Markerless Motion Tracking Enabling Motion-Compensated PET in Awake Rats

Andre Kyme, *Member*, *IEEE*, Stephen Se, *Senior Member*, *IEEE*, Steven Meikle, *Senior Member*, *IEEE*, Will Ryder, *IEEE*, *Member*, Kata Popovic, Roger Fulton, *Senior Member*, *IEEE*

Abstract*– Motion-compensated PET of awake animals has the potential to greatly improve translational neurological investigations by enabling brain function to be studied during learning tasks and complex behaviors. Previously we have demonstrated the feasibility of performing motion-compensated brain PET on rodents, obtaining the necessary head motion data using marker-based techniques. However, markerless motion tracking would simplify animal experiments and potentially provide more accurate pose estimates over a greater range of motion. Previously we have described a markerless stereo motion tracking system and associated algorithms and validated the approach in phantoms. In this work we performed a pilot study to demonstrate motion-compensated ^{18}F-FDG brain imaging in an awake, unrestrained rat using head pose measurements obtained from the markerless tracking system. Motion compensation clearly worked, resulting in easily identifiable structures in the head. However, it was also obvious that considerable residual error remained after correction. Post analysis of the motion estimates indicated that the residual error was the result of occasional spurious pose estimates, most likely caused by features on non-rigid parts of the head contributing to the pose estimation. Moreover, the line-of-response rebinning used for motion correction resulted in a large proportion of lost events, leading to noisy and inconsistent projection data. The latter is avoided by using a direct list mode reconstruction. In summary, markerless tracking continues to show promise for motion-compensated imaging of awake animals, but further optimization is required to match the accuracy and consistency of marker-based tracking.

I. Introduction

Motion-compensated positron emission tomography (PET) and single photon emission computed tomography has been shown to be a feasible and convenient method to obtain functional brain data from awake, unrestrained rodents [1, 2]. One of the key requirements of successful motion compensation is having accurately sampled motion estimates of an animal's head pose. Previously we have reported an in-house optical tracking system and associated algorithms enabling such estimates to be obtained without markers [3]. Markerless motion tracking has the potential to improve both the accuracy of motion measurements and the range of detectable motion compared to marker-based systems. It also simplifies motion-compensated imaging by avoiding irritation caused by the attachment of markers, and the need to acclimatize animals to the markers. In our previous study, the principles and potential performance of the markerless tracking system were demonstrated using a taxidermied rat head under robot control [3]. Here we describe what we believe is the first demonstration of motion-compensated brain imaging of an awake animal using motion estimates from a markerless tracking system.

This work was supported by Australian Research Council Grant No. DP 0663519.

André Kyme is with the Faculty of Health Sciences and Brain & Mind Research Institute, University of Sydney, NSW 2006, Sydney, Australia (e-mail: andre.kyme@sydney.edu.au). He was with the School of Physics, University of Sydney, NSW 2006, Sydney, Australia.

Stephen Se is with MDA Systems Ltd., 13800 Commerce Parkway, Richmond, B.C. V6V 2J3 Canada.

Steven Meikle, Will Ryder, Kata Popovic and Roger Fulton are with the Faculty of Health Sciences and Brain & Mind Research Institute, University of Sydney, NSW 2006, Sydney. Roger Fulton is also with the School of Physics, University of Sydney, NSW 2006, Sydney, Australia, and with the Department of Medical Physics, Westmead Hospital, NSW 2145, Sydney, Australia.

II. Methods

A. Motion tracking system

The tracking system consisted of four Flea2 Firewire cameras (Point Grey Research, Canada) fitted with 12 mm focal length lenses (GMN21214, Goyo Optical, Japan) (Fig. 1). Synchronized capture rates of up to 60 Hz were possible, however for the pilot study described below we sampled at 30 Hz, consistent with the >20 Hz sampling requirement for rat head motion inside the microPET (Preclinical Solutions, Siemens Healthcare Molecular Imaging, Knoxville, TN, USA) [4]. The cameras were arranged in two pairs viewing opposite sides of the head and were mounted on a frame that was secured directly to the scanner bed unit (Fig. 1). Calibration of the four-camera system was determined up to an unknown scale factor (i.e. a factor relating camera frame distances to world distances) using the method in [5]. Compensation for lens distortion was included in the calibration based on independent estimates obtained using the Matlab calibration toolbox [6]. The resulting calibration was transformed to the scanner frame using a method similar to that described in [4]. Scale was resolved by calculating the ratio of point-point distances obtained from each system. Mean calibration error was <0.15 pixels in x, y and z; this corresponded to <0.05 mm in the region of the head.

B. Feature detection and matching

Features were detected in all images using the scale-invariant feature transform (SIFT) algorithm [7]. SIFT features are represented by a 128-element descriptor. Descriptors obtained in different images were matched efficiently using a nearest neighbour approach [7]. Figure 2 shows an example of feature matching.

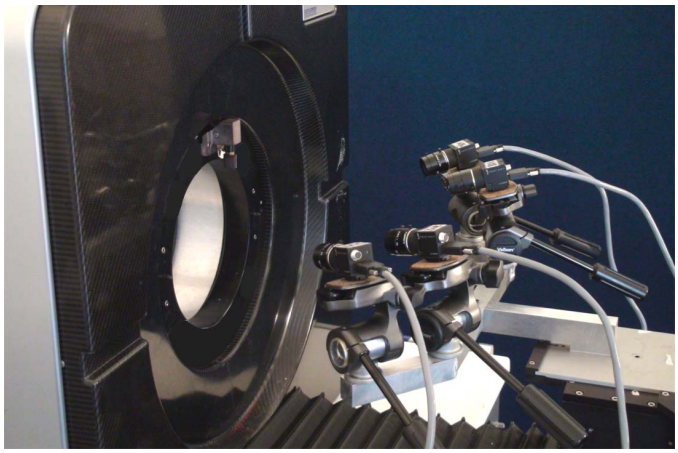

Fig. 1. Setup for markerless motion tracking. The four cameras were mounted to the microPET via an aluminium frame.

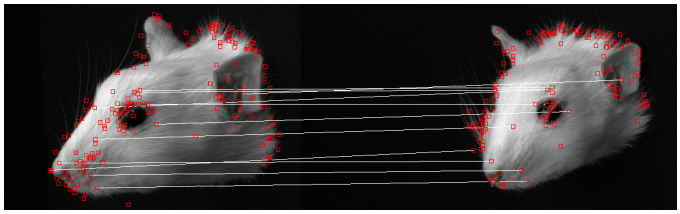

Fig. 2. Example of SIFT features (red) and matches (white lines) for a pair of images of a rat head.

C. *Landmark generation & pose estimation*

Landmark generation and pose estimation are summarised in Figs 2 and 3. Any feature matches between two images of a given frame (a frame being a set of four simultaneous images collected by the four cameras) were reconstructed as a three dimensional (3D) head landmark in the scanner frame using triangulation (Fig. 3, right-hand side). These 3D landmarks were stored in a database together with the associated SIFT descriptor. Pose estimation for a given frame (i.e. estimation of motion with respect to frame 1) required the detection of matches between images of the current frame and descriptors in the database (Fig. 3, left-hand side). Once such matches were found, pose estimation involved finding the rotation and translation minimizing the reprojection error (Fig. 4). This was performed very rapidly using an iterative Gauss-Newton approach which, instead of estimating overall pose, estimated the change from the previous pose. For a detailed description of the landmark generation and pose estimation process, see [3]. Each time the pose of a frame was estimated, any new landmarks arising from reconstructed matches in that frame were added to the database (referenced to the initial pose) and could be used for subsequent pose determination (Fig. 3). During pose estimation, outliers – that is, features associated with the background or with non-rigid parts of the head – were detected based on the reprojection error: for the first six Gauss-Newton iterations, outliers were rejected based on the deviation of the reprojection error from the median error; for the final four iterations, outliers were rejected based on the absolute value of the reprojection error.

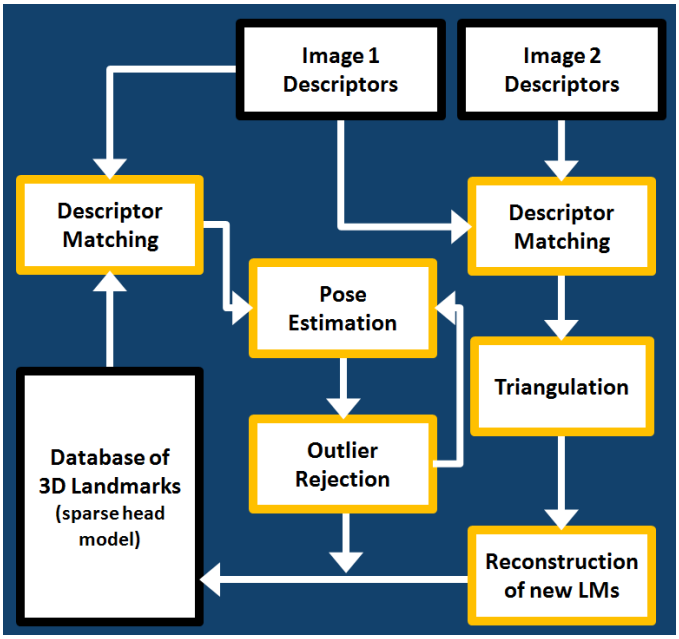

Fig. 3. Landmark (LM) generation and pose estimation. Feature (descriptor) matching between a pair of images enabled 3D LMs to be reconstructed using triangulation. Matching between an image and the LM database enabled pose estimation (see Fig. 4). Once pose had been estimated, new LMs were added to the database consistent with the initial (frame 1) pose.

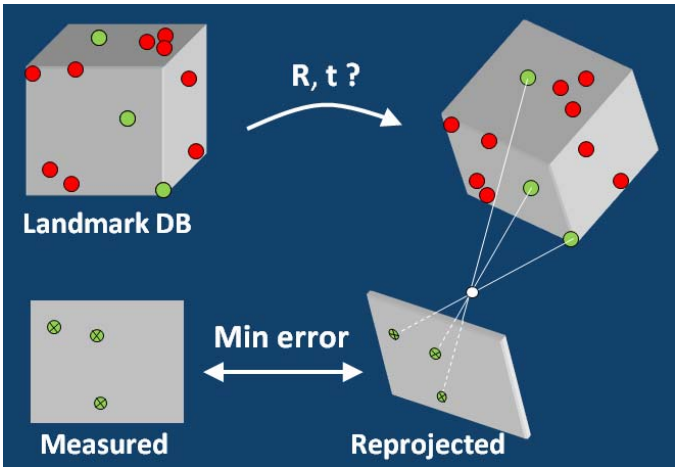

Fig. 4. Pose estimation. Once matches (here shown in green) had been found between a camera image and the descriptors in the database, pose estimation was the task of determining the rotation, **R**, and translation, **t**, minimizing the reprojection error. Instead of determining the pose directly, a correction to the previous pose was determined (see [3] for details).

D. Experimental validation

A 10-min ^{18}F-FDG PET emission scan of an awake adolescent male Sprague-Dawley rat inside an open-ended tube was acquired in conjunction with motion tracking at 30 Hz. For this study, indelible marks were drawn on rigidly-moving parts of the head to promote identification of reliable SIFT features, and lessen the effect of non-rigid structures (such as ears, whiskers and eyes) (Fig. 5). Although this approach was not strictly 'markerless', for all practical purposes it was equivalent to a markerless approach, taking just a matter of seconds to perform and having no apparent impact on the animal's behavior. Synchronization of the camera capture with the PET list mode acquisition was performed as described previously for marker-based motion tracking [4]. Data from the cameras were streamed directly to hard disk. Following the scan of the awake animal, a second 10-min emission scan was acquired with the rat under anesthesia (2% isoflurane). The tracking data from the awake rat study were processed offline to obtain motion estimates for each of the 17000 frames, relative to the first frame. The motion data were then used for line-of-response (LOR) rebinning-based motion correction as described in [8]. Reconstruction of the motion-corrected sinogram was performed using Fourier rebinning and 2D OSEM (16 subsets, 4 iterations).

III. Results and Discussion

Figure 6 shows the cumulative number of landmarks in the database as a function of pose number. Approximately 8500 landmarks were accumulated over the course of processing the 17000 frames. The rotation of the animal's head about the *x*-axis, estimated during the awake scan, is shown in Fig. 7. Figure 7 also shows six images of the rat (from camera 1) to indicate the range of motion during the scan. The image borders are color-coded with the vertical bars drawn on the graph to indicate where the particular pose occurred.

In order to check the veracity of the pose estimates, the 3D coordinates of the vertices of a triangle joining the eye centers and nose tip were obtained manually using the initial frame. For each subsequent frame these vertices were transformed according to the estimated pose and reprojected onto the camera images (Fig. 7). It is clear from Fig. 7 that although pose estimates were accurate for many of the frames, there were also periods of inaccurate pose estimation - representative examples are shown in the second and fifth images from the left. The most likely cause of this error is the inclusion of outliers in the pose estimation process. Therefore, in addition to the outlier rejection methods already being used (section IIC), further work is required to completely reject features on non-rigid parts of the head (e.g. whiskers, ears, eyes, nose and neck).

Figure 8 shows a reconstructed transverse and coronal slice from the awake rat study, with and without motion compensation. Corresponding slices for the anesthetized rat are shown for comparison. Motion compensation has clearly worked, although poorer spatial resolution and some missing data are apparent. The differences are attributable for the most part to the pose estimation errors identified. However, there were also many (~65%) lost events incurred during LOR rebinning which likely resulted in inconsistent and noisy projection data. Nevertheless, the results do demonstrate the feasibility of motion compensation of awake rat studies based on markerless motion estimation. We are currently in the process of implementing a fully 3D list mode reconstruction with motion compensation to reconstruct these data using all collected events. Improved outlier rejection techniques are also being considered.

IV. Conclusions

We have presented the first demonstration of motion-compensated awake animal PET based on markerless motion tracker estimates. A markerless approach to tracking considerably simplifies conscious animal experiments involving motion tracking and has the potential to improve both the accuracy of motion measurements and extend the range of detectable motion. Development of a fully 3D list mode reconstruction with motion compensation is currently underway to properly utilize all acquired events.

Acknowledgment

The work was supported by Australian Research Council Discovery Grant DP0663519.

References

[1] A. Weisenberger, B. Kross, S. Majewski, J. McKisson, V. Popov, J. Proffitt, A. Stolin, et al., "Awake Animal SPECT: Overview and Initial Results", *Proc. 2008 IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, pp. 5588-91, 2008.

[2] A. Kyme, V. Zhou, S. Meikle, K. Popovic, J.-P. Man, M. Akhtar, I. Karllsson, R. Fulton, "Motion tracking of fully conscious small animals in PET," *Proc. 2009 IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, pp. 2561-66, 2009.

[3] A. Kyme, S. Se, S. Meikle, C. Baldock, W. Ryder and R. Fulton, "Novel SLAM-Based Markerless Motion Tracking of Conscious Unrestrained Rodents in PET," Proc. 2011 IEEE Nucl. Sci. Symp. Med. Imaging Conf., pp. 3554-3557, 2011.

[4] A. Kyme, V. Zhou, S. Meikle, C. Baldock and R. Fulton, "Optimised Motion Tracking for Positron Emission Tomography Studies of Brain Function in Awake Rats," *PLoS ONE*, vol. 6, no. 7, 2011: e21727. doi:10.1371/journal.pone.0021727.

[5] T. Svoboda, D. Martinec and T. Pajdla, "A convenient multi-camera self-calibration for virtual environments," *PRESENCE: Teleoperators & Virtual Environments,* vol. 14, pp. 407-422, 2005.

[6] J-Y. Bouguet, *Matlab Calibration Toolbox*. URL: http://www.vision.caltech.edu/bouguetj/calib_doc/

[7] D. Lowe, "Object recognition from local scale invariant features," *Proc. Int. Conf. Comp. Vis.*, pp. 1150-1157, 1999.

[8] V. Zhou, A. Kyme, S. Meikle and R. Fulton, "An Event Driven Motion Correction Method for Neurological PET Studies of Awake Laboratory Animals," *Mol. Imag. Biol.*, vol. 10, pp. 315-324, 2008.

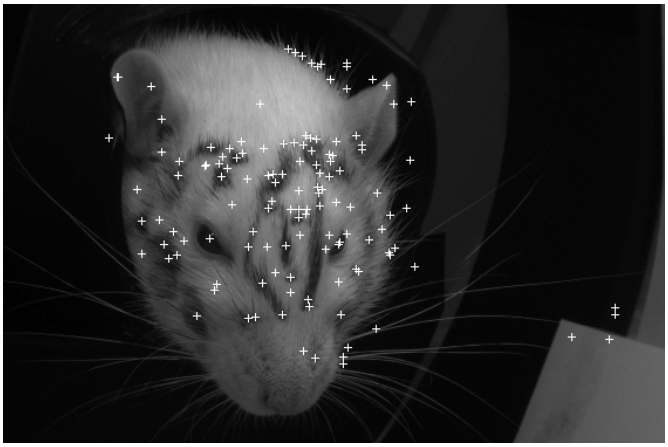

Fig. 5. SIFT features (indicated by the white crosses) detected in an image of the rat's head. Indelible markings on the face increased the number of features detected on rigid parts of the head.

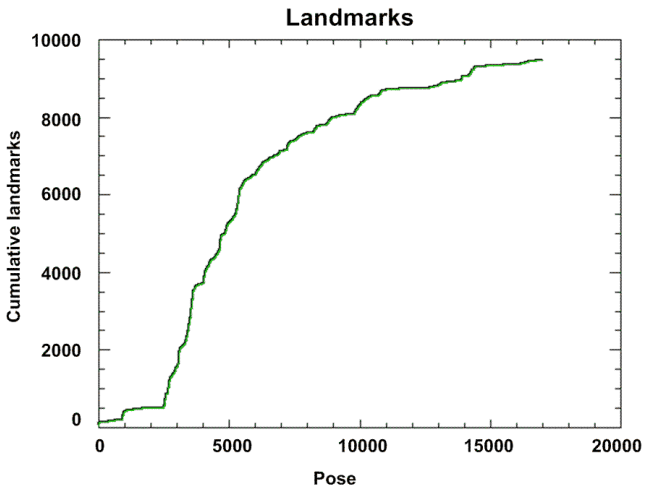

Fig. 6. Cumulative number of 3D head landmarks as a function of pose number.

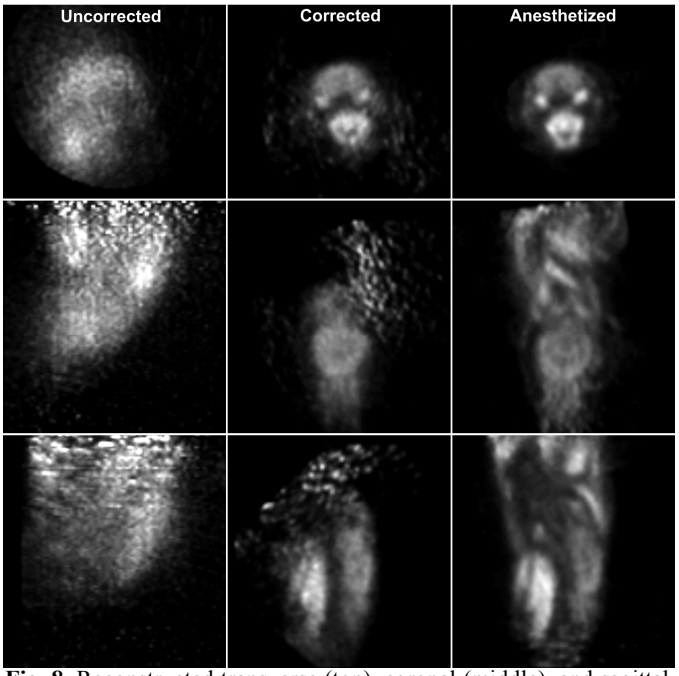

Fig. 8. Reconstructed transverse (top), coronal (middle), and sagittal (bottom) slices for the uncorrected (left), motion-compensated (middle) and no motion (i.e. anesthetized animal) (right) cases.

Fig. 7. Top: one component (*x*-axis rotation) of the estimated motion over the 17000 frames of the awake rat study. Bottom: images of the rat (taken using camera 1) at various stages in the study. The image borders are color-coded with the vertical bars drawn on the graph to indicate where in the sequence the image occurred. Overlaid on these images is a green triangle which, for reliable pose estimates, should connect the two eye centers and nose tip. See main text for details.