Learning Classifier to Evaluate Movement Quality in Unassisted Pick-and-Place Exercises for Post-Stroke Patients: A Preliminary Study

Hee-Tae Jung\textsuperscript{1,2}, Hwan Kim\textsuperscript{2}, Mi Young Oh\textsuperscript{3}, Taekyeong Ryu\textsuperscript{3}, Yangsoo Kim\textsuperscript{3}

\textsuperscript{1}College of Information and Computer Sciences, University of Massachusetts Amherst, USA
\textsuperscript{2}College of Rehabilitation Sciences, Daegu University, S. Korea
\textsuperscript{3}Heeyeon Rehabilitation Hospital, S. Korea

hjung@cs.umass.edu

Abstract—The use of service robots is gaining increasing interest in stroke rehabilitation research. In autonomous robot therapy sessions, it is important to monitor the qualitative motor performance of the patient and to evaluate if therapeutically desirable movements are reinforced as per the therapist’s instructions. Research on computational means to address this issue has received little attention, despite the acknowledged importance thereof. This paper poses the question of whether the subjective evaluation of the therapist on the qualitative motor performance of the individual patient can be modelled. In an attempt to answer this question, we frame it as a supervised machine learning problem and investigate if kinematic sensor readings to the subjective evaluation of the therapist can be correlated. We further examine if the personalized classifier trained with the data of the individual patient should be favored over the general classifier. Our preliminary results suggest that, given kinematic observation, trained classifiers may be able to imitate the evaluation of the therapist for the movement quality of the post-stroke patient. In our experiments, the general classifier outperformed the personalized classifiers, however, further investigation is necessary to generalize this finding.

I. INTRODUCTION

Strokes are one of the leading causes of death and disabilities in the elderly [1]. Even if a patient survives a stroke, they are often left with long-term motor disabilities [2]. In a conventional setting, the hospital assigns a team comprised of a doctor, nurses and therapists of different expertise for each patient to help them regain lost motor skills. A set of rehabilitation therapies is designed to induce a clinically more meaningful recovery from the patient. During therapy sessions, the therapist continuously examines whether the patient executes exercise movements as per instructions or not. Time and resources are often limited, and it is difficult for the patient to receive the quality and quantity of therapy they require. Hence, considering the various constraints, the hospital and the patient compromise the recovery goal and the corresponding therapies. Often, the patient is discharged with unsatisfactory recovery.

To overcome the traditional limitations and to automate part of the therapy, diverse robotic approaches have been investigated. Among them, the potential use of service robots is investigated for stroke rehabilitation in residential settings at the limited presence or absence of the therapist. Empirical studies suggest that therapy sessions may lead to observable improvement in motor function when the service robot induces greater ranges of motion by presenting progressively more difficult targets [3], [4]. In these studies, the robot evaluates the completion of the exercise movement by measuring the Euclidean distance between the target and the hand position of the patient, which introduces the discrepancy in the evaluation by the robot and the therapist [5]. Hence, even when the patient recruits a significant amount of compensation to reach the target, the robot may consider it as success and presents more difficult targets. This in turn may induce a greater amount of compensation.

The same problem exists in the socially assistive approach [6]. Such an approach focuses mostly on the engagement of the patient and stops short at evaluating the quality of the exercise movement by the patient. Hence, the service robot may encourage the patient to continue the exercise but not necessarily in a clinically meaningful manner. The same problem may be observed when using physically assistive robots [7] and rehabilitation games [8], unless the therapy session is supervised by an experienced therapist or an explicit computational means to imitate such a skill set is devised.

Addressing the functional limitation of contemporary approaches, we propose a supervised learning framework to correlate kinematic movements of the individual patient to the subjective evaluation of the therapist. Our results suggest that the proposed approach can successfully train classifiers that imitate the binary evaluation of the therapist given the kinematic data of the patient.

II. BACKGROUND

There exists an increasing amount of empirical evidence that acknowledges the importance of practicing exercise movements in a controlled manner while minimizing unnecessary compensation by the healthy side of the patient. When this principle is not upheld, observed functional gains may come from the improved motor compensation not from the motor recovery of the paretic limbs [9]–[12]. In practice, it is common that the patient goes astray from clinically meaningful movements and exploits task-particular compensatory behaviors that he has developed. Typical behavior in pick-and-place movements are illustrated in Figure 1. Hence, it is important to enable the robot to examine if the patient executes the movements per the instructions of the therapist.
An exercise movement is demonstrated by a therapist wearing IMUs.

(a) The items used for the study
(b) An exercise movement is demonstrated by a therapist wearing IMUs.

Fig. 2. The experiment setting

A patient should have had 70% or higher of the maximum range of motion, scored P- or higher at the muscle manual test [13]. The patient also should have been cognitively competent (24 ≤ mini-mental state examination [14]) to be able to comprehend and follow the instructions of the therapist during the data collection session. The recruited patients were both female, 81 and 69 years old respectively. The time from onset until the study participation were 10 months and 20 months respectively. The therapists should have had at least three years of practice experience to be included in the study. The informed consents were obtained from the patients and the therapists.

B. Therapeutic Exercise Task and Workspace

A task was to move a given item from an initial position to a target position specified by the therapist (Figure 3(a)). The therapist chose the targets that they deemed necessary to induce different types and levels of compensation and the performance. For the task, a 2-dimensional Cartesian space was defined on a table top, the size of which was approximately 1.7m × 0.9m. The grid was marked on the table so that both the therapist and the patient could easily identify target positions. The size of each grid cell was 0.1m × 0.1m. The initial positions were determined for patients A and B in accordance with their motor function. In general, they were near the center of the patient’s trunk. To gain an increased understanding, please see the actual setting (Figure 2(b)).

C. Data Collection

For each patient and therapist grouping, data was collected once a day for four consecutive days at the rehabilitation hospital. Before each session, a total of five MPU-9150 IMUs (InvenSense, USA) were fixed to the patient’s body using Velcro straps so as not to interfere with movements: one each on the torso, both upper and both forearms of the patient. During the session, the therapist placed an object at the initial position and pointed at a target position of their choice. Then the patient was to move the object to the specified target position. To induce a therapeutically desirable as deemed by the therapist, they gave necessary instructions, some of which are as follows.

1) Please, sit up straight as much as possible.
2) Please, place your intact arm on your knee naturally.
3) Please, use your entire arm.

III. PROPOSED APPROACH

During a therapy session, the patient practices with a set of targets \( \mathbf{p} = \{ \mathbf{p}_1, \ldots, \mathbf{p}_N \} \) within a therapeutic task that is prescribed by the therapist. For each given \( \mathbf{p}_n \), the patient executes an exercise movement, which is captured by sensors. We hypothesize that the kinematic information derived from the sensor readings may reveal the quality of patient’s movement, based on which the robot evaluates if the patient follows the instructions of the therapist and exerts his best effort. To correlate the kinematic information to the binary evaluation of the therapist, we take a supervised learning approach. Specifically, given the sensor readings \( \Xi \), we propose to compute the kinematic performance features \( X(\Xi) \). Therapist’s evaluation is specified binary \( y = \{0, 1\} \) where 0 indicates that the patient is following the therapist’s instructions in executing exercise movements and 1 indicates otherwise. Given these, we train a model \( f : X \rightarrow Y \).
Owing to the patient’s age and motor deficit, the patient was not always able to maintain instructed postures. In such cases, the therapist was not permitted to give physical assistance to the patient so as to minimize the noise in the collected kinematic data. This process was repeated for an average of approximately fifteen minutes.

During the therapy session, the movements of the patient were measured by the worn IMUs and video recorded in two different directions. The IMUs measured the orientation of the corresponding body segments in the form of quaternions and streamed to a laptop at 250Hz, which were stored in files real-time. After the completion of the session, the recorded videos were played and the performance of the patient was evaluated by the therapist in a binary manner $y \in \{0, 1\}$. In this paper, it was evaluated 0 if the patient could attain the given target in the instructed posture and with minimum compensatory movement (e.g. Figure 1(a)). It was evaluated 1 if the patient deviated from the therapist’s instructions and recruited too much compensatory behavior (e.g. Figure 1(b)-1(d)).

Since it is not likely to see significant changes in motor performance in just four days, the union of the datasets from four sessions were used as a single dataset for each patient (TABLE I). As result, for the patient A, we obtained 393 total movements where the qualitative performance of 195 movements was evaluated 0 and that of 198 was evaluated 1. Each kinematic observation lasted as short as 1.8 seconds and as long as 23.3 seconds. The average duration of all the movements was 5.7 seconds. For the patient B, we obtained 214 total movements where the performance of 90 movements was evaluated 0 and that of 124 movements was evaluated 1 by the therapist. The data for each exercise was as short as 4.0 seconds and as long as 69.1 seconds. The average duration for the movements was 17.20 seconds.

### D. Kinematic Performance Features

The sensor readings and the measured lengths of body segments were used to approximate relative joint angles between two consecutive body segments (in degrees) and the Cartesian position of body joints (in meters). They, in turn, were used to compute intermediate features $f$ that describe the qualitative traits of the kinematic performance of the patient. These features are known as *kinematic parameters* in the literature [15], [16], which are devised to complement the conventional standardized assessment tools and to objectively quantify the qualitative motor performance when the patient engages in the target-directed movement. Empirical studies provide evidence that, given the same task and targets, difference in the motor performance of stroke patients and normally functioning people can be observed by these parameters [17], [18]. It is suggested that some of these parameters can capture how stroke patients recruit different degrees of freedom to compensate for motor deficits that is not observed from normally functioning people [10], [12]. Of these parameters, we have chosen to use the following features.

1. **Peak velocity (PV)** - Greater Cartesian peak velocity of a limb is associated with greater muscular activation during motor action. In this work, PV for an impaired hand position ($x_1$) is used.
2. **Movement time (MT)** - Shorter movement time indicates greater efficiency of the movement. In this work, MT for an impaired hand position ($x_2$) is computed by the difference between the deactivation time and the activation time. Activation time is when the hand velocity exceeds $k \times PV$ for the first time while deactivation time is when the hand velocity drops below $k \times PV$. Following the convention by [19], $k = 0.05$ is used in this work.
3. **Total displacement (TD)** - Less total displacement implies more efficiency. To capture the quality of the movement that is robust to the varying distance of different targets, we use the relative TD of an impaired hand position with respect to the straight line distance to the target ($x_3$).
4. **Percentage of movement time at peak velocity (PPV)** - Greater PPV implies that the patient requires less sensory feedback for correcting target-directed movements. In this work, PPV of an impaired hand position ($x_4$) is used.
5. **Movement units (MU)** - A smaller number of movement units is associated with smoother movement while accomplishing a target. A single movement unit consists of a pair of one acceleration and one deceleration. Their counts for an impaired hand position ($x_5$) between the activation time and the deactivation time is used.
6. **Angle Amplitude (AA)** - Joint angle indicates the amount of contribution the joint makes to the exercise movement. In this work, AA for each arm is computed by summing the difference between the maximum joint angle and the minimum joint angle of shoulder and elbow ($x_{6:7}$). AA of a trunk ($x_8$) is also used.

### E. Classifier Training

For the current study, we employed *logistic regression*. Training was conducted using the *SimpleLogistic* function of Weka 3.8.1, where the logistic function is defined as follows.

$$f(x) = \frac{1}{1 + e^{-w^Tx}}$$

### F. Evaluation

Each patient may exhibit varying patterns of motor movements, and it can be hypothesized that the classifier trained for the individual patient may perform better than the general one. Hence, we first trained personalized classifiers (P) using the data of the patient A and the data of the patient B.
(a) The personalized classifier (left) and the general classifier (right) tested on the testing data of the patient A and B respectively. Separately, the general classifier (G) was trained using the data from both patients with the size of the training data controlled. The performance of the trained classifiers was subsequently computed in F-measures using 10-fold evaluation. Specifically, the trained classifiers (P and G) were applied to the held-out testing data for the patient A and B (here after P_A, G_A, P_B, and G_B).

V. RESULTS

The F-measures of the classifiers P_A, G_A, P_B, and G_B are 85.16%(±3.05), 87.30%(±1.39), 89.69%(±3.62), and 91.19%(±3.70) respectively. These results suggest that, with high accuracy, the trained classifiers can examine if the patient executes exercise movements per the instructions of their therapist. The performance of the classifiers G_A and G_B was marginally better than their counterparts. In other words, from our findings, the general classifier performed better than the personalized classifiers trained for the specific individual patients. However, owing to the limited number of the patients in this study, it is difficult to determine if the general classifier should be favored over the personalized classifier.

VI. CONCLUSION

With the goal of enabling the service robot in examining the quality of the exercise movement by the patient during the autonomous robot-mediated therapy, we proposed to take a supervised learning approach to correlate the kinematic information to the binary evaluation of the therapist. Specifically, we employed logistic regression to train the personalized classifiers and the general one. The findings suggest that both the personalized and general classifiers can evaluate the quality of the patient’s movement with high accuracy. Against our initial hypothesis, the general classifier trained using the data of both post-stroke patients outperformed the personalized classifiers trained using the data of the individual patients. Considering that the data was collected only from two patients, however, it is difficult to conclude that such results would still hold true when the general classifier is trained on a larger number of post-stroke patients. With further research, the proposed approach may enable the service robot in administering the therapy session with higher quality.

ACKNOWLEDGMENT

The data collection and the experiment of this particular study was conducted while Hee-Tae Jung was away from University of Massachusetts. The authors thank Namho Gong and Yu-Kyong Choe for their help.

REFERENCES