X. GAO[1], D.G. SHAPOSHNIKOV[2], S. ANISHENKO[2], S. BATTY[1]
[1]School of Computing Science, Middlesex University, United Kingdom
X.Gao@mdx.ac.uk
[2]A.B. Kogan Research Institute for Neurocybernetics, Rostov State University
nisms@krinc.ru

HUMAN FACE DETERMINANTS TO DETECT HEAD MOVEMENTS DURING POSITRON TOMOGRAPHY.

Abstract

The method for noninvasive detection of head motion during PET procedure based on two biologically motivated models namely CIECAM'97 [1-3] and BMV [4,5] is proposed. These models are used for colour segmentation of initial pictures to detect the facial area and identification of face determinants (i.e. the eye, nose, and mouth regions) correspondingly. Initial pictures (n=13) of PET procedure are classified into 4 groups according to lightness and viewing conditions. The range of CIECAM model parameters for each group are determined. High performance of the developed method to detect facial area and face determinants are revealed during computer simulation. Future steps to develop the proposed method have been considered.

1. Introduction

Positron emission tomography (PET) data acquisition is a relatively lengthy procedure (typically ~1 hour) and it is difficult for the subject to stay still during the data acquisition period. Consequently, head motion can significantly degrade the quality of PET studies of the brain [6-12]. 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. Hence, motion tracking and correction is necessary to preserve image resolution and to ensure that quantitative data corrections are applied as precisely as possible.

Known post-acquisition methods [6,8-12] to reduce the degrading effects of motion fall into two categories: image realignment; and raw data reorientation prior to image reconstruction. Optimal correction for motion requires accurate determination of the motion parameters and accurate reorientation of the raw data based on these parameters. Up to now the development of effective noninvasive method is actual for medicine practice because the most of known methods to detect head movements are based on use of devices or markers mounted on patient head or have low efficiency in real conditions [11].

In this study, noninvasive method to monitor the head movement is proposed. It is based on two biologically motivated models. One model is CIECAM97 [1,2] for measuring colour appearance invariantly to illumination conditions. The other vision model, Foveal System for Face Images (FOSFI), is developed based on BMV model (Behaviour Model of Vision) imitating some mechanisms of the real visual system for perceiving shapes [4,5]. These models are used for colour segmentation of facial area on initial pictures and for detection of face determinants (i.e. the eye, nose, and mouth regions) correspondingly.

2. Methodology

2.1 Human face detection

For colour segmentation of initial pictures of PET scanning procedure and detection of face area is used CIECAM97, the colour appearance model recommended by CIE (Commission Internationale de l'Eclairage) [1-3]. It imitate the human colour perception and can correctly identify different colour images invariant of lighting conditions and viewing angles due to the adaptation to the environment. This model is based on a simplified theory of colour vision for chromatic adaptation together with a uniform colour space. It can predict colour appearance as accurate as an average observer. This colour appearance model is expected to extend traditional colorimetry (e.g., CIE XYZ, and CIELAB) to the prediction of the observed appearance of coloured stimuli under a wide variety of viewing conditions, which is accomplished by taking into account the tristimulus values (X, Y, and Z) of the stimulus, its background, its surround, the adapting stimulus, the luminance level, and other factors such as cognitive discounting the illuminant.

Parameters of CIECAM97 model are relative tristimulus values of white ($X_W Y_W Z_W$), luminance of the adapting field (La, cd/m*m), relative luminance of the background (Y_b) and surround parameters (c, Nc, F_{LL}, F). Input array of CIECAM97 are relative tristimulus values of colour stimulus (*XYZ*) described in Eq. 1. The output of CIECAM97 includes mathematical correlates for perceptual colour attributes that are brightness (Q), lightness (J), achromatic response (A), colourfulness (M), chroma (C), saturation (S), and hue including hue composition (H) and hue angle (h).

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., $[X_w, Y_w, Z_w] = [0.95045\ 1.0\ 1.088754]$.

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

2.2 Detection of eye, nose, and mouth regions

After colour segmentation, the positions of facial determinants (i.e. eye, nose, and mouth regions) are estimated. Feature description of each facial area is provided by space-variant input window (IW, Fig. 1) and is represented by multidimensional vector s $\vec{F}$.

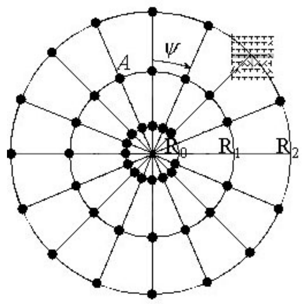

Fig. 1. Structure of space-variant input window with detailed description of context area around each node (one of them is shown as rectangular grid).

Feature vector $\vec{F}$ is based on edge orientation α in the vicinity of each of 49 sensor nodes A_i , i=0, 1…48. Orientation and contrast of local segments are determined by means of calculation of the difference between two oriented Gaussians with spatially shifted centers. Let $x_0 = X_0$, $y_0 = Y_0$ be co-ordinates of the central sensor node, then co-ordinates (x_i, y_i) of peripheral sensor node A_i , i=1, 2…48 can be determined as follows:

$$x_i = X_0 + R_l \cos\psi_k,$$

$$y_i = Y_0 + R_l \sin\psi_k$$

where R_l , l=0,1,2 is the radius of l-th concentric circle of the sensor (R_0 = 3 pixels, R_1 = 9, R_2 = 15) and $\psi_k = k{\cdot}22.5°$, k= 0,1…15 is the angle of the radiating line corresponding to the i-th sensor node. Here, R_l simulates a space-variant resolution level which is emulated by Gaussian convolution with different kernels depending on distance from the IW center (5x5 pixels for the central part of the IW, 7x7 - for the immediate, and 9x9 - for the peripheral part). Each sensor node is characterised by edge orientation α that dominates in the node context area (7x7 pixels) and its density ρ as follows:

$$\begin{aligned} &\rho(A_i) = \max_{\varphi} \rho_{\varphi}(A_i) \\ &\alpha(A_i) = \varphi \ \ if \ \ \rho_{\varphi}(A_i) = \rho(A_i) \end{aligned} \tag{2}$$

where $$\rho_{\varphi}(A_i) = \rho_{\varphi}(x_i, y_i) = \frac{1}{S(x_i, y_i)} \sum_{m,n} Sg_{\varphi}\left(Or(m + x_i, n + y_i)\right) \tag{3}$$

$$Sg_p(x) = \begin{cases} 1; \text{ if } x = p \\ 0; \text{ otherwise} \end{cases} \tag{4}$$

and $Or(x,y)$ is a detected edge orientation (not only dominating one) in the vicinity of the image element with co-ordinates (x,y); $S(x_i, y_i)$ is the square (see Fig. 1) of the context area for i-th sensor node equal to 49 pixels; m,n = -3,…0,…+3; φ = 0,1…15. The resulting feature vector $\vec{F}(\vec{\alpha}, \vec{\rho})$ is therefore formed as Eq.5

$$\vec{F}(\vec{\alpha}, \vec{\rho}) = (\alpha(A_0)\ldots.\alpha(A_{48}), \rho(A_0)\ldots. \rho(A_{48})) \tag{5}$$

Contrary to the cascade method for detection of informative facial regions developed earlier [5] specific prototype description for each region (eye, nose and mouth) is received by operator single positioning IW in region centers on initial image. After that IW scanned all image and location of image points feature description of which are similar to the prototype feature vectors are determined. Incoming and prototype feature vectors are compared by Eq. 6:

$$K^b = \sum_{i=0}^{i<49} \left[\operatorname{sgn}\left(Or_i^b - Or_i^{rw}\right) \cdot \left(1 - abs\left(\rho_i^b - \rho_i^{rw}\right)\right) \right] \tag{6}$$

where: $\operatorname{sgn}(x) = \begin{cases} 1, & \text{if } x = 0; \\ 0, & \text{in other case;} \end{cases}$

Or_i – dominating orientation of contrast segment in context area of the given IW node (orientation is determined by step 22.5° and denoted as 0, 1, 2, 3…., 15); indexes b denote prototype vectors, indexes rw – current vectors; ρ - density of dominating contrast oriented segment in the context area of each IW node.

3. Computer simulation

The pictures (n=13) with a subject lying down in the PET scanner with known head positions and illumination conditions are studied. Fig. 2 illustrates four examples of scanning situation.

Image classification is very important for choosing CIECAM97 model parameters to detect facial area on PET images. All pictures were preliminary classified by operator into four groups which are differed on lighting conditions and points of subject viewing:

group 1 – left side viewing, lightness is equal to 32.7 cd/m^2
group 2 – left side viewing, lightness is equal to 353 cd/m^2
group 3 – frontal viewing, lightness is equal to 32.7 cd/m^2
group 4 – frontal viewing, lightness is equal to 353 cd/m^2

CIECAM parameters adaptation to the different viewing and lightness conditions is necessary to develop automatic classification procedure. CIECAM parameters were obtained by calculation of mean values for saturation (S) and colourfulness (M) in squares (20x20 pixels) located in the left and right upper corners of initial images. These values and their combinations are different for the above groups. Automatic classification procedure included three stages:

the first stage: if S in the left square $\in(23;27)$ and S in the right square $\in(12;15)$, then an image is related to "group 2", otherwise go to the second stage;

the second stage: if S in the left square $\in(35;38)$ and S in the right square $\in(15;18)$, then an image is related to "group 4", otherwise go to the third stage;

the third stage: if M in the left square is more then M in the right square, then an image is related to "group 1", otherwise an image is related to "group 3".

After image group classification, colour characteristics for CIECAM model parameters were obtained for facial areas in each group. These parameters are presented in the Table.

Group	CIECAM97 parameters			
	A	H	J	Q
1	28-34	55-95	ignore	ignore
2	ignore	89-100	4-80	ignore
3	28-34	94-110	ignore	35-46
4	30-37	89-123	ignore	ignore

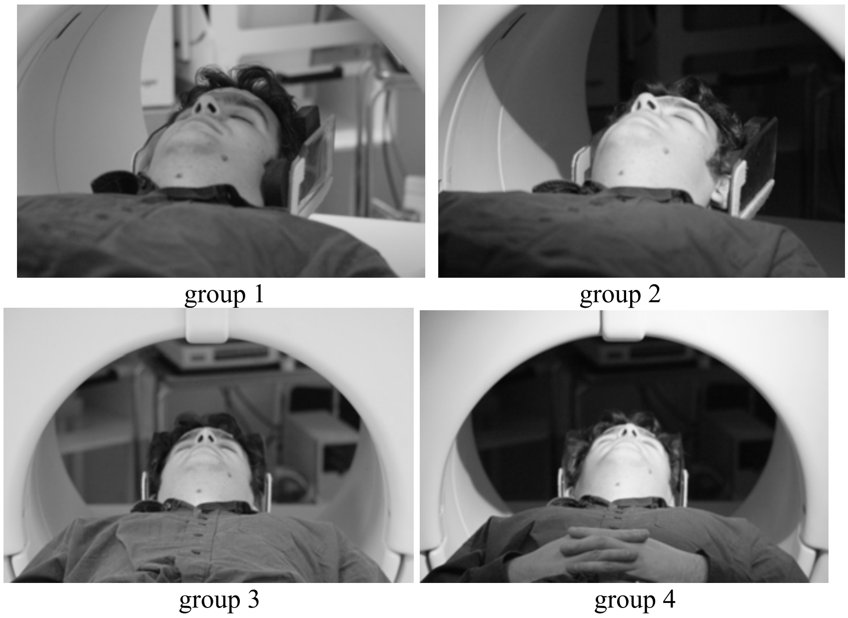

Fig. 2. Four examples of pictures for PET procedure with different illumination and viewing conditions.

The tests shown that CIECAM model with parameter adaptation provided face segmentation without false areas on all 13 images. Examples of segmented facial areas for four groups are shown in Fig. 3.

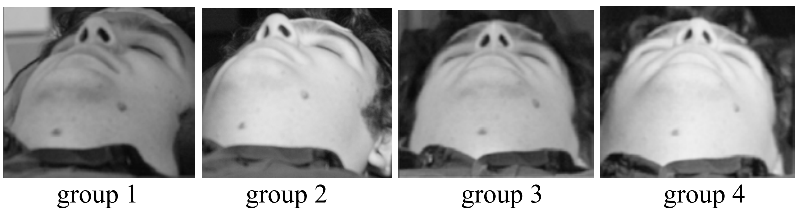

Fig. 3. Examples of segmented facial areas.

In