Digitized Video Subject Positioning and Surveillance System for PET

Y. Picard and C.J. Thompson
Medical Physics Unit and Montreal Neurological Institute
McGill University, Montreal, Quebec

Abstract

Head motion is a significant contribution to the degradation of image quality of Positron Emission Tomography (PET) studies. Images from different studies must also be realigned digitally to be correlated when the subject position has changed. These constraints could be eliminated if the subject's head position could be monitored accurately.

We have developed a video camera-based surveillance system to monitor the head position and motion of subjects undergoing PET studies. The system consists of two CCD (charge-coupled device) cameras placed orthogonally such that both face and profile views of the subject's head are displayed side by side on an RGB video monitor. Digitized images overlay the live images in contrasting colours on the monitor. Such a system can be used to (1) position the subject in the field of view (FOV) by displaying the position of the scanner's slices on the monitor along with the current subject position, (2) monitor head motion and alert the operator of any motion during the study and (3) reposition the subject accurately for subsequent studies by displaying the previous position along with the current position in a contrasting colour.

I. INTRODUCTION

A serious limitation to the ability to quantify PET studies is the ability to ensure the subject remains still [1,2,3,4]. Most PET studies are either dynamic ones lasting about an hour, or test-retest sequences. In either case the image variability can be reduced if the subject remains completely immobile or if he can be repositioned at exactly the same position for subsequent studies. Presently, most PET centres use head restraints. These grip the subject's head tightly and make head rolling or axial movements virtually impossible. Nevertheless, they do not prevent nodding movements and do not guarantee that the head will be at the same position for subsequent studies.

To monitor head movements, four video monitoring techniques for subjects undergoing PET studies have previously been reported [5,6,7,8,9,10]. None of these monitoring techniques can be used to position the subject's head within the FOV or to reposition it accurately for subsequent studies. Algorithms to align PET images [11,12] are thus required to correlate the different studies when the subject's head position has changed. A video subject repositioning system has also been reported previously [13] but it was not designed to monitor subject movements.

We have developed and built a video camera-based surveillance system that can both position and monitor the head of subjects undergoing PET studies. The positioning of the subject's head in the FOV of the scanner is achieved by displaying the position of the scanner's slices on the monitor along with the current head position. Repositioning is achieved by displaying the subject's previous position along with the current position in a contrasting colour. Position monitoring during the studies is achieved by tracking three light emitting diodes (LEDs) fixed on the subject's face. The video system is described in detail below.

II. METHODS

A. Layout of the Video System

Two black and white CCD cameras are positioned orthogonally on the gantry of the PET scanner, a Scanditronix PC-2048B [14], so that one camera gives a face view of the subject and the second, a profile view (figure 1). The face view camera (camera 1) is positioned such that the subject's face can be seen entirely on the

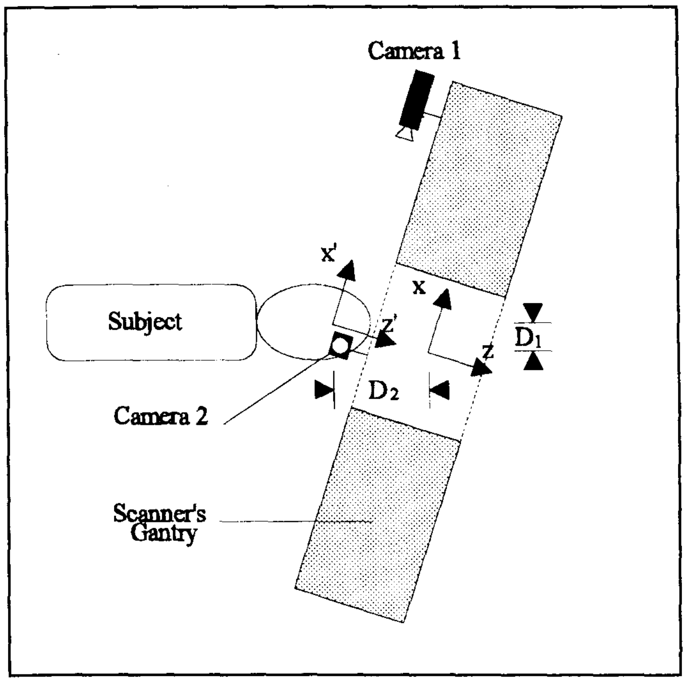

Fig. 1: Side view of the scanner gantry showing the coordinate system of the scanner (x-z) and the cameras (x'-z'). The y and y' axes emerge from the page. Camera 2 is behind the subject's head. The cameras are positioned orthogonally on the gantry such that the head is located in the FOV of the cameras. Subject positioning is done when the subject's head is outside the scanner FOV and the gantry is tilted to its scanning position as shown in this figure. Once the positioning is done, the couch is moved gently down by a distance D_1 and moved horizontally within the scanner by a distance D_2 such that the origin of the cameras' FOV coincides with the origin of the scanner.

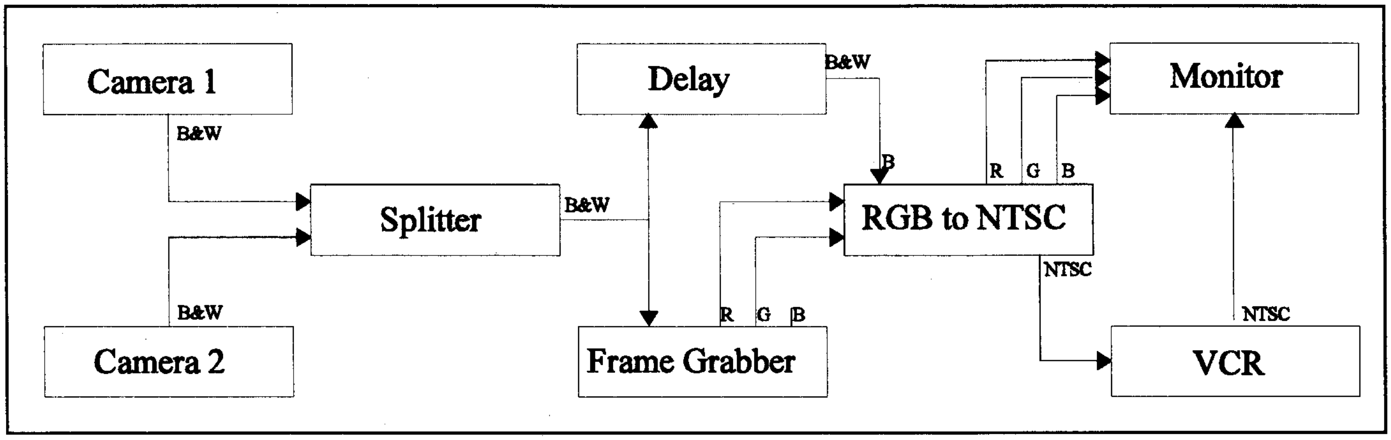

Fig. 2: Diagram of the video system. The connections are made using 75 ohm cables. R = red, G = green and B = blue.

left half of the camera output. The side view camera (camera 2) is positioned such that the head can be seen on the right half of the camera output. The FOV of both cameras are located outside the gantry, at a fixed distance from the FOV of the scanner.

Figure 2 shows a diagram of the video system connections. Both CCD cameras are connected to a screen video splitter/inserter which segments the screen images coming from the two cameras. The segmentation is set up such that the left half of the display gives the output of camera 1 and the right half, the output of camera 2. The splitter output is sent to a frame grabber board on a Vaxstation computer and also to the blue input of an RGB to NTSC (North America Television Standards Council) converter via a 1/60th of a second delay. The delay compensates for the time the frame grabber takes to digitize the incoming images. It produces a signal coming from the splitter which is synchronized with the output of the frame grabber. The frame grabber continuously digitizes the incoming images and outputs these images or user-selected digitized images into RGB signals. The red and green outputs of the frame grabber (its blue output is not connected) and the blue output of the splitter are then sent to the monitor inputs via the RGB to NTSC converter. The NTSC output of the converter is connected to the monitor via a video cassette recorder (VCR) used to record or view recorded images. A switch on the monitor allows viewing of the NTSC image from the VCR or the live RGB image.

Table 1: Equipment

Cameras	Panasonic B&W CCD WV-BL200
Lenses	Cosmicar auto-iris C6Z1218M2ESA F1.8/12.5-75mm
Lens controller	Pelco MLZ6DT
Screen Splitter	Pelco VSS100DT
Frame Grabber	Data translation DT2651
Computer	Micro VAX II
RGB to NTSC Converter	Panasonic UTP-2
VCR	Panasonic AG-1950
Monitor	Panasonic CT-1400MGC

Table I lists the equipment used. The cameras, the lenses and the splitter are the same as those used for surveillance systems in public places. The minimum requirements for the frame grabber are the following: be able to store digitized images in a 512x480 by 8-bit speed memory-mapped frame store buffer and have an output look-up table (LUT) and at least 2 8-bit RGB D/A (digital to analog) converters. Most high resolution frame grabbers available today meet these minimum requirements. The RGB to NTSC converter and the VCR are used only to confirm subject movement during a scan and are not essential components of the system.

The display on the monitor screen depends on the frame grabber setup. When the monitor displays RGB signals, the video system provides a continuous live image of the subject to the monitor and the frame grabber. Meanwhile, the frame grabber continuously digitizes the incoming live signal in a 512x480 8-bit buffer at a rate of 30 images per second. If the frame grabber is set to output a blank image, the resulting image on the monitor screen is a blueish live image of the subject; if the frame grabber is set up to output the digitized live image, the resulting image on the monitor screen is a black and white live image of the subject (assuming a proper output LUT); if the frame grabber is set up to output a stored digitized image, the resulting image on the monitor screen is then a mixture of the blueish live image of the subject with a reddish and/or greenish image of the digitized image depending on the LUT setting. This latter setup is the most useful setup for subject positioning and repositioning. Any image displayed on the monitor can easily be recorded on a VCR cassette. The monitor switch is set to NTSC signal only to view recorded images on the VCR cassette.

B. Calibration of the System

The system is calibrated by placing a block of known dimensions so that one corner of the block corresponds to the position of the origin in the cameras' FOV and the edges of the blocks are parallel to the cameras' right-handed coordinate system. We used a 10cm x 10cm x 10cm block for calibration. Figure 1 illustrates the coordinate system of the cameras' FOV. The z'-axis is parallel to the scanner axis and the x'- and z'-axes define a vertical plane which includes the origin of both the scanner and the cameras' FOV. The y'-axis is normal to that vertical plane. The distances D_1 and D_2 are then measured using the scanner's lasers, a ruler and a level. They correspond to the vertical and horizontal distances between the scanner's origin and the cameras' FOV origin.

Figure 3 shows the geometry used to calibrate the cameras for the x- and y-axes, that is to find out the relationship between the pixels on the monitor screen and the coordinates of the corresponding points in the cameras' FOV. Since the block is orthogonal to the cameras, the resulting image gives two rectangles corresponding to the surfaces of the block facing the cameras. All edges of both rectangles are located in a pixel line or column. This image of the block is first digitized by the frame grabber. The digital image is then treated as if it was located at an arbitrary distance (d_1 and d_2 in figure 3) from the origin. To calibrate the camera for the x- and y-axes, four imaginary planes are computed. Each of these planes is defined by a edge of the block (on the surfaces facing the cameras) and its image on the pixel plane as shown in figure 3. These planes are defined in figure 3 by the lines acg, adj, beg and bfh. The intersection of these planes gives the focal lines of the camera for the x- and y-axes (points a and b on the figure). Using similar triangles (acd≈agj and bef≈bgh), the focal distances f_1 and f_2 and the focal offset o_1 and o_2 (distance between the origin and the focal line in the x direction) are calculated. The same technique is used to calibrate the z-axis. At this point, for each point in the cameras' FOV it is possible to calculate the corresponding pixel pair on the monitor screen (one pixel for each camera) and vice-versa.

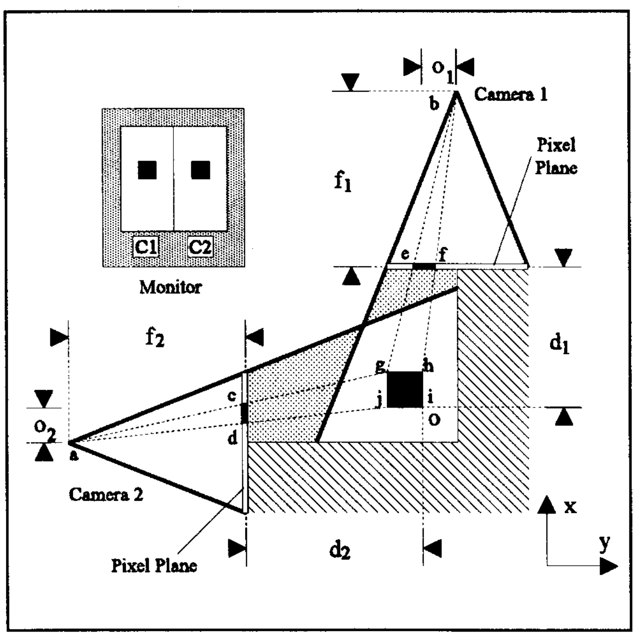

Fig. 3: Geometry used to calibrate the cameras. A block of known size is placed in the cameras' FOV. The resulting image is shown on the monitor. C1 and C2 correspond to images from camera 1 and 2 respectively. The hatched region is not displayed on the monitor (because of the video splitter) and is outside the FOV of the cameras. The dotted regions are seen by only one camera and thus are outside the FOV.

C. Subject Positioning

To position the subject in the scanner the frame grabber outputs a green grid corresponding to the location of the scanner's 8 direct slices which overlays the blueish live face and profile views of the subject's head. The grid is built up by determining the pixel pairs corresponding to a set of imaginary parallel rings positioned along the scanner's axis. Their origin coincides with the scanner's axis. Each ring lies axially in the centre of the corresponding slice. Each of the ring diameters can be adjusted so that they seem to be on the subject's head. The pixel pairs corresponding to all points along the ring are computed using d_1, d_2, f_1, f_2, o_1, o_2 and the ring diameter. These curves correspond to the half-ring closest to the camera. The other half is not drawn because it will likely be located behind the subject's head. With this green ring and the blueish live image on the monitor, the operator can easily move the subject's head to the desired position. The head position is then digitized by the frame grabber and stored in a file. The couch is afterward moved down and horizontally within the scanner's gantry by the distances D_1 and D_2 respectively (see figure 1).

D. Subject Monitoring

We assume that the subject's head is a solid body. It has six degrees of freedom (3 rotations and 3 translations) and three independent point coordinates on the body are necessary to specify its position in space at any time. Three LEDs are placed on the subject's head such that they can be seen by the two orthogonal CCD cameras: one on the tip of the nose, one between the two eyes and one between an eye and an ear for example. The frame grabber is set up to continuously digitize the live image and to refresh the frame buffer every 1/30 sec. The LEDs are tracked by extracting from the frame buffer 6 small 6x6 pixel matrices centred on the 6 LEDs positions of the previous frame. The centroid of each small matrix is then computed for the current frame and the matrix centre is updated if the position of LEDs has changed. The positions of the LEDs are computed using the calibration values (see above).

At the beginning of a scan the initial position coordinates of the three LEDs are stored in three vectors X_{0i}, that is:

$$X_{0_i} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}_i , \quad i = 1{,}2 \text{or} 3. \tag{1}$$

At any time during the scan, the positions of the three LEDs are stored in a matrix X_i, that is

$$X_i = \begin{pmatrix} x \\ y \\ z \end{pmatrix}_i , \quad i = 1{,}2 \text{or} 3. \tag{2}$$

The relationship between X_i and X_{0i} is

$$X_i = RX_{0i} + T \tag{3}$$

where T is the translation vector of the origin of the coordinate system fixed on the head,

$$T = \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} \tag{4}$$

in which d_x, d_y and d_z represent the translations along the scanner's coordinate system and R is a rotation matrix,

$$R = R_z R_y R_x \tag{5}$$

where

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix}$$
$$R_y = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \tag{6}$$
$$R_z = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In the rotation matrices, ϕ is the yaw angle, i.e. the rotation about the x-axis, θ is the pitch angle, i.e. the rotation about the intermediate y-axis (fixed on the head) and ψ is the roll angle, the rotation about the final z-axis (fixed on the head). The head displacement is completely described by the three translations of T and the three rotation angles of R. They are computed iteratively using a globally convergent Newton minimization routine [15] by minimizing the distances between the LEDs measured positions and the ones computed using equation (3).

E. Subject Repositioning

The subject's head can be repositioned to the position of the previous scan by retrieving the digital image of the previous scan position and storing it in the frame grabber's 512x480 8-bit buffer so that the pixel intensities are all between 0 and 254. The grid is then stored in the frame grabber buffer by setting the appropriate pixels to 255. The frame grabber output LUT is set up to produce a green pixel if the pixel intensity is 255 and to produce a yellow pixel of increasing intensity otherwise. With the incoming blueish live image of the subject's head, the resulting image on the monitor screen is a green grid overlaying a black and white image of the subject's head surrounded by yellow and blue fringes. The subject's head can be moved in space until the fringes disappear. The subject's head is then located at its original position.

F. Experiments

The system was tested by a healthy subject whose head was supported in a foam mould. These foam moulds are used routinely at the Montreal Neurological Institute. They are moulded on every subject undergoing PET studies. The subject had three LEDs attached to his face: one on the nose at about 1 cm from the tip of the nose, one between the eyebrows and one on the right temple. The subject was positioned in the FOV using the grid displayed on the monitor (see II.C). The scanner gantry was set to scan 20-degree sections of the brain. The subject was asked to remain still for 30 minutes while the video system was tracking the three LEDs to obtain the 3 translations and the 3 rotations corresponding to his motion. These were stored in a file every second. To test the efficiency of the head restraint, the subject was re-monitored for another 30 minutes but this time without the head foam. He was repositioned using the technique described in II.E.

To determined the noise of our system, the subject was replaced by a dummy. The dummy's head had also 3 LEDs positioned at approximatively the same locations as for the subject. The dummy's head was positioned at the subject initial position and monitored for 30 minutes.

III. Results and Discussion

Figure 4 shows two black and white pictures of the colour monitor of the video system when a subject is positioned in the FOV. The curves represent the half rings (their radius was set to 8.5cm) corresponding to the positions of the axial centres of the direct slices in the scanner. They are green on the colour monitor. Figure 4(a) corresponds to the monitor display when the subject is repositioned in the FOV. The image is an overlay of the digitized initial position image and the live image. Blue and yellow fringes (the fringes are grey on figure 4(a)) appear around and on the subject's head until his position coincides with his initial position as shown in figure 4(b). The resulting image on the colour monitor is a black and white image of the subject overlaid by a green grid. The repositioning operation required some skill since the head can be moved in 6 degrees of freedom. Nevertheless, when all the blue and yellow fringes vanish, the operator can be sure that the patient was repositioned at the initial position. Sometimes fringes are still visible on the monitor screen when the subject is back to its initial position. This is the case if the patient facial expression is different from when the initial position was digitized or if he had a haircut between the two scans. In these cases, the operator will have to rely exclusively on the fringes produced by the edges of the head and the nose, for example.

Displaying the grid corresponding to the location of the 8 direct slices of the scanner is also found to be useful to position the subject's head within the FOV of the scanner in order to decide on the slice position. During this operation, the resulting image on the monitor screen is similar to figure 4(b) except that on the colour monitor, the subject is in black and blue and the grid is yellow. The best results in subject positioning are obtained when the diameters of the rings corresponding to the slices are set to the diameter of the subjects' head (about 17cm).

Fig. 4: Black and white pictures of the colour monitor of the video system when a subject is repositioned in the FOV. The curves represent the half rings corresponding to the positions of the axial centres of the direct slices in the scanner. In (a), the subject is not at his initial position. Fringes appears around and on his face. In (b), the subject is back into his original position. All the fringes vanished. The fringes are easier to distinguish on the colour monitor than on the black and white picture of (a).

Figure 5 shows the RMS displacements per minute for different regions of the brain of the healthy subject who was monitored for 30 minutes with the head foam (white circles) and without the head foam (black squares). The RMS displacement is based on the average position of a point in the region during the first minute of each monitoring experiment. The initial position coordinates (in centimetres) of the brain regions using the coordinate system of figure 1 with the gantry tilted at 20 degrees were (0,0,0) for the thalamus, (7,0,3) for the superior frontal gyrus, (0,5,1) for the superior temporal gyrus and (-7,0,1) for the lateral occipital gyrus. The displacements were computed using the 6 rotation and translation data stored in the files. The solid line at the bottom of each graph corresponds to the maximum displacement measured using the dummy at each position. This is equivalent to the noise of the video system. On all graphs we notice that the subject remained remarkably still throughout the 30 minutes of monitoring since the RMS displacement is almost always less than 2mm. The RMS displacement has a tendency to increase with time. There is no significant difference in the amplitude of motion for the studies with and without the head foam to justify the cost of the head foam (about $35.00US) when the healthy subject's only task is to remain still. Motion for subject undergoing other tasks as well as remaining still are expected to be more important.

The pixel width of the video system is evaluated during the calibration procedure. Because the cameras use lenses, the pixel width changes with the position in the cameras' FOV. The pixel width is about 0.07cm in all x y and z directions on the surface of the subject's head. Thus the resolution of the video system is about 10 times better than the PET scanner, so feedback from the video system provides information about the possibility of blurring errors due to residual movements of the subject.

When the video system is set to monitor head motion with 3 LEDs attached to the head, the translation and rotations angles are computed every 60 msec on a microVAX II.

IV. CONCLUSION

We have described a video system which can both monitor head movements and position the head in the scanner FOV of a PET system. This system has applications in any imaging modalities where the scanning time is long compared with the normal time one can reasonably expect subjects to stay still. It is also useful for repositioning the head of the subject to its initial position. Images can thus be correlated without an algorithm for aligning and reslicing the images. We have also shown that if the only task of a healthy subject is to remain still for 30 minutes, a head restraint is not required. Further tests using a translation stage will be performed in a near future to establish the translational and angular precision of the video system.

V. ACKNOWLEDGMENTS

This work is supported by General Electric Medical Systems (Pet Engineering) and MRC grant #SP30. Thanks to Ray Clancy who volunteered to be the subject to test the system.

VI. REFERENCES:

[1] JM Links. "The Influence of Positioning on Accuracy and Precision in Emission Tomography." Editorial in: *J. Nucl. Med.* **32**:1252-3 (1991).

[2] M McCord, SL Bachararch, RO Bonow, V Dilsizian, A Cuocolo and N Freedman. "Misalignment between PET Transmission and Emission Scans: its effect on Myocardial Imaging," *J. Nucl. Med.* **33**:1909-14 (1992).

[3] BD Pate, BJ Snow, KA Hewitt et al.. "The Reproducibility of Striatal Uptake Data obtained with Positron Emission Tomography and F-18 fluro Dopa Tracer in non-human Primates." *J. Nucl. Med.* **32**:1246-51 (1992).

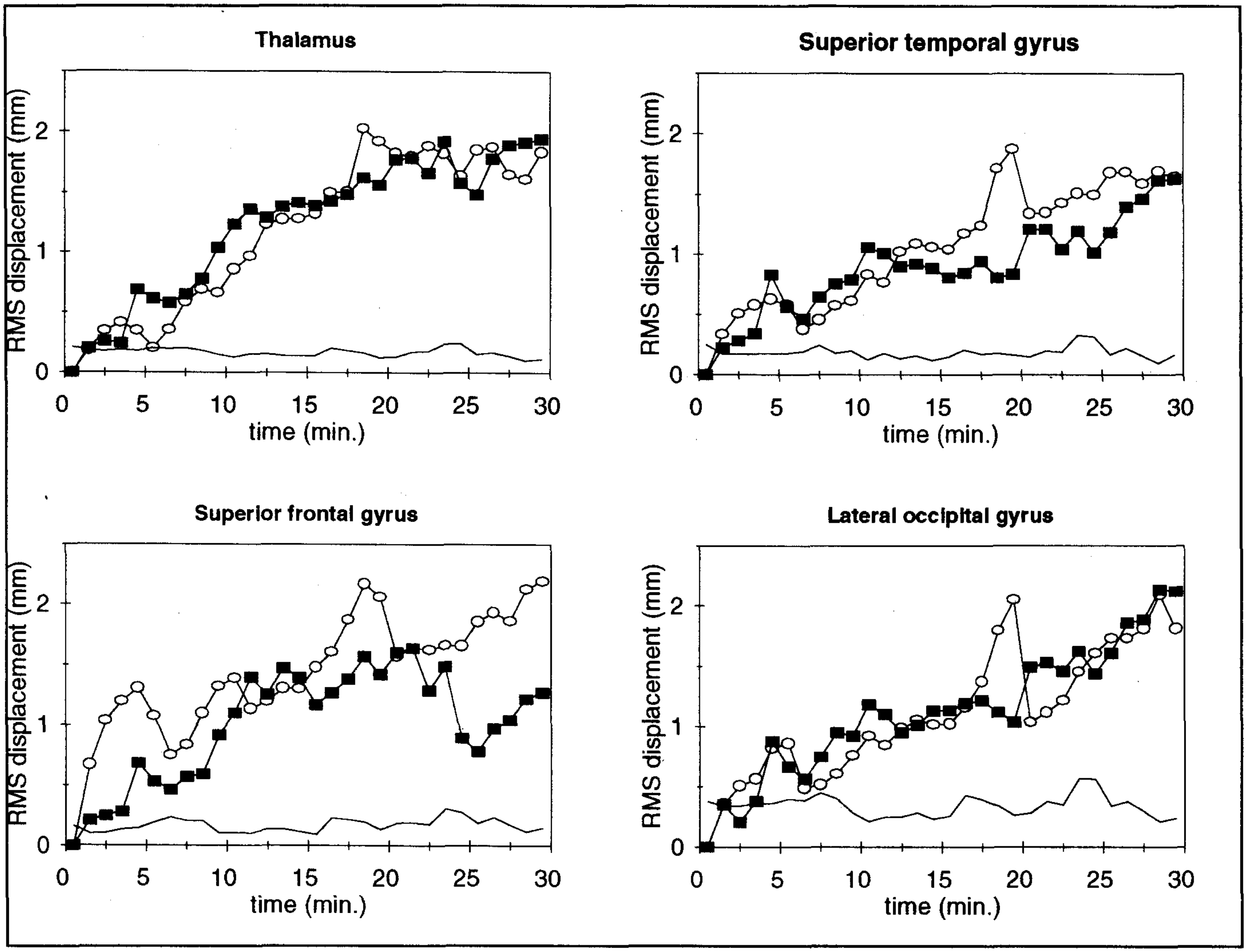

Fig. 5: RMS displacements per minute for different regions of the brain of a healthy subject who was monitored for 30 minutes with the head foam (white circles) and without the head foam (black squares). The solid line at the bottom of each graph is the noise.

[4] AN Bice, PJ Keller and BP Drayer, "Simulation Analysis of Movement Induced Artefacts in Neurological PET," Abstracts of the IEEE 1993 Nuclear Science Symposium and Medical Imaging Conference, San Francisco, p.142 (1993).

[5] S Watanuki, T Ido, T Matsuzawa, "Movement Monitor for Patient's on Positron Emission Tomography," CYRIC Annual Report, Tohoku University, Sendai Japan (1984).

[6] A Ranicar, Video substraction monitoring on the Siemens/CTI 931 and 953B, personal communication, July 1992.

[7] Brian Reece Scientific Ltd, Real time video subtraction system supplied to the MRC cyclotron unit, Hammersmith hospital, for patient position monitoring/recording, Advertising brochure, July 1992.

[8] M Daube-Witherspoon, YC Yua, MV Green and RE Carson, "Correction for motion distortion in PET by dynamic monitoring of patient position," *J. Nucl. Med.* **32**:816 (1991).

[9] SR Goldstein, ME Daube-Witherspoon and MV Green, "A Device for Continuous Measurement of Head Position during PET Brain Imaging," *J. Nucl. Med.* **35** (1994).

[10] M Menke, KR Buckley and MS Atkins, "A Software Compensation Method for Videometrically Detected Motion during Head PET Scans," paper number 8M10, presented at this conference.

[11] RP Woods, SR Cherry and JC Mazziotta, "Rapid Automated Algorithm for Aligning and Reslicing PET Images," *J. Comp. Assist. Tomogr* **16**:620-33 (1992).

[12] MA Mintum and KS Lee, "Mathematical realignment of paired PET images to enable pixel-by-pixel substraction," *J. Nucl. Med.* **31**:816 (1990).

[13] J Conti, MFD Deck, DA Rottenberg, "An Inexpensive Video Patient Repositioning System for Use With Transmission and Emission Computed Tomographs," *J. Comp. Assist. Tomogr* **6**:417-21 (1982).

[14] AC Evans, CJ Thompson, S Marrett and E Meyer, "Performance Evaluation of the PC2048, a new 15-slice encoded-crystal PET Scanner for Neurological Studies," *IEEE Trans. Med. Imag.* **10**:90-8 (1990).

[15] Press WH et al., *Numerical Recipies in FORTRAN: the art of scientific computing,* 2nd ed., Cambridge University Press, Cambridge, 376-85, (1992).