A Kinect-Based Rehabilitation Exercise Monitoring and Guidance System

Wenbing Zhao, Hai Feng, and Roanna Lun  
Department of Electrical and Computer Engineering  
Cleveland State University, Cleveland, Ohio 44115  
Email: wenbing@ieee.org

Deborah D. Espy and M. Ann Reinthal  
Department of Health Sciences  
Cleveland State University  
Cleveland, Ohio 44115

Abstract—In this paper, we describe the design and implementation of a Kinect-based system for rehabilitation exercises monitoring and guidance. We choose to use the Unity framework to implement our system because it enables us to use virtual reality techniques to demonstrate detailed movements to the patient, and to facilitate examination of the quality and quantity of the patient sessions by the clinician. The avatar-based rendering of motion also preserves the privacy of the patients, which is essential for healthcare systems. The key contribution of our research is a rule-based approach to realtime exercise quality assessment and feedback. We developed a set of basic rule elements that can be used to express the correctness rules for common rehabilitation exercises.

Keywords—Depth Sensing, Kinect, Motion Analysis, Virtual Reality, Realtime Feedback

I. INTRODUCTION

In preventive and rehabilitative healthcare, physical exercise is a powerful intervention. However, a program may require in the range of thousands of practice repetitions, and many people do not adhere to the program or perform their home exercises incorrectly, making the exercise ineffective, or even dangerous [1]. Exercise programs prescribed to address specific problems must be individually tailored by a clinician due to the presence of co-morbidities and additional impairments. The current state-of-the-art for exercise instruction and monitoring is usually limited to written instructions, exercise recording logs, and simple repetition counting devices. Unfortunately, this practice has a number of problems:

- The patient does not receive any feedback on the quality of the prescribed exercises.
- The clinician has no way of knowing whether or not the patient has carried out the prescribed exercises correctly and with the required number of repetitions.

Correct adherence to supplemental home exercise is essential for safe, effective, and efficient rehabilitation care [1]. The lack of correctness feedback on the in-home exercises is therefore a serious concern. The release of the Microsoft Kinect sensor, which is equipped with a depth camera capable of measuring 3 dimensional (3D) positions of the objects in its view, and the corresponding software libraries that enable the receiving of skeletal joint positions in realtime, have triggered tremendous interests in using the sensor to monitor in-home rehabilitation exercises because:

- The Kinect sensor can be programmed to record an exercise session and provide continuous feedback about correct exercise performance to the patient exercising at home.
- Kinect is an inexpensive device. The first generation of Kinect sensor is available commercially for around $100, which is about the cost of a single physical therapy session.

A Kinect-based system could facilitate proper performance of rehabilitation exercises at home, which would significantly minimize trips to a physical therapy center, which are costly and inconvenient for patients. Furthermore, such a system could increase patient accountability, allow the clinician to correct any errors in exercise performance, and allow program modification or advancement as needed. Indeed, several feasibility studies (including our own) have demonstrated that Kinect provides accurate measurement for exercises with good frontal view (e.g., [2]).

In this paper, we describe the design and implementation of such a Kinect-based system for rehabilitation exercises monitoring and guidance. The system demonstrates the correct way of doing an exercise via a 3D avatar on one side of the screen based on pre-recorded motion data. On the other side of the screen, another avatar is shown that reflects the actual patient movement. Furthermore, the system implements a set of correctness rules for each exercise and assesses the patient’s movement in realtime. The assessment results are incorporated in the patient avatar in the form of visual guides to help the patient perform the exercise correctly. The system also records vital data pertinent to the quality and quantity of exercises such as correct iterations as well as detailed motion data for realtime feedback and post-analysis.

This paper makes the following contributions:

- Our system is the first rule-based system for rehabilitation exercise monitoring and guidance with realtime visual feedback.
- A set of basic rule elements are developed such that they can be used to define correctness rules for common rehabilitation exercises. This enhances the adaptability of our system to accommodate new rehabilitation exercises and the adjustment of existing correctness rules when they are needed.
- Facilitated by the Unity framework, the motions are rendered via 3D avatars with frame-by-frame replay.
control and full 360-degree view. Hence, our system allows the patient to study the prescribed exercises, and the clinician to examine the patient sessions in great detail. Furthermore, our system preserves the privacy of its users while enabling the exercise session to be captured for full examination.

II. BACKGROUND

In this section, we provide more in-depth information regarding Microsoft Kinect. We also briefly introduce correctness rules for a few common rehabilitation exercises.

A. Microsoft Kinect

Microsoft Kinect was first released in 2010 as an addition to the Xbox 360 game console. It is equipped with an RGB camera, an infrared emitter, a depth sensor, and a microphone array. With the official Microsoft Software Development Kit (SDK) [3] or third party middleware toolkits, the 3D positions of skeletal joints can be obtained in streams of skeletal frames in realtime. Hence, Kinect has been used in areas far beyond console games [4].

The most interesting data supplied by Kinect (and its SDK) are the skeletal data, which can be obtained by enabling the skeletal data stream and by registering an event handler to receive and process skeletal frames. Each skeletal frame may contain data for up to 6 users, with up to 2 users fully tracked. For each fully tracked user, the 3D positions \((x, y, z)\) of up to 20 joints are reported, including head, shoulder center, left/right shoulder, left/right elbow, left/right wrist, left/right hand, spine, hip center, left/right hip, left/right knee, left/right ankle, and left/right foot.

B. Exercise Assessment

As we mentioned before, for rehabilitation, a patient must adhere to the instruction when performing the prescribed exercises. Hence, it is critically important to assess the patient’s movement in realtime while the patient is engaging in the rehabilitation exercises. There are several different ways of accomplishing this in a Kinect-based system, which we will elaborate in Section IV. In this paper, we choose to take a rule-based assessment approach because of the following benefits:

- The correctness rules for each exercise define the invariance of the exercise; hence, it is independent from the size and form of the person who performs the exercise. On the contrary, other approaches typically require scaling if the demonstrator is different from the patient, which can be complicated and difficult to implement.
- Compared with other approaches, the rule-based approach is much less computationally intensive and hence it fits well with the realtime assessment requirement.

In the rule-based approach, a set of rules are implemented and assessed in realtime for each exercise. Due to space limitation, we explain the correctness rules for two exercises, hip abduction and bowling.

For rehabilitation purposes, the hip abduction involves movement of the hip in which the abducting leg moves away from the body on the same plane as the rest of the body. Hence the primary rules for hip abduction includes:

- Rule 1. The abducting leg stays in the frontal plane during hip abduction. The frontal plane is the plane dividing the front and back half of the body.
- Rule 2. When abducting, the abducting leg moves from the midline of the frontal plane to the side of the body up. The abducting angle formed between the abducting leg and the initial position must be larger than a predefined value.

As the name suggests, the bowling exercise resembles the movement of regular bowling except without the bowling ball. The primary correctness rules for the bowling exercise include:

- Rule 1. The bowling arm must stay within the sagittal plane during the movements. The sagittal plane is the plane dividing the left and right sides of the body.
- Rule 2. The bowling arm must remain straight during the movement.

III. SYSTEM DESIGN AND IMPLEMENTATION

A. Design Rationale

Our system is designed to meet the following objectives:

O1 Provide a detailed, realistic visual guide on the correct movements for each exercise.

O2 Capture the actual movements of the patient for realtime visual display as well as for post-exercise review by the patient or the clinician.

O3 Provide intuitive feedback to the patient regarding the quality and quantity of the exercises.

O4 The system should not display images of the demonstrator or the patient, to conform to the privacy policy for human trial study and also to maximize the comfort level of the patients.

To satisfy objectives O1 and O4, we decided to use the Unity 3D game development framework [5]. By using the Unity framework, both the demonstrator and the patient are represented by separate 3D avatars, which automatically satisfies objective O4. Furthermore, by using 3D avatars, the movements of the demonstrator as well as the patient can be examined frame-by-frame in 360-degree view. This greatly facilitates the patient studying how to perform an exercise, and enables the clinician to examine the patient’s performance in maximum detail.

To satisfy objective O2, the 3D joint positions as well as the segment orientations of the demonstrator or the patient are captured and logged to files frame-by-frame. To satisfy objective O3, the correctness rules for each prescribed exercise are defined and implemented, and the patient’s movements during an exercise session are assessed against the rules in realtime. Furthermore, additional visual feedback objects are used to indicate the target positions for key joints during the exercise, and to indicate the quantity and quality of the exercise session.
B. Overview of the System

The system is implemented as a Unity project with the ZigFu plugin [6]. The ZigFu plugin provides a simplified interface to access the Kinect Application Programming Interfaces (APIs) within the Unity framework. The C# programming language is used to implement the system.

The project view of our system in Unity is shown in Figure 1. The main components (referred to as assets in Unity) are two predefined game objects (referred to as prefabs in Unity) provided by the ZigFu plugin:

- **DanaCoach.** This prefab is placed on the left side of the scene and its movement is controlled by a script (i.e., a C# file named Replay.cs) using the motion data collected previously when a clinician demonstrated an exercise, or when the patient performed the exercise under the supervision of a clinician. As such, this prefab is used to demonstrate correct exercise movement visually to the patient.

- **DanaSubject.** This prefab is placed on the right side of the scene and its movement is controlled by another script (i.e., a C# file named EtSkeleton.cs) using the motion data captured in realtime when a patient performs the prescribed exercise.

Other statically allocated components include the floor, directional light, main camera (placed at the center of the scene), status display, and an invisible object (called ETatHome) used to attach scripts to the ZigFu runtime for Kinect data collection. Furthermore, additional visual feedback objects are created for each exercise dynamically according to the correctness rule.

C. Correctness Rule Design

As we can see from Section II-B, the correctness rules for each exercise primarily specify the following:

- Movement restrictions on key joints and/or segments with respect to the frontal or sagittal plane.

- The amplitude of the movement for key joints.

- Restrictions on the angles formed by key segments, for example, the abducting leg must remain straight and the bowling arm must also remain straight.

Therefore, we developed a set of basic rule elements elaborated below. These elements can be used to specify the correctness rules for common rehabilitation exercises. The rules are represented in XML and they are loaded to the system at runtime so that exercises can be customized for each individual patient and new exercises can be prescribed without changing or compiling the code.

- **Target rule:** This rule specifies the target position for each key joint. In the XML representation, it is given as a sequence of Target elements. An example target rule for bowling is shown in Listing 1.

Listing 1. An example target rule for the bowling exercise.

```xml
<Target>
  <AnchorJoint>"RightShoulder"</AnchorJoint>
  <TargetJoint>"RightWrist"</TargetJoint>
  <TargetAngleXY>170</TargetAngleXY>
  <TargetAngleZ>45</TargetAngleZ>
  <ShowTarget>0</ShowTarget>
</Target>
```

The target position is specified using two angles:
- **TargetAngleXY:** The angle formed between the vertical axis pointing upward (i.e., the y axis) and the anchor-target line projected to the frontal plane.
- **TargetZ:** The angle formed between the z-axis (pointing away from the camera to the user) and the anchor-target line projected to the sagittal plane.

The example rule shown in Listing 1 specifies the initial target position of the bowling movement (i.e., the right hand moves backward to about 45 degrees). The ShowTarget is an auxiliary element informing the system whether or not a game object should be placed at the target position for visual aid. In this case, a value 0 means no object should be created.

- **RelativeAngle rule:** This rule specifies the desirable angle that should be formed between two adjacent segments during an exercise. For example, for bowling, the two segments are (right shoulder – right elbow) and (right elbow – right wrist), and the angle between the two segments should be 180 degrees, as shown in Listing 2. The XML elements used to describe the segments and the angle are self-explanatory. The ErrorBound element is used to explicitly indicate the tolerated variance to the ideal target angle.

Listing 2. An example relative angle rule for the bowling exercise.

```xml
<RelativeAngle>
  <CenterJoint>"RightElbow"</CenterJoint>
  <UpstreamJoint>"RightShoulder"</UpstreamJoint>
  <DownstreamJoint>"RightWrist"</DownstreamJoint>
  <TargetAngle>180</TargetAngle>
  <ErrorBound>10</ErrorBound>
</RelativeAngle>
```

- **MovingAngle rule:** This rule defines the restriction on the movement of a segment during an exercise in
terms of the angle formed between the segment and the frontal plane or the sagittal plane (but not both). An example rule is provided in Listing 3.

Listing 3. An example target rule for the bowling exercise.

```xml
<MovingAngle>
  <AnchorJoint>"RightShoulder"</AnchorJoint>
  <TargetJoint>"RightWrist"</TargetJoint>
  <TargetAngleXY>180</TargetAngleXY>
  <ErrorBound>10</ErrorBound>
</MovingAngle>
```

In the example, the rule dictates that the anchor-to-target joint line must move within the sagittal plane with a tolerated error of 10 degrees.

D. System Operations

The system allows several modes of operations:

- Recording demonstration session: A clinician or a patient under the supervision of the clinician can perform a designated exercise without visual guidance and the motion is recorded to files.

- Study demonstrated exercise: Before a patient engages in an exercise session at home, he/she could learn how to perform the exercise correctly by studying the recorded demonstration session. Our system allows the patient to review the recorded session frame-by-frame and in 360-degree view.

- Live exercise: Once the patient is confident in doing the exercise correctly, he/she can start doing exercises using our system. The patient can improve the quality of the exercise by observing the visual feedback provided in our system. In our current implementation, two forms of visual feedback are provided: (1) target object to indicate the amplitude of the movement, and (2) a counting object that changes color temporarily when the target is reached and when all the rules are obeyed for each iteration and the repetition count is incremented accordingly. A snapshot of this mode of operation for hip abduction is shown in Figure 2.

- Review of recorded exercise sessions: The patient or the clinician could review the performance of a patient’s session with the assessment details of every iteration (such as the actual angle measured).

IV. RELATED WORK

MotionMA [7] is a system closely related to ours. Similar to our system, MotionMA provides realtime feedback to the user regarding the quality of the exercise. However, MotionMA uses a different approach called Programming-by-Demonstration. In this approach, a model is derived from the recorded motion data of a demonstration by an expert. When the user performs the exercise, the quality is assessed against the model. For this approach to work, however, scaling must be applied carefully to accommodate the size and form differences between the demonstrator and the user.

Sun et al. [8] also proposed a system to facilitate in-home exercise assessment. The basic procedure is rather similar to that of MotionMA. The main difference is that statistical algorithms are employed to compare the motions of the demonstrator and those of the user. While potentially robust, this approach is significantly more computationally intensive and hence the assessment is done off-line instead of in realtime. Furthermore, the feedback contains only three categories: excellent, good, and bad. On the contrary, our system can provide specific feedback regarding which rule is violated.

V. CONCLUSION

In this paper, we described a rule-based system that facilitates in-home rehabilitation exercise monitoring and guidance. The core foundation enabling realtime assessment in our system is a rule-based approach. For each exercise, we define a set of correctness rules and they serve as the invariance of the exercise which is independent from the size and form of the users. We designed a set of basic rule elements that can be used to express the correctness rules for various common rehabilitation exercises. Our system may be operated in a number of modes. In addition to providing guidance and realtime assessment, our system can be used to study the demonstrated exercise and the recorded patient session frame-by-frame with 360-degree view. Hence, we believe that our system can be of practical use for rehabilitation exercises.

REFERENCES


