IEEE Journal of Translational Engineering in Health and Medicine

Received 21 February 2023; revised 12 June 2023; accepted 25 July 2023.
Date of publication 4 August 2023; date of current version 22 November 2023.

Digital Object Identifier 10.1109/JTEHM.2023.3300929

A Comparison of Approaches for Segmenting the Reaching and Targeting Motion Primitives in Functional Upper Extremity Reaching Tasks

KYLE L. JACKSON[1], ZORAN DURIĆ[1,2], SUSANNAH M. ENGDAHL[2,3,4], ANTHONY C. SANTAGO II[5], SIDDHARTHA SIKDAR[2,3], AND LYNN H. GERBER[2,6,7]

[1]Department of Computer Science, George Mason University, Fairfax, VA 22030, USA
[2]Center for Adaptive Systems and Brain-Body Interactions, George Mason University, Fairfax, VA 22030, USA
[3]Department of Bioengineering, George Mason University, Fairfax, VA 22030, USA
[4]The American Orthotic and Prosthetic Association, Alexandria, VA 22314, USA
[5]MITRE Corporation, McLean, VA 22102, USA
[6]College of Public Health, George Mason University, Fairfax, VA 22030, USA
[7]Inova Health System, Falls Church, VA 22042, USA

CORRESPONDING AUTHOR: K. L. JACKSON (kjacks21@gmu.edu)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board, George Mason University, under Approval Nos. 477548 and 492701.

ABSTRACT There is growing interest in the kinematic analysis of human functional upper extremity movement (FUEM) for applications such as health monitoring and rehabilitation. Deconstructing functional movements into activities, actions, and primitives is a necessary procedure for many of these kinematic analyses. Advances in machine learning have led to progress in human activity and action recognition. However, their utility for analyzing the FUEM primitives of reaching and targeting during reach-to-grasp and reach-to-point tasks remains limited. Domain experts use a variety of methods for segmenting the reaching and targeting motion primitives, such as kinematic thresholds, with no consensus on what methods are best to use. Additionally, current studies are small enough that segmentation results can be manually inspected for correctness. As interest in FUEM kinematic analysis expands, such as in the clinic, the amount of data needing segmentation will likely exceed the capacity of existing segmentation workflows used in research laboratories, requiring new methods and workflows for making segmentation less cumbersome. This paper investigates five reaching and targeting motion primitive segmentation methods in two different domains (haptics simulation and real world) and how to evaluate these methods. This work finds that most of the segmentation methods evaluated perform reasonably well given current limitations in our ability to evaluate segmentation results. Furthermore, we propose a method to automatically identify potentially incorrect segmentation results for further review by the human evaluator.

INDEX TERMS Time series analysis, movement segmentation, upper extremity, kinematic analysis, rehabilitation, motion primitives.

Clinical impact: This work supports efforts to automate aspects of processing upper extremity kinematic data used to evaluate reaching and grasping, which will be necessary for more widespread usage in clinical settings.

I. INTRODUCTION

Assessments of functional upper extremity (UE) movement quality are commonly used in the clinic and provide useful treatment outcome measures. These assessments typically require an individual to perform standardized functional tasks, such as moving small blocks over a partition under time constraint [1] and writing and drawing [2], [3]. Functional assessment measures also include self-reports, such as the Quick Disabilities of Arm, Shoulder, and Hand (QuickDASH) [4]. As a measure of functional ability, these assessments can capture an individual's perceived difficulty with the task [5], an observer's rating of an individual's ability

IEEE Journal of Translational Engineering in Health and Medicine

to perform the task [6], or the time required to complete the task [1].

While currently validated UE functional assessments (UEFAs) are essential to clinical practice, self-report measures can be biased and may not be sensitive [7], [8]. Furthermore, existing measures do not fully capture movement quality and efficiency, which are important for discerning between behavioral restitution and compensation during stroke rehabilitation [9] and evaluating UE prostheses [10], [11], [12], among other applications.

Using kinematics (e.g., position trajectory, velocity magnitude, acceleration, joint angles, etc.) to assess UE functional ability provides a more objective assessment of current skill and functional progression, which is challenging to capture when a domain expert (i.e., clinicians and biomechanists) relies solely on qualitative observational data [8]. Therefore, domain experts have begun using kinematic data to evaluate functional UE movement quality, such as smoothness [13] and efficiency [8], during standardized assessments.

Reaching, grasping, touching, pointing, or otherwise manipulating objects are essential in activities of daily living that incorporate the UE and are therefore of concern to clinicians and movement scientists. UE functional motions involving reach-to-point (RTP) [14] and reach-to-grasp (RTG) [15] are generally characterized by an initial reaching motion that covers most of the distance to the object (i.e., point of interest) followed by a deceleration into a targeted movement period [14]. Our definition of targeting considers the grasp primitive as a subset in targeting, where targeting can begin before grasping begins, e.g., when the grasp aperture increases [16]. Additionally, reaching typically requires gross UE motion, while targeting requires fine movements, therefore representing different functional challenges to the individual and involve potentially different neuromechanical pathways. Therefore, recent kinematic analyses have analyzed the reaching [3], [10], [12], [17] and targeting primitives [10], [17] separately for RTP and RTG motions.

Segmenting the reaching and targeting motion primitives for UEFAs using kinematic data is a relatively recent development. A variety of segmentation methods have been used, with relatively little discussion on how best to perform the segmentation. Furthermore, existing workflows used for segmentation involve applying segmentation algorithms and then manually reviewing the results [11]. As interest in FUEM kinematic analysis expands, such as in the clinic, the amount of data needing segmentation will likely exceed the capacity of existing segmentation workflows used by researchers, requiring new methods and workflows for making segmentation less cumbersome. A better understanding of the performance of different segmentation methods and how to incorporate segmentation automation into the kinematics analysis workflow [18] will support the translation of kinematics analyses to the clinic.

This paper provides insight into the performance of five different segmentation methods, including how to evaluate these methods and automatically identify questionable segmentation results that require further review. To our knowledge, this paper is the first to address the specific problem of segmenting reaching and targeting motion primitives for UEFAs. Our contributions are as follows:

- Comparing methods for segmenting the reaching and targeting motion primitives on data sets from two different domains (i.e., haptics simulation and real world), where we find most of the segmentation methods do similarly well.
- Proposing the novel use of the minimum jerk trajectory velocity profile as an indicator of potentially poor segmentation results, which can be used to direct a human reviewer's attention for further evaluation.
- Identification of challenges and opportunities associated with evaluating segmentation performance, which provides insight into the future development of segmentation methods.
- Modifications are made to the segmentation method from Jackson et al. [17] to adjust for edge cases.

Implementations of the methods discussed in this paper are available on GitHub.[1] The data used in this study are available at [19].

TABLE 1. Functional motion hierarchy modified from Schambra et al. [23].

Hierarchy Layer	Goals	Duration	Examples
Activities (broad)	Many	Minutes to hours	• Cooking dinner • Bathing • Getting dressed
Functional Movements or Actions	Few	Seconds	• Tasting soup • Zipping up jacket • Picking up cup
Functional Primitives or Movemes (granular)	One	Sub-seconds to seconds	• Reach • Reposition • Touch • Target • Grasp • Transport • Stabilize • Idle

II. BACKGROUND

A. MOTION PRIMITIVE SEGMENTATION

Methods for segmenting human functional movement are typically developed and evaluated for specific layers in the functional motion hierarchy described in Table 1. The reasoning for this is twofold. First, domain experts must be able to consistently extract the relevant subsections of an individual's movement to ensure they are performing controlled comparisons (e.g., when evaluating patient progress at various points during the rehabilitation process). Second, the developed segmentation algorithms are designed specifically for different components of the hierarchy (i.e., an action recognition algorithm will not do primitive segmentation). The definition of a movement segment varies across algorithms and applications but is generally considered to be a subsequence of the original time series sequence [20]. A potentially

[1] https://github.com/kjacks21/UE-reach-grasp-seg

confusing aspect of the literature is inconsistent references to the layers in the functional UE movement hierarchy in Table 1. For example, Lin [21]'s stated focus is on segmenting primitives for rehabilitation, however the labeled movement classes across the reviewed data sets include a mix of actions and primitives, and do not include functional motions used during RTP and RTG. Similarly, Kadu and Kuo [22] discuss action recognition, when instead activities are considered. To avoid confusion, this work follows Schambra et al. [23]'s UE functional motion hierarchy (see Table 1).

Activity [22], [27] and action [20], [28] recognition methods currently do not identify specific functional UE motion primitives, such as reaching and targeting, which are needed for some kinematics analyses. Additionally, UEFAs often have the activities and actions clearly defined. An example where the action is defined is the Southampton Hand Assessment Procedure's (SHAP) door handle task, where the participant must rotate the door handle until it is open and then release the handle [29]. An example of a defined activity is the SHAP food cutting task which consists of actions involved in cutting food (e.g., pick up knife, cut food), which would require a method for segmenting the actions before doing primitive segmentation. Activity and action recognition methods also often require templates, training data sets, or sufficient data for pattern mining [20], which are challenging to curate given the diversity in tasks, individual strategies for performing those tasks, and movement pathologies. Therefore, the methods used in this study are relatively simple and drawn from research papers performing kinematic analyses on functional UE motion.

Algorithms have been proposed for motion primitive segmentation by identifying points where the trajectory changes direction (i.e., strokes) [24], [25] and by using fixed kinematic thresholds [10], [12] (see Lin et al. [20] for expanded review). However, discrete strokes and directional changes do not necessarily indicate the end of the reaching motion primitive in UE RTP and RTG movements, and fixed kinematic thresholds can result in over-segmentation [20]. Another approach is to segment trajectories based on a percentage of total movement. For example, Li et al. [26] accounted for differences in participant kinematics while transporting objects by selecting 50% of movement time as when the hand reached a target position.

More complex methods, relative to simple kinematic and percent-of-movement thresholds, have been proposed recently [3], [17]. Jackson et al. [17] propose a segmentation point identification algorithm that does not rely on kinematic thresholds, and instead uses the shape of the velocity profile to identify the segmentation point. Sakai et al. [3] propose a multi-step segmentation method that combines kinematic thresholding with the segmentation method from Jackson et al. [17].

B. EVALUATING PRIMITIVE SEGMENTATION PERFORMANCE

The unavailability of motion primitive ground truth labels from RTP and RTG tasks makes algorithm evaluation difficult. While activities and actions (i.e., gross movements) are easier to visually identify, no method for definitively differentiating the reaching and targeting motion primitives currently exist. Variations in neuromuscular coordination used for RTP and RTG movements across different object configurations or movement pathologies are challenges in understanding where the true segmentation point is, assuming one exists. Therefore, all existing approaches to segmentation and validation are approximations, including the common approach of visually inspecting segmentation results or comparing with recorded video [12], [17].

Consequently, it is difficult to objectively compare algorithms for segmenting the reaching and targeting motion primitives. We propose a method (see Section III-D.2) for indicating when segmentation results may need further review by comparing the segmented reach motion primitive velocity profile to the minimum jerk trajectory (MJT) velocity profile [30]. We also visualize trajectories based on the distribution of errors computed using the MJT. While the error computed from comparing the two velocity profiles have been used in the robotics literature [30], its usage for evaluating reaching and targeting motion primitive segmentation performance is novel.

III. METHODS AND PROCEDURES

A. SEGMENTATION METHODS

1) 50% OF MOVEMENT

Percent-of-movement thresholds have been used to roughly indicate when a hand reaches a target during a functional UE task [26]. Although this approach may work well for small data sets and clean data where different thresholds can be visually inspected, it is used in this paper to demonstrate that it does not generalize well on more challenging data.

2) KINEMATIC THRESHOLDING

Variations of the kinematic threshold exist for segmenting the reaching motion primitive [3], [11]. We use the kinematic threshold described in Sakai et al. [3] (II.A.ii.a and b), which is a component of the segmentation method proposed in [3]. Restated here, the segmentation point is after the peak velocity magnitude when either of the following are first satisfied:

- The velocity magnitude reaches 5% of the peak velocity magnitude.
- The velocity magnitude is less than 20% of the peak velocity magnitude, and the acceleration is non-negative for the first time.

3) JACKSON ET AL. [17] (UPDATED)

The segmentation method from Jackson et al. [17] exploits the well-established property that RTP and RTG motions often have an initial, higher speed movement that covers a large distance followed by slower, finer movement to interact with the point of interest [14], [31]. The updated method does the following:

1) Identify the part of the trajectory that is close to the target and contains the reaching and targeting motion primitives.
2) Using this trajectory subset, identify the segmentation location, which we refer to as the "shoulder" of the velocity profile. The term "shoulder" in this context refers to the curved portion of the velocity profile between the reaching deceleration and targeting period.

First, a relative displacement threshold is found for position trajectory $Tr = [\vec{r_0}, \ldots, \vec{r_g}]$. This threshold is used due to potential re-adjustments in the trajectory that could result in multiple local maxima and minima. This calculation differs from that proposed in [17] to address an edge case where the trajectory between r_0 and r_g are not approximately linear (e.g., due to an individual adjusting the endpoint). Additionally, the displacement threshold works well if the start and end points are relatively distant from each other, which is not always the case.

Given the start $\vec{r_0}$ and end $\vec{r_g}$ coordinates of the trajectory Tr, at every index i along Tr, the displacement d_i is:

$$d_i = ||(\vec{r_i} - \vec{r_g})|| \tag{1}$$

With d_i computed for all points i on Tr, identify the earliest point d_s along Tr where $d_s \geq d_g(1-\alpha)$ and where $1 \geq \alpha > 0$. In other words, if $\alpha = 0.2$, then d_s will be at least 80% of the distance between $\vec{r_g}$ and $\vec{r_0}$ away from $\vec{r_0}$.

The velocity along the curve is obtained by the differentiation of the position vectors. We use the velocity magnitudes $v_i = \|\dot{\vec{r_i}}\|$ to segment the curve T_r. First, we find $p = argmax\{v_s, \ldots, v_g\}$, the index of the maximal velocity in T_r. We then compute orthogonal distances $d_{n,i}$ of points (i, v_i) from the line connecting (p, v_p) and (g, v_g) using the equation for the shortest distance from a point to a line [32],

$$d_{n,i} = \frac{-((i-p)(v_g - v_p)\text{-}(v_i - v_p)(g-p))}{\sqrt{(v_g - v_p)^2 + (g-p)^2}}. \tag{2}$$

The point (f, v_f) satisfying $f = argmax_i\{d_{n,i}\}$ is the segmentation point for the velocity profile. A consequence of this formulation is that local maxima cannot be flagged as the segmentation point, which could happen with the version of the method described in [17].

4) JACKSON ET AL. [17] (UPDATED), NO DISPLACEMENT THRESHOLD

To evaluate the utility of the displacement threshold described in the previous subsection, we evaluate the segmentation performance of the Jackson et al. [17] method without the displacement threshold. This method finds the "shoulder" in the velocity profile after the point indicating the peak velocity magnitude, regardless of distance to the grasp location.

5) SAKAI ET AL. [3]

The method proposed by Sakai et al. [3] was used for segmenting the targeting and reaching primitives in pen-point trajectories as part of an assessment that required participants to connect multiple dots on the surface of a digital tablet. It is unknown whether this method generalizes beyond pen-point trajectories, which is partially why it is included in this evaluation. This method combines kinematic thresholding (section III-A.2) and the Jackson et al. [17] method (section III-A.3), in addition to checks to verify whether the length of the trajectory is sufficiently long for applying the method from Jackson et al. [17]. The implementation of Sakai et al. [3] evaluated in this paper uses the updated version of the Jackson et al. [17] method, as described in section III-A.3.

The Sakai et al. [3] method follows multiple stages for segmentation, which we briefly describe. First, the method checks whether the velocity profile has a sufficiently long targeting period before using the Jackson et al. [17] segmentation method. This is done because the Jackson et al. [17] method can return improper results if the velocity profile does not have a sufficiently long tail after reaching. If the velocity profile is too short, where the length depends on a parameter c_1 (i.e., higher values for c_1 increase the targeting primitive length requirement), then the segmentation point is the earliest point after the peak velocity point at which either (1) the velocity magnitude reaches 5% of the peak velocity or (2) the velocity magnitude is less than q% of the peak velocity and the acceleration becomes non-negative for the first time [3]. The latter option (2) is a zero crossing method, where the crossing from negative to positive acceleration values indicates a local minimum in the velocity profile. If the velocity profile is determined to be sufficiently long for segmentation using the Jackson et al. [17] method, then segmentation is done using a portion of the trajectory that is within the length set by a parameter c_2, where the length found by using c_1 is less than c_2. A result of this is that the velocity profile could be shortened before the Jackson et al. [17] method is used, which the Jackson et al. [17] method does not do itself. Following Sakai et al. [3], $q = 20\%$, $c_1 = 2$, and $c_2 = 3$.

B. SEGMENTATION METHODOLOGICAL ASSUMPTIONS

Motion primitive segmentation methods make some assumptions about the kinematics being analyzed. The implementation of the methods considered in this paper assume the following criteria to be met or that the input data demonstrate these characteristics:

1) The trajectory must represent reaching towards one point of interest and must terminate once the point is touched for RTP and grasped for RTG tasks.
2) The trajectory has a reaching motion primitive followed by a targeting primitive.

For Condition 1, if there are multiple targets then the trajectories between each target must be pre-segmented.

Regarding Condition 2, depending on the assessment and movement quality, some trajectories may have no obvious targeting period (e.g., Fig. 7.D). These cases likely have overlapping reaching and targeting primitives. The methods considered in this paper identify a single point and do not capture this overlap, although some kinematic analyses may want to capture this overlap and will therefore require different segmentation method implementations. No method currently exists for automatically detecting where reaching transitions to targeting across a variety of functional tasks, although visual inspection of the kinematic data will help indicate when the targeting period begins. Additionally, a trajectory could include more than one reaching and targeting motion primitive (e.g., when the endpoint is adjusted while attempting to grasp an item).

C. EVALUATION DATA

The segmentation algorithms are evaluated on two data sets which involve RTP and RTG movements. The trajectories are preprocessed to include movement towards one point of interest at a time. Although the two data sets include transporting an object after grasp, our analysis focuses specifically on the reaching and targeting primitives before grasp.

1) HAPTIC VIRTUAL ENVIRONMENT

This George Mason University IRB-approved experiment (#477548, approved Jan. 29, 2014, informed consent obtained from participants) required participants to perform a simulated workbench clearing task within a haptic virtual environment (HVE) [17] (see Fig. 1), which involved grasping and mounting six tools on a pegboard. A convenience sample was used, with an invitation to participate by word of mouth, resulting in twenty-one non-disabled participants, aged 18 to 30, and comprising fourteen males and seven females. This task combines RTP and RTG movements because the manipulator is a single point in three-dimensional virtual space. The participant-manipulated stylus position, acceleration, and rotation were captured at 30 Hz and in three dimensions by the haptic device. Velocity magnitude was filtered using the fifth-order low-pass Butterworth filter with a cutoff frequency of 10 Hz. Trajectories were segmented into actions representing reaching to a tool and mounting a tool based on events recorded by the HVE software (i.e., the system tracks when objects are grasped and released). The mounting action is excluded from our analysis. The reaching actions, which contain reaching and targeting primitives, were used as input to the segmentation algorithms evaluated in this paper. Each participant performed three trials and a total of 305 trajectories were analyzed. $\alpha = 0.4$ for segmentation methods (3) and (5). The choice of α does not appear to drastically impact segmentation results if $\alpha \leq 0.5$, although it could be useful to test a few values for one's particular application.

2) THE TARGETED BOX AND BLOCKS TEST

For this George Mason University IRB-approved experiment (#492701, approved Oct. 24, 2013, informed consent obtained from participants), optical motion capture data were collected from three female participants, aged 22 to 29, performing the targeted Box and Blocks Test (tBBT) [1]. Two of the participants were non-disabled and the third participant performed the task using clinically-prescribed below-elbow myoelectric prostheses (see Fig. 2). The tBBT requires the individual to move wooden blocks over a partition in a predefined order and at predefined locations (see Fig. 2) and represents a RTG task when reaching to grasp a block. This assessment and variations of it are commonly used to assess functional UE movement quality for rehabilitation.

Trajectories were segmented into actions representing reaching for a block and transporting a block, although the transportation action is excluded from our analysis. Action segmentation was performed using heuristics and visual inspection of the data. This step is not the primary focus of this work, although an automated action segmentation method for tBBT could be future work. Specifically, our approach was to identify the velocity profile peaks using thresholds, followed by identifying local minima indicating a grasp or release, comparing the kinematics to the recorded video for context, and ensuring the correct number of actions were found. The segmented RTG actions are then used as input to the segmentation algorithms evaluated in this paper. Raw position information, captured from a marker on the wrist, is processed by removing spikes along each dimension (i.e., x, y, z), interpolating gaps due to marker occlusion via cubic spine interpolation, and applying the fifth-order low-pass Butterworth filter. We used $\alpha = 0.4$ for segmentation methods (3) and (5).

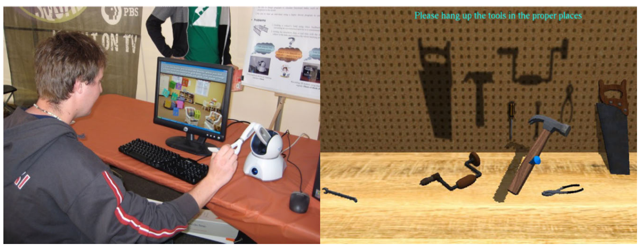

FIGURE 1. **(left) participant interfacing with a Touch™ haptic device used in our study and (right) the workbench clearing task.**

D. EVALUATION APPROACH

Ground truth labels for the segment location between the reaching and targeting motion primitives are not available due to there being no method currently for identifying the true segmentation point between reaching and targeting. While it would have been easier to simply label all trajectories where an acceptable label location would be, the segmentation workflow movement scientists use is to first apply a segmentation algorithm (e.g., a kinematic threshold) followed by visual inspection [11]. This is a result of difficulty in visually identifying from a graph the point at which, for example, the velocity magnitude reaches 5% of the peak velocity, necessitating some method to at least cue the rater's attention. Therefore, a few approaches are used to evaluate the segmentation results.

1) EXPERT EVALUATION OF SEGMENTATION RESULTS

Segmentation results for all methods were assessed and assigned one of the following labels as part of this analysis:

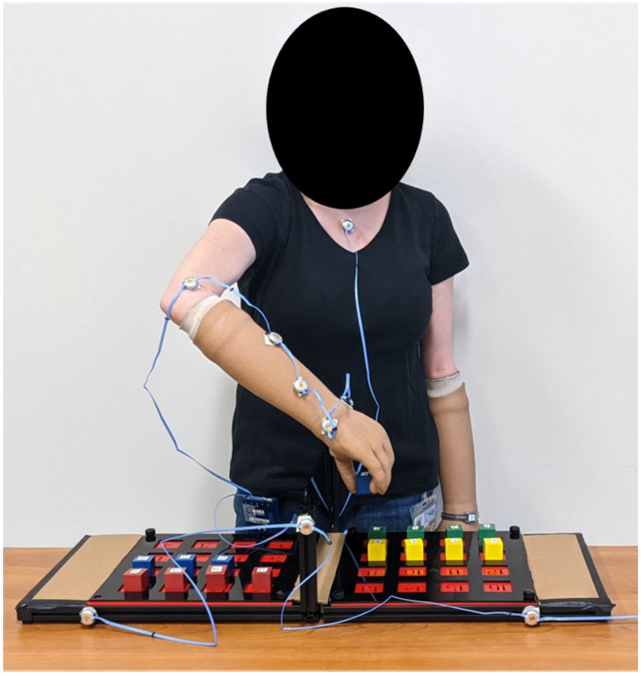

FIGURE 2. A participant performing the targeted Box and Blocks Test [1], outfitted with wired active optical motion capture markers.

- **Acceptable**: Segmentation result is acceptable; a human evaluator would likely not modify this result.
- **Questionable**: Segmentation result is questionable; a human evaluator would likely modify this result.
- **Ambiguous**: The trajectory and associated kinematics are sufficiently unclear that a human evaluator likely would not be able to segment or use the data for analysis. A trajectory labeled as ambiguous maintains that ambiguous label across all methods.

Note that these labels allow for variation in interpretations. This is intentional, as there is no method currently to identify the true point which separates reaching from targeting. For example, an acceptable segmentation result can take a range of points along the trajectory.

Two domain experts (S. Engdahl and A. Santago) labeled all segmentation results for the HVE data set on the kinematic thresholding method ($n = 305$) to establish labeling criteria and to better understand the challenges associated with evaluating segmentation results. To minimize bias, the domain experts did not know which method was used to perform the segmentation. For each trajectory, three plots indicating position, displacement, and the velocity profile overlaid with the segmentation result were used for evaluation, as is shown in the first column of Fig. 5. The raters had an initial percent agreement of 66.8% and a Cohen's kappa statistic of 0.18, indicating that it is challenging to have consistent evaluations across raters for segmentation results. Due to the inherent subjectivity in segmenting RTP and RTG motions, a follow-on meeting was held to further establish agreement on labeling criteria, resulting in an agreement on all except one segmentation result. These labeling criteria and labeled segmentation results were then used by a non-domain expert (K. Jackson) to label the remaining segmentation results for the HVE and tBBT data sets.

2) INDICATOR OF QUESTIONABLE SEGMENTATION RESULTS

Motion primitive segmentation results for the HVE and tBBT data sets are additionally evaluated by computing the mean absolute error (MAE) for each segmented trajectory's reaching primitive compared to the straight-path MJT velocity profile, referred to as MJT MAE. This is one of this work's novel contributions, where MJT MAE can be used as an indicator of potentially questionable segmentation issues that may need review by a domain expert.

The MJT velocity profile is smooth and has a unimodal bell-shape, which the segmented reaching velocity profile is expected to approximately follow. Higher MJT MAE values indicate a reaching velocity profile that deviates from a unimodal shape (see examples in Fig. 3), which could be due to an incorrect segmentation result or a complex reach that has multiple peaks and troughs.

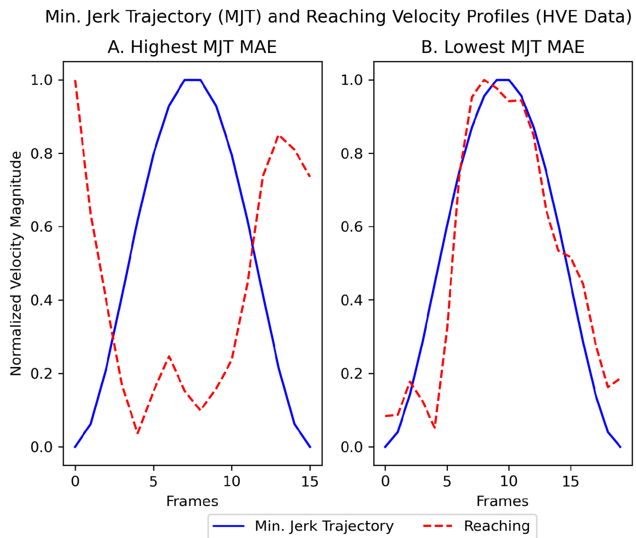

FIGURE 3. Two segmented reaching subsequences normalized velocity profiles from the HVE data set corresponding to the highest and lowest minimum jerk trajectory (MJT) mean absolute errors (MAE). The reaching segments are plotted with the normalized MJT velocity profile used to compute the MJT MAE. These two examples correspond to trajectories A and D in Fig. 5, respectively.

The straight-path MJT velocity magnitude time series is defined as:

$$V_{jerk} = \dot{x}(t) = x_f(30\tau^4 - 60\tau^3 + 30\tau^2) \tag{3}$$

where τ is normalized time equal to t/t_f and $0 \leq \tau \leq 1$, t_f represents the total duration of the reaching motion primitive, and x_f is the final position of the reaching motion primitive [30].

The segmented and MJT velocity magnitude values are normalized to be within the interval [0, 1] to evaluate the

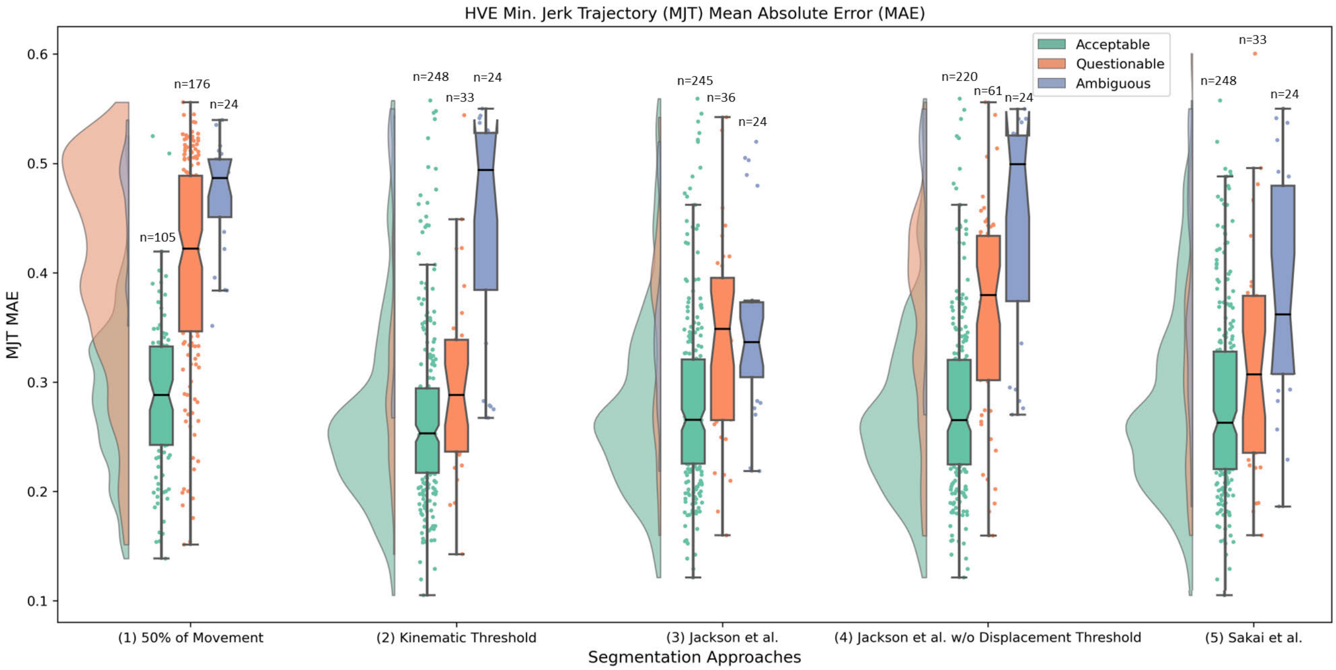

FIGURE 4. Rain cloud plots [33] depicting MJT MAE values for the five segmentation methods described in section III-A. The "clouds" to the left of the notch plots indicate the density of points for each label group, in the order of: acceptable, questionable, and ambiguous. The points represent each segmentation result's MJT MAE value based on a particular segmentation method.

shape of the profiles. The MAE for each segmented reaching motion primitive is calculated between all indices of the segmented normalized velocity profile $[v_0, \ldots, v_f]$ and the normalized MJT velocity profile:

$$\frac{(\sum_{i=1}^{t_f} |\frac{\dot{x}(i)}{V_{jerk,peak}} - \frac{v_i}{v_p}|)}{t_f} \tag{4}$$

where the denominator t_f is used because the numerator alone will generally result in higher values as t_f increases.

The MJT MAE for each trajectory in the HVE data set are reported as rain cloud plots [33] across all segmentation methods considered in this paper. Segmentation results for the trajectory with the highest, 75th percentile, 25th percentile, and lowest errors are visualized (see Figs. 5 and 7).

IV. RESULTS

A. HVE WORKBENCH CLEARING

The distribution of labels and MJT MAE values for all 305 trajectories analyzed are reported in Fig. 4. Example segmentation results for the Sakai et al. [3] method are visualized in Fig. 5 for different percentiles based on the distribution of MJT MAE values for the Sakai et al. [3] method depicted in Fig. 4.

The (1) 50% of movement and (3) Jackson et al. [17] (without the displacement threshold) methods did not perform as well as the other three methods. Methods (2), (4), and (5) performed similarly, although each method was able to acceptably segment some trajectories that were questionably segmented by the other two methods (i.e., no one method was able to handle all cases). One pattern observed in twenty-nine instances was that the (2) kinematic thresholding method had a tendency to segment slightly earlier compared to methods (3) and (5). However, these were still considered within the acceptable range given uncertainty about what point along the trajectory truly represents the transition from reaching to targeting.

With regards to the distribution of MJT MAE values, a general pattern is apparent in Fig. 4 that ambiguous and questionable segmentation results have a tendency to have higher MJT MAE values.

B. THE TARGETED BOX AND BLOCKS TEST

The tBBT segmentation performance results are in Fig. 6. Example trajectories from tBBT based on the distribution of MJT MAE values using the Sakai et al. [3] segmentation method are in Fig. 7. The results in Fig. 6 suggest that, like the HVE data, methods (2), (4), and (5) performed similarly on the tBBT data. While the segmentation results for methods (2), (4), and (5) were mostly labeled as acceptable, methods (2) and (5) tended to segment near the end of the trajectory (see Fig. 7 for examples). On the other hand, method (2) tended to segment earlier, usually at the "shoulder" of the velocity profile (e.g., approximately frame 65 of trajectory B in Fig. 7; additional examples are available at [19]).

V. DISCUSSION

A. WHICH SEGMENTATION METHOD IS BEST?

The (2) kinematic thresholding, (3) Jackson et al. [17], and (5) Sakai et al. [3] segmentation methods all performed similarly well. The HVE data presented complex trajectories and

IEEE Journal of Translational Engineering in Health and Medicine

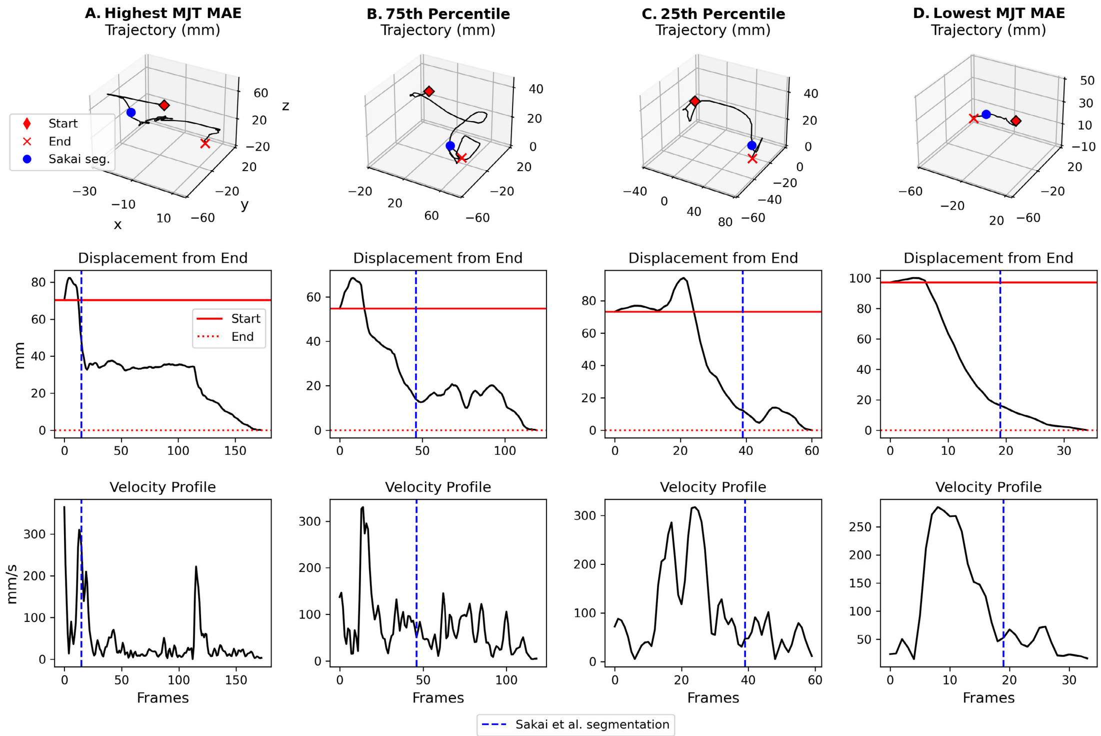

FIGURE 5. Workbench Clearing in Haptic Virtual Environment Visualized trajectories selected based on the distribution of MJT MAEs shown in Fig. 4's "Sakai et al." column. Each column corresponds to a trajectory from the HVE workbench clearing data set, with the second and third rows displaying results from the segmentation algorithm proposed by Sakai et al. [3]. The top row displays raw position trajectories, where the red diamond indicates the starting position, the blue circle indicates the location where the reaching and targeting primitives were segmented, and the red "X" indicates where the object was grasped. The second row shows where segmentation occurs along the displacement time series from the grasp location. The bottom row shows where segmentation occurs along the velocity profile.

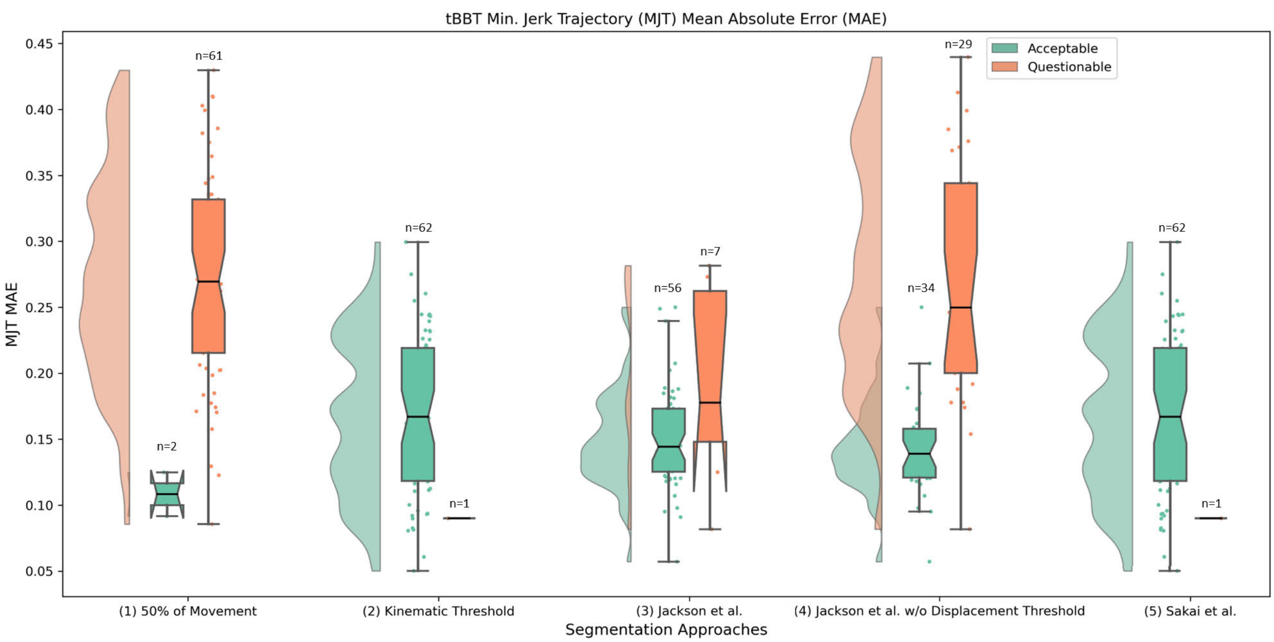

FIGURE 6. Targeted Box and Blocks Test MJT MAE values displayed for tBBT data across segmentation methods, as is done in Fig. 4. The notch plot order, from left to right, for each method are acceptable and questionable.

kinematics, which would be difficult for even a human evaluator to segment. These methods also performed similarly well on the tBBT data, which were generally less complex than the HVE data. The (5) Sakai et al. [3] method incorporates methods (2) and (3) into their method, and is therefore likely the most robust out of those considered in this work.

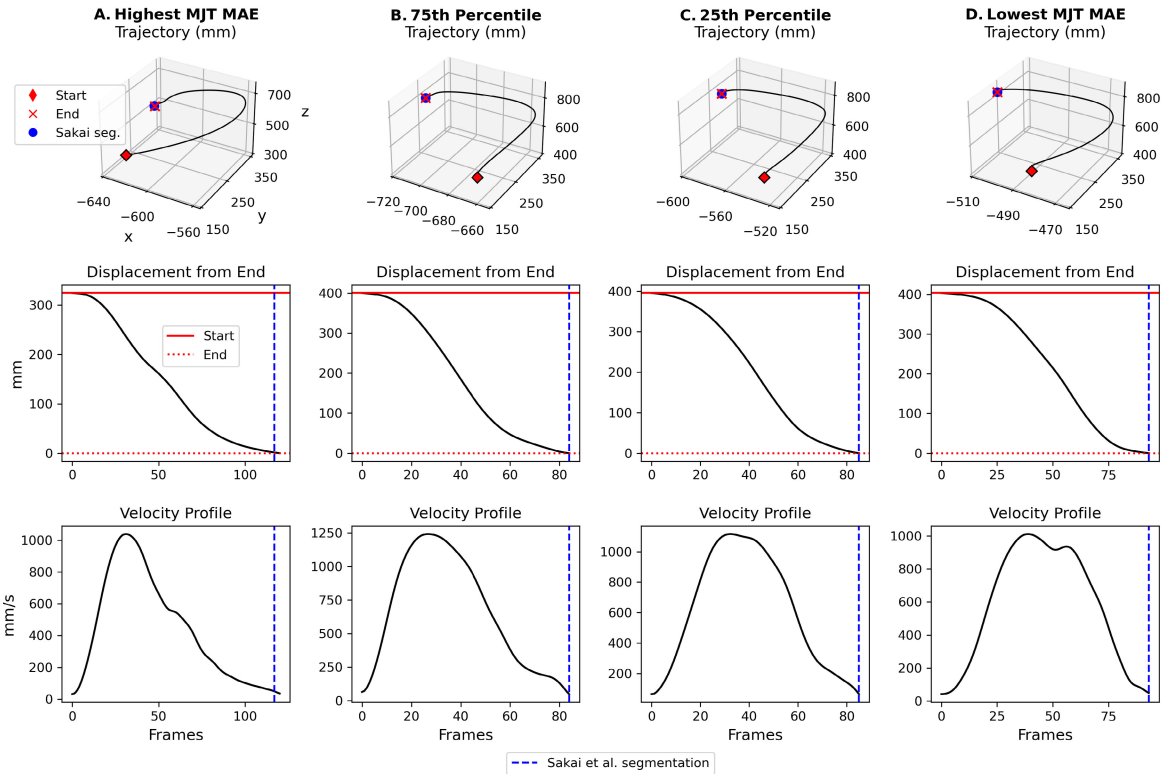

FIGURE 7. **Targeted Box and Blocks Test Four trajectories from the tBBT data visualized based on the distribution of MJT MAE values, as was done in Fig. 5.**

However, it is not clear what method is best for all applications given challenges with evaluating segmentation performance. Additionally, we use two data sets consisting of data from healthy, skilled individuals, where the tasks are relatively simple. Our HVE and tBBT also had an unbalanced number of participants (twenty-one and three, respectively), so the results could be biased towards motions in simulated environments. These data are not representative of all populations and tasks that clinicians work with, where tasks requiring greater precision may change the kinematic characteristics even in healthy populations. Therefore, future work would include evaluating the segmentation approaches on data from additional UEFAs and from persons with movement pathologies or disabilities.

The specific analysis being done will influence which methods and thresholds to use. For example, the Sakai et al. [3] and kinematic thresholding methods used a 5% of peak velocity threshold, which worked well for the HVE data but resulted in possibly late segmentation results for the tBBT data (see Fig. 5 and Fig. 7). A higher threshold, such as 10%, may provide better results on the tBBT data we used, although this may not be the case for all data collected from tBBT.

Nearly all our tBBT data had short or non-obvious targeting periods. Skilled RTG movements in non-disabled populations have been shown to adhere to stereotyped kinematic patterns, including tight coupling or overlap between the reaching and targeting primitives [16], [31], [34], which is relevant to the second segmentation methodological assumption listed in section III-B stating targeting must follow reaching. Similarly, the tBBT trajectories from the participant with the below-elbow myoelectric prostheses resembled the non-disabled trajectories due to a high level of prosthesis experience (26 years), but the trajectories may look different in other participants with disabilities (e.g., reaching and grasping are typically decoupled in UE prosthesis users [34]). Future work includes developing and evaluating segmentation methods that allow for overlapping motion primitives (i.e., where more than one segmentation point is provided between two primitives, which is in contrast with the methods considered in this study that provide only a single segmentation point), as there may be kinematic analyses that would benefit from more precise motion primitive segmentations.

B. SEGMENTATION PERFORMANCE EVALUATION

The most challenging part of our analysis was determining how to evaluate the segmentation methods. In the activity and action recognition literature, data sets typically come with labels that researchers can use to evaluate their methods. We did not have these for our data, nor does there exist a method currently for identifying the true location where reaching transitions to targeting, assuming it does exist. Our approach to evaluating segmentation results via domain expert review raised issues that could be addressed in future works.

As indicated in section III-D.1, initial agreement between our expert raters was not high for the HVE data. Additionally, the initial low agreement amongst our raters is partially due to the three label classes (i.e., acceptable, questionable, ambiguous) being vague, which was intentional. The consensus meeting resulted in agreement on all except one segmentation result, primarily due to one rater being more strict than the other on what was considered acceptable; however, other raters will likely have different opinions on the criteria for the evaluation labels. For this reason, the labels and segmentation results are available at [19]. We omit a statistical analysis comparing the varied segmentation methods and the MJT MAE distributions across the three labels due to the challenges associated with evaluation.

As kinematic analyses that require segmentation of the reaching and targeting motion primitives are used more frequently, researchers will need a method to more objectively evaluate segmentation methods. Specifically, more concrete labeling criteria and a better understanding of intra- and inter-rater variability associated with evaluating segmentation results will be needed. Improved evaluation approaches will support segmentation algorithm development and the potential usage of template and learning algorithms. While the targeting period may be difficult to segment using only the kinematics of an endpoint (e.g., wrist), additional data sources may help. For example, muscle activity measured by electromyography (EMG) [35] or sonomyography (SMG) [36] may assist with identifying the segmentation point between reaching and targeting when combined with kinematics. While these modalities have been used in other contexts such as detecting grasp intention (e.g., [38]) or movement onset [37], it is unclear if they could be used to identify the targeting primitive. If so, this approach could be used to help create ground truth data for developing and evaluating segmentation approaches that use kinematics alone. Similarly, detecting muscle activity in the hand or tracking finger aperture via motion capture could be useful for better delineating when grasping begins [16], [31], [34], [38].

C. MIN. JERK TRAJECTORY MEAN ABSOLUTE ERROR

The results suggest MJT MAE is a useful method for evaluating segmentation performance. In Fig. 5, higher MJT MAE values are associated with segmented reaching primitives that do not follow a clean bell-curve shape. For example, trajectory A in Fig. 5 has the highest MJT MAE score in the data set, indicating a questionable segmentation result that may need to be corrected. Note that MJT MAE should not be considered a definitive measure of segmentation performance given that reaching motions may have non-bell-shaped velocity profiles, which the MJT MAE penalizes.

Wider usage and automation of kinematic analyses that require segmentation of the reaching and targeting motion primitives would benefit from using the MJT MAE to identify segmentation results that need closer review. This approach is aligned with the segmentation workflow already used by domain experts (e.g., apply segmentation algorithms and then manually review [11]). However, the MJT MAE has limitations. Based on the results in Fig. 4, there is sufficient overlap over the MJT MAE distributions of the three evaluation labels that improvements to this approach are needed for identifying segmentation results that actually need additional review by a human evaluator (i.e., true positives). Establishing the range of acceptable MJT MAE values, or from a similar method, for specific UEFAs and populations is an area of future work.

D. CHALLENGING SEGMENTATION CASES

Our analysis of the HVE and tBBT data sets identified multiple kinematic profiles that the segmentation methods did not do well on. Examples of each are in the supplemental materials [19], which are briefly described below.

1) SLOW REACHING

Although the segmentation methods considered in this paper assume that velocity profiles exhibit the characteristic unimodal bell-shape, slow reaching motions did not demonstrate this. Segmentation results for these trajectories were therefore varied. Developing methods that can address this edge case are likely necessary, as slow reaching can be a viable strategy used by individuals for RTG and RTP motions.

2) RE-ADJUSTMENTS DURING REACH AND TARGETING

Particularly in the HVE data, participants would sometimes adjust the endpoint during reaching or targeting. If re-adjustments occurred during reaching, our evaluation approach was to consider that part of reaching. However, it was difficult to determine the segmentation point between reaching and targeting when re-adjustments occurred during targeting, where the endpoint was briefly moved away from the object and a small reach was used to move back to the object. Whether to include the re-adjustments during targeting as part of the segmented targeting time series will likely depend on the kinematic analysis being performed (e.g., it may be useful to include if assessing how much difficulty an individual is having in targeting an object for grasp or pointing).

3) MORE THAN ONE OBJECT TARGETED

There were some instances in the HVE data where a participant would reach and target for one object, then move to another object. As the HVE workbench clearing task does not specify an order of tools, this was valid for the test but made segmentation more challenging. Reaching and targeting more than one object is a violation of the assumption of the segmentation methods considered in this work, and would require additional processing. This issue may not be prevalent in some UEFAs where the order of objects to be grasped or pointed to is pre-defined (e.g., in tBBT). However, some UEFAs do not specify order (e.g., Box and Blocks Test). Careful observation of the participant and video recordings help provide context when evaluating segmentation results.

E. ENSEMBLE SEGMENTATION METHODS

Of the three best performing segmentation methods considered in this paper, each method acceptably segmented some trajectories where the other two failed. While it may be possible to develop one segmentation method that is robust to all applications, a promising direction is to use an ensemble of segmentation methods. Inspired by ensemble methods from the data mining and machine learning literature [39], ensembling segmentation outputs to all ''vote'' on a segmentation point would leverage the best of each segmentation method. Ensembles have demonstrated state-of-the-art performance on multiple problems in machine learning [39]. For example, ensemble learning has been used for activity recognition from wearable sensors [40]. One possible implementation of ensembling segmentation methods could be to have the final segmentation location be the average of all the segmentation locations from the multiple segmentation methods used in the ensemble, which is an approach used for fusing outputs of regression models [39]. Furthermore, high disagreement amongst the ensembled methods could signal a challenging trajectory that requires further review by a human evaluator, similar to what has been proposed for quantifying the uncertainty of deep learning models [41], providing an alternative or complement to the proposed MJT MAE.

VI. CONCLUSION

This paper provides an analysis of segmenting reaching and targeting motion primitives for RTP and RTG motions. The results suggest that recently proposed methods for segmentation do reasonably well, having been tested on HVE and tBBT data. However, our understanding of what indicates the precise point between reaching and targeting motion primitives from kinematics alone remains limited. A better understanding of where reaching transitions to targeting will help create ground truth data sets for more objective evaluation, along with enabling the development of learning-based methods that require training data. This work also proposes the MJT MAE to evaluate segmentation performance and indicate potentially questionable segmentation results, which could be incorporated into a segmentation workflow used by researchers and, eventually, clinicians. Mechanisms like the MJT MAE which flag questionable results may also help mitigate risks associated with domain experts using erroneous results in their decision making.

ACKNOWLEDGMENT

Approved for Public Release; Distribution Unlimited. Public Release Case Number 23-1479. Anthony C. Santago II affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions, or viewpoints expressed by the author.

REFERENCES

[1] K. Kontson, I. Marcus, B. Myklebust, and E. Civillico, ''Targeted box and blocks test: Normative data and comparison to standard tests,'' *PLoS ONE*, vol. 12, no. 5, May 2017, Art. no. e0177965.

[2] D. Impedovo and G. Pirlo, ''Dynamic handwriting analysis for the assessment of neurodegenerative diseases: A pattern recognition perspective,'' *IEEE Rev. Biomed. Eng.*, vol. 12, pp. 209–220, 2019.

[3] H. Sakai et al., ''Pen-point trajectory analysis during trail making test based on a time base generator model,'' in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 6215–6219.

[4] C. Gummesson, M. M. Ward, and I. Atroshi, ''The shortened disabilities of the arm, shoulder and hand questionnaire (Quick DASH): Validity and reliability based on responses within the full-length DASH,'' *BMC Musculoskeletal Disorders*, vol. 7, no. 1, pp. 1–7, Dec. 2006.

[5] B. M. Chesworth et al., ''Reliability and validity of two versions of the upper extremity functional index,'' *Physiotherapy Canada*, vol. 66, no. 3, pp. 243–253, Jul. 2014.

[6] (Apr. 4, 2016). *Action Research Arm Test*. Shirley Ryan AbilityLab. Accessed: Jan. 3, 2022. [Online]. Available: https://www.sralab.org/rehabilitation-measures/action-research-arm-test

[7] S. Wang et al., ''Evaluation of performance-based outcome measures for the upper limb: A comprehensive narrative review,'' *PM R*, vol. 10, no. 9, pp. 951–962, Sep. 2018.

[8] A. de los Reyes-Guzmán, I. Dimbwadyo-Terrer, F. Trincado-Alonso, F. Monasterio-Huelin, D. Torricelli, and A. Gil-Agudo, ''Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: A review,'' *Clin. Biomechanics*, vol. 29, no. 7, pp. 719–727, Aug. 2014.

[9] G. Kwakkel et al., ''Standardized measurement of quality of upper limb movement after stroke: Consensus-based core recommendations from the second stroke recovery and rehabilitation roundtable,'' *Neurorehabilitation Neural Repair*, vol. 33, no. 11, pp. 951–958, Nov. 2019.

[10] J. Cowley, L. Resnik, J. Wilken, L. S. Walters, and D. Gates, ''Movement quality of conventional prostheses and the DEKA arm during everyday tasks,'' *Prosthetics Orthotics Int.*, vol. 41, no. 1, pp. 33–40, 2017.

[11] S. M. Engdahl and D. H. Gates, ''Differences in quality of movements made with body-powered and myoelectric prostheses during activities of daily living,'' *Clin. Biomechanics*, vol. 84, Apr. 2021, Art. no. 105311.

[12] S. M. Engdahl and D. H. Gates, ''Reliability of upper limb movement quality metrics during everyday tasks,'' *Gait Posture*, vol. 71, pp. 253–260, Jun. 2019.

[13] S. Balasubramanian, A. Melendez-Calderon, A. Roby-Brami, and E. Burdet, ''On the analysis of movement smoothness,'' *J. NeuroEngineering Rehabil.*, vol. 12, no. 1, p. 112, Dec. 2015.

[14] E. D. Graham and C. L. MacKenzie, ''Physical versus virtual pointing,'' in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 1996, pp. 292–299.

[15] P. van Vliet, T. A. Pelton, K. L. Hollands, L. Carey, and A. M. Wing, ''Neuroscience findings on coordination of reaching to grasp an object: Implications for research,'' *Neurorehabilitation Neural Repair*, vol. 27, no. 7, pp. 622–635, Sep. 2013.

[16] H. Bouwsema, P. Kyberd, W. Hill, C. Van Der Sluis, and R. Bongers, ''Determining skill level in myoelectric prosthesis use with multiple outcome measures,'' *J. Rehabil. Res. Develop.*, vol. 49, no. 9, pp. 1331–1348, 2012.

[17] K. Jackson, Z. Duric, S. Engdahl, and L. Gerber, ''Characterizing functional upper extremity movement in haptic virtual environments,'' in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 3166–3169.

[18] K. L. Jackson, Z. Durić, S. M. Engdahl, A. C. Santago II, S. DeStefano, and L. H. Gerber, ''Computer-assisted approaches for measuring, segmenting, and analyzing functional upper extremity movement: A narrative review of the current state, limitations, and future directions,'' *Frontiers Rehabil. Sci.*, vol. 4, Apr. 2023, Art. no. 1130847.

[19] K. Jackson, Z. Duric, S. Engdahl, S. Sikdar, and L. Gerber, ''Haptics and targeted box and blocks data for evaluating reach and grasp segmentation methods,'' Distributed by Zenodo. Accessed: Apr. 21, 2023, doi: 10.5281/zenodo.7851399.

[20] J. F. Lin, M. Karg, and D. Kulic, ''Movement primitive segmentation for human motion modeling: A framework for analysis,'' *IEEE Trans. Hum.-Mach. Syst.*, vol. 46, no. 3, pp. 325–339, Jun. 2016.

[21] J. Lin, ''Temporal segmentation of human motion for rehabilitation,'' Doctoral dissertation, Dept. Elect. Comput. Eng., Univ. Waterloo, Waterloo, ONT, Canada, 2016.

[22] H. Kadu and C.-C. J. Kuo, ''Automatic human mocap data classification,'' *IEEE Trans. Multimedia*, vol. 16, no. 8, pp. 2191–2202, Dec. 2014.

[23] H. M. Schambra, A. Parnandi, N. G. Pandit, J. Uddin, A. Wirtanen, and D. M. Nilsen, ''A taxonomy of functional upper extremity motion,'' *Frontiers Neurol.*, vol. 10, p. 857, Aug. 2019.

[24] D. Popa, G. Simion, V. Gui, and M. Otesteanu, ''Real time trajectory based hand gesture recognition,'' *WSEAS Trans. Inf. Sci. Appl.*, vol. 5, no. 4, pp. 532–546, 2008.

[25] W. W. Kong and S. Ranganath, ''Automatic hand trajectory segmentation and phoneme transcription for sign language,'' in *Proc. 8th IEEE Int. Conf. Autom. Face Gesture Recognit.*, Sep. 2008, pp. 1–6.

[26] X. Li, A. C. Santago, M. E. Vidt, and K. R. Saul, ''Analysis of effects of loading and postural demands on upper limb reaching in older adults using statistical parametric mapping,'' *J. Biomechanics*, vol. 49, no. 13, pp. 2806–2816, Sep. 2016.

[27] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, ''Deep learning for sensor-based activity recognition: A survey,'' *Pattern Recognit. Lett.*, vol. 119, pp. 3–11, Mar. 2019.

[28] S. Herath, M. Harandi, and F. Porikli, ''Going deeper into action recognition: A survey,'' *Image Vis. Comput.*, vol. 60, pp. 4–21, Apr. 2017.

[29] C. M. Light, P. H. Chappell, and P. J. Kyberd, ''Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: Normative data, reliability, and validity,'' *Arch. Phys. Med. Rehabil.*, vol. 83, no. 6, pp. 776–783, Jun. 2002.

[30] A.-N. Sharkawy, P. N. Koustoumpardis, and N. Aspragathos, ''A neural network-based approach for variable admittance control in human–robot cooperation: Online adjustment of the virtual inertia,'' *Intell. Service Robot.*, vol. 13, no. 4, pp. 495–519, Oct. 2020.

[31] M. Jeannerod, ''The timing of natural prehension movements,'' *J. Motor Behav.*, vol. 16, no. 3, pp. 235–254, Sep. 1984.

[32] D. Sunday, ''Lines,'' in *Practical Geometry Algorithms: With C++ Code*. Independently Published, 2021, p. 34. [Online]. Available: https://www.amazon.com/Practical-Geometry-Algorithms-C-Code/dp/B094T8MVJP/ref=sr_1_1?crid=NN1J5U2I2NQ3&keywords=practical+geometry+algorithms&qid=1691203411&sprefix=practical+geometry+algorithms%2Caps%2C121&sr=8-1#detailBullets_feature_div

[33] M. Allen, D. Poggiali, K. Whitaker, T. R. Marshall, and R. A. Kievit, ''Raincloud plots: A multi-platform tool for robust data visualization,'' *Wellcome Open Res.*, vol. 4, p. 63, Apr. 2019.

[34] H. Bouwsema, C. K. V. der Sluis, and R. M. Bongers, ''Movement characteristics of upper extremity prostheses during basic goal-directed tasks,'' *Clin. Biomechanics*, vol. 25, no. 6, pp. 523–529, Jul. 2010.

[35] A. d'Avella, A. Portone, and F. Lacquaniti, ''Superposition and modulation of muscle synergies for reaching in response to a change in target location,'' *J. Neurophysiol.*, vol. 106, no. 6, pp. 2786–2812, 2011.

[36] A. S. Dhawan et al., ''Proprioceptive sonomyographic control: A novel method for intuitive and proportional control of multiple degrees-of-freedom for individuals with upper extremity limb loss,'' *Sci. Rep.*, vol. 9, no. 1, p. 9499, Jul. 2019.

[37] A. J. Tweedell, M. S. Tenan, and C. A. Haynes, ''Differences in muscle contraction onset as determined by ultrasound and electromyography,'' *Muscle Nerve*, vol. 59, no. 4, pp. 494–500, Apr. 2019.

[38] I. Batzianoulis, S. El-Khoury, E. Pirondini, M. Coscia, S. Micera, and A. Billard, ''EMG-based decoding of grasp gestures in reaching-to-grasping motions,'' *Robot. Auto. Syst.*, vol. 91, pp. 59–70, May 2017.

[39] O. Sagi and L. Rokach, ''Ensemble learning: A survey,'' *WIREs Data Mining Knowl. Discovery*, vol. 8, no. 4, p. e1249, Jul. 2018.

[40] J.-K. Min and S.-B. Cho, ''Activity recognition based on wearable sensors using selection/fusion hybrid ensemble,'' in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2011, pp. 1319–1324.

[41] M. Abdar et al., ''A review of uncertainty quantification in deep learning: Techniques, applications and challenges,'' *Inf. Fusion*, vol. 76, pp. 243–297, Dec. 2021.