Decision Support for Stroke Rehabilitation Therapy via Describable Attribute-based Decision Trees

Vinay Venkataraman, Pavan Turaga, Nicole Lehrer, Michael Baran, Thanassis Rikakis, and Steven L. Wolf

Abstract—This paper proposes a computational framework for movement quality assessment using a decision tree model that can potentially assist a physical therapist in a telerehabilitation context. Using a dataset of key kinematic attributes collected from eight stroke survivors, we demonstrate that the framework can be reliably used for movement quality assessment of a reach-to-grasp cone task, an activity commonly used in upper extremity stroke rehabilitation therapy. The proposed framework is capable of providing movement quality scores that are highly correlated to the ratings provided by therapists, who used a custom rating rubric created by rehabilitation experts. Our hypothesis is that a decision tree model could be easily utilized by therapists as a potential assistive tool, especially in evaluating movement quality on a large-scale dataset collected during unsupervised rehabilitation (e.g., training at the home), thereby reducing the time and cost of rehabilitation treatment.

I. INTRODUCTION

Stroke is the leading cause of disability in adults leaving millions disabled with chronic impairments [15], often left untreated due to insufficient coverage by insurance to undergo long-term rehabilitation therapy treatment. The traditional rehabilitation treatment composed of repetitive movement tasks under the supervision of a physical therapist can achieve motor recovery following stroke [8], [9], [12], [21]. To support long-term recovery, investigators in the stroke rehabilitation community believe that clinical intervention should be reinforced with home-based therapy, which a participant can possibly experience without much therapist supervision [1], [7], [19], and can also reduce the cost of long-term therapy treatment. Virtual and mixed reality environments have been incorporated in stroke rehabilitation to induce active learning by providing auditory and visual feedback based on automatic computational evaluations of movement. In this direction, a Home-based Adaptive Mixed Reality Rehabilitation (HAMRR) system (shown in Fig. 1) which integrates rehabilitation and motor learning theories with motion capture, activity analysis and multimedia feedback [5], [7], [10], has been shown as an effective rehabilitation system.

Researchers have been motivated to develop frameworks for quantification of movement quality [4], [6], [20], [22], [24] given its potential impact on disseminating interactive rehabilitation training to unsupervised contexts such as the home. Several automated approaches exist in literature to quantify movement quality based on complex models including nonlinear dynamical system theory [18], [20], [22], random forests [16], and SVMs [17]. While these approaches provide a computational framework for movement quality assessment showing high correlation with the clinical assess-
ment scores, it would be beneficial to have an interpretable framework which can be used as a decision support tool by physical therapists during rehabilitation treatment.

To assess the level of functional ability of a stroke survivor, therapists can employ validated rating rubrics such as the Wolf Motor Function Test [23], to systematically assign a movement quality assessment score after observing a participant perform a predefined set of functional tasks. Such a rubric imposes a hierarchical set of rules for a therapist to consider, in order to help evaluate a participant’s performance. Given this method of translating visual observation of movement to a quantitative score, we were motivated to investigate if a computational framework based on kinematic features can also be structured in a hierarchical form that can be easily understood by a therapist. We believe that such a framework would be useful in providing recommendations to physical therapists especially in the context of telerehabilitation, where a therapist reviews large amounts of movement performance data produced by a participant performing rehabilitation exercises without supervision (e.g., in the home). Large scale movement quality evaluation would greatly benefit from such systems by providing recommendations to therapists and also allowing them to check the reason for recommended movement quality score using describable attributes indicative of the impairments.

Contributions: We propose a hierarchical model using decision trees to simulate the results of the rating rubric created by rehabilitation experts to rate reach to grasp tasks across stroke survivors of various deficit. This is a step towards development of generalized models for knowledge representation of movement quality assessment of reach and grasp action based on previous work [10], [13], [14]. Within this experimental framework, we assume a simplified kinematic representation of reach and grasp action which focuses on a few specific elements of reach movement suitable for real-time monitoring and quantification of movement quality. The elements of the reaching movement chosen in our experiments include hand trajectory error in the horizontal and vertical planes, peak speed, jerkiness [6], velocity bellness [6] and torso rotation along XYZ axes. The main goal of this work is to learn a model that can simulate the resultant ratings of therapists using a rating rubric for movement quality assessment based on low-level kinematics indicative of the participant’s impairment which can be used as a decision support system to aid the therapist during supervised rehabilitation therapy.

II. METHODS FOR COLLECTING KINEMATICS AND THERAPIST RATINGS

A. Collection of Kinematics

The HAMRR system was designed to provide rehabilitation therapy to stroke survivors in a home-setting with reduced supervision by a physical therapist. This system was used as an apparatus to collect kinematics when participants perform movement tasks without any assistance of feedback. The HAMRR system has four Natural Point Opti-Track cameras facing down on a table to track a single reflective marker placed on the participant’s wrist (wrist marker) and four markers on the corners of a rectangular rigid plate placed on the participant’s left side of chest (Fig. 1 inset). The selection of the wrist marker was motivated by previous investigations indicating that the wrist trajectory as the most informative joint with respect to analyzing reach trajectory performance [6], [22]. In addition, we believe that it is important to monitor the torso compensatory strategies for efficient movement analysis.

The selection of the plate was motivated by efforts to capture body compensation. To compensate for the lack of extension during a reach, many stroke survivors use excessive shoulder movement (elevation and/or protraction) and excessive torso movement (flexion and/or rotation). Therefore, a system for rehabilitation training should monitor movement of the body to determine the extent to which a participant is utilizing pre-stroke movement strategies to advance his/her hand towards the target. The HAMRR system was designed for home-based use, and the sensing apparatus worn by the participant must be simple and easy to wear. Therefore, we are only using a single plate worn on the chest of the participant, which captures coarse torso movement as opposed to both shoulder and torso movement separately. The system is shown in Fig. 1 and detailed information of the system design can be found in [2].

B. Therapist Rating Protocol

Stroke rehabilitation experts have standardized means for systematically rating overall functional performance of a defined set of tasks (relevant to activities of daily living) included within the WMFT protocol. However, within the stroke rehabilitation community there lacks a consensus among physical therapists in defining an ontology of component-level labels for movement quality (i.e., methods for rating the movement components that contribute to completion of a functional task), thereby leading to lack of training datasets to develop algorithms for movement quality assessment. In other words, while kinematics can capture the component-level aspects of movement (trajectory, compensation) which are important for evaluating movement quality, there is not yet a corresponding rating system in the stroke rehabilitation community for these components. Therefore, our team has collaborated with rehabilitation experts to introduce a new rubric for physical therapists to rate movement quality for specific tasks trained by the HAMRR system. Movement quality is assessed in terms of trajectory, compensation, manipulation, transport of an object, and release. However, we limit our focus on trajectory and compensation in the context of reaching to grasp a stationary cone, as these movement components have established corresponding methods for quantifying performance using kinematics derived from 3D positions of reflective markers described in the section II-A. The rating rubric used by therapists to rate trajectory and compensation is provided in Table I. One should note that this rubric was designed given the constraints of the therapist viewing a single camera video of the participant while performing a task from the
C. Data Collection

The dataset used in our experiments consists of reaching tasks performed by a total of eight participants (refer Table II for demographics) to a cone on-table located at the participant’s midline. Each participant performed five reaches in each of four sessions (one session per week). These reaches were performed without any feedback from the system or therapist unless the participant was unclear on how to perform the task. During the task, each participant was seated at the HAMRR system and his/her movement was captured by the Opti-Track system. A custom designed iPad application was also concurrently used to capture video footage of a participant performing these tasks. These videos were randomized across participants and sessions before they were provided to therapists for evaluation. Therapists could only view one video at a time and were allowed to watch the videos as many times as they needed to form a decision on the ratings. However, therapists were not allowed to see or change responses to previous videos once they were submitted.

Trajectory performance was rated on a scale from 1 – 4 based on the therapist’s impression of the participant’s performance, where a 1 denotes that the participant could not complete the task and a 4 denotes that the participant performed the task with the same quality of performance as the therapist if he/she were to perform it. Compensation was rated on a scale from 1 – 4 based on the participant’s excessive use of the shoulder and/or torso and if compensation was used in single or multiple planes of movement. A 1 denotes that the participant used both excessive shoulder and torso movement in multiple planes of movement, while a 4 denotes that the shoulder and trunk are positioned in a manner similar to the therapist if he/she was performing the task.

III. Definitions of Kinematic Features

The following kinematic features were extracted to quantify the impairments of a participant while performing a reach to grasp a cone task.

1) Kinematic Features from Wrist Trajectory:

a) Trajectory Error: Trajectory error is a measure of spatial deviation of the wrist trajectory from the reference trajectory. The three-dimensional positions of the wrist marker \( p(t) = [x(t), y(t), z(t)] \), \( t = 0, \ldots, \tau \) were recorded from the start of the movement to the target grasp state. The coordinate system was rotated such that \( p(0) \) was the origin, \( X - Z \) plane was the horizontal plane and the straight line connecting \( p(0) \) and \( p(\tau) \) lies along the new \( Z \)-axis. This in effect re-parameterizes (after normalization) the trajectory \( x(t), y(t), z(t) \), \( t = 0, \ldots, \tau \) to \( {x}'(z), {y}'(z) \), \( z = 0, \ldots, 1 \). This re-parameterization works without introducing significant ambiguity in our experiments due to the strong directionality of the reach action. The \( Z \)-axis was further quantized into \( N = 50 \) bins, thereby transforming the trajectory to \( {x}'(n), {y}'(n) \), \( n = 0, \ldots, N - 1 \). We now have a vectorial representation of the trajectory suitable for real-time comparisons. For every point in the reach trajectory, horizontal error \( E_{\text{hor}} \) and vertical error \( E_{\text{vert}} \) were defined as

\[
E_{\text{hor}}(i) = x(i) - x_{\text{ref}}(i), \quad i = 0, \ldots, N - 1 \quad (1a)
\]
\[
E_{\text{vert}}(i) = y(i) - y_{\text{ref}}(i), \quad i = 0, \ldots, N - 1 \tag{1b}
\]

The horizontal trajectory error (\(\hat{E}_{\text{hor}}\)) and vertical trajectory error (\(\hat{E}_{\text{ver}}\)) were defined as (units in mm)

\[
\hat{E}_{\text{hor}} = \max_{0<i<N-1} (E_{\text{hor}}) \tag{1c}
\]
\[
\hat{E}_{\text{ver}} = \max_{0<i<N-1} (E_{\text{ver}}) \tag{1d}
\]

b) Jerkiness: The jerkiness (or smoothness) feature is a measure of variations in the velocity profile. An 'efficient' reach movement should have a smooth velocity profile with an accelerating followed by a decelerating pattern without any jerks. Jerkiness (in \(m/s^3\)) of a movement was computed using the definition given in [6] as

\[
J = \int_{t_{\text{som}}}^{t_{\text{eom}}} \sqrt{\left(\frac{d^3x}{dt^3}\right)^2 + \left(\frac{d^3y}{dt^3}\right)^2 + \left(\frac{d^3z}{dt^3}\right)^2} \, dt \tag{2}
\]

where \(x, y, z\) are 3-D coordinates of the participant’s wrist trajectory. \(t_{\text{som}}\) is the time index corresponding to start of the movement and \(t_{\text{eom}}\) is the time index of end of the movement.

c) Velocity Bellness: Ideally, the velocity profile of a reaching task should be a bell curve. Typically, stroke survivors throw their arm towards the target and then make fine adjustments to grasp the object. These adjustments show up as additional phases in the speed profile. It is believed that these occur during the deceleration phase and we use normalized area to evaluate velocity bellness (\(B_{NA}\)) given by

\[
B_{NA} = \frac{\int_{t_{v_{\text{max}}}}^{t_{\text{eom}}} v(t) \, dt}{\int_{t_{v_{\text{max}}}}^{t_{\text{eom}}} v(t) \, dt} \tag{3}
\]

where \(v(t)\) is the instantaneous velocity, \(t_{v_{\text{max}}}\) is the time index corresponding to maximum velocity, \(t_{\text{eom}}\) is the end of the first phase.

d) Peak Speed: An efficient reach movement is typically accomplished by a hand velocity between 0.4m/s and 0.6m/s. We use peak speed (in m/s) as a measure of deviation from this ideal range defined as the maximum velocity of each trial given by

\[
V_{\text{max}} = \max_{t_{\text{som}} < t < t_{\text{eom}}} [v(t)] \tag{4}
\]

2) Torso Features: Torso compensation is usually in the form of significant levels of torso leaning forward or torso twisting to the sides, which can negatively impact in long term functional recovery. As mentioned in section II-A, we use a rigid rectangular plate with four reflective markers on the corners, placed on the participant’s left side of chest to track the torso movements. Using the trajectory of rotation angles \(R_x\), \(R_y\) and \(R_z\) extracted from the centroid of the rigid plate (in radians), we compute a thresholded error function given by

\[
\hat{E}_{\text{hor}} > 73.05 \quad \hat{E}_{\text{ver}} > 76.32 \quad V_{\text{max}} > 0.54 \quad B_{NA} > 17.77 \quad J > 3.46 \quad 3.71 \times 10^{-3}
\]

Fig. 3: The decision tree model for movement quality assessment of wrist trajectory. The low-level kinematic features used were horizontal trajectory error (\(\hat{E}_{\text{hor}}\)), vertical trajectory error (\(\hat{E}_{\text{ver}}\)), peak speed (\(V_{\text{max}}\)), velocity bellness (\(B_{NA}\)) and jerkiness (\(J\)). The scores highlighted in blue are the decision tree outputs for wrist trajectory analysis.

\[
\hat{R}_x(i) = \begin{cases} R_x(i) & \text{if } R_x(i) > T_1 \\ 0 & \text{otherwise} \end{cases} \tag{5a}
\]

Similarly, thresholded error functions \(\hat{R}_y\) and \(\hat{R}_z\) are calculated using thresholds \(T_2\) and \(T_3\) respectively. The confidence values for torso compensation is then computed as

\[
C_x = \frac{\sum_{i<j} \hat{R}_x(i)}{\sum_{i<j} \hat{R}_x(i)} \tag{5b}
\]

Confidence levels \(C_y\) and \(C_z\) are similarly computed. The thresholds \(T_1\), \(T_2\) and \(T_3\) were selected to be 0.15 based on empirical analysis.

IV. EXPERIMENTAL RESULTS

Using the data collected from eight participants performing reach and grasp movements to a cone target, we learned a decision tree model (using CART [3]) for automated evaluation of movement quality. Each participant performed five reach movements in each of four sessions, corresponding to a total of 130 reaches (two participants could only complete one session each). 70% of reaches were selected as training samples, and the remainder were selected as testing samples.
The dataset was split into one 70 – 30 split at random and we do not provide any crossvalidation results in this paper.

The decision tree model for the wrist trajectory using low-level kinematic features like trajectory error, jerkiness, velocity bellness and peak speed is shown in Fig. 3. In order to evaluate the model for movement quality assessment of wrist trajectory, we compare the decision tree outputs with the component-level score for wrist trajectory given by therapists (shown in Fig. 4). The Pearson correlation coefficient between the decision tree model output and the therapist rating was found to be 0.8049. Similarly, using rotation angles, we learned a decision tree model for quantification of torso compensation as shown in Fig. 5. Comparison of decision tree outputs with the therapist rating for component-level scores for torso compensation (shown in Fig. 6) show high correlation with the Pearson correlation coefficient to be 0.9129.

Evident from the high correlations with the therapist ratings, our results indicate that the kinematic movement components such as trajectory error, speed profile deviation and torso compensation, when used as inputs to our decision tree models, are capable of simulating the ratings from therapists using the rating rubric described in Table I. These results lead us further postulate that the proposed framework can be used as an assistive tool to therapists during supervised rehabilitation to reduce the time spent on movement quality assessment.

From the wrist marker analysis, we find that the kinematic attributes indicative of impairments in spatial domain are near the top of the tree, while the kinematic attributes related to speed-profile deviation are near the bottom. One potential explanation for this is that these attributes are more easily perceived by therapists given the camera view of the videos being rated (see Fig. 2) and therefore are more important in the decision tree hierarchy. However, multiple hidden factors may also have influenced this result and more analysis or larger datasets reviewed by multiple therapists is needed before any conclusions can be drawn. Further, the torso analysis show that the rotation along Z-axis was not included in the decision tree model, indicating it to be an unnatural movement.

V. Conclusion and Future Work

In this paper, we present a computational framework capable of simulating the component-level movement quality assessment rubric with imposed hierarchical structure on physical therapists. This automatic assessment of movement quality framework can provide suggestions to physical therapists during supervised rehabilitation reducing the time spent on evaluating the quality of movements, thereby reducing the cost of long-term rehabilitation treatment.

Our results indicate that the kinematic components we chose (hand trajectory error in the horizontal and vertical planes, peak speed, jerkiness, velocity bellness and torso rotation along XYZ axes) combined with a decision tree model are capable of simulating the results of an imposed hierarchical structure used by trained therapists. The selected
low-level kinematic attributes are representative of the impairments in reach and grasp action and can collectively be used to generate a movement ‘component score’ showing high correlation with the therapist rating. These results also indicate that the proposed framework can be used as an assistive tool to therapists during supervised rehabilitation to reduce the time spent on movement quality assessment.

To more specifically qualify our findings: the rehabilitation experts were able to create an imposed hierarchy based on expert knowledge (presented in Table I). Given this hierarchy developed by expert knowledge and its careful implementations by highly trained therapists, we are able to replicate the results of their ratings through a decision tree approach. Our initial results support that these decision trees can help with semi-automated ratings when the therapist is absent, and assist therapists to provide ratings faster when they log-in remotely to fine tune a home-based training system for a participant. Since we achieved favorable results using this decision tree approach given a particular imposed hierarchy, when the hierarchy needs to be switched for different types of training, we propose that similar trees can be estimated based on different hierarchies across tasks, stages of therapy, and participants. Thus, our process is dependent on clear declarations of hierarchies by therapists and their consistent implementation.

Defining an ontology of component-level labels for movement quality assessment is seen as a difficult problem in the stroke rehabilitation community. While the current research was directed towards learning a simple decision tree model for knowledge representation of given physical therapists, our future goal is to extract a generalized knowledge representation for movement quality assessment using evaluations from multiple therapists. Similar problems have been discussed in the machine learning community [11]. We are currently collecting evaluation ratings from multiple therapists as different knowledge representations for movement quality assessment and will be used to estimate a generalized knowledge model using existing approaches for matching of knowledge structures.

VI. ACKNOWLEDGEMENTS

The authors thank the physical therapists Aimee Reiss and Marsha Bidgood at Emory University, and Meghan Buell at Rehabilitation Institute of Chicago for their contributions.

REFERENCES