Monitoring Intake Gestures using Sensor Fusion (Microsoft Kinect and Inertial Sensors) for Smart Home Tele-Rehab Setting

Hossein Mousavi Hondori, Maryam Khademi, Cristina Videira Lopes

Abstract—Smart home technologies help post-stroke patients complete activities of daily living (ADL) independently, while saving their time, money, and extra effort. The patients are otherwise required to visit rehabilitation clinics for formal care. Toward this goal, we present our approach to spot specific ADL’s of eating and drinking in a home setting. We fuse inertial and Microsoft Kinect sensors to monitor the patient’s intake gestures including fine cutting, loading food, and maneuvering the food to the mouth. For both sides of the body, we measured (i) position of the wrist, elbow, and shoulder; (ii) angular displacements at the elbow and shoulder joints; and (iii) acceleration of the spoon/fork/cup which are held by the subject. The use of Kinect allows distinguishing between healthy and paralyzed body sides which is a common problem in tele-rehab. The system was tested successfully on healthy subjects; because stroke-patients show slower motion in a shorter range, the system would serve them at least equally well.

I. INTRODUCTION AND BACKGROUND

STROKE represents a major health concern for the American public, ranking as a leading cause of severe disability. Approximately 7,000,000 Americans have suffered a stroke, and approximately 795,000 new strokes occur each year in the U.S. [1]. About 80% of acute stroke survivors lose arm and hand movement skills [2]. It is recommended to continue rehabilitation until maximum recovery has been achieved. However, due to cost factors of outpatient rehabilitation facilities, the extension of chronic stroke-patients’ care is limited. Besides, therapist’ information sources about the patients’ ADL performance at home are only limited to their self reports. The accuracy and efficacy of therapy could be improved by providing assessment technology, in the form of tele-rehabilitation, within patients’ homes.

One solution to the above problem is to have smart houses which include devices to monitor the patients’ health status by helping them not only function more independently but also receive feedback on their activities. These feedbacks can be also transferred to occupational/physical therapists to get informed about the patients’ level of progress.

Zheng et al. [3] developed a home-based telerehabilitation system for stroke patients in order for them to gain maximum rehabilitation in the home environment. This system relies mainly on wearable inertial sensors. The shortcoming of this method is the need for numeric integration of inertial measurement unit (IMU) data to obtain position and angular displacement of joints which introduces error to the output.

Yeong et al. [4] analyzed the minimal requirements that the workspace of a robotic device should have in order to promote training of principal activities of daily living, considering the shoulder movement’s limitations of subacute patients. They recorded the movements using a Vicon MX digital optical motion system which is expensive enough to limit its usage to a research lab setting. Moreover, this study was solely to understand the requirements of patients to train the robots and not directly monitoring them.

Amft et al. [5] proposed a two-stage recognition system for detecting arm gestures related to typical meal intake for automatic dietary monitoring in the domain of behavioral medicine. They used body motion sensors in the form of a jacket that the patient needs to wear in order to have his/her arm gestures monitored. Although it was a useful research prototype, the system needed to be replaced with less complex sensors in order to be used out of lab setting. Besides, our fusion sensor system outperforms this system in terms of accuracy as it takes advantage of both inertial and Kinect sensors simultaneously.

II. METHOD

A. Intake Gestures

The eating and drinking activities were selected as they are critical functional activities for recovery and are among commonly trained activities during conventional therapy sessions.

Food intake usually requires upper body movements including arms and trunk. We characterized these movements into food and beverage consumption such as holding the fork/spoon/knife/cup, loading the food, and actual food intake which is composed of motions to cut and maneuver the food piece to the mouth.

B. Setup

Our setup includes inertial and Microsoft Kinect sensors.

Inertial sensor: is a MEMS device that measures and reports on acceleration.

Microsoft Kinect: “is a motion sensing input device which is a horizontal bar connected to a small base with a motorized pivot and is designed to be positioned lengthwise
above or below the video display. The device features an "RGB camera, depth sensor and multi-array microphone running proprietary software" [6], which provide full-body 3D motion capture, facial recognition, and voice recognition capabilities" [7].

Sensor fusion: “is the combination of sensory data or data derived from sensory data from disparate sources such that the resulting information is in some sense better than would be possible when these sources were used individually” [8].

The setup also includes a dining table where the patient consumes his food using fork/spoon/knife and a cup. The inertial sensors are attached to each of these eating utensils which record their local movements made by the subject. While eating, the subject is also observed by the Kinect sensor which has been placed on the same table. Fig. 1 shows the system setup.

**Fig. 1.** The system setup including the dining table, sensorized tools (fork/knife/cup), and a Kinect sensor.

### III. EXPERIMENTAL DATA

#### A. Eating and drinking task

The subject was asked to perform several movements simulating eating soup with his right hand at comfortable speed. He was then requested to repeat the task of cutting a piece of stake with his right hand and maneuvering it to his mouth with his left hand for a couple of times. Afterwards, he was instructed to perform drinking water by picking the cup with his right hand, holding it still at the mouth for a few seconds (this helps segregate data more easily afterwards) and repeating the task again for several times.

Fig. 2 shows the 3D-trajectories of the subject’s head, left and right hand while doing the above eating and drinking tasks. As can be seen, the head is approximately at the same position during the task while the two hands move back and forth between the food plate on the table and the subject’s mouth.

**Fig. 2.** 3D-trajectories of head (blue), left hand (green), and right hand (red) movements while performing eating and drinking tasks.

#### B. Analyzing Kinect sensor data

Fig. 3 illustrates the changes in position of different upper-limb joints in a body skeleton including the subject’s head, center shoulder, right/left shoulders, right/left elbows, and right/left wrists while performing the eating and drinking tasks. As can be seen, the wrists and elbows of the subject move the most among the joints confirming the large maneuvering movements between the food plate on the table and his mouth while his center shoulder and head are still.

Fig. 4 demonstrates the change of angles between different upper-limb joints of the subject while eating and drinking. As can be seen, during the maneuvering movement, the right-elbow angle changes between approximately 50 to 110 degrees, whereas the left-elbow angle varies between approximately 65 to 115 degrees. In case of cutting the stake, these angles have smaller changes which are approximately between 90 to 110 degrees. Moreover, the right and left-shoulder angles follow the same pattern as right and left-elbow angles. The only difference is that their change is not as large as what is observed in the elbow angles.

#### C. Analyzing inertial sensor data

Fig. 5 demonstrates the data measured from inertial sensors while the subject performs different phases of eating and drinking. Due to gravitational accelerations, the signal of the sensors always carries a bias equal to 9.81 m/s². By determining the direction of gravity, this bias is removed from the x, y, and z inertial sensors of each measurement unit. In case of cutting the stake, frequency of the movements has the highest value, whereas while moving the food to the mouth, frequency and magnitude of the movement is steadier. Similarly, frequency and magnitude of the movements while drinking are fairly consistent.
Fig. 3. Change in position of different upper-limb joints in a body skeleton including the subject’s head, center shoulder, right/left shoulders, right/left elbows, and right/left wrists while performing the eating and drinking tasks.

Fig. 4. Change of angles between different upper-limb joints of the subject while performing the eating and drinking tasks.
experience, rehabilitation market prefers portable tools such as “Smart Mug” [11], [12]. Such designs can be combined easier with engaging and interactive technology such as Augmented Reality [13], [14].

Future work will include adopting the above-mentioned technologies in a smart home platform to assess a broad range of ADL’s. We also plan to test the system on stroke patients to evaluate their progress using their impaired limb.

REFERENCES


