Representation of continuous hand and arm movements in macaque areas M1, F5, and AIP: a comparative decoding study

To cite this article: Veera Katharina Menz et al 2015 J. Neural Eng. 12 056016

View the article online for updates and enhancements.

Related content
- A new hand-tracking method for primates
  S Schaffelhofer and H Scherberger
- Distinct neural patterns enable grasp types decoding in monkey dorsal premotor cortex
  Yaoyao Hao, Qiaosheng Zhang, Marco Controzzi et al.
- Continuous decoding of human grasp kinematics using epidural and subdural signals
  Robert D Flint, Joshua M Rosenow, Matthew C Tate et al.

Recent citations
- A.B. Goldring and L.A. Krubitzer
Representation of continuous hand and arm movements in macaque areas M1, F5, and AIP: a comparative decoding study

Veera Katharina Menz, Stefan Schaffelhofer and Hansjörg Scherberger

1 German Primate Center GmbH, D-37077 Göttingen, Germany
2 Department of Biology, University of Göttingen, D-37077 Göttingen, Germany

E-mail: hscherberger@dpz.eu

Received 9 February 2015, revised 5 August 2015
Accepted for publication 6 August 2015
Published 11 September 2015

Abstract

Objective. In the last decade, multiple brain areas have been investigated with respect to their decoding capability of continuous arm or hand movements. So far, these studies have mainly focused on motor or premotor areas like M1 and F5. However, there is accumulating evidence that anterior intraparietal area (AIP) in the parietal cortex also contains information about continuous movement. Approach. In this study, we decoded 27 degrees of freedom representing complete hand and arm kinematics during a delayed grasping task from simultaneously recorded activity in areas M1, F5, and AIP of two macaque monkeys (Macaca mulatta). Main results. We found that all three areas provided decoding performances that lay significantly above chance. In particular, M1 yielded highest decoding accuracy followed by F5 and AIP. Furthermore, we provide support for the notion that AIP does not only code categorical visual features of objects to be grasped, but also contains a substantial amount of temporal kinematic information. Significance. This fact could be utilized in future developments of neural interfaces restoring hand and arm movements.

Keywords: grasping, decoding, monkey, parietal cortex, premotor cortex, motor cortex

1. Introduction

The ability to perform dexterous hand manipulations is crucial for a person’s well-being (Anderson 2004). In order to restore the skill of everyday hand movements in patients who have lost this ability due to spinal cord injuries, loss of limbs or other motor diseases, the development of neuroprostheses could be very beneficial. The goal of these devices is to translate neural activity from brain areas involved in movement generation into motor output signals to control an external actuator as a hand and arm replacement. Initial successful attempts have been realized in both humans (Hochberg et al 2012, Collinger et al 2013, Afshar et al 2015) and monkeys (Wessberg et al 2000, Carmena et al 2003, Lebedev et al 2005, Velliste et al 2008) where neural activity from primary motor cortex (M1) and other cortical areas was used to control a robotic arm and hand, which enabled the subject to successfully perform several reaching and grasping tasks.

On the cortical level, the control of visually guided hand movements requires a visuomotor transformation to convert visual information of an object to be grasped or manipulated into a movement plan that can be executed. Several areas in the brain have been identified to be involved in this
visuomotor transformation (Jeannerod et al. 1995). The ante-
rior intraparietal area (AIP) in the parietal lobe is specialized
for the visual guidance of hand movements (Gallese et al. 1994, Murata et al. 2000, Baumann et al. 2009). It
receives visual information from areas like the caudal intra-
parietal sulcus (Sakata et al. 1995, Sakata et al. 1999) and
extracts visual and spatial characteristics of graspable objects.
The information is shared with the rostral part of the inferior
premotor cortex (area F5) (Sakata et al. 1995, Sakata et al.
1997). In area F5, a movement plan is selected based on
the information received from AIP (Murata et al. 1997) and
passed on to the primary motor cortex (M1) for execution.
From there, cortico-spinal projections activate the spinal cord
that excites the muscles (Bennett and Lemon 1996, Holdefer
and Miller 2002, Morrow and Miller 2003, Rathelot and
Strick 2006).

More precisely, F5 has been suggested to encode specific
grip types for grasping an object (Rizzolatti et al. 1987, Riz-
et al. 2006, Stark et al. 2007, Fluet et al. 2010) rather than
specific joint or muscle commands for executing a grip
(Rizzolatti and Luppino 2001, Umlita et al. 2007). Instead,
joint angle and muscle representations were reported to be
encoded in primary motor cortex (Thach 1978, Bennett
and Lemon 1996, Holdefer and Miller 2002, Morrow and
Miller 2003, Rathelot and Strick 2006, Umlita et al. 2007)
 altogether with more abstract kinematic features like movement
direction (Thach 1978, Georgopoulos et al. 1986, Ashe and
Georgopoulos 1994, Kakei et al. 1999), position of the wrist
(Thach 1978, Ashe and Georgopoulos 1994), and force

Therefore, many studies investigating the possibilities to
decode hand and arm movements from neural activity have
focused on utilizing neural activity from M1 (Ben Hamed
et al. 2007, Vargas-Irwin et al. 2010, Orsborn et al. 2014) and
F5 for their predictions (Bansal et al. 2012, Aggarwal
et al. 2013). Most of these studies recorded neural activity
from rostral regions of M1. Rathelot and Strick (2009) dis-
covered that M1 can be partitioned into two subdivisions with
different influences on motor output: one part (termed ‘old’
M1 and located in the rostral region of M1) contains cells that
contact spinal interneurons and that are therefore considered
to have a mainly indirect influence on motor commands. In
contrast, neurons making direct, monosynaptic connections
with motoneurons in the spinal cord (CM cells) are almost
exclusively located in the caudal region of M1 (termed ‘new’
M1), and these cells might be particularly relevant for dex-
terous finger movement control (Rathelot and Strick 2009).
Recordings from ‘new M1’ will likely include a larger fraction
of neurons located in the neighborhood of CM cells that
are strongly selective for finger movements and grip types,
e.g. as observed in Schaffelhofer et al. (2015a).

Although AIP is believed to carry information about visual
properties of the object to be grasped (Jeannerod et al. 1995, Sakata et al. 1995, Sakata et al. 1997, Sakata
et al. 1999, Murata et al. 2000), there is accumulating evi-
dence that units in AIP also carry information about grip types
and hand movement (Baumann et al. 2009, Aflalo et al. 2015).

Inactivation of area AIP in monkeys results in deficits of hand
prehshaping and mismatch of hand orientation when grasping
an object (Faugier-Grimaud et al. 1978, Gallese et al. 1994),
impairments that are very similar to those reported after
inactivation of F5 (Fogassi et al. 2001), and AIP neurons have
been reported to become active during movement execution
(Gallese et al. 1994, Jeannerod et al. 1995, Sakata et al. 1995,
Sakata et al. 1997, Gardner et al. 1999). Cells in AIP were
found to be tuned to grip type and target position in space
(Jeannerod et al. 1995, Murata et al. 2000, Baumann
et al. 2009, Lehmann and Scherberger 2013). It was suggested
that activity in AIP could already reflect a motor plan based
on visual and somatosensory information (Murata et al. 2000,
Debowy et al. 2001) instead of solely extracting visual cues
relevant for grasping.

So far, only very few studies have investigated the suit-
ability of AIP for decoding grasping features: Townsend et al.
(2011) as well as Lehmann and Scherberger (2013) were able
to decode two different grip types with an accuracy of ~70%
and 75% (averaged), respectively. However, the result was
considerably lower than for decoding from area F5 (perfor-
mance: >90%). A similar result was obtained by Schaffel-
hofer et al. (2015a) who predicted twenty different grip types
with both F5 and AIP and obtained higher decoding accuracy
with F5 than with AIP. In contrast, decoding of grip type
together with object position in space or object orientation
could be performed with higher accuracy from signals in AIP
than F5 (Townsend et al. 2011, Lehmann and Scherber-
ger 2013). Furthermore, signals from posterior parietal cortex
(PPC) such as the parietal reach region and area 5d have been
used to predict reach kinematics in 3D space (Wessberg

In this study, we recorded spiking activity simultaneously
from the three cortical areas AIP, F5, and ‘new’ M1. Using a
delayed grasping task with a large number of distinct objects,
we predicted 27 degrees of freedom (DOF) of the hand and
arm from these brain areas, which, to our knowledge, is the
most complete decoding study of finger, wrist, and arm joints
undertaken so far. We found that all 27 DOF could be pre-
dicted accurately from either of the three cortical areas with
M1 delivering the highest accuracy, followed by F5. AIP
yielded lowest decoding performance, but performed sig-
ificantly above chance. To our knowledge, this is the first
time that continuous decoding of grasp kinematics was
achieved using spiking activity recorded simultaneously from
areas AIP, F5, and M1.

2. Materials and methods

2.1. Basic procedures

For decoding hand movements, two purpose-bread macaque
monkeys (Macaca mulatta; animal Z: female, 7.0 kg; animal
M: male, 10.5 kg) were trained to grasp a wide range of
different objects (Schaffelhofer et al. 2015a) while wearing an
instrumented glove (Schaffelhofer and Scherberger 2012).
After training was accomplished, both animals were
implanted with head holders on the skull and subsequently with microelectrode arrays in cortical areas AIP, F5, and M1. In the following recording sessions, spiking activity was recorded from these electrodes together with the kinematics of the primate hand. Animal care and all experimental procedures were conducted in accordance with German and European law and were in agreement with the Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research (National Research Council 2003).

2.2. Experimental setup

During training and experimental sessions, animals were sitting upright in a customized animal chair with their head fixed. In order to protect the instrumented glove, we constrained the passive (non-grasping) hand with a plastic tube that encompassed the forearm in a natural posture. A capacitive switch at the side of the active (performing) hand allowed detecting the animal’s hand position at rest. All graspable objects (figure 1(b)) were placed at a distance of ~25 cm in front of the animal at chest level. We used a PC-controlled turntable (figure 1(a)) to pseudo-randomly present the objects. Light barriers and a step motor ensured precise object positioning. Additional light barriers beneath the turntable detected the time when the monkey lifted the displayed object.

Eight exchangeable turntables facilitated the presentation of a total of 48 objects. To obtain a high variation of hand kinematics, we designed the objects to have different shapes and sizes (figure 1(b)) including rings (outer diameter: 10, 20, 30, 40, 50 and 60 mm), cubes (length: 15, 20, 25, 30, 35 and 40 mm), spheres (diameter: 15, 20, 25, 30, 35 and 40 mm), cylinders (length: 140 mm; diameter: 15, 20, 25, 30, 35 and 40 mm), and bars (length: 140 mm; height: 50 mm; depth: 15, 20, 25, 30, 35 and 40 mm). Furthermore, a mixed turntable was used, holding mid-sized objects of various shapes for fast exchange (sphere 15 mm, horizontal cylinder 30 mm, cube 30 mm, bar 10 mm, ring 50 mm), as well as a special...
Prior to surgery, we performed a 3D anatomical MRI scan of the animal’s skull and brain to locate anatomical landmarks (Townsend et al. 2011). For this, the animal was sedated (e.g., 10 mg kg$^{-1}$ ketamine and 0.5 mg kg$^{-1}$ xylazine, i.m.), placed in the scanner (GE Signa HD or Siemens TrioTim; 1.5 Tesla) in a prone position, and T1-weighted images were acquired (iso-voxel size: 0.7 mm$^3$).

In an initial procedure, a head post (titanium cylinder; diameter 18 mm) was implanted on top of the skull (approximate stereotaxic position: midline, 40 mm anterior, 20 deg forward tilted) and secured with bone cement (Refobacin Plus, BioMed, Berlin) and orthopedic bone screws (Synthes, Switzerland). After recovery from this procedure and subsequent training with head fixation, each animal was implanted in a second procedure with six floating micro-electrode arrays (FMA: MicroProbes for Life Science, Gaithersburg, MD, USA). Specifically, two FMAs were inserted in each area AIP, F5, and M1 (see figure 1d). FMA consisted of 32 non-moveable monopolar platinum-iridium electrodes (impedance: 300–600 kΩ at 1 kHz) as well as two ground and two reference electrodes per array (impedance <10 kΩ). Electrode lengths ranged between 1.5 and 7.1 mm and were configured as in Townsend et al. (2011).

FMA implantation locations are depicted in figure 1d. In both animals the lateral array in AIP was located at the end of the intraparietal sulcus at the level of the parietal area PF, whereas the medial array was placed more posteriorly and medially at the level of the parietal area PFG (Borra et al. 2008). In area F5, the lateral array was positioned approximately in area F5a (Belmalih et al. 2009, Borra et al. 2010), whereas the medial array was located in F5p in animal Z and at the border of F5a and F5p in animal M. Finally, both arrays in M1 were positioned in the hand area of M1 (anterior bank of the central sulcus at the level of the spur of the arcuate sulcus and medial to it) (Rathelot and Strick 2009).

All surgical procedures were performed under aseptic conditions and general anesthesia (e.g., induction with ketamine 10 mg kg$^{-1}$ i.m., and atroptine 0.05 mg kg$^{-1}$ s.c., followed by intubation, isofluorane 1–2%, and analgesia with buprenorphene 0.01 mg kg$^{-1}$ s.c.). Heart and respiration rate, electrocardiogram, oxygen saturation, and body temperature were monitored continuously. Systemic antibiotics and analgesics were administered for several days after each surgery. To prevent brain swelling while the dura was open, the animal was mildly hyperventilated (end-tidal CO$_2$ <30 mmHg) and mannitol kept at hand. Animals were allowed to recover fully (~2 weeks) before behavioral training or recording experiments recommenced.

2.5. Hand kinematics

Previously, we developed an instrumented glove for small primates that allowed recording the animal’s finger, hand, and arm movements in 27 DOF (Schaffelhofer and Scherberger 2012). The glove was equipped with 7 electro-magnetic sensor coils that enabled hand and arm movement tracking at a temporal resolution of 100 Hz without depending on line of sight to a camera (figure 1c). This way, finger movements could be tracked continuously even when sensors were located behind or below objects. The method is described in detail in Schaffelhofer and Scherberger (2012); in short, it exploits the anatomical constraints of the primate hand and combines it with the 3D position and orientation information of the sensors located at the fingertips, the hand’s dorsum, and the lower forearm to obtain a full kinematic description of the animal’s hand and arm.

Recorded 3D positions of the finger, hand, and arm joints were compared with the animal’s anatomy. Rarely occurring
errors (e.g. due to short freezing of the kinematic tracking) were set to NaN. Then, each joint angle was calculated for every recorded time step and afterward linearly resampled to exactly 100 Hz. This way, missing values were linearly interpolated.

Specifically, the following 27 joint angles were extracted: finger flexion/extension in the carpometacarpal (CMC), metacarpophalangeal (MCP), and interphalangeal joint of the thumb and the MCP, proximal interphalangeal, and distal interphalangeal (DIP) joint of the index, middle, ring, and little finger (fingers 2–5); finger abduction/adduction in the thumb CMC and the MCP of fingers 2–5; radial/ulnar deviation (yaw), flexion/extension (pitch), and pronation/supination (roll) of the wrist; flexion/extension of the elbow; and shoulder abduction/adduction, flexion/extension, and internal/external rotation.

Furthermore, the velocity of every DOF at each time sample was calculated and added to the kinematics. This resulted in a 54x7-matrix with T as the total number of samples. Note, that only data that was recorded while the monkey was exposed to the handle or turntable task was used and concatenated into the matrix, including correctly and incorrectly performed trials, reward epochs, and the time epochs between trials (inter-trial intervals). Data recorded while turntables or the Grasping Box were exchanged by the experimenter were not included.

For illustration purposes, the 27 DOF were averaged during the hold epoch across all successful trials of one recording session (M111913), and a principal component analysis was performed. The first two principal components of the kinematics of each trial are illustrated in figure 1(e), demonstrating that we were able to cover a wide range of hand configurations together with sampling very small differences in hand kinematics.

2.6. Neural recordings

Extracellular signals were recorded simultaneously from six FMAs (6 × 32 channels) that were permanently implanted into area AIP, F5, and M1 (two FMAs in each area). All arrays provided sufficiently large numbers of recording channels within the specified electrode impedance (∼500 kΩ) across all electrode rows. Raw signals were sampled at a rate of 24 kHz with a resolution of 16 bit, and stored to disk together with the behavioral and the kinematic data using a RZ2 Biosignal Processor (Tucker Davis Technologies, FL, USA). Offline, raw data was bandpass filtered (0.3–7 kHz) and spikes were detected (threshold: 3.5 standard deviation). Spike sorting was performed first with WaveClus (Quiroga et al 2004) for automatic sorting, followed by manual inspection and revision using OfflineSorter (Plexon TX, USA). Spikes of single and multiunits were further analyzed.

2.7. Decoding

2.7.1. Decoding algorithm. For offline prediction of the hand and arm kinematics, we employed a Kalman filter (Kalman 1960) as described in Wu et al (2004, 2006). A Kalman filter assumes a linear relationship between the kinematics at time instant k and the following time instant k + 1 (step size: 10 ms), as well as between the kinematics and the neural data at time k:

\[ x_{k+1} = Ax_k + w_k, \]
\[ z_k = Hx_k + q_k. \]

Here, \( x_{k+1} \) is the 54 × 1-vector of the 27 angles and their velocities at time k + 1, \( A \) a 54 × 54-matrix relating the kinematics of one time step to the following one, \( z_k \) a vector of length N containing the neural data at time k where N denotes the total number of recorded units, \( H \) a N × 54-matrix relating the kinematics at time k to the corresponding neural information, and \( w_k \) and \( q_k \) noise terms that are assumed to be normally distributed with zero mean, i.e., \( w_k \sim N(0, W) \) and \( q_k \sim N(0, Q) \), with \( W \in \mathbb{R}^{54×54} \) and \( Q \in \mathbb{R}^{N×N} \) covariance matrices.

2.7.2. Decoding procedure. For decoding a complete recording session, we used 7-fold cross-validation. For this, the recording session was divided into seven data sets of equal length and composition: each task type (i.e., grasping box task and data from each turntable) was divided into seven parts of equal length and these parts were randomly attributed to the seven sets, so that eventually each set contained data of each turntable and the grasping box task.

To decode the kinematics of the ith set of the seven sets, the variables of the Kalman filter, namely \( A, H, W, \) and \( Q \), were calculated by using the neural and kinematic data of the remaining six sets (decoder training). Then, the kinematic data of the ith set (27 DOF and their respective velocities) were iteratively predicted by setting the initial value (and velocity) at the first time step of each DOF randomly within the observed range of the respective DOF/velocity during the recording (interval between the smallest and highest value). For each subsequent time step, the Kalman filter first calculates an a-priori estimation of the kinematics based on the kinematic state of the previous time step. Then, this prediction is updated based on the corresponding neural activity (a posteriori estimation) (Welch and Bishop 2006). This procedure was carried out for each of the seven data sets, hence leading to the complete decoding of the entire recording session. The decoding procedure was implemented in Matlab (MathWorks, Natick, MA, USA).

For illustration purposes, spatial trajectories of the recorded and decoded joints were used to drive a 3D skeletal model that was scaled to match the anatomy of each primate (Schaffelhofer et al 2015b). The model transformed the joint angles into the 3D positions of a skeleton and thus provided a 3D visualization of the complete hand and arm kinematics (see supplemental videos).
2.7.4. Evaluation of decoding performance. To assess decoding performance, we calculated both Pearson’s correlation coefficient (CC) and the relative root mean squared error (rRMSE) for each DOF across the entire decoded recording session to compare the measured trajectory of a joint angle with the decoded trajectory as predicted by the Kalman filter.

CC was defined as

$$CC = \left[ \frac{\sum_{k=1}^{N} (x_k - \bar{x})(\hat{x}_k - \hat{\bar{x}})}{\left( \sqrt{\frac{\sum_{k=1}^{N} (x_k - \bar{x})^2}{} \sum_{k=1}^{N} (\hat{x}_k - \hat{\bar{x}})^2} \right)} \right]$$,

where $N$ is the total number of samples, $x_k$ and $\hat{x}_k$ denote the true and decoded DOF at time $k$, respectively, and $\bar{x}$ and $\hat{\bar{x}}$ are the means across all $N$ samples of the true and decoded DOF, respectively.

rRMSE was determined by calculating the root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (x_k - \hat{x}_k)^2}$$

for each DOF and then normalizing it by the range of the respective DOF:

$$rRMSE = \frac{RMSE}{(d_{95} - d_{5})},$$

where $d_5$ and $d_{95}$ are the 5th and 95th percentile of all recorded data samples of the respective DOF. This normalization allowed comparing prediction errors of joint angles relative to their operational range.

For each neuronal dataset we determined the optimal lag and bin length combination that maximized CC or minimized rRMSE. Most of the time the optimal parameter combinations for CC and rRMSE matched. CC and rRMSE are both important performance measures, however, a trajectory prediction that captures the movement shape with an offset bias might have a high CC, but will also have a high rRMSE. In cases where the optimal parameter combination did not match for CC and rRMSE, we therefore favoured the parameter combination yielding better rRMSE, which generally also had a high CC. This best parameter combination was then applied to predict the movement kinematics of each dataset.

2.7.5. Chance decoding performance. Kalman filter decoding uses the previous kinematic state as the first part of its prediction (a-priori estimation). In case of a very regular movement, the filter might achieve a high performance by simply utilizing the regularity of the oscillatory kinematics, i.e., without relying on neural information at all. In such a case, a high decoding performance might be falsely attributed to the neural data.

To make such effects transparent and to avoid them as much as possible, we included both correct and incorrect trials (where the kinematics often did not follow the typical temporal structure of the correct trials) in the decoding. Furthermore, we calculated a chance level performance to illustrate prediction accuracy of a decoder relying solely on inherent kinematic information. By comparing the resulting decoding performance to the accuracy obtained with standard decoding, an estimation was obtained of how much the filter was actually utilizing neural data for prediction. To determine a chance performance level, we removed any potential task-relevant information contained in the spike train population by applying the following random shift method: spike sequences of individual units were shifted in time by a random shift length of at least 20 s in either positive or negative time direction. Neural data that was shifted beyond the end or the beginning of the recording interval was circularly reinserted at the beginning or end respectively. This procedure destroyed any potential relationship between the kinematics of the limb and the spike sequence while keeping the inherent temporal structure of the spike sequence intact. Importantly, the random time shifts were different for each
unit, which destroyed also any potential correlations in the neuronal population. Using this temporally reshuffled neural data, we carried out decodings and calculated the performance measures, CC and rRMSE, from which we determined the chance level performance as follows:

By repeating this surrogate method ten times, we established the probability distribution of CC and rRMSE during chance performance for each area and recording session. To test if the CC or rRMSE obtained from the respective standard decoding differed significantly from chance, we compared the respective CC or rRMSE value to the median of the chance probability distribution (two-sided sign test). If $p < 0.05$, we considered the performance of the standard decoding to be significantly different from chance.

When comparing the mean performance across all ten recording sessions to chance (as in the analysis of decoding performance for proximal and distal groups of joints, figure 6), a chance distribution across recording sessions was composed out of the means of the chance distributions of each recording. Then, a two-sided signed-rank test was performed to determine whether the median of that chance distribution differed significantly from the mean performance of the standard decoding across all ten recording sessions.

3. Results

We recorded 27 DOF of hand, wrist, and arm kinematics simultaneously with single and multiunit activity in the hand areas of primate motor (M1), premotor (F5), and parietal cortex (AIP) while animals performed the delayed grasping task. In total, ten recording sessions were analyzed for this study (monkey Z: six sessions, monkey M: four sessions).

3.1. Kinematic and neuronal data

Representative example trajectories of the 27 DOF from recording Z032012 are shown in figure 2 while the monkey grasped and lifted a horizontal bar, a cube, and a small ball. As can be clearly observed, the objects were grasped with different hand and arm configurations, which was an important goal of the task design. In total, about 50 distinct objects of various size and shape were tested (see experimental setup in materials and methods) covering a wide range of hand configurations. This allowed us to detect very small differences in hand kinematics (see also figure 1(e) in materials and methods).

Neural activity was recorded from six 32-channel FMAs that were implanted in M1, F5, and AIP (two each; see materials and methods). Table 1 provides an overview of the number of single and multiunits recorded in each cortical area and each session. An example population is illustrated in figure 2 (bottom) as a rasterplot of neural activity that was re-trained and decoding was restarted with new initial values within the range of the respective DOF (see decoding procedure in materials and methods). This led to brief discontinuities in most curves, since all decoding trajectories were randomly reset at this time instant. However, the Kalman filter managed to bring the trajectory back close to the true amplitude within 300–400 ms, hence illustrating the power of the algorithm.

For illustration purposes, the trajectories of both the recorded and the decoded joints were used to drive a 3D skeletal model (Schaffelhofer et al. 2015b) (see materials and methods). Supplemental video 1 shows the movements of the real arm and hand together with the decoded movements using activity from all three areas while the monkey was lifting three different objects from the mixed turntable. In contrast, supplemental video 2 compares the predicted movements when decoding was done from either of the three areas or from all areas combined.

3.2. Decoding of 27 DOF

Since a clear relationship between neural activity and kinematics could be observed, we decoded 27 DOF of the hand, wrist, and arm continuously over time using single and multiunit activity from M1, F5, and AIP. Figure 3 shows example trajectories (both true and predicted) from recording Z032012 of hand, wrist and shoulder joints over a time course of 80 s while the monkey was lifting objects of the mixed turntable. Decodings were performed with units from either area M1 (figure 3(a)), F5 (figure 3(b)), or AIP (figure 3(c)). For all three areas, the estimated trajectories followed the real one accurately: the timing of both movement onset and termination were captured very precisely. The movement amplitude seemed to be reproduced best when decoding with data from area M1 and was slightly diminished when predicting from areas F5 or AIP. Furthermore, in time periods without movement (inter-trial intervals) the predicted trajectory contained more jittering when decoding from AIP than from F5.

$t_{aw} = 1761.5$ s (marked with a pink triangle on top) is one of the time instants when the data was divided in order to carry out the 7-fold cross-validation: here the Kalman filter was re-trained and decoding was restarted with new initial values within the range of the respective DOF (see decoding procedure in materials and methods). This led to brief discontinuities in most curves, since all decoding trajectories were randomly reset at this time instant. However, the Kalman filter managed to bring the trajectory back close to the true amplitude within 300–400 ms, hence illustrating the power of the algorithm.

For illustration purposes, the trajectories of both the recorded and the decoded joints were used to drive a 3D skeletal model (Schaffelhofer et al. 2015b) (see materials and methods). Supplemental video 1 shows the movements of the real arm and hand together with the decoded movements using activity from all three areas while the monkey was lifting three different objects from the mixed turntable. In contrast, supplemental video 2 compares the predicted movements when decoding was done from either of the three areas or from all areas combined.
Figure 2. Joint angle kinematics of all recorded 27 DOF of recording Z032012 during 3 trials. Hand, wrist, and arm DOF are shown while the monkey grasped a horizontal bar, a cube, and a small ball (objects on top). Task epochs are marked by black vertical lines and labeled (‘C’, cue; ‘P’, planning; ‘G’, ‘go’ signal; ‘H’, hold epoch). Simultaneously recorded neural activity from M1 (79 units), F5 (75 units), and AIP (61 units) is presented below as rasterplots. Each unit is represented by one row of dots, and each dot represents the occurrence of a spike. Color code indicates the firing rate of a unit in Hz.
3.3. Decoding performance of different areas

To quantify the decoding performance of the different areas, we calculated the CC and rRMSE for each DOF over the entire recording session (Z032012) and averaged the CCs across all 27 DOF (figures 3(d) and (e)). As expected, decoding with neural signals from area M1 yielded the best performance in comparison to decoding from F5 or AIP. However, there was no significant difference in the mean between M1 and F5, neither for CC nor rRMSE (one-way ANOVA and Tukey-Kramer multicomparsion test, \( p > 0.05 \)). The similar CC demonstrated that the actual movement could be captured similarly precisely when decoding from M1 or F5. However, predictions made with F5 did not seem to capture the amplitude of the movement as well as when decoding from M1, and there was more noise during resting phases (see also figures 3(a) and (b)), which was reflected in a higher rRMSE for F5. AIP yielded good performance values as well; they did not differ significantly from those obtained from area F5, however, when comparing AIP and M1, a significant difference was found in the mean correlation coefficients (\( p < 0.05 \)).

In addition to predicting from one brain area alone, we also decoded from pairs of areas (AIP&F5 and M1&F5) and with all three areas combined. This allowed us to investigate whether decoding performance could be improved by combining information from different brain areas. Decoding from M1 and F5 in combination increased the mean CC and decreased the mean rRMSE slightly, but not significantly, in comparison to decoding from M1 alone. Similarly, there was no significant improvement when decoding from F5 and AIP in combination, as compared to decoding from F5 alone. As expected, decoding performance was best when all three areas were combined. However, there was no significant improvement in comparison to using only data from M1.

We were curious to evaluate which extent the decoding performance could be attributed to neural information, as opposed to movement information inherent in the kinematics, e.g., as present in rhythmic or oscillatory movements. In other words, we wanted to determine the decoding performance in the absence of any meaningful neuronal information with respect to movement. Therefore, we performed simulated decodings with randomly reshuffled spiking activity in which neural information was temporally uncoupled from movement kinematics to determine a chance level performance (random shift method, see materials and methods for further explanation). This method was repeated ten times for each neuronal dataset (area or combination of areas), and CC and rRMSE were calculated for each DOF and averaged across all 27 DOF. Mean and standard deviation (across ten repetitions) are shown in figures 3(d) and (e) (orange lines). Clearly, decoding with the original neural data significantly outperformed chance performance both in terms of CC and rRMSE (two-sided sign test, see chance decoding performance in materials and methods, \( p < 0.002 \)). This demonstrated that significant movement information was present in these brain areas, which the decoder was able to translate into accurate movement predictions.

Figure 4 illustrates the combined results across all recording sessions (six sessions for monkey Z, four sessions for monkey M; see also table 1). Overall, these results confirm the findings in the example recording of figures 3(d) and (e). When decoding with neural activity from a single area, M1 yielded best results, followed by F5 and AIP; all accuracies differed significantly from each other (one-way ANOVA and Tukey-Kramer multicomparsion analysis, \( p < 0.05 \)). Combining two or three areas for decoding did not improve the decoding performance significantly, as compared to using only the better performing single area.

Finally, we compared the performances between monkeys for each area. There were no differences in means or variances for M1 and F5 (two-sample \( t \)-test, \( p > 0.10 \), and Bartlett’s test, \( p > 0.05 \), respectively). However for area AIP, the variances differed significantly between animals (Bartlett’s test, \( p < 0.04 \)). Together, accurate movement prediction was possible with each of the three cortical areas M1, F5,
Figure 3. Examples of real and decoded trajectories of four example DOF (thumb MCP, index DIP, wrist yaw, and shoulder internal/external rotation) over a duration of 80 s (recording Z032012). Blue lines: recorded trajectories, red lines: decoded kinematics using neural activity from (a) area M1 (bin length 10 ms, lag = 50 ms, 79 units), (b) area F5 (bin length 10 ms, lag = 30 ms, 75 units), and (c) area AIP (bin length 10 ms, lag = 10 ms, 61 units). Decoding performance (CC and rRMSE) during these 80 s is indicated for each trajectory. Pink triangle in (a) indicates (re-)initialization of decoder for cross-validation. Decoding performance across the entire recording session and across all 27 DOF is shown in (d) for CC and (e) for rRMSE for individual areas and combinations thereof. Errorbars show standard deviation across 27 DOF. Orange lines and errorbars in (d)–(e) indicate chance level and standard deviation (see materials and methods).
and AIP, combining these areas increased the decoding accuracy not significantly, and decoding results were highly comparable between monkeys.

3.4. Influence of number of units on decoding performance

Because the number of units was lowest for area AIP (see table 1), an important control was to rule out that AIP’s lower decoding performance was simply due to a lower number of recorded units. To address this question, we randomly selected a fixed number of units from each area M1, F5, and AIP and decoded with this data set. This was done repeatedly (ten times) for each area. Resulting performances were averaged across all 27 DOF and repetitions. Figure 5 shows the mean decoder performance across all ten recording sessions for randomly selecting 5, 15, 25, and 35 units. In addition, the rightmost bars (‘max’) show the decoding performance when, for each recording session, the maximum amount of units available in all of the areas, i.e., the smallest of the three set sizes of units in M1, F5, and AIP, was randomly selected for each session (i.e., 45 units for session Z120511, 61 units for session Z011312, etc, see table 1).

Regardless of the actual number of units, M1 yielded the best decoding performance for both CC and rRMSE, followed by F5 and AIP. Decoding performance (both CC and rRMSE) differed significantly between M1, F5, and AIP, regardless of the amount of units chosen for decoding (one-way ANOVA and Tukey–Kramer multicomparison analysis, p < 0.05). The ranking of areas in terms of decoding performance therefore did not depend on the number of available units. Furthermore, as expected, decoding performance increased with the number of units included for decoding, and in individual recording sessions (data not shown) the standard deviation of CC and rRMSE decreased with growing number of units, hence demonstrating that results became more consistent.

Remarkably, performance was above chance even when using only a small number of units (e.g., five). For each recording, we compared the distribution of performance levels obtained by ten repetitions of randomly selecting five units with the distribution of chance levels obtained by ten repetitions of the random shift method obtained by using all available units in an area (Mann–Whitney U-test): when using M1 for decoding, the performances (both CC and rRMSE) differed significantly from chance for all ten recording sessions (p < 0.001). For F5, there was a significant performance difference to chance in 9 and 10 of 10 recordings for CC and rRMSE, respectively (p < 0.05). For area AIP, decoding with five units was able to yield a significantly better performance than chance in eight out of ten recordings for CC (p < 0.05) and in all recordings for rRMSE (p < 0.05).

We can therefore conclude that all areas carry substantial amount of information about hand and arm kinematics that enabled kinematic prediction with accuracy better than chance performance even with a small number of units. Decoding accuracy monotonically improved with increasing number of units used for decoding, whereas the ranking of areas in decoding performance was independent of the number of units used for decoding.

3.5. Decoding performance of proximal and distal movements

Decoding performance for different groups of joints was compared in figure 6. Although the electrode arrays were carefully placed in the hand area of motor, premotor, and parietal cortex, decoding of elbow and shoulder joints (‘arm joints’ in figure 6) generally performed best for all areas (significantly different from all other groups of DOF except for wrist angles when decoding with AIP and using rRMSE for judgement, one-way ANOVA and Tukey–Kramer multicomparison analysis, p < 0.05). This is likely due to the task...
design, which concentrated on eliciting a high variability in finger kinematics but showed a more stereotypical demand on elbow and shoulder movements. Therefore, these trajectories were easier to predict. We also observed less noise in the recording of those angles, which might have affected the decoding performance. However, since this group consisted of only 4 of all 27 total DOF, its bias towards higher prediction accuracy was only marginal in the overall result, as presented in figure 4.

Besides the difference between proximal and distal joint angles, there was only little variation in the decoding performance between hand and wrist joints (detailed results of significant differences are shown in figure 6). Thumb and little finger joints tended to be predicted with slightly less accuracy than the other DOF of the wrist and hand. Thumb performance could be attributed to the fact that the thumb was the finger with most variation in movement and therefore was also most prone to (kinematic) recording noise. Compared to middle and ring finger that tended to move in correlation (sometimes also together with the index finger), both thumb and small finger movements were more independent. In some grasping conditions, thumb and small finger were not actively involved in the grasp (i.e., they did not touch the object), but moved alongside the other fingers in a more passive way (e.g., when grasping an object that did not require contact with the thumb or little finger). These fingers could then exhibit movement patterns with higher variation across the tested grip types.

In addition, figure 6 shows the respective chance performance levels. As it is clearly visible, decoding performance of all joint groups was significantly better than chance level (obtained by the random shift method; see materials and methods) (two-sided sign test, see chance decoding performance in materials and methods, \( p < 0.002 \)).

3.6. Optimal decoding parameters

As mentioned above, different parameter combinations (varying lag and bin lengths) were tested for decoding and evaluated in terms of CC and rRMSE. Bin lengths of 10–250 ms were systematically combined with lags ranging from 0 to \( \pm 180 \) s and the best combination with respect to CC and rRMSE was determined. This procedure was repeated for all recording sessions and areas. In 95% of these optimal combinations a bin length of 10 ms was the best duration for counting spikes. Optimal lags, however, varied more strongly between areas and recordings. Figures 7(a) and (b) illustrate the decoding performance of M1, F5, and AIP as a function of lag lengths while bin length was kept constant at 10 ms. Furthermore, the cumulative distribution of optimal lag lengths (across recording sessions) for each area is provided.

![Figure 5](image.png)

**Figure 5.** CC and rRMSE was determined for decoding with a fixed number of randomly selected units (5, 15, 25, and 35 units) from each area and averaged across all DOF. This procedure was repeated 10 times for each recording session. Mean CC and rRMSE was determined across all repetitions and recording sessions (bars and errorbars: mean and standard deviation). Different colors indicate the three recording areas (M1, F5, and AIP). In the rightmost group of bars (‘max’), the number of randomly selected units was set to its maximal value for each session.
in figures 7(c) and (d). As described in section 2.7.3, materials and methods, negative lag lengths correspond to neural data that preceded the kinematics to be decoded, whereas positive lags indicate that neural activity followed the kinematics in time. To make this difference more obvious, the discontinuity between −0 and 0 was marked on the x-axis.

For area M1, the lag lengths yielding the best decoding performance for each recording clustered around durations from −70 to 10 ms. A two-sided sign test showed that the median of these lag lengths (−10 ms) was not significantly different from zero for CC (p = 0.07), however it was for rRMSE (median = −50 ms, p = 0.04). For area F5, the optimal lag lengths were shifted more into the negative range, indicating that neural data used for decoding preceded the kinematic prediction: for both CC and rRMSE, the median (−40 ms for CC, −60 ms for rRMSE) was significantly different from zero (p = 0.002 for both CC and rRMSE). However, this was not true for AIP: there, the median of optimal lag lengths for decoding was not significantly different from zero (p = 1 for CC, median = −0 ms; p = 0.45 for rRMSE, median = −15 ms). Together, optimal lag lengths for decoding with area M1 were close to zero with a shift into the negative range, which became even larger for decoding with F5. Optimal lags for decoding with AIP however clustered around zero. This becomes also apparent in figures 7(c) and (d) where the lines for F5 are shifted furthest to the negative range, followed by the lines for M1 and AIP when looking at lag lengths around zero.

Moreover, although medians of optimal lag lengths did not differ significantly between areas (Kruskal–Wallis test, p > 0.05), these plots demonstrate that there was much more variation in optimal lags for AIP than for the other two areas: for CC the respective standard deviation of optimal lag distributions were 30.62 ms (M1), 29.83 ms (F5), and 83.11 ms (AIP). This difference was significant (non-parametric Levene’s test, p < 0.03). Similar results were observed for rRMSE with standard deviations of 37.06 ms (M1), 38.89 ms (F5), and 452.86 ms (AIP). Here AIP was also significantly different (non-parametric Levene’s test, p < 0.01).

Furthermore, when decoding with data from AIP, a second but smaller peak in performance appeared for lags around −1700 ms (figures 7(a) and (b)), and the ranking of the three areas in terms of decoding performance was reversed for these long lags. Contrary to AIP, the performances for F5 and M1 improved steadily the closer the neural data used for decoding lay in time compared to the kinematics to be predicted. For lag lengths longer than approximately −600 ms, F5 still yielded higher decoding accuracy than M1. For positive lags, decoding performance monotonically decreased and there was no second peak in either area.

In conclusion, while consistently observing a bin length of 10 ms as optimal throughout areas, the impact of lags on decoding performance varied largely, and optimal lag lengths yielding highest decoding performance varied between M1, F5, and AIP.
4. Discussion

This paper investigated the possibility of decoding complete hand, wrist, and arm kinematics (represented by 27 DOF) from single and multiunit activity in the hand areas of motor (M1), premotor (F5), and parietal cortex (AIP). To our knowledge, this is the first study that combines and compares these three grasping areas for predicting versatile, continuous hand kinematics. Simultaneous recordings of population activity from these three areas with multi-electrode arrays made it possible to examine differences in decoding performance between these areas and evaluate the information content with respect to hand and arm kinematics.

4.1. Movement reconstruction with primary and premotor cortex

Continuous trajectories of 27 joint angles could be reconstructed accurately over time using single and multiunit activity from M1, F5, or AIP. However, decoding performance varied between these areas. Highest performance was achieved when using M1 for decoding, followed by F5 and AIP (figures 3 and 4). It is known that various parameters regarding hand and arm movements such as joint and muscle representations, wrist position, force, and grasp configurations are encoded in M1 (Thach 1978, Georgopoulos et al 1986, Ashe and Georgopoulos 1994, Taira et al 1996, Ashe 1997, Kakei et al 1999, Rathelot and Strick 2006, Umilta et al 2007) and F5 (Rizzolatti et al 1988, Kakei et al 2001, Raos et al 2006, Fluet et al 2010). Furthermore, F5 has both strong projections to the digit area of M1 (Muakkassa and Strick 1979, Matelli et al 1986, Dum and Strick 2005, Borra et al 2010) as well as direct connections to the spinal cord via cortico-spinal neurons originating from F5 (He et al 1993, Galea and Darian-Smith 1994). F5 has therefore been considered to operate—to some extent—at the same hierarchical level as M1 (Dum and Strick 2005). Based on these findings, it is not surprising that both M1 and F5 were able to predict joint kinematics quite precisely (figures 3 and 4).

The suitability of M1 and F5 for decoding arm and hand kinematics has previously been investigated (Carmena et al 2003, Lebedev et al 2005, Ben Hamed et al 2007, Vargas-Irwin et al 2010, Bansal et al 2012, Aggarwal et al 2013). However, different decoding algorithms, task types and decoding parameters have been employed that makes it difficult to compare these results with ours. Furthermore, the number of decoded DOF was much lower in most of these studies. However, the three following studies were closest related to ours: first, Aggarwal et al (2013) reported decoding performances obtained with activity from M1, PMd, and PMv that closely matched our results for...
predictions from M1 and F5 (PMv) both combined and separately. Second, Vargas-Irwin et al (2010) recorded single unit activity from the hand region of M1 and reported similar decoding accuracy levels for hand kinematics in an object-fetching task for activity from M1. Finally, third, Bansal et al (2012) decoded hand grasping movements separately from premotor and motor cortex and found no significant difference in decoding accuracy between M1 and PMv. However, the reported decoding performance (CC) was lower in their study as compared to ours.

All three studies included circumstances not present in our study that might have influenced the decoding performance in the positive direction: first, in all studies the number of objects used for grasping was substantially lower than in our task, therefore leading to less complex grasp kinematics. Second, these studies used only correctly performed trials or selected time periods of correct trials for prediction, whereas our study predicted all occurring kinematics, including incorrect trials and inter-trial intervals. We deliberately included these epochs in order to mimic a more natural behavior, where also unexpected movements occur in addition to the ones demanded by task design. Both points introduced more variance in the kinematics and increased the difficulty for precise decoding.

Third, the studies of Bansal et al (2012) and Vargas-Irwin et al (2010) used only a selection of best units for decoding and an optimal time lag between individual DOF and neural activity was determined separately for each neuron. In contrast, our study used only a general optimal bin and lag length for the entire neuronal population. Evaluation of unit contribution to kinematic features was left to the automated training of the Kalman filter. Despite all these points, our decoding accuracy was the same, or even higher, than reported in those studies.

Contrary to units in M1, a group of neurons in F5 already responds to object presentation (Murata et al 1997, Raos et al 2006, Lehmann and Scherberger 2013). This activity might have mistakenly been interpreted by the decoder as movement information. Hence, by not only decoding sequences of movement, but also including long periods of resting while the monkey was mentally preparing its grasp, overall decoding performance of F5 might have been decreased compared to M1.

Furthermore, prediction accuracy for using M1 might have benefited from the fact that our M1 electrodes were located in the anterior bank of the central sulcus, as opposed to the cortical surface anterior to the central sulcus, where direct, cortico-motoneuronal units are located (Rathelot and Strick 2009). Neurons in the anterior bank of M1 are therefore closely linked to these cortico-motoneurons, which could explain the suitability of this area for the decoding of hand kinematics and the strong neural selectivity for particular finger movements and grip-types (Schaffelhofer et al 2015a).

4.2. Movement reconstruction with parietal cortex

Decoding with AIP yielded the lowest decoding performance. It is possible that the signal-to-noise ratio of spikes versus background was lower in AIP than in M1 and F5, which might have influenced decoding performance. Despite this fact, decoding accuracy of AIP was significantly above chance (figures 3 and 4), suggesting that AIP carries substantial information about movement kinematics that can be decoded. These signals in AIP do not necessarily have to be movement commands, but could represent other information related to movement execution like object features relevant for grasping. However, the temporal movement characteristics appeared to be precisely captured by AIP, which supports the notion that motor related information in AIP is not exclusively categorical information, like visual features, but also includes temporal aspects of movement execution. In another decoding study where single unit activity from M1, F5, and AIP were used to predict kinematic states such as resting and movement, we found that AIP activity could clearly distinguish between these categories (Menz 2015). Furthermore, movement onset could be detected accurately with single units from AIP, therefore confirming the presence of temporal movement information (Menz 2015).

Predicted trajectories during resting periods often showed undesired jittering when decoding from AIP. This might be accounted for by the fact that AIP neurons respond to the visual presentation of objects and therefore exhibit higher activity already much before the actual movement execution (Sakata et al 1995, Murata et al 2000, Baumann et al 2009, Townsend et al 2011, Lehmann and Scherberger 2013). This visual activity could have affected the decoder and might have caused noisy movement residuals during periods of rest.

To our knowledge, this is the first study investigating area AIP as a potential brain region for the continuous decoding of hand and finger movements. Earlier studies using AIP and F5 only predicted categorical variables like grip types: Lehmann and Scherberger (2013), Schaffelhofer et al (2015a), and Townsend et al (2011) found that F5 was better suited for decoding of grip types than AIP. Furthermore, Townsend et al (2011) were able to observe that decoding from the combined activity of F5 and AIP did not increase decoding performance significantly in comparison to F5 alone. Both points correlate well with our results (figures 3 and 4). The fact that decoding performance did not improve when neural activities from F5 and AIP were combined suggests that AIP largely shares its kinematics-related information with area F5.

4.3. Influence of population size on decoding performance.

We demonstrated that differences in decoding performance between AIP, F5, and M1 were consistent between animals (figure 4) and did not primarily depend on the number of neurons available for decoding, but strongly reflected the type of information encoded in these areas (figure 5). Already a very small number of randomly selected units raised decoding accuracy above chance. As suggested previously (Vargas-Irwin et al 2010), neurons in M1 and F5 do not seem to reflect a specific kinematic variable but instead carry information about a broader range of movement parameters, which would
allow reasonable movement reconstruction even with a few neurons. This seemed to hold true also for units in AIP.

4.4. Reconstruction of proximal and distal joints

Although the task demanded a wide variety of finger movements and hand configurations, no substantial difference was found between decoding finger and wrist joints (figure 6). However, a slight tendency towards a higher decoding performance for index, middle, and ring finger was observed. These fingers also showed higher synergies in their movements. Since the Kalman filter used combined state information of all DOF for predicting a single joint (a priori estimate), more information was available for joints moving in a correlated fashion and therefore the algorithm could predict such DOF more accurately. In contrast, thumb movements were more independent from the other fingers and showed more variation and were therefore harder to predict by the decoder. Furthermore, the little finger was not actively used for grasping in many conditions but rather moved along with the other fingers, often without clasp the object properly. Since it was used more passively, its movement intention might have been underrepresented in the brain, which might explain the slightly reduced decoding performance. For a few objects this was also true for the thumb, like the ones from the ‘special’ turntable (see figure 1(b)), where the grasp was mainly carried out with fingers 2–4 (similarly to the ‘spherical power grasp’ described in Cutkosky and Wright 1986).

In contrast, shoulder and elbow movements were predicted with significantly higher accuracy. Since our task was designed to focus on and eliciting dexterous grasping kinematics, arm movement kinematics were more stereotypical than those of finger and wrist kinematics. This might have facilitated the prediction of shoulder and elbow movements. Nevertheless, by including incorrect trials that were often aborted in the middle or contained unexpected movements, we tried to prevent the occurrence of a stereotypical temporal structure in the kinematics. However, since we recorded from highly trained animals, the number of incorrect trials was generally low.

In conclusion, although the arrays were carefully placed in the hand areas of primary, premotor, and parietal cortex, both distal and proximal joints could be predicted with high accuracy. This finding is supported by the observation of Vargas-Irwin et al (2010) that individual neurons in primary motor cortex encode information about both proximal and distal kinematics. McKiernan et al (1998) showed that a large portion of cortico-motoneuronal cells have muscle fields in both distal and proximal muscles. Since we recorded in the ‘new’ part of M1 (Rathelot and Strick 2009), it is likely that our recorded population contained a considerable portion of such cells.

4.5. Optimal decoding parameters and their network implications

In our study, a bin length of 10 ms yielded the highest decoding performance in 95% of the tested data sets regardless of which brain area was used for prediction. Previous decoding studies used bin lengths of 100–150 ms, however without explaining their choice (Vargas-Irwin et al 2010, Bansal et al 2012, Aggarwal et al 2013). Long time windows will likely include the time point when information content is highest but they also act like low-pass filters. In contrast, with a short window, changes in firing rate can be detected with higher temporal precision, which is advantageous especially for brain areas with lower tuning strength, like AIP. However, it is crucial to combine the bin length with an adequate time lag to find the optimal relation between brain activity and actual movement execution. Indeed, we found the time point from which data is used for decoding to show a high impact on decoding performance, and the optimal time lag varied considerably for the three areas (figure 7).

When decoding from M1, we found the highest amount of information being present when neural data preceded the hand kinematics by 0–50 ms. Similar results were also reported by Morrow and Miller (2003), who decoded muscle activity from neurons in M1 with different lags. Their peak performance was obtained with a lag of –50 ms, it dropped faster than in our study when lags deviated from the optimal one, and it levelled off at a time lag of around ±250 ms. The faster decrease might be due to the use of a different decoding algorithm, or that sequentially recorded neuronal activity was used for decoding. Additional information encoded in the coherence or phase synchrony within the neuronal population (Averbeck and Lee 2004) was therefore not contained in their data but was available to our decoder due to simultaneous array recordings. This might have helped maintain a relatively high decoding performance for longer than optimal lag lengths. For a time lag of around ±500 ms however, their and our decoding performance was comparable.

Other studies determining the latency between spikes of cortico-motoneuronal cells in primary motor cortex and the onset of muscle activity of distal muscles reported an onset of postspike effects at about 10 ms after neuronal activity (Fetz and Cheney 1980, Kasser and Cheney 1985, Lemon et al 1986, McKiernan et al 1998). Since we recorded in the bank of M1 where cortico-motoneuronal cells are found (Rathelot and Strick 2009), our M1 population might be more directly linked to these cells, which might explain a tendency towards shorter lags in our, as compared to Morrow and Miller’s results.

Optimal time lags for F5 were longer than for M1. However, if neural activity preceded kinematics by long negative lags, F5 yielded higher decoding accuracies than M1. This indicates that grasp intentions are present in F5 earlier than in M1, in agreement with Umiltà et al (2007).

Furthermore, our time lag results are in line with the fronto-parietal grasp network hypothesis (Jeannerod et al 1995), which states that visual information about an object to be grasped is projected from AIP to F5 where a grip type is selected. Grip type information is then forwarded to neurons in M1 and translated into a motor command that is forwarded to the spinal cord. However, the role of AIP in this process seems to be more intricate. Decoding performance
peaked for lags close to zero, however, we also found an additional smaller peak for time lags of about −1700 ms. It is very unlikely that in natural grasping movements, AIP represents movement intentions more than 1.5 s before execution. Instead, this second peak could be caused by the visual task instruction: the object is illuminated approximately 1.5 s before movement onset, which is known to cause a visual activation of AIP (Sakata et al 1995, Schaffelhofer et al 2015a). Furthermore, Schaffelhofer et al (2015a) showed that object information (and to a less extent grip type) could be decoded from AIP activity during the cue epoch of the task. Clearly, object features and hand configurations were coupled in our task, therefore cue information could have been used by the decoder for movement prediction. However, since visual information could only be used as a rough guideline for the prediction of individual movement kinematics, decoding performance was lower for these long lags compared to the accuracy peak when decoding with very short time lags.

Although decoding performance was highest for AIP for lags around zero, it is unlikely that a direct movement command is encoded in that neural activity. Mulliken et al (2008) found similar short lag lengths to yield the highest amount of mutual information between units in the PPC and movement parameters. They argued that PPC neurons could encode a forward estimate of the movement and are therefore not only involved in movement planning, but also in online movement control that utilizes an efference copy from premotor and motor cortices (Mulliken et al 2008). The same is likely the case for AIP. It has been suggested that so-called motor neurons in AIP could represent a copy of the motor plan from F5 in order to compare it to the visual object properties and for subsequent fine-tuning of the motor command (Sakata et al 1995, Sakata et al 1997, Sakata et al 1999, Murata et al 2000). Our findings are compatible with this hypothesis.

5. Conclusion

This study shows for the first time that high-dimensional reaching and grasping movements could be decoded continuously over time with high precision not only from areas M1 and F5 but also from the parietal area AIP. Although the decoding performance from AIP was inferior to that from motor and premotor cortex, our results support the idea that AIP not only encodes categorical visual information, as shown previously by other studies, but that AIP also contains temporal movement information, at least to some extent. However, kinematic information in AIP seems to be redundant to movement intentions encoded in area F5.

Acknowledgments

We thank R Ahlert and N Nazarenus for assistance in animal training, M Dörge for technical assistance, and J A Michaels for comments on an earlier version of the manuscript. This work was supported by the BMBF (BCCN II, DPZ 01GQ1005C) and the European Union (FP7-611687; NEBIAS).

References

Anderson K D 2004 Targeting recovery: priorities of the spinal cord-injured population J. Neurotrauma 21 1371–83
Ashe J 1997 Force and the motor cortex Behav. Brain Res. 87 255–69
Averbeck B B and Lee D 2004 Coding and transmission of information by neural ensembles Trends Neurosci. 27 225–30
Ben Hamed S, Schieber M H and Pouget A 2007 Decoding M1 neurons during multiple finger movements J. Neurophysiol. 98 327–33
Cutkosky M R and Wright P 1986 Modeling manufacturing grips and correlations with the design of robotic hands Proc. of the IEEE Int. Conf. on Robotics and Automation vol 3 pp 1533–9
Dum R P and Strick P L 2005 Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere J. Neurosci. 25 1375–86
Faugier-Grimaud S, Frenois C and Stein D G 1978 Effects of posterior parietal lesions on visually guided behavior in monkeys Neuropsychologia 16 151–68
Galea M P and Darian-Smith I 1994 Multiple corticospinal neuron populations in the macaque monkey are specified by their unique cortical origins, spinal terminations, and connections between J. Cereb. Cortex 4 166–94
He S Q, Dum R P and Strick P L 1993 Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere J. Neurosci. 13 952–80
Hochberg L R et al 2012 Reach and grasp by people with tetraplegia using a neurally controlled robotic arm Nature 485 372–5
Holden R N and Miller L E 2002 Primary motor cortical neurons encode functional muscle synergies Exp. Brain Res. 146 115–23
Kalman R E 1960 A new approach to linear filtering and prediction problems J. Basic Eng-T ASME 82. 35–45
Lehmans J S and Scherberger H 2013 Reach and gaze representations in macaque parietal and premotor grasp areas J. Neurosci. 33 7038–49
Lemon R N, Mantel G W and Muir R B 1986 Corticospinal facilitation of hand muscles during voluntary movement in the conscious monkey J. Physiol. 381 497–527
Menz V K 2015 Continuous detection and prediction of grasp states and kinematics from primate motor, premotor, and parietal cortex Dissertation University of Göttingen
Muakkassa K F and Strick P L 1979 Frontal lobe inputs to primate motor cortex; evidence for four somatotopically organized ‘premotor’ areas Brain Res. 177 176–82
Murata A, Fadiga L, Gallese V, Raos V and Rizzolatti G 1997 Object representation in the ventral premotor cortex (area F5) of the monkey J. Neurophysiol. 78 2226–30
Quiroga R Q, Nadasey Z and Ben-Shaul Y 2004 Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering Neural Comput. 16 1661–87
Rizzolatti G and Luppino G 2001 The cortical motor system Neuron 31 889–901
Schaffelhofer S, Agudelo-Toro A and Scherberger H 2015a Decoding a wide range of hand configurations from macaque motor, premotor, and parietal cortices J. Neurosci. 35 1068–81
Schaffelhofer S and Scherberger H 2012 A new method of accurate hand- and arm-tracking for small primates J. Neural Eng. 9 026025
Stark E, Asher I and Abeles M 2007 Encoding of reach and grasp by single neurons in premotor cortex is independent of recording site J. Neurophysiol. 97 3351–64
Thach W T 1978 Correlation of neural discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum J. Neurophysiol. 41 654–76
Welch G and Bishop G 2006 An Introduction to the Kalman Filter (Chapel Hill, NC: University of North Carolina Press)