Musculoskeletal Representation of a Large Repertoire of Hand Grasping Actions in Primates

S. Schaffelhofer, M. Sartori, H. Scherberger, and D. Farina

Abstract—Reach-to-grasp tasks have become popular paradigms for exploring the neural origin of hand and arm movements. This is typically investigated by correlating limb kinematic with electrophysiological signals from intracortical recordings. However, it has never been investigated whether reach and grasp movements could be well expressed in the muscle domain and whether this could bring improvements with respect to current joint domain-based task representations. In this study, we trained two macaque monkeys to grasp 50 different objects, which resulted in a high variability of hand configurations. A generic musculoskeletal model of the human upper extremity was scaled and morphed to match the specific anatomy of each individual animal. The primate-specific model was used to perform 3-D reach-to-grasp simulations driven by experimental upper limb kinematics derived from electromagnetic sensors. Simulations enabled extracting joint angles from 27 degrees of freedom and the instantaneous length of 50 musculotendon units. Results demonstrated both a more compact representation and a higher decoding capacity of grasping tasks when movements were expressed in the muscle kinematics domain than when expressed in the joint kinematics domain. Accessing musculoskeletal variables might improve our understanding of cortical hand-grasping areas coding, with implications in the development of prosthetics hands.

Index Terms—Arm, grasping, hand, musculoskeletal model, nonhuman primates, upper extremity.

I. INTRODUCTION

Evolutions in primates has led to an organization of large sets of muscles around multiple joints enabling complex movements in the upper extremity [1]. Reach-to-grasp tasks have become a popular paradigm for studying the mechanisms underlying movement generation in the nervous and musculoskeletal systems [2]–[7]. Due to the high level of implantation complexity and ethical implications, invasive recordings of neuromuscular activity are rare in humans [8], [9]. Nonhuman primates (NHP) (i.e., macaca mulatta in this study) became the preferred research model due to their strong analogies to humans at the nervous and motor systems levels [10].

During reach-to-grasp actions, the dynamics of the musculoskeletal system have been studied with a main focus on limb and joint kinematic variables, including: hand velocity, direction, and joint angular position [5], [11]–[13]. These studies demonstrated strong correlations between the motor/premotor cortex neural activity and such kinematics variables in NHPs. In this context, movement was recorded using optical or electromagnetic methods, based on nonanatomical models of finger, hand and arm segments [5], [14]–[16]. Previously proposed methodologies have never taken into account the kinematic behavior of the underlying musculotendon units (MTUs) spanning the hand and upper extremity joints during reach-to-grasp movements.

This study uses musculoskeletal modelling and simulation for deriving MTUs kinematic variables noninvasively from the intact primate. This allows generating information that could not be easily accessed by experimental tests alone. These include the MTU instantaneous length, contraction velocity, and moment arms developed around multiple degrees of freedom (DOFs) in the finger, hand, and arm segments [17]–[21]. Current musculoskeletal models of the macaque monkey upper extremity [17], [21] do not account for the large number of DOFs and MTUs in the primates hand and are not suitable for studying complex reach-to-grasp tasks. In this study, a comprehensive musculoskeletal model of the human arm and hand was used as basis for exploring the complex NHP musculoskeletal function [18]–[20], [22]. This choice is further supported by extensive studies consistently confirming strong anatomical similarities between human and monkey primates [23], [24]. Scaling and morphing methods were then used to adapt each anatomical segment and muscle in the generic human model to match the specific anatomy of each primate [25].

In this study we assess the potentials of the scaled, animal-specific model for studying a large repertoire of grasping tasks in two macaque monkeys. For this, we studied how well features extracted at the joint level and at the MTU level, respectively, could represent large repertoires of grasping actions. Although in humans and NHPs, the number of MTUs is by far larger than the number of DOFs in the skeletal joints, the benefit associated with this high dimensionality is not yet fully explored. In
Fig. 1. Behavioral task. (a) Two macaque monkeys were trained to grasp a wide range of objects presented on a PC-controlled turntable. During a recording session the animals wore an instrumented glove holding electromagnetic sensor coils for tracking finger, hand, and arm movements. The magnetic field generator was placed below the turntable (blue box). (b) In total, 48 objects of different shape and size were presented to the animals, including rings (pink), cubes (blue), balls (orange), horizontal cylinders (grey), boxes (green), vertical cylinders (beige) and special forms (yellow). (c) Additionally, the monkeys performed precision (red) and power grips (blue) on a handle.

In this study we hypothesize that the information extracted at the MTU level can produce more compact representations of a large repertoire of hand grasping actions with an associated improvement in grasping decoding capabilities than it would be possible by extracting information at the joint level, as done traditionally in the literature. This would emphasize on the inherent nonlinearity between the skeletal and muscular systems. This could play a crucial role for understanding the efficient neuromuscular mechanisms underlying the control of large biological structures (i.e., muscles), which enable complex dexterous hand movement.

II. METHODS

A. Experimental Task

Two macaque monkeys (macaca mulatta; animal M and Z, 7 kg male and 11 kg female, respectively) were trained to grasp a wide range of different objects of equal weight while wearing an instrumented glove to track their hand and arm motion [42] [see supplementary movie 1–2]. Animal care and all experimental procedures were conducted in accordance with German and European laws governing animal care and were in agreement with the Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research [26].

Objects were presented to the monkeys on a PC-controlled turntable in front of them [Fig. 1(a)]. In total, each animal grasped 48 objects of different shape and size [Fig. 1(b)] including spheres, rings, cubes, horizontal cylinders, boxes, vertical cylinders, and six other objects having more complex forms. The turntable could be replaced by a handle [Fig. 1(c)] that the animals grasped with a precision grip (touching the center of the handle with index finger and thumb) or a power grip (enclosure of the handle using digits). This resulted in a total of 50 different grasping conditions, which the monkeys repeated for at least ten times during each experimental session. All repetitions were performed in a controlled trial sequence in which the monkey had to release a switch at the resting position, grasp for and hold the randomly selected object or the handle for 0.5 s.

All digital signals affecting and monitoring the animal’s behavior were controlled by a custom-written software implemented in Labview Realtime (National instruments; time accuracy: 1 ms) and were recorded together with the hand and arm kinematics.

B. Tracking of Hand and Arm Movements

The musculoskeletal simulations presented in this study required comprehensive motion data describing finger, hand, and arm movements. To record this kinematics we used an instrumented glove equipped with seven electromagnetic sensors [Fig. 1(a)] that was previously reported in detail [14], [27]. In short, the method accessed the orientation and the position of seven electromagnetic sensors that were placed on top of the fingernails of digits and thumb, the hand's dorsum and the wrist of the hand. In this way the position of the fingertips ($\hat{t}^\text{tIP}$), distal interphalangeal joints ($\hat{t}^\text{DIP}$), metacarpal phalangeal joints ($\hat{t}^\text{MC}$), and the wrist (W) could be derived from the sensors ($f \in \{1, 2, \ldots, 5\}$). Exploiting the manually measured segment length and the anatomical restrictions of the hand allowed computationally estimating the position of the proximal interphalangeal joints ($\hat{t}^\text{PIP}$), elbow (E), and shoulder (S) at high accuracies [14]. This gave a total of 18 joint positions in 3-D space (experimental markers).

C. Preprocessing of Kinematics

In addition to the 18 joint centers, two more markers (E' and W') were defined to better constraint the simulation throughout the task and were defined as function of the existing joint positions. Whereas vectors $\hat{E}W$ and $\hat{S}E$ allowed describing the orientation of forearm and upper arm respectively, the additional two markers defined their rotation. The E’ marker was placed...
Fig. 2. Kinematic drive of the musculoskeletal model. (a) Instrumented glove allowed extracting the position of 25 experimental markers (illustrated in blue). (b) Corresponding model markers (red) were added to the musculoskeletal model, placed at the matching positions at the model's joint centers and fingertips. (c) Inverse kinematics allowed finding the posture of the musculoskeletal model, giving the least mean error between all experimental markers (blue) and model markers (red). Marker name syntax: \( k^j_f \) joint name; \( f \in \{1, 2, \ldots, 5\} \), 1 = thumb, 2 = index, 3 = middle, 4 = ring, 5 = little; \( k \) marker name: \( k \in \{e, m\} \), e = experimental marker, m = model marker.

orthogonal on the plane \( P_1 \), described by \( W, E, \) and \( S \) and was placed \( d = 20 \) mm distant from its origin \( E \)

\[
E' = E + d \, n_1
\]  

where \( n_1 \) was the unit vector normal to \( P_1 \) pointing to the side of the medial epicondyle of the humerus [Fig. 2(b)]. The marker \( W' \) was located orthogonal to the sagittal plane (\( P_2 \)) of the hand (described by the reference sensor at the back of the hand) 20 mm distant from its origin \( W \)

\[
W' = W + d \, n_2
\]

where \( n_2 \) was the unit vector normal to \( P_2 \) pointing to the radial side of the hand [Fig. 2(b)].

Markers \( E', E \) and \( S \) provided sufficient information to describe the rotation, elevation and adduction/abduction of the humerus at the shoulder when applied to the driven musculoskeletal model. The marker \( W' \) gave information about the pronation/supination of the hand, which was directly linked to the rotation of the lower arm, e.g., ulna and radius. The 3-D trajectories of the 25 experimental markers (18 joints, five fingertips and two additional help points) were transformed into the coordinate system of the musculoskeletal model and taken as its input. We refer to them as the experimental markers driving the model [Fig. 2(a)].

All kinematic data were recorded at a sampling rate of 100 Hz. Three-dimensional coordinates of each experimental marker were then exported to our musculoskeletal model realized in OpenSim, a software system allowing users to develop models of musculoskeletal structures and to create dynamic simulations of movement [25].

D. Musculoskeletal Model

1) Model Description: The musculoskeletal model employed in this study was adapted from an anatomically correct musculoskeletal model of the human upper extremity. This incorporated a total of 15 DOFs in the joints of the upper extremity, thumb and index segments (Table I) [18]. For this study, we extended the original musculoskeletal model to include four additional DOFs in each of the remaining fingers (middle, ring, little) as reported in Table I. In this, the axes of rotation of the joints in each finger were determined as the long axis of cylinders fit to the articular surfaces of the metacarpal and phalangeal bones as previously described [18]. Our proposed extended model had a total of 27 DOF in the hand and arm segments.

The model incorporated a total of 50 MTUs [18]. Each MTU in the model was defined by a set of points and surfaces attached to the underlying anatomical segments. The MTU attachments were determined from digitized muscle insertions and anatomical description [18]. This allowed generating anatomically accurate MTU paths over the range of motion of the considered 27 DOFs. The path of each MTU corresponded to the instantaneous MTU length (\( L^\text{MTU} \)) measured from origin to insertion and accounting for intermediate MTU-to-bone wrapping points and surfaces. All included MTUs are listed in [18].

2) Model Adaptation: For accurate simulations, we scaled each skeletal segment and MTU in the generic model to match the primate-specific anatomy for each specific animal using the 25 experimental marker positions [Fig. 2(a)] tracked by the instrumented glove. The tracking technique explores the manually measured segment length, anatomical constraints of hand and arm, and the accurate spatial and rotational information of 7 sensor coils (see [14, ch. 2.5] for detailed information). As described previously, measured and computationally derived experimental markers provided high accuracies [14] and were therefore selected for scaling.
TABLE I

<table>
<thead>
<tr>
<th>Joint</th>
<th>Degrees of Freedom (DOFs)</th>
<th>DOF Acronym</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulder</td>
<td>elevation, rotation, and adduction/abduction</td>
<td>shELE, shROT, and shABD</td>
</tr>
<tr>
<td>Elbow</td>
<td>flexion/extension</td>
<td>elbFE</td>
</tr>
<tr>
<td>Wrist</td>
<td>flexion/extension, deviation, and pronation/supination</td>
<td>wristFE, wristDEV, and wristPS</td>
</tr>
<tr>
<td>Thumb (I) Carpometacarpal</td>
<td>adduction/abduction and flexion/extension</td>
<td>cmcABD and cmcFE</td>
</tr>
<tr>
<td>Thumb (I) Metacarpophalangeal</td>
<td>flexion/extension</td>
<td>mcpFE</td>
</tr>
<tr>
<td>Thumb (II) Interphalangeal</td>
<td>flexion/extension</td>
<td>ipFE</td>
</tr>
<tr>
<td>Fingers (II-V) Metacarpophalangeal</td>
<td>adduction/abduction and flexion/extension</td>
<td>mcpABD</td>
</tr>
<tr>
<td>Fingers (II-V) Proximal Interphalangeal</td>
<td>flexion/extension</td>
<td>pipFE</td>
</tr>
<tr>
<td>Fingers (II-V) Distal Interphalangeal</td>
<td>flexion/extension</td>
<td>dipFE</td>
</tr>
</tbody>
</table>

Symbol (II-V) refers to the index (II), middle (i.e., III), ring (IV), and little (V) fingers. Thumb finger is referred to as the first finger (I).
for both spaces (Pearson correlation coefficient, Matlab function: corrcoef).

Additionally, cluster quality was expressed in both representations as the Dunn Index (D) aiming to identify dense and well-separated clusters [29]. The index is defined as the ratio between the minimal inter-cluster distance to maximal intra-cluster distance

\[ D = \min \left\{ \frac{d(c_i,c_j)}{\Delta_{\text{max}}} \mid 1 \leq i < j \leq K \right\} \tag{4} \]

where \( d(c_i,c_j) \) represents the distance between clusters \( i \) and \( j \) and \( \Delta_{\text{max}} \) the maximum inter-cluster distance from all \( K \) clusters.

4) Object Classification Accuracy: The large repertoire of objects did not only cause a rich diversity of grip types, but also contained intended redundancies (different objects that were grasped in a similar fashion). These similarities among the different objects were used to evaluate and compare the information content carried in the MTU and the DOF space. As a measure, we determined how accurately the grasped objects could be decoded from the extracted DOF angles or, alternatively, from the \( \ell^{\text{MTU}} \). For this purpose, we implemented a likelihood decoder based on the Bayesian theorem [30]

\[ p(c \mid k_i) = \frac{p(k_i \mid c) \cdot p(c)}{p(k_i)} \tag{5} \]

In a first step, the decoder was trained with a segment of the recorded data to acquire \( p(k_i \mid c) \), which is the probability of observing the kinematic variable \( k \) (\( \ell^{\text{MTU}} \) or DOF) on channel \( i \) in condition \( c \). Probability distributions were modelled using Poisson distributions. In total, 50 different grasping conditions were decoded (48 objects plus precision and power grip in the handle task). In the DOF and the MTU space, 27 and 50 channels (number of DOF or MTU) were available, respectively. Since \( p(c) \) is uniformly distributed and \( p(k_i \mid c) \) does not depend on condition \( c \), both terms can be summarized by a normalization factor \( n_i \) that is independent of condition \( c \). This allows expressing the probability of \( c \) for given \( k_i \) as

\[ p(c \mid k_i) = n_i \cdot p(k_i \mid c) \tag{6} \]

In a second step (decoding), a different segment of the recording was used to find the most likely condition for the observed kinematic variable \( k \). To predict the decoded object, the likelihood function

\[ L(c) = \prod_{i=1}^{N} p(c \mid k_i) \tag{7} \]

was evaluated for all conditions \( \hat{c} \) and the condition selected as the predictor that maximized this function: \( \hat{c} = \arg\max_{c} L(c) \).

Since the factor \( n_i \) in (5) was constant across all conditions, it could be ignored for the maximization and the calculation reduced to

\[ \hat{c} = \arg\max_{c} \left( \prod_{i=1}^{N} p(k_i \mid c) \right) \tag{8} \]

Leave-one-out cross validation ensured that data sets for training were not used for testing.

For comparing the classification accuracy based on joint-angle or muscle patterns, we computed the decoding accuracy versus the number of included objects \( (n = \{1, \ldots, 50\}) \). To rank the objects according to their separability, we first extracted the kinematic variables \( (\ell^{\text{MTU}}, \text{DOF}) \) from the mean hold epoch and computed the Mahalanobis distance (MD) between each pair of objects groups (Matlab function: manova1). The two objects (groups) showing the largest MD to each other (most separated in \( \ell^{\text{MTU}} / \text{DOF} \) space) were added first to the decoder. Next, a new decoder was built by adding the object that maximized the performance in combination with previously selected objects. This “greedy” selection process was continued by gradually adding one more object at a time and rebuilding the decoder [5]. Using this procedure allowed us to rank the objects based on their separability in the DOF and MTU space. Objects that could be separated best by the decoder were ranked first.

III. RESULTS

The model was successfully used to provide fast, accurate, and reliable musculoskeletal descriptions for a large repertoire of grasping movements in primates.

A. Extracted Parameters

All extracted model parameters allowed visualizing the course of motion continuously over time. As an example, we present the kinematic and muscular variables of three subsequent grasping movements together in (Fig. 3) and as supplementary videos [see video 1 for DOF extraction and video 2 for MTU extraction].

The parameters clearly demonstrated a distinct approach for grasping the three objects: ring, sphere, and cylinder (i.e., Fig. 3(b), (c) at times \( t_1, t_2, \) and \( t_3 \), respectively). The ring was grasped using power grips, requiring a strong rotation (supination) of the wrist of about \( -70^\circ \) (i.e., see Wrist DOF variable in Fig. 3(b) at time \( t_1 \)). In macaque monkeys, the rotation of ulnar and radius is mainly achieved by the pronator teres (PT) and the pronator quadratus muscle (PQ, i.e., see Wrist \( \ell^{\text{MTU}} \) variable in Fig. 3(c) at time \( t_1 \)) for pronating the forearm, whereas the supinator (SUP, i.e., see Wrist \( \ell^{\text{MTU}} \) variable in Fig. 3(c) at time \( t_1 \)) and the biceps muscles (BIClong, BICshort; long and short head, respectively) are responsible for supination. In addition, the brachioradialis muscle (BRD) moves the forearm back to a neutral position, either from a supination or from a pronation posture. Results show how grasping the ring led to a supination of the forearm with a shortening of the SUP-muscle and a simultaneous lengthening of the PQ-muscle (i.e., see Wrist \( \ell^{\text{MTU}} \) variable in Fig. 3(c) at time \( t_1 \)). When the monkey grasped the
sphere or the cylinder, the rotation of the forearm was minimal, therefore affecting the corresponding muscles (i.e., PQ and SUP) less (i.e., see Wrist $\ell^{\text{MTU}}$ variable in Fig. 3(c) at time $t_2$ and $t_3$). Grasping the small sphere required a more precise control of index finger and thumb with longer grasping time needed in comparison to ring and cylinder. Furthermore, this grip type caused a strong flexion of index finger and thumb, which was also reflected in the corresponding length of thumb and index finger muscles, including the flexor pollicis longus (FPL), the extensor pollicis longus (EPL) as well as the flexor digitorum profundus index (FDPI) and the extensor indicis proprius (EIP) (i.e., see Thumb $\ell^{\text{MTU}}$ and Index $\ell^{\text{MTU}}$ variables in Fig. 3(c) at time $t_2$). In comparison to the other two objects, the cylinder was grasped with a whole-hand grip achieved by digits and thumb requiring less flexion of the fingers that caused less shortening of the muscles (i.e., see Thumb $\ell^{\text{MTU}}$ and Index $\ell^{\text{MTU}}$ variables in Fig. 3(c) at time $t_3$).

These findings demonstrate the ability of our musculoskeletal model to extract joint angle (DOF) and muscle kinematics ($\ell^{\text{MTU}}$) for describing finger, hand, and arm movements.

**B. Model Fit**

Results from the model fit test revealed an average RMSE error of $2.06 \pm 1.64$ mm and $2.33 \pm 1.50$ mm (mean ± standard deviation) across all pairs of experimental and model markers from all trials performed by animal Z and M respectively. The marker RMSEs in animal Z ranged from 0.77±0.13 mm (i.e., DIP marker of ring finger) to 5.66±0.31 mm (i.e., MCP marker of little finger). For animal M, marker RMSEs ranged from 0.93±0.22 mm (i.e., DIP of middle finger) to 8.26±0.10 mm (i.e., MCP of thumb marker). Across all trials and animals, the majority of average RMSE errors between experimental and model markers were smaller than 3 mm (for 40 out of 49 pairs) as shown in Fig. 4.

Importantly, these RMSE values are small in comparison to objects sizes that ranged from 15 to 60 mm in diameter or length (balls and cubes, respectively). Therefore, model-fit accuracies were well suited for investigating the musculoskeletal representations performed on the set of objects.

**C. Computation Speed**

The recording sessions from both animals produced extensive datasets that were used to assess the simulation speed of our musculoskeletal model. We found that performing inverse kinematics with OpenSim, despite providing a more complex musculoskeletal description of hand and arm, was significantly faster to compute than our previous, simpler kinematic model (KinemaTracks, [14]). The OpenSim's IK could process both sessions (i.e., 284795 and 342335 frames from animal Z and M, respectively) in 15:08 and 17:31 minutes, whereas KinemaTracks required 18:12 and 19:55 minutes. In this, the average time needed to extract one frame of DOF and MTU kinematics from all joints and muscles required 3.1 ms of computation time. The musculoskeletal model therefore allowed extracting the motion data faster than real-time and can be readily used to process extensive datasets such as the ones presented here.

**D. Principal Component Analysis of Muscle and Joint Angle Patterns**

We used PCA to interpret and visualize the large repertoire of grasping movements in both the DOF and $\ell^{\text{MTU}}$ space. As an example we plotted the first two PCA components in the hold epoch of animal Z for both, the DOF and $\ell^{\text{MTU}}$ representation (Fig. 5(a) and (b), respectively). The plots demonstrate a high variablity of joint angle and muscle patterns across the large number of objects and trials.

For animal Z, the first PCA component roughly represented the aperture of the hand, pointing from the condition of smallest (precision grip) to the condition of highest aperture (big boxes), whereas the second component approximately reflected the monkey's wrist orientation. We also used PCA to reveal similarities and dissimilarities between grasping movements. For example, smaller objects including the small ball, cube, or ring were grasped with a similar grip, using index finger and thumb, and were therefore located close to each other in PCA space. Even objects of largely diverse shape might be grasped very similarly such as the vertically aligned cylinders, big rings, and the handle. All of these objects were grasped with a kind of power grip (rotation of the forearm and enclosure with four digits).

Notably, the DOF and the $\ell^{\text{MTU}}$ space showed a strong similarity within PCA coordinates. Almost the same cluster overlaps were observed for the DOF and for the MTU representation [Fig. 5(a), (b)]. Computing the Pearson correlation coefficient between both multidimensional spaces further demonstrated the strong link between the DOF and the $\ell^{\text{MTU}}$ representation ($\rho = 0.85$ and 0.75 for animal Z and M, respectively). These similarities were expected, since the muscle patterns and the joint angle patterns are strongly linked to each other and both reflect the shape and size of the grasped objects.

Although the representation of grasping movements was similar in the DOF and in the MTU space, how many linearly uncorrelated components were actually engaged in the task remained an open question. The large number of natural grip types and the resulting large variability in our dataset allowed us to answer this question. We determined how many principal components were required to reproduce the individual grip types [Fig. 5(c)]. We found that eight PCA components were sufficient in both animals to explain more than 95% of variance across all conditions in the MTU-space, whereas in the DOF-space 11 (animal Z) and 12 components (animal M) were needed to explain the same amount of variance. The lower dimensional representation in the $\ell^{\text{MTU}}$ space is remarkable, since the number of MTUs (i.e., 50) involved in grasping strongly exceeded the number of DOF (i.e., 27).

In addition to these findings, we tested whether the 50 object conditions were also represented more compactly in the $\ell^{\text{MTU}}$ than in the DOF space by computing the Dunn index (D)[29]. In both animals, the condition clusters were represented more separated when expressed in the $\ell^{\text{MTU}}$ space than in the DOF space (D for animal Z: 1.17 DOF, 1.39 $\ell^{\text{MTU}}$; animal M: 0.57 DOF, 0.714 $\ell^{\text{MTU}}$).

These findings emphasize the importance of the musculoskeletal organization for the control of hand movements,
Fig. 3. Model features: (a) Musculoskeletal model, (b) allowed extracting joint angles, (c) as well as muscle kinematics. In total, the information of 27 DOFs (18 joints) and 50 MTUs could be accessed. (a) Hand and arm posture of the monkey while grasping three different objects in the hold period of sample trial t1 (ring), t2 (ball), and t3 (cylinder). In addition to the skeletal configuration, a selection of MTUs is illustrated for these trials in red [anterior deltoid (antDELT), posterior deltoid (postDELT), teres minor (TMIN), teres major (TMAJ), coracobrachialis (CORB), triceps long head (TRIlng), biceps long head (BIClng), brachialis muscle (BRA), extensor carpi radialis brevis (ECRB), flexor carpi radialis (FCR), FDPI (flexor digitorum profundus index), extensor indicis proprius (EIP), extensor pollicis longus (EPL), extensor pollicis brevis (EPB)]. (b) Example DOFsof the thumb, index finger, wrist, elbow, and shoulder while subsequently grasping three objects. (c) \( \mathcal{E}^{\text{MTU}} \) of example muscles linked to the joints in (b).

Fig. 4. Test of model fit accuracy. Histogram shows RMSEs (3) derived from all pairs of experimental and model markers averaged across all trials and animals (50 markers in total).

which might become highly relevant for the development of future hand prosthetics [31].

E. Object Classification Accuracy

As a measure of discriminability, we implemented a Bayesian classifier for decoding objects based on the extracted DOF or \( \mathcal{E}^{\text{MTU}} \) parameters (27 and 50 dimensions, respectively).

Starting with a decoder of including only two objects, we iteratively added more objects to the decoder in a “greedy” fashion based on their separability (see Section II). As shown in Fig. 6(a) (solid dots), the decoding accuracy remained stable until objects were included, that evoked similar DOF or \( \mathcal{E}^{\text{MTU}} \) representations with respect to other objects. As a consequence, the decoding accuracies dropped. Importantly, the \( \mathcal{E}^{\text{MTU}} \)-representation of grasping movements was better suited to distinguish individual objects than with the DOF representation. These findings were consistent in both animals and in agreement with the
cluster analysis results. To test, whether this effect remained with an equal number of dimensions, we decoded the same objects with a randomly selected set of ten dimensions (100 repetitions) in the $\ell^\text{MTU}$ and DOF space. As shown in Fig. 6(a) (dashed lines), $\ell^\text{MTU}$ still provided a better decoding in comparison to DOFs.

Further tests were performed on all dimensions. In the complete DOF space, the decoding accuracy was better than 95% for up to a maximum number of 20 objects in both animals. In comparison, the same decoding accuracy was achieved with a maximum number of 28 and 26 objects in the MTU space for animal M and Z, respectively. Beyond these thresholds the decoding accuracy decreased linearly with increasing number of objects (i.e., less separable objects).

To quantitatively estimate the point of accuracy drop, we implemented the L-method [32]. The algorithm searches for the two piecewise linear regression lines that fit the decoding results best [solid lines, Fig. 6(a)]. For this, the algorithm gradually changes the intersection points of the two lines and measures their fit. The RMSE error produced by this fit is plotted in Fig. 6(b) for variable intersection points. Results highlight a better resolution and separability between objects when using the $\ell^\text{MTU}$ description of grasping movements rather than the DOF description.

Additionally, we computed the PCA components accounting for 95% of the variance of the data in the $\ell^\text{MTU}$ and DOF space as a function of the number of included objects in the greedy selection. Fig. 6(c) shows a strong increase in variability when the most separable objects (ranked first) were added. However, the variability reached a plateau already for a small number of objects in both animals. This demonstrated that including similar objects could not increase the variability of grip types further and explained the significant drop of decoding accuracy for larger object sets. Classifying the objects based on $\ell^\text{MTU}$ and DOF demonstrated a higher separability and resolution in the MTU-space.

**IV. DISCUSSION**

In this study we presented a musculoskeletal modelling method that could account for the specific upper extremity anatomy in two primates. The animal-specific model was used to investigate grasp actions from a large repertoire of different objects. This study enabled observing and investigating larger kinematics variability in the primate’s arm and hand (Fig. 5) than it was possible in previous studies [5], [12]. Furthermore, it allowed for the first time extracting comprehensive kinematics information in 27 DOFs and 50 MTUs in the arm and hand segments simultaneously [17], [18]. This resulted in higher classification accuracy than reported in previously published methods that used a nine-DOF, nonanatomical kinematics model in conjunction with electromyography (EMG) signals from 24 muscles [16].

Importantly, results showed that the number of components extracted from MTU kinematics that describe 95% of the data variability across all grip types was substantially smaller than the number of components extracted from DOFs kinematics for reconstructing the same data variability [Fig. 6(c)]. This emphasizes the nonlinear relationship between MTU and DOF kinematics [18], [33] and proves its exploration to generate a more compact representation of grasping actions than it was possible with conventional joint kinematics. Further research is needed to investigate whether this may have related neural mechanisms. Supporting results in this direction [5] are providing evidence of direct relationships between single M1-neuron activity and the simultaneous actuation of proximal distal joints in the hand segments, such as wrist pronation/supination and thumb-index aperture. In this context, the relationship between single M1 neuron activity and simultaneous actuation of multiple DOFs might reflect the activation of the underlying muscles spanning proximal and distal joints. This hypothesis is further supported by the fact that M1-neurons have a direct influence on the spinal alpha motor neurons innervating MTUs [5]. The ability of extracting variables reflecting muscle behavior from reach-to-grasp movements, such as presented here, might help further explain the neural code of grasping.

The compact representation of grasping tasks provided by MTU kinematics also resulted in a decoding method that could account for a larger number of objects than observed using DOF kinematics-based decoders Fig. 6(a). In this, MTU redundancy may play an important role. Multiple MTUs spanning the same DOF may play different roles across grasping tasks. This may contribute to assuring good representation of a larger number of objects than it was possible with DOFs only.
Fig. 6. Object classification accuracy. (a) Classification accuracy of the Bayesian Decoder versus the number of gradually included objects sorted according to their separability (DOF in blue, $\theta^\text{MTU}$ in red). Decodings were performed on the complete spaces (solid dots; 27 DOF and 50 $\theta^\text{MTU}$) and on the reduced spaces (dashed line; 10 DOF and 10 $\theta^\text{MTU}$). To find the optimal number of objects for the decoder in the complete space, the L-method was applied (Section III-E), which provided two regression lines fitting the data best (solid black and green lines). (b) Edge position (intersection of both regression lines) was systematically varied in order to minimize RMSEs. Object numbers with minimal RMSE are marked with green boxes. (c) Number of principal components required to describe 95% of data variance were plotted against the number of included objects (DOF-domain in red, $\theta^\text{MTU}$-domain in blue). (d) For animal M, the set of objects that achieved decoding accuracies higher than 95% (28 in $\theta^\text{MTU}$ and 20 in DOF space) are illustrated. Their order corresponds to the x-axis of a–c. Large panels a–c: results for animal M; inlays: corresponding results for animal Z.

Some limitations in our methods should be recognized. The proposed primate-specific model was derived and extended from a human upper extremity model [18] (see Section II). This model was used as a basis for our study because of the extensive research consistently confirming strong anatomical and morphological similarities in the majority of the upper extremity joints and muscles between both species [23], [24], [34]. However, some differences should be acknowledged. While joints move about the same DOFs in both species, their range of motion can differ. To support and stabilize quadrupedal locomotion, the shoulder rotation in NHP is smaller than in humans, whereas the NHP wrist allows for a higher abduction and extension [23]. Differences are also reflected at the muscular level across proximal and distal joints [34]. This results in discrepancies in terms of insertion, origin, and size in muscle groups that are shared across the two species. Furthermore, this results in some muscles not being preserved across the two species [34]. Major differences in proximal muscles include the so-called “climbing muscle” (i.e., the dorsoepitrochlearis muscle), which is not present in humans [35], [36]. Major
differences in distal muscles relate to the macaque extensor digitorum communis group in the hand, which is equivalent to the human's extensor digitorum group. The macaque extensor digitorum communis group differs from the human's in that the four separate tendons are united as far distally as the metacarpal heads, after which they diverge and progress toward all fingers. Overall, morphological differences are related to the habitual quadrupedalism of the macaque, with the constant extension of the wrist and fingers and with larger moment arms in the shoulder muscles [34].

Therefore, the ability of fitting the generic model to each animal specifically played an important role in compensating for the reported morphological discrepancies. Fitting errors were small throughout each performed task, providing evidence that the scaled model not only matched a specific static pose in the monkey, but remained kinematically consistent throughout all performed motor tasks (see Section III-B). This resulted in a scaled model that, despite morphological discrepancies with respect to the primate anatomy, resulted in improved performances in dimensionality reduction and decoding with respect to conventional methods based on DOF kinematics analysis.

Future research will assess whether the availability of a more correct musculoskeletal model may substantially improve dimensionality reduction and decoding results. Future work will also use EMG-driven musculoskeletal modelling methodologies [37], [38] for reconstructing all the transformations from the neural drive onset (i.e., muscle-specific EMG onset) to the generation of functional movement in the upper extremity joints. Combining EMG-driven modelling [37] with in vivo recordings of neuronal activity will enable, for the first time, tracking and monitoring all the major steps taking place along the neuromuscular pathway in primates.

Findings from this study might have important implications in the context of neurorehabilitation technologies. The development of multi-articulated prosthetic hands with fingers and thumb individually actuated by artificial tendons is providing amputees with effective solutions to restore lost grasping capabilities [39], [40]. Our presented methodology could be used to determine the required length changes to be generated in the prosthesis artificial tendons to reconstruct large repertoires of grasping tasks accurately. Importantly, it allows controlling a large number of artificial tendons with a low-dimensional controller as suggested by our PCA results (i.e., Section III-D and Fig. 5). This could bring current PCA-based prosthesis controllers an important step forward [40]. Our proposed methodology could also have implications for the development of advanced neuroprostheses [8], [41]. In this scenario, recorded neuronal activity might better correlate to the coordinated behavior of multiple muscles than to the movement of individual DOFs, which might lead to the possibility of controlling tendon-driven prostheses with more intuitive control systems.

V. CONCLUSION

This study showed that a musculoskeletal model of the upper extremity can be scaled to match the specific anatomy and of individual primates. The scaled model provides a nonlinear transfer function from the joint domain to the muscle domain, which enables a more compact representation and a higher decoding accuracy of large repertoires of grasping actions than was possible using traditional joint kinematics recording. This has direct implications for improving the decoding of complex grasping movements and for the development of biologically inspired control systems for upper limb prostheses. A better understanding of the inherent nonlinearity between the skeletal and muscular systems is crucial for unravelling the nature of the neuromuscular control mechanisms in the primate's arm and hand movement.

ACKNOWLEDGMENT

The authors thank R. Ahlert, N. Nazarenus, and L. Burchardt for assistance in animal training, M. Dörge for technical assistance, L. Schaffelhofer for glove production, and J. Michaels for comments on an earlier version of the manuscript. A download of the musculoskeletal model is accessible under https://simtk.org/home/mskh.

REFERENCES


Authors’ photographs and biographies not available at the time of publication.