



Alejandro Díaz Centeno

**Evaluation of physical-motor status of people with reduced
mobility using motion capture with Microsoft Kinect**

Dissertação de Mestrado

Dissertation presented to the Programa de Pós-Graduação em Informática of the Departamento de Informática, PUC-Rio, as partial fulfillment of the requirements for the degree of Mestre em Informática.

Advisor: Prof. Alberto Barbosa Raposo

Rio de Janeiro
September 2016



Alejandro Díaz Centeno

**Evaluation of physical-motor status of people with reduced
mobility using motion capture with Microsoft Kinect**

Dissertation presented to the Programa de Pós-Graduação em Informática of the Departamento de Informática do Centro Técnico Científico da PUC-Rio, as partial fulfillment of the requirements for the degree of Mestre.

Prof. Alberto Barbosa Raposo

Advisor

Departamento de Informática – PUC-Rio

Prof. Hugo Fuks

Departamento de Informática – PUC-Rio

Prof. Simone Diniz Junqueira Barbosa

Departamento de Informática – PUC-Rio

Prof. Márcio da Silveira Carvalho

Coordinator of the Centro Técnico Científico da PUC-Rio

Rio de Janeiro, September, 30th, 2016

All rights reserved

Alejandro Díaz Centeno

Graduated in Computer Science from University of Havana (UH), Havana - Cuba in 2011. He joined the Master in Informatics at Pontifical Catholic University of Rio de Janeiro (PUC-Rio) in 2014.

Bibliographic data

Díaz Centeno, Alejandro

Evaluation of physical-motor status of people with reduced mobility using motion capture with Microsoft Kinect/ Alejandro Díaz Centeno; advisor: Alberto Barbosa Raposo. – 2016.

65 f. : il. (color.) ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2016.

Inclui bibliografia

1. Informática – Teses. 2. Sensor 3D. 3. Tecnologias Assistivas. 4. Avaliação fisioterapêutica. I. Raposo, Alberto Barbosa. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Acknowledgments

I would like to thank my family, for their support and encouragement during all these years of study.

To my wife for her constant support and help during my studies.

To my advisor Alberto Barbosa Raposo, for their time, patience and dedication to the preparation of this work

To PUC-Rio, CNPq and FAPERJ for funding my research.

To the Sinapse clinic and his staff for all their support.

To all my classmates, professors, colleges and staff from the Department of Informatics and from the Tecgraf Institute. Thanks to all for your help and for always being so accommodating.

Abstract

Centeno, Alejandro Díaz; Raposo, Alberto Barbosa (Advisor). **Evaluation of physical-motor status of people with reduced mobility using motion capture with Microsoft Kinect.** Rio de Janeiro, 2016. 65p. MSc. Dissertation – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

The evaluation of the motor status of stroke patients and elderly people is done by using qualitative scales without standardization or measuring instruments. The scales are most common because they are relatively inexpensive and accessible, but suffer the disadvantage of being subjective, variable, and require prolonged training time. Moreover, assessment instruments, although more accurate and objective, have the problem of being heterogeneous, usually very expensive, and focused on specific goals. The rise in recent times of 3D sensors with high accuracy and low cost, some of them well known as the Microsoft Kinect, allows the use of motion analysis to quantify the deficit or success of a physiotherapeutic or drug treatment in a quantitative and standardized way, enabling the automatic comparison with standards of healthy people, and people with the same stage of disease, or similar characteristics. The aim of this work is to create a system using Microsoft Kinect for capturing and processing motor status of patients with reduced mobility in a non-invasive way, providing clinical feedback that allows the conduction of a quantitative and objective evaluation of patients, enabling monitoring of disease progression and reduced rehabilitation time.

Keywords

Microsoft Kinect; 3D sensor; Assistive Technologies; physiotherapeutic evaluation.

Resumo

Centeno, Alejandro Díaz; Raposo, Alberto Barbosa. **Avaliação do estado físico-motor de pessoas com mobilidade reduzida usando captura de movimento com o Microsoft Kinect**. Rio de Janeiro, 2016. 65p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Na atualidade, a avaliação do estado motor de pacientes com AVC e idosos é realizada mediante o uso de escalas qualitativas e sem padronização ou instrumentos de medição. As escalas são mais comuns por serem relativamente baratas e acessíveis, mas sofrem a desvantagem de serem subjetivas, variáveis, e requerem tempo de treinamento prolongado. Por outro lado, os instrumentos de avaliação, embora mais precisos e objetivos, têm o problema de serem heterogêneos, geralmente muito caros, e focados em objetivos específicos. O surgimento nos últimos tempos de sensores 3D com alta precisão e baixo custo, alguns deles bem conhecidos, como o Microsoft Kinect, permite a utilização de análise de movimento para quantificar o déficit ou sucesso de um tratamento fisioterapêutico ou medicamentoso, de forma quantitativa e padronizada, possibilitando também a comparação de forma automática com padrões de pessoas sãs, do mesmo estágio de doença, ou de características similares. O objetivo desse trabalho é criar um sistema usando o Microsoft Kinect para a captura e processamento do estado motor de forma não-invasiva, fornecendo feedback clínico que permita realizar uma avaliação quantitativa e objetiva de pacientes com mobilidade reduzida, permitindo um acompanhamento da evolução da doença e redução do tempo de reabilitação.

Palavras-chave

Microsoft Kinect; Sensor 3D; Tecnologias Assistivas; Avaliação fisioterapêutica.

Table of Contents

1	Introduction	11
1.1	Contributions	12
1.2	Dissertation Structure	12
2	Background.....	14
2.1	Physical therapy assessment	14
2.1.1.	Berg Balance Scale.....	15
2.2	Motion Capture	17
2.2.1.	Marker-based Tracking.....	17
2.2.2.	Markerless Tracking.....	18
2.2.3.	Image Acquisition	18
2.3	Related Work	19
3	Methodology.....	22
3.1	Aim of the work.....	22
3.2	Materials	22
3.3	Set of Exercises.....	23
3.3.1.	Kinect Evaluation Exercises Set.....	23
3.4	Data Acquisition and Processing	25
3.4.1.	Microsoft Kinect Specifications	25
3.4.2.	Skeleton Coordinate Calibration	26
3.4.3.	Skeleton Scaling	28

3.4.4.	Computing the angle and velocity	30
3.4.5.	Constructing the dataset	31
3.4.6.	Dynamic Time Wrapping	31
4	Implementation	33
4.1	Architecture	33
4.2	User Interface.....	36
4.2.1.	Data Acquisition Module	37
4.2.2.	Editing Module	40
4.2.3.	Analysis and Visualization Module.....	41
5	Evaluation and Result.....	43
5.1	Experiment to find the dataset of healthy subjects	43
5.2	Evaluation using the Kinect-based system and the Berg Balance Scale	45
5.2.1.	Clinical scale assessment.....	46
5.2.2.	Kinect-based assessment	46
6	Conclusions	54
7	Bibliography	56
8	Annex A.....	60
A.1	Berg Balance Scale	60

List of Figures

Figure 1 - Kinect Evaluation Exercises Set	24
Figure 2 - Kinect Skeleton.....	26
Figure 3 - Kinect skeleton space.....	28
Figure 4 - Assessment System.....	34
Figure 5 - Data Acquisition Module.....	34
Figure 6 - Home Screen.....	36
Figure 7 - Exercise selection screen	37
Figure 8 - Exercise tutorial screen.....	38
Figure 9 - Calibration screen	38
Figure 10 - Ready to record screen.....	39
Figure 11 - Recording screen.....	39
Figure 12 - After record screen.....	40
Figure 13 - Editing Module screen.....	41
Figure 14 - Analysis and visualization screen	42
Figure 15 - Dataset vs ten normal subjects - axis X.....	44
Figure 16 - Dataset vs 10 normal subjects - axis Y	44
Figure 17 - Dataset vs 10 normal subjects - axis Z	45
Figure 18 - Dataset vs 5 unhealthy subjects - axis X.....	47
Figure 19 - Dataset vs 5 unhealthy subjects - axis Y.....	48
Figure 20 - Dataset vs 5 unhealthy subjects - axis Z.....	48
Figure 21 - Left Knee dataset average vs unhealthy subjects X-axis.....	49
Figure 22 - Subjects with low fall risk vs the dataset average X-axis.....	50
Figure 23 - Subjects with medium fall risk vs the norm X-axis.....	51
Figure 24 - Right Hand dataset average vs unhealthy subjects Z-axis.....	52
Figure 25 - Subjects with low fall risk vs the dataset average Z-axis	52
Figure 26 - Subjects with medium fall risk vs the dataset average Z-axis	53

List of Tables

Table 1 - Kinect XBOX 360 (Version 1) vs Kinect XBOX ONE (Version 2)	25
Table 2 - Participants' Characteristics	44
Table 3 - Participants' Characteristics	45
Table 4 - Berg Scale Results	46
Table 5 - Dynamic Time Wrapping cost- Standing unsupported with eyes closed .	48
Table 6 - Dynamic Time Wrapping cost- Standing on one leg	49
Table 7 - Dynamic Time Wrapping cost - Reaching forward with outstretched arms	53

1 Introduction

In recent years, life expectancy has been increasing steadily. Thus, motor deficits and reduced mobility are becoming more prevalent in our society, either due to advanced age or diseases of neurological origin. Virtual reality technologies already commercialized with recreational purposes, for example, for videogames are being used for the rehabilitation of stroke, Parkinson's disease, and patients with motor deficiencies, showing positive results (González, 2014) (Paavola, 2013) (Saposnik, 2010) (Stone, 2013).

However, we still find some loopholes in the use of such technology for rehabilitation of patients with motor deficits because this is not for what these technologies were developed. A simple example is the lack of applications of this technology to evaluate patients and their motor deficits. The correct evaluation is an essential part of rehabilitation. Recovery depends on an efficient evaluation to identify gaps and thus provide appropriate treatment (Gavim, 2012).

Currently, the assessment of patients' motor state is carried out by using qualitative scales and without standardization or measuring instruments. The scales are most common because they are relatively inexpensive and accessible, but suffer the disadvantage of being subjective, variable and require prolonged training time. Moreover, assessment instruments, although more accurate and objective, had the problem of being heterogeneous, usually very expensive and focused on specific goals.

The emergence of the 3D sensor with high accuracy and low cost, some of them well known as the Microsoft Kinect, brought the use of motion analysis to quantify the deficit or improvement of a physical therapy or drug treatment, in a quantitative and standardized way. They also allow the automatic comparison with standards of healthy people, people with the same stage of disease, or similar features.

This work was conducted in partnership with specialists in physical therapy of the SINAPSE clinic (Sinapse, 2016), Rio de Janeiro. The staff of the clinic gave the motivation for this work, the medical advice and participate actively in the construction of the solution.

1.1 Contributions

The main contribution of this work is to create a system using Microsoft Kinect for capturing and processing the motor state of patients, non-invasively, providing tools for the specialists in physical therapy that allow them to conduct a quantitative and objective evaluation of patients with reduced mobility, improving the usual assessment method based on scales.

We defined a Kinect evaluation exercises set based on a subset of the activities of the Berg Balance Scale, an accepted standard clinical criterion by which to measure balance (Jácome, 2016) (Saso, 2016) (Langley, 2007). We create a dataset for each exercise. We propose a system in which a specialist in physical therapy can record the performance of patients during the execution of this set of exercises, the system also provides tools to measure position, angle and velocity between different points of the body detected with Microsoft Kinect. These measures can also be compared with the dataset using dynamic time warping algorithm (Wang, 2013).

The experiments show that with a system like this, specialists in physical therapy can obtain more information about the patient's performance, not reachable using standard scales.

1.2 Dissertation Structure

This dissertation is structured as follows. Chapter 2 presents the basic concepts and summarizes related work. Chapter 3 introduces the methodology used in the research, describing methods, algorithms and materials. Chapter 4 describes the implementation of the system to capture and process the motor state of users. Chapter

5 exposes and discusses the results obtained in experiments using the software. Finally, Chapter 6 contains the conclusions and proposes future work.

2 Background

This chapter provides an overview of the main concepts related to this dissertation. Section 2.1 covers basic concepts about physical therapy assessment. Section 2.2 addresses motion capture definition and techniques. Finally, Section 2.3 presents additional related work.

2.1 Physical therapy assessment

Falls are a major contributor to injury with subsequent disability in people with motor deficits and reduced mobility, and poor balance is associated with increased risk of injurious falls. Falls are a threat to the quality of life and independence that is secondary to impaired mobility and loss of function. Emphasis on early identification for prevention and intervention in individuals who are at risk of falls is important in physical therapy and rehabilitation (Downs, 2014).

A plethora of clinical scales designed to assess balance has been developed over the last years. Traditionally, a great portion of physical therapy assessment of patients is based on a therapist's observation and judgment. Clinical scales consist of a variety of surveys, trials, and questionnaires that are administered to patients to capture basic assessment of their impairments, residual functional abilities, and daily life activities. The assessment methods, such as Berg Balance Scale (Berg, 1992)(Annex A1), Tinetti Test (Tinetti, 1986), Timed Up, and Go test (Podsiadlo, 1991), strongly rely on the therapist's visual assessment of how the patient is performing a standard task. This process needs a trained physical therapist or occupational therapist to spend some time with the patient. The assessment can be inaccurate for several reasons, one of which is the subjectivity of these behavioral and clinical assessments. Based on extensive clinical use and frequency of comparison with other balance measures in

both clinic and laboratory, and research settings, the Berg Balance Scale (BBS) is accepted as the standard clinical criterion by which to measure balance (Jácome, 2016) (Saso, 2016) (Langley, 2007).

2.1.1. Berg Balance Scale

The Berg Balance Scale (BBS) is intended to measure a subject's ability to maintain balance while doing functional tasks and has previously shown good measurement properties (Adam, 2016). It is based on the principle that a person's balance is challenged by diminishing his/her base of support. The test has 14 items that are common in everyday life. These tasks progress from sitting to comfortable standing, to standing with feet together, and, finally, to standing with one foot in front of the other and single-leg standing. Each item of the BBS is rated on a 5-point scale. Lower ratings are given for supervision, cueing, assistance, and/or whether time or distance requirements are met (Berg, 1992). When used as a standardized assessment, a cutoff score of 45/56 is an established criterion. When used to evaluate the risk of falling, the same cutoff score has the greatest specificity in identifying people who do not fall (Riddle, 1999). When calculating a total score for the BBS, a point of concern is that test items have different operational definitions for each of their rating categories. The BBS is detailed in the Annex A1. For example, for the item "transfer," a rating category of 2 indicates "able to transfer with verbal cueing or supervision," whereas for the item "standing on one leg" the same rating category of 2 indicates "able to lift leg independently and hold equal to or greater than three seconds." From this example, a question arises about the appropriateness of combining scores from ratings that have different meanings. A second concern with the rating scale is that a "passing" score on each item is not always the same numeric value. Although the BBS does not specifically define passing for each item, it could be defined as successful completion of the task, independent of qualifiers—for example, time and speed. When this definition of passing is used for the item "standing on one leg," a rating category of 2 must be achieved to pass successfully

(“able to lift leg independently and hold equal to or greater than 3 seconds”). In contrast, for the item “sitting unsupported,” a rating category of 4 must be achieved to pass (“able to sit safely and securely 2 minutes”). Using this interpretative framework, more points are earned by passing the item “sitting unsupported” than would be earned by passing the item “standing on one leg.” With this definition of passing, all items should share the same numeric rating. In addition to challenges in interpreting the rating scales, similar challenges arise in interpreting total BBS scores. Although it is well established that a BBS score of 45/56 or greater indicates a low probability of falls, it provides limited information about the functional capability of the patient. For example, a total score of 45/56 does not tell us whether a person can sit unsupported (the easiest item on the instrument) or stand on one foot (the most difficult item). Ironically, although the patient’s score indicates the probability of falls, it fails to describe the patient in functional terms. The total score provides limited information with which to guide clinical intervention (Diana, 2004). The development and implementation of effective and cost-efficient strategies to evaluate fall risk in patients with motor deficits and reduced mobility is, therefore, an urgent challenge for health care.

Clinical scales do not quantitatively determine the level of the impairment, but qualitatively associate a score with the patient’s performance. Therefore, despite providing a general picture of the patient clinical state, clinical scales alone, due to roof and floor effects, are not always sensible enough to determine the patients’ different functional profiles. Further, clinical scales reliability is partially dependent from the interobserver and intraobserver variability and, again, the sensibility is affected. Therefore, they often detect only remarkable improvements or aggravations (Jennifer, 2013).

The advancement of computer technology and the development of affordable devices capable of recording human body’s kinematic data with high accuracy and reliability enables research and development of tools in the physiotherapy area. The analysis of motion capture data results in better clinical assessment and more efficient therapeutic decision making accordingly.

2.2 Motion Capture

A motion capture system is capable of generating real-time information from the user's position. It can consist of one or more cameras for obtaining environmental information, processing the received data, and extracting features of interest, for example, joint position (Zhou, 2008). These types of systems can be divided into two groups: marker-based and markerless.

2.2.1. Marker-based Tracking

Marker-based tracking is a technique that applies markers in specific body areas, following the movement of these regions. Since the human body is highly articulated, it is sometimes difficult to correctly estimate the current position of each joint. Also, the image analysis is also dependent on the environment where the user is located, which may result in additional errors due to background noise. The use of markers decreases the difficulty of the estimation of position, as well as facilitates the detection and monitoring of movements, resulting in a need for lower processing. However, this technique requires that the user places the markers on specific points of the body, which sometimes might be something not feasible. There are also other problems resulting from the use of markers, such as the fact that they can accidentally move, giving rise to errors and variations in the readings obtained (Zhou, 2008).

These systems can be passive or active. A passive system uses a set of markers that do not produce light, only reflect the coming light against them. Systems of this kind can generate 3D views of the human body, from 2D information. In active systems, the markers produce light, for example infrared, which is then captured by the system. The advantage of these systems compared to passive ones is that they can achieve accuracies, for the position of the markers, in the order of millimeters (Zhou, 2008).

2.2.2. Markerless Tracking

High-resolution and high-speed cameras can be used to perform real-time motion detection, without the need to appeal to markers. The problem of tracking technique without the use of markers is that it requires a more sophisticated algorithm, comparing with the marker-based tracking, resulting in a greater need of processing power. This technique, like the previous one, has the problem of background noise, which may dramatically influence the final result. The markerless tracking method can be divided into two groups:

2D: Is a low-cost processing technique since it only takes into account the image plane. It makes use of a single camera for acquisition by reducing the amount of information available. However, this aspect can be limiting when it aims to analyze movements of the human body because they take place in three-dimensional spaces.

3D: These systems make 3D acquisition of the human body, providing another dimension to the analysis of movements. They then become more reliable systems when compared with 2D systems.

Comparing the markerless with the marker-based tracking methods, the first category brings many advantages for the user as it eliminates the need for physical markers. In the context of physical therapy assessment, this is certainly a point in favor, since there may be cases where it is impossible for the patient, due to locomotor limitations, to put on the markers. Simultaneously, this option would require knowledge of the patient, the anatomy of the human body, since the markers must be placed in very specific anatomical points to obtain a reliable reading of the acquired data (Zhou, 2008).

2.2.3. Image Acquisition

A conventional device for image acquisition generates files of type Red, Green, Blue (RGB) for later viewing and processing. This acronym represents the additive color space composed of the three primary colors. The file generated in each image pixel

contains intensity information of three values that when added produce a color of the visible light spectrum. Typical RGB devices are video cameras, scanners, and digital cameras. This format is most commonly used in image acquisition and processing systems since the equipment is quite affordable.

Another type of image acquisition devices are cameras Red, Green, Blue, Depth (RGB-D). These devices are capable of, in addition to capturing images in RGB format, obtaining the depth information of the pixels. There are two main methods for obtaining the depth information: active stereo or time-of-flight (Henry, 2012). The first one uses the overlapping of two images generated at the same time and in the same environment, with a relative displacement to each other, which are later combined. The second calculates the time that electromagnetic waves, generated by a transmitter, takes to travel a given distance. Both allow obtaining a spatial reading of the environment in which they are located (i.e., the distance values). These values are then related to the image in RGB format and a mapping between a certain pixel and its distance from the sensor is performed. The time-of-flight technique provides reasonably accurate information, and systems have a high acquisition speed. Some disadvantages are the detection limit, typically up to about 5m, noise contained in the readings, and their restricted field of view (Henry, 2012).

Among these devices, one of the most used and cited in the literature is the Microsoft Kinect (Mobini, 2014), a motion capture device based in markerless tracking, and in this work, we are focused on the utilization of the motion capture data obtained from the Kinect for the assessment of the motor state of physiotherapeutic patients.

2.3 Related Work

Motion capture using Kinect has become increasingly popular in physical therapy and rehabilitation (Wai, 2014); hence understanding limitations of the Kinect sensor is important. In the current section, we discuss the work focused on evaluating Kinect as a robust and reliable sensor.

There have been several attempts to evaluate the Kinect's measurements quantitatively. Obdrzalek et al. (2012), Mobini et al. (2014), and Fernández-Baena et al. (2012) investigated the accuracy of Kinect's joint tracking, specifically whether Kinect's joint localization and pose estimation is robust and reliable. For example, Obdrzalek et al. (2012) specified 6 physical exercises to examine pose accuracy of Kinect. The results show that Kinect is a good option to be used as an online motion capture device because of its low price. However, the Kinect skeleton tracking suffers from occlusion or having objects such as chairs in the scene (the chair's leg can be detected as an individual's leg in the seated mode). Therefore, developers should consider addressing issues such as occlusion, self-occlusion, and unconventional body postures or use of wheelchair/walkers. Mobini et al. (2014) first used a fabricated model of the upper body. Then, they estimated the displacement between various joints by Kinect and compared them with the actual values from the model. Fernández-Baena et al. (2012) compared precision in the computation of joint angles between Kinect and Vicon¹ which is one of the expensive commercial IR trigonometry motion capture systems. Based on the obtained results, Kinect is precise enough for most of the clinical rehabilitation treatments. In a similar attempt, Dutta (2012) compared Kinect with Vicon to investigate whether Kinect is sensitive enough to be used as a 3D motion capture device. Stone and Skubic (2011) validated Kinect's ability in the elderly fall monitoring and compared it against Vicon measurements, and showed that Kinect provides acceptable accuracy.

Kurillo et al. (2013) examined Kinect's accuracy by measuring the upper extremity's reachable workspace of 10 healthy individuals; movements were recorded using Kinect and an IR marker-based motion capture system simultaneously. The results showed that the Kinect-based system provides sufficient accuracy and reliable results compared to the motion capture system.

Bonnechère et al. (2014) validated range of motion (ROM) measurements using the Kinect with concurrent measurement performed by traditional marker-based stereophotogrammetry system; 48 volunteers were asked to perform shoulder

¹ <https://www.vicon.com/>

abduction, elbow flexion, hip abduction, and knee flexion motions in two sessions. Kinect and the maker-based system showed similar statistical trends in the recorded data but in some cases the measured ROMs were different. Clark et al. (2013) showed that through calibration with the 3D motion capture system, Kinect yields a significantly better accuracy.

The overall lesson from these studies is that Kinect is an acceptable and affordable depth sensor for physiotherapeutic purposes.

3 Methodology

This chapter describes the methodology for conducting the research related to this dissertation. Section 3.1 covers the aim of this research. Section 3.2 presents the materials used in the development. Section 3.2 describes the set of exercises for measuring the motion performance and Section 3.4 exposes the main ideas about the acquisition and processing of the data.

3.1 Aim of the work

The purpose of this work is to develop a system for capturing and processing the motor state of elderly and post stroke patients with reduced mobility using the Microsoft Kinect, with the expectation of using it to help the specialists in physical therapy to perform a quantitative and objective evaluation of those patients.

To achieve that goal, we constructed a dataset that describes the motion of a healthy subject, using the Kinect as the tracking sensor. It is intended that the main functionalities of the system allow comparing unhealthy subjects with that dataset of healthy subjects and with themselves along the time. Also, it is aimed to test the coherence between the Kinect-based assessment and standardized clinical assessment.

3.2 Materials

We used Microsoft Kinect for Xbox One with SDK 2.0. The system was written in C# using Unity3D (Unity, 2016) and Microsoft Visual Studio (Microsoft, 2016). The system is divided into three modules: Data Acquisition Module, Data Processing

Module, and Data Visualization Module. The specifications of the implementation will be described in the next Chapter.

This work was conducted in PUC-Rio in partnership with specialists in physical therapy of the SINAPSE clinic (Sinapse, 2016), located in Gávea, Rio de Janeiro. The staff of the clinic gave medical advice and allowed to use the space of the clinic to perform the tests.

3.3 Set of Exercises

In collaboration with the specialists in physical therapy, a set of tasks was designed, which allow measuring the motion performance of the participants, based on the exercises of the Berg Balance Scale described in Annex A1. It was not possible to include those activities of the Berg Balance Scale that have occlusion of joints because the Microsoft Kinect is not able to track those joints. For that reason, it was not possible to include in the set activities like "Turning to look behind", "Retrieving object from floor" and "Turning 360 degrees". The specialists in physical therapy corroborated that removing these exercises do not affect the patient's balance analysis. The resulting set is formalized below (Figure 1)

3.3.1. Kinect Evaluation Exercises Set

Before beginning the Kinect Evaluation Exercises Set (KEES) the subjects should understand that they must maintain their balance while attempting the tasks. The choices of which leg to stand on or how far to reach are left to the subject. The chair used during testing should have a reasonable height and must have arms. Either a step or a stool of average step height may be used for item 10.

1. **Sitting unsupported:** The subject must sit safely and securely with arms folded for 2 minutes.
2. **Sitting to standing:** The subject must stand up, trying not to use the hands for support and stabilize independently.

3. **Standing unsupported:** The subject must stand for two minutes without holding on.
4. **Standing to sitting:** The subject must sit down safely with the minimal use of hands.
5. **Standing unsupported with eyes closed:** The subject must close the eyes and stand still for 10 seconds safely.
6. **Standing unsupported with feet together:** The subject must place the feet together and stand without holding on one minute.
7. **Standing with one foot in front:** The subject must place one foot directly in front of the other. If the subject cannot put the foot directly in front, s/he must try to step far enough ahead that the heel of the forward foot is ahead of the toes of the other foot.
8. **Standing on one leg:** The subject must stand on one leg without holding on for 10 seconds.
9. **Reaching forward with outstretched arms:** The subject must lift the arms to 90 degrees, with the palms of the hands towards the Kinect. Stretch out the arm and reach forward as far as he can.
10. **Placing alternate foot on a stool:** The subject must place each foot alternately on the step/stool; each foot should touch the step/stool four times. The subject must try to complete eight steps in 20 seconds.

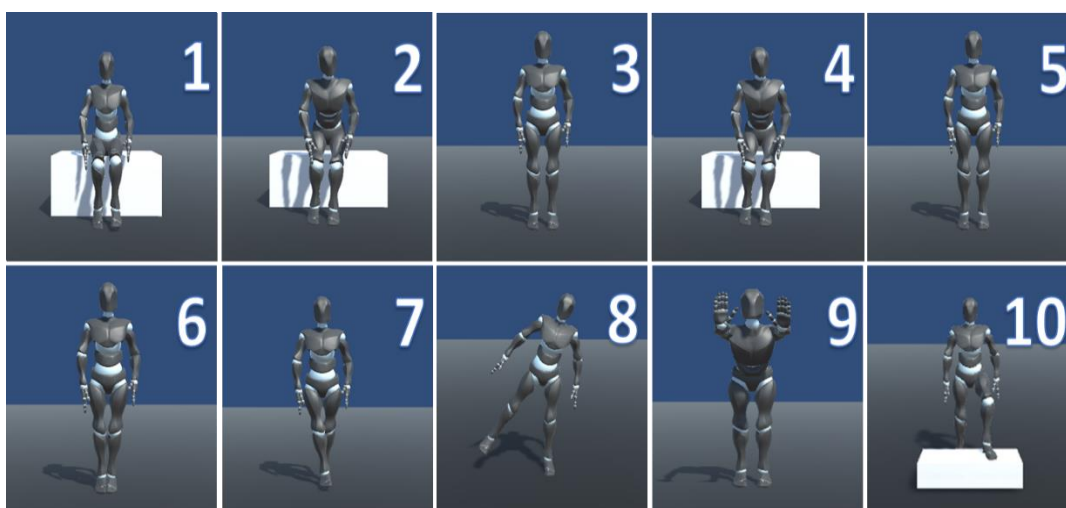


Figure 1 - Kinect Evaluation Exercises Set

3.4 Data Acquisition and Processing

For the data acquisition, users are recorded in front of the Microsoft Kinect performing the KEES. Before each exercise skeleton calibration parameters are obtained, which allow each user to be in the same coordinate system.

3.4.1. Microsoft Kinect Specifications

The Microsoft Kinect is a camera-based sensor mainly used to control computer games directly through body movement. Kinect tracks the position of the limbs and body, without the need for hand controllers or force platforms. The use of a depth sensor also allows Kinect to capture the three-dimensional movement patterns. Kinect was originally built for Xbox 360 video game console, but in September 2014 Microsoft launched a second standalone version of Kinect for Xbox One, introducing some improvements (Table 1), including: tracking up to 25 anatomically correct body points and orientations, a higher resolution, better accuracy, among other improvements.

	Version 1	Version 2
Depth range	0.4m → 4.0m	0.5m → 4.5m
Color stream	640×480@30fps	1920×1080@30fps
Depth stream	320×240	512×424
Infrared stream	None	512×424
Type of Light	Light coding	ToF
Audio stream	4-mic array 16 kHz	4-mic array 48 kHz
USB	2.0	3.0
# Bodies Traked	2 (+4)	6
# Joints	20	25
Hand Traking	External tools	Yes
Face Traking	Yes	Yes+Expressions
FOV	57° H 43° V	70° H 60° V
Tilt	Motorized	Manual

Table 1 - Kinect XBOX 360 (Version 1) vs Kinect XBOX ONE (Version 2)

3.4.2. Skeleton Coordinate Calibration

When recording their movements, users may get very different coordinates data for the skeleton positions, even if they make the same movement. Each user has different physical characteristics, such as size, and the length of body parts, and everybody is not going to be in the same position in front of the camera at recording time. It is hard to compare two user motions using these data. It is necessary to make a calibration of the user's skeleton recognized by Kinect to normalize the data.

Wei (2014) proposed an algorithm where they pull the user's skeleton to the center of the Kinect sensor and then rotate it to face the Kinect sensor, and the algorithm treats this position as the initial position in a universal coordinate system. In this way, all users' initial positions are normalized in the universal coordinate system, irrespective of their original standing position and angle. The calibration system in this work is based on that algorithm.

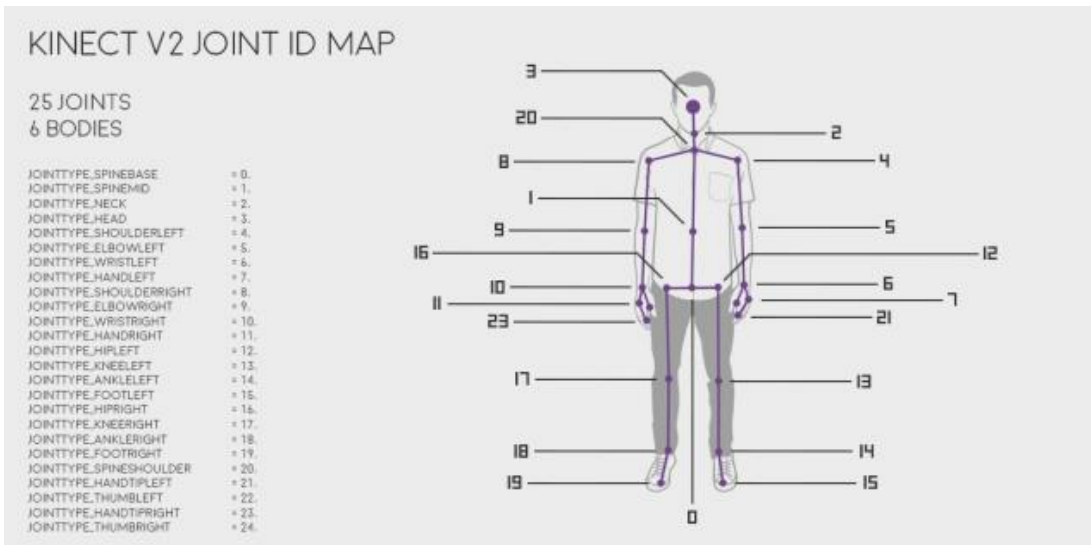


Figure 2 - Kinect Skeleton

Microsoft Kinect skeleton data contain 3D position data for human skeletons (Figure 2). Each joint position in the skeleton space is represented as (x, y, z) . The skeleton space coordinates are expressed in meters. As illustrated in Figure 3, it is a right-hand coordinate system that places a Kinect sensor at the origin. More specifically, the positive x-axis extends to the left of the Kinect, and the positive y-axis extends

upward. In order to calibrate the initial position of the user, the system collects the first 120 frames as initial data. Since the Kinect produces about 30 frames data per second, it suggests that the user remains standing still around four seconds. These initial data are used to calculate the initial angle θ and the initial center coordinate vector $C(X_c, Y_c, Z_c)$.

These two results are the foundation of the calibration. Then, any joint coordinates can be transformed and rotated to a coordinate system that places the initial center at the origin. The calibration process can be regarded as pulling the original skeleton to the center of the Kinect sensor, and rotating it to face the Kinect sensor. This is the initial position in an universal coordinate system. Consequently, all users' initial positions are the same in the coordinate system, wherever they stand at. After reconstruction in the universal coordinate system, all movements of the skeleton are in the universal coordinate system. Firstly, all joints are transformed to the origin of the universal coordinate system according to the initial center coordinate:

$$P_j(X_p, Y_p, Z_p) = (X_j - X_c, Y_j - Y_c, Z_j - Z_c) \quad j = 0 \dots 24$$

Where X_j, Y_j, Z_j are coordinates of joint j ; X_c, Y_c, Z_c are coordinates of the initial center.

Secondly, a quaternion rotation is used to rotate the coordinate vector $P_j(X_p, Y_p, Z_p)$ about the y-axis of the initial center by the initial angle θ . The quaternion rotation is a right handed rotation. The thumb points the direction of unit rotation axis vector R which is the y-axis of the initial center.

$$R = (X_R, Y_R, Z_R)$$

where $\|R\| = 1$. Thus, $X_R = 0$, $Y_R = 1$ and $Z_R = 1$. The rotation quaternion is defined to be:

$$Q = \cos\left(\frac{\theta}{2}\right) + X_R \sin\left(\frac{\theta}{2}\right)i + Y_R \sin\left(\frac{\theta}{2}\right)j + Z_R \sin\left(\frac{\theta}{2}\right)k$$

Finally, the coordinate of the new point is:

$$N_{PR} = (X_p \cos \theta + Z_p \sin \theta, Y_p, Z_p \cos \theta - X_p \sin \theta)$$

The system can utilize the calculated initial angle and initial center coordinate as long as the Kinect remains in the same position.



Figure 3 - Kinect skeleton space

3.4.3. Skeleton Scaling

After we have the skeleton in the right position, we need to scale it properly to a model body to be able to make any further comparison or to extract any standard measures. In every frame a factor will be used to scale the subject skeleton to the model, calculating the new positions of the joints. For each joint position $P1 = (x1, y1, z1)$ the scale factor will be the vector $F = (Fx, Fy, Fz)$, where each component Fx , Fy , and Fz are the scale factors for each component of the point $P1$. Once factor F is known, a scaling matrix Fm is built.

$$Fm = \begin{pmatrix} fx & 0 & 0 \\ 0 & fy & 0 \\ 0 & 0 & fz \end{pmatrix}$$

Then to obtain the new point scaled the scaling matrix is multiplied by the vector of the original point.

$$P2 = Fm * P1 = \begin{pmatrix} fx & 0 & 0 \\ 0 & fy & 0 \\ 0 & 0 & fz \end{pmatrix} \begin{pmatrix} x1 \\ y1 \\ z1 \end{pmatrix} = \begin{pmatrix} fx * x1 \\ fy * y1 \\ fz * z1 \end{pmatrix}$$

Since the proportions between the center and every joint of the skeleton are different, it is necessary to find the particular scaling matrix Fm for every joint. For that reason,

it is necessary to perform a setting step before the scaling process. In the configuration phase, the subject is recorded in standing position for four seconds, and the collected data is used to calculate the scale factor to transform the skeleton of the subject to a model skeleton recorded previously in the same position. For each joint position $P1 = (X1, Y1, Z1)$, the transformation factor F is calculated, for each dimension, to scale to the position $P2 = (X2, Y2, Z2)$ of the corresponding joint in the model skeleton. In order to find that factor the following algorithm was performed.

For both points $P1$ and $P2$, and for each dimension X , Y and Z of each, the Delta values from the center of the body $C = (Cx, Cy, Cz)$ to the points are calculated. In this moment of the capturing data process, both skeletons have the center of the body $C = (0,0,0)$, due to the calibration phase that took place before.

This is the process for calculating the scale factor for one joint:

Delta values for $P1 = (X1, Y1, Z1)$

$$D1 = |C - P1|$$

$$dx1 = |Xc - X1|$$

$$dy1 = |Yc - Y1|$$

$$dz1 = |Zc - Z1|$$

Delta values for $P2 = (X2, Y2, Z2)$

$$D2 = |C - P2|$$

$$dx2 = |Xc - X2|$$

$$dy2 = |Yc - Y2|$$

$$dz2 = |Zc - Z2|$$

Then the scale factor is calculated by dividing $D1$ by $D2$ as follows.

$$F = D1/D2 = (Fx, Fy, Fz)$$

$$F_x = dx_1/dx_2$$

$$F_y = dy_1/dy_2$$

$$F_z = dz_1/dz_2$$

Since this process is repeated for each joint, the result will be a set of 25 scale vectors $SF = F_1, F_2, \dots, F_{25}$, one for each joint.

3.4.4. Computing the angle and velocity

Assuming that the positions of the joints returned by the sensor are correct, the angle for each exercise is obtained by a simple calculation. Having three consecutive joints (e.g. hip, knee, and foot, respectively) define the vector between the joints 1 and 2, and the vector between 2 and 3. After that, calculate the scalar product of these two vectors and the product of its norms. Dividing these two values, and calculating the arc cosine, the angle between the three joints is obtained. The equation used is:

$$\alpha = \arccos\left(\frac{a \cdot b}{\|a\| \cdot \|b\|}\right)$$

where α represents the angle, a and b are vectors connecting joints 1 and 2, 2 and 3, respectively.

Having the value of the angle, the calculation of the angular velocity is straightforward. We chose the angular speed rather than linear speed because the result will be more intuitive, and thus there is only one value to show. To calculate the angular velocity, it is only necessary to subtract the current angle from the previous one and divide the difference of time between each measure. Since Kinect returns 30 frames per second, the time between each calculated value is 0.033s. The formula is:

$$v = \frac{\alpha_f - \alpha_i}{0.33}$$

Where v is the angular velocity, α_i is initial angle (in one frame) and α_f is the final angle (in the next frame).

3.4.5. Constructing the dataset

Each activity of the KEES will have a normalized value associated, i.e., the “correct” way to perform the activity. To construct the dataset ten healthy subjects were selected and recorded while performing each exercise. The subjects enrolled were six males and four females, with good mobility, and 27 ± 5 years. To obtain the average value of one joint j at a specific frame of the exercise the mean of the ten values of that joint at that instant is calculated.

$$Pnorm_j = \frac{1}{10} \sum_{k=1}^{10} (X_{c_j}, Y_{c_j}, Z_{c_j})_k \quad j = 0 \dots 24$$

where $(X_{c_j}, Y_{c_j}, Z_{c_j})_k$ means the position of joint j of subject k at that instant.

Thus, the average of an instant of time of one exercise will be the set of normalized joints of the skeleton at that moment m .

$$Nm = Pnorm_0, \dots, Pnorm_{24}$$

Therefore, the average of the complete exercise will be the set of normalized joints values in each moment of the exercise.

$$N = N_0, \dots, N_t$$

Where t is the number of frames captured for that exercise.

3.4.6. Dynamic Time Wrapping

All the data collected by the system are time series. In order to compare them it is necessary to use some algorithm that compares time series, based on similarity. In this way we can determine quantitatively when an exercise done by some user of the system resembles the dataset average.

Dynamic time warping (DTW) (Wang, 2013) is an algorithm for measuring similarity between two temporal sequences which may vary in speed. For instance, similarities

in walking could be detected using DTW, even if one person was walking faster than the other, or if there were accelerations and decelerations during the course of an observation. DTW has been applied to temporal sequences of video, audio, and graphics data, indeed, any data which can be turned into a linear sequence can be analyzed with DTW (Wang, 2013). The objective of DTW is to compare two (time-dependent) sequences $X := (x_1, x_2, \dots, x_N)$ of length $N \in \mathbb{N}$ and $Y := (y_1, y_2, \dots, y_M)$ of length $M \in \mathbb{N}$. These sequences may be discrete signals (time-series) or, more generally, feature sequences sampled at equidistant points in time.

4 Implementation

This chapter summarizes some of the details of the implementation of the application constructed to enable users to capture, process and analyze motion data of user doing the KEES. Section 4.1 describes the architecture of the application. Section 4.2 presents details about the user interface and exposes the functionality of each module that compose the system.

4.1 Architecture

The system was written in C# using Unity3D (Unity, 2016) and Microsoft Visual Studio (Microsoft, 2016). It is divided into three modules: Data Acquisition Module, Editing Module, and Analyses and Visualization Module (Figure 4).

The *Acquisition Module* is responsible for the recording of patients' data while they perform the evaluation exercises. Microsoft Kinect for Xbox One with SDK 2.0 was used to record the data. This module provides to the specialist options for start recording, stop recording, redo the exercise or save the recorded data. Also, it provides an explanation for the patient about how to do the exercise that is displayed before beginning to record.

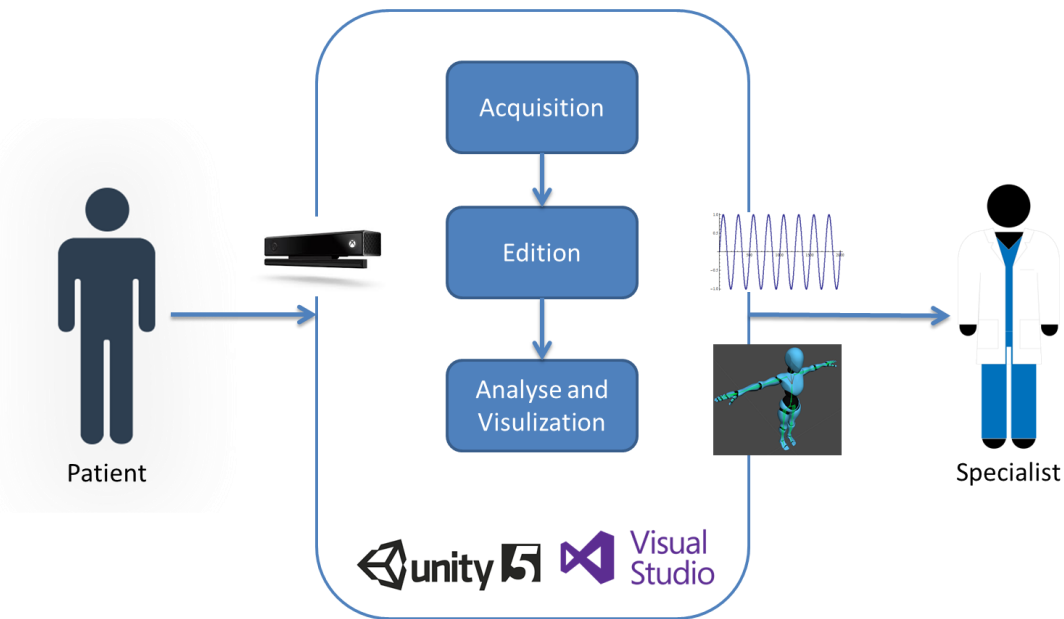


Figure 4 - Assessment System

Internally the acquisition process can be divided into four subprocesses, Data Recording, Skeleton Calibration, Skeleton Scaling, and Save the Data. The subprocesses are executed in the order presented in Figure 5. The objective is to prepare the data for the next phases.

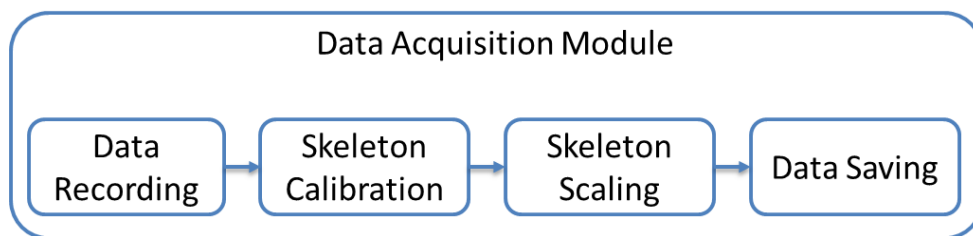


Figure 5 - Data Acquisition Module

The Data Recording process is made using the Unity Package for Kinect. The main class in the package is called *KinectSensor*, and represents the Kinect. At each frame, *KinectSensor* provides a “Body” that is an abstraction of the data that the sensor captures. That data includes a set of joints of the skeleton, their positions and orientation. We discuss the main classes involved in the process below.

We implemented the class *Exercise* that represents all the data of an evaluation exercise. It has the field *SkeletonData* that is a list of bodies obtained during the exercise. It also contains the field *CalibrationData* that has the values of Center and Angle necessary to calibrate the skeleton. *Exercise* contains another field called *ScaleData* that is a list of the 25 scale vectors, one for each joint. It also contains a function *Save* to store all the data of an Exercise.

The improvement from the Kinect Xbox 360 to the Kinect Xbox One brought improvements in all specifications of the Kinect, better camera, and infrared, more accurate depth map, which leads to a better definition of the skeleton obtained, but in consequence the device generates a huge amount of data, for example, all the information generated by the device within 10 seconds is about one gigabyte of information which makes it difficult the treatment exercises with minutes' durations. For this reason, the system only stores for each frame the information regarding the twenty-five joint that represents parts of the skeleton, significantly reducing the size of the information generated (the same 10 seconds is stored in only a few kilobytes). All reproduction and analysis of exercises and avatar animations is made based on this information.

The class *Recorder* is responsible for getting the data from *KinectSensor* and for creating an Exercise. Recorder has a function *Obtain()* that gets the *Body* of each frame, and adds it to the *SkeletonData* of the class *Exercise*. *Recorder* also contain two important functions *GetCalibrationData(List<Body> toCalibrate)* that computes the calibration angle and center that is used to fill the *CalibrationData* of the *Exercise*, and *GetScaleData(Body toScale)* that computes the scale vector of the joints.

In the *Edition Module*, the specialist can upload a recording of an exercise and make a temporal adjustment by modifying the interval of frames (or time) to be analyzed.

This module presents to the specialist an animation of the exercise done by the patient. To do that the class *Player* was implemented that uses the *Exercise* data to animate the avatar. If the specialist trims the recorded data another function of the class *Exercise* is used, *Trim(int initPos, int endPos)* that deletes all the *Bodies* from

SkeletonData from position 0 to *initPos* and from *endPos* to the *Length -1* position, and the *Exercise* is saved again.

The *Analysis and Visualization Module* provides to the specialist a set of tools to help her/him make the diagnosis more objectively. Among the tools it offers, there is the computing and plotting of the angle, the angular speed and the position of joints, as well as the comparison with the norm, obtained in the experimental stage of this work as presented in the previous chapter, implemented in the class *Plotter*.

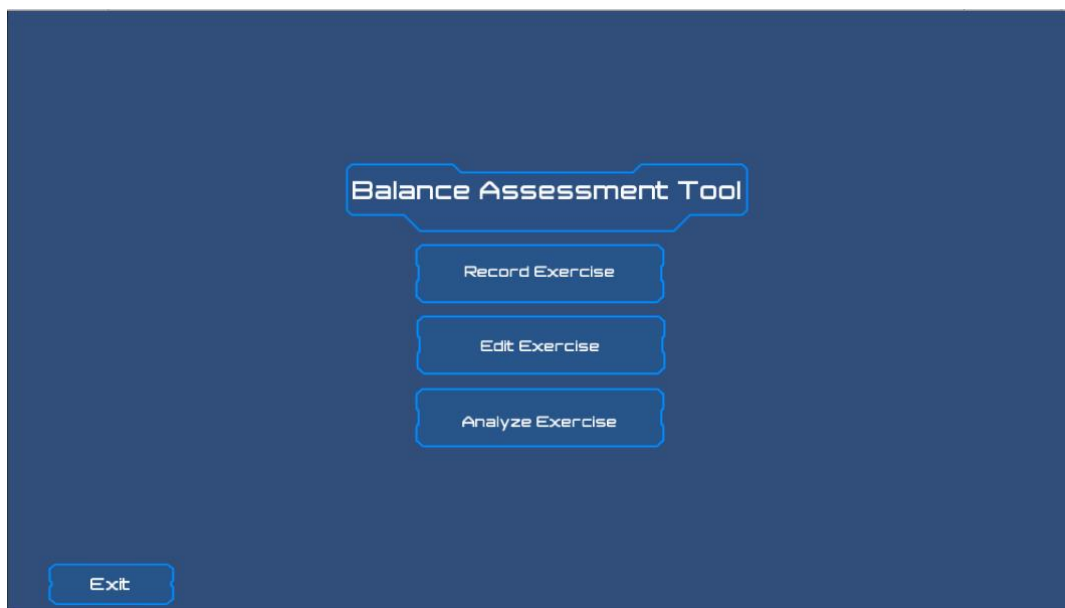


Figure 6 - Home Screen

4.2 User Interface

It was necessary to develop a user interface to interact with the system. The first screen (Figure 6) that appears when initializing the application is the gateway to all application modules. The acquisition module will be used by patients with the help of specialists, and editing and analysis modules will be used only by specialists.

4.2.1. Data Acquisition Module

If the user chooses to record an exercise, a dialog will be prompted in the screen, where the user needs to fill general data of the patient (e.g., id, name, age, sex, disease), and then passes to the exercise selection screen (Figure 7). In that screen the user can select an exercise to record and when s/he has finished s/he can return to the selection screen in order to select another exercise. In this way, the same patient can be recorded doing the entire KEES.



Figure 7 - Exercise selection screen

After selecting an activity, the system presents to the user a screen (Figure 8) with a tutorial that contains animations and text-based explanation, indicating all the details of the exercise.

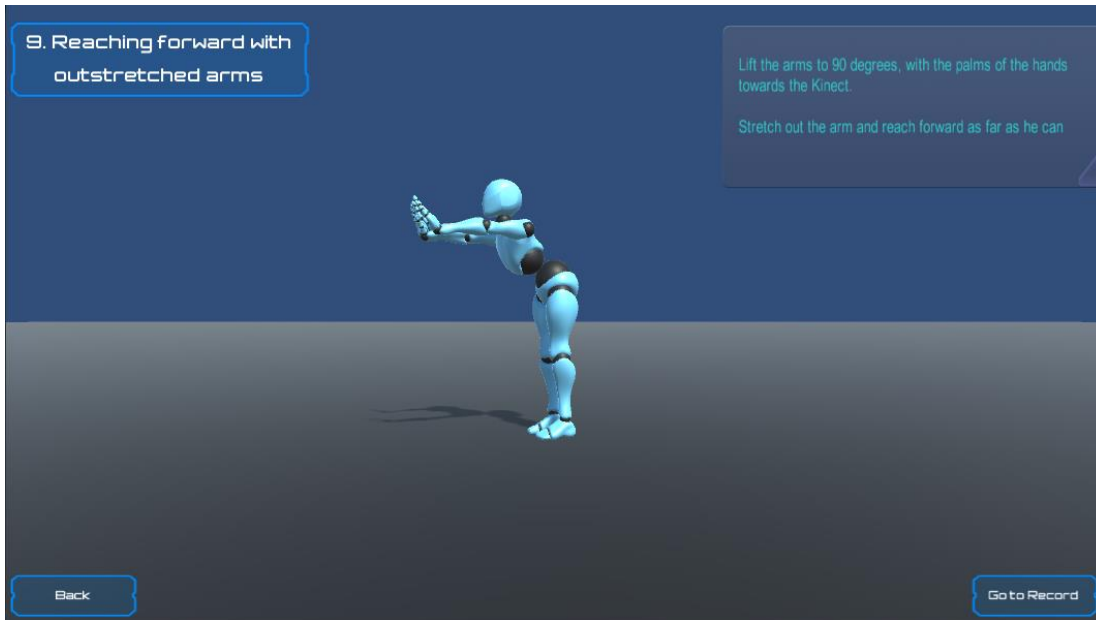


Figure 8 - Exercise tutorial screen

Then the user needs to be calibrated (Figure 9): the user must stay in a fixed position for approximately four seconds in order to obtain original position and angle in reference to the Kinect.



Figure 9 - Calibration screen

When the calibration process is finished, the specialist must initialize the recording (Figure 10) and when the user finishes to perform the exercise, the recording process should be manually stopped (Figure 11).

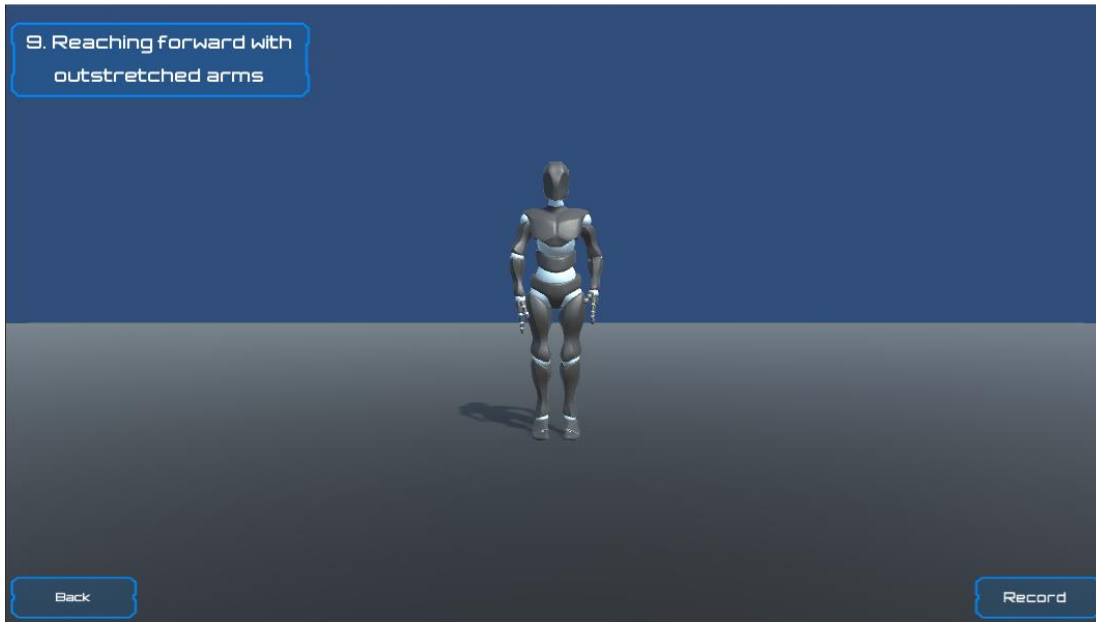


Figure 10 - Ready to record screen



Figure 11 - Recording screen

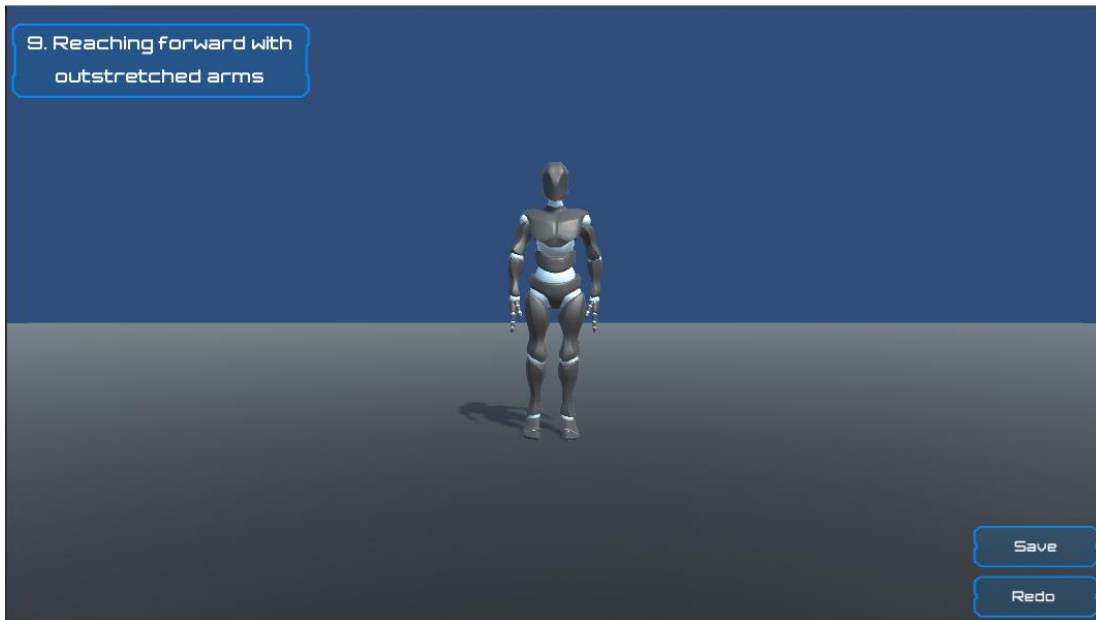


Figure 12 - After record screen

The specialist has the opportunity to redo the recording operation (Figure 12) if anything goes wrong or just save the record and go back to the exercise selection screen.

4.2.2. Editing Module

In the editing module (Figure 13) the specialist has the possibility to load a preview of the exercise recorded. This module has tools for replaying the recording, watching it from different views (e.g. Front view, left view) in the form of animation in an avatar, going step by step, and deleting non useful data that is recorded, because sometimes there is a delay between the user actual beginning and end of the exercise, making the information cleaner for posterior analysis.

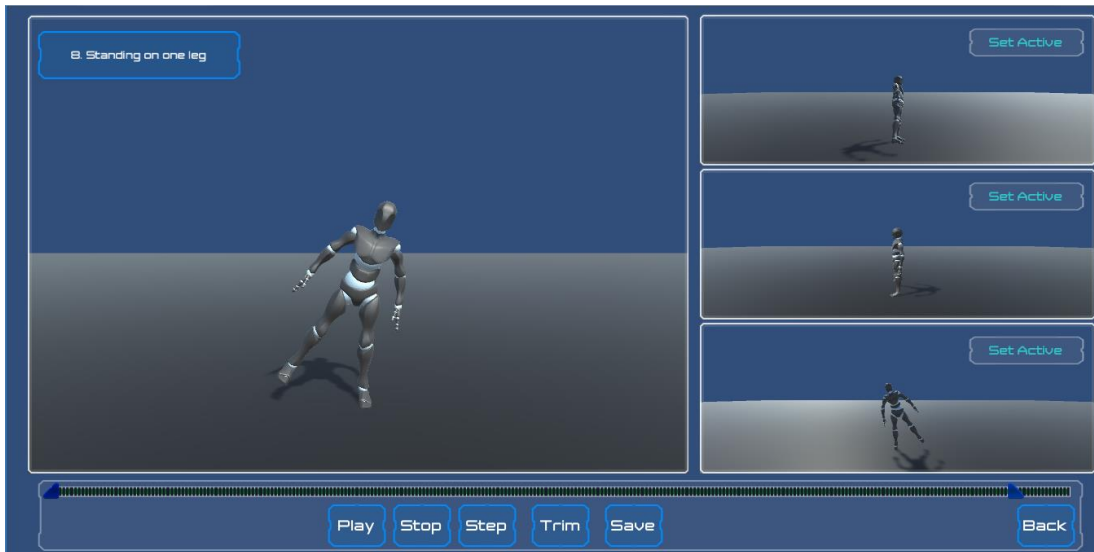


Figure 13 - Editing Module screen

4.2.3. Analysis and Visualization Module

The analysis and visualization module (Figure 14) has a plotting area, a multiview component, and a toolbar. As in the Editing Module, the multiview allows the specialist to have a view of the record from four different perspectives, and to see the real position of the body while at the same time analyze the value of a metric like the angular velocity of a joint of interest that can be plotted using the options of the toolbar in the module. It also allows to plot the same metric over the dataset of healthy subjects or previous data from the same patient, in order to compare with the current exercise.



Figure 14 - Analysis and visualization screen

5 Evaluation and Result

This chapter presents the experiments performed in this study. The first experiment conducted consists in finding the norm of each exercise in the KEES and is detailed in section 5.1. The second experiment consists in comparing the evaluation using the Kinect-based system and the Berg Balance Scale and is detailed in section 5.2.

5.1 Experiment to find the dataset of healthy subjects

In this experiment, we enrolled 10 healthy voluntary individuals, six males, and four females, age 27+- 5, with no mobility issues reported (Table 2). Users were instructed to perform the ten exercises of the KEES. The exercises were recorded using the Acquisition Module of this system. Then, the ten samples of the exercises were processed according to the method presented in section 3.4.5, obtaining for each exercise the “correct” way to perform the activity. The average of each exercise was saved for posterior use in the software, providing to the specialist another feature.

Participants	Age	Sex
Subject 1	25	F
Subject 2	27	M
Subject 3	32	M
Subject 4	26	F
Subject 5	29	M
Subject 6	31	F
Subject 7	26	F
Subject 8	29	M
Subject 9	23	M
Subject 10	22	M

Table 2 - Participants' Characteristics

As an example, we present a comparison between the dataset obtained and the healthy user using the exercise “Standing unsupported with eyes closed” of the KEES. The position of the left shoulder was plotted in the three axes as shows the Figures 15, 16 and 17. The movement in the X-axis means that the subject moves the shoulder to the left or to the right during the exercise. The movement in the Y-axis means that the subject moves the shoulder up or down. The movement in the Z-axis means that the subject moves the shoulder forward or backward.

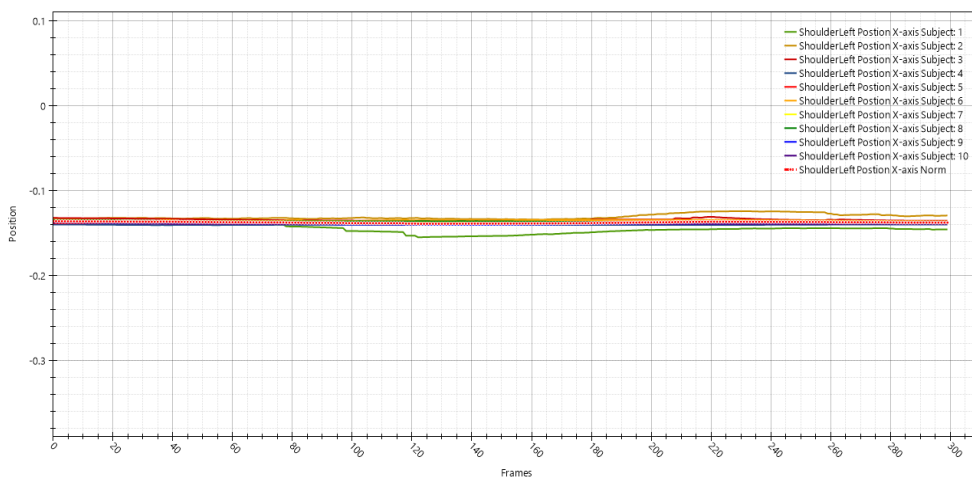


Figure 15 - Dataset vs ten normal subjects - axis X

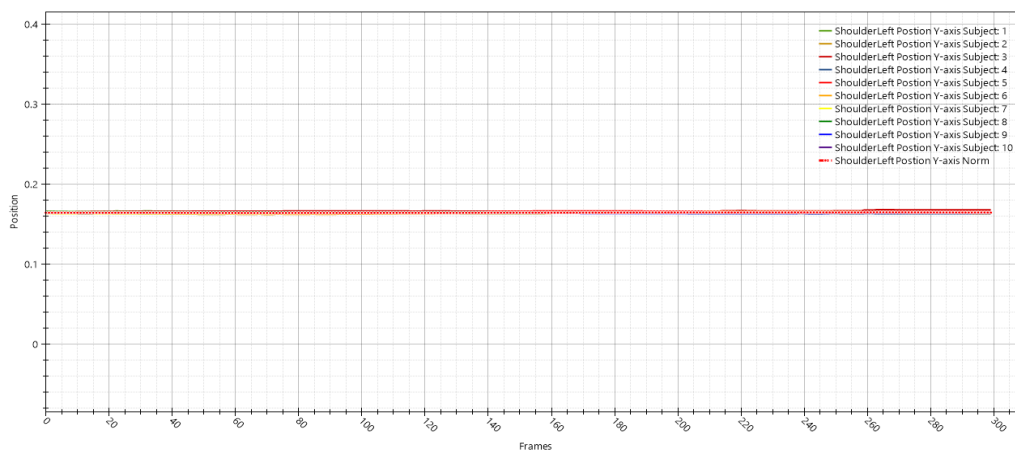


Figure 16 - Dataset vs 10 normal subjects - axis Y

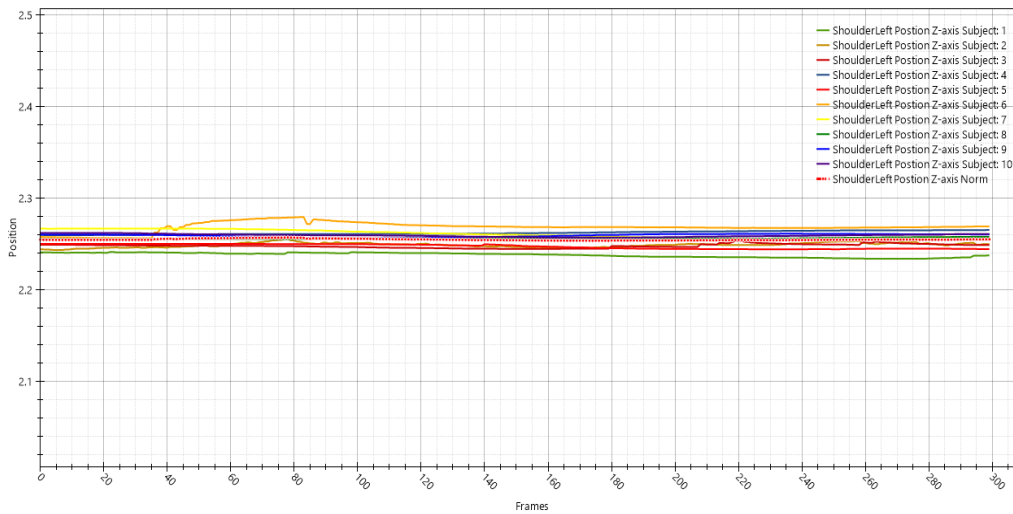


Figure 17 - Dataset vs 10 normal subjects - axis Z

5.2 Evaluation using the Kinect-based system and the Berg Balance Scale

This experiment enrolled five patients of the Sinapse clinic with reduced mobility. The experiments were conducted by the specialists in physical therapy at the clinic using the software and the Berg Balance Scale. They collected the data and then provided us the data. The subjects were three females and two males, age 60+-15 years. Subject 1 and subject 5 were elderly people, subject 2, subject 3 and 4 were post-stroke patients, all of them with balance problems (Table 3).

Patients	Problem	Age	Sex
Subject 1	Elderly people	75	F
Subject 2	Post-stroke	60	F
Subject 3	Post-stroke	50	M
Subject 4	Post-stroke	56	F
Subject 5	Elderly people	73	M

Table 3 - Participants' Characteristics

5.2.1. Clinical scale assessment

Users were instructed by the specialist to perform the 14 exercises of the Berg Balance Scale, in which they were evaluated. The score result and the interpretation of each evaluation, according to the Berg Scale, is presented in the table 3. Three of the five subjects were evaluated as “low fall risk” and two were evaluated as “medium fall risk” (Table 4). Even when they were correctly evaluated using the Berg Balance Scale, since 8 points is required to reveal a genuine change in function between 2 assessments, it turns out to be very hard to make a precise comparison between two subjects that have the same risk of fall according to the scale. Furthermore, like in this case, there may exist patients that still have issues in the movement and problems with balance and yet they obtain the “low fall risk” evaluation.

	Total	Interpretation
Subject 1	49	low fall risk
Subject 2	26	medium fall risk
Subject 3	38	medium fall risk
Subject 4	43	low fall risk
Subject 5	50	low fall risk

Table 4 - Berg Scale Results

5.2.2. Kinect-based assessment

After performing the evaluation using the Berg Scale, the specialists in physical therapy conducted this experiment also in the Sinapse clinic, using the system. The participants were the same. The Kinect was placed at 0.7 m of height. The patients were placed at 2 m of distance from the Kinect. The subjects were recorded while performing the ten exercises of the KEES.

For the analysis, we choose the exercises with more significant movement for the data collected from the patient, and the position of the body recommended by the specialists in physical therapy that has more importance in the observation.

If we use the same measures analyzed before in healthy users (the position of the left shoulder during the exercise “Standing unsupported with eyes closed”) to compare

the data obtained in this experiment of the five unhealthy subjects with the computed norm, we can notice a deviation from the dataset average even in patients with low fall risk (Figures 18,19,20). This noticeable difference in the movements' graphs can be used to help the specialist in getting a more precise assessment of the patient's situation. Also, we apply DTW to achieve a quantitative analysis for each patient (Table 5) including orientation values close to zero indicates more similarity between the average dataset.

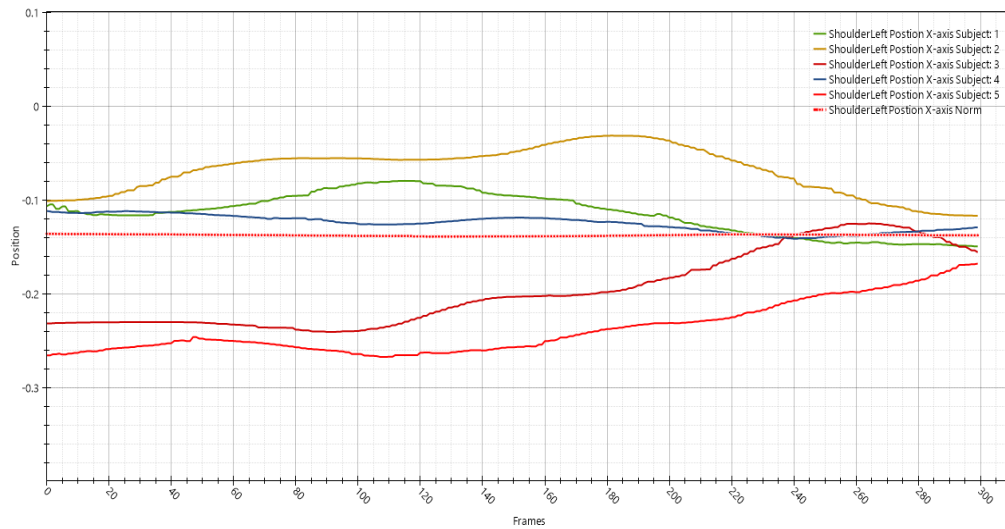


Figure 18 - Dataset vs 5 unhealthy subjects - axis X

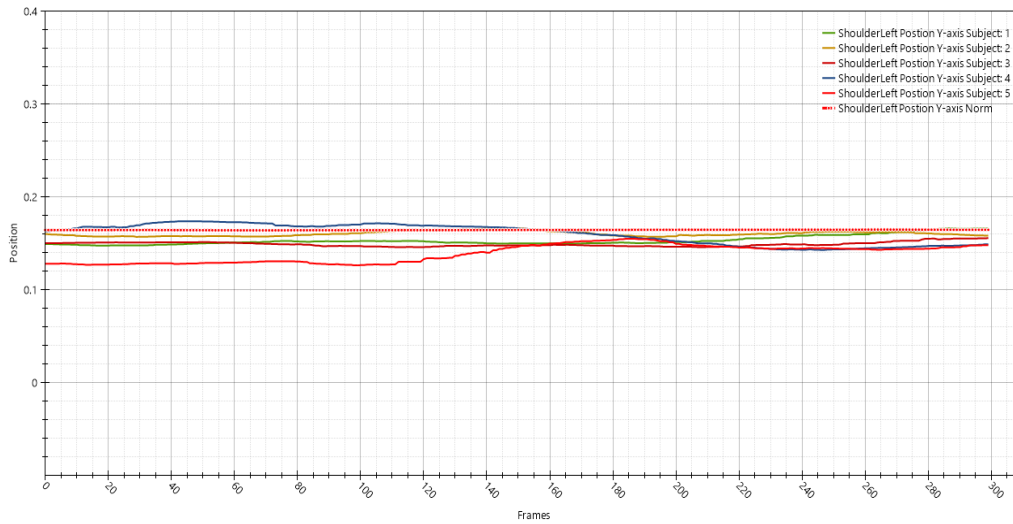


Figure 19 - Dataset vs 5 unhealthy subjects - axis Y

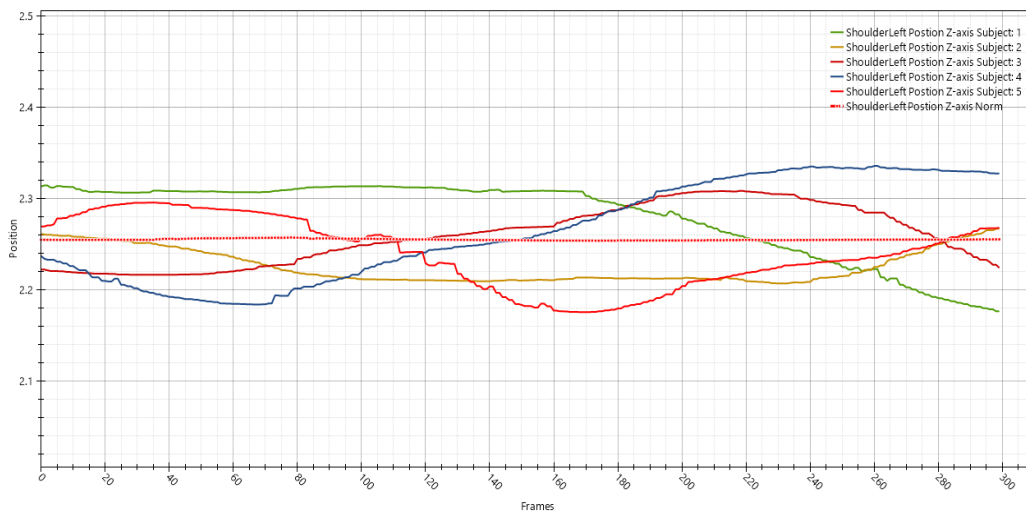


Figure 20 - Dataset vs 5 unhealthy subjects - axis Z

PUC-Rio - Certificação Digital Nº 1412706/CA

DTW cost vs Dataset Average						
Standing unsupported with eyes closed – Left Shoulder						
Patients	Position X	Position Y	Position Z	Orientation X	Orientation Y	Orientation Z
1	1.99	0.15	1.67	0.01	0.03	0.01
2	4.75	0.39	3.74	0.11	0.02	0.00
3	3.13	0.30	1.27	0.03	0.01	0.01
4	1.10	0.12	2.21	0.01	0.01	0.00
5	2.10	0.34	2.00	0.01	0.01	0.01

Table 5 - Dynamic Time Wrapping cost- Standing unsupported with eyes closed

We can observe similar results in the measurements of other joints in other exercises. In some cases, even patients top ranked in the scale can have deviation from the norm, if s/he is not really a patient with good balance. One example of this is their performance during the recording of the exercise “Standing on one leg”. Figure 21 shows the position in the X-axis of the joint of the left knee during the exercise of the five subjects comparing to the dataset average and Table 6 shows the DTW results for all the axis. In this measure a movement in the X-axis means that the subject moves the left knee to the left or to the right.

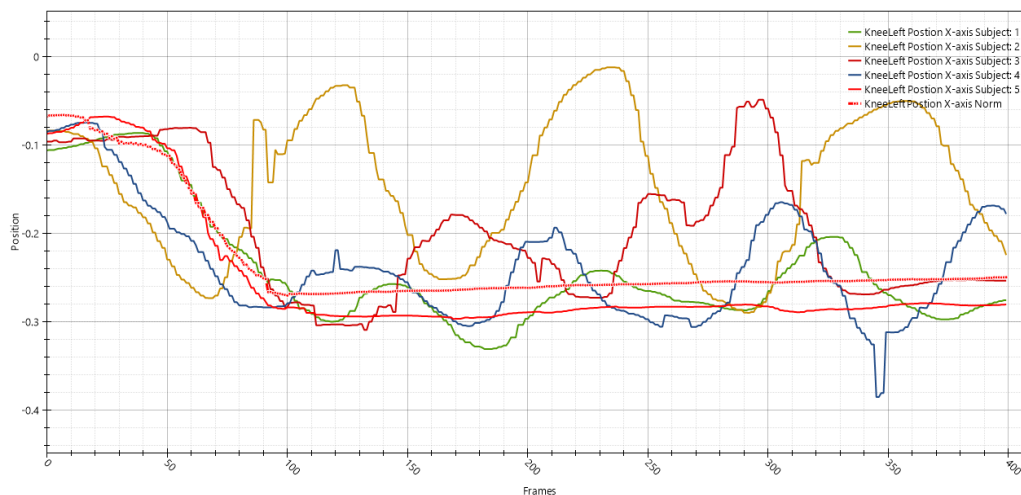


Figure 21 - Left Knee dataset average vs unhealthy subjects X-axis

The curve described by the dataset average represents that in an ideal performance where the subject raise the leg, the left knee moves from the start position to the left and it is sustained in that position.

DTW cost vs Dataset Average Standing on one leg – Left Knee						
Patients	Position X	Position Y	Position Z	Orientation X	Orientation Y	Orientation Z
1	4.35	0.23	0.77	0.17	0.71	0.15
2	16.98	5.31	2.50	1.22	0.53	0.31
3	8.67	7.39	3.66	0.19	0.04	1.07
4	5.29	3.68	1.55	0.30	0.06	0.22
5	3.04	1.67	0.80	0.03	0.21	0.39

Table 6 - Dynamic Time Wrapping cost- Standing on one leg

The subjects 1, 4 and 5 have the performance closer to the dataset average. Those subjects also obtained the “low fall risk” classification with the Berg Balance Scale that is the top classification, even when they performed the exercise with some difficulties. Figure 22 shows in detail a comparison between the norm and those subjects. Subject 1 got 49 in the scale and subject 5 got 50, according to the scale it necessary a difference of 8 point to shows that there is an improvement, so the subject 1 and 5 have almost the same classification, and the subject 4 that obtained 43 has a worst classification. However, in this exercise the performance of the subject 1 was close to the performance of subject 4 than the subject 5, and subject 5 was closest to the norm. Also, the subject 5 achieves the objective of raise the leg and maintain it in the air in a better way. The curves of the subjects 1 and 4 show periodic fluctuations consistent with their constant efforts for maintain the leg in the air.

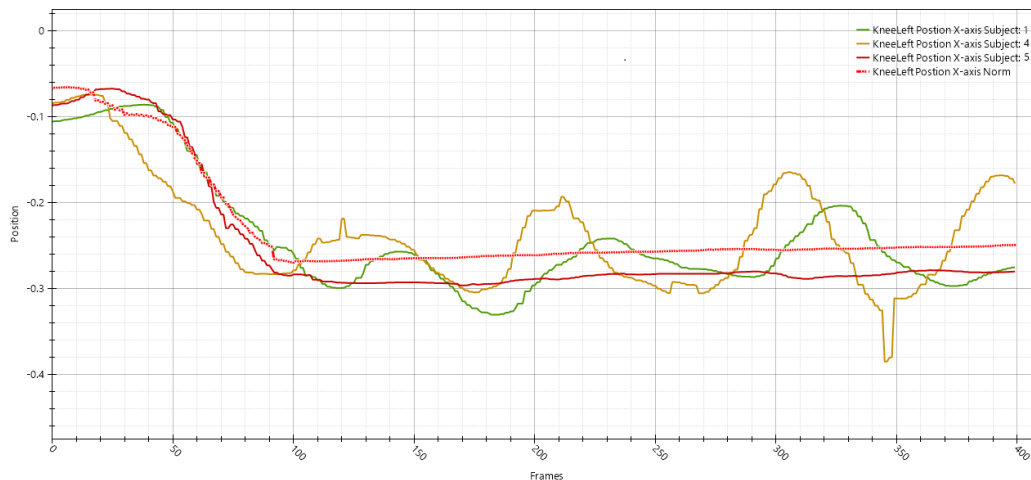


Figure 22 - Subjects with low fall risk vs the dataset average X-axis

Secondly, the subjects 2 and 3 obtained 26 and 38 points respectively, in the assessment with the scale. Both subjects were classified as “medium fall risk”. According to the scale there is a significant difference between them, but the graphic of their performances provides another type of understanding about their problems. Subject 3 presents a few fluctuations but still manages to keep the leg more time in the air, this subject must put the foot on the floor one time causing the knee return to the initial position in the X-axis (Figure 23). On the other hand, subject 2 presents

periodic fluctuations where the position of the knee returns to the initial position, as result of being struggling to raise the leg, he needs to put the foot on the floor many times. So, according to the Berg Balance Scale is clear that, even when they both were classified as “medium fall risk”, they present a significant difference in the scale, but we find that the uses of the measures provided by this system provide more information about the performance in the exercises.

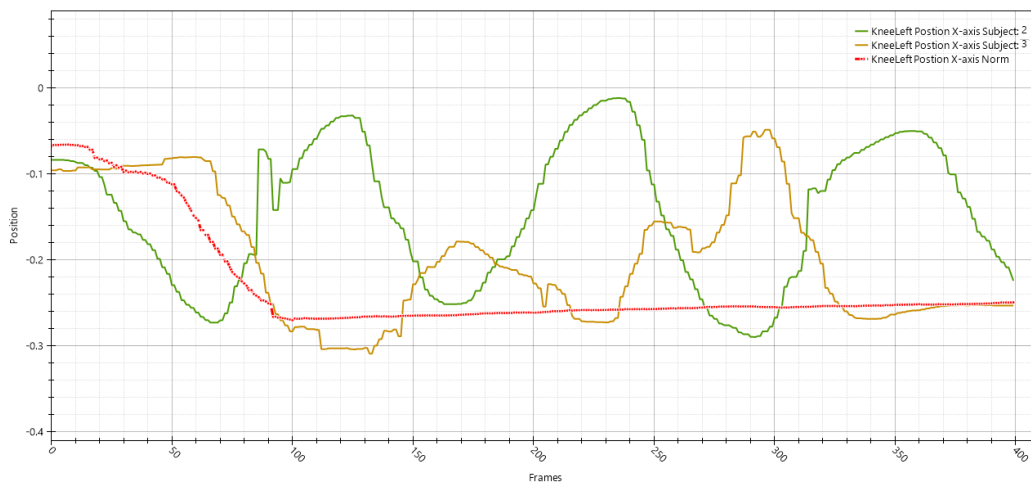


Figure 23 - Subjects with medium fall risk vs the norm X-axis

This case is a good example of how we can make an analysis of the motor state of the patients using the system, and how this new information can help the specialist to get a more effective assessment.

Another interesting result was observed during the analysis of the exercise “Reaching forward with outstretched arms”. The measure used was the position of the right hand in the Z-axis in order to provide information about the movement of the subject forward with extended arms. Figure 24 shows a comparison between the dataset average and the subjects and Table 7 shows the result of DTW. In this exercise the subject was oriented to extend and raise the arms, move the arms forward as much as they can. Figure 25 shows in detail the comparison between the norm and the subjects classified as “low fall risk” by the scale and Figure 26 shows the comparison between the norm and the subjects that obtain a “medium fall risk” classification with the scale. The distance reached by each subject it will be represented in meters by the Z-

axis, it can be measured by difference between the initial position and the final position in that axis.

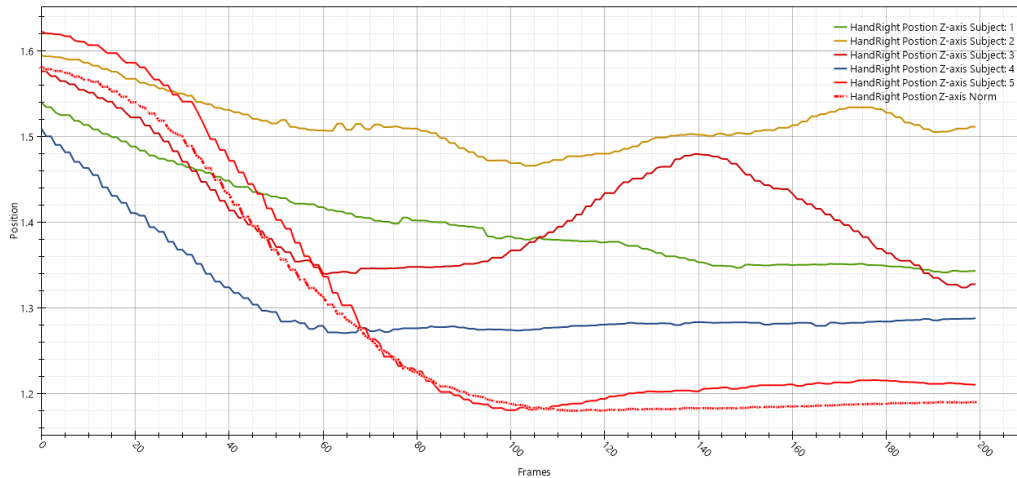


Figure 24 - Right Hand dataset average vs unhealthy subjects Z-axis

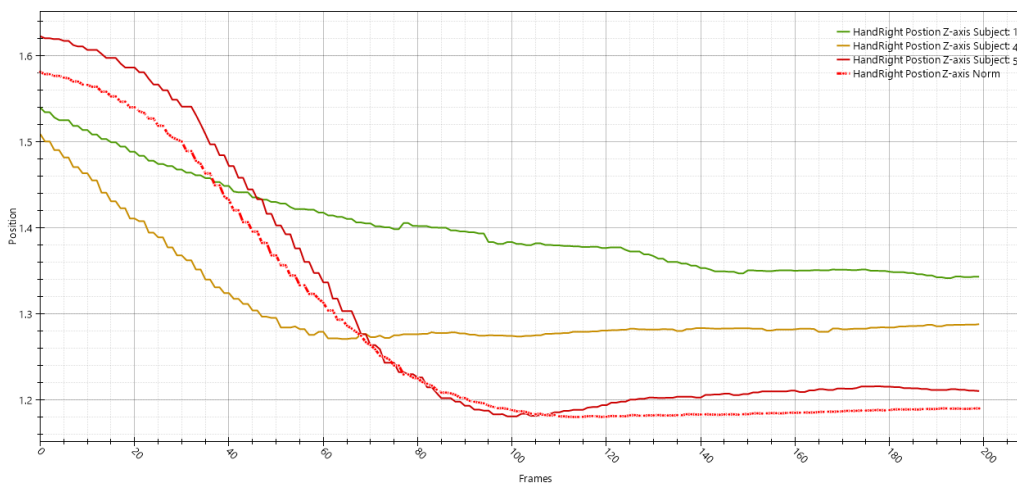


Figure 25 - Subjects with low fall risk vs the dataset average Z-axis

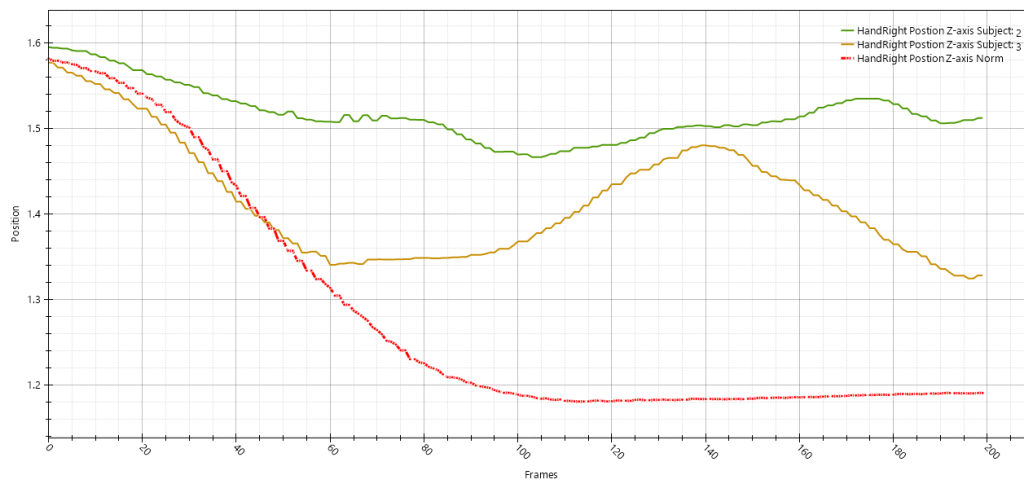


Figure 26 - Subjects with medium fall risk vs the dataset average Z-axis

DTW cost vs Dataset Average						
Reaching forward with outstretched arms – Right Hand						
Patients	Position X	Position Y	Position Z	Orientation X	Orientation Y	Orientation Z
1	0.07	0.24	3.44	0.01	0.03	0.01
2	0.91	0.81	8.12	0.23	0.01	0.00
3	0.66	0.90	6.53	0.03	0.01	0.01
4	0.52	0.73	2.07	0.01	0.01	0.00
5	0.03	0.56	0.53	0.01	0.01	0.01

Table 7 - Dynamic Time Wrapping cost - Reaching forward with outstretched arms

As shown in these experiments we can appreciate that classic clinical scales as the Berg Balance Scale provide limited information about the functional capability of the patients, and do not quantitatively determine the level of the impairment, but only qualitatively associate a score with the patient's performance. Systems like the one discussed in this work can bring tools for the specialists in physical therapy to evaluate a patient with more objectivity.

6 Conclusions

In this work, we presented the implementation of a system using Microsoft Kinect for capturing and processing the motor state of patients, non-invasively, providing tools for the specialists to allow them to conduct a quantitative and objective evaluation of patients with reduced mobility, improving the usual assessment method based on scales.

The vast majority of work related to the Microsoft Kinect in the area of physiotherapy are focused to assist in the rehabilitation of patient, but very few are related to the evaluation of the patient. This work brings a collaboration to this less studied area.

We defined a set of exercises based on a subset of the activities of the Berg Balance Scale. We created also a dataset of exercises using the data of 10 healthy individuals. We propose a system in which a specialist can record the performance of patients during the execution of this set of exercises, the system also provide tools to measure position, angle and velocity between different points of the body detected with Microsoft Kinect. These measures can also be compared with the norm, and the user can see an avatar animation based on that data that allows a 3D view of a patient in any moment.

We perform an experimental analysis of the performance of 5 patients of a physiotherapeutic and neurology clinic, showing that a system as the one implemented in this work can give more information about patients than classical scales.

After the tests,we asked for feedback from the specialists in physical therapy. In the opinion of the specialists a system like this greatly improves the work with the patients since it overcomes the limitations of the most used evaluation methods,

allows to analyze in a quantitative way the patients, the exchange of information between specialists, improves the history of the progress of the patient during the treatment, and generates a large amount of data for future investigations, allowing better separation between different pathologies. It changes the interaction with patients because patients themselves can observe their performance during rehabilitation, which contributes to the patient's psychological state, something very important during the process of rehabilitation. Because the system is compact and low cost, it may also facilitate the remote evaluation of patients who cannot get to hospitals or clinics.

As future work, we plan:

- Implement automatic temporal adjustment of the data recorded
- Increment the number of subject in the dataset, with more pathologies
- Include automatic analysis based on pattern recognition; for that it is necessary to create a big database of patients, more specialized in age ranges, localities, diseases.
- Include functionalities for remote assessment

7 Bibliography

- Berg, M. B.-D. (1992). Clinical and laboratory measures of postural balance in an elderly population. *Arch Phys Med Rehab.*
- Berg, K. O.-D. (1991). Measuring balance in the elderly: validation of an instrument. *Canadian journal of public health*, 57-11.
- Clark, R. A. K. J.-H. (n.d.). Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables,. *Journal of Biomechanics*, 46, 2722–2725.
- Clark, R. A. Y.-H. P. (2013). Validity of the Microsoft Kinect for providing lateral trunk lean feedback during gait retraining. *Gait and Posture*, 38, 1064–1066.
- Diana L. Kornetti, M. P.-P. (2004). Rating Scale Analysis of the Berg Balance Scale. *Arch Phys Med Rehabil.*
- Downs, S. J. (2014). Normative scores on the Berg Balance Scale decline after age 70 years in healthy community-dwelling people: a systematic review. *Journal of physiotherapy*, 85-89.
- Dutta, T. (2012). Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. *Applied Ergonomics*, 43, 645–649.
- Fernández, A. A. S. (2012). Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments. *Proceedings of the 4th International Conference on Intelligent Networking and Collaborative Systems*, 656–661.

- Gavim, A. E. (2012). A influência da avaliação fisioterapêutica na reabilitação neurológica. *Epilepsia*, 90.
- González, D. D.-P.-Z.-R. (2014). A Kinect-based system for cognitive rehabilitation exercises monitoring. *Computer methods and programs in biomedicine*, 620-631.
- Henry, P. M. K. (2012). Rgb-d mapping: Using kinect-style depthcameras for dense 3d modeling of indoor environments.
- Jácome, C. C. (2016). Validity, reliability, and ability to identify fall status of the Berg Balance Scale, BESTest, Mini-BESTest, and Brief-BESTest in patients with COPD. *Physical therapy*.
- Jennifer K Harrison, K. S. (2013). Assessment scales in stroke: clinimetric and clinical considerations. *Clin Interv Aging*, 8, 201–211.
- Kurillo, G. A. C. (2013). Evaluation of upper extremity reachable workspace using Kinect camera. *Technology and Health Care*, 21, 641–656.
- Langley, F. &. (2007). Functional balance assessment of older community dwelling adults: A systematic review of the literature. *The Internet Journal of Allied Health Sciences and Practice*.
- Microsoft. (2016). Retrieved from <https://www.visualstudio.com/>
- Mobini, A. S. B. (2014). Accuracy of Kinect's skeleton tracking for upper body rehabilitation applications. *Disability and Rehabilitation: Assistive Technology*, 9, 344–352.
- Obdrzalek, S. G. K. (2012). Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1188–1193.
- Paavola, J. M. (2013). Use of X-box Kinect gaming console for rehabilitation of an individual with traumatic brain injury: A case report. *Journal of Novel Physiotherapies*, 1-6.

- Podsiadlo, D., & Richardson, S. (1991). The timed 'Up & Go': A test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society*, 142–148.
- Riddle, S. P. (1999). Interpreting validity indexes for diagnostic. *Phys Ther.*
- Saso, P. M.-N. (2016). Responsiveness of the Berg Balance Scale in patients early after stroke. *Physiotherapy Theory and Practice*, 32, 251-261.
- Saposnik, G. T. (2010). Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation a pilot randomized clinical trial and proof of principle. *Stroke*, 1477-1484.
- Saso, A. M.-N. (2016). Responsiveness of the Berg Balance Scale in patients early after stroke. *Physiotherapy theory and practice*, 251-261.
- Sinapse. (2016). Retrieved from Sinapse: <http://sinapseclinica.com/>
- Skubic, E. E. (2011). Passive in-home measurement of stride-to-stride gait variability comparing vision and Kinect sensing. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society* , 6491–6494.
- Stone, E. E. (2013). Unobtrusive, continuous, in-home gait measurement using the Microsoft Kinect. *Biomedical Engineering, IEEE Transactions on*, 2925-2932.
- Tinetti, M., Williams, T. F., & Mayewski, R. (1986). Fall risk index for elderly patients based on number of chronic disabilities. *American Journal of Medicine*, 429–434.
- Unity, T. (2016). Retrieved from Unity3D: <https://unity3d.com>
- Wei, T. Q. (2014). Kinect Skeleton Coordinate Calibration for Remote Physical Training. *Proceedings of the International Conference on Advances in Multimedia*, (pp. 23-27).
- Wang X., A. M. (2013). Experimental comparison of representation methods and distance measures for time series data. *Data Mining and Knowledge Discovery*, 275–309.

Zhou, H., & Hu, H. (2008). Human motion tracking for rehabilitation—a survey. *Biomedical Signal Processing and Control*, 1–18.

8 Annex A

A.1 Berg Balance Scale

In this annex, we present the Berg Balance Scale (Berg, 1991):

Berg Balance Scale

Name: _____ Date: _____

Location: _____ Rater: _____

ITEM DESCRIPTION	SCORE (0-4)
Sitting to standing	_____
Standing unsupported	_____
Sitting unsupported	_____
Standing to sitting	_____
Transfers	_____
Standing with eyes closed	_____
Standing with feet together	_____
Reaching forward with outstretched arm	_____
Retrieving object from floor	_____
Turning to look behind	_____
Turning 360 degrees	_____
Placing alternate foot on stool	_____
Standing with one foot in front	_____
Standing on one foot	_____
Total	_____

GENERAL INSTRUCTIONS

Please document each task and/or give instructions as written. When scoring, please record the lowest response category that applies for each item.

In most items, the subject is asked to maintain a given position for a specific time. Progressively more points are deducted if:

- the time or distance requirements are not met
- the subject's performance warrants supervision
- the subject touches an external support or receives assistance from the examiner

Subject should understand that they must maintain their balance while attempting the tasks. The choices of which leg to stand on or how far to reach are left to the subject. Poor judgment will adversely influence the performance and the scoring.

Equipment required for testing is a stopwatch or watch with a second hand, and a ruler or other indicator of 2, 5, and 10 inches. Chairs used during testing should be a reasonable height. Either a step or a stool of average step height may be used for item # 12.

- 1- SITTING TO STANDING INSTRUCTIONS: Please stand up. Try not to use your hand for support.
 - () 4 able to stand without using hands and stabilize independently
 - () 3 able to stand independently using hands
 - () 2 able to stand using hands after several tries
 - () 1 needs minimal aid to stand or stabilize
 - () 0 needs moderate or maximal assist to stand

- 2- STANDING UNSUPPORTED INSTRUCTIONS: Please stand for two minutes without holding on.
 - () 4 able to stand safely for 2 minutes
 - () 3 able to stand 2 minutes with supervision
 - () 2 able to stand 30 seconds unsupported
 - () 1 needs several tries to stand 30 seconds unsupported
 - () 0 unable to stand 30 seconds unsupported

If a subject is able to stand 2 minutes unsupported, score full points for sitting unsupported. Proceed to item #4.

3- SITTING WITH BACK UNSUPPORTED BUT FEET SUPPORTED ON FLOOR OR ON A STOOL INSTRUCTIONS: Please sit with arms folded for 2 minutes.

- 4 able to sit safely and securely for 2 minutes
- 3 able to sit 2 minutes under supervision
- 2 able to sit 30 seconds
- 1 able to sit 10 seconds
- 0 unable to sit without support 10 seconds

4- STANDING TO SITTING INSTRUCTIONS: Please sit down.

- 4 sits safely with minimal use of hands
- 3 controls descent by using hands
- 2 uses back of legs against chair to control descent
- 1 sits independently but has uncontrolled descent
- 0 needs assist to sit

5- TRANSFERS INSTRUCTIONS: Arrange chair(s) for pivot transfer. Ask subject to transfer one way toward a seat with armrests and one way toward a seat without armrests. You may use two chairs (one with and one without armrests) or a bed and a chair.

- 4 able to transfer safely with minor use of hands
- 3 able to transfer safely definite need of hands
- 2 able to transfer with verbal cuing and/or supervision
- 1 needs one person to assist
- 0 needs two people to assist or supervise to be safe

6- STANDING UNSUPPORTED WITH EYES CLOSED INSTRUCTIONS: Please close your eyes and stand still for 10 seconds.

- 4 able to stand 10 seconds safely
- 3 able to stand 10 seconds with supervision

- 2 able to stand 3 seconds
- 1 unable to keep eyes closed 3 seconds but stays safely
- 0 needs help to keep from falling

7- STANDING UNSUPPORTED WITH FEET TOGETHER INSTRUCTIONS:

Place your feet together and stand without holding on.

- 4 able to place feet together independently and stand 1 minute safely
- 3 able to place feet together independently and stand 1 minute with supervision
- 2 able to place feet together independently but unable to hold for 30 seconds
- 1 needs help to attain position but able to stand 15 seconds feet together
- 0 needs help to attain position and unable to hold for 15 seconds

8- REACHING FORWARD WITH OUTSTRETCHED ARM WHILE

STANDING INSTRUCTIONS: Lift arm to 90 degrees. Stretch out your fingers and reach forward as far as you can. (Examiner places a ruler at the end of fingertips when arm is at 90 degrees. Fingers should not touch the ruler while reaching forward. The recorded measure is the distance forward that the fingers reach while the subject is in the most forward lean position. When possible, ask subject to use both arms when reaching to avoid rotation of the trunk.)

- 4 can reach forward confidently 25 cm (10 inches)
- 3 can reach forward 12 cm (5 inches)
- 2 can reach forward 5 cm (2 inches)
- 1 reaches forward but needs supervision
- 0 loses balance while trying/requires external support

9- PICK UP OBJECT FROM THE FLOOR FROM A STANDING POSITION

INSTRUCTIONS: Pick up the shoe/slipper, which is in front of your feet.

- 4 able to pick up slipper safely and easily

- 3 able to pick up slipper but needs supervision
- 2 unable to pick up but reaches 2-5 cm(1-2 inches) from slipper and keeps balance independently
- 1 unable to pick up and needs supervision while trying
- 0 unable to try/needs assist to keep from losing balance or falling

10- TURNING TO LOOK BEHIND OVER LEFT AND RIGHT SHOULDERS

WHILE STANDING INSTRUCTIONS: Turn to look directly behind you over toward the left shoulder. Repeat to the right. (Examiner may pick an object to look at directly behind the subject to encourage a better twist turn.)

- 4 looks behind from both sides and weight shifts well
- 3 looks behind one side only other side shows less weight shift
- 2 turns sideways only but maintains balance
- 1 needs supervision when turning
- 0 needs assist to keep from losing balance or falling

11- TURN 360 DEGREES INSTRUCTIONS: Turn completely around in a full

circle. Pause. Then turn a full circle in the other direction.

- 4 able to turn 360 degrees safely in 4 seconds or less
- 3 able to turn 360 degrees safely one side only 4 seconds or less
- 2 able to turn 360 degrees safely but slowly
- 1 needs close supervision or verbal cuing
- 0 needs assistance while turning

12- PLACE ALTERNATE FOOT ON STEP OR STOOL WHILE STANDING

UNSUPPORTED INSTRUCTIONS: Place each foot alternately on the step/stool. Continue until each foot has touched the step/stool four times.

- 4 able to stand independently and safely and complete 8 steps in 20 seconds
- 3 able to stand independently and complete 8 steps in > 20 seconds
- 2 able to complete 4 steps without aid with supervision

- 1 able to complete > 2 steps needs minimal assist
- 0 needs assistance to keep from falling/unable to try

13- STANDING UNSUPPORTED ONE FOOT IN FRONT INSTRUCTIONS:

(DEMONSTRATE TO SUBJECT) Place one foot directly in front of the other. If you feel that you cannot place your foot directly in front, try to step far enough ahead that the heel of your forward foot is ahead of the toes of the other foot. (To score 3 points, the length of the step should exceed the length of the other foot and the width of the stance should approximate the subject's normal stride width.)

- 4 able to place foot tandem independently and hold 30 seconds
- 3 able to place foot ahead independently and hold 30 seconds
- 2 able to take small step independently and hold 30 seconds
- 1 needs help to step but can hold 15 seconds
- 0 loses balance while stepping or standing

14- STANDING ON ONE LEG INSTRUCTIONS: Stand on one leg as long as you can without holding on.

- 4 able to lift leg independently and hold > 10 seconds
- 3 able to lift leg independently and hold 5-10 seconds
- 2 able to lift leg independently and hold ≥ 3 seconds
- 1 tries to lift leg unable to hold 3 seconds but remains standing independently
- 0 unable to try of needs assist to prevent fall
- TOTAL SCORE (Maximum = 56)