Development of a new 3D tracking system for multiple marmosets under free-moving conditions

NEW TRACKING SYSTEM FOR FREE-RUNNING CONDITIONS MARMOSETS

Terumi Yurimoto[1], Wakako Kumita[1], Kenya Sato[1], Rika Kikuchi[1], Yusuke Shibuki [1], Rino Hashimoto[1], Michiko Kamioka[1], Yumi Hayasegawa[1], Eiko Yamazaki[1], Yoko Kurotaki[2], Norio Goda[3], Junichi Kitakami[4], Tatsuya Fujita[5], Takashi Inoue[1], Erika Sasaki[1*]

[1]Department of Marmoset Biology and Medicine, Central Institute for Experimental Animals Kawasaki, 210-0821 JAPAN

[2]Center of Basic Technology in Marmoset, Central Institute for Experimental Animals Kawasaki, 210-0821 JAPAN

[3]Public Digital Transformation Department, Hitachi, Ltd., Shinagawa-ku, 140-8512 JAPAN

[4]Vision AI Solution Design Department Hitachi Solutions Technology, Ltd., Tachikawa, 190-0014 JAPAN

[5]Engineering Department Eastern Japan division, Totec Amenity Limited, Shinjuku-ku, 163-0417 JAPAN

*Corresponding author: Erika Sasaki (esasaki@ciea.or.jp)

19

Abstract

Accurate assessment of social interactions in mammalian models are necessary to elucidate the pathogenesis of psychiatric and neurodegenerative diseases. The common marmoset (*Callithrix jacchus*) is a useful model for understanding the mechanisms of these disorders. However, behavioral measurements of free-moving group animals, including assessment of sociality have not been conducted. Here, we developed a new animal behavioral analysis system for three-dimensional (3D) trajectory and behavior of free-moving individuals. Each marmoset was identified using deep learning facial recognition (accuracy ≥ 97%). Their trajectories were tracked independently by combining video, 3D coordinates detected by light detection and ranging (lidar) systems, and facial identification. Location preferences and distances between individuals were calculated using the 3D trajectories and grooming behavior was detected by deep learning. This novel system will allow the quantification of individual captive animals within groups, facilitating the quick and automatic measurement of social behavior, and enhancing development of objective behavioral indices.

Introduction

Neurological disorders, such as neurodegenerative diseases and psychiatric disorders, are associated with declines in cognitive function and abnormal behaviors, which fundamentally affect social life[1, 2, 3, 4]. Mouse models, including genetically modified mice, have contributed to our understanding of brain structures, elucidating mechanisms of the onset and progression of neuronal disease, and behavioral changes caused by diseases. However, due to the anatomical and physiological differences between rodents and humans, mouse models do not fully recapitulate the pathogenesis of neuronal diseases. Therefore, there are limitations to using mouse models to address questions around human brain function[5, 6, 7]. In contrast, nonhuman primates do share anatomical and physiological features with humans, especially in specialized functions acquired through unique cerebral expansions, such as tool use, language, and self-awareness[7, 8, 9, 10]. As these functions are particularly affected by neurodegenerative diseases and psychiatric disorders, nonhuman primate models are crucial for understanding the development of neuronal and psychiatric disorders.

The common marmoset (marmoset, *Callithrix jacchus*) is a small nonhuman primate with behavioral and social characteristics like those of humans: they are diurnal, form monogamous family units, engage in altruistic behaviors, such as food sharing, and cooperate with all family members to raise their youngest babies[9, 11, 12]. As marmoset body size and social units are small, maintaining complete social units for research infrastructure in a laboratory is easier than for other primates, such

53 as macaques. Therefore, maintaining marmosets in families as social units is ideal for social behavioral
54 study[9]. Furthermore, the recent development of genetic modification techniques, including gene-
55 editing technologies in nonhuman primates, made it possible to produce several genetically modified
56 disease model marmosets, such as neurodegenerative models[13, 14, 15]. The use of the marmoset as an
57 animal model is attracting attention because of its advantages as a model animal over rodents. Using
58 a marmoset model may elucidate basic gene functions in brain and molecular mechanisms in
59 genetically modified animals and aid in implementing cognitive neuroscience techniques in primate
60 research[7, 16], particularly national brain projects using marmosets have been conducted for these
61 purposes[17, 18].

62 Behavioral analyses are important for the validation of marmoset disease models. Various
63 behavioral task tests have been conducted for marmosets, including those using touch panels, tool
64 tasks, and evaluation of muscle motility[8, 19, 20]. However, when a novel disease model marmoset is
65 developed, it is unclear when and how behavioral changes and social interactions will arise due to the
66 disease. Even if multiple types of exhaustive behavioral task tests are conducted, it is not always
67 possible to accurately capture behavioral changes through these tests. Further, it is often challenging
68 to detect behavioral changes, such as increased appetite, hypoactivity, aggression, irritability, and
69 apathy, using behavioral task tests. It is therefore important to clarify various behavioral changes,
70 including changes in social interactions, in animal disease models based on their behavior during free

activity. In mice, there are some systems for analyzing animal behavior during free-ranging activities under group-housed conditions[21, 22, 23]. However, unlike mice, marmosets move quickly in three dimensions (3D), thus mouse-specific systems cannot be directly applied to marmosets. A novel system that can automatically and accurately detect and quantify marmoset behavior is required. Marmoset behavior tracking measurement systems under free-moving activities have been previously reported, such as a sleep study using an activity meter, head and eye tracking, and tracking of a single individual using a depth sensor[24, 25, 26]. However, the activity meter and the head and eye tracking are limited because they cannot track animal trajectories. Additionally, the behavior tracking measurement systems using the Kinect system depth sensor cannot analyze intra-familial marmosets under free-moving conditions because it does not support simultaneous analysis of multiple animals and it requires the use of a dedicated analysis cage. Another method using infrared motion sensors has been reported[27]; however, it is limited to two-dimensional motion and cannot detect the position and behavior of multiple marmosets under free motion.

Automatically detecting marmoset trajectories, various behaviors, and interactions between individuals under free-moving conditions in multiple animals in a cage is especially challenging, as there are many issues not present in other behavioral task tests or the quantification of behaviors in single animals. For example, marmosets are often close to each other when grooming, mating, or attacking other individuals[28]. This proximity make it difficult to detect individual bodies,

which is a complication to individual tracking and behavior detection. Furthermore, when individual tracking and behavior detection are performed in an animal's home cage, movable objects, such as toys for environmental enrichment and food dishes, can also obstruct tracking. Although there are many challenges, analysis in the home cage environment has the advantage of avoiding the stress caused by an unfamiliar experimental environment when observing family behavior. There have been previous attempts of pose estimation and tracking of multiple marmosets. However, individual identification is difficult, and tracking of grooming individuals remains a challenge, thus, successful long-term simultaneous tracking of multiple animals without labeling has not yet been achieved[29].

Here, we develop a 3D tracking technique using a combination of a light detection and ranging (lidar) system and deep learning to simultaneously track a marmoset family of three in the same cage. To match the tracking trajectory with the individual animal information, facial recognition was performed to identify individuals. Furthermore, as an example of social behavior detection, we attempted to automatically detect grooming behaviors using deep learning. The technology developed in this study will have applications in detecting important behavioral changes caused by a variety of diseases in small nonhuman model primates.

Results

Facial recognition

Five marmosets of the same generation (1-2 years old) were used to examine whether deep learning for face identification could be adapted to marmosets. These marmosets were kept in different cages in a conventional breeding colony and facial images were collected individually. To evaluate the trained models, 500 new facial images (100 for each animal) were obtained from the videos used for training. From the evaluation, the evaluation metrics for deep learning were as follows: specificity of 99.8%-100%, accuracy of 99.4%-99.8%, recall of 98.0%-100%, precision of 99%-100%, and an F-measure of 98.5%-99.5% (Table 1). The overall identification accuracy (total number of correct images/total number of images) of the five marmosets was 99.2%.

Next, face identification accuracy was examined in a family of three which was raised in the set-up home cage. The trained model was evaluated using data that was not used for training. For the evaluation, 5,600 facial images (1,400 for each class), three classes including each individual, and an unknown class were used. The unknown class was used to exclude inappropriate images for individual identification, such as unclear images, images of the back of the head and faces in profile, and images not showing faces. The evaluation metrics for deep learning for individual identification were as follows: specificity of 98.7%-99.7%, accuracy of 98.6%-99.2%, recall of 96.1%-98.8%, precision of 98.4%-99.1%, and an F-measure of 97.3%-98.3% (Table 2). The overall percentage of

correct identifications was 98.0%.

3D tracking

To optimize the laser measurement using lidar, a part of the cage made of metal mesh was replaced with acrylic panels, after which good measurement results were obtained (Fig. 1a). The laser beam distance needed to be sufficiently narrow and cover the entire cage, thus, the lidars were placed 1000 mm from the front of the cage (Fig. 1b). All 3D tracking and behavioral tests were conducted using a single family consisting of three marmosets (father, mother, and juvenile) in a cage.

The 3D coordinate information of marmosets was obtained using video tracking and lidar (Fig. 2a-f, and Mov. S1). The center of gravity points were assigned sequential tracking ID numbers during background processing, and each time the trajectory was broken, a new number was assigned. The pixel positions of the marmosets were then detected using an object detection algorithm in the video frame (Fig. 2c). The same process was applied to the four camera frames to determine the pixel positions of the marmosets for each camera. The 3D positions of each marmoset were calculated using pre-calibrated camera positions and the pixel coordinates of the marmosets (Fig. 2d). The heads of the marmosets extracted by the object detection algorithm were identified using the trained VGG19 model, and individual IDs were assigned to each marmoset by facial recognition (Fig. 2e, f). In cases where the marmosets' motion was too fast or was behind an object, the marmoset was not recognized by image processing, thus, lidar was used to interpolate the movement trajectory (Mov. S2, S3). Fig. 3

141 shows the 5 minutes of tracking trajectories without face identifications (Fig. 3a, b) and with face
142 recognition (Fig. 3c). When lidar alone or lidar and video tracking were performed, the tracking serial
143 numbers changed frequently, and it was impossible to continuously track the individual (Fig. 3a, b).
144 However, by applying face recognition to lidar and video tracking, the trajectory of each animal was
145 detected, and marmoset IDs were assigned (Fig. 3c).

146 To visually check the accuracy of tracking multiple marmosets, marmosets wore a color belly
147 band to identify each animal, and individual tracking was conducted for 60 min. During this
148 experiment, the marmoset IDs were switched nine times for more than 5 s, and the maximum time
149 length for accurate individual tracking was 19-minutes. Seven of the nine switching events were
150 resolved within 15 s. The video Mov. S4 shows a part of this experiment, and confirms that the system
151 could track the marmosets accurately even when they crossed. Although the marmoset cage had a
152 swinging hammock as enrichment (Mov. S4), these movements did not affect the tracking accuracy
153 of the marmosets when using video tracking and lidar. The location preference of each marmoset was
154 analyzed from the location information for an hour, and these results were compared with visual
155 observations (Fig. 4). The male (father) preferred the lower floor of the cage, while the female (mother)
156 and juvenile actively moved in a variety of areas in the cage (Table 3). For example, in the visual
157 observations of an hour, the juvenile was on the upper wooden bed 9.4% of the time, and on the middle
158 floor 63.8%, on the lower wooden bed 25.1%, on the lower floor 0%, and 1.7% of the time in the other

159 areas. In contrast, the system-analyzed data showed that the upper wooden bed was 11.7%, the middle
160 floor was 58.6%, the lower wooden bed was 24.1%, the lower floor was 0.4%, and the other areas was
161 5.2%. A strong significant correlation was found between the visual observation and system data of
162 this family ($R^2 = 0.96$, $P < 0.01$; Supplemental Fig. 1).

163 Furthermore, using the location information, the distances between the juvenile and the father
164 or mother were analyzed using position data (Supplemental Fig. 2). As a result, it appeared that mother
165 and juvenile were more closely bonded than father and juvenile were.

166 **Grooming behavior detection**

167 Grooming behavior, a representative social behavior of marmosets, was detected using video data
168 from the two upper cameras in front of the cage, excluding self-grooming behavior (Fig. 1b). The
169 grooming behaviors were determined from the video data of skilled animal technicians to obtain the
170 training data. The model was trained by annotating the grooming positions in each video frame. The
171 accuracy of grooming behavior detection by the machine was compared to that of visual observation
172 for one hour (Fig. 5). Twenty-four grooming units were visually detected and 21 units were detected
173 by the machine. Among these, 19 units were detected by visual observation and machine (recall:
174 79.1%). In the system, 19 of 21 units were confirmed as actual grooming behaviors (precision: 90.5%).
175 The five missed units had difficulty detecting behavior because of the placement of animals behind an
176 object: four had marmosets grooming behind the feeder, and one had grooming in the blind spot of the

bed. On the other hand, the two units were grooming behaviors detected by the machine but not by direct observation. At that time, the marmoset nuzzled its face into the other's tail or placed its hand on the other's shoulder. These actions seemed to be part of grooming behavior; however, observers did not judge grooming behaviors because they were not typical grooming actions.

Discussion

This is the first study of simultaneous 3D tracking of multiple unmarked marmosets using a novel approach achieved by lidar, video tracking, and facial recognition through deep learning. Furthermore, it is the only study of animal detection using lidar in laboratory animal applications. The detection speed and resolution of lidar are suitable for capturing quickly moving animals and is suitable for tracking small primates such as marmosets. In this study, the front and back of the home cage were replaced with acrylic panels, and lidar systems were placed 1000 mm from the front of the cage to cover the entire cage (Fig. 1a, b). This setup made it possible to obtain good lidar measurements. Three marmosets were kept in this home cage for more than six months, and they were healthy throughout the experimental period.

While lidar is suitable for detecting moving objects, it also falsely recognizes objects similar in size to a marmoset, such as the marmoset tail or feeder box. In contrast, video tracks do not have this issue because they use color image data to identify objects. Therefore, we used the information obtained from video tracking as a base for marmoset positions and complemented the

missing information with lidar to solve these false recognition issues. For example, video tracking lost detection when the marmoset motion was quicker than the shutter speed or when the marmosets were behind obstacles (see Mov. S2,3). In such cases, lidar is more effective than video-tracking. In addition, because both lidar and video tracking can be analyzed at high speed, without a time lag, no accumulation of data for analysis occurs, which is suitable for long-term analysis.

A similar system, the Kinect system, has been reported for animals similar in size to marmosets, such as rats[30] and pigeons[31], but the resolution of the Kinect system is not sufficient nor suitable for tracking fast-moving multiple animals[21]. When the Kinect system is used on multiple animals, a supplementary RFID is necessary[32]. Another animal behavioral analysis system, DeepLabCut, is a very powerful tool for estimating the posture of an animal using only a camera, but its accuracy depends on the camera's shutter speed[29, 33]. Although the images were captured at 10 Hz in this study, the subjects were blurred because the marmosets were very active and moved quickly. When using lidar, we were able to perform accurate tracking even during the jumping of the object, regardless of the camera performance.

Individual identification of marmosets is frequently performed using microchips, hair clipping, hair dyeing, and necklaces for marmosets[34]. In this study, facial images were used to achieve individual identification without artificial treatment because the face is the most characteristic part of marmoset bodies. The marmoset family used in this study had no blood relationships with other family

214 members because when making a new pair of marmosets, we selected unrelated marmosets to avoid
215 inbreeding to maintain genetic background diversity in the colony. Furthermore, the male was
216 vasectomized and the offspring (son) was obtained by embryo transfer from another pair of marmosets
217 to the mother's uterus in this family[15, 35]. The results of face identification showed that the most
218 accurate result for the family was that of the offspring (Table 2). The reason for the highest accuracy
219 in the offspring seemed to be that, although the offspring was beyond sexual maturity, young animals
220 could be easily identified because of their distinctive facial features. In addition, we found that using
221 five animals of the same age (1-2 years old), the animals could be identified; therefore, highly accurate
222 classification is possible by face recognition even when the animals are close in age (Table 1).

223 Incorporating facial recognition results into the trajectory made long-term tracking
224 possible, even for multiple animals in the same cage (Fig. 3). We examined the accuracy of marmoset
225 tracking using this system by color labeling the marmosets and comparing this to visual tracking (Mov.
226 S4, Table 3). This experiment confirmed that it is possible to track animals by face recognition, without
227 labeling marmosets or relying on color information. In addition, these data were able to calculate place
228 preference from the 3D trajectory of each marmoset (Fig. 4, Table 3, Supplementally Fig. 1),
229 locomotion, a distance between individuals (Supplementally Fig. 2), and an interaction between
230 individuals can be analyzed from the trajectory. Although some switching of marmoset IDs was
231 observed in the hour-long experiment, there was good correlation in the analysis of where and how

long each marmoset spent in the cage when comparing visual observations and by the system (Supplement Fig. 1). Therefore, although slight marmoset ID switching was observed, it did not affect the experimental results, and our tracking system was sufficient to analyze marmoset behaviors automatically. In animal models of autism spectrum disorder and Alzheimer's disease, abnormalities in social interaction have been reported, and differences in distance and position relative to other mice have been observed[36, 37, 38].. Therefore, in genetically modified marmoset models of Alzheimer's disease and other diseases that develop over long-term periods, the distance between marmosets and their location of preference may serve as early markers for the onset of the disease.

Environmental enrichment is necessary for captive marmosets. Since this tracking system aimed to analyze marmoset behavior in the home cage of the family group, it was necessary to make specifications that considered enrichment of the environment. For environmental enrichment, a hammock and colored balls were placed in the cage. Although these enrichments moved freely in the cage, marmoset tracking was not affected by these enrichments. Neither lidar nor video tracking detected the hammock motion because it was clearly different in size and shape from marmosets. However, the balls differed from the marmosets only in shape, and the balls were detected by lidar but never detected by video tracks (data not shown). Although lidar is a useful method for supplementing 3D information and creating trajectories, it is difficult to operate independently in a marmoset home cage with a variety of moving objects.

Grooming behavior, a typical social behavior, was detected through deep learning using video tracks. Although grooming behavior is difficult to continue tracking because marmosets are often close together, when combined lidar, video tracking, and facial recognition was used, it was possible to resume tracking when marmosets broke up after grooming behaviors (Mov. S4). The results of the comparison of grooming behavior detection by visual and mechanical detection showed that grooming could be detected with high precision (90.5%) and recall (79.1%; Fig. 5). In some cases, the time just before or just after the grooming was detected using deep learning. Specific behavioral characteristic behaviors might be unique to grooming before and after grooming. However, in some cases, grooming behavior could not be detected by the machine because the groomers were behind the feeding box or in other places where the camera had a blind spot. It is therefore necessary to update the training dataset to detect the behavior behind obstacles. Similar methods may also be used in the future to detect a variety of behaviors.

In conclusion, the combination of lidar, video tracking, and face recognition enabled the successful simultaneous tracking of each marmoset in a family without artificial labeling. It is possible to measure not only the 3D coordinates of individuals, but also other useful parameters, such as movement trajectory, locomotion, location preference, and distance between individuals, as well as to quantify the behavior of our developed system. These measurements will enable us to automatically analyze the activities and behaviors of marmosets under free-moving conditions. Importantly, these

268 data will be useful to identify early parameters for various diseases and to evaluate the pathology of
269 disease model marmosets. Furthermore, this system would be able to accurately quantify the behavior
270 over the entire lives of primates to discover behavioral changes before and after the onset of diseases
271 at an extremely early stage. These behavioral changes might be directly extrapolated to humans and
272 contribute to a better understanding of the mechanisms of complex diseases, such as
273 neurodegenerative and psychiatric disorders.
274

275 **Materials and Methods**

276 **Hardware Setup**

277 The home cage was made of acrylic panels on the front and back, and all food dishes, water jugs, and
278 bedding inside the cage were fixed to the wire mesh on the side to prevent them from moving. This
279 cage was designed to conduct behavioral analysis in a marmoset's home cage (Fig. 1a). Four poles
280 were set up in front of the cage at 1000 mm, and four lidars (VLP-16, Velodyne Lidar, San Jose, CA,
281 USA) and four cameras (C920, Logitech, Lausanne, Switzerland) were fixed to the poles (Fig. 1b).
282 The lidar was controlled by motion measurement software, and the path traced and velocity of the
283 marmosets were recorded (Hitachi, Ltd, Tokyo, Japan.). The laser beam of the lidar was irradiated
284 radially and adjusted to cover the entire inside of the cage. The recording parameters of the cameras
285 were 1080p and 60 Hz, and all cameras were adjusted to show the entire cage (Supplementary Fig. 3).

The positions of the cage, lidar, and camera were calibrated when the cage was moved and the interiors were changed. Lidar calibration was used to acquire background information, including the cage interior without marmosets. Calibration of the camera was performed by detecting alco-markers at the four corners of the cage.

Ethical approval

This study was conducted at the animal facility of the Central Institute for Experimental Animals (CIEA), Kawasaki, Japan. The animal experimental protocol was reviewed by the Institutional Animal Care and Use Committee and approved (approval no. 17051) according to the Regulations for Animal Experiments in CIEA, based on the Guidelines for the Proper Conduct of Animal Experiments by the Science Council of Japan (2006).

Animal facilities

The marmosets were kept in a family cage with dimensions of 820 × 610 × 1578 mm. Although the marmosets were kept in only one family in the room, they were able to communicate with the marmosets in other rooms by voice. All cages were equipped with a bunk, and the animal rooms were adjusted to 26-28°C, 40-60% humidity with a 12:12 h light/dark cycle. The animals were negative for Salmonella spp., Shigella spp., and Yersinia pseudotuberculosis on annual fecal examination.

Animals

Each marmoset family was housed in a home cage. The family consists of a father (I5072M, ten years

old), mother (I5894F, six years old), and a child (I940 gmM, one year old). Artificial or non-artificial physical features for individual identification were not used, such as hair dye, necklaces, piercings, or other visible signs, except during 3D tracking of marmoset-attached color marker experiments has been performed. In conventional animal care, a microchip is implanted subcutaneously for individual identification. The child was born by embryo transfer to the mother's uterus[15], and there was no blood relationship within the third degree of the family.

Five healthy marmosets around 1-2-year-old, one male and four female, were used in the preliminary study to develop deep learning facial recognition. (I774; one years old, I6695; one years old, I6708; one years old, I6677; one years old, I893; two years old). There were no artificial or visible physical characteristics, such as hair dye, necklaces, or piercings, and each marmoset was not related to any other marmoset within the third degree of kinship.

Facial recognition

The cage (W820 × D610 × H1578 mm) was divided into four compartments for video recording of individual animals. Each animal was housed in a separate cage, and videos were recorded from the cage front for 3-6 hours. The front door of the cage was then replaced with a clear acrylic cage. From these video frames, the face images were cropped only in the face area for the training data images, and these data were augmented by rotation. Then using 1400 face images for each class for training,

frontal images of the face were classified into four or five classes using the VGG19 model network. In the preliminary test, the five classes comprised five marmosets. For the family group, the four classes included those members and an unknown class to exclude inappropriate images for classification, and 1000 images were used for training.

Model training was performed by fine-tuning a pre-trained VGG19 model (learning rate 0.001, fixed number of layers 12, BATCH_SIZE 32)[39]. VGG19 is a network model proposed by the Visual Geometry Group at the University of Oxford in 2014 that obtained accurate classification performance on the ImageNet dataset. The pre-trained models used in this study were trained using the ImageNet dataset.

3D tracking

By lidar and video tracking, 3D tracking was performed. A lidar measures the distance to an object by irradiating a laser and analyzing the reflected laser. Marmoset detection using lidar involves the following steps: background information removal, clustering, and selection of marmoset clusters[40]. Background removal was performed during calibration. The observed points were moving points belonging to objects whose spatial positions changed and static points from the background environment. Marmosets were then selected from among all the moving clusters. The Euclidean cluster extraction algorithm was used for clustering, with a minimum cluster size of S1 and maximum

cluster size of S2 (S1: 100 mm, S2: 500 mm)[41]. The centroid of the cluster is regarded as the position of the marmoset for tracking. Three-dimensional video tracking was performed using four time-synchronized video cameras, 3D video tracking was performed. The positions of the whole body and head of the marmoset were detected in each video frame using Yolo v3, an object detection algorithm, and the 3D position was calculated from each camera information[42]. The YOLO model was trained using 4611 images in which the whole body and head area were annotated (learning rate: 0.0003, momentum: 0.9, batch: 64). For video tracking, 3D coordinates were used as primary information and supplemented with lidar coordinates when the primary information was missing. When the marmoset coordinates were detected outside the cage, they were considered missing values.

Detecting grooming behavior

Yolov3, an open-source deep learning software, was used to detect grooming behavior. As teacher images, 252 annotated image files in which grooming marmosets were surrounded by a minimum rectangle were prepared, and the model was trained using the dataset. The trained model analyzed the top two camera view images and detected grooming behavior in each video frame. To improve the accuracy by removing noise, a Gaussian weighted moving average of 50-second window and 5-second slide was applied to the frequency of grooming occurrence at each time to create the probability distribution of grooming occurrence. The same process was performed for the left and right cameras and the calculated probability distributions were averaged. When the average value exceeded a certain

threshold, grooming was considered to have occurred. To test the accuracy of this detection system, skilled breeding staff visually detected marmoset grooming in one hour of video that was not used for training data. This visual result was used as the true label, and the accuracy and repeatability of grooming detection were calculated in 30-second increments of time.

Statistical analysis

Pearson's correlation coefficient was calculated using GraphPad Prism version 9 (GraphPad Software, La Jolla, CA, USA).

Acknowledgments

This research was partially supported by the Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS), "Study of developing neurodegenerative model marmosets and establishment of novel reproductive methodology (JP19dm0207065)" from the Japan Agency for Medical Research and Development (AMED) and Grant-in-Aid for Scientific Research (A) "Development of an Automated Behavioral Analysis System for Evaluation of Genetically Modified Disease Model Marmosets" JSPS KAKENHI Grant Numbers JP21H04756 to ES and Grant-in-Aid for Early-Career Scientists "Effects of transient alcohol exposure to the fetus in early pregnancy on brain development." JSPS KAKENHI Grant Number JP20K16908 to TY.

Author contributions

T.Y. and E.S. conceived and designed the study. T.Y., N.G., J.K., and T.F. performed the setup of the measurement system. R.K., R.H., M.K., Y.H., and E.Y. performed the annotation of the training data. T.Y., Y.S., R.H, J.K. and T.F. processed data, trained the models, and analyzed results. T.Y. and R.K. performed the experiment. T.Y., W.K., K.S., N.G., T.I., and E.S. discussed and interpreted the results. T.Y., T.I. and E.S. wrote the manuscript. E.S. supervised the project. All authors commented on the manuscript.

Competing interests

N.G. is an employee of Hitachi, Ltd. J.K. is an employee of Hitachi Solutions Technology, Ltd. T.F. is an employee of Totec Amenity Limited. The remaining authors declare no competing interests.

Figure 1. Installation of hardware for the behavior analysis system.

(a) A home cage with acrylic panels on the front and back and metal mesh on the sides (b). Schematic diagram of light and ranging (lidar) systems and cameras installed in front of the cage. Lidars and cameras were installed at 1000 mm in front of the cage.

Figure 2. 3D tracking of marmosets using lidar and video tracking.

(a) Raw video frame. (b) Detected 3D coordinate of marmosets by lidar. White point: centroid of marmosets; white box: virtual space of the cage. (c) Detected location of marmosets in video frame by Yolo. Rectangle: detected marmosets. (d) Calculated 3D coordinate of marmosets by video tracking, colored point: coordinate of marmoset. (e, f) Individual information added to c and d using face identification, green: I5072M, red: I5894F, yellow: I940gmM.

Figure 3. 3D position trajectory of a marmoset in five minutes.

Marmoset trajectories analyzed by lidar (a) and lidar and video tracking (b). Different colors were assigned to each tracking serial number that numbered the background. (c) Lidar and video trajectory

tracking combined with face recognition. Green: father; red: mother; yellow: child.

Figure 4. Place preference of marmosets for 1 hour.

Color map of the time spent by marmosets in the cage for one hour. Locations where marmosets spent more than 1% of one hour are shown in red, and others are shown as color bars. (a) Father (I5072M), (b) mother (I5894F), (c) child (I940gmM).

Figure 5. Comparison of one hour of human visual and machine grooming detection.

Each cell represented 30 s. If a human detected grooming at least once in 30 s, it was painted blue. The system detected grooming and the cell was painted red when the intensity exceeded a threshold value.

References

1. Gauthier S, *et al.* Management of behavioral problems in Alzheimer's disease. *Int Psychogeriatr* **22**, 346-372 (2010).

2. Anderson KE. Behavioral disturbances in Parkinson's disease. *Dialogues Clin Neurosci* **6**, 323-332 (2004).

3. Finkel SI, Costa e Silva J, Cohen G, Miller S, Sartorius N. Behavioral and psychological signs and symptoms of dementia: a consensus statement on current knowledge and implications for research and treatment. *Int Psychogeriatr* **8 Suppl 3**, 497-500 (1996).

4. Levy SE, Mandell DS, Schultz RT. Autism. *Lancet* **374**, 1627-1638 (2009).

5. Webster SJ, Bachstetter AD, Nelson PT, Schmitt FA, Van Eldik LJ. Using mice to model Alzheimer's dementia: an overview of the clinical disease and the preclinical behavioral changes in 10 mouse models. *Front Genet* **5**, 88 (2014).

6. Cryan JF, Mombereau C. In search of a depressed mouse: utility of models for studying depression-related behavior in genetically modified mice. *Mol Psychiatry* **9**, 326-357 (2004).

7. Roelfsema PR, Treue S. Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. *Neuron* **82**, 1200-1204 (2014).

8. Yamazaki Y, Echigo C, Saiki M, Inada M, Watanabe S, Iriki A. Tool-use learning by common marmosets (Callithrix jacchus). *Exp Brain Res* **213**, 63-71 (2011).

9. Miller CT, Freiwald WA, Leopold DA, Mitchell JF, Silva AC, Wang X. Marmosets: A Neuroscientific Model of Human Social Behavior. *Neuron* **90**, 219-233 (2016).

10. Mitchell JF, Leopold DA. The marmoset monkey as a model for visual neuroscience. *Neurosci Res* **93**, 20-46 (2015).

11. Burkart JM, Fehr E, Efferson C, van Schaik CP. Other-regarding preferences in a non-human

primate: common marmosets provision food altruistically. *Proc Natl Acad Sci U S A* **104**, 19762-19766 (2007).

12. Abbott DH, Barnett DK, Colman RJ, Yamamoto ME, Schultz-Darken NJ. Aspects of common marmoset basic biology and life history important for biomedical research. *Comp Med* **53**, 339-350 (2003).

13. Sasaki E*, et al.* Generation of transgenic non-human primates with germline transmission. *Nature* **459**, 523-527 (2009).

14. Sato K*, et al.* Generation of a Nonhuman Primate Model of Severe Combined Immunodeficiency Using Highly Efficient Genome Editing. *Cell Stem Cell* **19**, 127-138 (2016).

15. Sato K*, et al.* A non-human primate model of familial Alzheimer's disease. *bioRxiv*, 2020.2008.2024.264259 (2020).

16. Kishi N, Sato K, Sasaki E, Okano H. Common marmoset as a new model animal for neuroscience research and genome editing technology. *Dev Growth Differ* **56**, 53-62 (2014).

17. Okano H, Miyawaki A, Kasai K. Brain/MINDS: brain-mapping project in Japan. *Philos Trans R Soc Lond B Biol Sci* **370**, (2015).

18. Network BICC. A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86-102 (2021).

19. Workman KP, Healey B, Carlotto A, Lacreuse A. One-year change in cognitive flexibility and fine motor function in middle-aged male and female marmosets (Callithrix jacchus). *Am J Primatol* **81**, e22924 (2019).

20. Yamazaki Y, Saiki M, Inada M, Watanabe S, Iriki A. Sustained performance by common marmosets in a delayed matching to position task with variable stimulus presentations. *Behav Brain Res* **297**, 277-284 (2016).

21. Hong W*, et al.* Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. *Proc Natl Acad Sci U S A* **112**, E5351-5360 (2015).

22. Redfern WS, *et al.* Automated recording of home cage activity and temperature of individual rats housed in social groups: The Rodent Big Brother project. *PLoS One* **12**, e0181068 (2017).

23. Sutoko S, *et al.* Early Identification of Alzheimer's Disease in Mouse Models: Application of Deep Neural Network Algorithm to Cognitive Behavioral Parameters. *iScience* **24**, 102198 (2021).

24. Mann TM, Williams KE, Pearce PC, Scott EA. A novel method for activity monitoring in small non-human primates. *Lab Anim* **39**, 169-177 (2005).

25. Yabumoto T, *et al.* MarmoDetector: A novel 3D automated system for the quantitative assessment of marmoset behavior. *J Neurosci Methods* **322**, 23-33 (2019).

26. Turesson HK, Conceição TBR, Ribeiro S. Head and gaze tracking of unrestrained marmosets. *bioRxiv*, 079566 (2016).

27. Ando K, *et al.* Differential effects of dopaminergic drugs on spontaneous motor activity in the common marmoset following pretreatment with a bilateral brain infusion of 6-hydroxydopamine. *Behav Pharmacol* **28**, 670-680 (2017).

28. Schultz-Darken N, Ace L, Ash H. Chapter 7 - Behavior and Behavioral Management. In: *The Common Marmoset in Captivity and Biomedical Research* (eds Marini R, Wachtman L, Tardif S, Mansfield K, Fox J). Academic Press (2019).

29. Lauer J, *et al.* Multi-animal pose estimation and tracking with DeepLabCut. *bioRxiv*, 2021.2004.2030.442096 (2021).

30. Matsumoto J, *et al.* A 3D-video-based computerized analysis of social and sexual interactions in rats. *PLoS One* **8**, e78460 (2013).

31. Lyons DM, MacDonall JS, Cunningham KM. A Kinect-based system for automatic recording of some pigeon behaviors. *Behav Res Methods* **47**, 1044-1054 (2015).

32. Chaumont Fd, *et al.* Live Mouse Tracker: real-time behavioral analysis of groups of mice. *bioRxiv*, 345132 (2018).

33. Mathis A*, et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21**, 1281-1289 (2018).

34. Susanne R, Ann-Kathrin O. Husbandry and Management of New World Species: Marmosets and Tamarins. *The Laboratory Primate*, 145-162 (2005).

35. Kurotaki Y, Sasaki E. Practical Reproductive Techniques for the Common Marmoset. *Journal of Mammalian Ova Research* **34**, 3-12 (2017).

36. Yasue M, Nakagami A, Nakagaki K, Ichinohe N, Kawai N. Inequity aversion is observed in common marmosets but not in marmoset models of autism induced by prenatal exposure to valproic acid. *Behav Brain Res* **343**, 36-40 (2018).

37. Wang L, Cai Y, Fan X. Metformin Administration During Early Postnatal Life Rescues Autistic-Like Behaviors in the BTBR T+ Itpr3tf/J Mouse Model of Autism. *Front Behav Neurosci* **12**, 290 (2018).

38. Filali M, Lalonde R, Rivest S. Anomalies in social behaviors and exploratory activities in an APPswe/PS1 mouse model of Alzheimer's disease. *Physiol Behav* **104**, 880-885 (2011).

39. Zhou J, Yang X, Zhang L, Shao S, Bian G, Shen C. Multisignal VGG19 Network with Transposed Convolution for Rotating Machinery Fault Diagnosis Based on Deep Transfer Learning. *Shock and Vibration* **2020**, 1-12 (2020).

40. Zhang J, Xiao W, Coifman B, Mills JP. Vehicle Tracking and Speed Estimation From Roadside Lidar. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **13**, 5597-5608 (2020).

41. Rusu RB. Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. *KI - Künstliche Intelligenz* **24**, 345-348 (2010).

42. Redmon J, Farhadi A. Yolov3: An incremental improvement. *arXiv preprint arXiv:180402767*, (2018).

Fig. 1

a

b

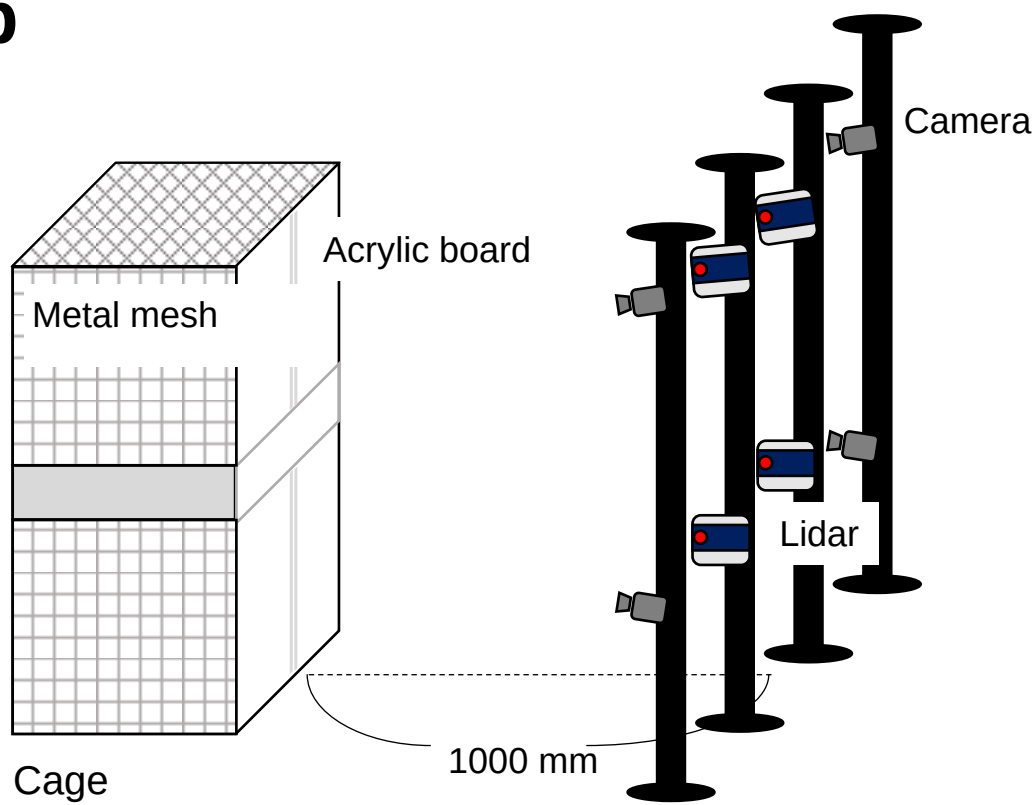


Fig. 2

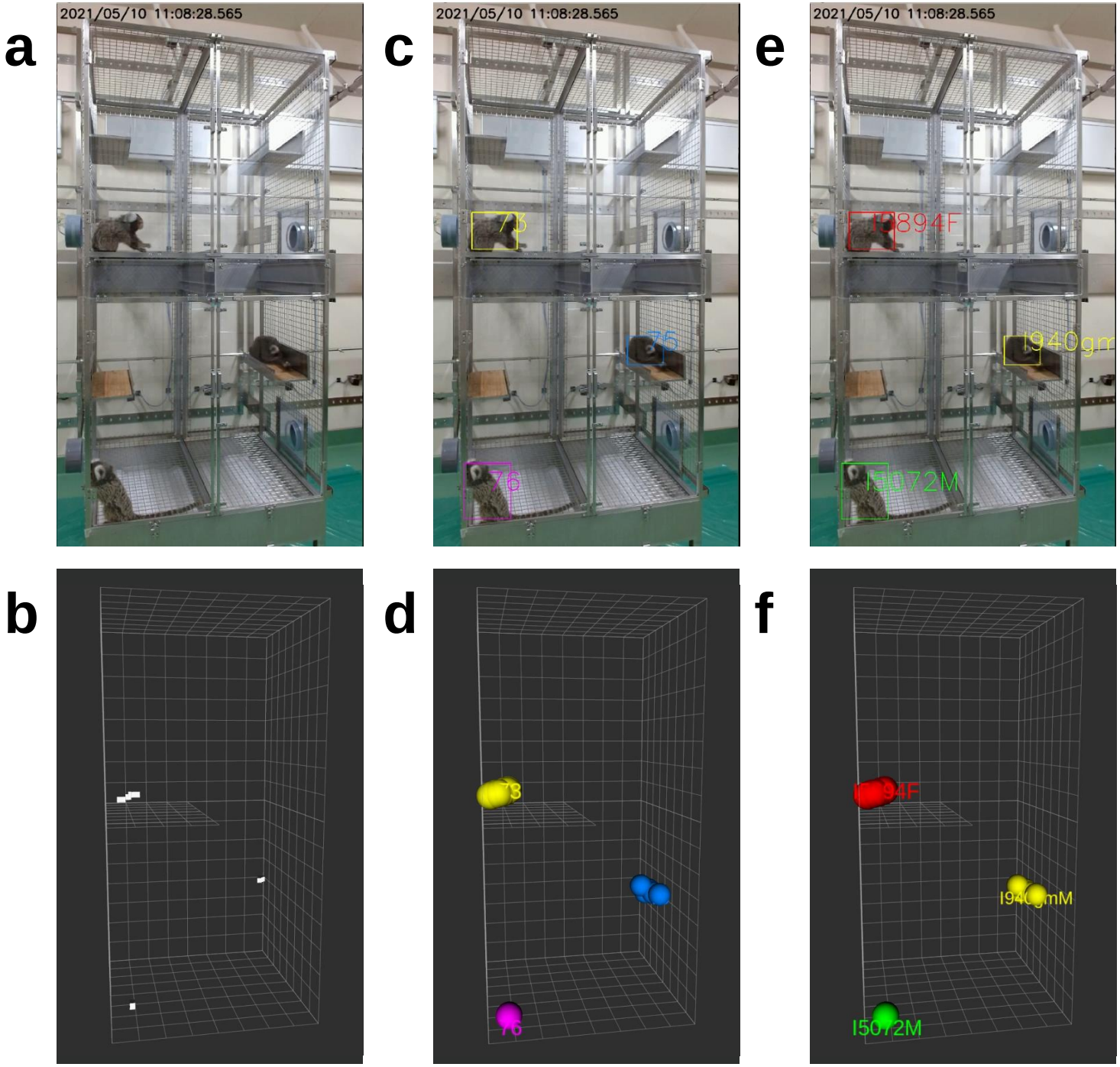
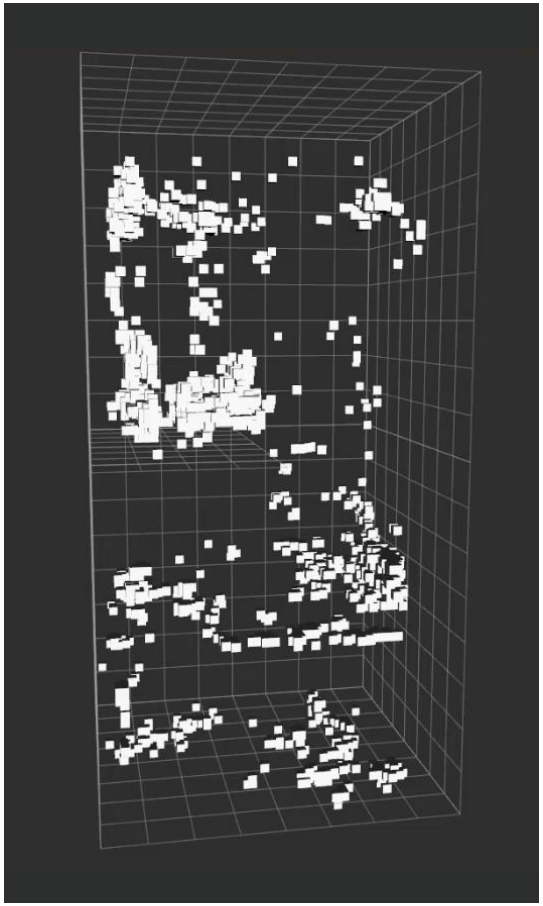
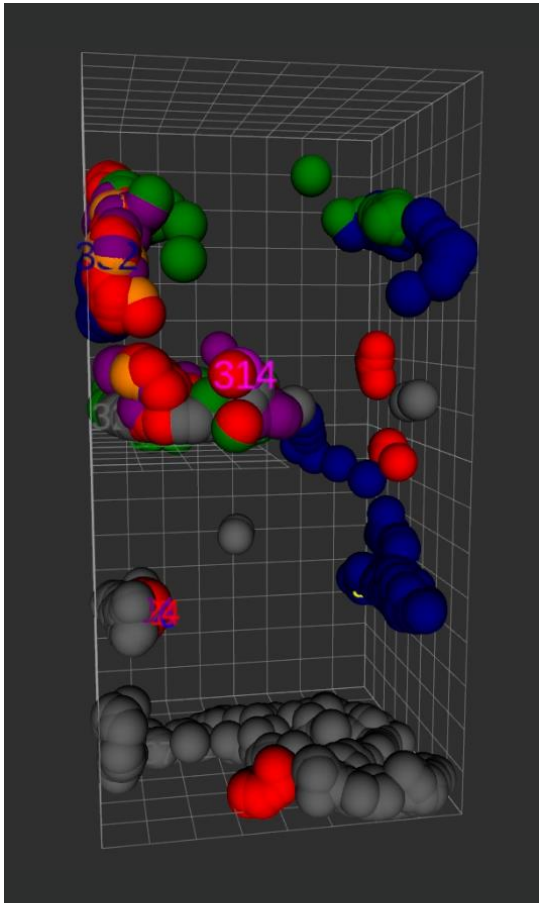

a
b
c
d
e
f
2021/05/10 11:08:28.565
2021/05/10 11:08:28.565
2021/05/10 11:08:28.565
73
76
76
I5072M

Fig. 3

a



b



c

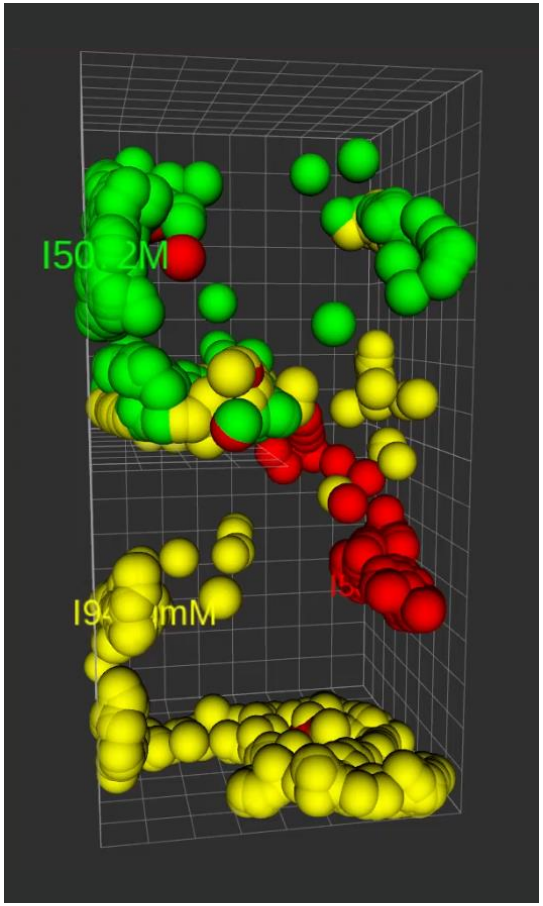
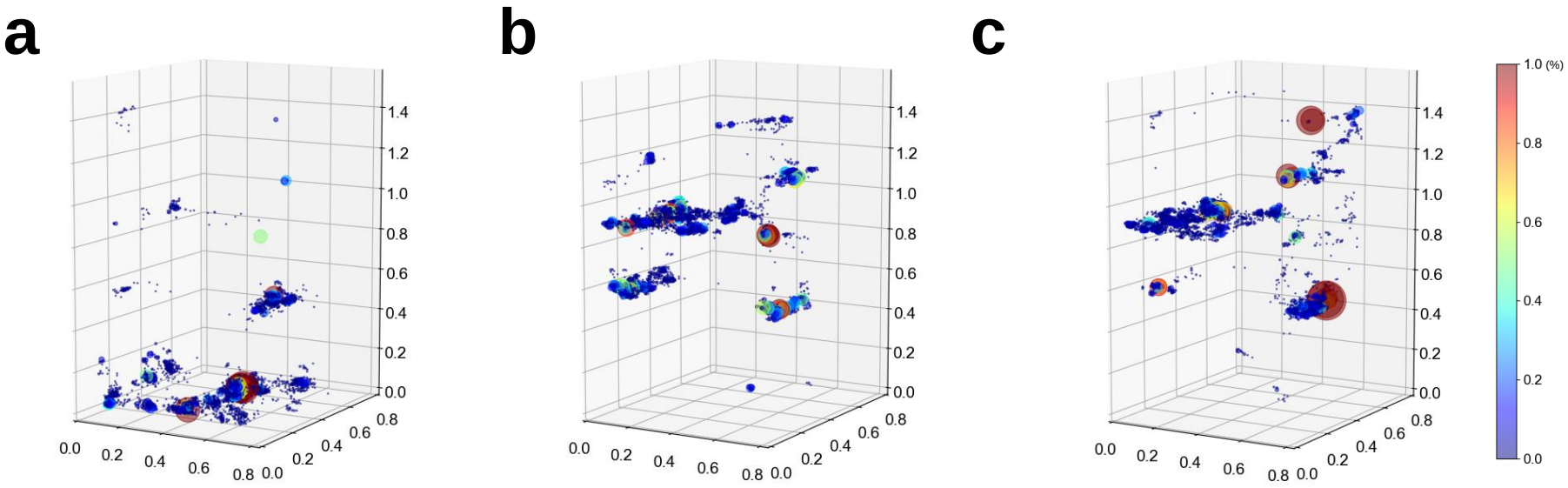


Fig. 4



Father (I5072M) Mother (I5894F) Child (I940gmM)

Fig. 5

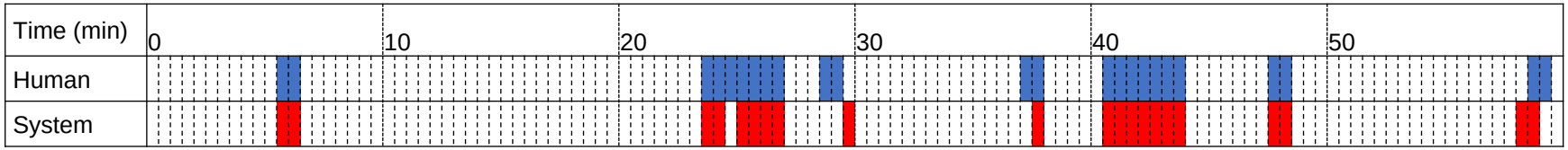


Table 1. Classification accuracy of face identification in five marmosets aged 1 to 2 years.

	Animal ID				
	I774	I6695	I6708	I6677	I893
Specificity	100.0%	99.8%	99.8%	99.8%	99.8%
Accuracy	99.8%	99.6%	99.4%	99.8%	99.8%
Recall	99.0%	99.0%	98.0%	100.0%	100.0%
Precision	100.0%	99.0%	99.0%	99.0%	99.0%
F-Measure	99.5%	99.0%	98.5%	99.5%	99.5%

Table 2. Classification accuracy of face identification in a marmoset family.

	Father (I5072M)	Mother (I5894F)	Child (I940gmM)	Unknown
Specificity	98.7%	99.5%	99.7%	99.5%
Accuracy	98.7%	98.6%	99.2%	99.5%
Recall	98.8%	96.1%	97.6%	99.6%
Precision	96.1%	98.4%	99.1%	98.5%
F-Measure	97.4%	97.3%	98.3%	99.0%

Table 3. Comparison of visual and the system analysis of time spent in some specific locations in the cage for an hour

	Father (I5072M)		Mother (I5894F)		Child (I940gmM)	
	Human	System	Human	System	Human	System
Upper bed	0.4%	0.1%	11.5%	12.1%	9.4%	11.7%
Middle floor	0.3%	2.2%	35.9%	49.1%	63.8%	58.6%
Lower bed	8.8%	15.1%	42.7%	34.8%	25.1%	24.1%
Lower floor	90.3%	80.2%	0.0%	0.4%	0.0%	0.4%
Other	0.2%	2.4%	9.9%	3.6%	1.7%	5.2%