CHAPTER 4

IMAGE RETRIEVAL SPECIFIC FAST FOURIER TRANSFORM IRS-FFT FOR EFFICIENT MEDICAL IMAGE RETRIEVAL

4.1 INTRODUCTION

Telemedicine is the delivery of expert opinion by expert health care professional using telecommunication technologies. Typical telemedicine applications need to support videoconferencing, transfer of medical database record including high quality large size medical images (Viswacheda et al 2007). The telecommunication medium could be either wired or wireless with wireless gaining popularity due to its extensive usage in e-health care and ambulatory systems despite being extremely lossy and error prone (Panayides et al 2008). Telemedicine provides medical expertise in emergency and disaster scenarios when the patient is geographically separated from the health professional by the use of advanced information technology services (Kyriacou et al 2003). Remote patient monitoring through wireless technology has also instigated research interest.

One of the major challenges faced in telemedicine, is the availability of an expert either online or offline to advice on the diagnosis, based on the medical image. This problem is addressed by the use of content based image retrieval (CBIR), as the user could retrieve diagnostic cases similar to the query medical image from the medical database. Medical image database contain a vast amount of data in visual format. Conventional
retrieval methods based on text and numerical data is cumbersome, as indexing the whole database is not possible. Moreover, the image content is more versatile than text. Content based image retrieval (CBIR) depends upon the ability of the algorithms to extract relevant image features and organize the same to represent the image (Samuel et al 2004). CBIR extracts the features like colour, texture, shape of the query image automatically and match them with the extracted features of the images in the database. (Henning Muller et al 2004). The images from the database are then retrieved based on the similarity measures. Figure 4.1 shows block diagram of CBIR process. CBIR is increasingly applied in medical image applications. A good number of CBIR algorithms, architectures and systems are reviewed in literature (Smeulders et al 2000). The CBIR systems could be also used for medical diagnostics and training of physicians (Lehmann et al 2003).

**Figure 4.1 CBIR General Block Diagram**

The limited bandwidth and a large amount of diagnostic data in image format need to be transmitted. This is a major issue in telemedicine which is addressed by image compression. Image compression reduces the
amount of data required to represent a digital image, thereby resulting in the reduction of storage and transmission requirements. Compression is obtained by removal of redundancies from the image, thereby yielding a compact representation of the image. The image compression is broadly classified as lossless or lossy techniques depending on whether the original image could be perfectly recovered from the compressed image or not. The amount of retained energy could be defined during compression, in order to maintain the fidelity of the data to be compressed. In order to reduce the bandwidth utilization in a telemedicine application, a lossy compression technique with maximum retained energy which is tolerable could be applied.

This chapter focuses on the image retrieval problem using compressed images and the impact of compression in the classification accuracy. This study aims to overcome the bandwidth limitations of wireless networks which are an essential part of telemedicine in an emergency department communicating either with an ambulance or with a remote disaster recovery team.

This chapter deals with the proposed methodology to extract features using a modified Fast Fourier Transform (FFT) named Image Retrieval Specific Fast Fourier Transform IRS FFT. Frequency domain analysis has produced accurate results for image classification problems and various methods are available in literature. The present chapter focuses on feature extraction using Fast Fourier transform FFT and the proposed IRS-Fourier transform. The extracted features from both these methods are used for image classification and a comparative study was presented. The classifiers considered to compute the classification efficiency are Naive Bayes Classifier, Radial Basis Function (RBF) and Multilayer Perceptron Neural Network (MLPNN).
4.2 HAAR WAVELET

Haar wavelet has been extensively used in image compression with good results. Wavelets could be described as a set of non linear bases which could be used to approximate a function by choosing appropriate wavelet basis function (Haar 1910). Wavelets use dynamic basis functions and hence the input function could be represented more efficiently compared to static basis function. The time series could be viewed in multiple resolutions, each with a different frequency. The average and differences of the signal is continuously computed breaking the signal into spectrum. Haar wavelet could be computed as follows for an array containing n samples:

1. The average of each pair of samples to produce n/2 averages was computed, in order to fill the first half of the array.
2. The difference between each average and the first samples was computed to produce n/2 differences which are filled in the second half of the array.
3. The above steps were repeated on the first half of the array.
4. The same process was repeated for each column for a two dimensional Haar wavelet.
5. All values less than or equal to the threshold value was eliminated.

4.3 FAST FOURIER TRANSFORM

A matrix of order $n \times n$ (n is the number of sample points) represents an image. To approximate the function using the discrete Fourier transform is computationally costlier (Cochran et al 1967). As $n^2$ arithmetic operations is required to multiply a $n \times n$ matrix by a vector, the computational problem gets worse with the increase in sample points n. But, in case of uniformly spaced
samples, Fourier matrix can be factored into a product of few sparse matrices. Thus, only nlogn arithmetic operations are required while multiplying the resulting factors with a vector. This operation is called Fast Fourier Transform (FFT).

Both the FFT and the Discrete Wavelet Transform (DWT) are linear operations, which on application generate a data structure containing $\log_2 n$ segments of various lengths and further fill and transform it into a different data vector of length $2^n$. Matrices formed by both the FFT and DWT have similar mathematical property, like the inverse transform matrix are the transpose of the original. Thus, both transforms can be seen as rotation in function space to a new domain. In FFT, the new domain consists of sines and cosines whereas for the wavelet transforms new domain consists of wavelets, mother wavelets, or analysing wavelets. Another similarity is that the basis functions are localized in frequency (Cooley et al 1965).

4.3.1 Proposed IRS-Fourier Transform

In order for learning algorithms to have higher classification, it is required to have the attribute value range as small as possible. However, Fast Fourier Transform coefficients has large values, which could affect the prediction capability. Though Fourier transform pairs using Gaussian, Sine, Cosine and Lorentzian have been proposed in literature they all face the problem of large scaling between coefficients. In order to overcome the large degree of variation in the attribute value, the Image Retrieval Specific abbreviated IRS-Fourier transform was proposed and implemented, which achieves dimensionality reduction.
Consider a matrix $U$ of dimension $N \times N$ with its $(m,n)^{th}$ element defined as

$$u[m,n] = \left( e^{-\frac{j2\pi}{N}} \right)^{mn}$$  \hspace{1cm} (4.1)

where, $(m,n = 0,1,2,...,N-1)$

where,

$$u_N = \frac{1}{\sqrt{N}} e^{-\frac{j2\pi}{N}}$$  \hspace{1cm} (4.2)

The IRS-FT could be defined as

$$X_{irs}[n] = \int_{-\alpha}^{\alpha} \frac{1}{\sqrt{2\pi} \sigma^2} e^{-(n-\mu)^2/2\sigma^2} X[n]$$  \hspace{1cm} (4.3)

where,

$$X[n] = \begin{bmatrix}
u[0,0] & u[0,1] & \ldots & u[0,N-1] & x[0] \\
u[1,0] & u[1,1] & \ldots & u[1,N-1] & x[1] \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
u[N-1,0] & u[N-1,1] & \ldots & u[N-1,N-1] & x[N-1] \end{bmatrix}$$  \hspace{1cm} (4.4)

And

$$\mu = \frac{1}{N} \sum_{n=0}^{N-1} X(n)$$  \hspace{1cm} (4.5)

The dataflow diagram of the proposed system could be represented using the butterfly diagram shown in Figure 4.2.
4.3.2 Wavelet Energy Features

The wavelet transforms reduce the correlations among the bands of the decomposition by exact covering of the frequency domain. It also improves the computational efficiency due to adaptive pruning of the transform. The images are decomposed on wavelet basis of resolution one and thus, at each resolution three detail coefficient matrices are obtained. The vertical, horizontal and diagonal structures of the image obtained are modified using magnitude operator. The first and second order moments of the resulting distribution is calculated. The wavelet sub bands are pre-processed by taking the absolute value of the coefficients and respective histograms of the coefficient matrix was computed. Average energy of each detail of the wavelet coefficient extracted is used as a feature. The average energy is computed by summing the squares of each detail image normalized for all the coefficients of the image. The energy is defined as:
where $D_{jl}$ is the detail coefficient image $l$ at resolution level $j$ of size $(M, N)$. This also gives an indication of total energy present at a particular spatial frequency and orientation.

### 4.3.3 Classification Algorithm

Once an appropriate set of features are computed, the next step is to adopt a suitable classification algorithm to assess the feature's discriminative power. In this chapter, the classifiers considered for classification are Naïve Bayes, Radial basis function and Multilayer Perceptron Neural Network.

Image classifiers could be divided as learning based or parametric classifiers and non-parametric classifiers. The former needs learning or training procedure of the classifier parameters for classification decisions. Boosting, decision trees, neural networks, support vector machine are all parametric classifiers. These are widely used for image classification. The non-parametric classifiers do not need training for classification decisions. The most commonly used non-parametric methods are Naive Bayes and nearest neighbour. The advantages of non-parametric classifiers over parametric classifiers are no learning/training phase required; any number of classes could be classified and over fitting of parameters is avoided. Thus, in this study for evaluating classification accuracy one non-parametric classifier (Naive Bayes Classifier) and two parametric classifiers (Radial basis function and Multilayer Perceptron Neural Network) are considered.
4.3.3.1 Naive Bayes Classifier

Probability based Bayesian theorem is the foundation for Naive Bayes Classifier which assumes mutually independent attributes. The naive Bayes classifier assigns an instance \( s_k \) with attribute values \( (A_1=v_1, A_2=v_2, \ldots, A_m=v_m) \) to class \( C_i \) with maximum \( \text{Prob}(C_i|(v_1, v_2, \ldots, v_m)) \) for all \( i \).

Likelihood of \( s_k \) belonging to \( C_i \) is given by

\[
\text{Prob}(C_i|(v_1,\ldots,v_m)) = \frac{P((v_1,\ldots,v_m)|C_i)P(C_i)}{P(v_1,\ldots,v_m)}
\]  

(4.7)

Likelihood of \( s_k \) belonging to \( C_j \) is given by

\[
\text{Prob}(C_j|(v_1,\ldots,v_m)) = \frac{P((v_1,\ldots,v_m)|C_j)P(C_j)}{P(v_1,\ldots,v_m)}
\]  

(4.8)

Therefore, when comparing \( \text{Prob}(C_i|(v_1, \ldots, v_m)) \) and \( \text{P}(C_j|(v_1, \ldots, v_m)) \), we only need to compute \( P((v_1,\ldots,v_m)|C_i) \text{P}(C_i) \) and \( P((v_1,\ldots,v_m)|C_j) \text{P}(C_j) \).

4.3.3.2 Radial Basis Function

A RBF network is composed of two layers: a hidden and an output layer. Network output is computed from the responses provided by the basis (or kernel) functions in the hidden layer nodes. Each of these nodes is characterized by a transfer function with a localized response to input stimulus. Thus, they produce a strong response only when the input vector falls within a small localized region (Hush 1993). The most common used basis function is the Gaussian kernel, which is defined as follows:
\[
\psi_h = \exp\left(-\frac{\|x - c_h\|^2}{2\sigma_h^2}\right)
\]  

(4.9)

where \(\psi_h\) is the output value of the hidden node \(h\), \(c_h\) is the centre of the Gaussian \(\sigma_h\) is the standard deviation (width) of the function, \(x\) is the network input vector and \(\|\|\) represents the Euclidean distance. Hence, the response of an output node \(m\) of the network could be expressed as

\[
y_m = g\left(\sum_{h=1}^{H} w_{hm} \psi_h \right)
\]  

(4.10)

where, \(y_m\) is the response of the node \(m\) in the output layer, the hidden layer contains \(H\) number of nodes, \(w_{hm}\) is the weight connecting hidden node \(h\) with output node \(m\) and \(g(.)\) is the activation function for nodes in the output layer. A linear or a sigmoid activation function could be considered.

4.3.3.3 MLP-Neural Network

Multilayer perceptron (MLP) Neural network are proficient in capturing and representing complex input/output relationships (Antani et al 2003). MLP neural networks are known as supervised network as it requires a desired output, in order to learn. MLP creates a model for mapping the input to the output by learning from training data (Santhanam et al 2011). MLPs consist of several layers of nodes or neurons (input layer, one or several hidden layer and an output layer), interconnected through weighted acyclic arcs. Figure 4.3 shows a graphical representation of an MLP Neural Network.
Figure 4.3 Graphical representation of an MLP Neural Network

Learning of MLP is done using back propagation algorithm. In back propagation, input data is repeatedly presented to the network and with each cycle the output obtained is compared with the desired output. The error is the difference between the output obtained and the desired output. The error is back propagated through the neural network and the weights are adjusted such that the error reduces with each cycle. This process is called training. On training, the MLP neural network is used for classification purposes.

4.4 RESEARCH METHOD

In this work, the image dataset comprising of the 176 MRI scan images with three class values were used. The images were compressed using Haar wavelet with a decomposition value of 1. For each image the overall threshold was selected so that the retained energy for all images was 99%. The threshold value is computed using

\[
\text{threshold}=n\sqrt{2\log(xy)}
\]  

(4.11)

where, n is the noise standard deviation and (x, y) is the size of the image
The flowchart of the experimental setup is given in Figure 4.4.

![Flowchart of the experimental setup](image)

4.5 EXPERIMENTAL RESULTS AND DISCUSSION

The obtained $M \times N$ matrix is used to extract features from the image using Fast Fourier Transform and proposed Image Retrieval Specific IRS – Fourier Transform. An appropriate class is assigned to the selected pixel for each image, in order to train the classification algorithm. The preprocessed data are classified using the classifiers Naive Bayes, Radial Basis Function and MLP neural Network using 10 fold cross validation. The results obtained are tabulated in Table 4.1.
Table 4.1 Classification accuracy by different algorithms on original and compressed image

<table>
<thead>
<tr>
<th>S.No</th>
<th>Type of Image</th>
<th>Feature Extraction Technique</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Naive Bayesian</td>
</tr>
<tr>
<td>1</td>
<td>Uncompressed Original</td>
<td>FFT</td>
<td>68.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRS-FFT</td>
<td>80.11</td>
</tr>
<tr>
<td>2</td>
<td>Wavelet based Huffman Encoding</td>
<td>FFT</td>
<td>69.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRS-FFT</td>
<td>79.54</td>
</tr>
<tr>
<td>3</td>
<td>Wavelet based proposed 3-pattern Huffman Encoding</td>
<td>FFT</td>
<td>72.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRS-FFT</td>
<td>80.11</td>
</tr>
</tbody>
</table>

It is evident from Table 4.1, that the proposed IRS-FT method for feature extraction is able to provide better features for the classification algorithms compared to Fast Fourier transform.

4.5.1 Performance Estimation for retrieval

The proposed system also provides good precision which becomes a desirable feature for retrieval of similar images for diagnosis. The image retrieval parameters are computed as follows

\[
\text{Classification Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Total Number of Tested Samples}} \times 100
\]

\[
\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}
\]
Recall = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in the Database}}

f \text{ Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}

The precision, recall and f-measure of various techniques used are given in the following Table 4.2. Precision and recall values are plotted in Figure 4.5.

**Table 4.2 Precision, Recall and f-measure on the 176 Image Dataset**

<table>
<thead>
<tr>
<th>Type of Image</th>
<th>Feature Extraction</th>
<th>Naive Bayes classifier</th>
<th>Radial basis function classifier</th>
<th>MLP NN classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>f-measure</td>
</tr>
<tr>
<td>Original images without compression</td>
<td>FFT</td>
<td>0.72</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>IRS FFT</td>
<td>0.84</td>
<td>0.80</td>
<td>0.82</td>
</tr>
<tr>
<td>Haar Wavelet based Huffman encoder compression</td>
<td>FFT</td>
<td>0.73</td>
<td>0.69</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>IRS FFT</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>Proposed 3 pattern Huffman encoder compression</td>
<td>FFT</td>
<td>0.73</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>IRS FFT</td>
<td>0.81</td>
<td>0.80</td>
<td>0.80</td>
</tr>
</tbody>
</table>
4.5 **Figure 4.5 Precision and Recall for various classifiers on the 176 Image Dataset**

4.6 **SUMMARY**

The image retrieval efficiency of the classifier with the proposed compression using the proposed IRS-FFT is investigated. Using 10 fold cross validation three different classifiers namely Naive Bayes classifier, RBF and MLP-Neural network classifier were trained and the classification accuracy measured. It is seen that the proposed method of feature extraction works consistently with different classifiers. The proposed compression along with the proposed feature extraction method improves the retrieval for both compressed and uncompressed medical images under study. Classification accuracies of over 90% are desirable for a medical image retrieval application. Hence, in order to improve the classification accuracy further, a neural network based classifier was proposed in the next chapter. In order to achieve higher compression, a hybrid compression was also presented in the next chapter.