Pós-Graduação em Ciência da Computação

THIAGO DE MENEZES CHAVES

A GESTURE RECOGNITION LIBRARY FOR THE THERAPY DOMAIN AND ITS APPLICATIONS

Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

RECIFE
2016

Thiago de Menezes Chaves

A Gesture Recognition Library for the Therapy Domain and Its Applications

A MASTERS THESIS PRESENTED TO THE CENTER FOR INFORMATICS OF FEDERAL UNIVERSITY OF PERNAMBUCO IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER IN SCIENCE IN COMPUTER SCIENCE.

ADVISOR: **Veronica Teichrieb**

RECIFE
2016

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

C512g Chaves, Thiago de Menezes
A gesture recognition library for the therapy domain and its applications / Thiago de Menezes Chaves. – 2016.
67 f.: il., fig., tab.

Orientadora: Verônica Teichrieb.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2016.
Inclui referências.

1. Visão computacional. 2. Reconhecimento de gestos. 3. Fisioterapia. I. Teichrieb, Verônica (orientadora). II. Título.

006.37 CDD (23. ed.) UFPE- MEI 2017-224

Thiago de Menezes Chaves

A Gesture Recognition Library for the Therapy Domain and its Applications

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação

Aprovado em: 11/03/2016.

BANCA EXAMINADORA

____
Prof. Dr. Giordano Ribeiro Eulalio Cabral
Centro de Informática / UFPE

____
Prof. Dr. Thales Miranda de Almeida Vieira
Instituto de Matemática / UFAL

____
Profa. Dra. Veronica Teichrieb
Centro de Informática / UFPE
(**Orientadora**)

ABSTRACT

The computational implementation of human body gestures recognition has been a challenge for several years. Nowadays, thanks to the development of RGB-D cameras it is possible to acquire a set of data that represents a human position in space. Despite that, these cameras provide raw data, still being a problem to identify in real-time a specific pre-defined user movement continuously which can then be applied in applications as, for example, the tracking of physiotherapeutic movements or exercises. This work presents two new techniques to identify gestures, both having physiotherapeutic concerns about the performed exercise; one is based on physiotherapeutic standards, the biomechanical planes, while the other aims to recognize the functional exercises and is based on a concept called checkpoints. Both these techniques were tested and validated by physiotherapists from the Physiotherapy Department at the Federal University of Pernambuco. The techniques were also integrated in a library which was then used in two case studies and two general applications where their applicability was tested in physiotherapeutic and non-physiotherapeutic domains obtaining good results and showing that they can be used on general applications as well.

Keywords: Gesture recognition. RGB-D. Physiotherapy.

RESUMO

Implementar um algoritmo computacional de reconhecimento de gesto tem sido um desafio por muitos anos. Hoje em dia, com o desenvolvimento das câmeras RGB-D, é possível adquirir um conjunto de dados que representa a posição de uma pessoa no espaço. Apesar disso, os dados adquiridos por estas câmeras ainda não são suficientes para identificar, em tempo real e de forma contínua, movimentos predefinidos dos usuários, os quais podem ser usados em aplicações como, por exemplo, a análise de movimentos ou exercícios fisioterapêuticos. Este trabalho apresenta duas novas técnicas de reconhecimento de gestos, ambas voltadas ao domínio de fisioterapia; a primeira é baseada em padrões da fisioterapia, chamados de planos biomecânicos, e a segunda tem como propósito reconhecer os gestos realizados durante os exercícios funcionais e é baseada num conceito chamado de checkpoints. Ambas técnicas foram testadas e validadas por fisioterapeutas do Departamento de Fisioterapia da Universidade Federal de Pernambuco. Essas técnicas foram integradas em uma biblioteca, a qual então foi utilizada para desenvolver dois estudos de caso e duas aplicações de propósito gerais, onde suas aplicabilidades foram testadas tanto no domínio fisioterapêutico como fora dele, obtendo bons resultados e mostrando que tais técnicas também podem ser usadas em aplicações gerais.

Palavras-chave: Reconhecimento de gestos. RGB-D. Fisioterapia.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Human motion analysis is an increasing field of research nowadays mainly due its wide range of applicability, such as natural interaction (Hopmann et al. 2011), accessibility, games (Kang, Lee, and Jung 2004), rehabilitation (A. Da Gama et al. 2012b), sports and training. This analysis consists of methods that enable a system to identify whenever or not a predetermined movement was performed by a user. For example, it can be used to check and inform when a movement is being performed properly or incorrectly during a sport training for athletes (Urtasun, Fleet, and Fua 2006), or as well, to trigger some in game action whenever a certain kind of gesture is detected.

Several proposals have been made in order to solve the problem of recording and identifying human motion and gestures. Firstly, in order to recognize the human body gestures, there is the need to identify the user in the environment and extract useful information about his/her body structure. Some approaches make use of inertial sensors like accelerometers and gyroscopes attached to the user. These sensors are used to retrieve data about the relative position of some body articulations (represented as points in the space). Another popular approach is to apply optical devices to detect a set of markers attached to the user body and an array of cameras to track them. Both approaches have the disadvantage that wearable devices need to be attached to the user. The use of markers or sensors can cause discomfort for the user, and also imply in errors case the markers are misplaced or displaced during the use.

A novel approach to exempt the need for attached devices is through the use of RGB-D cameras, which means a Red-Green-Blue camera with addition of a Depth sensor. Since the launch of the first Microsoft Kinect device (Microsoft Corporation 2016c), real-time depth sensors have been broadly explored and similar solutions have been shown by Asus Xtion (ASUS 2016) and Intel RealSense F200 (Intel Corporation 2016b), as well as the new version of the Microsoft Kinect (Microsoft Corporation 2016b) launched in 2013. For each depth sensor, usually, a tracking solution comes coupled on its Software Development Kit. Microsoft Kinect SDK 2.0 (Microsoft Corporation 2016a) offers a full body tracking solution, retrieving a skeleton containing 25 joints representing the user body, and also provides a face tracking solution. These depth sensors and their SDKs provide real-time output data representing the body skeletons of the sensed users.

The knowledge of the position of some body parts does not imply into recognizing motion or gestures. In order to perform the recognition, several approaches have been developed. Most of them uses machine learning techniques like Support Vector Machines (SVM), Random Forests (RF), neural network, among others.

This new approach in computationally recognizing gestures is very attractive to the therapy domain (more specifically, physiotherapy) since this area involves exercises and movement. Furthermore, the physiotherapy field has its own specific demands, which prevents the use of some already developed recognition techniques, or at least demands some modification or complementing these techniques.

The physiotherapy exercises can be classified into two groups, the biomechanical and the functional ones, each of them has its own peculiarities which are better presented at Chapter X. Both these types of gestures has to be continuously analyzed, not only classified but it has to be able of validating the gesture during its performance, for example, at the middle of it.

1.1 **Goals**

By joining the capabilities of these sensors and the physiotherapy needs, it is possible to build focused systems that can analyze and use the performed physiotherapy exercises as input to control interfaces. In order to achieve this goal it is required to recognize users' motions or gestures once the real-time skeleton data is provided by a particular sensor. Based on these factors, this work presents the development and evaluation of gesture recognition methods addressing the physiotherapy domain. Moreover, in order to apply the recognition solutions on the target scenarios, applications are developed as case studies.

The specific goals of this dissertation are as follows:

- Investigate related work in the area of gesture recognition with additional interest on the physiotherapy application domain;
- Develop gesture recognition tools to attend physiotherapists demands;
- Test and validate the developed tools;
- Develop applications as case studies using the recognition tools;
- Test and validate the developed applications.

1.2 **Contributions**

The main contributions of this work are listed as follows:

- A new skeleton representation was developed in this work, defined according to biomechanics standards and achieving invariance to different body scales, body or sensor orientations and translations;
- A movement recognition method was developped also according to biomechanic standards by considering the body planes as reference (A.E.F. Da Gama et al. 2014);
- Another movement recognition method was developed using a single-shot learning approach. The method uses checkpoints and is capable of rapidly recording and recognizing a wide range of gestures (Alana Da Gama et al. 2013). Although this thechnique does not bring any new feature, none of the searched techniques in literature could provide all of the following features joined in a movement analysis library:
    - Configurable precision;
    - Continuous output (as well as the discrete one, class output);
    - Does not suffer from time alignment issues (some thechniques require the user to tell when the gesture will be perfomed, or the input is already croped exactly where the gesture occurs);
    - Defines a gesture by a subset of the body segments leaving everything else free;
    - Works in real-time;
    - Fast registration of a new gestures.

- Both techniques introduce the feature of allowing the physiotherapist to control how precise the movement performed by the patient should be;
- A movement analysis library was developed to abstract the implementation details for the application which intends to use the recognition methods. The library also allows an easy and fast configuration of the desired behavior by the application;
- A VR based case study application on the physiotherapy domain was developed. The application was tested with therapists, programmers and unexperts to be evaluated and improved, which resulted on a second version of it showing high levels of satisfaction for the participants, and added value to physiotherapeutic playfulness and motivation according to the participants' perspective (Oliveira et al. 2013);
- A case study application on the physiotherapy domain that is the first to present an AR rehabilitation system based on ISB standards, which enables the system to interact and to be configured according to physiotherapeutic needs. The application was tested with patients, showing that the percentage of correct exercises, measured by the movement analysis method we developed, improved from 69.02% to 93.73% when users interacted with the mirrARbilitation. The number of exercise repetitions also improved from 34.06 to 66.09 signifying that our system increased motivation of the users. The system also prevented the users from performing the exercises in a completely wrong manner. Finally, with the help of our system the users worst result was performing 73.68% of the rehabilitation movements correctly. Besides the engagement, these results suggest that the use of biomechanical standards to recognize movements is valuable in guiding users during rehabilitation exercises (Alana Elza Fontes Da Gama et al., n.d.);

Also, two additional applications were developed to explore the recognition results on different domains such as digital art and entertainment.

1.3 **Structure Of The Document**

This dissertation is structured as follows. Chapter 2 discusses the main related work on the area of gesture recognition. Chapter 3 details the developed recognition solutions, explaining the development process and its results, beyond the tests performed in order to validate the developed solutions as feasible for the physiotherapy domain. Chapter 4 addresses the developed case studies that use the recognition methods detailed in Chapter 3. Finally, Chapter 5 presents the main conclusions and future work.

2 LITERATURE REVIEW

The process to recognize a gesture was an even harder task before the popularization of the RGB-D sensors. On most vision-based solutions, devices had to rely only on the RGB image where it was very difficult to segment the user from the background. This changed when the Kinect (Microsoft Corporation 2016c) came to market at a relatively low price. The depth map provided by the Kinect (among other depth sensors) allows the user to be segmented from the background. This information can be used to recognize a gesture using solely the depth image, or else the depth combined with the color image. At last the skeleton provided by the device can also be used on its own or combined with the other data sources.

One of the firsts works that came out in the same year the Kinect was launched was (Biswas and Basu 2011), which did not use the skeleton given by the device, but rather uses only the depth information in order to segment the user from the background. Then, it applies a supervised machine learning technique called Support Vector Machines (SVM) to learn the gesture from a grid of depth levels histograms around the user. These authors also focus on hand gestures and the samples only include the user from the waist up so the grid has a fixed size in a way that it gets exactly the region where the head and hands of the user are.

On the other hand, the work proposed by (Ming, Ruan, and Hauptmann 2012) uses the depth information together with the color image. It presents a new descriptor, 3D MoSIFT, which is an extension of the SIFT in order to incorporate the motion and 3D information. This technique uses also k-means to build a codebook, which is applied into a bag-of-features. This bag-of-features is then used as input to a SVM classifier. Tests were made using hand gestures and they showed progress in the recognition rate when compared to the previous works using SIFT and its improvements.

The Kinect also provides the information about the position of some body parts which includes joints of hands, feet and head. This set of points is also called skeleton and has the advantage that the body parts are already recognized, so the gesture recognition has more refined information to use. One of the first works that uses the Kinect skeleton to recognize gestures is (Alvarenga, Correa, and Osório 2011) which basically uses an artificial neural network to recognize some hand gestures using only the hand position given by the device. The system is limited to use only one of the points that the Kinect gives at a time. However, it got high accuracy in the tests made recognizing 83 out of 90 hand gesture samples.

(Suma et al. 2012) describe a tool which makes possible to describe a gesture using text. This text is composed by sentences which can describe a pose, e.g., “left hand above the head by 30 cm”, the speed of a gesture, e.g., “right hand forward by at least 5 cm/s”, or even body constraints as “turn left by 15 degrees”. This recognition can then be bind to a keyboard or mouse input so when a gesture is recognized, a key is pressed activating something in the application. This technique works in real-time and it is intuitive to use as the user describes the movement, but, although combining multiple descriptions can make it possible to recognize gestures that are not too simple, it still lacks the power to recognize complex gestures as these are hard to describe.

The work presented by (Gu et al. 2012) presents the use of a Hidden Markov Model (HMM) to recognize gestures using the skeleton. It trains a HMM for each gesture it needs to identify, so it is easy to further add a new gesture by training a new HMM and simply add it, while maintaining the others. Other advantages of this method are that it is person independent (i.e. it can be trained by one person and used by another, with different body proportions). It is also orientation and distance invariant, i.e., it does not matter where the user's body is facing to or if is closer or farther from the sensor, and it also recognizes if the gesture is performed faster or slower, within certain ranges, compared to the training data. However, it also has a noticeable delay in the recognition.

The method presented by (Wang 2012) proposes to recognize gestures that involve interaction with objects. It introduces two new features used in the proposed gesture recognition: 3D joint position feature and the Local Occupancy Pattern (LOP). The 3D joint position makes the data extracted from the skeleton invariable to the orientation and distance of the body related to the sensor, while the LOP extracts information about the environment, for example, if the user is touching a cup and thus he is performing the "drinking" action. Then, a Fourier Temporal Pyramid (FTP) is built in order to eliminate noise, and to solve the problem of time misalignment of the gestures. There is one pyramid for each joint and SVM is used in a process to select which of these features are relevant, and later used in the recognition step. Aside from the new introduced features and the whole process of recognition, it also introduced the partial gestures concept, which is to choose which body parts are relevant to each of the gestures and discard the influence of the unused body parts.

The biggest difference of the work of (Bigdelou et al. 2012) is that aside from recognizing a gesture, the proposed method also tells the state of the gesture, which was called as being the space-temporal information. They also propose a change in the representation of the skeleton information and test three ways of doing it. This new representation is normalized and a Principal Component Analysis (PCA) technique is applied to reduce its dimensionality, particularly to a unidimensional feature. This new measure resulted from the PCA will tell the space-temporal information, while the raw representation will be used to identify the gesture. The task of identifying is then to search between all the saved data and check which one is the closest to the current pose. They tested it in a user interface and got good recognition rates. One discussed issue was that the system could not tell when the gesture was starting, so the user should trigger the recognition by voice.

(Yang and Tian 2012) use a feature that has information about posture, movement and an offset. This feature has a dimension of about 2970 so PCA is used to reduce it. In order to classify the gestures, a Naïve-Bayes Nearest-Neighbor technique is applied, and it is shown that it performs better than SVM. One advantage of this technique is that it can recognize a gesture before its end. Tests showed that largely reducing the number of frames by 66% provokes a smaller decrease of the accuracy of recognition by only 6.6%.

An improvement to the Dynamic Time Warping (DTW) is proposed by (Celebi et al. 2013) as a result of applying weights to each of the joints that are being tracked. This way, if it is a hand gesture, the leg will have almost no influence in the result, supporting partial gestures and allowing the user to move freely his body parts that are not required for the gesture. DTW is good to solve the time alignment problem suffered by some of the methods.

Another way to recognize a gesture is proposed by (Jun et al. 2013a) where the movement is recorded, then there is a pre-processing step where the video is cut into several clips where each clip contains only one occurrence of a single gesture and also it is normalized regarding time so there are low speed discrepancies. PCA is applied to reduce its dimensionality and K-NN used to classify. (Wan et al. 2013) also use K-NN, beyond a particular feature called 3D EMoSIFT, an extension to the 3D MoSIFT proposed by (Ming, Ruan, and Hauptmann 2012). They use a codebook to generate histograms which are the input to the K-NN. Both of them have the problem of not working in real-time; the later has a delay of 887ms just to extract the features.

In the work of (Vemulapalli, Arrate, and Chellappa 2014) another body representation is presented using the information of the skeleton. For each joint there will be a basis centered on the joint and with the X-axis aligned to its parent joint. Then the representation of the body is the set of rotation and translation matrix that transforms a base from one joint to another, for each pair of joints. In addition, DTW and FTP are used to solve the problem of alignment in time and noise. Then, a SVM classifier performs the classification.

Also, the work of (Faugeroux et al. 2014) proposes a method of identifying several gestures performed in a single shot by identifying when the user briefly stops doing anything in order to segment this data to be the training data of the system. They also propose, in another work (Miranda et al. 2014), a technique to recognize gestures which uses SVM to identify key poses, which then, are used to construct a decision forest to perform the recognition. In this later method they also propose the use of a circular buffer to accumulate key poses so they do not need to segment the gestures. Also, their method can, optionally, consider time/speed constraints by storing one or more time vectors in each leaf of the forest.

Finally, (Chaaraoui et al. 2014) propose to use an evolutionary approach in order to select the minimum subset of the body segments needed to recognize the gesture. For its recognition, k-means is applied to extract some key poses for each of the gesture that will be recognized, forming a bag of key poses. For each frame of the sequence which the recognition is applied, K-NN is used to match to one of the key poses, so there will be a sequence of key poses in which, then, will be applied DTW to get a value of how similar is this entry to each of the trained data.

Regarding the recognition methods used specifically for physiotherapy, most of the papers focuses on the application and evaluation of the proposed system, just vaguely explaining what is used to recognize the gestures. Also, a systematic review focusing on the AR and VR applications for physiotherapy using Kinect can be found at (Alana Da Gama et al. 2015).

Among those which detail its recognizing method, (Lin, Hsieh, and Lee 2013) normalizes the skeleton using the torso size and it uses the joint position and angle between segments to compare what is being done to what was recorded. It uses a score to determinate the similarity between poses, where the final score is the sum of all joints scores, and a joint will have one point if its position or its angle exceeds its threshold, or two points if both thresholds are exceeded.

(Jun et al. 2013b) uses PCA to reduce the dimensionality of the training data, and further uses K-NN to classify new gestures. While (Leightley, Darby, and McPhee 2013) also applies PCA, but it pre-processes the data, aligning all the skeletons according to the hip-center of the first frame and then calculating the velocity and energy of each joint. Velocity is the displacement of the joint related to five frames before, and the energy is the sum of the square velocity in each dimension. So, the data consists of the position, velocity and energy of each joint, which dimensionality is reduced using PCA and then a SVM and a Random Forest are trained and evaluated, where the RF had better performance.

3 GESTURE RECOGNITION SOLUTION

Following the goal of providing a tool for developers and therapists to build and use interactive systems that support the physiotherapy activity we developed a gesture recognition solution. The solution development is described within the following sections.

3.1 **Body Tracking**

With the intention of improving interaction systems without the necessity of any markers attached to the user's body, Microsoft launched the Kinect sensor (Microsoft Corporation 2016c). This sensor works with RGB-D technology, which makes use of depth information to track skeletal data. A 3D human motion capturing algorithm makes it possible to create interactions for users to control an application, such as a game, without the need to touch/hold a controller.

Kinect and the other RGB-D devices available in the market are receiving a lot of attention thanks to their portability associated with a fast human skeleton recognition system developed on top of 3D measurement (Cui and Stricker 2011). With the appearance of this technology, several studies were developed trying to apply it in different fields, for example, games, human body tracking (Izadi et al. 2011), 3D reconstruction (Freedman et al. 2013; Lange et al. 2011) and rehabilitation (Espay et al. 2010; Han et al. 2013; Kato, n.d.). There arose also studies to evaluate this technology, such as depth information precision (Cui and Stricker 2011).

3.1.1 RGB-D Sensors

Similar technologies to the Microsoft Kinect sensor were also produced. The first generation of RGB-D sensors includes besides the Microsoft Kinect, the Asus Xtion PRO LIVE and PrimeSense Carmine. These are based on a depth map generated using a projected pattern, according to a patent developed by PrimeSense (Payne et al. 2014).

The technology of these sensors combines structured light with computer vision techniques. The principle is simple; the emitter projects a known infrared light pattern into the scene and an infrared camera captures the result of this projection. The distortion of this light pattern allows the 3D depth map computation (Payne et al. 2014; Sell and O'Connor 2014). The pattern is pseudo random, what reduces the interference effects of using multiple sensors since it is not the same pattern emitted. Figure 1 shows the technique setup of depth mapping using structured light pattern with a light emitter and camera, and a depth scene represented by a hand.

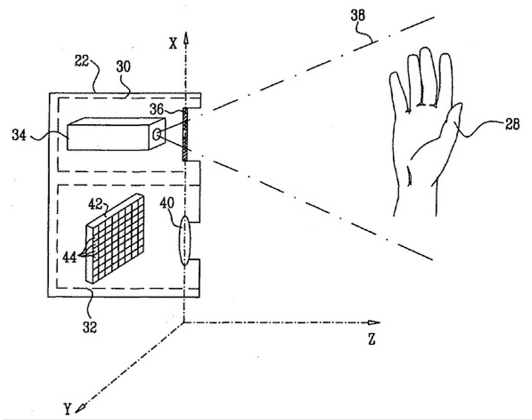


Figure 1. Technique of depth mapping using projected patterns (Payne et al. 2014)**.**

The second generation of RGB-D sensors combines the RGB camera with a time-of-flight (ToF) sensor that provides a depth image of the scene. The ToF camera computes depth information by measuring the time that a light pulse takes to travel from the camera to an object and back. Examples of this generation include the Microsoft Kinect second version (Kinect v2), SoftKinetic DepthSense (SoftKinetic 2016) and Intel Creative Senz3D (Intel Corporation 2016a). This technology was developed trying to provide high-resolution, low-latency, lighting-independent 3D image sensing.

3.1.2 Skeleton Estimation Software

The Kinect skeleton estimation software is an auxiliary library that receives information captured by the Kinect in order to perform skeleton estimation and provide skeleton tracking and joint positions. Actually there are two main tools to aid developers with Kinect sensor based implementation: OpenNI associated with Primesense's NITE software (PrimeSense 2016) and Microsoft Kinect SDK (Microsoft Corporation 2016c), which supports the first version of the Kinect through the SDK 1.X, and the second version of the device through its SDK 2.0.

One of the main differences between the two principal software is the platform in which they can be used. Microsoft Kinect SDK (MSSDK) is available only for Windows whereas OpenNI is a multiplatform and open-source tool. The number of joints tracked is also different: 15 joints with OpenNI, 20 joints with MSSDK 1.X (the five additional points are the two wrists, two ankles and the hip center) and 25 joints with MSSDK 2.0 (it also tracks thumb, hand tip and the neck). These three skeletons are shown in Figure 2. Additionally, MSSDK is able to track user´s upper limbs when lower body is not visible, allowing its use for scenarios where the user is sitting in a wheelchair. Despite these advantages, the MSSDK 1.X is more prone to false positives than the OpenNI, especially when the initial pose of a human body is too complicated, like squat. (Han et al. 2013).

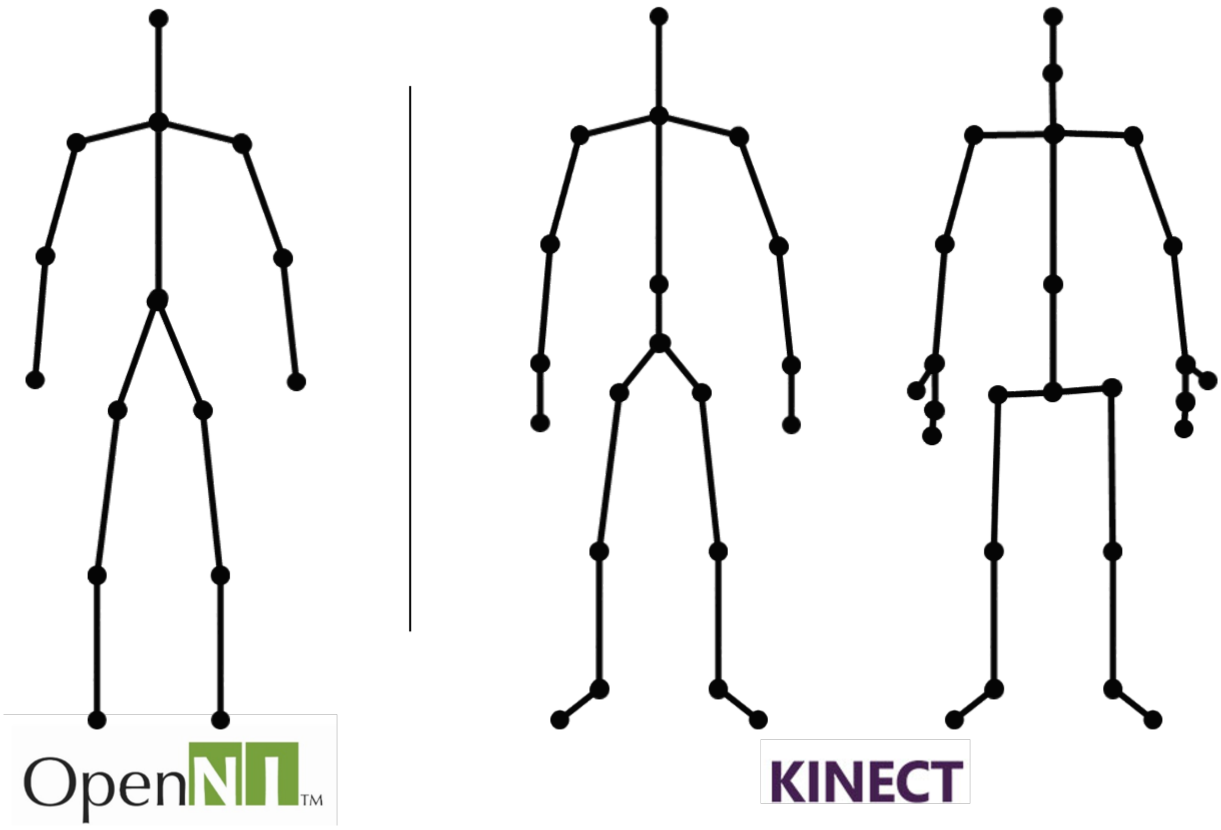


Figure 2. Devices skeletons. On the left the skeleton provided by the OpenNI, in the middle the one from Kinect SDK 1.X and on the right the one from Kinect SDK 2.0.

3.1.3 Kinect v1 Versus Kinect v2

This work started in 2011 just before the release of Microsoft Kinect v1 (Microsoft Corporation 2016b), which was November 2010. In July of 2014 Microsoft launched the second generation, Kinect v2 (Microsoft Corporation 2016b). During the period of development of this work, some limitations using Kinect v1 were detected, mainly related to biomechanical applications. Due to that, when the new generation of the sensor was acquired, we performed a preliminary test comparing Kinect versions analyzing these critical points in order to check the sensor improvements. The differences here presented are related to the skeleton tracking of Kinect v1 performed by the SDK 1.8 and Kinect v2 using the public preview 2.0 SDK. The comparison was performed with the aid of a physiotherapist in order to better identify the issues related to anatomy and biomechanics principles.

The first difference between the two Kinect versions is the number of joints. The new Kinect provides five extra skeletal points estimations: Neck, Fingers tip (right and left hand) and Thumb (right and left hand). The skeleton points tracked by both Kinect versions are presented in Figure 3. When using Kinect v1, one of the first problems detected was the behavior of the location of shoulders joints during movement of the arm, since the location is often misplaced configuring a tracking imprecision. This problem happens when the user moves his arm upper than 90 degrees being the location lower than the real shoulder position, as can be seen in Figure 3. The Kinect v2 shows a more accurate position and stable behavior of joint estimation during movement.

Observing the midline joints, there were some changes also. The head position for the Kinect v2 is located a little further than the other joints of the line; this can be visualized in the diagonal view in Figure 3. This makes the normal position recognized as the head to be tilted frontally. Only when the head is tilted back the position stays aligned. This is very important to be considered when performing motion analysis. This also happens with the Kinect v1; however, due the higher distance for the next point the inclination is smoothed. The spine center is also differently positioned. In the Kinect v1, it is positioned very low, at the lumbar spine. The new version presents a higher location for this point what seems to be a more adequate location. The first version provides a very distal reference resulting in a big gap of reference in all torso.

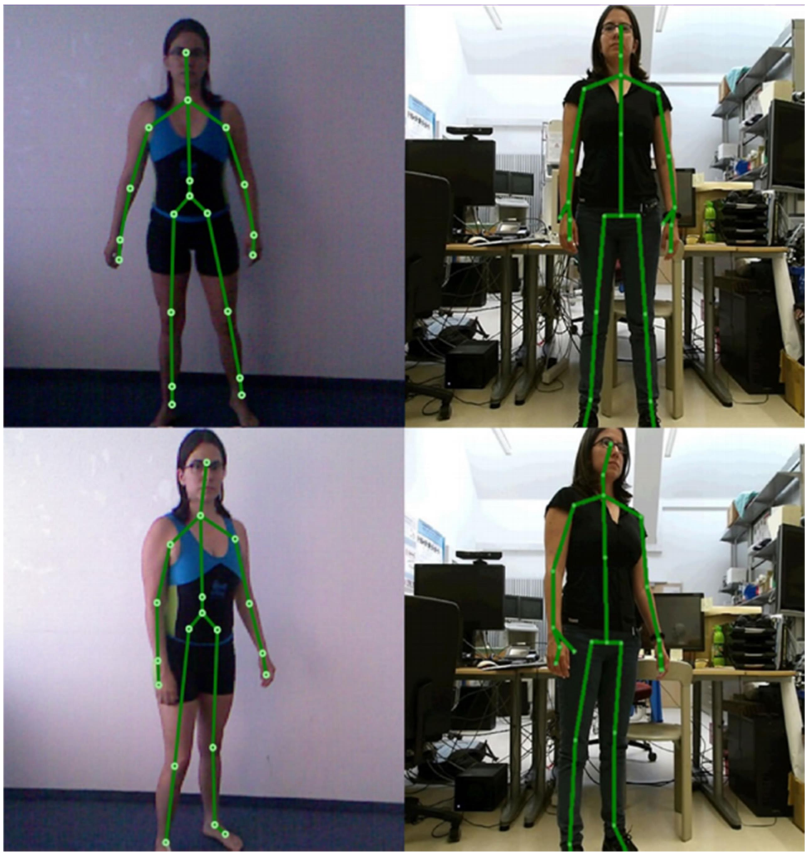


Figure 3. Kinect v1 (left) and Kinect v2 (right) skeletons tracking. Frontal (above) and diagonal (bellow) view.

By analyzing the upper limb region, another difference observed is that the scapular girdle movements can be detected using the new Kinect, what was not possible with the previous one. Figure 4 shows the skeleton tracking of Kinect v2 during scapular elevation. It is possible to notice that there is a change on joint's position during movement, which did not occur with the Kinect v1.

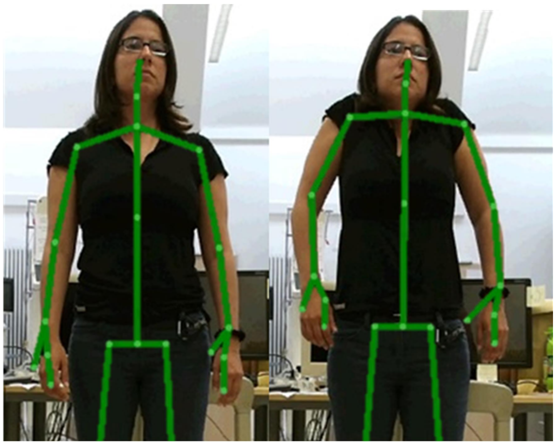


Figure 4. Skeleton tracking of Kinect v2 during scapular girdle elevation.

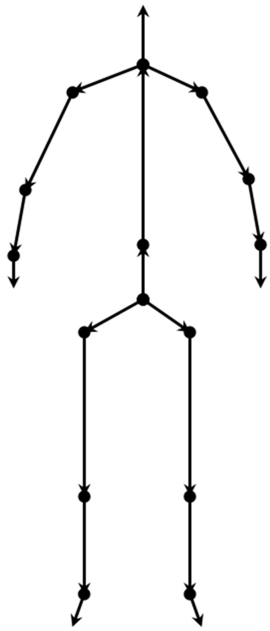
The spine center is also differently positioned. In the Kinect v1, it is positioned very low, at the lumbar spine. The new version presents a higher location for this point what seems to be a more adequate location. The first version provides a very distal reference resulting in a big gap of reference in the entire torso.

The same alignment problem that occurs in the new Kinect regarding the head position, also used to occur with the hip center estimation in the Kinect v1. The hip center in the old version is positioned a little frontally than the neighbor points. In the Kinect v2 this point is aligned with the others midline estimations. Figure 3 illustrates this difference. The performed tests show better skeleton recognition provided by the new Kinect generation, which solves some limitations of the first version (Alana Elza Fontes Da Gama 2015). However, anatomic accuracy is yet a limitation.

3.2 **Body Squeleton Representation**

For the first version of our system, the RGB-D device used was the Microsoft Kinect v1, and its SDK was used to recognize the skeleton of the user (the 3D points). Although, the data provided by the device is very sensitive to the user position and anatomy. If the user simply rotates his body, even maintaining fixed the relative positions between hands, arms, legs and other joints, the sensor will generate a completely new set of points' positions. This occurs because the skeleton 3D points are provided using as reference the sensor position and orientation. This problem also happens if two different users perform exactly the same movement, since the proportions of their body are not the same and thus, the position of the points will not be the same. This way, even if they were the same movement being performed, it would be a completely new set of information, thus, being a problem to recognize it.

To solve this problem, a new representation for the user's body data was needed. This was done by using normalized vectors and change of basis. Instead of using the 20 body points provided by the device, each segment of the body is represented by a vector, giving a set of 19 vectors as shown in Figure 5.



Vector	Basis
Neck	Torso
Shoulders	Torso
Thoracic Spine	Torso
Arms	Torso
Forearms	Forearm
Hands	Hand
Lumbar Spine	Torso
Hips	Torso
Thighs	Torso
Legs	Leg
Feet	Foot

Figure 5. The vectors that compose the skeleton representation and their respective basis.

Normalizing these vectors solves the problem of different users' anatomies but it is not enough to solve the spatial problem related to the position of both, the sensor and the user. The cause of this problem is in the origin of the coordinates used by the device. All coordinates are expressed considering the device as the center of the coordinate system. Therefore, a change of basis is needed but it cannot be the same basis for all vectors. For example, the position of the forearm depends on the arm's position, thus requiring its own basis based on the arm direction, which is different from the reference vector used for the leg. This way, it is needed a total of 9 different basis, which are: torso's, forearm's, hand's, leg's and foot's basis (the right members' basis is different from the left ones). Figure 5 illustrates the representation used in our recognition modules.

The following vectors compose the torso's basis: thoracic spine vector (I), the result (II) from the cross product between the thoracic spine and a vector from right to left shoulder, and the last vector is the cross product between (I) and (II). Also, the forearm's and leg's basis are alike, using the arm (thigh) vector (III), the cross product between (III) and a vector from left to right shoulder (hip) and the last vector being the result of the cross product between these two already calculated vectors that compose the basis. Lastly, the hand's and foot's basis are also alike and their basis is composed by the forearm (leg) vector (IV), the result (V) of the cross product between the forearm (leg) and arm (thigh) vectors, and finally, for the third vector, the cross product between (IV) and (V). Note that all these vectors are normalized so that they compose an orthonormal basis.

3.3 **Movement Analysis Library**

A movement analysis library was developed to provide developers and therapists the needed data for the respective applications. The library is designed to contain and provide easy access to the developed movement recognition techniques. It also provides functions to easy access sensor information. In order to use the library, it is necessary to set the movement recognition method and the skeleton tracking software that will be used. The skeleton tracking software can be chosen between OpenNI, Microsoft SDK (MSSDK) 1.8 and MSSKD 2.0.

Two movement recognition methods were proposed and are provided within the library. The first method is based on body-oriented planes, while the second is based on checkpoints (A.E.F. Da Gama et al. 2014). Both methods were designed with the aid of physiotherapists aiming to attend their needs while using a system capable of analyzing the patient motion. One of the raised requirements was to configure the precision of the recognition system, allowing the therapist to adjust it according to the patient needs. Also, understanding that motion should be focused on the exercises, dynamic gestures were the main target, and static gestures (namely, body postures) were not required.

Another interesting topic is that therapists were interested on providing continuous feedback during the gesture execution. Specifically, it was required for the system to provide the information of what percentage of the gesture was completed. Although most of related work is concerned to provide the output if a gesture is completed or not, in this case the path in between showed to be more relevant. Therapists showed to be more concerned about the repetition of the gesture, and the movement of going back and forth within a set of constraints. For instance, there are cases in which the patient does not complete the gesture, stops in the middle of it or just before the last intended position, but these cases must be tracked, in order to understand the patient difficulties and progress.

Moreover, the gestures are proposed not as a single execution through a trajectory, but as the trajectory itself. This means the patient can perform the gesture in both ways and that is still valid. This way the gesture is not placed as a unidirectional time performance, but as a trail, a path restriction in which the user is free to explore as long as his posture fits the gesture requirements regarding the set precision.

At last, two types of gestures were identified as requirements for the physiotherapy sessions. The first type of gestures is related to well-defined biomechanical movements, such as shoulder abduction. These movements are prescribed to exercise a particular set of muscles in a more controlled scenario to regain patient’s capabilities after an injury or stroke. The second type of gestures is related to functional activities, such as take a glass of water or unfasten the bra. These movements are prescribed mostly to allow the patient to regain functional capabilities, sometimes with less restriction to the way the gesture is performed but mostly concerned with the results (if the patient is capable to achieve that task).

The recognition methods are designed to attend these requirements. Two methods have been developed, being complementary to each other; the first is focused on the biomechanical movements while the second aims to allow the recognition of functional movements. Both recognition methods are better detailed on further subsections.

3.3.1 System Overview

Our system is structured as shown in Figure 6, composed by the following modules: Device, Checkpoints and Biomechanical Analysis which has an internal report module. . The Device module handles the RGB-D device functions, making it simple to change the used device. This is done by abstracting specificities from each device and standardizing the way data is provided. At the current version, the library supports any device that uses the OpenNI 2, the Kinect for windows SDK 1.8 or the Kinect for windows SDK 2.0. The Checkpoints module implements the technique for recognizing functional movements, while the Biomechanical Analysis is where the biomechanical recognition is performed and, as will be discussed later, it is also responsible for generating the physiotherapeutic session report, which is done by the Report module. Both this modules, Checkpoints and Biomechanical Analysis, also needs a configuration file as input, which are simple text files. Lastly, all the functionalities from all modules are abstracted in a layer of the library in order to deal with specificities so the application that uses it has easy access to the data provided by the device, the data from the recognition and the report functions. Aside from the data provided by the devices, the information the library gives to the application are: which gesture is being performed (since multiple gestures can be registered and tracked), the continuous value (i.e. a value that has the information of which part of the movement is being performed, the beginning, somewhere in the middle or if it is near the end of that gesture) and errors (which can tell about device problems like 'device not connected' or users problems like postural compensation detection, which will be explained later in this section).

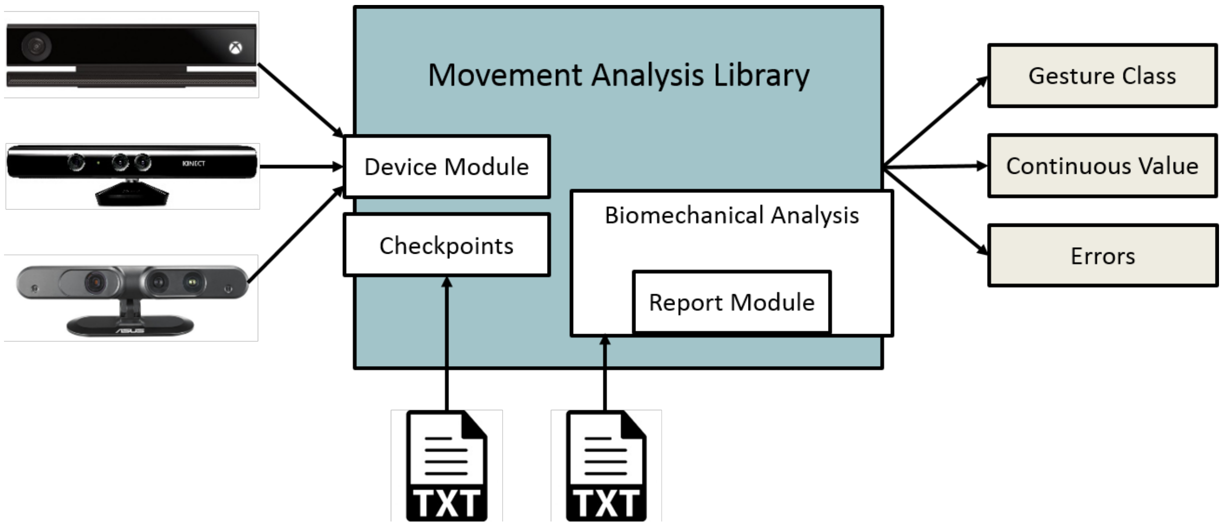


Figure 6. Motor rehabilitation support system architecture and its modules.

3.3.2 Movement Recognition By Planes

Applications for rehabilitation support require the movement to be analyzed according to physiotherapy needs. The description of a movement according to joint anatomy is provided by the International Society of Biomechanics (ISB) (Wu and Cavanagh 1995). The ISB standard uses bones, anatomical position and planes overlapped to the body in order to describe movement.

Human body movements occur in a three-dimensional space, which means that they can move through the three space planes. These planes are overlapped to the body, and receive specific names: frontal (XY), horizontal (XZ) and sagittal (YZ) (Wu and Cavanagh 1995). Movements on the frontal plane are named adduction or abduction, when the bone is approaching or moving away from the body, respectively; on the sagittal plane the movements are flexion (bones approaching) and extension (bones deviating), and the horizontal plane is where rotation movements occur (Clarkson 2005). In addition, each human joint has a Degree Of Freedom (DOF) associated to its movement, which indicates the number of planes where it is able to move. A single movement of a joint at one plane is a biomechanical movement and its angle is named Range of Motion (ROM).

In order to recognize the movements according to these requirements, the Biomechanical Analysis module gets as input the data from the calculated skeletal representation, as explained before, and calculates the ROM as being the angle between two vectors (Equation 1). This equation is derived of the dot product definition, just by rearranging the terms in order to isolate the angle. In this equation, V_1 and V_2 are vectors, the numerator is the dot product between them, the denominator is the product between their norms. However, this angle is independent of plane movement and to analyze the ROM it is necessary to guarantee that the movement is being executed within a specific plane.

$$\theta = \cos^{-1}\left[\frac{V_1 \cdot V_2}{||V_1|| * ||V_2||}\right] \tag{1}$$

Intending to classify a biomechanical movement in a certain plane, the angle between the moving and the normal vector of that plane should be 90 degrees, giving a Movement Tolerance Margin (MTM). The MTM can be specified by the therapist; it denotes the precision of the movement, meaning how far the user can deviate from the plane. Figure 7 shows a shoulder abduction on frontal plane, which is represented by the central position of hand, and the other two positions are the MTM which can be determined by the physiotherapist this way defining how far away from the plane the movement should still be acceptable.

Figure 7. Movement Tolerance Margim (MTM): how far away from the plane the movement can go.

With the information described before it is possible to classify each movement according to joint anatomy. The names of movements described by ISB standards are related to the direction where the bone moves during the movement and the number of movement described for each joint are dependent of its DOF. The movements that occur in the same plane, e.g. flexion and extension, are described with 1 DOF. The described movements for each joint are: for the shoulder, flexion, abduction and horizontal abduction are described. Horizontal abduction is an additional movement described for the shoulder where the arm moves on the horizontal plane (90 degrees from torso) going from the lateral to the front of the torso. Elbow is a joint with only 1 DOF, so only flexion is described for it. For the hip, flexion and abduction are described. The knee is a 2 DOF joint, nevertheless, only flexion can be described, since the rotation of this joint cannot be perceived externally. For the ankle, flexion is described.

All the information about how the movement should be performed is configured through a text file. In this file the physiotherapist can choose one from a list of biomechanical movements to be used for evaluation, where the application only checks if the performed movement is correct and saves this information, and/or interaction, i.e. when used as an input to, for example, an interface. The biomechanical movement is set by choosing the segment and the plane. The additional biomechanical and postural analyses are also defined in this file. The available configurations are:

- The body segment (vector) which will perform the movement and on what side;
- The plane of movement;
- The maximal angle that will correspond to the maximum movement on the application;
- The Movement Tolerance Margin (MTM);
- Postural acceptable angle: for torso, elbow and head;
- Axial rotation restrictions: arm or leg opening tolerance.

Figure 8 shows an example of a configuration file as well as the list of usable vectors and planes as explained before. It is important to notice that more than one movement can be described and used simultaneously. The difference is that the first described movement will be used as the controller of the application, and thus, it needs to have additional information about its maximum and minimum values, while from the second configured gesture to the last will be used only for evaluation. This can easily be further improved so all gestures have this information, if required by the application.

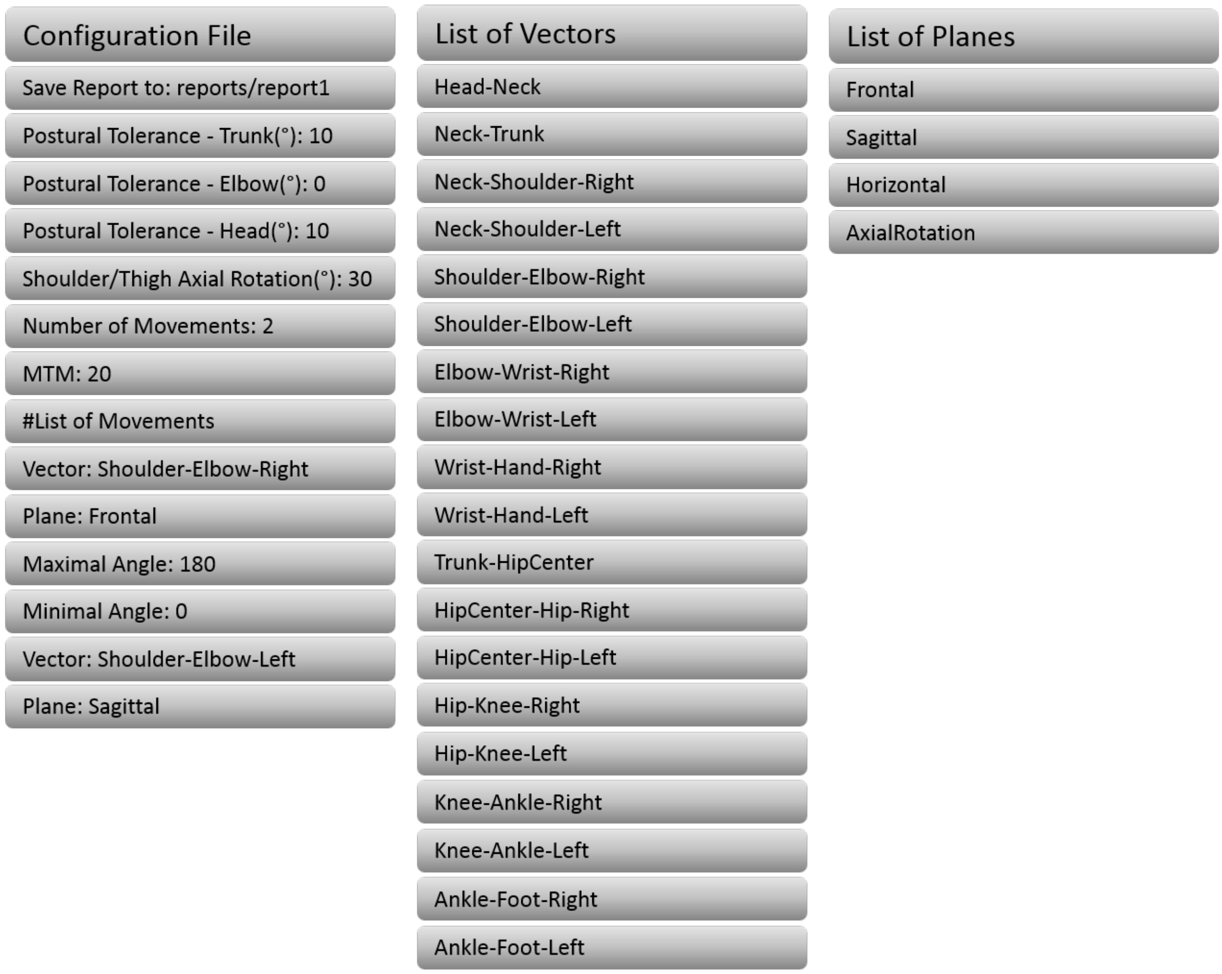


Figure 8. Configuration file and its lists of vectors and planes.

On the recognition solution based on planes, the percentage of performed gesture is given by the actual angle measured, in relation to the configured minimum and maximum angles, which are by its turn set by the therapist. These values are used to establish limits for interaction in order to enable the interactive system to be adaptable to patient limitations. For example, the maximum angle which will be required on the game. This way the system can support patient capabilities and the actions on the system can be limited to these values. Additionally, the Biomechanical Analysis module is responsible for reporting harmful postures as errors. The returned error includes the movement errors, which can be used by the application, for example, to exhibiting warnings. There are returned errors specific for the execution of biomechanics movements:

- Movement out of the respective plane: sagittal, horizontal and frontal;

- Arm or leg opened: for the axial rotations movements only;
- Plane not defined: when the plane set in the configuration is invalid;
- Angle not computed.

There are also generic errors related to general posture requirements for patients while performing both biomechanical and functional exercises:

- Compute postural analysis: check the torso alignment;
- Compute actual elbow straight: check if the elbow is stretched;
- Compute actual head straight: check if the head is aligned.

All these additional analyses are dependent on the value configured by physiotherapist. The acceptable angle for each of these positions is set in the configuration file. If the movement is being performed with these postures besides the configured the function returns the error for each of them: postural error, elbow not straight and head not straight, respectively.

The system provides a report after each session, which presents the biomechanical analysis results captured during the system execution in an accessible and documented way. The report provides physiotherapeutic information about the patient's performance during the use of the system. While the system is running, it is continuously receiving data from the biomechanical analysis and at the end of the session, the statistics measurements are computed including: maximal angle, percentage of time that the movement was executed incorrectly, if the movement was performed with postural compensation. Figure 9 shows the information provided by a game report after a patient uses the system for three minutes performing the shoulder frontal abduction of the right side movement. It can also provide a report of biomechanical movements even when functional movements are being used, meaning that biomechanical analysis and checkpoints technique can run simultaneously.

Figure 9. Physiotherapeutic session report file.

3.3.3 Movement Recognition By Checkpoints

There are several movements used in physiotherapy, usually functional exercises, which are not simple to describe using the planes concept. These movements consider one or more segments of the body, which moves not linearly, but describes curves, and each segment describes its own curve in space. Furthermore, there is a correlation between these curves, i.e., a point in a segment's curve must be bound to a specific point in the curve of another segment that together compound the movement. This is the problem that we tackle, describing the movement as a sequence of states (which was called checkpoint) saved as reference to be further compared to what the user is doing in real-time (Chaves et al. 2012) (this paper was awarded with the best paper price of the XIV Symposium on Virtual and Augmented Reality - SVR 2012).

In this method, a checkpoint is a set of one or more vectors of the skeletal representation in a slice of time. This way a set of checkpoints describes discretely the movement. The process can be split into two distinct steps as showed in Figure 10: recording and analysis. Recording is when the movement is registered by the system to further use. This is done by recording a user performing the movement and then, saving the checkpoints every time they are available, discarding the redundant ones. The analysis is the posterior step when the user is performing in front of the device and the system has to match the current movement to the one previously registered, evaluating if the current one is correct.

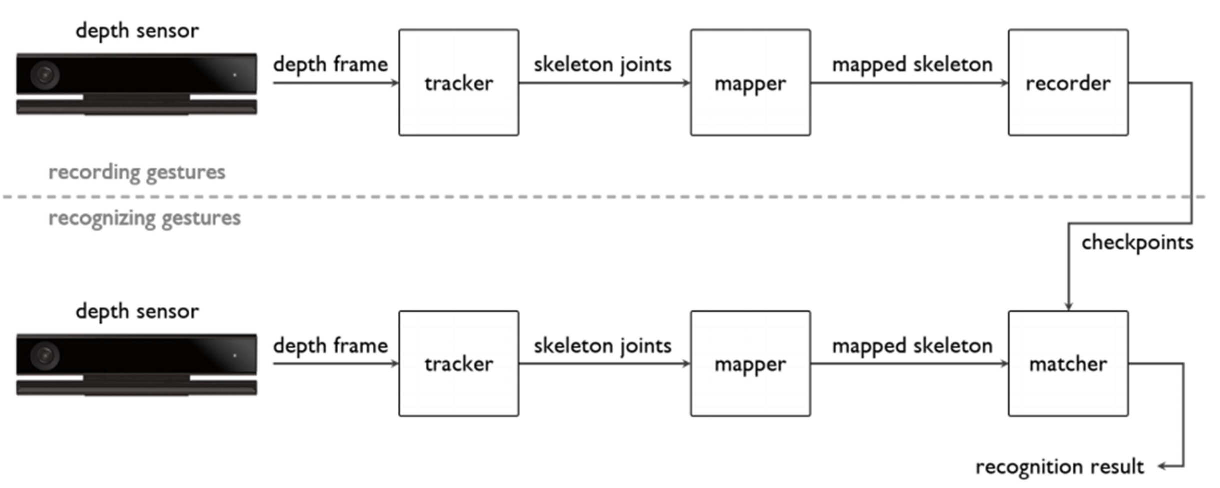


Figure 10. Overview of recording and movement analysis using the checkpoints method.

In order to do the analysis, the system tries to match the current state to one of the checkpoints saved before. It is important to notice that since the skeletal representation is a set of vectors, where each one has its own basis, a subset of the skeleton can be set to be relevant to the recognition. This way, to recognize a user waving his/her hand, all the system has to analyze the vectors from the hand and forearm, for example. To match two vectors, a range of acceptability is needed because almost never they will be at the exact same position; also, it is very hard to a human to perform the movement exactly the same way all the time. The saved states are discrete and each vector has its own range in a way that if the current movement's vector is within the range of a checkpoint, the current movement is validated and said to belong to that specific state. This way, the range defines the required precision of the gesture execution. Doing this, the information about the stage of the movement is gathered, i.e. if the user is performing the beginning of the movement, or its final, or it is in some point in the middle. In order to find the nearest state, a search algorithm is also performed by navigating through the set of checkpoints and checking if the current set of vectors matches the checkpoint that is being evaluated.

This can be enhanced to give a continuous state instead of a discrete one by comparing the current movement's vector to a line that links two subsequent checkpoints and finding the closest spot in this line that validates the movement. This requires a little change when comparing the current vector to the checkpoints. Instead of getting a single checkpoint and checking if the current vector is within its range, it is taken two checkpoints and calculated the minimum distance to the line that connects them. Figure 11 shows both methods, illustrating on the left the discrete checkpoints proposed in (Chaves et al. 2012), while on the right is the proposed improvement. Notice the area of recognition which better represents the path of the movement in the right part of the figure.

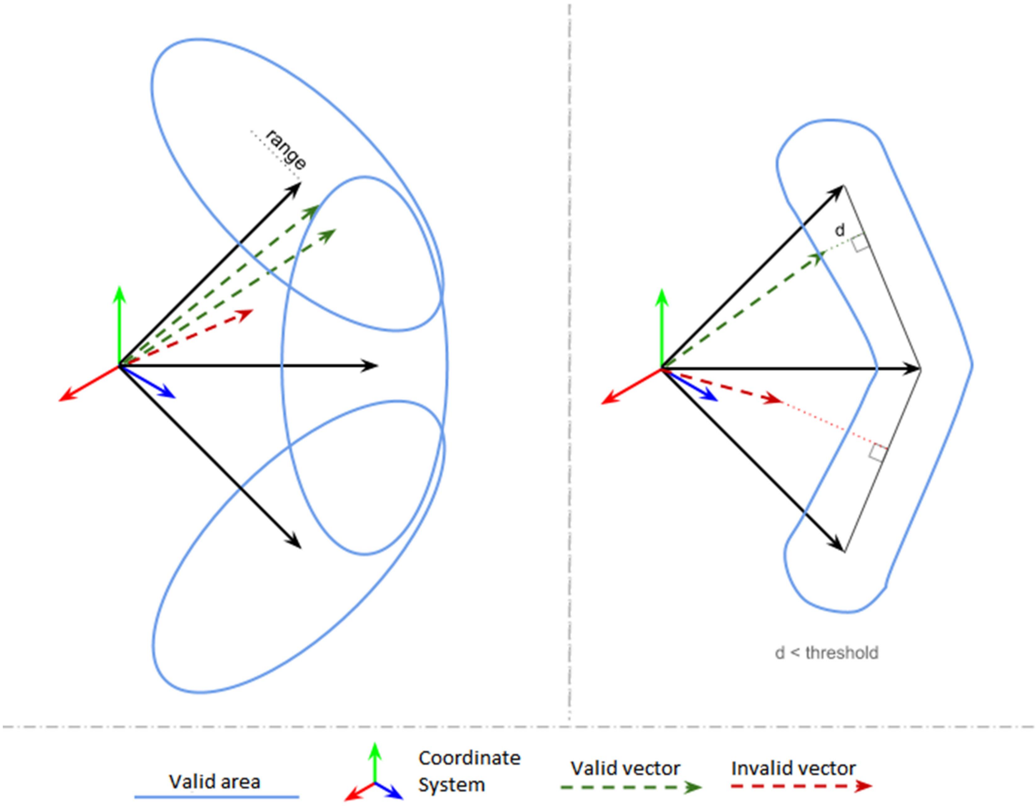


Figure 11. Difference between the discrete checkpoints and its new proposed method.

Figure 12 illustrates this process, where there are two checkpoints, C_1 and C_2, and also the vector from the current movement, P_1 (Figure 12A). In order to compute the minimum distance from P_1 to the line that connects C_1 and C_2, two other vectors are computed, C_1P_1 that is the subtraction between P_1 and C_1, and C_1C_2 that is the subtraction between C_2 and C_1 (Figure 12B). After this, it is calculated the projection of C_1P_1 on C_1C_2 (Figure 12C) using vector projection (Equation 2). This is the closest point between the vector of the current movement (P_1) and the line that connects the two checkpoints (C_1C_2).

Then, it is checked if the distance (Figure 12D) is greater than the range allowed. This distance is calculated by the norm (Equation 3) of the vector that is the subtraction between the projection and C_1P_1.

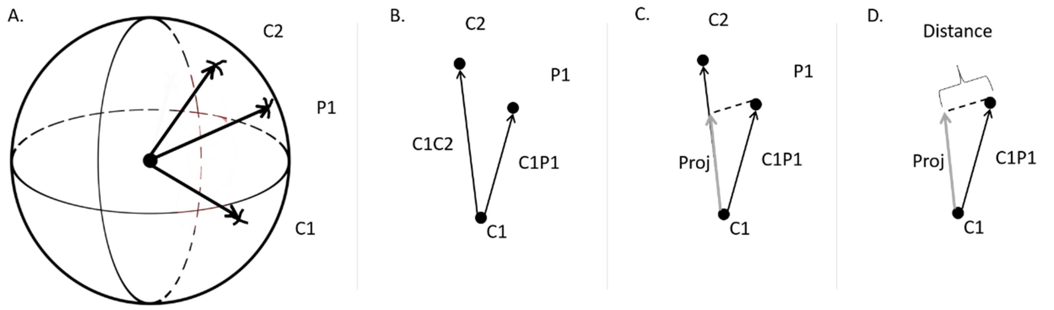


Figure 12. Calculating the shortest distance: A. Checkpoints and the current point; B. Auxiliary vectors; C. Projection vector; D. Shortest distance.

$$proj = \frac{C_1P_1 \cdot C_1C_2}{C_1C_2 \cdot C_1C_2} * C_1C_2 \tag{2}$$

$$norm = \sqrt{X^2 + Y^2 + Z^2} \tag{3}$$

For the functional exercises the configuration file is simpler and includes:

- Recording time;
- Checkpoints range;
- Segments which will be considered during movement execution;

Where the recording time is the time the user will have to perform the gesture in order to the system learn it, the range of the checkpoints as explained before and the segments that tell which part of the body will be used in that gesture. An example of this configuration file is given in Figure 13, where the list of vectors uses the same vectors as presented in the planes method in Figure 8.

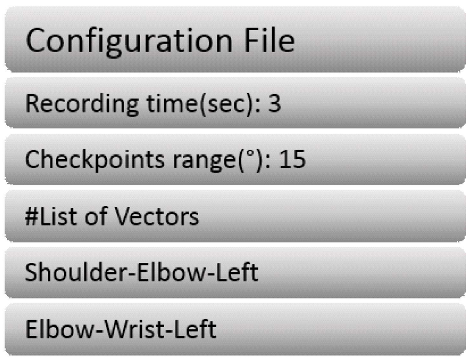


Figure 13. Configuration file of the checkpoints method.

3.4 Validation Tests

For both methods, we performed validation tests in order to verify if the recognition systems are capable of understand the user motion. In our case, the performer had no injuries or disabilities which could prevent the execution of proper physiotherapy movements. In addition, the performer was well instructed regarding the required motion. The precision of both methods (by planes and by checkpoints) was varied in order to understand which precision is suitable for a healthy user. Different gestures commonly used in physiotherapy were tested on both methods aiming to explore the range of applicability on each case. These tests, as also this work, was performed together with (Alana Elza Fontes Da Gama 2015), and are presented in the following subsections.

3.4.1 Biomechanical Moviment Tests

This section presents the experimental procedures performed to test the biomechanical movement recognition. The goal was to evaluate if the movement recognition technique is able to classify the biomechanical movements and this way detecting when they are being performed correctly or wrong. For that, the method was tested by performing all classified movements and checking the system capability to classify them correctly. It was measured by the percentage of correct exercises detected as right and wrong exercises as wrong ones.

To achieve that the movements were performed and analyzed using both Kinect versions, Kinect v1 and Kinect v2. The two Kinect were connected to a computer and the movements were performed by user standing in front of the sensor. The two Kinect were aligned in order to provide the nearest point of view as possible for the two sensors. The background was cleaned to allow skeleton tracking to work in the best condition. Figure 14 shows the setup used during the tests. Tests were performed on a computer with an Intel I7 4790k 4GHz processor, 32GB of RAM and Nvidia GTX 780 TI video chipset.

Figure 14. Setup for the biomechanical movement recognition tests.

In order to test the method, movements were performed in a correct and wrong way and the success rate of recognition by the system was scored. In order to guarantee that the movements were performed in a correct way the movements were executed by one physiotherapist specialized on biomechanics and with gymnastic preparation due to the fact that its practice and corporal conscience favors the performance of more precise movements. The person who performed the movements was carefully chosen and guided to perform them as perfect as possible since they would be interpreted as correct.

Each classified movement was performed 100 times: 70 times correct (35 at normal and 35 at fast velocity) and 30 times wrong (out of its respective plane). The tests were recorded using Kinect Studio 1.8 and 2.0, from both Kinect versions, enabled by the SDKs. The movements recorded could then be evaluated by the movement recognition method using different MTM guarantying the same movement in the different tolerances avoiding bias. The tests with the different MTM were performed trying to find the more adequate value for it in each biomechanical movement, where tracking and recognition have lower fail rate. During all movement performance the system was evaluating its execution in real time. To evaluate the data and compute success rate, graphics with the angles during movement were plotted and value of -20 was assigned when the movement was out of plane. When at any part of the movement this value was found the movement was computed as a wrong exercise. Movements performed with user facing the sensor. For the movements where occlusion could be a problem, such as the ones which occurs at Sagittal plane, axial rotations and shoulder horizontal abduction tests were also performed with user positioned rotated around 30 degrees to do not occlude joints during movement.

In order to analyze the data obtained from tests the success rates were computed. A descriptive analysis with percentage for each movement at different MTM was performed to present data. For the correct movements the value represent the number of correct exercises recognized as right and for the wrongs one the percentage of movement mistakes detected correctly. Since there was no different groups none comparative test were required.

The gesture recognition here proposed presented good capability to classify biomechanical movements for the majority of classified movements. Some joints presented limitation on its skeleton estimation being not possible to detect their movements. The method was also able to detect when the movement is being performed in a wrong way. This last feature is very useful for rehabilitation interactive systems which can make use of it to correct and guide patient during exercise (Alana Da Gama et al. 2012).

The use of movement angles measured using the segment and its JCS according with the ISB standards enabled to classify the movements according with the biomechanical concepts. It also allowed the recognition to work with user standing at different positions in relation to the sensor, including rotated and laterally displaced. This way, it is possible to provide the user a greater mobility during the use of the system, becoming one step closer to a natural interaction. Movements with different positions in relation to the sensor were also included in the correct performance test described before.

The results will present the recognition capability and success rate for each classified movements, and when necessary the limitations will be discussed. This section will first present the results by segment. For each movement, the results for the Kinect v1 and v2 will be presented. Just after discussion about the adequate MTM and an analytical analysis about the two sensors version will be provide.

CERVICAL SPINE

The movement recognition showed good results in classifying the cervical spine movements. Table 1 presents the success rate for the cervical classified movements using different MTM for both Kinects. The two classified movements, lateral flexion and the flexion and extension, presented small range of motion, around 33 degrees to movements at frontal plane and 20 degrees at sagittal plane, what can lead to a limited use of the movement as interaction control, which may be not so dynamic.

Table 1: Movement recognition for cervical spine. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2.

MOVEMENTS		CERVICAL SPINE 10 MTM KINECT		20 MTM KINECT		30 MTM KINECT	
		V1	V2	V1	V2	V1	V2
Lateral flexion	Normal	100	100	100	100	100	100
	Fast	100	100	100	100	100	100
	Wrong	100	100	66.7	90.0	0	23
Flexion / extension	Normal	71.4	100	100	100	100	100
	Fast	80.0	100	100	100	100	100
	Wrong	100	100	100	100	100	56.0

All movements at this joint were well recognized with 100% of success rate for all correct exercise at 20 degrees of tolerance. At 10 degrees of tolerance the Kinect v1 presented lower success rate than the new sensor version, however yet scoring higher than 70%. The lower scores for the wrong movements that occurs at 20 degrees of tolerance for the lateral flexion probably is consequence of the small range of motion at sagittal plane. If the movement at the opposite plane is lower than 20 degrees, the user cannot perform the movement out of the plane. This way it is possible to conclude that the lower success rate for these wrong exercises are caused by motion on the opposite plane not achieving the tolerance value.

SCAPULAR GIRDLE - CLAVICLE

Table 2 presents the results for the scapular girdle. The scapular girdle movements were not possible to track using the first version of the sensor. This fact is a consequence of an absence of change on shoulder joint estimation during scapular girdle movements, as shown in Figure 15.

Table 2: Movement recognition for scapular girdle. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2.

MOVEMENTS		SCAPULAR GIRDLE - CLAVICLE 10 MTM KINECT		20 MTM KINECT		30 MTM KINECT	
		V1	V2	V1	V2	V1	V2
Elevation and depression	Normal	0	68.7	0	100	0	100
	Fast	0	65.7	0	100	0	100
	Wrong	0	100	0	100	0	100
Protrusion and retraction	Normal	0	0	0	100	0	100
	Fast	0	0	0	94,3	0	100
	Wrong	0	100	0	23.3	0	0

Figure 15. Shoulder joint estimation during scapular girdle elevation.

Using the Kinect v2 it is possible to track the movements at this joint. The range of motion detected is small, mainly for elevation and depression, 15 degrees, and 27 degrees for protrusion and retraction. The short range of motion produces the same situation that occurred for cervical spine being almost impossible to perform wrong movements besides the tolerance value, since the amplitude of the wrong movement is shorter than the tolerance. So the recognition of wrong movements for protrusion and retraction cannot be performed. The use of 10 degrees of tolerance at this joint present low success rate for the elevation and depression, and zero for the protrusion and retraction. Using 20 degrees of tolerance 100% of success rate was found.

The movement dynamics when used for interaction can be limited due its low range of motion, mainly for the elevation, so the use of 20 degrees' tolerance is indicated. The correction, and this way accuracy of movement cannot be required when using the protrusion and retraction.

SHOULDER – ARM MOVEMENTS

The results for the shoulder movements' recognition are presented at Table 3. The movements at sagittal and frontal plane presented great success rate at 20 degrees. For the flexion and extension when using the Kinect v1 this result is better if the sensor is positioned 30 degrees from the user. Using 10 degrees tolerance the abduction and adduction continue working well, however with this tolerance the flexion movement works badly in both sensor. When using the sensor in diagonal the recognition for flexion and extension presents an improvement with the Kinect v2, but not for the first version.

Table 3: Movement recognition for the shoulder. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2 with sensor positioned frontally and diagonally.

SHOULDER – ARM MOVEMENTS WITH FRONTAL SENSOR							
MOVEMENTS		10 MTM KINECT		20 MTM KINECT		30 MTM KINECT	
		V1	V2	V1	V2	V1	V2
Adduc-tion / abduction	Normal	100	100	100	100	100	100
	Fast	80	77.1	100	100	100	100
	Wrong	100	100	100	100	100	100
Flexion and extension	Normal	14.3	60	100	100	100	100
	Fast	0	34.3	97.1	100	97.1	100
	Wrong	100	100	100	100	100	100
Horizontal adduction abduction	Normal	65.7	94.3	65.7	100	94.3	100
	Fast	11.4	80	60	91.4	80	97.1
	Wrong	100	100	100	100	80	86.7
Axial rotation	Normal	0	0	45.7	74.3	94.3	100

	Fast	0	8.57	57.1	82.9	97.1	100
	Wrong	100	100	100	100	100	100
SHOULDER – ARM MOVEMENTS WITH DIAGONAL SENSOR							
MOVEMENTS		**10 MTM KINECT**		**20 MTM KINECT**		**30 MTM KINECT**	
		V1	**V2**	**V1**	**V2**	**V1**	**V2**
Flexion and extension	Normal	0	100	100	100	100	100
	Fast	0	88.6	100	100	100	100
	Wrong	100	100	100	100	100	100
Horizontal adduction abduction	Normal	77.1	94.3	100	100	100	100
	Fast	91.7	94.3	100	100	100	100
	Wrong	100	100	100	100	86.6	96.7
Axial rotation	Normal	14.3	17.4	25.7	77.1	100	100
	Fast	0	8.6	17.1	74.3	57.1	100
	Wrong	100	100	100	100	100	100

For the movements at the horizontal plane the shoulder horizontal adduction and abduction present reasonable recognition using the Kinect v1 and good results with the new sensor version. The main limitation was in the use of Kinect v1 for fast movements. Using the diagonal sensor position, the movement is well recognized by the two sensors versions.

The axial rotation cannot be recognized with the 10 degrees of tolerance, probably due the difficulty of maintaining the arm on the parallel position without any additional reference. Using the 20 degrees tolerance only the Kinect v2 presented reasonable success rate on recognition with sensor in both position. Good recognition for both sensors was found using the 30 degrees tolerance.

The better results found for the Kinect v2 even in front of sensor in the occlusion situations show a better joint estimation performed by the new version on tracking movements when the joint is not being visible.

ELBOW – FOREARM MOVEMENTS

The elbow is a monoaxial joint and its anatomy enable movement only at the sagittal plane (Kisner and Colby 2012). This means that the joint cannot perform the movement out of the desired plane and it is naturally performed perfectly at the sagittal plane. Due that 100% of success rate was found for all situations tested, these results are presented at Table 4.

Table 4: Movement recognition for the elbow. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2.

ELBOW – FOREARM MOVEMENTS							
MOVEMENTS		**10 MTM KINECT**		**20 MTM KINECT**		**30 MTM KINECT**	
		V1	**V2**	**V1**	**V2**	**V1**	**V2**
Flexion / extension	Normal	100	100	100	100	100	100
	Fast	100	100	100	100	100	100
	Wrong	100	100	100	100	100	100

WRIST – HAND MOVEMENTS

The main problem in recognizing hand movements is the oscillation of this joint estimation. The results for this joint is presented at Table 5. For movements at the sagittal plane it is not possible to recognize a movement accordingly. The flexion and extension are reasonable recognized only using 30 degrees of tolerance since it is not being performed fast. At this tolerance it is not possible to detect if the movement is being performed in a wrong way. The wrong movements are performed at frontal plane and they have a range of motion smaller than the 30 degrees tolerance, what make not possible to detect the wrong movements. If the user wants to use the hand flexion and extension movement, it is only indicated using the 30 degrees tolerance in a normal velocity and mainly when just interaction with no control of movement is required.

Table 5: Movement recognition for the wrist. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2.

MOVEMENTS		WRIST – HAND MOVEMENTS 10 MTM KINECT		20 MTM KINECT		30 MTM KINECT	
		V1	V2	V1	V2	V1	V2
Radio and ulnar deviation	Normal	31.4	60	31.4	100	57.1	100
	Fast	0	11.4	5.71	74.3	45.7	85.7
	Wrong	100	100	100	100	83.3	100
Flexion / extension	Normal	0	0	0	0	54.3	77.1
	Fast	0	0	0	0	0	5.71
	Wrong	100	100	26.7	96.7	0	0

The movements at frontal plane present better recognition using the Kinect v2. The radial and ulnar deviation can be well recognized using this sensor at 20 and 30 degrees tolerance. The Kinect v1 presents limited capability in detecting these movements, probably due the higher instability of the joint estimation.

SPINE

The spine movements are well recognized using both sensors at 20 and 30 degrees tolerance. The results are presented at Table 6. The Kinect v2 also presented good recognition using 10 degrees tolerance except for the fast movements. The main limitation for spine movements recognition is in the correction of the lateral flexion, which once again is restricted due short the range of motion at the opposite plane. This way the wrong movement did not achieve the tolerance angle to be detected. For the spine lateral movements when correction is required the Kinect v2 is indicated since it can work at 10 degrees tolerance where the wrong detections have good success rate. The flexion and extension movements at 10 degrees tolerance works well only using Kinect v2 at normal velocity movements. When using Kinect v1 the fast movements at sagittal plane should be avoid.

Table 6: Movement recognition for the spine. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2.

MOVEMENTS		SPINE 10 MTM KINECT		20 MTM KINECT		30 MTM KINECT	
		V1	V2	V1	V2	V1	V2
Lateral flexion	Normal	42.9	88.6	100	100	100	100
	Fast	17.1	22.9	82.9	91.4	100	100
	Wrong	100	100	33.3	26.6	0	0
Flexion / extension	Normal	45.7	100	88.6	100	100	100

	Fast	14.3	57.1	54.3	100	100	100
	Wrong	100	100	100	100	86.7	96.7

PELVIS

The movements performed with the pelvis are very smooth. The joint estimation performed by the sensor is based on user pose. During the pelvis movements very small changes on the user poses occurs, this way the joint estimation does not change. Due that, no movement is detected at this joint. The result was similar for the two sensor versions and the joint estimation during the movement can be seen in Figure 16.

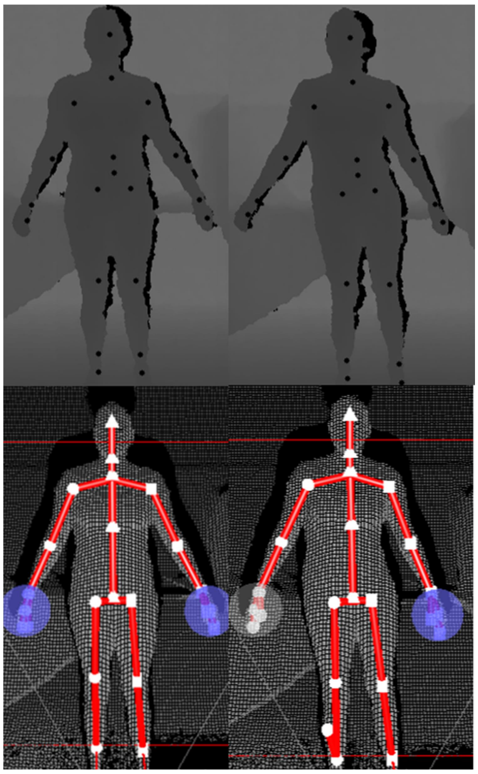


Figure 16. Skeleton estimation during pelvis elevation for Kinect v1 (above) and Kinect v2 (bellow).

HIP – THIGH MOVEMENTS

Results for this joint are presented at Table 7 including the tests with sensor positioned frontally and diagonally. Based on them it is possible to detect that the thigh movements can be all well recognized using the Kinect v2 at all MTM. For the recognition based on this sensor only the fast axial rotations at 10 degrees tolerance present low success rate. When using 30 degrees tolerance attention should be given to the lower movement accuracy required risking false positives.

Table 7: Movement recognition for the hip. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2 with sensor positioned frontally and diagonally.

HIP – THIGH MOVEMENTS WITH FRONTAL SENSOR							
MOVEMENTS		10 MTM KINECT		20 MTM KINECT		30 MTM KINECT	
		V1	V2	V1	V2	V1	V2
Adduction / abduction	Normal	31.4	100	100	100	100	100
	Fast	20	100	100	100	100	100
	Wrong	100	100	100	100	100	70
Flexion and extension	Normal	100	94.3	100	100	100	100
	Fast	94.3	97.1	100	100	100	100
	Wrong	100	100	100	100	100	100
Axial rotation	Normal	0	85.7	0	100	0	100
	Fast	0	42.9	0	100	0	100
	Wrong	100	100	100	100	100	100
HIP – THIGH MOVEMENTS WITH DIAGONAL SENSOR							
MOVEMENTS		**10 MTM KINECT**		**20 MTM KINECT**		**30 MTM KINECT**	
		V1	**V2**	**V1**	**V2**	**V1**	**V2**
Flexion and extension	Normal	97.1	100	100	100	100	100
	Fast	97.1	97.1	100	100	100	100
	Wrong	100	100	100	100	100	100
Axial rotation	Normal	0	85.7	0	100	25.7	100
	Fast	0	65.7	0	100	0	100
	Wrong	100	100	100	100	100	100

When using the Kinect v1 almost all movements can be well recognized at 20 degrees tolerance, except the axial rotation. The problem with this sensor version to recognize the axial rotation is the skeleton tracking with the position required – the 90 degrees hip flexion. The knee joint estimation at this position is located down on the leg and the hip in a higher position. This combination makes the segment of the thigh positioned diagonally for down, making the 90 position not achieved. Figure 17 shows the skeleton tracking at this position for both Kinect sensors. After observing this it was performed a test using 40 degrees tolerance and at this case all movements were recognized. However, the in general 30 degrees already present the problem of false positive, when using 40 degrees no accuracy can be guarantee.

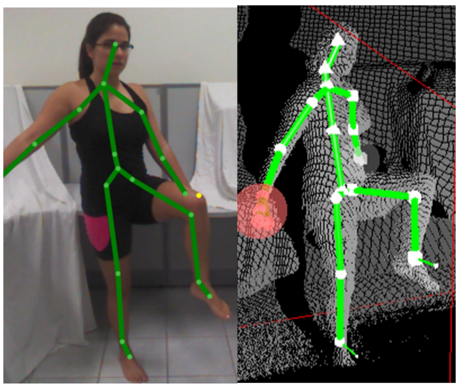


Figure 17. Kinect v1 (left) and Kinect v2 (right) skeleton tracking during axial rotation recquired position

KNEE – LEG MOVEMENTS

The same that occurs with the elbow, the knee joint do not perform movement at the frontal plane (Kisner and Colby 2012) resulting 100% success rate in all evaluated situations (Table 8). This anatomic characteristic makes impossible to perform movement at the wrong plane. Differently form the elbow, the knee is a biaxial joint, however the second movement that this joint perform is the axial rotation which is not detected during skeleton estimation.

Table 8: Movement recognition for the knee. Succsess rate (%) at different Movement Tolerance Margins (MTM) with kinect v1 and kinect v2 with sensor positioned frontally and diagonally.

KNEE – LEG MOVEMENTS WITH FRONTAL SENSOR							
MOVEMENTS		10 MTM KINECT		20 MTM KINECT		30 MTM KINECT	
		V1	V2	V1	V2	V1	V2
Flexion / extension	Normal	0	100	0	100	0	100
	Fast	0	100	0	100	0	100
	Wrong	0	100	0	100	0	100
KNEE – LEG MOVEMENTS WITH DIAGONAL SENSOR							
MOVEMENTS		**10 MTM KINECT**		**20 MTM KINECT**		**30 MTM KINECT**	
		V1	**V2**	**V1**	**V2**	**V1**	**V2**
Flexion / extension	Normal	100	100	100	100	100	100
	Fast	100	100	100	100	100	100
	Wrong	100	100	100	100	100	100

ANKLE – FOOT MOVEMENTS

The foot movements could not be recognized using any of the Kinect versions due absence of movement at joint estimation. When using the Kinect v1 during the dorsiflexion the point estimate for the foot goes to the leg and during the plantar flexion it returns to the foot. The vector which connects the ankle to the foot do not change. This situation can be seen on Figure 18. For the Kinect v2 the shift of foot joint estimation to the leg does not occur, however the movement yet cannot be detected due absence of joint changes during movement (Figure 18).

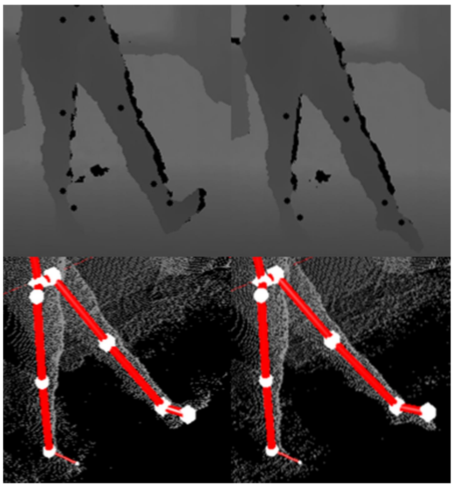


Figure 18. : Skeleton estimation during foot movements for Kinect v1 (above) and Kinect v2 (bellow). The left image shows foot during dorsiflexion and the right the plantarflexion.

ADEQUATE MOVEMENT TOLERANCE MARGIN

With the results presented it is possible to notice that the use of 10º MTM makes the recognition unstable. This occurs because although the movements are described in planes, the performance of them exactly at the plane during all trajectories is utopic. Besides that, the joint estimation performed by the sensor can have little oscillations even to the joints which have good tracking, what can also lead the segment to out of the plane if the tolerance is very restrict.

The use of 10º tolerance should be used when extremely accuracy is required during movement, since it involves perfection. The movements which present good success rate at this MTM are: Cervical movements, shoulder abduction and flexion, the last one with the sensor at diagonal, elbow flexion, hip flexion, and knee flexion. The ones which worked well only with the Kinect v2 at this tolerance was: shoulder horizontal abduction and the hip abduction,

In an opposite way, the 30º MTM presented great success rate when detecting correct movements, since it gives more movement freedom. However, it can in some case cause a false positive. When performing wrong exercises, the system failed detecting them as correct at this range. This means that this range starts to give excess of freedom and lower control of movement accuracy. The movement which showed this false positive situation in higher degree includes: cervical movements, scapular girdle protrusion, wrist flexion, and spine lateral flexion.

The difficult in detecting the wrong movements was also found in some case when using the 20 degrees tolerance. This was caused in the joints which present the opposite movement with range of motion smaller than the MTM. Since the wrong movement could not even achieve the tolerance angle it could not be detected. This situation happened with the

The ideal MTM is located at 20º MTM, presenting 100% success rate in detecting correct exercises for almost all classified movements. The success at this range is more frequent when using the Kinect v2. The first version of the sensor presents less success at this tolerance for movements where occlusion occurs, for example shoulder horizontal abduction. This fact shows improvements on the new sensor in relation to joint estimation when the joint is not visible.

It is important to notice that the success rate is related with the capacity of user to perform the movement in an accurate way. For example, shoulders axial rotation requires a fixed position which is difficult to maintain precisely without additional reference. In case where the movement is very difficult to be performed as standardized the use of larger MTM in interactive systems is suggested in order to provide more usability.

KINECT V1 AND KINECT V2

With the first Kinect launch in the end of 2010 a lot of studies applying this technology as a tool for develop interactive applications for motor rehabilitation started to be developed. Recently, in July 2014, Microsoft launched the new sensor version, the Kinect v2. The benefits of using such instrument which provides natural interaction associate with a portable and low cost characteristic is known. However, which of the physiotherapeutic movements work well or not when using this technology is not defined yet. Since this work developed a movement recognition which is capable to recognize the classified biomechanical movements for each joint and tested them it was possible to perform an analytical analysis about the Kinect applicability for rehabilitation purposes.

The use of the Kinect sensor as a skeleton tracking tool for biomechanical movements presented good results for the majority of movements classified, however limitations were found. Some of movements presented limited recognition when using the Kinect v1. However, with the improvements of the new version, the recognition using the Kinect v2 worked better for all movements being able even to recognize some movements that were not possible with the first version.

Movements which could not be detected using the Kinect v1 but are recognized using the new version includes clavicle, wrist and knee movements, being the restriction of this last one only if the user is frontally for the sensor. There are some movements which could be recognized however with bad detection. These include the shoulder horizontal abduction, shoulder rotation, spine flexion and hip rotation. These movements presented better recognition when using the Kinect v2. The pelvis and ankle movements could not be detected using any of the sensors versions.

LIMITATIONS

Limitations are mainly related to the sensor and the markerless technique capability. One of the problems occurs when there is occlusion; there is an inaccuracy due to indirect estimation performed by the sensor when one body part covers the view of another. The new version of sensor works better on these situations, but yet are the movements with more difficult on recognition. The occlusion problem may be improved with the use of multiple sensors.

Another limitation is the detection of axial rotation. Since the system detects joint position based on pose estimation, the bone rotation around its own axis does not change the visual pose and no change on joint location is found. So, it is not possible to detect these movements by the method proposed. The additional method developed using a fixed position as reference to measure the angle was used for the two limbs main joints, shoulder and hip. Unfortunately, the same could not be extrapolated to the other joints where axial rotation occurs due the absence of additional references.

3.4.2 Functional Moviments Tests

Initial tests using the checkpoints technique were performed and presented in (Chaves et al. 2012), these tests proved that the checkpoints were capable of identifying a gesture and also showed that it can give an output in less than 1ms which allows its use in real-time applications. But the checkpoints technique still needs to prove its efficacy when used for physiotherapy purposes. In order to do it, and as this technique was proposed to be complementary to the planes one, the new tests involve identifying functional gestures used in physiotherapy sessions.

The results of movement recognition by using checkpoints explore functional and multi-joint movements and are presented in Table 9. Due to the greater complexity of these movements compared to the biomechanical ones, the results showed the necessity of a higher range, being 20 to 30 degrees the range where there are better success rates. In order to complement this evaluation, and to check if these higher ranges will interfere on clinical applicability, five physiotherapists evaluated the executions by observation, similarly as the procedure performed on a clinical treatment. These results are presented on the subjective evaluation column showed at the right side of Table 9.

The performed movements used in the tests are the following:

- The Diagonal of Prorioceptive Neuromuscular Function (PNF) movement is similar to unsheathe a sword and point it up, so the hand starts at hip height on the opposite side (if using the right hand it is next to the left side of the hip) and goes to its final point that is pointing almost straight up (slightly to the right if using the right hand);
- Hand to head, which is similar to grabbing a comb and combing the hair;
- Hand to back is similar to a woman opening/closing her bra;
- Throw is simply toss an object;
- Glass to mouth is a gesture like taking a cup and drinking from it;
- Codman movement needs the user to be standing up but bent and supported on a table or chair in a way that the arm is freely hanging towards the floor but it makes almost a 90 degree to the chest;
- Kick;
- Squat;

All tests were performed on a Notebook Avell, 2.6 GHz processor (i7-3720QM), 8GB DDR3 of RAM and a video chipset GeForce GTX 670n. Tests were executed in an empty room, i.e., without any objects interfering in the Kinect tracking area, being only the user's body in the sensors field of view. The Kinect v2 was not launched at the time these tests were made.

Table 9. Success rate and subjective evaluation of movement recognition by checkpoints.

Movement	Success Rate			Subjective evaluation of clinical applicability		
Range	**10º**	**20º**	**30º**	**10º**	**20º**	**30º**
PNF Diagonal	20%	40%	100%	Very Bad	Medium	Good
Hand to head	90%	100%	-	Medium	Good	Very Good
Hand to back	0%	0%	70%	Bad	Bad	Medium
Throw	40%	80%	90%	Medium	Good	Very Good
Glass to mouth	60%	90%	100%	Bad	Good	Good
Codman	90%	100%	-	Medium	Good	Good
Kick	60%	100%	-	Medium	Good	Good
Squat	70%	100%	-	Medium	Good	Good

The most difficult movement to recognize was the hand to back, which presented a maximum of 70% of success rate at 30 degrees, being very difficult to track at lower ranges. This probably occurs due to the self-occlusion of the hand by the user body, which makes it difficult for the Kinect sensor to differentiate one from another. This result corroborates with physiotherapists evaluation, which gave this movement the lowest evaluation with medium applicability at 30 degrees range. This problem also may be solved by the use of two Kinects for skeleton tracking, and as said before without necessity of change movement recognition techniques.

The movements which presented better recognition rates, with greater success rate even at the lowest range, were the hand to head and Codman movements. However, according to therapists, using only 10 degrees of range is not the best for clinical applicability. At the range of 20 degrees, most movements are good to clinical application and presented good success rate, except for the hand to back that had problems and PNF Diagonal which only has a low recognition rate at 10 and 20 degrees and only has a good recognition rate (of 100%) at 30 degrees of range.

The two methods presented good potential for applicability on a rehabilitation system. According to the results, the tracking by planes showed good movement recognition with lower MTMs (10 to 20 degrees). Also, the recognition by checkpoints proved to be a good solution for functional and multi-joints movements that are not covered by the planes method. However, to be usable it requires tolerance ranges of 20 to 30 degrees.

4 CASE STUDIES

The aim of this work is the proposition and development of movement recognition methods that make feasible a rehabilitation system more clinical related in order to improve motor rehabilitation process. The movements recognition methods developed are capable to provide movement information in real-time. This knowledge about movement can be used in the systems to provide interaction. This way, the movements used on the physiotherapy, according to the methods recognition, can control an interactive system that at the same time entertains patients and leads them through exercises of a treatment protocol.

The case studies here developed consist of interactive systems that make use of the library established as a tool for movement recognition and analysis. The systems have two main goals: exercise performance according to clinical routine, including the biomechanical and the functional movements, and guidance and correction of the movement execution. This orientation can be useful as a complementary physiotherapy and mainly to home care therapies, where the physiotherapist cannot supervise the patient and patients have to execute the rehabilitation program alone.

The first case study is a Virtual Reality (VR) system, named reAIRabilitation, which will be presented in section 0. The second, mirrARbilitation is based on Augmented Reality (AR) and is described in section 4.2. Aiming to expand its applicability and demonstrate its range of use the recognition library is also explored on two additional case studies not related to the physiotherapy domain on sections 4.3 and 4.4.

4.1 **Case Study 1: Reairabilitation**

The first system developed to validate the findings of this research is an interactive game where the main character can be controlled by the user movement. The game starts with a first version named Dolphin´s Adventure and then it was updated to the reAIRabilitation version. It is worth to mention that the reAIRabilitation already was awarded twice, as First Place of the Games For Change Award and as Best Game Award in the Other Platforms Category, both of the 11th Indie Games Festival 2012 of the XI Brazilian Symposium on Computer Games and Digital Entertainment – SBGames 2012. The steps of the development process are presented in this section.

The game makes use of the recognition module, working with both gesture recognition methods, and the features of report and configuration. The library is accessed by the game, which can use the data extracted for interaction and/or to give the patient a corrective feedback.

4.1.1 Game Mechanics

The initial game concept was to enable the patient to control the main character of a game using physiotherapeutic (biomechanical) or functional movements, which are the same used during traditional physiotherapy. The game’s mechanic has been developed to induce the physiotherapeutic sequence of movements and repetitions.

In order to achieve that, game dynamics understands the patient’s movements to control the vertical motion of the main character. The patient has to make the main character catch some elements and avoid others, both coming from the opposite direction of the screen. Positive and negative feedbacks are given depending on the success of the user on performing these tasks. This way the user has a real motivation to perform the necessary moves.

For rehabilitation applications, one important characteristic is that the movement that controls the character of the game could be scaled and graduated, according to patient limitations. This way the maximal patient mobility will correspond to the maximal motion of the character. For example, the physiotherapist configures the game for shoulder abduction, which occurs in the frontal plane as explained before, and determines that the maximum ROM for the patient is 90 degrees. Using this configuration, the game will interpret and respond accordingly as a full movement when patient abduction is at 90 degrees. The game configuration, including the movement that will be used to control the game and the maximum and minimum ROM, is set with the configuration file described before.

4.1.2 Interface Evolution

The interface of the system was designed in two phases. Firstly, a prototype was developed and then, after the system and this interface passed through user tests, the second version was made. The first version of the game was focused on validating the hypothesis that a game specifically designed for physiotherapy rehabilitation with continuous feedback for the patient gestures is valid. It was defined a simple game and set of requirements, thereupon it was necessary to test if this concept had value to the patients and to the physiotherapist. With this goal the first version of the game was created, the Dolphin's Adventure. As the focus of this version was not specifically on the user's satisfaction with the graphics, the effort on creating high quality graphics, meaningful story and characters and other well-known characteristics accepted by the games market was not considered.

With this prototype developed, tests were made to evaluate it, in which all the characteristics of the system (technology and interface) were considered. After the results of these tests and all the user feedbacks being collected, synthesized and studied, the development of the final version was initiated and then tested to measure the improvements made in the system compared with the first version. In this subsection will be described how these project steps were conducted, focusing on the graphic features and interface of the system.

FIRST VERSION: DOLPHIN´S ADVENTURE

The theme of the first version of the game (Alana Da Gama et al. 2012) was chosen based on movement characteristics. As most of the moves to be made by the user should follow trajectories on the vertical axis, it was necessary that the character controlled by the user had its main moves on this axis too, making the system more intuitive. It was also important the use of a continuous movement enabling user to work all ROM during physiotherapy. Thus, an aquatic environment was designed for the game.

Knowing that the game is based on an underwater scenario, the main character was defined to be a dolphin, easily accepted as a friendly icon of this environment. The left part of Figure 19 shows this character and the scenario. As explained in the game mechanics, the user will be induced to catch some elements appearing on the screen in order to stimulate the movement. The characters chosen for this purpose were fish coins. In addition, to improve the interaction with the system and user motion, the element the user will have to avoid is a submarine and a piranha (left part of Figure 19). These objects will be moving on the X-axis from right to left direction.

Figure 19. On the left, Dolphin's Adventure scenario and graphic elements. On the right, catching (fuel box) and obstacles (stormy clouds) objects to induce user motion in the reAIRabilitation game.

The score, game time, user's movement angle and a virtual mirror were overlaid on the scenario to help the user understand and feel comfortable with the game mechanics.

SECOND VERSION: REAIRABILITATION

After the tests with the first version of the game, the main found issues were related to the graphic interface itself; mainly the feedbacks provided by the system, visual elements presentation and positioning. Users showed difficulty on understanding if they were performing the required motion correctly or not. Additionally, it was not clear for them which elements were positive (should be collided) and which were negative (should be avoided). At last, visual conflicts were observed between the background and the Heads Up Display (HUD) elements such as session time, punctuation and current angle. Given the need for improvements, the game was redesigned to be friendlier.

Brainstorming, sketching and refining the chosen alternative were the strategies used to define the new main character, scenario and additional elements. This conception was performed with an interdisciplinary group composed of designers, physiotherapists and programmers. After these sessions, an airplane was defined as the new main character, keeping the same restriction of vertical movement's freedom (Y-axis). The scenario has been made cleaner than the previous version and provides more space to the other elements. The creation of the other elements was given with the same necessities pointed in the first version: interaction for controlling the game and movement stimulation.

One important characteristic added to this second version of the game is the flexibility of the ring positions that are elements present on the scenario through which users should guide the airplane controlled by their movement. The physiotherapist can set the positioning and timing of these rings in the configuration file in order to make the patient do a specific sequence of movements. For example, if during the physiotherapy it is required an isometric contraction, where the patient has to maintain the movement by a certain amount of time, in the game the therapist can use a sequence of rings to induce that, as shown on the left of Figure 20.

To induce specific movement directions, rings were defined, as the main “must do” steps for the patient. Stormy clouds are now the elements to avoid. To improve the dynamics of the game, fuel boxes must be picked up in order to make the plane keep flying. All these elements are presented on the right part of Figure 19 and left part of Figure 20, and were chosen to make the user easily understand what to do without having to follow any instructions. Also, the interface shows the remaining time of the exercise and number of collected rings at the top of screen so the patient has the feedback of how well he is going through the game, and how much time until the end of the game. On the bottom of the screen it has the camera image, so the user can see him/herself and see that he/she is performing the movement wrongly, also it has the information about how much fuel the airplane has and the angle of the performed gesture is shown as a growing bar below the fuel information.

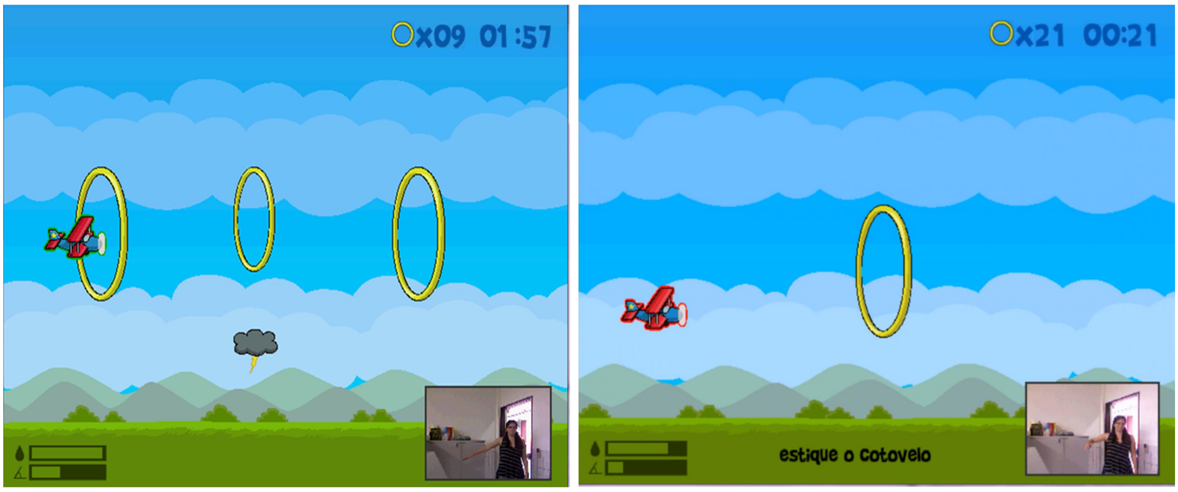


Figure 20. On the left, isometric contraction being induced by rings positions in the reAIRabilitation game. On the right, corrective feedback with instructional sentence (“Estique o cotovelo” (Straight your elbow)).

Based on some results extracted from the first prototype tests, there was a lack of feedbacks on the game. In the Dolphin´s Adventure, the user could not understand when he was doing the movements in a wrong way. To rectify this problem, visual and sonorous feedbacks were added to the game, both triggered when the patient does anything different from what has been planned by the physiotherapist. These elements can be seen on the right part of Figure 20 with warning and instructional messages on the center bottom.

The second version presents the same game mechanics however with a friendlier interface. The game also contemplates additional characteristics such as enabling therapist to define the elements positions, stimulating specific movements. The reAIRabilitation also provides corrective and instructional feedback, which is important for rehabilitation applications.

EVALUATION AND COMPARISON

In order to elicit and validate the improvements in the development of the game, it was submitted to user evaluation to receive feedback from them and, this way, comparing its characteristics and usability. The system was applied to three different population groups where each person's opinion and suggestions on how to improve it were collected. First, tests were applied with the primary prototype version, which was upgraded according to evaluation and necessities. Then, a second test was done with the new system's version.

The required population was composed of subjects from the physiotherapy area, computing area and general population. The therapists were included to enable suggestions about system physiotherapeutic effect and application, while computer specialists could give a more technical opinion. General population was added to evaluate general aspects of usability and motivation of system applicability. All users participated of two encounters, dedicating one encounter for each version of the system. In each encounter, all the users answered a survey consisting of nine questions. At the questionnaire end a space for suggestion was available. Here follow the applied questions:

1) Did you feel that you could control the game?
2) Do you perceive the physiotherapeutic function of the system?
3) Did you feel that the game helped you to correctly perform the movements?
4) Did you feel comfortable during the playing experience?
5) Did you find that the game is easy to play?
6) Do you think the game was fun?
7) Would the game improve your motivation to perform exercises?
8) Did you enjoyed the game scenario?
9) Did you feel challenged?

Each question could be answered, rating, according to a 1-5 Likert scale. In addition, a score was assigned for each question by considering the sum of all ratings of the respective question. This score allows a fast overview of the total of answers, considering all users. This measure also helps to achieve a fast comparison between two stages in which the same question was answered, this way giving a fast overview of the impact of the second tested version over the first one. The selected questions can be split among four major aspects, being each question related to one of the following core subjects:

- Control sensibility (question 1);
- Therapeutic domain (questions 2 and 3);
- Welfare (questions 4 and 5) and;
- Ludic value (remaining questions 6, 7, 8 and 9).

To validate differences a statistics analysis was performed with the Graph Prisma 5.0 software (GraphPad Software 2016). It has been used to verify the data distribution according to the Kolmogorov-Smirnov test. No normal distributions were found. Due to this fact, a comparison was performed with the Wilcoxon test for paired non-parametric data. The tests were considered with 95% of significance level and expressed through probability (p) value, where a p value lower than 0.005 means that the difference was significant.

In total 55 users participated in the two encounters, one for each version of the system, with a 30 days time interval between the encounters dedicated to implement the pointed improvements. In each encounter they answered the previous described questionnaire. In Figure 21 it is shown a chart for each one of the four aspects (grouping the respective questions of each aspect) and a final chart representing all questions together. Each chart presents the number of occurrences (vertical axis) of each rating (horizontal axis), presenting both the first and the second encounter results (labeled as 1st and 2nd time).

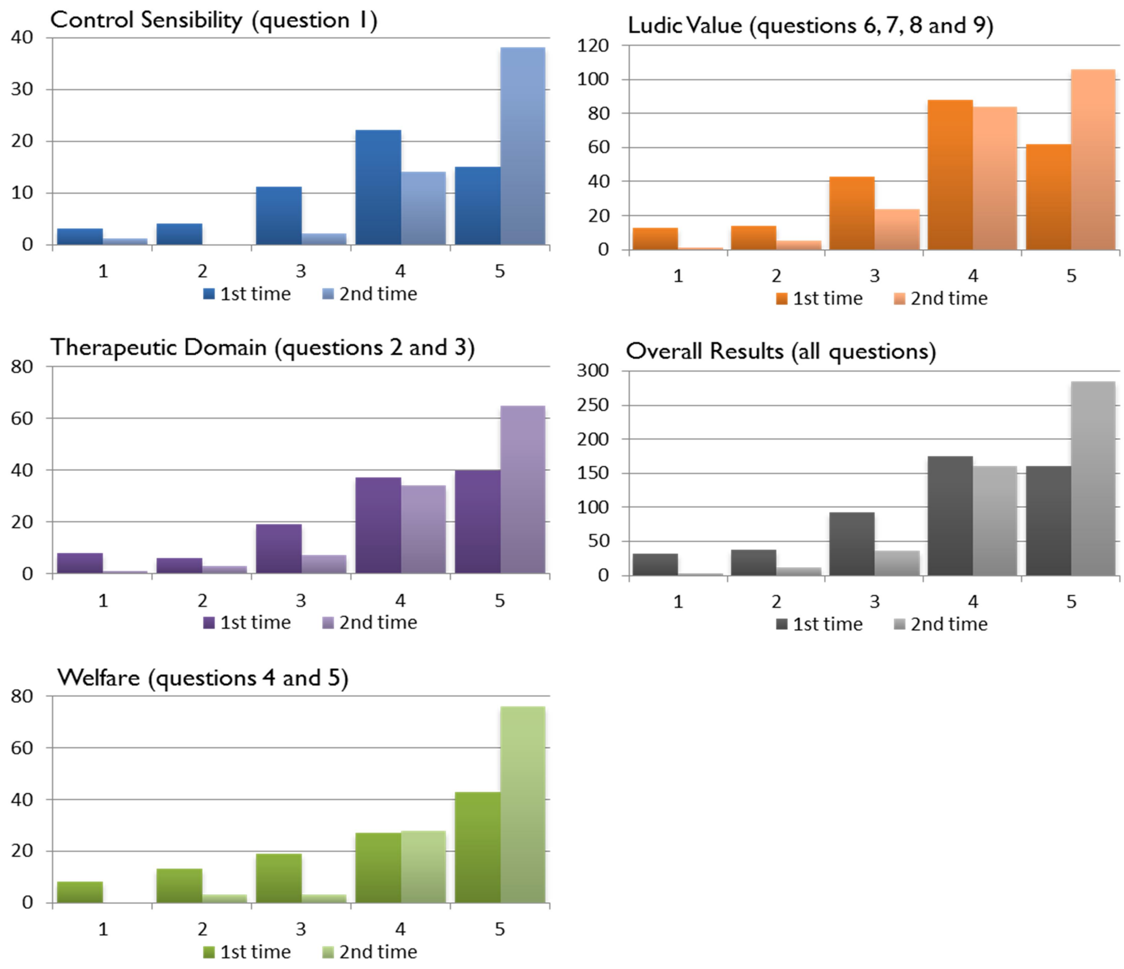


Figure 21. Ratings of each evaluation aspect plus the overall results regarding the comparison between the two versions of the VR system.

As an initial overview, it is noticed in Figure 21 that the users on the second encounter better evaluated all topics presented on the questions. It also can be seen that great part of the users migrated their ratings from a lower value to 5, in fact in the overall results the number of 5 ratings is 125 times greater in the second encounter. Independently of the first tested version, in a more absolute analysis, by considering that the total of answers of all questions is 495 and 445 of those, i.e. 89.9%, were a 4 or 5 rating, revealing a significant satisfaction from the users with the second version of the system.

As shown in Table 10, specifically, the questions 2 and 9 did not reach a significant growth in the second evaluation and so the second evaluation does not provide enough statistics data to declare that the second version of the system presents a better resolution for these topics. However, question 2 already presented a high score of 238 in the first evaluation thus, being understandable its low growth since the maximum limit was already too close. On the other hand, question 9 reveals that the game aspect of challenge still has a significant space for improvements since both evaluations of the users showed an intermediate score close to 200.

Table 10. Score of each topic in the 1st and 2nd encounter.

Aspect or Question	1st Time Score	2nd Time Score	2nd / 1st Score	p value
Question 1	207	253	122%	0.0002
Question 2	238	246	103%	0.3133
Question 3	187	243	130%	0.0001
Question 4	180	241	134%	0.0001
Question 5	234	266	114%	0.0005
Question 6	189	235	124%	0.0001
Question 7	223	253	113%	0.0006
Question 8	222	256	115%	0.0001
Question 9	198	205	104%	0.4049
Control Sensibility	207	253	122%	-
Therapeutic Domain	425	489	115%	-
Welfare	414	507	122%	-
Ludic Value	832	949	114%	-
Overall Results	1878	2198	117%	-

Furthermore, questions 3, 4 and 6 revealed the lowest result in the first evaluation and so, a major need for improvements compared to the other topics. Question 3 revealed the need for a feedback system that was implemented for the second version of the system, providing audio and visual information directed to assist the user during the execution. Question 4 by its turn revealed that the random criteria used to define whether an obstacle or a bonus coin should appear forced the users to perform too much isometric movements, e.g. keeping the arm raised for too much time. The second version of the system was prepared in a way that all positive and negative elements (e.g.: thunder clouds, gas and golden hoops) are set by the physiotherapist and, thus, they appear in the game inducing the user to switch the exercise mode between slow and fast movements as well as some rest time. One advantage of the new design of these elements is that it helped the user to visually recognize more quickly which elements he should avoid, which he should pick and which he should pass through its center. At last, question 6 revealed a space for improvement about the fun during the playing experience. As can be seen in Table 10 the interface improvements, plus some adjustments for the second version of the game solved partially this problem.

One of the reasons that may be responsible for the better results related to question 1 is that the version of the used Microsoft Kinect SDK was updated, and so, the precision of the tracking algorithm was increased. Besides that, the new design of the main character may have favored a better visual idea of control. Before, in the first version, the player controlled a pink dolphin, which was animated constantly moving in its own space, and so, its movement could confuse the user whether the movement was obeying his commands or just being performed by the game itself. The remaining questions (questions 5, 7 and 8) also revealed that the redesigned graphical interface had a good impact on users about the easiness of play, the motivation to play during the practice of physiotherapeutic exercises and the visual aspect of the presented scenario.

These results were published in (FREITAS et al. 2012) (which won the Third Best Paper Award in the Computing Track of the XI Brazilian Symposium on Computer Games and Digital Entertainment - SBGames 2012). A more detailed version of these tests can be found in (CARNEIRO et al. 2016; CARNEIRO et al. 2012; Oliveira et al. 2013).

4.2 Case Study 2: Mirrarbilitation

The second case study was based on AR, being the first to present an AR rehabilitation system based on ISB standards, which enables the system to interact and to be configured according to physiotherapeutic needs. The aim here is to establish a system where the instruction and motivation to perform the exercise would be provided overlaid on the real world. The idea was based on biofeedback concepts, where the patient auto visualization is shown to increase postural and movement control (Caudron et al. 2014; Thikey et al. 2012) improving their learning, performance and rehabilitation results (Thikey et al. 2012; Stanton et al. 2011; Richard et al. 2007; Hopmann et al. 2011). First, a prototype was created where the concepts proposed were validated. The first prototype was tested with three different populations in order to get different opinions: elderly, adults and physiotherapists. The results of these tests helped the definitions of the second AR version, the mirrARbilitation system.

The mirrARbilitation (Alana Elza Fontes Da Gama et al., n.d.) was developed by a colleague in cooperation with the NARVIS lab group (Kang, Lee, and Jung 2004) from the Technische Universität München, in Munich, Germany. This way, the mirrARbilitation system is described here as a third party system which successfully uses the developed movement analysis library of this work. These two systems are presented next.

4.2.1 First Prototype

For AR systems a basic principle is the presence of the real world overlapped with the augmented synthetic content. For the first prototype version, which purpose was to validate the idea, the depth image provided by the Kinect was used. This way there was no necessity to scale the skeleton positions extracted from the depth image to the respective position on the RGB image.

Guidance for correct movement execution is one of the important system principles. In order to do that it is necessary to know the movement direction and aim and find a way to show this information to the user. The movement in question is the shoulder abduction. During its performance, the user has to take the arm up while positioned laterally to the body. To induce this position a reaching object can be positioned laterally to the torso at a distance reachable by the hand. For the first prototype, a simple red square was used as a reaching object (Figure 22). The reaching object is positioned according to the shoulder position being the square height the shoulder Y-axis and the X position correspondent to the shoulder X-axis plus the arm size. Since the reaching object position is based on shoulder references, it is able to follow the user motion.

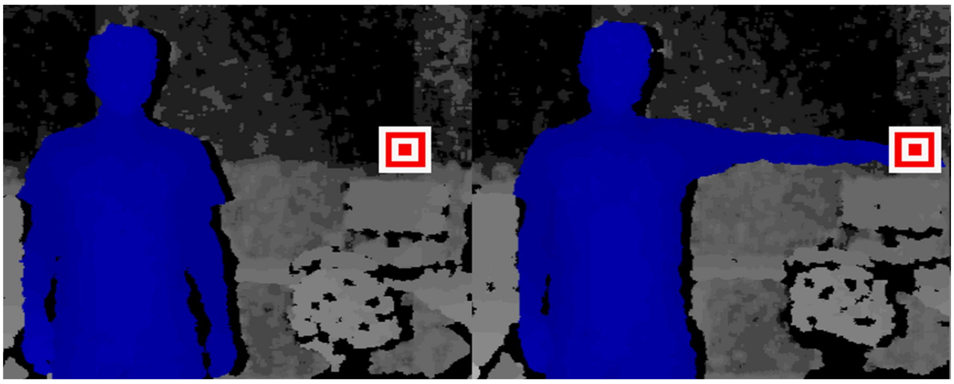


Figure 22. Reaching object to induce user movement.

The correctness of the movement execution in a rehabilitation process is essential for the treatment efficacy. Due to this, the system is programmed to punctuate whenever the user executes the movement correctly. Angles measurements as well as arms and torso alignment are used as criteria to describe the movement. Postural analysis and users' compensations during movement can also be controlled through the system.

Since this was the first system developed during this work, it did not make use of the movement recognition library and the movement description was made individually. So, aiming to recognize correctly the shoulder abduction execution, the following descriptors and requisites were used:

i. The shoulder abduction angle must be equal or greater than 90 degrees at the end of the movement;
ii. The elbow angle must be similar or higher than 160 degrees (to ensure that the arm is well stretched);
iii. The angle between the arm normal vector of the frontal plane must be within the range of 80 and 100 degrees in order to guarantee the lateral alignment of the arm;
iv. The right and left shoulder height (Y coordinate) must be similar, with a range of 10%;
v. The actual abduction angle must be higher than it was before;
vi. In order to keep punctuating, user needs to go down with his arm (the arm has to go down 30 degrees of shoulder abduction), and perform again the complete movement.

With the movement description it is now necessary to tell the user when he performed the movement correctly or not. In order to inform him, scores were created which increase each time the movement is executed correctly (Figure 23). In order to help the user to understand the movement dynamics, an additional instruction informing to return to the initial position was included when a score is achieved (Figure 23).

Figure 23. Feedbacks through scores return instruction and congratulation message.

In the motivational aspect, for each five points a congratulation message is given (Figure 23). The number of points where the feedback will be shown can be chosen by the user. Additionally, a movement status bar is presented and is loaded gradually according to the movement route (0 to 90 degrees) (left part of Figure 24). Knowledge about movement status helps the user to know if he is in the right way.

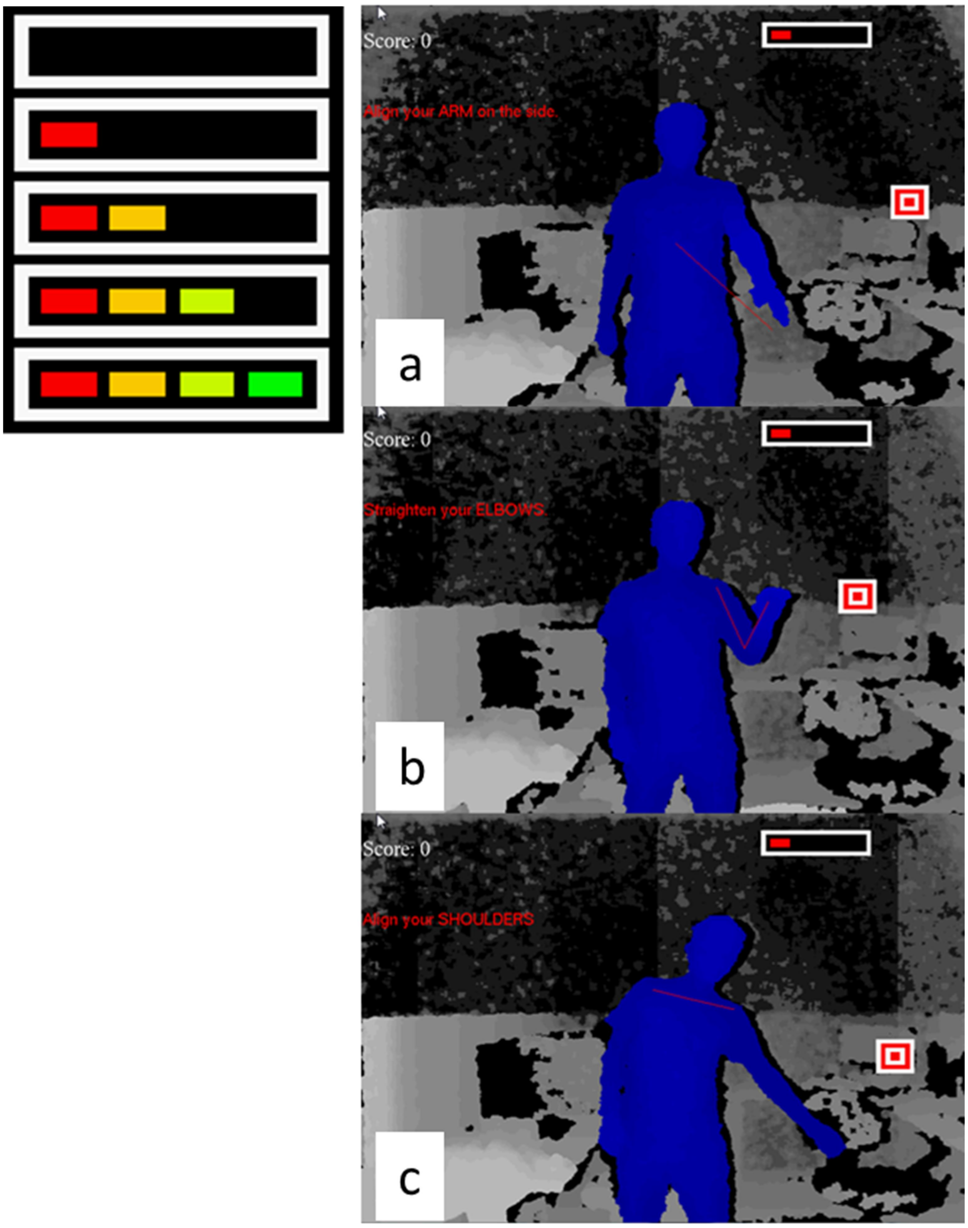


Figure 24. On the left, the movement status bar. On the right, warnings about wrong movement performance: body highlighting and instructions to correct it. A) Arm not aligned laterally. B) Elbow not straight. C) Postural inclination.

As discussed before, when performing rehabilitation exercises it is not only important to show how to perform the exercise but also avoid wrong execution. In order to help with that warning messages were created. When the movement is being done in a wrong way, a red text telling how to correct it is presented below the score area. Movement correction is also enabled highlighting body parts which should be corrected (right part of Figure 24). More information can be found in (A. Da Gama et al. 2012a; Alana Da Gama et al. 2012).

4.2.2 Mirrarbilitation

As for the first prototype, the movement chosen for interaction was the shoulder abduction. However, in the mirrARbilitation the movement recognition for interaction was performed using the movement analysis library developed. From the library the angle for the actual movement status and the wrong executions were extracted. This information was then used for interaction and feedback definition. More detailed information regarding the mirrARbilitation solution and its evaluation can be found in (Alana Elza Fontes Da Gama et al., n.d.).

INTERFACE

The first change made on the interface was the use of the real image with the RGB information provided by the Kinect. It was defined that in the new version the dynamic movement would be induced not only by a reaching object but also with a catching object making user to return to the start position. This way the movement flow will be maintained. The catching and reaching objects chosen for this version were a ball and a basket (Figure 25). These objects were used to induce the movement and also to add a lucid aspect making the system more attractive.

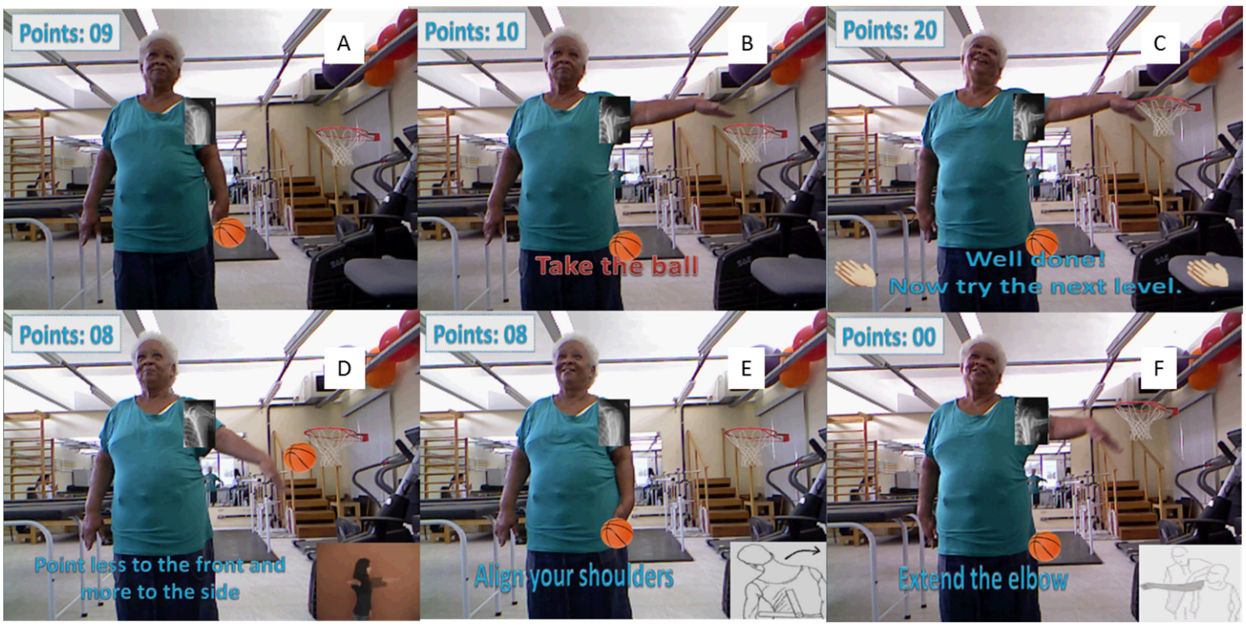


Figure 25. mirrARbilitation interface: A and B) Reaching and catching game dynamics; C) End of level with congratulation message; D to F) Warning and instruction for wrong movement's performance.

Additional warning to inform subject to return to start position was included with the message "Take the ball" (Figure 25). Despite the object definition, mirrARbilitation enables a certain freedom: the catching and reaching images and text can be easily changed any time by overwriting the files in the game folder with the new ones and renaming them. Differently from the first prototype, the mirrARbilitation system enables a reaching object to be set. The position can be configured according to the user's maximum ROM or according to the angle that the physiotherapist desires the patient to achieve. Since the position is depending on an angle, the reaching object is set at a position calculated according with shoulder position and arm size (Equation 4 and 5). Since the objects are located according to a body reference, the user is free to be in any place on screen. The only requirement is a full upper body view in order to allow appropriate movement analysis.

$$reachingObjectX = shoulderX + [\sin(aimAngle) * armSize] \quad (4)$$

$$reachingObjectY = shoulderY + [\cos(aimAngle) * armSize] \quad (5)$$

The score system and the congratulations message provided when achieving certain number of successful movements were both upgraded (Figure 25). The difference here is that when achieving the number of repetitions configured by the therapist the user crosses to the next level of the game. The next level is characterized by an increase on the position of the reaching object. The graduation of this improvement is also defined by the therapist, which chooses the number of degrees that will be added to reach the object position for the next level.

It is suggested that when the patient has conscience of his body function and mechanics he gets credulous about the physiotherapy (A. Da Gama et al. 2012b). This additional believe improves patient engagement and motivation on physiotherapy (Caudron et al. 2014; Thikey et al. 2012; Richard et al. 2007). Based on this concept, in order to show patient how joint mechanics is working and improving during movement, x-ray images taken from the patient were included. The x-ray images represent the shoulder anatomy during all ROM. The respective x-ray image is shown according to movement angle. All interface elements together can be seen in Figure 25.

MOVEMENT CORRECTIVE GUIDANCE

Instructions to help avoid the wrong movement execution are an important characteristic of these systems. For the mirrARbilitation system, the best way to show this information through images was defined by the specialists. During the definition process, it was established that the highlighting on body parts was not effective for patients. This kind of feedback brings a lot of information to the screen, which is not immediately understood requiring processing to achieve comprehension. A large number of rehabilitation system users are old or have cognitive or visual impairments associated. Due to that, it was suggested writing warnings accompanied with a picture showing how correction should be done, probably a clearer way to instruct patient (Figure 25 parts D to F). For the shoulder abduction the instructions for wrong performance were: point less to the front and more to the side (when arm is out of the biomechanical movement tolerance, out of the plane); extend the elbow; and align your shoulders.

The error tolerance is the last additional feature supplemented on our system. It was created in order to improve usability and make it more adaptable to users' interests. This value defines the number of seconds which the system enables the user to be at the wrong position. The error tolerance is set in seconds in the configuration file. If the user does not return to the correct position before the tolerance time, the system will ask him to reset the movement by returning to the start position.

EVALUATION AND COMPARISON

The mirrARbilitation evaluation was composed of three phases: exercise without the system, using the system and another time without it. At each phase, the users were asked to perform the movements until they feel tired. It included 33 participants from three groups: physiotherapists, developers and patients. The number of subjects was computed by using sample computation test at GPower 3.1 software (Faul et al. 2007). The main outcomes are the percentage of right exercises performances and the number of repetitions. These data were compared between the three phases by using ANOVA and Friedman test, being this last one used when there was at least one nonparametric data in the comparison (Alana Da Gama et al. 2013). These tests were performed with all subjects and also categorized into each group. Usability questionnaire was also applied and its results presented in average and standard deviation.

The user test was planned with the main goal to evaluate the effect in using the interactive system which uses the movement recognition technique that provides biomechanical analysis, and this way, information to corrective feedback in order to improve exercise quality. This was performed by checking if the user learns how to do the exercise correctly with the system help. In order to check it, it is necessary to ask the user to perform the exercise with and without mirrARbilitation for comparison. Knowing that, the following protocol was established in three phases as described below.

The phase 1 aimed to test subject's natural movement based only on therapist instructions. It simulated user performance at home after a therapeutic section. In order to do this, a physiotherapist instructed the user on how he/she should do the exercise and then he/she performs the exercise alone until gets tired.

The phase 2 tested the mirrARbilitation usage allowing to check the system capability to induce the correct movement. The same instruction about how to perform and the number of repetition was given (until feel tired).

The instruction to perform the exercise "as many times as possible" was given in order to check patient engagement on exercise, which is not the focus of this dissertation. The complete analysis of these tests can be found at (Alana Elza Fontes Da Gama 2015). The system capability to induce the correct movement is analyzed comparing the number of correct movements in both phases. In order to know if the movements were being performed correctly in the first phase, the movements performed were recorded with the Kinect Studio, provided by the Microsoft SDK. Further this record was used for the movement analysis.

An additional phase 3 was added, where the user performed the movement again without interacting with the system. This phase had two goals: i. to check if the number of repetitions changed before and after the phase 2 in order to evaluate if fatigue interferes on it; ii. to evaluate if the differences in the correctness of exercise between phase 1 and 2 occurs due to system help or because the user is learning how to do it correctly from repetition. The same Kinect studio recording for further analysis made at phase 1 was done at this phase.

All the acquired data was statistically analyzed using the IBM SPSS Statistics 20® (IBM Corporation 2011) and they are described at results section. To check if the system helped user to perform movements in a correctly way the success rate representing the percentage of right movements during section was used. Also, since it was hypothesized that differences between phase 1 and 2 would be found, a third phase as reference was added. Thus, the number of repetitions and the success rate between phase 1 and 3 were performed in order to check if the differences between phase 1 and 2 occurred due the system use or due others bias.

Statistical tests started with the Kolmogorov–Smirnov test in order to check data distribution and categorize the variables as parametric or non-parametric (Alana Da Gama et al. 2013). After this, it was detected that all variable presented parametric characteristics except for the number of repetitions at phase 3. For the parametric data the mean comparison between the phases were done with the ANOVA test for repeated measure (Alana Da Gama et al. 2013). The results are described in terms of mean and standard deviation. In order to provide more detailed analysis, the data was also described and compared at each user category: physiotherapists, patients and developers. The Kolmogorov-Smirnov test showed normal distribution of all these data at each group. So, all comparison inside the groups were done using the ANOVA for repeated measure. For the patients, since they have only two measure the comparison was made using the paired T-student (Alana Da Gama et al. 2013).

To evaluate the mirrARbilitation capability to induce the correct exercise performance the success rate was evaluated. This measure represents the percentage of exercises that were performed correctly. The results show a percentage improvement of correct movement exercises with the use of the system, p = 0.004 (Figure 26). No significant difference was found between the two moments without the use of system, phase 1 and 3 (p = 0.881). Additional result which is relevant to notice is the minimum value of success rate at each circumstance: phases 1 and 3 had 0%, i.e., there were at least one subject on those groups that did not perform a correct exercise not even a single time; on the other hand, phase 2 had a minimum of 73.68%.

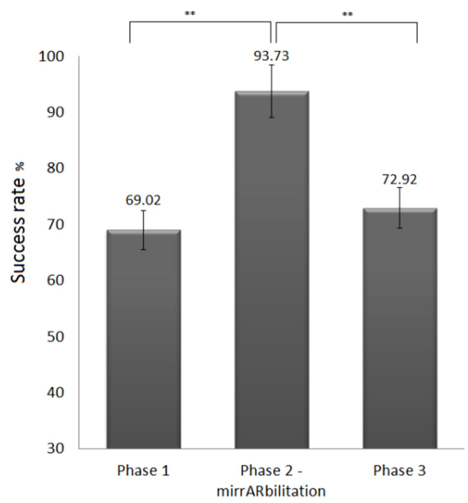


Figure 26. Average percentage of correct exercise at each phase. **p<0.01, test.

The average success rate by group followed the same tendency (Figure 27). However, for the physiotherapist group the difference between phase 1 and 2 (p = 0.099) and between 2 and 3 (p = 0.143) was not significant. Nevertheless, the minimum values continued drastically different between the phases even at this specialized group: phase 1: 6.67%, phase 2: 89.39% and phase 3: 9.09%.

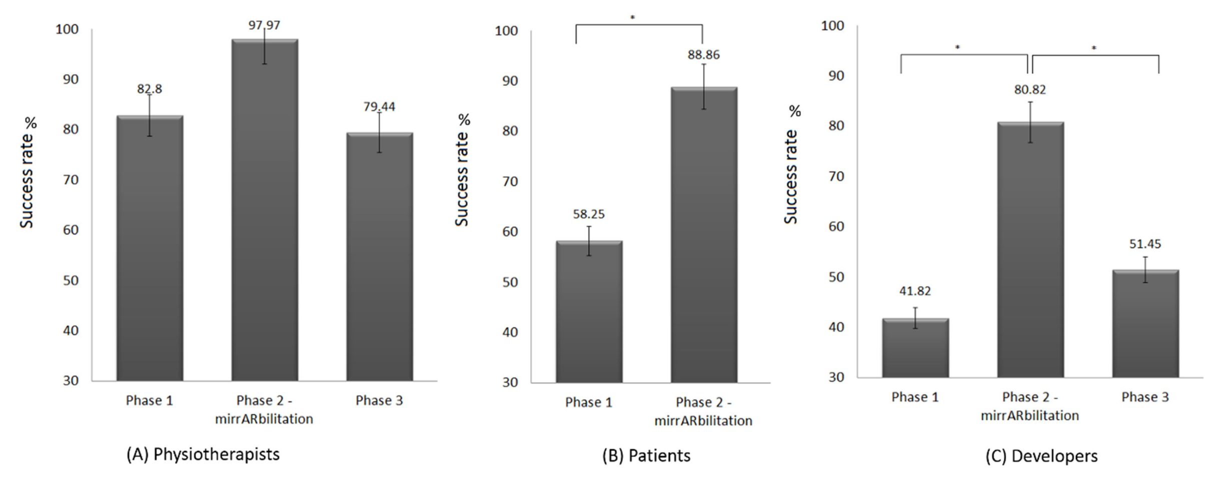


Figure 27. Percentage of correct exercise performed at each phase. (A) physiotherapist, (B) patients and (C) developers group. *p<0.05 at ANOVA test.

4.3 Case Study 3: À Sombra Leve Dos Anjos

The third case study was unrelated to the physiotherapy domain. It is described here as an additional use of the developed recognition methods and the library itself. It was an application to control an artistic installation. The installation was called "À Sombra Leve dos Anjos" and was placed at the Malakoff tower, which is a well-known touristic spot in Recife, Brazil. In this case, the developed application was also a third party development that used the movement analysis library developed in this work. The purpose of the installation was to allow visitors to navigate through a cartoonish art by using gestures. The installation was part of an exposition of the artist João Lin.

The art was projected on a wall in front of the user, and it represented the terrain and heavenly worlds, apart from one to the other by a large vertical distance. Only a small window of the art was shown to the user at each time, which started from the terrain part, and in order to ascend to the clouds, the user should perform a gesture that resembled the flapping of an angel wings. Here the metaphor pointed out is that the user arms are the angel wings. In this case, the checkpoints recognition method was used, a single execution of the wings flapping was recorded and later used for the recognition. Each time the gesture completed an up and down cycle, the projection was pushed downwards showing a higher part of the art on the visible window.

Figure 28 shows a picture taken from behind the user showing the projection in his front. On this picture, the user has his arms raised as an intermediary step of the gesture execution. The Microsoft Kinect v1 device was used in this installation; it was placed below the projection pointing to the user.

Figure 28. Moviment analysis library being used on an artistic instalation. On the left is the artistic instalation banner. On the right it shows a person using the developed system.

4.4 Case Study 4: Voxar Music

The fourth case study was also unrelated to the physiotherapy domain. It was focused on entertainment and user studies towards music experiencing and education. The developed library was used to develop an application. The goal of the application was to play a music title by activating a sound track containing the channel of a particular instrument. This way the user had to perform specific gestures to activate the percussion, and as well as other gestures for tracks containing bass, guitar, vocals and so on. The gestures were recognized using the checkpoints method. Figure 29 shows the application running.

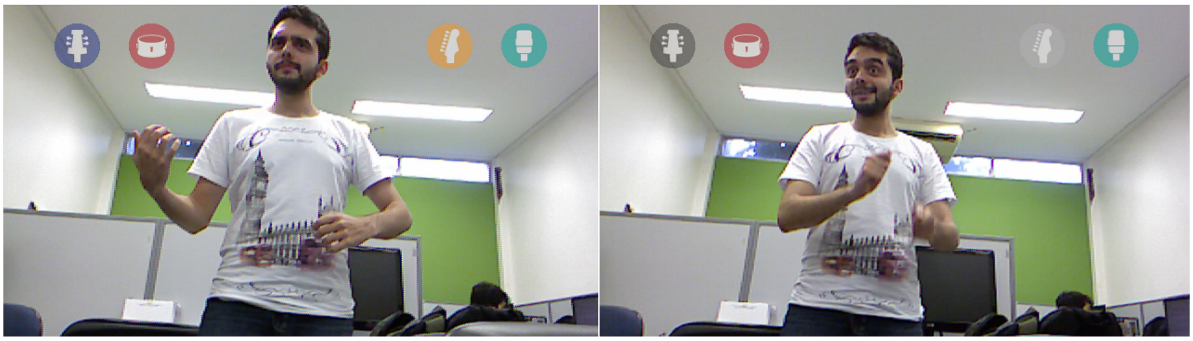


Figure 29. Music application using gestures recognized by the checkpoints method: the guitar gesture on the left and the drums gesture on the right.

5 CONCLUSION

This dissertation presents the development and validation of a movement analysis library. The library was built to aid applications on the physiotherapy domain which intend to use gesture recognition as an option to both, control some elements of the interface, as well as, analyze the performance of the gesture execution by the user. Two recognition methods were developed, the first targeting strict biomechanical movements such as shoulder abduction. The second method is for more freely defined movements, targeting functional gestures such as take a glass of water, or squat. Both methods are configurable regarding the required precision, and both were tested and validated showing to be potentially applicable on physiotherapy.

Later, both methods are tested and validated on two different case studies related to the physiotherapy domain, and explore the VR and AR environments, respectively. These case studies applications were developed to engage the patient while performing the gestures, contributing to the patient motivation on a game dynamic directly related to the treatment and to test the proposed gesture recognition techniques. Both the applications were well evaluated by the users, showing it has potential to be used in a treatment, but further and longer tests still must be performed in order to confirm this.

5.1 Publications

This dissertation also resulted on a set of publications on related venues to computer science and physiotherapy. Also, additional publications on related topics to Natural Interaction and Gesture Recognition were achieved. All the publications are listed below:

- Da Gama, Alana, Fallavollita, Pascal, Chaves, Thiago, Figueiredo, Lucas S., Baltar, Adriana, Meng, Ma, Teichrieb, Veronica, Navab, Nassir. **MirrARbilitation: a clinically-related gesture recognition interactive tool for an AR rehabilitation system.** Computers Method and Programs in Biomedicine (**under review**).
- Figueiredo, Lucas S., Pinheiro, Mariana, Neto, Edvar Vilar, Chaves, Thiago, Teichrieb, Veronica. **Sci-Fi Gestures Catalog.** 15th IFIP TC.13 International Conference on Human-Computer Interaction – INTERACT 2015, 395-411, 2015, Springer.
- Da Gama, Alana, Chaves, Thiago, Figueiredo, Lucas S., Teichrieb, Veronica. **Markerless Gesture Recognition according to Biomechanical Convention.** XXIV Brazilian Congress on Biomedical Engineering – CBEB 2014, 1-4, 2014.
- Figueiredo, Lucas S., Pinheiro, Mariana, Neto, Edvar Vilar, Chaves, Thiago, Teixeira, João Marcelo, Teichrieb, Veronica, Alessio, Pedro, Freitas, Daniel. **In-place natural and effortless navigation for large industrial scenarios.** Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments. Third International Conference – DUXU 2014, Held as Part of HCI International 2014, 550-561, 2014, Springer.
- Da Gama, Alana, Chaves, Thiago, Figueiredo, Lucas S., Teichrieb, Veronica. **Development and evaluation of movement recognition techniques for interactive rehabilitations**

support systems. International Conference on Virtual Rehabilitation – ICVR 2013,170-171, 2013.

- De Oliveira, Déborah Marques, Baltar, Adriana, Carneiro, Maíra, Cardoso, Ana Cláudia, Da Gama, Alana, Chaves, Thiago, Teichrieb, Veronica, Araújo, Cristiano Coêlho, Monte-Silva, Kátia Karina. **Desenvolvimento e aprimoramento de um sistema computacional- Ikapp- de suporte a reabilitação motora.** Motriz, Rio Claro, v.19 n.2, p.346-357, abr./jun. 2013.

5.2 Future Work

There are future works directed to two main points of this dissertation. The first one is to continue to iteratively develop the physiotherapy case studies, by applying them on real rehabilitation scenarios with partner therapists. The second direction is to improve the checkpoints recognition method and the library. Although the recognition showed to be satisfactory on the tested scenarios, its wide range applicability motivates its improvement in order to be validated among other recognition methods of generic purpose.

The development of these improvements is currently being conducted. The improvements are directed to three main topics: the smoothness of the trajectory described by the checkpoints; the possibility of learning a gesture from several samples in addition to the one-shot training; and on this several samples case the modulation of the error along the trajectory.

The checkpoints are intended to be modeled as control points of a set of Bézier splines as shown in Figure 30. The figure shows colored small circles that represent different positions of a joint during a training execution of the gesture. Each color represents a single execution. The executions are aligned in time by using DTW. Then the Bézier splines are estimated, rendered as a thick redline on the figure, which represents the gesture trajectory. This trajectory can be used itself as a representation of the trained gesture and a fixed range can be set. It is possible to calculate the distance between a point and the nearest corresponding point of the trajectory, and this way, if a point falls within the range it is considered as accepted by the gesture. In addition, Figure 30 shows experiments on dynamic ranges, i.e. the range is updated along the trajectory according to the deviation of each training execution, this way modelling the range according to the training behavior. This dynamic range is shown as circles centered on the trajectory. Each circle is attached to a checkpoint. This is an ongoing work and the experiments of dynamic ranges are not capable yet to correctly represent the training set, by containing the training trajectories within its range but not by too far. On the near future, these experiments will be conducted and the improved checkpoints method will be later ported to the movement analysis library.

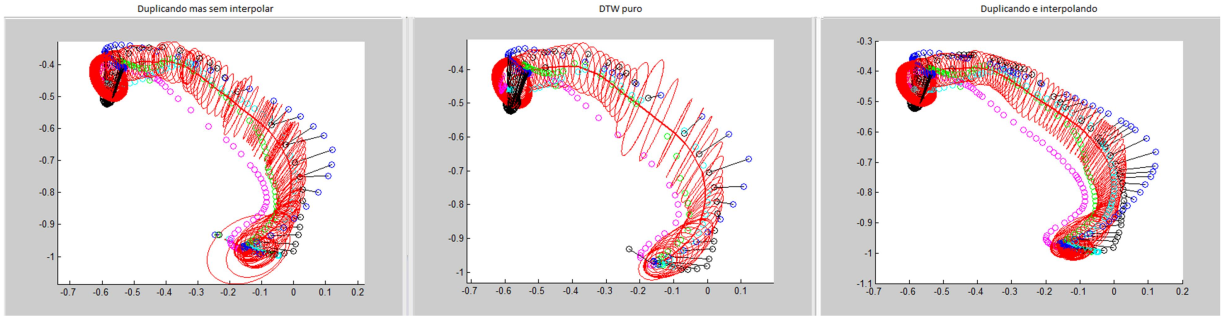


Figure 30. Experiments of an improvement for the checkpoints method.

REFERENCES

Alvarenga, Matheus Lin Truglio, Diogo S. Ortiz Correa, and Fernando Santos Osório. 2011. "Redes Neurais Artificiais Aplicadas No Reconhecimento de Gestos Usando O Kinect." http://osorio.wait4.org/publications/2012/COTB-MatheusLin-2012.pdf.

ASUS. 2016. "Asus Xtion-PRO." Accessed February 10. https://www.asus.com/3D-Sensor/Xtion_PRO/.

Bigdelou, Ali, Tobias Benz, Loren Schwarz, and Nassir Navab. 2012. "Simultaneous Categorical and Spatio-Temporal 3D Gestures Using Kinect." *IEEE Symposium on 3D User Interfaces 2012, 3DUI 2012 - Proceedings*, 53–60. doi:10.1109/3DUI.2012.6184184.

Biswas, K. K., and Saurav Kumar Basu. 2011. "Gesture Recognition Using Microsoft Kinect®." *The 5th International Conference on Automation, Robotics and Applications* 2: 100–103. doi:10.1109/ICARA.2011.6144864.

CARNEIRO, M., D. MARQUES, A BALTAR, A. CARDOSO, V. TEICHRIEB, and K. MONTE-SILVA. 2016. "APLICABILIDADE DE UM SISTEMA DE REABILITAÇÃO MOTORA EM PACIENTES PÓS-ACIDENTE VASCULAR ENCEFÁLICO." *Fisioterapia Em Movimento (PUCPR. Impresso)*.

CARNEIRO, M., D. MARQUES-OLIVEIRA, A. BALTAR, A. CARDOSO, T.M. DA GAMA, A. E. F. Chaves, D. FREITAS, C. ARAUJO, V. TEICHRIEB, and K. MONTE-SILVA. 2012. "Ikaap: A New Virtual Reality-Based System For Stroke Rehabilitation." In *8 Th World Stroke Congress*. Brasília.

Caudron, S, M Guerraz, A Eusebio, J-P Gros, J-P Azulay, and M Vaugoyeau. 2014. "Evaluation of a Visual Biofeedback on the Postural Control in Parkinson's Disease." *Neurophysiologie Clinique = Clinical Neurophysiology* 44 (1): 77–86. doi:10.1016/j.neucli.2013.10.134.

Celebi, Sait, Ali S Aydin, Talha T Temiz, and Tarik Arici. 2013. "Gesture Recognition Using Skeleton Data with Weighted Dynamic Time Warping." *8th International Conference on Computer Vision Theory and Applications (VISAPP 2013)*, Paper #79. http://www.visapp.visigrapp.org/Home.aspx?y=2013.

Chaaraoui, Alexandros Andre, José Ramón Padilla-López, Pau Climent-Pérez, and Francisco Flórez-Revuelta. 2014. "Evolutionary Joint Selection to Improve Human Action Recognition with RGB-D Devices." *Expert Systems with Applications* 41 (3): 786–94. doi:10.1016/j.eswa.2013.08.009.

Chaves, Thiago, Lucas Figueiredo, Alana Da Gama, Cristiano de Araujo, and Veronica Teichrieb. 2012. "Human Body Motion and Gestures Recognition Based on Checkpoints." In *2012 14th Symposium on Virtual and Augmented Reality*, 271–78. IEEE. doi:10.1109/SVR.2012.16.

Clarkson, Hazel M. 2005. *Joint Motion and Function Assessment: A Research-Based Practical Guide*. https://books.google.com/books?hl=pt-BR&lr=&id=3x6nlojg25AC&pgis=1.

Cui, Yan, and Didier Stricker. 2011. "3D Shape Scanning with a Kinect." In *ACM SIGGRAPH 2011 Posters on - SIGGRAPH '11*, 1. New York, New York, USA: ACM Press. doi:10.1145/2037715.2037780.

Da Gama, A., T. Chaves, L. Figueiredo, and V. Teichrieb. 2012b. "Poster: Improving Motor Rehabilitation Process through a Natural Interaction Based System Using Kinect Sensor." In *2012*

IEEE Symposium on 3D User Interfaces (3DUI), 145–46. IEEE. doi:10.1109/3DUI.2012.6184203.

———. 2012a. "Poster: Improving Motor Rehabilitation Process through a Natural Interaction Based System Using Kinect Sensor." In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, 145–46. IEEE. doi:10.1109/3DUI.2012.6184203.

Da Gama, A.E.F., T.M. Chaves, L. Figueiredo, and V. Teichrieb. 2014. "Markerless Gesture Recognition according to Biomechanical Convention." *XXIV Brazilian Congress on Biomedical Engineering*. http://www.canal6.com.br/cbeb/2014/artigos/cbeb2014_submission_598.pdf.

Da Gama, Alana, Thiago Chaves, Lucas Figueiredo, Cristiano Araújo, Veronica Teichrieb, Maíra Carneiro, Déborah Marques-Oliveira, Adriana Baltar, Ana Cardoso, and Kátia Monte-Silva. 2012. "Ikapp – A Rehabilitation Support System Using Kinect." In *2012 14th Symposium on Virtual and Augmented Reality*. http://www.lbd.dcc.ufmg.br/colecoes/svr/2012/0020.pdf.

Da Gama, Alana, Thiago Chaves, Lucas Figueiredo, and Veronica Teichrieb. 2013. "Development and Evaluation of Movement Recognition Techniques for Interactive Rehabilitations Support Systems." *International Conference on Virtual Rehabilitation*. ftp://128.125.133.25/arizzo/ICVR2013/data/papers/40_PID2886315.pdf.

Da Gama, Alana Elza Fontes. 2015. "Improvement of Computational Rehabilitation Support Systems Based on Virtual and Augmented Reality: Clinical Similarity and Biomechanical Movement Recognition." UFPE.

Da Gama, Alana Elza Fontes, Pascal Fallavollita, Thiago Menezes Chaves, Lucas Silva Figueiredo, Adriana Baltar, Ma Meng, Veronica Teichrieb, and Nassir Navab. n.d. "MirrARbilitation: A Clinically-Related Gesture Recognition Interactive Tool for an AR Rehabilitation System." *Computers Method and Programs in Biomedicine*.

Da Gama, Alana, Pascal Fallavollita, Veronica Teichrieb, and Nassir Navab. 2015. "Motor Rehabilitation Using Kinect: A Systematic Review." *Games for Health Journal* 4 (2). Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: 123–35. doi:10.1089/g4h.2014.0047.

Espay, Alberto J., Yoram Baram, Alok Kumar Dwivedi, Rakesh Shukla, Maureen Gartner, Laura Gaines, Andrew P. Duker, and Fredy J. Revilla. 2010. "At-Home Training with Closed-Loop Augmented-Reality Cueing Device for Improving Gait in Patients with Parkinson Disease." *Journal of Rehabilitation Research and Development*. http://search.proquest.com/openview/1662beceedb93364be350f2bbf0aad8a/1?pq-origsite=gscholar.

Faugeroux, Romain, Thales Vieira, Dimas Martinez, and Thomas Lewiner. 2014. "Simplified Training for Gesture Recognition." In *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*, 133–40. IEEE. doi:10.1109/SIBGRAPI.2014.46.

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39 (2): 175–91. doi:10.3758/BF03193146.

Freedman, Barak, Alexander Shpunt, Meir Machline, and Yoel Arieli. 2013. Depth mapping using projected patterns, issued July 23, 2013. https://www.google.com/patents/US8493496.

FREITAS, D., A. E. F. DA GAMA, L. FIGUEIREDO, T.M. CHAVES, D. MARQUES-OLIVEIRA, V. TEICHRIEB, and C. ARAUJO. 2012. "Development and Evaluation of a Kinect Based Motor

Rehabilitation Game." In *XI Simpósio Brasileiro de Games E Entretenimento Digital (SBGames)*. Brasília.

Gama, Alana Da, Thiago Chaves, Lucas Figueiredo, and Veronica Teichrieb. 2012. "Guidance and Movement Correction Based on Therapeutics Movements for Motor Rehabilitation Support Systems." In *2012 14th Symposium on Virtual and Augmented Reality*, 191–200. IEEE. doi:10.1109/SVR.2012.15.

GraphPad Software. 2016. "Graph Prisma." Accessed February 20. http://www.graphpad.com/scientific-software/prism/.

Gu, Ye, Ha Do, Yongsheng Ou, and Weihua Sheng. 2012. "Human Gesture Recognition through a Kinect Sensor." *2012 IEEE International Conference on Robotics and Biomimetics, ROBIO 2012 - Conference Digest*, 1379–84. doi:10.1109/ROBIO.2012.6491161.

Han, Jungong, Ling Shao, Dong Xu, and Jamie Shotton. 2013. "Enhanced Computer Vision with Microsoft Kinect Sensor: A Review." *IEEE Transactions on Cybernetics* 43 (5). IEEE: 1318–34. doi:10.1109/TCYB.2013.2265378.

Hopmann, Mathieu, Patrick Salamin, Nicolas Chauvin, Frédéric Vexo, and Daniel Thalmann. 2011. "Natural Activation for Gesture Recognition Systems." In *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '11*, 173. New York, New York, USA: ACM Press. doi:10.1145/1979742.1979642.

IBM Corporation. 2011. "IBM SPSS Statistics for Windows, Version 20.0." Armonk, NY.

Intel Corporation. 2016a. "Intel Creative Senz3D." Accessed February 10. http://us.creative.com/p/web-cameras/creative-senz3d.

———. 2016b. "Intel RealSense Technology." Accessed February 10. http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html.

Izadi, Shahram, Andrew Davison, Andrew Fitzgibbon, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, et al. 2011. "KinectFusion." In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology - UIST '11*, 559. New York, New York, USA: ACM Press. doi:10.1145/2047196.2047270.

Jun, Seung Kook, Suren Kumar, Xiaobo Zhou, Daniel K. Ramsey, and Venkat N. Krovi. 2013a. "Automation for Individualization of Kinect-Based Quantitative Progressive Exercise Regimen." *IEEE International Conference on Automation Science and Engineering*, 243–48. doi:10.1109/CoASE.2013.6654038.

———. 2013b. "Automation for Individualization of Kinect-Based Quantitative Progressive Exercise Regimen." *IEEE International Conference on Automation Science and Engineering*, 243–48. doi:10.1109/CoASE.2013.6654038.

Kang, Hyun, Chang Woo Lee, and Keechul Jung. 2004. "Recognition-Based Gesture Spotting in Video Games." *Pattern Recognition Letters* 25 (15): 1701–14. doi:10.1016/j.patrec.2004.06.016.

Kato, Pamela M. n.d. "Video Games in Health Care: Closing the Gap."

Kisner, Carolyn, and Lynn Allen Colby. 2012. *Therapeutic Exercise: Foundations and Techniques*. F.A. Davis. https://books.google.com/books?hl=en&lr=&id=udY-AAAAQBAJ&pgis=1.

Lange, Belinda, Chien-Yen Chang, Evan Suma, Bradley Newman, Albert Skip Rizzo, and Mark Bolas. 2011. "Development and Evaluation of Low Cost Game-Based Balance Rehabilitation Tool Using the Microsoft Kinect Sensor." *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference* 2011 (January). IEEE: 1831–34. doi:10.1109/IEMBS.2011.6090521.

Leightley, Daniel, John Darby, and Jamie S. McPhee. 2013. "Human Activity Recognition for Physical Rehabilitation." *2013 IEEE International Conference on Systems, Man, and Cybernetics*, 261–66. doi:10.1109/SMC.2013.51.

Lin, Ting Yang, Chung Hung Hsieh, and Jiann Der Lee. 2013. "A Kinect-Based System for Physical Rehabilitation: Utilizing Tai Chi Exercises to Improve Movement Disorders in Patients with Balance Ability." *Proceedings - Asia Modelling Symposium 2013: 7th Asia International Conference on Mathematical Modelling and Computer Simulation, AMS 2013*, 149–53. doi:10.1109/AMS.2013.29.

Microsoft Corporation. 2016a. "Kinect for Windows SDK 2.0." Accessed February 10. https://www.microsoft.com/en-us/download/details.aspx?id=44561.

———. 2016b. "Kinect for Xbox One." Accessed February 10. http://www.xbox.com/pt-BR/Xbox-One/accessories/kinect-for-xbox-one.

———. 2016c. "Kinect Sensor for Xbox 360." Accessed February 10. http://support.xbox.com/en-US/xbox-360/accessories/kinect-sensor-components.

Ming, Yue, Qiuqi Ruan, and Alexander G. Hauptmann. 2012. "Activity Recognition from RGB-D Camera with 3D Local Spatio-Temporal Features." *Proceedings - IEEE International Conference on Multimedia and Expo*, 344–49. doi:10.1109/ICME.2012.8.

Miranda, Leandro, Thales Vieira, Dimas Martínez, Thomas Lewiner, Antonio W. Vieira, and Mario F. M. Campos. 2014. "Online Gesture Recognition from Pose Kernel Learning and Decision Forests." *Pattern Recognition Letters* 39 (April): 65–73. doi:10.1016/j.patrec.2013.10.005.

Oliveira, Déborah Marques de, Adriana Baltar do Rêgo Maciel, Maíra Izzadora Souza Carneiro, Ana Cláudia de Andrade Cardoso, Alana Elza Fontes da Gama, Thiago de Menezes Chaves, Verônica Teichrieb, Cristiano Coêlho de Araújo, and Kátia Karina do Monte-Silva. 2013. "Desenvolvimento E Aprimoramento de Um Sistema Computacional- Ikapp- de Suporte a Reabilitação Motora." *Motriz: Revista de Educação Física* 19 (2). Motriz: revista de educação física (Online): 346–57. doi:10.1590/S1980-65742013000200012.

Payne, Andrew, Andy Daniel, Anik Mehta, Barry Thompson, Cyrus S. Bamji, Dane Snow, Hideaki Oshima, et al. 2014. "7.6 A 512×424 CMOS 3D Time-of-Flight Image Sensor with Multi-Frequency Photo-Demodulation up to 130MHz and 2GS/s ADC." In *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 134–35. IEEE. doi:10.1109/ISSCC.2014.6757370.

PrimeSense. 2016. "OpenNI." Accessed February 10. http://structure.io/openni.

Richard, E., V. Billaudeau, P. Richard, and G. Gaudin. 2007. "Augmented Reality for Rehabilitation of Cognitive Disabled Children: A Preliminary Study." In *2007 Virtual Rehabilitation*, 102–8. IEEE. doi:10.1109/ICVR.2007.4362148.

Sell, John, and Patrick O'Connor. 2014. "The Xbox One System on a Chip and Kinect Sensor." *IEEE Micro* 34 (2). IEEE Computer Society: 44–53. doi:10.1109/MM.2014.9.

SoftKinetic. 2016. "SoftKinetic DepthSense." Accessed February 10. http://www.softkinetic.com.

Stanton, Rosalyn, Louise Ada, Catherine M Dean, and Elisabeth Preston. 2011. "Biofeedback Improves Activities of the Lower Limb after Stroke: A Systematic Review." *Journal of Physiotherapy* 57 (3): 145–55. doi:10.1016/S1836-9553(11)70035-2.

Suma, Evan A, Belinda Lange, Albert Rizzo, David M. Krum, and Mark Bolas. 2012. "FAAST-R: Defining a Core Mechanic for Designing Gestural Interfaces." *Chi'12*, 5–10.

Thikey, Heather, Madeleine Grealy, Frederike van Wijck, Mark Barber, and Philip Rowe. 2012. "Augmented Visual Feedback of Movement Performance to Enhance Walking Recovery after Stroke: Study Protocol for a Pilot Randomised Controlled Trial." *Trials* 13 (1). BioMed Central: 163. doi:10.1186/1745-6215-13-163.

Urtasun, Raquel, David J. Fleet, and Pascal Fua. 2006. "Temporal Motion Models for Monocular and Multiview 3D Human Body Tracking." *Computer Vision and Image Understanding* 104 (2-3): 157–77. doi:10.1016/j.cviu.2006.08.006.

Vemulapalli, Raviteja, Felipe Arrate, and Rama Chellappa. 2014. "Human Action Recognition by Representing 3D Skeletons as Points in a Lie Group." *2014 IEEE Conference on Computer Vision and Pattern Recognition*. doi:10.1109/CVPR.2014.82.

Wan, Jun, Q Ruan, W Li, and S Deng. 2013. "One-Shot Learning Gesture Recognition from RGB-D Data Using Bag of Features." *The Journal of Machine Learning …* 14: 2549–82. http://dl.acm.org/citation.cfm?id=2567743.

Wang, Jiang. 2012. "Mining Actionlet Ensemble for Action Recognition with Depth Cameras." *Computer Vision and Pattern Recognition (CVPR)*. doi:10.1109/CVPR.2012.6247813.

Wu, Ge, and Peter R. Cavanagh. 1995. "ISB Recommendations for Standardization in the Reporting of Kinematic Data." *Journal of Biomechanics* 28 (10): 1257–61. doi:10.1016/0021-9290(95)00017-C.

Yang, Xiaodong, and Yl Tian. 2012. "Eigenjoints-Based Action Recognition Using Naive-Bayes-Nearest-Neighbor." *Computer Vision and Pattern Recognition …*, 14–19. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6239232.