

Summer 2014

Using the Microsoft Kinect to assess human bimanual coordination

Joshua James Liddy
Purdue University

Follow this and additional works at: https://docs.lib.purdue.edu/open_access_theses

 Part of the [Biophysics Commons](#), [Computer Sciences Commons](#), and the [Medicine and Health Sciences Commons](#)

Recommended Citation

Liddy, Joshua James, "Using the Microsoft Kinect to assess human bimanual coordination" (2014). *Open Access Theses*. 645.
https://docs.lib.purdue.edu/open_access_theses/645

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

**PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance**

This is to certify that the thesis/dissertation prepared

By Joshua J. Liddy

Entitled

Using the Microsoft Kinect to Assess Human Bimanual Coordination

For the degree of Master of Science

Is approved by the final examining committee:

Jeffrey M. Haddad

Shirley Rietdyk

Laura Claxton

Jessica Huber

To the best of my knowledge and as understood by the student in the *Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Jeffrey M. Haddad

Approved by Major Professor(s): _____

Approved by: Timothy P. Gavin

07/16/14

Head of the Department Graduate Program

Date

USING THE MICROSOFT KINECT TO ASSESS HUMAN BIMANUAL
COORDINATION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Joshua James Liddy

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2014

Purdue University

West Lafayette, Indiana

ACKNOWLEDGEMENTS

To start off, I would like to acknowledge my significant underestimation of the amount of time, effort, and energy that goes into completing a thesis. I don't believe I will make the same mistake in regards to my dissertation, or at least, let's hope not. Without the help of the following people I would most certainly not have completed this project.

First and foremost, I would like to thank my adviser, Jeff Haddad for all the time and energy he has put into my work, which ranged from hashing out details regarding experimental design to reading (and re-reading) numerous drafts with deadlines encroaching. I have learned so much in the past two years from Jeff and hope that I can continue to do so in the coming years. Unfortunately for us, I don't think there will be a Wings Over West Lafayette by the time I graduate.

Second, I would like to thank my committee members, Shirley Rietdyk, Laura Claxton, and Jessica Huber, for their invaluable input to my project. All three have been great mentors and have truly helped me to tweak the finer details of the project which both strengthened it and made it more manageable, in my opinion.

Next, I would be remiss to leave out Howie Zelaznik. Howie had conceived the original idea for the project, using Kelso's bimanual coordination paradigm, after a seminar presentation in which I had struggled (to say the least) to brainstorm ideas for how to move forward. He also provided the code for the metronome and the algorithm for finding each hand movement cycle, both of which were critical to the project.

I would also like to thank (in no particular order) the undergraduate students who helped me to collect and process all of the data: Trey DeHart, Jonathan Stewart, Erin Mehringer, Lauren Ashenbrenner, and Mara Schmidt. Without them I would be nowhere close to finished collecting data and would likely be suffering through tracking the more than 12 hours of hand wiggling that was collected.

Last, but certainly not least, I want to thank Rachel Wichlinski for her unrelenting support. I don't know if I can thank her enough for her help copying and pasting analog data in Excel while I was frantically analyzing the data the week the document was due. She has been my biggest supporter throughout this process and has been confident in my abilities even as I have wavered.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	viii
ABSTRACT	x
CHAPTER 1. INTRODUCTION	1
1.1 Commercial Motion Capture Systems	1
1.2 Microsoft Kinect	2
1.3 Coordination.....	5
1.4 Purpose.....	6
CHAPTER 2. REVIEW OF LITERATURE.....	8
2.1 Motion Capture Technology	8
2.1.1 Marker-Based, Optical Motion Capture.....	9
2.1.1.1 Passive (Reflective) Marker Systems	9
2.1.1.2 Active (Optoelectronic) Marker Systems	10
2.1.2 Shared Limitations of Passive and Active Systems	10
2.1.2.1 Marker Placement.....	10
2.1.2.2 Influences on Movement Patterns	11
2.1.2.3 Controlled Collection Environment.....	12
2.1.2.4 Soft Tissue Artifact.....	12
2.1.3 Summary of Marker-Based Motion Capture Limitations	13

	Page
2.2	Markerless, Optical Motion Capture..... 14
2.2.1	General Model for Human Pose Estimation 15
2.2.1.1	Model and Scene Initialization 16
2.2.1.2	Tracking..... 17
2.2.1.3	Pose Estimation 20
2.2.1.4	Pose Recognition 21
2.2.2	Limitations of Markerless Motion Capture..... 22
2.3	Repurposed Technology..... 23
2.4	Microsoft Kinect 24
2.4.1	How the Kinect works..... 25
2.4.2	Xbox vs. Microsoft Kinect..... 28
2.4.3	Current Research..... 31
2.4.3.1	Spatial Capabilities..... 32
2.4.3.2	Skeletal Tracking..... 35
2.4.3.3	Postural Control..... 40
2.4.3.4	Gait Assessment 45
2.4.3.5	Workplace Ergonomics 47
2.5	Open Questions 49
2.6	Coordination in Human Movement..... 51
2.6.1	Dynamical Systems Theory 51
2.6.2	Attractors..... 51
2.6.3	Bernstein’s Degrees of Freedom Problem 53
2.6.4	Absolute and Relative Coordination 54
2.6.5	Relative Phase of Human Movement..... 55

	Page
2.6.6	Relative Phase Assessment Techniques.....57
2.7	Movement System Variability 59
CHAPTER 3.	METHODS..... 61
3.1	Participants 61
3.2	Apparatus 61
3.3	Setup..... 62
3.4	Procedure..... 62
3.5	Data Analysis 63
3.6	Statistical Analysis 64
CHAPTER 4.	MANUSCRIPT 66
4.1	Introduction 66
4.2	Methods..... 69
4.2.1	Participants.....69
4.2.2	Microsoft Kinect69
4.2.3	Vicon.....69
4.2.4	Apparatus and Procedure70
4.2.5	Data Processing and Analysis71
4.2.6	Statistical Analysis.....72
4.3	Results..... 74
4.3.1	Mean CRP74
4.3.2	Mean CRP Variability.....76
4.4	Discussion 79
REFERENCES 84

LIST OF TABLES

Table	Page
Table 4.1. Mean CRP and standard error (in parentheses) of the Kinect and Vicon.....	75
Table 4.2. Standard deviation of mean CRP and standard error (in parentheses) of the Kinect and Vicon	77

LIST OF FIGURES

Figure	Page
Figure 4.1. Normalized phase planes of the right and left hands for each device at a frequency of 1.67 Hz from an exemplar participant: a) Vicon right hand, b) Vicon left hand, c) Kinect right hand, d) Kinect left hand.....	73
Figure 4.2. Ensemble CRP curves created by normalizing cycles to 100 data points and averaging across all cycles of a given frequency. The CRP curves in this figure are calculated from the phase planes in Figure 4.1, above. The solid lines represent ensemble CRP, the dotted lines represent ensemble CRP plus or minus the standard deviation of CRP.....	74
Figure 4.3. Mean CRP measured for in-phase and anti-phase patterns of coordination for the Kinect and Vicon. The Kinect significantly underestimated the anti-phase pattern and significantly overestimated the in-phase pattern. *p < .0001	75
Figure 4.4. Variability of CRP for the Kinect and Vicon across both patterns of coordination. The Kinect overestimated variability for both patterns. Both systems were able to capture the coordination stability differences between patterns. *p < .0001	77

Figure	Page
Figure 4.5. Relative phase variability for both devices across the various movement frequencies for the in-phase pattern. Variability measured by the Kinect was higher at all frequencies ($p < .0001$). Kinect variability was similar at all frequencies up to 2.0 Hz with the highest frequency being significantly higher than the rest ($p < .0001$). All Vicon frequencies were similar except for the highest frequency ($p < .0001$).....	78
Figure 4.6. Relative phase variability for both devices across the various movement frequencies for the anti-phase pattern. Variability measured by the Kinect was higher at all frequencies ($p < .0001$). Kinect variability was similar at all frequencies. All Vicon frequencies were similar except for the highest two frequencies ($p < .0001$).	78

ABSTRACT

Liddy, Joshua J. M.S., Purdue University, August 2014. Using the Microsoft Kinect to Assess Human Bimanual Coordination. Major Professor: Jeffrey M. Haddad.

Optical marker-based systems are the gold-standard for capturing three-dimensional (3D) human kinematics. However, these systems have various drawbacks including time consuming marker placement, soft tissue movement artifact, and are prohibitively expensive and non-portable. The Microsoft Kinect is an inexpensive, portable, depth camera that can be used to capture 3D human movement kinematics. Numerous investigations have assessed the Kinect's ability to capture postural control and gait, but to date, no study has evaluated its capabilities for measuring spatiotemporal coordination. In order to investigate human coordination and coordination stability with the Kinect, a well-studied bimanual coordination paradigm (Kelso, 1984, Kelso; Scholz, & Schöner, 1986) was adapted.

Nineteen participants performed ten trials of coordinated hand movements in either in-phase or anti-phase patterns of coordination to the beat of a metronome which was incrementally sped up and slowed down. Continuous relative phase (CRP) and the standard deviation of CRP were used to assess coordination and coordination stability, respectively.

Data from the Kinect were compared to a Vicon motion capture system using a mixed-model, repeated measures analysis of variance and intraclass correlation coefficients (2,1) ($ICC_{(2,1)}$).

Kinect significantly underestimated CRP for the the anti-phase coordination pattern ($p < .0001$) and overestimated the in-phase pattern ($p < .0001$). However, a high ICC value ($r = .097$) was found between the systems. For the standard deviation of CRP, the Kinect exhibited significantly higher variability than the Vicon ($p < .0001$) but was able to distinguish significant differences between patterns of coordination with anti-phase variability being higher than in-phase ($p < .0001$). Additionally, the Kinect was unable to accurately capture the structure of coordination stability for the anti-phase pattern. Finally, agreement was found between systems using the ICC ($r = .37$).

In conclusion, the Kinect was unable to accurately capture mean CRP. However, the high ICC between the two systems is promising and the Kinect was able to distinguish between the coordination stability of in-phase and anti-phase coordination. However, the structure of variability as movement speed increased was dissimilar to the Vicon, particularly for the anti-phase pattern. Some aspects of coordination are nicely captured by the Kinect while others are not. Detecting differences between bimanual coordination patterns and the stability of those patterns can be achieved using the Kinect. However, researchers interested in the structure of coordination stability should exercise caution since poor agreement was found between systems.

CHAPTER 1. INTRODUCTION

1.1 Commercial Motion Capture Systems

Optical, marker-based motion capture systems are the gold-standard for capturing three-dimensional (3D) human kinematics (Best & Begg, 2006; Corazza, Mündermann, & Andriacchi, 2006; Mündermann, Corazza, & Andriacchi, 2006; Visser, Carpenter, van der Kooij, & Bloem, 2008). These systems are accurate, reliable, and capable of tracking a variety of movements in multiple domains, including posture and gait assessments, clinical diagnostics, physical rehabilitation, and workplace ergonomics. The most common systems use passive (reflective) or active (optoelectronic) markers to track movements of anatomical segments. However, optical, marker-based motion capture has various drawbacks including the time consuming placement of markers, marker placement variability, the possibility of markers altering movement patterns, the need for carefully controlled collection environments, and soft tissue artifact (Andriacchi & Alexander, 2000; Cappozzo, Catani, Leardini, Benedetti, & Della Croce, 1996; Della Croce, Leardini, Chiari, & Cappozzo, 2005; Leardini, Chiari, Della Croce, & Cappozzo, 2005; Mündermann, Corazza, & Andriacchi, 2006). Additionally, and perhaps most importantly, these systems are often prohibitively expensive and non-portable.

Advances in the field of computer vision have led to the development of vision-based, markerless motion capture systems (Moeslund & Granum, 2001; Moeslund, Hilton, & Krüger, 2006; Poppe, 2007; Wang, Weiming, & Tieniu, 2003). Markerless kinematic systems typically use pose estimation to analyze dynamic movements. Human pose estimation creates a virtual kinematic skeletal model based on the orientation of the body segments. However, pose estimation is a complex process due to highly variable body shapes and sizes (including clothing), and potential complexities of the collection environment, including moving objects and environmental clutter (Poppe, 2007). Thus, markerless motion capture systems have operational assumptions that can ultimately influence accuracy. These assumptions require users to carefully control the environment, or restrict the types of movements that can be captured (Moeslund & Granum, 2001).

Both markerless and marker-based commercial motion capture systems utilize multiple cameras which limit portability and raise cost. Additionally, these systems do not allow human movement to be captured outside of the laboratory. As a result of these limitations, recent research has focused on repurposing gaming technology to capture movement in more naturalistic conditions.

1.2 Microsoft Kinect

The Microsoft Kinect is a hands-free gaming peripheral designed for the Xbox console. The Kinect creates a natural user interface through gesture and speech commands that facilitates real-time user interaction. Considerable interest has developed in researching the Kinect's utility for 3D kinematic data acquisition, as it is a portable, inexpensive, markerless, depth camera. The Kinect features an RGB camera, IR emitter

and camera, and a microphone array. Microsoft has provided a software development kit (SDK) to allow access to the Kinect's various data streams. With the SDK, the Kinect has potential to address many of the aforementioned issues with commercial systems.

The Kinect's pose estimation algorithm, described by Shotton et al. (2012), constructs a 3D, virtual kinematic skeleton of the user by investigating body part distributions from raw depth data and proposing joint configurations. The Kinect's pose estimation technique is beneficial in that it provides a sophisticated solution to the problem of robust pose estimation across variable body morphologies given the lack of an *a priori* model. Additionally, joint and limb orientation proposals are made without references to previous frames, thus minimizing error propagation. Nonetheless fitting a virtual skeleton frame-by-frame creates inconsistencies in joint placement and segment lengths. Imprecise anatomical landmark estimation can have profound impacts on joint kinematics (Della Croce et al., 2005), severely restricting the Kinect's usefulness in biomechanical applications unless this issue can be resolved *a posteriori* or measures concerned with joint displacement magnitudes and amplitudes are considered (Mobini, Behzadipour, & Foumani, 2013) as opposed to absolute positions.

There are two ways to extract kinematic data from the Kinect. First, the computed depth map can be accessed and custom algorithms to identify and quantify human movement can be written. Stone & Skubic (2011a, 2011b) demonstrated the accuracy and reliability of the Kinect to extract spatial and temporal parameters of gait as well as the associated variability from the computed depth map. This technique, although promising, requires extensive technical knowledge. Alternatively, the Microsoft SDK offers a

skeletal tracking feature which can be used to capture joint kinematics. This allows researchers to collect movements of interest and perform their own analysis.

Recent gait research utilizing the skeletal tracking feature suggests the Kinect is capable of extracting basic spatial measures such as stride length (Clark, Bower, Menitplay, Paterson, & Pua, 2013; Stone & Skubic, 2011a). However, temporal measures of gait are more difficult to collect potentially less reliable due to difficulties distinguishing anatomical landmarks of the foot from the ground as well as the lack of anatomical landmarks on the foot to distinguish gait events (Clark, Bower, Mentiplay, Paterson, & Pua, 2013; Stone & Skubic 2011a). However, it is unclear how well the Kinect captures the stride-to-stride variability of these parameters using the skeletal tracking feature.

In addition to gait assessments there have been numerous investigations of the Kinect's ability to accurately perform postural assessment across a variety of movements and patient populations. Studies have shown that the Kinect has potential for accurately collecting human movement in older adults and neurological populations (Galna, Barry, Jackson, Mhiripiri, Olivier, & Rochester, 2014; Obdrzalek, Kurillo, Ofli, Bajcsy, Seto, Jimison, & Pavel, 2012). Other work has assessed the validity and reliability of the Kinect joint positions during reaching and standing posture (Clark, Pua, Fortin, Ritchie, Webster, Denehy, & Bryant, 2012). These studies have provided evidence that the Kinect can accurately capture temporal kinematics but is slightly less accurate at capturing spatial kinematics than commercial motion capture systems (Galna et al., 2014). The current body of research utilizing the Kinect for biomechanical applications is promising

but ongoing research needs to address how well the device can capture complex, dynamic patterns of movement.

1.3 Coordination

Previous Kinect research has focused exclusively on spatial or temporal outcome measures such as joint angles, joint displacement, and gait parameters. The ability of the Kinect to examine human coordination has not been examined. Coordination focuses on the on cooperative action of multiple body segments to realize task-specific goals. The body segments of interest are treated as a single, functional unit, or coordinative structure, which is flexibly and temporarily assembled given the environmental, individual, and task constraints imposed on an organism (Kelso, 1995; Newell, 1986).

Turvey (1990) highlighted the ubiquitous nature of rhythmic movement in nature and its importance in understanding the temporal evolution of patterns. Examinations of rhythmic bimanual coordination have demonstrated that humans can coordinate their fingers (or hands) in either an in-phase (0°) or anti-phase (180°) pattern at low frequencies (Kelso, 1984, Kelso; Scholz, & Schöner, 1986). As the frequency of oscillation is increased, movements started in-phase remain in-phase. When the movements are initiated in the anti-phase pattern and frequency is increased a spontaneous transition occurs after passing through a specific movement frequency, termed the transition frequency (f_t). Relative phase is the collective variable used to assess coordination that describes the spatiotemporal relation of two segments. Pattern stability can be determined by examining the variability of relative phase. Immediately preceding f_t , the variability of relative phase increases dramatically. The structure of

variability in Kelso's paradigm denotes destabilization of one pattern and prescribes oncoming change to another pattern.

Advances in the field of non-linear dynamics have demonstrated that variables such as relative phase may better capture the true dynamics of motor tasks. Relative phase measures provide insights into spatiotemporal changes in the human movement system that may be undetectable by examining spatial or temporal information alone. Relative phase provides information about how segments are being coordinated and the associated variability provides information about the stability of coordination. Additionally, there have been numerous examinations of variability in the control and coordination of human movement which suggest the potentially functional role of variability. These investigations have suggested that variability facilitates postural flexibility and adaptability in response to perturbations and changes in environmental, individual, or task constraints (Riccio, 1993, van Emmerik & van Wegen, 2002).

1.4 Purpose

This investigation provides a novel approach for assessing the Kinect's capabilities as a low-cost, portable motion capture device in biomechanical applications. To date no study has investigated the ability of the Microsoft Kinect to accurately and reliably measure human coordination. We will assess the ability of the Kinect capture the relative phase dynamics using a paradigm similar to Kelso (1984). Additionally, if the Kinect proves capable of measuring the variability of coordination, it could potentially be used as a tool for diagnosis of developmental disorders (Volman, Laroy, & Jongmans, 2006), age-related and pathological changes (Plotnik, Giladi, & Hausdorff, 2007; Serrien,

Swinnen, & Stelmach, 2000; van Emmerik, McDermott, Haddad, & van Wegen, 2005), and gait related asymmetries and injuries (Haddad, van Emmerik, Whittlesey, & Hamill, 2006a ; Hamill, van Emmerik, & Heiderscheit, 1999) as often times these changes are accompanied by subtle changes to the structure of movement variability.

CHAPTER 2. REVIEW OF LITERATURE

2.1 Motion Capture Technology

The origins of motion capture and the analysis of human movement can be traced back to the 1800s. In the 1830's, the Weber brothers studied spatial and temporal parameters of human gait (Weber & Weber, 1836). They used basic tools available at the time including a chronograph, meter stick, and a diopter to examine stride length, gait velocity, the support and swing phases of gait, and the relationship between stride time and length (Medved, 2002). In the 1870s, Marey and Muybridge began to use photographic techniques, such as chronophotography, to analyze animal movement (Marey, 1874; Muybridge, 1887). These rudimentary techniques provided a platform for the development of modern motion capture technology. With the introduction of modern computers in the 1970's, motion capture systems with a high spatial and temporal resolution were developed. The current gold-standard technology for collecting three-dimensional (3D) kinematic data during human movement requires markers to be affixed to anatomical landmarks or body segments of interest (Mündermann, Corazza, Chaudhari, Andriacchi, Sundaresan, & Chapella., 2006). The positions of these markers are then captured using a variety of technologies. The most common systems are optical-based and utilize either passive (reflective) or active (optoelectronic) markers

(Best & Begg, 2006; Corazza, Mündermann, Chaudhari, Demattio, Cobelli, & Andriacchi, 2006; Mündermann, Corazza, & Andriacchi, 2006; Visser et al., 2008). The following sections discuss the characteristics of these systems, including their advantages, and disadvantages.

2.1.1 Marker-Based, Optical Motion Capture

2.1.1.1 Passive (Reflective) Marker Systems

Passive marker systems are frequently used in the examination of human movement coordination, posture, and gait. Multiple cameras flood the collection space with pulsed infrared (IR) light, which is reflected back to the cameras by the markers. Each camera records a two-dimensional (2D) image of the markers from the recorded reflection. Data from a minimum of two cameras is then used to determine the 3D coordinates of the markers within the calibrated space. Post-processing is necessary to identify specific trajectories since individual markers are not identified during the collection process. To improve marker visibility and 3D construction accuracy, six to eight cameras are generally recommended when collecting kinematics (Best & Begg, 2006). There are several disadvantages of passive marker systems including ghost markers (i.e. ambient environmental reflections), marker fallout (i.e. poor reflective properties or occlusion) or poor camera exposure settings, and the need for post-processing to identify trajectories (Best & Begg, 2006). One advantage of passive marker systems is that they do not have wires attached to each marker. Thus, the possibility of impeding movement is reduced as compared to active marker systems. Passive marker systems remain popular in both

clinical and research settings due to their relatively high accuracy, ease of marker placement, and the ability to accommodate large collection spaces (Best & Begg, 2006).

2.1.1.2 Active (Optoelectronic) Marker Systems

Active marker systems utilize infrared emitting diodes (IREDs) that pulse in sequence to identify each marker. When a particular IRED pulses, its 3D coordinates are recorded by a rigid bank of cameras (Best & Begg, 2006; Winter, 2009). Active marker systems alleviate many of the aforementioned issues associated with passive marker systems. The primary disadvantage of active systems is that the markers are typically wired. Thus, subject preparation time is generally longer and natural movement may be impeded (Payton & Bartlett, 2007). This concern is particularly problematic in gait studies where the wires have the potential to trip participants.

2.1.2 Shared Limitations of Passive and Active Systems

2.1.2.1 Marker Placement

Active marker systems share many of the same disadvantages as passive markers systems, including difficulty attenuating ambient reflections, marker slippage, markers outside of the collection space, poor camera coverage of the collection space, calibration errors, marker occlusion, high cost, and low portability (Best & Begg, 2006). The biggest disadvantage of both passive and active marker systems is the time consuming process of placing markers on participants (Mündermann, Corazza, & Andriacchi, 2006; Mündermann, Corazza, Chaudhari, et al., 2006; Simon, 2004; Winter, 2009). Placement time varies depending on the number of markers required to suitably capture the motor behavior of interest. Additional time may be required to replace markers prone to

chaffing or contacting body segments, such as markers medial to the knee joints during walking. In clinical settings, marker placement time is particularly inconvenient (Mündermann, Corazza, Chaudhari, et al., 2006). Minimizing marker placement time can maximize the number of assessments performed by clinicians and also reduce the time burden for patients. In addition, all marker-based systems introduce error due to placement variability (both within and between participants) (Della Croce et al., 2005; Mündermann, Corazza, & Andriacchi, 2006). Errors can be introduced through difficulties palpating anatomical landmarks, soft tissue thickness variability, and the palpation method used (Della Croce et al., 2005). Inter-participant error tends to be higher than intra-participant error due to the morphological differences between individuals (Della Croce et al., 2005). Marker placement errors lead to errors defining anatomical frames, such errors may be magnified in the calculation of joint kinematics (Cappozzo, Della Croce, Leardini, & Chiari, 2005; Della Croce et al., 2005).

2.1.2.2 Influences on Movement Patterns

Markers can often induce unnatural movement patterns. For example, to avoid knocking markers off, participants may adopt alternative patterns of movement. This issue is especially troublesome when subtle changes in movement are associated with pathology (Andriacchi & Alexander, 2000; Mündermann, Corazza, Chaudhari, et al., 2006). For instance, marker placement medial to the knee joint could produce changes in step width, which in turn lead to higher step width variability, a variable generally associated with a greater risk of falling in older adults (Hausdorff, Rios, & Edelberg, 2001; Maki, 1997). Thus, in order to minimize errors in data interpretation, investigators

should utilize marker setups that minimally impede participants (Andriacchi & Alexander, 2000; Chiari, Della Croce, Leardini, & Cappozzo, 2005; Mündermann, Corazza, & Andriacchi, 2006).

2.1.2.3 Controlled Collection Environment

The collection environment required to use either active or passive marker systems must be strictly controlled to avoid ambient reflections and objects obscuring camera views (Best & Begg, 2006; Cappozzo et al., 1996). The number of cameras needed requires very large spaces for data collection. Dead space, the area of view that will not contribute to the collection volume, should be minimized, as it reduces spatial resolution (Payton & Bartlett, 2007). Insufficient consideration for environmental concerns may negatively affect the quality of the data.

Additionally, static and dynamic calibration should occur before every collection session to ensure accuracy (Payton & Bartlett, 2007). Calibration of optical motion capture systems is time intensive and requires significant knowledge of the equipment. Each time the system is calibrated certain precautions must be taken including adjusting camera sensitivity and minimizing stray reflections.

2.1.2.4 Soft Tissue Artifact

Skin movement artifact can greatly influence the validity and reliability of kinematic data (Cappozzo, 1991; Cappozzo et al., 1996; Leardini et al., 2005; Mündermann, Corazza, & Andriacchi, 2006; Robertson, Caldwell, Hamill, Kamen, & Whittlesey, 2006). Markers placed on the body are meant to estimate the relative movements of the skeletal segments. These segments are assumed to be rigid, meaning that their length is

assumed constant (Robertson et al., 2006). However, marker movement may occur due to movement of the surface to which it is affixed, such as clothing or skin, this is called marker movement artifact. These errors have the potential to be higher in older populations who tend to have less elastic skin and more body fat. Movement artifact in marker-based systems has been examined by comparing data obtained from markers to more invasive techniques that better estimate anatomical segments such as a bone pins, external fixators, percutaneous skeletal trackers, and Roentgen photogrammetry (Leardini et al., 2005). Although more accurate, invasive techniques subject participants and patients to unnecessary harm and are therefore rarely used.

2.1.3 Summary of Marker-Based Motion Capture Limitations

Marker-based motion capture systems are currently considered the gold-standard for capturing 3D human movement. Despite the drawbacks, marker-based systems are frequently used due to their accuracy, reliability, and ability to capture a variety of complex movements in research, clinical, and entertainment domains. However, marker-based systems are often expensive, which may restrict access to institutions with large budgets. In addition to cost, marker based systems are not very portable, which restricts their use to laboratory or clinical settings. It is also important that precautions are taken to minimize systematic, environmental or subject-related error to ensure the integrity of the data restricting the use of these systems to highly trained, usually PhD-level investigators (Cappozzo et al., 1996; Della Croce et al., 2005; Leardini et al., 2005). Finally, both passive and active marker-based systems may restrict natural movements, introduce marker artifact and placement errors, require extensive amounts of time to attach markers,

and restrict the type of movement which can be examined due to the need for sufficient collection space (Cappozzo, 1991; Cappozzo et al., 1996; Della Croce et al., 2005; Leardini et al., 2005; Mündermann, Corazza, & Andriacchi, 2006; Mündermann, Corazza, Chaudhari, et al., 2006). Some of these issues can be addressed with recent developments in motion capture by utilizing advances in the field of computer vision.

2.2 Markerless, Optical Motion Capture

Recently, markerless computer vision techniques have been used to capture and recognize human movement (Moeslund & Granum, 2001; Moeslund et al., 2006; Poppe, 2007; Wang et al., 2003). Markerless motion capture technology is applicable to a large variety of domains which can be broadly categorized into surveillance, control-based interfaces, and motion-based analysis (Moeslund & Granum, 2001; Moeslund et al., 2006). Surveillance-based applications monitor the environment with the explicit purpose of interpreting human behavior. For example, such a system may be implemented in a senior dwelling community to detect and notify medical personnel when some negative event such as a fall occurs (Garrido, Penichet, Lozano, & Valls, 2013; Rantz et al., 2012; Stone & Skubic, 2012). Control interfaces interpret human movement for the purpose of interacting with gaming devices, virtual reality environments, or applications that involve gesturing as a means of issuing commands. Interactive rehabilitative programs that provide feedback to patients about postural orientation are becoming more common. These applications aim to create more autonomous training programs for both aging and neurologically impaired patients (Chang, Chen, & Chuang, 2011; Chang, Chen, & Huang, 2011; Chen, Chiang, Liu, & Chang, 2012; Garcia, Navarro, Schoene, Smith, & Pisan,

2012; Lange et al., 2011; Lange et al., 2012; Venugopalan, Cheng, Stokes, & Wang, 2013). Finally, motion analysis applications can be used to assess movement kinematics and as diagnostic tools (Clark, Pua, Bryant, & Hunt, 2013; Stone, Butler, McRuer, Gray, Marks, & Skubic, 2013). Markerless based systems work by estimating human pose. Assessing multiple poses over time can then be used to assess dynamic movements. Human pose estimation and relevant computer vision techniques will be discussed in the following sections.

2.2.1 General Model for Human Pose Estimation

In biomechanical applications, it is necessary to create a 3D kinematic skeletal model to perform movement analysis. This is known as pose estimation. Human pose estimation is a complex process due to the high variability in body sizes and appearances (Poppe, 2007). To add to the morphological complexity, the vantage point of the camera(s) or the collection environment may impact the estimation since a single pose can have many different observations and different poses may appear to be the same (Poppe, 2007). The details of pose estimation vary based on the application of the system but a general model was described by Moeslund & Granum (2001). The basic model for attaining kinematic data from markerless motion capture systems is as follows: model initialization, tracking, pose estimation, and pose recognition (Moeslund & Granum, 2001). Pose recognition is not necessarily required for attaining kinematics, however it is used to interpret human movement in control or feedback driven applications. Each area will be discussed separately as a discrete process but in some cases these processes may occur in parallel.

2.2.1.1 Model and Scene Initialization

When markers are removed from the motion capture process, the system has no identifiers to determine the individual's height, shape, or segment lengths. Initialization of the system may occur by providing a human model, a representation of the environment, or both. Model initialization creates an anthropometric model of the individual using either an initial pose or an *a priori* model which specifies fixed segment lengths, a known number of joints, and assigns a known number of degrees of freedom to each joint (Moeslund & Granum, 2001; Moeslund et al., 2006; Poppe, 2007). These models may also impose restrictions on joint angle motion, only allowing those possible for human movement to reduce the number of estimable postures (Moeslund & Granum, 2001). Recently, there has been an increased interest in utilizing kinematic data from marker-based systems to train and provide movement constraints for markerless systems (Moeslund et al., 2006). The purpose of training the system is to reduce the number of possible postures which simplifies the pose estimation process.

Scene initialization refers to a calibration of the camera(s) to the collection volume (Moeslund & Granum, 2001; Poppe, 2007). If the parameters of the cameras are unknown, it may be necessary to calibrate to ensure that the scene is correctly represented. Often times the purpose of scene initialization is to allow the system to observe the environment without the human present. This reduces the complexity of distinguishing the human from objects in the environment. Therefore, if the environment is dynamically changing scene initialization will not be as useful for reducing the complexities of pose estimation. Model and scene initialization may generally be considered a calibration process similar to those required by marker-based systems.

2.2.1.2 Tracking

Once the environment and human model have been initialized, human tracking can begin. Tracking is a means of segmenting and separating an individual from the environment in preparation for pose estimation or recognition (Moeslund & Granum, 2001; Wang et al., 2003). There is some overlap between tracking and pose estimation; however, they will be discussed as separate processes. There are three common practices seen in the tracking phase: figure-ground separation, creating a representation of the human, and frame-to-frame tracking definitions (Moeslund & Granum, 2001).

Figure-ground segmentation separates the human from the background using temporal and spatial techniques. Temporal segmentation utilizes subtraction or flow methods to separate a moving human from the scene. Subtraction methods take the current image and subtract prior images or use background substitution where an empty background image is subtracted from the current image (Moeslund & Granum, 2001; Wang et al., 2003). Flow refers to motion that occurs between images by identifying points, edges, or blob representations of anatomical segments (Moeslund & Granum, 2001; Wang et al., 2003). It is assumed that the environment is static in order to detect the human in this fashion.

Spatial segmentation relies on color thresholding, where the human is considered to be a different color than the background. Statistical methods are then used to characterize groups of pixels such as colors and edges (Moeslund & Granum, 2001). Statistical approaches are more robust compared to background subtraction methods, but increase the complexity of figure-ground segmentation because they are dynamically updated (Moeslund & Granum, 2001; Wang et al., 2003).

The representation process attempts to describe human segments by object or image-based representations. Object-based representations typically utilize figure-ground segmentation whereas image-based representations are often created directly from the image (Moeslund & Granum, 2001). Object representations attempt to distinguish between human segment motion and environmental object motion (Wang et al., 2003). Boxes, silhouettes, and blobs are considered to be object-based representations. Box representations fit boundary boxes to pixels belonging to the human from the figure-ground segmentation. Each box outlines a corresponding segment. Blob representations represent the entire person or their segments as less rigid shapes similar to box representation. These types of representation are more commonly intermediate steps and may be transformed before pose estimation (Moeslund & Granum, 2001).

Silhouette representations use figure-ground segmentation methods to find the edges of the individual, analogous to tracking the shadow of the individual (Poppe, 2007). Voxel representations use multiple cameras to capture the silhouette from different angles, creating a volumetric model of the individual that can be used to heal self-occlusion and depth calculation issues (Poppe, 2007). Additionally, contours or silhouette outlines may be used in the tracking process (Moeslund & Granum, 2001; Poppe, 2007; Wang et al., 2003). Silhouette representations may be used to track a human form over time or may be reprocessed in pose estimation (Moeslund & Granum, 2001). Recently, surface-based methods have been introduced which utilize a mesh consisting of polygons that are deformed over the human form (Poppe, 2007). Determining which object-based representation to use depends on the purpose of the application, in particular the accuracy required for joint location and segment orientations.

Image-based representations use image pixels and may be utilized independently or combined with one of the representations above (Moeslund & Granum, 2001). Often times the images are transformed into a non-Cartesian space using Fourier, principle component analysis (PCA), discrete cosine transforms, or wavelets (Moeslund & Granum, 2001). Other common representations use edge detection or features that depend on the area occupied by the human. These techniques rely on the human being a different color than the background (Moeslund & Granum, 2001; Poppe, 2007). Most of these representation examples attempt to identify segments for the purpose of applying a specific pose estimation algorithm; some, however, may only identify the human as a single object. The type of representation or combination of representations chosen is dependent upon the type of pose estimation algorithm.

Last, tracking of the representations over time requires identification a particular object, for instance a blob representing a segment, in consecutive frames. This process can be increasingly difficult if the environment is complex or contains objects in motion (Moeslund & Granum, 2001). Additionally, complexities in the human form may add to the difficulties of tracking, for example, clothing deformation. Alternatively, if an environment contains objects that closely resemble human segments, such as the legs of chairs or tables, it becomes difficult to track true body segments over time. Tracking across multiple images therefore depends on the ability of the algorithms to separate the human from the environment. It is simpler to utilize a fixed, static environment and well-defined human forms to reduce the complexity of identifying and tracking a human in the environment (Poppe, 2007). Another option is to estimate pose for each captured image. This removes the need for tracking a segment across images; however, estimating pose

for every frame can be computationally expensive if improperly implemented (Shotton et al., 2013). This section has touched upon a handful of the methods used to address the problem of separating, tracking, and representing a human in single or consecutive images. These methods can be combined and serve the purpose of preparation for pose estimation or pose recognition (Moeslund & Granum, 2001; Poppe, 2007).

2.2.1.3 Pose Estimation

Tracking captures the movements of particular segments of the individual from scene to scene. Pose estimation is distinct from tracking because one or more tracked representations are combined to determine the orientation of individual segments or the whole body. Pose estimation may occur during tracking or alone as a post-processing step where the positions and orientations of joints and limbs are estimated (Moeslund & Granum, 2001). There are a variety of methods to estimate joints and limbs, many of which require an *a priori* model as a calibration. Pose estimation can occur with or without models. Systems that use models may opt to use *a priori* modeling or create temporary models from lookup tables (Moeslund & Granum, 2001). Model-free pose estimation can build intermediate models directly from the image representations; they do not depend on information from a calibration pose or pre-defined human model. The issue with model-free methods is that extensive training is required to teach the system how to recognize poses and joint locations. Indirect model estimations use a table of human characteristics and attempt to match the individual to anthropometric data within the model. A best fit representation of the individual is thus chosen based on parameters such as height, joint positions, or aspect ratios of the limbs. Finally, direct model

estimations take an *a priori* model which is updated based on observed motion of the segment representations (Moeslund & Granum, 2001).

Pose estimation can be considered a top-down or bottom-up process. Top-down processing requires an initial pose and matches a skeleton to the representation created during the tracking process. Top-down estimations typically suffer from self-occlusion errors, where one segment hides another from view, and inaccuracies in the location of one joint may propagate to other joints (Poppe, 2007). Bottom-up processing identifies segment orientations and then combines them to create a full skeletal model. Bottom-up estimations have difficulties when there are limb-like objects in an image (Poppe, 2007). Alternatively, the two methods can be combined to reduce errors.

The type of model estimation chosen depends on system requirements. If there is only one person being tracked, it is beneficial to provide an *a priori* model. However, for a surveillance application, it may not be feasible to have predefined models available and as such indirect or model free approaches are best. Since pose estimation must be accurate in posture or gait assessments, the person being tracked by a system must adhere to the system's operational assumptions. If that is not possible, robust implementation that minimizes noise as well as environmental and human complexities is required.

2.2.1.4 Pose Recognition

Pose recognition occurs by taking the pose estimation, or human kinematic skeleton, and classifying it based on segment orientation (Moeslund & Granum, 2001). It may be unnecessary for certain applications to know the precise location of joints in order to classify the pose; such as in surveillance applications. However, for the purpose of

capturing and analyzing human movement, it is common to estimate pose based from skeletal models (Moeslund & Granum, 2001). Static recognition uses only a single frame while dynamic recognition, also known as gesture recognition, requires recognition of a sequence of poses (Moeslund & Granum, 2001).

2.2.2 Limitations of Markerless Motion Capture

Moeslund et al. (2001) described three critical performance parameters of all markerless motion capture systems: robustness, accuracy, and processing speed. The value of each parameter varies based on the application. Robustness relies on minimizing the operational assumptions of the system (environment-, subject-, or movement-related) (Moeslund & Granum, 2001). Accuracy refers to minimizing error between the actual and captured movements. Processing speed refers to how often the system can carry out a single pose estimation and may also be related to whether the system utilizes real-time or offline processing (Moeslund & Granum, 2001).

The operational assumptions of a system are inversely related to the complexity of the system (Moeslund & Granum, 2001). Therefore, the optimal markerless system has no operational assumptions but is highly complex. Incrementally simpler systems have more assumptions. When using a particular motion capture system, it is necessary to understand the operational assumptions because violations can lead to inaccuracies in pose estimation. Additionally, large latencies in pose estimation (i.e. low sampling frequency) may lead to difficulties in motion analysis due to poor temporal resolution. Most markerless systems operate under a number of the assumptions. Despite the fact that markerless systems may have slightly fewer operational assumptions than their marker-based counterparts they do not currently provide the same quantitative accuracy

(Mündermann, Corazza, Chaudhari, et al., 2006). In order to improve the current methods for analyzing human movement low-cost, portable, widely available, accurate, and robust kinematic measurement tools are necessary. This technology is currently unavailable, but a potential solution to this issue may be found through re-purposing entertainment technology.

2.3 Repurposed Technology

Recent innovations in interactive gaming have created technologies with the potential to address many of the issues with commercial vision-based motion capture systems. These technologies were created for entertainment and gaming but are now being tested for research and occupational purposes. For example, the Nintendo Wii Balance Board (WBB) (Nintendo, Kyoto, Japan) functions similarly to a laboratory-grade force plate. Specifically, the WBB can determine center-of-pressure (COP) movements to assess standing balance (Clark et al., 2010). The WBB has been used in a variety of applications including determination of weight-bearing asymmetry (Clark, McGough, & Paterson, 2011; Foo, Paterson, Williams, & Clark, 2013; McGough, Paterson, Bradshaw, Bryant, & Clark, 2012); balance training in older adults (Koslucher et al., 2012), children (Mombarg, Jelsma, & Hartman, 2013), and individuals with Multiple Sclerosis (Prosperini, Fortuna, Gianni, Leonardi, Marchetti, & Pozzilli, 2013); assessment of postural control (Holmes, Jenkins, Johnson, Hunt, & Clark, 2013; Howells et al., 2012; Young, Ferguson, Brault, & Craig, 2011); and COM estimation (González, Hayashibe, & Fraisse, 2012a). The WBB provides an excellent example of how gaming technology can be repurposed for biomechanical applications. However, the WBB is no longer being

produced by Nintendo and thus access to this technology is restricted. In addition, the WBB is not outfitted with a customizable software package to aid in the development of applications desired by clinicians and researchers.

Another device recently investigated for human movement assessment is the Microsoft Kinect. The Kinect was originally designed as a video-game peripheral for the Xbox 360 console. It functions as a hands-free controller and facilitates user interaction through gesture and speech commands. The interface created by the Kinect can be used to provide real-time feedback about body orientation for interactive gaming. The Kinect could potentially alleviate issues observed in commercial vision-based motion capture systems since it is a portable, inexpensive, markerless, depth camera. The current body of research surrounding the Kinect will be discussed in the following sections.

2.4 Microsoft Kinect

The Kinect is a hands-free gaming controller, originally created to rival the Wii and PlayStation Move (Menna, Remondino, Battisti, & Nocerino, 2011). The device captures human gestures and speech to create a more natural interaction in real-time (Microsoft, 2013a). The Kinect features an RGB camera which can capture color images and video. It also features an IR emitter and IR depth sensor. The IR emitter projects a pattern onto the environment and the reflected light is picked up by the IR camera. A multi-array microphone allows audio to be captured from multiple directions to distinguish between sources of sound, such as game-related sound or speech commands. A three-axis accelerometer allows the device to determine its orientation with respect to gravity. The Kinect's coordinate system originates in the center of the IR camera. The x-

and y-axis correspond to the horizontal and vertical coordinates of the image while the z-axis provides measurement of depth. The Kinect features a motorized base which can tilt the device $\pm 27^\circ$ in the vertical plane (Microsoft, 2013a). The Kinect's field of view is 57° in the horizontal plane and 43° in the vertical plane. The following section will discuss the functions and characteristics of the Kinect.

2.4.1 How the Kinect works

Many markerless approaches to motion capture utilize standard video cameras; however, the introduction of depth cameras has improved the process of estimating human pose (Shotton et al., 2013). Depth cameras are color and texture invariant, thereby reducing difficulties associated with background substitution. The approach used to estimate pose by the Kinect is discussed in the work of Shotton, Sharp, Kipman, Fitzgibbon, Finocchio, Blake, Cook, & Moore (2013).

The pose estimation technique used by the Kinect takes a single depth image, infers body position, proposes 3D joint positions and orientations, and then outputs the proposed kinematic skeleton (Shotton et al., 2013). The novelty of this approach is that it uses only the current image to infer human pose until the skeleton is fit, at which point information from previous images is incorporated. This allows for better recovery from tracking errors (Shotton et al., 2013). Markerless systems that estimate human pose from a single image reduce the opportunity for error propagation from previous poses, but require extensive training to accomplish this task. Inferring the position and orientation of the individual segments to build the pose estimate provides a solution to the propagation

of joint misplacement error. This technique is robust to a variety of body shapes and sizes and can have the potential to handle self-occlusion (Shotton et al., 2013).

The Kinect estimates pose in a two-stage process by computing a depth map and inferring body position from an intermediate parts representation (MacCormick, 2011; Shotton et al., 2013). To compute depth information, the Kinect projects a speckle pattern into the environment and determines depth by observing deformations of the pattern, this technique is termed structured light (MacCormick, 2011). Structured light is combined with two other common computer vision techniques, depth from focus and depth from stereo (MacCormick, 2011). Depth from focus dictates that objects further from the camera will be more blurry because they are represented by fewer pixels. Depth from stereo uses parallax, which is a change in the apparent position of an object due to a change in viewing perspective. The IR emitter is a fixed distance from the IR camera; therefore, shifts in the speckle pattern can be interpreted due to perspective differences. This allows the Kinect to determine depth by assessing the changes in the projected speckle pattern due to deformation by objects in the environment. Once the depth map has been computed, the body segments can be inferred.

In the second stage of pose estimation, an intermediate representation of body segments is determined by transforming a 3D surface model of the human into 31 distinct segments (Shotton, et al., 2013). Joint location proposals are created and assigned confidence weighted values based on the orientation of the inferred body segments. In order to ensure robust pose estimation the Kinect was trained with over 500,000 frames of human movement captured by commercial motion capture systems. The motion capture data facilitates the process of recognizing each segment of the body given the

kinematic skeleton. Additionally, to ensure robustness to morphological variability the Kinect was trained with synthetic models of 15 base characters, both male and female, children and adults (Shotton et al., 2013). Synthetic model parameters such as pose, rotation and translation of the individual, hair and clothing, weight and height, camera position and orientation, and camera noise were altered to ensure that the Kinect could handle a wide variety of potential users and environmental contexts (Shotton et al., 2013). The computer-generated models help to reduce color and texture variability of the individual; however, complexities associated with clothing deformation still exist. Thus, the Kinect should be able to accurately estimate joint location if tighter clothing is worn. The Kinect uses powerful hardware to collect movement but the hardware does have limitations. The camera produces a 640x480 array of depth values (i.e. z coordinate) at approximately 30 frames per second along with the x and y values that can be described by pixel location in the image. The Kinect may lose pixel information when the IR pattern is projected onto surfaces that absorb or reflect light away from the IR camera. It has also been observed that ambient sunlight prevents its ability to sense depth, therefore rendering the device incapable for outdoor use. The IR emitter has a limited range, requiring relevant movement to occur a maximum of 4 meters from the camera. There is also random depth error, known as quantization error, which is associated with increasing distance from the camera. This error can be as large as a few centimeters. Last, the edges of objects are sometimes vague and may switch between foreground and background (Shotton et al., 2013).

2.4.2 Xbox vs. Microsoft Kinect

The first generation of the Kinect featured two releases: the *Kinect for Xbox 360* and the *Kinect for Windows*. The *Kinect for Xbox 360* was introduced for gaming as a natural user interface with gesture and speech commands. The *Kinect for Windows* was released shortly afterward to allow software developers to create custom applications driven by natural interaction (Macknoja, Chavez-Aragon, Payeur, & Laganier, 2012).

One of the differences between the devices is software implementation. The *Kinect for Windows* was specifically designed for research and thus Microsoft released the software developer's kit (SDK) to interface with the device (Microsoft, 2013a). The Microsoft SDK provides developers with the tools needed to access the various features of the device (i.e. depth stream, RGB camera, microphones). The *Kinect for Xbox 360* does allow custom applications to be built using OpenNI SDK, a framework designed for 3D sensing devices (OpenNI, 2013).

There are two operational modes in the *Kinect for Windows*, default and near mode. In default mode the *Kinect for Windows* is functionally identical to the *Kinect for Xbox 360* between a range of 80 cm and 800 cm, but near mode allows for data to be collected as close as 40 cm (Microsoft, 2013b). The practical limits for obtaining accurate depth data suggest that users be between 80 cm and 400 cm from the camera in default mode. In near mode, the practical limits span from 40 cm to 300 cm (Microsoft, 2013b). Choosing an implementation is dependent on the range necessary for the application. The OpenNI SDK allows access to the default range of depth data.

Skeletal tracking is a feature offered by both software implementations. Differences arise in the number of joints available in each SDK. The Microsoft SDK joint

skeleton has twenty joints compared to OpenNI SDK which has only fifteen. The Microsoft skeleton includes wrists, ankles, and a hip center marker not available in the OpenNI skeleton. Differences in skeleton configurations can be seen in Figures 2.5.

Rather than using the built-in skeletal tracking, some studies have utilized the depth map to calculate 3D kinematics of interest (Stone & Skubic, 2011a, 2011b). These studies apply computer vision techniques such as background substitution to separate the human from the environment. Unfortunately, these techniques add additional complexity to human movement analysis. In addition, they are less accessible for clinicians or others without technical expertise. However, these methods can provide accurate and reliable spatiotemporal parameters of gait from the Kinect (Stone & Skubic, 2011a, 2011b, 2012, 2013).

Both software implementations share operational assumptions that may impact accuracy if violated. In both implementations, it is assumed the individual is facing the device. If the Kinect is capturing images from the sagittal plane of view it is more difficult to estimate pose (Microsoft, 2013a). The Kinect also requires (similar to commercial motion capture devices) that self- and environmental occlusion be minimized. OpenNI SDK employs NITE middleware developed by PrimeSense, the company that designed the hardware design and chip used in the Kinect. NITE has cited known issues with skeletal tracking that include less stable arm tracking near the torso, unstable leg tracking, issues tracking fast movements, and variable segment lengths from frame to frame that can negatively impact joint calculations (PrimeSense, 2013). It is unclear whether these issues exist within the Microsoft SDK. In addition, unlike the Microsoft SDK, OpenNI requires the individual to perform a calibration pose before tracking.

A performance evaluation of the two versions of the Kinect was performed by Macknoja et al. (2012) using the default depth ranges. The Microsoft SDK was used for Kinect for Windows while the Kinect for Xbox 360 used OpenNI. This study investigated depth performance and sensitivity to reflective surfaces. Depth performance evaluation showed that random depth error, also known as quantization error, reaches approximately 2 cm at distances of approximately 2.5 m from the camera which suggests that collecting human movement with the sensors should occur between approximately 0.5 m and 2.5 m. Both devices performed similarly in this domain and quantization error was seen to increase quadratically with distance from the device (Macknoja et al., 2012). Additionally, objects with poor reflective properties or dark colors substantially degraded the field of view as their distance from the Kinect increases (Macknoja et al., 2012). In other words the edges of the image frame were more prone to error. To address this issue the Kinect should be placed parallel to dark or non-reflective objects to reduce depth error. Depth errors may propagate to skeletal tracking if the individual is far from the Kinect or is wearing non-reflective clothing. Both devices were equally susceptible to depth errors from dark colored or non-reflective objects and increasing distance from the camera.

Last, more than one Kinect can be used with the Microsoft SDK; however, there are certain precautions that must be taken. Microsoft suggests that only one Kinect sensor be used in a collection area. Interference may occur due to overlap of infrared light patterns which may create difficulties interpreting the pattern distortions used for depth calculation (Microsoft, 2013b).

In conclusion, there are different depth ranges and skeleton tracking features permitted depending on the software implementation. Both models of Kinect perform similarly in depth performance and sensitivity to color and reflective objects. The Microsoft SDK has the advantage of near and seated mode options, additional skeletal joints, and the ability to utilize more than one sensor. In either implementation it is important to be aware of the operational assumptions before collecting kinematic data.

2.4.3 Current Research

The introduction of the Microsoft SDK in 2011 led to substantial interest in researching the capabilities and utility of the Kinect in a variety of domains. First, this review aims to cover spatial capabilities and the skeletal tracking accuracy. Second, biomechanical applications of the Kinect will be discussed. These include posture and gait assessments and workplace ergonomics. These applications all focus on the biomechanical assessment of human movement. Numerous studies have proposed rehabilitative programs for improving range of motion and balance utilizing gesture driven games and visual feedback of patient movement (Chang et al., 2012; Chang, Chen, & Chuang, 2011; Chang, Chen, & Huang, 2011; Lange et al., 2011; Lange et al., 2012; Luna-Oliva et al., 2013; Metcalf et al., 2013; Venugopalan et al., 2013). Rehabilitative applications will not be discussed as most do not directly investigate the Kinect's capabilities for accurately and reliably capturing human movement. It is important to note that, unless optimization procedures are implemented, these applications are likely to share the benefits and suffer from any of the limitations discussed below.

2.4.3.1 Spatial Capabilities

Much of the initial research surrounding the Kinect has focused on determining the spatial accuracy and depth resolution which determine acceptable ranges for data acquisition. Microsoft had released physical and practical field of view limits but verification of these limits was necessary. A variety of applications including surveillance and biomechanical applications rely on the accurate depth representation of the environment.

Menna et al. (2011) tested the precision of the depth values by placing the Kinect orthogonal to a flat, white wall. The distance between the wall and Kinect was manipulated to compare depth value standard deviation at 750 mm and 2750 mm. The standard deviation was expected to be 1.6 mm and 22 mm respectively given the camera specifications. However, standard deviation was found to be 40 mm and 300 mm at the edges of the image. As a solution, the outer 20% of the image was removed which brought the standard deviation within an acceptable range (Menna et al., 2011). It seems that the Kinect behaves as expected within the center of its field of view. Thus, researchers should take caution when attempting to capture movements that occurring near the edges of the device's depth range.

Khoshelham (2011) performed a similar manipulation using a planar surface at intervals between 0.5 m and 5.0 m from the device to determine quantization error at varying distances. The quantization error increased quadratically with increasing distance from the Kinect similar to Menna et al. (2011) (Khoshelham, 2011; Khoshelham & Elberink, 2012). At the maximum distance recommended for skeletal tracking, 3.5 m from the camera, the error was found to be approximately 2.5 cm. It was also found that

the depth resolution decreased with increasing distance, up to 7 cm at 5 m from the camera (Khoshelham, 2011; Khoshelham & Elberink, 2012). Therefore, it is important to remain within the device's functional boundaries between approximately 1m to 3 m in order to minimize quantization error and avoid losing spatial resolution as the individual approaches distances greater than or equal to 3.5 m from the device.

Along with quantization error, sometimes termed axial noise, the Kinect also exhibits lateral noise in the directions perpendicular to the z-axis (i.e. in the x- and y-axes) (Nguyen, Izadi, & Lovell, 2012). Lateral noise was first described by (Menna et al., 2011). Nguyen, Izadi, & Lovell (2012) also utilized a planar surface placed incrementally farther from the Kinect. Additionally the surface was rotated around the y-axis at each distance. Lateral noise is extracted from the edges of the depth map whereas the axial noise is determined from differences in the Kinect depth values and the plane's distance from the device (Nguyen et al., 2012). The Microsoft SDK was used in near mode and the distance from the plane to the device was increased from 0.5 m to 2.75 m. Similar to Khoshelham (2011) & Menna et al. (2011), quantization error, or axial noise, was found to increase quadratically with increasing distance from the Kinect. Additionally, it was found that lateral noise increased linearly with increasing distance from the Kinect. In regard to the angle of the plane to the camera, it was found that lateral noise increased slightly with inflection angle until about 70° while axial noise remained approximately constant until about 60° but increased rapidly from 60° to 90° (Nguyen et al., 2012). Thus, maintaining a location centered in the Kinect's view within the 1-3 m depth range is critical for minimizing noise into depth measurements. Preferably individuals should

sit or stand facing the camera, although slight deviations in inflection angle should not introduce large errors in depth measurement.

Dutta (2012) conducted an investigation to determine range, field of view, and accuracy of the *Kinect for Xbox 360* as a preliminary assessment of the device's capabilities for workplace evaluations. RMS error and standard deviation (in parentheses) was calculated in the x, y, and z directions. RMS errors of 1.69 cm (2.99 cm), 3.48 cm (7.65 cm), and 1.41 cm (2.5 cm) were found in the x, y, and z directions respectively (Dutta, 2012). The Kinect displayed large quantization errors at large z distances, similar to previous findings (Khoshelham, 2011; Khoshelham & Elberink, 2012; Menna et al. 2011, Nguyen et al., 2012). However, accuracy increased when the outer 25 pixels were removed from the depth image. This solution is similar to Menna et al. (2011), where 20% of outer pixels were removed to reduce lateral noise. Field of view findings suggested that the field of view of the Kinect's depth camera was similar to the ranges provided by Microsoft; 57° and 43° in the horizontal and vertical planes, respectively. When the outer 25 pixels were removed to increase accuracy, field of view ranges were reduced to 54° in the horizontal plane and 39.1° in the vertical plane when the outer pixels of the image were clipped (Dutta, 2012). Additionally, object detection assessment was performed and revealed that dark colors and shiny surfaces did not reflect light back to the Kinect, creating errors in spatial accuracy. The Kinect also appeared to have trouble detecting edges.

Overall, results suggest the Kinect's spatial accuracy is an order of magnitude lower than commercial motion capture systems such as Vicon (Vicon Motion Systems Ltd, Oxford, UK) (Dutta, 2012; Livingston, Sebastian, Ai, & Decker, 2012). However, skin movement

artifact errors of similar magnitude occur in marker-based systems which suggest the potential for comparable spatial accuracy given carefully controlled environmental conditions (i.e. avoidance of absorptive, reflective surfaces, minimizing inflection angle to objects, and distance to camera). Unfortunately, environmental control is one of the drawbacks associated with commercial motion capture systems. The Kinect does not appear to alleviate this issue. Considerations must also be taken to ensure the individual is between 1 m to 3 m from the camera to minimize axial and lateral noise. Additionally, the individual should remain in the center of the IR camera's field of view for accurate extraction of kinematic parameters. As Dutta (2012) pointed out, some biomechanical applications may sacrifice accuracy for portability.

2.4.3.2 Skeletal Tracking

For the Kinect to be a viable alternative to commercial motion capture, it needs to be capable identifying joint positions with a high degree of accuracy and reliability. This is a complex problem for markerless motion capture system for two reasons. First, system error is introduced from inferring depth, specifically due to the depth noise mentioned above. The spatial capabilities of the device dictate that quantization error increases quadratically and that lateral noise increases linearly with increasing distance from the camera (Dutta, 2012; Khoshelham, 2011; Khoshelham & Elberink, 2012; Menna et al., 2011; Nguyen et al., 2012). Second, error from the pose estimation algorithm can occur when inferring joint positions (Mobini, Behzadipour, & Saadat Foumani, 2013).

Obdrzalek et al. (2012) performed one of the first comparisons of Kinect joint locations to a commercial marker-based motion capture system, Impulse (PhaseSpace

Inc., San Leandro, CA). Two calibrated skeletons were generated from the commercial system's data, using PhaseSpace Recap software and Autodesk MotionBuilder (Autodesk, Inc., San Rafael, CA). Six movements were performed: knee lifts while sitting, hands above head while sitting, quiet sitting, leg swinging laterally, sit-to-stand, and line-tapping with toes (Obdrzalek et al., 2012). Additionally, these movements were recorded in 30-degree increments from the frontal to sagittal plane views of the individual.

Comparisons of the Kinect and PhaseSpace human models assumed fixed differences. The two virtual skeletons have different joint locations, but positional differences should remain constant during movement. The Euclidean distance between the joint locations was used to measure accuracy. First, the results suggest the accuracy of the Kinect decreases as the individual turns away from the camera (i.e. worst when viewing in the sagittal plane) in agreement with the findings of Nguyen et al. (2012). Thus, frontal viewing of the individual is recommended. Second, seated tasks also presented difficulty for the Kinect due to self-occlusion, which created lower limb segment length variability and joint misplacement. Variability of the segment lengths occurs because the pose estimation algorithm updates each frame with no consideration of previous poses. Limb length variability was found as high as 10 cm (Obdrzalek et al., 2012). This was thought to occur primarily due to the fact that lower extremity joint placement is most difficult when the leg is straightened.

One drawback of this investigation was the movement selection. The Kinect has self-occlusion issues and pose estimation difficulty from objects in the environment that resemble human limbs (Shotton et al., 2013). For these reasons, seated leg movements

could negatively impact the accuracy of lower extremity joint position and limb orientation. Despite violating some of the Kinect's operational assumptions this study provides evidence of segment length and self-occlusion issues described by Shotton et al. (2013) and highlights the specific movements that cannot be accurately captured with Kinect.

In another investigation, Clark et al. (2012) examined the skeletal tracking accuracy of the Kinect during lateral and forward reaches, and a single leg-balancing stance. Comparisons were made with a passive marker system, Vicon (Vicon Motion Systems Ltd, Oxford, UK). The outcome measures for the reaching trials were hand and sternum displacements and trunk angle. During single leg standing the ankle, knee, pelvis, sternum, and trunk angle were assessed. The reliability of the Kinect's distal (i.e. hand and ankle) and trunk (i.e. sternum and hip center) joints was generally lower for lateral reaching and single leg stability. The Kinect also displayed both fixed and proportional biases of absolute joint location. A fixed bias represents a constant positional difference between Kinect and VICON, where a proportional bias refers to changes in error related to the magnitude of joint displacement. For example, the sternum marker was found to be different between the two systems, and this difference increased with larger displacements of the trunk (Clark et al., 2012). It is possible to correct for these biases, but this assumes the differences are solely due to errors in the Kinect's joint locations. Soft tissue artifact is prominent in commercial marker-based motion based systems (Della Croce et al., 2005; Leardini et al., 2005). Thus, differences found between the systems do not necessarily indicate that the Kinect skeletal tracking feature is inaccurate. The Kinect's capability to reliably and accurately capture reaching and

standing posture seems promising for postural assessment. Although, the reliability of certain joint locations may limit its ability to accurately assess angular kinematics.

Performance measurements of the Kinect skeletal tracking feature were evaluated by Livingston et al. (2012) using the Microsoft SDK. Latency of the Kinect skeleton was found to be approximately 146 ms when the program was running at 30 Hz with a single user being tracked. However, latency increased up to a maximum of 500 ms as more users were tracked. Latency is a concern in applications where feedback about body orientation is updated in real time. For example, if the Kinect were used to provide rehabilitative feedback and the clinician was in view, latency could increase enough to perturb the individual (Miall & Jackson, 2006). Additionally, the sampling rate is irregular and can drop to ~20 Hz with multiple users in view. It seems that under single user performance conditions the skeletal tracking feature has a high enough temporal resolution and small enough latency to be used in biomechanical applications. When multiple users are tracked by the Kinect the temporal resolution decreases and the latency increases dramatically (Livingston et al., 2012).

Mobini et al. (2013) assessed the skeletal tracking capabilities of the Kinect by investigating joint center displacements of a fabricated wooden model of the upper body with known joint locations. The model was placed perpendicular to the z-axis and was incrementally moved from 0.95 m to 2.5 m from the Kinect. Error was defined as the absolute difference in displacement for each joint center separated into x, y, and z directions. RMS error of the joint displacements was found to be approximately 1.5 cm for the shoulder, elbows, and hand joints (Mobini et al., 2013). This investigation provides evidence that the Kinect can provide accurate tracking of upper extremity joint

displacements. Use of the Microsoft SDK seated mode is likely the best option for applications uniquely interested in the upper extremities. It is unclear how accurate lower extremity and trunk joint displacements are by comparison. One drawback with this paradigm is that tracking accuracy of a planar model may not transfer to human body and clothing morphologies.

Additionally, given that the Kinect estimates pose frame-by-frame, high-speed dynamic movements may result in a loss of accuracy. The 2013 NITE Algorithm 1.5 guide suggests that users avoid fast motion as it may cause tracking failure (PrimeSense, 2013). It is unclear whether the Microsoft SDK has similar issues. Given the lower and variable frame rate, it should be assumed that high-speed, dynamic movements could create tracking difficulties regardless of software implementation.

The skeletal tracking feature accurately quantifies gross movement trends without an anthropometric model (Obdrzalek et al., 2012). However, this may be improved in two ways. First, if the skeleton model can be calibrated to have rigid segments, joint position variability would likely decrease. This would require either changing the pose estimation algorithm which is a complex task, or applying a technique proposed by Weber et al. (2012) that utilizes a damped least-squares algorithm to fix the segment lengths based on *a priori* measurements. Additionally, there are fixed and proportional biases that may need to be addressed depending on the amount of precision needed (Clark et al., 2012). For example, if only gross movement trends are of interest, biases are less troublesome. Latency issues may be of concern for feedback applications or in settings where clinicians may need to provide support to patients during movements. Last, relative comparisons of joint displacement may provide a better criterion for accuracy assessment.

Absolute positional differences are not necessarily reflective of error solely from the Kinect as commercial systems are prone to soft tissue artifact (Leardini et al., 2005).

2.4.3.3 Postural Control

The spatial capabilities and functionality of the skeletal tracking feature limit the applicability to short, close range movement. The following sections discuss the functional capabilities of the device across three domains: postural assessment, gait assessment, and workplace ergonomics. Additionally, these applications highlight other benefits and limitations of the Kinect.

Postural assessments in clinical settings typically include qualitative, timing-based physical functioning or balance tests. These tests are simple and easy to use but do not provide insight into the postural strategies adopted by individuals as part of aging or disease. Three-dimensional kinematic assessment may provide information that better describes postural strategies or adaptations, thereby leading to the creation of more customized treatment strategies.

Obdrzalek et al. (2012) assessed the skeletal tracking accuracy of the Kinect for six movements to assess postural control in elderly individuals. The Kinect displayed high variability in joint location and segment lengths compared to a commercial motion capture system. Some of the movements were performed seated, which could introduce errors due to occlusion or difficulty separating the individual from the environment. However, the Kinect may still be useful in postural assessments that do not require the individual to be seated. Seated posture can be assessed if the Microsoft SDK is used and seated mode is activated, but only the joints of the upper extremities can be extracted.

Clark et al. (2012) assessed the validity and reliability of the Kinect joint positions compared to a Vicon during reaching and standing posture. Both fixed and proportional biases in joint locations were found between systems. For example, during lateral reaching the displacement (standard deviation in parenthesis) of the sternum for the Kinect was 305.2 (50.5) mm while the measured displacement for the Vicon was 290.0 (42.3) mm. Most of the biases reported by Clark et al. (2012) are proportional and the Kinect overestimates the joint displacement. Fixed biases present less of an issue in postural assessment because they are constant and can be accounted for; however, proportional biases in anatomical landmarks present a larger problem because the error changes with displacement magnitude (Clark et al., 2012). The Kinect is unable to calculate internal/external rotation, which limits movement assessment to flexion/extension and adduction/abduction. Difficulty in tracking rotational movements is likely due to the single, frontal perspective of the camera. This study demonstrates the Kinect's potential for assessment of postural control strategies in both reaching and standing posture.

Another assessment by Clark et al. (2013) investigated lateral trunk lean during gait by comparing the orientation of the trunk (defined by the sternum and hip joints) with the global vertical axis. Participants were required to match a 10° lateral trunk lean angle within $\pm 2^\circ$. Direct comparisons of the Kinect angle to a Vicon system resulted in mean errors (standard deviation in parenthesis) of 3.2° (2.2°). Two calibrations were performed to increase accuracy which led to similar lateral lean angles. Thus, the Kinect is capable of accurately measuring lateral lean angle when two calibration procedures are

used. However, the gross movements of the trunk were exaggerated in this study and thus smaller, natural movements of the trunk may not be as easily identified.

In addition to joint positions and angular kinematics, tracking whole-body center of mass (COM) position, velocity, and acceleration with respect to the individual's base of support (BoS), can provide higher-level information about the dynamics of balance during standing and suprapostural tasks (Haddad, Gagnon, Hasson, Van Emmerik, & Hamill, 2006; Haddad, Ryu, Seaman, & Ponto, 2010; Slobounov, Slobounova, & Newell, 1997). Whole-body COM is defined as the point location of the weighted average of the COM of each individual body segment. COM is difficult to calculate because it requires tracking body segments over time and knowing the weight distribution of each segment.

Two techniques have investigated the ability of the Kinect to calculate COM, but both utilize the Kinect and a WBB to determine center of pressure (COP). One technique utilizes static COP measurements to estimate COM using a static equivalent serial chain (SESC). This technique creates a chain of n links determined from the orientation and mass distribution of the body's segments where the end-effector of the chain is the COM (González et al., 2012a; González, Hayashibe, & Fraisse, 2012b). Calibration requires static COP measurements are used to estimate COM which can then be translated to dynamic movements. Another technique proposed by Dutta, Banerjee, & Dutta (2013), more commonly employed using commercial 3D motion capture systems, segments the body and calculates the whole-body COM as the weighted average of the body's segments. One drawback to both methods is the complexity associated with calculating COM. Both OpenNI and Microsoft SDKs offer a COM measure which simplifies

computational complexity. However, no investigation has detailed the COM accuracy from these two software packages.

Recently, Schmitz, Ye, Shapiro, Yang, & Noehren (2014) assessed the ability of the Kinect to measure joint angles. A testing jig meant to represent the thigh and shank was created with a 3 degree of freedom ball and socket joint. Retroreflective markers were affixed to both segments to allow for comparison to a traditional motion capture system. A digital inclinometer was utilized to provide ground truth of flexion-extension and adduction-abduction angles. Internal and external rotation were also assessed but could not be compared to the inclinometer, thus, comparisons were made between systems. Kinect data was captured utilizing the depth map and 2D images of the scene, not the skeletal tracking feature. The Kinect joint angles were within 2° of the inclinometer for flexion-extension and abduction-adduction which suggests the Kinect may be capable of detecting subtle changes in lower extremity joint angles. However, the frontal plane movements, abduction and adduction, between the two systems were significantly different. Test-retest differences were similar suggesting that the reliability of the Kinect was similar to the marker-based system. One obvious limitation of both Schmitz et al. (2014) and Mobini et al. (2013) is the use of artificial models to measure the Kinect's ability to capture joint kinematics. The authors acknowledge that capturing dynamic and more complex movements at varying distances from the camera may alter the results discussed above.

Last, Galna, Barry, Jackson, Mhiripiri, Olivier, & Rochester (2014) assessed the accuracy of the Kinect for capturing dynamic movements in individuals diagnosed with

mild to moderate Parkinson's disease. Joint positions were extracted using the Microsoft SDK and skeletal tracking feature. A 10 camera Vicon motion capture system (Vicon Motion Systems Ltd., Edinburgh, UK) was used as the gold-standard comparison. The Kinect was placed perpendicular to the ground at a height of 1 m and viewed the participants in the frontal plane. The movements of interest included: quiet standing, forward and lateral leaning, stepping forward and laterally, and walking in place. In addition, six movements from the Unified Parkinson's Disease Rating Scale were investigated including: hand clasping, finger tapping, foot tapping and leg agility, sit-to-stand, and hand pronation. Kinect data were not upsampled or normalized and thus the results provide an absolute measurement of the device's accuracy. Biases between the systems were address using two sided t-tests. Pearson's r correlations were used to assess overall agreement between systems and intra-class correlation ($ICC_{2,1}$) was used to measure absolute accuracy with 95% limits of agreement. Nine individuals diagnosed with Parkinson's disease were included as well as 10 controls that were not age or gender matched because no group comparison was made.

The results suggest that the Kinect is capable of accurately capturing the timing of these movements with no significant biases between the systems and Pearson's r and ICC values all greater than 0.9. As for spatial accuracy, no errors were found to be related to the magnitude of the movement as reported by Clark et al. (2012), except for the sit-to-stand task. Previous research by Obdrzalek et al. (2012) had suggested that errors found in sit-to-stand tasks could occur due to issues differentiating the individual from the chair as opposed to the Kinect's ability to accurately capture the sitting-to-standing behavior. For all other movements the Kinect tended to either under- or overestimate the range of

motion. The results of this study are promising as they open the door for opportunities to utilize devices like the Kinect to alter rehabilitation and training paradigms for individuals with Parkinson's disease. Overall, it appears that Kinect is capable of accurately and reliably capturing angular kinematics as well as a variety of other movement patterns, more testing on humans and complex, dynamic movements needs to be performed to determine its' performance capabilities.

2.4.3.4 Gait Assessment

Quantitative gait analysis is rarely used for diagnostic purposes in clinical settings (Simon, 2004). Rather, gait analysis is typically requested to supplement clinical assessments and help prescribe treatment strategies for patients. Laboratory gait analysis may not accurately portray gait patterns exhibited in everyday life. In order to better monitor and understand gait in everyday contexts it may be necessary to employ devices that allow for 3D kinematic analysis of gait outside the laboratory. Recent research has examined the Kinect's ability to provide accurate and reliably spatiotemporal parameters of gait which could lead to more widespread use in clinical settings.

The first investigation of the Kinect's utility for gait analysis was performed by Stone and Skubic (Stone & Skubic, 2011a; Stone & Skubic, 2011b). Comparisons were made between a multi-camera Vicon system (Vicon Motion Systems Ltd., Edinburgh, UK), a multiple web-camera setup, and the *Kinect for Xbox360*. Additionally, the orientation of the Kinect with respect to the walkway was manipulated. One perspective was parallel to the direction of progression while the other was oriented at approximately 60 degrees to the walkway. The skeletal tracking feature was not used to extract joint

kinematics; instead, analysis of the depth map was utilized. The use of the depth mapping features allows the individual to walk at a constant speed across the walkway without the spatial concerns associated with using the skeletal tracking feature such as distance from the Kinect. Three spatiotemporal gait parameters were investigated including gait speed, stride length, and time. Variability of the three parameters was also compared. The results demonstrated that the Kinect can accurately acquire spatiotemporal gait parameters and variability which highlight the potential for gait analysis or in-home monitoring (Stone & Skubic, 2011a; Stone & Skubic, 2011b). The Kinect with a view parallel to the plane of progression performed better in both the absolute and variability measures.

One of the limitations of utilizing the depth map for gait analysis is the additional complexity. When opting for extracting movement parameters via the depth map, it is necessary to create custom algorithms to identify and quantify movements of interest. The skeletal tracking feature reduces this complexity by returning the joint positions which allows researchers to perform the analyses of their choice. This is not to say that movements quantified using the depth map are less accurate, rather the creation and implementation of the desired software solution are far more time consuming and should be considered for each investigation individually.

Clark, Bower, Mentiplay, Paterson, and Pua (2013) recently assessed the validity and reliability of spatiotemporal parameters of gait using the skeletal tracking feature compared to a Vicon motion capture system (Vicon Motion Systems Ltd., Edinburgh, UK). The protocol used a 2.5 m walkway ranging 1 m to 3.5 m from the device. This procedure required gait initiation at 3.5 m and abrupt halting of gait near the 1 m mark. This allowed for the collection of a single, uninterrupted gait cycle. Additionally, the foot

initiating gait was alternated each trial. Gait parameters included gait speed, step and stride length, step and stride time, and foot swing velocity. The results suggest that step/stride length and gait speed are most valid. Step and stride times were influenced by the Kinect's inability to detect anatomical landmarks at foot contact and toe-off (Clark, Bower, et al., 2013). Thus, the skeletal tracking feature does show promise for capturing these measures which dramatically reduces the complexity associated with previous investigations of gait (Stone & Skubic, 2011a). However, due to the difficulty identifying gait events and the Kinect's relatively small collection volume, its potential in clinical settings is currently limited. In order to extract multiple gait cycles, multiple trials while alternating the foot moving first in gait initiation are required.

2.4.3.5 Workplace Ergonomics

Commercial marker-based and markerless motion capture systems cannot easily monitor human movement or employee posture due to high cost and spatial requirements (Best & Begg, 2006). As computers become universal in the workplace, it is increasingly important to monitor employee's seated posture to avoid musculoskeletal problems. Recently, an application was developed to monitor seated posture by tracking head position and rotations (Uribe-Quevedo, Perez-Gutierrez, & Guerrero-Rincon, 2013). When posture fell outside of an accepted range, an alert was issued informing the individual of their deviation. This type of application could be used to monitor proper posture during the workday and providing feedback to improve behavior.

Another area of interest in workplace ergonomics is assembly operations in industrial settings. Haggag, Hossny, Nahavandi, and Creighton (2013) used Rapid Upper Limb

Assessment (RULA) to determine when a participant's posture and joint angles could lead to injury (Haggag et al., 2013). An important issue encountered by this implementation was difficulty in calculating the joint angles because of self-occlusion. This resulted in higher variability in angles (Haggag et al., 2013). This preliminary risk assessment implementation shows potential for monitoring posture in industrial workplaces. Camera placement in the frontal plane can minimize self-occlusion errors. Unfortunately, in some cases this placement may not be afforded by the environment.

Kurillo et al. (2012) calculated upper extremity reachable workspace, defined as the set of points relative to the torso that the individual can reach (Kurillo et al., 2012). In essence, a larger surface area is associated with greater upper body range of motion (ROM). The surface areas calculated were compared between Kinect and Impulse, an active marker system (PhaseSpace Inc., San Leandro, CA). The resultant surface areas were found to be comparable which suggests that the Kinect may provide the capability to accurately measure upper extremity ROM.

More recently, Ning and Guo (2013) assessed the utility of the Kinect to assess spinal loading to extend trunk kinematic analysis to the workplace. The Microsoft SDK was used to extract skeletal joint locations. Peak trunk flexion and peak lumbosacral joint moments were calculated for twenty movements. All subjects were assumed to be of the same weight and height to decrease the complexity of computation. The mass, COM, and acceleration of each segment was used in the computation of the lumbosacral joint moment. The results were found to be consistent with qualitative evaluation. Thus, it seems feasible that the Kinect could be used for spinal loading assessment in the workplace. One obvious drawback to this approach is that subject anthropometrics are

required. However, once worker anthropometrics are known it becomes easier to monitor behavior longitudinally.

Overall, the Kinect's utility in workplace ergonomics seems reasonable. The devices spatial capabilities could be used to monitor workplace posture or determine upper body ROM. However, there are issues with workplace use that need to be addressed before implementation. Workplace environments may have complex backgrounds, reflective surfaces, or color schemes that do not allow the Kinect to accurately measure depth. Careful consideration must also be taken to ensure that the device can view the individual without environmental or self-occlusion. Additionally, workers may exhibit complex behavior. Thus, applications would benefit from context awareness when interpreting human movement.

2.5 Open Questions

The previous sections have highlighted the Kinect's current capabilities, potential use in biomechanical applications, and the operational assumptions of the device. Despite these limitations the Kinect's skeletal tracking feature is undoubtedly a convenient solution to collecting 3D kinematic data in more ecological settings outside the laboratory. Other techniques that utilize computer vision protocols, such as background substitution, provide a customized method for investigating 3D kinematics (Stone & Skubic, 2011a, 2011b). By contrast, these methods are far more complex which limits the applicability of such an implementation. The skeletal tracking feature reduces the complexity of identifying joints and determining their relative orientation making it more convenient in biomechanical applications. Additionally, in most cases it is preferable to

use the Microsoft implementation because it is more accessible and provides a greater repertoire of features including multiple range and skeletal tracking modes.

Given the current body of literature it seems that the Kinect is capable of capturing the temporal components of human movement (Galna et al., 2014). Many investigations have looked at the spatial components of human movements (Clark et al., 2012; Clark et al., 2013; Galna et al., 2014; Mobini et al., 2013; Obdrzalek et al., 2012; Schmitz et al., 2014) and have found varying degrees of accuracy but overall there seems to be a consensus that the Kinect has the potential to be a useful tool for clinicians and researchers alike.

Advances in the field of non-linear dynamics have shown that variables such as relative phase may better capture the true dynamics of motor tasks. These measures provide insights into spatiotemporal changes in the human movement system (i.e. musculoskeletal, nervous, sensory systems) that may not be detected by examining only spatial or temporal information. Dynamical systems perspectives on variability in human movement have determined that motor variability can be functional depending on task, environmental, and individual constraints. Additionally, if the Kinect is capable of capturing variability of coordinative measures, it may hold clinical utility for distinguishing between healthy and pathological populations or identifying the potential for injury. No current investigation of the Kinect has assessed its ability to measure coordination during human movement. Certain measures of coordination may be less susceptible to joint position errors and do not require absolute positional comparisons between systems. Concepts relevant to human coordination will be discussed below.

2.6 Coordination in Human Movement

2.6.1 Dynamical Systems Theory

Dynamical systems theory deals with temporally evolving, non-linear systems. These systems can be described by simple, non-linear rules, that give rise to complex behavior depending upon initial conditions. Dynamical systems shift between states according to these rules. The benefit to dynamical systems theory is that it is context free and thus is applicable to many systems.

Complex systems are composed of a large number of subsystems which may consist of smaller subunits. The simple interaction of these subunits leads to emergent behavior not exhibited any single subunit. By nature, the behavior of a complex system relies on the interaction of the subunits and subsystems, not their individual actions. Chaotic systems have small numbers of constituents but can produce exceedingly complex behavior from deterministic rules.

Dynamical systems theory applied to human movement behavior arises from the complex network of co-dependent subsystems (i.e. musculoskeletal, neural, sensory, respiratory) each of which has a large constituency of interacting components (i.e. blood cells, neurons, muscle fibers) (Davids, Glazier, Araujo, & Bartlett, 2003). These subsystems and subunits interact to exhibit complicated, emergent behavior over time. Thus the human movement system can be viewed as a complex, dynamical system (Williams et al., 2000).

2.6.2 Attractors

An important feature of complex systems is their ability to self-organize. Self-organization allows a system exhibit spontaneous pattern formation between its various

components. Self-organization occurs as transitions between different organizational states which emerge due to internal or external constraints which pressure the system to change (Davids et al., 2003). Self-organization is thought to occur around the attractors of a complex system.

Attractors represent states of the system which attract nearby trajectories, or paths, through phase space (Kelso, 1995; Williams et al., 2000). The phase space contains all possible states of a system. Each point in the phase space describes a unique state of the system (Hilborn, 1994). Four kinds of attractors can be defined. Fixed point attractors are defined as static points in the phase space where the system is stable. A dampened pendulum is a simple form of a point attractor where the position that pendulum comes to rest is the point attractor. Limit cycle attractors describe an isolated periodic orbit. An example would be a dampened force pendulum. The pendulum oscillates around the same points regardless of its starting position. A quasiperiodic attractor is a less stable kind of limit cycle attractor where there are multiple frequencies in the periodic trajectory of the system. Last, strange or chaotic attractors are sensitive to the initial conditions of the system. Strange attractors have no steady state, and exhibit complex, fractal-like behavior. Fractals exhibit self-similar patterns that appear similar at vastly different spatial scales.

It is possible for systems and state spaces to have multiple attractors; this is termed multistability. Attractors have basins of stability which is the region in phase space where initial conditions converge to the attractor (Kelso, 1995). An attractor becomes unstable when a control parameter reaches a critical value which produces a qualitative change in the attractor. A control parameter is an input to the system that can

alter the order parameter of the system. This alteration is termed tuning (Ricklefs et al, 2007; Kelso, 1995). The order parameter is the macroscopic variable of a system. Systematically tuning the control parameter through a critical point causes spontaneous changes to the order parameter, these are known as a non-equilibrium phase transitions. If the direction of the control parameter tuning is reversed past the critical value the system may remain in its new state, this is termed hysteresis (Kelso, 1995). Hysteresis demonstrates overlapping regions of attractors. Both non-equilibrium phase transitions and hysteresis are trademark phenomena of complex systems.

2.6.3 Bernstein's Degrees of Freedom Problem

Coordinating the vast number of biophysical degrees of freedom found in the human body is a daunting task. Even more impressive is the ability of the human movement system to produce smooth, coordinated movements. This is made more challenging by the number of independent variables that must be controlled to perform these movements (Turvey, 1990). Controlling and coordinating these degrees of freedom is critical to successful task performance.

The problem of coordinating the human body's degrees of freedom was discussed extensively by Russian physiologist, Nikolai Bernstein. Humans are capable of successfully performing tasks utilizing a variety of movement patterns which incorporate various degrees of freedom. Thus variability in motor performance can be viewed as the movement system's ability to flexibly achieve certain outcomes by employing different biomechanical degrees of freedom during task performance (Davids et al., 2003).

Bernstein initially proposed the idea that the redundant degrees of freedom problem could be solved with temporary couplings between multiple degrees of freedom which results in a coordinative structure (Turvey, 1990). Coordinative structures reduce the complexity of the movement system by exploiting the interconnectivity of the body's segments (Davids et al., 2003). A key feature of coordinative structures is that they are task-specific and thus are adaptable to changes in task or environmental constraints. Coordinative structures are able to readjust after perturbations to preserve task goals (Turvey, 1990).

2.6.4 Absolute and Relative Coordination

In human movement, coordination between the body segments can be absolute, relative, or uncoordinated. Absolute coordination is defined by the rhythmic movement of two or more limbs at the same 1:1 frequency. Absolute coordination displays phase relations that are constant such that there is no temporal difference between discrete points in the cycle. Relative coordination differs in that body segments tend toward a phase relation. However, variability in this phase relation arises due to weaker coupling of the segments (Von Holst, 1973).

Variability in phase may be due to the intrinsic properties of the segments. When similar (homologous) limbs (i.e. fingers or hands) are coordinated, they have similar characteristic frequencies. Thus these segments exhibit absolute coordination. Conversely, when coordinating dissimilar (non-homologous) limbs such as hands and feet, the limbs exhibit different characteristic frequencies and therefore greater phase

variability. Deviation from absolute coordination varies depending on the magnitude of difference between the two segments (Rosenblum & Turvey, 1988).

Relative coordination is driven by competition of the maintenance tendency and the magnet effect (Kelso, 1995). For coupled oscillators the magnet effect is the tendency of one oscillator to impose its preferred phase relation the other Von Holst (1973). Conversely, the maintenance tendency is the tendency of each oscillator to maintain its natural frequency (Von Holst, 1973). The coordinative state of the system is defined by the competition of these tendencies. Absolute coordination will tend to be exhibited if the magnet effect dominates. Relative coordination arises if the maintenance tendency prevails. Thus, relative coordination is characterized by the competition of two coupled oscillators to maintain their natural frequency and simultaneously impose this frequency on the other (Von Holst, 1973).

2.6.5 Relative Phase of Human Movement

Numerous examinations of rhythmic bimanual coordination have demonstrated that humans can coordinate their fingers (or hands) in either an in-phase (0°) or anti-phase (180°) mode at low frequencies (Kelso, 1981; Kelso, 1984; Kelso et al., 1986). Movements that are initially in-phase remain in-phase when the frequency of the movements is increased. When the movements are initially anti-phase and the frequency is increased, the anti-phase attractor becomes unstable and the system switches to the more stable in-phase mode.

Kelso's 1981 and 1984 experiments led to the development of a theoretical model of these phase transitions to describe the observed phase transitions in hand movements. This model was termed the Haken, Kelso and Bunz (HKB) Model (Haken, Kelso, & Bunz, 1985). The model describes the features of bimanual coordination using mathematical equations. V was used to represent potential function. φ was used to represent the order parameter, relative phase. It was assumed that potential V was symmetric and periodic: $V(\varphi) = V(-\varphi)$. Next, potential V was re-written including two cosine functions: $V = -a \cos \varphi - b \cos 2\varphi$. This equation generates a potential function (Figure 2.6) that describes the possible coordination patterns at various frequencies (specified by the ratio of b/a). The model predicts that as frequency increases (the ratio, b/a decreases), a critical value (ω_c) is reached, and the ball transitions to a new well (Figure 2.6). Thus, the fingers transition from an asymmetrical mode (anti-phase) ($\varphi = \pm \pi$) to a symmetrical (in-phase) mode ($\varphi = 0$). The importance of the HKB model lies in

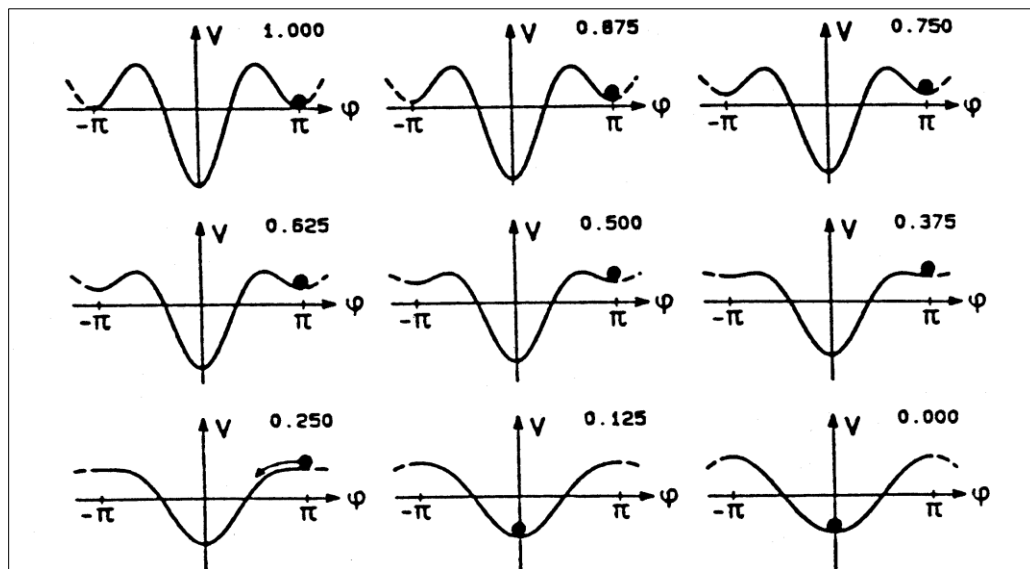


Figure 2.6. HKB model. The ball represents the current state of the systems. The numbers represent the ratio of b/a in the potential function V . As frequency increases, the ratio decreases and the potential wells become shallower, causing the ball to transition to a different well. Reprinted from Haken et al. (1985).

the demonstration of non-equilibrium phase transitions between coordination states which can be modeled using simple non-linear equations (Kelso, 1995). This model is supportive of dynamical systems theory in human movement and the self-organization of complex coordinative states.

2.6.6 Relative Phase Assessment Techniques

There are multiple measures that can be used to assess movement coordination between two coupled oscillators. The first method is discrete relative phase (DRP). DRP is a temporal evaluation of coordination that investigates the latency of discrete events in each cycle (i.e. maxima, minima, inflection point) (Hamill, Haddad, McDermott, 2000; Kelso, 1995). Formula 2.1 provides an example of DRP calculation where t_1 and t_2 are discrete key events and T represents the cycle time. DRP can range from 0° to 360° . However, 0° and 360° are equivalent. A DRP value of 360° indicates perfect in-phase while all other values denote that the oscillators are out of phase. DRP is a circular variable and thus mean DRP and DRP variability should be calculated over multiple cycles using circular statistics (Hamill et al., 2000). An advantage to using DRP is the simplicity of computation only temporal information is used to calculate the phase angle.

Formula 2.1
$$\varphi(t) = \frac{t_1 - t_2}{T} \times 360$$

To address the limitation of measuring coordination once per cycle, continuous relative phase can be used (CRP). CRP is calculated as the four-quadrant arctangent phase angle from a parametric position-velocity phase plot of a single oscillator (Formula 2.2) (Hamill et al., 2000). CRP of coupled oscillators is the parameterized difference in phase angle (Formula 2.3). CRP values range from 0° to 360° . However, there are

redundancies in the angles (i.e. 0° and 360° are the same). Thus, values generally range from 0° to 180° . Oscillators are perfectly in-phase at 0° and perfectly anti-phase at 180° . Any other value represents the relative amount of in-phase or anti-phase coordination. CRP can be used to assess movement variability in coordination. CRP demonstrates the changes in phase relations that occur within a cycle which cannot be examined using DRP. In addition, CRP variability can be calculated over the entire cycle as the point-by-point standard deviation.

Formula 2.2
$$\varphi(t) = \tan^{-1}\left(\frac{\omega(t)}{\theta(t)}\right)$$

Formula 2.3
$$CRP(t) = \varphi_{oscillator1}(t) - \varphi_{oscillator2}(t)$$

Normalization procedures are often used before the calculation of CRP. It has been suggested that CRP only be used when the joint motions are sinusoidal (Diedrich & Warren, 1995). Even in sinusoidal oscillators normalization procedures are recommended if the frequency of the oscillations is a value other than $0.5/\pi$ as the phase-plane will be stretched or compressed along the velocity axis (Peters, Haddad, Heiderscheit, van Emmerik, & Hamill, 2003). Normalization accounts for amplitude differences in the phase plane range of motion. This is relevant when intra-limb coordination is examined (i.e. right hand and right foot). However, normalization is still recommended for inter-limb coordination (i.e. right and left hands).

Hamill et al. (2000) examined different normalization procedures to highlight their effect on CRP. Two normalization procedures were examined and additionally normalization calculated on a cycle-by-cycle basis was compared with normalization

over multiple cycles. The first technique normalized the position-velocity phase portrait to a unit circle using maximum and minimum values (van Emmerik & Wagenaar, 1996). Second, position and velocity were normalized to +1 and -1 each respective axis based on the absolute maximum velocity location (Burgess-Limerick et al. 1993). Normalizing with the absolute maximum velocity method can present issues if the maximum and minimum velocities are not equal in magnitude as the phase portrait will contain dead space where no trajectory cross. The unit circle method does not suffer from this issue as it normalizes to the maximum and minimum. However, this method loses velocity information from the raw data as the origin no longer corresponds to zero velocity (Hamill et al. 2000).

Differences between normalization over a single or multiple cycles were found to be minimal. However, using the maximum of multiple cycles may lead to outliers becoming the reference for normalization distortion will propagate to all cycles. Additionally, normalization over multiple strides was found to better preserve the spatial layout of the original data (Hamill et al., 2000). These procedures can in turn affect variability. Thus decisions about normalization procedures and whether data should be normalized over single or multiple cycles should be made with respect to the type of movement coordination of interest (Hamill et al., 2000).

2.7 Movement System Variability

Recent perspectives on the functional role of movement variability have provided insights into adaptive and goal-directed postural control (Riccio, 1993). These perspectives generally take an ecological approach to human movement such that

environmental constraints shape the behavior of the individual based on the affordances of the postural control system. Traditional perspectives tended to equate noise and variability as harmful to task performance and indicative of pathology (Van Emmerik & Van Wegen, 2002). However, dynamical systems theory has provided tools to help understand the role of variability in coordinative human movement.

Variability in movement patterns is inherent given the large number of degrees of freedom that need to be controlled and coordinated, as stated above in the context of Bernstein's degrees of freedom problem. However, the abundance of biomechanical degrees of freedom provides flexibility and adaptability in that a single task can be successfully completed by coordinating multiple degrees of freedom.

Changes in coordination patterns may arise due to instability in the current state of the system. Relative phase provides information about spatiotemporal changes occurring in human movement. Additionally, variability of relative phase can be used to detect the overall stability of a pattern. If the Kinect is able to measure subtle changes in the variability of coordinated movement it may become a useful diagnostic tool for collecting kinematic data in more ecological settings.

CHAPTER 3. METHODS

3.1 Participants

Twenty healthy, adults, ages 18 to 25 were recruited from Purdue University. Participants will be free of learning, coordination, or neurological disorders. Participants signed an informed consent approved by the Purdue University Institutional Review Board.

3.2 Apparatus

Retroreflective markers were placed bilaterally on both wrists, on the ulnar and radial styloid processes, and on the 3rd metacarpo-phalangeal joint on the dorsal side of the hands. Six Vicon cameras captured 3D kinematic data at 120 Hz. A digital video camera (30 Hz) was synchronized with the Vicon. A Microsoft Kinect fixed to a tripod collected 3D kinematic data (irregularly at 30Hz) of the hand and wrist as defined by the Microsoft SDK skeletal tracking feature in seated and near mode using the Kinect Stream Saver application (Dolatabadi, Taati, Parra-Dominguez, & Mihailidis, 2013). The retroreflective marker placement locations were chosen to mimic the joint locations of the Kinect.

3.3 Setup

Participants sat in a chair with forearms resting on two modified armrests covered in sheets. Lightly colored sheets were used to minimize absorption of the IR light emitted by the Kinect and to hide any edges that might be mistaken for human limbs. Additionally, participants were asked to wear short sleeve shirts to minimize the influence of clothing deformation. The arm rests were adjusted to allow each participant's to sit as comfortably and naturally as possible with elbows flexed to approximately 90 degrees and forearms parallel to the ground. The participant's hands and wrists hung over the edge of the tables to allow freedom of movement. Their hands were prone to the ground to maintain view of the reflective markers. The Kinect will be aligned with the mid sagittal line of the participant in the frontal plane at a distance of 1.5 m. The Kinect was placed at a fixed height of 1.4 m for all participants.

3.4 Procedure

Participants performed coordinated movements of the hands. For the bimanual coordination tasks subjects performed rhythmic hand flexion/extension in the sagittal plane in either the in-phase or anti-phase pattern.

Participants performed a total of 13 trials begin all of these trials in anti-phase mode. A metronome was used to create a driving frequency for the oscillatory movements starting at 1.0 Hz and progressively increasing to 3.33 Hz and then progressively decreasing to 1.0 Hz. The metronome increased in 100 ms intervals every 10 s with the first frequency period lasting 20 seconds, allowing the individual to acquire the assigned pattern. Five trials were performed starting in the in-phase pattern and five

trials in the anti-phase pattern. Pattern order was randomized. Three additional anti-phase trials was conducted with the metronome increasing from 1.0 Hz to 5 Hz in steps of 100 ms and interval lengths of 3 s.

Before each trial the participant will be informed of the initial coordination pattern (i.e. in-phase or anti-phase) and will be asked to briefly produce the pattern to ensure that the coordination pattern is correct. Additionally, participants will be informed that they should not resist changes to the coordination pattern if it feels natural. Participants were also instructed to maintain a 1:1 ratio between the movements of their hands and the beat of the metronome. The first ten trials lasted for approximately two minutes and forty seconds each. The final three trials lasted approximately 25 seconds each.

3.5 Data Analysis

Positional data from each system was exported into a custom Matlab program for post-processing. The signal from the Kinect was interpolated to 120 Hz using a linear interpolation method, appropriate for irregularly sampled signals (Thiebaut & Roques, 2005). Both sets of data were filtered at 8 Hz. Data from the two systems were synchronized temporally using cross-correlation to correct the lag between signals (Li & Caldwell, 1999). A gross motor movement such as arm flexion was performed in each trial to use as an anchor point. Once the signals were synchronized in time, the time series was clipped to include only movement cycles. This ensured all coordination measures were performed when movement is at a steady state. Each frequency bin was analyzed separately which allows us to investigate how well the Kinect captures the relative phase

dynamics at different movement speeds. Additionally, values were averaged across similar conditions and frequencies for each subject (i.e. in-phase, 1.0 Hz).

The outcome measure was continuous (CRP) relative phase and relative phase variability. CRP measures are used to examine coordination of two coupled oscillators. CRP is calculated as the four-quadrant arctangent phase angle from a parametric position-velocity phase plot of a single oscillator (Formula 3.2) (Hamill et al., 2000). CRP of coupled oscillators is the parameterized difference in phase angle (Formula 3.3). CRP values range from 0° to 360°. However, there are redundancies in the angles (i.e. 0° and 360° are the same). Thus, values generally range from 0° to 180°. Oscillators are perfectly in-phase at 0° and perfectly anti-phase at 180°. Any other value represents the relative amount of in-phase or anti-phase coordination. CRP variability can be calculated over the entire cycle as the point-by-point standard deviation. Position and velocity will be normalized to +1 and -1 each respective axis based on the unit circle method for individual cycles (van Emmerik & Wagenaar, 1996).

Formula 3.2
$$\varphi(t) = \tan^{-1}\left(\frac{\omega(t)}{\theta(t)}\right)$$

Formula 3.3
$$CRP(t) = \varphi_{oscillator1}(t) - \varphi_{oscillator2}(t)$$

3.6 Statistical Analysis

Bimanual coordination and coordination stability were assessed between the two systems by comparing mean CRP and mean CRP standard deviation using a mixed model, repeated measures Analysis of Variance (ANOVA) and intraclass correlation coefficients (ICC) (2,1) using 95% limits of agreement. ICC provides information

about the strength of the relationship between two variables. When comparing Kinect and Vicon CRP values ICC can assess the agreement between the two systems regardless of whether measurement biases exist.

CHAPTER 4. MANUSCRIPT

4.1 Introduction

Optical marker-based systems are the gold-standard for capturing three-dimensional (3D) human kinematics (Best & Begg, 2006; Corazza, Mündermann, & Andriacchi, 2006; Mündermann, Corazza, & Andriacchi, 2006; Visser et al., 2008). These systems are accurate, reliable, and capable of tracking a variety of movements in multiple domains, including gait, posture, clinical diagnostics, physical rehabilitation, and workplace ergonomics. The most common systems require passive or active markers to be attached to the participant. Attaching markers is time consuming, and can lead to tracking errors due to placement variability and soft tissue movement artifact (Andriacchi & Alexander, 2000; Cappozzo et al, 1996; Della Croce et al., 2005; Leardini et al., 2005; Mündermann, Corazza, & Andriacchi, 2006). Additionally, marker based systems require trained personnel to operate and are prohibitively expensive and non-portable.

Repurposed gaming peripherals such as the Microsoft Kinect provide a promising alternative to commercial, marker-based motion capture systems. The Kinect is a markerless, inexpensive, and portable depth camera that can track the 3D kinematics of 20 skeletal joints. Recent research in the movement sciences has investigated the ability of the Kinect to accurately and reliability examine postural control and gait in a variety of settings (e.g. Clark et al., 2012; Clark, Bower, et al., 2013; Stone & Skubic, 2011a).

In the gait domain, the Kinect accurately captures step and stride parameters using the skeletal tracking feature. Temporal parameters of gait are however more difficult to determine due to the lack of anatomical landmarks on the feet (Clark, Bower, et al., 2013). In the postural domain the Kinect can accurately determine linear and angular displacements of joints and body segments in aging and pathological populations (Galna et al., 2014; Obdrzalek, et al., 2012). It appears the Kinect can accurately capture temporal postural kinematics better than spatial kinematics (Clark et al, 2012; Galna et al., 2014). In general, both posture and gait studies have found good agreement between the Kinect and commercial motion analysis systems.

To date, the ability of the Kinect to capture spatiotemporal human coordination has not been examined. Examinations of the cooperative action of multiple body segments have provided fundamental insights into how human movement is controlled based on the constraints of the task, individual, and environment (Kelso, 1995; Newell, 1986). In general, coordination research from the dynamical systems perspective has suggested that coordination arises through the self-organization of the multiple degrees of freedom inherent in the human body. Thus, coordination in a human is not unlike the coordination of other complex non-linear systems found in nature. For example, humans can coordinate their fingers (or hands) in either an in-phase (0°) or anti-phase (180°) pattern at low frequencies (Kelso, 1981; Kelso, 1984, Kelso; Scholz, & Schöner, 1986). As the frequency of oscillation is increased, movements started in-phase remains in-phase. When movement is initiated in the anti-phase pattern, and frequency is increased, the anti-phase pattern destabilizes and a spontaneous transition occurs at a critical frequency (Kelso, 1984). This complex process has been modeled using simple difference

equations, suggesting this coordinative behavior is not governed by complex cognitive processes.

Relative phase is a collective variable that has been extensively used to assess spatiotemporal coordination and coordinative stability (i.e. variability of relative phase) of two body segments. Relative phase measures provide insights into spatiotemporal changes in the human movement system that may be undetectable by examining spatial or temporal information alone (Kelso, 1995). Relative phase provides information about how segments are being coordinated and the associated variability provides information about the stability of coordination.

To date, only one study has used the Kinect to capture the variability of human movement (Stone & Skubic, 2011b). Assessment of gait parameter variability including stride length, stride time, and stride velocity revealed that the Kinect was able to accurately capture stride velocity variability. Given that human movement coordination and coordinative variability involves the spatiotemporal interactions of multiple segments and can occur over small time scales, its accuracy cannot be inferred from previous posture and gait validations. If the Kinect is indeed valid in assessing spatiotemporal coordination, new avenues of research can be opened that were not previously possible given the cumbersome nature of commercial marker systems. For example, Volman et al. (2006) demonstrated that children with developmental coordination disorder (DCD) less stably perform in-phase and anti-phase patterns of bimanual coordination. Consequently, the Kinect must be able to accurately capture coordination and coordination stability to be used to this effect. Thus, in this study, we assess the ability of the Kinect capture the relative phase dynamics of bimanual movements. We specifically adopted a paradigm

similar to Kelso, Scholz, & Schöner, (1986). Individuals performed bimanual hand movements to the beat of a metronome. Kinematics were collected using a Kinect and a Vicon motion analysis system.

4.2 Methods

4.2.1 Participants

Twenty-seven college-aged individuals participated. Procedures were approved by the University Institutional Review Board. Individuals were free of any neurological disorders.

4.2.2 Microsoft Kinect

The Microsoft Kinect was used to collect 3D kinematics utilizing the Kinect Stream Saver application (Dolatabadi, Taati, Parra-Dominguez, & Mihailidis, 2013). This application allows users to choose the depth range, skeletal tracking mode, data streams to be collected, as well as the camera tilt angle. Three-dimensional kinematics using the skeletal tracking feature were collected in near and seated mode at approximately 30 Hz at a resolution of 640 x 480 pixels. No audio or depth data were captured.

4.2.3 Vicon

Three-dimensional kinematics were collected with a 6-camera Vicon motion analysis system (Vicon Motion Systems, Oxford, United Kingdom). Six retro-reflective markers were placed bilaterally on the wrists, the ulnar and radial styloid processes, and the 3rd metacarpo-phalangeal joint on the dorsal side of the hands. This marker set-up was used to match the skeletal landmarks of the Kinect. Markers were tracked at 120 Hz. A

digital video camera (30 Hz) was synchronized with the Vicon. A single analog channel, collected at 16200 Hz, was used to determine the timing of metronome beeps.

4.2.4 Apparatus and Procedure

Participants performed thirteen trials of either in- or anti- phase hand movements to the beat of a metronome. For the first ten trials, five trials were started in either the in-phase or anti-phase pattern and were randomly ordered. Each of these trials lasted approximately two minutes and forty seconds. Metronome frequency ranged from 1.0 Hz to 3.33 Hz and was increased incrementally by 100 ms every 10 seconds until the maximum frequency was reached. Upon reaching the highest frequency (3.33 Hz), the metronome was then scaled back down to the starting frequency using the same time steps. Frequency intervals were chosen to be 10 seconds in order to increase the number of cycles for analysis. The first frequency step (1.0 Hz) lasted 20 seconds to allow participants to synchronize with the metronome. The last three trials were started in the anti-phase pattern with frequencies ranging from 1.0 Hz to 5.0 Hz. Movement frequency was increased by 100 ms every 3 seconds. The shorter time intervals served to reduce relaxation time, the time an individual has to stabilize the current pattern (Scholz, Kelso, & Schöner, 1987), thus increasing the likelihood of inducing a phase transition.

Participants were seated in a chair with custom armrests that allowed unrestricted sagittal plane wrist movements. Participants were instructed what pattern of coordination to perform and to their best to remain synchronized with the metronome. They were also instructed to try and resynchronize with the metronome if they noticed they were no longer synchronized using whatever coordination pattern felt most natural. Participants

were allowed to practice synchronizing with the metronome prior to data collection in each coordination pattern. Once completed, the metronome contained a 5 second pause and one loud final beep to signal the end of the trial. For half the participants, the Kinect was placed 1.5 m from the individual at a height of 1.4 m. For the other half of participants, the Kinect was moved forward along a 24 cm track so that the view of the participant was optimized in the monitor. This procedure was utilized because it was observed during pilot testing that even when within the operational range of the Kinect, the participants' distance relative to the Kinect influenced the quality of skeletal tracking. Thus, this procedure allowed us to determine how camera placement impacted accuracy of the Kinect.

4.2.5 Data Processing and Analysis

Nineteen individuals were included in this analysis; seven in the fixed position group and 12 in the adjusted position group. This analysis focuses exclusively on the long trials. Vicon kinematics were tracked (Vicon Motion Systems, Oxford, United Kingdom) and exported with the analog metronome data. The Kinect Stream Saver application was used to export the skeletal trajectories and frame time stamps. Both systems collected data from each trial. Data from both systems were then processed using a custom Matlab program (Mathworks, Natick, MA, USA). The Kinect data was up-sampled to 120 Hz using linear interpolation so that the collection frequency of both systems was the same. Kinematic data from both systems were then filtered using a fourth order Butterworth zero lag filter at 8 Hz. Sagittal plane segment angles and angular velocities were calculated relative to the horizontal. Cross-correlation analysis was used to temporally

align the kinematic trajectories from both systems. Cycles were identified from the right hand kinematics from the Vicon system. Each trial was segmented into each of the movement frequency bins using the metronome. The first cycle at each frequency interval was excluded. Each movement cycle was normalized to one hundred data points. Angular positions and velocities were normalized to a unit circle for each cycle (van Emmerik & Wagenaar, 1996) (Figure 4.1). Phase angles were then calculated for each hand by taking the arctangent of the angular velocity over the angular position (see Formula 4.1). Continuous relative phase (CRP) was calculated as the difference between the phase angles of the two hands (see Formula 4.2 & Figure 4.2). Relative phase variability was calculated as the standard deviation of each normalized data point across all cycles of a given movement frequency. Mean CRP and mean CRP variability was calculated for each movement frequency.

Formula 4.1
$$\varphi(t) = \tan^{-1}\left(\frac{\omega(t)}{\theta(t)}\right)$$

Formula 4.2
$$CRP(t) = \varphi_{oscillator1}(t) - \varphi_{oscillator2}(t)$$

4.2.6 Statistical Analysis

Bimanual coordination and coordination stability were assessed between the two systems by comparing mean CRP and mean CRP standard deviation using a mixed model, repeated measures Analysis of Variance (ANOVA) and intraclass correlation coefficients ($ICC_{(2,1)}$). The ANOVA model included device (i.e. Kinect or Vicon), distance (i.e. fixed or adjusted), pattern (i.e. in-phase or anti-phase), and frequency with pattern and frequency nested within device as repeated factors. For significant effects,

Tukey HSD tests were used to determine which comparisons were significantly different.

The alpha level was $p < 0.05$. The agreement between systems was evaluated using

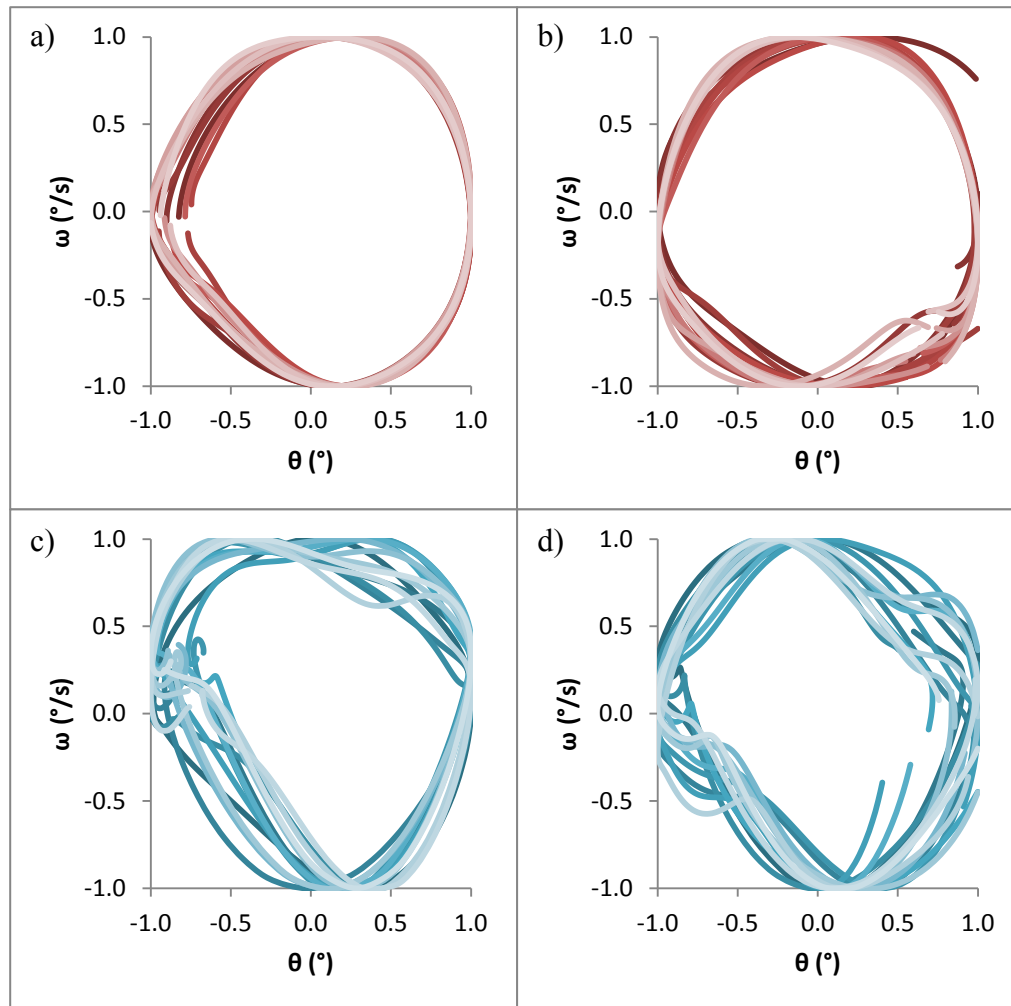


Figure 4.1. Normalized phase planes of the right and left hands for each device at a frequency of 1.67 Hz from an exemplar participant: a) Vicon right hand, b) Vicon left hand, c) Kinect right hand, d) Kinect left hand.

intraclass correlation coefficient ($ICC_{(2,1)}$) using 95% limits of agreement. ICC provides information about the agreement between the two systems regardless of whether measurement biases exist.

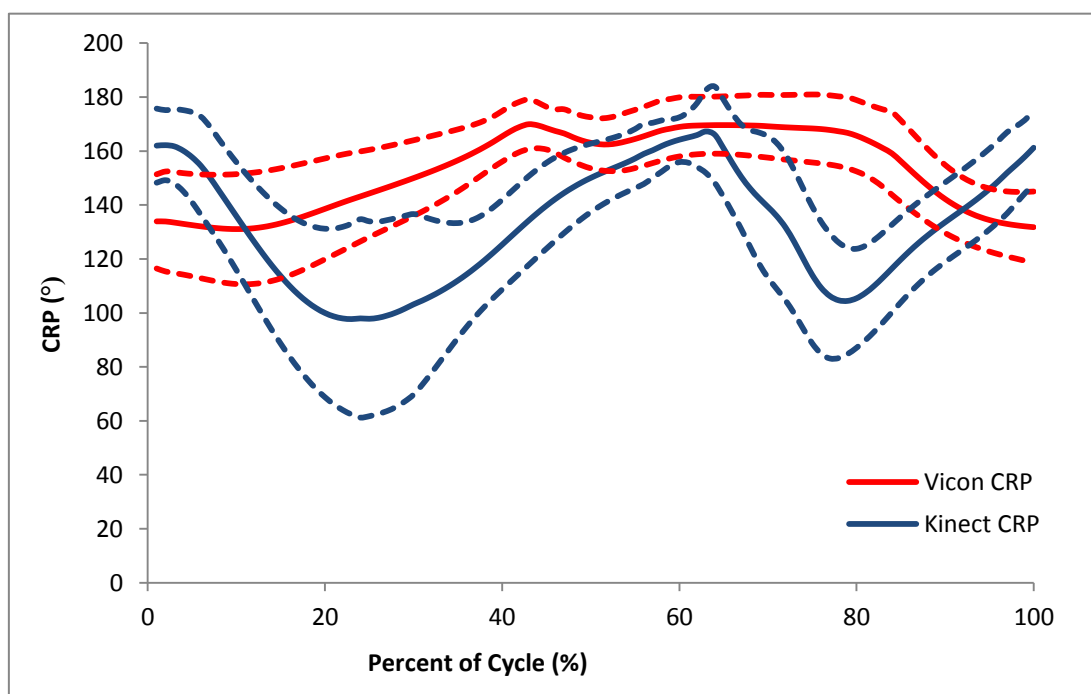


Figure 4.2. Ensemble CRP curves created by normalizing cycles to 100 data points and averaging across all cycles of a given frequency. The CRP curves in this figure are calculated from the phase planes in Figure 4.1, above. The solid lines represent ensemble CRP, the dotted lines represent ensemble CRP plus or minus the standard deviation of CRP.

4.3 Results

4.3.1 Mean CRP

As expected, significant differences between patterns of coordination, in-phase (0) and anti-phase (180), were found ($p < .0001$). The mean CRP and standard error for both devices across the two patterns of coordination is displayed in Table 4.1. Significant differences in mean CRP were found between devices ($p < .0001$). Specifically, the Kinect underestimated CRP by approximately 4 degrees. No main effects of Kinect distance or movement frequency was observed. Additionally, a significant device by pattern interaction was found. Interestingly, the Kinect underestimated the anti-phase pattern by 21.21° ($p < .0001$) and overestimated the in-phase pattern by 12.70° ($p <$

.0001; Figure 4.3). Despite these findings, a high intra-class correlation coefficient ($ICC_{(2,1)}$) for mean CRP was observed ($r=0.97$).

Table 4.1. Mean CRP and standard error (in parentheses) of the Kinect and Vicon

	Kinect	Vicon
In Phase CRP (°)	17.3 (1.6)	29.9 (1.6)
Anti-Phase CRP (°)	131.4 (1.6)	152.6 (1.6)

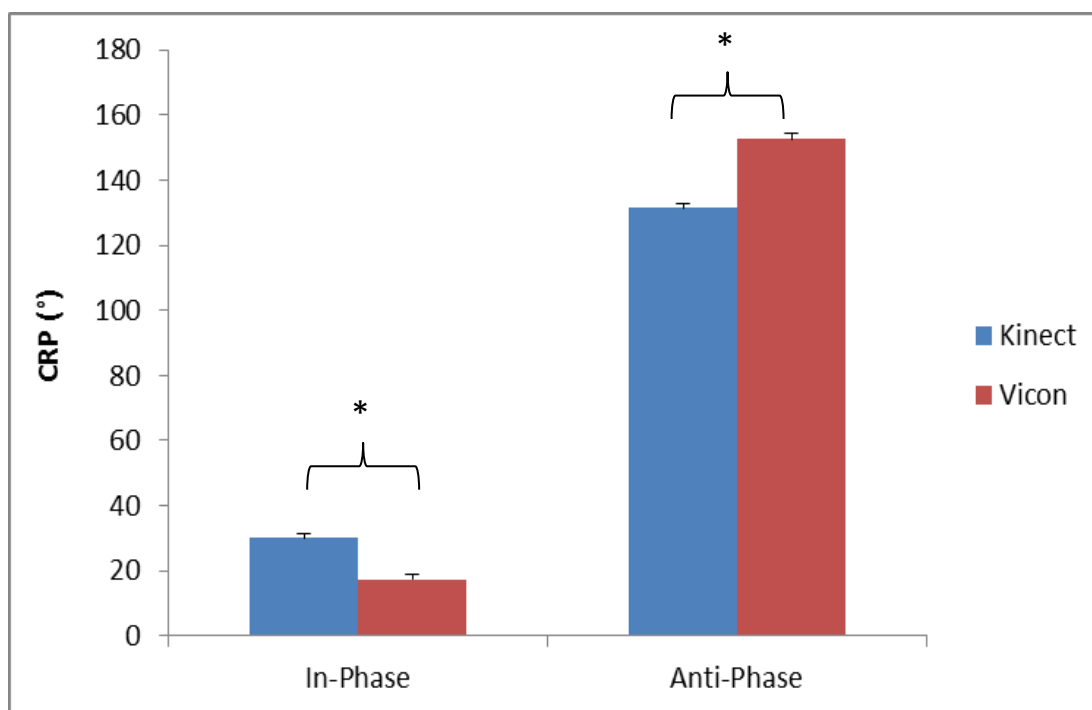


Figure 4.3. Mean CRP measured for in-phase and anti-phase patterns of coordination for the Kinect and Vicon. The Kinect significantly underestimated the anti-phase pattern and significantly overestimated the in-phase pattern. * $p < .0001$

4.3.2 Mean CRP Variability

Variability of CRP, a common measure to assess coordinative stability, was also examined. Variability was calculated as the standard deviation of mean CRP. A significant device by pattern interaction was found. The Kinect exhibited significantly higher variability than the Vicon ($p < .0001$) and between patterns of coordination with anti-phase variability higher than in-phase ($p < .0001$) (Figure 4.4). Differences were assessed using Tukey adjusted p-values of the device by pattern interaction. The Kinect displayed higher variability than the Vicon for both patterns ($p < .0001$). Both systems measured higher variability in the anti-phase pattern than in the in-phase pattern ($p < .0001$) (Table 4.2).

In addition, the coordination stability for the Kinect was significantly smaller in the adjusted position (26.51°) than the fixed position (21.00°) ($p < .0035$). The coordination variability was found to be significantly higher in the Kinect than the Vicon at all movement frequencies ($p < .0001$). No significant differences were found between coordination variability in the in-phase pattern at the various movement frequencies for the Vicon except for the highest frequency (3.33 Hz) which was significantly higher than all other frequencies ($p < .0001$). However, the Kinect in-phase coordination variability was similar at lower frequencies (1.0 - 2 Hz) with variability increasing dramatically at the highest two frequencies (2.5 Hz and 3.33 Hz) (Figure 4.5). For the anti-phase pattern, variability was similar at frequencies up to 2.0 Hz in the Vicon, with higher variability at 2.5 Hz and 3.33 Hz. Interestingly, variability was similar at all frequencies when measured using the Kinect while in the anti-phase pattern (Figure 4.6). The intra-class

correlation (ICC(2,1,)) calculated for CRP variability between devices found poor agreement ($r=.37$).

Table 4.2. Standard deviation of mean CRP and standard error (in parentheses) of the Kinect and Vicon

	Kinect	Vicon
In-Phase CRP SD (°)	20.3 (0.8)	9.2 (0.8)
Anti-Phase CRP SD (°)	27.1 (0.8)	14.0 (0.8)

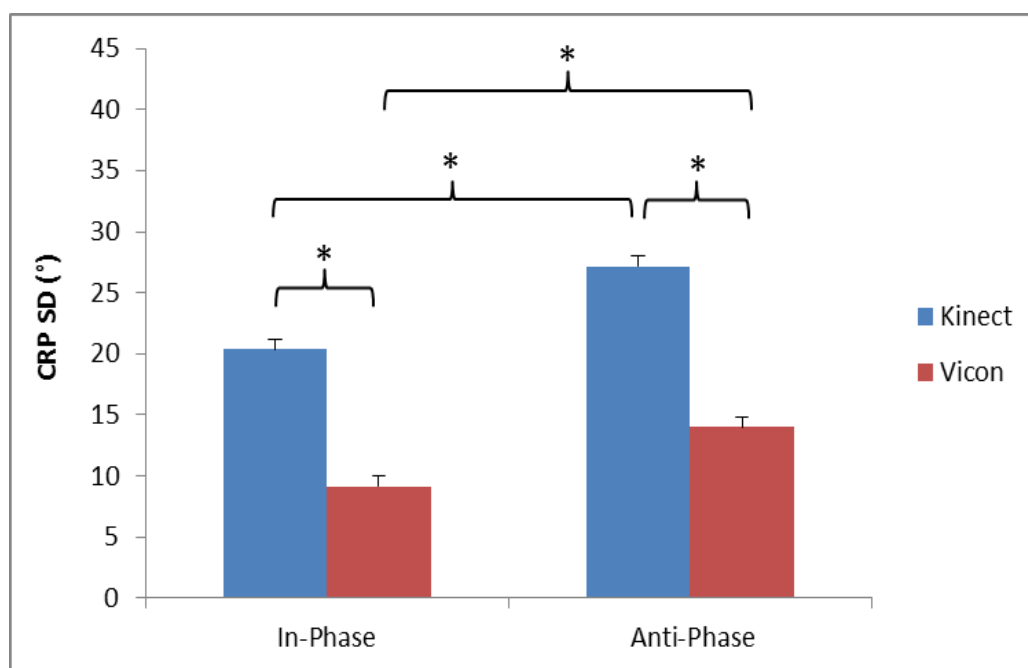


Figure 4.4. Variability of CRP for the Kinect and Vicon across both patterns of coordination. The Kinect overestimated variability for both patterns. Both systems were able to capture the coordination stability differences between patterns. $*p < .0001$

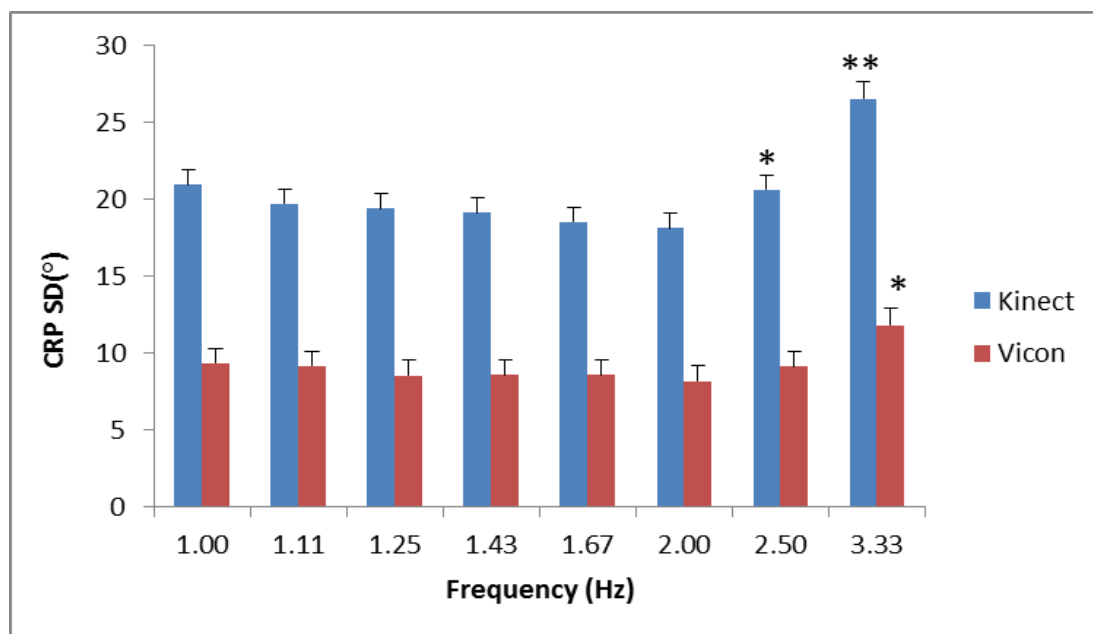


Figure 4.6. Relative phase variability for both devices across the various movement frequencies for the in-phase pattern. Variability measured by the Kinect was higher at all frequencies ($p < .0001$). Kinect variability was similar at all frequencies up to 2.0 Hz with the highest frequency being significantly higher than the rest ($p < .0001$). All Vicon frequencies were similar except for the highest frequency ($p < .0001$).

* $p < .0001$ ** $p < .0001$

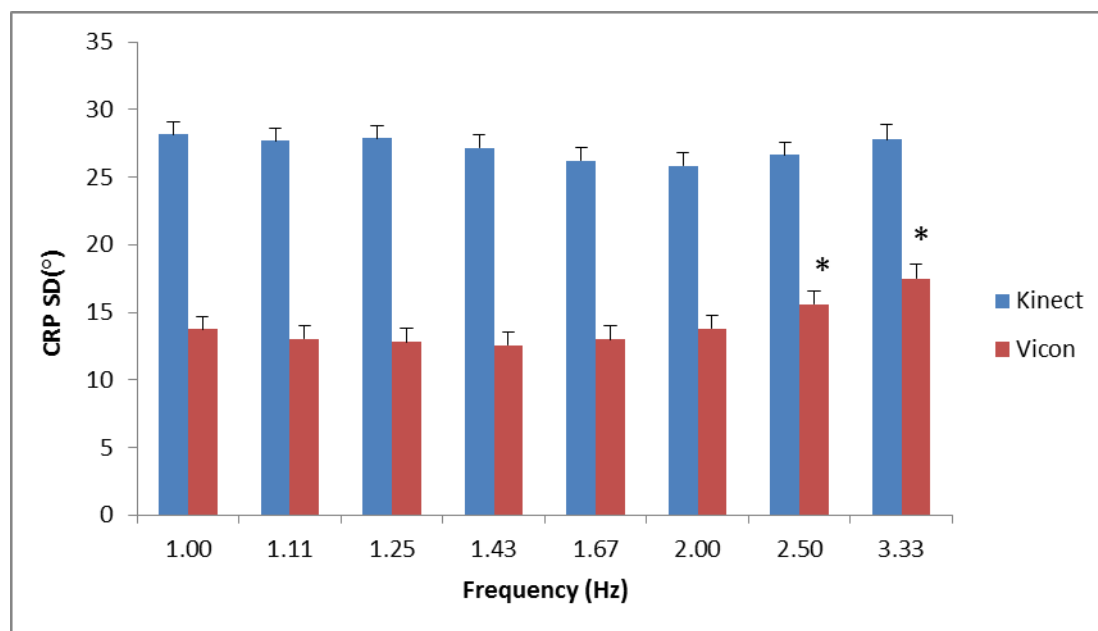


Figure 4.5. Relative phase variability for both devices across the various movement frequencies for the anti-phase pattern. Variability measured by the Kinect was higher at all frequencies than the Vicon ($p < .0001$). Kinect variability was not significantly different across frequencies. All Vicon frequencies were similar except for the highest two frequencies ($p < .0001$). * $p < .0001$

4.4 Discussion

Recent investigations of the Microsoft Kinect, a markerless, inexpensive, and portable depth camera, for uses in biomechanical applications have yielded generally positive results compared to commercial motion capture systems. To our knowledge, no research involving the Kinect has focused on spatiotemporal measures of human coordination. Our results demonstrate that the Kinect is capable of capturing certain features of human bimanual coordination but has trouble assessing the structure of coordination stability.

The primary purpose of this investigation was to determine the Kinect's capacity for capturing the spatiotemporal dynamics of human coordination using Kelso's (1984) bimanual coordination paradigm. Although the Kinect can capture basic postural and gait variables, common coordination variables such as relative phase are time dependent and evolve over short time scales. The lower temporal and spatial resolution of the Kinect may therefore not be suited to capture human coordination. Based on our results, the Kinect is less sensitive than the Vicon at assessing CRP. The Kinect significantly underestimates the anti-phase pattern of coordination by approximately 20° . For the in-phase pattern, the Kinect overestimates CRP by 12° (Figure 4.3). It is unclear why the Kinect's error for the anti-phase pattern is higher than in-phase. However, anti-phase variability, particularly at higher movement frequencies is less stable than in-phase (Kelso, 1984, Kelso; Scholz, & Schöner, 1986). This could lead to difficulties in accurately measuring relative phase. It seems likely that the inherently higher variability in the anti-phase pattern compared to the in-phase pattern made it difficult for the Kinect to accurately assess CRP. This in turn could limit the Kinect's use in capturing subtle

differences in coordination. However, the CRP differences between, although they might seem quite large, are not necessarily indicative that the Kinect cannot be used to assess coordination. The Kinect could prove useful in identifying generalized patterns of coordination since it is sensitive enough to distinguish between in-phase and anti-phase patterns and the ICC values when examining mean CRP was high. The high ICC suggests that although the Kinect CRP values are significantly different than the Vicon, there is a consistent bias between the two systems. With future research, it may be possible to remove this measurement bias. Overall, despite differences in mean CRP, the Kinect could still prove useful in biomechanical applications focused on assessing general patterns of human coordination.

The second focus of this study was to assess the Kinect's ability to capture coordination stability. The structure of coordination variability is a critical component of the relative phase dynamics as it specifies instability in one pattern of coordination and the emergence of a new coordination pattern (Scholz, & Schöner, 1986). Accurately assessing coordination stability is more difficult given the small spatiotemporal movement fluctuations that evolve over time. The in-phase pattern is inherently more stable than the anti-phase pattern and both devices were capable of detecting this difference (Figure 4.2). The Kinect did overestimate this variability compared to the Vicon for both patterns, but it is nonetheless important that the Kinect is sensitive to these differences. Another important feature of anti-phase coordination is that coordination stability decreases (i.e. higher variability) as movement speed is increased. The Vicon was able to capture this phenomenon at the highest two frequencies. However, there were no differences found between any frequencies for the Kinect in the anti-phase pattern.

The high variability found at all anti-phase movement frequencies for the Kinect likely influenced the large bias in mean CRP. In-phase coordination variability should remain relatively constant since this pattern is stable at all movement frequencies. As predicted, CRP variability was similar at all but the highest movement frequencies for the Vicon. The Kinect exhibited higher variability at the highest two frequencies. Although CRP variability measured by the Kinect was higher at all movement frequencies and patterns of coordination, the Kinect was able to detect differences between the two patterns of coordination. The ICC calculated for coordination stability highlights the poor absolute agreement between systems. Thus, the Kinect appears to be able to collect global changes in stability between two coordination modes, however, the lack of agreement between systems needs to be better understood.

In addition, it was found that Kinect placement relative to the participant is a critical aspect of the experimental set-up. When using the fixed distance, the Kinect skeleton was poorly fit to the shorter participants. Interestingly, no significant differences were found between the fixed versus adjusted groups for mean CRP. However, there was a significant improvement in coordination stability ($\sim 5^\circ$) in the adjusted group. There are two possible explanations for the improvement in skeletal fit when the camera was moved. First, it could be that adjusting the Kinect led to subtle improvements in capturing relative phase because the spatial resolution of the hands and forearms was improved. Second, moving the Kinect closer to the participant reduced depth related noise and thus improved the variability of relative phase. It should be noted, the best and easiest way to optimize the Kinect distance is to visually inspect the skeleton within the Stream Saver application until the image filled the window on the computer screen and there is minimal

noise in the joints of the upper extremity. However, more participants need to be added to the fixed distance group to determine the effect of Kinect placement. Additionally, an empirical investigation could better determine the best way to position the device relative to the participant.

There are a few considerations for future investigations of the Kinect for assessing human coordination. First, it is important to examine the ability of the Kinect to capture lower limb coordination during gait (e.g. thigh and leg CRP). Previous investigations of lower limb coordination have provided important insights into the mechanisms of injury (Hamill et al., 1999), the walk-run transition (Dietrich & Warren, 1995), and lower limb asymmetries (Haddad, van Emmerik, Whittlesey, & Hamill, 2006). The Kinect could possibly allow for gait coordination to be captured in a variety of environments not previously possible with traditional motion capture systems. Caution should be taken when applying the results from the current study to gait. Hand coordination may be different from leg coordination because the hands are the last segment on the Kinect's kinematic chain and may be subject to higher variability since internal and external rotations are not recognized.

Second, coordination needs to be assessed for larger movement amplitudes. The vertical displacements of the hands were smaller at higher movement frequencies. The Kinect may be less able to pick up such small changes in position. The Kinect for Xbox One, which is to be released soon, has a higher resolution and may therefore better capture bimanual coordination.

Last, future research should systematically adjust the Kinect's distance relative to each participant. From the current results, it is difficult to determine whether at the

adjusted Kinect distances were still too far to assess CRP. Previous work has demonstrated that depth related noise increases quadratically with increasing distance from the camera (Khoshelham, 2011; Menna et al, 2011; Nguyen et al., 2012). If a larger proportion of the image is taken up by the participant there is greater spatial resolution (i.e. pixels covering body area) which should improve accuracy.

In conclusion, the Kinect was unable to accurately capture mean CRP. However, the high ICC between the two systems is promising and the Kinect was able to distinguish between the coordination stability of in-phase and anti-phase coordination. However, the structure of variability as movement speed increased was dissimilar to the Vicon, particularly for the anti-phase pattern. Thus, taken together, whether or not the Kinect can be used to collect human coordination heavily depends on the exact research question being asked. Some aspects of coordination are nicely captured by the Kinect while others are not. Detecting differences between bimanual coordination patterns and the stability of those patterns can be achieved using the Kinect. However, researchers interested in the structure of coordination stability should exercise caution since poor agreement was found between systems.

REFERENCES

REFERENCES

- Andriacchi, T. P., & Alexander, E. J. (2000). Studies of human locomotion: past, present and future. *J Biomech*, 33(10), 1217-1224.
- Best, R., & Begg, R. (2006). Overview of movement analysis and gait features. *Computational intelligence for movement sciences: neural networks and other emerging techniques*, 1, 1-69.
- Burgess-Limerick, R., Abernethy, B., & Neal, R. J. (1993). Relative phase quantifies interjoint coordination. *J Biomech*, 26(1), 91-94.
- Cappozzo, A. (1991). Three-dimensional analysis of human walking: Experimental methods and associated artifacts. *Hum Movement Sci*, 10(5), 589-602.
- Cappozzo, A., Catani, F., Leardini, A., Benedetti, M. G., & Della Croce, U. (1996). Position and orientation in space of bones during movement: experimental artefacts. *Clin Biomech*, 11(2), 90-100.
- Cappozzo, A., Della Croce, U., Leardini, A., & Chiari, L. (2005). Human movement analysis using stereophotogrammetry. Part 1: theoretical background. *Gait Posture*, 21(2), 186-196.
- Chang, C. Y., Lange, B., Zhang, M., Koenig, S., Requejo, P., Somboon, N., ... & Rizzo, A. A. (2012, May). Towards pervasive physical rehabilitation using Microsoft Kinect. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on* (pp. 159-162). IEEE.
- Chang, Y. J., Chen, S. F., & Chuang, A. F. (2011). A gesture recognition system to transition autonomously through vocational tasks for individuals with cognitive impairments. *Res Dev Disabil*, 32(6), 2064-2068.
- Chang, Y. J., Chen, S. F., & Huang, J. D. (2011). A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. *Res Dev Disabil*, 32(6), 2566-2570.
- Chen, S. T., Chiang, I. T., Liu, E. Z. F., & Chang, M. (2012). Effects of improvement on selective attention: Developing appropriate somatosensory video game interventions for institutional-dwelling elderly with disabilities. *TOJET*, 11(4), 409-417.
- Clark, R. A., Bower, K. J., Mentiplay, B. F., Paterson, K., & Pua, Y. H. (2013a). Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. *J Biomech*, 46(15), 2722-2725.

- Clark, R. A., Bryant, A. L., Pua, Y., McCrory, P., Bennell, K., & Hunt, M. (2010). Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. *Gait Posture*, *31*(3), 307-310.
- Clark, R. A., McGough, R., & Paterson, K. (2011). Reliability of an inexpensive and portable dynamic weight bearing asymmetry assessment system incorporating dual Nintendo Wii Balance Boards. *Gait Posture*, *34*(2), 288-291.
- Clark, R. A., Pua, Y. H., Fortin, K., Ritchie, C., Webster, K. E., Denehy, L., & Bryant, A. L. (2012). Validity of the Microsoft Kinect for assessment of postural control. *Gait Posture*, *36*(3), 372-377.
- Clark, R. A., Pua, Y. H., Bryant, A. L., & Hunt, M. A. (2013). Validity of the Microsoft Kinect for providing lateral trunk lean feedback during gait retraining. *Gait Posture*, *38*(4), 1064-1066.
- Corazza, S., Mündermann, L., & Andriacchi, T. (2006a). Markerless motion capture methods for the estimation of human body kinematics. In *9th international symposium on the 3D analysis of human movement*.
- Corazza, S., Mündermann, L., Chaudhari, A., Demattio, T., Cobelli, C., & Andriacchi, T. (2006b). A markerless motion capture system to study musculoskeletal biomechanics: Visual hull and simulated annealing approach. *Ann Biomed Eng*, *34*(6), 1019-1029.
- Davids, K., Glazier, P., Araújo, D., & Bartlett, R. (2003). Movement systems as dynamical systems. *Sports Med*, *33*(4), 245-260.
- Diedrich, F. J., & Warren Jr, W. H. (1995). Why change gaits? Dynamics of the walk-run transition. *J. Exp. Psychol.-Hum. Percept. Perform.*, *21*(1), 183.
- Della Croce, U., Leardini, A., Chiari, L., & Cappozzo, A. (2005). Human movement analysis using stereophotogrammetry Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics. *Gait Posture*, *21*(2), 226-237.
- Dolatabadi, E., Taati, B., Parra-Dominguez, G.S., & Mihailidis, A. (2013). A markerless motion tracking approach to understand changes in gait and balance: A Case Study, the Rehabilitation Engineering and Assistive Technology Society of North America (RESNA 2013 Annual Conference). Student Scientific Paper Competition Winner).
- Dutta, T. (2012). Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. *Appl Ergon*, *43*(4), 645-649.
- Dutta, A., Banerjee, A., & Dutta, A. (2013). Low-cost visual postural feedback with Wii Balance Board and Microsoft Kinect - a feasibility. *2013 IEEE Point-of-Care Healthcare Technologies (PHT)*, 291-294.
- Foo, J., Paterson, K., Williams, G., & Clark, R. (2013). Low-cost evaluation and real-time feedback of static and dynamic weight bearing asymmetry in patients undergoing in-patient physiotherapy rehabilitation for neurological conditions. *J Neuroeng Rehabil*, *10*(1), 74.

- Galna, B., Barry, G., Jackson, D., Mhiripiri, D., Olivier, P., & Rochester, L. (2014). Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait Posture*, *39*(4), 1062-1068.
- Garcia, J. A., Felix Navarro, K., Schoene, D., Smith, S. T., & Pisan, Y. (2012). Exergames for the elderly: Towards an embedded Kinect-based clinical test of falls risk. *Stud Health Technol Inform*, *178*, 51-7.
- Garrido, J. E., Penichet, V. M. R., Lozano, M. D., & Valls, J. A. (2013). Automatic Detection of Falls and Fainting. *J. UCS*, *19*(8), 1105-1122.
- González, A., Hayashibe, M., & Fraise, P. (2012a, October). Estimation of the Center of Mass with Kinect and Wii balance board. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (pp. 1023-1028). IEEE.
- González, A., Hayashibe, M., & Fraise, P. (2012, August). Three dimensional visualization of the statically equivalent serial chain from kinect recording. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (pp. 4843-4846). IEEE.
- Haddad, J. M., Gagnon, J. L., Hasson, C. J., Van Emmerik, R. E., & Hamill, J. (2006). Evaluation of time-to-contact measures for assessing postural stability. *J Appl Biomech*, *22*(2), 155.
- Haddad, J. M., Ryu, J. H., Seaman, J. M., & Ponto, K. C. (2010). Time-to-contact measures capture modulations in posture based on the precision demands of a manual task. *Gait Posture*, *32*(4), 592-596.
- Haddad, J. M., van Emmerik, R. E., Whittlesey, S. N., & Hamill, J. (2006). Adaptations in interlimb and intralimb coordination to asymmetrical loading in human walking. *Gait Posture*, *23*(4), 429-434.
- Haggag, H., Hossny, M., Nahavandi, S., & Creighton, D. (2013, April). Real Time Ergonomic Assessment for Assembly Operations Using Kinect. In *Computer Modelling and Simulation (UKSim), 2013 UKSim 15th International Conference on* (pp. 495-500). IEEE.
- Hamill, J., McDermott, W. J., & Haddad, J. M. (2000). Issues in quantifying variability from a dynamical systems perspective. *J Appl Biomech*, *16*(4), 407-418.
- Hamill, J., van Emmerik, R. E., Heiderscheit, B. C., & Li, L. (1999). A dynamical systems approach to lower extremity running injuries. *Clin Biomech*, *14*(5), 297-308.
- Hausdorff, J. M., Rios, D. A., & Edelberg, H. K. (2001). Gait variability and fall risk in community-living older adults: a 1-year prospective study. *Arch Phys Med Rehabil*, *82*(8), 1050-1056.
- Hilborn, R. C., Coppersmith, S., & Mallinckrodt, A. J. (1994). Chaos and nonlinear dynamics: an introduction for scientists and engineers. *Comput Phys*, *8*(6), 689-689.

- Holmes, J. D., Jenkins, M. E., Johnson, A. M., Hunt, M. A., & Clark, R. A. (2013). Validity of the Nintendo Wii(R) balance board for the assessment of standing balance in Parkinson's disease. *Clin Rehabil*, 27(4), 361-366.
- Howells, B. E., Clark, R. A., Ardern, C. L., Bryant, A. L., Feller, J. A., Whitehead, T. S., & Webster, K. E. (2012). The assessment of postural control and the influence of a secondary task in people with anterior cruciate ligament reconstructed knees using a Nintendo Wii Balance Board. *Brit J Sport Med*, 47, 914-919.
- Kelso, J.A.S. (1981). On the oscillatory basis of movement. *Bull Psychon Soc*, 18, 63.
- Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. MIT press.
- Kelso, J. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *Am J Physiol*, 246(6 Pt 2), R1000-R1004.
- Kelso, J. A. S., Scholz, J. P., & Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Phys Lett A*, 118(6), 279-284.
- Khoshelham, K. (2011, August). Accuracy analysis of kinect depth data. In *ISPRS workshop laser scanning*, 38(5), W12.
- Khoshelham, K., & Elberink, S. O. (2012). Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors*, 12(2), 1437-1454.
- Koslucher, F., Wade, M. G., Nelson, B., Lim, K., Chen, F. C., & Stoffregen, T. A. (2012). Nintendo Wii Balance Board is sensitive to effects of visual tasks on standing sway in healthy elderly adults. *Gait Posture*, 36(3), 605-608.
- Kurillo, G., Han, J. J., Obdržálek, Š., Yan, P., Abresch, R. T., Nicorici, A., & Bajcsy, R. (2012). Upper Extremity Reachable Workplace Evaluation with Kinect. *Stud Health Technol Inform*, 184, 247-253.
- Lange, B., Chang, C. Y., Suma, E., Newman, B., Rizzo, A. S., & Bolas, M. (2011, August). Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (pp. 1831-1834).
- Lange, B., Koenig, S., McConnell, E., Chang, C., Juang, R., Suma, E., ... & Rizzo, A. (2012, March). Interactive game-based rehabilitation using the Microsoft Kinect. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE* (pp. 171-172).
- Leardini, A., Chiari, L., Della Croce, U., & Cappozzo, A. (2005). Human movement analysis using stereophotogrammetry Part 3: Soft tissue artifact assessment and compensation. *Gait Posture*, 21(2), 212-225.
- Li, L., & Caldwell, G. E. (1999). Coefficient of cross correlation and the time domain correspondence. *J Electromyogr Kines*, 9(6), 385-389.

- Livingston, M. A., Sebastian, J., Ai, Z., & Decker, J. W. (2012, March). Performance measurements for the Microsoft Kinect skeleton. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE* (pp. 119-120). IEEE.
- Luna-Oliva, L., Ortiz-Gutiérrez, R. M., Cano-de la Cuerda, R., Piédrola, R. M., Alguacil-Diego, I. M., Sánchez-Camarero, C., & Martínez Culebras, M. D. C. (2013). Kinect Xbox 360 as a therapeutic modality for children with cerebral palsy in a school environment: A preliminary study. *NeuroRehabilitation*, 33(4), 513-521.
- MacCormick, J. (2011). *How does the Kinect work?* PowerPoint Presentation. Dickinson College <http://users.dickinson.edu/~jmac/selected-talks/kinect.pdf>.
- Macknoja, R., Chávez-Aragón, A., Payeur, P., & Laganière, R. (2012, November). Experimental characterization of two generations of Kinect's depth sensors. In *IEEE International Symposium on Robotic and Sensors Environments* (pp. 150-155).
- Maki, B. E. (1997). Gait changes in older adults: predictors of falls or indicators of fear. *J Am Geriatr Soc*, 45(3), 313.
- Marey, E. J. (1874). *Animal Mechanism a Treatise on Terrestrial and Aerial Locomotion by EJ Marey* (Vol. 11). Henry S. King & Company.
- McGough, R., Paterson, K., Bradshaw, E. J., Bryant, A. L., & Clark, R. A. (2012). Improving lower limb weight distribution asymmetry during the squat using Nintendo Wii Balance Boards and real-time feedback. *J Strength Cond Res*, 26(1), 47-52.
- Medved, V. (2002). *Measurement of human locomotion*. CRC press.
- Menna, F., Remondino, F., Battisti, R., & Nocerino, E. (2011, June). Geometric investigation of a gaming active device. In *SPIE Optical Metrology* (pp. 80850G-80850G). International Society for Optics and Photonics.
- Metcalf, C. D., Robinson, R., Malpass, A. J., Bogle, T. P., Dell, T. A., Harris, C., & Demain, S. H. (2013). Markerless motion capture and measurement of hand kinematics: validation and application to home-based upper limb rehabilitation. *Biomedical Engineering, IEEE Transactions on*, 60(8), 2184-2192.
- Microsoft. (2013a). Kinect for Windows features. <http://www.microsoft.com/en-us/kinectforwindows/discover/features.aspx>. Retrieved November 18, 2013.
- Microsoft. (2013b). Kinect for Windows Dev Center. <http://www.microsoft.com/en-us/kinectforwindowsdev/default.aspx>. Retrieved November 18, 2013.
- Mobini, A., Behzadipour, S., & Saadat Foumani, M. (2013). Accuracy of Kinect's skeleton tracking for upper body rehabilitation applications. *Disabil Rehabil Assist Technol*, 9(4), 344-352.
- Moeslund, T. B. & Granum, E. (2001). A Survey of Computer Vision-Based Human Motion Capture. *Comput Vis Image Und*, 81(3), 231-268.

- Moeslund, T. B., Hilton, A., & Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Comput Vis Image Und*, 104(2-3), 90-126.
- Mombarg, R., Jelsma, D., & Hartman, E. (2013). Effect of Wii-intervention on balance of children with poor motor performance. *Res Dev Disabil*, 34(9), 2996-3003.
- Mündermann, L., Corazza, S., & Andriacchi, T. P. (2006). The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. *J Neuroeng Rehabil*, 3, 6.
- Mündermann, L., Corazza, S., Chaudhari, A. M., Andriacchi, T. P., Sundaresan, A., & Chellappa, R. (2006, February). Measuring human movement for biomechanical applications using markerless motion capture. In *Electronic Imaging 2006* (pp. 60560R-60560R). International Society for Optics and Photonics.
- Muybridge, E. (1969). *Animal locomotion*. New York: Da Capo Press.
- Newell, K. M. (1986). Constraints on the development of coordination. *Motor development in children: Aspects of coordination and control*, 34, 341-360.
- Ning, X., & Guo, G. (2013). Assessing spinal loading using the kinect depth sensor: A feasibility study. *Sensors Journal, IEEE*, 13(4), 1139-1140.
- Nguyen, C. V., Izadi, S., & Lovell, D. (2012, October). Modeling kinect sensor noise for improved 3d reconstruction and tracking. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on* (pp. 524-530). IEEE.
- Obdrzalek, S., Kurillo, G., Ofli, F., Bajcsy, R., Seto, E., Jimison, H., & Pavel, M. (2012, August). Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population. In *Engineering in medicine and biology society (EMBC), 2012 Annual International Conference of the IEEE*, (pp. 1188-1193).
- Payton, C., & Bartlett, R. (2007). *Biomechanical Evaluation of Movement in Sport and Exercise: The British Association of Sport and Exercise Sciences Guide*: Routledge.
- Peters, B. T., Haddad, J. M., Heiderscheid, B. C., Van Emmerik, R. E., & Hamill, J. (2003). Limitations in the use and interpretation of continuous relative phase. *J Biomech*, 36(2), 271-274.
- Plotnik, M., Giladi, N., & Hausdorff, J. M. (2007). A new measure for quantifying the bilateral coordination of human gait: effects of aging and Parkinson's disease. *Exp Brain Res*, 181(4), 561-570.
- Poppe, R. (2007). Vision-based human motion analysis: An overview. *Comput Vis Image Und*, 108(1-2), 4-18.
- PrimeSense. (2013). PrimeSense™ NITE Algorithms 1.5. <http://primesense.com>. Retrieved November 18, 2013.

- Rantz, M., Skubic, M., Abbott, C., Pak, Y., Stone, E. E., & Miller, S. J. (2012, October). Automated Fall Risk Assessment and Detection in the Home: A Preliminary Investigation. In *AAAI Fall Symposium: Artificial Intelligence for Gerontechnology*.
- Riccio, G. E. (1992). *Information in movement variability about the qualitative dynamics of posture and orientation*. University of Illinois at Urbana-Champaign.
- Rickles, D., Hawe, P., & Shiell, A. (2007). A simple guide to chaos and complexity. *J Epidemiol Commun H*, 61(11), 933-937.
- Rosenblum, L. D., & Turvey, M. T. (1988). Maintenance tendency in co-ordinated rhythmic movements: relative fluctuations and phase. *Neuroscience*, 27(1), 289-300.
- Schmitz, A., Ye, M., Shapiro, R., Yang, R., & Noehren, B. (2014). Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. *J Biomech*, 47(2), 587-591.
- Serrien, D. J., Swinnen, S. P., & Stelmach, G. E. (2000). Age-related deterioration of coordinated interlimb behavior. *J Gerontol B-Psychol*, 55(5), P295-P303.
- Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., . . . Moore, R. (2013). Real-time human pose recognition in parts from single depth images. *Commun ACM*, 56(1), 116-124.
- Simon, S. R. (2004). Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems. *J Biomech*, 37(12), 1869-1880.
- Slobounov, S. M., Slobounova, E. S., & Newell, K. M. (1997). Virtual time-to-collision and human postural control. *J Motor Behav*, 29(3), 263-281.
- Stone, E. E., Butler, M., McRuer, A., Gray, A., Marks, J., & Skubic, M. (2013, July). Evaluation of the Microsoft Kinect for screening ACL injury. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (pp. 4152-4155). IEEE.
- Stone, E. E., & Skubic, M. (2011a, May). Evaluation of an inexpensive depth camera for passive in-home fall risk assessment. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on* (pp. 71-77). IEEE.
- Stone, E. E., & Skubic, M. (2011b, August). Passive in-home measurement of stride-to-stride gait variability comparing vision and Kinect sensing. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (pp. 6491-6494). IEEE.
- Stone, E. E., & Skubic, M. (2012, August). Capturing habitual, in-home gait parameter trends using an inexpensive depth camera. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (pp. 5106-5109). IEEE.

- Stone, E. E., & Skubic, M. (2013). Unobtrusive, Continuous, In-Home Gait Measurement Using the Microsoft Kinect. *IEEE transactions on bio-medical engineering*.
- Thiebaut, C., & Roques, S. (2005). Time-scale and time-frequency analyses of irregularly sampled astronomical time series. *EURASIP Journal on Advances in Signal Processing*, 15(1900), 2486-2499.
- Turvey, M. T. (1990). Coordination. *Am Psychol*, 45(8), 938.
- Uribe-Quevedo, A., Perez-Gutierrez, B., & Guerrero-Rincon, C. (2013, May). Seated Tracking for Correcting Computer Work Postures. In *Biomedical Engineering Conference (SBEC), 2013 29th Southern* (pp. 169-170). IEEE.
- Van Emmerik, R. E. A., McDermott, W. J., Haddad, J. M., & Van Wegen, E. E. H. (2005). Age-related changes in upper body adaptation to walking speed in human locomotion. *Gait Posture*, 22(3), 233-239.
- van Emmerik, R.E.A., & Wagenaar, R.C. (1996). Effects of walking velocity on relative phase dynamics in the trunk in human walking. *J Biomech*, 29(9), 1175-1184.
- van Emmerik, R. E., & van Wegen, E. E. (2002). On the functional aspects of variability in postural control. *Exercise Sport Sci R*, 30(4), 177-183.
- Venugopalan, J., Cheng, C., Stokes, T. H., & Wang, M. D. (2013, July). Kinect-based rehabilitation system for patients with traumatic brain injury. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (pp. 4625-4628). IEEE.
- Visser, J. E., Carpenter, M. G., van der Kooij, H., & Bloem, B. R. (2008). The clinical utility of posturography. *Clin Neurophysiol*, 119(11), 2424-2436.
- Volman, M. J. M., Laroy, M. E., & Jongmans, M. J. (2006). Rhythmic coordination of hand and foot in children with developmental coordination disorder. *Child Care, Hlth Dev*, 32(6), 693-702.
- Von Holst, E. (1973). *The behavioural physiology of animals and man: the collected papers of Erich von Holst* (Vol. 1). University of Miami Press.
- Wang, L., Weiming, H., & Tieniu, T. (2003). Recent developments in human motion analysis. *Pattern Recogn*, 36, 585-601.
- Weber, W., & Weber, E. F. (1836). *Mechanik der menschlichen Gehwerkzeuge: eine anatomisch-physiologische Untersuchung* (Vol. 1). Dietrich.
- Williams, A. M., Davids, K., & Williams, J. G. P. (Eds.). (1999). *Visual perception and action in sport*. Taylor & Francis.
- Winter, D. A. (2009). *Biomechanics and Motor Control of Human Movement* (4 ed. Vol.). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Young, W., Ferguson, S., Brault, S., & Craig, C. (2011). Assessing and training standing balance in older adults: A novel approach using the 'Nintendo Wii' Balance Board. *Gait Posture*, 33(2), 303-305.