Toward a Robust System to Monitor the Head Motions During PET Based on Facial Landmarks Detection: a New Approach

Sergey Anishenko[1], Vladislav Osinov[1], Dmitry Shaposhnikov[1],
Lubov Podlachikova[1], Stephen Batty[2], Richard Comley[3], Xiaohong W. Gao[3]

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

[2]*MRC Prion Unit 3rd Floor, ION - Queen Square House, Queen Square, London, WC1N 3BG s.batty@prion.ucl.ac.uk*

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

Abstract

A new approach for detection of head motions during PET scanning is presented. The proposed system is based on two vision models (CIECAM'97 and BMV) and included 4 modules: (1) input module consisting of two calibrated cameras; (2) face segmentation module; (3) local facial landmarks detection module; (4) module for head movements estimation. The developed system is arrived on the pictures monitoring a subject's head while simulating PET scanning (n=12) and face images of the human with various skin colour (n=31). It is shown that the centres of chosen facial landmarks (eye corners and middle point of nose basement) have been detected with high precision (1±0.64 pixels). The results on 2D images with known moving parameters show that movement parameters of rotations and translations along X, Y, and Z directions can be obtained very accurately via the developed methods. Future steps to develop the proposed system have been considered.

1. Introduction

Positron emission tomography (PET) data acquisition is a relatively lengthy procedure (~ 1 hour) and it is difficult for a subject to stay still during brain scanning. Consequently, head motion can significantly degrade the quality of PET studies [1-10]. Even relatively small motions may significantly compromise the resolution, and hence the quantitative accuracy of the image data. In addition, head motion also causes misalignment between the emission and transmission scan data, leading to erroneous correction for photon attenuation [10]. Hence, motion tracking and correction is necessary to preserve image resolution and to ensure that quantitative PET data corrections are applied as precisely as possible. The modern methods [5, 7, 8] to reduce the degrading effects of motion fall into several categories: image realignment in frame-mode or list-mode acquisition, optical tracking systems by devices mounted on a patient's head, and their combinations. Up to now, the development of effective method to detect head movements based on new concepts is actual for medicine practice [5] because the known methods and systems produce a set of artefacts, operate off-line (after PET scanning), and demand high computational power.

As one of the prospective approach to solve the similar problems is concerned with the artificial systems based on simulation of biological vision mechanisms, especially the Foveal Systems [11-13]. The Foveal Systems imitate space-variant visual acuity, changing from the centre of the retina (the fovea) to its periphery, and controlling attention mechanisms for human gaze while image viewing. In our study, the robust method to monitor the head movement is proposed. It is based on two biologically motivated models. One model is CIECAM97 [14-16] for measuring colour appearance taking into account of illumination conditions. The other vision model, Foveal System for Face Images (FOSFI), is developed based on BMV model (Behaviour Model of Vision) simulating some mechanisms of the real visual system for perceiving shapes [13, 17]. These models are used for colour segmentation of facial area on initial pictures and for detection of local facial landmarks (i.e. eye corners and middle point of nose basement) correspondingly. Recently, we have developed and initially tested the basic algorithms [18, 19]. Here the system architecture to estimate head motion based on the developed algorithms is presented. Besides, future steps to develop the proposed system have been considered.

2. System architecture

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

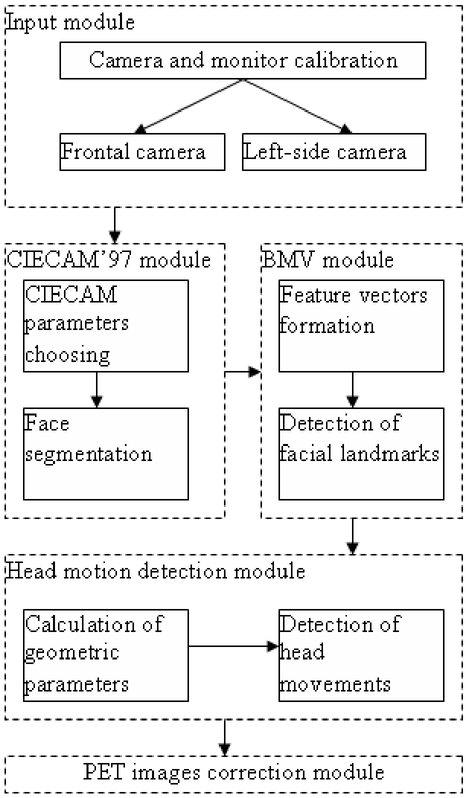

Figure 1. The overview of the system for head motion detection.

The following sections details each of the module in Figure 1.

2.1. Input module

Input module consists of two off-the-shelf digital cameras, Canon EOS-1D Mark II, to monitor the head movement on a subject performing a PET brain scan. Before the shooting, cameras and colour monitors are calibrated because both main algorithms (i.e. face segmentation and facial landmarks detection) is based on face skin colours. Figure 2 illustrates the steps of camera calibration.

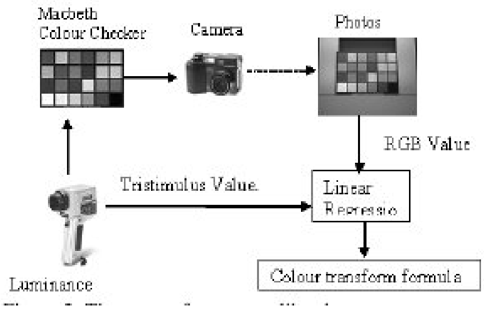

Figure 2. The steps of camera calibration.

Firstly, the camera is applied to take pictures of 24 colour patches from Macbeth Colour Checker. Then these images are transferred to a colour monitor that has been calibrated in advance under the same illuminant. 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 obtain camera calibration model by the least square method.

In order to make sure that the same colour appears the same all the time, colour monitor needs to be calibrated. This is done by using optical software with Spyder sensor. When an image is downloaded to a computer, it is presented in a RGB space. To convert RGB space to CIE standard XYZ space, the following equations are applied as shown in Eq (1) for average daylight with CIE standard illuminant D65 as reference white, i.e., [X_W, Y_W, Z_W] = [0.95045 1.0 1.088754].

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

2.2. Face segmentation module

This module consists of two submodules, i.e. CIECAM parameters choosing and face segmentation. Basic algorithms performed in this module are described earlier in detail [18, 19]. In computer implementation, choosing of CIECAM parameters to be characteristic for face skin colour may be performed by three methods:

1) Use of the averaged ranges of parameters to be determined for the human with the given skin colour.

2) Use of the colour range for each of CIECAM attributes (hue, chroma, and lightness) to be obtained for one (initial) image of the given subject. These colour ranges are then utilised to segment other pictures for the same subject.

3) Colour attributes are calculated by CIECAM model for each picture of a given subject.

Detection of potential face area approximated by a square is performed by the nearest neighbouring method [20]. The result of face segmentation module work is presented in Figure 3.

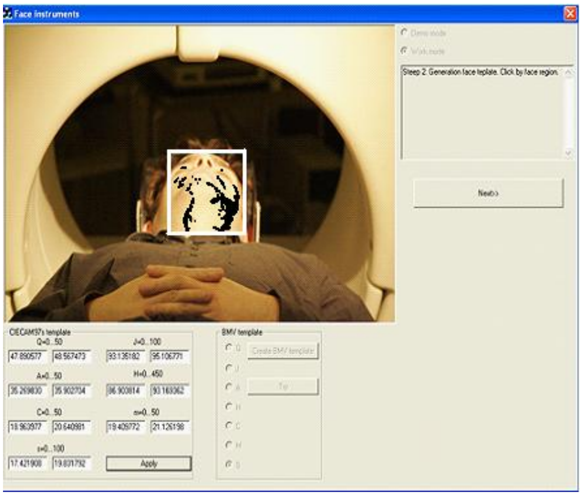

Figure 3. The module for face segmentation.

2.3. Facial landmarks detection module

Segmented fragments of initial pictures with face regions are used then in facial landmarks detection module. This module consists of two submodules, i.e. feature vector formation and detection of local facial landmarks. Basic algorithms performed in this module are described earlier in detail [18, 19]. In computer implementation, feature description of each facial landmark (Figure 4) is formed by space-variant input window and represented by multidimensional vector. The vector components are values of primary features detected in the vicinity of each of 49 input window nodes A_i , i=0, 1…48. According to the preliminary tests, eyes corners and middle point of nose basement are chosen as facial landmarks in the consideration that they have a set of relatively constant local features and are visible on all head pictures. Each component of feature vectors is in line with the orientation of local “colour” edge detected from colour attributes (lightness, chroma, and hue) which are extracted by convolving a map with a set of 16 kernels. Each kernel is sensitive to one of 16 chosen orientations. The whole set of 16 kernels are determined by differences between two oriented Gaussians with shifted kernels.

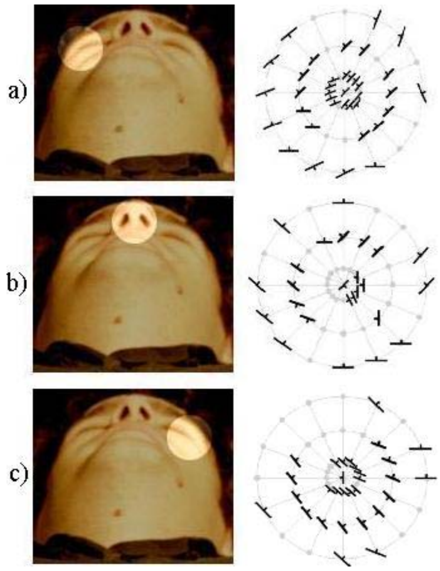

Figure 4. Etalon feature vectors of facial landmarks for a subject: (a) right eye corner, (b) middle point of nose basement, (c) left eye corner.

Detection of local facial landmarks is performed by comparison of feature vectors. Initial estimation of each chosen landmark is obtained by single positioning input window in the corresponding region centers on one image only for each subject, and works as template feature vectors. Then all consequent images of the same subject are scanned by input window to search image points with feature vectors similar to the template feature vectors. The result of facial landmarks detection module work is presented in Figure 5.

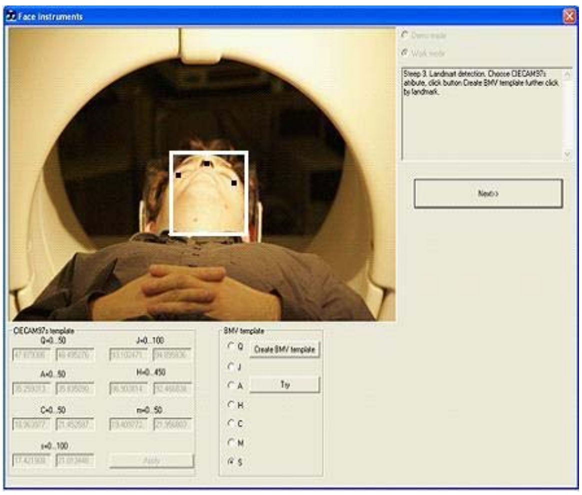

Figure 5. The module for local facial landmarks detection.

2.4. Head motion estimation module

Head motion estimation module is mainly realized as a set of algorithms and in part in computer implementation. This module consists of two submodules, i.e. calculation of geometric parameters between local facial landmarks and detection of head movements. In the first one, position of local facial landmarks on pictures X and Y axes as well as linear and angular relations between them are calculated as geometric parameters for each picture to be processed. In the second submodule, comparison of geometric parameters between facial landmarks identified on consecutive images is performed. The scheme for head motion detection is presented in Figure 6.

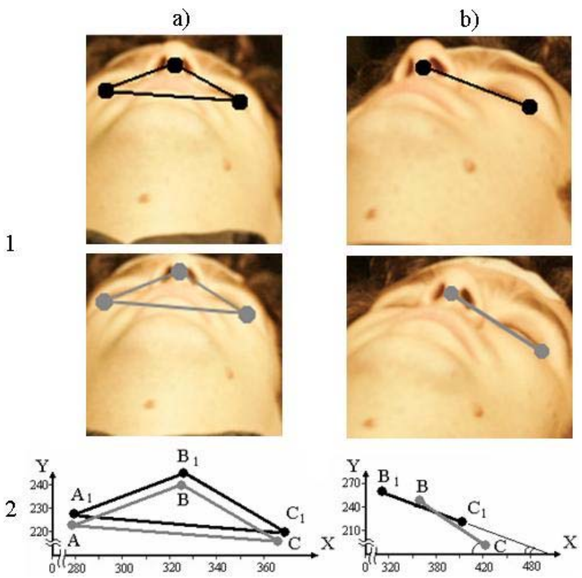

Figure 6. Spatial relations (1) between facial landmarks detected on consecutive pictures and their scheme (2): a) pictures captured by frontal camera (A, B, C and A_1, B_1, C_1 - landmarks identified); b) pictures captured by the left camera. In (2) the pixels of initial pictures (upper row) are presented on X and Y axes.

3. Computer simulation

Two sets of image were collected to estimate the system performance. The first image collection (n=12) with a subject lying down in the PET scanner includes the pictures with known subject head positions (measured by the build-in red laser beam of the PET scanner) and illumination conditions. All pictures monitoring a subject's head while simulating PET scanning had the same dimension (640x427 pixels) and were received by calibrated cameras (see Section 2.1). The second set of images includes the pictures of the human faces with various skin colour. A part of this collection was homemade (n=11) and another one (n=20) was take in Internet, mainly on three WEB sites (www.landentwicklung-muenchen.de , www.phys.cwru.edu , www.nymphetamine.nnm.ru). The most of the last pictures had various dimension and were received by cameras without calibration.

During computer simulation, very high system performance was revealed while processing the pictures from the first collection. In particular, face areas were correctly segmented on all pictures without false positive regions. Besides, all the visible landmarks (i.e. eye corners and middle point of nose basement, n=30) have been identified (3 landmarks on pictures captured by the front camera and 2 by the side camera, see Figure 6) 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 figured out by comparison of landmarks from two different images (normally two consecutive images) of the same subject doing the same scan. Estimations on 2D images with known moving parameters show that movement parameters can be obtained very accurately (less then 10 degrees for a sample).

While processing the pictures from the second collection, the system performance was relatively low. In particular, many false positive regions were identified (n=50) beside face areas (n=31) during face segmentation. Besides, only 56 facial landmarks among 88 visible ones were correctly detected.

Software module simulating the PET scanning situation was developed in first approximation for quantitative estimation of different factors contribution to the system performance. It includes 3D model of a head that may be rotated in any axes with one degree precision. The simulator will permit an opportunity to develop correct algorithms for integration of processing data for the pictures simultaneously received by the frontal and the left side cameras.

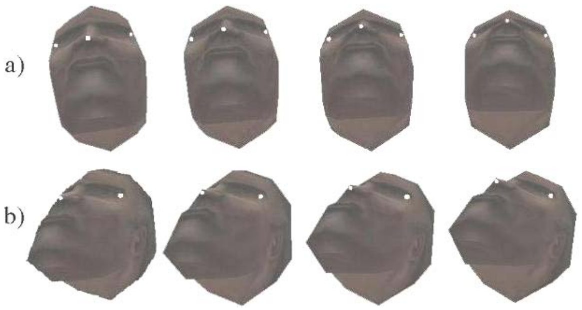

Figure 7. Head movement simulation, a) frontal camera, б) left side camera.

4. Conclusion

The state of the art methods [5, 7, 8] to reduce the degrading effects of head motion are based on image realignment in frame-mode or list-mode acquisition, optical tracking systems by devices mounted on a patient's head, and their combinations. In our study, a new noninvasive approach to monitor the human head motions during PET scanning based on two biologically motivated vision models, i.e. CIECAM97 [15, 16] and BMV [13, 17] has been developed. These two models are applied for colour-based segmentation of face regions and for detection of local facial landmarks respectively. Computer simulation results while processing the pictures captured by calibrated cameras shown very high performance of the system under development. In particular, centres of chosen facial landmarks, including eye corners and middle point of nose basement, are detected with very high precision with deviation about 1 pixel. Preliminary estimations on 2D images with known moving parameters 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.

Let us consider possible future steps to improve the system performance and toward its application in medicine practice. First of all the system must be tested on many pictures of the subjects with different range of skin colour. Then, the images taken from another camera should be analyzed together with the images from the first camera, which is still under investigation at the moment. Thirdly, a method for motion identification and reconstruction (for exemples by estimation of spatial relations between facial landmarks, see Figure 6) will be developed. It is supposed that face contour detection [20, 21], method face superposition [22], and comparison of two symmetrical facial lines of two consecutive images [23] in addition to facial landmarks detection can improve determination of head movements with demanded exactness. Finally, the detected motion parameters will be utilized in the process of PET brain image reconstruction, which requires list-mode (event-by-event data acquisition) data. It is foreseen that the the implementation of the all steps listed above will be used software simulator (see Figure 7) as additional tools and built a automatic motion correction system.

Acknowledgments: This work is part of TIME (Tele-imaging in Medicine) project funded by European Commission under Asia-IT&C program (CN/ASIA-IT&C/009(91450)) and BRIDGE Research Cooperation Award: Head Motion Detection for Positron Emission Tomography (PET) Image Scanning. In Russia, the work is supported in part by Russian Foundation for Basic Research, Grant No. 05-01-00689.

5. References

[1] P.M. Bloomfield, T.J. Spinks, J. Reed, L. Schnoor, A.M. Westrip, L. Livieratos, R. Fulton, T. Jones, "The Development of a Motion Correction System in Neurological PET Using List Mode Data Acquisition and a Polaris Tracking System", Neuroimage, no. 16, 2002, pp. S15.

[2] P. Bühler, U. Just, E. Will, J. Kotzerke, and J. van den Hoff, "An Accurate Method for Correction of Head Movement in PET", IEEE Transaction on Medical Imaging, v. 23, iss. 9, 2004, pp.1176-1185.

[3] M.V. Green, J. Seidel, S.D. Stein, T.E. Tedder, K.M. Kemper, C. Kertzman, T.A. Zeffiro, "Head Movements in Normal Subjects During Simulated PET Brain Imaging With and Without Head Constraint", J Nucl Med, no. 35, 1994, pp.1538-46.

[4] 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., no. 49, 2002, pp. 188-194.

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

[6] U.E. Ruttiman, P.J. Andreassen, D. Rio, "Head Motion During Positron Emission Tomography: Is it Significant?", Psychiatry Res: Neuroimaging, no. 65, 1995, pp. 41-43.

[7] 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.

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

[9] R.P. Woods, S.R. Cherry, J.C. Mazziotta, "Rapid Automated Algorithm for Aligning and Reclining PET

Images", J Comput Assist Tomogr, no. 16, 1992, pp. 620-633.
[10] H. Zaidi, M.-L. Montandon, and S. Meikleb, "Strategies for Attenuation Compensation in Neurological PET Studies. Review", NeuroImage, no. 34, 2007, pp. 518 – 541.
[11] I. Ahrns, H. Neumann, "Space-Variant Image Processing" Proc. in Artificial Intelligence, no. 9, 2000, pp. 203-206.
[12] C. Bandera, P. Scott, "Machine Vision for Pursuit. The Foveal Alternative"? Journal of Electronic Defense, v. 14, iss. 10, 1991, pp. 61-66.
[13] I. Rybak, V. Gusakova, A. Golovan, L. Podladchikova, N. Shevtsova, "Attention-guided Recognition Based on «What»; and «Where» Representations: A Behavioral Model", The first encyclopedic volume on Neurobiology of Attention., Eds. L. Itti, G. Rees, and J. Tsotsos. Elsevier, Academic Press, 2005, pp. 663-670.
[14] http://www.colorcal.com/
[15] CIE, The CIE 1997 Interim Colour Appearance Model (Simple Version), CIECAM97s, CIE TC1-34, April, 1998.
[16] 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.
[17] M.V. Petrushan, A.I. Samarin, D.G. Shaposhnikov, "FOSFI: A System for Face Image Recognition", Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, v. 15, no. 2, 2005, pp.425-427.
[18] X. Gao, S. Batty, L. Podlachikova, D. Shaposhnikov, and J. Clark, "Detection of Head Motions Using a Vision Model", Proc. of International Conference on Telehealth, Canada, 2007, pp. 167-171.
[19] 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", Journal of Computer Science, v. 3, no. 7, 2007, pp.528-532.
[20] X.W. Gao, L. Podladchikova, D. Shaposhnikov, K.Hong, N. Shevtsova, "Recognition of Traffic Signs Based on Their Colour and Shape Features Extracted Using Human Vision Models", J. Vis. Commun. Image R., no. 17, 2006, pp. 675-685.
[21] D.G. Shaposhnikov, L.N. Podladchikova, X. Gao, "Classification of Images on the Basis of the Properties of Informative Regions", Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, v. 13, 2003, pp.349-352.
[22] Kalinin K.B., Samarin A.I. "Gaze Fixation Points in Context-depended Visual Perception Tasks while Face Recognition", Proc. of "Neuroinformatics 2008", v. 2, 2008, pp.117-123 (in Russian).
[23] X.W. Gao, S. Batty, J. Clark, T. Fryer, A. Blandford, "Extraction of Sagittal Symmetry Planes from PET Images", Proc. of the IASTED Int. Conf. on Visualization, Imaging, and Image Processing (VIIP'2001), ACTA Press, 2001, pp 428-433.