Volumetric Congruent Local Binary Patterns for 3D Neurological Image Retrieval

Sidong Liu¹, Weidong Cai¹, Lingfeng Wen^{1,2}

¹ BMIT Research Group, School of IT,

University of Sydney, Australia

²Department of PET and Nuclear Medicine,
Royal Prince Alfred Hospital, Sydney, Australia

David Dagan Feng^{1,3,4}

³CMSP, Department of EIE,

Hong Kong Polytechnic University, Hong Kong

⁴Med-X Research Institute,

Shanghai Jiao Tong University, China

Abstract—The high thought put and increasingly accumulated data size of the 3D neuroimaging datasets have posed great challenges for neuroimaging data retrieval. To efficiently manage such large datasets, we proposed a volumetric congruent local binary pattern (vcLBP) algorithm for 3D neurological image retrieval. The vcLBP-based feature descriptor could describe the volumetric imaging data with higher robustness and meanwhile effectively compress the feature space by using the unique rotation, reflection and translation invariant patterns. We evaluated the proposed vcLBP algorithm using 132 sets of 3D positron emission tomography (PET) brain imaging data and the preliminary results suggested that our approach could effectively reduce the feature dimensions while achieving better results than other 3D feature descriptors. This vcLBP algorithm has a potential to be widely used in many other applications, such as image classification, content analysis and data mining.

Keywords-3D image analysis, image retrieval, vcLBP

I. Introduction

Functional neuroimaging, such as positron emission tomography (PET) and single photon emission computed tomography (SPECT), provides important insights into neuroscience and is a fundamental component of clinical neurological diagnosis. Instrumentation advances have led to the introduction of PET-CT scanners into clinical practice and with these advances there has been a marked increase in the size of the neuroimaging datasets. Such large neuroimaging datasets pose great challenges in image management and retrieval. Traditional keyword-based image retrieval is time-consuming and prone to errors. In addition, the visual and physiological features cannot be comprehensively described by keywords. Therefore, efficient management and analysis of such large datasets have prompted research in the field of content-based image retrieval (CBIR) [1].

Various studies have been reported on neuroimaging data retrieval. Kim *et al.* [2] proposed a dynamic brain PET retrieval system using the volume of interest (VOI) based pixel-wise tissue time activity curve (TTAC) clustering approach. Their approach focused on dynamic data and relied heavily on VOI segmentation. Unay *et al.* [3] developed a structure-based region of interest (ROI) retrieval system for brain Magnetic Resonance (MR) images. They used the local binary patterns (LBP) descriptors in their study and the results suggested that the dominate LBPs were more robust and efficient than the conventional LBPs. Batty *et al.* [4] designed a brain PET image

retrieval system based on the texture feature extracted by Gabor filters from ROIs. Both Unay's and Batty's systems [3]-[4] were based on the 2D feature descriptors. However, the neuroimaging data are commonly collected in 3D format and these 2D-oriented approaches are not able to capture the volumetric variations in 3D imaging data. There are also reports on 3D neuroimaging retrieval systems. Liu et al. [5] proposed a 3D gray level co-occurrence matrices (3D-GLCM) based feature extraction approach for the disorder-orientedmask (DOM) based retrieval of multiple types of neurological disorders. Recently, Qian et al. [6] investigated four 3D feature descriptors, namely, 3D-GLCM, 3D-wavelet, modified 3D-LBP and 3D-Gabor filters, and evaluated them using 100 MR brain images. They found that 3D-LBP could achieve more accurate retrieval than the other three approaches and was also computational efficient when using a small number of sampling voxels in the local neighbourhood.

We believe the 3D-LBP algorithm has two drawbacks and can be further improved. First, the feature dimension increases exponentially as the number of sampling voxels grows. This bottleneck makes 3D-LBP impractical when using a large number of sampling voxels. Second, 3D-LBP is subject to rotation, reflection and spatial translation. To overcome these drawbacks, in this paper, we proposed a volumetric congruent LBP (*vc*LBP) algorithm which could derive the unique rotation, reflection, translation invariant patterns based on the spatial relationships between voxels. Our approach could use much less patterns to describe the 3D imaging data meanwhile enhance the robustness of the LBPs compared to the conventional 3D-LBP algorithm.

In Section 2, we described the conventional LBP and our proposed vcLBP algrithms. Section 3 introduced the vcLBP-based neuroimaging retrieval framework, followed by the experiments and results in Section 4. Finally, we concluded in Section 5.

II. LOCAL BINARY PATTERN

A. LBP for 2D Images

LBP feature descriptor for 2D image data was designed by Ojala *et al.* [7]. For a monochrome texture image, a texture T in a local neighborhood can be defined as the joint distribution of the gray value of P(P>1) pixels, as in

$$T = t(g_c g_0, ... g_{P-1})$$
 (1)

where g_c is the gray value of the centre pixel of the local neighbourhood and $g_p(p=0,...,P-1)$ represents the gray values of P samples around the central pixel. All the values in T are then subtracted by the value of the central pixel, g_c , giving

$$T = t(g_{c}, g_{0} - g_{c}, ... g_{P-1} - g_{c}).$$
 (2)

Assuming the difference between $g_p - g_c$ is independent of g_c , and $t(g_c)$ describes the overall gray tone which is not related to T, thus T can be approximated by

$$T \approx t(g_0 - g_c, ... g_{P-1} - g_c).$$
 (3)

Note that the signed differences, $g_P - g_c$ are not affected by luminance, so to achieve gray-scale invariant, only the signs of $g_P - g_c$ rather than their exact values are considered, as in

$$T \approx t(s(g_0 - g_c), ..., s(g_{P-1} - g_c)),$$
 (4)

where s(x) = 1 if $x \ge 0$; and s(x) = 0 if otherwise. T is then converted into a unique decimal number that characterizes the spatial structure of the local binary pattern. Fig. 1 shows an example of an eight-sample local binary pattern. To achieve rotation invariant, all the unique patterns need to be identified (when P = 8, there are 36 unique patterns [7]), so we can find exact one match for any extracted LBP and compute the histogram based on these unique patterns.

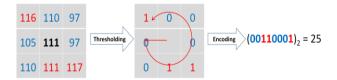


Figure 1. Extract and encode a 2D LBP.

B. State-of-the-art LBP Algorithms for 3D Imaging Data

Extending the 2D LBP algorithm to 3D imaging data seems to be very straightforward. For a local neighborhood, one can simply take P equidistant voxels around the central voxel, and then apply thresholding and encoding to identify the local binary patterns. However, there are two problems. First, as the number of sampling voxels grows, the length of the histogram, i.e., the dimension of the feature vector, increases exponentially. For example, if we take 18 voxels equidistant to the central voxel, we need 218 bins to record all the possibilities of the LBPs. If we take rotation, translation and reflection invariance into consideration, the possibilities can be greatly reduced, yet here comes the second problem which is how to achieve translation, rotation and reflection invariant LBP. For 2D LBP, it is very trivial to identify the unique rotation invariant patterns because the rotation degree of freedom is only 1. However, for 3D images, the rotation degree of freedom is 3 and the number of sampling voxels for 3D-LBP is also much larger than that for 2D-LBP.

Many attempts have been made to break the feature dimension bottleneck. Unay *et al.* [3] cut the 3D data into 2D planes. They then applied the 2D-LBP algorithm to derive the histograms from each of landmark planes and finally

concatenated them into one histogram. This approach can greatly reduce the feature space, but it remains 2D-oriented, since the LBPs extracted from landmark planes cannot capture the inter-plane volumetric variations which are the essential to characterize 3D images.

Zhao et al. [8] proposed the LBP-TOP (three orthogonal planes) algorithm which decomposed the local neighborhood volume into three orthogonal planes intersecting at the centre. The 2D local binary patterns were extracted from the three orthogonal planes and three histograms based on each of the planes were computed individually and then concatenated into a longer histogram as the final representation of input image. Fig. 2 illustrates the process of computing the final histogram using LBP-TOP. LBP-TOP can effectively reduce the feature dimension. However, the LBP-TOP algorithm has several drawbacks. First, it is not rotation/reflection/translation invariant. Second, the connection information between individual planes which is very important to characterize the local patterns is missing. Furthermore, LBP-TOP also contains redundant information. The intersect points of these three orthogonal planes, as marked by the yellow stars in Fig. 2, were counted twice for computing the histogram.

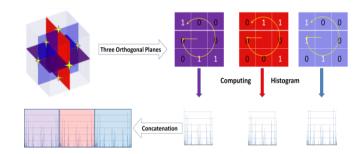


Figure 2. Histogram concatenation using the LBP-TOP algorithm.

Zhao *et al.* [8] also proposed a one-axis rotation invariant 3D-LBP algorithm for dynamic texture analysis of facial expressions. However, since this algorithm is designed for video, it allows the frames to only rotate around the timeline-axis. Fehr *et al.* [9]-[10] introduced a frequency transform based approach for the approximate computation of the uniform LBPs. However, their uniform LBPs are data-dependent and cannot provide complete gray-scale invariance [9]. To overcome these problems, they further proposed the *r*LBP which assigned the LBP to a template by calculating the correlation of the LBP and the fixed LBP templates of all angles. These templates were based on the spherical harmonic coefficients which also required the pre-computed discrete approximations of the harmonic base functions [10].

C. Volumetric Congruent LBPs

The abovementioned LBP approaches for 3D data [8]-[10] either lose too much information or not fully support rotation/reflection/translation invariant LBPs. To our knowledge, we are the first to propose the fully 3D rotation/translation/reflection invariant LBP algorithm. We name this algorithm the volumetric congruent LBP, denoted by vcLBP.

LBP encodes a pattern into a unique number, so each bin of the LBP histogram represents a unique pattern. It makes no sense to arbitrarily integrate two consecutive bins to generate a bigger bin. For example, (01011111) is next to (01100000) in the histogram, but the patterns they represent are far from similar. However, it makes perfect sense to put these patterns, as shown in Fig. 3, into one bin. All these patterns in the braces (Fig. 3-b) are generated by rotating, flipping or spatially translating the mother pattern on the left (Fig. 3-a). In other words, they are congruent and they can be abstracted by the pattern template on the right (Fig. 3-c).

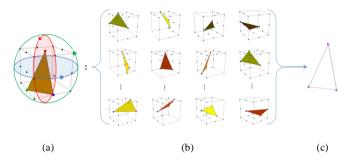


Figure 3. Congruent pattern abstraction.

Now the problem becomes how to find the congruent patterns. To prove any two geometric patterns constructed by voxels are congruent, we need to show that they were related by one isometry [11], which is a distance-preserving map between two metric spaces. Therefore, we need to compare all the edges in the first pattern and those in the second pattern. If all of the edges in first pattern could be mapped perfectly to the edges in the second pattern, then we can conclude these two patterns are related by an isometry hence they are congruent.

For one pattern A with n vertexes, we can represent it by an $n \times n$ symmetric, distance-preserving matrix, D_A , whose entry $d_A(i, j)$ is the distance between the i^{th} vertex to the j^{th} vertex of A. If there is another pattern B with same number of vertex as A, and satisfying

$$d_B(p(i), p(j)) = d_A(i, j)$$
(5)

for all $i, j \in A$, where p(x) is the mapping function from A to B, then A and B are related by

$$D_A = PD_B P^{-1} \tag{6}$$

where P is the permutation matrix which rearrange the order of vertexes of A to match A up with B. We do not know P, but we know the eigenvalues of A and B will always be the same [12]. To find P, we can first decompose A and B as:

$$D_A = Q_A E Q_A^{-1} \tag{7}$$

$$D_B = O_B E O_B^{-1} \tag{8}$$

where E is the diagonal matrix of eigenvalues of both D_A and D_B . P can be computed as:

$$P = Q_B Q_A^{-1} \tag{9}$$

If there exists a P, then A and B are congruent. For all entries of P(x, y) equal to 1, p(x) maps the x^{th} vertex of A to the y^{th} vertex of B, i.e.:

$$p(x) = y \tag{10}$$

Assuming n sampling voxels were considered, we should exhaustively test all 2ⁿ possibilities of the LBPs to identify the unique patterns. When we take 18 equidistant neighbours around the central voxel (radius of the sampling sphere is 1 voxel distance, and the values of the neighbours that do not fall exactly on voxels are approximated by the value of the nearest voxel), we can identify 6,426 unique patterns.

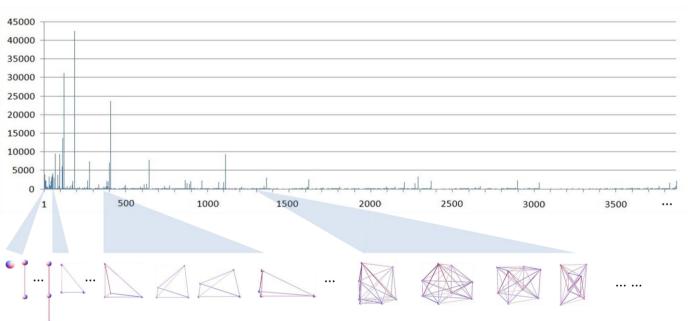


Figure 4. The illustration of a vcLBP-based histogram and the congruent patterns.

Unlike the 3D-LBP, we no longer translate the LBP into a number but built a lookup table which stores the unique values of the unique pattern templates. Fig. 4 shows one example of a histogram computed from a 3D neurological image. The vertical axis is the number of occurrences of the patterns and the horizontal axis stands for the indices of the patterns. Several *vc*LBP templates and their corresponding bins are illustrated at the bottom of Fig. 4.

III. VCLBP-BASED RETRIEVAL FRAMEWORK

A. Neuroimaging Dataset

The PET brain datasets, as summarized in Table I, were acquired on a CTI ECAT 951R whole body PET scanner, at the Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Sydney. Two types of neurodegenerative disorders, Alzheimer's disease (AD) and frontal-temporal dementia (FTD), were investigated in this study. We divided the dataset into two groups according to the ages of the patients. The early-onset group contained the patients whose ages were less than 65. Patients who were 65 and over were classified into the late-onset group.

TABLE I. P.	ATIENT DATASET SUMMARY
-------------	------------------------

	EARLY-ONSET		EARLY-ONSET	
Disorder	NO. of Cases	Age	NO. of Cases	Age
	(M:F)		$(\mathbf{M}:\mathbf{F})$	
AD	37 (11:26)	34-64	16 (9:7)	66-79
FTD	20 (13:7)	40-64	30 (16:14)	65-82
Normal	18 (5:13)	27-64	11 (9:2)	66-75
Controls				
Total	75 (29:46)	27-64	57 (34:23)	65-82

B. Data Pre-processing

To reduce the impact of the unreliable standard uptake values in PET, the cerebral metabolic rate of glucose (CMRGlc) parameters were first derived from raw static 3D FDG ([18F]2-fluoro-deoxy-glucose) PET images with the autoradiographic (ARG) algorithm [13]. Arterialized-venous blood samples were taken at 10 min and 45 min post injection to calibrate a population-based input function [14]. To eliminate the inconsistencies between different cases, the generated CMRGlc parametric image volume with dimensions of 128×128×56 were spatially normalized to a PET brain template with standard dimensions of 91×109×91, using the SPM2 package (Wellcome Trust Centre for Neuorimaging, London, U.K.) [15].

C. Feature Extraction

The *vc*LBP algorithm was used to extract the histograms from the spatially normalized CMRGlc images. Fig. 5 shows the algorithm of the extracting a *vc*LBP-based histogram from a given image. We iterated the voxels in the whole image to extract the local binary patterns at each voxel location. For an extracted local binary pattern, we computed its eigenvalues and then mapped it to the corresponding unique pattern template and read the ranking value of that unique pattern template from the look up table. Subsequently, the histogram element of that ranking value would be added by 1.

We extracted the histograms from all the neuroimaging data in the dataset and then stored them in the feature database. When we performed the retrieval of a query, we first extracted the feature from the query in the same way as we did for imaging data in the dataset and then compared it to all the histograms in the feature database.

```
FUNCTION feature_extraction (f(x))

SET x, the 3D coordinates of a voxel in the image f(x)

SET b, the local binary pattern

SET u, the unique pattern template

SET t, the value of the unique pattern in the lookup table

SET t, an empty array storing the vcLBP-based histogram

FOR x = x_0, ..., x_N, N is the total number of voxel in f(x)

b = B(x); extract the local binary pattern at x

u = U(b); map the pattern to a unique pattern template

t = T(u); read the value of the template from the lookup table

h[t] + +; compute the histogram

ENDFOR

RETURN h
```

Figure 5. The Pseudo Code of vcLBP-based Feature Extraction.

D. Performance Evaluation

The retrieval was conducted by the leave-one-out strategy on the whole dataset using query by example paradigm. The similarity was calculated by the Euclidean distance and the performance was evaluated by the average precision of the top-5 retrieved results. We used the diagnosis from the imaging studies as the ground truth. For two different neurological disorders, there can be a degree of overlap because certain brain regions could be sensitive to both disorders, for example, late-stage AD and FTD both exhibit a degenerative pattern in frontal cortex area. To balance the interference of these common regions shared by different disorders so as to objectively describe the retrieval results, the following relevance criteria were used. If the query is a dementia case and the retrieval result is also a dementia case but diagnosed as a different type, we then set the relevance score to 0.25. If the retrieval result belongs to the same group as the query, then the relevance score is set to 1.0. If the query is a dementia case and the retrieval result is a normal case, or vice versa, the relevance score is set to 0.

IV. EXPERIMENTS AND RESULTS

Table II shows the length of the histogram for each LBP algorithm using different number of sampling voxels. In this study, we used 18 equidistant sampling voxels around the central voxel to characterize the 3D local patterns. Both the vcLBP and the LBP-TOP could effectively reduce the feature dimensions. The histogram extracted by vcLBP is 8 times longer than LBP-TOP but 40 times shorter than 3D-LBP. Note the LBP-TOP algorithm takes 8 voxels from each of the three orthogonal planes, so literally the total number of sampling voxles is 24. However, these three planes intersect at 6 voxels which are counted twice when computing the histogram. Therefore, these voxels are redundant and only 18 voxels are actually used by the LBP-TOP algorithm.

TABLE II. THE NUMBER OF SAMPLING VOXELS AND THE LENGTH OF HISTOGRAM FOR DIFFERENT LBP ALGORITHMS

Algorithm	3D-LBP	LBP-TOP	vcLBP
r = 1, s = 6	$2^6 = 64$	$3 \times 2^4 = 48$	10
r = 1.4, $s = 12$	$2^{12} = 4,096$	$3 \times 2^4 = 48$	22
r = 1. s = 18	$2^{18} = 262.144$	$3 \times 2^8 = 768$	6.426

r: the radius of the sampling sphere *s*: the number of sampling voxels

To evaluate the effectiveness of our proposed *vc*LBP algorithm, in this study, we compared our proposed *vc*LBP to the LBP-TOP and the conventional 3D-LBP algorithms, as well as a baseline 3D-GLCM algorithm. We applied all these algorithms to both the early-onset and the late-onset patient groups.

TABLE III. EXPERIMENTAL RETRIEVAL RESULT SUMMARY

	3D-LBP	LBP-TOP	vcLBP	3D-GLCM	
	EARLY-ONSET				
AD	61.49%	62.84%	68.78%	65.14%	
FTD	61.50%	59.50%	62.00%	56.75%	
	LATE-ONSET				
AD	50.63%	51.56%	55.00%	41.88%	
FTD	68.00%	66.17%	68.67%	63.17%	

Table III summarizes our preliminary results. The retrieval performances of different disorders in different age groups were very different. For AD cases, the retrieval precision in the early-onset group was much higher (10.9% - 13.8%) than the late-onset group, while it was opposite for FTD cases. This could be explained by that the proportion of the AD cases in early-onset group was higher than that in the late-onset group, which gave more chance to retrieve an early-onset AD case. On the contrary, the higher proportion of FTD cases in the late-onset group led to a higher retrieval precision (6.5% - 6.7%) of the late-onset FTD cases.

Among all these four algorithms, we found *vc*LBP had the best performance and all the LBP algorithms achieved better results than the baseline 3D-GLCM algorithm in most cases, except for the 3D-LBP and LBP-TOP algorithms for the early-onset AD retrieval, 3.6% and 2.3% lower than 3D-GLCM, respectively. This might be explained by that 3D-GLCM could preserve the gray-scale contrast while LBPs could not, thus 3D-GLCM could outperform the LBP algorithms in rare occasions. However, the retrieval precisions for *vc*LBP were always higher than 3D-GLCM, and the largest difference was 13.1% for late-onset AD retrieval.

When we compared the three LBP algorithms with each other, we found that 3D-LBP was superior to LBP-TOP in FTD case retrieval and both of them had very similar performance to the vcLBP. However, for AD case retrieval, vcLBP achieved much higher precision than the other two.

V. CONCLUSION

In this paper, we presented a new feature descriptor for 3D imaging data, namely *vc*LBP. Our proposed *vc*LBP algorithm resolves the two drawbacks of the conventional 3D-LBP

algorithm. Firstly, our vcLBP fully supports the rotation/reflection/translation invariant pattern templates which can enhance the robustness of LBPs. Secondly, the vcLBP algorithm can effectively reduce the feature dimensions. In conclusion, the vcLBP algorithm is an advanced feature descriptor for 3D imaging data compared to the conventional 3D-LBP and LBP-TOP algorithms, and it also performs better than the 3D-GLCM algorithm in the application of neuroimaging retrieval. We believe that our approach has a great potential to be generally applied to other applications, such as 3D content analysis, data mining and image classification.

REFERENCES

- [1] H. Muller, N. Michoux, D. Bandon and A. Geissbuhler, "A Review of Content-Based Image Retrieval Systems in Medical Applications – Clinical Benefits and Future Directions," *Int. J. of Med. Inf.*, vol. 73, pp.1-23, 2004.
- [2] J. Kim, W. Cai, D. Feng and H. Wu, "A New Way for Multi-Dimensional Medical Data Management: Volume of Interest (VOI)based Retrieval of Medical Images with Visual and Functional Features," *IEEE Trans. on Inf. Tech. in Biomed.*, vol. 10, issue. 3, pp. 598-607, 2006.
- [3] D. Unay, A. Ekin, and R. S. Jasinschi, "Local Structure-based Region of Interest Retrieval in Brain MR Images," *IEEE Trans. on Inf. Tech. in Biomed.*, vol. 14, issue. 4, pp.897-903, 2010.
- [4] S. Batty, J. Clark, T. Fryer, and X. Gao, "Prototype System for Semantic Retrieval of Neurological PET images," *Proc. of the int. Conf. on Med. Imaging and Informatics (MIMI 2007)*, pp. 179-188, 2008.
- [5] S. Liu, W. Cai, L. Wen, S. Eberl, M. J. Fulham, & D. Feng, "A Robust Volumetric Feature Extraction Approach for 3D Neuroimaging Retrieval," *Proc. of the Engineering in Med. and Bio. Society Annual Conf.*, pp. 5657-5660, 2010.
- [6] Y. Qian, X. Gao, M. Loomes, R. Comley, B. Barn, R. Hui and Z. Tian, "Content-based Retrieval of 3D Medical Images," *Proc. of the 3d Int. Conf. on eHealth, Telemedicine and Social Medicine*, pp. 7-12, 2011.
- [7] T. Ojala, M. Pietikaien, and T. Maenpaa, "Multiresolution Gray Scale and Rotation Invariant Texture Analysis with Local Binary Patterns," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, issue 7, pp. 971-987, 2002.
- [8] G. Zhao and M. Pietikainen, "Dynamic Texture Recognition Using Local Binry Patterns with an Application to Facial Expression," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, issue 6, pp. 915-928, 2007.
- [9] J. Fehr, "Rotation Invariant Uniform Local Binary Patterns for Full 3D Volume Texture Analysis," *Proc. of Finnish Signal Processing*. Symposium, 2007.
- [10] J. Fehr and H. Burkhardt, "3D Rotation Invariant Local Binary Patterns," *Proc. of 19th Int. Conf. on Pattern Recognition*, 2008.
- [11] F. S. Beckman and D. A. Quarles, Jr., "On Isometries of Euclidean Space," Proc. of American Mathematics Society, vol. 4, pp. 810-815, 1053
- [12] J. H. Wilkinson, The Algebraic Eigenvalue Problem, pp. 54, Clarendon Press, Oxford, 1965.
- [13] G.D. Hutchins, J.E. Holden, R.A. Koeppe, et. al., "Alternative approach to single-scan estimation of cerebral glucose metabolic rate using glucose analogs with particular application to ischemia", *J. of Cerebral Blood Flow and Metabolism*, vol. 4, pp.35-40, 1984.
- [14] S. Eberl, A. R. Anayat, R. Fulton, P. K. Hooper and M. J. Fulham, "Evaluation of Two Population-based Input Functions for Quantitative Neurological FDG PET Studies," *Euro. J. of Nuclear Medicine*, vol. 24, issue. 3, pp. 299-304, 1997.
- [15] R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. T. Ashburner and W. D. Penny, Human Brain Function. Amsterdam; Boston: Elsevier Academic Press, 2004.