A Robot-based Application for Physical Exercise Training

Michalis Foukarakis, Ilia Adami, Danae Ioannidi, Asterios Leonidis, Damien Michel, Ammar Qammaz, Konstantinos Papoutsakis, Margherita Antona and Antonis Argyros
Foundation for Research & Technology – Hellas (FORTH), Institute of Computer Science, Vassilika Vouton, 70013 Heraklion, Greece

Keywords: Physical Exercise, Elderly People, Assistive Robot, Human Motion Capture, Gesture Recognition.

Abstract: According to studies, performing physical exercise is beneficial for reducing the risk of falling in the elderly and prolonging their stay at home. In addition, regular exercising helps cognitive function and increases positive behaviour for seniors with cognitive impairment and dementia. In this paper, a fitness application integrated into a service robot is presented. Its aim is to motivate the users to perform physical training by providing relevant exercises and useful feedback on their progress. The application utilizes the robot vision system to track and recognize user movements and activities and supports multimodal interaction with the user. The paper describes the design challenges, the system architecture, the user interface and the human motion capturing module. Additionally, it discusses some results from user testing in laboratory and home-based trials.

1 INTRODUCTION

Physical well-being affects a great deal the overall quality of life of any person, but even more so that of a senior, as it may be one of the determining factors for allowing that person to continue living an independent life at home. The multiple positive effects of regular physical exercise on maintaining and improving the overall physical and mental well-being of the elderly are well known and documented. Physical fitness is associated with better balance, motor skills, and improved reaction times for the elderly (Buchner et al, 1997). As data acquired from multiple studies have demonstrated (Laurin et al., 2001; Heyn et al., 2004), it has also been associated with a lower risk of developing cognitive disorders like dementia. Improving upper limb function and increasing trunk endurance and trunk strength has been proven to improve balance (Suri et al., 2009) and stabilize posture (Kaminski et al., 1995), thus reducing the risk of falls. However, in order for any exercise routine to be beneficial to the elderly, it is important that it is designed specifically for their needs, while taking into account all the potential physical limitations that may be present due to an ailment or due to age related dysfunctions (e.g., limited flexibility of the joints, shorter ranges of motion, etc.). In addition, supervision by a trainer is essential for ensuring that the exercises are executed correctly, thus limiting the risk of injury, and for providing appropriate encouragement or additional instructions when necessary. Many nursing and care facilities for the elderly offer exercise classes instructed by trainers. However, access to such structured classes is not always easy or possible for elderly people for many reasons, such as lack of transportation, inability to leave the house easily due to health conditions, geographic isolation, schedule conflicts, etc.

A potential solution to this problem has been explored in the context of the EU funded project HOBBIT[1]. For this purpose, a fitness application has been designed and developed, which is integrated in the HOBBIT robotic platform for the elderly to use on demand. More specifically, this application is a guided fitness routine especially designed for the needs of the elderly that can be practiced in a safe and pleasant way. The application runs as a component of the HOBBIT robot system, a low cost mutual care assistive robot targeted to reduce the risk of falling at home through a variety of functions.

[1] www.hobbit-project.eu

Foukarakis, M., Adami, I., Ioannidi, D., Leonidis, A., Michel, D., Qammaz, A., Papoutsakis, K., Antona, M. and Argyros, A.
A Robot-based Application for Physical Exercise Training.
In *Proceedings of the International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2016)*, pages 45-52
ISBN: 978-989-758-180-9

2 RELATED WORK

In recent years, researchers have explored ways to exploit the technological advances of gesture recognition and motion tracking systems in the field of rehabilitative and physical therapy, with very positive results. Virtual reality gaming using commercially available systems, such as Microsoft's Kinect and Nintendo's Wii platform, has been shown to improve measures of upper limb function in stroke survivors and balance in elderly people in community and hospital settings (Davies et al., 2014). A quick literature review on the subject of Kinect and physiotherapy will produce many results of such research projects which aim to develop applications that can provide real time virtual assistance to patients performing specific exercises, while also helping therapists to track and monitor the progress of their patients even from a remote location (Saposnik & Levin, 2011; Chang et al., 2011; Wei-Min et al., 2013). The subject of robotic assistants that facilitate physical training is addressed in (Fasola & Mataric, 2012; Werner et al., 2013).

Systems in these studies as well as commercial applications use a similar setting; the user stands in front of a large screen that displays two avatars, one that represents the trainer and the other the user. The motion-tracking camera reads the movements of the user and transfers them on the respective avatar. In most of these systems, the movements of the patient are displayed on the screen, against or in parallel to the movements of the trainer avatar. In that way, the user has immediate visual feedback on how accurate his or her movements are compared to the ones demonstrated by the trainer avatar. In other systems, the user has to perform specific movements in order to score points in the context of a fun video game. Using video gaming platforms in physical therapy has been shown to be an effective method that increases the engagement and motivation of patients to follow and complete their physical program (Davies et al., 2014; Unnikrishnan et al., 2013).

3 DESIGN CHALLENGES

The main objective of the fitness application is to promote and facilitate physical fitness for the users of the robot in their own home environment. To achieve this, the application provides a guided exercise program especially designed for the elderly, which can be initiated from the robot's main interface at any time. The application is also able to track the user's performance during the exercise and provide constructive feedback when needed. The application is not to be intended as a therapeutic or rehabilitation tool, but as a tool to help the users maintain or improve their motor skills and balance. The application offers exercises that are presented sequentially through a video trainer guiding the user throughout the fitness program. It uses a human tracking system in order to detect and display user's movements, through the avatar, while executing an exercise. The user's avatar is depicted right next to the virtual trainer for direct visual comparison. Feedback to the user (i.e., instructions, corrections, and general encouragement), is provided through HOBBIT in a multimodal manner, auditory and textual.

3.1 User Requirements

The user requirements for the fitness application were gathered and elaborated through brainstorming and observation sessions at a local day-care and activity center for the elderly, with the involvement of geriatric occupational therapists and HCI experts. The center offers a 30 minutes fitness program twice a week, where a group of 25 members have the opportunity to exercise the upper body while seated. The elaboration of the user requirements was based on two videotaped sessions presenting the potentials of the elderly in executing upper body exercises, on information gathered through an anonymous short questioner, as well as on valuable brain-storming sessions with the elder members and the center personnel.

The following requirements were derived:

- **Overall Experience.** The application must be pleasant and engaging.
- **Usable Interaction Paradigm.** The user interface should be designed in a way that the interaction paradigm is close to that of a real fitness class setting,
- **Continuous Flow of Interaction.** The user should not wait for long periods of time between exercises or application dialogues to avoid frustration or even fatigue.
- **Designated Exercise Area and Position.** All the exercises must be performed from a seated position. Exercising from a sitting position improves safety, since the base of support is greater than that of standing, while also increases the ability to perform motor activities with arms and hands (Greene & Roberts, 2005). The user's chair must be sturdy with no bulky legs and arms to allow unconstrained body movements and avoid tracking misreading and miscalculations. An area in front of the

user with appropriate space for the robot to stand is also needed.

- **Elderly Specific Exercise Routines.** The exercise routines must be designed for the average 70+ years of age user that is not experiencing any major or fitness forbidding health and physical problems. The exercise routines should be designed in close consultation with health professionals who have expertise for the elderly (e.g. occupational and physical therapists).
- **Instructions.** The instructions given by the trainer avatar must be precise, short, and clear.
- **GUI Visibility.** The avatars must be large enough to be clearly seen from required 2 meters distance that the user has to keep from the camera sensor of the robot.

3.2 Technical Constraints

The design of the fitness application had to overcome various challenges that derived from limitations and constraints of the specific robot-based set-up. These include:

- **Operation Distance.** HOBBIT uses a Microsoft Kinect visual sensor[2] to detect the user and track his/her movements. This means that the sensor requires the user to be within its field of view and at a certain distance in order to work properly. As a consequence, the robot needs to stand approximately 2 meters away from the user at all times during the execution of the fitness application. Figure 1 shows the recommended user/robot setup.

Figure 1: The fitness application setup. HOBBIT is standing at least two meters in front of the user.

- **Screen Size.** HOBBIT has a tablet on the front side where the application is shown. The screen's diagonal size is 12.1 inches, which is considered rather small for long distance viewing and poses significant constraints on the user interface design.
- **Limited Interaction.** The required distance between the robot and the user renders the touch screen interaction method and textual feedback inefficient, making other modes of interaction (i.e., gestures, voice commands, verbal output) essential for communication.
- **Sensing Constraints.** Despite the fact that state-of-the-art, patent pending technology is used to support human motion tracking and activity recognition (Michel & Argyros, 2014), certain body poses cannot be recognized (e.g., side views that lead to severe self-occlusions). Therefore, the exercise routine should exclude such poses to avoid human pose estimation errors.
- **System Complexity.** HOBBIT is a complex system that provides multiple functionalities which need orchestration. The fitness application is only part of this system and needs to cooperate fully with the rest of the components. In most cases (emergencies, low battery power, etc.) the application needs be able to terminate gracefully and transparently to the user.

4 SYSTEM ARCHITECTURE AND USER INTERFACE

Figure 2 shows the architecture of the fitness application, its internal components and its communication with the overall HOBBIT system. The application is divided into three modules: (i) The graphical user interface (GUI), which takes care of the visual part of the application and accepts user input when required; (ii) the sensor module, which is responsible for recognizing user movements and tracks the user skeleton so that it can be shown in the user interface; (iii) the controller module, which acts as an intermediate communication layer between the other two modules.

The controller module acts as an intermediary among the other components. It is responsible for translating and passing any messages that come from the sensor module to the GUI module and vice versa. This data includes information about which exercise is currently tracked, if a repetition was successful or not, if an error has occurred, and detected user skeleton joint coordinates. It also receives messages from HOBBIT's multimodal User Interface (MMUI) to facilitate robust integration with the rest of the user interface system.

The sensor module is the one responsible for detecting specific (correct or erroneous) user move-

[2] https://dev.windows.com/en-us/kinect

ments while he or she is doing an exercise. To this end, four specific physical exercises were designed for the elderly, including repetitive movements for lateral or/and bilateral abduction (or 90° elevation) of shoulders with upper extremities extended and lateral and bilateral elbow flexion with shoulder elevation in 90°. The exercises follow instructions and examples from the official website of the National Health Service in England[3]. This module is described in more details in section 5.

The GUI module is responsible for presenting audiovisual information to the user based on the approach described below.

Each of the exercises is demonstrated to the user through a video displayed on the touch screen (tablet) of the platform while the platform is placed towards the user at a distance of at least 2 meters. HOBBIT prompts the user to execute a specific number of repetitions for a specific exercise providing auditory guidance. During exercise execution, the GUI is split in the following 3 major areas: (i) Trainer area where the screen shows a video presenting a human trainer executing the exercise; (ii) User area where the user avatar (a visualization of current body configuration of the observed user) is shown; (iii) Feedback area displaying messages in the bottom center.

The system is able to recognize the repetitive movements of the arms corresponding to the known ongoing exercise and as a result to count them as executed repetitions. Moreover, any errors on the execution of an exercise compared to its known model are evaluated in order to provide relevant feedback to the user on the quality of his/her performance and potential correction to be applied in the next round of repetitions. For example, one of the four exercises refers to six repetitions of left arm movement from vertical to horizontal position and back to vertical. Some type of errors that can be assessed involve, usage of a wrong arm (e.g. the right arm instead of the left one) of the user or the arm should be extended closer to the horizontal position or the arm should be lowered closer toward the vertical position, etc. Auditory feedback toward the user can be provided at the time of execution or after the completion of the exercise. If the perceived repetitions were completed without any errors, the user will receive a positive message; otherwise relevant auditory feedback is provided letting the user know what kind of movement error was assessed during the repetitions.

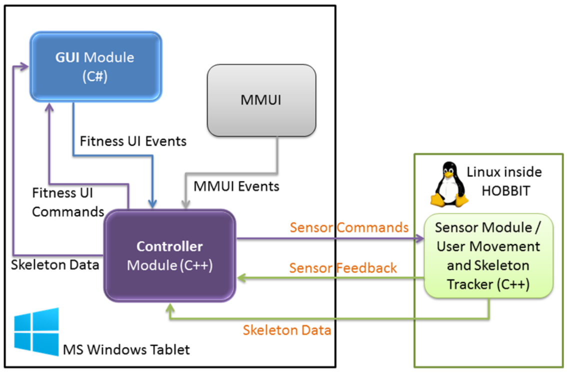

Figure 2: Fitness application communication architecture.

5 HUMAN MOTION CAPTURE AND GESTURE RECOGNITION

The developed fitness application relies on a computer vision method (Michel & Argyros, 2014) that estimates the 3D position, orientation and full articulation of the upper human body from markerless visual observations obtained by an RGBD camera. This method takes into consideration the high dimensionality and the variability of the tracked person regarding appearance, body dimensions, etc. It performs unobtrusive, markerless tracking that does not interfere with the environment, the subject and its actions. The method achieves real time performance on a conventional computer and relies on an inexpensive sensory apparatus (RGBD or depth camera). It exhibits robustness in a number of challenging conditions (illumination changes, environment clutter, camera motion, etc). In contrast to other methods, it performs automatic human detection and automatic tracking initialization and offers high tolerance with respect to variations in human body dimensions, clothing, etc.

The interaction of the user with the robot is supported by vision-based gesture recognition (Michel et al., 2014). The considered gestural vocabulary consists of five user specified hand gestures that convey important messages in the context of human-robot dialogue. Despite their small number, the recognition of these gestures exhibits considerable challenges. Aiming at natural, easy-to-memorize means of interaction, users have identified gestures consisting of both static and dynamic hand configurations that involve different scales of observation (from arms to fingers) and exhibit intrinsic ambiguities. The gestures are recognized regardless of the multifaceted variability of the human subjects performing them. Recognition

[3] http://www.nhs.uk

is performed online, in continuous video streams containing other irrelevant/non-modeled motions. All the above are achieved by analyzing information acquired by the possibly moving RGBD camera, in cluttered environments with considerable light variations. The invariance to light variations is due to the fact that gesture recognition is based solely on the depth information provided by the RGBD cameras. Several instances of RGBD cameras have been tested successfully including Kinect and Xtion (structured light emission) as well as Kinect 2 (time of flight).

5.1 Methodology

A 3D skeletal-based body model is used to estimate body pose and track body parts, providing the required representation to model any observed body configuration for each frame. A common exercise model for the fitness application has been defined that incorporates the following data fields: (i) a collection of sequential 3D body configurations that correspond to manually selected key-point configurations of the physical movements for an ideal execution of the exercise; (ii) a unique label; (iii) the expected number of iterations of the complete exercise; (iv) maximum allowed time intervals between consecutive key-point body configurations of an exercise; (v) Maximum allowed duration for a complete exercise to be executed.

The type of supported exercises and the key-point body configurations of each of them have been manually selected with the help of physiotherapists and physical trainers. The experts have indicated a valid set of appropriate, meaningful and safe body configurations to be considered as key-points for an exercise and the transition between them with respect to physical conditions of elderly. Moreover, reasonable maximum allowed time intervals between consecutive key-point body configurations have also been set by the experts.

An example of a simple fitness exercise, that is extending the arms from a seated idle body pose, is demonstrated in figure 3, based on its key-point body configurations. A pose-similarity function has been defined to evaluate the proximity between two body configurations, given the 3D positions and angles of the tracked body joints with respect to the detected body center, extracted by the human detection and tracking module. The similarity score is defined based on a weighted sum between the sum of absolute 3D Euclidean distances and the sum of absolute differences of angles between corresponding joints given that the two body centers are aligned. It is maximized in case the 3D positions and angles of the tracked joints coincide with the corresponding joints of the exercise model. Moreover, the absolute differences between corresponding body joints is stored for subsequent use. The defined human posture similarity metric is invariant to different user body sizes and dimensions as joint positions are normalized with respect to the body center and the relevant body dimensions.

The final step is to assess the similarity between two activities in terms of body configurations provided by the exercise model to be executed, given as input by the GUI module of the application, and the time series of body configurations computed by the vision-based user observation module. To this end, an activity similarity function has been defined. It calculates a confidence value reflecting which of the observed body configurations were highly similar to

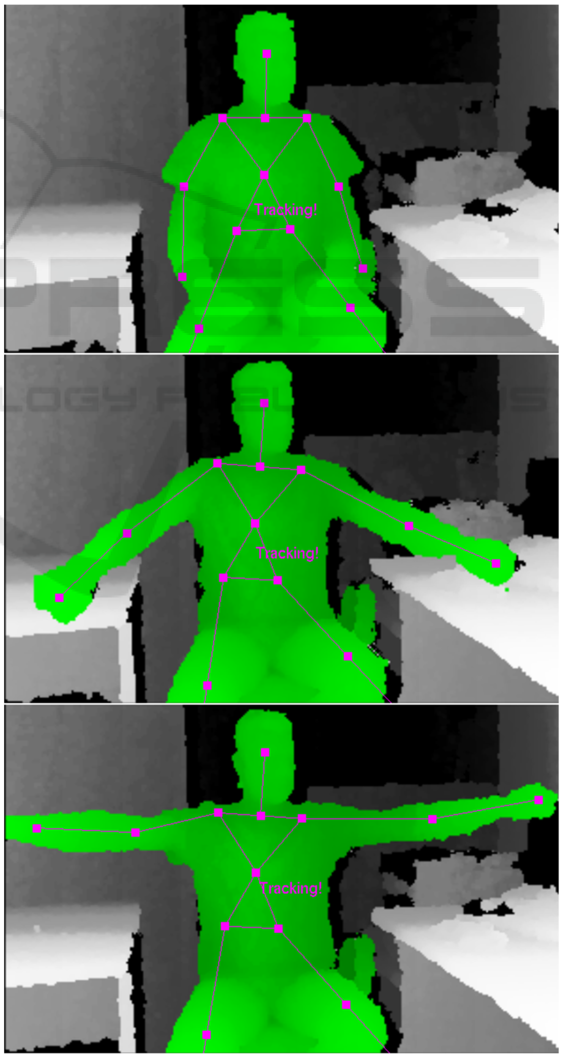

Figure 3: An example fitness exercise is demonstrated based on its three key-point body configurations.

which of the exercise key-point body configurations, how similar they were based on the pose-similarity score and if the temporal order of those matches follows the requirements set by the exercise model. The time intervals between matched body configurations are also calculated and taken into account.

The activity similarity score for an exercise and the time intervals between matched key-point configurations constitute the output of the developed module. Moreover, statistics on the time series of absolute residuals between corresponding skeletal joints in order to identify those joints that were not matched properly to the positions of the model are calculated. Therefore, it is assumed that those joints are mainly responsible for the possible failure to perform appropriately the intended exercise.

The output information facilitates the UI controller module of the application to select and provide appropriate instructions and prompts to the user as guidance or feedback in order to repeat an exercise and learn it successfully.

6 ITERATIVE DEVELOPMENT AND USER TESTING

The application went through various iterative development phases. The first priority was to establish robust communication architecture as described in section 4. After that, gradual improvements were made to the user interface and the controller module to accommodate user comments, usability concerns, sensor data changes and bug fixes.

A preliminary version of the fitness application has been tested both in the lab and in a user's home environment (see figure 4).

The first lab tests were performed by interaction and user interface experts and an older adult user. The users were instructed to activate the fitness option from the robot's menu and follow the instructions. The program included three exercises where the user was prompted to raise first the left arm to the level of the shoulder, then the right arm and finally both arms.

Figure 4: User testing the first application prototype.

From the feedback received, a number of changes were deemed necessary. First, a new trainer video was recorded, shortening the introduction and removing the lengthy trainer demonstration of each exercise to reduce distractions and ensure that the user will start training immediately and will not be confused about when to start. Second, exercise feedback was redesigned to be presented immediately after the user has done something that is detected as either wrong or deviates from the correct exercise form. This way, the user will have instant feedback on errors that he/she is doing during the exercise. Third, the user avatar was made more visible and as a mirror to the user's movements. Trainer instructions as to which arm is going to be used during each exercise, were added in order to eliminate user confusion. Fourth, two more arm exercises and two more exercise programs were added. The three exercises of the first prototype application were included in the 'easy' program, while two more programs, 'medium' and 'hard' were added, using combinations of the available exercises. The three new programs had different numbers of repetitions to allow more variations and avoid repetitive sessions and additional resting intervals to prevent user fatigue. Finally, the repetition counter in the user interface was found to be too small, thus it was redesigned to be larger and more visible.

This new version of the fitness application was tested again by lab personnel and by two elderly users in their home environment. This improved version was better accepted by the elderly users. They gave positive comments about the overall application performance and layout. One of the complaints was that the trainer feedback was unnecessarily repetitive and frequent. To improve the user experience, the frequency and content of this feedback were adjusted accordingly. Some examples of the fitness application integrated into HOBBIT's MMUI are shown in figure 5.

Finally, the fitness application has been tested in the context of the overall home-based trials of the HOBBIT robot, involving a total of 18 users of age 75+, with various degrees of vision, hearing and mobility impairments, in 3 European countries (Austria, Sweden and Greece). Each user had a prototype HOBBIT robot at home for a period of three weeks and could freely use all functionality of the robot, including the fitness application. Results were obtained through a mix of quantitative and qualitative methods, cultural probes and logging. Regarding the fitness function, the results were mainly collected through interviews with the users. 17 users said that they had tried the fitness function. The study revealed diversity with regard to how the fitness feature was perceived, with the Greek users being more satisfied than the users in Austria and Sweden. However, the results confirm that fitness was perceived by the users in all countries as an important feature, pointing to the need for wider selection of fitness exercises for a robot for the elderly. One participant who did not use the fitness function had severe visual impairment and was unable to read text on the icons or see what the fitness instructor did when the robot was placed 2m away from her. This points to the fact that the elderly have very diverse needs and the interface and functionality of an assistive robot needs to be adapted to the individual user and his or her needs. Finally, according to their comments, the users enjoyed the fitness function, and some mentioned that their level of exercise increased when they had HOBBIT due to the fitness program.

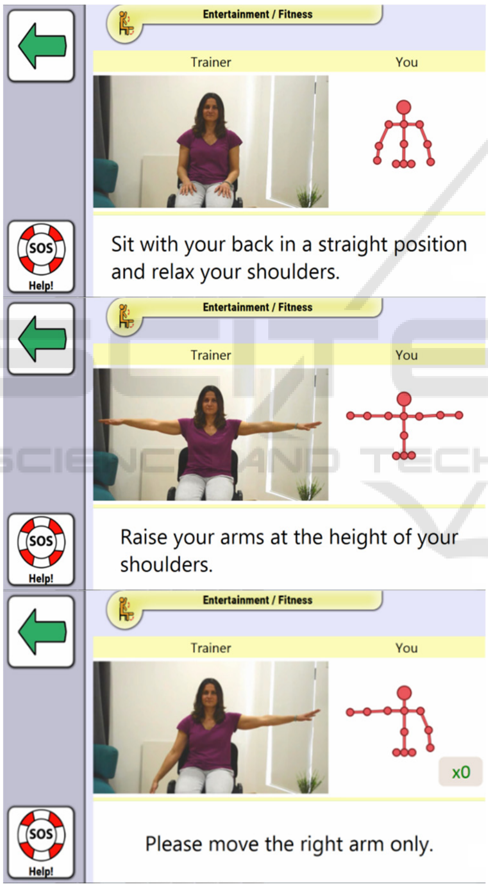

Figure 5: Final version of the application.

7 CONCLUSION AND FUTURE WORK

This paper has presented a physical exercise application meant to be used by the elderly in the context of a mutual care robot developed by the HOBBIT project. The paper has presented the design process, the user requirements, the technical constraints, the human motion capture and gesture recognition methods and the development and testing process of the fitness application. The need to develop this application was originated from the user requirements of the HOBBIT project. One of the goals of the project was to prevent falls and apart from some of the robot's functionalities implemented for this purpose, a fitness application was suggested as a means for fall prevention.

Future improvements include user interface adaptation features to increase accessibility and usability, but also motivation and confidence. The following types of adaptation are being considered: (i) suggest the appropriate level of exercise based on past activity; (ii) adaptive motivational messages based on tracked performance; (iii) adaptive feedback based on success rate when performing an exercise; (iv) personalized information at the end of each exercise section; (v) Adaptable pause behavior including pause and automatically resume on external intervention (e.g., phone call) and self-termination on long inactivity; (vi) Identify inactivity and motivate participation (during exercises).

In addition to being integrated in the HOBBIT platform, the fitness application can potentially also be used independently as a standalone application in older adult care facilities or in homes. The available exercises focus on upper-body tracking of a seated user. In an improved future version, a wider range of exercises can be programmed, including leg exercises for both seated and standing users. The user interface can also be redesigned and improved since the robot's limitations (e.g. small screen size, inexpensive sensors) would not apply in this case.

ACKNOWLEDGEMENTS

Part of this work has been conducted in the context of the Project ICT-HOBBIT "HOBBIT The Mutual Care Robot", funded by the European Commission under the 7th Framework Programme (Grant Agreement 288146). The authors would like to thank the TALOS Community Care & Active Ageing Center of Heraklion Municipality for their significant contribution to the selection of the exercise program.

REFERENCES

Buchner, D.M., Cress, M.E., de Lateur, B.J., Esselman, P.C., Margherita, A.J., Price, R., Wagner, E.H.: The Effect of Strength and Endurance Training on Gait, Balance, Fall Risk, and Health Services Use in Community-living Older Adults. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences*, 52(4), M218-24 (1997).

Chang, Y.J., Chen, S.F., Huang, J.D.: A Kinect-based System for Physical Rehabilitation: A Pilot Study for Young Adults with Motor Disabilities. *Res. Dev. Disabil* 32, 2566–2570 (2011).

Davies, T. C., Vinumon, T., Taylor, L., Parsons, J.: Let's Kinect to Increase Balance and Coordination of Older People: Pilot Testing of a Balloon Catching Game. *International Journal of Virtual Worlds and Human-Computer Interaction*. Vol. 2, 37-46 (2014).

Fasola, J., Mataric, M. J.: Using Socially Assistive Human–robot Interaction to Motivate Physical Exercise for Older Adults. *Proceedings of the IEEE* 100(8), 2512-2526 (2012).

Greene, D.P., Roberts, S.L., 2005. *Kinesiology: movement in the context of activity*, Elsevier Mosby. St. Louis, 2nd edition.

Heyn, P., Abreu, B. C., Ottenbacher, K. J.: The Effects of Exercise Training on Elderly Persons with Cognitive Impairment and Dementia: A Meta-analysis. *Archives of physical medicine and rehabilitation* 85(10), 1694-1704 (2004).

Kaminski, T.R., Bock, C., Gentile, A.M.: The coordination between trunk and arm motion during pointing movements. *Experimental Brain Research.* 106(3), 457-466 (1995),

Laurin D., Verreault R., Lindsay J., MacPherson K., Rockwood K.: Physical activity and risk of cognitive impairment and dementia in elderly persons. *Arch Neurol* 58, 498-504 (2001).

Michel, D., Argyros, A.: Apparatuses, methods and systems for recovering a 3-dimensional skeletal model of the human body, *U.S. provisional patent application No. 62/053,667*, 22/09/2014.

Michel, D., Papoutsakis, K., Argyros, A.: Gesture Recognition Supporting the Interaction of Humans with Socially Assistive Robots. *Advances in Visual Computing - 10th International Symposium*, ISVC 2014, 793-804, Springer (2014).

Saposnik, G., Levin M.: Virtual Reality in Stroke Rehabilitation: A Meta-analysis and Implications for Clinicians. *Stroke* 42, 1380–1386 (2011).

Suri P., Kiely D.K., Leveille S.G., Frontera W.R., Bean J.F.: Trunk muscle attributes are associated with balance and mobility in older adults: a pilot study. *PM R.* 1, 916–924 (2009).

Unnikrishnan, R., Moawad, K., Bhavani, R. R.: A physiotherapy toolkit using video games and motion tracking technologies. *In: Global Humanitarian Technology Conference: South Asia Satellite* (GHTC-SAS), pp. 90-95, IEEE (2013).

Wei-Min, H., Chih-Chen, C., Shih-Chuan W., Yu-Luen C., Yuh-Shyan H., Jin-Shin L.: Combination of the Kinect with Virtual Reality in Balance Training for the Elderly. *Engineering* 5, 171-175 (2013).

Werner, F., Krainer, D., Oberzaucher, J., Werner, K.: Evaluation of the Acceptance of a Social Assistive Robot for Physical Training Support Together with Older Users and Domain Experts. *Assistive Technology: From Research to Practice*, band 33, 137-142, IOS Press (2013).