Statistical Lesion Classification on Brain SPECT Images

Yi-Ying Wang Nan-Tsing Chiu[1] Tiee-Jian Wu Chwin-Min Weng[2]
Wen-Feng Kuo[3] Yung-Nien Sun[3,*]

Institute of Statistics, National Cheng Kung University Hospital, Tainan, Taiwan, 701, ROC
[1]*Department of Nuclea Medicine, National Cheng Kung University Hospital, Tainan, Taiwan, 701, ROC*
[2]*Department of General Education, Tainan Woman's College of Art and Science, Tainan, Taiwan, 710, ROC*
[3]*Department of Computer Science and Information Engineering, National Cheng Kung University Hospital, Tainan, Taiwan, 701, ROC*

Received 13 January 2003; Accepted 4 May 2003

Abstract

Recent researches [1] have shown that individuals with more years of education have a more advanced development in Alzheimer's disease (AD) than those who with fewer years of education. In this study, two groups of fifty-eight AD patients with different education levels were examined. We utilize two statistical methods to detect the significant difference in brain region perfusion between the high and low education level patients with AD. The contextual clustering method, similar to the one for MR lesion detection in [2], is designed and applied to the 99mTc-hexamethyl propylenamine oxime (HMPAO) brain SPECT images. The other uses the SPM99 software that is commonly applied in voxel by voxel analysis of SPECT data and is popularly adopted in neuroimaging society over the past few years [3, 4]. Our finding suggests that there are significantly reductions in region perfusion in the group of high education level patients compared with the group of low education level patients. These reductions were observed in the left temporal region.

Keywords: Single-photon emission computed tomography (SPECT), Alzheimer's disease, Statistical parametric mapping, Contextual clustering

Introduction

Brain perfusion Single Photon Emission Computed Tomography (SPECT) is widely used in the study for Alzheimer's disease. In this paper, two groups of SPECT images were used to study the influence of education level to the disease progression of AD. In this evaluation, the subjects with a history of prior neurological illness were first excluded. There are totally fifty-eight patients with AD diseases examined in this study. 99mTc-hexamethyl propylenamine oxime (HMPAO) brain SPECT images were acquired from each subject and evaluated by our proposed system.

Before the statistical analyses, the accurate registration and grey level normalization of images to a standard template are necessary. These were done by using the SPM 99 modules provided by the Wellcome Department of Cognitive Neurology, Institute of Neurology, London [5]. SPM is a powerful technique and popularly used in the neuroimaging society. Here, it is adopted for the comparison of functional imaging data sets between two groups of patients.

In the statistical analysis, fifty-eight medical images in two groups were summarized to a t-maps image by using the two-sample t-test. A single active region usually consists of numerous voxels. Hence, it may be a good idea to incorporate the contextual information in detecting the active regions. Therefore, a method similar to the Eero Salli's contextual clustering method [1] was applied to display the significant differences regions. Unlike the conventional method of voxel by voxel testing, the contextual clustering algorithm not only computes intensity value itself but also takes the neighboring voxels into account.

This paper is organized as follows. Section 1 describes the SPECT image preprocessing including registration and normalization. The two statistical analyses are described in Section 2. These methods are used to detect the active brain region from the Single-Photon Emission Computed Tomography (SPECT) images. The results are presented in Section 3, and finally the conclusions are given in Section 4.

Image Preprocessing

SPECT images preprocessing were firstly manipulated with the SPM 99 software implemented in MATLAB (Mathworks Inc., Sherborn, MA). Images from fifty-eight subjects were analyzed together by first normalizing them all to the same space. These SPECT images were transformed into a

* Corresponding author: Yung-Nien Sun
Tel: +886-6-2757575 ext. 62526; Fax: +886-6-2747076
E-mail: ynsun@mail.ncku.edu.tw

Table 1. Summary of basic information for the two education groups

	Mean (SD)	
	Group A Below High School	Group B Above High School
N	29	29
Age (yr)	76.62(8.76)	77(6.84)
Education (yr)	7.2(2.2)	15.2(2.9)
Female/Male	18/11	18/11

(•) indicates the standard deviation

standard anatomical space, which conformed to the space described in the atlas of Talairach and Tournoux (1988). We use the 12-parametered affine transformation to register these brain images with respect to a template image [6, 7]. The spatial normalization procedure was achieved using the SPECT template provided by the SPM 99 software, which had been modified in order to be symmetrical. The matrix size of the SPECT image template provided with the SPM99 software was 79×95×68 voxels with an isotropic voxel size of 2 mm. The images were then smoothed using a Gaussian kernel (6×6× 6-mm FWHM). This step will increase the signal to noise ratio and improve the performance of the subsequent statistical analysis.

The fifty-eight images thus constructed were then tested in this study. Fifty-eight patients (36 women, 22men, aged 46-86 years, mean: 76.81) clinically diagnosed with AD were evaluated, and we divide these patients into two groups: one is with high education level, and the other is with low education level. Table 1 summarizes the fifty-eight subjects of the following two education groups.

Statistical Analysis

After the image processing step mentioned above, we obtain two mean images from the two groups of patients. The mean image from higher education group is $\overline{Y}_A$, and the one from lower education group is $\overline{Y}_B$. We then apply the two statistic methods that are the contextual clustering and the SPM99 to detect the regions with significant deviations in blood perfusion. The two methods are illustrated as follows:

(i) Contextual Clustering Algorithm

The contextual clustering method was first developed by Eero Salli for detecting the MR lesions [1]. The method uses contextual information to increase the detection rate of weakly activated regions. In this study, we present a new approach modified from [1] to process the SPECT classification problem. We use the simulated annealing (SA) procedure to maximize the posterior probability [8] and apply it to the HMPAO SPECT brain images. At the first step, the voxels in 3D SPECT images are classified into the activated and non-activated voxels by thresholding the standard normal distribution maps. Then, based on the concept of contextual information, the voxels are re-classified by using the distribution maps and the classification information of adjacent voxels simultaneously. The contextual clustering algorithm is described as follows:

A.Statistical Parametric Mapping

The first step in the contextual clustering algorithm is in fact the same procedure adopted in the statistical parametric mapping module (SPM99). The image is denoted by $y = (y_i, i \in I)$, where y_i is the image intensity at the voxel site indexed by i. Let $I = \{i_1, i_2, \ldots, i_n\}$ represent the set of corresponding voxel pairs on the two mean images, for every voxel pair $i \in I$, the test statistic can be computed with equation (1)

$$t_i = \frac{\overline{Y}_{iB} - \overline{Y}_{iA}}{\sqrt{\frac{((n_A-1)\hat{\sigma}^2_{iA} + (n_B-1)\hat{\sigma}^2_{iB})}{(n_A+n_B-2)} * ((1/n_A)+(1/n_B))}} \quad (1)$$

where n_A and n_B are the sizes of group A (subindex A) and group B (subindex B)； $\overline{Y}_{iA}$ and $\overline{Y}_{iB}$ are the means of the observed intensity values of the corresponding voxels in group A and B；and $\hat{\sigma}^2_{iA}$ and $\hat{\sigma}^2_{iB}$ are the estimated variances.

The group A represents the high education level group and the group B represents the low education level group here. From equation (1), we can compute t values, which follow the t distribution with $n_A + n_B - 2$ degree of freedom. If the degree of freedom is large enough, t distribution will approximate to the standard normal distribution. Since the degree of freedom is 56 in our case, we can transform the resulting distribution from equation (1) to the standard normal one with a new statistics z_i computed by using equation (2),

$$z_i = q_{norm}[p_t(t_i, df)] \quad (2)$$

where $q_{norm}(x)$ is the normal inverse distribution, $p_t(y,df)$ is the cumulative distribution function for the t distribution with the degree of freedom of df at y.

B.Contextual Clustering

After obtaining the z distribution, the concept of contextual clustering can then be applied. A neighborhood sphere of order r at a given voxel i is defined as

$$G_i^r = \{j, 0 \le \|i-j\| \le r, j \in I\} \quad (3)$$

where $\sqrt{r}$ is referred to as the radius of the neighbor. In this study, the third-order neighborhood is used. Some parameter definitions and concept of MRF are presented as follows:

We assume that the total number of classes in the image is k, $k=0$ is the background class, $k=1$ is the activation class. Let $\mathbf{a}_i$ is the random variable representing the true classification value at voxel i. Then $a_i = 0$ represents that the voxel i belongs to the nonactive region, and $a_i = 1$ represents that the voxel i belongs to the active region.

Let $p(z_i \mid \mathbf{a}_i)$ be the conditional probability density that the voxel i has value z_i given that the true segmentation is $\mathbf{a}_i$. Also let $\mathbf{a}_{\partial i}$ represent that the classification levels of the 26 neighboring voxels of voxel i. Thus we can regard $P(a_i \mid \mathbf{a}_{\partial i})$ as the conditional prior probability of the class $\mathbf{a}_i$ at voxel i, and let the random variable **a** belong to a MRF.

The segmentation process combines the contextual clustering and hypothesis testing. And $p(z_i \mid \mathbf{a}_i = 0)$ represents the probability obtained from the z maps under the null hypothesis of no activation. Similarly, $p(z_i \mid \mathbf{a}_i = 1)$ represents the probability obtained under the alternative

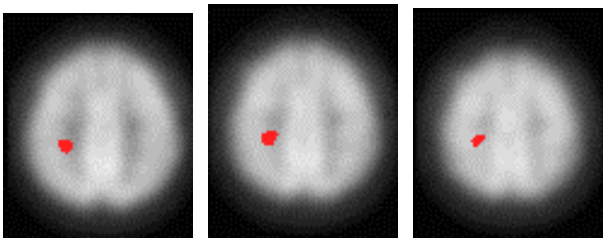
Initialize z ;
Repeat
Generate $z_{i'} \in N(z_i)$ ($z_{i'}$ belongs to one of the 26 neighborhood of z_i)

$$\Delta g = g(\mathbf{a})' - g(\mathbf{a}) \quad \text{where} \quad g(\mathbf{a})' = \frac{1}{2\sigma^2_{k,i'}}(z_{i'} - \mu_{k,i'})^2 - \gamma\, m_{i'}(\mathbf{a})$$

$$p = \min\{1, e^{-\Delta g/T}\}$$

then generate a random number r, $r \in [0,1)$
if $r < p$ then $z_i \leftarrow z_{i'}$
and change the class at i to i'
until convergence

Figure 1. Metropolis sampling of Gibbs distribution



(a) Slice 50 (b) Slice 52 (c) Slice 54

Figure 2. Brain regions with significantly reduced perfusion by Contextual Clustering Method: Mean images of all 58 AD patients of (a) slice 50; (b) slice 52; (c) slice 54

hypothesis of activation. Hence, the testing problem is equivalent to a classification problem.

Let **z** represent the vector of all z_i's, the likelihood function can be expressed by equation (4)

$$l(\mathbf{z} \mid \mathbf{a}) = \prod_{i=1}^{n} p(z_i \mid a_i) \tag{4}$$

By applying the Bayes' rule, we can get the posterior density approximated to equation (5). And

$$p(\mathbf{a} \mid \mathbf{z}) \propto l(\mathbf{z} \mid \mathbf{a}) p(\mathbf{a}) \tag{5}$$

where $p(\mathbf{a})$ is the Markovian prior.

Then, the ICM algorithm [9] can be used to find the maximum of equation (5) through the following steps:

1. Initial classification is chosen so that $p(z_i \mid a_i)$ is maximized at each voxel *i*.
2. Given **z** and the information of neighborhood of voxel *i*, we find a suitable classification to make the conditional probability maximum.

$$p(a_i \mid z, \hat{\mathbf{a}}_{\partial i}) \propto p(z_i \mid a_i) p(a_i \mid \hat{\mathbf{a}}_{\partial i}) \tag{6}$$

3. step (2) is repeated until the classification does not change.

The ICM algorithm maximizes the conditional density of the image data at each site *i* given the image **z** and the current estimate of **a** at all other sites. The procedure offers a good segmentation with a good starting configuration [10].

The prior model is based on a 3D MRF. Through the Hammersley-Clifford theorem [11], an MRF can be written as a Gibbs field as follow:

$$p(\mathbf{a}) = \frac{1}{Z}\exp[-U(\mathbf{a})/T] = \frac{1}{Z}\exp[-\sum_{c \in C} V_c(\mathbf{a})/T] \tag{7}$$

where U is the energy function, C is the set of cliques, $V_c(\bullet)$ is the potential associated with the clique, the temperature T is a control parameter, Z is the normalization constant, and the summation is taken over all the cliques in the image.

From equation (7), the conditional prior probability of class *k* at voxel *i* is derived as equation (8)

$$p(\mathbf{a} = k \mid \mathbf{a}_{\partial i}) = \frac{e^{\gamma m_i(a)}}{\Sigma e^{\gamma m_i(a)}} \tag{8}$$

In equation (8), parameter γ defines the weight for contextual information, $m_i(a=k)$ is the number of neighbors of the voxel *i* having class k.

The class density are defined to be Gaussian, i.e.,

$$p(z_i \mid \mathbf{a}) = \frac{1}{\sqrt{2\pi\sigma^2_{k,i}}}\exp[\frac{-(z_i - \mu_{k,i})^2}{2\sigma^2_{k,i}}] \tag{9}$$

Substituting equation (5) from (8) and (9), the posterior probability $p(\mathbf{a} \mid z_i)$ is given by equation (10). We find the optimal segmentation **a** by maximizing the equation (10).

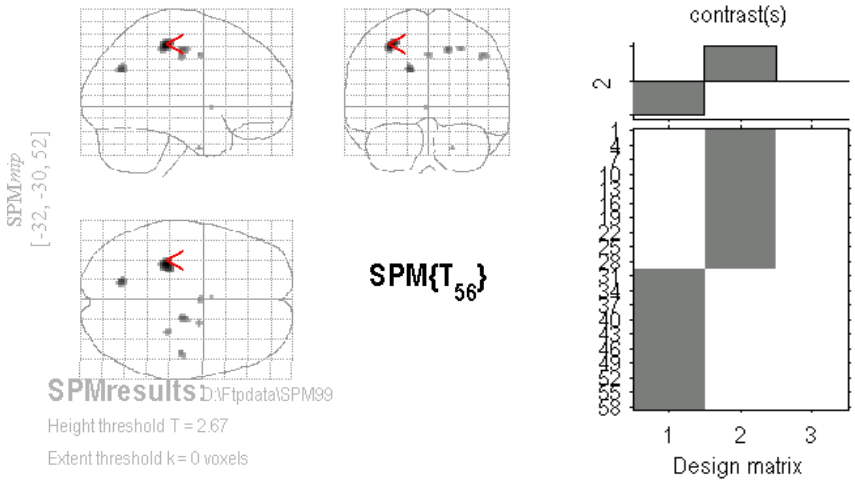
$$p(\mathbf{a} \mid z_i) \propto e^{-g(a)} \tag{10}$$

where $g(\mathbf{a}) = \frac{1}{2\sigma^2_{k,i}}(z_i - \mu_{k,i})^2 - \gamma\, m_i(\mathbf{a})$

Our aim is to find **a**, which represents the correct class at each voxel site given by z_i. Thus, we want to find the MAP estimation from the z_i maps. The optimal segmentation a^* can be written as

$$a^* = \arg\max p(\mathbf{a} \mid z_i) \tag{11}$$

Different from Eero Salli's contextual clustering rule, we use a simulated annealing (SA) scheme to minimize the



Statistics: ***volume summary (p-values corrected for entire volume)***

set-level		cluster-level			voxel-level				x,y,z {mm}
p	c	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{corrected}$	T	$(Z_=)$	$p_{uncorrected}$	
0.995	8	0.993	79	0.285	0.976	3.45	(3.27)	0.001	-30 -30 52
		1.000	24	0.568	0.999	3.12	(2.99)	0.001	-16 -66 32
		1.000	20	0.606	1.000	3.02	(2.89)	0.002	16 -16 48
		1.000	16	0.649	1.000	2.88	(2.77)	0.003	46 -18 44
		1.000	9	0.745	1.000	2.87	(2.76)	0.003	28 -30 48
		1.000	9	0.745	1.000	2.86	(2.75)	0.003	0 -4 44
		1.000	6	0.799	1.000	2.82	(2.72)	0.003	20 -4 -36
		1.000	2	0.898	1.000	2.72	(2.62)	0.004	-2 6 -2

table shows at most local maxima > 8.0mm apart per cluster

Height threshold: T = 2.67, p = 0.005 (1.000 corrected) Degrees of freedom = [1.0, 56.0]
Extent threshold: k = 0 voxels, p = 1.000 (1.000 corrected) Smoothness FWHM = 14.2 14.5 15.2 {mm} = 7.1 7.2 7.6 {voxels}
Expected voxels per cluster, <k> = 74.828 Search volume: S = 1498640 mm^3 = 187330 voxels = 433.5 resels
Expected number of clusters, <c> = 17.15 Voxel size: [2.0, 2.0, 2.0] mm (1 resel = 391.22 voxels)

Figure 3. Comparison of patients with high and low education level using SPM99.

posterior energy. The initial segmentation are obtained by clustering the SPM{z} into two clusters: activated voxels and nonactivate ones. The ML estimates of the parameters $\mu_{k,i}$ and $\sigma^2_{k,i}$ are determined by taking the mean and variance from the **z** maps of the 26 neighborhood voxels of voxel *i* that belong to the class k.

The simulated annealing [12-14] is a very popular optimization method. The idea is that there is a cost function **g** which associates the cost with a state of the system, which is the 'temperature' T and is initially set high and gradually lowered to approaching zero. There are various ways to change the state of the system. The main advantage of the SA scheme is to accept the temporary increase in energy function thus to escape from the local energy minima. Here, we implement the algorithm based upon the *Metropolis* [15] criterion.

Starting with an initial configuration, a sequence of iterations is generated. At every iteration, a configuration z_i is randomly selected from the 26 neighborhood of the current configuration z_i and the change Δg between the cost functions of the two configurations is calculated. If the cost function decreases ($\Delta g < 0$) the new configuration z_i is accepted unconditionally; otherwise, it is accepted but only with probability $\exp(-\Delta g / T)$. Fig.1 illustrates the Metropolis algorithm. The simple code for segmentation algorithm is presented in Fig. 1, and the results by using Contextual Clustering are showed in Fig. 2.

(ii) Using SPM99 Software

The voxel-based statistical parametric mapping (SPM 99) was commonly used for statistical analyses. Several preprocessing steps such as the registration and normalization were described in section 1.

Our SPM modules use a classical two-sample *t*-test. Regional CBF differences were assessed as an inter-group comparison between the high and low education level groups by choosing the option "Compare-Population 1 scan/subject (two- sample t-test) ". In this comparison, one contrast was examined, i.e. low group minus high group.

In SPM99, the gray matter threshold was set at 0.8 as the

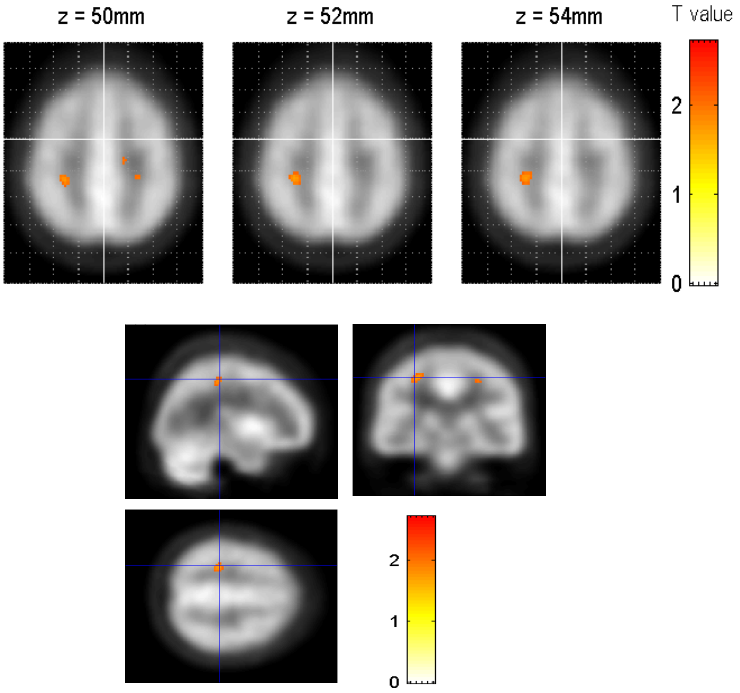


Figure 4. Main differences of the two groups are showed.

default value. Voxels exceeding the specified threshold will be examined by the software. However, Stamatakis's paper [16] suggested that the threshold should be set to 0.5 rather than 0.8 when there were significant perfusion region in the SPECT data. So, we set the gray matter threshold to 0.5. In addition, the effects of global CBF among scans were removed by using the proportional scaling with the default value of 50. In this study, the two-sample *t*-test we used is a special case of a general linear model:

Let Y_{sA} and Y_{sB} are two independent groups of random variables, where $s=1,\ldots,S$ indicates the observations, A represents the high education group, and B represents the low education group. Then the two-sample *t*-test assumes:

$$Y_{sq} \overset{iid}{\sim} N(\mu_q, \sigma^2) \qquad \text{for } q = \text{group } A, B,$$

and the null hypothesis (NH) is $\mu_A \geq \mu_B$. The model can be writing as follows:

$$Y_{sq} = \mu_q + \varepsilon_{sq}, \quad \text{and } \varepsilon_{sq} \overset{iid}{\sim} N(0, \sigma^2) \tag{12}$$

The q subscript of μ indicates there are q levels in the group effect. We can re-written the equation (12) using the two dummy variables I_{1sq} and I_{2sq} as

$$Y_{sq} = I_{1sq}\mu_A + I_{2sq}\mu_B + \varepsilon_{sq} \tag{13}$$

Therefore, the design matrix I with the two columns of dummy variables indicates the group membership, where I_{1sq} indicates that whether Y_{sq} belongs to the A group, i.e. $I_{1sq} = \begin{cases} 1 & if\ q = A \\ 0 & if\ q = B \end{cases}$, I_{2sq} is similarly to I_{1sq}.

We are looking for evidence that μ_A is less than μ_B. In other words, we have to check whether the difference between the two population means, $\mu_B - \mu_A$, is greater than zero. The model (13) leads to a design matrix I and the parameter vector $\beta = (\mu_A, \mu_B)^T$. Thus, the NH: $\mu_A \geq \mu_B$ is equivalent to NH: $c^T\beta \leq 0$ with $c = (-1,1)^T$. Through some derivation, we can get $t = \dfrac{\overline{Y}_B - \overline{Y}_A}{\sqrt{\hat{\sigma}^2(\frac{1}{n_A} + \frac{1}{n_B})}}$.

In the case of two independent random samples of size n_A and n_B, $\hat{\sigma}^2$ is the pooled sample variance, and the t distribution will have $n_A + n_B - 2$ degree of freedom. Thus, using the t test statistic, the rejection region for the one-tailed test is Reject NH, if $t > t_\alpha$.

The display format in Fig.3 provides the front, bottom, and the left-sided views of the brain data. This is a typical picture provided by SPM99. The bottom shows the statistical results associated with the SPM99. The results are presented for voxel clusters that survive the statistic tests. In these tests, the height threshold is set 2.67 in t statistic, and the extent threshold is set 0 pixels. The footnotes also give the other parameters related to this analysis.

The image shown in the upper left of Fig. 3 is a so-called "glass-brain" with the indications of classification results where voxels are higher than the threshold ($p<0.005$ or $t>2.67$). The red arrow indicates the current location relative to the acnterior commisure (-32, -30, 52).

The adopted design matrix for the linear models is also shown in the Fig.3 at the upper right corner, where the mid-gray represents 0, and the white represents 1. The design

matrix has one column for each parameter and a row for each subject. The first two columns contain variables indicating the two conditions of education levels: the first column contains n_A 1's and n_B 0's, represents the group A measurements, and the second column contains n_A 0's and n_B 1's. Here, n_A and n_B are the size in group A and B respectively (in our case, $n_A = n_B$=29). The gray box above the design matrix shows the particular contrast adopted in this study. We are interested in the reduction of perfusion in the corresponding regions between the groups A and B. Therefore, to examine group effects, we define "$\mu_A < \mu_B$" as (-1, 1). The t-statistic image was estimated at the spatial resolution of 14.2×14.5×15.2 mm, that is equivalent to 433.5 resels or 187330 intracerebral voxels. Voxels above the critical threshold are shown in the black regions at the top left diagrams of Fig.3.

The resulting SPM(t) was transformed to a normal distribution, i.e., SPM(Z). The resulting foci were then characterized in terms of the spatial extent and the peak height. The significances of region were estimated using distributional approximations in the theory of Gaussian fields (Friston *et al.,* 1995) [17]. The SPM(Z) maps were analyzed using the three corrected levels of statistical inference [18-19]. (1) At the voxel level inference, P(u, k=0, c=1) is the probability that the maximal statistical value in the interest region, which is at least one point in size and above the threshold u, could have occurred in the search volume. (2) At the cluster level inference, we relax the constraint variable k of spatial extent to a positive integer. In this case, c still takes the value 1 such that every interested region is inspected. (3) At the set level inference, the cluster number c is relaxed to a positive integer such that all the interested regions are evaluated simultaneously. The significance threshold of voxel size in this study was set at $p<0.005$. Fig.4 shows the sagittal, coronal, and transaxial planes with the SPM "activations" overlaid on the SPECT template. The significant regions are showing in each image. The color bar at the right shows the corresponding t-value corresponding to a particular color.

Experimental Study and Discussions

Two groups of AD patients with the age and sex matched were selected. There were 18 women and 11 men in the high education level group, and mean ages ± SD were (67.86, 85.38); there were also 18 women and 11 men in the low education level group, and mean ages ± SD were (70.16, 83.84). In this study, we used two statistical analyses to detect perfusion deficit regions, i.e., Contextual Clustering Method and SPM 99 Module. The results were showed in Fig.2、Fig.3 and Fig.4, respectively. Comparing the two different education level subjects, the left temporal perfusion deficit was significantly greater in the group with the high level of education (see Fig. 2). A significant rCBF decrease was found in the left temporal lobe (see Fig. 3) for the high education group. Regions in red and yellow represents the significant ($p<0.005$) decreases of activity in the high education group in comparison with the low education group (see Fig. 4). The simulated annealing procedure works iteratively to either accept or reject the change of states between the temperature extremes. We started with a high temperature and exploring the state space, and gradually decreasing the temperature to zero. The parameter γ with the empirical values was set as γ=0.03. The algorithm converged in 60-80 iterations in this study. The good results were presented settling on the global optimum. But its time consumption is the main drawback of this approach.

Conclusion

SPM99 is a very popular tool in neuroimage analysis, which is user friendly and strongly supported, and is chosen as the comparative method in this study. It offers the objective and voxel by voxel analysis for investigating regional cerebral blood flow (rCBF). The other method designed and used in this study is the contextual clustering algorithm. It optimizes the contextual information by using the simulated annealing procedure and also reports good lesion detection for the SPECT images. Although the results of the two different analyses were similar, the sensitivities obtained in the contextual clustering method were higher than SPM99 which utilized many user-specified parameters [21].

Of course, AD occurs independently of education. However, the level of education in some way will affect the cognitive capacity of the individuals. In other words, though the cognitive capacities of the two kinds of AD patients are the same, the patient with high education level suffers from a severer perfusion deficit than the corresponding low education level patient. The goal of the present study is to find the change areas of rCBF between the images from two different education level patients with AD. We hope our case report could provide valuable information for the abnormality studies of the brain SPECT imaging in the future.

Acknowledgements

The authors would like to express their appreciations for the supports under contract NSC90-2213- E006-126 from the National Science Council and by academic excellence project on Neural Plasticity under contract 89-B-FA08-1- 4 from the Ministry of Education.

Reference

[1] Yaakov Stern, Gene E. Alexander, Isak Prohovnik, and Richard Mayeux, "Inverse Relationship Between Education and Parietotemporal Prefusion Deficit in Alzheimer's Disease," *Ann Neuro,* 32: 371-375, 1992.

[2] Eero Salli, Hannu J. Aronen, Sauli Savolainen, Antti Korvenoja, and Ari Visa, "Contextual clustering for Analysis of Functional MRI Data," *IEEE Transactions on medical image*, 20(5): 403-414, 2001.

[3] Darius Soonawala, Tania Amin, Klaus P. Ebmeier,1 J. Douglas Steele, Nadine J. Dougall, Jonathan Best, Octave Migneco, Flavio Nobili, and Klemens Scheidhauer. "Statistical Parametric Mapping of 99mTc-HMPAO-SPECT Images for the Diagnosis of Alzheimer's Disease: Normalizing to Cerebellar Tracer Uptake," *NeuroImage,* 17: 1193-1202, 2002.

[4] O. Migneco, M. Benoit, P. M. Koulibaly, I. Dygai, C. Bertogliati, P. Desvignes, P. H. Robert, G. Malandain, F. Bussiere, and J. Darcourt. "Perfusion Brain SPECT and Statistical Parametric Mapping Analysis Indicate That Apathy Is a Cingulate Syndrome: A Study in Alzheimer's Disease and Nondemented Patients," *NeuroImage,* 13: 896-902, 2001.

[5] SPM has an extensive web site at: http://www.fil.ion.ucl.ac.uk/spm

[6] J. Ashburner, P. Neelin, D. L. Collins, A. Evans, and K. Friston. "Incorporating Prior Knowledge into Image Registration," *NEUROIMAGE*, 6: 344-352, 1997.

[7] Ming Zhu, "Volume Image Registration by Cross-Entropy Optimization," *IEEE Trans. on Medical Imaging*, 21(2): 174-180, 2002.

[8] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, 220(4598): 671-680, 1983.

[9] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Statist. Soc. B*, 48: 259-302, 1986.

[10] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Processing*, 40(4): 901-914,1992.

[11] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Statist. Soc. B*, 36: 192-236, 1974.

[12] Fox, G. C. "A review of automatic load balancing and decomposition methods for the hypercube," in M. Schultz, editor, *Numerical Algorithms for Modern Parallel Computer Architectures*, 63-76. Springer-Verlag, New York, 1988.

[13] Hajek, B. "Cooling schedules for optimal annealing," *Mathematics of Operational Research*, 13: 311-329, 1988.

[14] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *Equation of State Calculations by Fast Computing Machines, J. of Chem. Phys*., 21(6): 1087-1092, 1953.

[15] Metropolis, N., A. Rosenbluth, M., Rosenbluth, A. Teller, and E. Teller. "Equations of State Calculations by Fast Computing Machines," *J. Chem. Phys.,* 21: 1087-1092, 1958.

[16] E.A. Stamatakis, M. F. Glabus, D. J. Wyper, A. Barnes, and J. T. L. Wilson, "Validation of Statistical Parametric Mapping (SPM) in Assessing Cerebral Lesions: A Simulation Study," *NEUROIMAGE,* 10: 397-407,1999.

[17] Siegmund, D.O and Worsley, K.J.. "Testing for a signal with unknown location and scale in a stationary Gaussian random field". *Annals of Statistics*, 23: 608-639, 1995.

[18] J-B. Poline, K.J. Worsley, A.C. Evans, and K.J. Friston, "Combining Spatial Extent and Peak Intensity to Test for Activations in Functional Imaging," *NEUROIMAGE,* 5: 83-96, 1997.

[19] Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. "Comparing functional (PET) images: The assessment of significant changes". *J. Cereb. Blood Flow Metab.,* 11: 690-699, 1991.

[20] K. J. Friston, A. Holmes, J-B. Poline, C. J. Price, and C. D. Frith, "Detecting Activations in PET and fMRI: Levels of Inference and Power," *NEUROIMAGE,* 40: 223-235, 1996.

[21] Jagath C. Rajapakse, Senior Member, and Jayasanka Piyaratna, Student Member, "Bayesian Approach to Segmentation of Statistical Parametric Maps," *IEEE Transactions on Biomedical Engineering*, 48(10): 1186-1194, 2001.