



An infrared range camera-based approach for three-dimensional locomotion tracking and pose reconstruction in a rodent

Tai-Hsien Ou-Yang^a, Meng-Li Tsai^b, Chen-Tung Yen^c, Ta-Te Lin^{a,*}

^a Department of Bio-Industrial Mechatronics Engineering, National Taiwan University, 1 Roosevelt Road, Section 4, Taipei, Taiwan

^b Department of Biomechatronic Engineering, National Ilan University, Ilan, Taiwan

^c Department of Life Science and Institute of Zoology, National Taiwan University, Taipei, Taiwan

ARTICLE INFO

Article history:

Received 17 February 2011

Received in revised form 20 July 2011

Accepted 21 July 2011

Keywords:

Locomotion

Depth image

Behavior

Pose reconstruction

ABSTRACT

We herein introduce an automated three-dimensional (3D) locomotion tracking and pose reconstruction system for rodents with superior robustness, rapidity, reliability, resolution, simplicity, and cost. An off-the-shelf composite infrared (IR) range camera was adopted to grab high-resolution depth images ($640 \times 480 \times 2048$ pixels at 20 Hz) in our system for automated behavior analysis. For the inherent 3D structure of the depth images, we developed a compact algorithm to reconstruct the locomotion and body behavior with superior temporal and solid spatial resolution. Since the range camera operates in the IR spectrum, interference from the visible light spectrum did not affect the tracking performance. The accuracy of our system was $98.1 \pm 3.2\%$. We also validated the system, which yielded strong correlation with automated and manual tracking. Meanwhile, the system replicates a detailed dynamic rat model in virtual space, which demonstrates the movements of the extremities of the body and locomotion in detail on varied terrain.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Locomotion analyses are widely applied in neuroscience research, animal welfare, and pharmacological studies on different animal species (Tecott and Nestler, 2004), including rodents (Denenberg, 1969; Clarke et al., 1985; Pan et al., 1996; Rousseau et al., 2000; Twining et al., 2001; Spruijt and deVisser, 2006), pigs (Lind et al., 2005), insects (Noldus et al., 2001; Reynolds and Riley, 2002), arthropods (Aagaard et al., 1991), and fish (Nilsson et al., 1993; Kato et al., 1996).

Manually recording animal behavior is economical but inherently suffers from subjectivity and fatigue (Lind et al., 2005; Tort et al., 2006). Otherwise, automated locomotion recording and analysis have obvious advantages including repeatability and reliability with lower labor costs (Tecott and Nestler, 2004; Spruijt and deVisser, 2006). Early automated measuring systems were usually based on photo-cell sensors (Clarke et al., 1985; Ericson et al., 1991) or electronic touch panels (Kao et al., 1995) which are relatively more reliable than manual recording but are expensive and have low resolution (Pan et al., 1996). Since the 1980s, computer vision-based systems were introduced into the field of animal locomotion analysis (Godden and Graham, 1983), and they provided greater accuracy and flexibility than other systems (Noldus et al., 2001).

Diverse methods were presented to track animal behavior based on computer-vision, including monochrome video thresholding (Godden and Graham, 1983; Nilsson et al., 1993; Hoy et al., 1996; Kato et al., 1996; Noldus et al., 2001; Lind et al., 2005; Tort et al., 2006), color space conversion and filtering (Noldus et al., 2001; Obradovic et al., 2009), background image subtraction (Kato et al., 1996; Noldus et al., 2001; Publicover et al., 2009), and model recognition (Rousseau et al., 2000; Twining et al., 2001). Some of systems are commercially available (Spruijt and deVisser, 2006). Some efforts on three-dimensional (3D) tracking using multiple cameras (Kato et al., 1996; Fry et al., 2000) or video cameras with mirrors were presented (Chràsková et al., 1999). Three-dimensional human behavior reconstruction has been widespread and widely studied based on various approaches (Moeslund and Granum, 2001). Correspondingly, all sorts of three-dimensional rodent behavioral phenotyping systems were presented, HomeCageScan® (CleverSys, Reston, VA, USA) provides rodent behavior phenotyping modules based on body part feature recognition to analyze extensive rodent's behaviors, which recognize rodent behavior with video clips recorded from video cameras from limited top and side view-points. Otherwise, LABORAS® (Metris, Hoofddorp, the Netherlands) reconstructs rodent three-dimensional behavior with physical vibration and force signal with specific sensor instruments, which would not be affected by light interference. But the flexibility of measurement is rather constrained and sensitive to physical noise (Spruijt and deVisser, 2006). Various qualitative validity assessment procedures were also proposed (Van de Weerd et al., 2001;

* Corresponding author. Tel.: +886 2 33665331; fax: +886 2 23929416.
E-mail address: m456@ntu.edu.tw (T.-T. Lin).

Lind et al., 2005; Tort et al., 2006). However, applications of vision-based systems were often limited by color-related factors including contrast, brightness, animal color, complicated backgrounds, and poor tolerance to interference. Some studies further used expensive infrared (IR) cameras (Noldus et al., 2001; Lind et al., 2005; Obradovic et al., 2009).

Microsoft Kinect® (Microsoft Corporation, Redmond, WA, USA) is a low-cost (<US\$ 150) commercial composite sensor device with an IR range camera, charge-coupled device (CCD) camera, microphone array, and accelerometer, and was designed for human gesture and pose recognition for the consumer market. The IR range camera and video camera composite module comes with PrimeSensor® technology (PrimeSense, Tel Aviv, Israel), which integrates IR depth images and color video acquisition in a system-on-a-chip design for high-resolution depth information registration at greater than a 30-Hz theoretical sampling rate. In this study, we developed an automatic locomotion analysis system for rats using a depth map directly acquired from the IR range camera of Kinect, which combines advantages of an IR camera and a laser range finder at a reasonable cost. Therefore, the IR range camera is capable of retrieving video in low-brightness conditions, and interprets distance (depth) information of each point in the scene as corresponding pixel values.

By interpreting depth information in the captured frames, our method is able to reconstruct continuous high-resolution 3D location and measurements of a rat, without limitations of indoor ambient light conditions, contrast, invasive labeling tags, or varying colors of the rat and background. For independence of the analysis, our system was divided into two parts: a recording part and an analysis part. The depth image recording part retrieves and compresses depth and color image streams from an IR range camera and CCD camera, respectively. The recorded depth images are device independent; any similar device which produces depth images can be used as the source for the depth image stream. The analysis part performs off-line 3D measurement reconstruction and locomotion analysis for further assessment and post hoc tests.

To validate our method, we recorded range camera frames and CCD camera frames configured horizontally. By directly accessing memory instead of using built-in functions in a computer vision library, our system achieved a sampling rate of >20 Hz with a resolution of 640 × 480 pixels.

2. Materials and methods

2.1. Animals

To validate our system, 21 animals in total were used, including 5 Long Evans (378.0 g, 286.0 g, 242.5 g, 241.5 g, 241.0 g), 5 Wistar rats (386.0 g, 395.0 g, 398.0 g, 400.0 g, 417.5 g) and 5 C57BL/6 mice (34.5 g, 31.0 g, 30.5 g, 29.0 g, 27.5 g) for assessment of the effects on the system performance caused by varied fur colors and sizes of the animals. The other 2 Long Evans (285.0 g and 286.0 g), 2 Wistar rats (278.0 g and 300.0 g) and 2 C57BL/6 mice (22.5 g and 23.0 g) for lightless environment performance examination. All experimental procedures were approved by the Institutional Animal Care and Use Committee of National Ilan University, and were in accordance with guidelines specified in the “Codes for Experimental Use of Animals” of the Council of Agriculture of Taiwan, based on the Animal Protection Law of Taiwan.

2.2. Hardware

The hardware setup included a behavior apparatus, Microsoft Kinect® composite sensor device, and a personal computer. The behavior apparatus consisted of aluminum

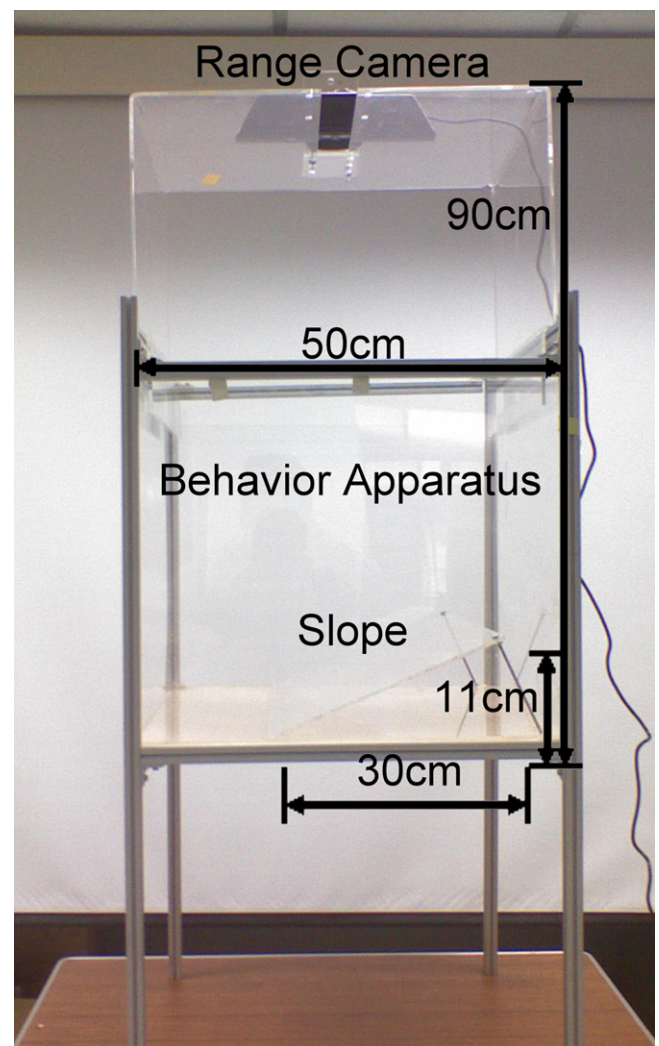


Fig. 1. Photograph of the experimental configuration including the behavior apparatus, transparent slope, and Microsoft Kinect® composite sensor. The apparatus with the transparent slope consisted of aluminum extrusions and acrylic plates for convenience of adjustment. The rat roamed freely on the floor. The Kinect® sensor was mounted 90 cm above the floor for a comprehensive view of the apparatus.

extrusions (120 cm × 52 cm × 52 cm), four acrylic side plates (50 cm × 50 cm × 0.2 cm), and an acrylic floor plate (50 cm × 50 cm × 0.5 cm) with a grid pattern consisting of 5-cm × 5-cm squares for performance validation. The sensor device was mounted 90 cm above the bottom of the apparatus to record stereo dorsal images of the rat. To assess the 3D tracking performance in a complex topographical configuration, an acrylic slope (50 cm × 30 cm × 11 cm) was set up in the behavior apparatus (Fig. 1).

The Microsoft Kinect® composite sensor device was connected to a personal computer for 3D rat locomotion and pose reconstruction (Fig. 2A). The Kinect® device acquires depth images by continuously projecting coded IR light and decodes the light reflected which contains depth information of the scene, which provides a Video Graphics Array (VGA) depth image with 3-mm spatial and 1-cm depth resolution at a rate of 30 frames/s (30 Hz) under indoor lighting conditions. Considering stability, the sampling rate is set at a rate of 20 frames/s (20 Hz) while recording the series of depth image. Respectively, the recorded depth image series were processed by the method proposed in this paper for locomotion analysis at about 20 Hz rate. The driver for the sensor is a Code Laboratories® NUI Platform (Code Laboratories, Henderson,

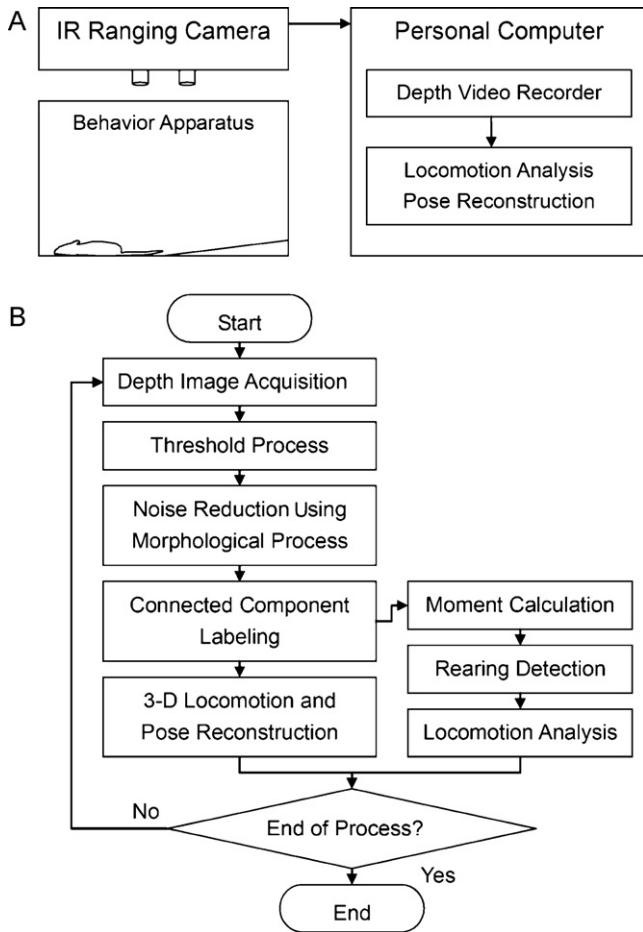


Fig. 2. Hardware and software diagrams of our approach. (A) This diagram shows the hardware setup. The range camera above the behavior apparatus was connected to a personal computer. The depth image stream was recorded for off-line locomotion measurements and pose reconstruction. (B) This diagram is a flowchart of the software process of our system. The acquired frame was processed by threshold, morphological processes for noise reduction, and connected component labeling. Then the detected rat's behavior was reconstructed in a solid space and calculated for the locomotion analysis.

NV, USA). Both the PrimeSense® and Code Laboratories® provide open-source or free standard development kits and drivers for depth image acquisition.

The personal computer was equipped with an Intel i5-760® 2.8-GHz processor (Intel, Santa Clara, CA, USA), 4.0-GB RAM, and an NVIDIA GeForce GT320® graphic card (NVIDIA, Santa Clara, CA, USA).

2.3. Software

We developed a collection of algorithms to analyze the acquired depth frames for non-invasive 3D rat locomotion reconstruction without constraints of labeling, fur color, background color, or ambient light conditions. A flow chart describing our approach is given in Fig. 2B.

Our software was developed with Microsoft Visual Studio 2010 Express® on Microsoft Windows 7®. We applied an OpenCV computer vision library (Willow Garage, Menlo Park, CA, USA) for image processing and an OpenGL graphic library (Khronos Group, Beaverton, OR, USA). OpenCV is an open-source library which provides functions of image processing, object segmentation, recognition, and tracking. OpenGL is a graphic library providing 3D graphic rendering application interfaces for solid rat pose reconstruction. The depth and color video were simultaneously recorded and com-

pressed using a Microsoft Video 1® algorithm. The recorded color and depth frames were analyzed post hoc with a resolution of 640 × 480 pixels.

2.4. Depth image recording

Depth image frames and color video frames were both retrieved from the Kinect® sensor device, and recorded at a synchronized step in VGA resolution for validity assessment. The performance of depth image retrieval was not affected by low-brightness conditions, when the color video turned black. To achieve a sampling rate above 20 Hz for electrophysiological studies while simultaneously recording two VGA video streams, our system maps the retrieved frames to the memory to optimize the recording performance instead of using functions provided by standard libraries.

2.5. Locating and tracking

The acquired raw depth frame was trimmed according to the 3D field of view assigned manually to locate the rat in an assigned region. Since the apparatus' floor occupies only 57.6% in the center of camera's field of view, the camera's distortion could be ignored.

The rat identification algorithm using a threshold is described in (1) below; each pixel $p(i,j)$ containing depth information, z_{ij} , was preserved if z_{ij} was above the estimated floor, τ , otherwise, z_{ij} was replaced by 0. The floor position could automatically be estimated, but for validity assessment, it was user-defined during the experiments. In the present experimental configuration, the depth values from the range camera ranged 70–90 cm.

$$p(i,j) = \begin{cases} z_{ij} & \text{for } z_{ij} > \tau \\ 0 & \text{for } z_{ij} \leq \tau \end{cases} \quad (1)$$

After applying morphological processes, including erosion and dilation, to reduce the noise, the connected component labeling algorithm was applied to identify the rat in the sectioned depth image. The connected component labeling algorithm can be considered a process of labeling all connected subsets, C_i , in a binary image set, I . A pixel, $p(x,y)$, is given a label value, i , if a labeled neighboring pixel belonging to C_i exists. If no neighboring labeled pixels exist, it is assigned a new label, C_{i+1} . The OpenCV library provides a state-of-the-art connected component labeling algorithm packaged as the function, `cvFindContours()`, which was adopted in our system to identify rats in the field of view.

The mass center of the identified rat was calculated using (2)–(4), which produce a 0th moment, and two 1st moments of the rat, and $p(x,y)$ contains pixels belonging to the identified rat.

$$m_{00} = \sum_x \sum_y p(x,y) \quad (2)$$

$$m_{10} = \sum_x \sum_y xp(x,y) \quad (3)$$

$$m_{01} = \sum_x \sum_y yp(x,y) \quad (4)$$

The mass center of the rat, p_{MC} , is presented by (5):

$$p_{MC} = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (5)$$

The labeled connected component subset is stored as sequences of chain codes representing identified rats. The size, shape, and mass center of the rat in each frame are preserved in data structures for further locomotion analysis. The vertical position of the rat is referred to as its vertical mass center.

The locomotion tracking result can be accessed by the reconstructed virtual rat with multiple viewports, raw depth images,

color video, and its path. For performance validation, our system provides a manual rat locomotion recording interface. In validation mode, the system pauses at each frame until the rat is located by a mouse click with the confirmation key pressed. Both the automatically measured path and manually located path are recorded in a pairwise manner in Microsoft Excel® spreadsheet format.

2.6. 3D locomotion reconstruction

Locomotion of the rat was reconstructed by an algorithm that directly accesses the depth image matrix according to the rat's mass center, size, and solid shape extracted by the locating and tracking algorithm. The reconstruction algorithm was developed using the OpenGL graphic library to render a detailed locomotion model of the rat by projecting a depth image point cloud according to the requested angle of view.

2.7. Procedures

Validity assessment methods of locomotor measurements were proposed in previous efforts on animal behavior measurements (Lind et al., 2005; Tort et al., 2006). The behavior apparatus with a transparent slope and the range camera were placed in the experiment room with ambient light and light-proof curtains.

The system noise level could be accessed by tracking a stationary object (Lind et al., 2005). A Wistar rat (286.0 g) was anesthetized with sodium pentobarbital (50 mg/kg; i.p.) and placed into the behavior apparatus. The anesthetized Wistar rat was recorded by the range camera for 3 min both in lightened and lightless conditions to assess the basic system noise.

The experiments for validation were performed with 15 animals: 5 Long Evans rats and 5 Wistar rats to assess the confounding effect of fur color of the animals; 5 rats and 5 C57BL/6 mice to assess the effect of animal's size.

In each experiment, the range camera records the 15 animal in the condition with ambient light for 10 min. The animal free roamed spontaneously during the experiment. If the animal did not present spontaneous behavior, sounds were produced by tapping the apparatus to provoke its behavior before the recording.

Another experiments with 6 animals were performed in lightless conditions to assess the system's performance in experiments with circadian rhythm or longterm home cage behavioral phenotyping with lighting changes. Therefore, we recorded 21 video clips of free-roaming animals and 2 clips of anesthetized animals in total to assess the validity of the system.

Each video clip with 10-min length was analyzed in validation mode with our system to compare automatically and manually recorded measurements. 5 observers manually scored all 15 video clips respectively. The frames with obvious manual tracking errors were discarded and not included in the further analysis. Therefore, 75 pairs of tracking results in lightened conditions were analyzed. Other 30 pairs recorded in lightless conditions were examined to report path omissions by the observers because the color camera could only capture black frames. According to a published method (Tort et al., 2006), we used Pearson's correlation and Student's *t*-test to assess results of automated measuring and manual recording of a rat's horizontal locomotion. Due to the rearing behaviors would affect the vertical tracking results, the validity assessment focused on horizontal tracking performance.

Considering that rearing behavior may statistically affect vertical locomotion measurements, both the raw results and rearing-excluded results were analyzed in Section 4. In a lightless condition, we recorded a depth image stream and color video, but the CCD camera was incapable of sensing the scene.

Locomotion measurements of the 15 animals produced by our off-line tracking system and manual recording were analyzed by

Table 1

For each individual of the animal, we listed the mean values and standard errors (mean $\pm$ S.E.) of the Pearson's correlation, mean absolute error (MAE), and root mean square (RMSE) of the automated and manual tracking sessions ($n = 5$).

Animal	Weight (g)	Correlation (%)	MAE (%)	RMSE (pixel)
Wistar				
A	386.0	94.3 $\pm$ 1.7	4.4 $\pm$ 2.4	10.6 $\pm$ 1.6
B	395.0	93.6 $\pm$ 2.8	6.4 $\pm$ 4.1	9.4 $\pm$ 1.2
C	398.0	90.7 $\pm$ 0.3	17.4 $\pm$ 12.7	11.1 $\pm$ 0.9
D	400.0	94.5 $\pm$ 1.6	13.9 $\pm$ 9.2	10.5 $\pm$ 0.5
E	417.5	93.9 $\pm$ 0.9	6.7 $\pm$ 3.0	8.4 $\pm$ 0.2
Group average	399.3 $\pm$ 11.5	93.4 $\pm$ 1.5	9.8 $\pm$ 6.3	10.0 $\pm$ 0.9
Long Evans				
A	378.0	91.5 $\pm$ 1.2	23.0 $\pm$ 15.0	10.4 $\pm$ 2.6
B	286.0	92.6 $\pm$ 2.1	10.5 $\pm$ 6.3	15.3 $\pm$ 1.6
C	242.5	93.6 $\pm$ 1.0	15.7 $\pm$ 18.3	13.4 $\pm$ 1.8
D	241.5	92.6 $\pm$ 1.6	7.8 $\pm$ 2.8	10.2 $\pm$ 4.0
E	241.0	91.2 $\pm$ 0.7	9.4 $\pm$ 3.6	9.9 $\pm$ 1.8
Group average	277.8 $\pm$ 59.2	92.3 $\pm$ 1.3	13.3 $\pm$ 9.2	11.8 $\pm$ 2.3
C57BL/6				
A	34.5	94.5 $\pm$ 2.5	10.1 $\pm$ 5.4	12.9 $\pm$ 0.7
B	31.0	94.3 $\pm$ 1.4	14.3 $\pm$ 6.1	10.4 $\pm$ 2.1
C	30.5	94.5 $\pm$ 2.5	16.8 $\pm$ 3.6	12.1 $\pm$ 2.1
D	29.0	92.6 $\pm$ 2.2	17.6 $\pm$ 12.8	8.9 $\pm$ 2.7
E	27.5	93.9 $\pm$ 2.0	10.7 $\pm$ 5.8	10.7 $\pm$ 2.1
Group average	30.5 $\pm$ 2.6	94.0 $\pm$ 2.1	13.9 $\pm$ 6.7	11.0 $\pm$ 1.9

Pearson's correlation coefficient and paired Student's *t*-test with $p < 0.05$ accepted for statistical significance, mean absolute error, and root mean square error in Microsoft Excel®. The system noise level was analyzed by calculating the distance between the estimated mass center of the anesthetized rat in each frame and the mass center in the first frame.

To assess the accuracy of our system, we adopted a previously published method (Lind et al., 2005). We designed a turntable according to the method, which consisted of a gear box, a transparent cantilever, and a plastic box as a model rat. One side of the transparent acrylic cantilever was fixed to the gear box, and a plastic box (50 cm $\times$ 30 cm $\times$ 11 cm) was fixed 25 cm from the motor shaft. Therefore, the plastic box travelled in a circular path at a mean angular speed of 2.5 radians/s for 10 rounds. The ratio of the depth image pixels to centimeters was defined as 5.0. The actual distance travelled was compared to the tracking result of our system.

3. Results

3.1. Validation

Assessments of the automated tracking and the manual tracking yielded correlation coefficients, the mean absolute errors, and the root mean square errors derived from individual means are presented in Table 1. In addition, examples of correlations between automated tracking and manual tracking of the 3 strains of rodents are shown in Fig. 3, which represented the strong correlations between automated and manual tracking. The 75 pairs of tracking results were further analyzed by one-way ANOVA, and all of the *p*-values testing whether the factors including fur colors, sizes, and strains are consistent with the null hypothesis ($p > 0.05$). Therefore, the correlations of animal's colors, sizes, and strains to system performance were not significant. Significance of different observers is also consistent with the null hypothesis ($p > 0.05$) through one-way ANOVA. All 6 clips of raw depth-map frames and automated tracking results of the experiments in lightless conditions were compared by the observers and no path omissions reported.

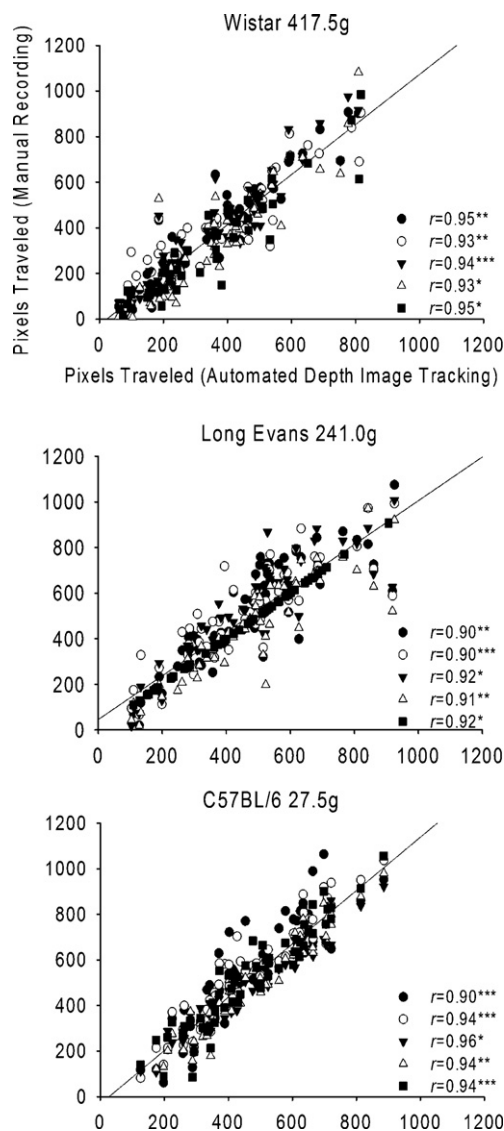


Fig. 3. Examples of correlations between the automated locomotion measurements (x-axis) and manual recording (y-axis); * $p < 0.05$ compared with the automated measurements; ** $p < 0.01$; *** $p < 0.001$. The symbols (black and white circles, triangles, and squares) represented experiment sessions completed by each observer, and the correlation coefficients (r) were presented in the right of the symbols. The results presented high correlation coefficients with significance ($r = 0.90$ – 0.96). According to the regression lines, the results presented small deviations between correlation coefficients.

Since the absolute error of the turntable test was 74.2 to total 3927.0 pixels (14.8–785.4 cm) and the relative error was 1.9%, the accuracy of our system was 98.1% with a standard deviation of $\sigma = 3.2\%$. The system noise levels were 0.07 cm per frame in lightened condition and 0.05 cm per frame in lightless condition, and the noise level were limited to 2 pixels in both clips.

For proper rapidity and resolution, the raw depth image stream and color video were both recorded at a 640×480 resolution at about a 20-Hz frequency. The frame rate of the off-line analysis was about 10–20 Hz. If the system recorded the depth image only, the sample rate could exceed 20 Hz for 1.8 MB per frame. With the present hardware setup, around 5 horizontal pixels represented 1 cm on the floor of the apparatus, which is very close to the hardware specification. Our system grabbed and interpreted all 12-bit (2048) vertical resolutions from the IR range camera for $640 \times 480 \times 2048$ solid space resolution without sensor fusion techniques.

3.2. 3D pose reconstruction

With our off-line analysis software, the system simultaneously demonstrates a reconstructed rat in solid space from three angles of view. For clarity, locomotion indices including travel distance, velocity, acceleration, margin occupation, and stereotypic movement are available for analysis in the spreadsheet file generated after each tracking session. Statistical assessments were described in the validation section. In lightless conditions, the depth image was not perceptibly affected. On the contrary, the color video window displayed only dark blank frames.

The OpenGL-based 3D pose reconstruction displays the detailed rat body and behavior including ears, forelimbs, and tail with continuous locomotion in solid space (Fig. 4). For the locomotion analysis, the tail was not reconstructed on the viewport. The system provides comprehensive rat behavioral information including crouching, stretching, grooming, sniffing, climbing, and changes in the head positions.

In the present system, all of the data were grabbed from a single sensor (the color video was for a validity assessment only); a dynamic solid replica of the scene could be observed in the virtual space without stitching. For example, in Fig. 4, the rat scrambled onto the slope in 10 frames from frame 15 to frame 25. The reconstructed path (black line) was distinctively displayed above the model slope (gray grids) along with the movement of the rat with body bending and vertical head movement. The rat explored and scrambled down from frame 200 to frame 242, with the path rendered in detail in three viewports. Then, rearing behavior was recorded within 10 frames of around 0.5 s. The rat reared its right forelimb and touched the wall of the apparatus in frame 295. Finally it drew its forelimb back in frame 297.

4. Discussion

The behavioral measurement method using IR range camera depth images in the present study provides obvious advantages on rodent behavioral studies: robustness, rapidity, reliability, and resolution (Publicover et al., 2009). Moreover, our approach is also simpler and cheaper than previously published methods.

Measuring a rat's locomotion using depth images grabbed from a single IR range camera preserves detailed behaviors on account of its spatial ($640 \times 480 \times 2048$ pixels) and temporal (20 Hz) resolution. With a high spatial resolution, the system is capable of detecting minor movements and avoids underestimating locomotion (Lind et al., 2005). Though Ethovision XT® (Noldus, Wageningen, The Netherlands) and other commercial systems provides a resolution of 768×576 pixels (Noldus et al., 2001), the resolution of the proposed method is not restricted to Kinect®'s 640×480 pixels resolution. The depth map could be acquired from other appropriate instruments including industrial infrared ranging camera or through stereo vision technique, which could possibly provides resolution up to 1920×1080 pixels (Full HD 1080p) using off-the-shelf digital cameras. For cost efficiency, we used Kinect® for depth map acquisition in the experiments.

To avoid omitting necessary information, we utilized a relatively simple compression format. About 360 MB/min was used to record the depth images only, and with a color video, about 1.4 GB/min was used, which provides high throughput and detailed behavioral information. Our system is also able to record and analyze data in other compression formats if the memory capacity is limited.

Since the IR range camera utilizes the IR region of the electromagnetic spectrum, the proposed approach is immune to interference from the visible-light spectrum. During the experiments, light variations and shadows were not optimized, which made tracking rats using a CCD camera very difficult (Lind et al.,

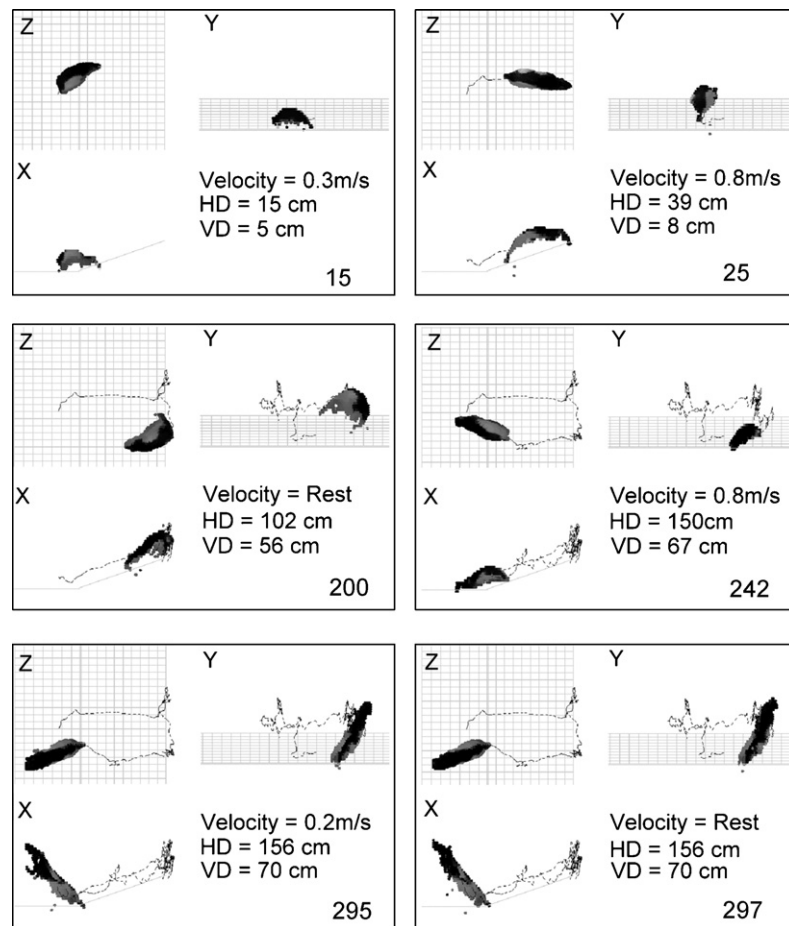


Fig. 4. Representative results exemplified by a dynamically reconstructed rat pose viewed from different angles in solid space. The accumulated horizontal distance (HD), accumulated vertical distance (VD), and velocity were shown in each frame. “Rest” indication replaced the velocity value if the subject was at rest. From frame 15 to frame 25, the rat scrambled along the slope with motion reconstructed in detail with observable features of the head and ears within 1 s. From frame 200 to frame 242, the rat explored the slope with a combination of behaviors, and scrambled down to the floor, which can be observed by the path (black line). Limited by the range sensor’s field of view, the lower part of the rat was not retrieved and not interpolated. Dispersed points are noise pixels which were intentionally not filtered out for the performance assessment. In frame 295, the rat reared up on the floor and touched the wall of the behavior apparatus with its right forelimb. In frame 297, it drew the forelimb back with a subtle head movement.

2005; Publicover et al., 2009). However, the IR depth images were not retrieved within the visible light spectrum. In addition, liquid or solid excretion of the rat during experiments did not affect the tracking performance. During the experiment in ambient-lighted condition, shadows, animal’s excretions, contrasts, and light variations were not strictly controlled to simulate a tough phenotyping setup in a common laboratory. The experiments in lightless condition performed in a light-proof room, were not able to be analyzed by the operators because the color camera recorded no more than a black screen, and other active infrared cameras may interfere the Kinect®. We examined the raw depth-map frames, and the recorded clips did not present target omission or obvious error. Thus, our system is robust to factors in the visible light spectrum and which suggests reliability of locomotion measurements. Our approach performed as well as in tough setups with low foreground-to-background contrast, changing background (animal’s excretion or color changes during floor drying), and varied lighting conditions in experiments related to circadian rhythms and cognition (Publicover et al., 2009). The system is also robust to various colors and sizes of the animal, and is capable of tracking Wistar rats (white, about 386.0–417.5 g), C57BL/6 mice (black, 27.5–34.5 g), and Long Evans rats (black hooded, about 241.0–378.0 g). Tracking results were tested through one-way ANOVA of all 3 strains of rodents, and the results presented the accuracy of the system were not affected by the varied fur color

and body sizes. The Pearson’s correlation coefficients and the root mean square errors of automated and manual tracking also validated the tracking performance of the system. According to the inspections on the raw records, the deviations of mean absolute errors in specific sessions were suffered from the errors made by human observers.

Moreover, data omission or noises were subtle during the experiments (Fig. 4), the locomotion parameters were recorded directly without smoothing algorithms. But if the system retrieves depth images from other instruments, for instance, stereo vision, further smoothing on the three-dimension coordinates of locomotion parameters may be necessary. Moreover, the system includes a composite sensor device, an apparatus, and a computer that was assembled for < US\$1500, which is very reasonable compared to off-the-shelf systems using IR photocells and IR cameras. Since similar range-sensing technologies are still in development, alternatives of the hardware at lower costs will become available in the future.

Assessments of the validity of the system yielded significant statistical results of automated locomotion measurements compared to manual recording and topographical models even with relatively complicated setups. The automated tracking results correlated to the manually recorded results with evident significance and correlations for horizontal tracking 15 of rodents, and results were close to the linear regression line (Fig. 3). The horizontal tracking correlation coefficient was similar to published research (Tort et al., 2006),

at $r = 0.976$, $p < 0.001$. Minor deviations in the regression line were possibly caused by an inherent inaccuracy of the manual recording technique (Tort et al., 2006). An underestimation of the topographical model indicated that using only 2D information may cause considerable omission of locomotion measurements. For accuracy, our system yielded a 98.1% accuracy with a 640×480 -pixel resolution and a 25-cm-radius circular path, which is superior to a published method using a video camera (Lind et al., 2005), which had a 9% error with a 384×288 -pixel resolution and a 14-cm-radius circular path. The results suggest that higher hardware resolution is advantageous for greater accuracy. The three-dimensional system noise level was limited to 0.7 cm per frame in both lightened and lightless condition, and the variation was limited to 2 pixels, which suggest the system still has a noise caused by the instruments and the experiment condition and similar to the published paper (Lind et al., 2005).

Since deviations in tracking may result from subjectivity and fatigue (Lind et al., 2005; Tort et al., 2006), the results suggest greater accuracy and reliability of the present method for 3D locomotion measurements.

For 3D tracking, multiple cameras or camera-mirror systems require sensor fusion techniques including registration, mapping, distortion compensation, and frame synchronization, which limit the resolution, field of view, and accuracy of behavior detection. The depth image provides the possibility of reconstructing a subject in behavioral experiments as a dynamic 3D model, which provides a 3D view of the subject's behavior in detail. Massive spatial information of depth images was applied by our system to reconstruct 3D poses of a rat, including vertical head movements and body bending. Observing detailed body part movements with a single video camera is challenging, even with multiple-camera systems, due to constraints of rectilinearity of the light, and only the overall body locomotion is able to be analyzed, while body-part and extremity movements are inevitably omitted.

The reconstructed dynamic rat model was not refined to reveal detailed features of body parts, and was interpreted in the OpenGL 3D model in a point cloud fashion. Researchers can simultaneously view the rat from different angles through the graphical user interface. The solid reconstruction result distinctively demonstrates the temporal path of the rat in a relatively complicated topographical setup. These subtle movement detections are valuable for neuroscience studies, and suggest the potential for application to gait analysis, and even discovering new behaviors (Tecott and Nestler, 2004). Meanwhile, the system also presented detailed movements of body parts without constraints of the viewing angle (Fig. 4). The extracted results suggest the evident potential of developing more-exact and -qualitative measurements of a rat's behavior at the body-part scale and in experimental setups with complicated topographical features and light conditions. These important behaviors are not easily analyzed in detail by computers with only horizontal information. We also noted that even when the rat was standing still, minor horizontal and vertical movements existed, which were usually smoothed over by previous studies.

Published studies on qualitative comparisons of behavioral analysis systems proposed that most off-the-shelf systems require specific hardware and provide mainly pharmacological validation (Spruijt and deVisser, 2006) but qualitative validation reports. Moreover, not all commercial systems provide behavior recognition ability. For example, Ethovision XT® (Noldus, Wageningen, The Netherlands) provides an IR camera setup option and limited behavior recognition (Noldus et al., 2001; Spruijt and deVisser, 2006), and tracks insects and other animals in a 3D space using a multi-camera sensor fusion technique. But our system provides a more-versatile and -efficient configuration and approach. With a single off-the-shelf range camera, our approach provides 3D locomotion measurements, and ver-

satility of tracking and reconstructing detailed 3D behaviors of larger vertebrates. Both LABORAS® (Metris, Hoofddorp, The Netherlands) and HomeCageScan® (CleverSys, Reston, VA, USA) provide advanced behavior recognition capabilities for various rat behaviors. LABORAS® reconstructs the animal behaviors physical vibration and force signal, but its sensitive to physical noise. HomeCageScan® is based on body part feature recognition recorded from video cameras from limited viewpoints (Spruijt and deVisser, 2006).

However, further behavior detection capabilities using depth images are still in development.

The primary limitation of our approach is reconstruction of shaded parts of the rat. Features on the shaded abdomen side of rat were omitted since the IR range camera still followed the rectilinearity of light. Limbs of the rat are challenging to detect when shaded by the body, which limits the application to gait analysis. Vertical limb movements are subtle for a 1-cm vertical resolution of Kinect®. Depending only on depth images without color images also limits performance of tracking with color features, including tracking limbs or multiple rats. Using sensor fusion techniques with more range cameras or video cameras will be introduced for more-comprehensive dynamic rat pose reconstruction and gait analysis. Considering the complexity of rat behavior classification, automated body part behavior tracking is not discussed in this paper. More-reliable behavior detection algorithms for depth images are required for sniffing, grooming, rearing, and stereotypical movements for accurate automatic pose reconstruction, which can be assessed by qualitative methods (Van de Weerd et al., 2001) for behavioral classification.

For wider applications in behavioral studies, tracking multiple animals was tested, but more-complex interaction patterns are still being analyzed for accurate tracking algorithms. Automatic floor detection was tested in early experiments and yielded fine results; we are introducing opaque topographical feature modeling to enhance the floor detection algorithm. With inspiring potential for subtle behavior and gait analysis, it is reasonable to introduce feature-based body-part and extremity recognition algorithms to track rat's behavior to the body-part scale.

5. Conclusions

In conclusion, we developed a locomotion measurement and pose reconstruction system based on IR range camera depth images for locomotion analysis of rodents. The system provides superior robustness, rapidity, and reliability, and high temporal and spatial resolution of 3D tracking with low-cost hardware and relatively simpler and faster algorithms than previous studies. The validity of the system was assessed using proposed methods, and yielded superior accuracies for horizontal tracking, vertical tracking, and behavior detection in non-optimized experimental conditions. The system also reconstructs a detailed dynamic virtual rat in the computer with subtle body poses available, which suggests the potential of qualitative studies on subtle behaviors at the body-part scale in free-roaming experimental setups.

Acknowledgements

This study was supported by a grant (NSC98-2313-B-197-003-MY3) from the National Science Council, Taiwan.

References

- Aagaard A, Andersen BB, Depledge MH. Simultaneous monitoring of physiological and behavioural activity in marine organisms using non-invasive, computer-aided techniques. *Mar Ecol Prog Ser* 1991;73:277–82.
- Chršková J, Kaminsky Y, Krekule I. An automatic 3D tracking system with a PC and a single TV camera. *J Neurosci Methods* 1999;88:195–200.

- Clarke RL, Smith RF, Justesen DR. An infrared device for detecting locomotor activity. *Behav Res Methods Instrum Comput* 1985;17:519–25.
- Denenberg VH. Open-field behavior in the rat: what does it mean? *Ann N Y Acad Sci* 1969;159:852–9.
- Ericson E, Samuelsson J, Ahlenius S. Photocell measurements of rat motor activity: a contribution to sensitivity and variation in behavioral observations. *J Pharmacol Methods* 1991;25:111–22.
- Fry SN, Bichsel M, Müller P, Robert D. Tracking of flying insects using pan-tilt cameras. *J Neurosci Methods* 2000;101:59–67.
- Godden DH, Graham D. 'Instant' analysis of movement. *J Exp Biol* 1983;107:505–8.
- Hoy JB, Koehler PG, Patterson RS. A microcomputer-based system for real-time analysis of animal movement. *J Neurosci Methods* 1996;64:157–61.
- Kao SD, Shaw FZ, Young MS, Jan GJ. A new automated method for detection and recording of animal moving path. *J Neurosci Methods* 1995;63:205–9.
- Kato S, Tamada K, Shimada Y, Chujo T. A quantification of goldfish behavior by an image processing system. *Behav Brain Res* 1996;80:51–5.
- Lind NM, Vinther M, Hemmingsen RP, Hansen AK. Validation of a digital video tracking system for recording pig locomotor behaviour. *J Neurosci Methods* 2005;143:123–32.
- Moeslund TB, Granum E. A survey of computer vision-based human motion capture. *Comput Vis Image Und* 2001;81:231–68.
- Nilsson GE, Rosen P, Johansson D. Anoxic depression of spontaneous locomotor activity in crucian carp quantified by a computerized imaging technique. *J Exp Biol* 1993;180:153–62.
- Noldus LPJJ, Spink AJ, Tegelenbosch RAJ. Ethovision: a versatile tracking system for automation of behavioral experiments. *Behav Res Methods Instrum Comput* 2001;33:398–414.
- Obradovic Z, Pantovic S, Rosic G, Selakovic Z, Rosic Mirko. The new experimental model for behavioral investigations in animal studies. *Gen Physiol Biophys* 2009;28:77–86.
- Pan WHT, Lee CR, Lim LH. A new video path analyzer to monitor travel distance, rearing and stereotypic movement of rats. *J Neurosci Methods* 1996;70:39–43.
- Publicover NG, Hayes LJ, Guerreroc LF, Hunter KWJ. Video imaging system for automated shaping and analysis of complex locomotory behavior. *J Neurosci Methods* 2009;182:34–42.
- Reynolds DR, Riley JR. Remote-sensing, telemetric and computer-based technologies for investigating insect movement: a survey of existing and potential techniques. *Compute Electron Agric* 2002;35:271–307.
- Rousseau JBI, Van Lochem PBA, Gispens WH, Spruijt BM. Classification of rat behavior with an image-processing method and a neural network. *Behav Res Methods Instrum Comput* 2000;32:63–71.
- Spruijt BM, deVisser L. Advanced behavioural screening: automated home cage ethology. *Drug Discov Today Technol* 2006;3:231–7.
- Tecott LH, Nestler EJ. Neurobehavioral assessment in the information age. *Nat Neurosci* 2004;7:462–6.
- Tort ABL, Neto WP, Amaral OB, Kazlauskas V, Souza DO, Lara DR. A simple webcam-based approach for the measurement of rodent locomotion and other behavioural parameters. *J Neurosci Methods* 2006;157:91–7.
- Twining CJ, Taylor CJ, Courtney P. Robust tracking and posture description for laboratory rodents using active shape models. *Behav Res Methods Instrum Comput* 2001;33:381–91.
- Van de Weerd HA, Bulthuis RJA, Bergman AF, Schlingmann F, Tolboom J, Van Loo PLP, et al. Validation of a new system for the automatic registration of behaviour in mice and rats. *Behav Process* 2001;53:11–20.