Retrieval of 3D Medical Images via Their Texture Features

Xiaohong Gao, Yu Qian, Martin Loomes, Richard Comley, Balbir Barn, Alex Chapman, Janet Rix
Middlesex University, London, NW4 4BT, UK
{x.gao, y.qian, m.loomes, r.comley, b.barn, a.chapman, j.rix}@mdx.ac.uk

Rui Hui, Zengmin Tian
General Navy Hospital, Beijing, P.R. China
huirui2002@163.com, tianzengmin@vip.sina.com

Abstract **-- While content-based image retrieval has been researched for more than two decades, retrieving 3D datasets has been progressing considerably slower, especially in the application to the medical domain. This is in part due to the limitation of processing speed while trying to retrieve high-resolution datasets in real-time. Another barrier is that most existing methods have been developed based on 2D images instead of 3D, leaving a gap to be filled. At present, a significant number of exploitations are focusing on the extraction of 3D shapes. As it happens, it appears that, to a large extent, the remaining information tends to be equally important in the task of clinical decision making. With this in mind, in this paper, a texture-based online system, MIRAGE, has been developed to facilitate CBIR for 3D images. Specifically, four texture-based approaches stemming from 2D forms are studied extensively through the application to 3D images using a collection of MR brain images and are implemented, which include 3D Local Binary Pattern (LBP), 3D Grey Level Co-occurrence Matrices (GLCM), 3D Wavelet Transforms (WT) and 3D Gabor Transforms (GT). Based on the nature of the content, each approach has its own advantages and disadvantages. For example, in terms of retrieval precision of tumours and processing speed, LBP not only achieves precision rate of up to 78% but also can perform retrieval in real time with sub-second processing speeds, outperforming the others.**

Keywords – CBIR; 3D image retrieval; 3D texture extraction; MIRAGE system; 3D visualization.

I. INTRODUCTION

Due to the advances of medical imaging techniques, more and more images are in three (or higher) dimensional forms, allowing a coherent and collective view. Since many of these images are comprised of 2D slices, most current databases archive and index them in 2D form, especially for the systems that are indexed by their content. As a result, a number of limitations have arisen with the most significant one being that the information extracted from a single 2D slice cannot be representative due to the fact that slices are getting thinner and thereafter resolutions are getting higher.

On the other hand, at present, content-based retrieval for three dimensional (3D) images has been researched primarily to meet the demand for 3D pictures available over the internet. In this way, the main challenge facing the extraction of features from 3D images is that these features have to be invariant of viewing angles, i.e., invariant of rotation, in order to achieve higher retrieval hit of similar objects, even though sometimes they may not be visible from all the viewing angles. For example, if a query image is a 3D rabbit with a head facing the view, a good retrieval system should bring back relevant objects including those showing only its tails as an exact match. In other words, even if the view angle is at the back of the object, the matched objects can still be found. In addition, in 2D cases, the viewing angle is always at 0^o, being normal to the computer screen, by which most existing algorithms can fulfill this request. Also, many of the other characteristics of content-based image retrieval (CBIR) are shared between 2D and 3D, including scaling and translation of regions of interest. This has led to the shift of many current studies to focusing on the invariance of transformations (including rotation, scaling and translation) of objects, which has more to do with shapes.

A. 3D CBIR for Non-medical Images

Since the emerging of the internet in 1990, coupled with the advance of computer hardware, vast amounts of textual and imagery data are available online, prompting the creation of an array of text-based search engines, such as Google and Yahoo, in an attempt to filter the relevant data. For image data, however, thanks to their embedded information being inside the pictures, a text-based approach has its limitations, especially for those images that are not properly, if not at all, labelled. Consequently, CBIR has been researched both horizontally and vertically. As the trend continues, the progress in the last century (1994-2000) has been very well documented in [1], whereas the state of the art in the last decade (2001-2008) was reviewed by [2] with a number of future directions being identified. Generally, a CBIR system follows the procedures of development as shown next. Firstly, it extracts features of images in terms of their global visual information, such as colour, texture, and shape. Then, these features are represented using mathematical vectors that, in turn, are employed to index each image. Finally, when a query image is submitted, the system needs to extract these features from the query image and to perform the comparison with the feature database that has been stored in advance. In this way, the retrieval process of an image can be as fast as that in a text-based system since the similarity calculation is based on numerical data.

For 3D online images, the majority of approaches concerns with the features of shapes as an indexing key. For example, in [3], 3D Zernike descriptors have been developed to describe

shapes of objects, by taking advantages of polynomial representations, on which these descriptors are based, being invariant of transformations. To this end, a database has to constitute objects differentiated by shapes, such as airplanes, chairs, etc.. Similarly, in order to achieve transformation invariance, a graph-based shape descriptor is created in [4] in an effort to determine the way to calculate a similarity between 3D objects. Recently, the retrieval of 3D objects has been attempted using impact descriptors [5] attempting to capture the surrounding areas of a 3D shape in order to offer a histogram of time-space curvature, which are invariant of rotation and translation. Elsewhere, other shape-based 3D models are included in [6-9]. Because shape-based approaches only describe the surface of a 3D object, they tend to ignore the content inside that object. Depth based descriptors therefore have been developed as demonstrated in [10], which is however in principle, still capture the outliner of a shape at each depth (z-buffer).

More recently, the approach of scale invariant feature transformation, commonly known as SIFT, has attracted substantial attention. Originally developed for 2D images [11], SIFT has been extended to 3D spaces in an attempt to perform action recognition [12] in a video sequence and object recognition for an airport security checking [13].

B. 3D CBIR for Medical Images

Progress on CBIR for 3D images have been reviewed by many researchers [14] with several developed systems demonstrated, which are summarized in Table 1 and described in details below.

TABLE 1: 3D CBIR SYSTEMS OF MEDICAL IMAGES

Name/ Feature	Imaging Modality	Domain	Reference
QBISM / intensity-based	MRI/PET	Brain	Arya [15]
Pre-defined-semantic-based	CT	Brain	Liu [16]
MIMS / ontology-based	All	All	Chbeir [17]
Knowledge-based	All	All	Chu [18]
ILive – modality-based	All	All organs	Mojsilovic [19]
2D Texture-based	MR	Heart	Glatard [20]
FICBDS / Physiological information – based	Functional PET	Brain	Cai [21]
3D PET / lesion-based	PET	Brain	Batty [22]
MIRAGE / 3D texture-based	MR	Brain	Gao [23], Qian[24]

In the system of QBISM, 3D functional brain images are queried and visualized [15], by which intensity-based volume data are stored for spatial references, whereas Talairach brain atlas [25] is employed to construct a region-based retrieval. The key to this system is the application of volumetric data type, i.e., the Region or Volume being expressed as <x, y, z, value>, in the representation of image data, which in some cases, might be prone to noise.

In other cases, the retrieval task of 3D images can work well based on feature extraction [16] from 2D slices, whose success to a great extent, is dependent on the application fields of the created databases.

On the other hand, semantics based retrieval remains acceptable to images of all dimensions as evidenced by [19]. The strength of this work therefore lies in the approaches employed for categorization of images that bear semantically well-defined data sets. This task itself however in most cases poses greater challenges than semantic representation itself. Nevertheless, semantics based retrieval of medical images offers one of the current trends. Likewise, ontology-based [17] and knowledge-based approaches [18] can shorten the semantic gap to a certain extent between low level features and high level semantics, which in turn requires skilful expertise, i.e., in-depth knowledge, to interpret images and convert contents into textual descriptions.

For subject-based images that bear centralised characteristics, local features can play an important part in indexing and retrieving images. For example, the system of FICNDS [21] employs physiological kinetic features for retrieving images. Similarly, Batty and Gao [22] have employed binding potential (BP) values to index functional PET images. Although effective, this method is very discipline-defined and relies heavily on the additional supply of extra information. For example, in FICNDS, to define a tracer kinetic model, plasma time activity (PTA) curves should be obtained from a series of blood samples, which are not easily available for most of the images in a database. Although PTA can still be modeled by the application of control regions as applied by Gao et al, a sequence of images acquired over a period of time, say 90 minutes, are still needed, which again is not readily available in most of image repositories. Additionally, in essence, the establishment of kinetic models stems from the data of 2D slices, which may lose information in between slices. For a system that warehousing images of variety of domains, more general approach appears to be in demand in order to be sustainable.

A texture-based approach for retrieving of 3D+ cardiac images has been applied by Glatard [20] with the employment of a 2D Gabor filter. While working on 4D (3D + time) heart images, their adoption of a Gabor filter is again, in essence, a 2D form based on regions coupled with an extra parameter dedicated to myocardium features.

For application to medical images, a Volume of Interest (VOI) consists of not only boundary shapes, but also inside textures representing tissue properties of the VOI. The information extracted from these textures equally plays an important role in describing the VOI and is important to medical doctors at most of the time. Therefore these texture features should be taken into consideration in the representation of an object as well.

Apparently, it is possible to represent texture in 2D-based form, since a 3D dataset constitutes a stack of 2D slices.

However, using a slice-by-slice 2D approach suffers from the disappointment that some important information inter-laced within the volumetric data is missing. Thus, in terms of a 3D form of texture, this spatial structural information should be extracted from a cube instead of a surface or a square. Towards this end, while working on images of 3D brain, Gao et al [23] and Qian et al [24] have furthered four texture-based approaches into 3D form to the domain of medical image retrieval to extracting texture information that is subsequently utilized for indexing them in their developed system MIRAGE [26].

Specifically, in this study, the approach of Local Binary Pattern (LBP) [27] is addressed first because of its discriminative power and computational simplicity, and applied to a collection of 3D MR brain images for extracting texture information that is subsequently utilized for indexing them. Three other well-known methods in texture representation are also investigated, including Grey Level Co-occurrence Matrices (GLCM), Wavelet Transforms (WT) and Gabor Transforms (GT). The novelty of this work demonstrates the feasibility of 3D texture-based approaches for image retrieval while maintaining real time operation. This is achieved by the introduction of a pre-processing stage of a selection of potential VOIs into query datasets; by which, through the use of statistically analysis of the bilateral symmetry of a brain MR image, a potential VOI of a query can be detected in real time, preceding the extraction of 3D texture features and the calculation of similarities.

The remaining of the paper is hence structured in the following pattern. Section II explains the methods employed in the study, which is followed by Section III that shows the experimental results. The interface design is detailed in Section IV, which is succeeded by Section V providing conclusion and discussion. The last two sections give acknowledgment and references respectively.

II. METHODOLOGY

The development of a repository requires two main phases, which are ingestion and retrieval. In this investigation, at the phase of ingestion of the data, the collected data firstly undergo a pre-processing stage to normalize them into the same resolution before the indexing stage, as shown in the flow chart in Figure 1. As illustrated in the diagram, after spatial normalization of volumetric brain data into a standard template, the data are then divided into 64 non-overlapping equally sized blocks, from which, 3D texture features can be extracted to create a feature database. On the query side, a pre-processing stage is introduced to detect a potential VOI after spatial normalization from a query image. As a result, 3D texture features from a query can only extracted from these potential sub-blocks of VOIs, which, in the retrieval stage, are compared with the corresponding features in the feature database to obtain retrieval results. Details are elaborated in the following sub-sections.

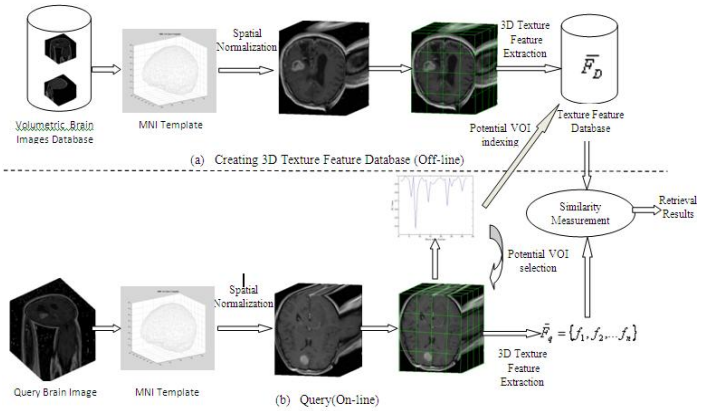


Figure 1. Framework of 3D MR image retrieval.

A. Spatial Normalization

By nature, data are collected from different sources, leading to the fact that brain images vary in both shape and size. In order to make inter-brains comparable, it is necessary to transform the dataset of each individual brain into a standard brain template. In this regard, the software of Statistical Parametric Mapping (SPM5) [28] is employed to spatially normalize a brain image into a template of either MNI T1 or T2 [29] depending on whether an image is acquired by an MR scanner of either T1 or T2 type. In this way, all the images in the database are of the same size with 157×189×69 voxels.

B. Extraction of Volumetric Textures

In order to describe local features from different parts of a brain, a 3D volumetric brain is divided into 64 non-overlapping equally sized blocks, giving 4 blocks along each of *x*, *y*, *z* axes respectively, as shown in Figure 1. Texture features are then extracted using 3D LBP to create a feature database, upon which image searching and retrieval are performed.

C. 3D Local Binary Pattern

The Local Binary Pattern (LBP) operator is derived from a general definition of texture in a local neighborhood (e.g., 8 × 8 pixels). In a 2D form, for each pixel in an image, a binary code is produced by thresholding its value with the value of a centre pixel. A histogram is then generated to calculate the occurrences of different binary patterns. To extend this approach to 3D images, similarly to [30], a 3D dynamic texture is recognized by concatenating three histograms obtained from the LBP on three orthogonal planes. When applied to our normalized brain images, they are in the plane of Left-Right (LR), Anterior-Posterior (AP), and Superior-Inferior (SI) respectively, as depicted in Figure 2.

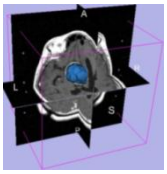


Figure 2.An example of three orthogonal planes in a 3D brain.

These three orthogonal planes intersect in a centre voxel. By selecting 8 neighbours as a local neighbourhood with the radius length being one voxel, fifty-nine uniformed LBP codes are subsequently extracted from the planes of SI, LR and AP respectively, again as illustrated in Figure 2, producing a 59 bin histogram for each plane by accumulating 59 binary patterns. Finally, the three histograms are concatenated to generate a 3D texture representation, giving the size of a feature vector being 177 (=59×3) elements.

D. Lesion Detection

The initial goal of the development of this 3D CBIR system is to search images with lesions of similar location, size or shape (all the collections of images are with lesions). Although a feature database has been implemented in advance, the processing of a query has to be conducted in real time. In other words, after a query is submitted to the system, 3D texture features should be extracted from its 64 sub-volumetric spaces together with the calculation of similarity distances. To this end, while maintaining the overall performance of retrieval, the detection of candidate lesions from sub blocks is carried out first to highlight the abnormalities, such as tumours, with an intension to speed up the retrieval process.

To do this, the characteristics of bilateral symmetry of a brain along its mid-plane (parallel to SI direction as shown in Figure 2) remain assumed. Similarly to [31], by comparing the left half with the right counterpart of a hemisphere along this middle symmetry plane, the abnormality can be envisaged to be singled out. Since a normalized brain image has been divided into 64 blocks, statistical features (e.g., mean, standard deviation, etc.) of each sub-block together with its mirror block are then calculated and compared to establish potentially abnormal sub-blocks.

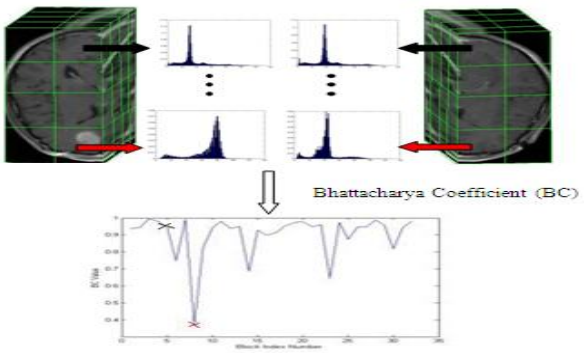


Figure 3.Potential VOI selection

As demonstrated in Figure 3, a normalized brain is divided into left *(L)* and right *(R)* parts by a sagittal plane, leading to 32 sub-blocks each, within which a grey level histogram is obtained. The Bhattacharya Coefficient (*BC*) [32] is thereafter computed between two normalized histograms H_L and H_R, which are obtained from two mirror symmetric sub-blocks as defined in Eq. (1).

$$BC(H_L, H_R) = \sum_i \sqrt{H_L(i) * H_R(i)} \quad (1)$$

The more similar H_L and H_R are, the closer to 1 the *BC* value is. On the other hand, less similar histograms tend to have smaller *BC* values. In total, 32 *BC* values are calculated from 32 paired mirrored symmetric sub-blocks that are plotted at the bottom of Figure 3. The horizontal axis points to the index numbers of sub-block pairs, whereas the vertical axis represents the corresponding *BC* values. Also shown in the figure are the *BC* values presenting the top normal sub-block pair marked with a black 'x', whereas the bottom abnormal sub-block pair marked with a red cross. Therefore, the mean value of the *BC* range works as a threshold to be applied to detect the potentially abnormal sub-block, i.e., where *BC* < *Threshold*.

After the affirmation of a lesioned VOI from a query is established, 3D texture features are extracted exclusively from this VOI of the query, and are later compared with the features from similar blocks of those images in the feature database in an attempt to search images with similar lesions in terms of textures.

E. Similarity Measurement

To measure the degree of similarity between two images *Q* and *I*, a distance function should usually be in place calculating the distance between features of the two images. For a 3D LBP, the histogram intersection is applied to measure features of histograms and is given in Eq. (2),

$$D(Q, I) = \sum_i \min(Q_i, I_i) \quad (2)$$

where *i* represents each bin in a histogram. The more similar they are between a query (Q) and an image (*I)*, the bigger the value of the *D* is. Therefore, the retrieved results are ranked in descending order based on the value of D.

III. EXPERIMENTAL RESULTS

A. Data Collection

In this study, the database contains over 100 3D MR brain images with lesions (e.g., tumour, biopsy) and detailed diagnosis. Each dataset has a resolution in a range between 256 × 256 × 22 mm^3 and 256 × 256 × 44 mm^3, and is in DICOM (Digital Imaging and Communications in Medicine) format with 16 bit grey-level resolution.

B. Results on Detection of Lesions

Since the location of a lesion region plays an important part in retrieving relevant datasets, the evaluation on the detection of lesion positions is carried out first. In Table 2, the first row is the labelling number of the location of a VOI assigned by the authors for the convenience of calculations, e.g., '1' refers to the abnormal part in the front top left part of the brain. The second row is the total number of images containing VOIs in the same positions in the database, whilst the number of correctly detected images by the approach of lesion detection as explained in Section II.D is given on the third row. Therefore, the overall performance in terms of VOI locations is calculated as the number of detected positive VOIs divided by the total positive VOIs and is 91.3% (=168/184).

TABLE 2 VOI DETECTION RATE

VOI Location	1	2	3	4	5	6	7	8	**Total**
Number of images	24	46	18	38	24	12	14	8	**184**
Correctly detected images	24	42	16	34	24	8	12	8	**168**
Correct Detection Rate (%)									**91.3**

C. Comparison with the Other Texture-based Approaches

The other three methods widely employed in texture representations are also exploited in this investigation by the extension to 3D, including Grey Level Co-occurrence Matrices (GLCM), Wavelet Transforms (WT), Gabor Transforms (GT), which are summarized next.

In 3D form, GLCM [33, 34] are defined as three dimensional matrices of a joint probability of occurrence of a pair of grey values separated by a displacement *d = (dx, dy, dz)*.

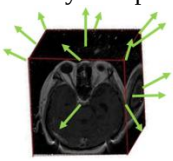
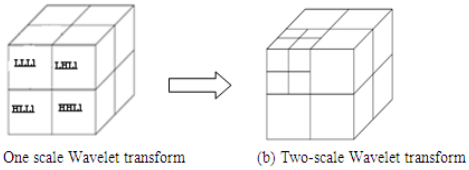


Figure 4.Thirteen directions in 3D GLCM.

For example, four distances with 1, 2, 4, and 8 voxels respectively and thirteen directions, as depicted in Figure 4, which are chosen in this study, will produce 52 (=4×13) displacement vectors, and thereafter 52 co-occurrence matrices. As a result, four Haralick texture features [35], being energy, entropy, contrast and homogeneity, are computed from each matrix, generating a feature vector of 208 components (=4 (measures) × 52 (matrices)).

On the other hand, the 3D WT provides a spatial and frequency representation of a volumetric image, which can be achieved by applying both high-pass (H) and low-pass (L) filters along all three dimensions. This is then followed by a 2 to 1 sub-sampling of each output volumetric image [36], giving rise to eight wavelet coefficients sub-bands (one low frequency sub-band and seven high frequency sub-bands) at each scale, as schematically presented in Figure 5(a). The process is subsequently repeated in the lowest frequency sub-band (LLL_1), generating a 3D wavelet transform of two scales as shown in Figure 5(b).



(a) One scale Wavelet transform (b) Two-scale Wavelet transform

Figure 5.One scale and two scales of 3D WT.

In this investigation, 2 scales of 3D WT, as shown in Figure 5(b) are chosen. The measurement of mean μ and standard deviation σ are then extracted for each sub-band. So that there are 30 features, i.e., 2 (scales) *7 (sub-bands in each scale) *2 (measures) +2 (measures in the lowest resolution) =30, derived from a Wavelet transform of 2 scales, yielding the dimension of a vector being 30.

With respect to Gabor Transforms, in order to extend GT into three dimension, a set of 3D Gabor filters are generated similar to [37, 38] to detect spatial orientations and scale tunable edges and lines (bar), which can be formulated as Eq. (3).

$$g(x,y,z,F,\theta,\phi) = \hat{g}(x,y,z)\exp\left[j2\pi(F\sin\theta\cos\phi x + F\sin\theta\sin\phi y + F\cos\theta z)\right] \quad (3)$$

where $\hat{g}(x,y,z)$ is a 3D Gaussian function, together with radial centre frequency F and orientation parameters (θ and ϕ), determining a Gabor filter in three dimensions.

In this study, the following parameters are defined, including four centre frequencies with F = {0.0442, 0.0625, 0.0884, 0.125} circle/voxel respectively, six orientation angles, i.e.., $\theta = \{0^0, 30^0, 60^0, 90^0, 120^0, 150^0\}$ and six values of ϕ, i.e., $\phi = \{0^0, 30^0, 60^0, 90^0, 120^0, 150^0\}$, which leads to the number of 144 (= 4*6*6) Gabor filters that are employed to extract texture features. Given a 3D volumetric texture $f(x,y,z)$, its 3D Gabor transform GT_i is defined by

$$GT_i = f(x,y,z) * g(x,y,z,F_i,\theta_i,\phi_i) \quad i = 1,2,3...144 \quad (4)$$

The mean μ and standard deviation σ of GT coefficients are then calculated which act as a representation of texture features from 144 Gabor transforms respectively. Therefore a feature vector includes 288 elements (= 4 (scales) *36 (orientations) *2 (measures)).

To calculate similarity distances from these three methods, a normalized Euclidean distance is employed to compare two 3D patterns in a feature space, as defined by Eq. (5).

$$D(Q,I) = \sqrt{\sum_i \left(\frac{Q_i - I_i}{\sigma_i} \right)^2} \quad (5)$$

where σ_i refers to the standard deviation of a set of representative features over the entire database and is therefore utilized to normalize each individual feature component. The retrieved 3D images are ranked in an ascending order of feature distances.

In summary, the above three 3D texture approaches together with LBP are applied to extract texture features from each sub-volumetric block. Furthermore, the dimension of a feature vector for a 3D brain remains to be the size of local features multiplied by 64, the number of the blocks each volumetric image is divided into, yielding 13312, 1920, 9216 and 11328 components for the approaches of 3D GLCM, 3D WT, 3D GT and 3D LBP respectively.

Subsequently, the performance of image retrieval is evaluated based on the measures of Precision (P) and Recall (R). Precision is defined as the fraction of retrieved images relevant to a query whilst recall is the fraction of relevant images retrieved. Precision and recall values are usually presented together in a Precision-Recall (P-R) graph that demonstrates the retrieval performance at each point in the ranking. In a P-R graph, the horizontal axis refers to a recall whereas the vertical axis shows the corresponding precision at each of the usual recall points, i.e., 10%, 20%,…,100% or 0.1, 0.2, …, 1. A single value, usually, the Mean Average Precision (MAP) value is employed to assess the overall performance for all queries and is calculated as

$$\text{Mean Average Precision(MAP)} = \frac{1}{M} \sum_{i=1}^{M} AP_i \quad (6)$$

where M is the total number of the queries, AP_i is the average precision for the i[th] query that is formulated as Eq. (6),

$$\text{Average Precision(AP)} = \frac{1}{N_r} \sum_{j=1}^{N_r} P_j \quad (7)$$

where N_r is the total number of relevant images in a dataset for a query, p_j is the precision when retrieving the j[th] relevant image.

Figures 6 and 7 depict the average Precision Recall Graph for ten queries across the whole datasets with Figure 6 showing the results without a pre-processing stage of VOI selection whilst Figure 7 with the pre-processing stage.

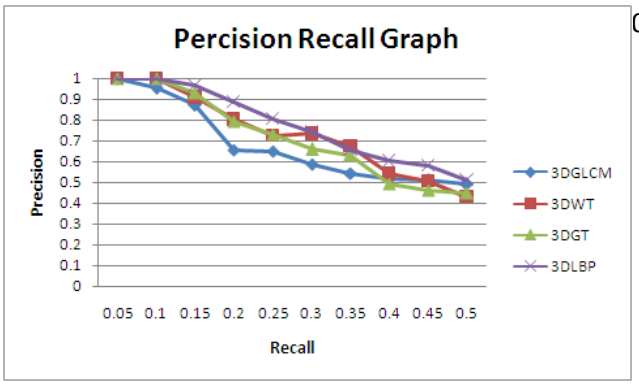


Figure 6. Average precision recall graph for ten queries without VOI selection.

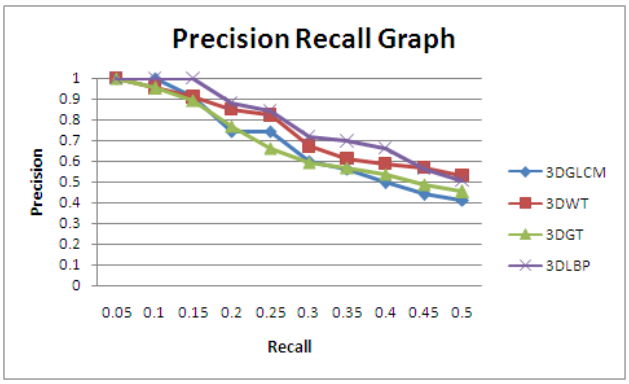


Figure 7. Average precision recall graph for ten queries with VOI selection.

Overall, the mean average precision (MAP) at 0.5 recall rate for ten queries cross the whole database by using the approaches of 3D GLCM, 3D WT, 3D GT and 3D LBP are shown in the following table.

TABLE 3 VALUE OF MEAN AVERAGE PRECISION

Methods	Without VOI selection	With VOI selection
3D GLCM	0.677	0.690
3D WT	0.731	0.749
3D GT	0.714	0.691
3D LBP	0.774	0.786

Comparing the value of MAP with and without potential VOI selection, the methods of 3D GLCM, 3D WT and 3D LBP with potential VOI selection show a slightly improved performance with bigger MAPs.

Figure 8 visualizes the retrieved results by using the four approaches with a pre-processing stage of VOI selection. The query image with a tumour in the middle is displayed in 3D fashion and 3 slices appearing in 3 orthogonal planes on the top row, i.e., in axial, sagittal, and coronal directions. The retrieval results are visualized by using an open source software 3D Slicer [39].

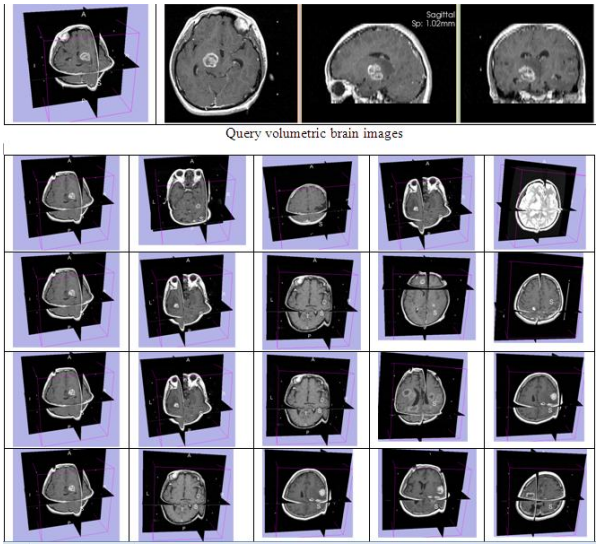


Figure 8. Retrieved results in top 5 ranking from
3D GLCM (row 1), 3D WT(row 2), 3D GT (row 3), and 3D LBP (row 4).

D. Query Time

It is understandable that retrieving images in 3D form might not be performed in real time, one of the drawbacks in the development of CBIR systems for images of higher dimensions. Table 4 demonstrates the average querying time, amounting to the period spent on both feature extraction and retrieval. The second column is the averaged querying time without a pre-processing stage while the third column is with VOI selection, i.e., with a pre-processing stage. All methods are programmed in software of Matlab R2009a running with an Intel P8600 1.58GHz CPU and 3.45GByte RAM.

TABLE 4 QUERY TIME

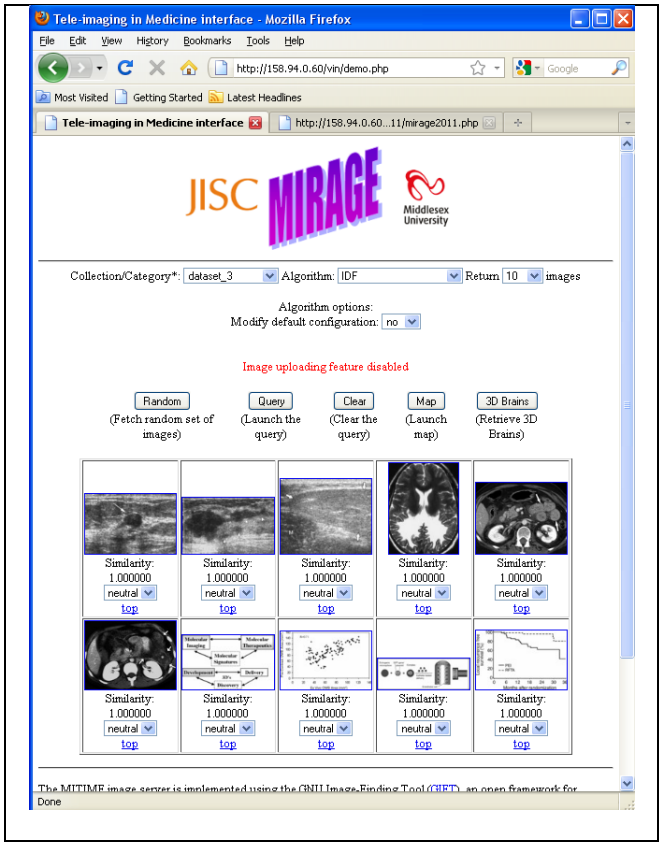
Methods	Without VOI selection	With VOI selection
3D GLCM	43.37s	10.96s
3D WT	4.46s	1.22s
3D GT	38.79m	10.77m
3D LBP	0.74s	0.21s

As can be seen in Table 4, the query time with VOI selection offers 4 times faster operation than that without. In particular, the query time for 3D GT takes much longer than for the other methods spending 38 minutes, due to the employment of 144 times of 3D convolutions for each block, whereas the query time for the other methods are completed in the space of few seconds. The table also illustrates that the 3D LBP approach outperforms the other three with sub-second retrieval time and the highest precision rate of 78%, as given in Table 3. However, this conclusion is very much content-based. In this case, the retrieval performance is based on the retrieval of images with similar lesion positions. Further studies are in need to explore whether any other contents, such as tumour shape, might be in favour of any of the other methods. All in all, all these four methods are implemented in the developed CBIR system that is addressed below.

IV. INTERFACE DESIGN

An online CBIR system, MIRAGE, acronym for Middlesex medical Image Repository with CBIR ArchivinG Environment, for both 2D and 3D images has been developed and is online at [26]. Figure 9 demonstrates the interface of the system, whist Figure 10 illustrates the flowchart of the architecture of interface. It consists of three modules with components of image classification, 2D image retrieval and 3D image retrieval respectively.

In Figure 9, the top picture displays a random selection of ten images from the collection of ‘dateset_3’ chosen from the dropdown menu of Collection Category, which can be achieved by simply pressing the ‘Random’ button. The last button on this figure gives the choice of the number of images to be displayed, which can be up to 140. Obviously more images will take longer to show up. Upon these shown images, users can pick one or more as query image or images by changing the status of each one from ‘neutral’ to ‘rel’ that refers to relevant, or ‘non-rel’ to eliminate the like of that image. By clicking the ‘Query’ button, the screen will show the retrieved images that are similar to the chosen query image or images.



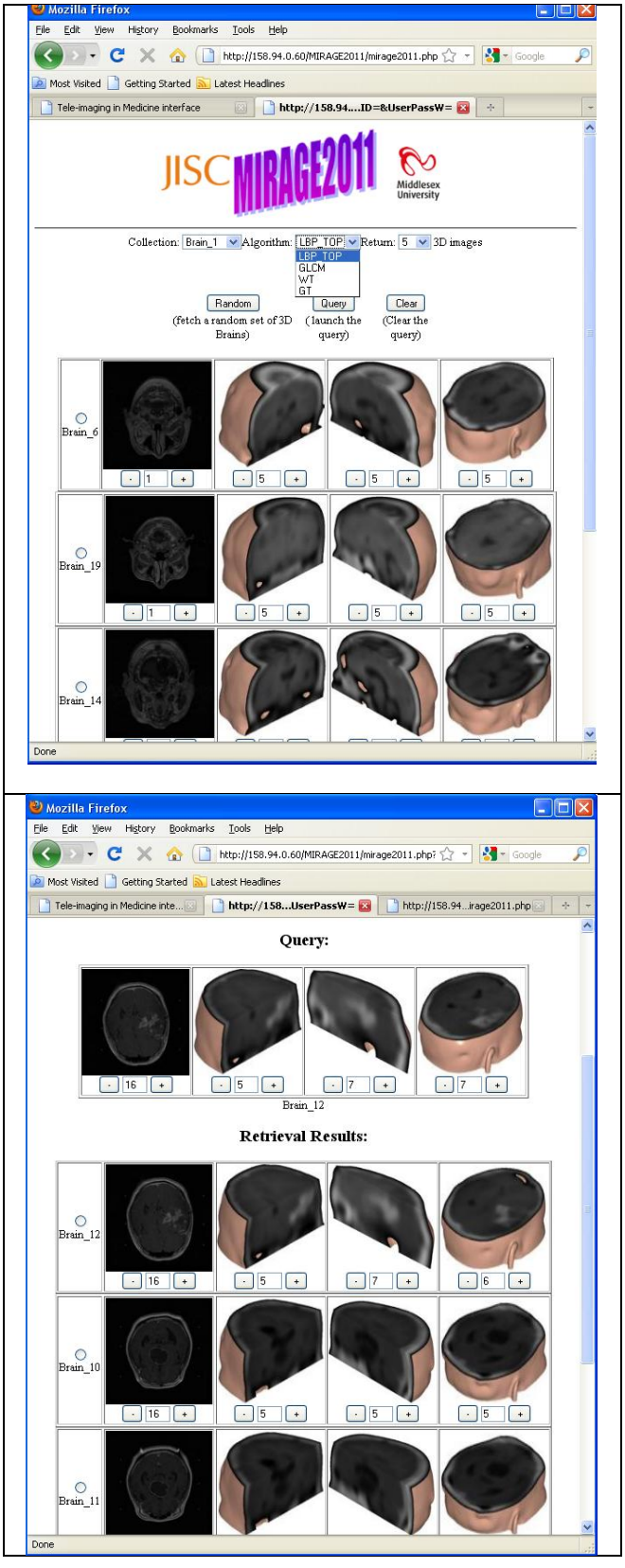


Figure 9. The interface of MIRAGE. Top: 2D images; Middle: 3D images; Bottom: retrieved results for 3D query.

The novelty of this work is the implementation of retrieval for 3D images which are demonstrated in the middle and bottom figures of Figure 9. The middle picture illustrates the implementation of the four aforementioned algorithms that can be applied to the retrieval process. There are four ways to view each 3D dataset. Among the 5 columns in the figure, on the second left (the leftmost column lists the name of the data), 3D data are shown in 2D form. By clicking the '-' or '+' button at the bottom, users can view the 3D brain images slice by slice from the top of the head to the neck. In order to refer each slice to the 3D brain, the three columns on the right hand side showcase the mapping from 2D to 3D in the direction of back-front (coronal, column 3), left-right (sagittal, column 4), and top-bottom (axial, column 5). Similar to the 2D form of the top figure, a query image can be selected by ticking the image name and then pressing the 'Query' button. The retrieved images are then given as demonstrated in the bottom figure of Figure 9. Again, with the consideration of speed, only 5 datasets are shown at each time.

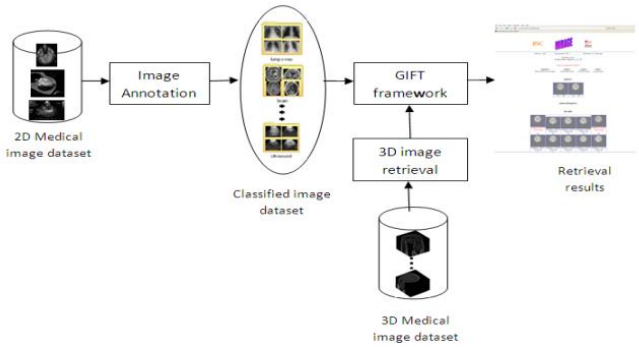


Figure 10. The framework for MIRAGE.

Built on the open source GNU Image Finding Tool (GIFT [40]), the online database is based on the Query-by-Example (QBE) paradigm coupled with a facility of user-relevance feedback whereby retrieved images most closely resemble a query image in appearance (i.e., the content that an image is carrying).

For 2D images, two algorithms have been implemented for indexing image collections, which are IDF (Inverse Document Frequency) and Separate Normalisation. The IDF is a classical method and is based on counting the number of documents in the collection being searched, which contain (or are indexed by) the terms in question [41], and has been applied in text retrieval systems, giving rise to the efficiency when employed in an image system. Conversely, feature normalisation refers to the compensation of scale disparity between the feature components that are defined in different domains.

On the client side, a web page based interface is given. Whilst the client-server communication is achieved using the XML-based Multimedia Retrieval Markup Language (MRML). All client-server communication, including queries from the client or results returned by the server, is realized through message passing. Consequently, the client can be implemented in any programming language. The current MIRAGE client is implemented using PHP (Personal Home Programming) language to generate dynamic web pages for the client web browser.

With respect to 3D interface, Figure 11 schematically illustrates a flowchart of the development.

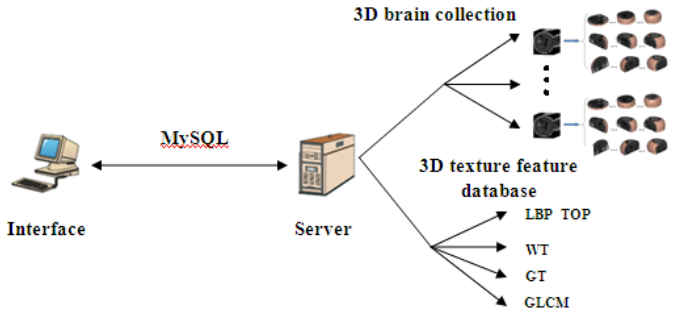


Figure 11. Framework of 3D brain retrieval system.

As illustrated in Figure 11, the visualization of 3D images relies on a Client-Sever architecture with MySQL communication protocol.

In order to display 3D brain as a whole instead of a pile of 2D slices, the skull of a brain is generated first from 3D volume data by using the method of iso-surface extraction, which is then followed by setting the step forward or backward with end-caps in three directions (i.e., X, Y, Z) respectively to show the inside structure of a brain from bottom to top, back to front and left to right, as schematically illustrated in Figure 12. All processing procedures including the step of images of 3D inter-section (to be controlled by '-' and '+' buttons in Figure 9) and the extraction of texture features from aforementioned four approaches can be performed offline in advance. In this way, the created 3D inter-sections of brains with its original 3D brain image are then stored in the image database in a server, whilst texture features are stored in a feature database.

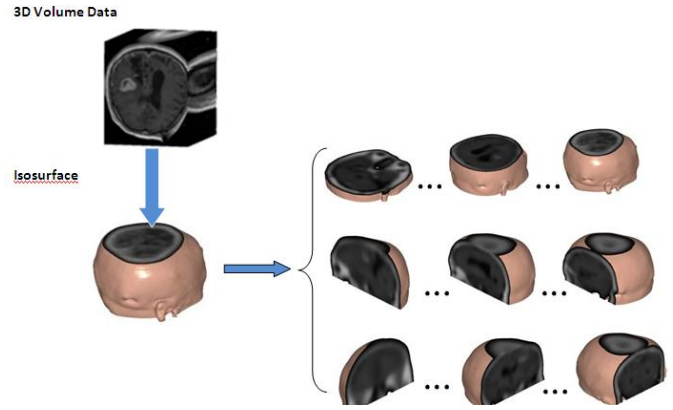


Figure 12 Creation of iso-surface and inter-sections of 3D volume data.

On the client side, interface relies on HTML, which is dynamically generated by means of hypertext preprocessor and therefore can be displayed in any internet browser. As shown in Figure 9, the user can view not only the original 3D brain image shown in first column of each row but also its cutaway in three different directions on the basis of slice by slice.

V. CONCLUSION

This paper is an extended version of [24], which introduces an online image retrieval system, MIRAGE, with a facility of CBIR for 3D images. Four texture based approaches that draw on the techniques of Local Binary Pattern, Grey Level Co-occurrence Matrices, Wavelet Transforms and Gabor Transforms have been exploited through the extension into 3D format, to retrieve lesioned MR brain images in this system. The results are very encouraging showing that not only higher precision rates can be achieved, but also that can it be done in real time. In comparison with each other, LPB outperforms the other three to a great extent whereas the 3D wavelet approach also performs well with similar retrieval accuracy, although slightly under-performed in terms of time. In terms of processing speed, it appears the pre-processing stage of detection of potential VOIs is essential to highlight lesions, the regions of interest that retrieved images should contain.

Because of the time required in the establishment of a feature database in 3D form, i.e., normalization, feature extraction, etc., in particular by using the approach of 3D GT (up to several minutes are needed for each brain), only ~100 datasets are included in this study. The very next step is to process more datasets. In addition, although the precision rate of 78% is very promising, a better rate may be possible by the combination of a few of these texture descriptors, while maintaining the short processing time. Comparison with shape based approaches is also in the pipeline, with the aim of developing CBIR systems for higher dimensional datasets.

ACKNOWLEDGMENT

This research is financially funded by UK JISC. Their support is gratefully acknowledged.

REFERENCES

[1] Smeulders A.W., Worring M., Santini S., Gupta A., and Jain R., Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell* 2000, 22 (12): 1349–1380.

[2] Datta R., Joshi D., Li J., and Wang J., Image retrieval: ideas, influences, and trends of the new age, ACM Computing Surveys, 2008, 40(2), pp.5-1:60.

[3] Novotni M. and Klein R., 3D Zernike Descriptors for Content Based Shape Retrieval, *Proceedings of the 8th ACM Symposium on Solid Modelling and Applications*, Seattle, Washington, USA, 2003, pp. 216-225.

[4] Bustos B. Keim D., Saupe D., and Schreck T., Content-based 3D Object Retrieval, *IEEE Transactions on Computer Graphics and* Applications, 2007, 27(4), pp.22-27.

[5] Mademlis A., Darasb P., Tzovarasb D., and Strintzis M.G., 3D Object Retrieval Using the 3D Shape Impact Descriptor, *Journal of Pattern Recognition*, 2009, 42 (11), pp.2447-2459 .

[6] Cao L., Liu J., and Tang X., 3D Object Retrieval Using 2D Line Drawing and Graph Based Relevance Feedback, *Proceedings of the 14th Annual ACM International Conference on Multimedia*, Santa Barbara, CA, USA, 2006, pp. 105 – 108.

[7] Ichida H., Itoh Y., Kitamura,Y., and Kishino F., Interactive Retrieval of 3D Shape Models Using Physical Objects, *Proceedings of the 12th Annual ACM International Conference on Multimedia*, New York, NY, USA, 2004, pp. 692 – 699.

[8] Gong B., Xu C., Liu J., and Tang X., Boosting 3D Object Retrieval by Object Flexibility, *Proceedings of the 7th ACM International Conference on Multimedia*, Beijing, China, 2009, pp. 525-528.

[9] Bustos B., Keim D., Saupe D., and Schreck T., Content-Based 3D Object Retrieval, *IEEE Computer graphics and Application*s, 2007, 27(4), pp. 22-27.

[10] Vajramushti N., Kakadiaris I.A., Theoharis T., and Papaioannou G., Efficient 3D Object Retrieval Using Depth Images, *Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval*, New York, NY, USA, 2004, pp. 189 – 196.

[11] Lowe D. G., Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, 2004, 60(2), pp. 91-110.

[12] Scovanner P., Ali S., and Shah M., A 3-Dimensional SIFT Descriptor and Its Application to Action Recognition. *ACM Conference on Multimedia*, 2007, pp. 357-360.

[13] Flitton G., Breckon T., and Megherbi N., Object Recognition using 3D SIFT in Complex CT Volumes, *British Machine Vision Conference (BMVC)*, 2010, pp.1-12.

[14] Gao X. W., Qian Y., and Hui R., The state of the art of medical imaging technology: from creation to archive and back, *The Open Medical Informatics Journal*, 2011, 5 (Suppl 1), pp.73-85.

[15] Arya M., Cody W., Faloutsos C., Richardson J., and Toya J., QBISM: Extending a DBMS to Support 3D Medical Images, *Proceedings of the Tenth International Conference on Data Engineering*, 1994, pp.314-325.

[16] Liu Y. and Dellaert F., A classification based similarity metric for 3D image retrieval, Computer Vision and Pattern Recognition, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'98)*, 1998, pp.800-805.

[17] Chbeir R., Amghar Y., and Flory A., MIMS: a Prototype for medical image retrieval, *In RIAOCID*, 2000, pp.846-861.

[18] Chu W., Hsu C., Cardenas A., and Taira R., Knowledge-based image retrieval with spatial and temporal constructs, *IEEE Transactions on Knowledge and Data Engineering*, 1998, 10(6), pp. 872–888.

[19] Mojsilovic A. and Gomes J., Semantic based categorization, browsing and retrieval in medical image databases, *Proceedings of image Processing*, 2002, pp.III:145-148.

[20] Glatard T., Montagnat J., and Mgnn I.E., Texture based medical image indexing and retrieval: application to cardiac imaging, *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval*, 2004, pp.135-142.

[21] Cai W. and Feng D., Content-based Retrieval of dynamic PET functional images, *IEEE Transactions on Information Technology in Biomedicine* 2000, 4(2), pp.152-158.

[22] Batty S., Fryer T., Clark J., Turkheimer F., and Gao X.W., Extraction of Physiological Information from 3D PET Brain Images, *VIIP'2002 (Visualization, Imaging and Image Processing) 2002*, pp. 401-405.

[23] Gao, X.W., Qian Y., Loomes M., Comley R., Barn B., Chapman A., and Rix J., Texture-based 3D image retrieval for medical applications, *IADIS e-Health2010*, Germany, 2010, pp 29-31.

[24] Qian Y., Gao X., Loomes M., Comley R., Barn B., Hui R., and Tian Z., Content-based retrieval of 3D medical images, *The Third International Conference on eHealth, Telemedicine, and Social Medicine, eTELEMED 2011*, France, IARIA, XPS Press, 2011, pp.7-12.

[25] Talairach J. and Tournoux P, Co-planar stereotactic atlas of the human brain, Thieme, Stuttgart, 1988.

[26] http://image.mdx.ac.uk/vin/demo.php. Retrieved on 26/1/2012.

[27] Unay D., Ekin A, and Jasinschi R.S., Medical Image Search and Retrieval using Local Patterns and Kit Feature Points, *Proceedings of the International Conference on Image Processing*, San Diego, California, USA, 2008, pp. 997-1000.

[28] http://www.fil.ion.ucl.ac.uk/spm/. Retrieved on 26/1/2012.

[29] Montreal Neurological Institute, http://www.mni.mcgill.ca/. Retrieved on 26/1/2012.

[30] Zhao G. and Pietikainen M., Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 9 (6) , pp. 915-928.

[31] Gao, X.W., Batty, S., Clark, J., Fryer, T., and Blandford, A., Extraction of Sagittal Symmetry Planes from PET Images, *Proceedings of the IASTED International Conference on Visualization, Imaging, and Image Processing (VIIP'2001*), 2001, pp. 428-433.

[32] Bhattachary A., On a Measure of Divergence between Two Statistical Populations Defined by Their Probability Distribution, *Bulletin of the Calcutta Mathematical Society*. 1943, 35, pp. 99-109.

[33] Kovalev V.A., Kruggel F., Gertz F.J., and Cramon D. Y., Three-Dimension Texture Analysis of MRI Brain Datasets, *IEEE Transactions on Medical Imaging*, 2001, 20 (5), pp. 424-433.

[34] Philips C., Li D., Raicu D., and Furst J., Directional Invariance of Co-occurrence Matrices within the Liver, *Proceedings of IEEE International Conference on Biocomputation, Bioinformatics, and Biomedical Technologies*, Bucharest, Romania, 2008, pp.29-34.

[35] Haralick R.M., Shanmugam K., and Dinstein I., Textural Features for Image Classification, *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, 3 (6), pp.610-621.

[36] Mallat S. G., A Theory for Multiresolution Signal Decomposition: the Wavelet Representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11 (7), pp. 674-693.

[37] Feng M. and Reed T.R., Motion Estimation in the 3-D Gabor Domain, *IEEE Transactions on Image Processing*, 2007, 16 (8), pp.2038-2047.

[38] Wang Y. and Chua C., Face Recognition from 2D and 3D Images Using 3D Gabor Filters, *Journal of Image and Vision Computing*, 2005, 23 (11), pp. 1018-1028.

[39] www.slicer.org. Retrieved on 26/1/2012.

[40] http://www.gnu.org/software/gift/. Retrieved on 6/1/2012.

[41] Robertson S., Understanding Inverse Document Frequency: On theoretical arguments for IDF, *Journal of Documentation*, 2004, 60 (5), pp.503–520.