A system for automated tracking of motor components in neurophysiological research

Palmér, Tobias; Tamté, Martin; Enqvist, Olof; Halje, Pär; Petersson, Per

Published in:
Journal of Neuroscience Methods

DOI:
10.1016/j.jneumeth.2012.01.008

2012

This manuscript is a Research article. It contains 31 pages, whereof 5 are figures and 1 is tables.

A system for automated tracking of motor components in neurophysiological research

Tobias Palmér[a,1], Martin Tamtè [a,*,1], Pär Halje[a], Olof Enqvist[b] and Per Petersson[a]

[a]Neuronano Research Center, Department of Experimental Medical Science, BMC F10, Lund University, 221 84 Lund, Sweden

[b]Center for Mathematical Sciences, Lund University, 22100 Lund, Sweden

[1] These authors contributed equally to this work.

All authors e-mail addresses:

*Martin Tamtè: Martin.tamte@med.lu.se

Tobias Palmér: Tobias.palmer@med.lu.se

Pär Halje: Par.Halje@med.lu.se

Olof Enqvist: Olofe@maths.lth.se

Per Petersson: Per.petersson@med.lu.se

∗ Corresponding author at:

Neuronano Research Center (NRC)

Department of Experimental Medical Sciences

Lund University

BMC F10

221 84 Lund

Sweden

Phone: +46-46 222 14 64

Fax: +46-46-222 05 78

E-mail address: Martin.Tamte@med.lu.se

The authors report no conflict of interest

Abstract

In the study of motor systems it is often necessary to track the movements of an experimental animal in great detail to allow for interpretation of recorded brain signals corresponding to different control signals. This task becomes increasingly difficult when analyzing complex compound movements in freely moving animals. One example of a complex motor behavior that can be studied in rodents is the skilled reaching test where animals are trained to use their forepaws to grasp small food objects, in many ways similar to human hand use. To fully exploit this model in neurophysiological research it is desirable to describe the kinematics at the level of movements around individual joints in 3D space since this permits analyses of how neuronal control signals relate to complex movement patterns. To this end, we have developed an automated system that estimates the paw pose using an anatomical paw model and recorded video images from six different image planes in rats chronically implanted with recording electrodes in neuronal circuits involved in selection and execution of forelimb movements. The kinematic description provided by the system allowed for a decomposition of reaching movements into a subset of motor components. Interestingly, firing rates of individual neurons were found to be modulated in relation to the actuation of these motor components suggesting that sets of motor primitives may constitute building blocks for the encoding of movement commands in motor circuits. The designed system will, thus, enable a more detailed analytical approach in neurophysiological studies of motor systems.

1. Introduction

The central nervous system fundamentally deals with the control of actions. Consequently behavioral studies have often been a natural starting point for investigations aimed at understanding its functions. For the same reason, the search for new therapies for neurological and psychiatric diseases largely depend on animal models designed to mimic certain aspects of the disease that cause observable changes in the behavior of the subject. With the more recent development of techniques allowing for simultaneous recording of neuronal activity in many parts of the central nervous system in freely behaving animals, the electrophysiological processes underlying such changes in behavior - or even the generation of specific components of observed actions - have the potential to be investigated in much greater detail (Nicolelis, 2008). The access to neuronal data with sub-millisecond temporal precision in turn further increases the need for more detailed documentation of movement patterns displayed by freely behaving animals. However, because natural behavior typically involves chains of movement sequences incorporating many partially overlapping motor components, an extra challenge in this respect is to reliably identify and isolate the execution of these individual motor elements. The most common approach for behavioral recording in neurophysiological research is probably the use of digital video techniques, where image sequences are obtained from different camera angles and specific behaviors are either manually identified off-line (Cenci and Lundblad, 2007; Whishaw et al., 1999) or, when clearly visible in any of the cameras, automatically identified and quantified from this viewing angle (Peikon et al., 2009; Vorhees et al., 1992). In situations where movements involve several joints the problem of tracking motor components involving for example small angle changes in distally located joints becomes increasingly complex. A well-studied and functionally very important example of movement sequences involving parallel movements in multiple joints in humans is the skilled arm and hand movements involved in reaching for and grasping of different objects (de Bruin et al., 2008; Gentilucci et al., 1997; Jeannerod, 1984). Perhaps surprisingly, rodents can after extensive training also perform reaching and grasping movements using their forepaw in many ways similar to a human hand, hence making this behavior particularly well suited for studies on skilled motor control in translational research (Peterson, 1934; Sacrey et al., 2009).

To be able to track forelimb movements with high fidelity in the skilled reaching task we designed a system that uses a three-dimensional (3D) model of the paw for which movements are reconstructed based on image sequences recorded from multiple viewing angles. Each paw pose is estimated by an optimization procedure that maximizes a matching quality measure in order to retrieve the best approximation of that pose. The matching quality is measured as the discrepancy between projections of the 3D model onto the image planes and the actual images, using edges and silhouettes as cues. We here describe how this system allows us to correlate single unit activity of neurons in corticostriatal circuits in rats to different motor components in the reaching-grasping sequence, opening up for significantly more detailed analyses of skilled movement control.

2. Materials and Methods

2.1. Animals

One adult female Sprague-Dawley rat (230g; Taconic Inc.) was used in the study. The animal was kept on 12:12 h light cycle and received food and water *ad libitum* except for a 22 h-period prior to each testing session during which no food was provided. After each testing session the animal had free access to food for 1 h. All experiments were approved in advance by the Malmö/Lund ethical committee of animal experiments.

2.2. Training protocol

The rat was trained during a three week period prior to the implantation of recording electrodes according to the protocol described by Whishaw and colleagues (Whishaw et al., 2008). Training entailed habituation of the rat to the apparatus, habituation to the food reward (pellet) and establishment of paw dominance. The training continued until performance no longer improved between sessions, reflecting a fully learned behavior (Hermer-Vazquez and Moshtagh, 2009). See Appendix A for details.

2.3. Electrodes

Formvar-insulated 33 µm tungsten wires (CFW Inc.) were arranged into four separate 4x5 arrays with 250 µm spacing between adjacent wires. Each array consisted of 16 recording channels and two reference channels, as well as two blind channels. The wires of each array were cut to the appropriate length for the corresponding recording site (cortical or striatal). Reference wires were cut ~1 mm shorter than the recording wires and de-insulated ~300 µm at the tip, positioning them dorsally to the recording site (at the cortical surface and within the corpus callosum, for cortical and striatal arrays, respectively). The wires were attached to board-to-board-connectors (Kyocera 5602) with conducting epoxy (Epotek EE 129-4), and linked to the acquisition device via a board-to-Omnetics connector adapter (Kyocera 5602; Omnetics). A 200 µm silver wire was used as animal ground via direct connection to four screws inserted into the cranium.

2.4. Surgery

Implantations were performed under Fentalyl/Medetomidine anesthesia (0.3/0.3 mg/kg, i.p.). Electrodes were implanted in the forelimb area of the primary motor cortex (center coordinates: AP: +1.5, ML: ±2.8, DV: -1.0 from cortical surface, Donoghue and Wise, 1982) and of the striatum, (center coordinates: AP: +0.2, ML: ±3.8, DV: -3.5 from cortical surface, West et al., 1990) in both hemispheres. The implant was fixated with dental acrylic attaching to screws in the skull. After surgery the anesthesia was reversed by Atipamezole hydrochloride (5 mg/kg, i.p.). Buprenorphine (0.5 mg/kg, s.c.) was administered as postoperative analgesic. The animal was allowed to recover for a week after implantation before testing commenced.

2.5. Experimental set-up

The testing apparatus for the reaching task consisted of a 450x140x350 mm (l/w/h) transparent Plexiglas cage with a 13 mm wide aperture at the middle of one of the short sides (vertical position of aperture: 40 to 150 mm above the ground). Outside the

aperture a 30 mm deep shelf was positioned. To facilitate the placement of food pellets, three separate hemispherical indentations (5 mm in diameter) were made in the shelf 15 mm from the outer edge of the slit. The middle pocket was positioned right in front of the aperture with the other two pockets centered 6.5 mm more lateral on each side. This configuration prevented the rat from using its tongue to acquire the pellet. Furthermore, it permitted the experimenter to decide which paw the rat had to use, as this geometry allows only reaches with the paw contralateral to the side pockets (for further details, see Whishaw and Pellis, 1990). At the center of the cage was a 40 mm high solid obstacle that enforced a forelimb stepping movement similar to the actual reaching and grasping movement, for comparison of similar movements with different purposes.

2.6. Reaching task

In the behavioral task the rat was placed in the reaching apparatus from where it could obtain 45 mg food pellet rewards, positioned in the indentation on the reward shelf, via controlled reaches through the aperture of the wall. A trial ended either if the rat acquired the pellet after one or several reach attempts, or if the pellet at any time was moved from its original position, in which case the food pellet was manually removed by the experimenter. In order to produce discrete reaching trials, the rat was trained to return to the end of the cage opposite to the reaching slit before it was presented with another food pellet, requiring the animal to reposition before every trial. Moreover, by semi-randomly withholding food the rat was prompted to identify the presence of a pellet before each reach attempt, yielding maximal accuracy of each skilled reach (Appendix A).

2.7. Acquisition of neurophysiological signals

Extracellular neuronal recordings were acquired using a multichannel recording system with Cheetah software (Neuralynx Inc.) and digitized at 32 kHz per channel. Local field potentials were bandpass filtered between 0.1-300 Hz (not used in this study) and action potential waveforms between 600-9000 Hz.

2.8. Video acquisition systems

The details of the paw movement during the reach and grasp behavior were captured by two front-view cameras (CMOS, 640x480 pixels; Dalsa Inc.) positioned close to the aperture. Additionally, three mirrors positioned along the edges of the reward shelf gave two extra viewing planes for each front-view camera. Thus, the front-view cameras and mirrors were mounted such that six complementary viewing planes covered the region of interest where the rat forelimb was moving. Extra care was taken to avoid uneven light conditions or reflexes from surfaces. To ensure that inadvertent variations in the camera positions would not influence the 3D reconstruction a calibration procedure with an object of known measures was performed at the start of each recording. This allowed for off-line determination of the mathematical function that best described the projections of 3D objects onto the image planes. The apparatus also included one side camera recording the movements of the rat within the box (Stingray, 640x480 pixels; Allied Vision Technologies).

The two front-view cameras were triggered by an external pulse generator (Master 8, A.M.P.I) acquiring images at 200 Hz during a time period manually controlled by a switch during the ongoing experiment, while the side-view camera continuously acquired images at 50 Hz. A set of images acquired from this set of cameras at the same moment in time is in the following referred to as a *multiframe*. In order to ensure perfect temporal alignment between the acquired images and the neuronal recordings, the timestamp for each multiframe was also stored in the Neuralynx multichannel recording system. Uncompressed image data was sampled from the three cameras on three separate computers, using Common Vision Blox software (Stemmer Imaging, Gmbh) and FireView (Allied Vision Technologies), for the front-view and side-view cameras, respectively. Post-acquisition compression of stored data was made for reduction of data size using standard DivX-codecs.

2.9. Software implementation

The analysis software was implemented using an object-oriented approach in Matlab (MathWorks) along with mex-functions (C code compiled in Matlab) that handled a few frequently used low-level functions [*the program can be obtained from the authors on request*]. For derivation of equations used in projective geometry see Appendix B.

2.10. Paw model

The rat paw was modeled as a set of 13 elements of which one element represents the palm of the paw and the other 12 represent the phalanges (three for each of the four long digits - the most radial digit, corresponding to the human thumb, was not included in the model due to its particularly reduced size in rodents). Each element was described as an ellipsoid joined to the adjacent elements at defined points, corresponding to anatomical bones and joints. The size and position of the paw elements was manually adjusted to fit a high resolution 3D-image of the paw. To allow for individual differences in paw size, overall scaling of the model paw was performed for each animal (system robustness to choice of parameter values are presented in Appendix C). A few logical constraints based on physiological limitations were introduced to narrow down searches for possible paw poses. For example, the most distal phalanx of the paw is especially small which makes the tracking of movements in the third (distal interphalangeal; DIP) joint of the digit less reliable. This joint was therefore assumed to have a flexion/extension parameter determined by the physiological influence of the second (proximal interphalangeal; PIP) joint; in detail, the DIP-joint was set to have the same angle as the PIP-joint, with a flexion limit of 60 degrees, measured from its fully extended state (Landsmeer, 1963).

After decreasing the degrees of freedom, the included parameters were the following:

i) Three parameters for wrist position in space

ii) Three parameters for wrist movements

iii) One parameter for opposition of the palm due to movement of the metacarpal bones

iv) One parameter for adduction/abduction for each digit

v) Two parameters for flexion/extension of the first two joints of each digit (the metacarpophalangeal; MCP-joint, and the PIP-joint).

These parameters add up to a total of 19 degrees of freedom that together define the movements of the model paw. The most likely paw pose in 3D space could then be estimated by adjusting the orientation of the ellipsoid elements around their respective movement axes and projecting the composite pose onto calculated image planes corresponding to the camera and mirror positions (Fig.1).

2.11. Paw pose estimation

For each multiframe, an iterative optimization procedure was employed to estimate the paw pose in terms of the parameter set described in 2.10. In order to find the initial set of parameters, we employed a database containing the parameter sets corresponding to the most commonly encountered poses (see 2.11.3). The search algorithm then generated a number of hypothetical poses and chose the one that best matched the captured video in each iterative step. The following section describes how matching quality was measured and Section 2.11.2 how hypothetical poses were generated in the search algorithm.

2.11.1. Matching quality

To measure the matching quality of a given pose with respect to the video images we used a combination of two matching quality measures - first, the silhouette matching quality, measuring how well the video image silhouette is explained by the given pose, and second, the edge matching quality, measuring how well the video image edges are explained by the pose (Oikonomidis et al., 2011).

Prior to each reaching session, an image of the static background was captured. This image was then used to detect foreground objects by so called background subtraction. Pixels deviating significantly from the background image are likely to belong to foreground objects such as the forelimb, the food pellet or the snout. Hence, pixels

with a deviation over a predefined threshold are classified as foreground. This information was stored in a binary silhouette image $\boldsymbol{S}$, where $S_{ij} = 1$ means foreground and $S_{ij} = 0$ means background for the pixel at (i, j).

If we then consider a given pose, defined by the values for all the 19 parameters denoted in section 2.10, and project the 3D model of the pose onto a given image plane, we obtain a predicted silhouette image $\widehat{\boldsymbol{S}}$. If the model parameters are correct, matrices $\boldsymbol{S}$ and $\widehat{\boldsymbol{S}}$ should be similar. To quantify this we used the Jaccard index, being the number of pixels that are equal to 1 in both of the matrices $\boldsymbol{S}$ and $\widehat{\boldsymbol{S}}$, divided by the number of pixels equal to 1 in at least one of the matrices $\boldsymbol{S}$ or $\widehat{\boldsymbol{S}}$. Thus, the silhouette matching quality was defined as

$$q_S = \frac{\sum_{i,j} S_{ij}\hat{S}_{ij}}{\sum_{i,j} S_{ij} \| \hat{S}_{ij}} \tag{1}$$

where || is the logical OR-operator.

To measure the edge matching quality, an edge image $\boldsymbol{E}$ was computed from $\boldsymbol{S}$ in the following way. First, the 8-neighborhood (horizontal, vertical and diagonal) of each foreground pixel is analyzed. If at least one of the pixels in the neighborhood is background, the pixel is considered to be an edge pixel. The edge orientation in that point is approximated as the mean angle from the center pixel to all background pixels in the neighborhood. Depending on this angle, the corresponding element in the edge image $\boldsymbol{E}$ is set to either -1 (for angles from 0 to π) or +1 (for angles from π to 0). If there are no background pixels in the neighborhood the corresponding element is set to 0. Finally, the measured edge image $\boldsymbol{E}$ is filtered with a Gaussian filter to increase robustness.

Analogously, a predicted edge image $\widehat{\boldsymbol{E}}$ was constructed from $\widehat{\boldsymbol{S}}$, but without the Gaussian filtering. Edge matching quality was then defined as:

$$q_E = \frac{\sqrt{max\left\{0, \sum_{i,j} E_{ij}\hat{E}_{ij}\right\}}}{\sum_{i,j} |\hat{E}_{ij}|} \tag{2}$$

Both q_S and q_E yield values between 0 and 1, where 1 corresponds to a perfect match. The combined matching quality for each view was then obtained as

$$q = q_S q_E \quad (3)$$

Finally, the total matching quality was defined as the average of the combined matching quality over all viewing planes.

2.11.2. Generating hypotheses

The 19 parameters describing paw poses were divided into five subsets which were processed consecutively, as an exhaustive search over all combinations of the 19 parameters would be computationally unfeasible. At each step of the search algorithm, optimization was performed over one of these subsets. To illustrate how this optimization is carried out we here consider two of the subsets. The others follow the same principles.

Search algorithm I - Flexion/extension optimization

The algorithm searches for the best combination of flexion/extension parameter values for a pose with their current values as the starting point.

1. Create n_1 hypotheses for the flexion at the first joint (MCP) of digit 2 in an interval centered at the current flexion value.

2. Similarly, create n_2 hypotheses for the second joint (PIP).

3. For each of the $n_1 n_2$ pairs of hypotheses, compute the fraction of the major axes of the three ellipsoids that overlaps with areas classified as foreground when projected onto all the different camera image planes. If below some threshold, discard the hypothesis.

4. For the remaining hypotheses, generate silhouette and edge images for digit 2 from the 3D model and evaluate how much of the silhouette and edges that correspond to the observed counterparts. Keep the m best hypotheses.

5. Repeat step 1-4 for digits 3, 4, and 5.

6. For all m^4 combinations of flexion parameters, evaluate the total matching quality and assume the best combination as the result of optimization.

Search algorithm II - Wrist position optimization

The algorithm searches for the best position of the wrist with the current position as the starting point.

1. Sample n values for each of the three spatial dimensions in intervals centered on current position values. The possible combinations of these values lead to n^3 hypothetical poses.

2. For each of the n^3 hypothetical poses, compute the fraction of the major axes of the ellipsoids of the digits that overlaps with areas classified as foreground when projected onto all the different camera image planes. If below some threshold, discard the hypothesis.

3. For all remaining hypotheses, evaluate the total matching quality and assume the best combination as the result of the optimization.

Optimization was iterated over the five subspaces until no further improvement of matching quality was found for any of the different search spaces. Control experiment confirmed that differences in matching quality of final pose estimates depending on the specific search order used were small (Appendix C).

2.11.3. Pose estimation initialization

The database employed to initialize the optimization procedure for paw pose estimation contained the parameter sets corresponding to the most commonly encountered paw poses. In addition, poses that had previously proven difficult to estimate were also included (such poses were identified by poor matching quality in earlier tracking results and were semi-automatically re-estimated and manually verified to have high matching quality before being added to the database). As a starting point, each search was initialized through selection of a set of parameters from the database based on the optimal parameter set found in the previous step. The entire

database was used if no such parameter set existed. Heuristic functions were employed to decrease the number of parameter sets to a more limited number and the set with best matching quality according to Eq. 3 was chosen for initialization. This procedure contributed to an increase in robustness and speed of the subsequent search algorithms.

2.12. Analysis of single unit activity

Spiking activity in each channel of the extracellular recordings was separated into single unit (SU) or multi unit (MU) activity through manual spike feature based sorting techniques (Offline Sorter, Plexon Inc.). To isolate action potentials from a single unit, waveforms plotted in principal component feature space (PC1-PC3) were required to form a cluster well separated from both noise and other units. A refractory period of 1.6 ms was assumed and was used as a control criterion for isolated SUs. SUs encountered in the same channel across days were assumed to be the same if they had similar waveform and firing dynamics. The video tracking gave us an extensive parametrical description of the whole reaching and grasping movement, allowing flexibility in how to align and compare different reaching trials. We chose to use the time of maximal forelimb extension as our temporal reference since that permitted us to include and compare also those attempts that occurred after a failed first attempt. This was important for the subsequent neurophysiological analysis, where a main goal was to study how variations in the reaching and grasping movements were reflected in neural activity, and these higher-order attempts added substantial variation to our data set. Once the times of maximal extension were determined standardized firing rates were estimated by convoluting the spike trains with a Gaussian kernel (σ=30 ms, step size=10 ms, "*psth*" function in the Chronux toolbox, Mitra and Bokil, 2008) and normalizing them to the standard deviation of the baseline (-1500 ms to -500 ms). Paired, two-sided Wilcoxon signed rank tests were used to test for significant deviations from baseline activity ($p<0.05$, "*signrank*" function in Matlab 2010b)

3. Results

3.1. Acquisition of experimental data

Behavioral and neuronal data were acquired in eight sessions on different days, where each session continued as long as the rat showed interest for the task. Each of the eight sessions contained on average 52±37 (mean±SD) reaching trials, which in unsuccessful trials often included several additional reach attempts.

3.2. Reliability of pose estimation

It was initially confirmed through visual inspection of a large number of reaching sequences that the calculated paw poses corresponded very well to the subjective estimates of the poses as judged by the good fitting of the pose projection onto the video images from the different cameras. A typical image sequence from two camera views and the projection of the calculated pose of the 3D model onto the corresponding image planes are shown in Fig. 2A. Obtaining a more quantitative measure of tracking performance is however a more challenging task since it essentially requires *ground truth* information on the exact paw pose; that is, direct measurements on the physical paw (Erol et al., 2007; Ho et al., 2011) which is not possible in freely moving animals. We therefore instead evaluated the performance of the system quantitatively using a paraformaldehyde fixated rat forepaw with static joint angles that was manually positioned in many different spatial locations typical of normal reaching movements in the set-up. The 'true' joint angles of the fixated paw were determined in a separate procedure from a large number of photos taken from different angles that were used to create a visual hull of the paw surface to which the 3D model could be manually fitted with high precision (Fig. 2B and C)

The joint angles of the pose estimates generated by the system for the fixated paw (one for each paw positioning) was shown to deviate on average only 2.4±2.4% of the full movement range for each joint compared to the reference pose obtained by manual fitting of the 3D model to the visual hull (in Cartesian coordinates this error corresponds to a spatial shift of the joint positions of 0.29±0.22 mm [being in the same range as the camera resolution for an individual image plane: ~5 pixels/mm]). In the histogram in Fig. 2D, the mean error of all the estimated poses compared to the reference pose are summarized and an example of a relatively poor estimate (worse

than 95% of the estimates) is shown to illustrate the degree of resemblance in this situation (Fig. 2E). Thus, it appeared that at least in this simplified testing paradigm the system provided accurate pose estimates for all the positions tested. It was also confirmed that the matching scores generated by the quality function used to assess the goodness of pose fitting (2.11.1) had a strong inverse correlation with the calculated distance from the reference pose [Pearson's correlation coefficient: $\rho = -0.85$, $p \ll 0.001$].

Furthermore, computer simulations were employed to evaluate the robustness of the search algorithm to variations of free parameter values (such as foreground/background thresholds or size scaling of the model paw) and sensitivity to input noise (for example light reflections or variations in illumination). Such noise will typically result in segmentation errors. In the simulations performed the system overall displayed a stable performance in situations resembling actual recording conditions and using parameter values within the normally used range. Outside a given interval, system performance would then gradually deteriorate. These data are summarized in Appendix C.

3.3 Identification of motor components

In most types of natural behavior several, partially overlapping, movement components are combined into compound motor actions. In previous studies using manual classification of movements, a limited number of heuristically defined motor components have been proposed to make up the reaching and grasping movements of rats performing this task (Gholamrezaei and Whishaw, 2009; Whishaw and Pellis, 1990). As a starting point for identification of motor components we therefore divided a subset of recorded reaching trials into movement sequences incorporating four of the previously suggested phases: 1) *Advancement*, 2) *Arpeggio*, 3) *Grasping* and 4) *Retraction*. Due to variations in reaching speed, the number of multiframes in each phase often varied slightly between different trials. However, the start and end points of each phase were generally easily identifiable. For all the 19 degrees of freedom, the parameter values obtained from the multiframe corresponding to the start point of each phase were subtracted from that corresponding to the end point. Averaging these

parameter differences thus generated four 19-dimensional movement vectors representing the joint movements of each phase. By considering Retraction as backward Advancement, the Retraction vector could be replaced by the negative Advancement vector. The remaining three vectors were subsequently used as base vectors of the motor components, spanning a 3D motor component space. To eliminate tracking noise and to further separate the different motor components, any non-zero vector element corresponding to movement in a joint that was judged not to be part of a given motor component was set to zero - for example wrist movements were not considered to be part of grasping. Examples of calculated motor components are shown in Fig. 3. Note that even though each motor component was extracted from movements of a certain part of the compound action sequence, they were often flexibly combined with other motor components during the reaching and grasping movement.

3.4 Reaching movements described in motor component space

By reducing the 19 degrees of freedom to this much smaller set of motor components, each reach attempt could be relatively well described as a trajectory in 3D motor component space. The reduction was achieved through a least-squares approximation of each 19-dimensional vector to a linear combination of the three base vectors, which results in a unique point on the trajectory in the 3D motor component space. Interestingly, in spite of the reduced dimensionality this compact 3D representation of the compound movements proved sufficient to separate different reaching and grasping patterns in most reaching trials. In fact, in many reaching trials the three motor components were executed in close to sequential order giving rise to line segments of movement trajectories running almost parallel with the axes in 3D motor component space. Ten individual reaching trials recorded during a single session are shown in Fig. 4. Note that although no single trial is identical to the others, trajectories from successful trials (green) are spatially separated from failures (red) in this representation (shadow images on the sides of the box denote one standard deviation from the mean trajectory of each of the two groups).

3.5 Modulation of neuronal firing correlated with actuation of motor components

The neural recordings included single- and multiunit activity from the primary motor cortex and dorsolateral striatum. As expected, several units displayed a wide range of different task related firing rate modulations. This included phasic increase in activity before, during or after the task, but also partial or total suppression of activity during the task. Using a comparatively long time window of -500 ms to 200 ms (relative to the time of maximal forelimb extension) when comparing firing rates to baseline (-1500 ms to -500 ms) we found that the overall fraction of task related SU neurons was 35 % (29/83). The corresponding number for MUs was 13% (5/38). To search for neuronal modulation specifically linked to the actuation of the identified motor components, the reaching attempts were sorted with respect to the maximum value of each component in each attempt and data were divided into three equally sized groups (low/middle/high). That meant that each reach attempt was assigned to either low, middle or high three times (once for each component) so that a specific reach attempt could be, for example, low in the advancement, middle in the arpeggio and high in the grasp component. Group dependent rate modulations were subsequently tested for by comparing differences between the medians of the three groups (Kruskal-Wallis test, 2 d.f.) in 10 ms time steps (see 2.12 for details). Using the distinction that a cell was considered significantly modulated in relation to a given motor component if the null hypothesis that the medians were identical could be rejected at the $p=0.05$ level in at least eight consecutive time steps, it was found that 11%, 13% and 10% of the SUs were significantly modulated in relation to Advancement, Arpeggio and Grasp, respectively. For MUs the corresponding numbers were 21%, 5% and 21%. Examples of three cells [SUs recorded in striatum (top/middle) and motor cortex (bottom)] showing specific modulations in relation to actuation of the three different motor components are presented in Fig. 5. Thus, the system proved to allow for the correlation of neuronal recordings to kinematic data representing sub-components of compound reaching movement thereby opening up for future studies on how the central nervous system encodes motor commands.

4. Discussion

To perform detailed motion tracking of movements around multiple joints in freely moving animals using traditional semi-manual movement tracking is a daunting task requiring many hours of analyses by well-trained observers (Hermer-Vazquez et al., 2004; Hyland and Jordan, 1997; Whishaw and Pellis, 1990). For some motion tracking applications, fully automatic systems have been developed to facilitate this procedure, typically by the use of reflective markers on concerned body parts (Peikon et al., 2009; Zakotnik et al., 2004). However, in the case of hand pose estimation in humans, improved motion tracking algorithms not requiring body markers have been developed recently (Erol et al., 2007). To track fine movements in subjects not suitable for tracking with markers such as freely behaving rodents, this latter solution is clearly preferable. Consequently, a primary objective of the current study was to develop a marker-less motion tracking system and evaluate the system using reaching and grasping movements performed by rats. A further requirement was that the system must provide kinematic data on forepaw movements in this behavioral task at a level of detail that allow us to study minor differences between reaching trials to correlate this variability to parallel neuronal recordings in motor control circuits.

In the data presented here, it is shown that high-resolution tracking of fine movements involving multiple joints can be performed automatically by the system from recorded video sequences of a rat grasping food pellets, with reliable outcome. Although the accuracy of the pose estimates generated by the system could not be directly evaluated in the real reaching situation the control experiments on a fixated paw indicated that mean errors were typically only a few tenths of a millimeter. This precision was judged to be sufficient for the intended purposes and compares well to systems designed for human hand tracking (Oikonomidis 2010, Nirei et al. 1996). The system also proved to be relatively robust to addition of artificial noise or manipulation of free parameters over a wide range. Thus, under the variations of normal experimental conditions the pose estimates provided by the system appears to be reliable enough for the desired applications. It is also illustrated how this more detailed kinematic description can be used as a basis for new analytical approaches in neurophysiological investigation of motor signals. In fact, several examples of specific firing rate modulations in single cells relating to the actuation of the individual motor components were found. Although a more comprehensive analysis of motor coding in corticostriatal circuits would require a significantly larger data set the current results

nevertheless clearly indicate that the obtained kinematic data opens up for very detailed neurophysiological studies of motor control circuits. These findings are encouraging since the demand for detailed tracking of movement kinematics with high temporal precision is steadily increasing in neurophysiological research to permit interpretation of the rapidly growing neuronal data sets of freely behaving animals (Hoffman and McNaughton, 2002; Kruger et al., 2010; Nicolelis, 2001). Hence, automated methods like the one described in this paper could potentially become an important tool in basic and applied neuroscience. However the described method could likely also improve the quality of behavioral experiments in many other fields and partly reduce the need for the elaborate and time consuming process of manual scoring that, in addition, often requires extensive training to ensure reliable and reproducible identification of movement patterns (Cenci and Lundblad, 2007; Eshkol, 1958; Whishaw and Pellis, 1990; von Laban, 1980).

In the current study previously identified motor components (advancement/retraction, arpeggio and grasping) was used as a starting point for analysis. Nevertheless, the decomposition of any movement into motor components can naturally be done in many different ways once the detailed movements over individual joints have been extracted by the system. This flexibility in the choice of bases functions and spatial coordinate system for the description of movements may prove particularly important in investigations of supraspinal motor control systems given the ongoing controversy regarding the primary coding strategy used by these systems (see for example; Georgopoulos et al., 1986; Loeb et al., 1996; Todorov and Jordan, 2002).

For experiments requiring an even higher level of detail in motion tracking a few improvements to the current system could be made. In our current model design, the distal interphalangeal joint of the digits is assumed to have the same angle as the proximal interphalangeal joint due to the small size of the 3rd phalanx (up to a maximum distal joint flexion angle of 60 degrees). While this is a reasonable approximation when considering the anatomical constraints of the muscles and tendons of the digits (Landsmeer, 1963) and was deemed sufficient for the current purposes, it may be desirable in other applications to have independent tracking of movements around this joint. Similarly, additional parameters such as separate metacarpal bones, a shoulder and an elbow joint could be beneficial to add to the

model. Indeed, in a recently published study using x-ray video techniques to track skeletal bones during reaching movements in rats, complex coordination of the proximal parts of the limb and trunk were found (Alaverdashvili et al 2008). In relation to our data it is worth noting that the x-ray images highly resemble the projections of our 3D paw model onto these viewing planes, although in our study images of the paw were also obtained from an additional viewing plane. Finally, as neural data is sampled in magnitudes of tens of kilohertz it would in some experiments be appropriate to have a comparable rate of acquisition of images for the behavioral tracking. In the current set-up an image acquisition rate of 200 Hz appears to meet basic criteria, but a higher acquisition rate would likely remove some of the tracking noise since consecutive poses would be more alike.

As previously pointed out, the reaching and grasping behavior in rodents bears many similarities to the corresponding behavior in primates and other larger animals indicating that this model lends itself well to translational research, although the very rapid execution of movements in small animals obviously constitutes an extra challenge. This difference between primate and rodent behavior must also be taken into account when analyzing neuronal correlates of different movement components since the neuronal activity related to specific sub-components of the movement as well as the potential use of sensory feedback for corrections of ongoing movements is more limited in the latter group (typical minimum latencies of tactile cortically evoked potentials in rat/man are ~10/20 ms whereas typical reaching times in this task are ~250/500 ms, see; Allison and Hume, 1981; Sacrey et al., 2009). It is also worth pointing out that while rats primarily depend on odor cues to locate the position of the food pellets humans rely on visual guidance in a similar reaching task when the experimental conditions allow for it (de Bruin et al., 2008; Whishaw and Tomie, 1989). However, in spite of the overall slower actuation of reaching and grasping movements in humans, many of the individual motor components involved are undoubtedly executed very rapidly – thus the automated image analysis approach developed here may also open up for new studies of reaching in humans. The technique could potentially provide a new tool for more detailed movement tracking both in the lab (for example in studies of motor learning and control) and in the clinic - diagnosing diseases which affect the motor system at an early time point, giving guidance in selection of optimal treatment strategies (Doan et al., 2010; Jenkins et al.,

2010) or helping to assess the effectiveness of rehabilitation programs after for example trauma or ischemic events affecting motor systems (Massie et al., 2009; McCrea et al., 2005).

Appendix A. Training Protocol

1. Every day for a few days prior to the experiments, the rats were accustomed to the food reward after 20-24 hours food deprivation. After each training session the animal received free supply of their regular food for one hour.
2. The teaching of the reaching behavior began by making the rats aware that there will be food placed on the shelf. This was done by placing multiple pellets on the shelf nearby the aperture for the rats to reach for in any way they can.
3. The acquisition of food was subsequently made harder by moving the pellets further away from the opening slit, forcing the rats to use their forepaws to reach for the pellets.
4. As the rats developed a preference of paw, only one pellet was placed in the socket contralateral to their paw of choice to further promote the use of this paw.
5. In the next training step the rat was taught to move to the back of the cage after each reaching trial, this was achieved by simply not placing any food rewards on the shelf until the rat had reached the opposite end of the cage. This was needed to properly separate different trials.
6. To verify that the displayed behavior was indeed goal directed reaching for food; no food pellet was placed on the shelf at certain occasions. If the rat made a reach attempt anyhow, no pellet was placed on the shelf the consecutive trial either, thereby forcing the rat to identify the presence of food before attempting to reach for it.

Appendix B. Projective geometry

B1. The camera model

A camera matrix describes the transformation of 3D points to 2D points in a pinhole camera: $\boldsymbol{X} \rightarrow \boldsymbol{PX}$, where $\boldsymbol{P}$ is a 3×4 camera matrix and $\boldsymbol{X} = [X \quad Y \quad Z \quad 1]^T$ a 3D point represented in homogeneous coordinates. This transformation is expressed in the *camera equation*,

$$\lambda \mathbf{x} = P\mathbf{X}, \quad \text{(B.1)}$$

where $\boldsymbol{x} = [x \quad y \quad 1]^T$ is the 2D image point $[x \quad y]^T$ represented in homogeneous coordinates, and $\lambda > 0$ is the *depth*, i.e., the distance from the 3D object point $\boldsymbol{X}$ to the image plane (Hartley and Zisserman, 2000). The estimation of the camera matrix **P** is performed using Direct Linear Transformation (DLT) (Abdel-Aziz and Karara, 1971). Using an object whose 3D coordinates are known and manually entering the image coordinates of some points of the object gives the user a set of 3D coordinates and their corresponding image coordinates which is used for DLT.

B2. Quadric surfaces

A quadric surface is a surface in a 3D space defined by a second order implicit equation,

$$\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{Q}\boldsymbol{X} = 0,$$

where $\boldsymbol{X} = [X \quad Y \quad Z \quad 1]^{\boldsymbol{T}}$ is a homogenous coordinate-vector and $\boldsymbol{Q}$ a 4x4 symmetric matrix. Quadrics can be used to describe a number of geometrical shapes, where the ellipsoid and the elliptical paraboloid are the ones used in the present study. In detail, each of the phalanges, the palm of the paw, the forearm, and the food pellet are modeled as ellipsoids, while the snout is modeled as an elliptical paraboloid. Through defining all geometrical shapes of the tracked objects as quadrics, analytical solutions and efficient implementations are possible for the majority of the employed mathematical operations. For instance, the projection of an ellipsoid onto an image plane as well as the silhouette outline of the projection can be computed analytically, and the quadrics $\boldsymbol{Q}_0$ and $\boldsymbol{Q}_1$ occupying the same space can be detected by analyzing the eigenvalues of $\boldsymbol{Q}_0^{-1}\boldsymbol{Q}_1$. This is described further by (Stenger, 2001).

Appendix C. Robustness to choice of free parameter values and measurement noise

C1. Assessment of algorithm robustness to variation in free parameter values

In the first stage of the algorithm the paw is separated from the background based on the light intensity in individual pixels (2.11.1). Proper foreground/background segmentation is therefore sensitive to choice of threshold value. In (Table C.1) the precision (SD of the

distribution of pose estimate errors) and accuracy (mean pose estimate error) of pose estimates compared to the fixated reference pose (Fig. 2D) are presented for a range of threshold values (10-70% of dynamic range) showing a relatively stable performance. The error in each joint angle is normalized to the observed full range of motion for that joint angle and the average value over all joint angles is presented.

Because the size of the image of the paw may vary slightly depending on interindividual differences or changes in the camera set-up, the paw model is manually scaled to the video images for each animal. In (Table C.2) the precision and accuracy of pose estimates compared to the fixated reference pose are similarly presented for a range of scaling values (80-120% of actual value). Note the comparatively stable performance for sizes deviating < 20% from original size.

As described in (2.11.2) the search algorithm was based on an iterative process consecutively searching the five parameter subspaces. Because methods based on stepwise local optimization run the risk of entering local minima we evaluated to what extent different search orders generated different final pose estimates. The video sequence of a reaching trial containing 25 multiframes was searched iteratively using ten different randomly chosen orderings of the five subspaces. Comparing the variation in final pose estimate accuracy for the different search orders showed that choice of search order had a measurable but in practice negligible effect on final pose estimate quality [SD was on average only ±0.13% of full joint movement range (corresponding to ±0.01mm) which should be compared to the mean accuracy of 2.94% (or 0.31mm) for this data set].

C2. Assessment of algorithm robustness to measurement noise

Measurement noise, such as fluctuations in illumination conditions, focusing errors, light reflexes and shadows will ultimately result in errors in paw segmentation (2.11.1). Hence, to allow for a quantitative evaluation of the robustness of the pose estimates to acquisition noise, computer simulations were performed where errors in segmentation were introduced in a controlled manner through addition of artificial noise to paw images. Synthetic silhouette images were first generated from the geometric paw model by projection onto the image planes and segmentation errors were then modeled by distortion of these silhouette images. To simulate segmentation errors a number of foreground pixels were chosen at random and all

foreground pixels at a distance varying from 5 to 10 pixels from these points were labeled as background. Similarly, areas around a number of randomly chosen background pixels were changed to foreground pixels. The effect on pose estimation quality as a function of degree of distortion of the self-generated images is shown in Table C.3. The segmentation error is expressed as the number of erroneously labeled pixels divided by the number of foreground pixels in the original image. The error in each pose parameter is normalized to the observed full range of motion for that parameter and the average value over all parameters is presented.

Acknowledgments

Grants: The Neuronano Research Center is supported by grants from the Swedish Research Council/Linneus grant (project#: 60012701, 80658701) and from the Knut and Alice Wallenberg Foundation (project#: KAW 2004.0119). This project was also sponsored by grants from the Olle Engkvist, Jeanssons, Magnus Bergvall, Kock, and Segerfalk Foundation and a project grant from the Swedish Research Council (project#: K2010-62X-21400-01-3).

The authors are thankful to Lars Clementz for building the experimental set-up and to Dr. Ulrike Richter for thoughtful comments on earlier versions of the manuscript.

References

Y.I. Abdel-Aziz and H.M. Karara. Direct Linear Transformation into Object Space Coordinates in Close-Range Photogrammetry. Proc Symp Close-Range Photogrammetry 1971; 1-18.

Alaverdashvili M, Leblond H, Rossignol S, Whishaw IQ. Cineradiographic (video X-ray) analysis of skilled reaching in a single pellet reaching task provides insight into relative contribution of body, head, oral, and forelimb movement in rats. Behav brain res 2008; 192: 232-47.

Allison T, Hume AL. A comparative analysis of short-latency somatosensory evoked potentials in man, monkey, cat, and rat. Exp Neurol 1981; 72: 592-611.

Cenci MA, Lundblad M. Ratings of L-DOPA-induced dyskinesia in the unilateral 6-OHDA lesion model of Parkinson's disease in rats and mice. Curr Protoc Neurosci / editorial board, Jacqueline N. Crawley ... [et al.] 2007; Chapter 9: Unit 9 25.

de Bruin N, Sacrey LA, Brown LA, Doan J, Whishaw IQ. Visual guidance for hand advance but not hand withdrawal in a reach-to-eat task in adult humans: reaching is a composite movement. J Mot Behav 2008; 40: 337-46.

Doan JB, Whishaw IQ, Pellis SM, Suchowersky O, de Bruin N, Brown LA. Challenging context affects standing reach kinematics among Parkinson's disease patients. Behav Brain Res 2010; 214: 135-41.

Donoghue JP, Wise SP. The motor cortex of the rat: cytoarchitecture and microstimulation mapping. J Comp Neurol 1982; 212: 76-88.

Erol A, Bebis G, Nicolescu M, Boyle RD, Twombly X. Vision-based hand pose estimation: A review. Comput Vis Image Underst 2007; 108: 52-73.

Eshkol NW, A. Movement Notation Weidenfeld and Nicolson: London, 1958.

Gentilucci M, Toni I, Daprati E, Gangitano M. Tactile input of the hand and the control of reaching to grasp movements. Exp Brain Res 1997; 114: 130-7.

Georgopoulos AP, Schwartz AB, Kettner RE. Neuronal population coding of movement direction. Science 1986; 233: 1416-9.

Gholamrezaei G, Whishaw IQ. Individual differences in skilled reaching for food related to increased number of gestures: evidence for goal and habit learning of skilled reaching. Behav Neurosci 2009; 123: 863-74.

Hartley R, Zisserman A. Multiple view geometry. Cambridge university press 2000.

Hermer-Vazquez L, Hermer-Vazquez R, Moxon KA, Kuo KH, Viau V, Zhan Y, Chapin JK. Distinct temporal activity patterns in the rat M1 and red nucleus during skilled versus unskilled limb movement. Behav Brain Res 2004; 150: 93-107.

Hermer-Vazquez L, Moshtagh N. Rats' learning of a new motor skill: insight into the evolution of motor sequence learning. Behav Processes 2009; 81: 50-9.

Ho M-F, Tseng C-Y, Lien C-C, Huang C-L. A multi-view vision-based hand motion capturing system. Pattern Recogn 2011; 44: 443-53.

Hoffman KL, McNaughton BL. Coordinated reactivation of distributed memory traces in primate neocortex. Science 2002; 297: 2070-3.

Hyland BI, Jordan VMB. Muscle activity during forelimb reaching movements in rats. Behav Brain Res 1997; 85: 175-86.

Jeannerod M. The timing of natural prehension movements. J Mot Behav 1984.

Jenkins ME, Johnson AM, Holmes JD, Stephenson FF, Spaulding SJ. Predictive validity of the UPDRS postural stability score and the Functional Reach Test, when compared with ecologically valid reaching tasks. Parkinsonism Relat Disord 2010; 16: 409-11.

Kruger J, Caruana F, Volta RD, Rizzolatti G. Seven years of recording from monkey cortex with a chronically implanted multiple microelectrode. Front Neuroeng 2010; 3: 6.

Landsmeer JM. The Coordination of Finger-Joint Motions. J Bone Joint Surg Am 1963; 45: 1654-62.

Loeb GE, Brown IE, Scott SH. Directional motor control. Trends Neurosci 1996; 19: 137-8.

Massie C, Malcolm MP, Greene D, Thaut M. The effects of constraint-induced therapy on kinematic outcomes and compensatory movement patterns: an exploratory study. Arch Phys Med Rehabil 2009; 90: 571-9.
McCrea PH, Eng JJ, Hodgson AJ. Saturated muscle activation contributes to compensatory reaching strategies after stroke. J Neurophysiol 2005; 94: 2999-3008.
Mitra P, Bokil H. Observed Brain Dynamics. Oxford University Press: New York 2008.
Nicolelis MA. Actions from thoughts. Nature 2001; 409: 403-7.
Nicolelis MAL. Methods for neural ensemble recordings. CRC 2008.
Nirei K, Saito H, Mochimaru M, Ozawa S. Human hand tracking from binocular image sequences. Industrial Electronics, Control, and Instrumentation, 1996, Proc of the 1996 IEEE IECON 22nd Intl Conf 1996: 297-302 vol.1.
Oikonomidis I, Kyriazis N, Argyros A. Markerless and Efficient 26-DOF Hand Pose Recovery. In Kimmel R, Klette R, Sugimoto A, editors. Comp Vis. – ACCV 2010. Springer Berlin / Heidelberg, 2011: 744-57.
Peikon ID, Fitzsimmons NA, Lebedev MA, Nicolelis MA. Three-dimensional, automated, real-time video system for tracking limb motion in brain-machine interface studies. J Neurosci Methods 2009; 180: 224-33.
Peterson G. Mechanisms of handedness in the rat, Comp. Psychol. Monogr 1934; 9: 1–67.
Sacrey LA, Alaverdashvili M, Whishaw IQ. Similar hand shaping in reaching-for-food (skilled reaching) in rats and humans provides evidence of homology in release, collection, and manipulation movements. Behav Brain Res 2009; 204: 153-61.
Stenger B. Model-Based 3D Tracking of an Articulated Hand. In Mendon PRS, Cipolla R, editors, 2001: 310-.
Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. Nat Neurosci 2002; 5: 1226-35.
West MO, Carelli RM, Pomerantz M, Cohen SM, Gardner JP, Chapin JK, Woodward DJ. A region in the dorsolateral striatum of the rat exhibiting single-unit correlations with specific locomotor limb movements. J Neurophysiol 1990; 64: 1233-46.
Whishaw IQ, Haun F, Kolb B. Analysis of behavior in laboratory rodents. Modern techniques in neuroscience. Springer, Berlin, Germany, 1999: 1243-75.
Whishaw IQ, Pellis SM. The structure of skilled forelimb reaching in the rat: a proximally driven movement with a single distal rotatory component. Behav Brain Res 1990; 41: 49-59.
Whishaw IQ, Tomie JA. Olfaction directs skilled forelimb reaching in the rat. Behav Brain Res 1989; 32: 11-21.
Whishaw IQ, Whishaw P, Gorny B. The structure of skilled forelimb reaching in the rat: a movement rating scale. J Vis Exp 2008.
von Laban R, editor. The mastery of movement. Macdonald & Evans: London, 1980.
Vorhees CV, Acuff-Smith KD, Minck DR, Butcher RE. A method for measuring locomotor behavior in rodents: contrast-sensitive computer-controlled video tracking activity assessment in rats. Neurotoxicol Teratol 1992; 14: 43-9.
Zakotnik J, Matheson T, Dürr V. A posture optimization algorithm for model-based motion capture of movement sequences. J Neurosci Methods 2004; 135: 43-54.

Figures with legends

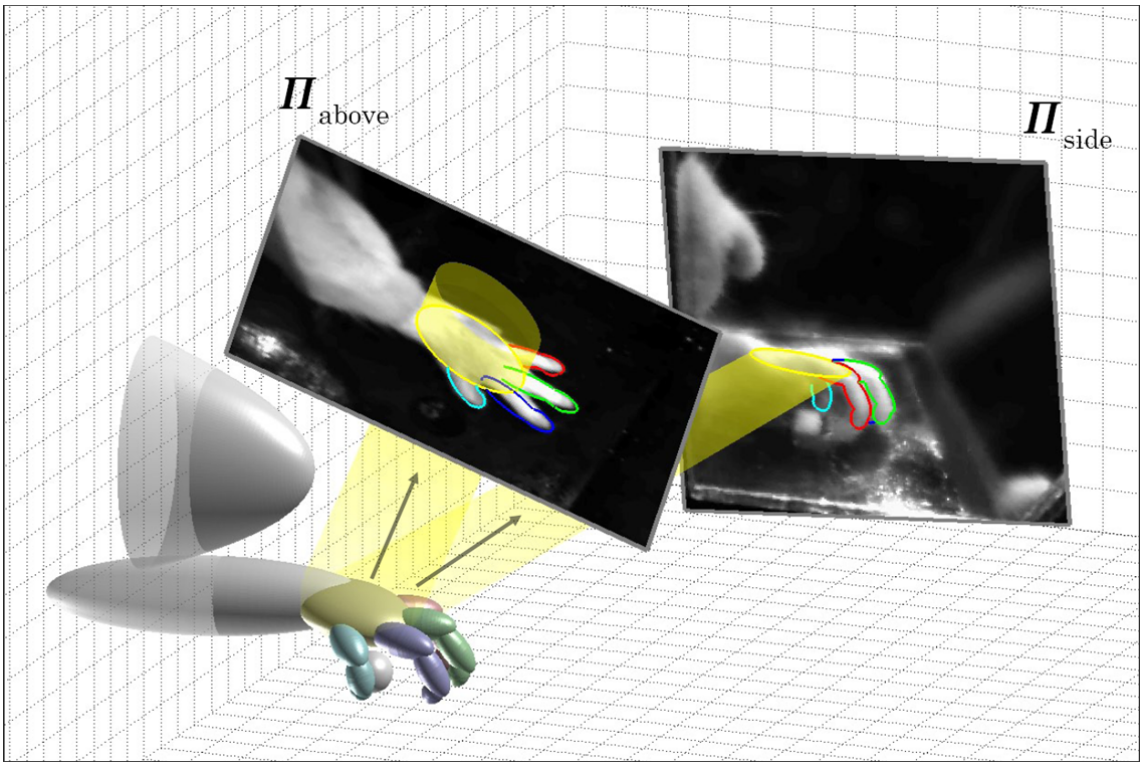

Fig. 1. Computational principle for estimation of paw poses. The pose of a 3D forepaw model is inferred by fitting of the calculated 2D projections onto the different image planes. An estimated 3D paw pose and two of the six image planes (Π_{above} and Π_{side}) are shown. Colored silhouettes outlined on the video images denote the calculated projections.

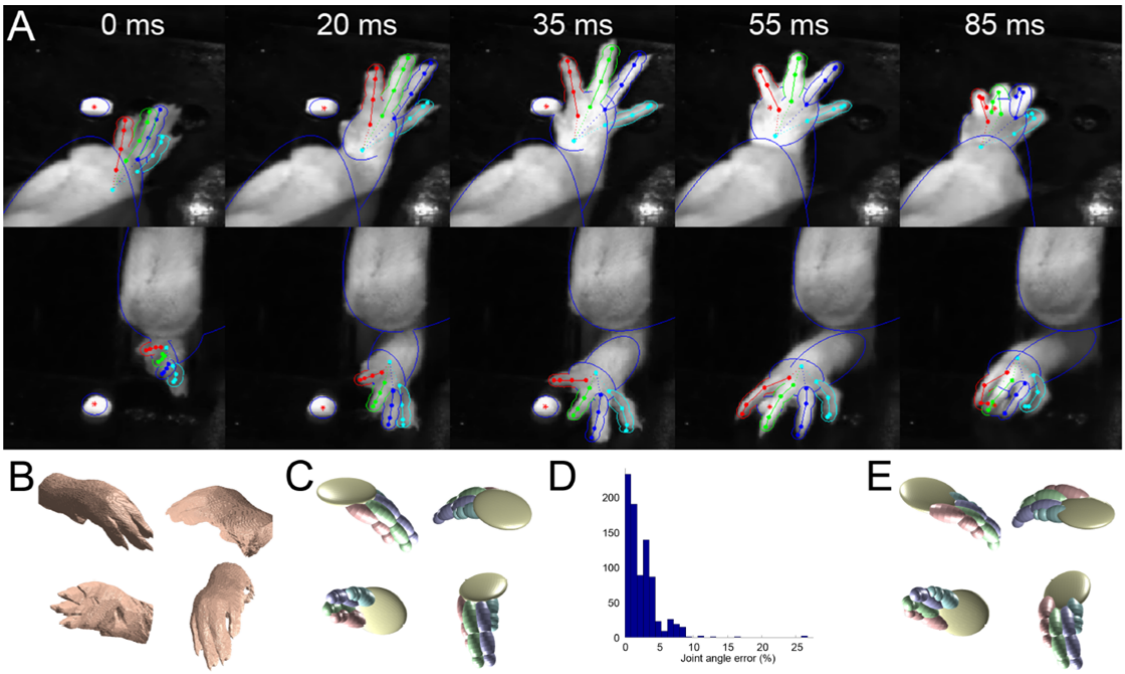

Fig. 2. Tracking of skilled reaching movement sequences. (A) The tracking of the paw in a typical reaching and grasping movement for two of the six viewing planes. The silhouette outline of the estimated paw pose is superimposed on the images for illustration of tracking performance (colored dots/lines indicate joints/bones of the model). (B) Visual hull of the fixated paw shown from four different viewing angles. (C) Reference pose obtained by manual fitting of the 3D model to the visual hull. (D) Histogram of differences in joint angles between estimates provided by the system and the manually fitted reference pose (errors are denoted in percentage of total movement range; n=76 poses). (E) Example of a less accurate pose estimate provided by the system (mean angle error = ~3.3%). Note that for 95% of the estimates the mean error is less than 3.3% and that even for less accurate pose estimates great resemblances to the manually fitted reference pose (C) is evident.

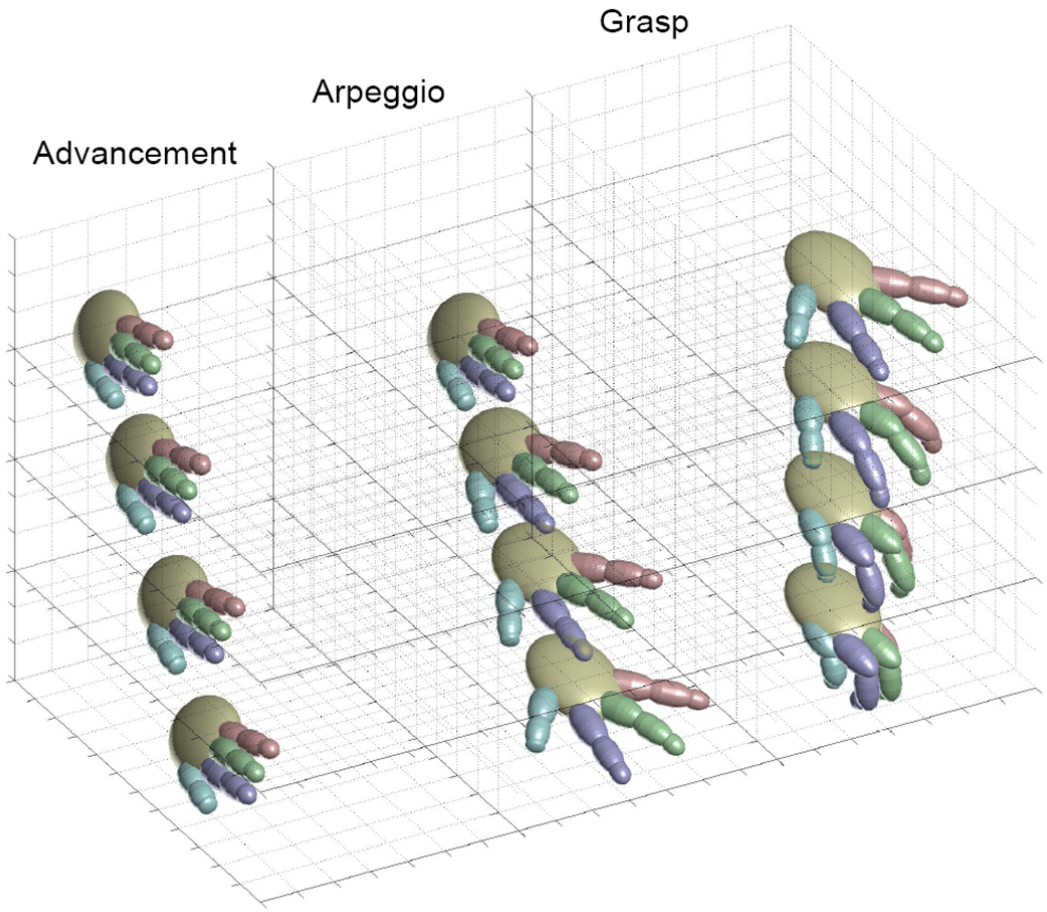

Fig. 3. Decomposition of skilled reaching into motor components. Examples of heuristically defined motor components of compound reaching and grasping movements. The Advancement motor component (left) captures translation along the axis ranging from the aperture of the box to the food pellet. Arpeggio (middle) captures the pronation of the forelimb and adduction of the digits preceding grasping and Grasp (right) is associated with changes in the angles of the joints of the paw and digits.

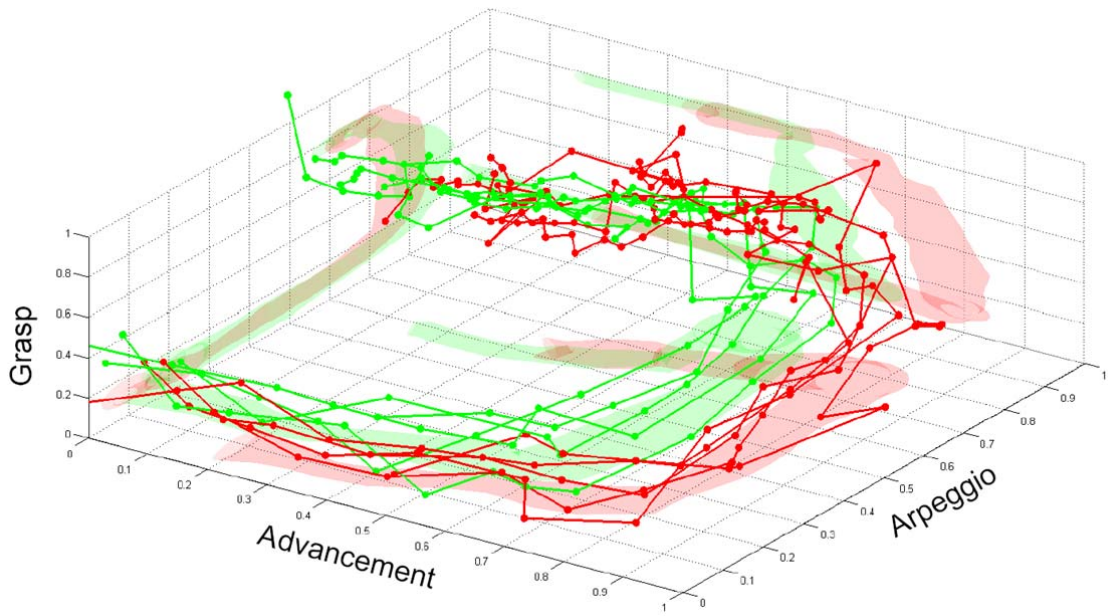

Fig. 4. Movement trajectories in motor component space. Example of reaching trials represented in motor component space. Five individual successful reaching trials (green) are shown together with five unsuccessful trials (red). Dots indicate the time points of paw pose estimations (5 ms intervals) and colored shadows on box walls indicate one SD from the mean trajectory for each group. For ease of comparison, the plot only contains first reach attempts.

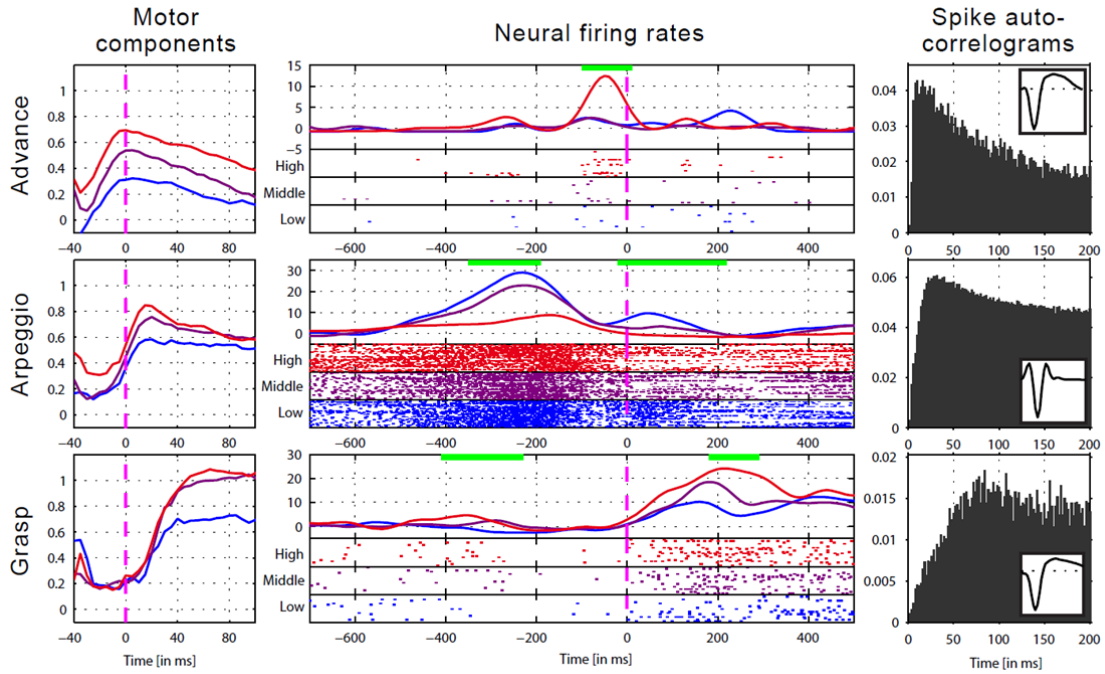

Fig. 5. Neural correlation to motor components. Examples of neurons correlated to individual motor components. Reach attempts are sorted depending on the peak value for a motor component during the attempt, as shown in the left column. The top row is sorted for advance so that the shortest extensions are pooled into the blue group, the intermediate extensions into the purple group and the longest into the red group. In the same manner, the middle row is sorted for arpeggio and the bottom row is sorted for grasp. The middle column show standardized peri-event firing rates and raster plots of individual trials for each group, aligned to time point of maximum paw extension and averaged over all attempts in each group. A green line indicates a significant difference in firing rates between any of the three groups ($p<0.05$). The right column shows the autocorrelogram and waveform for each presented neuron.

Table C.1 Robustness to light intensity thresholding

Threshold value	Accuracy Mean error (%)	Precision SD of errors (%)
0.10	2.33	0.76
0.20	2.40	1.25
0.30	2.05	0.70
0.40	2.01	0.61
0.50	2.27	0.85
0.60	2.49	0.94
0.70	2.44	0.97

Table C.2 Robustness to paw size variation

Scaling factor	Accuracy Mean error (%)	Precision SD of errors (%)
0.80	2.83	0.91
0.85	2.51	0.87
0.90	2.28	0.67
0.95	2.34	0.67
1.00	2.44	1.18
1.05	2.29	0.75
1.10	2.52	1.03
1.15	2.35	1.03
1.20	2.85	1.77

Table C.3 Robustness to image distortion

Noise level	Accuracy Mean error (%)	Precision SD of errors (%)
0-5 %	4.9	0.51
5-10 %	5.3	0.72
10-15 %	5.9	0.81
15-20 %	5.9	1.50
20-25 %	6.2	1.22
25-30 %	6.7	1.87
30-35 %	7.2	0.62
35-40 %	7.6	1.16
40-45 %	10.1	3.25