

Kinect-based rehabilitation exercises system: Therapist involved approach

Li Yao^{a,*}, Hui Xu^a and Andong Li^b

^a*Computer Science and Engineering Department, Southeast University, Nanjing 211189, China*

^b*Computer Science Department, New York University, NY, USA*

Abstract. The Kinect-based physical rehabilitation receives increasing recognition as an approach to provide convenience for the patients who need the therapy usually from the health professions. Most of the previous studies were driven from the patients' point of view. This paper proposes a system aiming to simplify the recovery instruction from therapists, increasing patients' motivation to participate in the rehabilitation exercise. Furthermore, the architecture for developing such rehabilitation system is designed by motion capture, human action recognition and standard exercises prototype with Kinect device.

Keywords: Rehabilitation, kinect, motion capture, therapist-oriented

1. Introduction

Recent studies have shown that the repetitive training using human computer interfaces and virtual reality over a long period of time is helpful for both physical and mental recovery [1,2]. This kind of training process is supervised with the continuous movement control in an objective way. Moreover, with the help of programmable physical exercises, the rehabilitation could be achieved anywhere, even at home.

Using Kinect to capture movement is a feasible way to instruct the patients to do rehabilitation exercises. As a key role in physical recovery session, but often ignored by current related research, the rehabilitation therapists devote their essential efforts to scheduling treatment, action evaluation and curative effect assessment. Since the similar systems are patient-oriented, the patients are only able to follow the rehabilitation exercises efficiently. The rehabilitation program has to be accurately studied and prepared by therapists beforehand. Therefore, the rehabilitation system should provide a tool for fast deployment therapy sessions based on the therapist background, and it depends on patients' cases.

This paper introduces a system based on motion capture with the Microsoft Kinect device. Kinect is a line of motion sensing input devices by Microsoft for Xbox video game consoles and Windows PCs. The system integrating functions of real-time display, exercises modification and action evaluation is developed for therapists to prototype rehabilitation exercise.

*Corresponding author: Li Yao, Computer Science and Engineering Department, Southeast University, Nanjing 211189, China. Tel.: 86-25-52090988; Fax: 86-25-52090981; E-mail: yao.li@seu.edu.cn.

2. Related work

Body action tracking has a wide range of applications. Professional camera-based motion capture systems, which required special markers or suits, were needed in these applications. This kind of systems has high precision without a segmentation step. Positions and poses can be derived from the sensors installed on the device. An ad hoc sensors can easily be added, thus efficiently solving occlusion problems [3]. However, their limited usability, portability and adaptability make these systems expensive and obsolete. In addition, they are not recommended in rehabilitation because their encumbrance greatly limits user movement [4,5]. Recent research has focused on marker-less camera-based systems. Such systems regard image processing as the most core part and largely depend on how the scene is captured. Applied algorithm extracts specific data from the depth image or something similar generated by the camera.

Microsoft Kinect, which contains special cameras and sensors, was initially launched as a commercial extra device on Xbox video game console. Later on, developers and researchers took advantage of it in various areas [6], e.g. head pose and facial expression tracking, hand gesture recognition, human activity recognition, and healthcare applications.

A system proposed by Danilo Avola [7] has considered implementing a tool for therapists, using a programming language called eXtensible Markup Language (XML) as the script to create rehabilitation exercise task. However, it is very difficult for therapists to program rehabilitation exercises.

3. System design

Our system is specially designed for therapists to prescribe recovery rehabilitation exercises in an easier way. Figure 1 demonstrates the fundamental modules of our system.

The device layer acts as an infrastructure management tool, which includes Kinect device and other kinds of video input sources. Drivers for Kinect are deployed between the device and higher level APIs, in order to ensure Kinect can work properly. The above device layer is APIs which could be used by developers for enhancing features. Plugins can be added into the system and are visible on the graphic user interface. The user interface provides a friendly interaction for users. The topmost level is

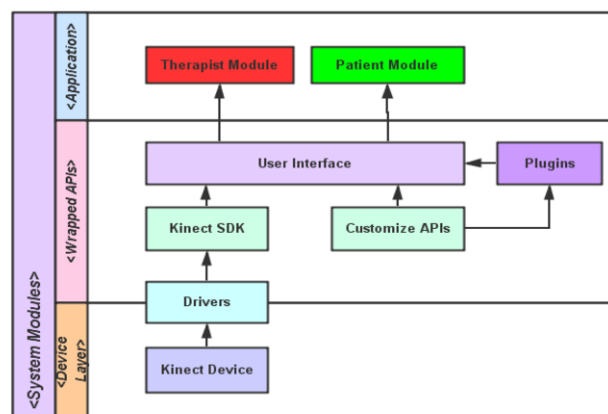


Fig. 1. Fundamental modules.

application layer, containing two modules, i.e., the patient module and the therapist module, to meet with two types of users. This paper mainly focuses on the therapist module. Therapist module has three main functions: standard rehabilitation exercise recording, action evaluation and rehabilitation curative effect assessment.

3.1. Standard rehabilitation exercise recording

In order to expedite recovery, patients need to practice during rehabilitation training sessions. In [3,4], the actions are relatively simple and proper for most patients. In most cases, the system designers create the actions by just referring to their common sense, which means that the exercises may not always be effective for every patient. In [7], the designed system, with conceived actions carried out by therapists, is too difficult to be efficient: in order to design the envisaged action, therapists should have programming skills and be literate in XML.

Our system uses Kinect's motion capture functions to obtain raw skeleton data. With the standard movement pattern performed by the actors or the therapists being recorded as video data stream, our system compares and evaluates the patients' movement by providing the feedbacks on the screen in real time. According to the feedbacks and suggestions, patients can adjust their body immediately. By parsing prototype rehabilitation exercises into data stream, workout session video files could be stored in XML files during the preprocessing period. These files are classified by different rehabilitation achievements, then packed and provided to the patients. Then, the patients can practice through the rehabilitation system.

3.2. Action evaluation

In traditional rehabilitation training, therapists need to correct patients when they are not right performed. Each therapist can only take care of one patient at the same time, which means that resource utilization has been greatly restricted. The experiments in [6,8] are carried out through virtual reality technology. The feedback screen displays real-time movements of patients to help them adjust the actions. This can reduce the intervention of therapists in the rehabilitation procedure and improve time efficiency for therapists. However, the completion status of the patients' actions can only be assessed by themselves to decide whether they need repeated practice.

Considering the importance of the therapist intervention in the rehabilitation system, the patients could have real-time feedback by the comparison of standard exercise data and processed frame stream from Kinect in our system. Intuitive evaluation of specific figures determines whether the patient needs to repeat the action. More details would be stored locally on the patients' machine and therapists can request the details anytime through the system.

3.3. Rehabilitation curative effect assessment

In traditional rehabilitation training, therapists and patients need face-to-face contact. Therapists would check the patient's recovery state by conversation and observation. The judgment made by therapists is based on the action performed by patients.

Whereas, if the system collects data of movements and uploads completion rate of action timely, therapists can get the information on the patient and know the patient's recovery state without face-to-face contact. Similarly, patients can ask for the assessment from therapists to check recovery state.

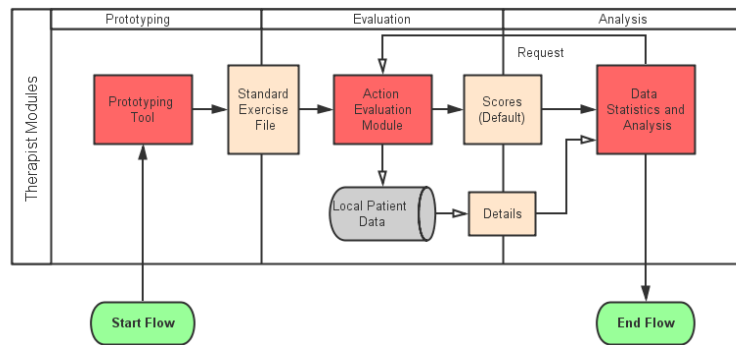


Fig. 2. Therapist module.

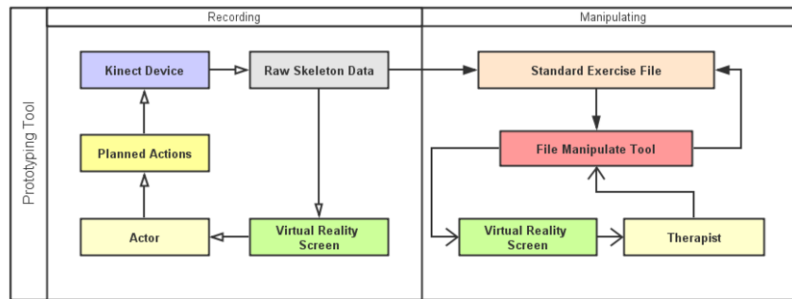


Fig. 3. Flow chart of prototyping tool.

The three functions above are respectively corresponding to three modules: exercises prototyping tool, action evaluation module, data statistics and analysis module. Figure 2 shows the relationship among these modules.

1) Prototyping tool: Therapists can use this tool to create rehabilitation exercise task by recording the standard action sequences using Kinect. Recorded action sequences can be transferred to an XML file, namely the Standard Exercise File, and vice versa. Figure 3 demonstrates the flow chart of the prototyping tool. An actor does the planned actions and Kinect device captures the poses and then generates the raw skeleton data. By real-time processing, the actor can see the actions that he is doing in the virtual reality screen. Raw data also have to be processed and converted to Standard Exercise File. When the therapist wants to manipulate the file, he/she should load it into the File Manipulate Tool. Through the virtual screen, therapists can manipulate the file in a visible way by drag-and-drop operations. All the XML files belonging to one exercise can be packed together and dispatched to specific patients.

Kinect device generates two kinds of frame data: the depth data and color data. Figure 4(a) is the depth data. According to the depth data, the system can get skeleton data using Kinect SDK. The skeleton can be described as the following joints: head, neck, right shoulder, left shoulder, right elbow, left elbow, right hand, left hand, torso, right hip, left hip, right knee, left knee, right foot, and left foot [6]. To extract these joints, the user has to stand about two meters away from the camera so that the entire body is visible. Upper body mode can be used for those individuals with a disability of standing up. Figure 4(b) demonstrates the skeleton, and all its body part has spatial coordinates in the virtual space. The XML files record the coordinates which can be reused when manipulating.

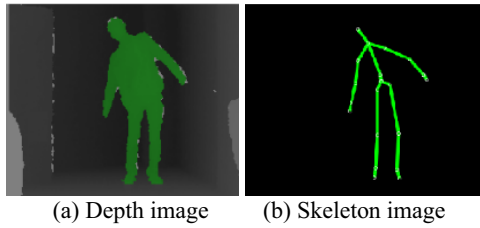


Fig. 4. Depth data and generated skeleton image.

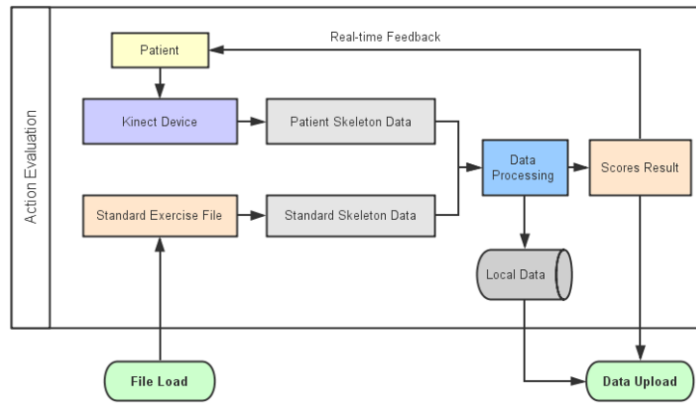


Fig. 5. Data flow of action evaluation.

2) Action evaluation module: this module scores the perfection of the patients’ exercise. The standard exercise file is originally in XML format. Figure 5 is a data flow of the action evaluation procedure. While the patient does the exercise, the system loads the standard exercise file to generate the standard skeleton data. At the same time, the Kinect device captures the poses and generates the patient’s skeleton data. In the data processing procedure, the scores are given based on the comparison of patient skeleton and standard skeleton data.

As the patients would not do the actions as the same rate as the standard skeleton frame sequences, a method is developed to make the comparison possible. Dynamic time warping (DTW) is a well-known technique to find an optimal alignment between two given (time-dependent) sequences under certain restrictions [9]. Our data are based on the time sequence but in image format which means a more suitable algorithm is expected. [10] introduces a method using cross correlation, which is well known in detecting common periodicities between two signals of interest, to obtain the coefficient of cross correlation (r_{xy}) of two time series x and y , each with N data points:

$$r_{xy}(k) = \frac{c_{xy}(k)}{\sqrt{c_{xx}(0)c_{yy}(0)}} \tag{1}$$

where

$$c_{xy}(k) = \begin{cases} \sum_{t=1}^{N-k} (x_t - \bar{x})(y_{t+k} - \bar{y}) + \sum_{t=N-k+1}^N (x_t - \bar{x})(y_{t-N+k} - \bar{y}), & k = 1, 2, \dots, N \\ \sum_{t=1}^N (x_t - \bar{x})(y_t - \bar{y}), & k = 0 \end{cases} \quad (2)$$

$$c_{xx}(0) = \sum_{t=1}^N (x_t - \bar{x})^2 \quad c_{yy}(0) = \sum_{t=1}^N (y_t - \bar{y})^2 \quad (3)$$

where k is a number indicating a time shift of one signal with respect to the other; $k = 0$ for the original two time series synchronized in time. The correlation $r_{xy}(0)$, calculated from $c_{xy}(0)$, will give an indication of pattern similarity between the two sets of data. This is just a description of the method and it has to be implemented in the system to mark the joints as the variables for measuring actions. A joint as the variable from the exercise file and the raw data in an interval predefined are extracted, and the data from exercise file are marked as x while raw data as y . Then a sequence of $r_{xy}(k)$ calculated can be obtained to make an evaluation from these r values by an average score.

3) Data statistics and analysis module: This module is based on request and response between the patients and therapists. At the patient side, many kinds of data are recorded locally and only the scores gotten in module 2 will be submitted to therapists by default. Therapists can assess the current state of patients according to the uploaded data. Many types of visual data help therapists make analysis. Therapists may return the assessment result to the application. If the therapist needs more details about the patient, he/she can send a request to the server side and the patient's machine would collect local database and give a response.

3.4. Preliminary system evaluation

The evaluation of the proposed system is directed at testing the ability of a group of therapists in prototyping rehabilitation tasks. Eight therapists were recruited for the test. All of them declared themselves to be able to use a personal computer and to manage an electronic device. None of them had programming skills.

In order to make them use the system skillfully, the system developers spent half an hour in showing how to use the system. Figure 6 shows the system interface. The result of this test is summarized in Table 1 (where: "Y" and "N" are used to indicate the therapist "has" or "does not have" the corresponding ability, respectively).

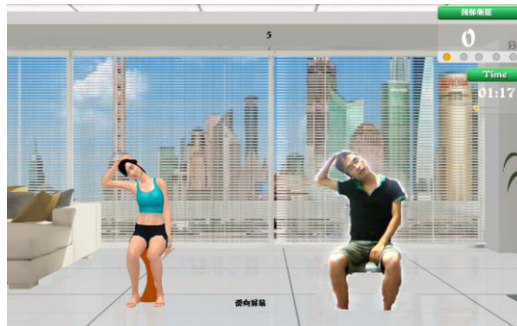


Fig. 6. System interface

Table 1
Preliminary system evaluation

Therapist #	1	2	3	4	5	6	7	8
Create rehabilitation exercises	Y	Y	Y	Y	Y	Y	Y	Y
Modify rehabilitation exercises	Y	N	N	Y	Y	Y	N	Y
Check the patients' recovery states	Y	Y	Y	Y	Y	Y	Y	Y
Overall difficulty	1	3	3	2	2	1	3	1

Therapists #1, #4, #5, #6, #8 were capable of creating a complete new rehabilitation exercise, and using data statistics and analysis tools to check the patients' recovery states, and modify the rehabilitation exercises.

Therapists #2, #3, #7 couldn't use our system to modify the exercises because they found that the drag-and-drop operations in the 3D space are not easy.

At the end of the test, therapists were asked to express a subjective overall opinion regarding the difficulty of usage of our system (in Table 1, their opinion on the overall difficulty is indicated by a scale from 1, the minimum degree of difficulty, to 5, the maximum degree of difficulty). Our system was considered not particularly difficult to use.

Although still preliminary, due to the low number of trials and the limited number of consulted therapists, these results can be considered very encouraging. Most of the therapists can create and modify the rehabilitation exercises even though the therapists involved were beginners.

4. Conclusion and future work

Our system is proposed for therapists to realize quick rehabilitation exercises task creation, accurate action evaluation as well as objective rehabilitation curative effect assessment. Our system can greatly improve the work efficiency and objectiveness of therapists' treatment.

In the future, there will still be some work for further study. For example, how to define the completion rate in the action evaluation to make it more scientific, in other words, how to judge the accuracy of patients' action. Our implementation is simply based on therapists' experience like angles or positions of head or other specific body part. More scientific method should be studied by collaboration of therapists, patient and engineers to assess the rehabilitation curative effect.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 61273300) and the Open Project Program of the State Key Lab of CAD & CG (No. A1429), Zhejiang University.

References

- [1] P.L. Weiss, D. Rand, N. Katz and R. Kizony, Video capture virtual reality as a flexible and effective rehabilitation tool, *Journal of Neuro Engineering and Rehabilitation* **12** (2004), 1–12.
- [2] M.R. Golomb, B.C. McDonald, S.J. Warden, J. Yonkman, A.J. Saykin, B. Shirley, M. Huber, B. Rabin, M.A. Baky, M.E. Nwosu, M. Barkat-Masih and G.C. Burdea, In-home virtual reality videogame telerehabilitation in adolescents with hemiplegic cerebral palsy, *Archives of Physical Medicine and Rehabilitation* **91** (2010), 1–8.

- [3] M.A.R. Ahad, J.K. Tan, H.S. Kim and S. Ishikawa, Analysis of motion self-occlusion problem due to motion overwriting for human activity recognition, *Journal of Multimedia* **5** (2010), 36–46.
- [4] C.G. Burgar, P.S. Lum, P.C. Shor and H.F.M.V. der Loos, Development of robots for rehabilitation therapy: The palo alto stanford experience, *Journal of Rehabilitation Research & Development* **37** (2000), 663–673.
- [5] L.E. Kahn, P.S. Lum, W.Z. Rymer and D.J. Reinkensmeyer, Robot-assisted movement training for the stroke-impaired arm: Does it matter what the robot does, *Journal of Rehabilitation Research & Development* **43** (2006), 619–630.
- [6] D. González-Ortega et al., A kinect-based system for cognitive rehabilitation exercises monitoring, *Compute Methods Programs Biomed.* (2013). doi: 10.1016/j.cmpb.2013.10.014
- [7] D. Avola and M. Spezialetti, Design of an efficient framework for fast prototyping of customized human-computer interfaces and virtual environments for rehabilitation **110** (2013), 490–502.
- [8] Y.-J. Chang, W.Y. Han and Y.-C. Tsai, A kinect-based upper limb rehabilitation system to assist people with cerebral palsy, *Research in Developmental Disabilities* **34** (2013), 3654–3659
- [9] M. Müller, *Dynamic Time Warping, Information Retrieval for Music and Motion*, Springer Press, Berlin, 2007, pp. 69–84.
- [10] L. Li and G.E. Caldwell, Coefficient of cross correlation and the time domain correspondence, *Journal of Electromyography and Kinesiology* **9** (1999), 385–389.