components of the adaptive exergame architecture

Below this layer we have the Game Engine itself. This component encapsulates all aspects of a typical game engine, including graphics rendering, physics emulation, and everything else related to managing the game variables and rules to respond to user's actions. This component by itself does not implement anything related to dynamic adaptation, instead it interacts with the adaptation module by "registering" a set of variables that will be managed by the adaptation module. After registering these variables and specifying some rules for them, the game engine does not need to know anything else of what is going on a lower level. It simply trusts that the "registered variables" will be changed according to the user level of exert to have the game change as needed.

4.3.2 Adaptation Engine

The Adaptation Engine on the other hand is the algorithm that implements the adaptation strategy. The adaptation engine receives a set of variables from the game engine as input. These variables come with a set of rules that define the way they should be changed to increase or decrease the level of challenge and exertion demanded from the player. The adaptation engine takes the rules and applies game changes accordingly, taking into account the level of exert reported by the exertion estimation algorithm. This module can be perceived as a loop which constantly makes a decision to make the game more or less physically demanding, according to how the user is doing and the target level of exertion for the session.

4.3.3 Exertion estimation module

The exertion estimation module is responsible for combining 3 kinds of inputs: The raw data readings coming from the sensor management module, the context information of the player, and finally the knowledge base used as a reference to estimate the level of exertion based on the other two inputs. Using these three inputs, it makes an informed decision as to the level of exert that the player is currently subject to, deciding whether the user should exert more or less.

This module also implements a loop for constant reading from the sensors. It combines these with the information retrieved from the context and knowledge base to make a decision about the next level. Two approaches for this kind of estimation are further discussed in section 4.5.

4.3.4 User context and Knowledge Base

The user context and knowledge base would typically be implementations of a database statically storing all the reference information, although the proposed architecture makes no such assumption and these could easily be some sort of dynamic entity where the data can change all the time or be fed from remote locations, depending on the specific implementation.

4.3.5 Sensor Management Module

At the lowest level we have the Sensor Management module. This is an abstraction that handles anything related to interacting with each one of the physiological sensors that are attached to the player. Normally each sensor would provide a different API for communication and this module makes an abstraction each of those API's and provides a unified interface to read information from them. Even if one particular heart rate sensor is changed to another type, the upper level components need not to change its implementation or even know about the change.

4.3.6 Natural interface applications

This component addresses the requirement to retro-fit new natural interfaces into pre-existing computer games. The premise around this component is that the computer game engine can receive input from the natural interface as if it was coming from supported peripherals such as mouse and keyboard. To support this, the proposed architectures contemplate the use of two components that need to be in place.

The first component is a remote application that is running in close contact with the natural interface to be coupled with the game. This application makes use of whatever API's are provided by the device or sensor that is being used as a natural interaction interface. The application constantly reads the sensory data and, depending on the application logic and the type of interface being implemented, it detects command events that ought to be sent to the host application as input for the videogame. To illustrate this with an example, if the natural interface is based on motion detection with a depth enabled camera. The remote application is the logic that process low level interaction with the camera, and it detects the bodily gestures that are meaningful in the context of the game. Those gestures are then to be transmitted to the host application (i.e. "jump", "duck", "tilt sideways" etc. are examples of actions detected by the remote application).

The second process is the host application. It runs side by side with the game engine. In our proposed architecture, the host application does not directly interface with the game engine; instead, the data coming from the sensors (i.e. attached to a smartphone) are mapped to typical input commands that the game engine can respond to, such as commands from the mouse and keyboard. The host application sends input interruption commands to the operating system so that the game engine can respond to those commands as if they were coming directly from the attached peripheral devices.

This approach allows for the integration of natural exergame interfaces based on a variety of sensors and devices such as motion detection cameras, accelerometers, exertion bicycles and other vascular machines, smartphones, haptic devices and any other kind of device that is required to extend the game with a natural interface.

Although the communication between devices and host application is based on a wireless network, this is only for improving the usability and ubiquity of the derived natural interface. No assumptions are made as for the type of protocol to be employed as long as both ends support the same protocol. Depending on the nature of the interface being implemented, performance may or may not be of concern. Examples of wireless network communication include WiFi, Bluetooth, or even wide area networks such as 3G. However such communication could be established by any other wired means.

4.4 Components interaction

To better illustrate how the components of the architectural design come to realize the end system Figures 4.2 and 4.3 represent the interaction between the components as the game plays.

Figure 4.2 addresses the interactions that happen for the adaptation aspects. The game engine initiates the sequence at the beginning of the game session by registering a series of game variables with the adaptation engine, including the rules that define how those variables should be modified. The game engine then initializes the exertion estimation algorithm, which retrieves context information from the database. Once the initialization is complete, the adaptation engine runs in a synchronized loop where it constantly checks with the exertion estimation process if there is a need for changing the game dynamics. The exertion estimation makes this decision by retrieving reference date from the knowledge database and physiological data from the sensor manager. Using these inputs it processes an estimation of the current level of exertion for the player and makes a decision regarding the need for a change in the game dynamics. Such decision is taken by the adaptation engine, and if a change is needed it applies the adaptation strategy, according to the rules, by modifying the values for the registered variables of the game. Finally the game engine applies the changes to the game dynamics that would prompt the desired physiological response from the player.

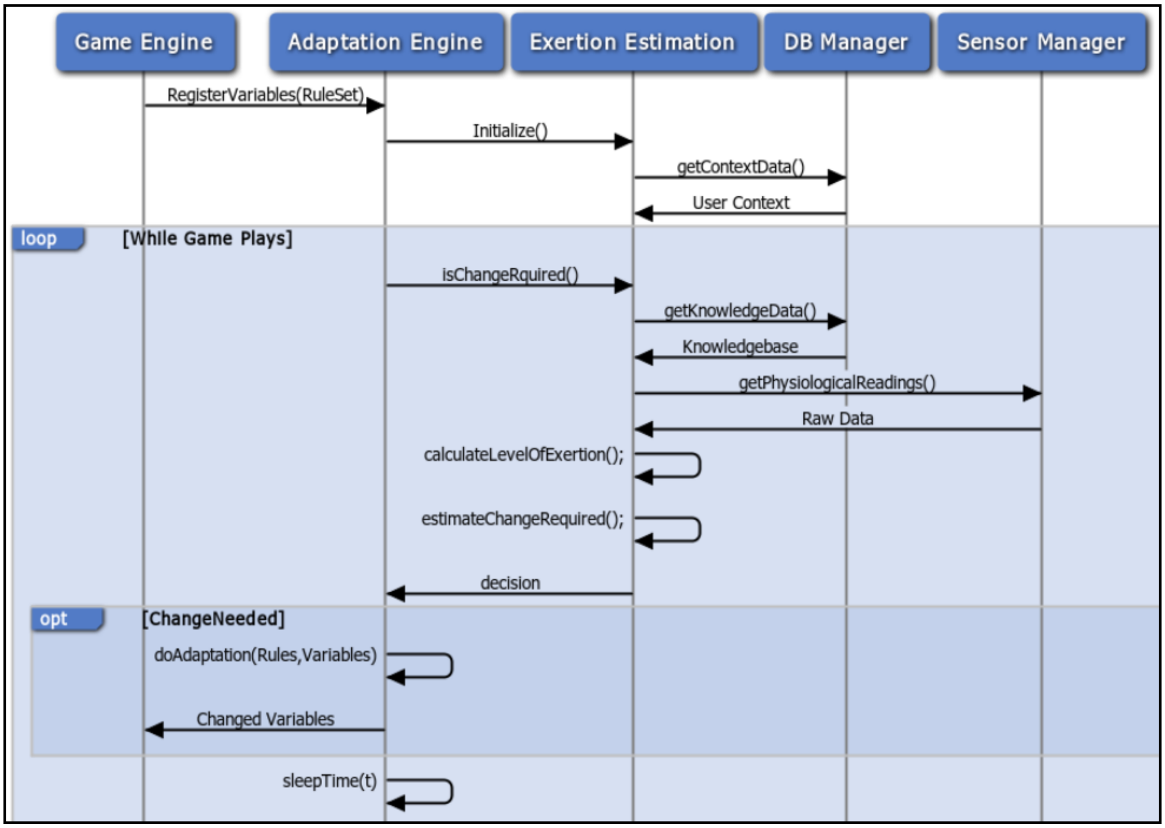


Figure 4.2 Interaction between components
to implement game adaptability

Figure 4.3 depicts a sequence diagram showing the interactions that happen between the different components for the natural interface support. In a typical interaction scenario, the remote application initiates the sequence, requesting the most up to date reading of a particular sensor. The device's API would query the device and return a raw data reading to the remote app. The remote app would then analyze this data and decide if the reading reflects an action as defined for the particular interface implementation, and if this is the case then it would send the "action" information across the network to the host application. If no action was detected or the previous action was already sent, it goes back to continue reading from the sensors and the process is repeated.

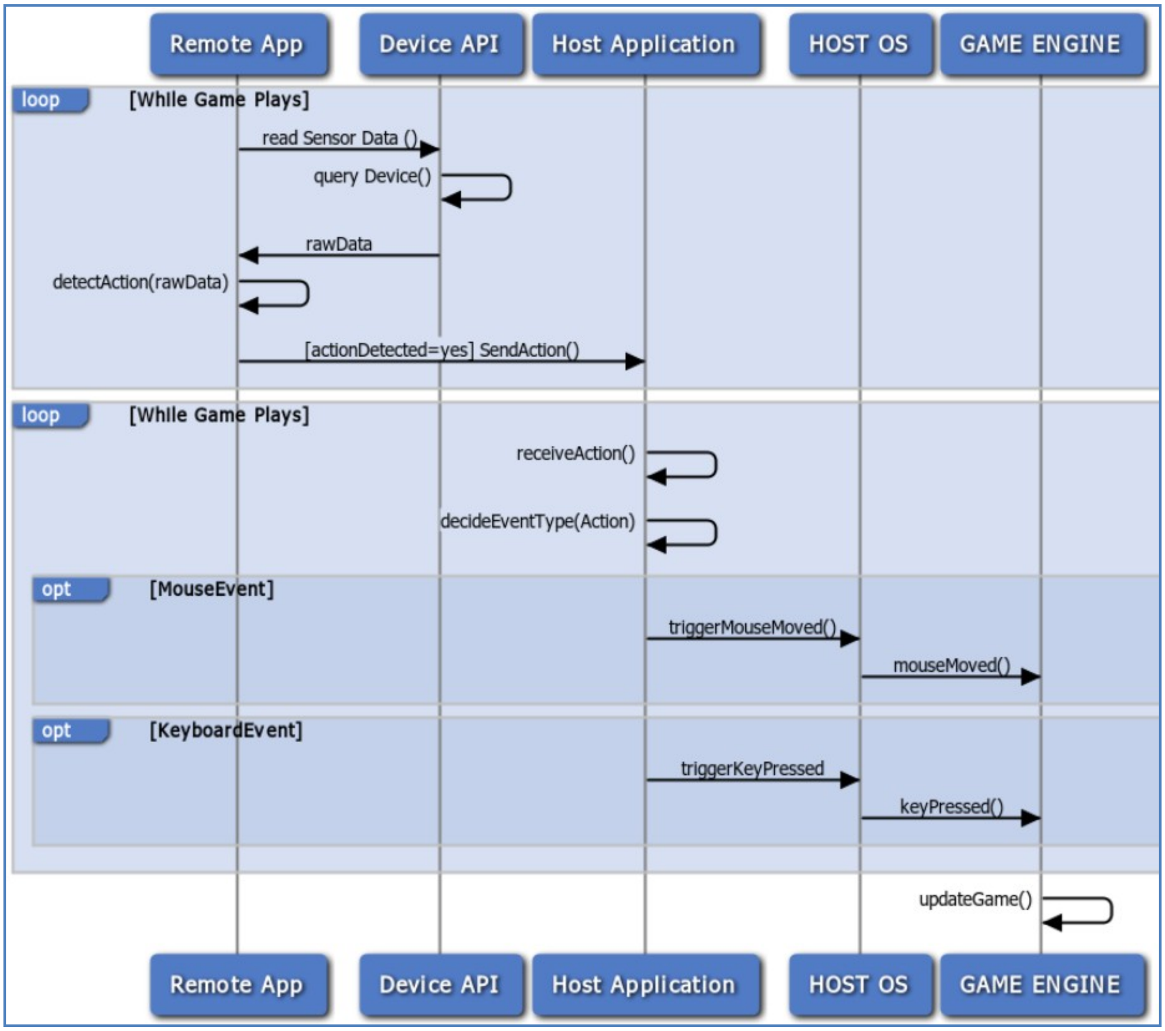


Figure 4.3 Interactions between components to implement a natural interface

The host application expects action messages to come from the remote application. It would then verify the action performed and map it to a specific keyboard or mouse event, triggering the related event in the context of execution of the operating system. Then, the operating system would send an appropriate message to the running game engine, which would respond accordingly, as if the input actually came directly from the keyboard or mouse. This process would be repeated for any action message received by the host application.

From the sequence diagram, it can be seen that there is no need for the game engine to be aware of the newly implemented exertion interface, as it would simply operate with commands coming from the keyboard and mouse, as it was originally programmed. The type of "actions" that the remote application can support is fully flexible with this approach, and only limited by the imagination of the game developer and the sensors available to support the interactions. The Host Application can support also multiple configuration files that define the kind of mapping that should occur for a given triggered action to the intended keyboard/mouse event. These files are meant to be configuration profiles, each of which is linked to a game title, so that the same exertion interface can be used by more than one game.

4.5 Algorithms for estimation of level of exertion

In this section we describe in detail the two algorithms of the exertion estimation module presented as part of the architecture. These algorithms compute the magnitude and nature of change required in the context of the exertion interface to adapt to the context of the player The first algorithm is based on a linear approach, where all game variables are changed in fixed step sizes (up or down) until the player reached the desired target exertion zone. The linear approach is easier to implement, but given the fixed size of the steps it is less adaptable when the player is already close to the target zone. The second one employs a fuzzy logic control technique, which varies the magnitude of change for each variable depending of the error level with respect to the target. This approach is more complex to implement, but it delivers a smoother rate of change of difficulty as the player approaches the target zone.

4.5.1 Linear algorithm for exertion estimation

The linear version of the exertion estimation algorithm makes the assumption that the changes to be made to the registered variables will happen in fixed step sizes. The inputs for the algorithm are the current level of exertion and the target exertion zone values (upper and lower limit).

Linear Algorithm for estimation of level of exertion

```
Initialization:
    1. Let $s_1, s_2, s_3 \dots s_n$ be the readings from sensors used to estimate exertion level
    2. Let $TZ_{min}$ be the lower boundary of the target exertion zone
    3. Let $TZ_{max}$ be the upper boundary of the target exertion zone
    4. Let $\Delta$ be a fixed time that defines the frequency of execution
Begin
    Loop while <game plays>
        For s = 1 to n
            $X_t \leftarrow$ Add weighted value of reading $s_i$
        End
        If $X_t <= TZ_{min}$ Then
            Return 1
        If $X_t > TZ_{min}$ AND $X_t <= TZ_{max}$ Then
            Return 0
        If $X_t > TZ_{max}$ Then
            Return 1
        Sleep for time $\Delta$
        $t = t + \Delta$
    End loop
End
```

Figure 4.4 Steps of the linear algorithm used to drive
the level of exertion of the player to the target zone

The output of the algorithms is defined by a function used to compare the current level of exertion against the target zone identified from the knowledge base. The possible outputs of this function are described in the following equation:

$$\mathrm{CH}\,(X) = \begin{cases} 1 & \text{if } x < \mathrm{TZ}_{\min} \\ 0 & \text{if } \mathrm{TZ}_{\min} < x < \mathrm{TZ}_{\max} \\ -1 & \text{if } x > \mathrm{TZ}_{\max} \end{cases} \qquad (4.1)$$

Where $\mathrm{TZ}_{\min}$ is the lower boundary of the Target Zone, $\mathrm{TZ}_{\max}$ is the upper boundary of the target exertion zone and X represents the current exertion level. The resulting value of CH(X) is one of three possible values (-1, 0 or 1), depending whether the current exertion level is below, above or within the target zone. A positive value indicates an increase in the exertion interface demands and a negative value implies making the game easier. Figure 4.4 shows the general steps for the linear exertion estimation algorithm.

4.5.2 A fuzzy logic control algorithm for an exertion game

Fuzzy logic was first introduced in the field of control systems by Zadeh [Zadeh_72]. The premise is that the introduction of fuzzy sets provides for a non-analytic approach to the traditional analytic control theory. In general, fuzzy control systems are good for situations where there is some uncertainty in the response of the process being controlled.

Controlling an exertion interface (which is composed of multiple actuators and game variables), in such a way that we can obtain a desired physiological response from the player, is one of those cases where uncertainty exists. For any change introduced in the system, the player’s body might respond differently and in an unpredictable way. Furthermore one player’s physical condition is different to another, which introduces more unknown variables to the control mechanism.

In this context, a fuzzy control algorithm provides a mechanism to compensate for such uncertainty about the behaviour of the controlled system (exergame interface and player’s exertion level). Figure 4.5 depicts the block diagram of the fuzzy logic control for an exertion game interface.

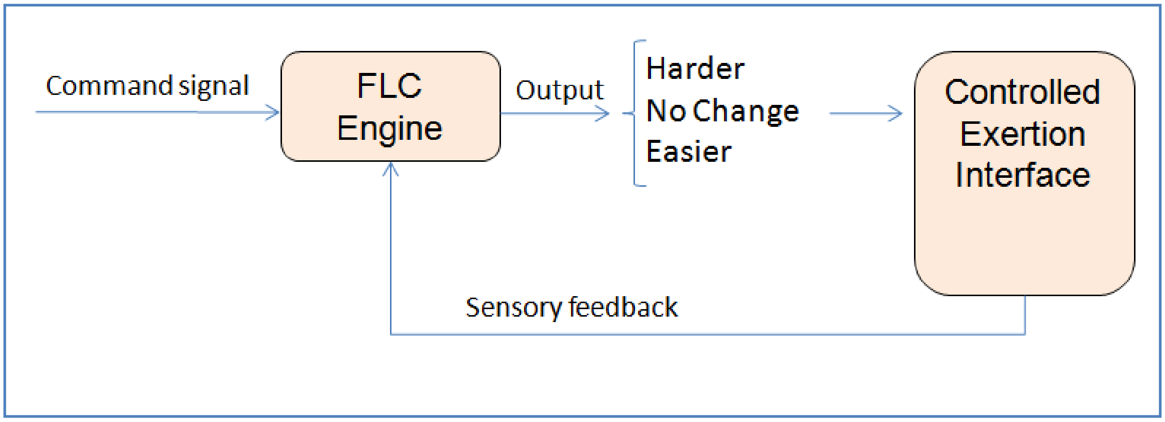


Figure 4.5 Block diagram of the fuzzy logic control for an exertion game interface

To the right of Figure 4.5 we find the system being controlled: the exertion interface. On the left there is the Fuzzy Logic Control (FLC) engine. The FLC engine receives a command signal, namely the target exertion zone for the player. It also receives sensory feedback coming from the controlled system, which is the current exertion level of the player, as a reading of physiological sensors. It then computes an output signal that tells the game interface to make the game harder, easier or remain unchanged.

The process of computing the output based on the input signals is represented in figure 4.6, which depicts the internal elements of the FLC engine.

The whole process consists of the following four general steps:

1) Input processing: Take the command and feedback signals and compute the control 'error' and 'error rate' values.
2) Fuzzification: Take the crisp values of error and error-rate and calculate the weighted equivalent values for each fuzzy set, applying the membership functions.
3) Inference: Take the membership values from step 2 and execute the fuzzy rules, generating the fuzzy output value.
4) Defuzzification: Apply a method to convert the fuzzy output into a crisp value.

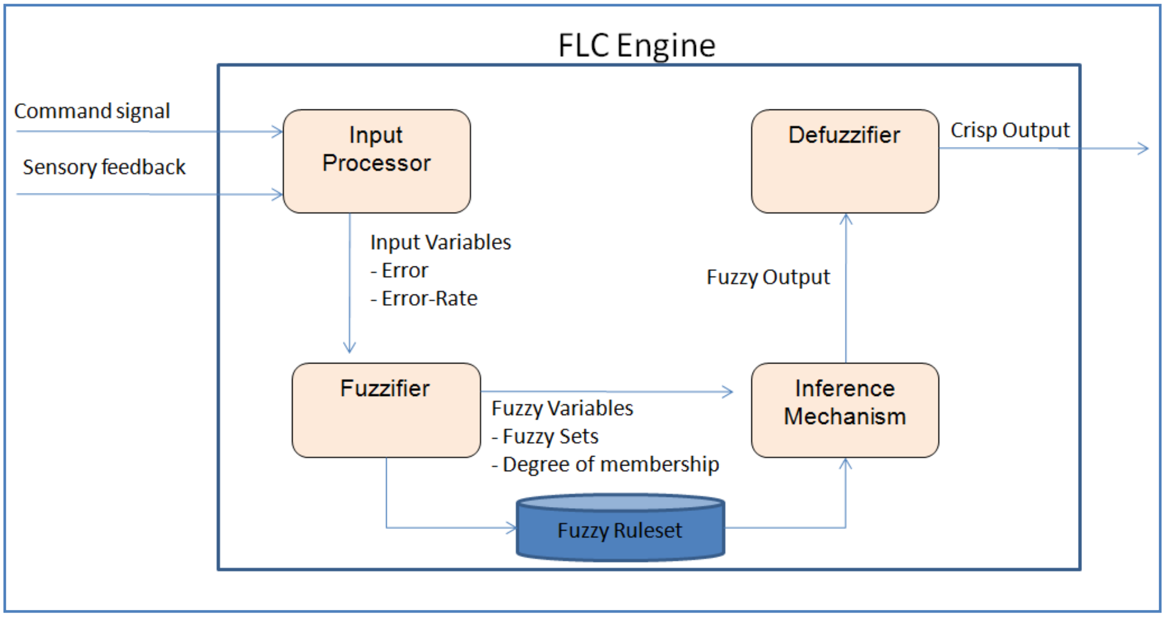


Figure 4.6 Components of the Fuzzy Logic Control (FLC) engine

Input Processing

The input signals ‘command’ and ‘feedback’ computed into an ‘Error’ variable, which is defined as the difference between the values of the command input (CMD) and the feedback value (FDBK) (See equation 4.2).

$$Error = CMD - FDBK \tag{4.2}$$

This error represents how far from the exertion target the player is currently exercising. The value of the error can be positive (the player is below target) or negative (the player is above target) as depicted by Figure 4.7.

Additionally, the ‘error rate’ is computed. The error rate is the time derivative of the error value (Equation 4.3).

$$\mathrm{Rate(Error)} = \frac{\mathrm{CMD} - FDBK}{dt} \tag{4.3}$$

This variable tells the system whether the value of the error function is increasing or reducing. By knowing the error’s change rate we can tell if the player’s level of exertion is increasing or decreasing. The value of the error rate can be negative (player's exertion is reducing), positive (increasing) or no change.

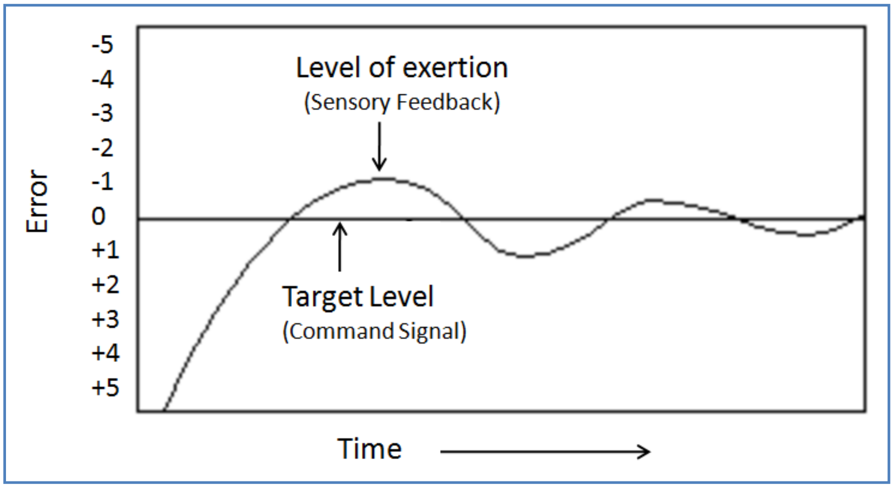


Figure 4.7 Command signal and Error function in the exertion control system

Fuzzification

After the error and error-rate are computed, the fuzzifier converts the crisp values into grades of membership for linguistic terms of fuzzy sets. For this purpose a membership function is used to associate a grade for each linguistic term.

Figures 4.8 and 4.9 present the membership functions for input variables, error and error-rate.

As can be seen in Figure 4.8, the error input is mapped into degrees of membership of five fuzzy sets described by five linguistic terms: Large negative (LN), small negative (SN), zero (Z), small positive (SP) and large positive (LP).

Fuzzification of the input value for error is carried out by selecting the corresponding input parameter from the horizontal axis and projecting upwards until reaching the upper boundary of the membership function. This point projected horizontally to the vertical axis determines the degree of membership, which is defined as a value between 0 and 1.

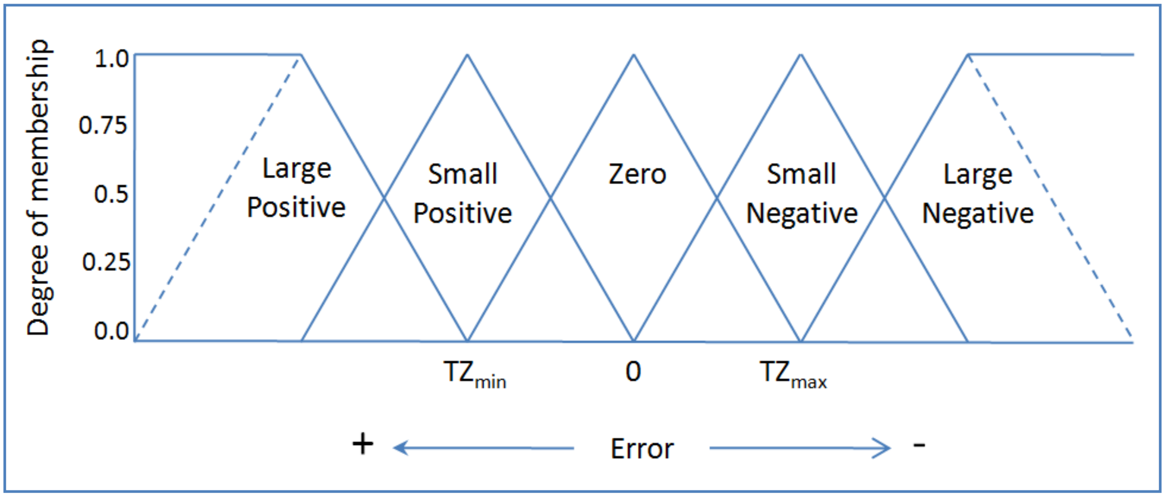


Figure 4.8 Error membership function

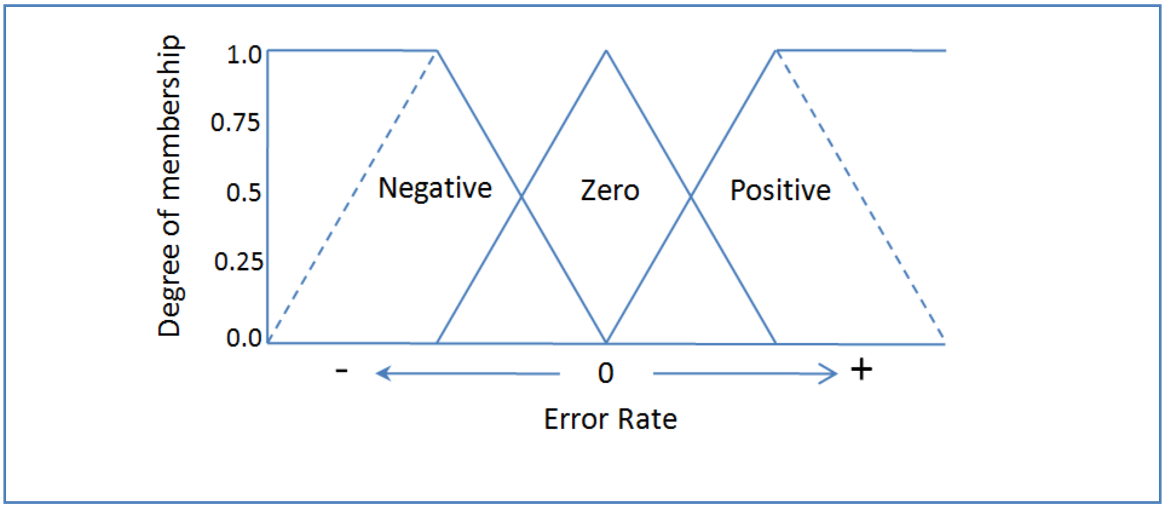


Figure 4.9 Error-rate membership function

Figure 4.9 shows the membership function for error-rate. In this case only 3 linguistic variables describe the fuzzy sets: negative (N), zero (Z) and positive (P). This membership function has only those three linguistic variables since our only interest is to know the direction (and not the magnitude) of the error change rate for the estimation algorithm.

From the linguistic terms described in the membership functions above, we can then derive a set of rules that will be used in the inference process. Table 4.1 presents the rule matrix derived from combining all the linguistic terms for both input parameters.

	LN	**SN**	**Z**	**SP**	**LP**
N	SD	NC	SI	SI	LI
Z	SD	SD	NC	SI	SI
P	LD	SD	SD	NC	SI

Table 4.1 Rule matrix for the linguistic terms

The columns in the rule matrix represent the possible output of the membership function for the error parameter, expressed as linguistic terms that represent each fuzzy set. Similarly, the rows represent the possible outputs of the error-rate membership function. The cells in the center of the matrix are the conclusions that can be reached, based on the conditions given by rows and columns. Each conclusion is a fuzzy representation of the output control signal for that particular condition. In the next section we explain how those conclusions are combined considering the weighted values of the degrees of memberships by the inference mechanism. The following set of rules is derived from the matrix:

1) **IF** error = LN **AND** error-rate = N **THEN** SD
2) **IF** error = LN **AND** error-rate = Z **THEN** SD
3) **IF** error = LN **AND** error-rate = P **THEN** LD
4) **IF** error = SN **AND** error-rate = N **THEN** NC
5) **IF** error = SN **AND** error-rate = Z **THEN** SD
6) **IF** error = SN **AND** error-rate = P **THEN** SD
7) **IF** error = Z **AND** error-rate = N **THEN** SI
8) **IF** error = Z **AND** error-rate = Z **THEN** NC
9) **IF** error = Z **AND** error-rate = P **THEN** SD
10) **IF** error = SP **AND** error-rate = N **THEN** SI

11) **IF** error = SP **AND** error-rate = Z **THEN** SI
12) **IF** error = SP **AND** error-rate = P **THEN** NC
13) **IF** error = LP **AND** error-rate = N **THEN** LI
14) **IF** error = LP **AND** error-rate = Z **THEN** SI
15) **IF** error = LP **AND** error-rate = P **THEN** SI

Inference

The inference mechanism consists of evaluating each one of the conditions that form the rules. The ‘IF’ part of a rule is called the “antecedent” and the ‘THEN’ part is called the “consequent”. To evaluate a particular condition in an antecedent, the input parameters crisp values are mapped to the linguistic counterparts with the corresponding degree of membership, as explained in the fuzzification process. Those weighted fuzzy variables are then used to evaluate the antecedent of a rule. Whenever the antecedent is satisfied, then the rule is fired and the consequent is taken.

For each consequent that is taken, a firing strength for the rule is assigned. This is done applying an aggregation operator for the weights of the individual conditions that satisfied the rule’s antecedent. For this research we have chosen the fuzzy logical AND operator. The AND operator simply chooses the minimum weight from among all the conditions of the antecedent.

The degree of truth (firing strength) of each rule is calculated then by evaluating the non-zero values applying the AND operator (minimum non-zero value is taken). Once all the rules are evaluated, the resulting weighted consequences are combined into a single fuzzy set. The method selected for the aggregation of the fuzzy results for each rule is the Root Sum Square (RSS). This method has been selected because it takes into consideration the weighted contributions of all of the rules that were fired. This is important when the number of input parameters and derived rules is relatively small. RSS is described by equation 4.4:

$$\sqrt{\sum R^2} = \sqrt{\left(R_1^2 + R_2^2 + R_3^2 + \dots R_n^2\right)} \tag{4.4}$$

Where: $R_1^2 + R_1^2 + R_1^2 + \ldots . R_n^2$ represent the truth values of a set of rules that share the same conclusion. Hence, the outcome of applying the RSS aggregation method is a fuzzy set of output membership function strengths: one for each possible linguistic variable in the output membership function, which is shown in Figure 4.10.

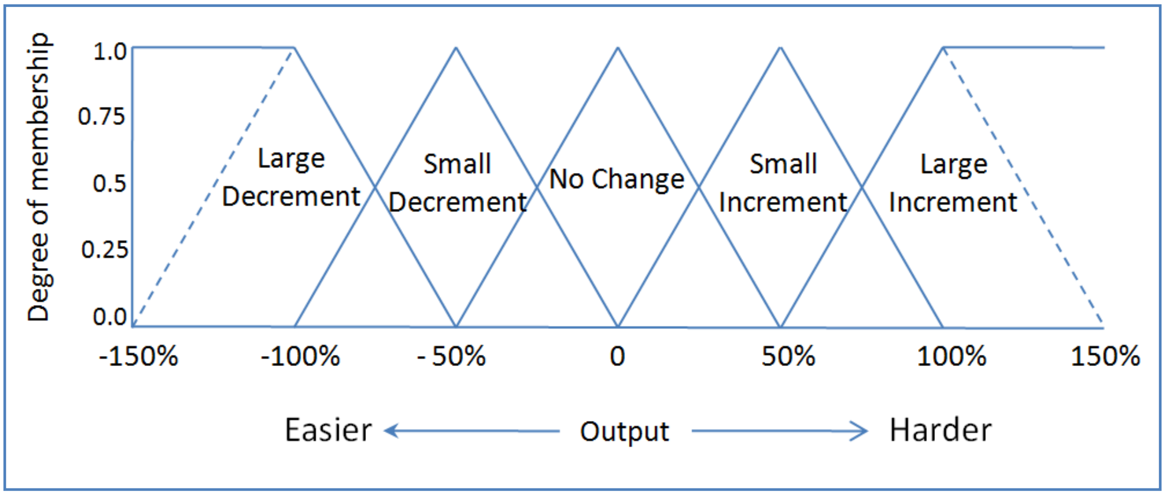


Figure 4.10 Output membership function

Defuzzification

The last step in the process is the conversion of the output fuzzy set obtained from the inference into a crisp value that can be used to drive the difficulty change on the game engine.

Several techniques exist for the process of defuzification [Abraham and Nath_00]. However, we have selected the (widely used) technique of calculating the "Fuzzy Centroid" (FZ). This approach is more accurate to represents the fuzzy output region [Cochran and Chen_05]. Equation (4.4) shows the formula for the Fuzzy Centroid calculation:

$$\mathrm{FZ}(Y') = \frac{\sum \mu_Y(x_i) x_i}{\sum \mu_Y(x_i)} \tag{4.5}$$

Where: $\mu_Y(x_i)$ is the membership value in the membership function and X_i is the center of the membership function. With this technique, the weighted strengths of each membership function output is multiplied by their respective function center points and summed. This operation renders the effective fuzzy area, which is then divided by the sum of the weighted member function strengths. The result of this is a crisp output used in the adaptation process of the exergame interface.

The crisp value to be obtained is a percentage value; it represents how much out of the step size in difficulty should be applied in the next iteration of the game engine adaptation loop.

All four steps described before are integrated in the algorithm shown in Figure 4.11.

Fuzzy Algorithm for estimation of level of exertion

```
Initialization:
    1. Let $s_1, s_2, s_3 ... s_n$ be the readings from sensors used to estimate exertion level
    2. Let $TZ_{min}$ be the lower boundary of the target exertion zone
    3. Let $TZ_{max}$ be the upper boundary of the target exertion zone
    4. Let $\Delta$ be a fixed time that defines the frequency of execution
Begin
    Loop while <game plays>
        $CMD \Leftarrow TZ_{min} + (TZ_{max} - TZ_{min}) / 2$
        For s = 1 to n
            $FDBK_t \leftarrow$ Add weighted value of reading $s_i$
        Error = CMD – FDBK
        $dt = t_0 - (t_{-1})$
        Error-Rate = (CMD – FDBK) / dt
        Map crisp values (Error and Error-Rate) into fuzzy sets F
        Calculate degree of membership W for each (parameter $\Leftrightarrow$ set)
        For each rule in R do
            If (antecedent(F) of $r_i$ == true) Then
                Set firing strength $r_i$ applying AND operator $\Rightarrow$ R'
        For each element in R' do
            SQ-Sum += $r_i^2$
```

```
        Y' = SQ-Sum^0.5
        Output = Fuzzy Centroid FZ(Y')
        Return Output
        Sleep for time Δ
        t = t + Δ
End
```

Figure 4.11 Steps of the fuzzy algorithm used to drive the level of exertion of the player to the target zone

4.6 Summary

The above described components form the basis for our architecture. Our design aims at providing developers with a guideline on how to implement exergames where a meaningful exercising session is necessary and can be achieved by dynamically monitoring the user performance and adapting and changing the game dynamics accordingly. The architecture makes no assumptions of the type of programming languages or specific technologies that must be used for technical implementations, and the loosely coupled design allows for it to be used as a whole or omitting some of its components.

In the next Chapter, we present a particular implementation as a proof of concept. Such implementation can be particularly useful to further understand how the components generally described above are realized in an actual adaptive exertion game with natural interfaces.

5 An adaptive exergame with multimodal natural interfaces

5.1 Introduction

The previous chapter presented the design principles of an architecture that can be used as guidelines for the development of exertion game engines that adapt to the physical exertion context of the player.

Based on the proposed architecture, we have implemented a specific exergame that implements adaptation as a proof of concept. For this game, we have also replaced the standard computer input devices (mouse and keyboard) by a more natural interface, leveraging the sensors embedded in smartphones. We designed two smartphone controllers that allow interaction with our test bed game engine.

Following this same approach we propose another natural interface using hand gestures recognition by means of a pre-existing algorithm that combines Bag-of-features and Support Vector Machine (SVM) to achieve user-independent and real-time hand gesture recognition.

Finally in order to improve the accessibility of the game we propose a haptic interface that allows visually impaired individuals to play the game. This approach employs sensory substitution to transform spatial information of game objects from the visual to the haptic channel.

5.2 A Target-Shooting Adaptive Game

We developed a game that follows the pattern described by our proposed architecture as a proof of concept. In the following subsections we present the implementation of such a game, clearly identifying the specifics for each component of the architecture.

5.2.1 User Interface and Game Engine

The game is based on a typical target-shooting scenario. The user interface shows a 3D Island Environment where the camera viewport has a first person shooter perspective showing the direction where the player is looking at. The player's character is standing on a platform facing a field, and in his/her line of view there are three targets located at different distances which the player must knock down by throw coconuts at them. The player plays the game while sitting on a stationary bicycle which acts as the input controller for the game. The computer screen is positioned in front of the player at eye level as it can be seen on figure 5.1.

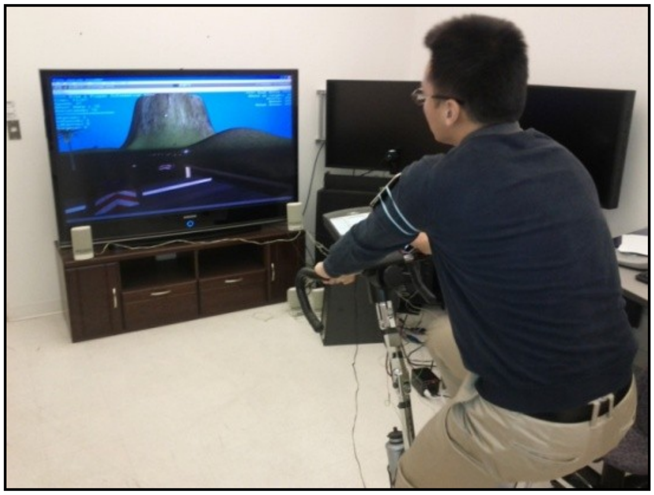


Figure 5.1 User playing the game using a stationary bike as input controller

The Game Engine handles the rules of the game as follows: to gain points, the player must knock down all 3 targets within a 10 seconds time frame between knocking down the first and the third target. After 10 seconds of being knock down, a target stands up again by itself. The player gets feedback about the direction of the throw by seeing a crosshairs icon on the screen. See figure 5.2.

The cross hairs move in a horizontal direction in an autonomous way, moving left and right at a constant speed uncontrolled by the player and affecting the direction of the throw in the x axis. The vertical movement of the crosshairs are dictated by the stationary bicycle used as an input controller. The RPM (revolutions per minute) reading from the stationary bike pedals is mapped into the Y coordinates that drive the vertical displacement of the crosshairs, hence affecting the trajectory and distance (depth) to which a coconut may be thrown by the player. The higher the RPM reading is, the higher the throw and vice versa. As the game plays the player must adjust slightly the speed of his/her pedaling to reach the closer or farther targets. Then he can shoot by pressing a button adapted to the handle bar of the bike.

```
==== Bike & Player Information ====
Speed: 35
Pulse : 116
```

Figure 5.2 Screenshot of the game showing coconuts being thrown at targets

Coconuts thrown describe a parabolic path according to the simulated gravity rendered by the physics engine, hence the crosshairs are just a general guideline for the direction of the throw, but does not act as a sniper's riffle where you must locate the crosshairs exactly on the intended target. This makes the game a bit more challenging and improves the likelihood that the player must estimate this shift on direction of the throw, which increases the physicality of the game given the type of input controller.

The game continues for a fixed amount of time that can be preset at which point the game is over and the total score is calculated.

5.2.3 Managed Variables

The game engine as described above would register some game variables that are key to increase/decrease the playfulness of the game in terms of physical exertion. In this case the variables are described as follows:

<u>Bike resistance factor</u>: This variable defines the degree of resistance applied to the pedals of the bike. The greater the value for this factor, the more strength is required to complete a revolution when pedaling.

<u>RPM to height ratio</u>: This variable defines the factor at which the RPM reading affects the height at which the crosshair is pointing. The greater the factor is, the more the height of the crosshairs is affected given a constant RPM reading. Higher ratio values will make the game easier (less effort) to play, because the user can reach the farthest targets at a lower pedaling speed.

<u>Crosshairs horizontal speed</u>: This determines the speed at which the pointer autonomously moves horizontally. Although this factor is not directly linked to the physicality of the game, it does makes it harder to knock down the targets by giving the user less time to move the vertical aim from one point to another, which results in demanding a more drastic change in speed, and hence greater physical effort.

<u>Inter-Target Distance</u>: The distance at which the targets are located from one another (in the z axis). The farther apart the targets are, the greater is the difference of speed required to reach them which results in greater physical effort.

For each one of these managed game variables, the game specify some change rules. That is, how the variables should be changed depending on whether we want to increase or decrease the physical challenge of the game. The amount of change for each step and the upper and lower boundaries for each variable (the point at which the variable should not be further affected in a given direction) is given in the Table 5.1 which summarizes the change rules for the variables described above.

Variable	Min/Max Values	Change Interval (Units)	Change to make the game more difficult	Change to make the game easier
Height Ratio	0/3	1 (Unit)	Reduce	Increase
Bike resistance	10/200	10 (Watts)	Increase	Reduce
Pointer Speed	1/10	0.5 (Meter/s)	Increase	Reduce
Inter-Target Distance	5/20	1 (Meter)	Increase	Reduce

Table 5.1 Change rules for the game engine managed variables.

As it can be seen, for each variable we specify the direction of change (+/-) for each type of decision (more/less difficult), along with the amount of change (Interval) and the boundaries (Min/Max values). These rules will be used by the adaptation module according to the given strategy.

5.2.4 Adaptation Strategy

In our adaptation strategy we run a loop in the program that continuously iterates executing the steps of the exertion estimation algorithm and using the algorithm outcome to make the appropriate changes to the managed variables. Figure 5.3 illustrated the steps involved in the adaptation algorithm strategy.

This strategy basically modifies the variables accordingly while the player is out of the exertion target zones, whereas while the player remains within the boundaries of the zone, then the variables are left unmodified.

Adaptation Algorithm for management of registered variables

```
Initialization:
    1. Let AL be a reference to the Linear Algorithm listed in Figure 4.4
    2. Let AF be a reference to the Fuzzy Algorithm listed in Figure 4.11
    3. Let V be the set of managed variables registered by the game engine
    4. Let $\Delta$ be a fixed time that defines the frequency of execution
Begin
    Loop while <game plays>
        Set Change to outcome of  AL() or AF()
        If (Change is not 0) Then
            For each var $v_i$ in V
                $v_i = v_i + Interval(v_i) * (1 + Change)$
        Sleep for time $\Delta$
        $t = t + \Delta$
    End loop
End
```

Figure 5.3 Algorithm for management of registered variables

5.2.5 Context, Knowledge Base and Exertion Estimation

The knowledge base for this implementation is a chart that lists the target heart rate zones for performing a reasonable good exercising session (by age) as estimated by the American Heart Association [AHA_12]. Table 5.2 shows the chart with the range for the optimal zone (50 to 85 % of maximum capacity) along with the average maximum values for each age group. The age of the player is used then to obtain the appropriate target HR zone.

In terms of context, besides the age of the player the adaptation engine relies on a constant monitoring of the current exertion level at which that the player is performing.

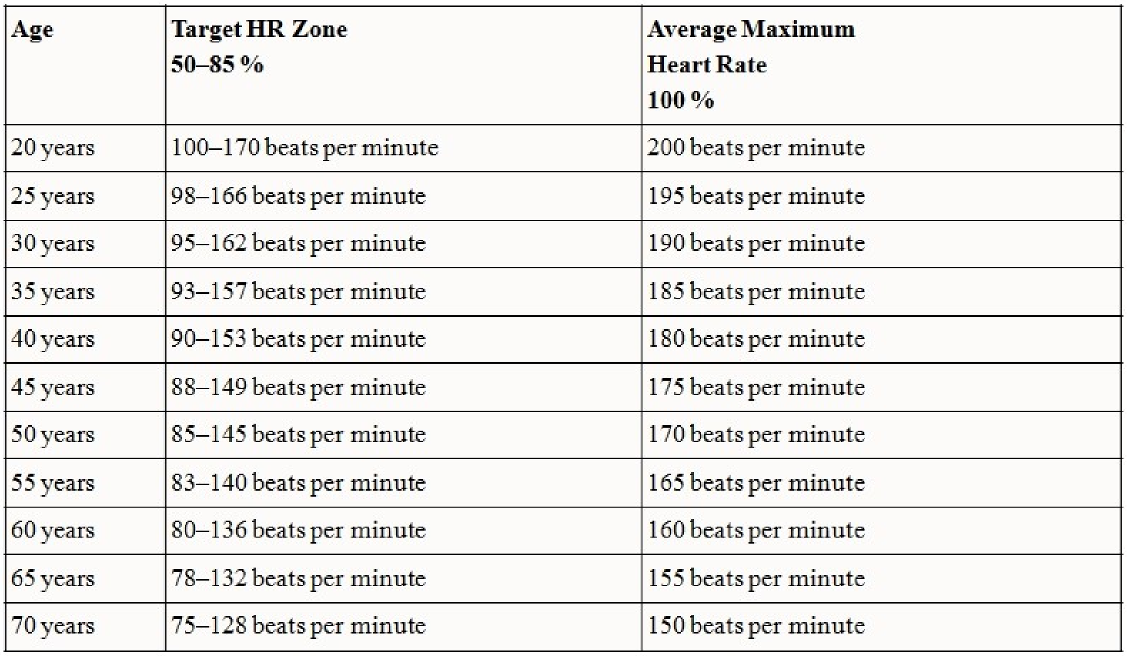
For this purpose our implementation makes use of a heart rate monitor attached to the player. The heart rate reading is taken as a source for context that then drives decisions to be made in the game engine.

The estimation of the changes required to drive the level of exertion to the target is based on calculations that follow the algorithms described in section 4.5. In this implementation two separate algorithms are considered, one linear and one based on fuzzy logic. For the linear approach a simple function is used to compare the raw reading from the heart rate against the target zone identified from the knowledge base (See equation X in section 4.1). In the case of the fuzzy implementation, the target heart rate is used to estimate the error and error-rate input variables, which in turn are mapped to fuzzy sets. The inference process then estimates the fuzzy output set and the defuzzification process renders a crisp output value (see section 4.2). In either case, the output of the exertion estimation is used by the adaptation module to make decisions and change the managed variables accordingly (Make the game harder or easier).

As it can be seen, this game implementation clearly follows the architectural principles described in the previous chapter. Because of this, the game can be easily expanded in several ways. By adding more sensors, or by implementing more sophisticated approaches for the adaptation strategy, or the exertion estimation. In any case the implementation for the remaining of the components remains unchanged.

5.2.5 Summary

In this section we presented the design and implementation of a specific exergame based on the proposed architecture. This initial proof of concept serves as a good working example on how the concepts discussed in this work can be applied in a practical application as presented in the next sections. Chapter 6 includes the results of a user study in regards to this implementation.



Age	Target HR Zone 50–85 %	Average Maximum Heart Rate 100 %
20 years	100–170 beats per minute	200 beats per minute
25 years	98–166 beats per minute	195 beats per minute
30 years	95–162 beats per minute	190 beats per minute
35 years	93–157 beats per minute	185 beats per minute
40 years	90–153 beats per minute	180 beats per minute
45 years	88–149 beats per minute	175 beats per minute
50 years	85–145 beats per minute	170 beats per minute
55 years	83–140 beats per minute	165 beats per minute
60 years	80–136 beats per minute	160 beats per minute
65 years	78–132 beats per minute	155 beats per minute
70 years	75–128 beats per minute	150 beats per minute

Table 5.2 Knowledge base with the target heart rate zones per age group

5.3 Smartphones as a natural form of interaction with exertion games

Exergames come in many modalities. However, a common feature of exergames is a novel user interface which may incorporate one or more sensors such as accelerometers, gyroscopes, laser pointers, video cameras, and even global positioning systems. These sensors and devices have mostly been used to track and assess the movement of the game players as they perform a physical activity, so that the game can respond accordingly [Exergames_12].

On the other hand, with the advent of modern mobile phones with higher processing power and memory capacity, so came the availability of advanced sensors embedded into such devices. Nowadays a typical smart phone in the market may be loaded with an array of sensors and embedded devices such as cameras, accelerometers, gyroscope, digital compass, GPS, sensitive multi-touch screens, proximity and ambient light sensors, and connectivity capabilities such as Wi-Fi or wide area data network. A recent market study about mobile phones and sensors reports that, by 2015, 44% of mobile phones will be smartphones [MEMS_11], while another report indicates that by 2013, 85% of smartphones will come with an integrated GPS, more than 50% will have accelerometers, and close to 50% will ship with gyroscopes [ABI_10].

All these capabilities have been used by application developers to provide a rich user experience and enhance the level of interaction that a user can have with the device. Although this wide range of capabilities has made smartphones remarkably versatile devices, with a broad scope of uses [Falaki *et al*_10], one particular type of application that benefits a great deal from these rich interactions is videogames.

There are a considerable number of videogames available to be played on the device itself, using the embedded screen [Kimbler_10]. Furthermore, researchers have extensively been investigating the possibilities of using the smartphone as a tool to interact with remote computers, displays and other devices [Boring *et al*_09] [Scheible *et al*_08] [Kray *et al*_10]. However, little research has been done to investigate the use of smartphones as game controllers, leveraging their sensor capabilities, and applying them to play exertion videogames.

We believe that smartphones have a significant potential to be used as game controllers. The capabilities provided by the array of sensors that can be found in a state of the art smartphone matches and, in some instances, even surpasses those that can be found in modern videogame consoles. Furthermore, smartphones have the additional capacity to be programmed to perform very specific tasks, which provides a truly flexible platform for the development of customized user interfaces.

This flexibility is certainly evident when we look at the nature of exertion games. When a player is engaged in an exergame session, they are likely to be moving their body, or otherwise performing physical exertion, which makes the usage of traditional hand held game controllers difficult. As it was explained earlier in this section, commercial videogame consoles that support exergaming usually have some sort of controller loaded with sensors that perceive the actions of the player and interprets them as input signals for the videogame in a natural way. These kinds of sensors are very similar to those found in smartphones. In particular, the use of the accelerometer and gyroscope of smartphones has been explored to track hand or body gestures that are used to direct actions in a videogame [Gilbertson *et al*_08] [Kiili and Merilampi_10] [Mottola *et al*_06].

In this section we present a proof-of-concept application which consists of an exergame interface designed using the guidelines outlined by our architecture. The exergame interface consists of two smartphone applications used as input controllers for a target-shooting exertion game that is played while exercising on a stationary bicycle.

5.3.1 A natural interface using smartphones as input controllers

In order to validate the design of the components of our architecture, we designed and implemented two smartphone input controllers that work in different modalities. For this purpose, we make use of a previously implemented exertion PC game that is played while exercising on a stationary bike (section 5.2). This game provides a convenient testing platform for the new smartphone controller interface, since it normally responds to input from the computer keyboard, which is awkward to do while exercising on the bike, and hence, a more natural form of interaction is required.

Game description

The game interface, rules and dynamics are described in section 5.2.1. The game was originally programmed so that the player can control the horizontal movement of the crosshairs by using the left and right arrows of the keyboard. The control is limited to changing the direction of the crosshair's displacement (left or right), but the speed and movement of the crosshairs actually is preset. This feature, although makes the user to exercise more, it prompts the user to make constant and frequent changes of direction. Throwing the coconuts is triggered by pressing the 'control' key of the keyboard. The player may shoot as often as he wants.

Figure 5.4 depicts the 3 input methods to control the game. This is a good example of a game that requires a novel and multimodal way of interaction, since handling the keyboard while exercising on the bicycle is an awkward challenge. In the next section, we explain how the mobile applications replace the keyboard input to provide a more natural interface.

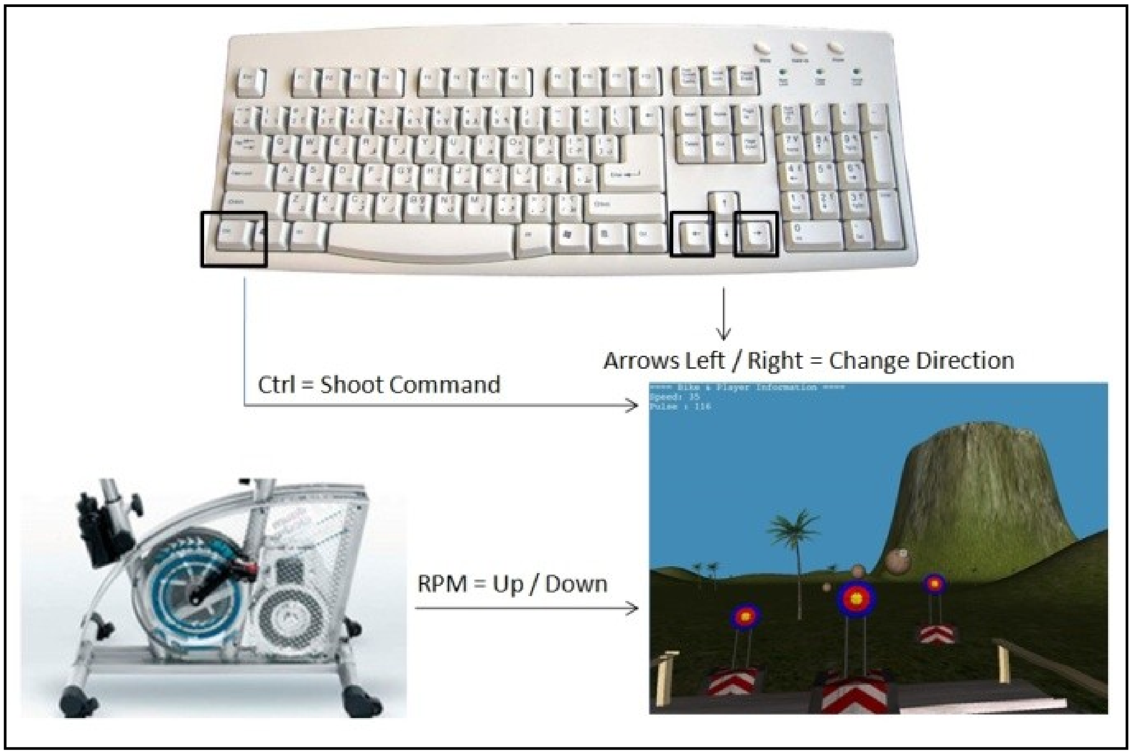


Figure 5.4 Game input methods (Without smartphone controller)

From the game dynamics as explained in the previous section, we have identified two areas of opportunity to support game interaction using a smartphone controller interface. The first interaction to support is the “Fire” game action. For this, we have made use of a soft button application running on a smartphone attached to the bike’s handle bar to fire quickly and comfortably while playing the game. Second, it is necessary to have a more natural way to change the direction in which the crosshairs icon is moving in the game. For this purpose, we have implemented an application that makes use of the accelerometer embedded in the device, to allow the player to “tilt” his body left and right to change directions as required.

The ‘Fire button’ interface

The first mobile application is a fire button. This is an interface that displays a soft button that occupies the entire screen of the smartphone. The smartphone is on the top of the handle bar, which allows the player to press the button with their thumbs, while holding the handle bar. See Figure 5.5.

Figure 5.5 “Fire button” application running on the smartphone attached to the bike

Figure 5.6 shows a sequence diagram where the interaction and the messages exchanged between the different components of the architecture can be seen in the context of this particular interface implementation. These interactions follow the general pattern described by the interaction diagram presented in Chapter 4 (Figure 4.3).

The interaction begins with the mobile application by registering an event listener for the ‘button tap’ action of the device’s interface. From there, the operating system monitors the state of the button, and when a ‘tap’ action is detected, an event notification is passed to the mobile application. The application interprets this tap as a “Shoot” action and sends such action message to the host application over the network. The host application in turn receives this action and decides what the related input event type is. In this case, the “Shoot” action should trigger a “key pressed” event from the keyboard with the code for the “Ctrl” key. This key pressed event is triggered in the context of the operating system, notifying the game engine, which is listening for this type of events. Finally, the game engine responds by triggering a ‘coconut throw’ within the game.

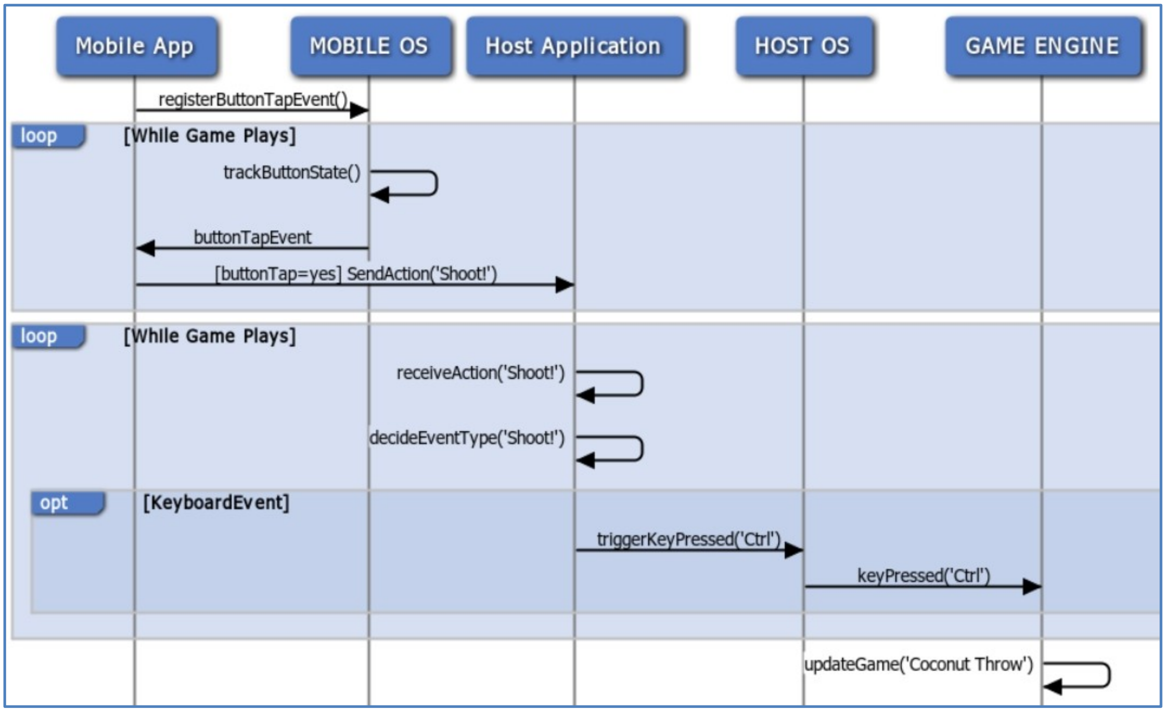


Figure 5.6 Sequence diagram showing messages and interactions of the components for the "Fire button" application

The 'Move-tilt' interface

The second custom interface's purpose is to aid in the control of the changes of direction of the crosshairs in the game.

For this, the smartphone is attached to the arm of the player and the player tilts his body left and right while sited on the bicycle (Figure 5.7). The application uses the smartphone's accelerometer to detect changes in the balance of the player. Accelerometers are devices that allow tracking of applied forces to the device in all 3 dimensional axis (x, y, z). This includes the force of gravity; hence by looking at the direction from which the force of gravity is being applied to the accelerometer, we can calculate the smartphone's orientation in space.

Every time there is a change by some degree in the player's body position (and attached smartphone), the application sends information about this action to the host application depending on the direction that the player is moving their body. The application will trigger a change of direction only when the "tilt" has occurred in a magnitude greater than a pre-calibrated threshold expressed in ‘G’ forces. This is because the accelerometer is acutely sensitive, and since the player’s body will inevitably move at all times while exercising, this could cause unintended changes of direction.

When the application is first launched, it requires some calibration. During the calibration process, the player must assume a natural position on the bike, sitting straight and pedaling for a few seconds. The smartphone application will then measure the range of movement that the player naturally has when pedaling the bike. This initial range is used as the positive and negative thresholds used to trigger a change of direction. Once the calibration is done, the player may start the game session, while the application tracks his movements and sends the proper messages to the host application when the player tilts their body left and right.

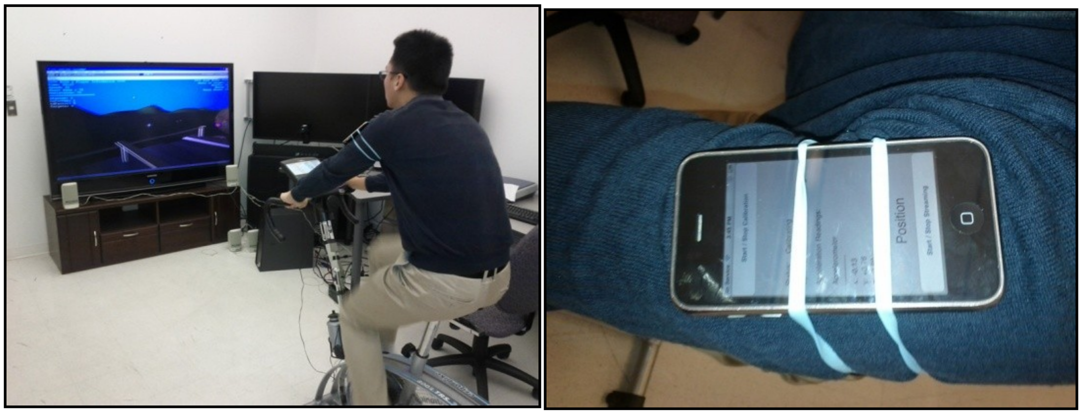


Figure 5.7 “Move-tilt” application running on the smartphone,
attached to the player’s arm

Figure 5.8 depicts the sequence of interaction between the different components to implement the functionality of the “Move-tilt” interface.

The mobile app runs in a loop requesting the accelerometer readings from the operating system, and then analyzes the forces vector to decide if the player body position is beyond the calibrated threshold. If this is the case, it sends a tilt message (“right” or “left” depending on the calculated forces) action, to the host application. In the same way it works for the fire button interface, the host application finds the correct mapping for this action, which is a “right arrow” keyboard event and triggers it in the operating system context. Finally, the game engine gets the notification and changes the direction of the crosshairs on the screen.

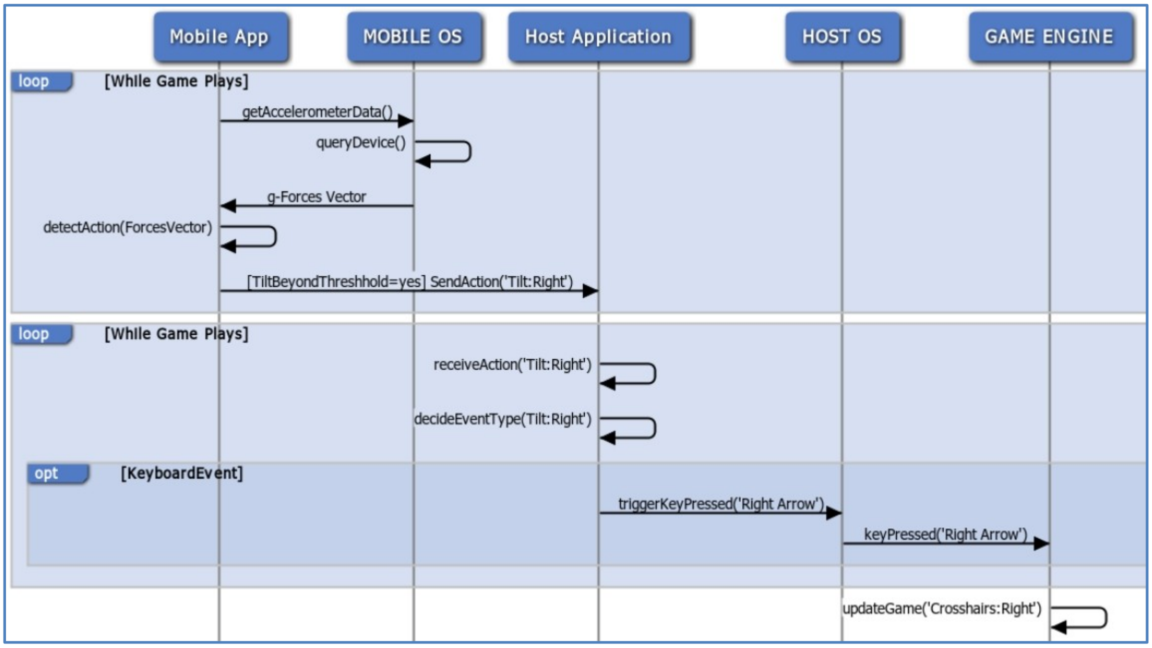


Figure 5.8 Sequence diagram showing messages and interactions of the components for the "Move / Tilt" application. Based on the proposed interaction guidelines described in section 4.4

With the mobile application replacing the functionality of the keyboard, we aimed at having the player more engaged in the gaming experience using a more natural interface. Figure 5.8 depicts the input methods to control the game with the smartphone controller interfaces.

With the implementation of the smartphone input controllers, we have an exertion game interface that does not require the use of traditional input devices for computer games such as mouse and keyboard. Although the game chosen for this implementation was part of previous work, the approach outlined by our architecture allows the integration of these interfaces into any pre-existing or of-the-shelf computer videogame (Figure 5.9).

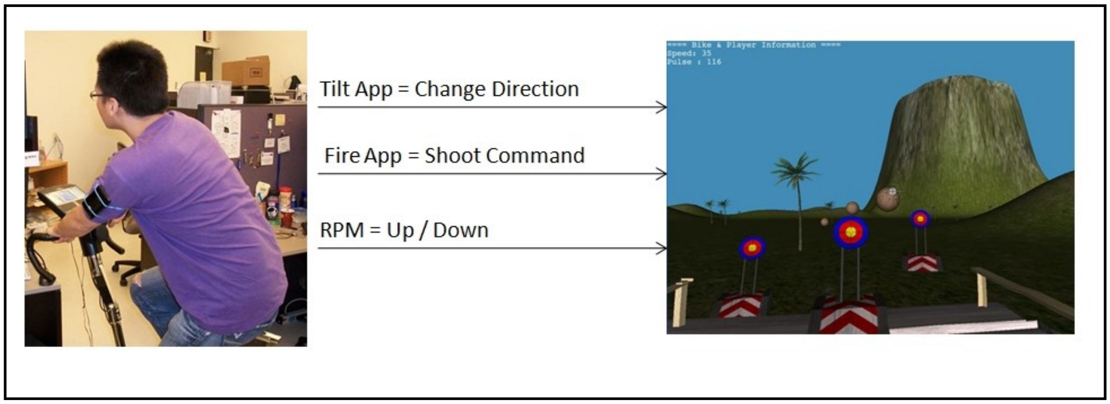


Figure 5.9 Game input methods with the smartphone controllers

Technical set up

For a host computer, we used a PC running Windows XP operating system. The configuration of the system was a 2.0GHz processor and 2GB of RAM. The exercise bike was an Ergo-bike 8008 TRS 3 model connected to the host computer by a serial cable.

The game was developed using Unity Game Engine and programmed using javascript and C#. No special modifications were made to the game to be compatible with the smartphones interfaces, as it simply responded to the keyboard input events received from the operating system.

The host application consisted of a console application developed using Microsoft C#. The smartphone (iPhones 3G) applications connected to the host computer thru a Wi-Fi local area network, using a wireless g router attached by a network cable to the host computer. The router was setup to use the 802.11g standard and the network environment was clean; only the two smartphone applications and the wired host computer were being served by the router.

The communication between the mobile apps and the host computer was established using network sockets based on the Berkley standard and using the Transmission Control Protocol (TCP). The host application acted as a server waiting for clients (mobile apps) to connect.

5.3.2 Summary

We presented the results of our work exploring the usage of smartphones as input controllers for exertion computer games. We have reviewed the properties that make smartphones ideal platforms for the development of this type of user interfaces for videogames in general and for exertion games in particular.

In this implementation we used the design guidelines outlined on our architecture. No assumptions are made of the underlying programming languages, communication protocols, and operating systems. It offers support for the design of configurable and flexible exertion interfaces that can be adapted to control pre-existing videogame titles without the need to modify their code base.

We have made a proof of concept implementation. The implemented system included two different smartphone interfaces to control an exercising videogame. A “fire button” application based on a soft button and a “move-tilt” application based on accelerometer data, which allowed players to interact with a target-shooting exertion game based on a stationary bicycle using an 802.11g wireless network communication link.

5.4 Hand gestures as a natural interface for exertion games

Natural interaction proposes the notion that users of a computer system should not concern with devices such as the mouse and keyboard. Instead humans can interact naturally with computer as they would do with other humans by employing gestures or speech to convey their intentions and possibly cognitive or emotional status. In particular, using “Hand Gestures” as an interaction method is gaining interests in human-computer interaction context in recent years. The main objective of gesture recognition research is to build a system which can recognize human gestures and utilize them to control an application or a videogame instead of traditional input devices.

The utilization of hand gesture interaction with videogames enables players to interact with computer environments in a natural, immersive, and intuitive manner. Another benefit of utilizing hand gestures in this context is that users need no longer to be in close physical contact with the controlled device or computer, but they can interact from a distance or without the need to have a display on sight.

Vision-based hand gesture recognition has been an active research area recently with applications such as sign language recognition, socially assistive robotics, directional indication through pointing, and control through facial gestures among others [Fang *et al*_07] [Nickel and Stiefelhagen_07] [Baklouti *et al*_08]. Within the broad range of application scenarios, hand gestures can be categorized into at least four classes: controlling gestures, conversational gestures, communicative gestures, and manipulative gestures [Wu and Huang_99].

By using the guidelines provided by the natural interface components proposed by our architecture (Chapter 4) we designed and implemented a gesture interface that can work with any pre-existing computer games (without a need to change their source code). The controlling gesture interface allows players of an exertion game to direct the game character's actions by using a set of hand gestures that are recognized by the computer.

5.4.1 Hand gesture algorithm and grammar

To detect hand gestures we used Darda's algorithm for real time hand gesture detection and recognition using bag-of-features and support vector machine techniques [Dardas and Georganas_11].

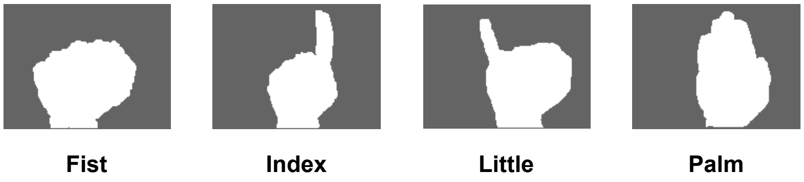


Figure 5.10 Templates of hand posture

This algorithm detects gesture commands as actions, which consist of a sequence of hand postures. It can detect 4 postures: fist, index, little, and palm as in figure 5.10. Although in our interface development we used only palm and fist.

The gesture commands can be generated by two ways. First, by observing the gesture transitions from posture to posture. For example palm to fist as depicted by Figure 5.11.

Figure 5.11 Transitions from fist to other postures

Second, by observing the direction of movement for each posture: up, down, left or right (Figure 5.12).

Figure 5.12 Movement direction cases of palm posture

5.4.2 The hand gesture interface

From the description of the game in section 5.2.1, the two main interaction schemes to address are the “Shooting” Action and the “Change Direction” action, which would normally be done by using the keyboard.

To address the problem of moving the crosshairs left and right and throwing coconuts while pedaling the bike, we used three hand gestures that generate commands, two of which allow the player to change direction (left/right) and a third one to shoot a coconut.

The derived interaction commands for the game are presented in Table 5.3. Figure 5.13 shows the interface with the player on the bike making hand gestures to play the game.

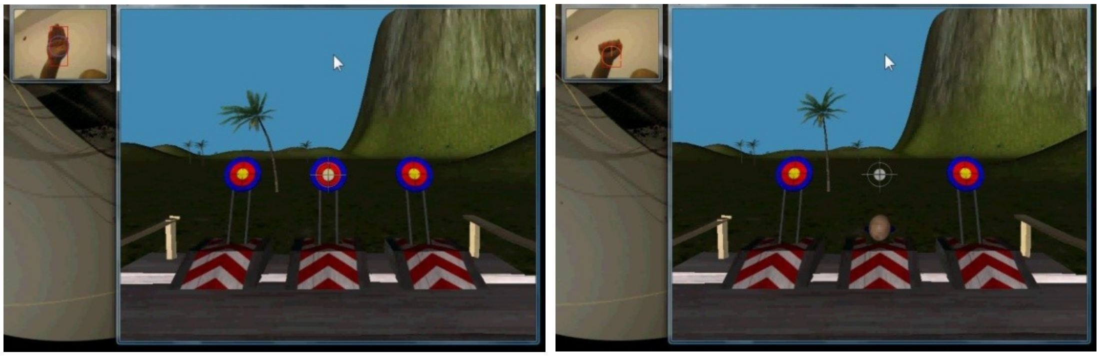
Current State	Next State	Direction	Send Key	Event
Palm	Palm	Right	{RIGHT}	Move Right
Palm	Palm	Left	{LEFT}	Move Left
Palm	Fist	Hold	^	Shoot

Table 5.3 Interaction commands performed by hand gestures

Figure 5.13 Game input method with a hand gesture recognition

Shooting hand gesture

The first custom gesture represents a “fire button” action. When the “Palm to Fist” gesture is detected by the gesture recognition process, it generates a Key-pressed event with the “Ctrl” key as an identifier. This event is then picked up by the game, which responds by throwing a coconut. This gesture can be observed in Figure 5.14 when the player changes her/his hand posture from palm posture to fist posture to shoot the target by throwing a coconut. The palm posture of the player is shown in the small screenshot in the top left corner of Fig. 5.14 (A) while in the next frame, the player changes her/his hand posture to fist posture to shoot the target as shown in the small screenshot in the top left corner of Figure 5.14 (B).



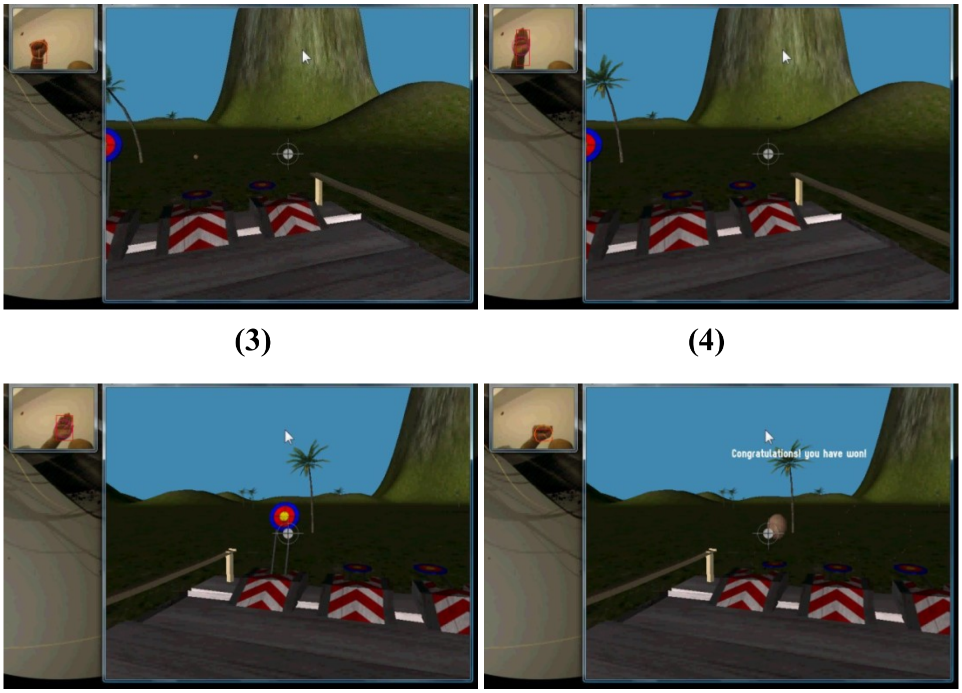
(A) (B)

Figure 5.14 “Shooting” action using hand gesture for the exertion game.
(A) Player with palm posture. (B) Player changes to fist posture for shooting

Changing direction hand gesture

The other two gestures represent the action of moving the crosshairs left or right. Again, when either of these two gestures is detected, a key-pressed event is generated with the “Left arrow” or “Right arrow” identifiers, accordingly. These gestures can be observed in Fig. 5.15 when the player moves her/his hand palm posture left or right.

(1) **(2)**



(3) (4)

(5) (6)

Figure 5.15 Change Direction hand gestures.

(1) Player with palm posture (2) Player moves their palm posture right to move crosshairs in that direction. (3) Player changes to fist posture for shooting. (4) Player changes to palm posture. (5) Player moves their palm posture left to move crosshairs left. (6) Player changes to fist posture for shooting.

Technical Setup

For a host computer, we used a PC running Windows XP operating system. The configuration of the system was a 2.0 GHz processor and 2GB RAM. The exercise bike was an Ergo-bike 8008 TRS 3 model connected to the host computer by a serial cable. The web cam was a low-cost Logitech QuickCam connected to the host computer by a USB port. The game was developed using Unity Engine and programmed using javascript and C#. C# was used to integrate our DLL file generated from C-based gesture recognition component. With this framework, the C# methods can call applications and libraries in C/C++. The program in C# was used to send events to the keyboard using the SendKeys.SendWait method. Those events or commands are generated using our hand gestures.

5.5 A haptic interface for the visually impaired

Engaging in any form of physical activity can be challenging for those individuals with visual impairments. Because of the challenges faced, often these individuals are subject of greater risks of health problems related to inactivity [Skaggs and Hopper_96] [Lieberman and McHugh_01] [Surakka and Kivela_08].

Exergames are more physically demanding than regular computer games [Graves et al_07] [Luke_05] and for the visually impaired they seem in essence a good alternative; They can be played alone, indoors, and do not require the supervision or assistance that is commonly needed for blind people to engage in such activities [Lieberman *et al*_06]. However, playing exergames can be equally challenging, since most of the time their interfaces and interaction approaches are typically based on the acquisition of visual cues which prompt some input from the user at a specific time and of a particular nature. Often playing the game without perceiving such cues becomes impossible. This limits the ability of visually impaired people to engage in traditional exergames.

Although computer generated haptics has the potential to be a valuable technology to help gamers in having a more immersive and realistic experience, it is also a very well suited type of interaction to address the problem of accessibility to games for the visually impaired. There are currently different technologies that have been proven to improve the accessibility of electronic media to the blind people, such as automated reading software, voice synthesis, tactile displays with Braille and speech recognition. Yet, those are mostly focused on the input/output of text and there is still some media which is difficult to represent in a non-visual modality, particularly graphics (2D or 3D), which form the basis for any computer game.

There have been two main approaches to enable visually impaired individuals into computer games. The first one is to design games inherently designed and implemented with the purpose of not relying on visual feedback. That is, games whose playability is completely subject to interaction based on audio or haptic clues. The second approach is sensory substitution. In this case the cues that would normally come from the visual channel are replaced by haptic stimuli. This allows for the modification of currently existing games that were originally designed for non disabled individuals and adapted to be played by the blind.

In our work we focus on sensory substitution, we aim at providing the ability to play more complex games that normally would rely on visual cues. We propose an interface to interact with an exergame that is played with a stationary bicycle as the main input controller. Interaction is supported by using two haptic displays, using a funnelling effect that allows acquisition of haptic perception of the same spatial cues necessary to play the game visually.

In this section we present a description of our exertion game implementation, we discuss the implications of sensory substitution to enable spatial perception through haptics and describe in detail the haptic interface that has been implemented.

5.5.1 Visual to haptic sensory substitution of game objects

For this implementation we used the game described in section 4.2.4.1: "User Interface and Game Engine". From the description of the game in that section, we can observe that there are two basic visual cues that enable a user to play the game. The first one is the location of the nearest target in the direction where the crosshairs are moving. And the second one is the height at which the crosshairs are pointing at any given time. With those two pieces of information a player can adjust the pedaling speed to match the location of the target with the location of the crosshairs pointer, so that they know when to shoot. A final cue is the feedback on whether the shooting was successful in knocking down the targets or not, in the case of the haptic enabled game, this information is conveyed by an auditory clue.

The visual feedback is so important that the game could actually still be played by sight only. In order to allow a person who can not perceive the visual cues of the game to be successful at playing it, we have made a substitution of some of the visual elements to a haptic feedback. For this purpose we have first observed and identified the visual reference elements or cues that a player typically uses in order to achieve the goals of the game, which are knocking down the targets and gaining points.

By analyzing the dynamics of the game, we have identified 6 game cues that the user should be aware of to fully play the game (Figure 5.16):

1. Relative position of the nearest target
2. Position of crosshairs on the Y axis
3. Position of crosshairs on the X axis
4. Coconut thrown and flying towards the target
5. Event of coconut hitting a target and knocking it down
6. Message alerting the user whenever they have won bonus points

Figure 5.16 Visual cues that a typical user has as a reference to play the game

We have classified each of these cues as being either primary or secondary, depending on the level of importance in terms of how useful they are to locate and knock down targets (which is the primary objective of the game). Thus, we consider cues 1 and 2 (relative position of the nearest target and the crosshairs position in the Y axis) to be primary cues. Cues 3 to 6 are then considered to be secondary as they do add value to the game experience but are not an absolute requirement in order to achieve the primary goal. Inf particular, for the sake of simplicity we have obviated cue number 3, since the horizontal movement of the crosshairs is autonomous and the user has no control over it. Plus, it is less relevant as long as the user knows where the next target is located.

Since the human sense of touch is very sensitive, and there is evidence that suggest a limitation in terms of the amount of information that can be conveyed by this sense through sensory substitution [Enriquez *et al*_06], we have restricted the haptic feedback to only the primary cues of the game. This is of particular importance since our approach to sensory substitution requires a certain level of accuracy to be able to locate and knock down the targets. Secondary cues are left only as auditory elements.

Primary cues are then to be presented to the user by haptic feedback. For this purpose we have created two haptic displays which are worn on the player's forearms as arm bands, each using an array of vibrotactile motors that are activated accordingly to represent the fixed position of the nearest target in one arm, and the moving position of the crosshairs (using a funnelling effect) in the other arm.

The right arm holds an array of vibrotactile actuators that are activated in sequence with an algorithm that gives a continuous haptic feedback with a funnelling illusion (described in detail in section 5.5.2). This funnelling effect is mapped to the movement of the crosshairs pointer as it moves up and down the screen, whose spatial representation is represented in the space between the first and the last actuator located in the player's arm. The left arm holds a similar array of actuators, except that those are activated in a discrete manner to represent the fixed position of the nearest target on the display. The target may be at 3 different fixed heights, each one is mapped to the locations of each one of the three actuators in the player's arm.

As the game plays, the player gets haptic feedback about both visual cues: The location of the target in the display, and the location of the crosshairs aiming at the target. The player then is able to identify where the crosshairs are located in relation to the target (above or below) and adjust the speed of the bicycle accordingly. When the player perceives that the haptic feedback is in the same location in both of their arms, they know that the crosshairs are pointing to the target and they can then shoot to knock it down.

The exergame experience is completed by the use of auditory cues, which give additional information to convey a sense of accomplishment when the targets are hit, whenever bonus points are won and representing the shooting action (cues 4 to 6).

In the next two sections we describe in detail the implementation of the haptic displays, including technical details, the funnelling effect algorithm and how it all ties up to the game.

5.5.2 Funnelling illusion to represent distance information with a haptic display

The main goal of the proposed exertion game is hitting the targets on the screen. The main sensory device that players use for this purpose is "vision", which allows the player to estimate the location of the game objects (targets, pointer and other elements), and the distance between them in relation to the screen dimensions.

To perform sensory substitution we need a way to map those locations and distances from the screen into the haptic display. There are different methods to represent distance information, both visually and tactually in a display, and previous work has made an analysis of them for its applications in Telepresence and Tele-operation [Cha *et al*_10]. In our case we needed an effective way to represent relative distances in a continuous way without the need for the player to know the range with precision beforehand. We also need for a rapid change of the perceived distance in both directions. For this purpose we have chosen to use a vibrotactile array with motors that are activated with a "funnelling illusion" as proposed by Cha and others [Cha *et al*_10].

This technique overcomes issues with previous approaches by allowing a lower sensor density and increasing the physical resolution of the display. It also reduces the cost, size and complexity of the vibrotactile displays, while still rendering a vibrotactile sensation that is perceived as smooth and continuous movement.

The "funnelling illusion" is a concept developed in the field of psychophysics, and it can be described as a phantom sensation perceived in between two adjacent points of the skin where real vibration stimuli are applied [Bákésy_58]. The location where this phantom sensation is perceived can be manipulated by modulating the intensity of the two stimuli being rendered on the skin. [Alles_70].

Figure 5.17 depicts this funnelling effect. The dots that can be seen on the arm represent the vibration stimuli being applied to two points in the skin and the size of the dot represents the level of intensity of the vibration. The triangle shows the location where the perceived phantom sensation would be located. As it can be seen from (a) and (c), when the intensities are modulated, the funnelled stimuli shifts toward the more intense vibrotactile stimuli.

By modulating the intensities of two adjacent vibrotactile actuators and by adding more actuators to the array, we can build a haptic display capable of rendering the location of the moving crosshairs whose location would normally be perceived visually. The height at which the crosshairs are displayed at any given time and the entire range of movement of the crosshairs are mapped into the space between the first and last actuator of the array located in the arm of the player. In the case of the location of the nearest target, this mapping is more direct, and each one of the three positions of the targets are mapped into each one of the three actuators in the array for the second arm. These mappings are depicted in Figure 5.18.

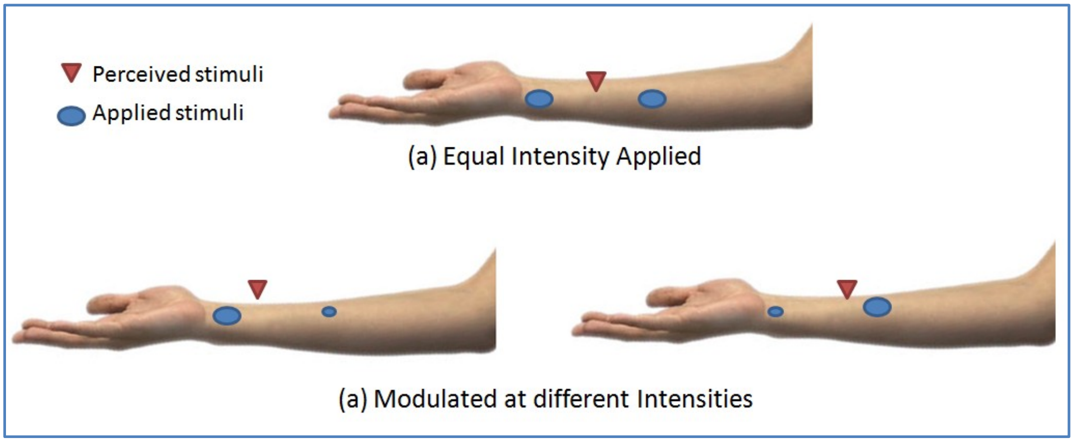


Figure 5.17 Funnelling effect showing the location of the phantom sensation and the locations of the modulated vibration stimuli

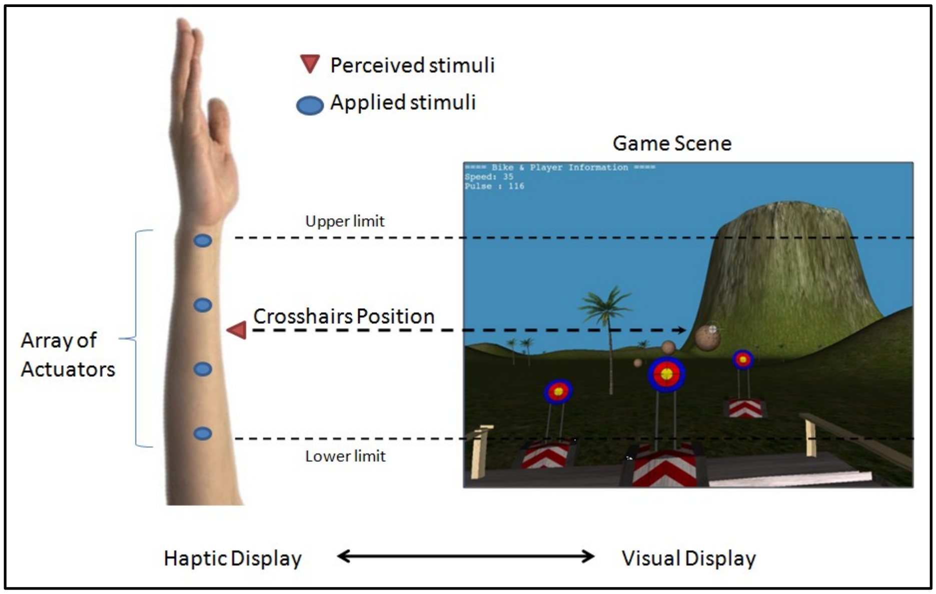


Figure 5.18 Mapping of the game objects location from Visual to Haptic displays

In the next section we present the implementation of two such displays that form the haptic interface for our exertion game.

5.5.3 Armband Haptic Displays

In previous section we have reviewed the concept of “funnelling illusion” which is the basis of our haptic display implementation. Here we discuss the details of the hardware used to build the displays and the related software that enables their use as output devices for the exertion game.

The displays were constructed using an array of actuators, each consisting in a pancake-type vibrating DC motor commonly used in cell phones and other portable devices. This type of motor is inexpensive, lightweight, with low power consumption and easy to deploy and arrange in different configurations. They operate with a 3.6 voltage and have a frequency range of operation up to 220Hz.

The array of vibrators is attached with Velcro stickers to a piece of neoprene fabric which forms the armband that then can be attached to the player’s arm with Velcro straps. The flexibility of the neoprene fabric allows for a firm contact of the actuator with the skin, but leaves room so that it is not uncomfortable. When the armband is folded around the player’s arm the first actuator rests on the back of the wrist and the last one is close to the elbow as shown on Figure 5.19.

Each one of the actuators is connected to the end of a ribbon type cable, which connects on the other end to the discrete control pins of a microcontroller. The microcontroller was interfaced to a laptop computer using a usb port. For our implementation we have used the Arduino Mega microcontroller which was programmed to generate a pulse width modulation signal (PWM) to each actuator as commanded from the game engine. The PWM signal allows us to set variable levels of intensity of vibrations for each motor.

Figure 5.19 The arm bands with actuators in position
and user (top right) blindfolded playing the game

Two arm band haptic displays were implemented. One rendered a funnelling effect along the path of the array of actuators to represent the moving crosshairs used for aiming in the game. The other arm band simply activated the actuators in a discrete way to represent the fixed position of the next target. We describe the differences of implementation for each of them next.

“Crosshairs” haptic display

This haptic display was built to represent the location of the moving crosshairs. For this purpose it redered a funnelling effect. Based on results from previous work, it has been found that in order to obtain the best results in terms of a continuous and smooth stimulus with the funnelling effect, the distance between two actuators should be somewhere between 40 and 80 mm [Cha *et al*_08]. In our particular implementation we chose a separation of 50mm, which is within the suggested range, and which worked out well for having a continuous feeling across the entire array and was appropriate for our type of application. This gives a total length of 150mm for the entire area covered by the array, which is well below the 252mm that is the average length of an adult’s forearm [Alves *et al*_04].

In the Arduino platform the range of values allowed for PWM signals is between 0 and 255. In a previous psychophysical experiment [Cha *et al*_08], it has been investigated the relationship between applied levels of intensities and the perceived intensity of a vibrotactile stimuli by humans. From this work it has been found that the full range of PWM intensities can be narrowed down to 12 levels of intensity that humans can actually perceive and discern when utilizing the same type of motors employed in this implementation. Based on those findings we have used a subset of 12 values evenly spread across the entire range from 0 to 255 to be used in our implementation.

The location of the generated “phantom sensation” was calculated by assuming a linear relationship with the vibration intensities of two adjacent actuators. The first actuator would start at maximum intensity while the second would not vibrate at all, then as the first one decreases the level of intensity of the second one increases, rendering a perceived sensation that moves from the first actuator to the second as intensities vary inversely. Using 4 actuators in the array, each one taking 13 possible values of intensity, including the value of “0” (no vibration at all), gives a total of 39 possible perceived positions across the forearm, starting with the first actuator in the back of the wrist and ending at the location of the last actuator at a distance of 150mm from the first one.

“Targets” haptic display

This haptic display was built to represent the location of the next target in the game. No funnelling effect was necessary in this case, and the actuators were simply activated at full intensity as the reference point against the player is to match the moving crosshairs.

5.5.4 Summary

In this section we have presented an approach to improve accessibility of our game for the visually impaired. We achieved this by employing a funnelling effect algorithm and two haptic displays to do sensory substitution of reference cues from the visual to the haptic realm.

This sensory substitution allows players to perceive relative spatial information about game objects in the game scene, and use this information to achieve the goal of the game in the same way than a visually paired individual would do it.

Our technical implementation consisted of two haptic displays, one plotting the location of the “targets” objects in the player’s arm and another one applying continuous vibrotactile feedback in the other arm to represent the moving crosshairs pointer in the screen. Players could infer the right time to shoot by aligning the crosshairs tactile sensation in one arm with the relative position of the next available target in the other arm.

This approach demonstrated how techniques that were used in the fields of telepresence and tele-operation can be applied to videogames. In addition, although this particular implementation was aimed at improving accessibility, the same concepts can be used in other games for sighted individuals to represent additional information, when the visual channel is already close to being overloaded.

5.6 Conclusion

Our research work has produced an architecture for design and development of adaptive, multimodal, exertion games that primarily support an increased level of exercising during game sessions, while still providing a rich multimedia entertaining experience.

We have shown the feasibility of systems of these characteristics by making the implementation of an exertion game that is driven by an adaptive algorithm used to control the heart rate of an exertion-game player, modifying the rules and properties of the videogame in order to prompt the desired physiological repose.

Additionally we have demonstrated the multimodal capabilities of the cited architecture and derived game. We have done this by implementing three different kinds of exertion interfaces.

The first one was based on a specialized framework for rapid integration of smartphones as exertion interfaces into pre-existing videogames, without the need to modify the videogame’s source code. The interface consisted of two smartphone applications, used as interfaces to drive actions in an exertion game.

The second one provided a gestural interface to control a videogame using natural hand gestures. And the third one was based on a specialized hardware videogame device implemented by two haptic displays that enable spatial perception of game objects in order to improve accessibility of computer games for the visually impaired.

6 Architecture validation and user studies

6.1 Introduction

The evaluation process was carried out considering two scopes.

In first instance we wanted to validate the main premise of the proposed adaptive architecture. The primary goal of the proposed architecture was the definition of guidelines for the development of exertion games that would adapt to the different conditions of physical fitness of the individuals engaged in game play. Based on this premise, a game derived from the architecture design should be able to drive the heart rate of an individual to the desired target zone. Additionally the gaming experience should remain as an enjoyable and entertaining activity.

Second, we wanted to validate the design principles and usability of the interfaces that were implemented for the exertion games. Our goal here was to obtain feedback from players engaging on the game, to assess their perception of how natural each interface felt and if they seemed to experience any usability annoyances, or performance issues while interacting with them. Finally we wanted to assess the users' intention of use of such technologies.

In the following sections we describe the methodology of evaluation, respective results and discussion, for both of the scopes described above.

6.2 Validation of the adaptive architecture components

With the purpose of assessing the effectiveness of our adaptive game implementation, we designed an experimental evaluation for the system. In our experiment we wanted to assess how effective was the adaptive version of the game to take the player's level of exertion to a pre-defined target zone, when compared to a non-adaptive version of the same game. In this section we present a description of the protocol for our experiment.

6.2.1 Experiment variables and hypothesis

The experiments were carried out with the game modality (non adaptive vs. linear adaptation vs. fuzzy adaptation) as the independent variable, and the level of exertion of the player as the dependent variable. The modality of the experiment was within subjects, where all participants were subjected to both conditions of the independent variable.

We established the hypothesis that players would have higher levels of exertion when playing the adaptive game, compared to their own game sessions playing the non-adaptive version of the game. We expected that the level of exertion should remain generally within the boundaries of the specified target zone for the session, which was estimated using the 3-minute step test. Furthermore the fuzzy adaptation mechanism is expected to take the player to the target zone and keep them there in a more steady change rate with less fluctuation.

6.2.2 Participants and experimental setting

For this experiment we had 2 volunteers to participate as subjects. We will refer to them as subject 'A' and subject 'B'. Both subjects were male individuals, ages 41 and 35, with reportedly general good health conditions. 'A' was a non smoker and 'B' was a smoker. Neither of the participants reported being an avid video game player or practicing regularly any particular sport.

The experiments were carried out in an isolated area of the computing lab to avoid distractions. The equipment consisted of the stationary bicycle, the attached heart rate monitor, a 21 inch wide screen LCD monitor, and stereo speakers located in front of the bike (Figure 5.1). Additionally a step bench with a height of 12 inches, a stopwatch and a metronome were used for the 3 minute step test.

6.2.3 Assessment of participant's fitness and target heart rate zones

6.2.3.1 The 3-minute step test

Before the actual experiment, participants were subjected to the 3-minute step test, which is aimed at estimating the overall fitness level of the participant (cardiovascular fitness).

In this test, the individual steps up and down a 12-inch bench for 3 minutes, keeping a steady peace that is given by a metronome. At the end of the 3 minutes the heart beats are counted for a whole minute and the resulting heart rate tells generally the level of fitness of the individual. In our experiment the results were 90 and 112 for subjects 'A' and 'B' respectively, which according to the guidelines published by the YMCA have an equivalency of "Above Average" and "Below Average" [Golding_02].

6.2.3.2 The Karvonen method

The adaptive version of the game was pre-programmed with an estimated target heart rate zone specific to each participant. The target zone was estimated using the Karvonen method, which in turn is based on the maximal heart rate (HR_{max}) value of an individual. There have been a lot of studies to determine the best way to calculate the HR_{max} factor [Robergs and Landwehr_02]. In our experiment we have chosen the formula defined by Inbar and others which has reportedly been the most used general equation [Inbar *et al*_94]:

$$HR_{max} = 205.8 - 0.685 * (Age) \qquad (6.1)$$

Once the HR_{max} has been estimated, we calculate the Target Heart Rate (THR) zones using the well known Karvonen method, which considers the resting heart rate [Karvonen_11]:

$$THR = ((HR_{max} - HR_{rest}) * y) + HR_{rest} \qquad (6.2)$$

Where HR_{max} is estimated as per equation 6.1, HR_{rest} is the heart rate reading of the player at rest and 'y' is the intensity desired for the exercise expressed as a percentage. In our experiment we used an intensity of 65% to 85% of the HRmax which resulted in target zones of 138-161 and 141-164 for subjects 'A' and 'B' respectively.

6.2.4 Procedure

The experiment was executed one participant at a time. At the beginning of the session, the participant was introduced to the overall procedure of the experiment, and a demonstration of the game and the general instructions to play were presented. Then the participant was given 5 minutes to do some warm up and stretching off the bike, then they did another five minute warm-up in the bicycle alone without having the game on. Once the warm-up period ended the game started and the player was left to play the game for 5 minutes. At the end of the 5 minutes of game-play, the game would be over and the participant was allowed to cool down pedaling on the bike.

This procedure was executed three times with each participant. One time with the adaptive engine disabled (leaving all the game variables at their default values). A second time with the linear adaptive engine enabled, and a third one with the fuzzy adaptation approach.

6.2.5 Data Collection

There were two methods for data collection. One was the embedded logic in the video game to collect data from the bicycle interface and from the game variables and parameters. The second one was a subjective method to get the participant's perceived level of exertion.

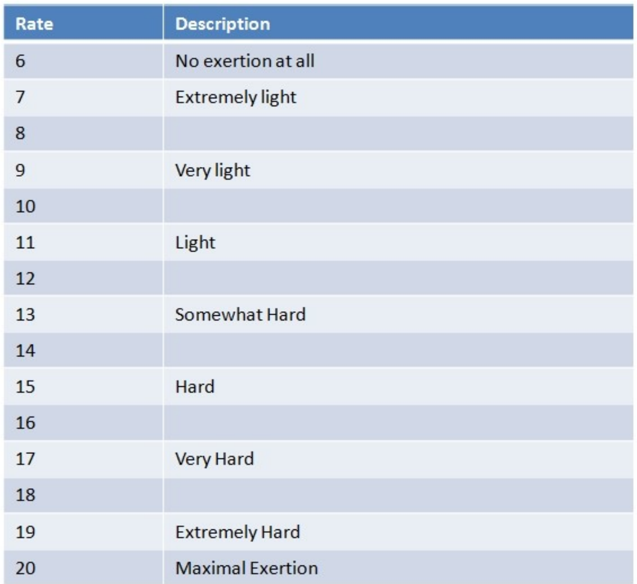
The game was programmed to keep a log of several game variables and input variables coming from the bicycle interface. For every second of the game session we logged the following information:

1. Elapsed time (in seconds)
2. Heart rate
3. Current speed
4. Current resistance setting of the bike (in watts)
5. Height ratio variable value
6. Horizontal cross hairs speed
7. Inter-Target Distance
8. Shots on target (in that second)
9. Shots on Target (accumulated)
10. Bonus (in that second)
11. Bonus(Accumulated)

From the above, items (1) to (3) are considered environment readings, and in particular the heart rate is the objective measurement of the current level of exert of the subject. Items (4) to (7) correspond to the variables managed by the adaptation engine as described in the previous section, and the final four are meant to assess the performance of the user with respect to the game itself.

In addition to these objective measurements of the game and player conditions, we wanted to gather feedback from the participant's themselves about how hard they felt they were working out. For this purpose we used Borg's Rating of Perceived Exertion scale [Borg_70]. We used the 15 point version of the scale as it is shown on table 6.1.

Participants were asked to rate their perceived exertion level three times during the game session, at elapsed minutes one, three and five.



Rate	Description
6	No exertion at all
7	Extremely light
8	
9	Very light
10	
11	Light
12	
13	Somewhat Hard
14	
15	Hard
16	
17	Very Hard
18	
19	Extremely Hard
20	Maximal Exertion

Table 6.1 Borg's Rate of Perceived Exertion (RPE) scale

6.2.6 Results and discussion

For the objective part of our evaluation we wanted to observe how the hart rate signal behaved over the course of the game session, by comparing the non-adaptive vs. the two version of the adaptive game.

Figures 6.1 and 6.2 show the heart rate readings for every second, over the course of the five minute game session for subjects 'A' and 'B' respectively. The 'X' axis shows the elapsed time in minutes, and the 'Y' axis shows the recorded heart rate value in beats per minute (bpm). Each one of the series corresponds to each version of the game, adaptive (linear and fuzzy) and non-adaptive. The bold horizontal lines depict the boundaries of the Target Heart Rate zone for the subject.

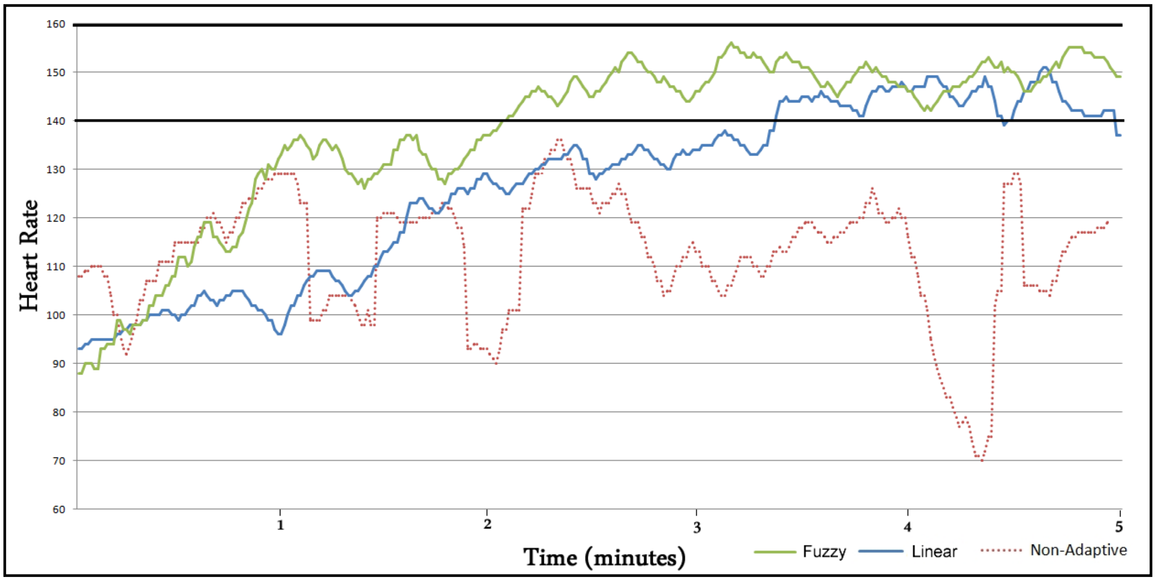


Figure 6.1 Subject 'A' heart rate readings

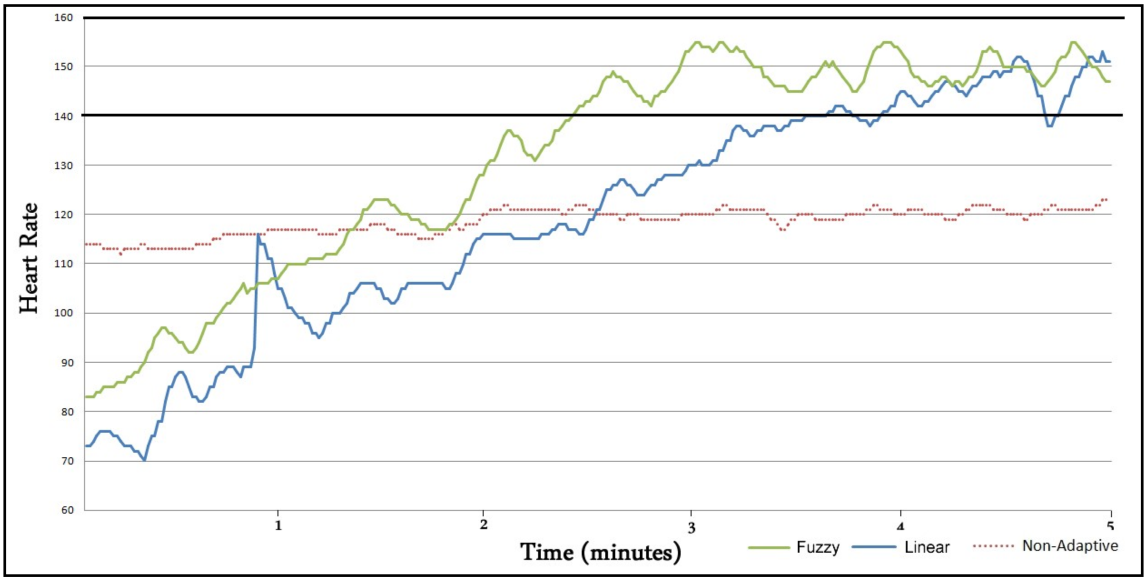


Figure 6.2 Subject 'B' heart rate readings

Although there are natural differences between one individual to another in terms of the heart rate pattern, it can be seen that in both cases the non-adaptive version of the game did not caused the heart rate to reach the target zone. Whereas the adaptive version made subject ‘A’ reach the zone a few seconds after the third minute for the linear algorithm and just before minute 2 for the fuzzy approach. Subject ‘B’ reached the target at about minutes 2.5 and 4 for fuzzy and linear algorithms respectively. It can also be observed that the non-adaptive game keeps the heart rate around the same levels during the session (even if fluctuating), while with the adaptive version, the heart rate starts low but continues to climb constantly to reach the target zone.

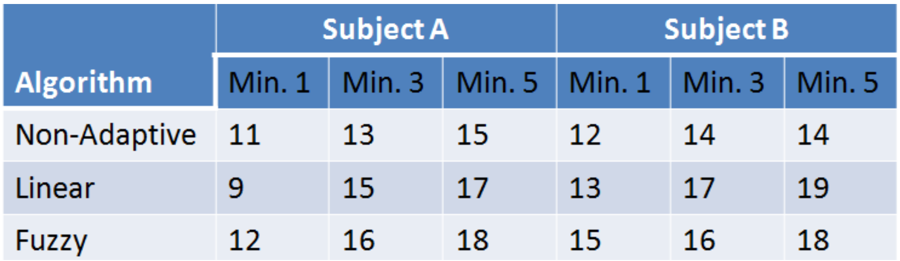
When examining the heart rate pattern derived from both adaptive algorithms, it can be observed that the fuzzy implementation made the heart rate rise faster to the target zone. And, as opposed to the linear adaptation, the fuzzy implementation kept the heart rate in a smoother pattern, closer to the middle point of the target heart rate zone.

Although it is clear that the target zone was reached, our current experiment was unable to show the effect on a session of longer period, so we could confirm that the heart rate stays below the upper boundary of the target zone as it should be according to the adaptation algorithm.

Another observable difference is the variability on heart rate in subject ‘A’ when compared to subject ‘B’, particularly on the non-adaptive version of the game. We attribute this constant change in heart rate of subject ‘A’ to his better fitness level and to his particular technique to tackle the game objectives, which allowed him to have periods of aerobic recuperation, where his heart rate would fall to lower levels. However even with those qualities, subject ‘A’ still has a climbing pattern on his target readings for the adaptive version of the game.

The speed recordings of the game session are consistent with this greater level of effort from the subjects when playing the adaptive game. Subject ‘A’ average speeds were 30km/h for the non-adaptive game, 35km/h for the linear version of the adaptation and 33km/h for the fuzzy algorithm. Subject ‘B’ speed readings were 26km/h, 32km/hand 36km/h respectively.

For the subjective measure of the level of exertion of the subjects, Table 6.2 shows the reported rates according to the Borg's scale presented in the previous section. For each subject we can see the reported rates on each modality of the game for minutes 1, 3 and 5 of the session.



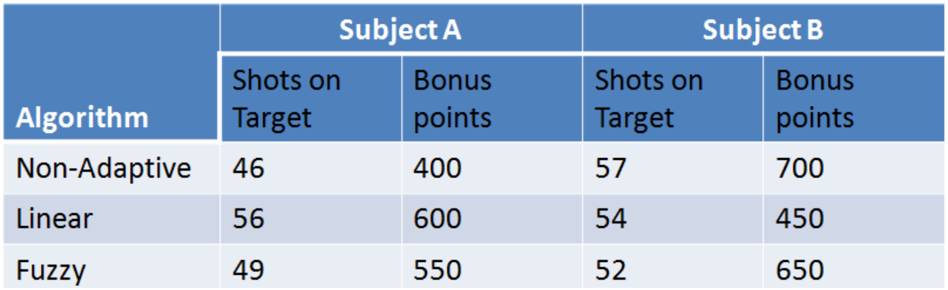
Algorithm	Subject A			Subject B		
	Min. 1	Min. 3	Min. 5	Min. 1	Min. 3	Min. 5
Non-Adaptive	11	13	15	12	14	14
Linear	9	15	17	13	17	19
Fuzzy	12	16	18	15	16	18

Table 6.2 Rates of perceived exertion for the game sessions.

We find these reported rates to be consistent with the objective measurements of the heart rate of both individuals. In both cases, the ratings for the non-adaptive session are lower than the ones of the adaptive session, particularly those reported at minute five. Also the difference between the lowest and the highest rating is much greater in the adaptive version of the game. In the case of subject 'A', the differences were 4, 8 and 6 for 'non-adaptive', 'linear' and 'fuzzy' respectively, while for subject 'B' the differences were 2, 6, 3. It can be observed that in the case of the fuzzy algorithm the range of those values is less than the linear algorithm. This can be explained because of the nature of the fuzzy algorithm to demand more from the user at the beginning when the heart rate is the farther from the target heart rate zone.

Although objective level of exertion was well within the target zone, the perceived exertion levels reported by the subjects were significantly high by the end of the five minute session (17, 18 and 19). This numbers correspond to descriptions between "Very hard" and "Extremely hard" in the Borg's scale. This is a point of concern, since users should exercise, but they should not feel like they are working out in an extreme manner if they are to sustain that level of exertion for more than the five minutes we allocated for our experimental session. The reason for this is that the adaptive approach causes the bike resistance to increase at all times, giving little time for the heart rate to reflect the real level of effort of the player, and so, by the time the resistance is set to be reduced, the player is already exerting very hard. Too overcome this, the step size of the bike resistance could be tweaked to increase and decrease at a lower rate.

In terms of performance on the game itself, table 6.3 shows the values for cumulated shots on target and bonus points gained during the game.



Algorithm	Subject A		Subject B	
	Shots on Target	Bonus points	Shots on Target	Bonus points
Non-Adaptive	46	400	57	700
Linear	56	600	54	450
Fuzzy	49	550	52	650

Table 6.3 Shots on target and bonus points acquired during the game sessions.

Subject ‘A’ did better when playing the adaptive game, while subject ‘B’ did better on the non-adaptive version. From the data we are unable at the time to conclude that the game adaptation had any particular effect on the performance of users on the game.

6.3 Usability studies of multimodal interfaces

6.3.1 Methodology

We evaluated the exertion interfaces through user studies. For each type of interface, a set of users participated in gaming sessions, answered a questionnaire and were interviewed. The questionnaires consisted of a 5-point Likert scales with 10 items including reversal statements [Likert_32]. Likert scales are typically used to measure psychometric response to assess the attitude that a user has towards some topic or experience (in this case with technology), and are applied often as the items in a questionnaire. It has been suggested that Likert scales are among the most common methods to estimate usability [Dumas_98].

In our case, Likert items were concerned with the player perception of ease of use, accuracy and overall performance of the interface being evaluated. Particularly we wanted to assess how natural the interactions were to them and if they perceived any performance issues such as difficulty to perform the tasks in hand or delays in those changes.

Responses to the individual items of the questionnaire are considered as ordinal data, which is a common practice with Likert scales. This is due to the fact that, although there are response levels for each question, it cannot be guaranteed that the participants perceive the difference between two adjacent levels to be equal. For example the difference between ‘agree’ and ‘strongly agree’ does not necessary equals to the difference between ‘agree’ and ‘neutral’ in the mind of the participant. Hence, we present the results for these questionnaires in terms of descriptive statistics that when analyzed can help to draw a picture of the perception that the users got for each topic covered by the items of the questionnaire.

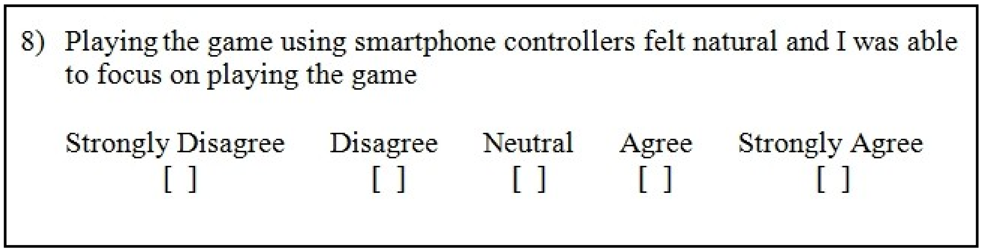
The interviews were guided by five open questions that would cover closely the same topics of the questionnaire except that they were meant to encourage the participant to elaborate verbally about their overall experience. All questionnaires and interview templates can be found on the appendix section.

6.3.2 User Study 1: Smartphone controllers

6.3.2.1 Procedure

We conducted a testing session where 10 participants played the game using the implemented smartphone interfaces. At the end of the game sessions participants answered a questionnaire and were interviewed to collect their impressions and their suggestions.

The sessions were carried out with one participant at a time. At the beginning of the session, the participant was introduced to the overall procedure of the experiment, and a demonstration of the game and the general instructions to play were presented. Then the participant was given 5 minutes to do some warm up in the bicycle alone without having the game on. Once the warm-up period ended the game started and the player was left to play the game for 5 minutes. At the end of the 5 minute, the game was over and the participant was allowed to cool down pedaling on the bike. Figure 6.3 shows a sample question from the full questionnaire that can be found on Annex I.



8) Playing the game using smartphone controllers felt natural and I was able to focus on playing the game

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
[ ]	[ ]	[ ]	[ ]	[ ]

Figure 6.3 Sample Likert item from the questionnaire.

The items of the questionnaire were divided into four main areas. Questions 1 to 3 were all related to the 'fire button' application and each one addressed a specific topic: ease of use, perception of delay, and accuracy to perform the task. Questions 4 to 6 similarly addressed the same topics but for the "Tilt-change direction" application. Questions 7 and 8 were in general about the whole smartphone interface to the game to assess if it was comfortable and felt natural and intuitive. Finally questions 9 and 10 were aimed at assessing the willingness of the participant to use a system like such in their daily lives.

6.3.2.1 Results and discussion

In this section we present the results of the questionnaire and the interviews carried out during the evaluation. Table 5.4 shows a summary of the Median, Mode and Range on the data set pertaining to each of the items of the questionnaire.

From this data summary, it can be seen that the lowest value for the median across all questions was 3.5 (between 'neutral' and 'agree') while the highest was 5 ('strongly agree') with a similar tendency for the most repeated responses (mode) with values between 3 and 5. The range on the data sets was never beyond 2 points, which can be interpreted as a sign of a cohesive response pattern.

Table 6.5 shows an aggregation of responses for each of the Likert levels in each question. For each question, we see the total number of responses for each particular level. In Figure 6.4, it can be appreciated the same values in a stacked bar chart with the total number of responses expressed in percentage for each level; in this case levels for which the total number of responses is 0 are not shown.

Question:	**Median**	**Mode**	**Range**
1.- Fire Button –Easy	4	3	2
2.- Fire Button – No Delay	5	5	1
3.- Fire button – Accurate	5	5	1
4.- Tilt – Easy	4	4	2
5.- Tilt – No Delay	3.5	3	2
6.- Tilt – Accurate	5	5	1
7.- Smartphone – Comfortable	4	4	2
8.- Smartphone –Natural	3.5	3	2
9.- Use Intention	4.5	5	2
10.- Use Intention (reversed)	5	5	1

Table 6.4 Median Mode and Range for each Likert item

In terms of comparing the two types of applications, it can be seen that the fire button application got higher rankings. This is particularly true for the “no delay” question. As it was confirmed during the interviews, 3 users expressed that the tilt application was a bit difficult to get used to at first, since there were some instances where they would not be able to make the crosshairs change direction or it would change when it was not intended. Althoug some subjects perceived this lack of response, a close revision of the videos and application logs shows that the application was responding accurately and according to the calibration parameters. Still they do not seem to have perceived the problem as a reason to lose accuracy on knocking down the targets, since the median and mode both show 5 as the rating for that topic.

By looking at the time stamps of the logs, and matching them with the video, at the time were this sort of latency was perceived, we realized that in fact, the users were not tilting their body enough to trigger the action in the game, or accidentally triggered the action by making sudden moves that would increase the g-forces that the application was tracking to calculate the tilt threshold. As for the presence of real network latency, we cannot draw any definitive conclusions since our evaluation was focused on user experience and not on network performance. However, we estimate the likelihood of having latency in the transmission of commands to be relatively low, based on the fact that we were using the 802.11g standard (rated at 54Mbps) in a clean environment (no other traffic) sending messages that were not larger than 16 ASCII characters in length.

Question:	**Strongly Disagree [1]**	**Disagree [2]**	**Neutral [3]**	**Agree [4]**	**Strongly Agree [5]**
1.- Fire Button - Easy	0	0	4	4	2
2.- Fire Button - No Delay	0	0	1	2	7
3.- Fire button - Accurate	0	0	0	4	6
4.- Tilt - Easy	0	0	2	7	1
5.- Tilt - No Delay	0	0	5	3	2
6.- Tilt - Accurate	0	0	1	2	7
7.- Smartphone - Comfortable	0	0	3	6	1
8.- Smartphone -Natural	0	0	5	2	3
9.- Use Intention	0	0	2	3	5
10.- Use Intention (reversed)	0	0	1	3	6

Table 6.5 Aggregation of responses of each Likert level across all 10 questions.

During the interviews, four participants expressed some concern about the fire button application, arguing that in a longer exercising session, it would be uncomfortable, and should probably be replaced by some sort of hardware attached to the bike handles.

When judging the overall experience with the smartphones interface, 60% of the participants expressed a level of 4 (agree that it is comfortable). In terms of how natural and intuitive the interface was, responses were spread across values 3 to 5, with 50% of the respondents staying ‘neutral’ and the other 50% expressing ‘agree’ and ‘strongly’ agree. This was again explained in the interviews, where they expressed that, since they were judging both applications at the same time, the awkwardness of the fire-button application made them give a lower overall rating.

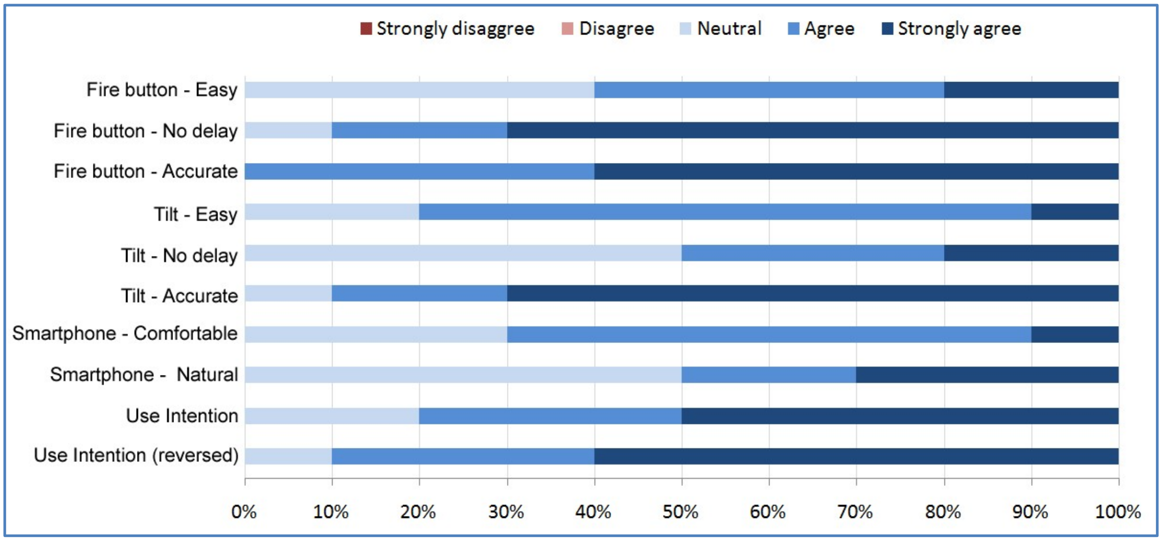


Figure 6.4 Distribution of responses for each Likert level as a percentage of the total number of responses

During the interviews, we could see that participants were highly motivated by the whole idea of exercising while playing a game with a natural interface that does not require traditional input controllers such as game pads. They clearly stated that, if available, they would like to use a system like this on their daily lives. This is consistent with the answers that were given in the questionnaire.

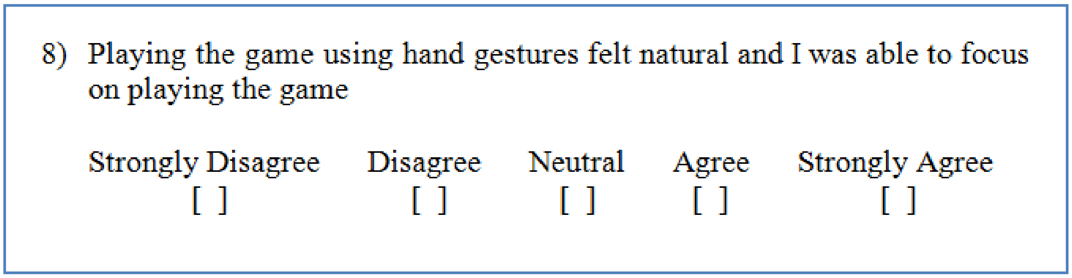
The evaluation presented here is only limited to the particular exertion interface implemented. Our goal was to validate and receive feedback about this initial implementation conceived by using the precepts of our proposed architecture.

6.3.3 User Study 2: Hand gesture control

6.3.3.1 Procedure

With the purpose of getting feedback from actual users about our interface implementation, we conducted a user study, where 5 participants played the game using the hand gesture interface. At the end of the game sessions participants answered a questionnaire and were interviewed to collect their impressions and their suggestions regarding how natural was for them to play using hand gestures to control the actions in the game.

The sessions were carried out with one participant at a time. At the beginning of the session, the participant was introduced to the overall procedure of the experiment, and a demonstration of the game and the different hand gestures were explained. Then the participant was given 5 minutes to do some warm up in the bicycle alone without having the game on. Once the warm-up period ended the game was started and the player was left to play the game for 5 minutes. At the end of the 5 minute, the game would be over and the participant was allowed to cool down pedaling on the bike. Fig. 6.5 shows a sample question from the full questionnaire (Appendix 1).



8) Playing the game using hand gestures felt natural and I was able to focus on playing the game

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
[ ]	[ ]	[ ]	[ ]	[ ]

Figure 6.5 Sample Likert item from the questionnaire

6.3.3.2 Results and discussion

Table 6.6 shows a summary of the Median, Mode and Range on the data set pertaining to each of the items of the questionnaire.

From this data summary it can be seen that the lowest value for the median across all questions was 3 (‘neutral’) while the highest was 5 (‘strongly agree’) with a similar tendency for the most repeated responses (mode) with values between 3 and 5. The range on the data sets was never beyond 2 points, which shows a cohesive response pattern.

Table 6.7 shows an aggregation of responses for each of the Likert levels in each question. For each question we see the tally of responses and a second row shows the same values expressed in percentage.

Question:	Median	Mode	Range
1.- Left/Right Gesture – Easy	4	4	2
2.- Left/Right Gesture – No Delay	3	3	1
3.- Left/Right Gesture – Accurate	4	4	2
4.- Shoot Gesture – Easy	5	5	2
5.- Shoot Gesture – No Delay	3	4	2
6.- Shoot Gesture – Accurate	4	4	2
7.- Hand Gestures - Comfortable	3	3	1
8.- Hand Gestures -Natural	4	3	2
9.- Use Intention	4	4	1
10.- Use Intention (reversed)	5	5	1

Table 6.6 Median Mode and Range for each Likert item

From the details of the distribution of responses in Tables 6.6 and 6.7, it can be seen that both modalities of hand gestures (‘Left/Right’ and ‘Shoot’) had a low rating (between 2 ‘disagree’, and 3 ‘neutral’) in response to the statement of ‘not perceived delay’. According to the responses, users did not generally agreed, and presented a ‘neutral’ perspective to the statement that said that there was no-delay. However, in spite of this perception, users did not seem to be affected in terms of their perceived accuracy to knock down the targets (the goal of the game) which they rated in the ‘agree’ level with 80% of the responses for the ‘left/right’ gesture and with 80% of the responses spread between ‘agree’ and ‘strongly agree’ for the ‘shoot’ gesture.

In terms of ease of use, both types of hand gestures were rated generally in the same order, between ‘agree’ and ‘strongly agree’. This trend was reinforced during the interviews by the users, who expressed that the method of interaction was simple and intuitive.

When judging if the use of the hand gestures was comfortable for directing the actions on the exertion game, users expressed some reserves. The median value was at 3 with a range of 1, and 60% of the responses expressing a ‘neutral’ position and 40% in ‘agree’. During the interviews users expressed that, even for an exertion game, they think that keeping the hand raised at all times for a longer game session would be quite uncomfortable, which is consistent with their responses on the questionnaire. Still, in terms of how natural the interaction was to them, the median and mode values were at four with a low range of 1, which shows that the mode of interaction was intuitive.

Question:	**Strongly Disagree [1]**	**Disagree [2]**	**Neutral [3]**	**Agree [4]**	**Strongly Agree [5]**
1.- Left/Right Gesture - Easy	0	1	1	3	0
		20%	20%	60%	0%
2.- Left/Right Gesture - No Delay	0	0	3	2	0
			60%	40%	0%
3.- Left/Right Gesture - Accurate	0	1	0	4	0
		20%	0%	80%	0%
4.- Shoot Gesture - Easy	0	0	2	0	3
			40%	0%	60%
5.- Shoot Gesture - No Delay	0	2	1	2	0
		40%	20%	40%	0%
6.- Shoot Gesture - Accurate	0	0	1	3	1
			20%	60%	20%
7.- Hand Gestures - Comfortable	0	0	3	2	0
			60%	40%	0%
8.- Hand Gestures -Natural	0	0	2	1	2
			40%	20%	40%
9.- Use Intention	0	0		3	2
			0%	60%	40%
10.- Use Intention (reversed)	0	0	0	2	3
			0%	40%	60%

Table 6.7 Aggregation of responses of each Likert level across all 10 questions

Overall from the interviews we could see that users were highly motivated to use the hand gesture control and this can also be seen on the answers to the last two items of the questionnaire, which show an intention of use with a rating of 5 “strongly agree” in the Likert scale.

6.4 Conclusion

In this chapter we have presented the results of our evaluation. First, we have validated the primary goal of the proposed architecture. In our evaluation of the game we have assessed the effectiveness of the adaptation strategy to reach a predefined target heart rate zone when compared to a traditional non-adaptive version of the same game.

From our preliminary results we have found that the adaptation strategy of the game was able to take the players to the desired target zones, in contrast to the non-adaptive version which kept the users heart rate below that range. We found this trend consistent with the perceived exertion level that the subjects reported during the game sessions. Results at this stage of our research are encouraging and future work could include further evaluations including a greater number of subjects and longer game sessions.

In terms of evaluating the usability of the exertion interfaces, we performed user studies to get feedback from users about three main factors: usability, overall performance of the interface and people’s willingness to adopt the technology. Based on the findings from questionnaires and interviews, we can argue that users where very kin about trying new ways to interact with the exertion game. They regarded the technology favourably and they did not have any major complications to control the game through the proposed interfaces. Finally they expressed clear intention to use the technology if it was available to them.

7 Conclusion and Future Work

7.1 Conclusion

In this thesis we have presented research work done towards the design and development of architecture and related systems that support a shift in the exertion game paradigm. The contributions of this work are enablers in the design and development of exertion games with a strict serious game approach. Such games should have “exercising” as the primary goal, and a game engine that has been developed under this scheme should be aware of the exertion context of the player. The game should be aware of the level of exertion of the player and adapt the gaming context (in-game variables and exertion interface settings) so that the player can reach a predefined exertion rate as desired.

To support such degree of adaptability in a multimedia, multimodal system, we have proposed a system architecture that lays down the general guidelines for the design and development of such systems. The architecture proposes three main components to support the scenario described before.

An adaptation module provides for the registration of ‘managed variables’ from the underlying game engine. Such variables are then modified automatically by the adaptation process as required by the exergame and according to the context information of the player. This provides transparent access to the exertion interface and provides free functionality of adaptation to the game engine.

An exertion estimation module employs two algorithmic approaches to calculate the magnitude and nature of change required from the game engine to meet the exertion requirements of the interface. One of the algorithms involved in this module uses a linear approach that is easy to implement and mathematically simple. A second fuzzy logic control algorithm provides a more robust approach to estimation and control of the level of exertion of the player. Its fuzzy nature requires a more complex mathematical model, but helps to ease the ‘overshooting’ effect of the linear algorithm.

Finally a natural interface processing module provides a middle layer of communication for game engines to support multimodal natural interfaces.

We have tested our proposed architectural guidelines by realizing a practical implementation of an exertion game and validated our implementation with an experiment where users played the game in both modalities: adaptive and non-adaptive.

The system incorporates multimodal exertion interfaces. In particular we have researched the usage of smartphones as input controllers, leveraging on the internal sensors available in such devices. We have also explored the use of natural interfaces with hand and body gesture recognition to control the actions in the game. We also have improved the accessibility of the system by implementing a haptic interface that enables visually impaired individuals to play the game by doing sensory substitution and spatial representation of game objects in the haptic realm.

We conducted user studies to better inform and evaluate our interface design decisions. The studies were based on experimental testing of the interfaces by users, who then were interviewed and answered questionnaires aimed at evaluating the ease of use, performance of the interfaces and intention of use from the participants.

Our findings have been positive and encouraging. We have demonstrated with an experiment that the adaptive version of the algorithm can drive the heart rate of an individual to reach a target zone that is above the normal heart rate that the same individual would have when playing a non-adaptive version of the same game. In particular, our fuzzy logic implementation has demonstrated that the level of the pattern of heart rate of the player is more stable than when employing a linear algorithm. Feedback given by users during the interviews of the studies is also positive. The valuable input from our participants has been recorded and discussed in the results section of this work.

7.2 Future Work

The results obtained are promising, and although control of the heart rate of the player is pretty unpredictable, the fuzzy implementation has proven to overcome the uncertainty.

However, experiments performed so far included only short exertion sessions, as the intention was only to prove the feasibility of adapting the dynamics of the game engine to the player's exertion context. In future work experiments with longer exertion sessions could be key in detecting possible adjustments to the rule set used in the fuzzy model. In particular the error-rate could include more linguistic variables that describe more accurately the speed of change happening in the heart rate. This additional input for the inference process could help the control algorithm to better adapt and render a more smooth function of control.

In terms of estimating how much the player is exercising at any given time, more physiological sensors could be included, such as respiration rate, transpiration, oxygen level measurements and others. Although the architecture design proposed here does account for additional sensors, in practice, an intelligent algorithm could be built to combine all signals into a crisp output, or even account the eventuality of losing some of the readings for a particular sensor.

Another possible path for further research could be to compare the use of different exertion interfaces. All of the experiments performed in this research were done using an exertion bike. However, other cardiovascular machines could be employed (i.e rower, treadmill etc.) and results compared. Similarly the adaptation capabilities of this development could be tested with other videogame titles to assess any possible shortcomings to the current implementation that are related to the nature of the computer game being supported.

Finally the fuzzy logic model proposed here could be taken from software to firmware. In this case, the adaptation capabilities of exertion games as proposed here could be incorporated into off-the-shelf videogame consoles and not only on computer videogames.

References

[Abraham and Nath_00] Abraham A. and Nath B., Hybrid intelligent systems: A review of a decade of research, School of Computing and Information Technology, Faculty of Information Technology, Monash University, Australia, Technical Report Series, 5(2000), 1-55.

[ABI_10] "Mobile Device User Interfaces. Touchscreens, Software, Accelerometers, Haptics, and Sensors". 2010. ABI Research. Technology Market Repor.

[Abt_70] Abt C. C., Lanham MD, Serious Games.: University Press of America, 1987. (Reprint. Originally published: New York: Viking Press, 1970.)

[Adelstain et al._03] Adelstein B. D., Begault D. R., Anderson M. R., and Wenzel E. M. 2003. Sensitivity to haptic-audio asynchrony. In Proceedings of the 5th international Conference on Multimodal interfaces. ICMI '03. ACM, New York, NY, 73-76.

[AHA_12] American Heart Association. Recommended heart rate target zones. http://www.heart.org/HEARTORG/GettingHealthy/PhysicalActivity/Target-Heart-Rates_UCM_434341_Article.jsp

[Alamri *et al*_08] Alamri A., Eid M., Iglesias R., Shirmohammadi S., and El Saddik A., "Haptic virtual rehabilitation exercises for poststroke diagnosis," IEEE Trans. Instrum. Meas., vol. 57, no. 9, pp. 1876–1884, Sep. 2008.

[Allan and Kristofferson_74] Allan L. G. and Kristofferson A. B. 1974. Successiveness discrimination: Two models. Perception & Psychophysics. 15,1, 37-46.61-72.

[Alles_70] Alles, D. S. (1970). Information transmission by phantom sensations. IEEE Transactions on Man-Machine Systems, MMS-11(1), 8–91.

[Alves *et al*_04] Alves, N., Cndido, P. L., & Frazao, R. (2004). Innervation of the pronator quadratus muscle. International Journal of Morphology, 22(4), 253–256.

[America's army] http://www.americasarmy.com/ Accessed April, 2010.

[Audiogames_12] Audio Games. Games based on sound. http://audiogames.net/. Retrieved January, 2012

[Bákésy_58] Bákésy, G. V. (1958). Funneling in the nervous system and its role in loudness and sensation intensity on the skin. The Journal of the Acoustical Society of America, 30(5), 399–412.

[Baklouti *et al*_08] Baklouti M., Monacelli E., Guitteny V., Couvet S., “Intelligent assistive exoskeleton with vision based interface”, in “Proceedings of the 6th international conference on Smart Homes and Health Telematics”, Vol. 5120. 123–135, 2008.

[Basori *et al.*_08] Basori, A. H., Daman, D., Bade, A., Sunar, M. S., and Saari, N. 2008. The feasibility of human haptic emotion as a feature to enhance interactivity and immersiveness on virtual reality game. In *Proceedings of the 7th ACM SIGGRAPH international Conference on Virtual-Reality Continuum and Its Applications in industry* (Singapore, December 08 - 09, 2008). VRCAI '08. ACM, New York, NY, 1-2. DOI= http://doi.acm.org/10.1145/1477862.1477910

[Bergen_98] Bergen, D. 1998. *Play as a Medium for Learning and Development, p. 7*.

[Berkovsky et al_09] Berkovsky S., Bhandari D., Kimani S., Colineau N., and Paris C., “Designing games to motivate physical activity”, Proc. ACM International Conference on Persuasive Technology, NY, USA, 2009, pp 1 – 4.

[Björk_08] Björk, S. 2008. Games, gamers, and gaming: understanding game research. In *Proceedings of the 12th international Conference on Entertainment and Media in the Ubiquitous Era* (Tampere, Finland, October 07 - 09, 2008). MindTrek '08. ACM, New York, NY, 64-68. DOI= http://doi.acm.org/10.1145/1457199.1457213

[Blunt_09] Blunt, R. 2009. Do Serious Games Work? Results from Three Studies. *eLearn* 2009, 12 (Dec. 2009) DOI= http://doi.acm.org/10.1145/1661377.1661378

[Bogost_07] Bogost, I. 2007 *Persuasive Games: the Expressive Power of Videogames*. The MIT Press.

[Bogost_10] CNN Campaign Rush Serious Game
http://www.persuasivegames.com/games/game.aspx?game=campaignrush Accessed April, 2010

[Borg_70] Borg, G, "Perceived Exertion as an indicator of somatic stress", Scandinavian journal of Rehabilitation Medicine 1970, 2(2), 92-98

[Boring *et al*_09] Boring S., Jurmu M., and Butz, A. 2009. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7 (OZCHI '09). ACM, New York, NY, USA, 161-168.

[Broeren et al_08] Broeren, J., Bjorkdahl, A., Claesson, L., et al. Virtual rehabilitation after stroke. Studies in Health Technology and Informatics 136, (2008), 77-82.

[Cha *et al*_08] Cha, J., Rahal, L., & El Saddik, A. (2008b). A pilot study on simulating continuous sensation with two vibrating motors. In Proc. IEEE Int. Workshop on Haptic Audio Visual Environments and Games (pp. 143–147).

[Cha *et al*_09] Cha, J.; Eid M.; Barghout A.; Mahfujur Rahman ASM; El Saddik A.. (2009). HugMe: synchronous haptic teleconferencing. In Proceedings of the 17th ACM international conference on Multimedia (MM '09). ACM, New York, NY, USA, 1135-1136.

[Cha *et al*_10] Cha J., Barghout A., Kammerl J., Steinbach E., and El Saddik A., “Improving Spatial Perception in Telepresence and Teleaction Systems by Displaying Distance Information through Visual and Vibrotactile Feedback”, Presence: Teleoperators and Virtual Environments, Vol 19(5), 2010

[Chen et al_08] Chen Q., Georganas N., Petriu E., “Hand Gesture Recognition Using Haar-Like Features and a Stochastic Context-Free Grammar,” IEEE Transactions on Instrumentation and Measurement -2008.

[Chen *et al*_11] Chen S., Pan Z., Zhang M. and Shen H., "A case study of user immersion-based systematic design for serious heritage games", Multimedia Tools and Applications, 2011.

[Clarke and Duimering_06] Clarke, D. and Duimering, P. R. 2006. How computer gamers experience the game situation: a behavioral study. *Comput. Entertain.* 4, 3 (Jul. 2006), 6. DOI= http://doi.acm.org/10.1145/1146816.1146827

[Cochan and Chen_05] Cochran J. K., J.K. and Chen, H. (2005). Fuzzy multi-criteria selection of object oriented simulation for production system analysis, Computers and Operations Research, 2(2005), 32, pp. 153-168.

[Comai et al 2010] Comai S., Mazza S., and Mureddu L.. 2010. A haptic-based framework for chemistry education. In Proceedings of the 5th European conference on Technology enhanced learning conference on Sustaining TEL: from innovation to learning and practice (EC-TEL'10), Martin Wolpers, Paul A. Kirschner, Maren Scheffel, Stefanie Lindstaedt, and Vania Dimitrova (Eds.). Springer-Verlag, Berlin, Heidelberg, 614-619.

[Connected to the Future_03] Connected to the Future: A Report on Children's Internet Use, Corporation for Public Broadcasting,

[Conti et al._05] Conti F, Barbagli F, Morris D, and Sewell C. 2005. CHAI 3D: An Open-Source Library for the Rapid Development of Haptic Scenes. IEEE World Haptics.

[Corti_06] Corti K.. Games-based Learning; a serious business application. WhitePaper. http://www.pixelearning.com/docs/games_basedlearning_pixelearning.pdf Accessed April 2010.

[Csikszentmihalyi_75] Csikszentmihalyi M., Beyond Boredom and Anxiety: Jossey-Bass: San Francisco, CA, 1975.

[Dardas and Georganas_11] Dardas N., Georganas N., "Real Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine

Techniques", IEEE Transactions on Instrumentation and Measurement, VOL. 60, no. 11, pp. 3592 - 3607, Nov 2011.

[Dixon and Mood_48] Dixon W.J.. and Mood A.M., "A method for obtaining and analyzing sensitivity data", Journal of the American Statistical Association. Vol. 43, No. 241, Mar., 1948

[Dixon and Spitz_80] Dixon N. F. and Spitz L. 1980. The detection of auditory visual desynchrony. Perception. 9(6) 719 – 721

[Dumas_98] Dumas, J. (1998b). Usability Testing Methods: Subjective Measures: Part II - Measuring Attitudes and Opinions. October issue of Common Ground, The newsletter of the Usability Professionals' Association, 4-8.

[Egenfeldt-Nielsen_07] Egenfeldt-Nielsen, S. (2007). Att skapa ljuv musik: Det pedagogiska anvandandet av datorspel. In: Linderoth, Jonas (Ed). Datorspelandets Dynamik. Studentlitteratur.

[Eid et al_07] Eid M. A., Mansour M., El Saddik A. H., and Iglesias R. 2007. A haptic multimedia handwriting learning system. In Proceedings of the international workshop on Educational multimedia and multimedia education (Emme '07). ACM, New York, NY, USA, 103-108.

[El Saddik_07] El Saddik A. The Potential of Haptics Technologies IEEE Instrumentation Measurement, February, 2007

[Enriquez *et al*_06] Enriquez M., MacLean K., and Chita C.. 2006. Haptic phonemes: basic building blocks of haptic communication. In Proceedings of the 8th international conference on Multimodal interfaces (ICMI '06). ACM, New York, NY, USA, 302-309.

[Exergames_12] Wikipedia: Exergames. http://en.wikipedia.org/wiki/Exergaming Retrieved April 2012

[EyeToy Kinetic_10] http://en.wikipedia.org/wiki/EyeToy:_Kinetic. Accessed on March, 2010

[EyeToy_11] EyeToy Kinetics. http://en.wikipedia.org/wiki/EyeToy:_Kinetic. Retrieved April 2011

[Facts & Research_10] Facts & Research, Entertainment Software Association, www.theesa.com

[Falaki *et al*_10] Falaki H., Mahajan R., Kandula S., Lymberopoulos D., Govindan R., and Estrin D.. 2010. Diversity in smartphone usage. In Proceedings of the 8th international conference on Mobile systems, applications, and services (MobiSys '10). ACM, New York, NY, USA, 179-194

[Fang *et al*_07] Fang G., Gao W., Zhao D., "Large-vocabulary continuous sign language recognition based on transition-movement models", Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on 37(1), 1–9, 2007.

[FAS_10] http://www.fas.org Accessed March, 2010

[Fjeld and Voegtli_02] Fjeld M. and Voegtli B. 2002. Augmented Chemistry: An Interactive Educational Workbench. In Proceedings of the 1st International Symposium on Mixed and Augmented Reality (ISMAR '02). IEEE Computer Society, Washington, DC, USA, 259-.

[Fujimoto et al_08] Fujimoto T. Ishibashi Y., and Sugawara S. 2008. Influences of inter-stream synchronization error on collaborative work in haptic and visual environments. In Proc. Haptics'08.

[Gamepolitics.com_10] "Korean Govt. Gets Behind Serious Games". http://www.gamepolitics.com/2009/05/20/korean-govt-gets-behind-serious-games. Accessed March, 2010.

[Games for Change_10] http://www.gamesforchange.org. Accessed April, 2010

[Games for Health_10] http://www.gamesforhealth.org. Accessed March, 2010

[GamesHealth_11] Games For Health Initiative. http://www.gamesforhealth.org/. Retrieved April 2011

[Garcia and Coulton_08] Garcia Wylie C. and Coulton P.. 2008. Mobile exergaming. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology* (ACE '08). ACM, New York, NY, USA, 338-341.

[Garcia and Coulton_09] Garcia W.C.; Coulton, P. 2009. Mobile persuasive exergaming. In Proc. ICE-GIC 2009. International IEEE Consumer Electronics Society's Games Innovations Conference, 25--28 Aug. 2009, 126--130.

[Gilbertson *et al*_08] Gilbertson P., Coulton P., Chehimi F., and Vajk T.. 2008. Using “tilt” as an interface to control “no-button” 3-D mobile games. Comput. Entertain. 6, 3, Article 38 (November 2008), 13 pages.

[Göbel *et al*_10] Göbel S., Hardy S., Wendel V., Mehm F, and Steinmetz R. Serious games for health: personalized exergames. In *Proceedings of the international conference on Multimedia* (MM '10). ACM, New York, NY, USA, 1663-1666.

[Golding_02] Golding L.A.. YMCA Fitness Testing and Assessment Manual-4th Edition. Human Kinetics Pub., Inc.; 1 edition (Dec 31 2002)

[Graves et al_07] Graves L, Stratton G, Ridgers ND, Cable NT (2007). "Comparison of energy expenditure in adolescents when playing new generation and sedentary computer games: cross sectional study". BMJ 335 (7633): 1282–4. doi:10.1136/bmj.39415.632951.80. PMID 18156227.

[Gutschmidt *et al*_10] Gutschmidt R., Schiewe M., Zinke F., and Jurgensen H.. 2010. Haptic emulation of games: haptic Sudoku for the blind. In Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '10), Fillia Makedon, Ilias Maglogiannis, and Sarantos Kapidakis (Eds.). ACM, New York, NY, USA, , Article 2 , 8 pages.

[Henrysson *et al*_05] Henrysson A, Ollila M., and Billinghurst M.. 2005. Mobile phone based AR scene assembly. In Proceedings of the 4th international conference on Mobile and ubiquitous multimedia (MUM '05). ACM, New York, NY, USA, 95-102.

[Hogan *et al*_07] Hogan, M., Sabri, H., and Kapralos, B. 2007. Interactive community simulation environment for community health nursing. In *Proceedings of the 2007 Conference on Future Play* (Toronto, Canada, November 14 - 17, 2007). Future Play '07. ACM, New York, NY, 237-240. DOI= http://doi.acm.org/10.1145/1328202.1328248

[Hoysniemi_06] Hoysniemi J. 2006. International survey on the Dance Dance Revolution game. Comput. Entertain. 4, 2, Article 8 (April 2006)

[Huber et al_08] Huber, M., Rabin, B., Docan, C., et al. PlayStation 3-based tele-rehabilitation for children with hemiplegia. Virtual Rehabilitation, 2008, (2008), 105-112.

[Huizinga_38] Huizinga, J. Homo L. Boston, MA: The Beacon Press, 1955.

[Hurst and Wezel_12] W. Hurst, C. Wezel, “Gesture-based interaction via finger tracking for mobile augmented reality”, Multimedia Tools and Applications, 2012.

[Inbar *et al*_94] Inbar, O. Oten, A., Scheinowitz, M., Rotstein, A., Dlin, R. and Casaburi, R. Normal cardiopulmonary responses during incremental exercise in 20-70-yr-old men. Med Sci Sport Exerc 1994;26(5):538-546.

[Jack et al_01] Jack, D., Boian, R., Merians, A.S., et al. Virtual reality-enhanced stroke rehabilitation. Neural Sys. And Rehab. Engr., IEEE Trans. 9, 3 (2001), 308-318.

[Jay et al._07] Jay, C., Glencross, M., and Hubbold, R. 2007. Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment. ACM Trans. Comput.-Hum. Interact. 14, 2, 8.

[Johnston and Whitehead_09] Johnston, H. and Whitehead, A. 2009. Distinguishing games, serious games, and training simulators on the basis of intent. In *Proceedings of the 2009 Conference on Future Play on @ GDC Canada* (Vancouver, British Columbia, Canada, May 12 - 13, 2009). Future Play '09. ACM, New York, NY, 9-10.

[Jukes and Dosaj_06] “Understanding Digital Children (DKs)”. Ian Jukes and Anita Dosaj, The InfoSavvy Group, February 2006

[Just *et al*_06] Just A., Rodriguez Y., Marcel S.. "Hand posture classification and recognition using the modified census transform". In: Proceedings of the Seventh IEEE international conference on automatic face and gesture recognition (FG'06), University of Southampton, UK, pp 351–356, 2006.

[Kameyama and Ishibashi_06] Kameyama S. and Ishibashi Y. 2006. Perception of synchronization errors in haptic and visual communications. In Proc. SPIE Optics East.

[Karvonen_11] Wikipedia- Karvonen. http://en.wikipedia.org/wiki/Heart_rate#Karvonen_method. Retrieved June 2011.

[Kiili and Merilampi_10] Kiili K. and Merilampi S. 2010. Developing engaging exergames with simple motion detection. In Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek '10). ACM, New York, NY, USA, 103-110.

[Kimbler_10] Kimbler, K. App store strategies for service providers. In proceeding of the 14th International Conference on Intelligence in Next Generation Networks (ICIN), 2010

[Kinect_11] Microsoft Kinect. http://en.wikipedia.org/wiki/Microsoft_Kinect. Retrieved April 2011

[KinectSurgeons_12] "Toronto Hospital Using Kinect To Help Surgeons Call Up Images During Operations" http://www.gamasutra.com/view/news/33598/

[Kostomaj and Boh_09] Kostomaj M., Boh B. "Design and evaluation of user's physical experience in an Ambient Interactive Storybook and full body interaction games". Multimedia Tools and Applications, 54 (2), pg. 499-525, August 2011.

[Kotick_09] Web link. "Kotick: World games industry worth $55bn by 2012". http://www.mcvuk.com/news/34469/Kotick-World-games-industry-worth-55bn-by-2012. Accessed March, 2010

[Kray *et al*_10] Kray C., Nesbitt D., Dawson J., and Rohs M.. 2010. User-defined gestures for connecting mobile phones, public displays, and tabletops. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10). ACM, New York, NY, USA, 239-248.

[Levitin et al._00] Levitin, D. J., MacLean, K., Mathews, M., Chu L, and Jensen, E. 2000. The perception of cross-modal simultaneity. In Proceedings, Computing Anticipatory Systems. AIP Conf. Proc. 517, 323-329.

[Levitt_70] Levitt, H. 1970. Transformed up-down methods in psychoacoustics. J. Acoustical Soc, of America. 49, 2, 467-477.

[Li *et al*_11] Li K., Du Y., and Fu Z.. “TreeHeaven: a table game using vision-based gesture recognition”. In Proceedings of the 2011 ACM symposium on The role of design in UbiComp research & practice (RDURP '11). ACM, New York, NY, USA, 11-14. 2011.

[Lieberman and McHugh_01] Lieberman, L.J. & McHugh, B.E. (2001). Health related fitness of children with visual impairments and blindness. Journal Of Visual Impairment and Blindness, 95(5), 272-286.

[Lieberman *et al*_06] Lieberman, L.J., Modell, S., Jackson, I. (2006). Going PLACES: A transition guide to physical activity for youth with visual impairments. Louisville, KY: American Printing House for the Blind.

[Lieberman_01] Lieberman, D. A. 2001. Management of Chronic Pediatric Diseases with Interactive Health Games: Theory and Research Findings. Journal of Ambulatory Care Management, 24(1), 26--38.

[Likert_32] Likert, R. (1932). "A Technique for the Measurement of Attitudes". Archives of Psychology 140: 1–55.

[Lode et al._10] Lode V., De Boeck J. and Coninx K.. The Phantom versus the Falcon: Force Feedback Magnitude Effects on User’s Performance during Target Acquisition. Haptic and Audio Interaction Design. Lecture Notes in Computer Science, 2010, Volume 6306/2010, 179-188

[Luke_05] LUKE, R. C. 2005. Oxygen cost and heart rate response during interactive whole body video gaming. Masters Thesis, California State University, Fresno

[Macedonia_01] Macedonia, M. Games, Simulations, and the Military Education Dilemma (2001).

[MacLellan et al_09] G. MacLellan, L. Baillie, and M. Granat. “The application of a physical activity and location measurement system to public health interventions to promote physical activity”, Proc. ACM International Conference on Pervasive Technologies Related to Assistive Environments, New York, USA, 2009, pp. 1 - 8.

[Marsh et al_05] Marsh T., Wong W. L., Carriazo E., Nocera L., Yang K.,Varma A., Yoon H., Huang Y-L, Kyriakakis C, Shahabi C. All User Experiences and Lessons Learned from Developing and Implementing an Immersive Game for the Science Classroom Proceedings of HCI International 2005, Las Vegas, Nevada, USA.

[Marshall et al_04] Marshall, S.J., Biddle, S.J.H., Gorely, T., Cameron, N., and Murdey, I. 2004. Relationships between media use, body fatness and physical activity in children and youth: a meta analysis. International Journal of Obesity, 28, 1238-1246.

[Mayer_09] Mayer, I. S. 2009. The Gaming of Policy and the Politics of Gaming: A Review. *Simul. Gaming* 40, 6 (Dec. 2009), 825-862. DOI= http://dx.doi.org/10.1177/1046878109346456

[McGrath and Summerfield_85] McGrath, M. and Summerfield Q. 1985. Intermodal timing relations and audio-visual speech recognition by normalhearing adults. Journal of the Acoustical Society of America. 77.

[MEMS_11] “MEMS & Sensors for Smartphones”. 2011. Yole Developppement. Market report.

[Michael and Chen_05] Michael, D. R. and Chen, S. L. 2005 *Serious Games: Games that Educate, Train, and Inform*. Muska & Lipman/Premier-Trade.

[Miller *et al*_07] Miller D., Aaron P., and Douglas S. A.. 2007. Finger dance: a sound game for blind people. In Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility (Assets '07). ACM, New York, NY, USA, 253-254.

[Mitchell and Savill-Smith_04] Mitchell, A., & Savill-Smith, C. (2004). *The Use of Computer and Video Games for Learning: A Review of the Literature*. Learning and Skills Development Agency, London: Ultralab. Retrieved April 19th, 2008, from http://www.lsda.org.uk/files/PDF/1529.pdf

[Miyasato et al._96] Miyasato T. Noma H. and Kishino F. 1996. Subjective evaluation of perception of delay time between visual information and tactile information. IEICE Trans. Fundamentals of Electronics, Communications and Computer Science, E79-A, 5,.

[Mokka et al_03] Mokka S., Väätänen A., Heinilä J., and Välkkynen P. 2003. Fitness computer game with a bodily user interface. In Proceedings of the second international conference on Entertainment computing (ICEC '03). Carnegie Mellon University, Pittsburgh, PA, USA, 1-3

[Morelli *et al*_10] Morelli T., Foley J., Columna L., Lieberman L., and Folmer E.. (2010a). VI-Tennis: a vibrotactile/audio exergame for players who are visually impaired. In Proceedings of the Fifth International Conference on the Foundations of Digital Games (FDG '10). ACM, New York, NY, USA, 147-154.

[Morelli *et al*_10b] Morelli T., Foley J., and Folmer E.. (2010b). Vi-bowling: a tactile spatial exergame for individuals with visual impairments. In Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '10). ACM, New York, NY, USA, 179-186.

[Mottola *et al*_06] Mottola L., Murphy A. L., and Picco G.P. 2006. Pervasive games in a mote-enabled virtual world using tuple space middleware. In Proceedings of 5th ACM SIGCOMM workshop on Network and system support for games (NetGames '06).

[Mueller and Agamanolis_08] Mueller F. and Agamanolis S., “Exertion interfaces”, Proc. ACM CHI extended abstracts on Human factors in computing systems, New York, USA, 2008, pp. 3957 - 3960.

[Mueller et al_08] Mueller F., Gibbs M.R., and Vetere F.. 2008. Taxonomy of exertion games. In Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat (OZCHI '08). ACM, New York, NY, USA, 263-266.

[Mueller *et al*_11] Mueller, F., Edge, D., Vetere, F., Gibbs, M. R., Agamanolis, S., Bongers, B., & Sheridan, J. G. (2011). Designing Sports: A Framework for Exertion Games. In CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, Canada.

[My Health Couch_10] http://myhealthcoach.uk.ubi.com/. Accessed March, 2010.

[Narayanasamy et al_06] Narayanasamy V., Wong K.W. , Fung C.C. , Rai S., Distinguishing games and simulation games from simulators, Computers in Entertainment (CIE), v.4 n.2, April-June 2006

[Nickel and Stiefelhagen_07] Nickel K., Stiefelhagen R., "Visual recognition of pointing gestures for human-robot interaction", Image Vision Computing 25(12), 1875–1884, 2007.

[Novint_12] Novint Falcon Haptic Device. Novint Technologies. http://home.novint.com/products/novint_falcon.php. Last accessed May, 2012

[Ntalianis *et al*_10] Ntalianis K., Doulamis A., Tsapatsoulis N. and Doulamis N., "Human action annotation, modeling and analysis based on implicit user interaction", Multimedia Tools and Applications, Volume 50, Number 1, Pages 199-225, 2010.

[Palluel-Germain et al._07] Palluel-Germain R., F. Bara, A. Hillairet de Boisferon, B. Hennion, P. Gouagout, and E. Gentaz. 2007. A Visuo-Haptic Device - Telemaque - Increases Kindergarten Children's Handwriting Acquisition. In Proceedings of the Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC '07). IEEE Computer Society, Washington, DC, USA, 72-77.

[Pan *et al*_10] Pan Z., Li Y., Zhang M., Sun C., Guo K., Tang X., Zhou S.. "A real-time multi-cue hand tracking algorithm based on computer vision". In: Virtual reality conference (VR). IEEE, Piscataway, pp 219–222, 2010.

[Parker et al._09] Parker, J. R., Sorenson, N., Esmaeili, N., Sicre, R. S., Gil, P., Kochlar, V., Shyba, L., and Heerema, J. 2009. The Booze Cruise: impaired driving in virtual spaces. *IEEE Comput. Graph. Appl.* 29, 2 (Mar. 2009), 6-10. DOI= http://dx.doi.org/10.1109/MCG.2009.34

[PHAC_10] Obesity in Canada – Snapshot. Public Health Agency of Canada. http://www.phac-aspc.gc.ca/publicat/2009/oc/index-eng.php Retrieved May, 2010

[Phantom_12] Phantom Haptic Device. Sensable Technologies. http://www.sensable.com/index.htm Last ccessed May, 2012

[Prensky_01] Prensky M., Digital Game-Based Learning. McGraw-Hill, 2001. Chapter 10.

[Rapeepisarn *et al*_06] Rapeepisarn, K., Wong, K. W., Fung, C. C., and Depickere, A. 2006. Similarities and differences between "learn through play" and "edutainment". In *Proceedings of the 3rd Australasian Conference on interactive Entertainment* (Perth, Australia, December 04 - 06, 2006). ACM International Conference Proceeding Series, vol. 207. Murdoch University, Murdoch University, Australia, 28-32.

[Raybourn and Bos_05] Raybourn, E. M. and Bos, N. 2005. Design and evaluation challenges of serious games. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA, April 02 - 07, 2005). CHI '05. ACM, New York, NY, 2049-2050. DOI= http://doi.acm.org/10.1145/1056808.1057094

[Robergs and Landwehr_02] Robergs, R, and Landwehr, R, "The Surprising History Of The 'HRmax= 220-age' Equation', Journal of Exercise Physiology Online, 5(2), 2002

[Romand and Dough_07] Roman, P. A. and Brown, D. 2007. Constructive simulation versus serious games: a Canadian case study. In *Proceedings of the 2007 Spring Simulation Multiconference - Volume 3* (Norfolk, Virginia, March 25 - 29, 2007). Spring

Simulation Multiconference. Society for Computer Simulation International, San Diego, CA, 217-224.

[Roskos_00] Roskos, K.A., & Christie, J.F. (Eds.). (2000). *Play and literacy in early childhood: Research from multiple perspectives.* Mahwah, NJ: Lawrence Erlbaum Associates

[Ruiz *et al*_09]. Ruiz Jr., C. R., Salimi, N. H., Chodnok, S., and Ha, P. T. 2009. Exploring Traditional Art and Culture in Modern Computer Games. In *Proceedings of the 2009 Fifth international Joint Conference on inc, IMS and IDC* (August 25 - 27, 2009). NCM. IEEE Computer Society, Washington, DC, 1311-1316. DOI= http://dx.doi.org/10.1109/NCM.2009.178

[Salmon et al_05] Salmon, J., Ball, K., Crawford, D., Booth, M., Telford, A., Hume, C., Jolley, D., and Worsley, A. 2005. Reducing sedentary behaviour and increasing physical activity among 10-year-old children: overview and process evaluation of the 'Switch-Pay' intervention. Health Promotion International,20, 7-17.

[Sato et al_08] Sato M., Liu X., Murayama J., Akahane K., and Isshiki M. 2008. A haptic virtual environment for molecular chemistry education. In Transactions on edutainment I, Abdennour El Rhabili (Ed.). Lecture Notes In Computer Science, Vol. 5080. Springer-Verlag, Berlin, Heidelberg 28-39.

[Scheible *et al*_08] Scheible J., Ojala T., and Coulton P.. 2008. MobiToss: a novel gesture based interface for creating and sharing mobile multimedia art on large public displays. In Proceeding of the 16th ACM international conference on Multimedia (MM '08). ACM, New York, NY, USA, 957-960.

[Schreiber et al_09] Schreiber M., Wilamowitz-Moellendorff M., and Bruder R. 2009. New Interaction Concepts by Using the Wii Remote. In Proceedings of the 13th International Conference on Human-Computer Interaction. Part II: Novel Interaction Methods and Techniques, Julie A. Jacko (Ed.). Springer-Verlag, Berlin, Heidelberg, 261-270

[SeriousGames_11] Serious Games Initiative. http://www.seriousgames.org/. Retrieved April 2011

[Sicart, 2008] Sicart, M. 2009. Newsgames: Theory and Design. In *Proceedings of the 7th international Conference on Entertainment Computing* (Pittsburgh, PA, September 25 - 27, 2008). S. M. Stevens and S. J. Saldamarco, Eds. Lecture Notes In Computer Science, vol. 5309. Springer-Verlag, Berlin, Heidelberg, 27-33. DOI= http://dx.doi.org/10.1007/978-3-540-89222-9_4

[Sinclair *et al*_07] Sinclair J., Hingston P., and Masek M. 2007. Considerations for the design of exergames. In *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia* (GRAPHITE '07). ACM, New York, NY, USA, 289-295.

[Sinclair *et al*_09] Sinclair J., Hingston P., and Masek M. 2009. Exergame development using the dual flow model. In *Proceedings of the Sixth Australasian Conference on Interactive Entertainment* (IE '09). ACM, New York, NY, USA, , Article 11 , 7 pages.

[Skaggs and Hopper_96] Skaggs, S., & Hopper, C. (1996). Individuals with visual impairments: A review of psychomotor behavior. Adapted Physical Activity Quarterly,13(1), 16-26.

[Sliney and Murphy_08] Sliney, A. and Murphy, D. 2008. JDoc: A serious game for medical learning, achi. In Proceedings of the First International Conference on Advances in Computer-Human Interaction, 131–136.

[Song *et al*_11] Song Y., Tang S., Zheng Y., Chua T. and Zhang Y., “Exploring probabilistic localized video representation for human action recognition”, Multimedia Tools and Applications, 2011.

[Sparacino_08] Sparacino F., “Natural interaction in intelligent spaces: Designing for architecture and entertainment”, Multimedia Tools and Applications, 2008.

[Stanley et al_10] Kevin G. Stanley, Ian Livingston, Alan Bandurka, Robert Kapiszka, and Regan L. Mandryk. 2010. PiNiZoRo: a GPS-based exercise game for families. In

Proceedings of the International Academic Conference on the Future of Game Design and Technology (Futureplay '10). ACM, New York, NY, USA, 243-246.

[Steinmetz_96] Steinmetz, R. 1996. Human perception of jitter and media synchronization. IEEE Journal on Selected Areas of Communication. 14, 1,

[Suhonen *et al. 2008*] Suhonen, K., Väätäjä, H., Virtanen, T., and Raisamo, R. 2008. Seriously fun: exploring how to combine promoting health awareness and engaging gameplay. In *Proceedings of the 12th international Conference on Entertainment and Media in the Ubiquitous Era* (Tampere, Finland, October 07 - 09, 2008). MindTrek '08. ACM, New York, NY, 18-22.

[Surakka and Kivela_08] Surakka A, Kivela T. Motivating visually impaired and deaf-blind people to perform regular physical exercises. Br J Visual Impairment 2008;26:255-68.

[Sussi *et al*_07] Susi T., Johannesson M., and Backlund P.. Serious games - an overview. Technical Report HS-IKI-TR-07-001, School of Humanities and Informatics, University of Skovde, Sweden, February 2007.

[Svhan_05] Svahn, M. 2005. Future-proofing advergaming: a systematisation for the media buyer. In *Proceedings of the Second Australasian Conference on interactive Entertainment* (Sydney, Australia, November 23 - 25, 2005). ACM International Conference Proceeding Series, vol. 123. Creativity & Cognition Studios Press, Sydney, Australia, 187-191.

[Takahashi *et al*_11] Takahashi M., Fujii M., Naemura M. and Satoh S., "Human gesture recognition system for TV viewing using time-of-flight camera", Multimedia Tools and Applications, 2011.

[Tashiro_09] Tashiro J. 2009. What really works in serious games for healthcare education. In *Proceedings of the 2009 Conference on Future Play on @ GDC Canada* (Future Play '09). ACM, New York, NY, USA, 3-4. DOI=10.1145/1639601.1639604 http://doi.acm.org/10.1145/1639601.1639604

[Tatematsu et al_10] Tatematsu A., Ishibashi Y., Fukushima N., and Sugawara S. 2010. QoE assessment in haptic media, sound and video transmission: Influences of network latency. In Proc. IEEE 2010 International Communications Quality and Reliability (CQR) Workshop.

[Teh *et al*_08] Teh, Soon J.K.; Cheok A.D., Peiris R.L., Choi Y., Thuong V., and Lai S.. (2008). Huggy Pajama: a mobile parent and child hugging communication system. In Proceedings of the 7th international conference on Interaction design and children (IDC '08). ACM, New York, NY, USA, 250-257.

[Testa_08] Testa, B.M. "'America's Army' provides an enterprise platform for Army training". Defense Systems. http://defensesystems.com/articles/2008/05/much-more-than-a-game.aspx. Accessed April, 2010.

[Thompson_09] Thompson M. When religion and games intersect—and how it often goes badly. Web article: http://arstechnica.com/gaming/news/2009/12/ when-religion-and-games-intersectand-why-it-often-goes-badly.ars
Accessed April, 2010

[TN Games_11] TN Games. The 3rd Space Vest. http://tngames.com/products. Retrieved June, 2011

[Top Ten Industry Facts_10] Top Ten Industry Facts, Entertainment Software Association. http://www.theesa.com/facts/index.asp Accessed May, 2010

[Tropico_10] http://www.tropico3.com/ Accessed april, 2010

[Vajk *et al*_08] Vajk, T., Bamford, W., Coulton, P., and Edwards, R., Using a Mobile Phone as a “Wii-like” Controller for Playing Games on a Large Public Display, International Journal of Computer Games Technology, vol. 2008, Article ID 539078, 6 pages, 2008.

[Varona *et al*_09] Varona J., Jaume-i-Capo A., Gonzalez J., and Perales F.. “Toward natural interaction through visual recognition of body gestures in real-time”. Interact. Comput. 21, 1-2 (January 2009), 3-10. 2009.

[Vogels_04] Vogels, I.M.L.C. 2004. Detection of temporal delays in visual-haptic interfaces. Human Factors.46, 1, 118-134.

[Vossen_04] Vossen, D. The Nature and Classification of Games. Avante, 2004, Vol. 10, No. 1, pp. 75

[Wang et al_07] Wang Q.; Tan W.; Song B.; , "Research and Design of Edutainment," *Information Technologies and Applications in Education, 2007. ISITAE '07. First IEEE International Symposium on* , vol., no., pp.502-505, 23-25 Nov. 2007

[Warren_06] Warren J. (November 25, 2006). "A Wii Workout: When Videogames Hurt". Wall Street Journal. http://online.wsj.com/public/article/SB116441076273232312-3nPirhZn20_L2P7m_ROtFUkh6yA_20071124.html Retrieved 2008-01-16.

[Weng *et al*_10] Weng C., Li Y., Zhang M., Guo K., Tang X., Pan Z. "Robust hand posture recognition integrating multi-cue hand tracking". In: Proceedings of the entertainment for education, and 5th international conference on E-learning and games. Springer, New York, pp 497–508, 2010.

[WHO 10] WHO Obesity and overweight fact sheet. http://www.who.int/mediacentre/factsheets/fs311/en/index.html Retrieved May, 2010

[Wii Sports_10] http://en.wikipedia.org/wiki/Wii_sports. Accessed on March, 2010

[Wii_11] Wikipedia: Wii Consol. http://en.wikipedia.org/wiki/Wii . Retrieved April 2011-04-25

[Wisdom tree_10] http://www.wisdomtreegames.com/. Accessed on April, 2010

[Wong *et al*_07] Wong W.L., Shen C., Nocera L., Carriazo E., Tang F, Bugga S., Narayanan H., Wang H., and Ritterfeld U.. 2007. Serious video game effectiveness. In *Proceedings of the international conference on Advances in computer entertainment technology* (ACE '07). ACM, New York, NY, USA, 49-55.

[Wu and Huang_99] Wu Y., Huang T. S., “Human Hand Modeling, Analysis and Animation in the Context of HC1”, in IEEE International Conference Image Processing". Kobe, Japan, 1999.

[Xu et al_10] Xu Z., Yu H., and Yan S. 2010. Motor rehabilitation training after stroke using haptic handwriting and games. In Proceedings of the 4th International Convention on Rehabilitation Engineering \& Assistive Technology (iCREATe '10). Singapore Therapeutic, Assistive \& Rehabilitative Technologies (START) Centre, Kaki Bukit TechPark II,, Singapore, , Article 31 , 4 pages.

[Yuan and Folmer_08] Yuan B. and Folmer E. 2008. Blind hero: enabling guitar hero for the visually impaired. In Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility (Assets '08). ACM, New York, NY, USA, 169-176.

[Yuan *et al*_10] Yuan B. and Folmer E., and Harris, F. C. Game accessibility; a survey. Universal Access in the Information Society 10, 1 (2010), 88-111.

[Yuan_09] Yuan B. 2009. Towards Generalized Accessibility of Video Games for the Visually Impaired. Ph.D. Dissertation. Univ. of Nevada, Reno, Reno, NV, USA. Advisor(s) Frederick C. Harris and Eelke Folmer.

[Zadeh_72] Zadeh L.A., “A rationale for fuzzy control,” J. Dynamic Syst. Meas.Control, vol.94, series G, pp.3-4,1972.

[Zayda_05] Zyda, M.; , "From visual simulation to virtual reality to games," *Computer* , vol.38, no.9, pp. 25- 32, Sept. 2005