

Received March 6, 2020, accepted April 17, 2020, date of publication May 21, 2020

Development of a Biomechatronic Device for Motion Analysis Through a RGB-D Camera

By F. Pristerà^{1,3}, A. Gallo^{2,3}, S. Fregola^{2,3}, A. Merola¹

¹ Department of experimental and clinical medicine, Università Magna Graecia di Catanzaro

² EthosLab S.r.l

³ IanusLab.

ABSTRACT

This work investigates the validity and reliability of a novel biomechatronic device providing an interactive environment in Augmented Reality (AR) for neuromotor rehabilitation. An RGB-depth camera and telemonitoring/remote alert module are the main components of the device, together with a PC-based interface. The interactive environment, which implements some optimized algorithms of body motion capture and novel methodologies for human body motion analysis, enables neuromotor rehabilitation treatments that are adaptable to the performance and individual characteristics of the patient. The RGB-Depth camera module is implemented through Microsoft Kinect, ORBBEC ZED2K devices; the telemonitoring module for teleassistance and therapy supervision is implemented as a cloud service.

Within the module of body motion tracking, the abduction and adduction movements of the limbs of the full-body structure are tracked and the joints angles are measured in real-time; the most distinctive feature of the tracking module is the control of the trunk and shoulder posture during the exercises performed by the patient. Indeed, the device recognizes an incorrect position of the patient's body that could affect the objective of the exercise being performed. The recognition of an incorrect exercise is associated with the generation of an alert to both the patient and the physician to maximize the effectiveness of the treatment based on the user's potential and to increase the chances of getting better biofeedback.

The experimental tests, which have been carried out by reproducing several neuromotor exercises within the interactive AR environment, show that the feature recognition and extraction, both of joints and segments of the musculoskeletal structure and wrong postures of the patient can achieve good performance in several experimental conditions.

The developed device is a valid tool for patients affected by chronic disability, but it could be extended to neurodegenerative diseases in the early stages. Thanks to the enhanced interactivity in augmented reality (AR), the patient can overcome some difficulties during the interaction with the most common IT tools and technologies; also she/he can perform rehabilitation at home. The physician can also check the therapeutic results while customizing the care pathway in real-time. The enhanced interactivity, provided by the device during rehabilitation sessions, increases the patient's motivation and the continuity of care, as well as supporting low-cost remote assistance and telemedicine which optimizes therapy costs.

The key points of the developed devices are:

1. Making rehabilitation motivating the patient to become an active "player."
2. Optimization of therapy effectiveness and costs.
3. The possibility of low-cost remote assistance and telemedicine.

Keywords – *Body Motion analysis, Smart rehabilitation, Home rehabilitation, Biomechatronic device.*

INTRODUCTION

Gesture recognition refers to the recognition of significant expressions of a motion made by a person using hands, arms, head, or other parts of the body.¹ This gesture recognition provides a wide range of applications such as:

- Development of aids for the hearing impaired.
- Support for children interacting with computers.
- Monitoring of emotional states or stress levels of patients.
- Navigation and/or interaction in virtual environments.
- Communication in video conferencing.
- Support for patients with specific physical interaction difficulties with machines and computers.

In the last few years, to make more natural and intuitive the environmental interactions with computers new research topics are exploring the direct use of hand gestures, without the use of mouse or joystick, to communicate with machines. The use of controllable interfaces through hand gestures can provide:

- a more natural interaction with the machine since the gestures are a natural form of communication and easy learning;
- a more powerful and effective interaction mediated by the device that acquires both hand position and the trajectories of the extremities of the upper limbs. A single gesture can be acquired by the interface to identify both a target object and the action to perform on it; and
- direct interaction from a cognitive point of view where the hand becomes the input device, without needing to intermediate transducers.

Some studies on the use of gestures for human communication have detected that 70– 80% of verbal messages during a dialog may be expressed exclusively through gestures that involve all parts of the body.²

Systems that exploit interactions mediated by gesture-recognition technologies can also provide a valuable tool for people with limited motor skills by allowing an efficient human-machine interaction based on a limited set of gestures or body movements. AR could also return to prominence by becoming a promising form of investment in military, entertainment and medical industries.

Especially in clinical rehabilitation, AR can improve the experience of patients by increasing the effectiveness of treatment.³

The main feature of AR systems is the ability to adapt the experience of a patient to his real physical ability.⁴ Furthermore, the arrangement of the joints and the measured value of joint angles must be taken into account. This allows the patient to objectively assess the effectiveness of treatment with the possibility of increasing the biofeedback.⁵ The real-time monitoring and measurement of the user's performance can provide biofeedback and, consequently, also aid in the evaluation of improvements or deterioration of the patient's performance. The evaluation of the performance can be achieved based on some performance indices and clinical protocols see.⁴

There is evidence that most patients can benefit from virtual reality rehabilitation. This includes patients who have had strokes,⁵ patients who need to recover limb motor skills in general,⁶ patients who need to perform neurorehabilitation in the early stages of recovery,⁷ as well as the elderly, children, and anyone who needs to work on posture or balance.⁴

This study aims to develop and test a rehabilitation device that motivates the patient to become a "player," by optimizing both the effectiveness and costs and provides the possibility of implementing, in a clinical and/or domestic context, low-cost and remote assistance and telemedicine services.

The current study is part of a more complex project that takes into account the following main steps:

Step-1: Identification and preliminary testing of different commercial devices using the RGB-D camera for rehabilitation purposes and analyzing different rehabilitation scenarios to conceptually represent the entire rehabilitation process.

Step-2: Building the first prototype using commercial hardware and implementation of dedicated software for image acquisition and processing.

Step-3: Testing of the prototype and the software in a simulated but real context. At this point, all the components of the device are globally tested and compared to standard clinical practice (process and tools).

Step-4: The rehabilitation protocol implemented is extended into the prototype.

Step-5: The system moved to the experimental stage and was tested on large scale involving different clinical partners to produce data for assessment of the device performance, not just on hardware but considers the whole process of rehabilitation mediated by the biomechatronic device.

Currently, we are working on the Step-3 of the project and this paper describes the experimental results obtained on the performance during the interaction of the patient-biomechatronic device within the AR environment.

MATERIALS AND METHODS

Body motion capture and motion analysis in an environment for interaction in AR is based on the use of a RGB-Depth camera (Microsoft Kinect) and video output devices (PC monitor or projector); the interaction environment is conceived to support adaptive and customized neuromotor rehabilitation during some exercises that promote the interaction between the patient and the device.

The system, as shown in Figure 1, can be divided into 3 distinct macrophases: (1) A phase of pre-processing carried out on each captured frame, which allows to segment one or more human figures; (2) A conversion phase that allows converting the obtained image into a segmented model, functional to the next step and achieved at low computational burden required for the extraction of the points in order to track the different parts of the body (arms, legs, and head); (3) A post-processing phase that allows the extraction of the movement.

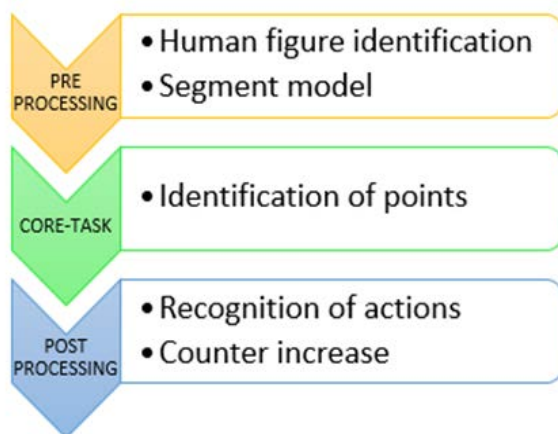


FIGURE 1. Main phases of the system.

The flow diagram shown in Figure 2 summarizes the steps point by point. Steps 1, 2, 3, and 4 are made by recalling the Microsoft's Kinect for Windows SDK functions, while the following steps (5, 6, and 7) have been developed specifically for the functions made available in Processing, a real open-source programming language that has enabled the acquisition and the elaboration of the data stream from Microsoft Kinect and other compatible devices.

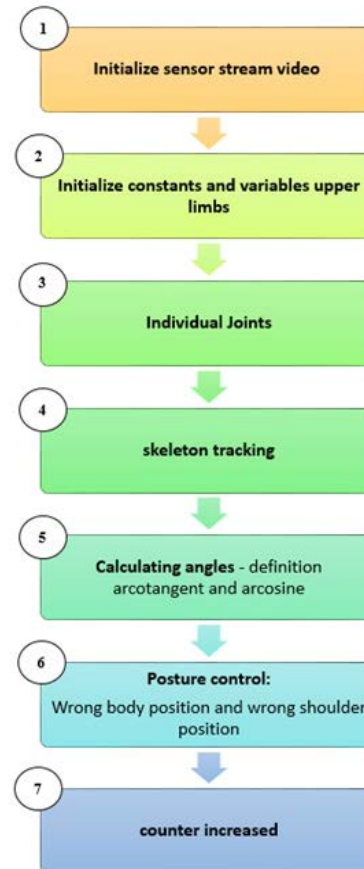


FIGURE 2. Flow diagram of the algorithm used to process the patient data during the rehabilitation session.

The presentation and the discussion of the results obtained for the rehabilitation of the upper limb are discussed in the next sections. The main features of the device are (1) tracking of the body, (2) calculation of detected angles; (3) posture control, and (4) performance acquisition.

1. The basic algorithm can able to identify the body joints and to track their position. The tracked movements of abduction and adduction of the upper limbs are visualized in real-time and in AR and the visual support to the patient improves the execution of movements and the patient's performance.
2. The calculation of the joint angle detected on the tracked limbs during in the different opening and closing positions and the visualization of the values of the angular displacements are carried out in real-time. In this step, the variables and constants are initialized for the right and left upper limbs.
3. Posture control during the exercise execution is determined by 2 methods "Shoulder Check" and "Body Check."
4. The device starts to acquire the patient's performance and then checks the correct posture at a constant acquisition rate during the rehabilitation session. Finally, the output is stored and transferred to the cloud service.

To develop our rehabilitation prototype, we used a Microsoft Kinect. It consists of an RGB camera with a resolution of 640×480 pixels to 1280×1080 at 30 Hz that can be increased at the expense of a drop in frame-rate. The same device is equipped with a depth camera consisting of an infrared projector and a monochrome CMOS camera with a resolution of 320×240 pixels. Finally, an array of 4 microphones for listening to voice commands is integrated into the Microsoft Kinect. For both cameras, the viewing angle is 57.5 degrees in horizontal line and 43.5 degrees vertically, with the possibility of extending the last one by 54 degrees thanks to the inclination platform that is equipped with a motor that rotates the sensor to automatically center the user.

Other features are 3 optical devices for visual recognition of the moving body, 2 video cameras and an additional infrared sensor, and a Kionix KXSD9 three-axis accelerometer.⁸

Each patient is analyzed separately. This choice allows the focus to be on the blob of each patient that can be extracted from the background. In the developed model, the main steps of the proposed system have been identified as solutions to the following problems listed below.

- What technique is adopted for segmenting the human figure?
- Which parts of the body are recognized and the related movements to be tracked, with low computation burden, by optimizing the precision/performance ratio of the recognition and tracking algorithm implemented into the device?

Once the above problems are solved, the selection and the development of the algorithms for detection and tracking of the target movements of the body have been performed.

The identification of the points of the body segments as target features, which are important to describe an action and to track body motion, takes into account all the points of the image and provides an estimate of all points in the form of a "line" according to the following 5 variables below.

1. The expected position of the patient (for example standup).
2. What point is detected and its anatomical name.
3. The environmental characteristics of the scene such as the illumination level of the room.
4. The patient's position is centered into the scene.
5. The Cartesian coordinates of the joints of the tracked musculoskeletal structure.

For the study of movement, the approach used is based on an algorithm that calculates the opening angle of the upper limb. After the identification of the skeletal segments and joints, the next step is to calculate the joint angles of the musculoskeletal structure from the Cartesian coordinates extracted from the segmentation image data.

After presenting the principles functional features of the rehabilitation system and before evaluating the motion tracking performance, the approach taken for implementing the overall system, made by the user interface and body motion analysis module, is presented here.

Since the software is intended for a target group of patients with motor problems but also for a target group that is halfway between rehabilitation and neurodegenerative diseases, the model developed was designed to be as intuitive and easily manageable as possible. The

font size of the text was chosen according to the reading from a certain distance and red was the color chosen to make it immediately visible to the patient.

First of all, we note the division of the rehabilitation interface into 2 areas regarding the exercises to be performed with the right arm (top right), with the left arm (top left). Also, for each limb, the visual output of the detected angle is given.

In the middle of the window, the mirror image of the patient detecting depth is displayed on the screen in real-time. The mirror image allows the patient to better coordinate body movement and also to identify the target joints and segments of the skeletal structure on which the therapy is focused. Once the recognition by the Kinect has been performed, a sequence of segments and a series of ellipses in a yellow, red, green, or blue color is used to highlight the joint junctions.

The procedure that allows us to understand where the joints inside the human skeleton are located is called pose recognition in parts, and it is realized starting from the depth image. The approach used refers to modern and robust techniques in object recognition based on the principle of subdivision of objects into parts.

The Kinect obtains 3-D information from the analyzed scene by creating a depth map within it. This map is normally obtained through a stereovision system but the Kinect is not such a system as it is equipped only with a color camera, a depth camera, and an infrared emitter.⁹⁻¹⁰ The solution adopted was that the infrared emitter projects a large number of light spots into the environment whose distribution, at first sight, seems random. The emitted pattern is visible by turning off the lights and framing the environment with a digital camera.

The optical sensors contribute by providing the PrimeSense PS1080-A2 chip with the necessary data to create an image containing depth information related to the observed scene. This image also contains a certain amount of information related to the distortion of the spots with respect to their ideal position. In this way, the Kinect determines the distance of the objects in the scene and their conformation.

After recording a depth image of the observed scene, the next step of the process involves the software execution of the tracking algorithm that identifies the number, the position, and the skeletal joints of the human skeletal structure that are to be tracked.

Microsoft's tracking algorithm¹¹ is the result of 500,000 samples of recorded data concerning different human behaviors (dancing, moving, greeting, etc.). The tracking data is processed in real-time and provided to the computer where the Microsoft Kinect is interfaced. A result of this tracking is the skeleton data will be available as a Skeleton object, obtained by calling the `getSkeleton()` function. The position and orientation of each articulation are stored into the SkeletonJoint object, which can also be obtained by using the `Skeleton.getJoint()` function. To obtain information about the various joints, such as position and speed, to measuring user attention, or to draw the skeletal model (joints), is it possible to use the function `JointType joint = user.getSkeleton().getJoint(JointType.(anatomical part of interest))` which will return an object of JointType type giving information about the target joint.

This provides a general method for the representation of open kinematic chains, and the attachment of reference systems to the joints to determine their characteristic parameters as shown in Figure 3.

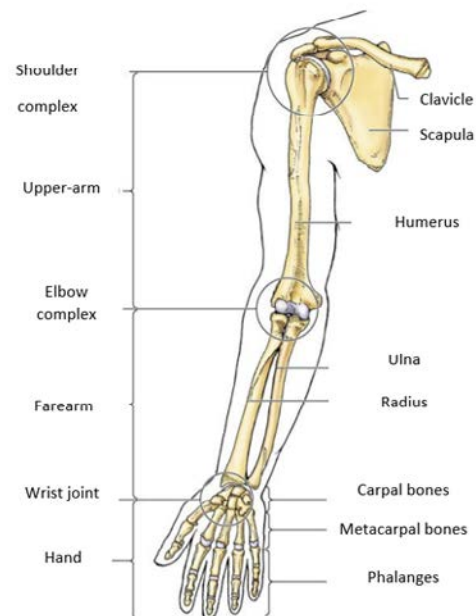


FIGURE 3. Arm joints used for the kinematic model.

Once defined the reference systems and assigns constants representing the lengths of the various links:

- A) the length of the shoulder
- B) the length of the arm
- C) the length of the forearm

Detection of the Position of the Arm Joints

While previous problems are more theoretical, the first practical problem concerns the angles measuring the arm joints and then the data output for the display device and remote monitoring. To solve this problem, we reference the joints J1 and J2 as shown in Figure 4.

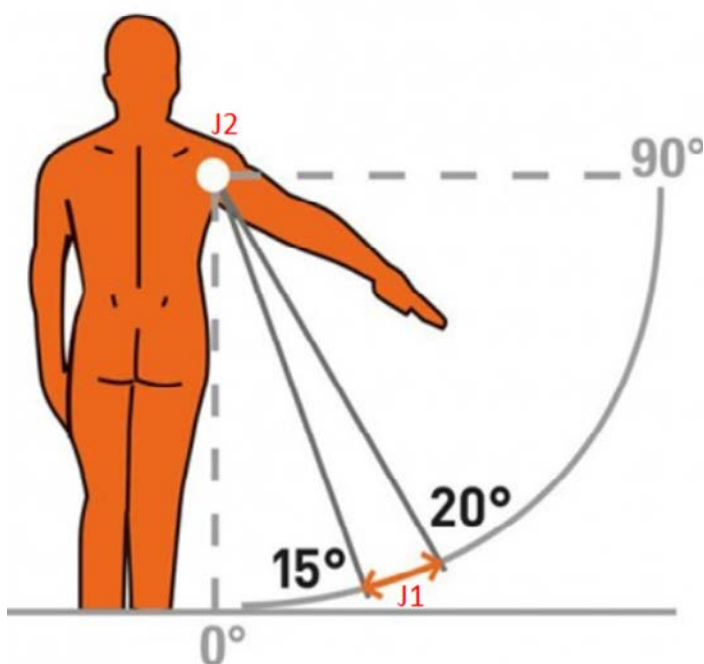


FIGURE 4. Variation of the angle in the abduction-adduction movements.

To calculate the angles of the arm joints, the position in the Cartesian space of the joint are taken from the image data stream. We consider 3 basic elements of the arm (shoulder, elbow, and wrist) to obtain the tracked angles.

The calculation of these 2 angles of the shoulder and elbow is based on trigonometric transformations (atan and asin functions) from the position in the Cartesian space of the joints identified by the tracking module.

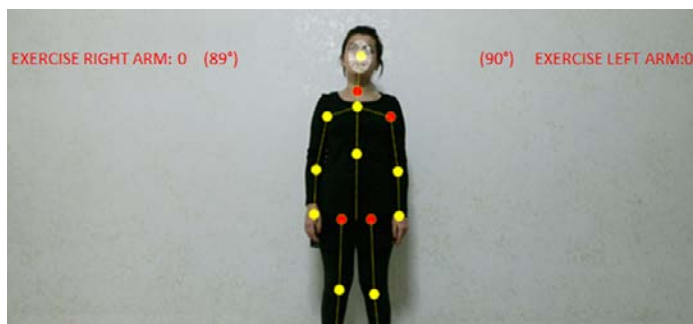


FIGURE 5. The image shows the initial values of the patient.

To calculate the angles of the various arm joints, a count_step method has been developed that uses the Kinect libraries to detect the position in Cartesian space of the 3 fundamental elements of the arm (shoulder, elbow, and wrist) and then algebraically obtain the angles.

From the trigonometric projection of Figure 6, the angles are derived from the position in Cartesian space of the points identified by Microsoft Kinect device.

```

Float m1=0;
Float m2=0;
Float P1X = joints[jointType1].getX();
Float P2X = joints[jointType2].getX();
Float P1Y = joints[jointType1].getY();
Float P2Y = joints[jointType2].getY();
m1 = (P1X-P2X);
m2 = (P1Y-P2Y);
Int A1= (int) Math.abs(((Math.atan(m2/m1))*100 / Math.PI));
    
```

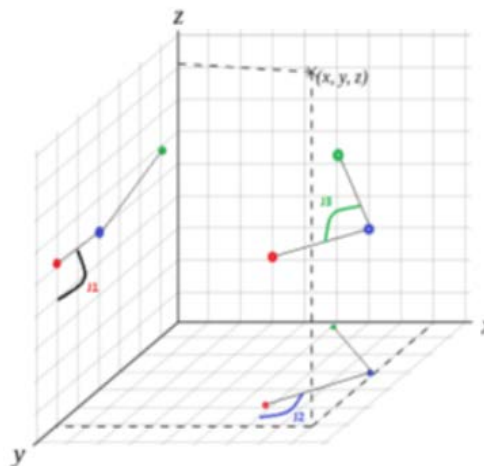


FIGURE 6. Projection and angle calculation.

The algorithm involves the calculation of the arguments of the atan function in terms of the increments m_1 and m_2 along x-axis and y-axis respectively. The values of the Cartesian coordinates of the 2 joints on the same musculoskeletal segments are used to calculate the angle. The results show the device can measure a joint angle with a margin of error of $\pm 1^\circ$. The margin error has been estimated by comparison with a set of repeated measurements performed on the shoulder through a set of protractors commonly used by physiotherapists.

The experimental tests have shown that the system is also capable of checking the patient's correct posture. An alert is generated if the patient's shoulders or body are in an incorrect position, by taking also into account a threshold interval such as around a reference angle of 0° achieved by the shoulder during the horizontal position of the upper arm. The device displays a warning on the AR interface by giving the patient information about the type and side of the wrong position. The same information is transmitted by email to the therapist at the end of the session by the telemonitoring module through a report collecting the number and types of the patient's errors. If the subject's position is incorrect, performance will be paused until the subject resumes the correct posture to perform the exercise. Examples of wrong positions detected are shown in Figure 7 and Figure 8.



FIGURE 7. Wrong position detection for right shoulder.



FIGURE 8. Wrong position detection for body.

RESULTS

The device performance has been tested in:

- a controlled environment with only people who want to interact with the system present in the field of view of the webcam;
- a heterogeneous background for each test performed where the background has heterogeneous characteristics (shades, shadows); and
- short distances where the distance between the patient and the location of the webcam (and screen) does not exceed 3 meters.

During the tests performed during rehabilitation sessions on the upper limbs, the points that identify the joint junctions are almost always detected correctly, obtaining good accuracy. The device uses a detection algorithm of a wrong posture during rehabilitation sessions. The algorithm has taken into account 2 reference points for the body and the shoulder, and calculate a gradient. If the slope is greater than a threshold, the device does not consider the wrong exercise. Table 1 shows the values measured in a patient session.

TABLE 1. Values Measured in a Patient Session

	Starting angle	End angle
Right arm	76°	2°
Left arm	80°	5°

ABDUCTION-ADDUCTION OF THE RIGHT AND LEFT ARMS

The movement starts with the arm in the rest position and the counter is increased only when the arm is at 90° relevant to the bust and returns to the initial position. If the patient performs the movement "by half," meaning it reaches an opening lower than the established ones, the counter is not increased.

The actual repetition count only occurs if the patient completes the movement such as if he starts from a rest position, performs abduction, until he reaches 90° position, and then it returns to the rest position through the adduction of the arm (Figure 9). Since the system has

tolerance thresholds of $\pm 1\%$, the count is considered valid if the measured value falls within this threshold.

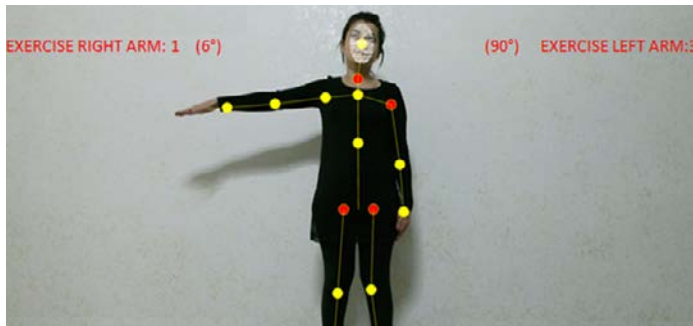


FIGURE 9. Counter increase. Movement begins with the arm in the rest position, the counter is increased only when the arm is J1/J2.

SHOULDER REHABILITATION RESULTS

The proposed isotonic exercises aim to recover joint mobility and optimize joint function. The goal of the execution of the exercise, from 0° to 60° , is to restore a normal joint function with an opening angle up to 180° .

After this first evaluation, related to the movements of the trunk (and therefore the pelvis), the patient's hands were placed along the axis of the trunk. The patient was then asked to tilt their torso. The test was successful and just as expected, when the tilting exceeded the permitted tolerance limit, the patient was notified of the incorrect position of the trunk.

The results obtained and the sequence of exercises used in the rehabilitation path, with reference to the rehabilitation standards taken as a reference,¹² show that the values obtained compared to those measured in normal clinical practice and are more than satisfactory.

The device is also equipped with a telemonitoring module that is connected to the cloud and can send the patient's performance to the physician (or physiotherapist) in real-time. Thanks to the telemonitoring system, the physician can optimize the rehabilitation path according to the patient's performance.

DISCUSSION

From the experimental findings it possible to conclude that the exercises mediated by the AR interface can effectively support the main rehabilitation protocols both

for the recovery of mobility following trauma and for the case of surgery.

In shoulder rehabilitation protocols, in the case of surgery involving the rotator cuff, or even in the case of surgery for proximal fractures (near the shoulder) of the humerus, isotonic exercises (including abduction and adduction of the arm), must be performed standing in front of a mirror, taking care not to contract the upper beam of the trapezius muscle, that is, avoiding the elevation of the shoulder. The objective is to recover joint mobility and to optimize joint function; the exercise is performed from 0° to 60° , until the joint function is reached that allows an opening up to 180° .

It is emphasized that the movement during the exercise should be performed slowly and should not be painful and the exercises should be avoided if the joint is sore or swollen, since the rehabilitation session not only aims to strengthen the muscles but also increase the amplitude of the joint movement, while improving the precision and safety of the movement.

It is also useful to perform the exercises with the joint that not affected. As shown by the results obtained, the system can be easily integrated into standard clinical practice and, at the same time, the device easily customizable to guarantee, personalized rehabilitation-functional path.

CONCLUSIONS

In this study, the Microsoft Kinect V2 has been tested by assessing the performance of the motion tracking of a patient in rehabilitation. The experimental tests show that the tracking algorithm implemented is very robust and the system performance has been characterized over several operating conditions. The estimated accuracy during tracking is a few millimeters in most cases. The tracking algorithm follows the full-body figure of the individual as long as they remain within the field of vision of the RGB-D camera. The tracking performance is strongly affected by the following circumstances: a limb covering another one, an object placed into the scene, and a camera point of view not perpendicular to the frontal plane of the body. In these cases, the visualization of the tracked trajectories and body elements into the virtual reality scene was affected by some drift between the 3-D

visualization and the real points on the patient's musculoskeletal system. Accuracy is the main requirement needed evaluate the sensor performance as objective as possible after considering all the performance constraints involved in the rehabilitation tasks used for testing the device. In this respect, the achieved maximum accuracy of 1° is satisfactory.

At the current development stage, the estimation of system performance has been carried out by evaluating the tracking accuracy of the joint angles from 0–90°. The accuracy is measured as the angle of error between the output of the mechatronic device and the protractor measurements carried out manually on the patient during abduction–adduction exercises.

In conclusion, the strong points of the device and resulting rehabilitation model are its accessibility and usability. Following the preliminary phase of definition of the design constraints based on the needs of standard rehabilitation protocol, the obtained results are the design, testing of a device, and the assessment of its therapeutic applicability that support an intuitive interaction together with the adaptation to the specific performance and individual characteristics of the "target" user.

The device provides a valid tool for people with chronic disabilities but also for the treatment of neurodegenerative diseases, especially in the early stages. Thanks to the conceived interactive model, patients can improve their quality of life by overcoming difficulties in interacting with the most common digital tools and new technologies (information technology) that can be introduced in healthcare facilities as well as in everyday environments, to improve therapeutic performance and optimize cost.

Therefore, patients can carry out the rehabilitation exercise sessions independently but at the same time they can be supported and encouraged and, above all, the physiotherapist can "control" the results in real-time, allowing a better evaluation of the evolution of the treatment path together with greater personalization and higher frequency of treated patients over time providing the advantages of low-cost teleassistance and telemedicine in the context of "at-home rehabilitation,"

As a further development of the device, a voice recognition module will be implemented, as the voice signal can

be used to start the rehabilitation session or to interrupt and then resume the exercise later. Another important development concerns facial recognition. The ability to customize the "list" of exercises that the patient has to perform and the extension of the dataset of exercises made available to the patient would be most useful. This could expand the possibility of therapy and to make the rehabilitation model more complete and effective.

At the current stage of the development, we are testing the use of hardware platforms, like the Kinect (Azure, Orbec 3D), that support the Processing software. By exploiting the enormous progress in the field of machine learning, further tests are being addressed through the use of a normal web camera for tracking movements that at present; however, it still guarantees a suboptimal accuracy compared to those obtained in this study.

REFERENCES

1. Du W and Li H. Vision Based Gesture Recognition System with Single Camera. CAD Laboratory Institute of Computing Technology, Chinese Academy of Sciences, 100080, Beijing, China; 2000.
2. MCNEILL, D. & LEVY E. T. Speech, Gesture, and Discourse, in *Discourse Processes*, 15, (1992) 277-301.
3. Rose FD, Brooks BM, Rizzo AA. Virtual reality in brain damage rehabilitation, *Cyber Psychol Behav* 2005;8:241–262.
4. Paraskevopoulos IT, Tseklevs E, Craig C, et al. Design guidelines for developing customised serious games for Parkinson's disease rehabilitation using bespoke game sensors, *Entertain Comput* 2014;5:413–24.
5. Abate AF, Acampora G, Ricciardi S. Augmented Tour of Archaeological Site by Means of Adaptive Virtual Guide. *DMS*; 2008.
6. Steven B, Feiner K. Augmented reality: a new way of seeing. *Sci Amer* 2002.
7. Larsona EB, Feigonb M, Gagliardod P, Dvorkina AY. Virtual reality and cognitive rehabilitation: A review of current outcome research. *NeuroRehabil* 2014;34.
8. Rahman M. *Beginning Microsoft Kinect for Windows SDK 2.0: Motion and Depth Sensing for Natural User Interfaces*, APress; 2017

9. Roccetti M, Marfia G. Recognizing Intuitive Pre-defined Gestures for Cultural Specific Interactions: An Image-based Approach”, Proc. 3rd IEEE International Workshop on Digital Entertainment, Networked Virtual Environments, and Creative Technology (DENVECT’11) - 8th IEEE Communications and Networking Conference (CCNC 2011), Las Vegas (USA), IEEE Communications Society, January 2011.
10. Wren C, Azarbayejani A, Darrell T, and Pentland A. Pfunder: Realtime Tracking of the Human Body.” IEEE Trans. on Patt. Anal. And Machine Intell 1997;19.
11. Kinect. Microsoft Corporation, 2014. Accessed: 2014-10-01. Available at: <http://www.kinect.com/>.
12. Linee guida per le attività di riabilitazione. Conf. Stato-Regioni, Gazzetta Ufficiale N. 146 del 24 Giugno; 2002.