Toward a Robust System to Monitor Head Motions during PET Based on Facial Landmark Detection: a New Approach

Sergey Anishchenko[1], Vladislav Osinov[1], Dmitry Shaposhnikov[1],
Lubov Podlachikova[1], Richard Comley[2], Xiaohong W. Gao[2*]

[1] A.B.Kogan Research Institute for Neurocybernetics, Southern Federal University,
Rostov-on-Don, Russia, dima@nisms.krinc.ru

[2] School of Computing Science, Middlesex University, London, NW4 4BT, UK,
** Corresponding Author, x.gao@mdx.ac.uk*

Abstract

A new approach for the detection of head motions during PET scanning is presented. The proposed system includes 4 modules, which are: input module, face segmentation, facial landmark detection, and head movement estimation. The developed system is tested on pictures monitoring a subject's head while simulating PET scanning (n=12) and face images of subjects with different skin colours (n=31). Experimental results show that the centres of chosen facial landmarks (eye corners and middle point of nose basement) can be detected with high precision (1±0.64 pixels). Processing of 2D images with known moving parameters demonstrates that the parameter movement in terms of rotation and translation along X, Y, and Z directions can be obtained very accurately via the developed methods.

1. Introduction

Positron emission tomography (PET) data acquisition is a relatively lengthy procedure (~ 1 hour), making it difficult for a subject to stay still during the entire brain scanning. Head motion can degrade the quality of PET studies, because even relatively small motions may significantly compromise the resolution [1-5]. Furthermore, head motion also causes misalignment between the emission and transmission of scan data, leading to erroneous correction for photon attenuation [5]. Hence, motion tracking and correction is necessary to ensure that quantitative PET data are obtained as precisely as possible. The existing methods [2-4] to reduce the degrading effects of motion fall into several categories: image realignment in frame-mode or list-mode acquisition, optical tracking systems, and their combinations. Until now, the development of an effective method to detect head movements is an active research area for medical practice [2].

In this study, a robust method to monitor head movement is proposed. It is based on two biologically motivated models. One is CIECAM97 [6] for measuring colour appearance invariantly to illumination conditions. The other is a Behaviour Model of Vision (BMV) simulating some mechanisms of the human visual system for perceiving shapes [7]. These models are used for colour segmentation of the facial area on the original recorded pictures and for the detection of facial landmarks (i.e. eye corners and middle point of nose basement), respectively. The basic algorithms have been developed and undergone preliminary testing [8]. Their integration and overall system architecture are presented in this paper.

2. System architecture

The proposed system includes four basic modules: (1) input module; (2) face segmentation module; (3) facial landmarks detection module; (4) module for head movement estimation. These modules are implemented as a set of devices (the first one), software packages (the second and the third ones), and a set of algorithms (the fourth one).

2.1. Input module

The input module consists of two off-the-shelf digital cameras, Canon EOS-1D Mark II, to monitor the head movement of a subject undergoing a PET brain scan.

Before shooting, the cameras and colour monitors are calibrated (Figure 1).

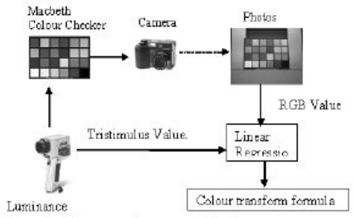

Figure 1. The steps of camera calibration.

First, the camera takes pictures of 24 colour patches from a Macbeth Colour Checker. These images are then transferred to a calibrated colour monitor. The 24 colour patches are then measured using a colour meter, CS-100A, to obtain their CIE XYZ values. The two sets of XYZ values are then used to determine the camera calibration model using the least squares method.

In order to ensure colour consistency, the colour monitor needs to be calibrated. This is done using optical software with a Spyder sensor. When an image is downloaded to a computer, it is represented in an RGB space. To convert an RGB space to the CIE standard XYZ space, equations (1) are applied for average daylight with CIE standard illuminant D65 as white reference, i.e., $[X_W, Y_W, Z_W] = [0.95045\ 1.0\ 1.088754]$.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (1)$$

2.2. Face segmentation module

The colour range in terms of CIECAM attributes of hue, chroma, and lightness are obtained first from different subjects of various skin colours. These ranges are then applied as thresholds to segment the head from the rest of the picture.

2.3. Facial landmarks detection module

A feature description of each facial landmark (Figure 2) is formed by a space-variant input window and represented by a vector in multi-dimensions. Each component of the feature vector is in line with the orientation of a local “colour” edge detected from the colour attributes. This is done by convolving local image fragments with a set of 16 kernels that is determined by differences between two oriented Gaussians with shifted kernels [8].

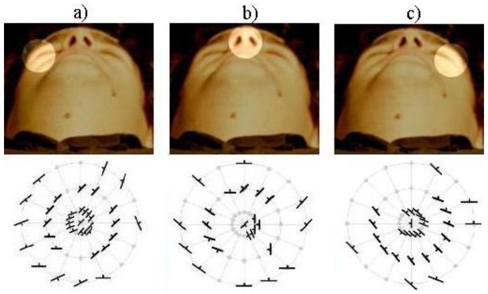

Figure 2. Feature vectors of facial landmarks for a subject: (a) right eye corner, (b) middle point of nose basement, (c) left eye corner.

Initial estimation of each chosen landmark is obtained by a single positioning input window in the corresponding region, one image for each subject, and works as a seed feature vector.

2.4. Head motion estimation module

In this module, positions of facial landmarks, as well as the linear and angular relations between them, are calculated for each picture of the whole video clip. In addition, comparison of geometric parameters between landmarks identified in consecutive images is performed, which is shown in Figure 3.

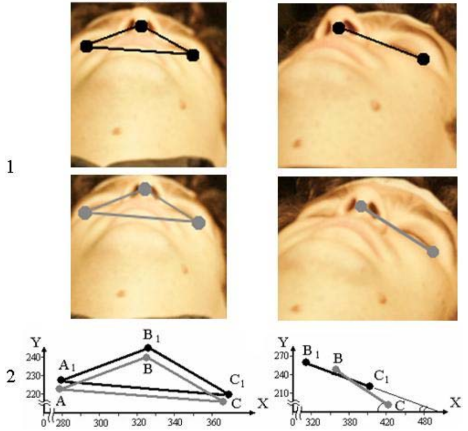

Figure 3. Spatial relations (1) between facial landmarks detected on consecutive pictures and their schemes (2): a) pictures captured by frontal camera (A,B,C and A_1,B_1,C_1 - landmarks identified); b) pictures captured by the left camera.

3. Computer simulation

Two sets of image data have been employed in the computer simulations. The first image collection (n=12) with a subject lying down in the PET scanner includes the pictures with known subject head positions and illumination conditions. All pictures monitoring a subject's head have the same dimension (640x427 pixels) and are recorded by calibrated cameras. The second collection contains pictures of human faces with different skin colours (n=31). This collection has various dimensions and is recorded by cameras without any calibrations.

During the computer simulation, face areas were correctly segmented on all images in the first collection without false positive regions. Furthermore, all the visible facial landmarks (n=30) have been identified with very high precision (1±0.64 pixels). After the detection of facial landmarks, the movement parameters including translation and rotation along each of X, Y, and Z axes can be determined by comparison of landmarks from two different images of the same subject undergoing the same scanning. Estimations on 2D images with known moving parameters show that movement parameters can be obtained very accurately. For example, the rotation angle for the two pictures presented in Figure 3b is equal to 10°. While processing the pictures from the second collection, many false positive regions were identified (n=25) outside face areas (n=31), arriving at only 56 of 88 visible facial landmarks being correctly detected.

A software module simulating the PET scanning situation has been developed for quantitative estimation of different factors contributing to the system performance (Figure 4). It includes a 3D model of a head that may be rotated along any axis with arbitrary angle. Preliminary tests indicate that it is possible to achieve one degree precision of rotated angle based on the estimation of facial landmark parameters. The simulator will give an opportunity to develop correction algorithms for the simultaneous processing of the pictures taken by the frontal and the left side cameras.

Figure 4. Head movement simulation, a) frontal camera, б) left side camera.

4. Conclusion

In this study, a new robust approach to monitoring human head motion during PET scanning has been presented. Computer simulation results have shown very high performance of the developed system in processing the pictures captured by calibrated cameras. In particular, centres of chosen facial landmarks, including eye corners and middle point of nose basement, are detected with very high precision. Preliminary estimations on 2D images with known motions show that movement parameters (i.e. rotation and translation along X, Y, and Z axes) can be obtained very accurately via analysis of spatial and angular relations between identified facial local landmarks. It is anticipated that future modifications will improve the overall system performance.

Acknowledgments: This work is supported under the TIME project, funded by the European Commission Asia-IT&C program (CN/ASIA-IT&C/009(91450)), and a BRIDGE Research Cooperation Award funded by the British Council, on Head Motion Detection.

5. References

[1] B.F. Hutton, A.Z. Kyme, Y.H. Lau, D.W. Skerrett, R.R. Fulton, "A Hybrid 3-D Reconstruction/registration Algorithm for Correction of Head Motion in Emission Tomography", IEEE Trans Nucl Sci., 2002, pp.188-194.

[2] A. Rahmim. "Advanced Motion Correction Methods in PET. Review Article", Iran J. Nucl. Med., v. 13, no. 241, 2005, pp.1-17.

[3] L. Tellmann, R. Fulton, U. Pietrzyk, I. Nickel, I. Stangier, O. Winz, H. Herzog, "Concepts of Registration and Correction of Head Motion in Positron Emission Tomography", Z Med Phys., v. 16, no. 1, 2006, pp. 67-74.

[4] H. Watabe, K. Koshino, P.M. Bloomfield, R.R. Fulton, H. Iida, "Development of Motion Correction Technique for PET Study Using Optical Tracking System", Int. Congress Ser., 1265, 2004, pp. 31 – 38.

[5] H. Zaidi, M.-L. Montandon, S. Meikleb, "Strategies for Attenuation Compensation in Neurological PET Studies. Review", NeuroImage, 2007, pp.518 – 541.

[6] M.R. Luo, and R.W.G. Hunt, "The Structure of the CIE 1997 Colour Appearance Model (CIECAM97s)", Color Res. Appl., no. 23, 1998, pp. 138-146.

[7] I. Rybak, V. Gusakova, A. Golovan, L. Podladchikova, N. Shevtsova, "Attention-guided Recognition Based on «What»; and «Where» Representations: A Behavioral Model", Neurobiology of Attention, 2005, pp. 663-670.

[8] X.W. Gao, S. Anishenko, D. Shaposhnikov, L. Podlachikova, S. Batty, and J. Clark. "High-precision Detection of Facial Landmarks to Estimate Head Motions Based on Vision Models", J. of Comp. Sci., v. 3, no. 7, 2007, pp.528-532.