A GAMING PLATFORM FOR ENHANCED PHYSICAL AND COGNITIVE
ACTIVITY TRAINING IN OLDER ADULTS

BY

EDISON F. MUÑOZ RECUAY

A Thesis Submitted to the Graduate Faculty of
WAKE FOREST UNIVERSITY GRADUATE SCHOOL OF ARTS AND SCIENCES
in Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE
Computer Science
August 2011
Winston-Salem, North Carolina

Approved By:
V. Paúl Pauca, Ph.D., Advisor
David John, Ph.D., Chair
Anthony P. Marsh, Ph.D.
Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Dr. V. Paúl Pauca for supporting me along in my studies, for his motivation, patience and immense knowledge. He encouraged me to continue with this research. Special thanks to Dr. Pauca’s family who helped take care of me in the process.

I would also like to thank my thesis committee, Dr. Anthony Marsh and Dr. David John, for their support, recommendations, comments and suggestions.

My sincere thanks go to the team of the Brain Boot Camp project: Dr. Anthony Marsh, Santiago Saldana, Paco Saldana, Dr. Jack Rejeski, Dr. Dale Dagenbach, Dr. Janine Jennings, Julie Sorensen, and Charlie Pashayan. I really enjoyed working with you and learned a great deal from everyone.

To my advisors in Perú, Mrs. Illa Rocconi and Ms. Lurka Ramirez, who encouraged me to pursue graduate studies in the United States and to Dr. Silverio Bustos and all my undergrad professors and friends who supported me on this venture.

Also, I would like to thank the friends I met along these two years: Mathyou Koval, Ranjan Banerjee, Erika Bechtold, Saurav Sarma, Changkun Xia, Ye Zheng, Percy Campos, Mike Forkin, David Sontheimer, Phillip Luke, Annie An, Clara Lucas, François Baranne, Vanessa Torres, Emma Cruz and Tommy Guy. I shared great moments with you guys, thank you for your friendship.

A very special thanks to my family, my mom Elsa and my sister Miriam, they are always in my heart, thank you for your love and kindness; to my cousin Polo Recuay and nephew Arturo Valdez for your help and advice. Finally thanks to God for guiding my life and placing good people along the way.
# Table of Contents

Acknowledgements ................................................................. ii

List of Figures ................................................................. v

Abstract ........................................................................... vii

Chapter 1  Introduction .......................................................... 1

Chapter 2  Physical Activity ..................................................... 5
  2.1 What is Physical Activity? ................................................. 5
    2.1.1 Why is Physical Activity Important? ............................ 5
  2.2 Physical Activity in Aging .............................................. 6
    2.2.1 Determinants of physical activity ............................... 7
    2.2.2 Promoting Physical Activity ..................................... 8
    2.2.3 Physical Rehabilitation .......................................... 9
    2.2.4 Walking in Older Adults ....................................... 10

Chapter 3  Human Computer Interaction ................................. 12
  3.1 Computer Controls ....................................................... 12
    3.1.1 Primary Interaction .............................................. 12
  3.2 Game Controls ........................................................... 12
    3.2.1 Joysticks and Gamepads ....................................... 13
    3.2.2 Motion Control .................................................. 13
    3.2.3 Body Control .................................................... 16
  3.3 HCI for Rehabilitation ............................................... 16

Chapter 4  Computer Vision .................................................. 18
  4.1 A Brief History ......................................................... 18
  4.2 Object Detection and Tracking ..................................... 19
    4.2.1 Data Preprocessing Approaches ............................. 20
    4.2.2 Detection of Human Shapes ................................. 22
  4.3 Hardware-Aided Detection of Human Shapes .................... 25
Chapter 5  Brain Boot Camp Game Prototypes ................................................. 27
  5.1 Prototype I: Detection via the Nintendo Wii ........................................... 27
    5.1.1 Prototype System Architecture ..................................................... 28
    5.1.2 Implementation ................................................................. 30
    5.1.3 Analysis and Evaluation ......................................................... 32
  5.2 Prototype II: Detection via Kinect ....................................................... 33
    5.2.1 The Unreal Development Kit (UDK) ............................................. 34
    5.2.2 Components .............................................................................. 34
    5.2.3 Game Prototypes ........................................................................ 37

Chapter 6  Gaming Principles for Brain Boot Camp ............................................. 41
  6.1 Usability Principles ............................................................................... 41
  6.2 Player involvement .............................................................................. 42
  6.3 Principles of Gaming Therapy for Health and Rehabilitation Games ........ 42
    6.3.1 Environment .............................................................................. 42
    6.3.2 Body involvement ...................................................................... 43
    6.3.3 Player adaptability, high customization ...................................... 43
    6.3.4 Feedback information for player and therapist ............................ 43
    6.3.5 Feedback of movements with use of tactile vibrators .................. 43
    6.3.6 Social interaction - Two players competition or cooperation ...... 43
  6.4 Proposed Evaluation for Brain Boot Camp ........................................... 44
  6.5 Survey Results ..................................................................................... 45

Chapter 7  Gesture Recognition: Detection of Marching Motion ....................... 49
  7.1 Approaches for Marching Recognition .................................................. 49
    7.1.1 A Simple Marching Detection Algorithm .................................... 49
    7.1.2 Chaos Theory and Embedding ..................................................... 51
  7.2 Experimental Results ........................................................................... 54
    7.2.1 Marching and Non-Marching Datasets ........................................ 54
    7.2.2 Recognition Results of the Simple Marching Detection Algorithm .... 55
    7.2.3 Results of the Embedding Method .............................................. 57

Chapter 8  Conclusions and Future Work ......................................................... 62
  8.1 Conclusions ......................................................................................... 62
  8.2 Future Work ......................................................................................... 62

Bibliography ................................................................................................. 64

Vita .................................................................................................................. 70
List of Figures

2.1 Post Stroke Rehabilitation ........................................ 9
3.1 Atari 2600 ............................................................ 13
3.2 Early Gamepad device ............................................... 14
3.3 Analog Stick ......................................................... 14
3.4 Wii Remote controls .............................................. 15
3.5 Play station motion control ....................................... 15
3.6 Microsoft Kinect Camera .......................................... 16
4.1 (a) Geometrical head-torso model, (b) template for calculating fitness coefficients [18] ........................................ 22
4.2 The process to go from the binary image to the star skeleton [14] . . . . 24
4.3 Simulation of the star skeletonization process. ..................... 25
4.4 Distance thresholding with the Kinect camera reveals the shape of a person standing in front of the camera. ...................... 26
5.1 Game Architecture .................................................... 28
5.2 Wii-based game console. ............................................ 31
5.3 Box of motion sensors (Wii Remotes). ............................. 32
5.4 Distance thresholded data provided by the Kinect camera. The center of gravity is shown with a red oval with coordinates $X = 299$ and $Y = 306$. ................................................................. 34
5.5 UDK Game engine structure. ....................................... 35
5.6 Database access schema ............................................. 36
5.7 Database schema .................................................... 37
5.8 Original N-Back schema from 0 to 3 [20]. ......................... 40
6.1 Mapping of the heuristics and our proposed evaluation ............. 44
7.1 Knee angle classification into three groups: 0. transition (yellow), 1. knee up (red), and 2. knee down (blue). The lines represent the thresholds of maximum and minimum angles to classify the angle samples . . 50
7.2 Steps of the Algorithm. ............................................. 54
7.3 Embedding of the left knee angle for marching motion with $m = 3$. 57
7.4 Embedding of left knee angle, non-marching motion with $m = 3$. . . 58
7.5 Invariant Features for Left Knee Angle Marching (left) and Non-Marching (right), for $m = 36$. .......................... 59
Abstract

The number of adults in the U.S. over the age of 85 is expected to triple by 2050. Maintaining function and mobility in this population will be of important consideration with significant economic implications. Regular physical activity, such as walking, is an important factor helping improve cardiovascular function, muscle strength, and balance in older adults. However, the benefits can differ depending on the level of function of the person. Recently, it has been found that simple walking routines combined with more complex movement requiring focused attention and decision making can lead to improved outcomes among older adults with poor lower extremity function compared to walking alone. In this thesis, we combine methods of gaming, computer vision, and novel camera technology to explore the role of computer-based virtual environment gaming for delivering complex physical and cognitive training among older adults. We investigate and evaluate current health game design principles based on feedback received by a small set of participants. In addition, we propose and evaluate two methods for detection of marching motion from a time sequence belonging to a fifteen point skeletonization of the human body.
Chapter 1: Introduction

Advancement in medicine and related areas have led to increased life expectancy over the last century, at least among people in developed countries. It is estimated that in 2050 there will be three times more people in the U.S. over the age of 85 than there are today [12]. However advancement in technology and engineering has also led to less physical activity and increased sedentary behavior [16]. The combination of these two (and the number of baby boomers expected to retire over the next few years) means that a significant increase in medical attention and personal assistance might be necessary.

Not surprisingly, regular physical activity is an important factor in the disablement process among older adults. A number of studies have shown that physical activity improves cardiovascular function, muscle strength, and balance, delaying the onset of functional limitations and mobility disability, see e.g. [26, 32]. Mobility disability further impairs strength, speed, endurance, coordination and dexterity, dramatically increasing the risk for dependency [17].

Walking is the most common mode of physical activity in older adults and studies showing the benefits of traditional walking programs have been published in the literature, see e.g. [7, 10]. However, according to Frank and Patla [13], traditional walking interventions lack the perception and decision making complexity of real-world environments [13]. This observation has led to a number of studies exploring the role of decision making complexity and cognitive challenge in physical activity. In particular, Marsh et al. [30] compared a traditional walking program with an innovative walking program requiring participants to walk for short periods of time, engaging between bouts in activities that challenged balance and involved complex
movements requiring focused attention and decision making. Their study showed that older adults with better lower extremity function (short physical performance battery (SPPB) scores of 9-12) had better outcomes with the traditional walking program. In contrast, older adults with poor lower extremity function (SPPB scores of 3-8) benefited most from the innovative walking program. These results are also consistent with the findings of the LIFE-P study [34] where participants with low SPPB scores at baseline did not benefit from a standard walking program to the same degree as those with high SPPB scores.

Overall, these findings suggest that physical activity programs for older adults with poor lower extremity function should train focused attention and decision making on complex motor tasks requiring the integration of information from multiple sensory inputs. The Brain Boot Camp pilot project sponsored by the Translational Science Center at Wake Forest University is investigating precisely this hypothesis as well as the effect of such training in the brain. As part of the Brain Boot Camp project, in this thesis, we explore the role of computer-based virtual environment gaming in the process of delivering innovative physical activity interventions for adults with poor lower extremity function.

Computer gaming has changed dramatically over the last decade, spurred primarily by changes in the enabling hardware, optical, and sensor technology. Whereas a player was confined to using the mouse, keyboard or joystick for game interaction, recent advancements in sensor and processing hardware have facilitated the use of the player’s own body to control the game. Nintendo’s Wii platform made prevalent the use of motion detector sensors (the Wii-remote) that the user could hold in his hands and wave around in specified known motions, e.g. hitting a ball with a racket, punching, etc. The recent introduction of Microsoft’s Kinect has taken this interaction further, freeing the user from the need to hold any motion sensors. The
Kinect, consisting of several cameras, infrared illumination, multi-array microphones, and actuators, is able to detect, in real time, the relative position \((x, y, z)\) of a number of points in three dimensional space. The introduction of powerful mobile devices are also an enabling technology having the potential to dramatically facilitate human computer interaction.

Recent advancements in computer vision have also had a tremendous impact on the development of computer games. The aim of computer vision is to better understand the three-dimensional structure of the world, as seeing through the lens of a camera (or cameras), obtaining high level information of objects of interest in the scene. Combined with computer vision and powerful gesture recognition, Kinect is able to detect complex body positioning in space, facilitating interaction of the user with a virtual environment.

The types of games being developed are also undergoing significant changes. Though the large majority of the computer gaming industry is still devoted to entertainment only, e.g. World of War Craft, recently, a great deal of resources are being devoted to so-called games for health. These types of games are aimed at improving health and health care and include exergaming, physical therapy, disease management, health behavior change, biofeedback, rehabilitation, epidemiology, training, cognitive health, nutrition and health education. There are a variety of factors involved in the development of such a games and how they might be combined to obtain the desired effects is of current interest. The principles underpinning the design of such games is a subject of the current investigation.

The remainder of this thesis is structured as follows. We describe various aspects of physical activity in older adults in Chapter 2. Advancements in human computer interaction for gaming purposes are described in Chapter 3. We explore some basic concepts of computer vision as it applies to human shape detection and tracking in
Chapter 4. Efforts spent on the development of prototypes implementing innovative physical activity and cognitive interventions for adults with poor lower extremity function are described in Chapter 5. In particular, we present three games that have been developed as part of the Brain Boot Camp research project. Gaming principles for health related games are explored in Chapter 6. Here, we present preliminary results on game design principles based on feedback received from a small set of participants. In Chapter 7, we present two methods for detection of a marching motion. The first method uses a simple phase and magnitude analysis of the knee angle. The second method is based on Chaos Theory and uses the concept of embedding to map one-dimensional signals into higher-dimensional spaces where the type of motion can be more easily classified. Preliminary results are given for both approaches. Conclusions and future work are outlined in Chapter 8.
Chapter 2: Physical Activity

In this chapter, we briefly explore the role of physical activity (or the lack of) in older adults and describe current efforts by the research community to help maintain or improve their functional ability and quality of life in this population.

2.1 What is Physical Activity?

According to the World Health Organization, physical activity is defined as any bodily movement produced by skeletal muscles requiring energy expenditure. Physical inactivity, or the lack of physical activity, is an independent risk factor for chronic diseases, and overall is estimated to cause 1.9 million deaths globally per year. Physical activity can be done purposefully or unconsciously in our daily routines at home, work, etc. and can be performed moderately or vigorously [16]. According to Di Pietro [7], physical activity is, however, a complex behavior that is often difficult to describe and deal with it. Many variables must be considered, such as sex and race, as they involve significant differences in reported physical activity patterns, which are evident even in older adulthood [7].

2.1.1 Why is Physical Activity Important?

Physical activity is an important factor for development; it involves responsibility with personal care, it helps prevent diseases, and promotes overall well-being. According to Kramer et al. [23], exercise increasing aerobic fitness may selectively improve cognitive processes involving frontally mediated activities such as planning, scheduling, and working memory.

Physical activity can also help control pain and other symptoms. For example,
in people with osteoarthritis, physical activity can help reduce pain and prevent the risk of developing osteoarthritis in a joint that has been injured [9]. Many researchers also believe that physical activity can bring significant psychosocial benefits, such as social interaction and distraction from daily stressors.

There is also evidence of a relationship between level of physical activity and depressive symptoms [9]. For instance, Blumenthal et al. [2] published results of a controlled trial of endurance exercise (walking or jogging three times a week), antidepressant medication (sertraline), or their combination in older adults who met diagnostic criteria for major depressive disorder. After 16 weeks, all three groups had clinically significant improvements in depression scores. After 10 months, remitted participants in the exercise-only group had significantly lower relapse rates than participants in the medication groups.

### 2.2 Physical Activity in Aging

Physical activity is especially important among older adults. Without regular exercise, seniors can often experience accelerated loss of skills necessary for their every-day life activities. According to Carlson et al. [3], based on the 2008 Physical Activity Guidelines for Americans, the prevalence of physical activity among adults over 65 were 18% highly active, 11.6% sufficiently active, 18.9% insufficiently active, and 50.9% inactive. Doing physical activity helps to prevent obesity, cardiovascular disease, stroke, hypertension, colon and breast cancer, as well as diabetes and osteoporosis [42].

The economic impact of the loss of functional ability among older adults, is also a problem. According to a study by the U.S Department of Health and Human Services developed in 2000 [33], in adults over 18 years of age, the cost (in billions of dollars) of treatment for diseases resulting from lack of non-physical activity was:
Heart Diseases $183, Cancer $157, Diabetes $100 and Arthritis $65. It is important to mention that 60% of the population included in this study were people over 65 years of age and those diseases listed previously have more prevalence in this segment of the population.

2.2.1 Determinants of physical activity

Among older adults, the influence of weight-bearing, strength, and flexibility aspects of physical activity on bone and lean mass preservation and balance takes highest priority with regard to maintaining functional ability and independence [7]. However, physical activity in older age tends to be of lower intensity and highly variable over time. As a result, accurately measuring or estimating activity level becomes difficult. For this reason, DiPietro et al. [7] has proposed specific determinants to better understand the variability in activity patterns in older adults. These determinants are summarized next.

Physiological Factors

Adequate classification of the target population is an important consideration since heredity, gender, and genetic predisposition are key components of physical fitness or functional capacity. In the case of adults, and specifically older adults, speed, flexibility, balance, and strength may also be important determinants of participation in particular activity as simple as walking.

Psychosocial Factors

Self-efficacy or confidence in one’s abilities is a factor strongly associated with the adoption and adherence to physical activity, especially in older adults. Moreover, stress tolerance and self-motivation have consistently correlated with physical activity
level in several adult populations.

**Social Support**

Older adults may require third party intervention to encourage participation in physical activity. Social influence, peer reinforcement, and support from family and friends are especially important to physical activity patterns in older adult populations.

**Safety and Accessibility**

Walking in safe and well-lit bicycle paths and recreational areas are very important for some older adults, especially for those living in urban environments where sidewalks often do not exist.

**2.2.2 Promoting Physical Activity**

There are two types of motivation affecting physical activity: intrinsic motivation, which comes from within a person, and extrinsic motivation, which comes from sources external to a person. However, intrinsic motivation may often be affected by external factors. A physical instructor can impact intrinsic motivation through informative feedback. Praise, for example, is considered a form of positive feedback having the potential to increase intrinsic motivation for performing a specific task. Criticism, on the other hand, is a form of negative feedback that often tends to negatively impact intrinsic motivation [40, 41]. We will see later on that intrinsic motivation in older adults can be effectively tapped in novel game-based training routines to increase desired performance.
2.2.3 Physical Rehabilitation

Physical rehabilitation aims to enhance and restore functional ability and quality of life to the elderly as well as to those with physical impairments or disabilities. Rehabilitation for acquired impairment or disability due to, for example, osteoarthritis, arthrosis, and stroke, is common among older adults. This practice however is not straightforward.

According to Garrison and Felsenthal [15], the practice of geriatric stroke rehabilitation is difficult because elderly patients may be severely upset that they have sustained a stroke, appearing passive about their situation. At this age the effect of other diseases (such as the common cold) increases, making it more complicated to focus only in stroke symptoms. Significant attention and support in caring for these patients and their families may be needed because of secondary afflictions such as depression [24].

Figure 2.1: Post Stroke Rehabilitation
2.2.4 Walking in Older Adults

Physical rehabilitation to enhance or restore functional ability in older adults for non-acquired conditions is also an important topic of current research. Older adults are often at risk due to the lack of regular physical activity in their daily routines. Walking appears to be the most common mode of physical activity in older adults [7, 10], but the benefits may vary depending on the physical conditioning of the individual. One of the reasons for this difference may be the lack of real-world perception and decision making complexity in traditional walking routines [13]. To test this theory, Marsh et al. [30] compared a traditional walking program (WALK) with an innovative walking program (WALK+) requiring participants to walk for short periods of time, engaging between bouts in activities that challenged balance and involved complex movements requiring focused attention and decision making. In the case of WALK, participants walked two laps at a low intensity before walking for up to 25 minutes at a moderate-intensity walking pace. On the other hand, WALK+ included additionally four obstacle stations along the track. The total time of the walk/obstacle session was a maximum of 25 minutes of which 8-10 minutes were assigned to the obstacles stations. The more significant results that from this study include:

- There were no significant differences between the WALK and WALK+ groups in any demographic or health measure at baseline.

- Participants with low baseline function (short physical performance battery (SPPB)\(^1\) score of 3-8) assigned to the WALK group showed only small improvement in their SPPB score after the intervention, while those with low baseline function assigned to the WALK+ group showed substantial improvement in their SPPB score.

\(^1\)SPPB is an objective assessment tool for evaluating lower extremity functioning in older persons. It was developed by the National Institute on Aging.
Higher functioning older adults with SPPB of 9-12 may have benefited more from WALK than WALK+.

WALK and WALK+ might not be applicable to older adults who cannot walk short distances or who have SPPB scores between 0 and 2. In particular, very frail older adults might be at increased risk for falls in the absence of adequate intervention-staff supervision.

According to this study, physical activity programs for older adults with poor lower extremity functions should be focused on complex motor tasks that require the integration of information with multi sensory inputs. In this thesis, we aim to fill this gap by leveraging computer-based virtual environment gaming, computer vision, and modern imaging hardware. Computer gaming may allow us to tap into an individual’s intrinsic motivation by positive re-enforcement through a fun and interactive virtual environment, encouraging older adults to engage in physical activity.
Chapter 3: Human Computer Interaction

Human computer interaction (HCI) is a vibrant field of research within applications ranging from planning and design of the interaction between people and computers. In this thesis, we are interested in HCI for enhancing function among at risk older adults. We start with a brief look at the history of computer and gaming controls.

3.1 Computer Controls

Computer controls have played a significant role over the last two or three decades, facilitating the interaction with the computer to perform a specific activity.

3.1.1 Primary Interaction

Perhaps the first real user-interface available for computer interaction was the video display terminal. Video display terminals used the cathode ray tube, commonly found in television, with an electric typewriter to enable users to visualize, type, and edit text. These electronic typewriters quickly gave way to the electric keyboards we use today. The computer mouse quickly followed. It was developed by Engelbart, who received a patent for his invention [31], describing it as an ”X-Y position indicator for display systems.”

3.2 Game Controls

The real-time nature of the computer controls quickly led to the development of computer-based games and dedicated gaming controls. Gaming is today one of the
largest industries in the world and games are in constant innovation, not only in the graphical aspect, but also in the hardware.

### 3.2.1 Joysticks and Gamepads

Atari, a pioneer of the game industry, developed several controls between 1970 and 1982, including joysticks and driving controllers for the Atari 2600 (see Figure 3.1), mice, trackballs, and keyboards for the Atari 7800, and gamepads for the Atari 7800 [8]. The Atari 7800 gamepad consisted of a D-Pad and 2 buttons on a base (see Figure 3.2) attached to the console by a wire. The evolution of the standard game control scheme from joystick to gamepad was necessary for a variety of reasons, including sensitivity to small movements [21].

![Figure 3.1: Atari 2600](image)

Later on, these devices evolved to a control combining a version of the joystick (analog stick) with gamepads. The Sony gamepad (see Figure 3.3) is a primary example [6].

### 3.2.2 Motion Control

Motion controls are one of the lastest innovations in gaming controls in the last few years. In contrast to gamepads and joysticks, motion control sensors permit interact
with the game console not only through a gamepad, but also through movement of the human body itself. (standing up and moving lower and upper extremities).

The first two companies offering the advantage of motion controls were Nintendo and Sony with "Nintendo Wii” and "Sony Playstation 3”, respectively. In the first case, the Wii controller is known as the Wii Remote (see Figure 3.4) and is shaped liked a television remote control composed by multiple accelerometers, motion sensors, and light sensors. The Wii remote identifies movement by sensing light emitted by the so-called ”sensor bar”- a stick with infrared light emitting LEDs. In the second case,
the control is called PlayStation Move (see Figure 3.5). This control also has motion sensors, but unlike the Wii, it projects an ultrared light that is captured by a stationary camera. Development of the Wii and PlayStation Move were accompanied by the development of new games oriented to physical activity and interaction with other players, e.g. Wii Sports (bowling, tennis, etc).

These games become popular among retirement communities, because of the simplicity of the interface, low cost, and prevalence of this technology. Several studies related to physical activity and the Wii have been publicized in the literature.

Figure 3.4: Wii Remote controls

Figure 3.5: Play station motion control
3.2.3 Body Control

Originally known by the code name Project Natal and launched on November 2010, Kinect for Microsoft Xbox 360 is a controller-free gaming and entertainment system. Unlike the Wii or Playstation Move, the user does not need to hold any gamepads or sensors to interact with the system.

The Kinect allows a more natural user interface using gestures and spoken commands. Technologically, Kinect is composed of a RGB camera, depth sensor and multi-array microphone running a proprietary software; these components provide full-body 3D motion capture, facial recognition and voice recognition capabilities.

![Microsoft Kinect Camera](image)

Figure 3.6: Microsoft Kinect Camera

3.3 HCI for Rehabilitation

Human computer interaction is a field that is undergoing tremendous changes with the new emerging technology. Today, Microsoft’s Kinect camera and the Nintendo Wii have enabled the players to use their body as the interface with the computer and gaming system. Researchers are currently exploring the use of Kinect for human computer interaction in specific fields, such as medical imaging. For the Brain Boot Camp project, this technology offers the opportunity to develop low-cost virtual envi-
ronment games specifically tailored for older adults, which may be easily customized and used in research and clinical settings.
Chapter 4: Computer Vision

According to Szeliski [39], computer vision attempts to describe the world that we see in one or more images and to reconstruct its properties, such as shape, illumination, and color distributions. Computer vision is being used today in a wide variety of real-world applications, including optical character recognition, machine inspection, medical imaging, automotive safety, motion capture, monitoring for intruders, analyzing highway traffic, monitoring pools for drowning victims, fingerprint recognition and biometrics. In this chapter, we briefly describe computer vision algorithms application to detection and tracking of the human body.

4.1 A Brief History

Computer vision first started out in the 1970s. At that time, it was believed by some of the early pioneers of artificial intelligence and robotics from MIT, Stanford, and CMU that solving the visual input problem would be an easy step along the path to solving more difficult problems such as higher-level reasoning and planning [39].

What distinguished computer vision from digital image processing was a desire to recover the three-dimensional structure of the world. Early attempts at scene understanding which involved extracting edges and then inferring the 3D structure of an object were developed by Roberts in 1965. In 1971, several line labeling algorithms were developed by Huffman, Clowes in 1975, and in 1980 by Kanade. The topic of edge detection was also an active area of research; a survey of contemporaneous work can be found in 1975 by Davis. Three-dimensional modeling of non-polyhedral objects was also being studied by Baumgart in 1974 and by Baker in 1977. A qualitative approach to understanding intensities and shading variations and explaining them by
the effects of image formation phenomena, such as surface orientation and shadows, was presented by Barrow and Tenenbaum in 1981. The 90’s saw many interesting developments in image filtering, 3-D range data processing, physics-based vision, and energy-based segmentation. The turn of the century saw significant advancement higher-level analysis such as face recognition and detection, texture synthesis and inpainting, computational photography, and feature-based recognition.

Nowadays, computer vision interacts heavily with other fields, such as pattern recognition and machine learning. In these areas, different approaches have been developed for human identification, motion and tracking, shape representation and matching, video analysis, face analysis, document analysis, etc. In this thesis we are interested in methods of computer vision allowing feature detection and extraction in real-time.

4.2 Object Detection and Tracking

Object detection and tracking are problems treated differently in computer vision, depending on the application area. There are different approaches addressing this problem. In our specific case, we are interested in approaches for detecting and tracking different points in the human body, such as arms, legs, and head. This small set of connected points representing the human body as a skeleton are followed over time, enabling real-time detection and tracking.

For the purpose of interactive virtual environment games for older adults, the following conditions on the methods for human body tracking appear to be necessary:

- The system should detect and track a human silhouette in about 30 frames per second. This is required for implementing a more natural human computer interface (see Chapter 3).
A skeletonization (extraction of a small set of feature points) of the human silhouette must be easily available as input for other components of the gaming platform. This enables rapid customization and parameter tuning.

The skeleton detection and tracking should be robust to noise, i.e. illumination changes as well as background objects. This step is fundamental to providing a natural interface.

The underlying model should be strong enough to recover from wrong calculations due to occlusions or unexpected changes in illumination.

All these aspects are very important for fast, accurate, and robust tracking of the human body.

4.2.1 Data Preprocessing Approaches

Data preprocessing is an important and necessary component previous to feature detection and tracking. Methods like background subtraction [22] find and remove objects in a scene deemed to be unimportant (background), thus avoiding details and other unnecessary distractors of the scene. Other methods like motion detection methods, such as optical flow [37] can track points while there is an object in motion in the scene. Higher level methods for detecting patterns and objects of interest can then be applied.

Background Subtraction

One of the most popular methods of background subtraction is the Codebook representation model of Kim et al. [22]. In this method, sample background values at each pixel in the acquired image are quantized into so-called codebooks. A codebook represents a compressed form of background model information, acquired over a long
image sequence. A codebook representation can handle scenes containing moving backgrounds or illumination variations, and can achieve robust background detection for different types of video data [22]. This method has the potential to deal with multiple changing backgrounds and can be applied both in RGB or grayscale space.

**Optical Flow**

Optical flow refers to the pattern of apparent motion of objects, surfaces, and edges in a scene caused by the relative motion between an observer and the scene [43]. Methods of optical flow attempt to estimate the motion between two image frames $f_t$ and $f_{t+\delta t}$, taken at times $t$ and $t + \delta t$, respectively, at certain image positions. In particular, assuming that a pixel in $f_t$ at position $(x, y)$ moved by $(\delta x, \delta y)$ at time $t + \delta t$, the noise-free intensity of the pixel is expected to be the same, that is,

$$f_t(x, y) = f_{t+\delta t}(x + \delta x, y + \delta y). \quad (4.1)$$

Assuming the movement to be small, a Taylor series expansion provides a relationship between both image frames as:

$$f_{t+\delta t}(x + \delta x, y + \delta y) = f_t(x, y) + \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial t} \delta t + \ldots \quad (4.2)$$

As a result, it follows that,

$$\frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial t} \delta t = 0 \quad (4.3)$$

$$\frac{\partial f}{\partial x} V_x + \frac{\partial f}{\partial y} V_y + \frac{\partial f}{\partial t} = 0, \quad (4.4)$$

where $V_x, V_y$ are the $x$ and $y$ components of the velocity or optical flow of $f_t$ and $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, and $\frac{\partial f}{\partial t}$ are the derivatives of $f_t$ along the $x$, $y$ and $t$ directions. This equation is known as the aperture problem in optical flow.
The Lucas-Kanade algorithm [27] is a widely used method for solving Eq. (4.4). This method assumes that optical flow is constant in the local neighborhood of a pixel, solving the over-determined systems of equations in Eq. (4.4) in the least squares sense. In practical applications, a training process is needed to determine pixels for which optical flow is to be computed. Pixels of interest must belong to a moving object in the scene. Background illumination must also be constant. Occlusion and background changes are also problematic and may require preprocessing via background subtraction.

4.2.2 Detection of Human Shapes

Shape detection algorithms look for specific patterns in a scene. In practice, these methods require preprocessing the image to select points of interest. We describe two methods for human shape detection in real-time.

**Torso and Hand Segmentation.** Huo *et al.* [18], propose an approach to detect body movement. In this approach, features resembling body parts, such as the torso and hands, are detected and tracked over time. Their method uses a simplistic model of the geometrical structure of the head-torso (see Figure 4.1 a), having parameters $P = (x, y, scale)$ for position and scale, combined with a particle filter method [19]

![Geometrical head-torso model and template](image)

**Figure 4.1:** (a) Geometrical head-torso model, (b) template for calculating fitness coefficients [18]
for efficient shape detection and tracking over time. The particle filter provides the probability of detection of a person in the image in terms of a set of fitness coefficients. The template used to calculate these fitness coefficients is shown in Figure 4.1(b).

While this approach can be implemented in real-time, it is primarily limited to extraction of the torso and cannot be easily modified to track the whole body. In addition, recognition performance can be significantly degraded if the person is not facing the camera, thus mismatching the head-torso model shown in Figure 4.1(a).

**The Star Skeleton.** The star skeleton is a fast skeletonization approach where a 5-point skeleton is constructed by connecting the centroid of a target human silhouette to its contour extremes [14]. In particular, after background subtraction and smoothing, a binary image containing the human silhouette is obtained by Canny edge detection followed by dilation and erosion operators. The 5-point skeleton is then found with the following steps:

- Calculate the center of gravity \((x_c, y_c)\) in the target image as an average of points along the silhouette,
- Calculate the distance \(d_i\) from the center \((x_c, y_c)\) to each silhouette point \((x_i, y_i)\).
- The signal \(d_i\) is smoothed out to reduce noise,
- The local maxima of \(d_i\) are taken as the extremal points (corresponding to the gross extremities of the target).

The star skeleton is constructed by connecting the local maxima of \(d_i\) with the centroid \((x_c, y_c)\). See Figure 4.2.

The main advantages of the star skeleton approach are its low computational complexity, allowing for detecting and tracking in real-time and the fact that it does not
require a priori model of the human shape. However, some of the image filtering steps, such as low-pass Gaussian filtering, need careful parameter selection. Figure 4.3(a,b)) shows the distance $d_i$ obtained for the silhouette model shown in Figure 4.3(c).
4.3 Hardware-Aided Detection of Human Shapes

The introduction of new imaging technology, such as the Wii and Kinect, into the imaging and gaming markets has greatly facilitated the process of shape detection. In the Wii system, detection is a trivial process as the player is forced to hold the camera sensor in his hand. The system is thus able to track up to four points in space in real-time. Player movement is translated to the avatar, facilitating interaction with the virtual environment. As we will see in Chapter 5, this setup can be modified by placing the light-emitting LEDs on the player while leaving the camera
sensor stationary, for a hands-free interface. The downside of this approach is that additional light-emitting LEDs would be required on the user’s body to determine shape.

The Kinect camera is able to detect not only lateral placement \((x, y)\) of points in the scene, but also distance from the points to the camera \((z)\). This is an important advantage compared to traditional cameras for a number of reasons. In particular, for shape detection, thresholds on the distance can be set in order to select points of interest in the scene. Figure 4.4 shows the result of distance thresholding to obtain a binary map for an object of interest. A skeleton can then be easily obtained from this binary map using a method such as the star skeleton.

![Figure 4.4: Distance thresholding with the Kinect camera reveals the shape of a person standing in front of the camera.](image)
In this chapter, we present a system architecture framework for the design of virtual environment type games. We present two prototypes as well as games developed as part of the Brain Boot Camp project. Key aspects considered as part of our team-approach development effort include:

1. Integration of hardware components. Several hardware pieces, including cameras, LED-based illuminators, vibration devices, and processing units needed to be combined wirelessly to enable a hands-free interaction environment.

2. Virtual environment game design. This aspect required understanding and customization of gaming technology and related software development kits, such as Microsoft’s XNA Engine and the Unreal Engine 3.

3. Data processing and database design. This aspect required development of functions for proper access and storage of basic player information. A number of markers related to a player’s performance in the game also needed to be collected and properly stored. Additionally, storage of the player’s skeleton data over time is required for later processing and analysis.

4. Integration of research-oriented goals (e.g. physical activity) into the game logic. This aspect required significant customization of the game to ensure that research goals were met.

5.1 Prototype I: Detection via the Nintendo Wii

Our first prototype was designed and built around the Nintendo Wii remote (see Chapter 3). These devices are capable of detecting infrared light emitted by the...
sensor bar, transmitting the relative position of the remote via bluetooth to the Wii game console.

In our design, we followed the ideas introduced by Johnny Lee [25], and placed LEDs on the player, while keeping the Wii-remote stationary. This approach allows for a hands-free interface where gross and fine motor movement is elicited based on the distance of the player with respect to the Wii-remote. The further the player is made to stand from the Wii-remote, the longer the movement required from the player to move a specified distance. This approach was employed by Smartt, Saldana and Pauca [38] to elicit fine motor control (finger movement) in children with disabilities.

5.1.1 Prototype System Architecture

The system architecture for our first prototype is shown in Figure 5.1. The main features of this prototype include:

Figure 5.1: Game Architecture
**Gee Nodes.** A Gee node is a device containing a LED for emitting infrared light. The player wears the Gee Node in some part of his or her body (basically arms, legs and chest).

**Sensing Platform: Wii Remotes.** The Nintendo Wii Remote contains a camera able to detect the light emitted by the Gee Nodes. It can provide monitoring of a cone-shaped space of about 40 degrees wide.

**Output: Screen and Tactile Vibrator Devices.** The screen displays the virtual environment, 3D objects and the Avatar (a representation model of the player in the game) reflecting player movement. The player must interact with the game based on what he/she sees on the screen. Additionally, the player can have tactile vibration devices on arms and legs, for haptic feedback.

**Game Application.** This component contains the logic for handling signals received from the Gee Nodes and Wii-remote, modifying accordingly the virtual environment showed in the screen.

**Game Engine.** The game engine is the software framework that contains all the libraries providing functionality for 3-D rendering and gaming logic. The actual games run under the game engine environment.

**Wii Library.** This component contains the functionality that provides the interface between the Wii Remotes and the game application.

**Database: 3D Models.** This database stores the different 3D models or so-called assets the developer may want to use in his or her game.
**Database: User Profiles.** This database stores player information, such as, levels, scores, age, current progress, and overall performance. Additionally information regarding the player’s position in space over time may be saved.

## 5.1.2 Implementation

The system architecture shown in Figure 5.1 was implemented with the following key elements:

- **Game engine.** We used the XNA Microsoft Engine Gaming Framework that is oriented to 3-D graphics and provides a variety of game functionality tools.

- **Game console/system.** The initial implementation of the system (see Figure 5.2(a)) contains the game engine (software), motion sensors (see Figure 5.3 (a) and (b)) and the Gee Nodes (see Figure 5.2(b)).

- **Sensing platform.** The sensing platform was modified to include two Wii-remotes (see Figures 5.3) sitting parallel to each other, to detect 3-D movement of the player within the system’s field of view (a cone-shaped area starting at the box and extending outwards towards the player). The Wii remotes have power supplied by USB cable connected to the computer. The box is placed in front of the player and should be at the player’s chest level for optimal viewing of the avatar in the screen.

- **Gee nodes.** In this implementation, we used a single Gee Node to detect player movement. Figure 5.2(b) shows the player wearing the Gee node at chest-level. The player interacts with the virtual environment shown in the screen by his/her body movement (see Figure 5.2(c)).
Figure 5.2: Wii-based game console.
5.1.3 Analysis and Evaluation

The Wii-based prototype was fully developed and tested internally as part of our research work. The main benefits provided by this prototype include:

- Motion detection and preprocessing (see Chapter 4) requires minimal effort. The computational unit receives a very small amount of data from the Wii-remotes over time.

- Extremely fast user interface. The system is able to detect 3-D movement at roughly 200 frames per second. Processing of this data to adjust the virtual environment is also done very efficiently.

- Low computational footprint. The small computational requirements mean that a low-end laptop can be used for processing.

Key drawbacks of the Wii-based prototype include:

- Need for intrusive elements on the player’s body. Though small, a Gee Node need to be carefully placed on the body. A player must wear the Gee Nodes while interacting with the system.
• Modeling a person with a skeleton would require several Gee Nodes to be placed on the body, making the system much less convenient to use.

• Due to the required calibration of Gee Nodes and Wii-remotes (for optimal illumination), the system may be prone to errors induced by movement or occlusion.

5.2 Prototype II: Detection via Kinect

A prototype was designed and built around Microsoft’s Kinect camera [35] which was recently introduced in December of 2010. As previously discussed, distance thresholding helps remove unnecessary detail from the scene, leaving only the detection of the player position to be completed.

To determine and track a human skeleton with Kinect, we use the OPEN NI framework. This framework provides a number of classes to deal with detection and tracking. Specifically a class called ModuleSkeletonInterface provides methods for determining 15 points (head, neck, left shoulder, left elbow, left hand, right shoulder, right elbow, right hand, torso, left hip, left knee, left foot, right hip, right knee, and right foot) on a human silhouette in 3-D space. In addition, this class allows us to form and track a skeleton over time. The 15-point skeleton can be obtained at a rate of 30 frames per second. With this information it is possible to analyze body movements in real-time. Figure 5.4 shows the calculation of the center of gravity using the star skeleton approach.

1OPEN NI is a not-for-profit organization formed to certify and promote the compatibility and interoperability of Natural Interaction (NI) devices, applications and middleware.
Figure 5.4: Distance thresholded data provided by the Kinect camera. The center of gravity is shown with a red oval with coordinates $X = 299$ and $Y = 306$.

5.2.1 The Unreal Development Kit (UDK)

For this prototype we opted to replace the XNA of Microsoft Engine Gaming Framework for the Unreal Engine 3 from Unreal Development. This is a production-level system used by professional game developers. Its software development kit is totally oriented to game development and provides various tools for developing object-oriented game designs. The structure of UDK is given in Figure 5.5.

5.2.2 Components

Dynamic Link Libraries (DLL). These are external libraries that can be accessed from other programing languages such as C/C++ or Visual C++. These libraries are used for capturing the player skeleton data and for subsequent processing, such as the calculation of the center of gravity and joint angles.

Controllers. These are the main classes in the game engine. They implement the game loop and allow programming of the game logic. Also, the states of objects, such
as the avatar (figure representing the player in the game) are kept and modified by these classes.

**Interfaces with Flash.** These are classes programmed in the Unreal Script programming language. They can be used by external systems such as Macromedia Flash to implement menus and other visual effects.

**Pawn and general information classes.** These are classes programmed in the Unreal Script language, enabling the interaction of three dimensional objects in the
virtual environment. For example, the `Pawn` class can be used to represent the player, scenario, and other 3-D objects.

**Database access classes.** These are classes programmed in the Unreal Script language. They provide access to the physical database layout, providing operations such as insertion, update, and elimination of records. We use SQLite which is a database engine oriented to portable applications. Figure 5.6 shows the interaction of UDK classes with the database engine and its interface.

[Image: UDK database access schema]

**Figure 5.6: Database access schema**

**Database schema.** The database schema contains four tables: player, score, level, and program as can be seen in Figure 5.7. We store the main information of the player,
levels and their score details such as the number of objects avoided, objects touched, game time for each level, the last level played, percent accomplished and quantity of movement (both are calculated during the game).

5.2.3 Game Prototypes

Three game components were implemented as part of this work. These components are based on the interaction of the player with a virtual environment consisting of a hallway facing a door at the distance. The player moves through this hallway accomplishing specific tasks.
Go and No Go.

The purpose of the “Go” component of the game is to avoid obstacles coming from the door towards the player. The player’s goal is simply to avoid as many of the obstacles as possible, moving side to side or up and down away from the obstacles. A score is calculated based on how many objects the player successfully avoided in a specific amount of time. If the player successfully avoided at least 90% of the obstacles, then the player is given the choice to move up a level of difficulty. Subsequent levels differ by the number and speed of the obstacles coming towards the player. The “Go” component enforces primarily physical activity and the number of obstacles and their speed are used as variables controlling the amount of physical activity elicited from the player.

The “No Go” component of the game involves the opposite of the “Go” component: the player is instructed via visual and/or auditory clues to touch obstacles of specific color coming towards the player. Objects of the “No Go” type are shown at random with those of the “Go” type. A different score is calculated based on how many of the “No Go” obstacles the player was able to touch. If the player was successful in avoiding or touching at least 90% of the total number of “Go”/“No Go” obstacles then the player is given the choice to move up a level of difficulty. Subsequent levels differ by the speed, number, and color of the “No Go” obstacles (controlling variables). The introduction of the “No Go” component introduces a basic level of cognitive training and more complex decision making into the game.

Ts and Ls.

The purpose of this game is to determine, within a pre-specified amount of time, the side of the hallway containing a letter “L.” In this game, the hallway is split in halves by virtual doors containing either a set of “T”s randomly placed and oriented
within the door, or a set of Ts and one L in the other. The player must choose the correct door to go through (the one containing the L) and move towards it within the pre-specified amount of time. This game enforces primarily a level of cognitive challenge and decision making into the game, involving low physical activity. The number of Ts and Ls are used as variables to control the level of cognitive challenge in this part of the game. This game is base on the studies of Wolfe [44] about “Visual Search,” where it is studied the particularities of finding an object between other similar distractors. In case of the “T”s and one “L” game, one cannot fully process all the visual information in the field of view at once. With aging the limitations of peripheral visual processing decreases. Studies indicate that it is more difficult for older adults to differentiate between a target object and distractors [29].

N Back.

The purpose of this game is to move through a maze of colored doors, by carefully remembering the color of the doors presented in the past. One-back ($N = 1$) means the player must move through a door of a specific color, making the decision based on the color of the doors just presented to him. This is a very challenging cognitive game requiring minimal physical activity. It is based in the experiment of Jonides et al. [20] which evaluates the effect of variations in memory load that are involved in verbal working memory. Players have to manipulate saved and current information. A variation of this task is the inclusion of additional sets of neutral doors, appearing between the sets of colored doors. According to Jonides et al. [20], as we increment the value of $N$, there are magnitude increments in brain activation in a large number of places in the brain, which are identified as verbal working-memory processes. In Figure 5.8 it is shown the schema of Jonides’ proposal.
Figure 5.8: Original N-Back schema from 0 to 3 [20].
Chapter 6: Gaming Principles for Brain Boot Camp

Computer game design is a challenging and creative art form. Game designers spend a great deal of time trying to determine what a player may want out of a game, attempting to understand the balance between what is enjoyable and challenging about the game experience. Determining a set of principles to follow is a very important aspect of a game design. The key principles of game design include: challenge, socialization, emotional experience, exploration, fantasy and interactivity [11]. Players expect a consistent world, reasonable solution, direction, immersion and incrementation.

In this chapter we present an analysis of game design principles applicable to health games, such as those developed under the Brain Boot Camp project.

6.1 Usability Principles

For video games, D. Pinelle, N. Wong and T. Stach [36] developed a set of heuristic evaluations based on previous studies of Clanton and Federoff of design principles for games. These heuristics are oriented to the evaluation of performance of the game and how usable it is. Some important usability principles include consistency, game customization (video, audio, speed and difficult settings), capacity to skip non-playable sequences during the game, information of the status and statistics of the game, game instructions and help, training, and visual representation that is easy to understand [36].
6.2 Player involvement

Clantons [5] evaluated game design based on the level of engagement of the player into the game. This helps to determine if pressure during the game is fun and if it is better to avoid linear and plain sequences because they could be monotonous. This approach does not emphasize usability aspects, emphasizing that confusion is not fun. Key aspects to consider include: entertainment level, the game should give the perception that the player has the control at all times and provide an emotional connection between the player and his/her environment during the play time. The game should assure immersion with audio and visual content, making sure status information and scoring during the game time does not interfere with game play.

6.3 Principles of Gaming Therapy for Health and Rehabilitation Games

There are some features that a game program oriented to rehabilitation should consider, in particular the following set of principles were designed for patients recovering from stroke.

6.3.1 Environment

It is important to simulate daily or familiar activities and movements that a player can develop. These activities must be meaningful for them [4]. Additionally, if the game is a 3D virtual simulator, it must simulate well familiar aspects such as collision, gravity, and friction force [28].

42
6.3.2 Body involvement

The game should involve the whole body and natural movements which encompasses not only sensory organs such as eyes and ears, but also arms, legs, knees, neck and head.

6.3.3 Player adaptability, high customization

Every person has different needs; for that reason, the game must be designed to be adaptable to the height and width of a person. Another important issue is that this design should allow customization, so it would be easy to tailor different sessions for patients according to the type of injury, (for instance, people after suffering a stroke could undergo left/right hemiplegia, then they would need exercise more their right/left extremeties) and their capabilities [28].

6.3.4 Feedback information for player and therapist

It is necessary to record as many matrices as needed for subsequent feedback, then we can track the progress of every patient and propose more challenges and motivations.

6.3.5 Feedback of movements with use of tactile vibrators

Players need to realize what movements are right or wrong, for that reason the game should have devices which could be used in arms and legs, so they will get automatic feedback during the game [4].

6.3.6 Social interaction - Two players competition or cooperation

One specific aspect of seniors therapy after a stroke is social interaction while they are performing a task [28]; it can enhance self-esteem and also reduce loneliness.
6.4 Proposed Evaluation for Brain Boot Camp

Based on the principles mentioned before, we propose a set of principles to evaluate a game design. We want to consider usability principles such as game settings, status information of the player, how easy it is to learn how to play the game and clear goals. For player involvement, we are interested in knowing if the user had fun and if knowing that the game would improve his/her physical condition, that the player would be motivated to play the game very frequently. Figure 6.1 shows how these principles are present in the proposed heuristic evaluation for the game-based physical activity program that we are developing.

![Diagram showing the mapping of the heuristics and our proposed evaluation](image-url)

Figure 6.1: Mapping of the heuristics and our proposed evaluation
6.5 Survey Results

Based on the principles that we chose to do a evaluation survey of the Kinect-based prototype, we organized two sessions with a small number of participants. Users were around 77 years old on average.

The console was prepared with a 46-inch TV with a Kinect Camera and a computer controlled by the operator. Every person played for about 15 minutes. The operator also helped register their information into the database. Two of the participants mentioned they had vision problems and one of them highlighted that he had hearing problems as well. Results from the survey are summarized in Tables 6.1 and 6.1.

The following conclusions can be made from the results of this survey:

- Most players thought that the game was moderately easy to very easy to play. In particular, they were able to learn the game rules after a few minutes of playing.

- When we tested the visibility of objects, they said that objects were fairly easy to recognize. Colors of the objects were however somewhat problematic, requiring a change in the transparency of the objects, making them more opaque.

- Sounds were useful in helping them to recognize if a mistake was made when avoiding objects or touching them.

- Music appealed to the women, but not to the men.

- Information shown after finishing each level was mostly clearly understood. The font size of the message was large enough to see.

- Level of enjoyment was given fairly high scores, especially among the women, only one player chose a moderate score. They were constantly helped by the
researchers who were monitoring the game, explaining the plot of the game and giving instructions about how they should move. One participant thought that he could catch the objects with their hands, unfortunately, this feature is not implemented yet.

- Most people found the game to be interesting and stimulating.

- Regarding physical activity, older people of the group exhibited physical exertion, including an increase in respiration rate and sweating. It was also interesting that one of the gentleman, who described most of the aspects of interaction of the game as moderate, had the highest score when we asked about physical activity; interestingly this gentleman was the oldest in the group.

- With respect to cognitive challenge, older people thought it was highly challenging, while younger people from the group gave a lower score.

- When asked if knowing that this game was beneficial to their health, they stated that they would continue playing it. We obtained most positive answers to this question. With respect to the time that they would play the game, their responses were varied, but the older people agreed to play it at least 20 minutes daily.
<table>
<thead>
<tr>
<th>Participants</th>
<th>Person 1</th>
<th>Person 2</th>
<th>Person 3</th>
<th>Person 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>84</td>
<td>73</td>
<td>86</td>
<td>68</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>Female</td>
<td>Male</td>
<td>Male</td>
</tr>
<tr>
<td>Possible affections</td>
<td>Vision Problems</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Wear corrective lenses</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Difficulty hearing</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Own a computer or other device</td>
<td>computer</td>
<td>laptop</td>
<td>yes, but do not use</td>
</tr>
<tr>
<td></td>
<td>Like new technology</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Feedback:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How easy was it for you to understand and play the game?</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>How easy was it for you to see the objects on the screen?</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>How easy was it for you to hear the music and sounds in the game?</td>
<td>10</td>
<td>9</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Did you enjoy the music in the game?</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Were you satisfied with the feedback related to your performance?</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Was the feedback clear?</td>
<td>9</td>
<td>8</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Did the feedback help you stay interested in the game?</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Was the game enjoyable?</td>
<td>9</td>
<td>10</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Was the game interesting?</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Was the game stimulating?</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>How physically challenging was the game for you?</td>
<td>9</td>
<td>1</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>How cognitively challenging was the game for you?</td>
<td>7</td>
<td>2</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6.1: Test Results First Part
<table>
<thead>
<tr>
<th>Participants</th>
<th>Person 1</th>
<th>Person 2</th>
<th>Person 3</th>
<th>Person 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback: How likely is it that you would play this game everyday if you could and you knew that it would help your physical and cognitive function?</td>
<td>0 - Not at all, 5 - Moderately, 10 - Extremely</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>What is the maximum amount of time in minutes that you would play the game at any one time?</td>
<td>20-30 min</td>
<td>0-5 min</td>
<td>30-40 min</td>
<td>10-20 min</td>
</tr>
<tr>
<td>Do you have any additional comments?</td>
<td>Repetition would very helpful, a game suitable for wheelchair folk - perhaps using arm movements - certainly good for concentration practice</td>
<td>Bad hearing a handicap</td>
<td>Interesting early version - quite good, will be better as new version are developed</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.2: Test Results Second Part
Chapter 7: Gesture Recognition: Detection of Marching Motion

In this chapter, we explore methods for recognition of specific body movements. In particular, we are interested in the recognition of marching motion, which is a topic key interest for the Brain Boot Camp project.

7.1 Approaches for Marching Recognition

7.1.1 A Simple Marching Detection Algorithm

This method is based on a simple observation about the knee angle, which can be easily calculated from the knee, foot and hip points in 3-D space. The algorithm is described as follows:

Input:

- \( signal = \{a_1, a_2, \ldots, a_n\} \): knee angles
- \( maxTol \): specifies the minimum knee angle
- \( minTol \): specifies the minimum knee angle
- \( stdTol \): specifies a tolerance for subsection standard deviation

Output:

- out: true if signal corresponds to marching and false otherwise.

Steps:

1. Label each angle \( a_i \) as pertaining to one of three groups (0, 1, or 2). See Figure 7.1:
0. transition \((\text{minTol} \leq a_i \leq \text{maxTol})\),
1. knee up \((a_i < \text{minTol})\), and
2. knee down \((a_i > \text{maxTol})\)

Figure 7.1: Knee angle classification into three groups: 0. transition (yellow), 1. knee up (red), and 2. knee down (blue). The lines represent the thresholds of maximum and minimum angles to classify the angle samples.

2. Sum the number of points in each identified subsection, producing an output of the form \(\{n_j, \ell_j\}\), where \(n_j\) specifies the number of points in subsection \(\ell_j\). For example, the output produced for the data in Figure 7.1 would be:

<table>
<thead>
<tr>
<th>num-Points</th>
<th>5</th>
<th>7</th>
<th>7</th>
<th>7</th>
<th>8</th>
<th>6</th>
<th>8</th>
<th>6</th>
<th>7</th>
<th>10</th>
<th>6</th>
<th>4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>...</td>
</tr>
</tbody>
</table>

3. Calculate the mean, \(\mu_i\), and standard deviation, \(\sigma_i\), for each of the three identified groups, \(i = 0, 1, 2\).

4. For each group, \(i = 0, 1, 2\), count the number of subsections that are within one standard deviation, \(\sigma_i\), of the group mean, \(\mu_i\).

5. The input signal is said to correspond to marching, \((\text{out} = \text{true})\), if the following criteria are true:
(a) For each group, at least \( stdTol \, (\%) \) of the subsections are within one standard deviation of the group mean, and

(b) The sum of subsections belonging to groups 1 and 2 (knee up or down) is between 40\% and 60\% of the total number of subsections. Ideally, 50\% of the subsections belong to the transition group and the other 50\% to knee up and down.

6. If any of these conditions is not met, then the input signal corresponds to non-marching, \((out = false)\).

### 7.1.2 Chaos Theory and Embedding

The chaos method is based on the framework proposed by Ali, Basharat, and Shah [1] that uses the theory of non-linear dynamics and chaos theory to analyze human actions. With chaos theory one can study non-linear functions sensitive to initial conditions that appear to be in a chaotic state. Chaos theory analysis allows us to find deterministic and global stability in seemingly chaotic data. Many dynamical systems, such as the movement of the human body, can be represented as functions describing trajectories in a period of time. Specifically, dynamical systems can be represented by state-space models, where state variables \( X(t) = [x_1(t), x_2(t), \ldots, x_n(t)] \in \mathbb{R}^n \) define the status of the system at a given time, \( t \). State-space variables are located in \( n \)-dimensional manifolds in Euclidean space. The space of the state variables is called phase space. The path traced by a state variable over time is called a trajectory.

**Attractors**

An attractor is defined as the region of the phase space to which all the trajectories settle down as time limit approaches infinity. Invariant features of a system’s attractor can be measured in various ways. These measures quantify properties that
are invariant under smooth transformations of phase space. Invariants are classified into three classes: metric, dynamical, or topological. According to Ali, Basharat, and Shah [1], it is possible to take many invariants measures to describe a particular dynamical system, however in their experiments as well ours we take three: correlation dimension, correlation sum, and the Lyapunov exponent. Additionally we also consider the variance of the signal over time.

**Embedding**

Embedding refers to the process of mapping a one-dimensional signal into a \( m \)-dimensional signal. Embedding allows us to extract information required to identify and classify attractors of a dynamical system. In our case, the state variable describing the 15-point skeleton over time can be defined as \( V(t) = [p_1(t), \ldots, p_{15}(t)]^T \) for \( i = 1 : 15 \). In this case, \( p_i(t) = [x_i(t), y_i(t), z_i(t)] \), is the position of point/joint \( i \) in the skeleton at time \( t \). The trajectory of joint \( b \) can be denoted as:

\[
V_b(t) = [p_b(0), p_b(1), \ldots, p_b(t - 1)]. \tag{7.1}
\]

A one-dimensional signal can be obtained by selecting only a particular dimension. For example,

\[
[V_b(t)]_x = [x_b(0), x_b(1), \ldots, x_b(t - 1)] \tag{7.2}
\]

denotes all \( x \) coordinates of joint \( b \) over the trajectory (see e.g. Figure 7.2(a)).

An embedding of the one-dimensional signal \( (7.2) \) into \( m \)-dimensional space with delay \( \tau \) results in a matrix of the form:

\[
X_b = \begin{bmatrix}
x_b(0) & x_b(\tau) & \ldots & x_b((m - 1)\tau) \\
x_b(1) & x_b(1 + \tau) & \ldots & x_b(1 + (m - 1)\tau) \\
x_b(2) & x_b(2 + \tau) & \ldots & x_b(2 + (m - 1)\tau) \\
\vdots & \vdots & \ddots & \vdots
\end{bmatrix} \tag{7.3}
\]
Each row of this matrix is a point in $m$-dimensional space (see e.g. Figure 7.2(d)). The embedding parameters $\tau$ and $m$ are obtained using a method of mutual information minimization between $x_\theta(t)$ and $x_\theta(t + \tau)$ over $\tau$ and the false nearest neighbor method described in [1]. The method of embedding a one-dimensional signal into $m$-dimensional phase space is described bellow:

**Steps**

**Input:**

- $signal = \{a_1, a_2, \ldots, a_n\}$: any 1-D component of a joint over time (see Figure 7.2(a)).

**Output:**

- The Invariant features vector for the input signal.

$$F_a = [Lyapunov \ Corr.Sum \ variance \ Corr.Dimentions]$$

**Steps:**

1. Determine the first minimal of the $\tau$ function (see Figure 7.2(b)).

2. Choose a stable value for the embedding dimension $m$ (see Figure 7.2(c)).

3. Apply the chaos method to get the phase space of the reconstructed signal (see Figure 7.2(d)).

4. Calculate the invariant feature vector.

For our problem, we execute the embedding algorithm on each component $x, y$ and $z$ of the left-shoulder and left-knee joints as well as on the left-knee angle defined by the left-hip, left-knee, and left-foot joints.
7.2 Experimental Results

7.2.1 Marching and Non-Marching Datasets

A set of marching and non-marching datasets were obtained using the Kinect prototype of the game described in Chapter 5. Specifically, 11 time sequences belong to a marching action and 5 belong to non-marching. We asked people to perform a marching movement in front of the camera for 6 seconds on average while the operator recorded the signals of all joints. This task consisted of a typical marching movement alternating legs and arms. The axis coordinates that had more variation during this
task were $y$ and $z$. On the other hand, we also asked to people to play the game, performing random movements. In this task, we had different variations in all the axis coordinates $x, y$ and $z$.

### 7.2.2 Recognition Results of the Simple Marching Detection Algorithm

Input parameters for the simple marching detection algorithm were experimentally determined from the data itself. The $maxTol$, $minTol$ and $stdTol$ parameters were set to $175^\circ$, $165^\circ$, and $90\%$ in the first try and $80\%$ in the second try respectively. Recognition performance is particularly sensitive to $stdTol$, the threshold on the number of subsections that are one standard deviation away from the group mean.

#### Recognition Results

To present the results we generated a confusion matrix to show the recognition performance. Tables 7.2 and 7.1) show the recognition performance obtained for $stdTol = 90\%$ and $stdTol = 80\%$, respectively. Notice that with $stdTol = 90\%$ we get only $31.24\%$ of accuracy, while $stdTol = 80\%$ results in $100\%$ recognition accuracy.

<table>
<thead>
<tr>
<th></th>
<th>Marching</th>
<th>Non Marching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marching</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Non Marching</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 7.1: Confusion matrix for $stdTol = 90\%$.

<table>
<thead>
<tr>
<th></th>
<th>Marching</th>
<th>Non Marching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marching</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Non Marching</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 7.2: Confusion matrix for $stdTol = 80\%$. 
Analysis

The following remarks can be made about recognition of marching movement using the simple marching algorithm.

Advantages:

• This algorithm is linear, $O(n)$, on the size of the input signal, requiring very low computational work,

• A real-time implementation can be readily obtained using a sliding window approach. In this case, the algorithm is applied at time $t$ to a segment of the knee-angle signal $a_{t-m+1}, a_{t-1}, a_t$ of length $m$,

• Parameter selection can be done at run time during training or with the help of a game operator

Drawbacks:

• This method considers knee angle information only.

• The algorithm can be fooled by a person moving only his left knee in a marching motion.

• This method is hard to generalize to include additional points in the 15-point skeleton. Adding additional joints, such as the right knee angle, is difficult, since these are not independent variables. The algorithm would need to carefully encode and characterize this dependence.
7.2.3 Results of the Embedding Method

Embedding of the Marching Dataset. The input signal for left knee angle is shown in Figure 7.3(a). The result of embedding this signal into 3 dimensions is shown in Figure 7.3(b). Embedding into higher dimensions while maintaining a constant $\tau$ results in similar plots if the signal is projected into its first three dimensions. The invariant features parameters calculated from the embedded signal change are shown in Figure 7.5(a,c,e). The Lyapunov exponent appears to change the most.

Embedding of the Non Marching Dataset. The input signal for left knee angle is shown in Figure 7.4(a). The result of embedding this signal into 3 dimensions is shown in Figure 7.4(b). As can be observed, the pattern is significantly different than those obtained for marching in Figure 7.3.

The invariant features obtained from this embedding are shown in Figure 7.5(b,d,f). Notice the significant difference between the correlation sum of marching and non-marching signal. On the other hand, the correlation dimension between marching and non-marching appear to be very similar for the maximum Lyapunov exponent.
Recognition Results

For recognition of the marching motion, we construct a larger feature vector consisting of invariant features for all 3 dimensions of the left-shoulder and left-knee joints as well as the invariant features of the left-knee angle. As suggested in [1], we used the $k$-means clustering algorithm (in which any set of invariant features belong to one cluster with the nearest mean) to find two clusters of feature vectors and the leave-one-out cross validation method for evaluating recognition performance. The resulting similarity matrix shows that we obtained 75% of total accuracy (see Table 7.3).

<table>
<thead>
<tr>
<th></th>
<th>Marching</th>
<th>Non Marching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marching</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Non Marching</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 7.3: Confusion Table of Embedding Method
Figure 7.5: Invariant Features for Left Knee Angle Marching (left) and Non-Marching (right), for $m = 36$. 
Analysis

The following remarks can be made about recognition of marching movement using the embedding algorithm.

Advantages:

- The embedding algorithm is easily generalizable. Additional points of the skeleton can be added without any modifications to the code.
- The embedding approach is able to learn the underlying dependence between variables, without additional effort from the programmer.
- The embedding algorithm has the capability to distinguish additional motion patterns. In this experiment, we have considered only a marching motion and classify others as non-marching. Other motions such as lateral movement, jumping, catching, etc. may also be considered without modifications to the code.

Drawbacks:

- Recognition performance was not as good as that obtained with the simple marching algorithm.
- We can see 36.3% of the signal was misclassified.
- The computational complexity of embedding is higher than that of the simple approach. Computation of the feature vector values takes more time than in the simple marching approach. However, if the size of the input signal is kept small (using a sliding window approach), real-time calculation is possible.
- Classification via $k$-means is not sufficiently reliable. As implemented, classification requires input from the operator to determine which cluster might
correspond to marching or non-marching. Training may be used to find these centers ahead of time.
8.1 Conclusions

In this thesis, we explore the role of computer gaming as a way to elicit physical activity with complex motion patterns and decision making, often involved in real-world scenarios for an at risk group of older adults. Walking as a physical activity is an important and common mode of exercise for older adults and we focus our game development efforts on this form. In addition, we explore and evaluate the games developed as part of the Brain Boot Camp project using a set of gaming principles determined from the literature. We consider various different aspects, such as cognition, physical performance, evaluation of the senses, and enjoyment of the game.

We presented two games prototypes oriented to cognitive and physical measurement. These prototypes involved new gaming technology based on the Nintendo Wii and Microsoft Kinect. They allowed the player to interact with the game environment without using controls that for the majority of older people are difficult to use. We also presented and evaluated two methods for detection of a marching motion and point out key advantages and disadvantages between them. Additionally, we believe both methods can be implemented in real time.

8.2 Future Work

Future work may include the following topics:

- Implementation of marching recognition in real time as part of the Kinect prototype. This feature would allow the game to better determine quality of move-
ment and improvement over time.

- Adaptation of the game to particular needs of specific players. For example, people with right or left hemiplegia may need obstacles to be directed more towards a particular direction.

- Evaluate and refine game principles for older adults with a larger population. Evaluation of these principles on various disabilities would also be of interest.

- Social integration is a key aspect to consider for people after suffering a stroke. For this reason, the implementation of a game for two players or more players at the same time with a meaningful task would be helpful.

- Implement the game prototype as a complete game framework to allow the operator or therapist to create new games without the need to reprogram the code.
Bibliography


EDISON F. MUÑOZ

211 Harmon Court
Winston Salem, NC 27106
Tel: 919.886.2346
munoef9@wfu.edu

EDUCATION
Wake Forest University MS in Computer Science, August 2011
Ricardo Palma University BS in Computer Science (honors), December 2004

WORK EXPERIENCE
Wake Forest University, Translational Science Center May 2010 to Present
Research Assistant
Developed a novel virtual reality-based system using the Microsoft Kinect and the Unreal Development Kit for physical and cognitive training of adults with poor lower extremity functions. Technology used: XNA, Open CV, UDK, SQLite, Visual C++, C#.

Branddy.com June 2010 to Present
Software Architect and Project Manager
Led the development team of a start-up venture that seeks to use social media to enhance product commercialization. Technology used: ASP.NET, PHP, MySQL, JQuery, Push Technology and HTML5.

Department of Transportation and Communication Ministry of Peru 2007–2008
Project Leader
Led the software development team in charge of building the "National Transportation Sanctions System," which automates all the sanctioning processes for infractions and offences in public and private transportation around the country. Technology used: ASP.Net, C#, and ORACLE databases.

Costamar Group 2006–2007
Software Development Chief
Led a team of seven software engineers in nine projects involving website development, software integration, and deploying systems. All the projects were aligned with the business strategies of the company. Technology used: Microsoft .Net, SQL Server Databases, SQL Reporting Services, and J2EE.

Consultant
Provided technical expertise for client companies in Peru related to software development, databases administration, and processes design. Technology used: Microsoft .Net, SQL Server, and CMMI.

Belcorp and Nexolink 2004–2005
Analyst and Software Developer
Designed and implemented software solutions to manage sales and consolidation processes as well as mechanisms for communication between different web-based technologies. Technology used: Microsoft .Net, SQL Server Databases, C#, Progress Server Databases, Java.

AWARDS
Full academic scholarship by Wake Forest University for a Masters in Computer Science.

LANGUAGES
English and Spanish

70