www.fit4work-aal.eu

FIT4WORK

SELF-MANAGEMENT OF PHYSICAL AND MENTAL FITNESS OF OLDER WORKERS

CO-FUNDED BY

AAL PROGRAMME

The National Centre for Research and Development

GOBIERNO DE ESPAÑA · MINISTERIO DE INDUSTRIA, ENERGÍA Y TURISMO

ZonMw

uefiscdi · EXECUTIVE AGENCY FOR HIGHER EDUCATION, RESEARCH, DEVELOPMENT AND INNOVATION FUNDING · INNOVATION AND CREATIVITY

REPUBLIC OF SLOVENIA MINISTRY OF HIGHER EDUCATION, SCIENCE AND TECHNOLOGY

PARTNERS

PSNC

AWF · Poznan University of Physical Education

"Jožef Stefan" Institute

UNIE KBO

SGS

Teamnet transforming technology

OSM

3D Motion Sensing

Ambient Assisted Living Joint Programme project no. AAL-2013-6-060

Deliverable 4.3.1, version 1.0

Deliverable 4.3.2, version 1.0

Lead author: Maria Saornil, Fundacion CTIC (subcontracted by SGS Tecnos)

Co-authors: Oscar Franco, Other Side Mirror
Michał Kosiedowski, Poznań Supercomputing and Networking Center

Project coordinator: Poznań Supercomputing and Networking Center, ul. Jana Pawła II 10, 61-139 Poznań, Poland, email: fit4work@fit4work-aal.eu

Published on 28th of April, 2016 (deliverable D4.3.1), on 17th of November, 2017 (deliverable D4.3.2)

The Fit4Work project is co-financed though the AAL Joint Programme by:

- European Commission
- National Centre for Research and Development, Poland
- Ministry of Industry, Energy and Tourism, Spain
- Executive Agency for Higher Education, Research Development and Innovation Funding, Romania
- Ministry of Higher Education, Science and Technology, Slovenia
- The Netherlands Organisation for Health Research and Development (ZonMW), The Netherlands

Table of contents

1. Introduction

Fit4Work aims at monitoring and improving healthy lifestyle of the elderly workers. To do that the developed system monitors the users' physical activity, mental stress and the ambient conditions at the workplace.

This document describes the work done for the development of the 3D Motion Sensing Module responsible for fitness monitoring and training of system users. The document is structured in the following sections:

- **Problem Analysis (section 2)**, aimed at the evaluation of the exercises that have been selected for the users of the Fit4work platform, in terms of technical feasibility for the creation of a module based on depth sensors for monitoring the execution of those exercises.
- **Module Design (section 3)**, that presents the architecture and main components of the 3D Motion Sensing Module.
- **3D Motion Sensing module (section 4)**, which describes the resulting module developed for the automatic evaluation of the exercises executed by the Fit4work users.
- **Evaluation and future work (section 4.4.3)**, that analyzes the module developed taking into account the required criteria defined at the initial stage of the project, and identifying the aspects for improvement.

The current report stands for the final report concerning the developments related to the 3D sensing module of the Fit4Work solution and stands for deliverable D4.3.2. This deliverable was built on the basis of deliverable D4.3.1 and includes the contents of that report (this refers mainly to section 2 of this document). Deliverable D4.3.1 was then extended with new contents, including updates to section 2 as well as (most) contents of further sections of this report.

2. Problem Analysis

2.1. General overview

The module described in this document is aimed at monitoring the physical activities performed by the users under the training program suggested by the Fit4Work system (see Figure 2.1). This is part of the training support sub-module of the system.

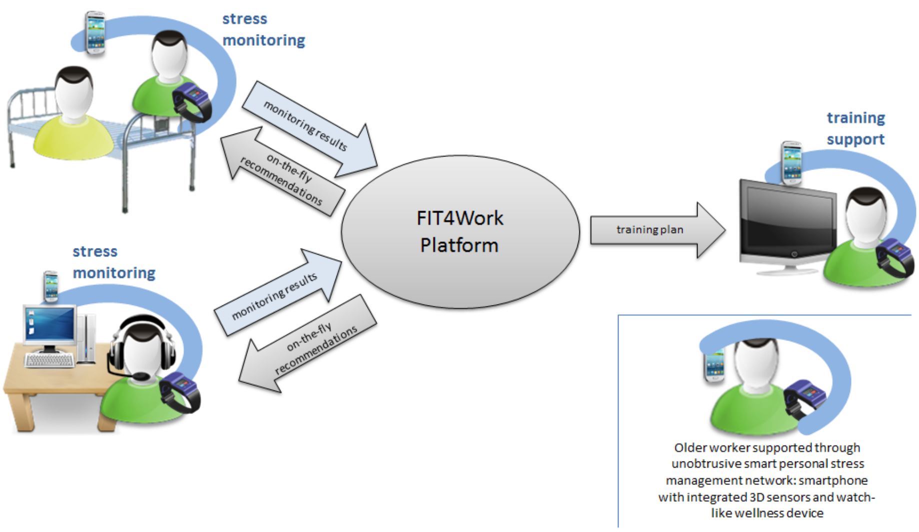


Figure 2.1 General overview of the Fit4Work system

The process followed by the system to recommend the performance of physical exercise was defined as follows:

- The stress monitoring modules check periodically the physical and mental state of the user, at home or at work;
- Taking into account the data collected from a set of wearable sensors and ambient sensors, the system checks the user state and it generates recommendations to improve the quality of life of the user;
- These recommendations may include the realization of a set of physical exercises adapted to the specific needs of the user;
- The user consults the recommendations in his/her smartphone through the Fit4Work application, loading a personalized and guided training program that is adapted to his/her physical skills.
- The user then is able to execute the suggested exercises wherever he/she is (e.g. at work, at home).
- The execution of the exercises is monitored by the 3D Motion Sensing Module that captures and analyses the user movements in order to check if the exercise is properly performed.

Criterion 1: Monitor the execution of physical exercises
The 3D Motion Sensing Module should be able to monitor the execution of a physical exercise, checking if the user is doing something wrong and providing feedback.

Project coordinator: Poznań Supercomputing and Networking Center, ul. Jana Pawła II 10, 61-139 Poznań, Poland, email: fit4work@fit4work-aal.eu

- The user gets feedback from the application that indicates if the exercise routine is correct, at the same time that it checks the health of the user.

2.2. Technology analysis

2.2.1. Sensors

Taking into account the scenarios presented, the 3D Motion Sensing Module should be easily adapted to a smartphone, for example, using an external accessory. This will facilitate the ubiquitous use of the system, wherever the user needs it.

Criterion 2: Adaptation to smartphones
The 3D Motion Sensing Module should be easily adapted to smartphones to facilitate its use from any location.

Actually, there are only a few technologies for 3D motion capture that can be adapted to a smartphone, among which these two options stand out:

- **Google Tango**: Tango is an augmented reality computing platform developed by Google, which uses computer vision techniques to enable mobile devices to extract 3D information from the environment, allowing 3D mapping, indoor navigation and 3D object recognition among other functionalities. This technology has been integrated in several Android devices (see Figure 2.2), such as the Lenovo Phab 2 and it includes an SDK for developers that enables the creation of mobile applications based on its technology. It is based on the use of IR projection to extract depth information, which is combined with high resolution RGB-IR sensors to capture the variations in the IR projection pattern and get RGB information from the environment.

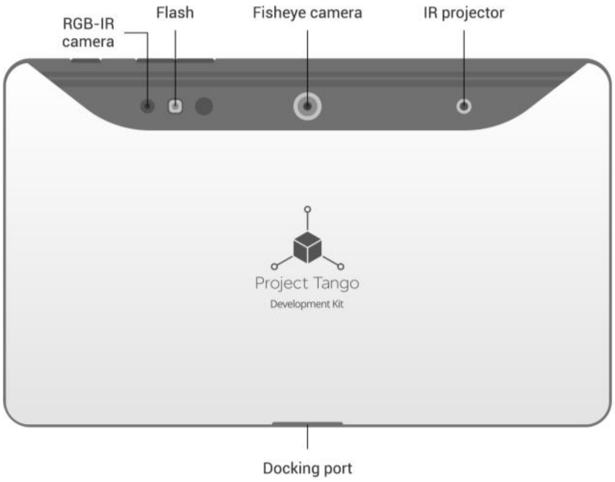


Figure 2.2 . Hardware Diagram for the Tango Tablet Development Kit

- **Structure Sensor**: It is a mobile structure light system in the form of a small and lightweight external device that provides depth sensing capabilities to collect 3D information of the environment, and instantly register it as point cloud data, making the 3D recognition and reconstruction of indoor spaces more affordable. The depth sensing module is based on the Prime Sense infrared depth-sensing technology, consisting of a laser-emitting diode, infrared radiation range projector, and an infrared sensor that send data to a system on a chip for processing. It can be used standalone, connected to a PC and providing depth information of a given scenario, or connected to a smartphone or Tablet combined with the RGB camera of the device (see Figure 2.3).

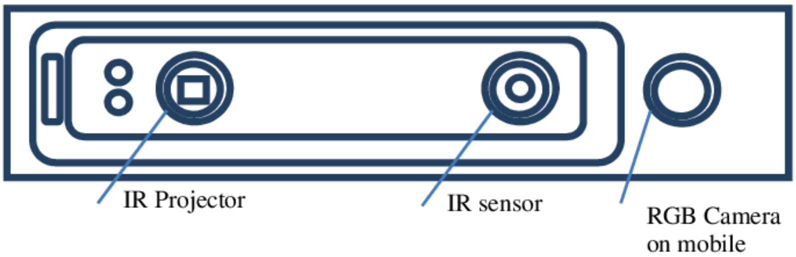


Figure 2.3 . Hardware Diagram for the Structure Sensor device

In terms of costs and performance, there are also two other options that could be taken into account, although they are not directly prepared to work with mobile devices, but they could be used in a fixed station:

- **Kinect**: This is a low-cost and easy to use device that is also based on the Prime Sense light coding technologies to get RGB-D data from the environment. It was developed by Microsoft as an accessory for gaming, but it has become one of the most common devices used for 3D scanning and 3D recognition experiments, mainly due to its reduced cost. It uses structured infrared light patterns for depth estimation, and also includes an RGB sensor (see Figure 2.4).

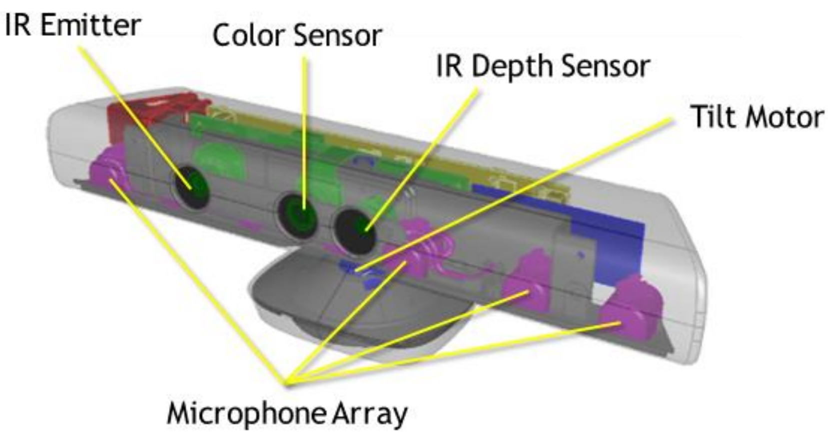


Figure 2.4 . Kinect sensor

- **Kinect 2**: This is an evolution of the Kinect device that performs Time of Flight techniques for depth sensing, which is faster and more precise than the first version of Kinect (see Figure 2.5). It is able to recognize and track 26 body joints, while Kinect only recognizes 20. Both cameras (RGB and IR) have also more resolution than Kinect's sensors.

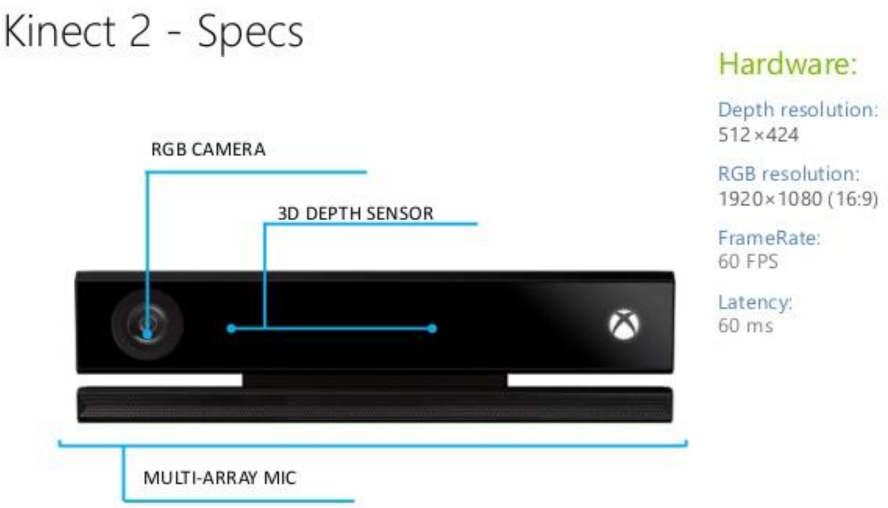


Figure 2.5 . Kinect 2 sensor

Table 2.1 presents a comparison between the mentioned technologies, taking into account the technical features and the cost:

Table 2.1 . Comparison between 3D sensor devices

Technology/ Feature	Platforms	3D Technology	Measuring Range (m)	Error	Resolution RGB	Resolution Depth	FPS	FoV	PL/SDK	Cost
Google Tango	Android	Structured Light, Time of Flight, Stereo	0,5 – 4,0	-	4MP	320x180	90	120 ° - 180 °	C/C++	~450€
Structure Sensor	iOS, Android, Windows, Mac OS, Linux	Structured Light	0,4 – 3,5	1%	640 x 480	640 x 480	30-60	58° x 45°	C/C++	~305€
Kinect	Windows, Mac OS, Linux	Structured Light	0,4 – 3,5	<4cm	640 x 480	640 x 480	15-30	57° x 43°	C/C++, C#,Java, Python, VB	~160€
Kinect 2	Windows, Mac OS, Linux	Time of Flight	0,5 – 4,5	-	1920x1080	512x424	60	70° x 60°	C/C++, C#	~160€

Another important aspect to take into account is the 3D imaging technology used in each case for the recognition of the environment in three dimensions. The available technologies are:

- Time of flight
- Stereoscopic vision
- Structured light (fixed or programmable)

The main difference among these technologies is explained in Table 2.2:

Table 2.2 . Comparison between 3D imaging technologies[1]

	Time of flight	Stereoscopic vision	Fixed structured light	Programmable structured light
Operational principle	**IR pulse, measure light transit time**	**Two 2D sensors emulate human eyes**	**Single pattern visible or IR illumination, detects distortion**	**Multiple pattern visible or IR illumination, detects distortion**
Point cloud generation	Direct out of chipset	High SW Processing	Medium SW processing	SW processing scales with # of patterns
Latency	Low	Medium	Medium	Medium
Active illumination	Yes	No	Yes	Yes – customizable spectrum
Low light performance	Good	Weak	Good	Good
Bright light performance	Medium	Good	Medium / weak Depends on illumination power	Medium / weak Depends on illumination power
Power consumption	Medium/high Scales with distance	Low	Medium	Medium Scales with distance
Range	Short to long range Depends on laser power & modulation	Mid range Depends on spacing between cameras	Very short to mid range Depends on illumination power	Very short to mid range Depends on illumination power
Resolution	QQVGA, QVGA -> Roadmap to VGA	Camera Dependent	Projected pattern dependent	WVGA to 1080p -> Roadmap to WQXGA
Depth accuracy	mm to cm Depends on resolution of sensor	mm to cm Difficulty with smooth surface	mm to cm	µm to cm
Scanning speed	Fast Limited by sensor speed	Medium Limited by software complexity	Medium Limited by SW complexity	Fast / medium Limited by camera speed
Applications				
Location	✔	✔		✔
Identification	✔	✔	✔	✔
Measurement & Inspection	✔	✔	✔	✔
Biometrics				✔
UI Control/ gaming	✔		✔	
Augmented reality	✔	✔		✔
	Learn more	Learn more	Learn more	Learn more

[1] Texas Instrument – Comparison of 3D imaging technologies, available here: http://www.ti.com/sensing-products/optical-sensors/3d-time-of-flight/technology-comparison.html

Taking into account the requirements of the Fit4Work system, we have to select a technology adapted to smartphones. As the Structure Sensor is a more robust and stable technology, we select it as the principal 3D technology for providing 3D imaging capabilities to the Fit4Work system. Besides, the Structure Sensor has proven to provide a higher accuracy for smaller rooms at short distance (1.5-2m), and it is possible to use several Structure sensors in the same area without interference[2].

Although the Structure Sensor is mainly oriented to iOS devices, it can work in other platforms, for which a set of Open Source Drivers are available[3]. There is an official Structure SDK mainly prepared for the creation of native applications in XCode (iOS), but there are also plugins that allow integrating the sensor in a Unity (C#) project which is a game development platform that enables the creation of high quality 3D/2D multiplatform applications. As Unity is also compatible with the Kinect Sensor, the 3D Motion Sensing Module can be implemented for both technologies using the same project (no extra effort required).

2.2.2. Capturing movement with RGB-D sensors

On the one hand, the use of depth data combined with RGB information simplifies the task of tracking human movements compared with the use of traditional RGB image sensors. The depth data is invariant to changes in the colours or the illumination of the scene. Besides it allows filtering the elements by the distance to the sensor, which allows detecting the users in front of the sensor with a higher accuracy for human tracking applications. Thus, the use of depth information will improve the results of gesture and movement recognition.

On the other hand, the RGB data captured can be displayed to the user during the exercises as if in a mirror, enabling the users to see and control their poses while doing the exercises. As the data is also being processed and analysed, users can also get real-time feedback of the good or bad realization of the selected exercise. This is one of the advantages of using image sensors for exercise monitoring instead of other movement sensors, such as wearables with accelerometers.

Furthermore, virtual elements can be displayed on the screen used as a mirror to guide the users in the realization of the exercises as in videogames. The use of gamification techniques has been proved to increase the engagement of users with the realization of activities.

Finally, the RGB videos could be recorded and sent to rehabilitation or medical specialists so that they can check the performance of the exercises and the physical evolution of the users.

As an example, the Spanish company VirtualWare is an existing success story with their product *VirtualRehab*[4] that makes use of these technologies for the rehabilitation of patients (see Figure 2.6).

---

[2] G. Leslie, S. Spann, et al.3D Scanning Technology – A Comparative Analysis. Part of the MCAA's Construction Technology Research Series, performed by JBKLABS, May 2016.
[3] Open Source drivers for the Structure Sensor: https://forums.structure.io/c/open-source-drivers
[4] http://www.virtualrehab.info/

Figure 2.6 . VirtualRehab application

2.2.3. Computer Vision libraries

Actually, there are several open source libraries that can be used for the recognition of human movements and the analysis of physical exercises using RGB-D information. These are the most relevant:

- **OpenNI2 Framework**: set of APIs for the development of applications that use natural interactions between the users and the devices. This library provides body motion tracking, hand gesture and voice recognition, and can be used with both the Kinect sensor and the Structure sensor.
- **OpenCV library**: this is the most popular computer vision library, providing a wide range of image processing utilities and computer vision algorithms that can be used for the design of modules for the recognition of human movements. It has a very big community of developers using openCV and providing support for any technical issue or problem that may occur, and there are also many computer vision problems already solved that can be used. It is possible to process RGB and depth data with this library, thus it is only required to get the drivers for the sensors selected, that allow extracting the RGB/depth streams. Libfreenect already solves this issue for Kinect and Kinect 2.

2.3. User profile

The target users for the system are healthcare workers with these profiles:

- Persons aged 55-65, working in homecare or stationary care services;
- Persons aged 65+, working as volunteering informal caregivers in a hospital;
- Persons aged 65+, working as volunteering informal tele-assistants.

Criterion 3: Adaptation to user profile
The 3D Motion Sensing Module should be suitable for users over 55 years old.

2.4. Exercises proposal

The system is mainly aimed at maintaining the physical and mental capabilities of adults over 55 years old, considering the particular case of the people working for the health sector.

Regarding the physical state, the system should take into account the musculoskeletal disorders (MSDs), which are one of the most common physical problems among healthcare workers. Particularly, low back pain and musculoskeletal pain in neck and shoulders are commonly suffered by people working in the health sector.

Criterion 4: Prevention of MSDs
The 3D Motion Sensing Module should be able to evaluate physical exercises that help preventing musculoskeletal disorders, particularly those related with problems in low back, neck and shoulders.

With this criterion, and taking into account the state-of-the-art on 3D imaging technologies for the recognition and evaluation of physical movements, a set of exercises was selected as potential, example exercises to be evaluated with the Fit4work tools. Literature analysis suggests the following as exercises that should be considered when developing the 3D sensing module[5]:

- The study of Freimann et al[6], who investigated the effects of an 8 week home exercise therapy programme in people suffering from moderate MSDs in the neck or lower back.
- The Low Back School method[7 8], which teacher care and body protection mechanisms.
- General fitness training recommendations for the realization of shoulder exercises.

Table 2.3 describes some of these exercises.

Table 2.3 . Examples of physical exercises that should be considered when developing the 3D sensing module

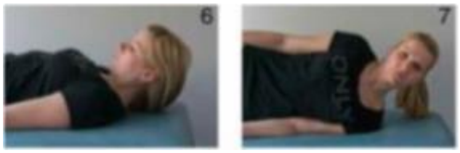
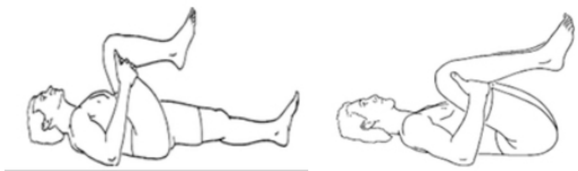
Type	Part of the body	Visual description
ROM exercises	neck	

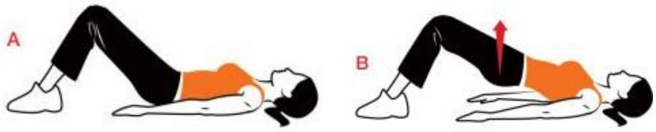
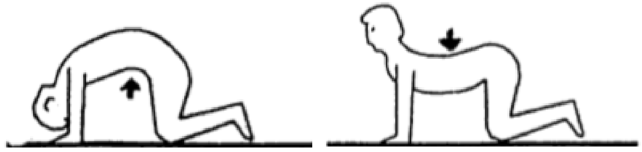
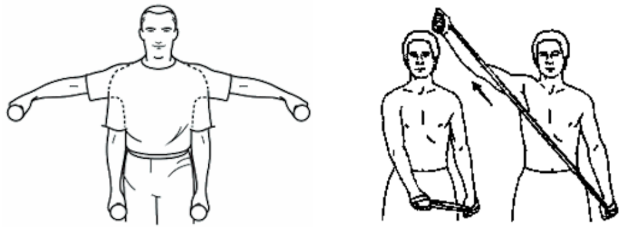
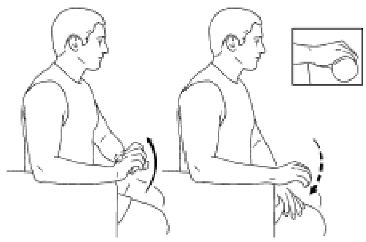
[5] See deliverable D4.5.1/4.5.2 for discussion on actually selected program of physical exercises

[6] Freimann T, Merisalu E, Pääsuke M. Effects of a home-exercise therapy programme on cervical and lumbar range of motion among nurses with neck and lower back pain: a quasi-experimental study. BMC Sports Science, Medicine and Rehabilitation. 2015;7:31. doi:10.1186/s13102-015-0025-6.

[7] Josa, R. M.; Derm, J.: La prevención de las lesiones de espalda en el trabajo hospitalario. Salud trab. 1989.

[8] Santandreu, M. E.; Sánchez, J. ; González, J.; Gómez, A.: Dolor vertebral en el personal hospitalario. Rehabilitación 1994.

Type	Part of the body	Visual description
Stretching exercises	neck	
Strengthening exercise	neck	
ROM back exercises	core, back	
Strengthening exercise	core, back, lumbar muscles, glutes	
Unilateral stretching exercise	back, lumbar muscles	
Breathing exercises	Low back	
Unilateral strengthening and abdominal toning	Low back, core	

Type	Part of the body	Visual description
Glutes and abdominal strengthening	Glutes, core	
Lumbo-pelvic flexibilization exercises	Low back	
Shoulder strengthening exercises	shoulders	
Wrists strengthening exercises	Wrists	

2.5. Evaluation of the exercises from a technological perspective

As mentioned in *section 2.2 Technology analysis*, the Fit4work system will integrate a 3D Motion Sensing module based on the use of depth sensors, such as the Structure Sensor or the Microsoft Kinect sensor, to analyze the exercises performed by the users in an indoor environment. These sensors, combined with computer vision algorithms, should be able to capture the user movements and determine if an exercise previously selected is correctly executed by the user.

The objective of this section is the evaluation of the use of depth sensors combined with computer vision algorithms for monitoring users while they execute the physical exercises proposed, trying to determine the best approach for this project.

There are many situations that can hinder the recognition of the user gestures: loose clothing, occlusions, realization of movements not easily visible even by the human eye, etc. Not every exercise can be properly monitored using cameras, but there are a big variety of them that can be recognized with a good accuracy, and, as explained, we think that the use of depth cameras can be an advantage in terms of user experience. Therefore, the exercises already proposed to achieve the goals of the Fit4Work project have been reviewed and classified in terms of technical feasibility:

- **Full body frontal exercises**: Exercises in which the user is in front of the sensor, the full body appears in the image and there are not occlusions. In this case, the exercises could be analysed using skeleton tracking algorithms.

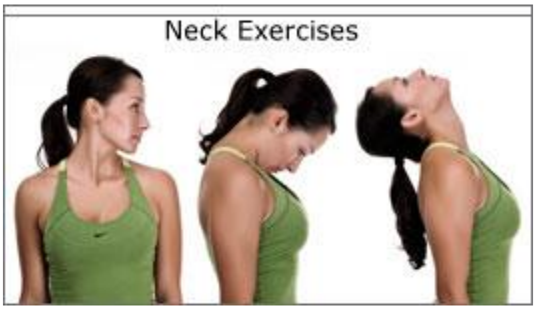
Examples: arms extensions, lateral steps, hip abductions, knee lifts…


- **Full body lateral exercises**, in which the camera is also placed in front on the user capturing the whole body, but the exercises should be reviewed laterally, so part of the body is not seen by the sensor. This case is solved also with skeleton tracking algorithms but the performance can be reduced as part of the body is hidden. There are state-of-the-art algorithms that deal with this problem with a more exhaustive calibration and more complex tracking techniques.

Examples: knee lunges, frontal steps, abdominal toning exercises…


- **Partial body exercises**, in which just one area of the body is captured or analysed. In this case, it will be necessary to create specific computer vision algorithms to track the part of the body under evaluation. The OpenCV library already provides a set of algorithms for the detection of faces, eyes, hands, and other elements that could simplify the design of these algorithms.

Examples: neck exercises, back exercises, hand movements, etc.



- **Exercises using external elements** (dumbbells, chairs, elastic bands, etc.) that could have visible markers that facilitate monitoring the correct realization of the exercise. This type of exercises can make use of marker recognition algorithms for object detection and tracking.

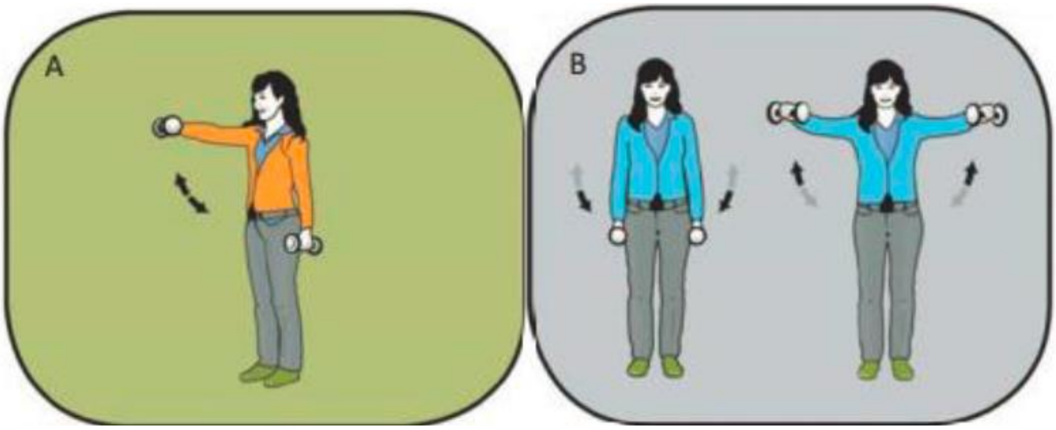


Table 2.4 summarizes the differences in each case, regarding the technological requirements and the extensibility to new exercises of the same category not taken into account at the design stage.

Table 2.4 . Categorization of exercises from a technological point of view

ID	Type	Sensor	Algorithms	CV Libraries	Technical Complexity	Extensibility	Maintainability
Fit-Ex-1	*Full body frontal exercises*	Depth or RGB-D	Skeleton tracking	OpenNI2	+	++++	++++
Fit-Ex-2	*Full body lateral exercises*	Depth or RGB-D	Skeleton tracking	OpenNI2	++	+++	+++
Fit-Ex-3	*Partial body exercises*	RGB or RGB-D	Custom tracking	OpenNI2 and OpenCV	++++	+	+
Fit-Ex-4	*Exercises using external elements*	RGB or RGB-D	Marker tracking & body tracking	OpenNI2 and OpenCV	+++	++	++

Although it would be interesting to research the integration of all type of exercises in the Fit4Work platform, the efforts will be focused on those exercises that can be evaluated with a depth sensor, that are *Fit-Ex-1* and *Fit-Ex-2*. Besides, it would be easier to add new exercises in the categories that require skeleton tracking algorithms, as it is simpler to characterize and exercise in relation to the body parts involved in its movements.

3. Module design

3.1. Technical requirements

Apart from the criteria described in the previous section, there are also other requirements for the module from a technical point of view, taking into account the functionalities expected from the module, as listed in Table 3.1.

Table 3.1 . Technical requirements for the 3D Motion Sensing module

ID	Description
Functionality 1	**The module shall be able to recognize the physical exercises executed by users**
FR_1.1	The module captures data from users while they execute a specific exercise using a depth or RGB-D sensor.
FR_1.2	The data captured from users shall be sufficient to recognize their movements.
FR_1.3	The module allows to characterize and model an exercise.
FR_1.4	The module is able to compare a modelled exercise with the movements of a user.
FR_1.5	The module is capable of recognizing some of the exercises proposed.
Functionality 2	**The module shall be able to extract information of the body of users**
FR_2.1	The module is able to understand automatically the information provided by the depth or RGB-D sensors (depth and RGB data).
FR_2.2	The module is able to identify the different body parts of a user.
FR_2.3	The module recognizes a specific pose of the user based on the relative position of different body parts.
FR_2.4	The module recognizes a specific exercise as a sequence of poses.
Functionality 3	**Registration of new exercises in the system**
FR_3.1	The module allows to registrate a new exercise to be evaluated and recognized.
FR_3.2	The module stores all the exercises to be recognized.
FR_3.3	The module provides a list of exercises already registered in the system.
Non-functional requirements	
NFR_1	The module is compatible with the Structure Sensor and the Kinect sensor.
NFR_2	The module is embeddable in a Unity project (C#).
NFR_3	The user interface should be adapted to users aged 55+.
NFR_4	The module should be adapted to smartphones (iOS/Android).
NFR_5	Optionally, the module should also work in desktop versions (Windows platforms).

3.2. General overview

Figure 3.1 presents the main components that configure the system, and the interaction between them. Table 3.2 describes those modules in more detail.

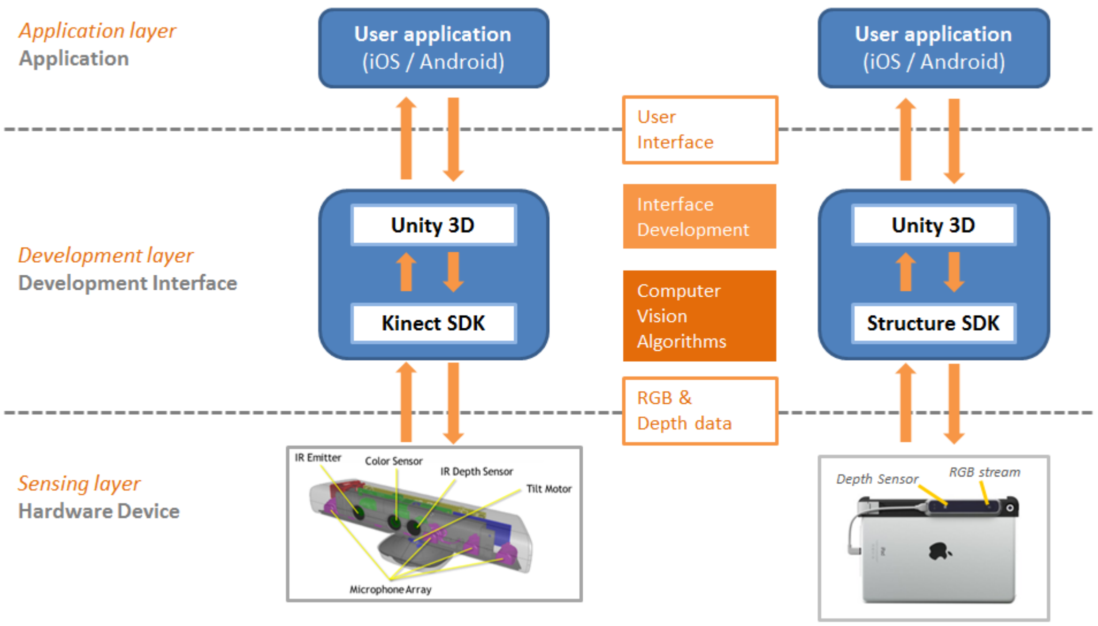


Figure 3.1 . Module diagram for Kinect and Structure sensors

Table 3.2 . Components of the module

Layers	Description
Sensing layer	This layer is responsible for capturing the depth and RGB data from users while they perform the physical exercises. Two type of sensors will be tested: Kinect and Structure sensor.
Development layer	This is the intermediate layer, which is aimed at processing the data provided by the sensing layer to extract relevant information related to the user movements. Computer vision algorithms will be used for body recognition and skeleton tracking, in order to detect: • The body parts (body joints) • The user pose • The execution of a specific sequence of poses In this layer, the comparison of the user poses with the reference poses of the exercises expected is also carried out. To simplify the creation of a solution for multiple platforms (iOS, Android and Windows), a Unity project will be created, in which the logic of the module will be integrated. For the use of both sensors, two development kits will be used: Kinect SDK and Structure SDK.
Application layer	The objective of this layer is the presentation of information to users, and the provision of an interaction channel, from which users can load an exercise, start the automatic evaluation of exercises and get feedback to know if they are moving correctly.

The functioning of the module is as follows (see Figure 3.2):

- The user starts the Fit4work application and sees the customized recommendations of exercises proposed by the platform.
- The user selects one of the exercises.
- The 3D motion sensing module is the initiated.
- The module captures the RGB and depth data provided by the sensor.
- The data is pre-processed to remove noise and enchance the features of interest for the body joint detection.
- The user is detected by using segmentation techniques, identifying the body joints.
- After this, the selected exercise is loaded. This exercise consists of a sequence of poses to be checked automatically.
- Once the exercise is loaded, the module analyses the body joints of the user in order to determine if the first pose of the exercise is performed.
- After the first pose is completed, the module looks for the next pose, and so on and on until the sequence of the exercise is fully completed.
- For each pose, and also at the finalization of the exercise, the module provides feedback to the user giving information of the good or bad performance of a movement.

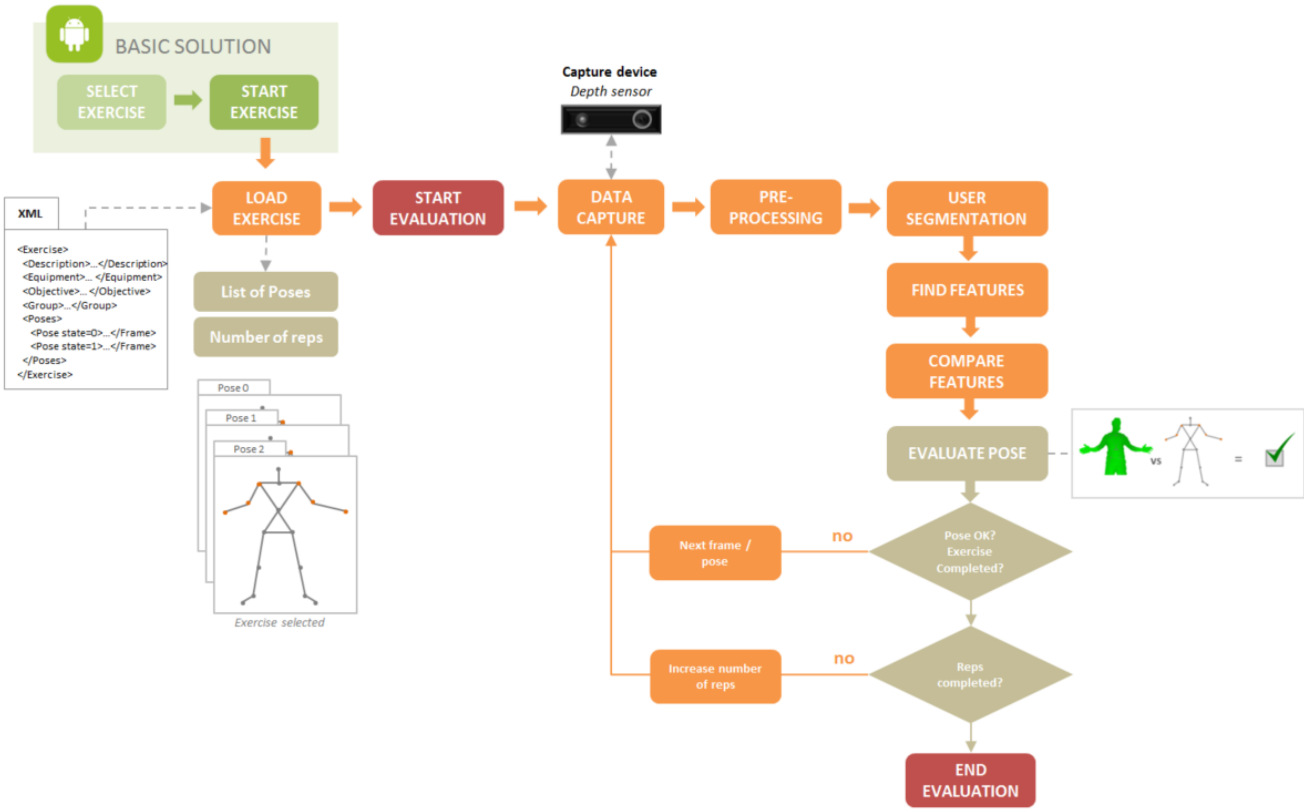


Figure 3.2 . Functioning of the 3D Motion Sensing module

Functionally, the module is composed by three elements or sub-modules:

- Capture sub-module
- User analysis sub-module
- Exercise evaluation sub-module

3.2.1. Capture sub-module

This component is responsible for obtaining the RGB and depth images from the capture sensor (Structure or Kinect).

It should make the following checks before processing any information:

- The communication with the sensor has been correctly established.
- The computer vision libraries are correctly installed and ready for detecting and tracking the user. For the recognition of the two groups of exercises selected, *Fit-Ex-1* and *Fit-Ex-2*, the OpenNI2 library will be used, which provides skeleton tracking capabilities.

After that, the module starts requesting data from the sensor and sends it to the User analysis sub-module for processing (**Data Capture**).

3.2.2. User analysis sub-module

This part is aimed at processing the RGB & depth data provided by the capture module to detect the user and extract his pose, in order to determine if it is similar to the pose indicated in the exercise selected.

These are the processing tasks executed by this sub-module:

- **Pre-processing**: the depth and RGB data is pre-processed to enhance the relevant features, related with the body gesture of the user.
- **User segmentation**: the user is located in each frame and the **body joints** are detected (see Figure 3.3). The body joints are the different parts of the body of a user that are going to be tracked to determine if a movement is correctly executed. The Kinect SDK and the Structure SDK allow to work with 20 body joints, which are:
    - Right and left hand.
    - Right and left wrist.
    - Right and left elbow.
    - Right and left shoulder.
    - Right and left hip.
    - Right and left knee.
    - Right and left ankle.
    - Right and left foot.
    - Shoulder center and hip center.
    - Spine.
    - Head.

  For each of this body points, the skeleton tracking libraries provide this information:
    - Color RGB
    - Position (X, Y and Z values).

- Orientation (X, Y, Z and W values).

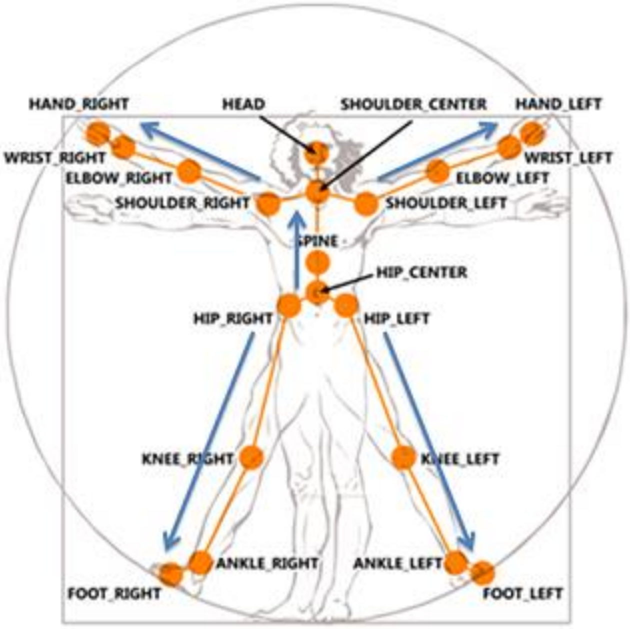


Figure 3.3 . Common body joints detected by skeleton tracking libraries

- **Find features**: which consists of extracting the body joints of the user at a certain moment.

This sub-module provides information of the movements and poses to the Exercise evaluation sub-module, so that it can compare the data obtained from the user with the data of the exercise loaded. To simplify the explanation of the functioning of this module, we will use the term „user pose" to refer to the configuration of the body joints of the user at a certain moment.

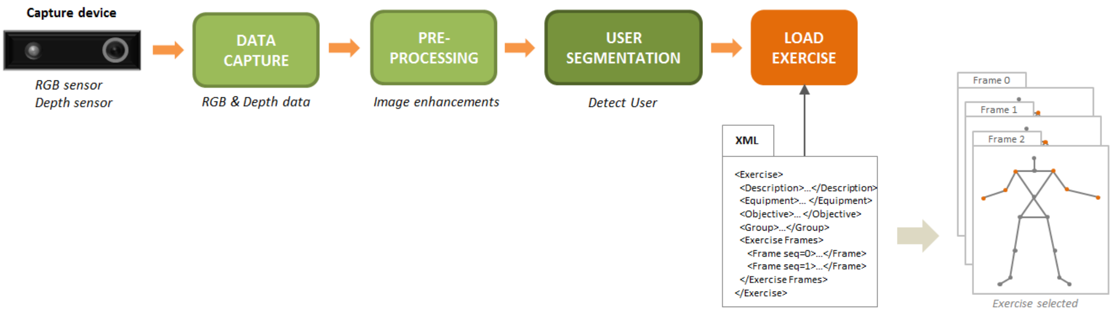


Figure 3.4 . Functional blocks of the 3D Motion Sensing module (I)

3.2.3. Exercise evaluation sub-module

It is responsible for loading an exercise selected by the user and verification that it is correctly performed. For this, the module will first load the model of the exercise selected in order to get the sequence of steps to be carried out by the user to consider the exercise completed. This module will produce a set of statistics obtained from the realization by the user.

This sub-module works as follows (see Figure 3.5):

- **Load exercise**: once an exercise has been selected for practice, this sub-module gets the model of that exercise, which consists of a set of poses to be executed in a certain order.
- **Request features**: the sub-module first load the first pose of the exercise, and request the current body pose to the User analysis sub-module.
- **Compare features and evaluate pose**: the evaluation sub-module compares the user pose with the pose expected. If it matches, a positive feedback is given to the user, and the next pose of the exercise is loaded. If it does not match, a negative feedback is sent to the user.
- **Exercise evaluation loop**: The evaluation sub-module continues comparing the user poses with the exercise poses until the whole exercise is completed.

It would be useful to add a timeout, to let users rest if they fail to perform the exercise correctly.

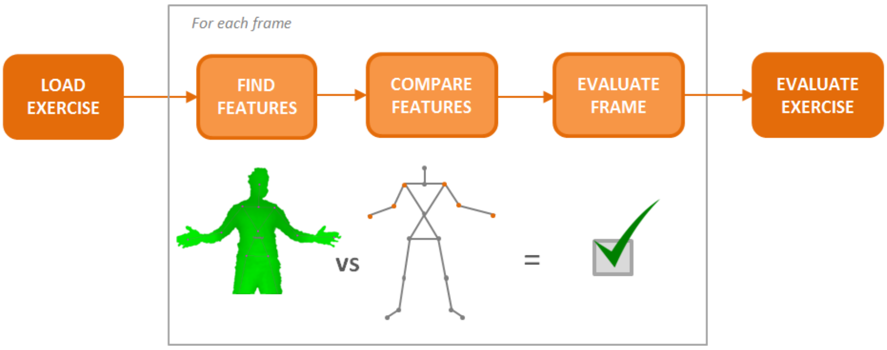


Figure 3.5 . Functional blocks of the 3D Motion Sensing module (II)

3.3. Exercise modeling

In order to evaluate if an exercise is being performed correctly by the user, it is necessary to analyse the different steps that have to be followed to consider an exercise completed.

The process of characterization of an exercise consists of the following steps:

1. Describe the exercise, including actions, objectives and equipment (if necessary).
2. Define the body features that will be tracked for the exercise (body joints).
3. Divide the exercise in different stages, which we will call “exercise frames”. In this way, an exercise can be conceived as a sequence of exercise frames.
4. Characterize each exercise frame, indicating the relative position of the body features in the frames, and other relevant information for the realization of the exercise, such as the period of time that the user should stay in that position or the accepted error rate.

A structured data format will be used to model and store the exercises that will be available in the application, for example XML. The template in Table 3.3 can be used to facilitate the collection of information of each exercise for the creation of the structured data models (example included).

Table 3.3 . Template for the definition of exercises

Exercise	One foot balance
Description	1. Stand up with your feet placed about hip-width apart 2. Maintain balance with one leg in abduction. You can use a wall or a chair as a support if necessary. 3. Return to the starting position 4. Repeat with the other leg
Equipment	None
Body features	Body joints (knee right, knee left, foot right, foot left…)
Objective	Static balance
Group	Resistance exercise
Frames	 **1**  **2**  **3**  **4**

Table 3.4 features an example of XML file for the definition of exercises following the presented structure.

Table 3.4 . XML file for the definition of exercises

```
<Exercise>
  <Description>…</Description>
  <Equipment>… </Equipment>
  <Objective>… </Objective>
  <Group>…</Group>
  <Exercise Frames>
    <Frame seq=0>…</Frame>
    <Frame seq=1>…</Frame>
  </Exercise Frames>
</Exercise>
```

The use of a structured format for the definition and storage of exercises facilitate the automatic evaluation of an exercise. Furthermore, this feature could be used to implement a tool for exercise registration by recording a session of a user executing correctly an exercise, in which all the user poses are stored as XML files that include information of the position of the body joints. This tool is not immediate, as it requires the correct filtering of data to remove any possible noise or unwanted information, thus it is proposed as future work.

Future Work: Tool for exercise registration
The execution of an exercise could be recorder in the form of an XML or a set of XML files registering every pose of the user during the movements. If the user is an expert, this exercise could be stored as a reference for future evaluations.

4. Implementation of the 3D Motion Sensing module

In this section, the resulting 3D motion sensing module is explained.

4.1. Technologies used

- Kinect SDK v1.8
- OpenNI2 v2.2
- NiTE v2.2
- Unity 3D v5.4

4.2. Module structure

Figure 4.1 depicts the structure of the Unity project created for the 3D image sensing module, which integrates a user interface for desktop devices.

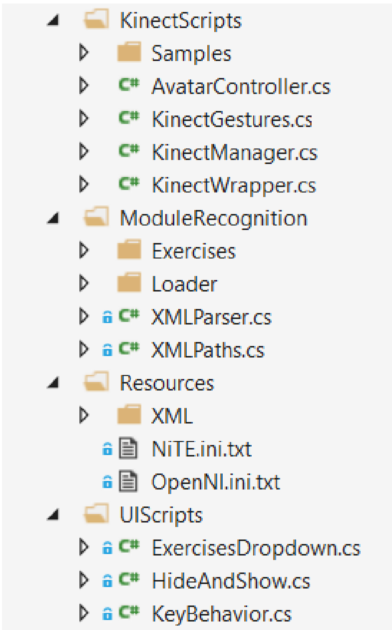


Figure 4.1 . Structure of the project of the 3D Motion Sensing module

- **KinectScripts**: this is the main module of the project, which integrates all the necessary logic for the coordination of the different processes. It includes the User analysis sub-module, and the classes for the representation of the movements of the user with an avatar in the user interface.
    - **KinectManager:** this is the coordinator of the application, that initiates the libraries and loads the different modules when needed.
    - **KinectWrapper**: this is the class responsible for the Data Capture from the sensor using OpenNI2 and NITE libraries. This class provides the user poses (position of the body joints).
    - **KinectGesture**: this is the class that compares the user poses with the pose expected in an exercise, showing feedback through the user interface with the results of the comparison.
    - **AvatarController**: this class is used by the user interface to move the avatar in the scene. It gets the user pose (position of the body joints) and updates with it the pose of the avatar.

- **ModuleRecognition**: this is the module responsible for the management of the exercises stored in the system. It provides the XML of an exercise selected and parses it to be interpreted by the KinectGesture class.
- **Resources**: this is a folder containing all the configuration files needed for the functioning of the 3D motion sensing module, and also the folder with the exercises stored in XML format.
- **UIScripts**: this is the folder containing all the views and logic of the user interface.

Once created, the project can be easily configured for its use with the Kinect or the Structure sensor changing one option in the configuration files of the OpenNI2 Library, specifically in the PS1080.ini file and the UsbInterface parameter:

- For **Kinect sensor**: UsbInterface=**2**
- For **Structure sensor**: UsbInterface=**0**

4.3. Exercise selection

In order to determine the complexity of the movements in each exercise (based on input of deliverable 4.5.1/4.5.2), videos of the exercises were recorded in an ideal scenario (see Figure 4.2).

Figure 4.2 . Recordings of the correct execution of the exercises

These videos were evaluated taking into consideration the technical constraints of each exercise proposed. For instance, exercises were the user is laterally placed (Fit-Ex-2) are more difficult to evaluate than those that are frontal to the sensor (Fit-Ex-1); in other cases the clothes hide the body parts hindering the visualization of the body parts (e.g. lumbar extension exercises).

The first version of the module was prepared for the recognition of these exercises:

- Lateral step (left/right)
- Double lateral step (left/right)
- Lateral jumps (two legs)
- One leg lateral jump (left leg/ right leg)
- Forward kick (left leg/ right leg)
- Knee raise (left/right)
- Hand raise (left/right)
- Deep squat

The idea was to develop the module considering the most attainable cases, and evaluate the added value of the module with those, leaving for later the integration of improvements. For each of these exercises, an XML was manually created, indicating the relative position of the body joints in each of the poses that conform an exercise (see Table 4.1 for an example).

Table 4.1 . Example of XML file for the Knee Raise exercise

```
<?xml version='1.0'?>
<Exercise>
        <Description>
                Ejercicio para levantar la rodilla izquierda hasta la altura de la cintura
        </Description>
        <Equipment>
                Ropa ligera que permita desplazarse sin dificultad
        </Equipment>
        <Objective>
                Levantar la rodilla izquierda hacia arriba
        </Objective>
        <Group>
                Personas mayores de 55 años
        </Group>
        <Poses>
                <Pose state="0">
                        <Tracked>
                                <Joints>
                                        <Joint key="8" value="HIP_CENTER"></Joint>
                                        <Joint key="11" value="LEFT_KNEE"></Joint>
                                </Joints>
                        </Tracked>
                        <Operations>
                                <Operation index="0" threshold="0.3">
                                        <Joints>
                                                <Joint key="8" value="HIP_CENTER"></Joint>
                                                <Joint key="11" value="LEFT_KNEE"></Joint>
                                        </Joints>
                                </Operation>
                        </Operations>
                </Pose>
                <Pose state="1">
                        <Tracked>
                                <Joints>
                                        <Joint key="8" value="HIP_CENTER"></Joint>
                                        <Joint key="11" value="LEFT_KNEE"></Joint>
                                </Joints>
                        </Tracked>
                        <Operations>
                                <Operation index="0" threshold="0.3">
                                        <Joints>
                                                <Joint key="8" value="HIP_CENTER"></Joint>
                                                <Joint key="11" value="LEFT_KNEE"></Joint>
                                        </Joints>
                                </Operation>
                        </Operations>
                </Pose>
        </Poses>
</Exercise>
```

The manual creation of theses XMLs may be hard work in some cases, particularly when the exercise is long including a wide variety of poses. Besides, it is important to check the poses with an expert that is able to determine which are the most important aspects in the realization of an exercise, taking into account that it is possible to define the position of the 20 body joints mentioned in the XYZ axes.

Note: Expert assessment for the definition of exercises
An expert on the practice of physical exercises should asses during the definition of an exercise, pointing to the most critical aspects in the movement.

4.4. Module evaluation

4.4.1. Setup

The Kinect and Structure sensors were connected to a computer with Windows 8 OS, taking into account the recommendations presented in Figure 4.3.

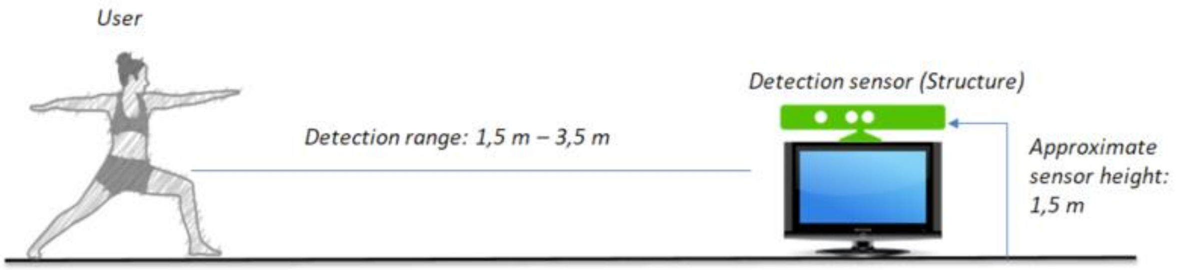


Figure 4.3 . Recommendations for the system setup

As it can be observed in the figure, the sensor works correctly if the user is positioned inside the recommended detection range (1,5 - 3.5 meters from the sensor). Regarding the height of the sensor, to capture the whole body of a person, the minimal distance to the floor should be 1,5 meters. Anyway, it is possible to check from the user interface if the user is completely being captured by the sensor.

4.4.2. User interface

To facilitate the use and the evaluation of the module, a user interface was developed (see Figure 4.4). This user interface shows an avatar that copies the movements of the user, which are captured by the Kinect or the Structure sensor. This way the user can see what the module is detecting, and he also can get feedback of the execution of an exercise.

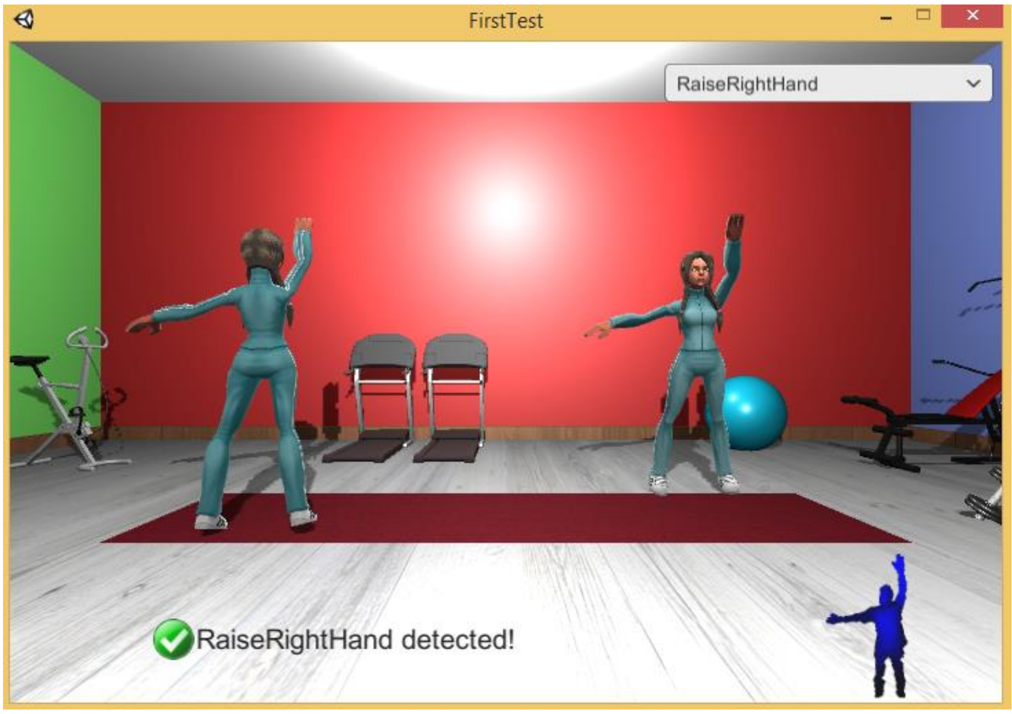


Figure 4.4 Screenshot of the user interface

These are the main components of the interface:

- **Exercise selection**: in the right up corner, there is a combo box that allows to select the exercise to be practiced from a list of available exercises, which have been previously modelled in the system with the corresponding XML file.
- **User shilouette**: in the right down corner, the shilouette of the user is shown, as it is segmented by the user analysis sub-module. Looking at this shilouette it is possible to check if the full body of the user is being captured by the sensor.
- **Central scene**: in the central part of the interface, there is a scene that contains two avatars, that are sincronized with the user movements, showing the front and the back of the avatar.
- **Feedback section**: in the left down corner, the feedback provided to the user is shown, indicating if the expected pose has been correctly detected.

4.4.3. Example of use

Figure 4.5 presents an example of use of the impelemented module. As it can be observed, it is possible to see the use of the module for the „leg raise” and the „hand raise” exercises.

Figure 4.5 Screenshots of the use of the module

In both cases the avatar is synchronized with the movements of the user, as it get information of the user body joint by the Structure sensor.

The use of avatars reproducing the movements of the user, make the application more intuitive and easier to use, so that the user can know what is being understood by the system.

5. Evaluation

5.1. Module validation

Table 5.1 summarizes the achievement of objectives, taking into account the functionalities expected from the module and the requirements defined for it.

Table 5.1 . Evaluation of the results

ID	Description	Validated
Functionality 1	**The module shall be able to recognize the physical exercises executed by users**	
FR_1.1	The module captures data from users while they execute a specific exercise using a depth or RGB-D sensor.	True
FR_1.2	The data captured from users shall be sufficient to recognize their movements.	True
FR_1.3	The module allows to characterize and model an exercise.	True
FR_1.4	The module is able to compare a modelled exercise with the movements of a user.	True
FR_1.5	The module is capable of recognizing some of the exercises proposed.	To be improved
Functionality 2	**The module shall be able to extract information of the body of users**	
FR_2.1	The module is able to understand automatically the information provided by the depth or RGB-D sensors (depth and RGB data).	True
FR_2.2	The module is able to identify the different body parts of a user.	True
FR_2.3	The module recognizes a specific pose of the user based on the relative position of different body parts.	True
FR_2.4	The module recognizes a specific exercise as a sequence of poses.	True
Functionality 3	**Registration of new exercises in the system**	
FR_3.1	The module allows to registrate a new exercise to be evaluated and recognized.	True
FR_3.2	The module stores all the exercises to be recognized.	True
FR_3.3	The module provides a list of exercises already registered in the system.	True
Non-functional requirements		
NFR_1	The module is compatible with the Structure Sensor and the Kinect sensor.	True
NFR_2	The module is embeddable in a Unity project (C#).	True
NFR_3	The user interface should be adapted to users aged 55+.	To be improved
NFR_4	The module should be adapted to smartphones (iOS/Android).	True
NFR_5	Optionally, the module should also work in desktop versions (Windows platforms).	True

5.2. Conclusions and future work

We have developed a 3D motion sensing module capable of comparing the user movements with a specific exercise, previously registered in the system, to check if the exercise is correctly executed.

The module is currently prepared for the execution of very simple exercises, and can be improved with the following features, that remain as future work:

- Recognition of more complex exercises, properly defined with the assessment of an expert in the performance of physical activities.
- Implementation of a tool for the automatic registration of exercises, by recording the session of a user.
- Use of additional avatars as coaches showing the correct execution of an exercise.

Following the recommendation of Midterm Review experts this module has not been included in the minimum viable product as developed as the final Fit4Work system prototype (see Deliverable 5.4.2). It is foreseen as a possible extension of the system in the future, replacing also the currently developed, basic functional exercise component as presented in deliverable 4.5.2.