Silhouette-Based Markerless Motion Estimation of Awake Rodents in PET

Andre Z. Kyme, *Member*, *IEEE*, Paul Strenge, Felicity Lee and Steven R. Meikle, *Senior Member*, *IEEE*

***Abstract*– The ability to image the brain of a freely moving rodent using motion-compensated PET presents many exciting possibilities for exploring the links between brain function and behavior. Markerless optical approaches for pose estimation have several potential advantages over marker-based methods: improved accuracy and increased range of detectable motion; no 'decoupling' of marker and head motion; and no acclimatization of the animals to attached markers. Our aim in this work was to describe and validate a silhouette-based multi-camera method for estimating the pose of a rat.**

Random-walk and K-means clustering approaches were very adaptable to uneven lighting and generally provided excellent object segmentations. In obtaining a high quality rat model, shape-from-silhouette and laser scanning both resulted in useful models; laser scanning provided sub-millimeter resolution with very few artifacts and was the method of choice. In our experimental validation, the 3D-2D (model-silhouette) optimization clearly converged to sub-degree and sub-millimeter alignment of the measured and estimated silhouettes. The average discrepancy between test points transformed using the estimated versus ground-truth poses was 0.94 mm ± 0.51 mm.

This investigation focused on rigid motion of a rat phantom as a proof-of-principle of the technique. Future work will focus on investigating the potential of designing a non-rigid rodent body model in order to apply the method to a freely moving animal during PET imaging.

I. Introduction

The ability to image the brain of a freely moving rodent using motion-compensated positron emission tomography (PET) presents many exciting possibilities for exploring the links between brain function and behavior [1].

A key requirement of this approach is obtaining accurate estimates of animal pose throughout a scan. Markerless optical approaches for pose estimation have several potential advantages over marker-based methods: improved accuracy and increased range of detectable motion; no 'decoupling' of marker and head motion; and no acclimatization of the animals to attached markers [2].

Here, our aims are (i) to describe and validate a calibrated multi-camera setup for capturing image streams of a freely moving rat; (ii) to describe techniques for extracting silhouettes from these image streams; and (iii) to demonstrate how the silhouettes can be used in conjunction with an *a priori* rat model to extract pose estimates in close agreement with ground-truth measurements.

Manuscript received Nov 10, 2017.

A. Z. Kyme is with the School of AMME, Faculty of Engineering & IT, and the Brain & Mind Centre, University of Sydney, Sydney, Australia (telephone: +612 93512260, e-mail: andre.kyme@sydney.edu.au).

P. Strenge is with the Institute of Medical Engineering, Universitat zu Lübeck, Lübeck, Germany (e-mail: p-strenge@t-online.de).

F. Lee is with the Australian National University, Canberra, Australia (e-mail: u5558958@anu.edu.au).

S. R. Meikle is with the Faculty of Health Sciences and Brain & Mind Centre, University of Sydney, Sydney, Australia (e-mail: steven.meikle@sydney.edu.au).

II. Methods

A. Image Segmentation

The accuracy of silhouette-based methods depends strongly on the accuracy of segmentations. We investigated several segmentation approaches: simple thresholding, region-growing, a hybrid thresholding and random walk algorithm, and K-means clustering. Algorithms were compared with respect to (i) adaptability to uneven lighting; (ii) susceptibility to segmentation artifacts; (iii) efficiency. Although thresholding and region-growing methods were simple and fast, both were prone to leaking artifacts in shadowed parts of the image. The random-walk and K-means clustering approaches were very adaptable to uneven lighting and generally provided excellent segmentations. The K-means algorithm occasionally resulted in artifacts due to the randomization step. Overall the method of choice was the random walk approach. The pipeline for implementing this approach and an example segmentation are shown in Fig. 1.

B. Multi-Camera Setup

The multi-camera setup comprised monochrome 1280x1024 resolution CMOS cameras (Flea3, FLIR) fitted with either 3.5 mm low distortion wide-angle lenses (Edmund Optics) or 12 mm narrow-angle lenses (Goyo Optical). Intrinsic camera parameters (principal point, focal length, distortion coefficients) were determined for each camera, then all cameras were spatially calibrated using a multi-camera self-calibration method [3]. We synchronized and cross-calibrated the camera network with either a 6-axis robot (Epson Corp, Japan) or marker-based optical motion tracking (MicronTracker, ClaronNav) used previously for rat head tracking [4]. The robot and the marker-based system provided ground-truth motion estimates for comparison with estimates from the markerless approach.

C. Space Carving

The space carving algorithm [5] provides an estimate of 3D shape given a set of silhouettes of the object from multiple views. It is well known, however, that when the number of views, n, is limited there is insufficient information to properly reconstruct the object. Using many views to improve the space carving reconstruction is impractical for small animal PET

978-1-5386-2282-7/17/$31.00 ©2017 IEEE

systems. An alternative is to include an *a priori* rat model and determine pose by maximizing similarity between the model and measured silhouettes.

D. Rodent Model

The furry texture of rodents makes it challenging to obtain a robust 3D model. We investigated 3 different approaches to obtain a model of a taxidermal rat specimen: (i) microCT; (ii) many-view shape-from-silhouette; and (iii) 3D laser scanning.

E. Pose Estimation

A point $\mathbf{X}=(x, y, z)$ in 3D space can be projected to image coordinates for camera i according to:

$$\mathbf{x}_i = \mathbf{P}_i \mathbf{X} \tag{1}$$

where $\mathbf{x}_i = (u_i, v_i)$ are the image coordinates and $\mathbf{P}_i$ is the camera calibration matrix for camera i (i = 1, 2, 3, …). Let $\mathbf{S}_i$ denote all points in the segmented silhouette extracted from camera i and let $\mathbf{S}_i'$ be the silhouette obtained by reprojecting the model points $\mathbf{X} = [\mathbf{X}_1\ \mathbf{X}_2\ \ldots\ \mathbf{X}_N]$ after transformation by an arbitrary rigid-body transformation:

$$\mathbf{S}_i = \mathbf{P}_i \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{X} \tag{2}$$

where $\mathbf{R}$ and $\mathbf{t}$ are the rotation and translation components, respectively, of the transformation. We can then estimate the pose of the rat by solving:

$$\underset{\mathbf{R},\mathbf{t}}{\arg\min}\ f(\mathbf{R},\mathbf{t}) \tag{3}$$

where

$$f(\mathbf{R},\mathbf{t}) = \sum_{i=1}^{4} \mathrm{SSD}(\mathbf{S}_i, \mathbf{S}_i') \tag{4}$$

and SSD is the sum of squared difference operator.

F. Experimental Validation

We performed two types of phantom experiments to validate the method using the taxidermal rat specimen. In the first experiment, 3 cameras were set up to view the head of the phantom, which was rigidly attached to the robot end-effector. The head was moved to 500 discrete poses sampled from real rat head motion. At each pose we collected images from the three cameras for silhouette-based pose estimation. In the second experiment, we used four symmetrically arranged cameras to track 6 discrete poses of the complete rat phantom (head + body). A marker attached to the rat was tracked using the MicronTracker and provided ground-truth motion estimates (Fig. 2). The Nelder-Mead optimization algorithm was used to solve eqn (3) for both experiments.

III. Results

Figure 3 shows the results of the space carving algorithm (without color information) for our 4-camera setup. Artifacts are clearly apparent.

In obtaining a high quality rat model, microCT proved to be a poor option due to metal artifacts generated from wires used in the taxidermy procedure. Shape-from-silhouette (140 views) and laser scanning both resulted in useful models; laser scanning provided sub-millimeter resolution with very few artifacts (Fig. 4) and was the method of choice.

In the head-only experimental validation of the silhouette-based method, the optimization clearly converged to sub-degree alignment of the measured and estimated silhouettes (Fig 5). In the whole-body validation, the average discrepancy between test points transformed using the estimated versus ground-truth poses was 0.94 mm ± 0.51 mm (Fig 6).

IV. Discussion and Conclusion

A 3D-2D (model-silhouette) based registration approach for pose estimation of the rodent body has been developed which provided sub-millimeter accuracy in a phantom study. Optimization of the alignment between measured and estimated silhouettes was free of local minima using four views but may have been more limited using three views. However, a thorough characterization of the dependence of the accuracy on the number of views requires further investigation. We also demonstrated that an accurate rat model can be obtained relatively quickly and easily using a laser scanning approach without the fur causing artifacts.

This investigation focused on rigid motion as a proof-of-principle of the technique. Future work will focus on investigating the potential of designing a non-rigid rodent body model in order to apply the method to a freely moving animal during PET imaging.

Acknowledgement

This work was supported by the Australian Research Council (DE160100745). The authors wish to thank Dr Tzong-Tyng Hung from UNSW Biological Resources Imaging Lab, Matt Adcock and Stuart Anderson from Quantitative Imaging, CSIRO, and Mark Cray and Nicholas Schwabe from HiTech Metrology for help in testing various techniques to model a rodent. Paul Strenge was supported by a student internship allowance from Universitat zu Lübeck.

References

[1] Cherry 2011 Nat. Methods 8 301-03.
[2] Kyme et al. 2014 Trans. Med. Imaging 33 2180-90.
[3] Svoboa et al. 2005 PRESENCE 14 407-22.
[4] Kyme t al. 2012 J. R. Soc. Interface 3094-3107.
[5] Kutulaos and Seitz 2000 Int. J. Comput. Vis. 38 199-218.

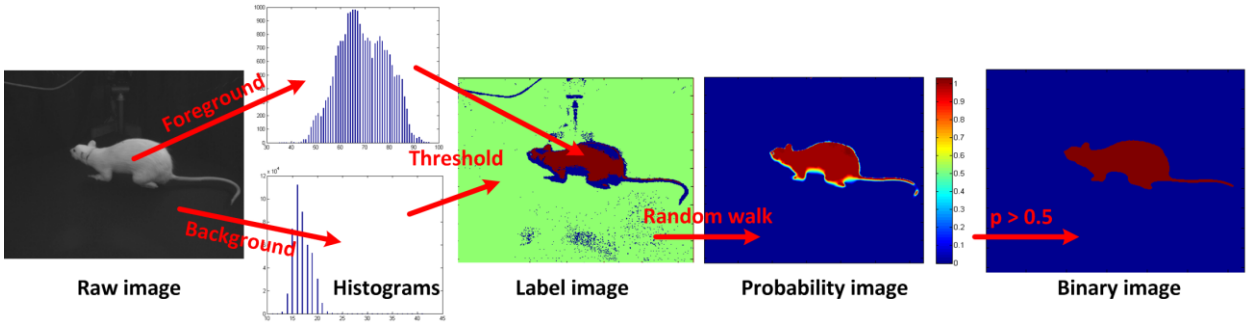

Fig. 1. Segmentation pipeline. Histograms of foreground / background regions are formed; a statistical threshold is applied to form a label image; the random walk algorithm is used to generate a probability image; a threshold is applied to obtain the segmentation.

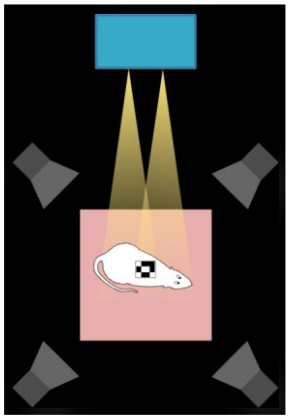

Fig. 2. For the experimental validation of the whole-body rat phantom, the phantom was tracked simultaneously using the markerless and marker-based systems.

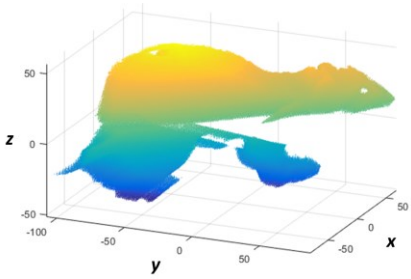

Fig. 3. Failure of space carving to reconstruct an accurate 3D rat model of a taxidermal rat specimen using our 4-camera setup. Without using colour information, space carving is equivalent to a visual hull reconstruction.

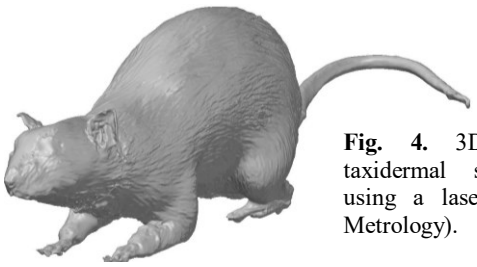

Fig. 4. 3D model of the taxidermal specimen obtained using a laser scanner (HiTech Metrology).

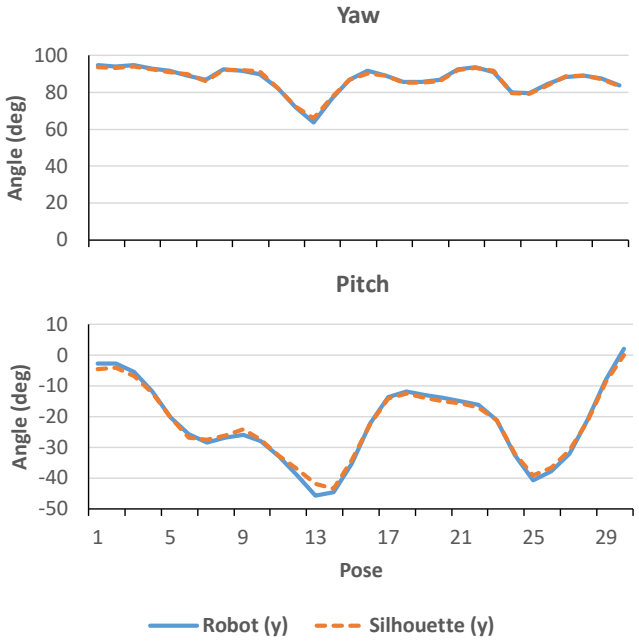

Fig. 5. Silhouette-based yaw and pitch motion estimates compared to ground truth (from robot) for 30 poses of the head.

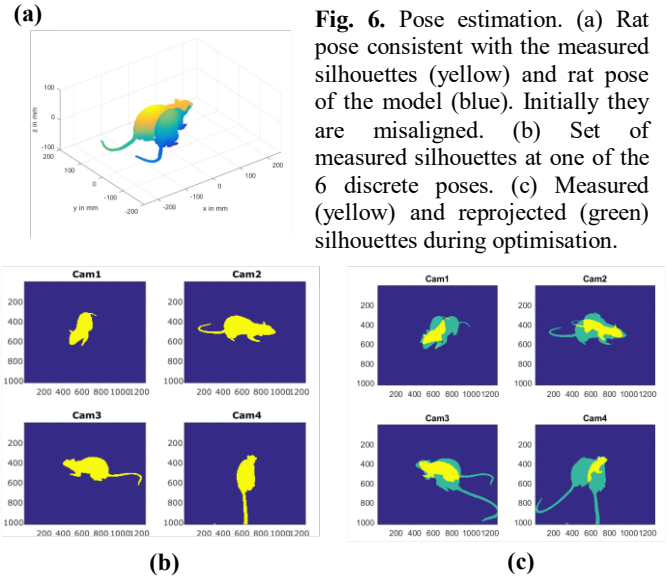

Fig. 6. Pose estimation. (a) Rat pose consistent with the measured silhouettes (yellow) and rat pose of the model (blue). Initially they are misaligned. (b) Set of measured silhouettes at one of the 6 discrete poses. (c) Measured (yellow) and reprojected (green) silhouettes during optimisation.