Combining Wireless Neural Recording and Video Capture for the Analysis of Natural Gait

Justin D. Foster [Student Member, IEEE],
Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA

Oren Freifeld [Student Member, IEEE],
Division of Applied Mathematics, Brown University, Providence, RI 02912 USA

Paul Nuyujukian [Student Member, IEEE],
Bioengineering and Stanford Medical School, Stanford University, Stanford, CA 94305 USA

Stephen I. Ryu,
Department of Neurosurgery, Palo Alto Medical Foundation, Palo Alto, CA 94301 USA

Michael J. Black[*] [Senior Memeber, IEEE], and
Max Planck Institute, 72012 Tübingen, Germany and the Department of Computer Science,
Brown University, Providence, RI 02912 USA

Krishna V. Shenoy[*] [Senior Memeber, IEEE]
Departments of Electrical Engineering and Bioengineering and the Neurosciences Program,
Stanford University, Stanford, CA, 94305 USA

Justin D. Foster: justinf@stanford.edu; Oren Freifeld: freifeld@dam.brown.edu; Paul Nuyujukian: paul@npl.stanford.edu;
Stephen I. Ryu: seoulman@stanford.edu; Michael J. Black: black@cs.brown.edu; Krishna V. Shenoy:
shenoy@stanford.edu

Abstract

Neural control of movement is typically studied in constrained environments where there is a reduced set of possible behaviors. This constraint may unintentionally limit the applicability of findings to the generalized case of unconstrained behavior. We hypothesize that examining the unconstrained state across multiple behavioral contexts will lead to new insights into the neural control of movement and help advance the design of neural prosthetic decode algorithms. However, to pursue electrophysiological studies in such a manner requires a more flexible framework for experimentation. We propose that head-mounted neural recording systems with wireless data transmission, combined with markerless computer-vision based motion tracking, will enable new, less constrained experiments. As a proof-of-concept, we recorded and wirelessly transmitted broadband neural data from 32 electrodes in premotor cortex while acquiring single-camera video of a rhesus macaque walking on a treadmill. We demonstrate the ability to extract behavioral kinematics using an automated computer vision algorithm without use of markers and to predict kinematics from the neural data. Together these advances suggest that a new class of "freely moving monkey" experiments should be possible and should help broaden our understanding of the neural control of movement.

[*]These authors contributed equally.

I. Introduction

An important long-term goal of neuroscientific studies is to understand the neural control of unconstrained behavior. The analysis of such movement is in contrast to previous work in which movement is narrowly constrained, the context is controlled, and the behavior is repetitive and highly trained. The hypothesis underlying this work is that the variability and complexity of "natural" movements may reveal new insights into the neural control of movement. The present study represents a first step towards this goal by using a head-mounted neural recording system with wireless data transmission and computer vision to analyze the walking motion of a non-human primate. It expands on previously reported studies that have explored the relationship between cortical neural activity and gait by allowing for untethered multichannel recording [1] and by allowing for unconstrained quadrupedal movement with no head or neck fixation [2]. In this way the present study helps move further toward the goal of establishing an animal model of freely moving humans, which is important for the advancement of basic systems neuroscience as well as neural prostheses [3], [4].

The analysis of natural movement in primates requires: 1) a mechanism to record neural population activity from a freely moving animal, 2) a method for recording behavior that does not encumber nor alter the movement, 3) a representation of the movement suitable for analysis, 4) a task that exhibits significant variability yet is sufficiently repeatable to facilitate analysis, and 5) a method for analyzing high dimensional neural and behavioral signals that gives insight into the neural coding of movement.

The system described here, and outlined in Fig. 1, uses a device for recording the simultaneous activity on multiple channels of a microelectrode array implanted chronically in the dorsal aspect of premotor cortex (PMd) in one rhesus monkey. The head-mounted device wirelessly transmits the neural activity to a host computer, freeing the animal to move without the constraints of standard recording methods, which use wire cables. We record the full-body motion of the animal using a video camera. Unlike previous work that has relied on placing reflective markers on the body [5], we describe a markerless system that simplifies the experimental setup. Our task requires the monkey to walk quadrupedally on a treadmill (see Fig. 1a). The speed of the motion can be varied and the monkey is allowed to move backwards and forwards, changing his location on the treadmill. The monkey is also free to rotate his head and vary his gait as desired. We then extract behavioral parameters using a computer vision method that segments the monkey from the background and models the behavior using a time-varying image-based signature. The behavioral measure is obtained by projecting an image patch onto a low dimensional subspace obtained using principal component analysis (PCA). Finally, we define a linear decoding model and show how it relates to behavior.

II. Wireless Neural Recording

A. Behavioral Task and Neural Recordings

All protocols were approved by the Stanford University Institutional Animal Care and Use Committee. We trained an adult male rhesus macaque to walk on a treadmill at speeds

ranging from 2.0 kph to 3.5 kph. Each session lasted approximately 20 minutes and was divided into blocks where the monkey walked continuously for up to 5 minutes before a break. For this study, we selected and analyzed one day's session where the monkey walked at a constant speed of 3 kph for 5 blocks, each about 2 minutes in length. Synchronization between the neural and video data streams was typically accurate to within +/− 500 ms; each data stream had a precise, independent clock of its own. The neural data was aligned to maximize the linear fit to the behavioral data, thereby correcting for any clock synchronization offset, as well as identifying the optimal lag between kinematics and neural activity. Broadband neural activity on 32 electrodes was sampled at 30 Ksamples/s and transmitted wirelessly using a HermesD system [6]. An OrangeTree ZestET1 was programmed to package the HermesD output datastream into a UDP Ethernet packet stream, which was saved to disk.

B. Neural Data Analysis

Each channel of neural recordings was filtered with a zero-phase highpass filter to remove the local field potential (LFP), since LFP is not the focus of the present study. We then determined the times of spikes on each channel by a threshold crossing method. Points where the signal dropped below $-3.0\times$ the RMS value of the channel were spike candidates. Occasional artifacts (likely due to static discharge) were automatically rejected from the candidate spike set based on the shape and magnitude of the signal near the threshold crossing point. Finally, firing rates of each channel were obtained by convolving the spike trains with a Gaussian filter with a standard deviation of 30 ms and binned every 42 ms, determined by the frame rate of the video.

III. Movement Analysis

A. Hand-Tagged Gait

Video image sequences were manually tagged to estimate the ground truth walking movement, at one moment per gait cycle. Tags marked frames where the contralateral arm (left PMd was implanted and right arm analyzed) was maximally retracted. The tags were then fit with a sinusoidal waveform with peaks corresponding to the hand tagged time. The phase of the arm between consecutive hand tags was assumed to precess at a constant rate (i.e., linear phase).

B. Image Analysis

Video was captured at 24 fps at a resolution of 1624×1224 pixels using a Point Grey Grasshopper GRAS-20S4M/C camera. Data acquisition was performed using a 4DViews 2DX Multi-Camera system.

The goal of the image analysis was to extract a behavioral descriptor that was correlated with the recorded neural data. In each frame, we first segmented the monkey from the background by employing a pairwise Markov random field approach [7]. The optimal segmentation was computed by graph cuts [8]–[10] using GCMex [11]. The segmentation in turn defined a contour that marked the outline of the animal as shown in Fig. 2a. Forearms and hands were not included in the contour because they did not contrast well with the black

tread of the treadmill. Next, among the points along the front of the animal, we automatically determined a contour point whose y-value was predefined to match the upper-arm region. This was defined to be the point of reference in the current frame. We further defined a small image patch adjacent to the reference point to be our region of interest (ROI), illustrated by the box in Fig. 2a, 2c, and 2d. Effectively, this procedure yielded a tracker whose output was a small rectangular frame that is "locked" to the upper arms. The resulting image patch was invariant to translation; it captured the essence of the arm movements without including nominally irrelevant information such as the absolute position of the animal on the treadmill. The size of the ROI color image was 25 by 31 pixels, a 2325-dimensional vector. Finally, to obtain a behavioral descriptor of the motion, we performed dimensionality reduction on the ROI image using PCA. We focused our attention on the first principal component (PC) plotted in Fig. 2b. The first PC captured much of the behavioral activity: it was roughly periodic with peaks corresponding to times of maximal contralateral retraction (Fig. 2c) and troughs corresponding to times of maximal contralateral protraction (Fig. 2d) as depicted by the arrows Fig. 2b.

IV. Data and Results

A. Decoding Hand-Tagged Kinematics

The hand-tagged data provided an approximation of the behavioral data. We linearly regressed the hand-tagged signal against neural firing rates. We made no assumption of the timing between the two signals. Instead, we fit the sinusoidal waveform at many different lag values and compared RMS values at each point. These fits of the hand-tag waveform were done with a 33-parameter linear fit (effectively, a linear filter as in [12] with only one time bin), one parameter per channel plus a constant offset. The results are shown in Fig. 3a, demonstrating an optimal lag for minimizing RMS error.

To gain an understanding of how well the neural data predicts the hand-tagged measure of gait, we performed a 10-fold cross validation of the neural fit for two blocks, employing this optimal lag value. A sample decode of the is shown in Fig. 3b, and the resulting R^2 value was 0.83 for this fit.

B. Decoding ROI Image PC coefficients

With the hand-tagged decode results as a baseline, we next explored decoding the first PC coefficients. As in the processing of the hand-tagged data, no assumptions were made regarding the alignment of the video and neural data. Minimizing the RMS error once again yielded the optimal lag (Fig. 4a). Similarly, neural decodes of the first PC coefficients were made using 10-fold cross validation for 5 blocks using the optimal lag. The R^2 value for this analysis was 0.72. A sample decode is shown in Fig. 4b.

V. Discussion and Conclusions

A. Neural Activity Predicts Walking Kinematics

We have shown that neural data recorded from the arm region of PMd can predict contralateral arm kinematics associated with walking on a treadmill. By sweeping the alignment between the neural and video data, an optimal latency between them was found.

A dip in the RMS error is observable in both hand-tagged and PC decoding around the value of optimal latency. These decreases in RMS arise from variance in the stride length consistent in both the behavioral data and the neural firing rate data. The characteristic dip in the dataset confirms that we are aligning our data at the correct location and that the neural data correlates well with stride timings. Furthermore, we were able to employ a simple linear fit and use 32 channel neural data to predict the phase of the arm movement reasonably well ($R^2 = 0.83$). Future experiments would be needed to compare quantitatively decode performance on this task with decode performance in traditional, more constrained settings (e.g., [12]).

B. Markerless Extraction of Relevant Behavior

This study demonstrated two methods for extracting behavior that were relevant to the experimental task without using markers to obtain physiological measurements. The first method was hand-tagging the video and the second was an automated method for processing the frames. Importantly, this study showed a proof-of-concept for automatically extracting key features of movement that are encoded in the firing rates of neurons. Thus, we were able to avoid the task of hand labeling.

C. Future Work

In the present study, hand-tagged images provided a relatively good ground truth for interpolating the phase of the gait. However, hand-tagging is not feasible for more complex studies of natural behavior for a number of reasons: 1) hand-tagging cannot provide a complete representation of posture, 2) it is often somewhat qualitative and subject to user error, and 3) it does not scale to large datasets.

It is promising that a relatively simple model for extracting behavioral measurements performed comparably to hand-tagging (R^2 of 0.73 vs. R^2 of 0.83). Future work should incorporate additional camera angles of the same space allowing for better models of behavior and, as a result, more precise estimation of kinematics. Detailed models that enable precise measurements of behavior will likely outperform hand-tagging methods and enable a new class of experiments focused on unconstrained movement.

Acknowledgments

References

1. Nakajima, K.; Mori, F.; Murata, A.; Inase, M. Single-Unit Activity in Primary Motor Cortex of an Unrestrained Japanese Monkey During Quadrupedal v.s. Bipedal Locomotion on the Treadmill. 2010 Neuroscience Meeting Planner; San Diego, CA. 2010; Society for Neuroscience;

2. Winans, JA.; Tate, AJ.; Lebedev, MA.; Nicolelis, MAL. Extraction of Leg Kinematics from the Sensorimotor Cortex Representation of the Whole Body. 2010 Neuroscience Meeting Planner; San Diego, CA. 2010; Society for Neuroscience;

3. Gilja V, Chestek CA, Nuyujukian P, Foster JD, Shenoy KV. Autonomous Head-Mounted Electrophysiology Systems for Freely-Behaving Primates. Current Opinions in Neurobiology. Jun. 2010 20:676–686.

4. Gilja V, Chestek CA, Diester I, Henderson JM, Deisseroth K, Shenoy KV. Challenges and Opportunities for Next-Generation Intra-Cortically Based Neural Prostheses. IEEE Transactions on Biomedical Engineering. In press.

5. Vargas-Irwin CE, Shakhnarovich G, Yadollahpour P, Mislow JMK, Black MJ, Donoghue JP. Decoding Complete Reach and Grasp Actions from Local Primary Motor Cortex Populations. Journal of Neuroscience. 2010; 30(29):9659. [PubMed: 20660249]

6. Miranda H, Gilja V, Chestek CA, Shenoy KV, Meng TH. HermesD: A High-Rate Long-Range Wireless Transmission System for Simultaneous Multichannel Neural Recording Applications. IEEE Transactions on Biomedical Circuits and Systems. Jun; 2010 40(3):181–191. [PubMed: 23853342]

7. Geman S, Geman D, Relaxation S. Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1984; 6(2):721–741. [PubMed: 22499653]

8. Boykov Y, Veksler O, Zabih R. Efficient approximate energy minimization via graph cuts. IEEE transactions on Pattern Analysis and Machine Intelligence. Nov; 2001 20(12):1222–1239.

9. Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE transactions on Pattern Analysis and Machine Intelligence. Sep; 2004 26(9):1124–1137. [PubMed: 15742889]

10. Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? IEEE transactions on Pattern Analysis and Machine Intelligence. Feb; 2004 26(2):147–159. [PubMed: 15376891]

11. Fulkerson, B.; Vedaldi, A.; Soatto, S. Class segmentation and object localization with superpixel neighborhoods. Proceedings of the International Conference on Computer Vision; October 2009;

12. Chestek CA, Batista AP, Santhanam G, Yu BM, Afshar A, Cunningham JP, Gilja V, Ryu SI, Churchland MM, Shenoy KV. Single-Neuron Stability During Repeated Reaching in Macaque Premotor Cortex. The Journal of Neuroscience. 2007; 27(40):10742–10750. [PubMed: 17913908]
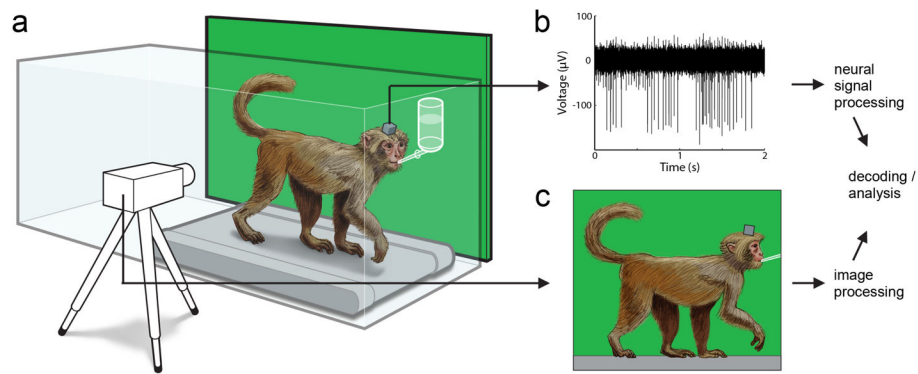
Fig. 1. System overview
(a) The motion of a monkey is passively observed with a video camera while neural activity is recorded and transmitted wirelessly. (b) An example electrode's filtered broadband signal obtained wirelessly from the HermesD recording system and filtered to remove the local field potential. (c) Video images are captured from a side view in order to extract information about posture.
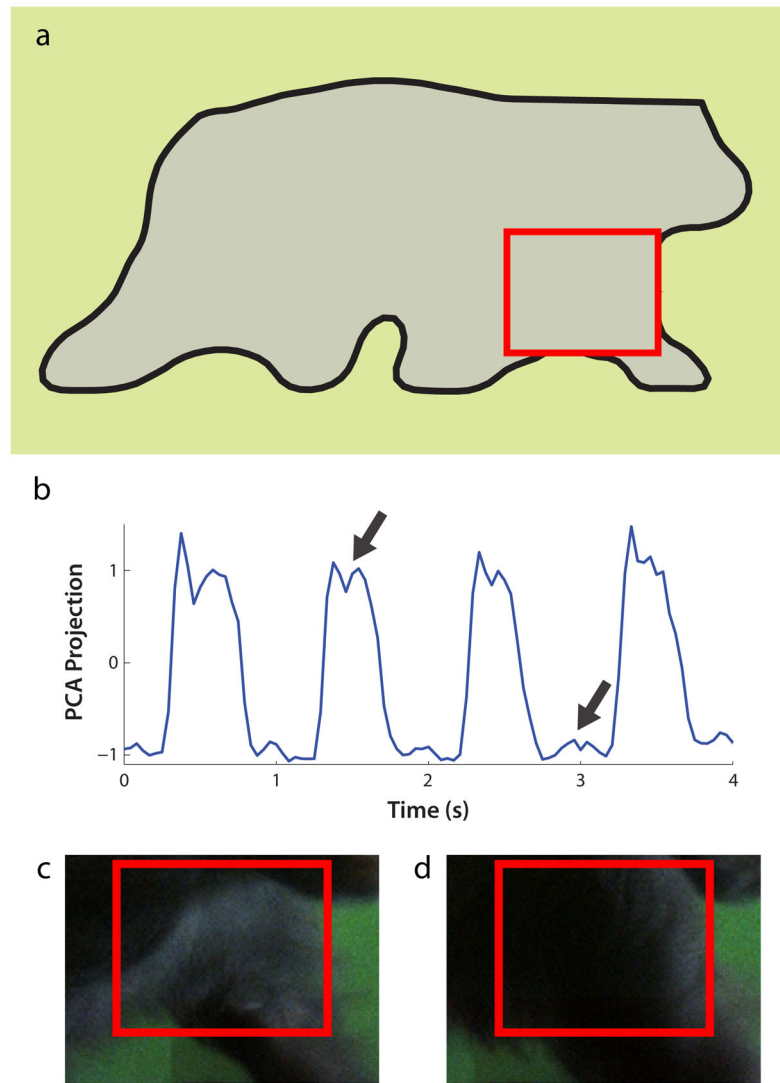
Fig. 2. Image processing
(a) Image segmentation of a video frame. Pixels interior to the outline were defined to be part of the animal and pixels exterior to the outline are deemed part of the background. The box represents the ROI. (b) Behavioral signal: coefficient of the first PC over time. Peak arrow indicates time of maximal retraction (typical ROI shown in panel c) of the contralateral arm. Trough arrow indicates time of maximal protraction (typical ROI shown in panel d).
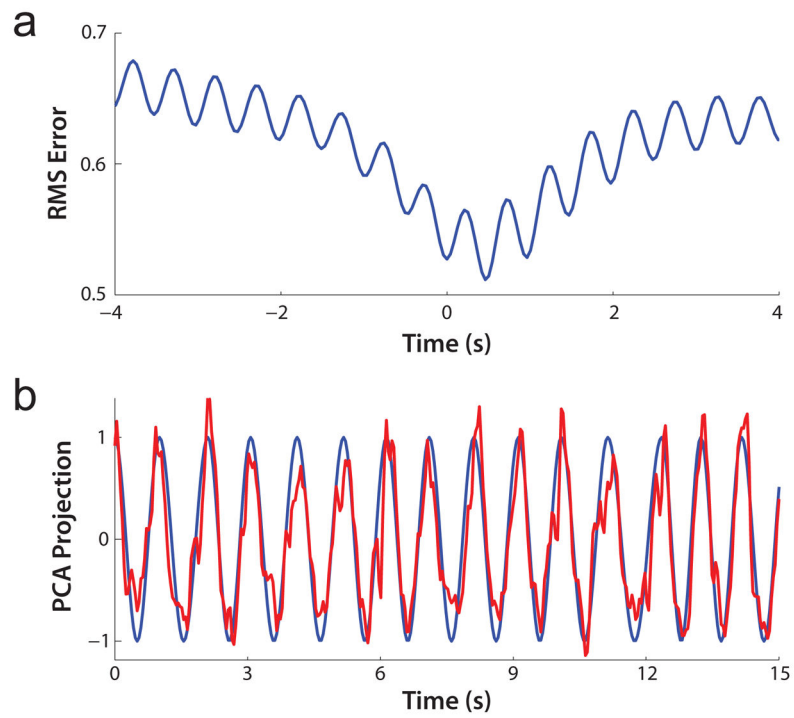
Fig. 3. Decode of hand-tagged signal
(a) Plot of RMS error versus lag for block 4, zero represents initial synchronization. (b) Sinusoidal signal plotted in blue with decoded signal in red.
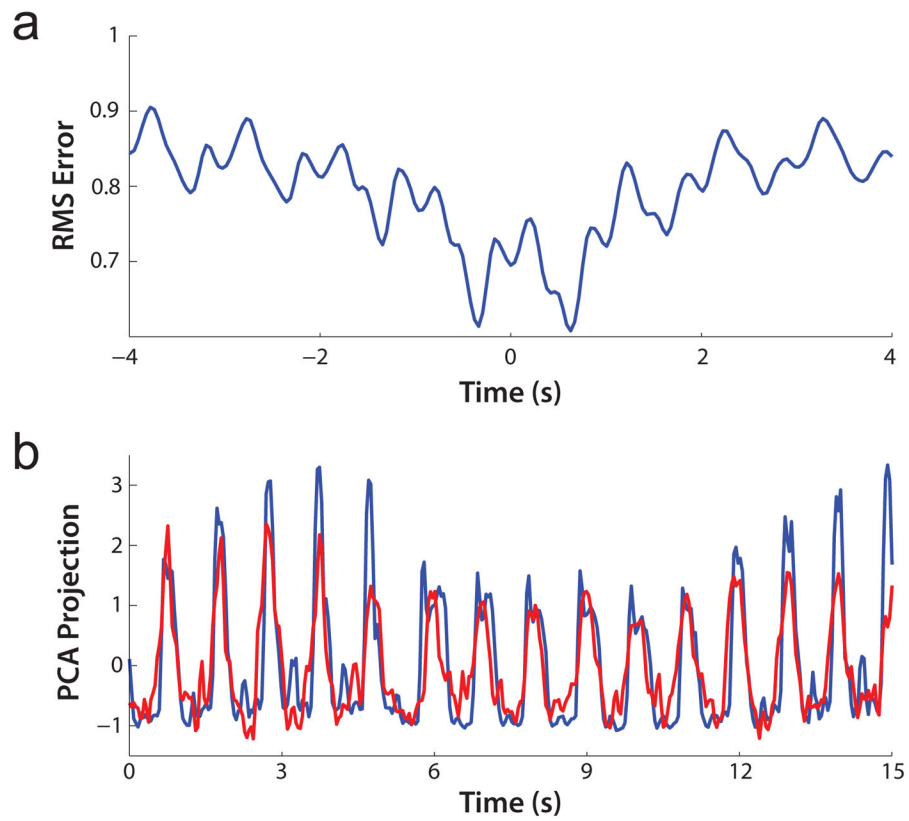
a

b

Fig. 4. Decode of first PC coefficients
(a) Plot of RMS error of decoder versus lag for block 4. (b) First PC coefficients plotted in blue with decode in red.