Endoscopic Image Retrieval System Using Multi-scale Image Features

Manish Chowdhury
Machine Intelligence Unit
Indian Statistical Institute
203 B. T. Road
Kolkata-108, India
st.manishc@gmail.com

Malay Kumar Kundu
Machine Intelligence Unit
Indian Statistical Institute
203 B. T. Road
Kolkata-108, India
malay@isical.ac.in

ABSTRACT

We present a novel Content Based Medical Image Retrieval (CBMIR) scheme for color endoscopic images using Multi-scale Geometric Analysis (MGA) of Nonsubsampled Contourlet Transform (NSCT) and the statistical framework based on Generalized Gaussian Density (GGD) model and Kullback-Leibler Distance (KLD). The subband images obtained from the NSCT decomposition are divided into number of blocks and then the coefficients of each block of each subband is modeled with GGD parameters and computing the similarity using the KLD among the model parameters. The retrieval performance of the proposed system is further improved using Least Square-Support Vector Machine (LSSVM) classifier. Extensive experiments were carried out to evaluate the effectiveness of the proposed system on endoscopic image databases consisting of 276 images. Experimental results show that the proposed CBMIR system performs efficiently in image retrieval paradigm.

Categories and Subject Descriptors

H.2.4 [**Information Storage and Retrieval**]: System—*Multimedia databases*; H.2.8 [**Information Storage and Retrieval**]: Database Applications—*Image databases*; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing—*Indexing methods*; I.4.7 [**Image Processing and Computer Vision**]: Feature Measurement—*Feature representation*; I.4.9 [**Image Processing and Computer Vision**]: Applications; I.4.10 [**Image Processing and Computer Vision**]: Image Representation—*Statistical*

General Terms

Algorithms, Design, Experimentation, Performance

Keywords

NSCT, Endoscopy, KLD, LSSVM, CBMIR, MGA, GGD

PerMIn '15, February 26 - 27 2015, Kolkata, West Bengal, India

http://dx.doi.org/10.1145/2708463.2709046.

1. INTRODUCTION

Over the last two decades, with the rapid and significant development in technologies and instrumentations, medical imaging has become an increasingly critical area of scientific research and plays a vital role in a large number of health-care applications. Now-a-days an increasing number of medical image modalities are being used to support physicians in their clinical diagnosis and treatment. In large hospitals and diagnostic centers, terabytes of digital medical images are being generated and stored daily for diagnosis of different types of diseases, research and education etc., [17, 29]. In today's health-care framework, these images are archived in Picture Archiving and Communication Systems (PACS) [22, 15]. Using PACS, the medical personals may want to browse through similar content images in this archive to ensure a proper diagnosis and treatment planning [2, 1]. Effective and efficient searching in these large image collections poses significant technical challenges and it raises the necessity of constructing intelligent Content Based Medical Image Retrieval (CBMIR) system. The current state-of-the-art in CBMIR approach has been presented in [1, 24, 4].

A modern CBMIR system consists of four main parts: feature extraction, classification, similarity measurement and relevance feedback, respectively. Relevance feedback is outside the scope of this paper. Generally, there are two types of feature extraction schemes (*i.e.*, spatial and spectral domain analysis) exist in the literature, to find the salient and significant information of the medical images [10]. It has been found that the spectral domain approaches are more robust to noise and capable of representing images more effectively than the spatial domain approaches [7, 25].

Among the various spectral approaches, Wavelet Transform (WT) has been used in CBMIR systems extensively [25, 21, 20]. But the problem with WT is its inherent non-supportiveness to directionality and anisotropy. To overcome these limitations, recently a theory called Multi-scale Geometric Analysis (MGA) for high-dimensional signals has been proposed and several MGA tools were developed like Curvelet Transform (CVT), Contourlet Transform (CT), Nonsubsampled Contourlet Transform (NSCT) etc., with applications in diversified problem domains [12, 3, 8]. CBMIR systems based on these MGA tools have been found to perform better than WT based schemes [27, 13, 5]. In this paper, we further generalize this framework into NSCT domain for natural color medical image retrieval.

After representing the image's content by its different low level features vector, the images are classified into various

groups by its visual content representation. An open challenge for automatic categorization of medical images is the inter-class versus intra-class variability problem: an image that belongs to a particular visual class might look very different from the other images of that class, while images that belong to different visual classes might look very similar [2, 1, 30]. Several CBMIR prototypes have been proposed to address this above mentioned problem using image's content representative features [23, 1, 16]. Moreover, many machine learning techniques (such as Multilayer Perceptron (MLP) [20], Support Vector Machine (SVM) [18], etc.) have been used to improve the retrieval accuracy, and to decrease the computational complexity of the retrieval system [23]. However, the training process of MLP is lengthy and performance is normally relies on the initial parameter setting. Similarly, training with SVM is computationally expensive for high dimensional data sets [28]. Therefore, to reduce the time complexity and prolonged training process for parameters optimization, we have used LSSVM in our proposed CBMIR system.

Traditionally, the most commonly used distance measures in image retrieval system are Manhattan distance (MD) and Euclidean distance (ED). But the major problem of these distances are their non-rotational invariance. Therefore, in contrast to these distance measures, Kullback Leibler Divergence (KLD), an information-theoretic measure is found to be more effective in image retrieval system [11, 19, 6].

Even though statistical modeling has been used extensively in texture analysis problem [11], the effectiveness of these approaches is not explored well in the field of natural color medical image retrieval. Therefore, in this article, we propose a statistical model based on GGD and KLD over NSCT subbands coefficients to investigate its performance in color endoscopy image retrieval problem. But, in addition to color information, our approach extends the feature extraction process by using a window, generating more local features. After feature extraction, subset of optimal features is selected using Principal Component Analysis (PCA) algorithm. Moreover, the computational complexity is further being reduced using LSSVM based classifier. Preliminary studies on a endoscopic color image database consisting of 276 images of 6 different types of endoscopic classes show promising retrieval result.

The paper is organized as follows: Section 2 describes the theoretical preliminary of the NSCT. The detail descriptions of the used feature extraction procedure, classification through LSSVM and KLD based similarity measure are explained in Section 3, Section 4 and Section 5, respectively. The summary of the proposed CBMIR system is described in Section 6. Section 7 discusses the experimental system and results. Finally, Section 8 concludes the article.

2. NON-SUBSAMPLED CONTOURLET TRANSFORM (NSCT)

In this section, we briefly describe the non-subsampled contourlet transform, which will be adopted in our system to devise a proper image representation.

NSCT is a fully shiftinvariant, multiscale, and multidirection expansion with fast implementability [8]. As opposed to the Contourlet Transform (CT), which is not shift-invariant due to the presence of down-samplers and up-samplers in both the Laplacian Pyramid (LP) and Directional Filter Bank (DFB) stages, NSCT achieves the shift-invariance property by using non-subsampled LP filter bank and non-subsampled DFB.

2.1 Non-Subsampled Pyramid (NSP)

NSP is a shift-invariant filtering structure that leads to a subband decomposition that resembles the Laplacian pyramid, which ensures the multi-scale property of the NSCT. As shown in Figure 1, it is constructed by using two-channel non-subsampled 2D filter banks, which produce low and high-frequency image at each NSP decomposition level. Filters at subsequent stages are obtained by upsampling the low-pass filters at the first stage. As a result, NSP can result in $\mathsf{k}+1$ sub-images, which consist of one low-frequency image and k high-frequency images whose sizes coincide with the source image, k being the number of decomposition levels. Figure 1(a) gives the NSP decomposition with $\mathsf{k} = 3$ levels.

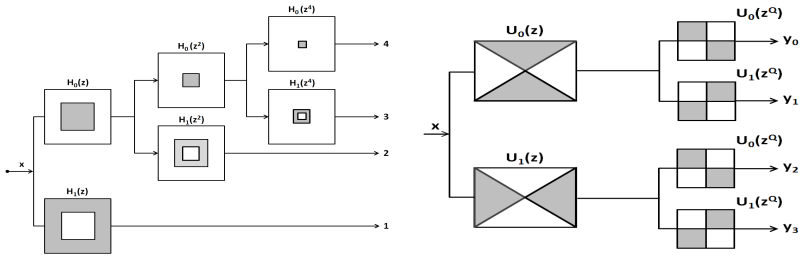
2.2 Non-Subsampled Directional Filter Bank (NSDFB)

The NSDFB is constructed by eliminating the downsamplers and upsamplers of the DFB and by upsampling the filters accordingly [8]. This results in a tree composed of two-channel NSFB, described in Figure 1(b) (4 channel decomposition). At each stage of the NSP, the NSDFB allows a decomposition into any number of 2^l directions, l being the number of levels in the NSDFB. This provides the NSCT with the multi-direction property and offers precise directional information. The combination between NSP and NSDFB is depicted in Figure 2(a). The resulting filtering structure approximates the ideal partition of the frequency plane displayed in Figure 2(b). Differently from the contourlet expansion, the NSCT has a redundancy given by $\mathsf{r} = 1 + \sum_{j=1}^{\mathsf{k}} 2^{\ell_j}$, where ℓ_j is the number of levels in the NSDFB at the jth scale. We refer to [8] for further details about NSCT.

Among the different MGA tools NSCT has better frequency selectivity and regularity, as well as it is a flexible multiscale, multi-directional, and shift-invariant image decomposition method [7]. NSCT coefficients are capable of capturing the fine details present in the image, which is essential in CBIR, as natural images are full of subtle image information. Moreover, it is widely accepted that the human visual system bases its perception on multiple channels that are tuned to different ranges of spatial frequencies and orientations. Measurements of the receptive fields of simple cells in the primary visual cortex revealed that these channels exhibit approximately a dyadic structure [14]. This behavior is well matched by NSCT decomposition. These facts motivate the use of NSCT to construct the image signatures in the proposed CBIR system.

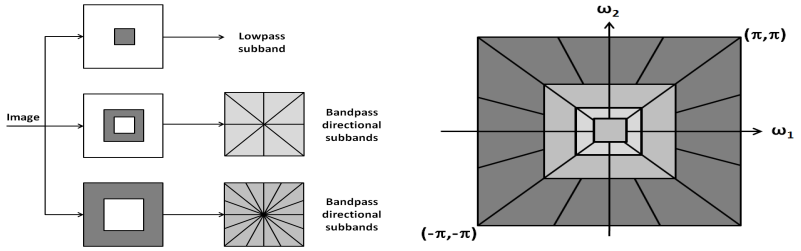
3. NSCT BASED LOCALIZED GGD FEATURES

The images in the database prior to NSCT decomposition, are transformed from RGB to CIELab color space. This ensures that the textural characterizations of the images are independent of the color characterization. NSCT decomposition over the intensity plane (L) characterizes the texture information, while the NSCT decomposition over



(a) (b)

Figure 1: (a) Non-subsampled Pyramid Filter Bank: Three-stage decomposition. (b) Four-channel NSDFB constructed with two-channel fan filter banks.



(a) (b)

Figure 2: Non-subsampled contourlet transform (a) NSFB structure that implements the NSCT. (b) Idealized frequency partitioning obtained with the proposed structure.

chromaticity planes (a and b) characterizes color. Texture and color information are extracted by using NSCT on each color plane with l level of decomposition (here, $l = 3$) which results in 10 high pass and 1 low pass sub-band images. Therefore, this configuration produces sb (= 11) subbands for each color plane. As, there are 3 color planes, so altogether we get S ($= 3 \times sb$) subbands for each image of the database. Each subband contains the transform coefficients of size $n \times n$, where n is the size of the NSCT subband images. Now, for the computation of the local feature vectors, each such subbands are sub-divided into N number of blocks using window size $w \times w$, where $N = \frac{n^2}{w^2}$. The distribution of the coefficients of each N blocks of each subband coefficients are then modeled with GGD which is defined as,

$$p(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta} \quad (1)$$

where x is the NSCT subband coefficients and $\Gamma(.)$ is the gamma function, i.e., $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$, $z > 0$. Here the scale parameter α models the width of the probability distribution function (PDF) peak (standard deviation), while the shape parameter β is inversely proportional to the decreasing rate of the peak. These two parameters need to be estimated for feature vector creation. As Maximum Likelihood (ML) estimator is best suited for estimating heavy-tailed distribution like GGD for both small and large samples, we have used ML estimator in our proposed scheme.

For the sample set $x = (x_1, x_2, x_3, ..., x_k)$, x_i is the NSCT coefficients at the i^{th} subband, and $i \leq L$, the ML estimator is defined as [11].

$$L(x; \alpha, \beta) = log \prod_{i=1}^{L} p(x_i; \alpha, \beta) \quad (2)$$

GGD parameters are defined with the following equations, which have a unique root in probability

$$\frac{\partial L(x; \alpha, \beta)}{\partial \alpha} = -\frac{L}{\alpha} + \sum_{i=1}^{L} \frac{\beta |x_i|^\beta \alpha^{-\beta}}{\alpha} = 0 \quad (3)$$

$$\frac{\partial L(x; \alpha, \beta)}{\partial \beta} = \frac{L}{\beta} + \frac{L\Psi(1/\beta)}{\beta^2} - \sum_{i=1}^{L} \left(\frac{|x_i|}{\alpha}\right)^\beta log\left(\frac{|x_i|}{\alpha}\right) = 0 \quad (4)$$

where $\Psi(.)$ is the digamma function, i.e.,$\Psi(z) = \Gamma'(z)/\Gamma(z)$. α has a unique, real, positive solution and can be obtained from Eq.(3) by fixing $\beta > 0$:

$$\hat{\alpha} = \left(\frac{\beta}{L} \sum_{i=1}^{L} |x_i|^\beta\right)^{1/\beta} \quad (5)$$

Substituting this into (4), the shape parameter β is the solution of the following *transcendental* equation

$$1 + \frac{\Psi(1/\hat{\beta})}{\hat{\beta}} - \frac{\sum_{i=1}^{L} |x_i|^{\hat{\beta}} \log |x_i|}{\sum |x_i|^{\hat{\beta}}} + \frac{log\left(\frac{\hat{\beta}}{L} \sum_{i=1}^{L} |x_i|^{\hat{\beta}}\right)^{1/\hat{\beta}}}{\hat{\beta}} = 0 \quad (6)$$

which can be solved numerically using the Newton-Raphson iterative procedure. Therefore, each block of each NSCT subband (*i.e.*, sb) is represented by two GGD parameters. Considering S number of subbands of each image I and N number of blocks for each subband, we obtain $2 \times S \times \frac{n^2}{w^2}$ dimension of feature vectors, where $n = 256$ and $w = 64$.

Finally, each image in the database is represented by 1056 dimension of localized feature vectors. However, this localized feature vectors actually reduced the discriminatory capacity of the classification accuracy because of the redundancy among different feature vectors. Therefore, we need to apply some dimension-reduction technique to reduce the

the feature vector dimension and improve the retrieval accuracy with minimum computational complexity. In this study, we have used the well-known Principal Component Analysis (PCA) for this purpose. PCA reduces the dimensionality of the search to a basis set of prototypes that best describes the images [26, 2]. This process results in an overall reduction of feature dimension by around 82% (from 1056 to 195). Therefore, each image is represented with 195 compact localized statistical feature vectors.

4. CLASSIFICATION USING LSSVM

Once the NSCT based GGD feature vectors are obtained, LSSVM classifier is used to classify images of the database. LSSVM avoids solving quadratic programming problem and simplifies the training procedure of conventional SVM [28]. Considering a linearly separable binary classification problem:

$$(x_i, y_i)_{i=1}^{n} \quad and \quad y_i = \{+1, -1\} \tag{7}$$

where x_i is an n-dimensional vector and y_i is the label of this vector. LSSVM can be formulated as the optimization problem:

$$\min_{w,b,e} \mathcal{J}(w, b, e) = \frac{1}{2} w'w + \frac{1}{2} C \sum_{i=1}^{n} e_i^2 \tag{8}$$

subject to the equality constraints

$$y_i[w'\varphi(x_i) + b] = 1 - e_i \tag{9}$$

where $C > 0$ is a regularization factor, b is a bias term, w is the weights vector, e_i is the difference between the desired output and the actual output and $\varphi(x_i)$ is a mapping function.

The lagrangian for problem of Eq.(8) is defined as follows:

$$\mathcal{L}(w, e_i, b, \alpha_i) = \min_{w,b,e} \mathcal{J}(w, b, e) - \sum_{i=1}^{n} \alpha_i \{y_i[w'\varphi(x_i) + b] - 1 + e_i\} \tag{10}$$

where α_i are Lagrange multipliers. The Karush-Kuhn-Tucker (KKT) conditions for optimality $\frac{\partial \mathcal{L}}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{n} \alpha_i y_i \varphi(x_i)$; $\frac{\partial \mathcal{L}}{\partial e_i} = 0 \rightarrow \alpha_i = C e_i$; $\frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0$; $\frac{\partial \mathcal{L}}{\partial \alpha_i} = 0 \rightarrow y_i[w'\varphi(x_i) + b] - 1 + e_i = 0$,is the solution to the following linear system

$$\begin{bmatrix} 0 & -Y \\ Y & \varphi\varphi' + C^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \overline{1} \end{bmatrix} \tag{11}$$

where $\varphi = [\varphi(x_1)'y_1, ..., \varphi(x_n)'y_n], Y = [y_1, ..., y_n], \overline{1} = [1, ..., 1]$, and $\alpha = [\alpha_1, ..., \alpha_n]$. For a given kernel function $K(,)$ and a new test sample point x, the LSSVM classifier is given by

$$f(x) = sgn[\sum_{i=1}^{n} \alpha_i y_i K(x, x_i) + b] \tag{12}$$

5. KLD BASED SIMILARITY MEASURE

After the image is classified using LSSVM according to the NSCT based GGD feature vectors, KLD is used for distance calculation between the query image and the database images.

Let $p_1(x)$ and $p_2(x)$ be two continuous probability distribution functions. KLD distance between these two probability distribution function is defined by [11]

$$D(p_1, p_2) = \int p_1(x) log \left(\frac{p_1(x)}{p_2(x)} \right) dx \tag{13}$$

We obtain the KLD between two GGDs substituting the GGD probability distribution function of Eq.(1) into Eq.(13), which is defined as [11]

$$D(p(,;\alpha_1, \beta_1) || p(,;\alpha_2, \beta_2)) = \log \frac{\beta_1 \alpha_2 \Gamma(1/\beta_2)}{\beta_2 \alpha_1 \Gamma(1/\beta_1)} + \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma((\beta_2 + 1)/\beta_1)}{\Gamma(1/\beta_1)} - \frac{1}{\beta_1} \tag{14}$$

NSCT coefficients of each subband are assumed to be independent, thus the computed localized feature vectors are also independent to each subbands. Therefore, KLD between two images is obtained by calculating the sum of all the Kullback-Leibler distances across all blocks of the NSCT subbands of these images. Let, α_j^i and β_j^i be the NSCT GGD features of each block of each subband j of image i. Using the definition of KLD between two GGDs shown in Eq. (14), the KLD between two images P and Q is defined by Eq.(15)

$$D(P, Q) = \sum_{j=1}^{\gamma} D(p(.; \alpha_j^{(P)}, \beta_j^{(P)}) || p(.; \alpha_j^{(Q)}, \beta_j^{(Q)})) = \sum_{j=1}^{\gamma} log \frac{\beta_j^{(P)} \alpha_j^{(Q)} \Gamma(1/\Gamma_j^{(Q)})}{\beta_j^{(Q)} \alpha_j^{(P)} \Gamma(1/\Gamma_j^{(P)})} + \sum_{j=1}^{\gamma} \left(\frac{\alpha_j^{(P)}}{\alpha_j^{(Q)}} \right)^{\beta_j^{(Q)}} \frac{\Gamma((\beta_j^{(Q)} + 1)/\beta_j^{(P)})}{\Gamma(1/\beta_j^{(P)})} - \sum_{j=1}^{\gamma} \frac{1}{\beta_j^{(P)}} \tag{15}$$

where γ is the total number of blocks in NSCT subbands of an image. For a given query image, KLD is measured between the query image and each of the database images using Eq.(15). Database images are ranked according to ascending order of distances and are retrieved from the database.

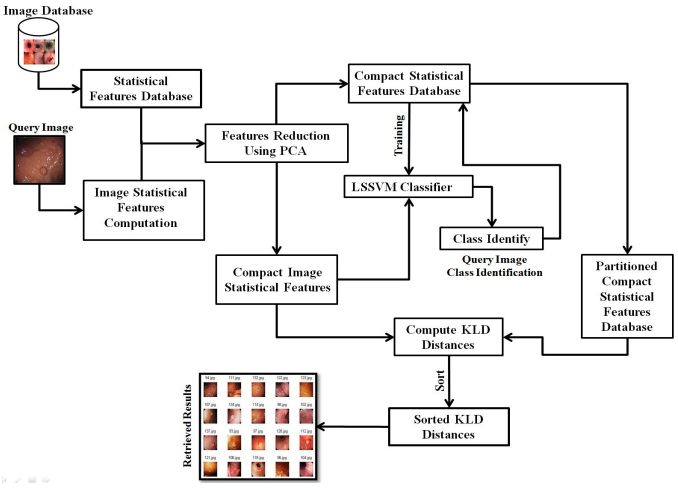


Figure 3: Block Diagram of the Proposed CBMIR System

6. THE SUMMARY OF CBMIR SYSTEM

Here, we outlines the salient steps of the proposed endoscopic based CBMIR system. The Fig. 3, shows the block diagram of the proposed CBMIR system.

Steps:

1. Input the query image.

2. Compute the GGD based compact image statistical features of the query image as described in Section 3.

3. Identify the class of the query image using LSSVM, depending on the GGD based compact statistical features as described in Section 4.

4. As describe in Section 5, after the pre-classification of the query image, find the KLD distances between the compact statistical features of the query image and the compact local statistical features of the image database stored in "Partitioned Compact Statistical Features Database".

5. Sort the distances in ascending order and display 20 images corresponding to the first 20 sorted distances.

6. Stop.

7. EXPERIMENTAL RESULTS

Extensive experiments were carried out to evaluate the performance of the proposed CBMIR system. We conducted several different experiments to evaluate the effectiveness of the proposed system both quantitatively and qualitatively.

7.1 Experimental setup

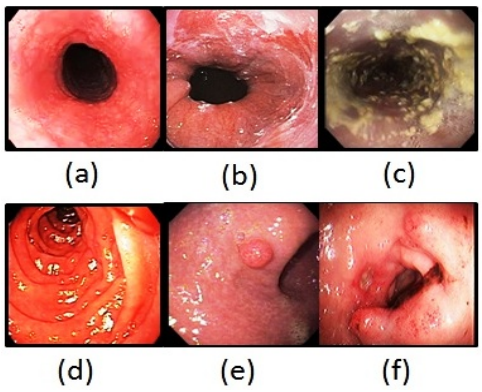


Figure 4: Examples of categories from endoscopy image datasets (a) Normal Esophagus (b) Barrett's Esophagus (c) Esophageal Inflammation (d) Normal Upper Gastrointestinal (e) Upper Gastrointestinal Polyps and (f) Upper Gastrointestinal Ulcers

In our experiments we have used endoscopy image dataset consist of 276 images which grouped into 6 categories (46 images per category). The datasets are heterogeneous as they have different sizes along with 4 different types of abnormality (diseases)and 2 normal types as seen in Fig. 4.

The number of decomposition level of NSCT, considered in this study is 3. The details about NSCT decomposition are discussed in [7, 8]. Moreover, in this experiment, we have considered the local features of the endoscopic images. Local features is important for image retrieval because the similarity between two images perceived by human observers strongly depends on the layout of the image as well as the spatial relations among depicted objects [9].

In LSSVM, we have used the Radial Basis Function (RBF), $K(x_i, x_j) = exp(-\gamma \parallel x_i - x_j \parallel^2), \gamma > 0$, as the kernel for training. There are two tunable parameters while using the RBF kernel in LSSVM classifier: C and γ. The kernel parameter γ controls the shape of the kernel and regularization parameter C controls the tradeoffs between margin maximization and error minimization. Hence, a Cross Validation (CV) is conducted, to choose the best pair of (C, γ) having the lowest CV error rate is picked. After finding the best values for the parameters C and γ, these values are used to train the LSSVM model, and the test set is used to measure the error rate of the classification system. In all the three classifier CV is done on 70% of the train dataset and remnant 30% is used as test dataset for evaluating the performance of the classifiers used in our study. We have achieved the classification accuracy of 71.45% for endoscopic image database.

Quantitative evaluation of the proposed CBMIR system is analyzed using two statistical measures: Average Precision (AP), and Average Recall (AR). We have computed the precisions and recalls considering all the images of the used databases as the query images, and then take the average of the obtained precision and recall values over all the images as the final evaluation result. The two statistical measures are defined as follows:

$$Precision\ (P) = \frac{N_{RIR}}{N_{RIR} + N_{IRIR}} \tag{16}$$

and

$$Recall\ (R) = \frac{N_{RIR}}{T_{RID}} \tag{17}$$

where, N_{RIR} is the number of relevant images retrieved, N_{IRIR} is the number of irrelevant images retrieved and T_{RID} is the total number of relevant images in the database. During the experiments, top 20 retrieved images were used to compute the precision and recall values.

7.2 Feature Evaluation

The graph of the Fig. 5 shows the effect of PCA based NSCT+GGD (*i.e.*, GGD based localized statistical features on NSCT subbands) feature representation schemes verses classification accuracy. From the graph, it can be inferred that with nearly 18% of NSCT+GGD based features have satisfactory classification accuracy as compared to 100% of their respective features. Therefore, we have selected 195 distinguished NSCT+GGD feature vectors for our proposed CBMIR system on endoscopic images.

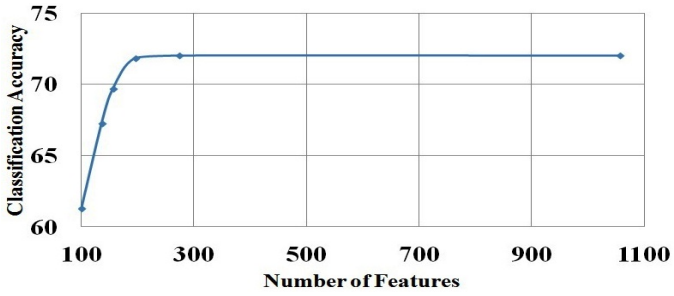


Figure 5: Classification Accuracy Vs Number of Features

7.3 Results and Discussion

The retrieval performance of the proposed CBMIR system (NSCT+GGD+LSSVM) is shown in the Fig.6, in terms

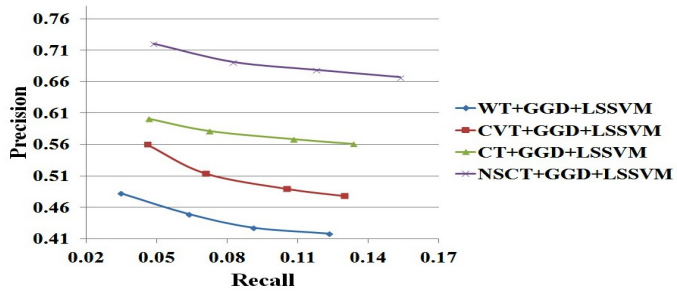


Figure 6: Performance comparison in terms of precision and recall curves

of precision and recall curve. The performance comparisons of WT, CVT and CT with NSCT are also shown in Fig.6 (WT+GGD+LSSVM, CVT+GGD+LSSVM and CT+GGD+LSSVM, respectively). Here, endoscopy image database were considered for performance evaluation. It is clear from the Fig.6, that NSCT performance much better than WT, CVT and CT in proposed endoscopic CBMIR system.

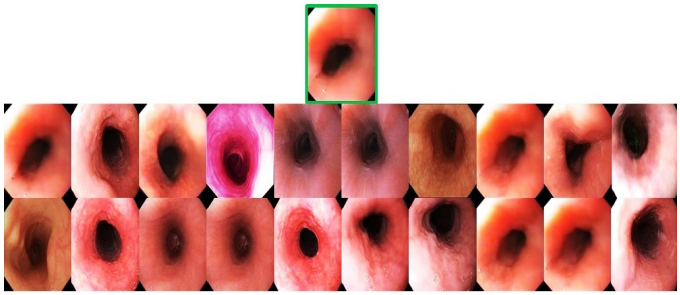


Figure 7: Visual results of the proposed CBIR system for Normal Esophagus (top image is the query image)

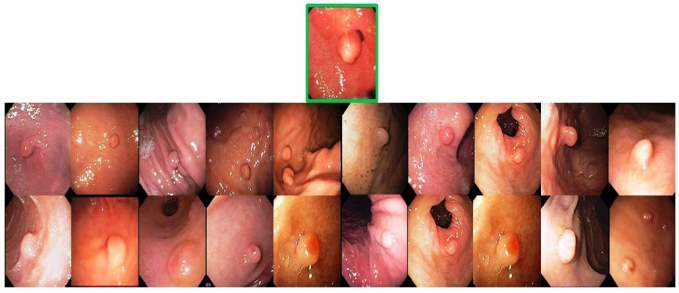


Figure 8: Visual results of the proposed CBIR system for Upper Gastrointestinal Polpys (top image is the query image)

Fig. 7, Fig. 8 and Fig. 9, shows three examples of the visual results obtained by the proposed CBMIR system on endoscopy image database using query images from "*Normal Esophagus*", "*Upper Gastrointestinal Polpys*" and "*Barrett's Esophagus*" classes, respectively. Fig. 8 and Fig. 9 shows the diseased endoscopic images where as Fig. 7 shows the esophagus or gastrointestinal tract of healthy persons. In "*Barrett's Esophagus*", a pink esophageal lining (mucosa) appear that extends a short distance (usually less than 2.5 inches) up the esophagus from the gastroesophageal junction and finding intestinal type cells (goblet cells) on biopsy of the lining whereas in "*Upper Gastrointestinal Polpys*" is a mass of tissue protruding abnormally from the mucous membrane (protective lining of epithelial cells for secretion and absorption) of the esophagus. From the given instances, it can be inferred that the local color, texture and geometrical invariant information can distinguished the normal esophagus or gastrointestinal tract endoscopic images from the abnormal ones. Our proposed NSCT+GGD features can represent those information efficiently and in turn increases the efficiency of the proposed CBMIR system. This fact is corroborated by all the instances of the results obtained from our system where we can clearly see that all the retrieved images are from the respective classes corresponding to the query images.

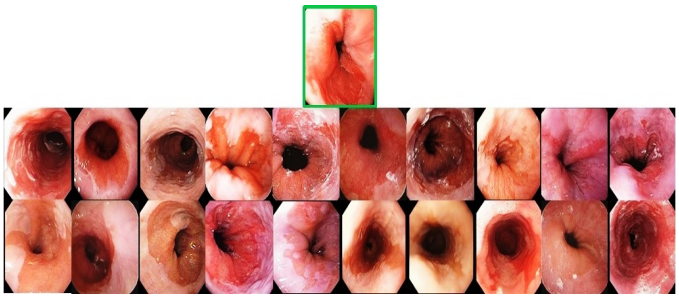


Figure 9: Visual results of the proposed CBIR system for Barrett's Esophagus (top image is the query image)

The proposed algorithm has been implemented using MATLAB R2014a on a Dell Precision T7400 workstation. The total processing time of the query images is computed by considering the feature extraction, classification and retrieval process. Average time taken for extraction of GGD features of each image is 19.770591 seconds where as processing time is 0.862763 seconds, considering 1056 features. However, the processing time of the proposed algorithm is further being reduced by using the PCA algorithm.

8. CONCLUSIONS

From our experiments, we have noticed that statistical modeling of NSCT based image coding is suitable for representing localized low level features of the endoscopic images. The proposed CBMIR system with LSSVM classifier based on compact GGD based localized statistical features on NSCT coefficients is able to achieve the satisfactory results with minimum computational cost. To overcome the problem of misclassification in pre-classification process, we are trying to introduce a neuro-fuzzy ranking approach. As the future scope of research, the proposed NSCT based local statistical features could be tested for other CBMIR system which consist of different medical image modalities.

9. ACKNOWLEDGMENTS

This work is mainly funded by Machine Intelligence Unit, Indian Statistical Institute, Kolkata-108 (Internal Academic Project) for providing facilities to carry out this work. Malay

Kumar Kundu acknowledges the Indian National Academy of Engineering (INAE) for their support through INAE Distinguished Professor fellowship.

10. REFERENCES

[1] C. B. Akgül, D. L. Rubin, S. Napel, C. F. Beaulieu, H. Greenspan, and B. Acar. Content-based image retrieval in radiology: Current status and future directions. *J. of Digital Imaging*, 2(24):208–222, 2011.

[2] U. Avni, H. Greenspan, E. Konen, M. Sharon, and J. Goldberger. X-ray categorization and retrieval on the organ and pathology level using patch-based visual words. *IEEE T Med. Imaging*, 30(3):733–746, 2011.

[3] E. Candes and D. Donoho. Continuous curvelet transform: I. Resolution of the wavefront set. *Appl. and Comput. Harmon. Anal.*, 19(2):162–197, 2005.

[4] M. Chowdhury, S. Das, and M. K. Kundu. Effective classification of radiographic medical images using LS-SVM and NSCT based retrieval system. *Proc. of 5th Int. Conf. on CODEC*, pages 1–4, 2012.

[5] M. Chowdhury, S. Das, and M. K. Kundu. Novel CBIR system based on ripplet transform using interactive neuro-fuzzy technique. *Elect. Letters on Comp. Vision and Image Anal.*, 11(1):1–13, 2012.

[6] M. Chowdhury, S. Das, and M. K. Kundu. A ripplet transform based statistical framework for natural color image retrieval. *Proc. of 17th Int. Conf. on ICIAP*, pages 492–502, 2013.

[7] M. Chowdhury and M. K. Kundu. Comparative assessment of efficiency for content based image retrieval systems using different wavelet features and pre-classifier. *Multimed. Tools Appl.*, 72(3):1–36, 2014.

[8] A. L. da Cunha, J. Zhou, and M. N. Do. The nonsubsampled contourlet transform: Theory, design, and applications. *IEEE T Image Process*, 15(10):3089–3101, 2006.

[9] A. Depeursinge, T. Zrimec, S. Busayarat, and H. Müller. 3D lung image retrieval using localized features. In *Medical Imaging 2011: Computer-Aided Diagnosis*, volume 7963, page 79632E. SPIE, Feb 2011.

[10] T. Deselaers, D. Keysers, and H. Ney. Features for image retrieval: an experimental comparison. *Inform. Retri.*, 11(2):77–107, 2008.

[11] M. N. Do and M. Vetterli. Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE T Image Process.*, 11(2):146–158, 2002.

[12] M. N. Do and M. Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *IEEE T Image Process*, 14(12):2091–2106, 2005.

[13] H. N. Duc, T. L. Tien, T. D. Honl, C. B. Thu, and T. N. Xuan. Image retrieval using contourlet based interest points. *Proc. of 10th ISSPA*, pages 93–96, 2010.

[14] A. M. Garcia-Perez. The perceived image: Efficient modelling of visual inhomogeneity. *Spatial Vision*, 6(2):89–99, 1992.

[15] H. Greenspan and A. T. Pinhas. Medical image categorization and retrieval for PACS using the GMM-KL framework. *IEEE T on Inform. Technol. Biomed.*, 11(2):190–202, March 2007.

[16] W. Hsu, S. Antani, L. R. Long, L. Neve, and G. R. Thoma. SPIRS: a web-based image retrieval system for large biomedical databases. *Int J Med Inform*, 78(1):S13–S24, 2009.

[17] K. H. Hwang, H. Lee, and D. Choi. Medical image retrieval: Past and Present. *Healthc Inform Res.*, 18(1):3–9, 2012.

[18] D. K. Iakovidis, D. E. Maroulis, and S. A. Karkanis. An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. *Comput Biol Med*, 36(10):1084–1103, 2006.

[19] D. H. Johnson and S. Sinanovic. Symmetrizing the kullback-leibler distance. *IEEE T Inf. Theory*, 1(1):1–10, 2001.

[20] S. Karkanis. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE T. Inform Tech Biomed*, 7(3):141–152, 2003.

[21] S. A. Karkanis, D. K. Iakovidis, D. A. Karras, and D. E. Maroulis. Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures. In *Proc of ICIP*, pages 833–836. IEEE, Oct 2001.

[22] T. M. Lehmann and et al. The IRMA reference database and its use for content-based image retrieval in medical applications. *Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie*, 2004.

[23] M. Liedlgruber and A. Uhla. A summary of research targeted at computer-aided decision support in endoscopy of the gastrointestinal tract. *Technical Report, University of Salzburg*, 2011.

[24] H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. *Int. J. Medical Inf.*, 73(1):1–23, 2004.

[25] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, and C. Roux. Wavelet optimization for content-based image retrieval in medical databases. *Medical Image Analysis*, 14(2):227–241, 2010.

[26] M. Rahman, P. Bhattacharya, and B. C. Desai. A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. *IEEE T Inf. Technol. Biomed.*, 11(1):58–69, 2007.

[27] I. J. Sumana, M. M. Islam, D. Zhang, and G. Lu. Content based image retrieval using curvelet transform. *Proc of 10th Workshop on MMSP*, pages 11–16, 2008.

[28] J. A. K. Suykens and J. Vandewalle. Least squares support vector machine classifiers. *Neural Process. Lett.*, 9(3):293–300, 1999.

[29] H. D. Tagare, C. C. Jaffe, and J. Duncan. Medical image databases: a content-based retrieval approach. *J. Am. Med. Inform. Assoc.*, 4(3):184–198, 1997.

[30] T. Tommasi, F. Orabona, and B. Caputo. Discriminative cue integration for medical image annotation. *Pattern Recogn. Lett.*, 29(15):1996–2002, 2008.