Middlesex Medical Image Repository with a CBIR ArchivinG Environment (MIRAGE)

Final Report

12/2010

Authors: **Dr. Xiaohong Gao**
Dr. Yu Qian

Contact: **Xiaohong Gao**

x.gao@mdx.ac.uk

School of Engineering and Information Sciences, Middlesex University, The Burroughs, London NW4 4BT

Tel: 020 8411 2252
Mobile: 07971 168267

Table of Contents

Contents

1. Acknowledgements

Middlesex Medical Image Repository with a CBIR ArchivinG Environment (MIRAGE) (4/2009-9/2010) was funded by JISC *Information Environment Programme (INF11)* under the scheme of *Repositories Start-up and Enhancement (SUE)* strand (02/09). Their support is gratefully acknowledged.

During the project, support, advice and input from several user groups, data providers, and technical supporting staff were given and were of great value. These groups included:

Academic staff at Middlesex University
Technical staff at EIS, CLQE, and CIE of Middlesex University
Medical informatics team at University Hospital of Geneva, University of Applied Sciences at Western Switzerland
Neurosurgery Department at China Navy General Hospital
JISC website with related information and documents, in particular, CETIS and CHERRI
JISC weekly Newsletters

The project team would also like to acknowledge the open source website GNU (www.gnu.org) from which the GIFT software was downloaded to start the repository.

Thanks also go to a group of MSc students who are on the programme of BioMedical Modelling and Informatics for conducting a survey on the usability, feasibility, and applicability of the developed system. A number of survey results are presented in this report.

2. Executive Summary

The main aim of MIRAGE was to develop a subject-based repository of medical images, in the immediate term, benefiting MSc students who are on the programme of Biomedical Modelling and Informatics (BMI). In the long term, the established repository would serve a wider community by providing an immeasurable rich supply of collections for data mining and would be embedded with the existing online teaching systems, such as OASIS+.

The project was structured into nine work packages with three stages of major software development including image deposit and retrieval, interfacing with OASIS and user evaluation. Project management was lightweight given the fact that the three partners of the project were all from the same university and were located at the same campus.

The project began with the ingestion of a large collection of medical images into the existing server that then had only archived a few hundreds of images of limited domains. Since many image data are without any textual labelling, archiving image data is different from that to text files that can be indexed using a few key words embedded in the files. This deposition stage hence included the establishment of both feature and image databases. A number of approaches in extracting features had been applied in pre-processing images, which lasted about 6 months. Then the project team faced a technical challenge when trying to add more storage into the server that was running out of space, which was the interoperability between the operation system and the software. Around four months was spent on solving this issue. The system is now up running and accommodates about 100,000 medical images in 2 dimensions. By building on from an open source software GNU GIFT, the online system (http://image.mdx.ac.uk/vin/demo.php) currently not only facilitates a means to search images by their contents, notably content-based image retrieval (CBIR), but also interfaces with OASIS+, the online teaching system at Middlesex University to ensure it can be located easily .

All of the project work has been completed successfully. Although at present, the interface of image deposition/ingestion is not external facing, this was done with an intention in order to restrict the access to the image server. It is anticipated that other security measures will be implemented in the future when the interface is open to the public. The MIRAGE system has been very well received by the user groups with 80% agreement achieved when it comes to the user expectations, retrieval speed, user friendly, and usefulness in teaching and research.

The project team has also developed an approach in indexing 3D images. However, when it comes to viewing them online, they are displayed in the form of an array of 2D images, giving less intuitive view. It is envisaged that visualisation plug-in software will be developed in the future to foster 3D view. Another issue that has been identified is to upload query images by users, which again will form one of the future work to ensure the security of the server is not compromised when the uploading is in place.

3. Background

In recent years, many online learning systems are available to students and have played an important part in helping them learning. These systems usually tend to be in general purpose in order to meet majority students' need. However, sometimes, subject-based databases are in demand by a number of groups. At Middlesex University, a new MSc programme was introduced in 2007 on BioMedical Modelling and Informatics (BMI), a collaborative venture between two schools, Engineering and Information Sciences (EIS) and Health and Social Sciences (HSS), which had been attracting an increasing number of students. During the course of studying and conducting final projects, a large number of images had been employed in addition to other form of data. Before the repository was established, those images had been communicated via portable media (e.g, USB sticks) between tutors and students, which had caused tremendous inconvenience, given the volume size of the data (>20,000 images). Further concerns included the time consumed in selecting the right sets of data and potential security risk ascribed to memory sticks (e.g, worms). Delaying was inevitable if the delivery was missed or the wrong set of data had been copied. On the other hand, although a small portion of images had been uploaded to OASIS+ for downloading, retrieving images is usually very different from text retrieval, i.e., browsing images one by one to find the right one was not practical due to their sheer size (> 1Mbytes per image).

Following a successful bid to JISC in the spring of 2009, an attempt to establish a subject-based repository started. It was planned to add more images into the existing server whilst improving the retrieval facilities. The project had full support from the senior executives in the university, especially Pro-Vice Chancellor Professor Martin Loomes for providing match funding to the project to support a Research Assistant working full time on the project.

At present, the server warehouses about 100,000 images. And more images are still being processed and are added to the server. Flexible interface has also been developed. With the system being based on the internet platform, images can now be accessed by the students anywhere in the world, which is in part in the consideration that many of our students are form overseas and are able to continue to do their projects while on holiday at their home countries, as some other online learning systems are university bound, i.e., they can only be accessed within university's networks.

All the collected images in the server comply with the informed consent requirement and consist of 2D retinal images, 3D brain images (CT, MR and PET) and 4D cardiovascular ultrasound images. MIRAGE adapts an open framework of GIFT (GNU Image-Finding Tool) for content-based 2D medical image retrieval. By introducing the automatic image annotation, MIRAGE provides the possibility of combining visual content with keywords to achieve the higher level of semantic search. In addition, MIRAGE has developed its own method for 3D brain images retrieval to complement the existing 2D medical image repository.

3.1 Text-based and content-based retrieval system for images

With the increasing use of digital methods of data capture and organisation, the sustainability and services offered by institutional repositories are in need more than ever. According to Sir Muir Gray, then chief knowledge officer of the NHS[1], the "application of what we know already will have a greater impact on health and disease than any drug or technology likely to be introduced in the next decade". Repositories are a typical example collecting large amount of information waiting to be exploited.

[1] MEDINFO 2007, *http://www.chi.unsw.edu.au/CHIweb.nsf/page/Conference%20Papers*

Traditionally indexing and retrieving images are based upon textual annotations manually extracted from an image, such as webpage of Medical Images and Illustration [2], However such indexing is not scalable and adequate in describing important features, such as spatial-temporal information which users may interest in medical images. On the other hand, manual textual annotations will lead to subjective and time consumption. Progress has been made to develop content-based image retrieval systems (CBIR) by querying the collections using a sample image to search for images with similar appearance. These text-based and CBIR systems are usually designed to work independently with few existing systems providing a retrieval facility accommodating metadata to allow retrieval to be performed via either text or query sample, or even more challenging both, especially in the medical domain where several features may be irrelevant, such as colour .

3.2 CBIR for Teaching and Learning

In terms of learning and teaching, CBIR can play an important role in helping students to classify images of variety domains, e.g., a lung image from a brain in terms of appearance, especially when students are at the early stage of getting familiarised with images. The inclusion of visual features into this stage is anticipated to achieve educational purposes.

[2] http://www.mic.ki.se/MEDIMAGES.html

4. Aims and Objectives

The overall aims of the project were to set up a multimedia repository of medical images and to facilitate a multi-modal retrieval tool, which will be met by the following objectives:

- Extract image contents following the MPEG-7 standard by using two descriptors. One is a contour-based shape descriptor (to describe lesions) and the other is a global feature of texture (to present a domain), which will be saved as XML document files to be compatible with the other web page standards;
- Annotate images with their diagnostic keywords, e.g. tumour, head injury, etc., based on the accompanying text of an image using the Unified Medical Language System (UMLS);
- Locate anatomical locations of diagnostic lesions in each image based on standard domain atlases, e.g. Talairach atlas for brain images;
- Build a lookup table to link keywords with their content features in terms of shape, texture, and anatomical location;
- Design a multi-functional interface to facilitate text-based, content-based and semantic-based image retrieval;
- Review issues on ethical, privacy and copyright in dealing with medical data;
- Promote an e-learning environment;
- Interface with OASIS+ based on SCORM.

5. Methodology

The project was structured into 2 stages representing two kinds of activities the repository would offer. Stage 1 deposited images into the repository with meaningful (semantic) tags, which allowed multi-modal retrieval to take place in Stage 2. To achieve this, the project had employed agile open source development strategies such as SCRUM, delivering weekly cycles of development. The project was built on the existing image server interfaced by GIFT, an open-source GNU image-finding tool that had been expended to incorporate both CBIR and text-based retrieval techniques.

The main focus of the development was the ingestion of the repository and the design of the infrastructure of an interface between MIRAGE and OASIS. The research teams in both EIS and CLQE (Centre of Learning and Quality Enhancement) supervised the design and evaluation of the repository and had been working together with the researcher (Dr. Qian) employed in the project for the development of the repository. Work on the extraction of image contents from the collections, development of the framework of lookup table for semantic retrieval, and annotation of images using keywords and anatomy information based on the vocabularies of UMLS was also undertaken by the researcher.

The existing image server (http://image.mdx.ac.uk) archiving PET brain images then has been built on, and its access interface GIFT, an open-source GNU image-finding tool, had been expanded to incorporate both CBIR and text-based retrieval techniques.

Figure 1 illustrates the infrastructure of MIRAGE. To address the problem of current text-based image retrieval systems, MIRAGE integrated the methods of content-based image retrieval (CBIR) for 2D and 3D medical image collection and automatic image annotation to label the images with its keywords, leading to a higher level of semantic search. MIRAGE therefore consisted of three modules as shown in Figure 1, and comprised of image annotation, 2D image retrieval and 3D image retrieval, which was built on the open framework and preceding research results. 2D medical images were firstly classified into different categories labelled with their domain names (e.g. lung, brain, etc) by using the method of image annotation. Based on GIFT framework, an open framework for CBIR, 2D medical images were retrieved from the same category as the query. MIRAGE has developed its own method of CBIR for 3D brain image retrieval to complement the medical image repository. The details of methods adapted and developed in MIRAGE are introduced in the following the Section of Implementation.

For the **quality management and standards**, the project had adopted the existing and emerging standards wherever possible and employ state of the art techniques in the fields of both content-based and text-based image retrieval, which detailed in Section 6. The ULMS acted as a gold standard to annotate images, whilst the MPEG-7 standard was applied to extract the content of images. Application of standard atlases for locating ROIs was practiced throughout the project as well as the adaptation of SCORM when interfacing with OASIS+ to be generic and compatible with Blackboard and other open source. The GIFT interface with a MySQL database had been employed together with OASIS techniques for the image server.

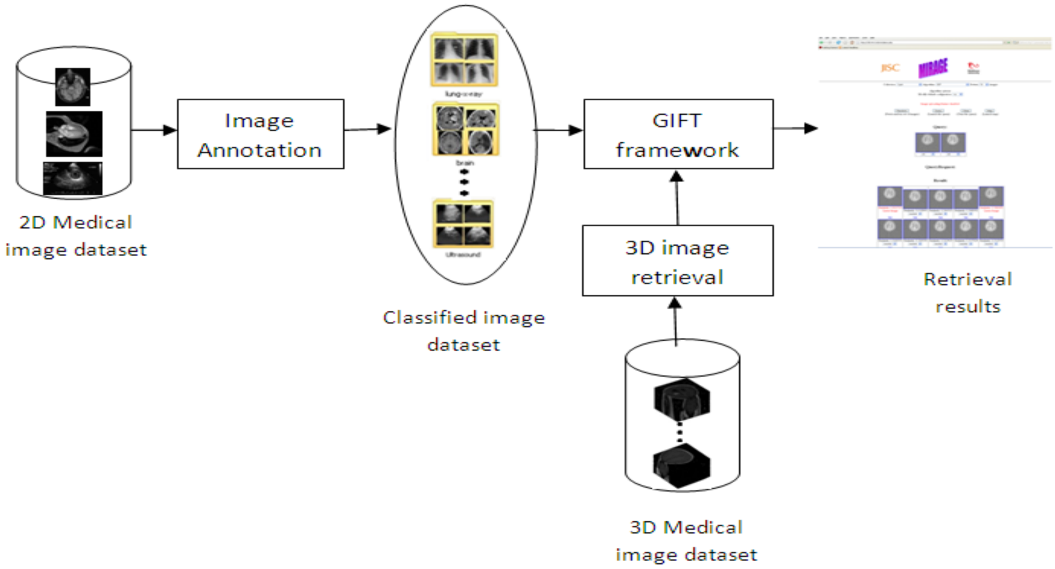


Figure 1. The Framework for MIRAGE.

6. Implementation

According to the methodologies described in Section 5, implementation was realised in the following details. The implementation followed JISC standards closely and open source infrastructure to enable global accessibility. The techniques included adopting existing standards of JISC, webpage technology, OASIS+, and open sources, such as SCORM [3], to ensure the developed repository being compatible with the other existing e-learning and teaching systems.

6.1 MIRAGE collection

MIRAGE server resided in a Linux operation system at MU and so far warehoused around 100,000 2D medical images and 100 3D MR brain images. 2D medical images including CT, MR, Ultrasound etc, were collected from Cross Language Image Retrieval (Image CLEF) [4]. These medical images had been utilized as benchmark images for many years. All images are in JPEG format. 3D brain images (around 100) including both normal and lesioned brains, were collected from Neuro-imaging Centre at Beijing General Navy Hospital, China, and were utilized to evaluate 3D CBIR. These images are in DICOM (Digital Imaging and Communications in Medicine) format.

6.2 Image Annotation

In order to achieve a higher level semantic search, automatic image annotation was applied in MIRAGE to organize and categorize images of interest from image repository. Automatic image annotation is the process by which a computer system automatically assigns metadata in the form of captioning or keywords to a digital image. At present, the Bag-of-visual-Words (BoW) paradigm is extremely popular and successfully applied for image categorization. This method is to transform images into a set of 'visual vocabulary' and to represent the images using the statistics of appearance of each word as feature vectors, upon which the learning of an image classification rule could be achieved as a classifier. The classification system based on BoW paradigm was composed of four phases in this project:

- Phase 1 -- the visual features were extracted from local patches of each image in the training archive. A visual dictionary called codebook was then constructed;
- Phase 2 -- to quantize the visual features of the image dataset into discrete "visual words";
- Phase 3 -- an image was represented as a unique distribution (e.g. a histogram) over the generated dictionary of words; and
- Phase 4 -- image representations of the training dataset obtained in Phase 3 were applied to train the classifiers using supervised machine learning methods. Finally, the trained classifier automatically assigned new images into corresponding categories.

The flowchart of visual dictionary construction and image representation are demonstrated in Figure 2. Whereas the procedure of image classification is displayed in Figure 4 with Figure 3 showing the coding stages.

[3] http://en.wikipedia.org/wiki/SCORM
[4] http://imageclef.org/

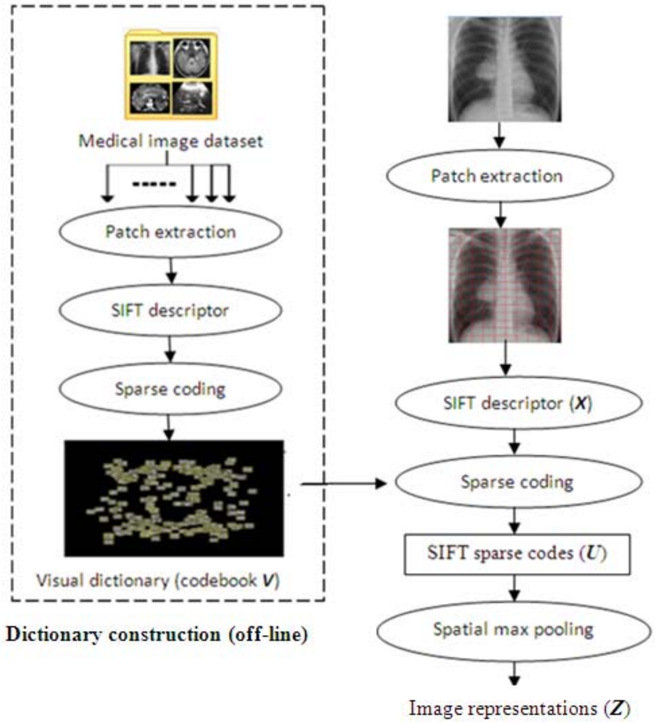


Figure 2. Dictionary construction and image representations

6.2.1 Phase 1: local feature extraction

In order to describe local features, an image was divided into non-overlapping equally sized patches which were considered as candidates for basic elements, or "words" in visual dictionary. Scale Invariant Feature Transform (SIFT) descriptors were applied to extract local features by computing a gradient orientation histogram within each patch. For each of 8 orientation planes, the gradient image was sampled over by a 4 by 4 grid, thus resulting in a vector of 128 dimensional features for each patch.

6.2.2 Phase 2: Visual dictionary construction

Sparse coding of model data vector as a sparse linear combination of a set of basic elements called dictionary was applied to construct visual dictionary. Encoding each descriptor of an image was done by solving the optimization problem as formulated in Eq.(1).

$$\min_{U,V} \sum_{m=1}^{M} \left\| x_m - u_m V \right\|^2 + \lambda \left| u_m \right| \tag{1}$$

where $X = [x_1, x_2, \ldots x_m]$ is a set of SIFT descriptors from train dataset, and $V = [v_1, v_2, \ldots v_k]$ is the codebook. Likewise, $U = [u_1, u_2, \ldots u_m]$ is sparse codes for images based on codebook V, and $\left\| \bullet \right\|$ denotes the L_2 distance.

In the training stage, 50,000 descriptors extracted from random patches of train dataset were applied to conduct an off-line training on the codebook V. Whilst in the coding stage, each image X was encoded as a set of U by inputting the trained codebook V that is as expressed in Eq.1.

6.2.3 Phase 3: Image representations

Considering spatial location of local features, a spatial pyramid structure based on SIFT sparse codes were adapted as the image representations, as demonstrated in Figure 3. Instead of the traditional histogram of 'visual word', the max pooling of SIFT sparse codes was calculated cross different locations and over different scales of a spatial pyramid structure of an image, which was concatenated to form a spatial pyramid representation of an image, denoted as Z in Figure 3.

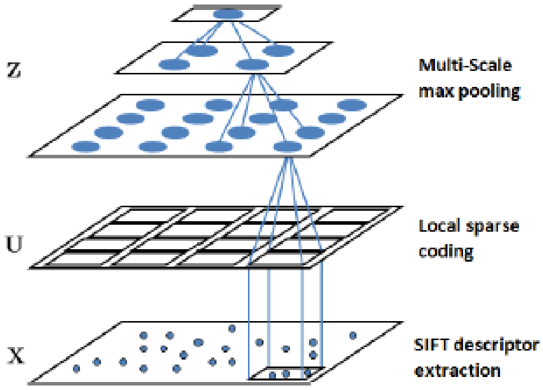


Figure 3: Spatial pyramid image representations based on SIFT sparse codes.

6.2.4 Phase 4: Image classification

As shown in Figure 4, after obtaining image representations from training dataset, image classification was performed using a multiclass Support Vector Machine (SVM) with a linear SPM (Spatial Pyramid Matching) kernel that was formulated as

$$k(z_i, z_j) = z_i^T z_j = \sum_{l=0}^{2} \sum_{s=1}^{2^l} \sum_{t=1}^{2^l} \langle z_i^l(s,t), z_j^l(s,t) \rangle \tag{2}$$

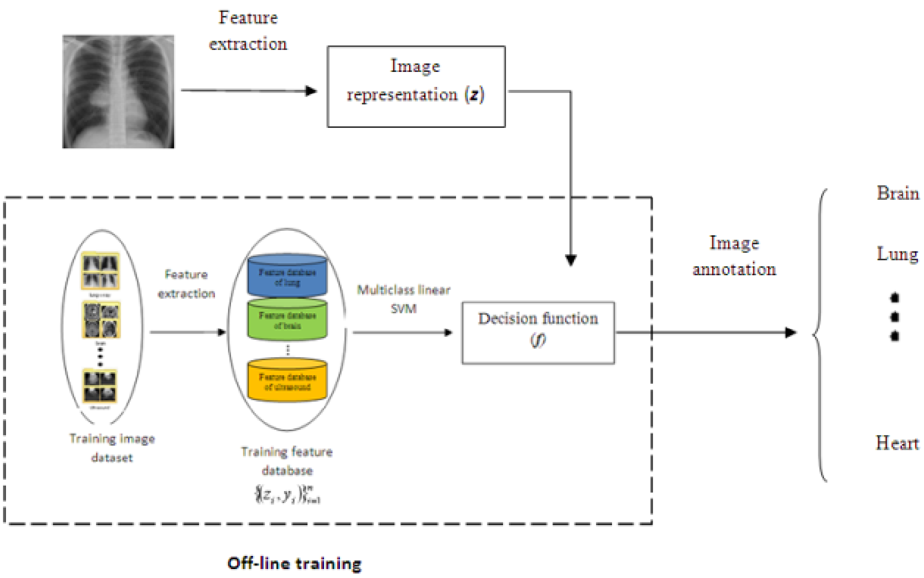


Figure 4. Image classification

A kernel function $k(z_i, z_j)$ is to transform features into a higher dimensional feature space in order to separate input data set by a linear hyper-plane. In Eq. (2), $z_i^l(s,t)$ is the max pooling of the SIFT sparse codes in the (s,t)-th region of image I_i in the scale level l. For binary classification, an SVM aims to learn a decision function based on the training dataset as defined in Eq.(3).

$$f(z) = \sum_{i=1}^{n} a_i k(z_i, z) + b \tag{3}$$

Where $\{(z_i, y_i)\}_{i=1}^{n}$ represents n^{th} training images, and $y_i \in \{-1,1\}$ denotes the label of training images. For image representation of z, if $f(z) > 0$ then the image is classified as 1, otherwise as -1. In order to classify multiple categories (L), the training images expressed as $\{(z_i, y_i)\}_{i=1}^{n}$, $y_i \in \{1,2...,L\}$. One-against-all strategy was applied to train $L(L-1)/2$ binary classifiers.

Unlike traditional BoW paradigm, sparse coding was employed in the project instead of vector quantization (VQ) to extract the SIFT descriptors of local image patches. Furthermore, instead of using histograms, multiple scales of max pooling was employed as an image representation by the use of simple linear SVMs. In comparison with the SVMs using nonlinear kernels, e.g. histogram intersection kernels, linear SVMs could dramatically reduce the training complexity while maintaining a good performance. This method was evaluated on repository MIRAGE in Section 7.1.

6.3 Content-based image retrieval

After image classification, 2D medical images were retrieved from the same category as the query using the approach of Content-Based Image Retrieval (CBIR). CBIR is performed in the following steps:

1). Feature extraction, such as colour, texture and shape, from an image in order to index it.
2). Query submission.
3). Query feature extraction
4). Comparison of features between query's and the feature database.
5). Similarity calculation.
6). Display retrieved results in the descending order of similarity distances.

In order to generate more meaningful retrieval results both perceptually and semantically, MIRAGE had been incorporated with a facility of users' relevance feedback to let users to modify the retrieval results in terms of similarity.

MIRAGE project adapted an open framework GIFT (GNU Image-Finding Tool) [5] that enabled the image retrieval based on visual similarity through Query by Example (QBE) on images. An inherent index system using inverted-file technique enabled to search a large dataset within seconds. On the other hand, GIFT has a distributed architecture (Client – Server) with its own XML (eXtensible Markup Language) based communication protocol MRML (Multimedia Retrieval Markup Language) as illustrated in Figure 5. The client initiates an event by sending a MRML message to the server and receives another MRML message replied with the details of the collections of images available and the associated algorithms. A subsequent client-server communication, including queries from the client or results returned from the server by using a stand of XML message passing. As a direct result, the client can be implemented in any programming language. The MIRAGE client interface was based on PHP to generate dynamic web pages. This separated interface and query engine ensured their independent development, whereas the standard XML parsers were utilised to process the communication messages and was able to be extended. In

[5] http://www.gnu.org/software/gift/

addition, GIFT provides a plug-in mechanism which permits to add any new ways of querying into GIFT without changing the interior working structure of GIFT.

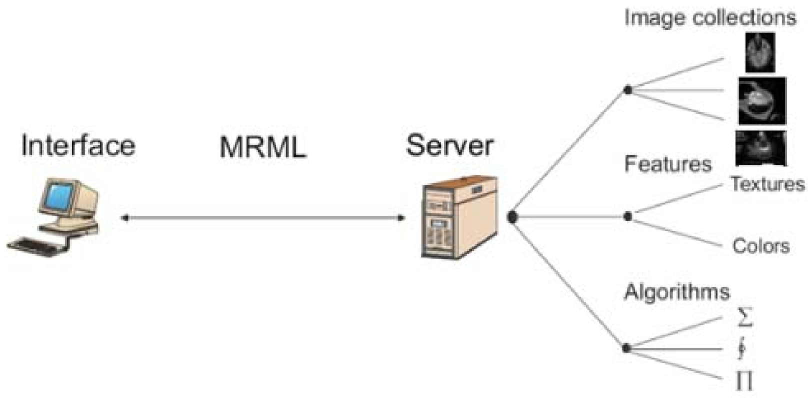


Figure 5. GIFT Framework

GIFT on medical image repository is evaluated in Section 7.2. The details on GIFT installation and usage are described in Appendix 1.

6.4 Content-based 3D brain image retrieval

MIRAGE also accommodates 3D images. Texture based features are extracted in this regards as detailed at [6] that is also given in Appendix 3.

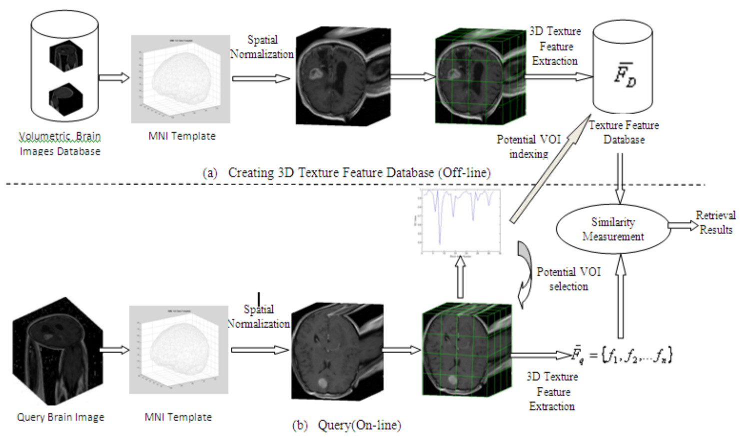


Figure 6. Framework for Content-based 3D Brain Image Retrieval

As shown in the diagram in Figure 6, the collection of 3D brain images firstly underwent a pre-processing stage to normalize them into the same resolution before the indexing stage. After spatial normalization of volumetric brain data into a standard template, the data were then divided into 64 non-overlapping equally sized blocks, from which, 3D texture features were extracted to create a feature database. On the query side, a pre-processing stage was introduced to detect the potential VOI of lesions after spatial normalization from a query image. Subsequently, 3D texture features from a query were only extracted

[6] X. Gao, Y. Qian, et al., Texture-based 3D image retrieval for medical applications, IADIS e-Health2010, July 2010.

from these potential sub-blocks, which, in the retrieval stage, were compared with the corresponding features in the feature database to obtain retrieval results. Figure 7 demonstrates an example retrieved using different texture approaches.

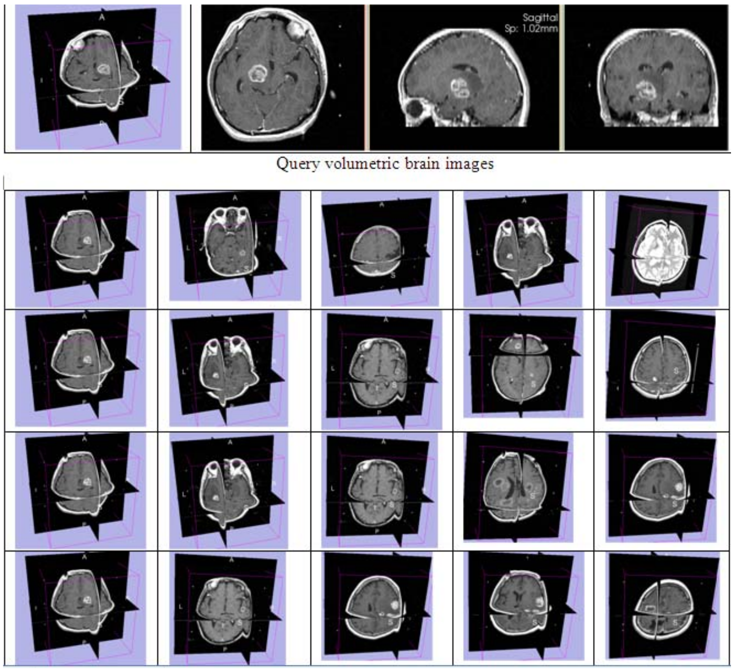


Figure 7. Retrieved results in top 5 ranking from 3D GLCM (row 1), 3D WT(row 2), 3D GT (row 3), and 3D LBP (row 4).

6.5 Evaluation

The system evaluation was carried out from both objective and subjective prospects. As an objective evaluation, a number of statistic measures are applied to evaluate the research methods, such as Average Accuracy Rate (AAR) for image classification and Mean Average Precision (MAP) for image retrieval. On the other hand, the subjective evaluation was accomplished by using a survey questionnaire conducted by the students and lecturers at MU who were using the medical image repository MIRAGE.

To assess the accuracy of image classification, a confusion matrix was firstly created as given in Table 1 by comparing the classification results with the ground truth information. Each row of the matrix represents the instances in ground truth, while each column refers to the instances in the classification results. The diagonal directional elements (e.g. a, e, and i) represent the correctly classified images whilst the rest elements represent misclassified ones. The Accuracy Rate (AR) for a given class was defined as the percentage of correctly classified images in a given class divided by the total number of images of ground truth in that class. The Error Rate (ER) for a given class was defined as the percentage of wrongly classified images in a given class divided by the total number of classified images in that class. In Table 1, the Average Accuracy Rate (AAR) for three classes is therefore defined as (a+e+i)/(a+b+c+d+e+f+g+h+i), and Average Error Rate(AER) is equal to 1-AAR.The strength of a confusion matrix is not only to demonstrate the accuracy of a classification but also to identify the nature of the classification errors. AAR and AER calculated from confusion matrix are applied to evaluate the method of image annotation on medical image repository, and evaluation results are shown in Section 7.1.

Table 1. Exam*+ple of confusion matrix

		Classification Results			Accuracy Rate
		Class 1	**Class 2**	**Class 3**	
Ground Truth	**Class 1**	**a**	**b**	**c**	**a/(a+b+c)**
	Class 2	**d**	**e**	**f**	**e/(d+e+f)**
	Class 3	**g**	**h**	**i**	**i/(g+h+i)**
Error Rate		**(d+g)/(a+d+g)**	**(b+h)/(b+e+h)**	**(c+f)/(c+f+i)**	

The performance of image retrieval was then evaluated based on the Precision (P) and Recall (R). Precision is defined as the fraction of retrieved images relevant to the query while recall is the fraction of relevant images retrieved. Precision and recall values are usually presented together in a Precision-Recall (P-R) graph that depicts the retrieval performance at each point in the ranking. In P-R graph, the horizontal axis expresses recall and vertical axis expresses the corresponding precision at standard recall points 10%, 20%,…,100%. By representing P-R graph using one value, Mean Average Precision (MAP) value is applied to access overall performance for all queries and calculated as Eq. (4).

$$\text{Mean Average Precision (MAP)} = \frac{1}{M}\sum_{i=1}^{M} AP_i \tag{4}$$

where M is the total number of the queries, AP_i is the Average Precision for the ith query and formulated as

$$\text{Average Precision (AP)} = \frac{1}{N_r}\sum_{j=1}^{N_r} P_j \tag{5}$$

where N_r is the total number of relevant images in a dataset for a query, whilst p_j is the precision when retrieve the jth relevant image. A P-R graph together with a MAP value are therefore applied to evaluate the performance of CBIR for 2D and 3D images in this project and results are displayed in Section 7.2.

In addition, an on-line questionnaire listed in Appendix 2 was designed in the hope to subjectively evaluate and further improve the system. The questionnaire consisted of 15 questions organized in three categories in order to collect a) the general information on the use of MIRAGE, b) the evaluation of system usability, and c) the comments/recommendations in regarding to the features of MIRAGE. The survey was carried out by students and lecturers at MU who used the medical image repository. More feedbacks from a wider community will be conducted in the future to further improve the system. The survey results are shown and further analyzed in Section 7.3.

7. Outputs and Results

The primary output of the project is an enlarged medical repository facilitating a content-based image retrieval. In order to provide a service to research, learning and teaching activities in the university, MIRAGE was firstly introduced to MSc students who were on the programme of Biomedical Modelling and Informatics through giving research seminars and lectures to them as well as by assigning related coursework to them to encourage the usage of MRIAGE. Meanwhile a survey was carried out by students and lecturers to help us evaluate and further improve the system. Further more, the project findings and the developed system were also promoted to a wider community through international conferences.

To benefit the other institutions of UK HE sector, integration with an on-line e-learning system OASIS+ was also implemented. The implementation followed JISC standards closely and open source infrastructure to enable global accessibility.

Figures 8 to 12 present a series of snapshot of the interfaces of MIRAGE repository.

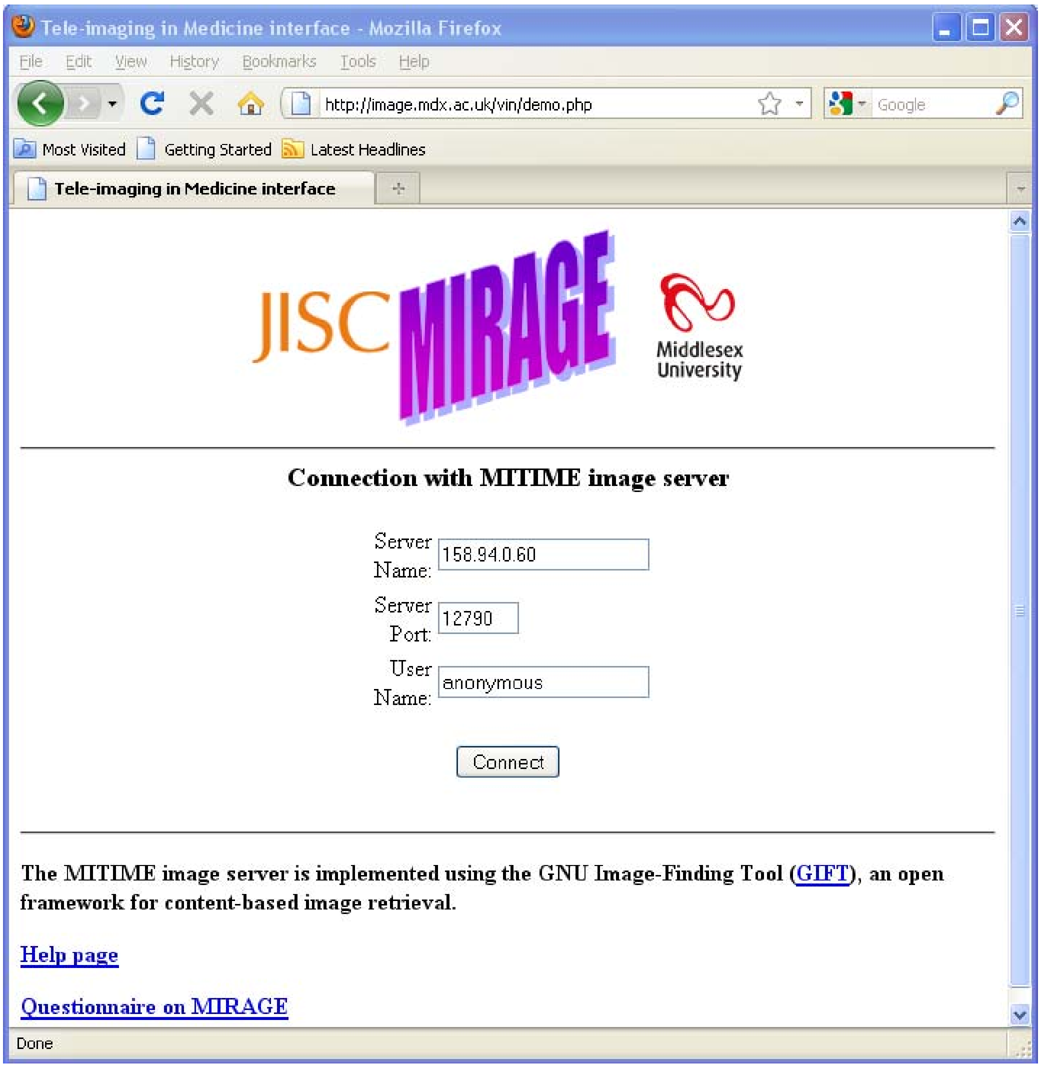


Figure 8. The interface of MIRAGE.

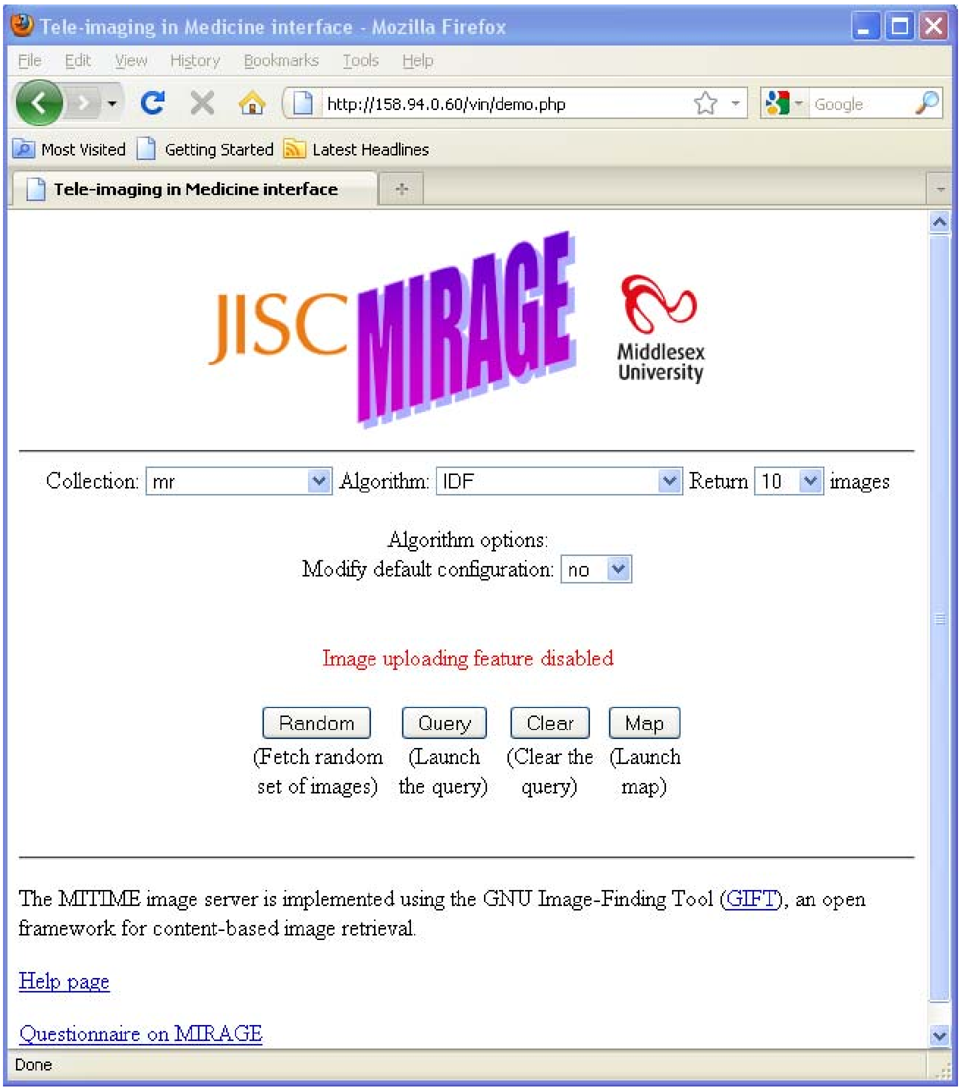


Figure 9. The query interface of MIRAGE.

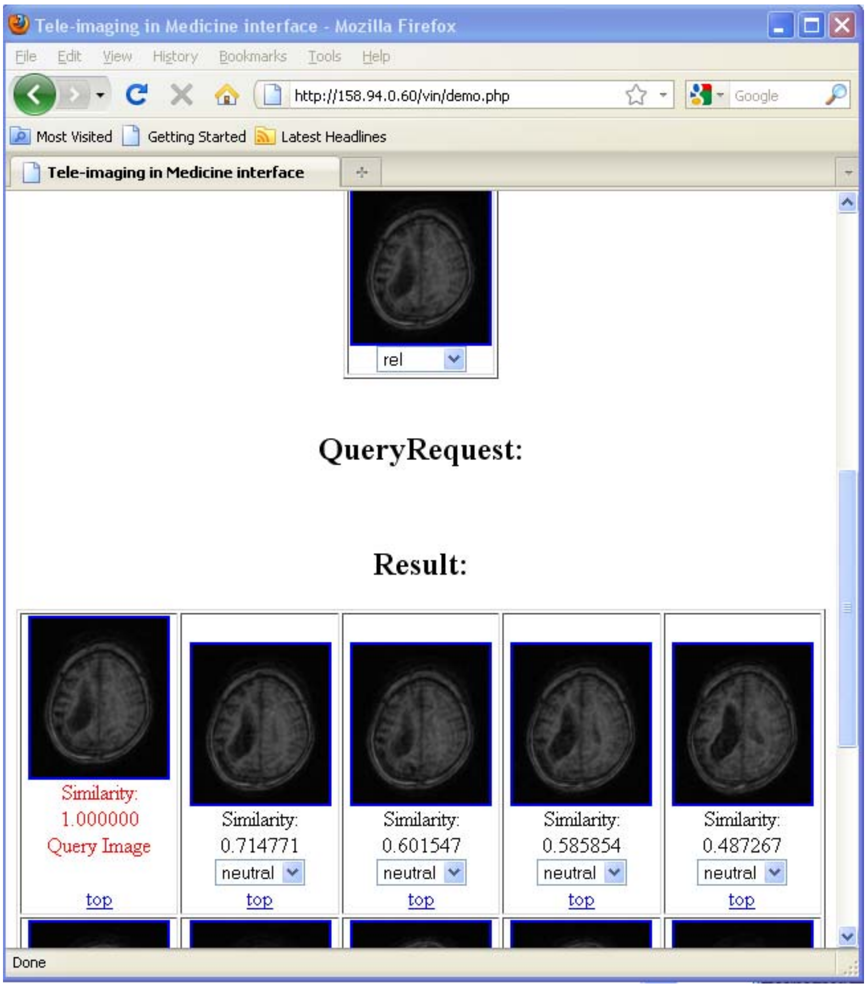


Figure 10. An example of a query and the retrieved results by CBIR.

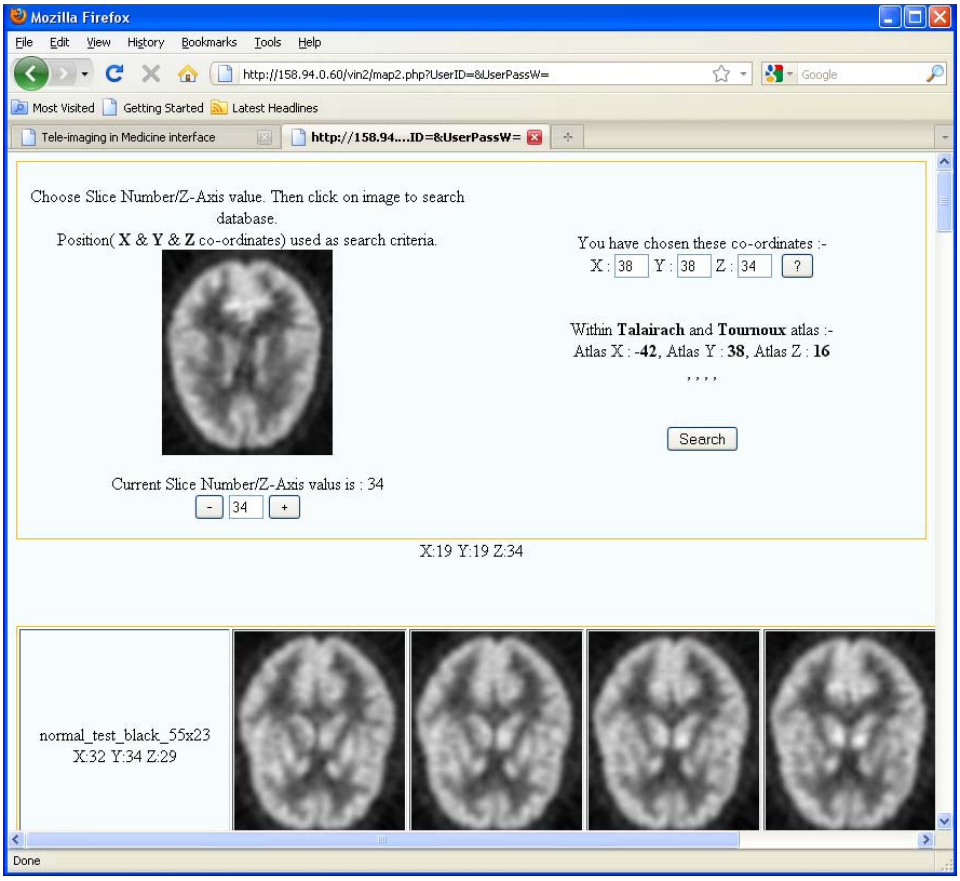


Figure 11. Retrieved results based on location.

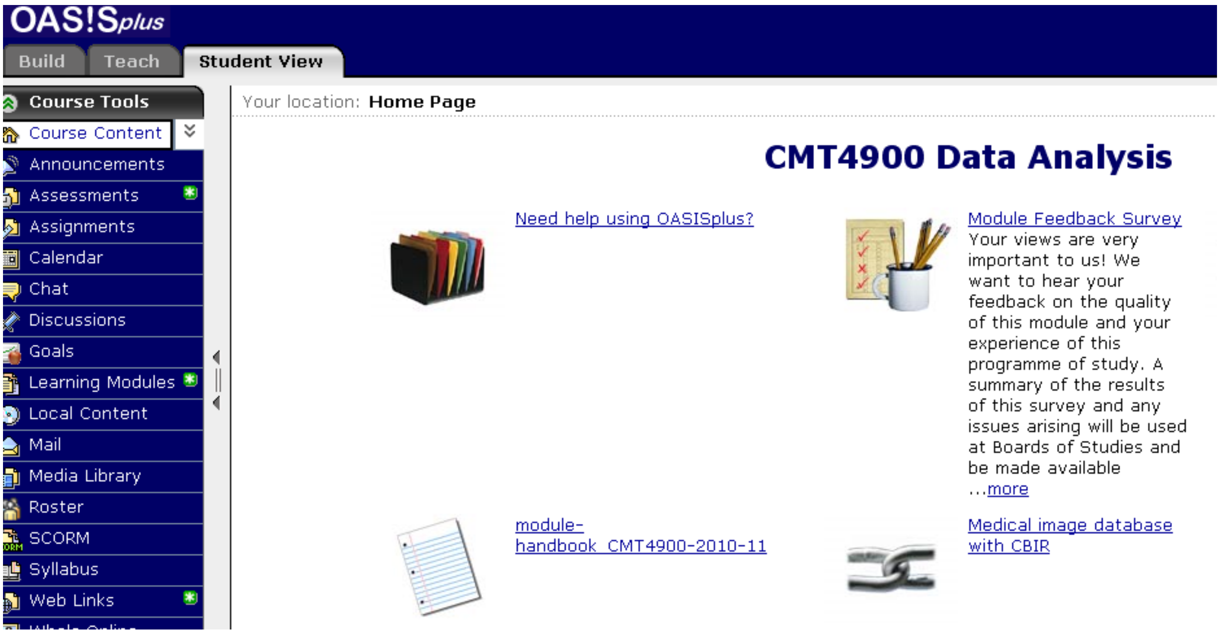


Figure 12. Interfacing with OASIS+ with a link at right bottom to the repository.

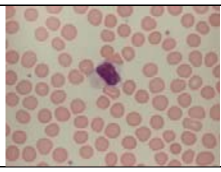
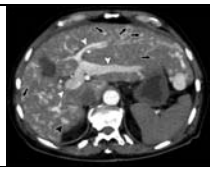
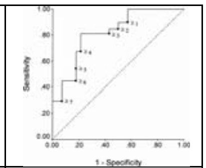
This project integrated the existing technologies, including GIFT framework for CBIR and image annotation using SIFT sparse codes. At the mean time, development of approaches for content-based 3D image retrieval was also attempted. All methods were tested on medical images and evaluation results are shown in the following sections.

7.1 Results on Image Annotation

In order to train a codebook, 1000 images were randomly selected from our medical image database and 200,000 random patches were obtained from these images. SIFT descriptors were then extracted from each patch, yielding a feature database with the size of 200,000*128 elements, which were subsequently applied to train a codebook with the size of 1024*128.

For assessing the method of image annotation, a set of ground truth data was created from the existing medical image repository. The ground truth was that the data were from six domains, including images of brain, lung (x-ray), microscopy, abdomen, ultrasound and graph respectively. There were 100 images in each category. The example images from each category are shown in Table 2.

Table 2: Confusion matrix for 6 medical image categories

					
Brain	**Lung(x-ray)**	**Microscopy**	**Abdomen**	**Ultrasound**	**Graph**

All the images in the ground truth were represented by SIFT sparse codes based on the trained codebook. 50 images per category were selected as the training dataset that were used to train classifiers by using the approach of multi-class SVM, whilst the rest 50 per category were applied as testing dataset. The classification results for the six categories are visualized in a confusion matrix in Table 3 as explained in Section 6.5.

Table 3: Confusion matrix for 6 medical image categories

		Classification Results						Accuracy Rate (AR)
		Brain	**Lung(x-ray)**	**Microscopy**	**Abdomen**	**Ultrasound**	**Graph**	
Ground Truth	**Brain**	48	0	2	0	0	0	96%
	Lung (x-ray)	0	50	0	0	0	0	100%
	Microscopy	0	0	49	0	1	0	98%
	Abdomen	0	0	0	50	0	0	100%
	Ultrasound	0	0	0	0	50	0	100%
	Graph	0	0	0	0	0	50	100%
Error Rate (ER)		0	0	3.92%	0	1.96%	0	

The values in the last column of Table 3 are the AR values for each class, whereas the values in the last row are the ERs for each class. The Average Accuracy Rate (AAR) for all classes is 99% (297/300), and Average Error Rate (AER) is 1% (3/300), demonstrating the approach of annotation being very accurate.

7.2 Results for 3D Image Retrieval with CBIR

The performance of 3D medical image retrieval by using LBP was evaluated in this section. Details are given in Appendix 3. Figure 13 demonstrates the average Precision Recall Graph for ten queries across the whole datasets (around 100 of 3D MR brain images). The Mean Average Precision (MAP) and average query time by using the approaches of 3D GLCM, 3D WT, 3D GT and 3D LBP are show in Table 4. Query time is the accumulated time spending on both feature extraction and retrieval. All programs were written in the software of Matlab R2009a running on a computer with specifications of Intel P8600 CPU of 1.58GHz and 3.45GB RAM.

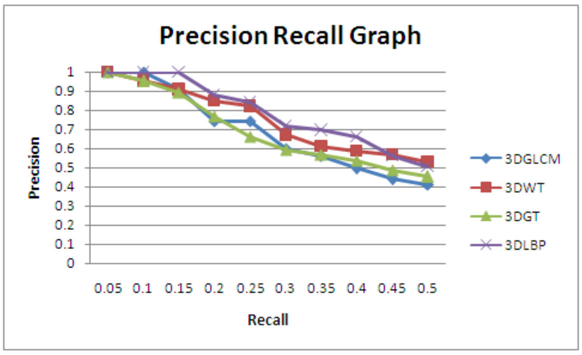


Figure 13. Average precision recall graph for ten queries.

Table 4. Mean Average Precision (MAP) and Query time for 4 texture representation methods

Methods	**Mean Average Precision (MAP)**	**Query time**
3D GLCM	0.690	10.96s
3D WT	0.749	1.22s
3D GT	0.691	10.77m
3D LBP	0.786	0.21s

The above results show the approach of LBP not only can achieve precision rate by up to 78% but also can perform retrieval in real time with sub-second speed. Comparison with the other three texture-based methods, including 3D Grey Level Co-occurrence Matrices, 3D Wavelet Transform and 3D Gabor Transform was also carried out with LBP outperforms over them in terms of both retrieval precision and processing speed.

7.3 Results on Subjective Evaluation

An on-line questionnaire shown in Appendix 2 was applied to subjectively evaluate and thereafter further improve the system. This questionnaire was consisted of three parts covering the general impression of the repository, system evaluation and comments on the system respectively. This survey was carried out by MSc students and lecturers at MU. They feedbacks are summarised below.

- **Subjects**

  15 subjects completed the questionnaire, 80% being on MSc programme of BioMedical Modelling and Informatics (BMI). The general information about respondents is displayed in the following graphs.

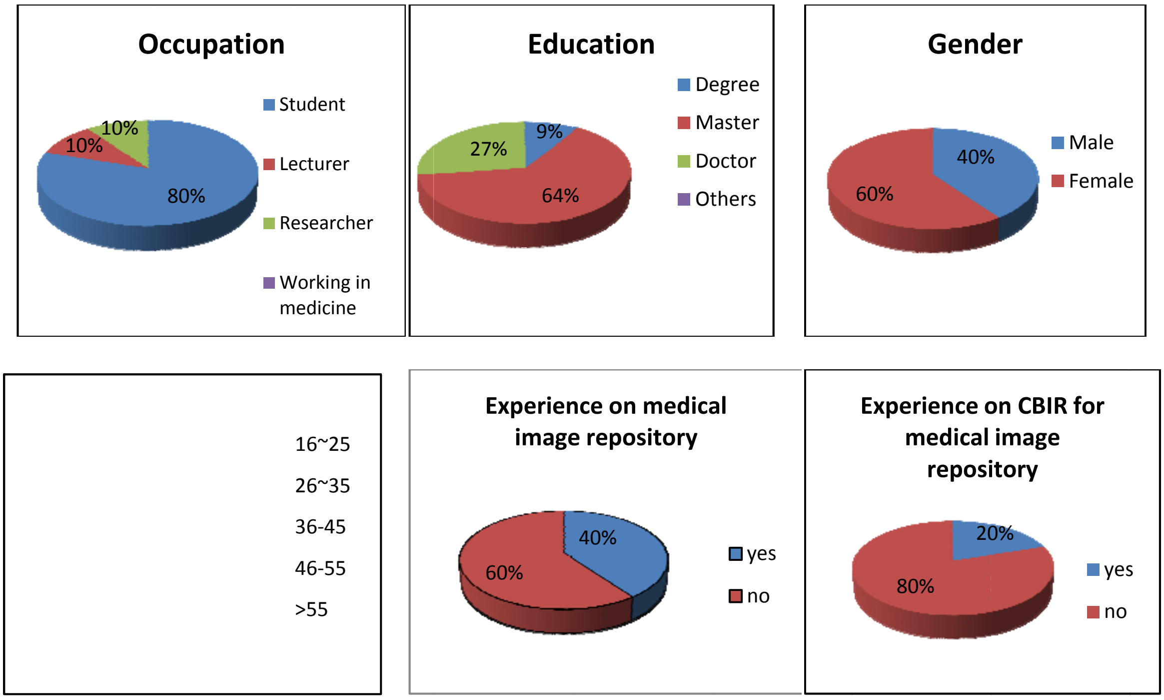
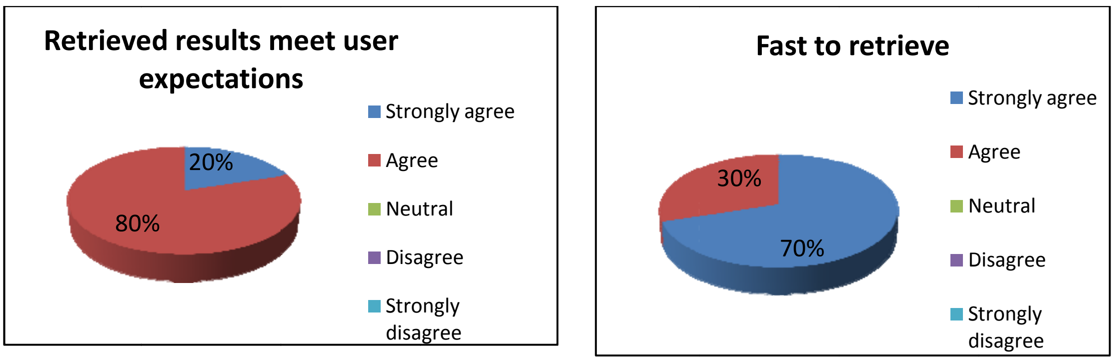


Figure 14: Information of respondents.

The respondents are nearly evenly split between females (60%) and males (40%), whereas their ages cover a wider range with the greatest percentage (70%) group being aged between 16~25. Nearly all the respondents are either of a postgraduate or with a postgraduate degree. 60% respondents have the experience of using medical image repositories. However only 20% have experience using CBIR for medical image retrieval.

- **System evaluation**

As shown in Figure 15, all respondents 'agreed' (80%) or 'strongly agreed' (20%) with the retrieved results, suggesting that the system meets users' expectations. They all 'agreed' or 'strongly agreed' that the system was fast and easy to use, and was useful to teaching and learning.



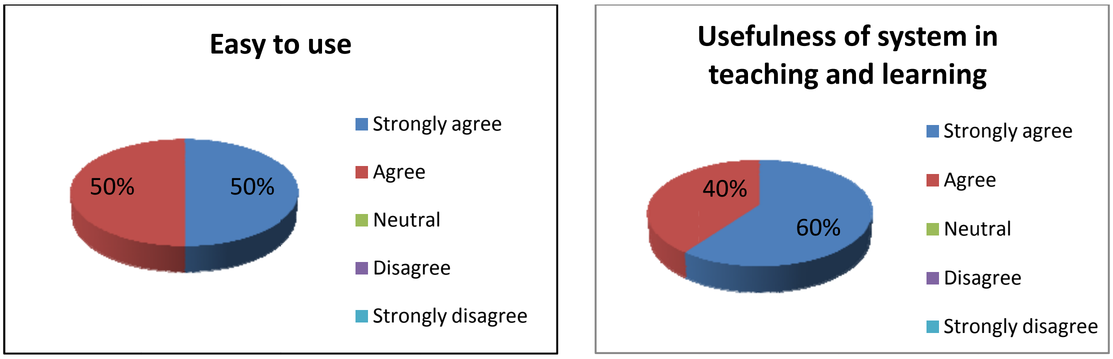


Figure 15: System evaluation

- **Comments**

  Over half of the respondents gave comments regarding to the features of the repository provided. These are summarized and cited as below.

  <u>Positive features</u>: fast, reliable, easy to use, accurate and interesting.

  For example:

> *"Using image as a query makes it easy to get the exact result especially for the students. Because of this system, we do not have to depend on our verbal expressions or some other features to find out the image we want and because of that I think it is also very time saving."*

> *"The system works quite fast. And the good thing about this system is that you can keep on refining your results. For example if you get images from first query then from within the result you can select further images as relevant or non-relevant one and in this way you can narrow down your research and get to the kind of image you want so this thing is quite flexible and gives lot of power at the hands of user."*

  <u>Negative features</u>: unfriendly interface and query mode.

  For example:

> *"The major weak features I think is its user interface graphics. The homepage is not that appealing. I think it is not very user friendly."*

> *"The thing that is bit restricting about this project is that in the beginning you just get random results and you might get random results that might be far different from the kind of image you are looking for and it would take you much longer time to drill down to the kind of image you are interested in."*

Drawing from these users' comments, a new proposal has been submitted to JISC (under Repository Embedment scheme with outcome coming in January 2011) in an aim to further improve user interface and query mode.

7.4 Engagement with the community

The developed tools would support image retrieval from collections of any form including file-based, relational and XML databases, making a step forward towards communication of medical images in terms of contents. All the source code from the project are available to the public as open source and instructions are given at Appendix 1. Talks were given at one international conference eHealth 2010 and will be given at eTelemed 2011 in the hope to publicise the project findings. Project webpage [7] has been and will be kept updating to present the latest development of the project.

7.5. Stakeholder analysis

The project partners three partners from Middlesex University. They are EIS (Engineering and Information Sciences), CLQE (Centre of Learning and Quality Enhancement) and CIE (Centre of International Education).

To EIS, The established repository provides improved access and comprehensive content that can be engaged with their MSc BMI programme, arriving at more successful teaching and learning outcomes. The research results gained from the project will enhance the research activities in the School and will attract wider publicity and higher number of MSc student intake. Whereas to CLQE, the online OASIS+ e-learning system is complemented by adding to its capacity and services without a corresponding increase in resource, leading to a higher quality of learning based on e-resources. In particular, it will benefit the students located at overseas to find rich supply of learning resources in one system. On the other hand, with the help of OASIS+, MIRAGE can serve the other campuses of MU, including those overseas, thereby accommodating equal accessibility.

Furthermore, to JISC, the experience obtained in this project in archiving medical images could be re-applied in similar projects, and the software (as well as the interface with OASIS) can be easily shared in the UK HE/FE for education and research purposes since ~80% UK universities are adopting Blackboard Vista.

8. Outcomes

MIRAGE has been successfully completed and has met all its objectives. In particular, MIRAGE has bore fruits in the followings:

- Establishment of a repository warehousing 100,000 2D images and 100 3D images.
- Multimodal retrieval facilities for the repository.
- Development of an approach for retrieval of 3D images.
- Interfacing with OASIS+.
- Project web page at http://www.mitime.org/mirage/
- Comprehensive evaluation of the developed system.
- A user study.
- Two publications:
    - X. Gao, Y. Qian, et al., Texture-based 3D image retrieval for medical applications, IADIS e-Health2010, July 2010.
    - Y. Qian, X. Gao, et al., Content-based Retrieval for 3D Medical images, eTELEMED 2011, France, *to appear*.

MIRAGE provides a platform to access and search medical images based on visual similarity. In terms of teaching and learning, lecturers and students can easily find a set of interesting images from the repository through the use of visual access methods. At present, this repository is being employed considerably by Middlesex MSc students who are on the programme of Biomedical Modelling and Informatics to conduct their coursework and final dissertations. Other PhD students who are interested in

[7] http://www.mitime.org/mirage

medical images are also trying the system. In the long term, embedding with OASIS+ is envisaged (not just interfacing as the case at the moment). Semantic retrieval, for example, the query could be '*find me brain images with a tumour of 2cm in diameter*', will also be facilitated.

During the development of MIRAGE, a technical issue was indentified, which might relate to many other developers who adapt open source software. That is compatibility. MIRAGE was built on an open source software GIFT on a server that was running an old version of operation system Red Hat (as the server was setup 7 years ago). Since the server has rapidly reached its limit of 40 GB storage space, replacing it with a newer one is an obvious solution. However, the new server can only come with a newer version of operation system (OS) (Linux Red Hat 5.0 or Debian 5.0) that appeared not being compatible with the current open source GNU GIFT code, i.e., by copying the old software to the new server did not work. Installing a newer version of GIFT (2005) is therefore in progress, which somehow posed more challenges than converting old version of GIFT to the newer OS. Communications with GIFT support service had been in progress, although took longer that initially thought. After discussion with our supporting staff, the decision was made to keep the old OS while increasing the hardware space. At the mean time, configuration of new OS with newer version of GIFT is still in progress by our supporting staff and hopefully the new server is configured successfully soon.

9. Conclusions

In summary, the development of subject-based repository of MIRAGE not only provided with students an addition to the online learning systems but also enriched the project team with an in-depth experience in the design and development of internet-based e-learning system. It is anticipated that the system will be widely used by the teaching, learning, and research communities. As a programme leader of MSs BMI, I have already seen that the arrival of MIRAGE benefits our teaching and learning considerably. It is expected that the students complete their course work and dissertations with a rich supply of relevant data in their hands.

To ensure the system MIRAGE extensible to live up to its full potentials, a number of improvements are identified. The main areas for further development include

- An interface for visualisation of 3D images
- A facility for uploading query images from users – at present, this function is disabled deliberately for the concern of server safety. Other security measures should be in place once this function is enabled.
- The combination of text-based and content-based retrieval – these two functions at the moment work independently. Communication with similar projects will help in this regard.

10. Recommendations (optional)

One minor recommendations stemming from this project is

- If time permits, JISC should invest more time in giving advices to start-up projects that had little experience to begin with – none of JISC staff visited this project team during the project life-time. However, JISC weekly newsletters (via email) provided very useful information on similar projects.

11. References

- MEDINFO 2007*, http://www.chi.unsw.edu.au/CHIweb.nsf/page/Conference%20Papers*
- http://www.mic.ki.se/MEDIMAGES.html
- http://en.wikipedia.org/wiki/SCORM
- http://imageclef.org/
- http://www.gnu.org/software/gift/
- X. Gao, Y. Qian, et al., Texture-based 3D image retrieval for medical applications, IADIS e-Health2010, July 2010.
- http://oasisplus.mdx.ac.uk
- http://www.surveymethods.com/
- http://www.mitime.org/mirage

12. Appendixes (optional)

Appendix 1: GIFT Instruction

1. Install GIFT in Linux

After the GIFT is downloaded from ftp://ftp.gnu.org/gnu/gift, the installation of GIFT on a Linux box is quire straightforward by using the following commands:

```
% ./configure  --enable-multithreading
%  make
%  make install
```

The major problem likely to arise is that of missing Perl modules. These can be searched at http://search.cpan.org and installed according to the accompanying instructions. Alternatively, they can be installed from the Linux shell prompt as the command:

```
% perl -MCPAN -e 'install XML::DOM'.
```

2. Add image collection into GIFT

The GIFT software is only installed on server side (http://image.mdx.ac.uk). The indexing and feature extraction function (gift-add-collection.pl being in /usr/local/bin/) is used to process off-line a collection of images stored on the server. Features are extracted from this collection of images and thumbnail pictures are generated for the user client interface. The GIFT server can subsequently be started on the server machine with a configuration file showing where to find the images, thumbnail images, and feature files. The commands below show images from the /home/tempfriend/brain_image folder that are added to the database under the collection name 'Brain' as shown in interface.

```
gift-add-collection.pl \
--gift-home /home/tempfriend/config/ \
--image-directory /home/tempfriend/brain_image/ \
--collection-name brain\
--url-prefix http://image.mdx.ac.uk/vin/brain_image/ \
--thumbnail-url-prefix http://image.mdx.ac.uk/vin/brain_thumbnails/ \
```

The step of indexing and feature extraction results in the following hierarchy:

- Thumbnail images stored in /home/tempfriend/brain_thumbnails
- Feature files saved at folder /home/tempfriend/config/gift-indexing-data/brain

Then, the image collection and the corresponding image thumbnails are copy to http://image.mdx.ac.uk/vin/ manually by using the commands:

- cp /home/tempfriend/brain_image/* /var/www/html/vin/brain_image
- cp /home/tempfriend/brain_thumbnails/* /var/www/html/vin/ brain_thumbnails

After closing any previously running session of the server, GIFT is started by using the command as below

```
gift  --datadir  /home/tempfriend/config  --port 12790
```

Note that any modifications to the configuration (changing in URL of images and thumbnails, the name of collection,etc) can be done by changing the configuration file in /home/tempfriend/config/gift-config.mrml and restart the server by using the command:

```
gift  --datadir  /home/tempfriend/config  --port 12790.
```

3. Client interface

This review is based on the PHP web interface on a Linux server at http://image.mdx.ac.uk/vin/demo.php. The login page allows the client to establish a connection with the server and to obtain information from the server. Once a session is started, the page displays the names of the collections available on the server together with the associated algorithms for assessing similarity.

To set up a query, random images first need to be fetched from a chosen collection by pressing the "Random" button. This operation displays the chosen number of images randomly selected from the collection. To start with, there is no preference for any of the retrieved images. By default, all the displayed images have their relevance set to 'neutral', but this can be changed to 'rel' (relevant) or 'non-rel' (non-relevant) using the drop-down menu beneath each image. The user can indicate his preference as 'rel' or aversion as ‘non-rel’ for as many images as he/she wants, and leave the others as ‘neutral’. By clicking the 'Query' button, the user activates the client to generate a message to be sent to the server concerning those images selected as related or not. The server returns a list of those images that most (least) similar to the images as a query in apppearance. A similarity score which depends on the algorithm chosen is also returned for each image. The search can be further refined by marking the images from the resulting list as 'rel' or non-rel’. A new search can be initiated by pressing the ‘Clear’ button.

Image uploading from the client to the server is disabled at present (shown in red) because of potential security reasons. This may change at a later stage if authentication and password verification are introduced for connecting to the server.

4. Links:

Here are some useful links:

- GIFT website: http://www.gnu.org/software/gift/
- GIFT download: ftp://ftp.gnu.org/gnu/gift
- GIFT communication protocol: http://www.mrml.net/
- GIFT help archive : http://lists.gnu.org/archive/html/help-gift/
- Viper group (publications and links) : http://viper.unige.ch/doku.php

Appendix 2: Questionnaire on MIRAGE

Part 1: General information on user

1. What institute are you from?

2. You are
   Student
   Lecturer
   Researcher
   Working in medicine
   Others

3. What subject are you current teaching, learning, researching and interesting?

4. Education
   Degree
   Master
   Doctorate
   Others

5. Gender
   Male
   Female

6. Age
   16~25
   26~35
   36~45
   46~55
   >55

Part 2: Evaluation for MIRAGE

7. Did you ever use any on-line medical image repository?
   Yes
   No

8. Did you ever use visual similarity method to search and browser medical images?
   Yes
   No

9. The retrieved results of this system meet your expectations
   Strongly agreed
   Agreed
   Averaged
   Disagreed
   Strongly disagreed

10. This system is fast to retrieve images.
    Strongly agreed
    Agreed
    Averaged
    Disagreed
    Strongly disagreed

11. This system is easy to use.
    Strongly agreed
    Agreed
    Averaged
    Disagreed
    Strongly disagreed

12. The visual image search is a good method of retrieving medical images for teaching, studying and researching.
    Strongly agreed
    Agreed
    Averaged
    Disagreed
    Strongly disagreed

Part 3: Comments

13. What are the positive features of MIRAGE?

14. What are the weak features of MIRAGE? Could you provide any suggestions to improve the system?

15. Any other comments?

Appendix 3. A publication on 3D Image retrieval at eTelemed 2011 (*to appear*).

Content-based Retrieval of 3D Medical Images

Y. Qian, X. Gao *, M. Loomes, R. Comley, B. Barn
School of Engineering and Information Sciences
Middlesex University
London, NW4 4BT, United Kingdom

R. Hui, Z. Tian
Department of Neurosurgery,
General Navy Hospital,
Beijing, P.R. China

**Corresponding author:* x.gao@mdx.ac.uk

***Abstract* -- While content-based image retrieval (CBIR) has been researched for more than two decades, retrieving 3D datasets has been progressing considerably slowly, especially in the application to medical domain. This is in part due to the limitation of processing speed when 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, significant amount of exploitations are focusing on the extraction of 3D shapes. However, it appears other information tends to be equally important in clinical decision making. In this paper, Local Binary Pattern (LBP), the texture based approach stemming from 2D forms, has been studied extensively through the application to 3D images from a collection of MR brain images in a content-based image retrieval system (CBIR). The initial results show LBP not only can achieve precision rate of up to 78% but also can perform retrieval in real time with sub-second processing speeds. Comparison with the other three popular texture-based methods, namely 3D Grey Level Co-occurrence Matrices, 3D Wavelet Transforms and 3D Gabor Transforms, is also carried out. The results demonstrate that LBP outperforms over them both in terms of retrieval precision and processing speed.**

Keywords – CBIR, 3D image retrieval, 3D texture extraction.

1. 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 can not be representative due to the fact that slices are getting thinner (i.e., 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 images 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 the retrieval of relevant 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, i.e., the view angle is at the back of the object. In addition, in 2D cases, the viewing angle is always 0^o, being normal to the computer screen, by which most existing algorithms can fulfill this request. 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 many current studies being focusing on the invariance of transformations (including rotation, scaling and translation) of objects, which has more to do with shapes. In [1], 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. In this way, a database has to be consisted of objects differentiated by shapes, such as airplanes, chairs, etc.. Similarly, in order to achieve transformation invariance, a graph-based shape descriptor is created in [2] to determine similarity between 3D objects. More recently, the retrieval of 3D objects has been attempted using impact descriptors [3] to capture the surrounding areas of a 3D shape in order to offer a histogram of time-space curvature that are invariant of rotation and translation. Other shape-based 3D models are included in [4][5][6][7]. 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 [8], which is however in principle, still capture the outliner of a shape at each depth (z-buffer).

For the 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.

One way to represent texture is 2D-based, since a 3D dataset is composed of a stack of 2D slices. However, using a slice-by-slice 2D approach suffers from the drawback 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. With this in mind, in this study, the approach of Local Binary Pattern (LBP) [9] is extended 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. Comparison with the other three popular methods in texture representation is also carried out, including Grey Level Co-occurrence Matrices (GLCM), Wavelet Transforms (WT) and Gabor Transforms (GT). The novelty of this work is the extension of 2D texture features into 3D while minimizing the calculation cost. This is achieved by the introduction of a pre-processing stage of a selection of potential VOIs into query datasets. In 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, before preceding with the extraction of 3D texture features and the calculation of similarities. This work forms part of our currently online CBIR system at [10]. The structure of the paper is in the following pattern. Section II explains the methods employed in the study, whist Section III shows the experimental results. The conclusion and discussion are given in Section IV, which is followed by Sections of Acknowledgment and References.

II. METHODOLOGY

In this investigation, at the phase of ingestion of the data (at least two phases should be in place including ingestion and retrieval in the system), the collected data firstly undergoes a pre-processing stage to normalize them into the same resolution before the indexing stage, as shown in the flow chart in Figure 1. 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. Thereafter, 3D texture features from a query are 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 explained in the following sub-sections.

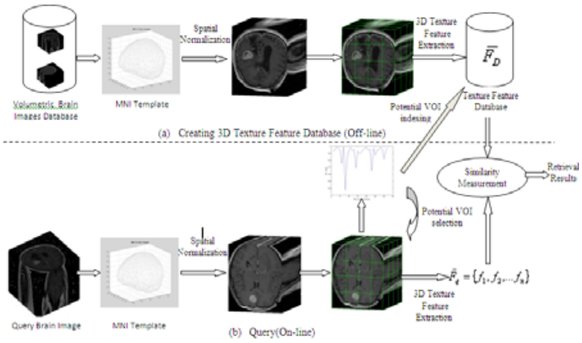


Figure 1. Framework of 3D MR image retrieval.

A. Spatial Normalization

In practice, data are collected from different sources, therefore brain images vary in both shape and size. To make inter-individual brains comparable, it is necessary to transform the dataset of each individual brain into a standard brain template.

Statistical Parametric Mapping (SPM5) [11] is used in this regard to spatially normalize a brain image to an MNI template [12]. In this way, all the images in the database are of the same size of 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 illustrated 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 (3D LBP)

The Local Binary Pattern operator is derived from a general definition of texture in a local neighbourhood (e.g, 8 × 8 pixels). In 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 method to 3D images, similar to [13], 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 left-right (LR), Anterior-Posterior (AP), and Superior-Inferior (SI), as shown in Figure 2.

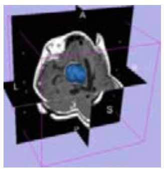


Figure 2. An example of three orthogonal planes in a 3D brain.

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 as being 177 (=59×3) elements.

D. Lesion Detection

The main purpose 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, i.e., after a query has been submitted, 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 potential lesions from sub blocks is carried out first to highlight abnormalities, such as tumours, to speed up the retrieval process.

To do this, the characteristic of bilateral symmetry of a brain along its mid-plane (parallel to SI direction as shown in Figure 2) is assumed. Similar to [14], by comparing the left half with the right half of a hemisphere along this middle symmetry plane, the abnormality is 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.

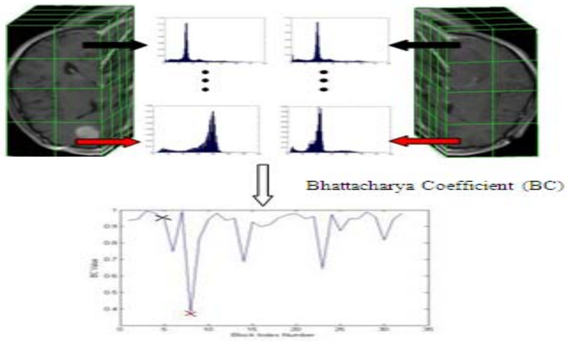


Figure 3. Potential VOI selection

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 calculated. The Bhattacharya Coefficient (*BC*) [15] is then computed between two normalized histogram H_L and H_R, which are obtained from two mirror symmetric sub-blocks as defined in Eq. (1).

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

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 and 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', and 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, the 3D texture features are extracted exclusively from this VOI of the query, and are later compared with the features from similar blocks of 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 two images. For a 3D LBP, the histogram intersection is applied to measure features of histograms that is given in Eq. (2),

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

where *i* represents each bin in the 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

The database contains over 100 MR brain images with lesions (e.g., tumour, biopsy) and detailed diagnosis. Each dataset is of resolution of $256 \times 256 \times 44$ 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 1, the first row is the labeling number of the location of a VOI assigned by us for the convenience of calculations, e.g., '1' refers to the abnormal part in the front top left-most part of the brain. The second row is the total number of images containing VOIs in such positions in the database, whilst the number of correctly detected images by the LBP is given on the third row. Therefore the overall performance of the LBP 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 1 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**

In terms of precision, the retrieved accuracy is 78% based on ten query images with the ground truth being diagnostic information relating to the locations and sizes of tumours, demonstrating very promising results.

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). These are summarized next.

In 3D form, grey level co-occurrence matrices [16][17] are defined as three dimensional matrices of the joint probability of occurrence of a pair of grey values separated by a displacement *d = (dx, dy, dz)*.

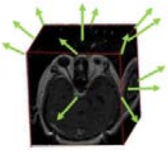


Figure 4. Thirteen directions in 3D GLCM.

For example, four distances with 1, 2, 4, and 8 voxels respectively and thirteen directions, as depicted in Figure 4, and chosen in this study, will produce 52 (=4×13) displacement vectors, and thereafter 52 co-occurrence matrices. As a result, four Haralick texture features [18], being energy, entropy, contrast and homogeneity, are computed from each matrix, generating a feature vector with 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, which is then followed by a 2 to 1 sub-sampling of each output volumetric image [19], 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), providing a 3D wavelet transform of two scales as shown in Figure 5(b).

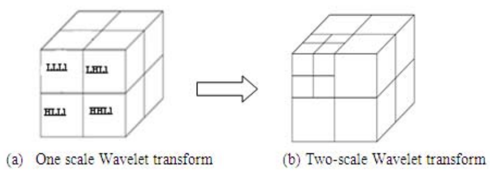


Figure 5. One scale and two scales of 3D WT.

With respect to Gabor Transforms, in order to extend GT into three dimension, a set of 3D Gabor filters are generated similar to [20][21] to detect spatial orientations and scale tunable edges and lines (bar), which can be formulated as Eq. (3).

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

To calculate similarity distances from these three methods, a normalized Euclidean distance is employed to compare two 3D patterns in a feature space, which is defined by Eq. (4).

$$D(Q,I)=\sqrt{\sum_i\left(\frac{Q_i-I_i}{\sigma_i}\right)^2} \quad (4)$$

where σ_i is the standard deviation of a set of representative features over the entire database and are utilized to normalize each individual feature component. The retrieved 3D images are ranked in 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 is the size of local features multiplied by 64, the number of 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.

The performance of image retrieval is evaluated based on the 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, which 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 standard recall points, i.e., 10%, 20%,…,100% or 0.1, 0.2, …, 1. To represent a P-R graph using a single value, usually, the Mean Average Precision (MAP) value is employed to assess the overall performance for all queries and is calculated as

$$\text{Mean Average Precision (MAP)} = \frac{1}{M}\sum_{i=1}^{M} AP_i \quad (5)$$

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)

$$\text{AveragePrecision (AP)} = \frac{1}{N_r}\sum_{j=1}^{N_r} P_j \quad (6)$$

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 being the results without a pre-processing stage of VOI selection and Figure 7 with the pre-processing stage.

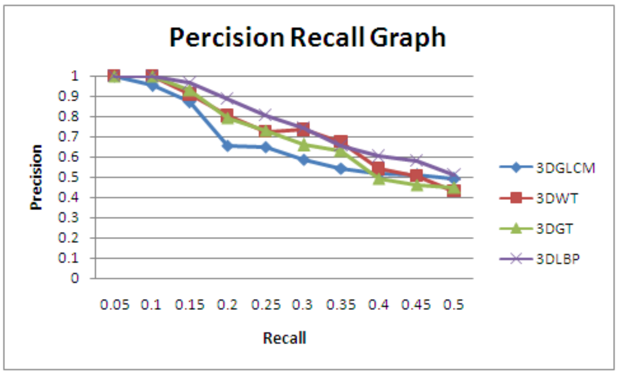


Figure 6. Average precision recall graph for ten queries without VOI selection.

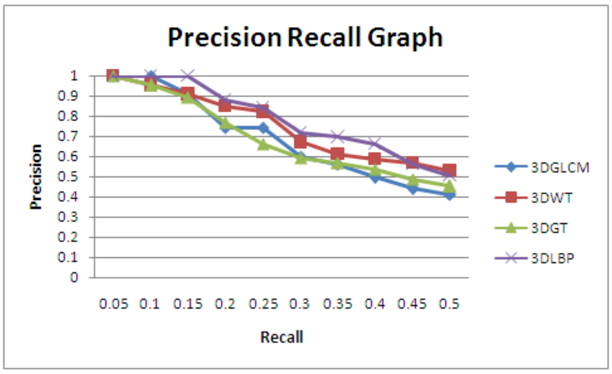


Figure 7. Average precision recall graph for ten queries with VOI selection.

In summary, 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 show in the following table.

TABLE 2 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 little improved performance.

Figures 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 [3].

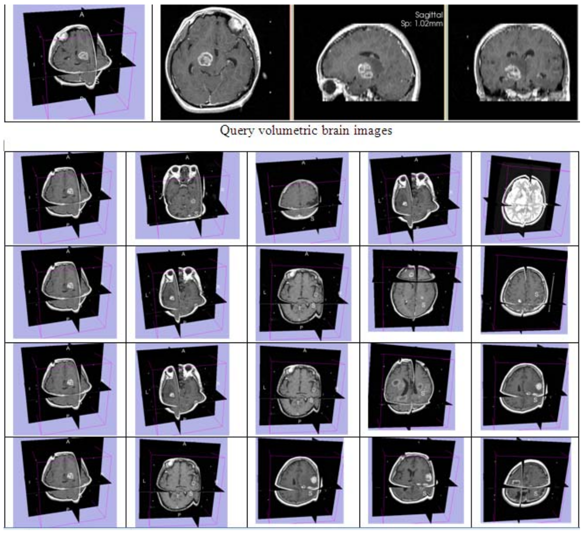


Figure 8. Retrieved results in top 5 ranking from
3D GLCM (row 1), 3D WT(row 2), 3D GT (row 3), and 3D LBP (row 4).

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 higher dimensions. Table 3 demonstrates the average querying time, amounting to the times 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. All methods run in Matlab R2009a with CPU of an Intel P8600 1.58GHz CPU and 3.45GByte RAM.

TABLE 3 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 3, the query time with VOI selection offers 4 times faster operation than that without. In particular, the query time for 3D GT takes a much longer time than the other methods spending 38 minutes, due to the employment of 144 times of 3D convolutions for each block, whereas the query times for the other methods are in the space of few seconds. The table also supports our choice of the 3D LBP approach with sub-second retrieval times and the highest precision rate of 78%, as given in Table 2.

IV. CONCLUSION

In this paper, a texture based approach that draws on the Local Binary Pattern technique has been employed through the extension into 3D format, to retrieve lesioned MR brain images in a CBIR system. The results are very encouraging showing that not only higher precision rates can be achieved, but also that it can be done in real time. In comparison with the other three texture based methods, 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 minutes are needed for each dataset), only ~100 datasets are included in this study. The very next step is to process more datasets. Although the precision rate of 78% is very promising, a better rate should 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. The authors would also like to thank Janet Rix and Alex Chapman at Middlesex University for their recommendations to the project.

REFERENCES

[1] 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.

[2] Bustos, B. Keim, D., Saupe, D. and Schreck,T., "Content-based 3D Object Retrieval", *In IEEE Transactions on Computer Graphics and* Applications, Vol. 27, No. 4, 2007, pp. 22-27.

[3] Mademlis, A., Darasb, P., Tzovarasb, D., and Strintzis, M.G., "3D Object Retrieval Using the 3D Shape Impact Descriptor", *Journal of Pattern Recognition*, Vol. 42 No.11, 2009, pp. 2447-2459 .

[4] 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.

[5] 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.

[6] 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.

[7] B. Bustos, D. Keim, D. Saupe, Tobias Schreck, Content-Based 3D Object Retrieval, IEEE Computer graphics and Applications, 27(4): 22-27, 2007.

[8] 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.

[9] 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.

[10] http://image.mdx.ac.uk.

[11] http://www.fil.ion.ucl.ac.uk/spm/.

[12] Montreal Neurological Institute, http://www.mni.mcgill.ca/ .

[13] Zhao, G. and Pietikainen, M., "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 6, 2007, pp. 915-928.

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

[15] Bhattachary A, " On a Measure of Divergence between Two Statistical Populations Defined by Their Probability Distribution", Bulletin of the Calcutta Mathematical Society.Vol.35, 1943, pp99-109.

[16] Kovalev, V.A, Kruggel, F., Gertz, F.J., and Cramon, D. Y., "Three-Dimension Texture Analysis of MRI Brain Datasets", *In IEEE Transactions on Medical Imaging*, Vol. 20, No. 5, 2001,pp. 424-433.

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

[18] Haralick, R.M, Shanmugam, K., and Dinstein, I., "Textural Features for Image Classification", *In IEEE Transactions on Systems, Man, and Cybernetics*, Vol.3, No. 6, 1973, pp. 610-621.

[19] Mallat, S. G., "A Theory for Multiresolution Signal Decomposition: the Wavelet Representation", *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No.7, 1989, pp. 674-693.

[20] Feng, M. and Reed, T.R.,"Motion Estimation in the 3-D Gabor Domain", *In IEEE Transactions on Image Processing*, Vol. 16, No. 8, 2007, pp. 2038-2047.

[21] Wang, Y. and Chua, C., "Face Recognition from 2D and 3D Images Using 3D Gabor Filters", *Journal of Image and Vision Computing,* Vol. 23, No. 11, 2005, pp. 1018-1028.