ANIMAL LOCOMOTION

A markerless platform for ambulatory systems neuroscience

Michael P. Silvernagel[1†], Alissa S. Ling[1†], Paul Nuyujukian[2,3,1,4,5]*, for the Brain Interfacing Laboratory[‡]

Motor systems neuroscience seeks to understand how the brain controls movement. To minimize confounding variables, large-animal studies typically constrain body movement from areas not under observation, ensuring consistent, repeatable behaviors. Such studies have fueled decades of research, but they may be artificially limiting the richness of neural data observed, preventing generalization to more natural movements and settings. Neuroscience studies of unconstrained movement would capture a greater range of behavior and a more complete view of neuronal activity, but instrumenting an experimental rig suitable for large animals presents substantial engineering challenges. Here, we present a markerless, full-body motion tracking and synchronized wireless neural electrophysiology platform for large, ambulatory animals. Composed of four depth (RGB-D) cameras that provide a 360° view of a 4.5-square-meters enclosed area, this system is designed to record a diverse range of neuroethologically relevant behaviors. This platform also allows for the simultaneous acquisition of hundreds of wireless neural recording channels in multiple brain regions. As behavioral and neuronal data are generated at rates below 200 megabytes per second, a single desktop can facilitate hours of continuous recording. This setup is designed for systems neuroscience and neuroengineering research, where synchronized kinematic behavior and neural data are the foundation for investigation. By enabling the study of previously unexplored movement tasks, this system can generate insights into the functioning of the mammalian motor system and provide a platform to develop brain-machine interfaces for unconstrained applications.

INTRODUCTION

Motor systems neuroscience explores how complex, interconnected neuronal networks in the brain elicit movement (*1*). Awake, behaving animal studies are a cornerstone of this field and have been crucial to understanding how the brain controls motor activity (*2*). Many types of animals have been used to study the neural control of movement, from simple neural circuits in worms and fruit flies to complex mammalian circuits in mice and rhesus macaques (*3*–*6*). Of these, macaques are a particularly appealing choice, both for their ability to execute dexterous movements and for their neuroanatomical similarities to humans (*7*, *8*).

Traditional macaque studies often constrain most bodily movements except the movement of interest to reduce confounding variables, allowing for tighter correlations between the measured behavior and brain activity (*9*, *10*). A common investigation is a reaching task where the macaque sits in a chair with its head, legs, and nonreaching arm at rest and gently constrained (*11*, *12*). Targets are presented on a screen while movement kinematics are measured by marker tracking (*13*), a manipulandum/joystick (*14*), or haptic devices (*15*), and neural data of brain-implanted electrodes are collected from a wired neural recording device (*9*, *16*, *17*). By correlating neural modulation with movement, conclusions can be drawn from the data generated by these classes of studies (*9*, *18*).

Growing evidence from freely moving studies, however, suggests that neural activity from classically constricted tasks may not generalize to other ambulatory behaviors. Research with freely moving worms found neuron activity that had not been previously observed in behaviorally constrained worms (*19*). Freely moving research in mice have arrived at similar conclusions (*20*, *21*). Given that neuronal activity differs for free and constrained behavior in other animals, questions arise as to whether this holds for higher-order mammals, such as macaques (*22*, *23*). Answering this requires macaque studies that incorporate ambulatory movement and unconstrained tasks.

Outstanding need

Although previous systems that pair kinematics with wireless neural recordings have laid the groundwork for understanding neural control of free behavior, existing platforms do not address three key components. First, to support unrestricted motor activity, observational enclosures must be sufficiently large. Second, given the potential for intermixed limb representations in motor cortex (*24*), full-body kinematics synchronized with wireless electrophysiology must be captured to gain a complete picture of freely moving neural activity. Previous markered and markerless systems for macaques, however, have limited tracking to one or two limbs. Last, for these setups to be adopted by the broader neuroscience community, they should be robust to noise and camera movement, be easily implemented, and support recordings on the time scale of hours.

Existing neuroscience platforms

A key advancement in realizing freely moving macaque studies has been the development of battery-powered, head-mounted systems. Wireless systems are capable of recording neural activity from hundreds of channels of chronic multielectrode implants, allowing

[1]Department of Electrical Engineering, Stanford University, Stanford, CA, USA. [2]Department of Bioengineering, Stanford University, Stanford, CA, USA. [3]Department of Neurosurgery, Stanford University, Stanford, CA, USA. [4]Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA. [5]Stanford Bio-X, Stanford University, Stanford, CA, USA.
*Corresponding author. Email: 21sciencemag@pn.stanford.edu
†These authors contributed equally to this work.
‡Individuals in the Brain Interfacing Laboratory are listed in the Acknowledgments.

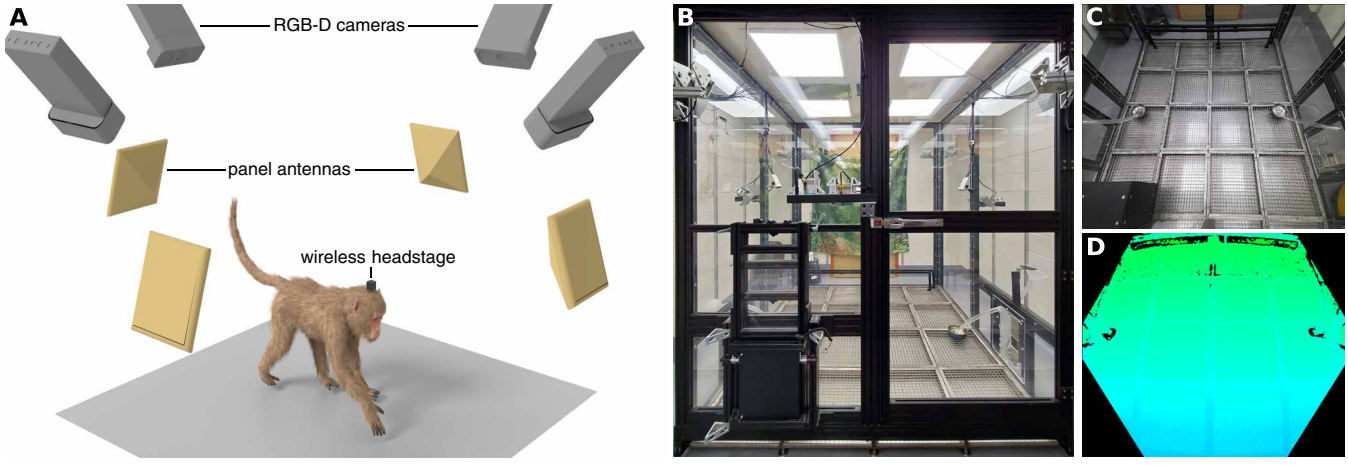


Fig. 1. Observational enclosure setup and raw camera data. (**A**) Rendered image of macaque in observational enclosure with RGB-D cameras (dark gray rectangles) and antennas (tan cubes) capturing behavioral depth data and wireless neural data. Four RGB-D cameras are placed inside the enclosure 90° apart to get a 360° view. Sixteen antennas are placed outside the enclosure, eight at a height of 1 m on the side walls and eight at a height of 2 m at the corners. A 96-channel battery-powered wireless neural transmitter (black cube on macaque head) is wirelessly transmitting neural data to the receiving antennas. (**B**) Photo of observational enclosure, which measures 2.1 m by 2.1 m by 2.4 m (length by weight by height) in size. A camera is placed in the center of each enclosure wall about 1.5 m from the floor. (**C**) Color image from one camera view. The field of view (FOV) for the RGB camera was 90° by 74°. (**D**) Heatmap from a depth image. The FOV for the depth camera was 75° by 65°.

for platforms where macaques are unconstrained and have a greater range of motion (*15*, *25*, *26*). Studies leveraged custom-built wireless systems (*15*, *26*–*33*) or used commercial offerings (*34*).

Initial unconstrained macaque studies allowed for free expression of a particular task, such as walking, by having the animal ambulate on a treadmill (*35*–*38*) or through a narrow corridor (*39*). Kinematics were captured from video data, and these easily repeatable behaviors required limited training and facilitated high trial counts. With these more naturalistic behaviors, studies developed brain-machine interfaces for a wider range of tasks, such as navigating a mobile wheelchair (*40*) or alleviating spinal cord injury (*39*).

Findings from these task-based freely moving platforms suggest that different behaviors give rise to different neural activity, such as engaging multiple limbs versus one limb (*38*, *41*). Furthermore, the variability of recorded neuronal activity may also increase with task difficulty (*22*, *23*). Despite this, questions still remain as to the extent neural activity differs between simple constrained tasks, task-based free movement, and fully unconstrained movement, and how that affects the performance of naturalistic brain machine interfaces. Answering these questions requires the ability to track three-dimensional (3D), full-body macaque kinematics in an open space.

Although physical markers provide accurate results, they are poorly suited for ambulatory macaque studies. Macaque fur is long, dense, and fast growing, allowing markers to easily become obscured and making reliable placement challenging. Moreover, because animals are unconstrained, there is a nontrivial risk that they will remove, damage, or eat the markers, rendering them useless. Marker-based outfits, such as jackets and bodysuits, are uncomfortable, require animal training and accommodation, and constrict natural movement. Any approach where the animal's skin is tattooed or painted requires regular shaving, which presents potential husbandry challenges (*42*). These procedures can negatively affect animal well-being and are advised against from an animal welfare perspective (*43*–*45*).

Limited platforms currently exist to capture unconstrained macaque behavior using markerless motion capture (*46*–*48*). Bala *et al.* tracked freely moving macaques in a large, unconstrained environment using OpenMonkeyStudio. Here, transfer learning from human pose data was used to track 2D RGB (red green blue) joint positions, which were then projected into 3D space (*46*). OpenMonkeyStudio required 62 RGB cameras to span its recording space. MacaquePose, a single-view, RGB, "in the wild" behavior dataset, used DeepLabCut to extract kinematics (*47*). DeepLabCut, which is built on top of the DeeperCut multi-person pose estimation model (*49*), facilitates accessible pose tracking of animals in controlled settings (*50*). Nakamura *et al.* created a markerless motion capture system for animals that operated on 3D surface images generated by four Microsoft Kinect depth cameras (*48*). Kinematics were calculated by fitting 3D hulls from point cloud data to an articulated skeleton using PyBullet (*51*). Although these approaches allowed for markerless tracking, these platforms only collected behavioral data in their studies.

Of platforms that capture macaque behavior markerlessly, few exist with synchronized neural data (*34*, *35*). One study recorded macaque locomotion and neural data on a treadmill with eight synchronized video cameras and used computer vision algorithms to detect the 2D locations of the right upper arm joints (*35*). That team used a custom-built wireless system to transmit wireless neural data (*52*). Another study created a modified home cage environment that enabled visual and haptic feedback. DeepLabCut was used to extract behavioral data of reaching and walking from four RGB cameras, and neural data were recorded using commercial hardware and floating microwire arrays (*34*).

Although both platforms recorded full-body movement, their recordings were limited to small enclosures, which only allowed for a subset of movement types. The first study constrained behavior to the plane of the treadmill (*35*), and the second included enough space to take a few steps to the task and return to a seated position (*34*). In addition, these studies only captured kinematic data for the upper extremities.

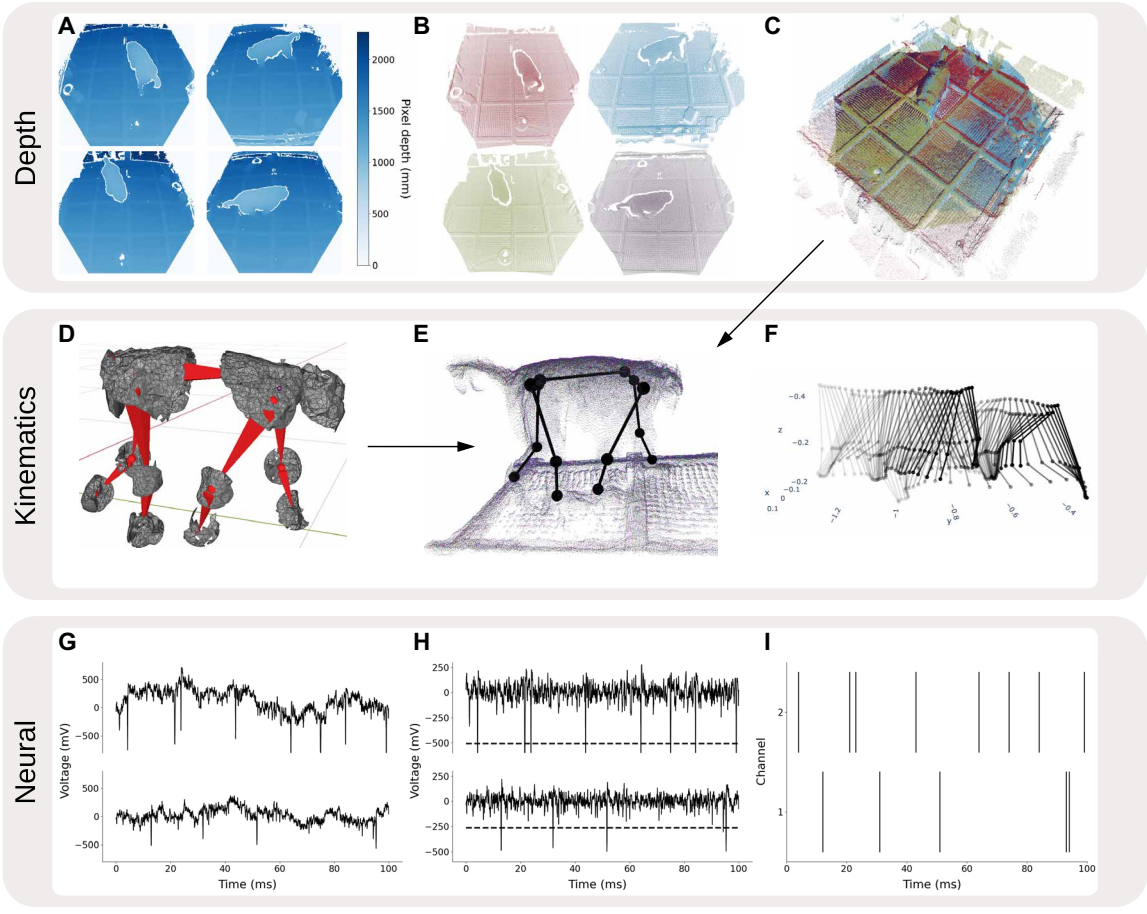


Fig. 2. Data collection and analysis pipeline of point cloud, kinematic, and neural data. (**A**) Raw depth data from four RGB-D cameras. (**B**) 3D point clouds generated from (A) using each camera's intrinsic and extrinsic properties. These provide views of the macaque's left side (blue), head (green), backside (red), and right side (purple). (**C**) The single-view point clouds from each camera were transformed into a unified coordinate system to create a complete view of the macaque. (**D**) Articulated skeleton for a rhesus macaque created in Blender. A mesh of the point cloud is created in Meshlab, and segments are modeled using Boolean intersections between the mesh and primitive spheres. Segments rotate with the bones, and the skeleton is structured in a bottom-up fashion, where the tail was designated as the center point with which all other segments move. (**E**) Monkey C's full-body kinematic fit overlaid on top of the point cloud. Arrows represent the inputs (point cloud and skeleton) to the optimization problem to output the kinematic fit. (**F**) Full-body 3D pose trajectories for Monkey C as he performs a walking task for 30 frames (1 s). (**G**) Raw neural data from two arbitrary channels sampled at 30 kHz. (**H**) High-pass filtered data from (G). The dashed lines represent each channel's baseline RMS voltage. (**I**) Spike rasters created from threshold crossings in (H). Vertical lines denote spiking events in each channel.

Table 1. Mean algorithm error in centimeters for the right elbow, wrist, knee, and ankle compared with existing platforms performing 3D estimation for fully unconstrained behavior. Error was calculated by the Euclidean distance between the algorithm joint locations and hand-labeled joint locations. Paper error and spread represent the mean error and SD for 300 frames of walking. Error for (*48*) was the mean error of the reported estimation errors. Error for (*47*) was aggregate error, so the SD could not be calculated. Error for the whole body was reported in one number for (*46*). Dashed lines indicate no error reported for the segment.

Segment	This work	Nakamura *et al.* (*48*)	Labuguen *et al.* (*47*)	Bala *et al.* (*46*)
Elbow (cm)	3.9 ± 2.2	6.4 ± 3.6	5.8	–
Wrist (cm)	5.0 ± 2.9	8.3 ± 6.9	6.9	–
Knee (cm)	3.7 ± 1.7	7.8 ± 5.1	6.7	–
Ankle (cm)	4.1 ± 1.9	11.2 ± 7.9	6.5	–
Body (cm)	–	–	–	6.8 ± 2.3

Platform overview

To address the outstanding needs, an end-to-end data acquisition framework was developed to extract synchronized wireless neural data and markerless, full-body kinematics for ambulatory macaques in a large environment (Fig. 1). The system used four commercial RGB-D (RGB-depth) cameras to capture point cloud data of the animal and a commercial wireless neural transmitter to provide ≈100 channels of neural data. An algorithm was developed using an articulated skeleton model and Bayesian optimization to perform pose tracking on the unified point cloud data. The accuracy of the algorithm, calculated by comparing the algorithm joint results with hand-labeled joint locations, exceeds that of comparable platforms. As a proof of concept, this approach is demonstrated with neural and kinematic data from one monkey. This system captures neural data and performs full-body pose tracking of unconstrained macaques (Fig. 2), serving as a tool to conduct freely moving neuroscience studies and develop brain-machine interfaces for naturalistic movement.

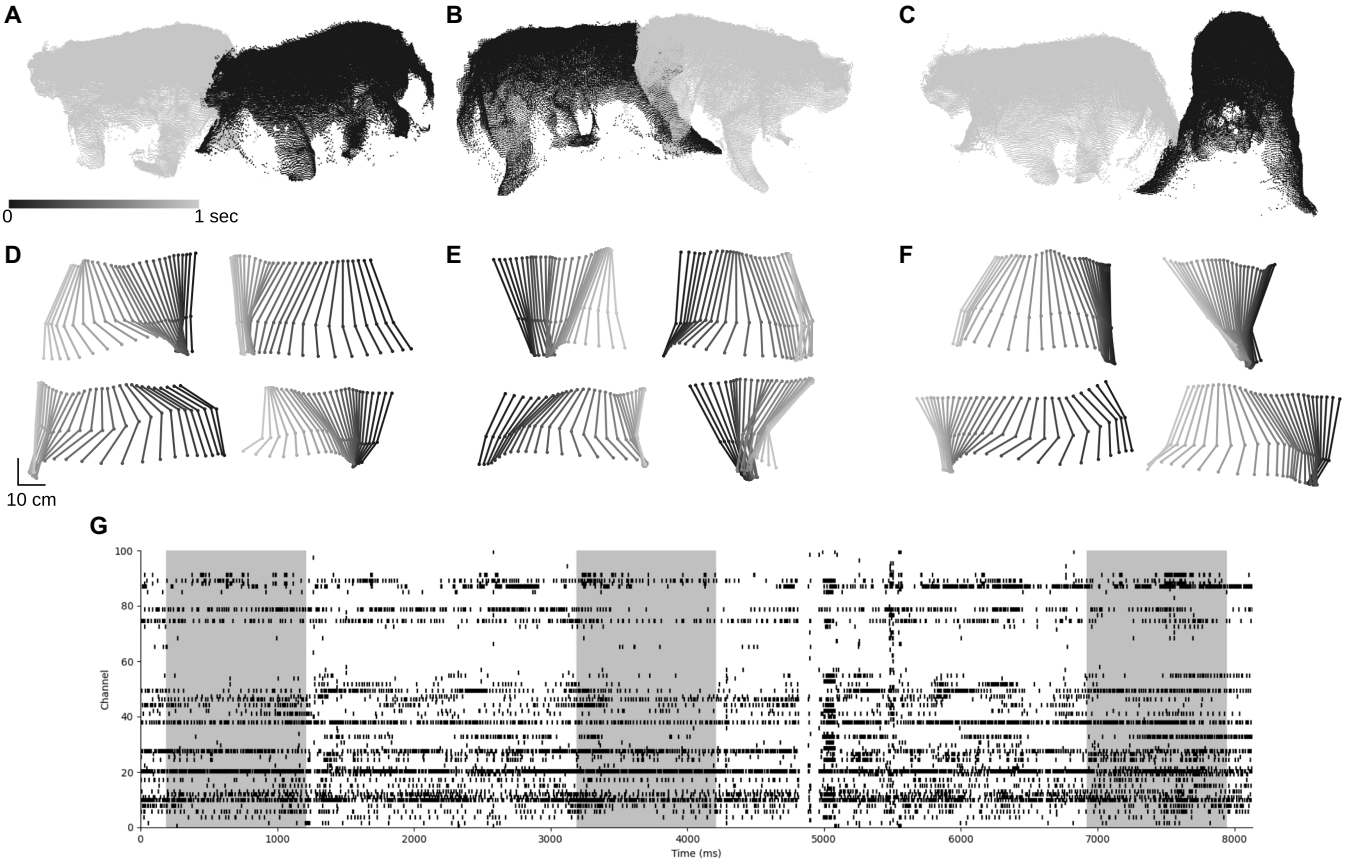


Fig. 3. Synchronized neural data and full-body kinematics. (**A** to **C**) Cropped point clouds of Monkey C across a 1-s interval, with darker shading marking the beginning of the interval and lighter shading marking the end (C210503_01). All views are from a fixed location. (**D** to **F**) Limb kinematics over 1-s interval shown in (A) to (C). In clockwise order starting from the top left: left hindlimb, right hindlimb, right forelimb, and left forelimb. Time progression is marked using the same shading scheme in (A) to (C). (**G**) Spike raster for 96 channels over 8000 ms. Highlighted sections denote the three 1-s intervals shown above.

RESULTS

Point cloud

An observational enclosure was constructed in which freely moving animal activity can be accurately characterized over a 2.1 m–by–2.1 m–by–2.4 m volume (Fig. 1B). This space, which is large enough to allow for diverse animal movements, is spanned by four depth (RGB-D) cameras. In this configuration, the entire floor space is visible to depth sensors of at least two cameras, and more than 85% of the enclosure's surface area can be detected by all four depth sensors; RGB frames from each camera capture the entire floor space. High-resolution depth and color images—measuring 640 pixels by 576 pixels and 2048 pixels by 1536 pixels, respectively—were acquired at a rate of 30 Hz (Fig. 1, C and D). Because the entire system generates data at a rate of 184 megabytes per second (MBps), all data acquisition could be implemented on a single desktop computer. Both depth and RGB data were acquired to highlight the full capabilities of the system (Fig. 2, A to C), but subsequent motion analysis only used depth data (Fig. 2, D and F).

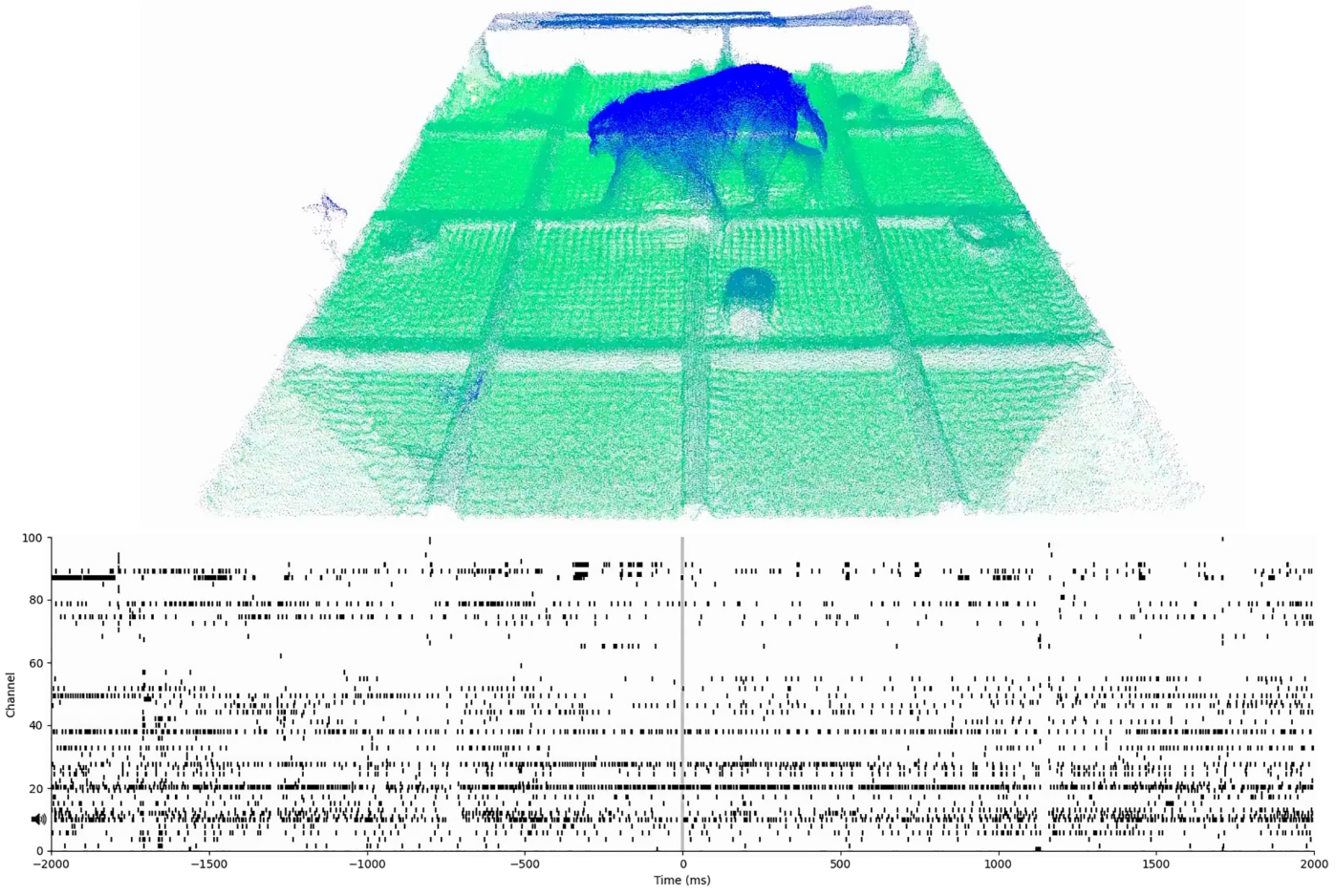
Using camera-specific intrinsic parameters, point clouds were generated by converting each depth pixel to a 3D location relative to the camera from which it was captured. Point clouds from each camera can be aligned to a single reference frame to recreate a detailed 3D representation of the enclosure.

Figure 3 shows an 8-s segment of a walking task, where the animal (Monkey C) traversed opposite sides of the enclosure for a treat reward. Figure 3 (A to C) consists of 1-s intervals, where aligned, cropped point clouds highlight Monkey C's stance during various walking stages. Taken from a fixed viewpoint, the point cloud shading depicts how the animal's position evolves during each 1-s interval. In these figures, Monkey C's head, torso, and limbs are clearly visible, allowing these body segments to be tracked over time.

Kinematics

3D full-body kinematics were extracted from the point cloud data using an articulated skeletal model and an optimization algorithm for two macaques. Full-body kinematics were extracted for locomotion and arm reaching during a walking task, two common behaviors studied in neuroscience.

Figure 3 shows full-body kinematics for three laps of a stereotyped walking behavior in 2D for ease of visualization. Monkey C walked 1.5 m between two bowls placed on opposite walls of the observational enclosure. Each lap is 30 frames (1 s) in duration and shows the transition between swing and stance phases for each limb as Monkey C walks forward. The lap ended as Monkey C reached for a treat reward. The trajectories for the right arm, left arm, right



Movie 1. Monkey C performing a walking task (C210503_01). The movie shows the synchronization of wireless electrophysiology and markerless pose tracking as Monkey C walks six laps back and forth between two bowls placed on opposite sides of the observational enclosure. Point cloud behavioral data, synchronized neural data, and audio generated from the raw voltage waveform of channel 11 are presented for the duration of the video, lasting 25 s (750 frames). Neural data are recorded from the arm region of the primary motor cortex (M1) for 96 channels. The kinematic overlay is shown for the last three laps. Kinematics for each lap last around 100 frames as the macaque walks and reaches for a food reward. The video rotates to show the full 360° view of the scene.

leg, and left leg are separated and shown in different panels for the three laps.

Kinematics were extracted for a second monkey for additional validation of the pose-tracking algorithm. Figure S1 shows full-body kinematics of monkey B for three laps of the same stereotyped walking behavior, with each lap lasting 60 frames (2 s). Although each monkey had unique movements and behaviors, the pose-tracking method could initialize an individualized skeleton, and the algorithm could track walking behavior for both. Table S1 quantifies each animal's walking movement using common gait parameters.

Algorithm accuracy

To validate the performance of the algorithm, error was calculated by comparing the algorithm results with hand-labeled joint locations. Hand-labeled data were collected by tagging the joint locations on the volumetric point cloud frames in Open3D. Because depth data captured the surface points of the macaque, both the point clouds and the resulting skeletal mesh segments modeled the surface contour of the animal's body. Therefore, hand-labeled and algorithmic joint locations could be directly compared and provide an accurate representation of the measurement error. Error was calculated for the right elbow, wrist, knee, and ankle because errors from these joints are representative of full-body errors. Four trials from the walking task (300 frames) in Fig. 3 were hand-labeled on the point clouds for these joint positions. The Euclidean distance between the algorithm's fit and the manually tracked position was calculated for each joint, and the mean and median were used to summarize the error. The median error and interquartile range were 3.6 ± 3.1 cm, 4.4 ± 3.5 cm, 3.6 ± 2.6 cm, and 3.9 ± 2.9 cm for the right elbow, wrist, knee, and ankle, respectively. For reference, Monkey C's arm (shoulder to wrist) is about 40 cm, and leg (hip to ankle) is about 45 cm.

Table 1 compares the mean error and SD for each joint to 3D errors reported in other free-behavior, markerless macaque platforms. Previous studies evaluated system performance against hand-labeled annotations (*35*, *47*, *48*). OpenMonkeyStudio calculated error from a single, head-fixed marker using an Opti-Track system (*46*). As they noted, rigid landmarks like the head are easier to track than extremities with high degrees of freedom, so this figure may not be representative of their limb error. The mean joint estimation error in this study ranged from 3 to 5 cm, generally outperforming existing platforms, potentially making this a relevant platform for motor

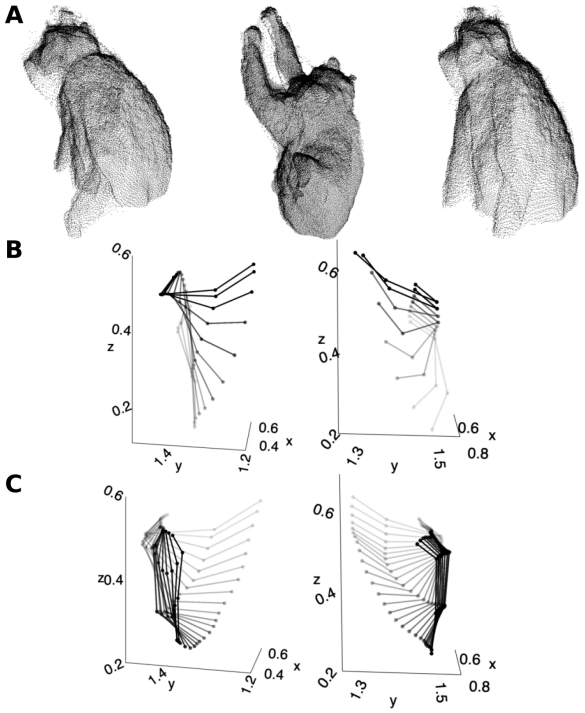


Fig. 4. Kinematics of grabbing a fly in the air. Point clouds framing the behavior are shown on the first row with 3D kinematics for each arm directly underneath. Kinematics tracking the chest, shoulders, elbows, and wrists are separated to show the trajectory for each movement. The time evolution of the kinematics is shown from lighter to darker. All coordinates are in meters and are in the same coordinate system with (0,0,0) at the far left corner of the enclosure (Fig. 1B). (**A**) Point clouds of Monkey C attempting to grab a fly showing the transitions from sitting, to reaching, and back to sitting (C210423_01). (**B**) Right arm up-reach (0.43 s) and left arm up-reach (0.27 s) (left to right). (**C**) Right arm down-reach (0.77 s) and left arm down-reach (0.93 s) (left to right).

systems neuroscience. Missing table values indicate that errors were not reported for those joints. Although a previous study reported smaller error, the error cannot be compared with this study because movement was only tracked at a fixed position in a small treadmill corridor (*35*). Figure S2 reports the error of the walking task per frame per joint in a histogram to show the error distribution. For all joints, the distribution is right-skewed, meaning that most measurement error was low.

Bodily or environmental occlusions introduced error that made initialization and algorithm fitting difficult. The monkey's fur created noise in the point cloud data as he moved, which also contributed to the error of fit.

Neural data collection

In addition to movement data from depth cameras, this setup allows for the simultaneous acquisition of wirelessly transmitted neuronal data. Through antennas mounted outside the observational enclosure, the setup supports up to two 96-channel wireless transmitters. This allows for the concurrent capture of hundreds of channels at any given time from multiple regions of the brain. These signals were acquired at 30 kHz, then filtered and thresholded to isolate spiking events. Given the absolute refractory period lasting on the order of 1 ms after a firing event, spiking data were resampled at 1 kHz without loss of information. From this, the system could calculate the firing rates of hundreds of neurons in the premotor and motor cortices as the macaque moved in the observational enclosure.

Synchronization

To extract meaningful correlations between kinematics and neural activity, these events must be precisely aligned in time. This was accomplished by generating a 30-Hz output signal using the same acquisition system that processed the wireless neural data. This signal was input to the cameras' 3.5-mm sync ports through a custom-built cable to initiate the capture of depth and RGB frames. Frame captures, along with the wireless neural signals, were time-stamped against the same system clock. Because the clock generated ticks at 30 kHz, timings were accurate to 33 μs, which is more than sufficient for analysis of spiking activity at millisecond time scales. Figure 3G displays a spike raster, where vertical lines denote spiking events across 96 channels in primary motor cortex (M1). The highlighted sections of the raster mark 1-s segments of neural data precisely aligned with the point clouds and kinematic tracings that lie in the figure panels above each highlighted segment. This alignment technique permits tight correlations between behavioral activity and the neural mechanisms that generate the activity.

Movie 1 further illustrates the same walking task depicted in Fig. 3 over a 25-s period. Aligned point clouds and synchronized neural data are visible as Monkey C ambulates across the enclosure, and full-body kinematic overlays are presented for the last three laps. Kinematics are not shown when Monkey C turns around and changes directions. Neural data are presented as spike rasters from 96 recording channels, and modulation of the spiking activity can be observed during periods where Monkey C is walking. The unfiltered waveform from channel 11 was converted into an audio signal and is presented with this recording. Spiking activity presents as popping noises in the audio. As Monkey C moves his left arm, neural modulation of the swing and reaching phase can be heard in the recording.

Generalization to arbitrary behaviors

Along with stereotyped walking behaviors, the platform can capture kinematics and neural data for arbitrary, untrained behavior. Kinematics were extracted for a variety of behaviors, such as manipulation of toys and pursuit of insects (Figs. 4 and 5). To visualize the distinct movements, the kinematics of each behavior were segmented into different panels. Shading of the trajectories denotes time evolution of the movement, with lighter traces marking the start of the movement and darker traces signifying the end of the movement. Kinematics were only shown for the upper limbs because they depict the most relevant movement in these segments.

Figure 4 shows Monkey C's arm kinematics as he attempts to catch a fly, spanning 36 frames (1.2 s). Point cloud frames depict the animal sitting, reaching, and returning to a resting position. During the upward reach, Monkey C moves his left arm at a speed of 2.0 m/s and his right arm at 1.2 m/s. This sequence shows kinematics for predatory behavior and swift movements that are otherwise challenging to capture in a constrained space. Movie 2 shows the point clouds and kinematic overlay of this behavior.

Figure 5A shows 150 frames (5 s) of arm kinematics as Monkey C walks over a tire (outer diameter, 71 cm; inner diameter, 46 cm) to forage in its center, highlighting the platform's ability to capture

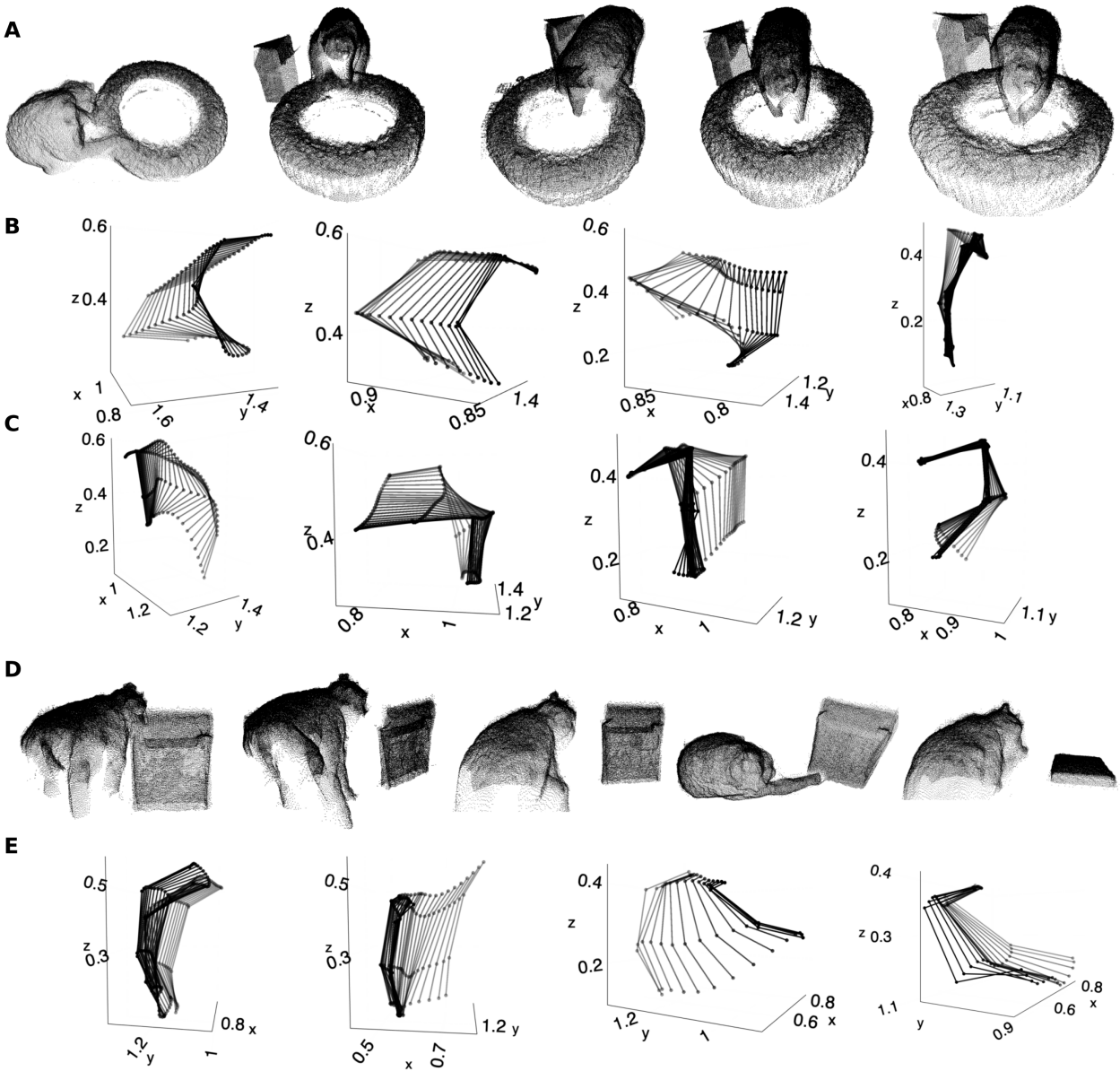
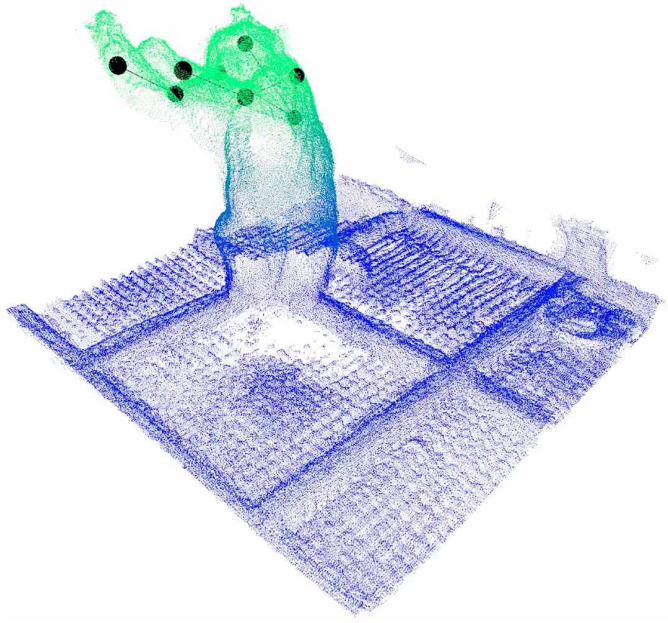


Fig. 5. Object interaction kinematics. Point clouds and kinematics have the same layout as Fig. 4. (**A**) Point clouds of Monkey C climbing into a tire and foraging for food in the center from a sitting position (C210430_02). (**B**) Right arm kinematics of standing up, moving his arm into, and resting his arm in the tire. Total time for each movement section was 0.66, 0.6, 0.7, and 2.03 s from left to right. (**C**) Left arm kinematics of standing up, bending his arm, moving his arm into, and foraging for food inside the tire. Total time for each movement section was 1.17, 0.83, 0.187, and 1.13 s from left to right. (**D**) Point clouds of circling and eating a cricket off of a box with his right arm (C210503_02). (**E**) Right arm kinematics of rearing up, moving to the side of the box, hitting the cricket off the box, and eating the cricket. Total time for each movement section was 1.33, 1.17, 0.5, and 0.5 s from left to right.

dexterous movements. Figure 5B provides further examples of Monkey C engaging in organic object interaction and predatory activity. Here, over a span of 105 frames (3.5 s), Monkey C circles a box to swipe at a cricket. Table S2 quantifies the velocities and distances of the hands and center of the macaque for these naturalistic behaviors. The simultaneous recording of kinematics and neural data for behaviors approximating those exhibited by macaques in their natural habitat highlights the neuroethological value of this platform.

DISCUSSION

This observational enclosure is a platform capable of recording 3D, full-body kinematics synchronized with wireless electrophysiological recordings in large animals. The size of the enclosure allows for a diverse range of naturalistic activities, and with the capability to sample hundreds of neural channels simultaneously, rich neural data can be acquired from multiple brain regions. The low camera count and ability to implement all data acquisition on a single desktop computer facilitates reproduction of this setup and



Movie 2. Monkey C attempting to capture a fly (C210423_01). The movie shows Monkey C swiftly reaching for a fly. The first section shows the point cloud frames of the animal sitting, reaching, and returning to a resting position. The next section shows the point cloud and the overlaid skeletal fit of Monkey C's arm kinematics for the fly reaching behavior. The last section shows the kinematics of Monkey C's skeleton alone. The video is 2 s in duration (60 frames), with the kinematic overlay for 36 of those frames. This sequence shows predatory behavior and swift movements that can be captured in this observational enclosure.

adoption by the broader neuroscience community. Point cloud data were used because they provide easy, robust calibration and accurate depth measurements. 2D depth data could be accurately transformed into 3D points, compared with a computationally intensive and error-prone transformation if RGB data were used. Furthermore, the kinematic extraction algorithm provides improved accuracy over previously reported systems that extract markerless 3D macaque kinematics.

This end-to-end system represents a robust tool through which neural control of free movement can be understood, a question that largely remains unanswered. Previous systems that pair kinematics with wireless neural recordings have laid the groundwork for understanding neural control of free behavior, yet this platform uses two key advantages. First, the open environment of the observational enclosure makes it highly reconfigurable. The platform not only supports a stereotyped, trial-based walking task that can be compared with similar studies (*34*, *35*, *39*) but also serves as a direct comparison with conventional constrained trial-based studies (*53*). In addition, it permits the capture of diverse, naturalistic behaviors that may provide a more complete understanding of the neural basis of movement (*54*). This flexibility allows task complexity to be adjusted, in both cognitive difficulty and physical movements required to complete the task objective. By varying task difficulty, relationships between task complexity and neural activity can be better understood. Second, this platform tracks kinematics from the animal's entire body as opposed to only one or two limbs. With traditional views of the motor homunculus recently challenged (*24*), the ability to quantify motion across different areas of the body allows for studies that probe the extent to which limb mappings are intermixed across motor regions of cortex. This research will build on previous, partially constrained animal studies to broaden our theoretical understanding of motor systems neuroscience and develop more accurate neural models of healthy movement. This is key to enabling improved brain-computer interfaces that restore movement and augment rehabilitation therapies after stroke and other brain injuries. Given the added insights generated by freely moving experiments, we anticipate that these platforms will quickly gain traction in the motor systems neuroscience research community.

Limitations

Point cloud errors were observed as Monkey C moved within 5 cm of the enclosure wall, possibly due to infrared (IR) interference from reflections off the acrylic panels. This was most noticeable when he turned around after collecting a food reward in the walking task. Placing IR-absorbing materials or coatings to the outside of the enclosure could address this issue. Moreover, because of camera positioning, lower limbs could become partially obscured during certain movements, such as when the animal sat down. Augmenting the enclosure with additional cameras placed near ground level may increase visibility of the lower limbs.

The pose-tracking pipeline is not fully automated, and some manual interventions were needed to reset the skeleton in the correct position when it drifted from a proper fit. For example, the initialization of the skeleton was a manual process that consisted of locating joint centers on the point cloud and creating a mesh segment around them. After this first step, the algorithm could extract joint kinematics without reinitializing the skeleton. Future work will consist of automating the initialization of the skeleton, eliminating the most manual aspect of the algorithm.

The algorithm could accurately and automatically fit the full walking movement for each limb (about 100 contiguous frames) when the full point cloud could be seen, and each segment was prominent because coverage from the four camera setup prevented most occlusions. However, some manual interventions were needed when limbs became occluded. Rare cases of occlusions were most noticeable for the lower limbs when the macaque sat down or if an object was blocking his limbs. Minor interventions consisted of rotating the mesh segments to better align with the point cloud data or changing the cost function constraint so that segments were weighted more heavily on the correct side of the body. These adjustments were commonly needed when the animal changed direction or his body blocked his limbs. Refinement of the algorithm's parameters would improve the reliability of the automated fit and diminish the frequency of manual interventions.

Future work

To extend recording capabilities, several improvements could be made to the platform. These include automating the food reward delivery system, increasing the hard disk recording capacity, and extending the battery life of the wireless headstage, which currently lasts 3 hours on a single charge. Longer recording times would enable the collection of repeated task trials, such as the stereotyped walking task presented here, which can be compared with similar motions performed in traditional, fixed-movement platforms. Sufficiently long (e.g., >12 hours) recording capabilities could permit overnight studies, including neural activity during sleep. The enclosure could also be equipped with a set of illuminated push buttons that provide visual cues and allow for the implementation

of decision-based tasks. Continued improvements will be made to improve the error of fit by leveraging other computer vision tools and techniques, such as the Bullet Physics Library (*51*) and DeepLabCut (*50*). In addition, incorporating RGB input into the algorithm for feature-based matching has the potential to further reduce error and eliminate the need manual model initialization.

MATERIALS AND METHODS

Observational enclosure

The observational enclosure, which measures 2.1 m-by-2.1 m-by-2.4 m (length by width by height), is composed of 0.375-inches-thick clear acrylic panels framed with aluminum extrusions. A small tunnel—measuring 30.5 cm–by–33 cm–by–35.5 cm (length by width by height) and constructed from opaque ABS (acrylonitrile butadiene styrene)—facilitates transfer between the enclosure and the home cage through a transfer box. Treats can be delivered to the monkey through 1-inch-diameter PVC (polyvinyl chloride) tubing that extends outside the enclosure through holes in the acrylic paneling. Food dropped in these tubes is deposited in metal food bowls secured to the enclosure floor. A pair of perch bars, raised 15 cm off the ground, lines the back wall of the enclosure and provides additional environmental enrichment for the animal. The observational enclosure can be equipped with food and water dispensing trays to make it compatible as a home cage (*55*, *56*) for long-term studies.

Camera setup and frame acquisition

Four Azure Kinect RGB-D cameras (Microsoft, Redmond, WA) were mounted at a height of about 1.5 m on each enclosure wall and angled downward to maximize floor coverage. In this configuration, the entire floor space was visible to depth sensors of at least two cameras, and more than 85% of the enclosure's surface area can be detected by all four depth sensors; RGB frames from each camera capture the entire floor space. High-resolution depth and color images—measuring 640 pixels by 576 pixels and 2048 pixels by 1536 pixels, respectively—were acquired at a rate of 30 Hz. Multiple camera frame synchronization was coordinated using the Kinect's 3.5-mm sync ports. Cameras were arranged in a daisy-chain configuration, as described in the device documentation (*57*). RGB frame acquisition occurred synchronously, and depth frame acquisition was offset by 160 μs between subsequent cameras to prevent IR laser interference. All cameras were powered by a 5-V, 12-A power supply (Phevos, Brighton, CO) via custom-built cables, and 16-m active USB A to C cables (Newnex, Santa Clara, CA) facilitated data transfer. An EliteDesk 800 G2 SFF desktop (Hewlett-Packard, Palo Alto, CA) running Ubuntu 18.04 (Canonical, London, UK) and equipped with a GeForce GT 1030 graphics card (EVGA, Brea, CA) was used to interface with cameras and store data. To maximize storage capacity, depth frames were stored in an HDF5 file format with the zlib compression filter enabled, reducing depth data volumes by a factor of 3. RGB frames were natively compressed as JPEGs and also stored in an HDF5 format. The four-camera setup generated data at a rate of 184 MBps, a 20× reduction over similarly sized, RGB-only platforms (*46*). To accommodate this data rate, frames were stored on a 2-terabyte MP600 M.2 solid-state disk (Corsair, Fremont, CA), providing sufficient storage for nearly 3 hours of continuous recording. If necessary, lower data rates and longer recording times can be achieved by adjusting RGB acquisition. Reducing RGB frame sizes to 1280 by 720 reduces the data rate to 76 MBps; collecting only depth data reduces this rate to 31 MBps.

Point cloud generation and alignment

The platform captures both depth and RGB frames, but only depth data were used for motion analysis. Cameras generated depth frames using an amplitude-modulated continuous-wave time-of-flight sensor (*58*), where the *Z*-coordinate value from the depth sensor's focal point is calculated for every pixel. Synchronized depth frames taken from cameras in our enclosure can be seen in Fig. 2A. Using the camera's intrinsic parameters, each pixel can be transformed into a 3D point with *X*-, *Y*-, and *Z*-coordinate values relative to each camera. In this way, camera-specific point clouds—as shown in Fig. 2B—are generated. Point cloud generation was facilitated by the Azure Kinect Sensor SDK 1.4.1. To create a unified point cloud (i.e., where points from all cameras are mapped onto a single common coordinate axis), an asymmetrical calibration tool was imaged in the center of the observation area. After a rough, manual alignment, an iterative closest point registration algorithm was used to construct the transformation matrices from three other cameras into the fourth's frame of reference (*59*). Points were then filtered on the basis of the physical dimensions of the enclosure. All point cloud alignment, filtering, and visualization were facilitated through the Open3D 0.12.0 Python package (*60*).

Neural data acquisition

Neural data are transmitted from the animal through a wireless headstage (CerePlex W, Blackrock Microsystems, Salt Lake City, UT). Outside the enclosure, 16 panel antennas (PA-333810-NF 3.3GHz-3.8GHz 10dBi Panel Directional Outdoor Antenna, FT-RF, Jhubei City, Taiwan) are evenly spaced to receive signals from two wireless headstages operating at distinct frequencies in the 3- to 4-GHz range. This allows hundreds of channels of neural data to be simultaneously recorded. Raw neural data were sampled at 30 kHz and saved to an SQLite database. At this sampling frequency, 192 channels generate a data rate of about 11 MBps.

Neural data processing

Raw neural data generated in our enclosure were sampled at a rate of 30 kHz. Two example channels of raw neural data can be seen in Fig. 2G, and additional channels are shown in fig. S3A. All raw channels were filtered using a fourth-order, high-pass, digital Butterworth filter with a cutoff frequency of 250 Hz (Fig. 2H and fig. S3B). Using a 60-s baseline of data, the root mean square (RMS) signal value was calculated on a per-channel basis. Multiplying this RMS value by a factor of −1.5 provides a threshold for neuron spiking detection. Thresholds for the example channels appear as dashed lines in Fig. 2H and fig. S3B. With the isolation of spiking events (fig. S3C), a spike raster—which depicts neural spiking activity across recording channels—can be generated. As shown in Fig. 2I, recording channels are stacked vertically on the *y* axis, and vertical lines represent spiking events recorded by each channel as they occur in time. These rasters depict how neural activity in a recorded region modulates over time (fig. S4). All raw and derived neural data were stored in an HDF5 file format with zlib compression enabled.

Data synchronization

To synchronize all point cloud and neural data, a 30-Hz, active high, 150-μs pulse width trigger signal was generated by the same

acquisition system (Cerebus, Blackrock Microsystems, Salt Lake City, UT) used to process all wireless neural data. This trigger signal was passed to a TLC2272 rail-to-rail operational amplifier (Texas Instruments, Dallas, TX) in a noninverting configuration, raising the output signal to 5 V. The signal was sent to the cameras' 3.5-mm sync port through a custom-made cable to initiate frame acquisition. These 30-Hz pulses were timestamped by the acquisition system and saved to an SQLite database, where they could be aligned to the timestamped neural data. Because the system clock generates ticks at 30 kHz, timings are accurate to 33 μs, which is more than sufficient for the analysis of spiking activity at millisecond time scales.

Animal procedures and behavioral task

All animal procedures and protocols were approved by the Stanford University Institutional Animal Care and Use Committee. One adult male rhesus macaque (*Macaca mulatta*, Monkey C), age 14, of Chinese origin, weighing 15.3 kg, was implanted with two 96-channel multielectrode arrays (Blackrock Microsystems, Salt Lake City, UT) in arm regions of primary motor (M1) and dorsal premotor (PMd) cortex determined by visual anatomical landmarks using standard neurosurgical techniques on 25 March 2021.

Monkey C was trained to enter the observational enclosure through a tunnel system (visible at bottom left of Fig. 1B). Once in the enclosure, he would be presented with one of two tasks. The first task was a simple walking task, where an auditory cue would prompt him to walk to one of two bowls placed on opposite sides of the enclosure and grab a treat. In this way, Monkey C was trained to walk laps between the two bowls. The second task consisted of placing toys or insects, such as fruit flies and crickets, and observing the monkey's unprompted interactions. For each session, Monkey C spent around 20 min in this space while the cameras captured his behavior, after which he exited the enclosure through the same tunnel, returning to his home cage. All participation in the tasks was voluntary and positively reinforced.

Articulated skeleton model

A customized, articulated skeleton was created to recover the pose of a macaque and used to fit each subsequent data frame. The skeleton was created in Blender from primitive meshes that represent the different segments (Fig. 2D) (*61*). A mesh of the point cloud was created in MeshLab using the ball pivoting algorithm, and meshes for each segment were created by using Boolean difference with primitive spheres to model the surface of the point cloud (*62*). The meshes were fit to the distal ends of each segment to weigh the position at the joint most heavily. This ensured that segments that depend on the previous segment's location were set up at the correct place. Bones were attached to each mesh so that rotating or translating a bone also rotated or translated the mesh to a given angle or position. For the first frame, the articulated skeleton was fit by hand to a model point cloud that distinctly showed all segments of interest. For the walking behavior, the skeleton was structured in a bottom-up fashion starting at the tail because the torso was the most stable and unchanging segment. For the general behaviors fitting the upper limbs, the skeleton was structured in a top-down fashion starting at the head for the same reason. The macaque skeleton consisted of 15 segments and 15 joints (head, chest, torso, left and right shoulders, elbows, wrists, hips, knees, and ankles). Segments were rotated around the joints at angles bound by anatomical constraints.

Pose-tracking algorithm

The pose-tracking algorithm was coded in Python and primarily used Open3D and Blender to solve for the kinematics of an ambulatory macaque (*61*). This problem is a nonconvex, expensive black-box function with no closed-form analytical solution. A global Bayesian optimization was chosen to minimize this cost function (*63*). It uses a Gaussian process before fitting the data and an acquisition function to guide the process of choosing further points. The Gaussian process models the function f as a probability distribution to compute the conditional distribution, formulated as follows (*64*)

$$P(f(x)\,|\,\mathbf{P}_{1:t}, x) = \mathcal{N}(\mu_t(x), \sigma_t^2(x)) \quad (1)$$

where $\mathbf{P}_{1:t}$, x is the set of observations $f(x_1), f(x_2), \ldots, f(x_t)$; $x_1, x_2, \ldots, x_t$ are parameter values of the observed variables; and $\mathcal{N}$ is the multivariate normal distribution. This conditional distribution is used by the acquisition function to estimate values of x to compute $f(x)$.

The acquisition function is the upper confidence bound (UCB) and calculates how desirable it is to sample at point x, formulated as (*64*)

$$\mathrm{UCB}(x) = \mu_t(x) + \kappa\sigma_t(x) \quad (2)$$

where $\mu_t(x)$ is the mean and $\sigma_t(x)$ is the variance at point x. The sum is weighted by κ, which can take on values between 10^{-3} and 10^3 and is a user parameter that tunes the trade-off between exploration and exploitation. A higher κ enforces unexplored solutions in the search space, and a lower κ enforces high-performing solutions. For this algorithm, κ was set to 2.5, weighing exploration slightly higher than exploitation.

To summarize, Eqs. 1 and 2 are the definitions of Bayesian optimization functions (*64*). A rotation angle is chosen by optimizing the acquisition function (Eq. 2) over the Gaussian process (Eq. 1). The objective function (Eq. 3) samples the chosen angle and outputs the number of points inside a mesh segment.

The algorithm optimized the following cost function: Maximize the number of points inside each segment to calculate the best-fit joint angle and location (Eq. 3). The inputs to the optimization problem were a specific pose—a set of points represented by 3D coordinates (x, y, z)—and a mesh primitive representing the body segment being fit to the point cloud. The resulting output for individual segments was a 6D array consisting of a joint angle (a, b, c) and location (x, y, z).

Ray casting was used to determine whether a point lies within a mesh object (*65*). Briefly, ray casting draws an infinitely long ray in the x, y, and z directions from a specified point. It counts the number of intersections between the ray and the object, and if the number is odd, the point lies within the mesh; if the number is even, the point lies outside the mesh. By iterating through all the points in the point cloud, the number of points within a mesh segment can be counted for a given angle rotation or translation. The joint angle or translation that yields the maximum number of points inside the mesh is the solution for that segment for that frame.

The cost function maximized the number of points inside the mesh of interest. Because this problem was nonconvex, the constraints of the cost function described the anatomical limitations of the body. The equation for the cost function is detailed as

$$\max.\mathbf{C} = 5\alpha(x,y,z) + 10\beta(x,y,z) + 10\gamma(x,y,z) - 10\delta(x,y,z) \quad (3)$$

where **C** is the number of points within the mesh, α is if bone location is within 5 cm of the previous bone location, β is if bone location is on the correct lateral side of the body, γ is if bone location is on the correct transverse side of the body, and δ is if bone location is inside another primitive segment. A graphical representation of Eq. 3 is provided in fig. S5 to exemplify how each cost function element operates in practice. For the tire movement, an additional cost function measure was added to negatively weigh if the bone location is within the tire, modeled as a primitive torus.

Functionally, this equation positively weighs continuity of movement and anatomical correctness, such as when limb segments are on the correct side of the body in the lateral and transverse plane. It penalizes anatomical impossibilities, such as when the end of a bone segment is inside other segments.

Blender was compiled into a Python module to integrate the pipeline into a singularity container for easy parallelization on the Sherlock cluster, managed by the Stanford Research Computing Center. Pool-based parallel execution ran on multiple central processing unit (CPU) cores for each segment. Each side of the body can be fit independently, allowing all limbs to be fit in parallel.

The algorithm results were then smoothed out using a low-pass seven-tap linear-phase FIR (finite impluse response) filter at a cutoff frequency of 5 Hz sampled at 30 Hz. The low-pass filter was applied to the *x*, *y*, and *z* dimensions separately.

SUPPLEMENTARY MATERIALS

www.science.org/doi/10.1126/scirobotics.abj7045
Figs. S1 to S5
Tables S1 and S2

REFERENCES AND NOTES

1. S. H. Scott, J. F. Kalaska, Voluntary movement: Motor cortices, in *Principles of Neural Science, Sixth Edition*, E. R. Kandel, J. D. Koester, S. H. Mack, S. A. Siegelbaum, Eds. (McGraw-Hill, 2021), ch. 34, pp. 809–853.
2. N. Kriegeskorte, Relating population-code representations between man, monkey, and computational models. *Front. Neurosci.* **3**, 363–373 (2009).
3. U. Castiello, The neuroscience of grasping. *Nat. Rev. Neurosci.* **6**, 726–736 (2005).
4. H. J. Bellen, C. Tong, H. Tsuda, 100 years of *Drosophila* research and its impact on vertebrate neuroscience: A history lesson for the future. *Nat. Rev. Neurosci.* **11**, 514–522 (2010).
5. B. Ellenbroek, J. Youn, Rodent models in neuroscience research: Is it a rat race? *Dis. Model. Mech.* **9**, 1079–1087 (2016).
6. D. S. Bassett, O. Sporns, Network neuroscience. *Nat. Neurosci.* **20**, 353–364 (2017).
7. B. McCowan, K. Anderson, A. Heagarty, A. Cameron, Utility of social network analysis for primate behavioral management and well-being. *Appl. Anim. Behav. Sci.* **109**, 396–405 (2008).
8. E. Bizzi, M. C. Tresch, P. Saltiel, A. d'Avella, New perspectives on spinal motor systems. *Nat. Rev. Neurosci.* **1**, 101–108 (2000).
9. A. P. Georgopoulos, R. Caminiti, J. F. Kalaska, J. T. Massey, Interruption of motor cortical discharge subserving aimed arm movements. *Exp. Brain Res.* **49**, 327–336 (1983).
10. E. V. Evarts, Motor cortex reflexes associated with learned movement. *Science* **179**, 501–503 (1973).
11. M. H. Schieber, A. V. Poliakov, Partial inactivation of the primary motor cortex hand area: Effects on individuated finger movements. *J. Neurosci.* **18**, 9038–9054 (1998).
12. S. H. Scott, Optimal feedback control and the neural basis of volitional motor control. *Nat. Rev. Neurosci.* **5**, 532–545 (2004).
13. A. Ramakrishnan, P. J. Ifft, M. Pais-Vieira, Y. W. Byun, K. Z. Zhuang, M. A. Lebedev, M. A. L. Nicolelis, Computing arm movements with a monkey brainet. *Sci. Rep.* **5**, 1–15 (2015).
14. E. V. Okorokova, J. M. Goodman, N. G. Hatsopoulos, S. J. Bensmaia, Decoding hand kinematics from population responses in sensorimotor cortex during grasping. *J. Neural Eng.* **17**, 046035 (2020).
15. D. A. Schwarz, M. A. Lebedev, T. L. Hanson, D. F. Dimitrov, G. Lehew, J. Meloy, S. Rajangam, V. Subramanian, P. J. Ifft, Z. Li, A. Ramakrishnan, A. Tate, K. Z. Zhuang, M. A. Nicolelis, Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nat. Methods* **11**, 670–676 (2014).
16. M. Matsumura, T. Cope, E. Fetz, Sustained excitatory synaptic input to motor cortex neurons in awake animals revealed by intracellular recording of membrane potentials. *Exp. Brain Res.* **70**, 463–469 (1988).
17. V. Mountcastle, B. Motter, M. Steinmetz, A. Sestokas, Common and differential effects of attentive fixation on the excitability of parietal and prestriate (V4) cortical visual neurons in the macaque monkey. *J. Neurosci.* **7**, 2239–2255 (1987).
18. M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, K. V. Shenoy, Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
19. J. P. Nguyen, F. B. Shipley, A. N. Linder, G. S. Plummer, M. Liu, S. U. Setru, J. W. Shaevitz, A. M. Leifer, Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1074–E1081 (2016).
20. K. Svoboda, N. Li, Neural mechanisms of movement planning: Motor cortex and beyond. *Curr. Opin. Neurobiol.* **49**, 33–41 (2018).
21. F. Helmchen, M. S. Fee, D. W. Tank, W. Denk, A miniature head-mounted two-photon microscope. *Neuron* **31**, 903–912 (2001).
22. P. Gao, S. Ganguli, On simplicity and complexity in the brave new world of large-scale neuroscience. *Curr. Opin. Neurobiol.* **32**, 148–155 (2015).
23. H. Sompolinsky, Computational neuroscience: Beyond the local circuit. *Curr. Opin. Neurobiol.* **25**, xiii–xviii (2014).
24. F. R. Willett, D. R. Deo, D. T. Avansino, P. Rezaii, L. R. Hochberg, J. M. Henderson, K. V. Shenoy, Hand knob area of premotor cortex represents the whole body in a compositional way. *Cell* **181**, 396–409.e26 (2020).
25. V. Gilja, C. A. Chestek, P. Nuyujukian, J. Foster, K. V. Shenoy, Autonomous head-mounted electrophysiology systems for freely behaving primates. *Curr. Opin. Neurobiol.* **20**, 676–686 (2010).
26. D. A. Borton, M. Yin, J. Aceros, A. Nurmikko, An implantable wireless neural interface for recording cortical circuit dynamics in moving primates. *J. Neural Eng.* **10**, 026010 (2013).
27. G. Santhanam, M. D. Linderman, V. Gilja, A. Afshar, S. I. Ryu, T. H. Meng, K. V. Shenoy, HermesB: A continuous neural recording system for freely behaving primates. *IEEE Trans. Biomed. Eng.* **54**, 2037–2050 (2007).
28. C. A. Chestek, V. Gilja, P. Nuyujukian, R. J. Kier, F. Solzbacher, S. I. Ryu, R. R. Harrison, K. V. Shenoy, HermesC: Low-power wireless neural recording system for freely moving primates. *IEEE Trans. Neural Syst. Rehabil. Eng.* **17**, 330–338 (2009).
29. H. Miranda, V. Gilja, C. A. Chestek, K. V. Shenoy, T. H. Meng, HermesD: A high-rate long-range wireless transmission system for simultaneous multichannel neural recording applications. *IEEE Trans. Biomed. Circuits Syst.* **4**, 181–191 (2010).
30. M. P. Maher, J. Pine, J. Wright, Y.-C. Tai, The neurochip: A new multielectrode device for stimulating and recording from cultured neurons. *J. Neurosci. Methods* **87**, 45–56 (1999).
31. A. Jackson, C. T. Moritz, J. Mavoori, T. H. Lucas, E. E. Fetz, The neurochip BCI: Towards a neural prosthesis for upper limb function. *IEEE Trans. Neural Syst. Rehabil. Eng.* **14**, 187–190 (2006).
32. S. Zanos, A. G. Richardson, L. Shupe, F. P. Miles, E. E. Fetz, The Neurochip-2: An autonomous head-fixed computer for recording and stimulating in freely behaving monkeys. *IEEE Trans. Neural Syst. Rehabil. Eng.* **19**, 427–435 (2011).
33. M. Yin, D. A. Borton, J. Komar, N. Agha, Y. Lu, H. Li, J. Laurens, Y. Lang, Q. Li, C. Bull, L. Larson, D. Rosler, E. Bezard, G. Courtine, A. V. Nurmikko, Wireless neurosensor for full-spectrum electrophysiology recordings during free behavior. *Neuron* **84**, 1170–1182 (2014).
34. M. Berger, N. S. Agha, A. Gail, Wireless recording from unrestrained monkeys reveals motor goal encoding beyond immediate reach in frontoparietal cortex. *eLife* **9**, e51322 (2020).
35. J. D. Foster, P. Nuyujukian, O. Freifeld, H. Gao, R. Walker, S. I. Ryu, T. H. Meng, B. Murmann, M. J. Black, K. V. Shenoy, A freely-moving monkey treadmill model. *J. Neural Eng.* **11**, 046020 (2014).
36. J. D. Foster, O. Freifeld, P. Nuyujukian, S. I. Ryu, M. J. Black, K. V. Shenoy, Combining wireless neural recording and video capture for the analysis of natural gait, in *2011 5th International IEEE/EMBS Conference on Neural Engineering* (IEEE, 2011), pp. 613–616.
37. A. Presacco, R. Goodman, L. Forrester, J. L. Contreras-Vidal, Neural decoding of treadmill walking from noninvasive electroencephalographic signals. *J. Neurophysiol.* **106**, 1875–1887 (2011).
38. N. Fitzsimmons, M. Lebedev, I. Peikon, M. A. L. Nicolelis, Extracting kinematic parameters for monkey bipedal walking from cortical neuronal ensemble activity. *Front. Integr. Neurosci.* **3**, 3 (2009).
39. M. Capogrosso, T. Milekovic, D. Borton, F. Wagner, E. M. Moraud, J. B. Mignardot, N. Buse, J. Gandar, Q. Barraud, D. Xing, E. Rey, S. Duis, Y. Jianzhong, W. K. Ko, Q. Li, P. Detemple, T. Denison, S. Micera, E. Bezard, J. Bloch, G. Courtine, A brain–spine interface alleviating gait deficits after spinal cord injury in primates. *Nature* **539**, 284–288 (2016).
40. S. Rajangam, P.-H. Tseng, A. Yin, G. Lehew, D. Schwarz, M. A. Lebedev, M. A. Nicolelis, Wireless cortical brain-machine interface for whole-body navigation in primates. *Sci. Rep.* **6**, 122170 (2016).

41. S. Mori, F. Mori, K. Nakajima, *Adaptive Motion of Animals and Machines* (Springer, 2006), pp. 53–65.
42. Y. Higurashi, M. A. Maier, K. Nakajima, K. Morita, S. Fujiki, S. Aoi, F. Mori, A. Murata, M. Inase, Locomotor kinematics and EMG activity during quadrupedal versus bipedal gait in the Japanese macaque. *J. Neurophys.* **122**, 398–412 (2019).
43. M. A. Novak, J. S. Meyer, Alopecia: Possible causes and treatments, particularly in captive nonhuman primates. *Comp. Med.* **59**, 18–26 (2009).
44. J. A. Kramer, K. G. Mansfield, J. H. Simmons, J. A. Bernstein, Psychogenic alopecia in rhesus macaques presenting as focally extensive alopecia of the distal limb. *Comp. Med.* **61**, 263–268 (2011).
45. A. Heagerty, R. A. Wales, K. Prongay, D. H. Gottlieb, K. Coleman, Social hair pulling in captive rhesus macaques (*Macaca mulatta*). *Am. J. Primatol.* **79**, e22720 (2017).
46. P. C. Bala, B. R. Eisenreich, S. B. M. Yoo, B. Y. Hayden, H. S. Park, J. Zimmermann, Automated markerless pose estimation in freely moving macaques with OpenMonkeyStudio. *Nat. Commun.* **11**, 4560 (2020).
47. R. Labuguen, J. Matsumoto, S. B. Negrete, H. Nishimaru, H. Nishijo, M. Takada, Y. Go, K. I. Inoue, T. Shibata, MacaquePose: A novel "in the wild" macaque monkey pose dataset for markerless motion capture. *Front. Behav. Neurosci.* **14**, 581154 (2021).
48. T. Nakamura, J. Matsumoto, H. Nishimaru, R. V. Bretas, Y. Takamura, E. Hori, T. Ono, H. Nishijo, A markerless 3D computerized motion capture system incorporating a skeleton model for monkeys. *PLOS ONE* **11**, e0166154 (2016).
49. E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, B. Schiele, DeeperCut: A deeper, stronger, and faster multi-person pose estimation model, in *European Conference on Computer Vision* (Springer, 2016), pp. 34–50.
50. A. Mathis, P. Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis, M. Bethge, DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
51. E. Coumans, Y. Bai, PyBullet, A Python module for physics simulation for games, robotics and machine learning (2016); http://pybullet.org.
52. H. Gao, R. M. Walker, P. Nuyujukian, K. A. Makinwa, K. V. Shenoy, B. Murmann, T. H. Meng, HermesE: A 96-channel full data rate direct neural interface in 0.13 μm CMOS. *IEEE J. Solid State Circuits* **47**, 1043–1055 (2012).
53. J. F. Kalaska, Emerging ideas and tools to study the emergent properties of the cortical neural circuits for voluntary motor control in non-human primates. *F1000Research* **8**, 749 (2019).
54. J. W. Krakauer, A. A. Ghazanfar, A. Gomez-Marin, M. A. MacIver, D. Poeppel, Neuroscience needs behavior: Correcting a reductionist bias. *Neuron* **93**, 480–490 (2017).
55. National Research Council (US) Committee for the Update of the Guide for the Care and Use of Laboratory Animals, *Guide for the Care and Use of Laboratory Animals* (National Academies Press, 2011).
56. National Research Council (US) Committee on Guidelines for the Use of Animals in Neuroscience and Behavioral Researh, *Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research* (National Academies Press, 2003).
57. T. Sych, Synchronize multiple Azure Kinect DK devices (2020); https://docs.microsoft.com/en-us/azure/kinect-dk/multi-camera-sync.
58. C. S. Bamji, S. Mehta, B. Thompson, T. Elkhatib, S. Wurster, O. Akkaya, A. Payne, J. Godbaz, M. Fenton, V. Rajasekaran, L. Prather, S. Nagaraja, V. Mogallapu, D. Snow, R. McCauley, M. Mukadam, I. Agi, S. McCarthy, Z. Xu, T. Perry, W. Qian, V.-H. Chan, P. Adepu, G. Ali, M. Ahmed, A. Mukherjee, S. Nayak, D. Gampell, S. Acharya, L. Kordus, P. O'Connor, IMpixel 65nm BSI 320MHz demodulated TOF Image sensor with 3μm global shutter pixels and analog binning, in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)* (IEEE, 2018), pp. 94–96.
59. C.-H. Chen, D. Ramanan, 3D human pose estimation = 2D pose estimation + matching, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 7035–7043.
60. Q.-Y. Zhou, J. Park, V. Koltun, Open3D: A modern library for 3D data processing. arXiv:1801.09847 [cs.CV] (30 January 2018).
61. B. O. Community, Blender—A 3D modelling and rendering package (Blender Foundation, 2018); http://blender.org.
62. M. Corsini, P. Cignoni, R. Scopigno, Efficient and flexible sampling with blue noise properties of triangular meshes. *IEEE Trans. Visual. Comput. Graphics* **18**, 914–924 (2012).
63. F. Nogueira, Bayesian Optimization: Open source constrained global optimization tool for Python (2014); https://github.com/fmfn/BayesianOptimization.
64. J. Snoek, H. Larochelle, R. P. Adams, Practical Bayesian optimization of machine learning algorithms, in *Advances in Neural Information Processing Systems 25 (NIPS 2012)* (Association for Computing Machinery, 2012).
65. S. D. Roth, Ray casting for modeling solids. *Comput. Graphics Image Process.* **18**, 109–144 (1982).

Acknowledgments: We thank K. Lebedev for animal care and K. Chin and M. Truong for administrative support. The members of the Brain Interfacing Laboratory are Michelle S. Wechsler, Mackenzie J. Risch, Stephen I. Ryu, Stephen E. Clarke, and Iliana E. Bray. M. S. Wechsler and M. J. Risch were responsible for animal care, surgical support, and guiding behavioral training. S. I. Ryu was responsible for nonhuman primate array implantation. S. E. Clarke provided surgical assistance and evaluation of the quality of microelectrode arrays. S. E. Clarke and I. E. Bray assisted in animal care. **Funding:** This work was funded by an NSF GRFP (DGE-1656518) to M.P.S. and the Stanford University Wu Tsai Neurosciences Institute.
Author contributions: M.P.S. and A.S.L. were responsible for collecting data with Monkey C (investigation). M.P.S. and A.S.L. constructed and validated the observational enclosure (methodology). M.P.S. was responsible for developing the data capture pipeline, which recorded from cameras and synchronizing neural data (methodology and software). A.S.L. was responsible for creating and implementing the markerless pose-tracking algorithm (methodology and software). Data curation and data visualization were performed by M.P.S. and A.S.L. M.P.S. and A.S.L. wrote the original draft and reviewed and edited the manuscript. P.N. was involved in all aspects of the project, conceptualized the work, supervised the effort, provided funding, and reviewed and edited the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

Submitted 29 May 2021
Accepted 10 August 2021
Published 8 September 2021
10.1126/scirobotics.abj7045

Citation: M. P. Silvernagel, A. S. Ling, P. Nuyujukian, for the Brain Interfacing Laboratory, A markerless platform for ambulatory systems neuroscience. *Sci. Robot.* **6**, eabj7045 (2021).

Science Robotics

A markerless platform for ambulatory systems neuroscience

Michael P. Silvernagel, Alissa S. Ling, Paul Nuyujukian, and for the Brain Interfacing Laboratory

Sci. Robot. **6** (58), eabj7045. DOI: 10.1126/scirobotics.abj7045

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.