A Survey on Medical Image Analysis using Image Descriptors Methods

Jyoti Patil
Research scholar,
Department of CSE
K.L.University, A.P.

G.Pradeepini, PhD
Associate Professor,
Department of CSE
K.L.University, A.P.

Sunayana Mane
Assistant Professor,
Department of IT
Sinhgad College of Engg, Pune

ABSTRACT
With the tremendous growth in the use of biomedical literature, practitioners have to analyze day to day generated 3D medical images of sonography, MRI scan, CT scan. The idea behind this survey is to study different 3D image concepts and processing methods which will help to create model of 3D image analysis, retrieval in the field of medical domain. A new technique like Hadoop'sMapReduce is used to achieve parallelism in Image processing methods which leads to quick and accurate diagnosis. This paper focuses on study of different 3D image processing methods which can be parallelized using Hadoop framework. This will help readers in choosing appropriate method or methods in their development or research activity.

General Terms
3D images, Content Based Medical Image Retrieval, Hadoop, MapReduce.

Keywords
(Scale Invariant Feature Transform (SIFT), Speeded Up Robust Feature (SURF), Bag-of-visual words (BoVW), Color and Edge Directivity Descriptor (CEDD), Fuzzy color Histogram.

1. INTRODUCTION
This is the era of Big Data. Due to enhanced 3D camera capturing techniques and hardware, lots of 3D data get generated per second. Especially in the medical field for analyzing 'similar type of disease' or 'disease relevant for different diagnosis' needs real time decision making in Operation Theater or in diagnosis. For quick decision making the patient's image should be automatically analyzed by comparing it with stored and indexed database. 3D clinical image data is typically characterized by diseases that affect local structures in the image data as opposed to the entire body. Analysis or retrieval is based on local regions of interest marked by user. Across the human body there is high variability so for capturing those visual differences with each image we need to study important features and descriptors associated with it.

This paper describes different local features and global descriptors of image processing. For low level local feature processing, the methods like Scale Invariant Feature transform (SIFT) [6], Speeded Up Robust Feature (SURF) [2] are used. For mid- level local feature and global descriptors like Bag of Visual Words (BoVW), Vector of Locally Aggregated Descriptors (VLAD) [6] are used. Recently, Riesz miniature transformation, Color and Edge Directivity Descriptor (CEDD), fuzzy color and grey level histograms, Local Binary Pattern (LBP) [7] and Texture bags methods are collectivelyused for image processing. Further sections from this paper will give more insights on each of the techniques.

1.1 3D Clinical Images
3-D (three-dimensional) Image describes an image that provides the perception of depth. In 3D camera two still or motion lenses are used which are slight apart, to take photograph of any three-dimensional object. The process effectively duplicates the stereoscopic vision of human eye.

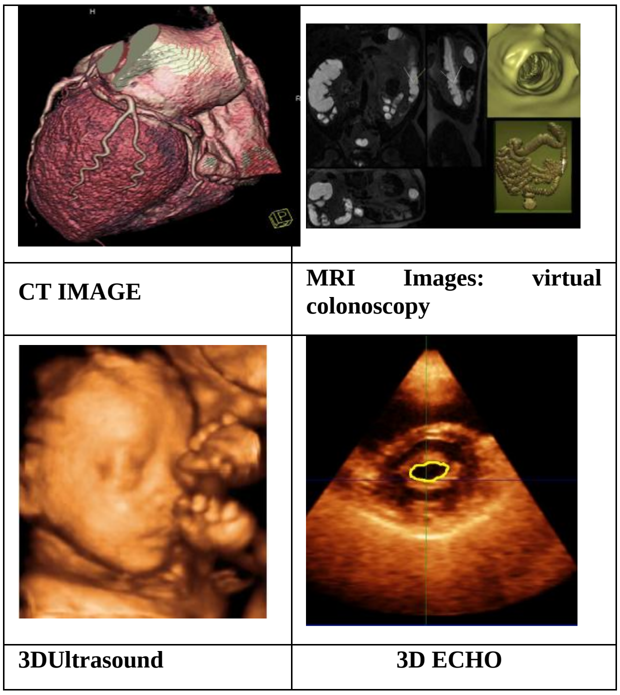

Figure 1. 3D Clinical Image Examples

The image obtained by merging two flat images that viewer's eye see separately; creates a visual illusion of depth as their brains relates and map the images into asingle one. The point at which the left and right images overlap is called as the point of convergence. This point describes the subject of the image as it is the clearest part of the image. Objects at the point of convergence appear on the surface screen. As viewer's go either closer or further away from the point of convergence the objects in 3D imaging appears either closer or further away from the viewer, creating the illusion of depth.

3D clinical images collected using the techniques like 3D Computerized Axial Tomography (CAT) or CT, 3D Magnetic Resonance Imaging (MRI), 3D ultrasound. Using techniques like Tactile imaging (Tactile imaging is a medical imaging modality that translates touch sensing into a digital image. The tactile image is a function of P(x, y, z), where P is measurement of the pressure on soft tissue surfaces under some applied deformation and x, y, z are coordinates where pressure P was measured) in Echocardiography 2D, 3D and Doppler images we can create pictures of the heart and plot the blood flowing through each of the four heart valves. Figure 1 illustrates some examples of 3D clinical Images.

1.2 How to analyze 3D clinical images

Clinical Image Datasets are created using computer database application. They are indexed according to particular disease in the database. Then practitioner gives input image like 3D ultrasound image as a query and this image is matched with the existing database. System returns the result in the form of most appropriate matched images. Based on images retrieved from database and their disease type, input image's disease can be diagnosed. For matching, analysing and retrieving images various image processing techniques are used. Let us discussed them one by one.

2. 3D IMAGE PROCESSING METHODS

We are focusing on methods based on local features and global features descriptors of an image. Local features are nothing but position, scale, orientation and Local Image Structure in Canonical Coordinates. While in global feature each image is represented by a single feature vector, capturing information from the whole image. The constituents of the image, such as individual regions or objects are remaining unattended.

2.1 LocalFeatures

In the processing of each local feature descriptor first interest points are selected then region points around each local descriptor are considered and then local descriptors from the region is computed and normalized. Finally local descriptors are matched. Local descriptors are needed as they detect the same points independently in each image. These are Invariant to translation, rotation, scale and affine transformation, presence of noise, blur etc. The selected region points should contain "interesting" structure and there should be enough points to represent the image. They must also be time efficient.

2.1.1 Scale Invariant Feature Transform (SIFT)

(SIFT)[5]is an image descriptor for image-based matching and recognition developed by David Lowe (1999, 2004). These descriptors as well as related image descriptors are used for a large number of applications in computerized analysis related to point matching between different views of a 3-D scene and view-based object recognition. In SIFT [5]; descriptor interest points are detected from a grey level image at which statistics of local gradient directions of all image intensities were accumulated. This method gives a summarized description of the local image structure in local neighborhood around each interest point. Finally this descriptor is used for matching corresponding interest points between different images. A Gaussian pyramid is constructed from the input image by repeatedly using smoothing and subsampling, and a difference-of-Gaussians pyramid is obtained from the differences between the adjacent levels in the Gaussian pyramid. Interest points are obtained from the points at which the difference-of-Gaussians values which assume extreme based on two values the spatial coordinates in the image domain and the scale level in the pyramid. The result of a Difference of Gaussians (DoG) is applied in scale-space to a series of smoothed and re-sampled images to detect the minima and maxima which are used in first part. Then this calculated Difference of Gaussians operator can be seen as an approximation of the Laplacian operator. These minima and maxima are candidates as interest points (key points). Edge response time and Low contrast candidate points along an edge are discarded for robustness to noise. After that the dominant orientations get assigned to localized key points which provide rotation–in-variance.

In the second part, a key point descriptor is formed by first computing the gradient magnitude

$$DOG(x, y^{\cdot}, s + \Delta s) - L(x, y^{\cdot}, s) \approx \frac{\Delta s}{2} \nabla^2 L(x, y^{\cdot}, s) \qquad (1)$$

This method for detecting interest points leads to scale-in-variance in the sense that first the interest points are preserved under scaling transformations and second the selected scale levels are transformed in accordance with the amount of scaling. Then these scale values are obtained from interest points which are used for normalizing local neighborhoods with respect to scaling variations and it is essential for the scale-invariant properties of the SIFT[5] descriptor.

STEPWISE SIFT Process:

Step 1: Detect Scale-space extrema – Detect interesting points (invariant to scale and orientation) using DOG.(Refer Figure 2)

Image DOG extrema

Figure 2. SIFT STEP 1

Step 2: Localize Key points – Determine the location and scale at each Candidate location and then select them based on their stability.(Reject the low contrast points and the points that lie on the edge) as shown in figure 3.

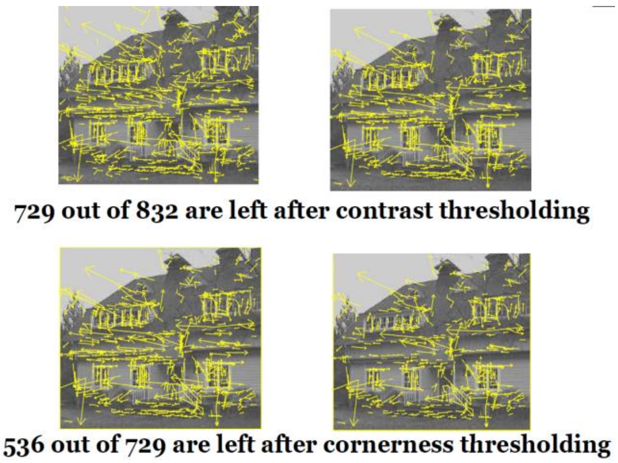

Figure 3. SIFT STEP 2

Step 3: Estimate Orientation – local image gradients are used in orientation assignment for each localized key point. Values of theta, scale and location for each feature are gettingpreserved. (Refer Figure 4)

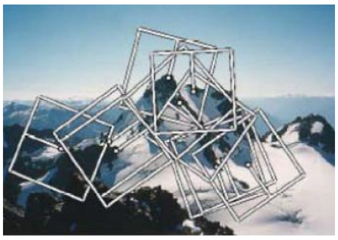

Figure 4. SIFT STEP 3

Step 4: Key point Descriptor - local image gradients at selected, scale around key point are extracted and then form a representation invariant to local shape distortion and illumination. (Refer Figure 5)

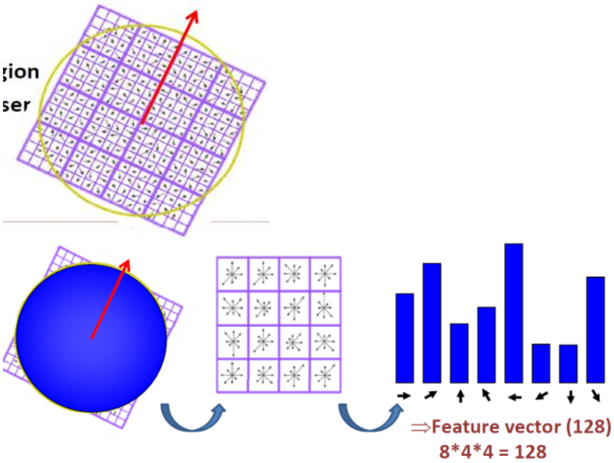

Figure 5. SIFT STEP 4

Create 8 bin gradient histograms for each sub-region weighted by magnitude and Gaussian window (σ is half the window size) finally, 128 dim vector is normalized to make it illumination invariant.

2.1.2 Speeded Up Robust Feature (SURF)

This local feature is similar to SIFT [6], only the main difference found in the way of detecting interest points. SURF[2] is used to create a stack without down sampling higher levels in the pyramid, which results in images of the same resolution. Because of usage of integral images, SURF [2] needs to filter the stack using a box–filter approximation of second-order Gaussian partial derivatives. Figure 6 shows box filters or images. For any given point x = (x, y) in an given image I, the Hessian matrix $H(X, \sigma)$ σ) in x at scale σ is defined as follows

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (2)$$

Here Lxx(x, σ) is the convolution of the Gaussian second order derivative with the image I in point x, and similarly for $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$.

The integral image IΣ(x, y) of an image I(x, y) represents the sum of all pixels in image I(x,y) of a rectangular region which is formed by (0,0) and (x,y).Using integral images, it takes only four array references to calculate the sum of pixels over a rectangular area of any size.

- Approximate L_{xx}, L_{yy}, and L_{xy} using box filters.
- Box filters shown are 9 x 9 matrix which is good approximations for a Gaussian with σ=1.2)

Due to using integral images, filters of different size can be applied at the same speed and instead of using a different measure for selecting scale of interest points and location (e.g., Hessian and DOG in SIFT[5], SURF[2] uses the determinant of H^{SURF}_{approx} to find both. To obtain a good approximation determinant elements must be weighted.

$$det(H^{SURF}_{approx}) = \hat{L}_{xx}\,\hat{L}_{yy} - (0.9\hat{L}_{xy})^2 \qquad (3)$$

Once interest points have been localized both in space and scale, the next steps are:

(1)Find feature direction i.e. Orientation assignment

(2) Calculate Keypoint descriptor which is same as SIFT [5].

STEPWISE SURF Process

STEP 1: Find image interest points -Use determinant of Hessian matrix.

STEP 2: Find major interest points in scale space -Non-maximal suppression on scaled interest point maps

STEP3. Find feature “direction” - compute rotationally invariant features.

STEP4. Generate feature vectors.

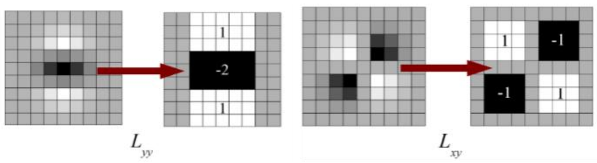

Figure 6. SURF Box Filters

2.2 Mid-Level Features And Global Descriptors

A mid-level descriptor is a numeric or symbolic, continuous or discrete measurement calculated after a global (i.e. collection-wise) analysis of low-level descriptors. While local features perform well in object recognition, image classification and CBIR, they are inefficient for large scale tasks. So for these reasons statistical image representations have been used, which is also called mid–level features, with Bag–of–Visual–Words (BoVW) being the most commonly used.

2.2.1 Bag–of–Visual–Words (BoVW)

The Bag of Visual Words approach provides a good general overview, and has shown promising results in medical image retrieval tasks. The main trend is the shift from purely Difference of Gaussians based interest points to dense grids of interest points - thus effectively eliminating the influence of any interest point detector, altogether. The stability of the interest point detector is crucial for obtaining discriminative descriptor prototypes in the clustering stage of the bags of visual words approach. While the most popular descriptors (e.g., SIFT[5], or derivatives such as SURF[2] are insensitive to small perturbations in the interest points, fine-grained, repetitive structures like those encountered in microscopic images or lung CTs prove difficult for repeatable interest point identification.

One of the most common approaches for image description using local features in large datasets is the BoVW representation. Selected training set of images is chosen and local descriptors are extracted from interest points of each image of this set. Then descriptors are clustered using a clustering method into k clusters and the centroids of the clusters are used as visual words. V represents visual vocabulary of all cluster centers.

$\{v_1, \ldots, |v_k\}, v_i \in R, i = 1, \ldots\ldots, k$(4)

Then, local features are extracted from stored images in the database and mapped with the cluster centers to create for each image a histogram of visual words. Images are indexed in terms of histograms of the visual words (bag–of–visual–words) by assigning the nearest visual word to each feature vector.

The final image descriptor of given image I, called Bag–of–visual–words, is defined as a vector

$$F(X) = \{v_1, \ldots, |v_k\} \quad (5)$$

Such that, for each local feature vector f(x) extracted from the image i

$$\bar{v}_i = \sum_{l=1}^{n_f} \sum_{j=1}^{k} g_j\ (f(x_l)), \forall_i = 1, \ldots \ldots k \quad (6)$$

Here n_f indicates the number of local features extracted from the image and mentioned mid–level feature,

$$g_j\left(f(x) = \begin{cases} 1\ if\ d_\epsilon\ \left(f(x), v_j\right) \le d_\in(f(x), v_l) \forall v_l \in V \\ 0\ therwise, \end{cases}\right.$$
(7)

As well as its variants, can be used in combination with any of the Low–level local features.

STEPWISE Bag-of -visual- words process:

STEP 1. Represent each training image as a vector- Use a bag of visual words representation

STEP 2. Train a classification method to discriminate vectors with respect to positive and negative training images - Use a Support Vector Machine (SVM) classifier

STEP 3. Apply the SVM based trained classifier to the test image

Figure 7. Bag of visual words representation

As shown in Figure 7 Visual words are 'iconic' image patches or fragments represent the frequency of occurrences of word but not theirposition. Thus BoVW methods largely unaffected by position and orientation of object in image having fixed length vector irrespective of number of detections and found Very successful in image classification according to the objects they contain.

2.2.2 Vector of Locally aggregated Descriptors (VLAD)

VLAD[6] is a recently introduced mid–level feature that has been shown to outperform the state–of–the–art of BoVW representation in several computerized image processing tasks. It differs from the BoW[11] image descriptor by recording the difference from the cluster center, varies the number of SIFTs assigned to the cluster. It inherits some of the invariances of the original SIFT[5] descriptor, like in-plane rotational invariance, and which is tolerant to other transformations such as image scaling and clipping. Another difference from the standard BoW[11] approach is that VLAD retrieval systems generally preclude the use of the original local descriptors. These descriptors are used in BoW[11] systems for spatial verification and re-ranking, but it requires more storage to be held in memory on a single machine for very large image datasets.

The image descriptor $F(X) = \{v_1, \ldots, |v_k\}$ is a concatenation of vectors $\overline{v_j}$ with elements $\overline{v_{ij}}$ defined as:

$$\bar{v}_{i,j} = \sum_{j=1}^{d} f\left(x_j\right) - v_{i,j}\ , \forall f(x)\ such\ that\ NN\left(f(x)\right) = v_i$$
(8)

Where,

d is the dimensionality of the feature space,

F(x) is a local feature vector extracted from the image I,$V_i \in V$ is the visual word,

$v_{i,j}$ is the j^{th} element of v_i,

NN(x) is the nearest visual word to x.

Figure 8 showsExample of perspective-based space transformation and surrogate text representation.

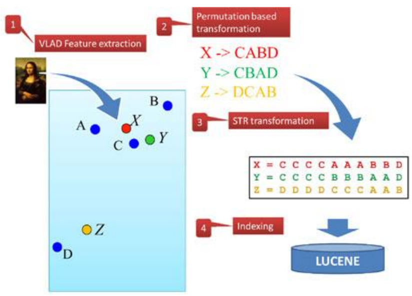

Figure 8.Example of perspective-based space transformation and surrogate text representation: 1) from the images we extract the VLAD features represented by points in a metric space. Blue points are reference features and colored points are data features. 2) The points are transformed into permutations of the references. 3) The permutations are transformed into text documents. 4) The text documents associated with the images are indexed.

Conventional search engines use methods like inverted indexing and file indexing to speed up the solution of user queries. We are studying a methodology which will support inverted files of standard text search engines to index vectors of locally aggregated descriptors (VLAD)[6] to deal with large-scale image search scenario.

3. GLOBAL DESCRIPTORS METHODS

Global descriptor describes the whole image. They are not very robust as some change in part of the image may cause it to fail as it will affect the resulting descriptor. A local descriptor defines a patch within an image. An image is matched with multiple local descriptors and this is more robust as there is no necessity to match all the descriptors for all the comparison to be made.

3.1 Riesz Miniature Transform

This descriptor represents the image as a single Riesz transform vector, which is a multidimensional extension of the Hilbert transform. A first step is performing the down sampling of the image to reduce the dimensionality of the descriptor. It uses a linear combination of Nth order Riesz templates at varying scales. The weights of the linear combination are obtained from one–versus–all support vector machines. Steer ability and multiscale properties of Riesz wavelets allow for scale and rotation covariance of the descriptor. Orientations are normalized by locally aligning the Riesz templates, which is carried out analytically. This approach has been used to model texture, shown to outperform state–of–the–art texture attributes in lung classification.

The Riesz transform is a multidimensional extension of the Hilbert transform, which maps any function I(x) to its harmonic conjugate and is a very powerful tool for mathematical manipulations of periodic signals. For an image $I_i(x)$ 2 R2, the different components of the Nth–order Riesz transform R are defined in the Fourier domain as

$$R(\widehat{n_1\, n_2})I_i\,(\omega) = \sqrt{\frac{n_1+n_2(-j\,\omega_1)(-j\,\omega_1)^{n_1}}{n_1!n_{2!}\;\|\omega\|^{n_1+n_2}}}\,\widehat{I_i}(\omega) \qquad (9)$$

For all combinations of (n1, n2) with n1 +n2 = N and n1,2 2 N. $\widehat{I_i}(\omega)$denotes the Fourier transform of $I_i(x)$, where the vector ω is composed by ω 1,2 corresponding to the frequencies in the two image axes. The multiplication by j ω1, 2 in the numerator corresponds to partial derivatives of f and the division by the norm of ω in the denominator results in only phase information being retained. Therefore, the 1st–order R corresponds to an all pass filter bank with directional (singular) kernels $h_{1,2}$:

R Ii(x) =

$$\mathcal{R}\mathbf{I}_i(x) = \begin{pmatrix} \mathcal{R}^{1,0} \\ \mathcal{R}^{0,1} \end{pmatrix} = \begin{pmatrix} h_1(x) * \mathbf{I}_i(x) \\ h_2(x) * \mathbf{I}_i(x) \end{pmatrix},$$

where

$$h_{1,2}(x) = \frac{x_{1,2}}{2\pi||x||^3},$$

and $x_{1,2}$ correspond to the axes of the image. The Riesz transform alters with translation, scaling or rotation. The orientation of the Riesz components is determined by the partial derivatives in Equation Where as 2N Riesz filters are generated by above equation; only N +1 components have distinct properties due to commutation of the convolution operators. The Riesz components yield a steerable filterbank allowing toanalyze different textures in any direction, which is an advantage when compared to classical Gaussian derivatives or Gabor filters. Qualitatively, the first Riesz component of even order corresponds to a ridge profilewhereas for odd ones we obtain an edge profile, but much richer profiles can be obtained by linear combinations of the different components. The templates of h1,2(x) convolved with Gaussian kernels for N=1,2,3 are depicted in Fig. 3. The Nth–order Riesz transform can be coupled with an isotropic multiresolution decomposition (e.g., Laplacian of Gaussian (LoG)) to obtain rotation–covariant (steerable) basis functions.Figure 9 shows Riesz representation of healthy verusus fibrosis pattern.

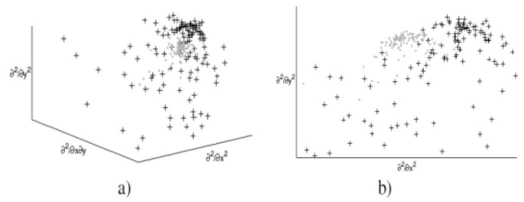

Figure 9.Riesz representation of healthy (gray dots) versus fibrosis (black crosses) patterns. a) Initial Riesz coefficients in 3D. b) The Riesz coefficients in 2D after having locally aligned the texture based on local prevailing orientation.

3.2 Color and Edge Directivity Descriptor (CEDD)

This global descriptor extracts the color information from regions of the image using a set of fuzzy rules and resulting in a HSV color space histogram. It includes texture information using the MPEG-7 Edge Histogram Descriptor rules. Finally, it uses Gustafson Kessel fuzzy classifier to binarize the histogram. The descriptors with more than one feature in a compact histogram can be considered as the family of Compact Composite Descriptors. The structure of CEDD[3] contains 6 texture areas and each texture area is separated into 24 sub regions, with each sub region describing a color. CEDD's color information retrieved from 2 fuzzy systems that which will map the colors of the image in a 24-color palette. CEDD[3] composed of a fuzzy version of the five digital filters proposed by the MPEG-7 which is useful to extract feature information. The CEDD extraction procedure is defined as follows: When an image block (rectangular part of the image) relates with the system that extracts a CCD, this section of the image concurrently goes across 2 units. The first unit known as the color unit, classifies the image block into one of the 24 shades used by the system. Consider classification be in the color m, m ∈ [0, 23] i.e total 24 shades. The second unit, known as the texture unit, which classifies the texture area a this section of the image where a ∈ [0, 5]. The image block is categorized in the bin a × 24 + m. The process is repeated for all the calculated image blocks of the given image. After completion of the process, the histogram is normalized within the interval value [0,1] and quantized for binary representation in a 3 bits per bin quantization

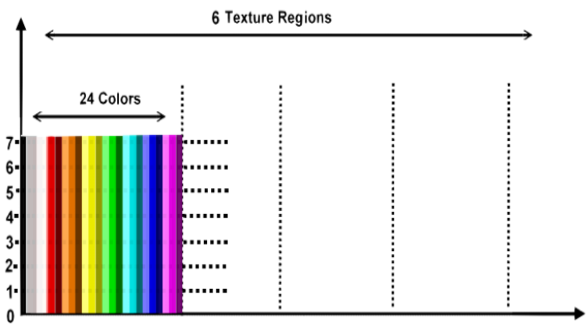

Figure.10 Structure of CEDD.

The most significant attribute of CEDDs is the achievement of best results that they bring up in various known benchmarking image databases. Another important attribute of CEDD [3] is it requires small size for indexing images. The CEDD length is 54 bytes per image.

3.3 Fuzzy Color Histograms

In contrast with conventional color histogram (CCH) which normally sets each pixel into one of the available bins only, where our FCH considers the color similarity information by assigning each pixel's total membership value to all the histogram bins. Very similar to the CEDD [3] feature, FCTH[4] mainly differs in using Haar Wavelet transform to model texture information. The color histogram is viewed as a set of color distribution from the probability viewpoint. As color space is given containing color bins, the color histogram of image containing pixels is represented as $H[I]=[h_1,h_2\ldots h_n]$, where $h_i=N_i/N$ is the calculated probability of a pixel in the image belonging to the i^{th} color bin, and is the total number of pixels in the i^{th} color bin. According to the probability theory, can be defined as follows

$$h_i = \sum_{j=1}^{N} P_{i\,|\,j} P_j = \frac{1}{N} \sum_{j=1}^{N} P_{i\,|\,j}$$

Where P_j is the probability of a pixel selected from image being the j^{th}

Pixel, which is1/N, and $P_{i|j}$ is defined as the conditional probability of the selected j^{th}pixel belonging to the i^{th} color bin.

3.4 Local binary pattern (LBP)

As we need another base feature extractor we use a three-dimensional adaptation of the Local Binary Pattern operator (LBP) [7] described by Ojala. It responds well to microscopic structure, is computationally cheap and invariant to the gray–scale range. Figure 11 shows The LBP [7] operator computes the local structure at a given pixel i by comparing the values of its eight neighboring pixels with the value of i. In some medical imaging modalities, such as CT scan method in which intensities of local regions are an important decision making instrument used by practitioners. As the LBP operator is defined as gray-scale invariant, we furthermore supplement a local average intensity measure DI of the 3x3x3 LBP cube to map the feature vector if absolute intensity is relevant.

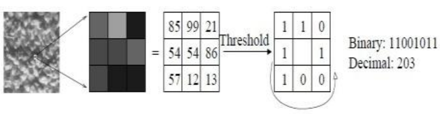

Figure.11 Structure of LBP operator

The LBP [7] operator was originally designed for obtaining texture description. The operator sets a label to every pixel of an image by making thresholding for the 3x3-8 connected region neighboring of each pixel with the center pixel value and converting the result as a binary number.

3.5 Text Bags

Texture bags learn features that represent a given set of training data by establishing a vocabulary of features. These features are based on e.g., local binary patterns and are optimal for describing the variability present in a specific region. The features extracted from training data are quantized by performing clustering in the feature space. The resulting k clusters constitute the texture vocabulary or texture words Wk. Each voxel is represented with its closest texture words k, i.e. with the index of the closest cluster center. Image or volume regions such as super voxels are represented by the histograms of texture words occurring in the region. The method has two phases. (1) During the learning phase the algorithm calculates descriptors and learns a three-dimensional texture vocabulary to grab the structure in the training data. (2) In the retrieval phase, the medical practitioners marks a region of interest in the query image, the algorithm searches for mapped regions in terms of texture, contrast, and intensity in the whole data set, and ranks images in the imaging repository accordingly. Figure 12 shows three examples of 3D texture words.

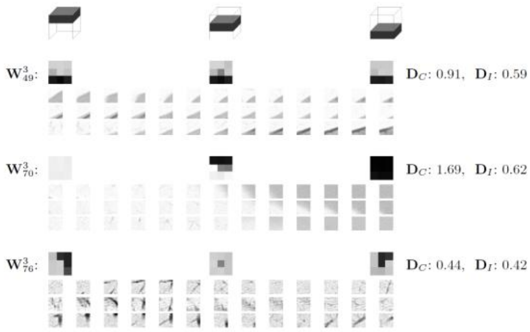

Figure .12 Three example 3D texture words w49, w70, w76 trained in an unsupervised manner on lung tissue on scale 3. The slices of the cube from top to bottom are depicted from left to right. Next to the cube are the measures for contrast (DC) and intensity (DI). Below each texture words are three examples of lung tissue belonging to this word. Note the similar structure of the example-tissues for each texture word.

4. ANALYSIS

Local feature descriptor methods are SIFT[5] and SURF [2]. It is found that the SIFT [5] has detected more number of features compared to SURF [2] but it has low speed. The SURF [2] is fast as compared to SIFT and has good performance as the same as SIFT. Midlevel descriptors are Bag of visual words and VLAD[6]. VLAD[6] was designed to be very low dimensional (e.g. total 16 bytes per image) so that all the descriptors for larger number of image datasets (e.g. 1 billion images) could still fit into main memory.VLAD[6], like visual word encoding, starts by vector quantizing a locally invariant descriptor such as SIFT. It differs from the BoW[11] image descriptor by recording the difference from the cluster center, rather than the number of SIFTs assigned to the cluster. BoW[11] inherits some of the invariances of the original SIFT[5] descriptor, such as rotational invariance in plane, and is somewhat tolerant to other transformations such as image scaling and clipping. Another difference from the standard BoW[11] approach is that VLAD[6] retrieval systems generally preclude the use of the original local descriptors. These are used in BoW[11] systems for spatial verification and re-ranking, but they require too much storage to be held in memory on a single machine for very large image datasets. In the global descriptors methods fuzzy color and texture histogram (FCTH) gives better performance than CEDD [3] and Bag of visual words method.

5. CONCLUSION

This paper surveys various image processing methods that are currently available and are used in medical image analysis. Comparisons between different methods are provided for the readers who want to develop system using content based retrieval. A combination of two or three methods can be used together for efficient performances.

The future work involves how these methods can be parallelized into Map-reduce framework and development of CBIR system for medical image analysis using Hadoop.

6. REFERENCES

[1] DimitriosMarkonis, Reńe Donner , LjiljanaDolamic,RogerSchaer,GeorgLangs, Ćelia Boyer, Henning Muller " *Report on and Prototype of final Image Retrieval and Analysis Framework" of KHRESMOI"Feb 2014.*Bowman, M., Debray, S. K., and Peterson, L. L. 1993.

[2] Herbert Bay, TinneTuytelaars, and Luc Van Gool. "*Surf: Speeded up robust features".* InComputer Vision–ECCV 2006, pages 404–417. Springer, 2006

[3] Savvas A. Chatzichristofis and Yiannis S. Boutalis. "*CEDD: Color and edge directivityDescriptor: A compact descriptor for image indexing and retrieval"*. In Lecture notes inComputer Sciences, volume 5008, pages 312–322, 2008.

[4] Savvas A. Chatzichristofis and Yiannis S. Boutalis. "*FCTH: Fuzzy color and texture histogram: A low level feature for accurate image retrieval"*. In Proceedings of the 9th InternationalWorkshop on Image Analysis for Multimedia Interactive Service, pages 191–196, 2008.

[5] Lindeberg,Tony (2012). *"Scale invariant feature transform"*. *Scholarpedia* **7** (5): 10491

[6] ReljaArandjelovi´c Andrew Zisserman, "*All about VLAD"*.University of Oxford.

[7] HaihongShen, Qishan Zhang1, Dongkai YangAdaptive Local Binary Patterns for 3D Face Recognition

[8] Andreas Burner1,?, Rene Donner1 , Marius Mayerhoefer2 , Markus Holzer1 , Franz Kainberger2 , Georg Langs1,3"*Texture Bags: Anomaly Retrieval in Medical Images Based on Local 3D-Texture Similarity* in Medical Content-Based Retrieval for Clinical Decision Support".springer Volume 7075 pp 116-127.

[9] AdrienDepeursinge, Antonio Foncubierta–Rodriguez,Dimitri Van de Ville and Henning Muller,*Multiscale Lung Texture Signature Learning Using The Riesz Transform"* in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012.Springer Volume 7512 pp 517-524.

[10] Rene Donner, Bjoern H. Menze , Horst Bischof , Georg Langs ,"Global localization of 3D anatomical structures by pre-filtered Hough Forests and discrete optimization" Elesivier,Journal of medical Image Analysis. 2013 December; 17(8): 1304–1314.

[11] Guillaume Lavoué, "Bag of Words and Local Spectral Descriptor for 3D Partial Shape Retrieval" Eurographics Workshop on 3D Object Retrieval (2011).

[12] Sebastian Haas, Ren´e Donner, Andreas Burner, Markus Holzer, and Georg Langs "*Superpixel-Based Interest Points for Effective Bags of Visual Words Medical Image Retrieval"*, MCBR-CDS 2011, LNCS 7075, pp. 58–68, 2012. Springer-Verlag Berlin Heidelberg 2012.

[13] Jyoti S. Patil, Sunayana A. Mane "3-D Image Analysis Using Mapreduce" ICPC 2015 IEEE conference on pervasive computing.

[14] Chris Sweeney Liu Liu Sean Arietta Jason LawrenceUniversity of Virginia "HIPI: A Hadoop Image Processing Interface for Image-based MapReduce Tasks"