Estimating Gesture Accuracy in Motion-Based Health Games

Christian Barrett, Jacob Brown, Jay Hartford],
Michael Hoerter, Andrew Kennedy,
Ray Hassan, and David Whittinghill
Purdue University
401 N. Grant St., West Lafayette, IN 47907
phone: 765-494-1353
email: dmwhittinghill@purdue.edu

Abstract

This manuscript details a technique for estimating gesture accuracy within the context of motion-based health video games using the MICROSOFT KINECT. We created a physical therapy game that requires players to imitate clinically significant reference gestures. Player performance is represented by the degree of similarity between the performed and reference gestures and is quantified by collecting the Euler angles of the player's gestures, converting them to a three-dimensional vector, and comparing the magnitude between the vectors. Lower difference values represent greater gestural correspondence and therefore greater player performance. A group of thirty-one subjects was tested. Subjects achieved gestural correspondence sufficient to complete the game's objectives while also improving their ability to perform reference gestures accurately.

1 Introduction

Traditionally, human interaction with computers has been overwhelmingly tactile: users move mice, tap keyboards, press buttons and, more recently, touch capacitive screens [BO10]. In 2009 Microsoft introduced the Kinect, a relatively inexpensive gaming peripheral that has seen widespread adoption [Eps13]. The Kinect uses a depth camera to enable markerless human-computer interaction using hand, arm, and full body gestures. This device opened up an entire new dimension of possibilities for computer-user interaction. Of particular consequence was the potential to create a new class of applications that would enable users with physical handicaps to interact with computers in ways that were never before possible [CCH11]. Pioneers have created a wide variety of novel applications using the Kinect [CCH11] [SHBS11] [BHP11] [CCWC13]. Though these applications impressively demonstrate the potential of the device, the specific development techniques and algorithms used to achieve these results are not immediately obvious. The approaches that have been implemented are necessarily idiosyncratic due to the wide-open nature of the Kinect itself. Further, optical human motion tracking is not an especially simple computational task. If depth-camera based applications are to reach their full potential impact, there must be more discussion of the algorithms and techniques that are the foundation of their development. Ideally, knowledge of human motion tracking should move beyond the periphery of user input design. Application developers do not have a reliable, established body of knowledge from which to draw in order to apply this technology to novel problems. True, the Kinect has democratized this functionality to some degree, but more work must be done.

Consider the example of two-dimensional GUI development. The idioms and algorithms behind two-dimensional GUI interface development are well known and have resulted in the proliferation of the GUI throughout our society. Kinect-based applications have a similar potential to proliferate but can never reach this full potential without broader dissemination of the algorithms and development techniques which make development of these applications possible. The loss to the fields of health and disability in particular, but also to entertainment, laboratory sciences, and other as yet unforeseen domains, could be considerable.

Hand and arm gesture tracking is a key function that many Kinect applications require. To address this need, we present a technique for capturing model gestures and assessing the similarity of all subsequent user attempts to recreate the initial model gesture. This technique has the advantage of not only informing the caller of a binary true-false gesture match, but also returns a quantification of the quality of the gesture match, thus allowing the caller to respond to varying match quality thresholds. By quantifying the degree of likeness, it becomes much easier to create applications for which similarity, or the lack there of, is of particular consequence. One prime example is the application about which this paper is centered, in which a physical therapy training game records how well a user has matched a prescribed therapeutic gesture. The observed similarity can then be used by a clinician to more precisely quantify to what degree a patient is adhering to his or her prescribed physical therapy.

In this manuscript, we describe the implementation of our novel arm gesture tracking technique and present the results that were gathered as part of an empirical health game study. Thirty-one test subjects played a physical therapy video game in which clinically significant gestures needed to be performed to complete the game successfully. The goal of this study was to create a game experience that coaches proper physical therapy technique while providing a measure of how well the player performed. An additional goal was to demonstrate that the algorithms and techniques our team developed could successfully drive a physical therapy game experience. Both goals were met. In addition, further discussion is offered documenting some practical challenges associated with Kinect development.

Note that this project was advised and supervised by pediatric orthopedists with the Peyton Manning Children's Hospital and St. Vincent's Health in Indianapolis. All physical therapy-related activities conducted in this study were directly derived from the clinics of these practicing physicians.

2 Previous Work

1) *Kinect and RGB-D Camera.* Though the Kinect is a relatively new device, markerless motion tracking using optical devices is not. Several sophisticated approaches exist [MHK06]. The arrival of RGB-D cameras (red, green, blue, & depth), like the one in the Kinect, has enabled a new class of research questions and accompanying applications.

Since its inception the Kinect has been used in a wide variety of contexts beyond entertainment. Lange et al. [LCS+11], discovered in their therapy games with the Nintendo Wii, that players had a tendency to "cheat" and perform movements that would trick the system into recognizing incomplete movements as valid. To combat this tendency they developed a therapy game based on the Kinect, that used the device's depth camera to more accurately determine what a player's limbs were actually doing during therapy.

Chang et al [CCH11] conducted a study in which children with cerebral palsy would raise their hands in a therapy gesture that would correspond to a whale's tail flapping on the screen. Greater gesture accuracy on the part of the player resulted in more dramatic flapping of the whale's tail on screen. Though the authors did not discuss their technical measure for gesture accuracy, they did report that patient engagement increased as well as patient performance in the intervention stage.

A variety of other uses for the device have also been implemented. For instance, Stowers et al [SHBS11] used the Kinect as a means of controlling a small quadrotor helicopter. The validity of the Kinect for use as a means of assessing posture and gait control was studied by Warburton et al. [CPF+12]. The Kinect was even used to evaluate the performance of a dancer [AKD+11].

2) *Health Games.* Health games are part of a burgeoning field of using technology to improve medical outcomes by improving behavioral health. Game technology has been especially helpful in the fields of rehabilitation and physical therapy. A study by Bania et al [BDT11] demonstrated that an intervention consisting of online educational and social support increased physical activity in sufferers of cerebral palsy

but only during the period of the intervention. When the intervention concluded, activity levels decreased to pre-intervention levels. This indicates that though technology can influence behavior, something more is needed to maintain it.

Games are one possible means of filling the motivation gap. Chang et al. [CCWC13] developed a Kinect-based vocational task training system for individuals with cognitive impairments that relied on operant condition techniques similar to those found in games. They discovered that in follow-up studies, subjects' acquisition of job skills were facilitated by the training.

Howcraft et al. [HKF+12] studied pediatric cerebral palsy patients who were instructed to play action video games (AVG), games that require physical activity to play, that simulated dance and boxing. The study was described as a "quantitative exploration of energy expenditure, muscle activation, and quality of movement". They discovered that patients exerted a moderate level of physical exertion during gameplay but, as significantly, the children reported high levels of enjoyment on the Physical Activity Enjoyment Scale (PACES) [KD91]. They noted the potential for AVG's to be enjoyable opportunities for the practice of motor activities.

Warburton et al. [WBH+07] conducted an experiment in which medical indicators of fitness such as maximal aerobic power, body composition, muscular strength, muscular power, flexibility and resting blood pressure were measured across groups of students who played either an interactive fitness video game or a standard training regimen. The results were interesting, as the game group demonstrated significantly superior measures on most items. The superior performance was a function of, more than any other factor, attendance and participation. Subjects who were in the game group showed up more often and therefore exercised more often. These findings demonstrate the power of games as an excellent motivator of behavior.

3 System Overview

3.1 Application Flow

Our application is a serious game called *Burnie* and has three distinct, sequential phases: *reference*, *game*, and *analysis*. In the *reference* phase, the user performs the series of gestures that are required to play the game. In the *game* phase, the user plays the game

and her or his performance is recorded by the game. In the *analysis* phase, the user's performance is analyzed.

The program first assembles a database of reference gestures from the user. For this study, four gestures are saved to the reference database. These serve as the template for a "correct" gesture - the ones that are compared against as the subject plays the game. Once the reference database is populated it can be used for multiple game sessions and does not need to be re-recorded. Once the reference database is assembled, subjects play the game. As they play, the computer is logging the orientations of their gestures into a log file. This log file forms the basis of the post-game analysis. The post-game analysis tells the player as well as the supervising clinician how well the player performed.
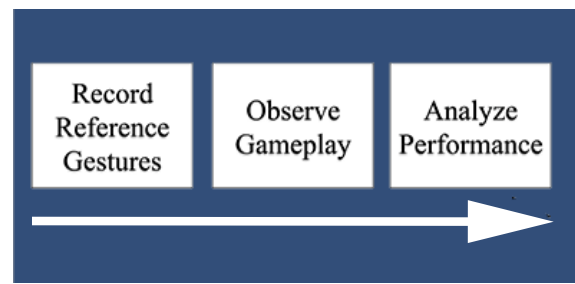

Figure 1: The functional pipeline. First, reference gestures are recorded. Second, the player plays the games and his motions are logged. Lastly, the logged motions are compared to the reference and analyzed.

3.2 Reference Recording

In order to achieve compliance with a user's prescribed therapy routine, it is important that each of a user's actions during therapy are performed as closely to the reference pose as possible. Therefore when recording the reference poses for each specific gesture in front of the Kinect, these performances must be the "ideal" version of these motions. To ensure accurate execution of the reference pose, the prescribing physical therapist should be present when the recording occurs in order to coach the patient toward recording the most ideal gesture possible.

A further benefit of recording the specific fashion in which each individual performs his or her reference gestures is to capture the idiosyncrasies of each individual player's body type and physical capacity. Van den Broeck et al. [BCM+10] demonstrated that treatment outcomes for diplegic children were improved

when treatment was individually tailored to the child versus general to the population. This application is intended to be used as an adjunct to physical therapy with pediatric cerebral palsy patients, thus the physical capabilities of the intended audience cannot be assumed to be uniform. Cerebral palsy is a multi-faceted disease that manifests differently in each individual [AARS+08] and thus requires an individualized treatment approach. Durstine recommends that, due to differences in disease expression, that treatments be tailored to each individual [DPF+00].

We define a pose as an array of Euler angles, one set for each observed joint. Euler angles were chosen primarily for practical reasons as the OpenNI API returns a steady stream of Euler angles for each joint the hardware observes, thus making overall software implementation simpler. Though quaternions would likely have been a more optimal choice due to their resistance to gimbal lock, we found that in this particular use case, Eulers worked reliably. The set of three Euler angles, α, β, γ, represents the orientation of a rigid body relative to a basis orientation (see *Figure 2*). A joint is characterized by the three values (α, β, γ), which are themselves calculated as the difference between the joint's orientation and the world's orientation. A pose is defined as the Euler angle orientations of all the joints of interest. All Euler angles in this manuscript are represented in degrees.

Each gesture within the game is characterized by the Euler angles of each elbow and each shoulder. (In this manuscript, we will refer to these points as *joints*). Thus the reference for each joint is recorded as the α, β, γ values of the joint's orientation in space at the time of recording. This creates a database of all possible joint combinations this particular user is capable of matching. A user profile is created in order to store each user's individual joint reference and performance data (see *Figure 3*).

The user must record four gestures which we refer to as: *flap*, *dive*, *strafe left*, and *strafe right*. These gestures are described in detail below.

Flap: starting with arms abducted in the neutral position, subject flexes the shoulder up however many degrees she is able in the frontal plane.
Strafe Left: subject keeps arms abducted to a neutral position while performing a lateral flexion of the spine by however many degrees she is able to her right side.
Strafe Right: subject keeps arms abducted to a neutral position while performing a lateral flexion of the spine
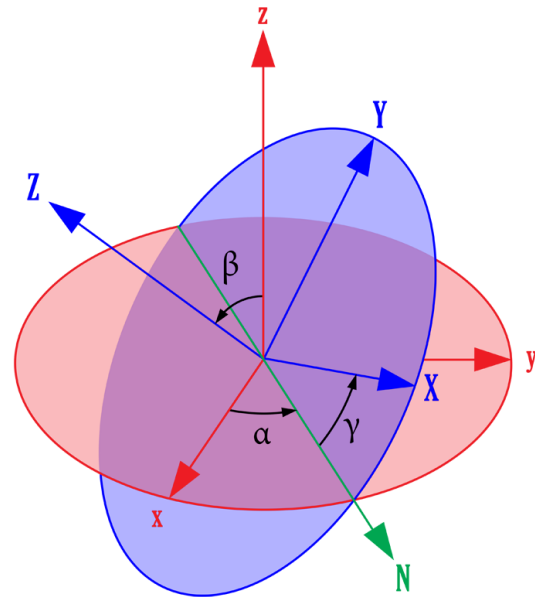

Figure 2: Visualization of Euler angles. The red circle represents the orientation of the world and is defined by three basis vectors. The blue circle represents the orientation of a given joint and is also defined by its three basis vectors. For each basis, α, β, γ, the difference (in degrees) between the reference and the joint's orientation is the Euler angle. It is this set of difference vectors that is the basis of our joint definitions.

by however many degrees she is able to his left side. *Dive*: Starting with arms abducted in the neutral position, subject finishes with arms abducted with inward rotation and flexed elbows and knees.

3.3 The Game

The game we created is called *Burnie* and is, in essence, a bird simulator. The player is represented on screen as a bird (named Burnie). Burnie's goal is to navigate through a virtual landscape by flapping, diving, strafing, and gliding. These character motions actions are triggered by the player performing flaps, dives, and strafes, respectively (gliding is the default position and does not require player motion). These actions cause Burnie to translate across the x and y plane of the world while forward (z) motion is "on rails" and is set at a pre-determined velocity (See *Figure 3*).

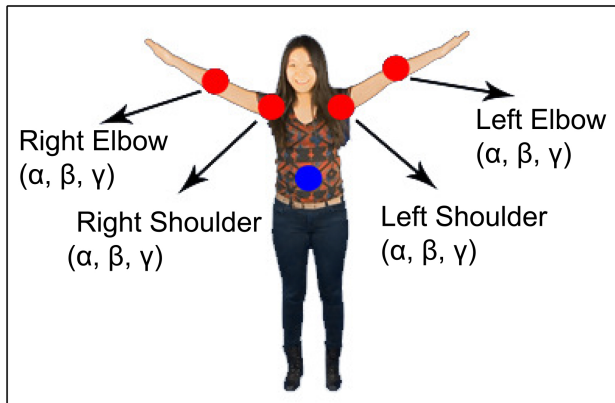The virtual game world is colorful, nature themed,

Figure 3: The four joints that are tracked by the game. The Euler angles for each joint are recorded as references and saved to that player's profile.
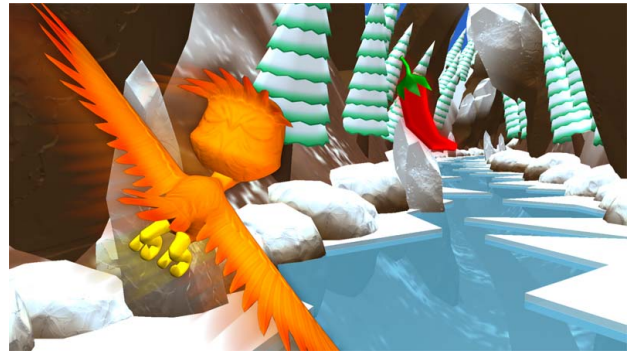

Figure 5: The game world varies across a variety of novel landscapes. Here, Burnie is diving for a hot pepper in the frozen mountain river area.
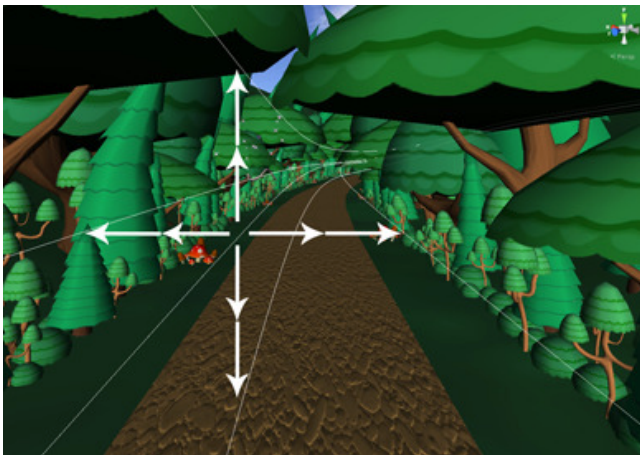

Figure 4: The player controls movement in the x and y planes. Forward velocity is "on rails" and controlled by the game.

and has a variety of novel landscapes (See *Figure 4*). Scattered throughout the world are a number of floating items the player can intercept in order to earn points, see amusing character animations, and hear fun sounds. A single play session of the game takes approximately fifteen minutes to complete.

3.4 Gesture Tracking

In motion capture systems, the user's body is "skeletonized" by the software and represented as a series of joints that are identified and connected via a stick-figure-like lattice. Obdrzalek et al. [OKO+12] studied the accuracy of the Kinect's ability to identify joints and limb lengths relative to the capabilities of PhaseS-

pace, Inc's [Pha14] marker-based motion capture system, Impulse. In their study, the Kinect's skeletonization was compared to one that the Impulse automatically generated (based on an extensive calibration routine), and a second that was manually calibrated using AutoDesk Motion Builder [Aut14]. A significant difference between the Impulse system and Kinect is that pose estimation on the Impulse requires a training or manual adjustment step whereas the Kinect makes a set of assumptions of "normal" body types based upon a large dataset of empirical observations and thus does not allow individual tuning [SSK+13]. Nonetheless, Obdrzalek et al. observes that for the Kinect, "the joint estimation is comparable to motion capture" [OKO+12], specifically for controlled body postures such as standing, and exercising arms.

Joint orientation is tested at a sampling rate of thirty times per second (the return rate of the Kinect). It should be noted that when conducting analyses on these data it is important to consider whether to sample at the native 30hz rate or to smooth the inputs in order to simplify calculations. The decision to sample should be a function of how kinetic the gameplay is. For games that require faster, sweeping motions, the higher sampling rate is preferred, while for steady, slow poses smoothed samples are typically sufficient.

With each sample, the application compares the current joint orientation against the user profile's database of known gestures. More specifically, the magnitude of the reference and observed Euler angles is calculated and compared. The smaller the magnitude, the greater the correspondence. As early developer tests estimated that a tolerance constant of < 155.0 applied to the reference comparison resulted in a level of appli-

cation responsiveness that felt "right" to most testers, this value was applied to all subjects. Further research on why this tolerance works best is being explored.

In the event that all four joints match one of the reference gestures an event is triggered that begins the logging of performance data. The magnitude between the α, β, γ values of the current joints from the α, β, γ values of the reference joints is calculated and stored in the log file. These values continue to be logged at a rate of 30 observations per second while the current gesture and the reference gesture are within the tolerance zone. When the gesture leaves the tolerance zone, logging of performance observations is paused until the next match occurs.

3.5 Gesture Quality Analysis

Gesture quality is determined along two dimensions: gesture *accuracy* and *time* on gesture. To achieve a more holistic assessment of the quality of a player's performance, both measures must be accounted for. These measures can serve as proxies for user stability and strength, respectively.

1) *Gesture Accuracy*. The first measure is the average difference of all observations across all joints unitized against the match tolerance. Or:

$$A = 1/N \sum_{i=1}^{J} \sum_{j=1}^{N} 1/t(\alpha_j, \beta_j, \gamma_j)$$

where:

- A = accuracy

- α_j, β_j, γ_j = difference between observed Euler components and reference Euler components

- J = number of joints in the pose

- N = number of observations

- t = magnitude tolerance from reference Euler angles

The resulting unitized value indicates the degree to which this gesture was performed perfectly according to reference. Lower values indicate higher accuracy: an accuracy value of 0.001 indicates perfect adherence to the reference, and an accuracy value of .999 indicates the poorest performance level possible while still maintaining the reference gesture (the algorithm does not allow for values of 0.0 or 1.0). Put another way, accuracy is quantified as, what percentage the observed pose matches the reference pose.

We would like to note that we tested using a uniform averaging of joints. It does seem likely however that this approach could be finessed by weighting the accuracy averaging such that a greater sensitivity is expressed toward the shoulders or the elbows, as a single degree of motion does not likely have a 1:1 relationship with gestural integrity. Put another way, a single degree of *shoulder* movement likely has a different impact than does a single degree of *elbow* movement when determining how accurately a user has performed a gesture. At present however, it is unclear the precise quantitative nature of this discontinuity. Future studies are planned to empirically determine the nature of this relationship but, at present, we suggest that uniform averaging is nonetheless sufficient for our immediate purposes.

2) *Gesture Time*. The time measure is an indication of the duration, on average, that a user was able to maintain a gesture that matched the reference. This is defined by a simple averaging of time held for each gesture instance.

$$d = 1/N \sum_{i=1}^{J} \sum_{j=1}^{N} 1/30$$

where:

- d = duration of gesture instance

- J = number of joints

- N = number of observations

- *1/30* = represents the sampling rate of the Kinect

The functional value of gesture time is somewhat subjective as the clinician who is supervising the therapy results may need to evaluate this on individualized basis. Not all gestures require the patient to hold the gesture for a long time, though some do. From a clinical perspective, the value of gesture time is determined by the idiosyncracies of each patient's treatment plan.

An aggregate calculation was nonetheless performed to assess the acuity of the software system. Prior to inclusion in our calculation, gestures recognized by the system must be a purposeful part of a contiguous sequence of observations. As such, long sequences of observations ($> 100ms$) were deemed to represent normal intended functioning of the software, whereas disconnected, brief observations were deemed to represent accidental or spurious gesture event triggering, and were excluded from the analysis.

4   Implementation

The absolute sensitivity and resolution of the Kinect are not published as the technology is proprietary to Primesense company. However, it is know that the Kinect uses an active form of depth information acquisition called structured light [YH10].

The game was developed using the Unity 3D 4.0 game engine. OpenNI drivers were used to communicate with the Kinect. Visual assets were created in Maya and Photoshop. User testing was performed using a large-screen HDTV with a hi-fidelity sound system on an Intel Quad Core i7 with 12GB's of RAM and an Nvidia Quadro 4000 graphics card. All code was written in C# from within Unity.

5   Results

Thirty-one subjects played the game as part of a pseudo experiment (a non-random population of convenience was used, recruited via flyers and word of mouth). Subjects ranged in age between 4 and 41.

Subjects played the game for approximately 15 minutes. During that time, they performed one of four gestures as close as possible to the reference gestures. *Flap*, *strafe left*, and *strafe right*, were of the most clinical significance so, in the interest of refining the analysis, only these are included. Of most interest was the performance of the elbow joints as these have the greatest potential range of motion and resultant clinical impact. Only elbow data for the three poses is analyzed.

The difference between the player's pose and the matching reference pose is represented on the accompanying charts. The difference is plotted on the y axis as *Gesture Accuracy* and time of gesture is plotted along x. Note that time is not absolute across x as different Euler elements could be in and out of tolerance simultaneously. Therefore x represents the performance trend *sequentially* but not in *temporally absolute* terms. Some linear interpolation was performed to fill gaps in the data.

This plotting scheme allows us to observe whether the player's ability to perform the required gestures increased, decreased, or stayed the same as they played the game.

The values in the accompanying charts plot the difference between the α, β, γ magnitude values of the observed gesture and the α, β, γ magnitude of the reference gesture. Note that there are data clustering

at differences less than < 155. These data represent observations that fall in our defined hit-detection tolerance range. These sub-tolerance observations represent a quasi-perfect match according to our schema and therefore, by definition, do not need improvement. It is the "problem" observations that are of interest as these can be used to specifically identify areas in which the subject needs to improve. As a result, the data in the range of greater than > 155 are of primary interest in this study. In the > 155 range, there is a small but clear trend across all gestures. The longer subjects played the game, the more precise their gestures became. Gesture correspondence improved as a function of time for all three gestures (see *Figures 6-11*).
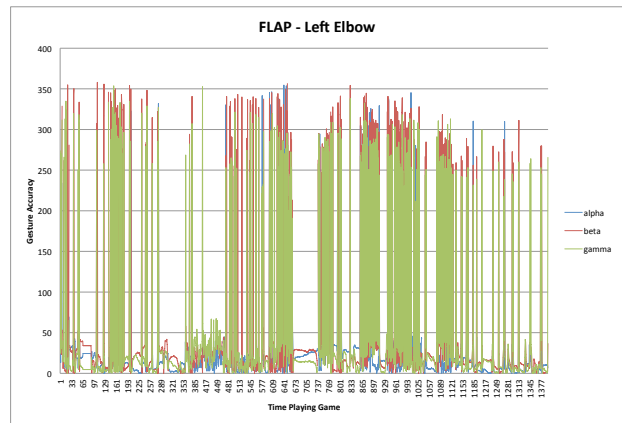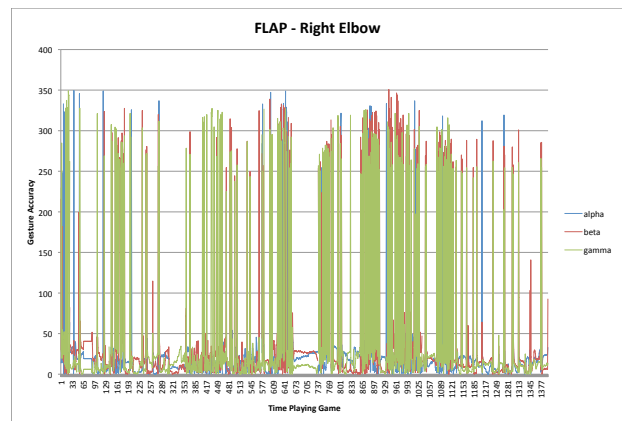

Figure 6: FLAP Gesture - Left Elbow

Figure 7: FLAP Gesture - Right Elbow

One of the primary objectives of the game is to train users to reliably perform physical therapy gestures more accurately. This study has demonstrated
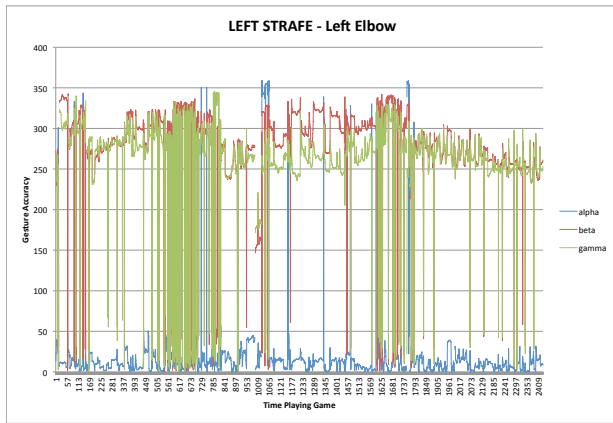
Figure 8: LEFT STRAFE Gesture - Left Elbow
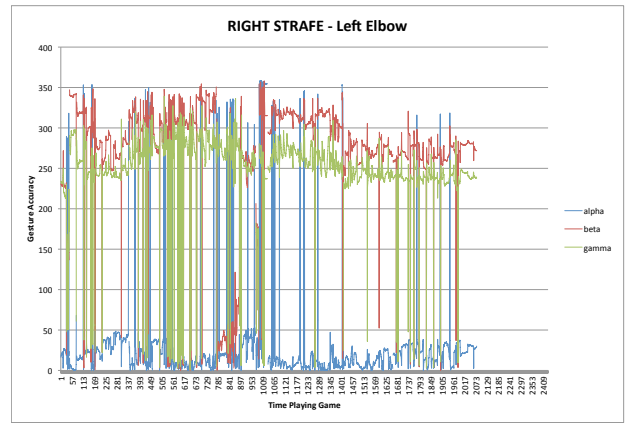

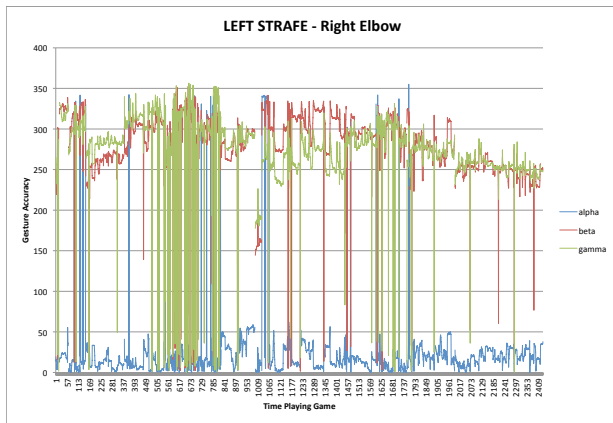Figure 10: RIGHT STRAFE Gesture - Left Elbow

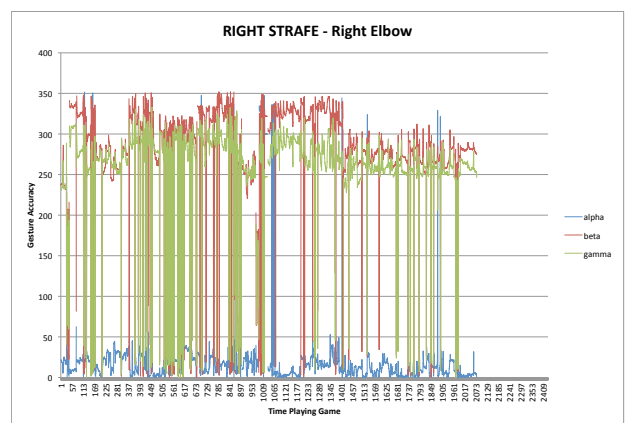Figure 9: LEFT STRAFE Gesture - Right Elbow

Figure 11: RIGHT STRAFE Gesture - Right Elbow

that, for the duration of a single play session, players were able to learn the gestures and improve their performance over the course of the play session. More research is required to determine to what degree this learning effect might persist across multiple play sessions and whether players are able to generalize this learning to non-therapeutic, real life scenarios.

6 Conclusions and Future Work

The results of this study demonstrate that our gesture tracking algorithm provides the player sufficient accuracy and responsiveness such that she is able to repeatedly perform the required physical therapy gestures. Further, we observed that as players continued to play the game, their ability to model the reference gestures increased slightly. Our approach is characterized by a comparison of the magnitude of the Euler angles (treated as a vector) of each of the player's significant

joints, versus the same measure of the player's joints as recorded in their earlier reference pose. We also demonstrated a method for quantitatively evaluating the player's performance by measuring and then averaging the user's closeness to the reference gestures, and recording the player's time on gesture. Taken together, these measures are loose proxies for precision and strength.

Though data were collected describing the motion of the shoulders, no discernible pattern could be gleaned from the analysis of these data. It is not entirely clear why a pattern did not emerge. It might be the case that our approach is most sensitive to either single degree-of-freedom observations, or is more attuned to observations in which the range of motion is large (as is the case with our study's elbow motions).

During the course of development and testing, the team encountered several aspects of Kinect development that potential developers should consider. First,

the Kinect functions best when not in sunlight as its depth camera, which functions in the infrared spectrum, easily becomes overwhelmed. Second, avoid any physical objects in the gameplay space if possible. It was not uncommon for the unit to pick up chairs and jackets as users. Third, the Kinect gets hot the longer it is used. We observed that the unit was more responsive to user input when it was not hot. Heat abatement of the unit should be considered when testing. Fourth, we observed a tendency for the Kinect to "prefer" different colors of garments on subjects. Given the manner in which the Kinect acquires depth information, this should not be an issue, nonetheless we discovered consistent trends in favor of dark solid colors with good contrast to the background. Lastly, frequent re-calibration of the Kinect unit was required. We achieved this by placing a hand over the sensor for five seconds. Upon removal, the unit performed well again.

There are several dimensions of the current study that will benefit from future work. At the time of development, OpenNI's Kinect drivers were regarded as the most compatible with Unity and thus strongly swayed our choice of drivers. Though the OpenNI drivers worked well enough for our needs, our development team would nonetheless have liked to experimented with either Microsoft's Kinect SDK, or Zigfu suite of Kinect drivers.

The last and most significant area for future work is in the area of further validation. For instance, the precision and strength proxies are only general indicators and would benefit from more validation via medical studies. Additionally, expanding this study to examine pediatric populations would greatly improve the breadth and impact of this work as the subjects in this study were predominantly college-age, healthy students. Repeating the study with children or with a population suffering from a motor handicap such as cerebral palsy would further validate this study's findings. Another promising future research direction is to align gesture recognition with the recently published Observed Movement Quality (OMQ) metric [JDD+12] or Neural Task Training (NTT) [SNRSE03] in order to more closely align the game's gestures with a broader taxonomy of physical therapy gestures.

7 Acknowledgments

References

[AARS+08] Heidi Anttila, Ilona Autti-Rämö, Jutta Suoranta, Marjukka Mäkelä, and Antti Malmivaara, *Effectiveness of physical therapy interventions for children with cerebral palsy: a systematic review*, BMC pediatrics **8** (2008), no. 14, ISSN 1471-2431.

[AKD+11] Dimitrios S. Alexiadis, Philip Kelly, Petros Daras, Noel E. O'Connor, Tamy Boubekeur, and Maher Moussa, *Evaluating a dancer's performance using Kinect-based skeleton tracking*, Proceedings of the 19th ACM International Conference on Multimedia (New York, NY, USA), ACM, 2011, pp. 659–662, ISBN 978-1-4503-0616-4.

[Aut14] Autodesk Inc., *Motion builder*, www.autodesk.com/products/motionbuilder/overview, 2014, Last visited October 17th, 2014.

[BCM+10] C. Van Den Broeck, J. De Cat, G. Molenaers, I. Franki, E. Himpens, D. Severijns, and K. Desloovere, *The effect of individually defined physiotherapy in children with cerebral palsy*, European Journal of Paediatric Neurology **14** (2010), no. 6, 519–525, ISSN 1090-3798.

[BDT11]   Theofani Bania, Karen J. Dodd, and Nicholas Taylor, *Habitual physical activity can be increased in people with cerebral palsy: a systematic review*, Clinical rehabilitation **25** (2011), no. 4, 303–315, ISSN 1477-0873.

[BHP11]   Antonio Padilha Lanari Bó, Mitsuhiro Hayashibe, and Philippe Poignet, *Joint angle estimation in rehabilitation with inertial sensors and its integration with Kinect*, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011, pp. 3479–83, ISBN 9781424441228.

[BO10]   Gary Barrett and Ryomei Omote, *Projected-Capacitive Touch Technology*, Information Display **26** (2010), no. 3, 16–21, ISSN 0362-0972.

[CCH11]   Yao-Jen Chang, Shu-Fang Chen, and Jun-Da Huang, *A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities*, Research in developmental disabilities **32** (2011), no. 6, 2566–70, ISSN 1873-3379.

[CCWC13] Yao-Jen Chang, Li-Der Chou, Frank Tsen-Yung Wang, and Shu-Fang Chen, *A kinect-based vocational task prompting system for individuals with cognitive impairments*, Personal and Ubiquitous Computing **17** (2013), no. 2, 351–358, ISSN 1617-4909.

[CPF+12] Ross A. Clark, Yong-Hao Pua, Karine Fortin, Callan Ritchie, Kate E. Webster, Linda Denehy, and Adam L. Bryant, *Validity of the Microsoft Kinect for assessment of postural control*, Gait & posture **36** (2012), no. 3, 372–377, ISSN 1879-2219, DOI 10.1016/j.gaitpost.2012.03.033.

[DPF+00] J. L. Durstine, P. Painter, B. A. Franklin, D. Morgan, K. H. Pitetti, and S. O. Roberts, *Physical activity for the chronically ill and disabled*, Sports Medicine **30** (2000), no. 3, 207–219, ISSN 0112-1642, DOI 10.2165/00007256-200030030-00005.

[Eps13]   Zach Epstein, *Microsoft says xbox 360 sales have surpassed 76 million units, kinect sales top 24 million*, bgr.com/2013/02/12/microsoft-xbox-360-sales-2013-325481/, 2013, Last visited October 16th, 2014.

[HKF+12] Jennifer Howcroft, Sue Klejman, Darcy Fehlings, Virginia Wright, Karl Zabjek, Jan Andrysek, and Elaine Biddiss, *Active Video Game Play in Children With Cerebral Palsy: Potential for Physical Activity Promotion and Rehabilitation Therapies*, Archives of Physical Medicine and Rehabilitation **93** (2012), no. 8, 1448–1456, ISSN 0003-9993, DOI 10.1016/j.apmr.2012.02.033.

[JDD+12] Anjo J. W. M. Janssen, Eline T. W. Diekema, Rob Van Dolder, Louis A. A. Kollee, Rob A. B. Oostendorp, and Maria W. G. Nijhuis-van der Sanden, *Development of a Movement Quality Measurement Tool for Children*, Physical Therapy **92** (2012), no. 4, 574–594, ISSN 0031-9023, DOI 10.2522/?ptj.20100354.

[KD91]   D. Kendzierski and K. DeCarlo, *Physical activity enjoyment scale: Two validation studies*, Journal of Sport and Exercise Psychology **13** (1991), no. 1, 50–64, ISSN 0895-2779.

[LCS+11] Belinda Lange, Chien-Yen Chang, Evan Suma, Bradley Newman, Albert Skip Rizzo, and Mark Bolas, *Development and evaluation of low cost game-based balance rehabilitation tool using the microsoft kinect sensor*, 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, IEEE, 2011, DOI 10.1109/IEMBS.2011.6090521, pp. 1831–1834.

[MHK06]   Thomas B. Moeslund, Adrian Hilton, and Volker Kruger, *A survey of advances in vision-based human motion capture and analysis*, Computer Vision and Image Understanding **104** (2006), no. 2-3, 90–126, ISSN 1077-3142, DOI 10.1016/j.cviu.2006.08.00.

[OKO+12]   S. Obdrzalek, G. Kurillo, F. Ofli, R. Bajcsy, E. Seto, H. Jimison, and M. Pavel, *Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population*, 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012, DOI 10.1109/EMBC.2012.634614, pp. 1188–1193.

[Pha14]   PhaseSpace Inc., *Impulse motion capture*, www.phasespace.com/impulse-motion-capture.html, 2014, Last visited October 17th, 2014.

[SHBS11]   John Stowers, Michael Hayes, and Andrew Bainbridge-Smith, *Altitude control of a quadrotor helicopter using depth map from Microsoft Kinect sensor*, 2011 IEEE International Conference on Mechatronics, IEEE, 2011, DOI 10.1109/ICMECH.2011.5971311, pp. 358–362, ISBN 978-1-61284-982-9.

[SNRSE03]   M. M. Schoemaker, A. S. Niemeijer, K. Reynders, and B. C. M. Smits-Englesman, *Effectiveness of neuromotor task training for children with developmental coordination disorder: a pilot study*, Neural Plasticity **10** (2003), no. 1-2, 155–163, ISSN 2090-5904, DOI 10.1155/NP.2003.15.

[SSK+13]   Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore, *Real-time Human Pose Recognition in Parts from Single Depth Images*, Communications of the ACM **56** (2013), no. 1, 116–124, ISSN 0001-0782, DOI 10.1145/2398356.2398381.

[WBH+07]   Darren E. R. Warburton, Shannon S. D. Bredin, Leslie T. L. Horita, Dominik Zbogar, Jessica M. Scott, Ben T. A. Esch, and Ryan E. Rhodes, *The health benefits of interactive video game exercise*, Applied physiology, nutrition, and metabolism **32** (2007), no. 4, 655–63, ISSN 1715-5312, DOI 10.1139/H07-038.

[YH10]   Sooyeong Yi and David Hwang, *Active Ranging System Based on Structured Laser Light Image*, Proceedings of SICE Annual Conference 2010, 2010, pp. 747–752, ISBN 978-1-4244-7642-8.