# SEGMENTATION OF REGIONS IN JPEG COMPRESSED MEDICAL IMAGES

*Pramod K. Singh*

School of Computer Science and Engineering
University of New South Wales, Sydney, NSW - 2052, Australia

**ABSTRACT**

A novel algorithm for the segmentation of medical images using features derived directly from JPEG compressed domain is proposed in this paper. The algorithm uses features extracted from DCT coefficients without its inverse transform and the Rule based Fisher Discriminant K-means (FDK) technique for clustering image pixels based on derived feature vectors. In this study, we extract features for each 2x2 DCT block of compressed image. The extracted feature vector is used by an extended version of the adaptive K-means clustering algorithm for the classification of image pixels.

## 1. INTRODUCTION

In modern digital systems images are stored in compressed format. Among the compression methods, Discrete Cosine Transform (DCT) based JPEG is the most popular technique. DICOM (Digital Imaging and COmmunication in Medicine) standard also permits the lossy compression of some medical diagnostic images via a JPEG baseline system. Image segmentation partitions an image into units, which are homogeneous with respect to one or more characteristics. It is performed on the images either before they are compressed or after decompression. Extraction of features for segmentation directly from compressed image has obvious advantage of eliminating some or all of the processing related to decompression.

DCT is one of the best filters for extraction of features in the spatial frequency domain. Several researchers have explored the use of DCT for image segmentation in the pixel domain [3][5][6], whereas, segmentation in compressed domain is not enough investigated, however, Feng et.al [4], recently, reported a segmentation algorithm in JPEG compressed domain. Their algorithm is based on a region-growing technique, which is a promising technique if the homogeneous region is spatially unscattered but it is unable to provide satisfactory results in those applications, for example, medical diagnostic images, where regions of similar pattern are scattered at different places in an image. The DCT coefficients provide some gray scale and textural information but more can be derived from their statistical derivatives. Unlike the Feng et.al our alogorithm uses statistical derivatives of DCT coefficients as feature vector for the better segmentation.

In the statistical segmentation algorithms, feature vectors are submitted to a classifier for the classification of image pixels. K-means is an efficient and robust algorithm for unsupervised classification. Several versions of the K-means algorithm [1][6] have been proposed for processing the feature vectors of images. In this study the Fisher Discriminant K-means (FDK) clustering technique is used for clustering spatial pixels based on feature vectors extracted directly from DCT domain.

The rest of the paper is organized as follows: in Section 2, feature computation in DCT domain is described. The segmentation using FDK clustering is described in Section 3. Results of the segmentation of the test images are presented in Section 4. The paper is concluded in Section 5.

## 2. FEATURE COMPUTATION IN DCT DOMAIN

In most of the segmentation algorithms the DCT coefficients and their derivatives are calculated on pixel by pixel basis. In [3] feature vectors were extracted using 2D-DCT of 4x4 pixel blocks for segmentation of ultrasound images. Ng et al. [5] report on using the local variance of DCT coefficients to segment images. Their method involves generating a 3x3 DCT at each pixel location, using the surrounding points. The local variance of each DCT coefficient is then computed using 15x15 sliding window. Jie Wei [6] used a DCT descriptor along with positional coordinates for each pixel of the image in his segmentation algorithm. Our features are based on the local mean and standard deviation (SD) of each DCT coefficients of 2x2 block using 12x12 sliding window.

In this section, we explain the methods used for the extraction of features in DCT domain. First we transformed 8x8 JPEG baseline compressed image into 2x2 block JPEG image and a local mean and SD value of coefficients for each 2x2 block is computed using 12x12 window centred at

Author acknowledges the financial support in part by the Australian Research Council through a Linkage grant (No LP0212081), with Medical Imaging Australasia as clinical and Philips Medical Systems as Industrial partners. Special thanks to Prof Paul Compton for his valuable suggestions.

each 2x2 block in the transformed JPEG image.

### 2.1. The Compressed Domain

The images that are fully compressed are passed through a set of transformation steps which together define the standard. The JPEG standard consists of following six steps: (1) shifting of pixel range (2) DCT transformation on 8x8 pixel blocks (3) quantization (4) Zig-Zag ordering (5) run-length coding (6) entropy coding. Our feature extraction algorithm operates on JPEG baseline system data, which is the data after step (3). Each coefficient of an 8x8 block, in JPEG baseline compressed image, determines the bit allocation in the spatial frequency domain.

### 2.2. Computation of the DCT Coefficients of a 2x2 Block

In our approach we derive discriminatory information about the gray scale and textural distribution of spatial frequency components by calculating local energy coefficients of the DCT of a sub-image. To derive the feature set we first transform the JPEG baseline compressed image into 2x2 block compressed image. We use the algorithm of block transformation developed by Feng et.al [4]. With this algorithm a given block of NxN DCT coefficients could be split directly into four smaller sub-blocks of N/2xN/2 data without involving any inverse DCT transform. The DCT sub-blocks transformation method can be described as follows:

$$\begin{pmatrix} C_s^{0,0} & C_s^{0,1} \\ C_s^{1,0} & C_s^{1,1} \end{pmatrix} = \mathbf{2DC_bD^T} \tag{1}$$

where $C_b$ represents a matrix of DCT coefficients of NxN pixels, and it is split into four sub-blocks represented by four matrices, i.e. $C_s^{i,j}$ $(i, j \in [0, 1])$; D is an orthogonal matrix with NxN elements. The matrix D for a block 8x8 being decomposed to 4x4, and for 4x4 decomposed to blocks of 2x2 are given as Table I and II respectively. Considering the sparse nature of matrix D and non significant value of elements of $C_b$ at the lower right corner of N/2xN/2, we can deduce that $C_b^{i,j} \approx 0$ for $i \neq 0$ or $j \neq 0$, if we divide $C_b$ and D into 4 sub-squared matrices and represent these sub-squared matrices as $C_b^{i,j}$ and $D^{i,j}$ (i,j = 0,1). Thus we can derive:

$$C_s^{0,0} = 2D^{0,0}C_b^{0,0}{D^{0,0}}^T \tag{2}$$

$$C_s^{0,1} = 2D^{0,0}C_b^{0,0}{D^{1,0}}^T \tag{3}$$

$$C_s^{1,0} = 2D^{1,0}C_b^{0,0}{D^{0,0}}^T \tag{4}$$

$$C_s^{1,1} = 2D^{1,0}C_b^{0,0}{D^{1,0}}^T \tag{5}$$

Now we can compute the coefficients of 2x2 blocks from 8x8 blocks of JPEG baseline image using the value of D given in Table I and II in the above equations.

Table I: Matrix D for transferring 8x8 to 4x4

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.5 | 0.4531 | 0 | -0.1591 | 1 | 0.1063 | 0 | -0.0901 |
| 0 | 0.2079 | 0.5 | 0.3955 | 0 | -0.1762 | 0 | 0.1389 |
| 0 | -0.0373 | 0 | 0.2566 | 0.5 | 0.3841 | 0 | -0.1877 |
| 0 | 0.0114 | 0 | -0.0488 | 0 | 0.2452 | 0.5 | 0.4329 |
| 0.5 | -0.4531 | 0 | 0.1591 | 0 | -0.1063 | 0 | 0.0901 |
| 0 | 0.2079 | -0.5 | 0.3955 | 0 | -0.1762 | 0 | 0.1389 |
| 0 | 0.0373 | 0 | -0.2566 | 0.5 | -0.3841 | 0 | 0.1877 |
| 0 | 0.0114 | 0 | -0.0488 | 0 | 0.2452 | -0.5 | 0.4329 |

Table II: Matrix D for transferring 4x4 to 2x2

| | | | |
|---|---|---|---|
| 0.5 | 0.4619 | 0 | -0.1913 |
| 0 | 0.1913 | 0.5 | 0.4619 |
| 0.5 | -0.4619 | 0 | 0.1913 |
| 0 | 0.1913 | -0.5 | 0.4619 |

### 2.3. The Feature Vector Computation

With a view to obtain a feature set containing discriminative information of textures we arranged the DCT coefficients in a one dimensional array, indexed from 0 to 3, which represent the 2x2 block, scanned row wise. Let $C_i(u, v)$ be the DCT coefficient i at (u,v) in the frequency domain, where i = 0 to 3. So the local mean in the window of size NxN is:

$$M_i = \frac{1}{N^2} \sum_{u=0}^{N} \sum_{v=0}^{N} C_i(u, v) \tag{6}$$

local SD of each coefficient $SD_i$,

$$SD_i(u, v) = \sqrt{\frac{\sum_{u=0}^{N} \sum_{v=0}^{N} (C_i(u, v) - M_i)^2}{N^2}} \tag{7}$$

In our experiment N=12 provides the best results. The mean and SD in the local coefficient of NxN window is computed for each 2x2 block of the compressed image. Thus we get a feature vector of 8 features for each 2x2 block.

## 3. IMAGE SEGMENTATION VIA FISHER DISCRIMINANT K-MEANS (FDK) CLUSTERING

In order to create segments of separable gray scale and texture regions the extracted features are clustered. It is assumed that the features obtained at different points in the image are not identical and they can form a cluster in the multi-dimensional feature space if two 2x2 DCT blocks are of coherent gray scale and texture. The adaptive K-means clustering algorithm is an efficient method for clustering data which has a hyper-spheroidal distribution and non overlapping clusters. This algorithm, however, has limitations with those applications where clusters are non hyper-spherical and overlapping. To overcome the limitations of adaptive K-means we use the rule based Fisher Discriminant K-means (FDK) algorithm, which is an extended version of the recently proposed binary hierarchical K-means Iterative Fisher (KIF) algorithm by Clausi [1]. Clausi's algorithm is able to derive separable clusters in non hyper-spherical data but fails to provide clear distinguishable classes in overlapping clusters data set.

Similar to Clausi we assume following constraints: (1) no a priori knowledge of cluster distribution, (2) no a priori knowledge of number of samples belong to each cluster, (3) no a priori knowledge of actual number of clusters in the sample data but the maximum possible clusters $K_{max}$ in the domain data is known. For instance, in our experiment we assume that the maximum possible distinct patterns $K_{max} = 5$ can be in a High Resolution Computer Tomography (HRCT) image of lung.

In FDK, we use fast adaptive K-means clustering [2] to generate $K_{max}$ clusters followed by Fisher Linear Discriminant (FLD) to improve the classification by obtaining appropriate separable clusters K ($<= K_{max}$) using the estimated cluster covariances. The adaptive K-means clustering generates clusters using the minimum Euclidean distance function, which assumes the distribution of samples is hyperspherical and non-overlapping, but this is rarely satisfied assumption in real data sets. The approach of Fisher Linear Discriminant (FLD) can generate more appropriately separated classes using the estimated covariances of each class. The FLD measures the separation between two projected n-dimensional cluster classes on one dimensional vectors. This provides the optimal separation of two clusters [1]. The following steps are used to determine actual clusters:

*Step 1:* Initially entire set of feature vectors are divided into $K_{max}$ clusters using adaptive K-means algorithm which provides centroid of each of these clusters.

*Step 2:* In order to produce the Fisher linear discriminant the class parameters (mean and covariances) are derived from the clusters obtained from step 1. A Fisher distance matrix (FDM) is generated using the Fisher linear inter-cluster distance. The FDM is used for determining the actual number of clusters presented in the data set.

*Step 3:* A set of rules are applied on a FDM table for determination of actual separable clusters. The cluster determination rules are described as follows:

Here, a dangling cluster(DC) concept is introduced, which is an overlapping region between two non-clearly separable clusters. For instance, region C in Fig.1 can be considered a dangling cluster. The dangling clusters are detected by applying following rule (1) on FDM. Consider that there are three clusters generated, using adaptive K-means method, in the feature vector set, represented in two dimensional space in Fig.1, and the FDM is produced using Fisher distance between these clusters. Suppose the explicitly set threshold of Fisher distance is $\bar{\delta}$ for measuring the appropriate separation of clusters. Let $\delta_1$ be the Fisher distance between clusters A and C and similarly $\delta_2$ and $\delta_3$ are between B and C and between A and B respectively. The values of $\delta_1$, $\delta_2$ and $\delta_3$ can be obtained from the FDM. The actual separable clusters determination rules are:

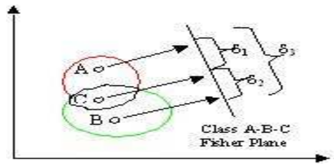



**Fig. 1**. Splitting of clusters into three classes using adpative K-means algorithm

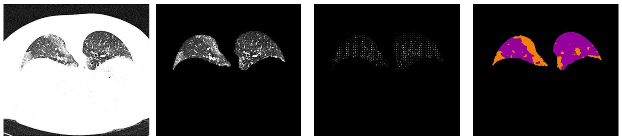
- Rule (1): If $\delta_1 < \bar{\delta}$ and $\delta_2 < \bar{\delta}$ and $\delta_3 \geq \bar{\delta}$ then cluster C is dangling cluster.
- Rule (2): If cluster C is dangling cluster and $\delta_1 < \delta_2$ then C is merged with A, else C is merged with B.
- Rule (3): If $\delta_1 < \bar{\delta}$ and $\delta_2 < \bar{\delta}$ and $\delta_3 < \bar{\delta}$ then clusters A,B and C are part of single cluster and merged.

*SUMMARY*: First 2x2 block DCT is generated from JPEG baseline image, then feature vector is extracted using local statistical measures on 2x2 block of compressed image. Initial $K_{max}$ clusters are created using the adaptive K-means clustering algorithm and then FDM is generated. The cluster determination rules are applied on FDM to find the actual separable segments in the image. The determined segments are labelled with distinct colours. We use labelling algorithm to label each 2x2 block falling in one class with single colour even if they are in spatially isolated regions.

## 4. SEGMENTATION OF IMAGES AND RESULTS

We apply the proposed segmentation algorithm as described in Section 3, on JPEG baseline compressed High Resolution Computer Tomography (HRCT) gray scale lung images of 512 by 512. In our experiment the compressed images contains only parenchyma regions of the lung. In other words, before compression, the HRCT lung image is processed and the background, which is a non significant part for diagnostic use, is removed from the image. The preprocessed image is compressed using a JPEG baseline system. In this study we used 56 labelled HRCT lung images collected from 24 patients. Three radiologists have labelled normal and abnormal tissue patterns in these images. The abnormal patterns are Emphysema, Ground Glass Opacity, Honey Combing, and Consolidation. Most of the images contain one normal and one abnormal pattern. Few images contain more than one abnormal patterns.

To illustrate the processing we use image of Fig.2(a) as an example. This image contains ground glass opacity pattern labelled by a radiologist. Fig.2(b) is the preprocessed image and Fig.2(c) is a JPEG baseline compressed form of this. Next we obtained a feature vector of 8 features for each 2x2 blocks of the compressed image. These feature vectors



(a) (b) (c) (d)

**Fig. 2**. (a) Original Image of Ground Glass Opacities pattern (b) Preprocessed image (c) JPEG Baseline Compressed Image (d) Segmented compressed Image

are used by adaptive K-means with a setting $K_{max} = 5$. Further, we calculate FDM table, as shown in Table III. Finally, actual clusters are deduced using cluster determination rules. In the given example we obtained only two clusters, shown in Fig.2(d), clusters $C_1$ and $C_3$ are merged together to be one cluster and similarly $C_2$, $C_4$ and $C_5$ are combined as second cluster using the rules noted above.

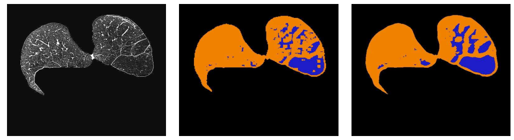
The results of segmentation on JPEG compressed images were compared with the segments obtained using the same features obtained from the DCT of 2x2 block of original image for each pixel and same segmentation method. The results of the proposed algorithm seem quite competitive to pixel based segmentation. Some results of the proposed segmentation method are given in Fig.3. Although we have only presented some images, the segmentation of all the 56 images are of similar quality. Note that there is no ground truth images so the segmentation results are verified by a radiologist.

Table III: Fisher distance matrix

| | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ |
|---|---|---|---|---|---|
| $C_1$ | 0 | 21.33 | 4.78 | 33.62 | 27.85 |
| $C_2$ | - | 0 | 7.93 | 4.16 | 6.78 |
| $C_3$ | - | - | 0 | 18.97 | 16.26 |
| $C_4$ | - | - | - | 0 | 2.62 |
| $C_5$ | - | - | - | - | 0 |

## 5. CONCLUSION

In this paper, we proposed a method for segmentation of JPEG baseline compressed medical images. In this work, a feature vector based on local statistical measures of the DCT coefficients for each 2x2 block of the JPEG compressed image is computed, which appears useful in capturing the gray scale as well as texture characteristics of the image for effective segmentation. In this algorithm, we directly use the DCT coefficients of JPEG compressed images for feature extraction, which provides a computationally competitive advantage over the pixel based segmentation algorithm. Secondly, segmentation solution using adaptive K-means along with the rule based FDM technique appears to be an appropriate for non hyper-spherical and overlapping data sets. In this algorithm we need to set the threshold parameter $\bar{\delta}$ which is based on the quality of the feature vectors. Based on our experiments we found a common value of $\bar{\delta} = 11$ was appropriate for segmentation of HRCT lung images.



(a) (b) (c)

(d) (e) (f)

**Fig. 3**. Segmentation Results (a) Image of Emphysema pattern (d) Image Containing Emphysema and Honey Combing patterns (b)&(e) Segmented JPEG Compressed Image (c)&(f) Segmented Original Image

## 6. REFERENCES


[1] D.A. Clausi, "K-means Iterative Fisher (KIF) Unsupervised Clustering Algorithm Applied to Image Texture Segmentation", *Pattern Recognition*, vol.35, 2002, pp.1959-1972.

[2] C.J. Darken, and J. Moody, "Fast Adaptive K-means Clustering: Some Empirical Results" *Proc. Intl. Conf. on Neural Networks*, vol.II, 1990, pp. 233-238.

[3] Z. Dokur, and T. Olmez, "Segmentation of Ultrasound Images by Using a Hybrid Neural Network", *Pattern Recognition Letters*, vol.23, 2002, pp.1825-1836.

[4] G. Feng, and J. Jiang, "Image Segmentation in Compressed Domain", *Journal of Electronic Imaging*, vol.12, no.3, 2003, pp.390-397.

[5] I. Ng, T. Tan, and J. Kittler, "On Local Linear Transform and Gabor Filter Representation of Texture", *Proc. 11th Intl. Conf. Pattern Recognition*, vol.III, 1992, pp.627-631.

[6] J. Wei, "Image Segmentation Based on Situational DCT Descriptors", *Pattern Recognition Letters*, vol.23, 2002, pp.295-302.