

Effective & High Performance Optimized Techniques For Analysis Of Echocardiographic Image In Bioinformatics

Sanjay N. Patil ¹, Dr. Athar Ali ²

¹ Ph D Research Student, Maharishi University of Information Technology, Lucknow, India

² Research Guide, Maharishi University of Information Technology, Lucknow, India

Abstract - To address the problem of manually analysis of images, in this research work we attempted to present efficient, robust and especially automatic framework for the analysis of echo images based on computer aided diagnosis (CAD) using image processing. Now days, CAD is regular routine in many big hospitals for detection of different diseases such as cancer, brain tumour etc. The proposed research work is aimed to make the accurate medical decisions using echo images. This work is based on different steps such as appropriate echo images acquisition, its pre-processing, segmentation, feature extraction and then detection in order to perform analysis. For each step, we designed efficient algorithms. For echo images, the extraction ROI (Region of Interest) is challenging. The ROI extraction is done by using the space and time efficient image segmentation algorithm. This research work addresses such problem for color Doppler and 2D images by designing efficient segmentation algorithm. The analysis of color Doppler images is done by using color histogram; texture, edge density, as well as other statistical techniques are used. For classification process, we used naïve bayes classifier. The extensive simulation results for echo images are presented and discussed during this research work.

Key Words: Optimization technique, Image In Bioinformatics, Analysis Of Echocardiographic, Image processing.

1. INTRODUCTION

Here we explore flowing of the blood through vessels, heart valves and chambers. The movement of the blood reflects sound waves to a transducer. Then the ultrasound computer measures the direction and speed of the blood flowing through the heart and blood vessels. Doppler measurements may be displayed in black and white or in color [7].

Doppler Echocardiography is a technique for identifying the direction and velocity of moving blood within the heart. Figure shows Doppler echocardiographic image. It identifies the cardiac and stenosis and other abnormal flows. Doppler test gives us quantitative information needed in the decision-making process pertaining to the patient with heart disease. It also provides evaluation of normal and abnormal flow conditions [10]. A special

microphone has been engaged by Doppler Echocardiography to allow the technician for analyzing the speed and direction of blood flow in the vessels. Thus it enables to know the direction and evaluation of regurgitation. The obstruction can be precisely located by evaluating the speed of blood flow at different locations around the obstruction. Now Doppler Echocardiography has become the most consistent, noninvasive, sensitive and specific technique for diagnosing Valvular Regurgitation [10]. The amassing of color Doppler enables mapping of the regurgitant flow and guides while placing the Doppler beam parallel to the flow. Doppler Echocardiographic Systems can be broadly divided into two types. They are continuous wave and pulsed wave. The main difference between the two systems is the design and operating features of the transducer and signal processing procedures and the type of information supplied.

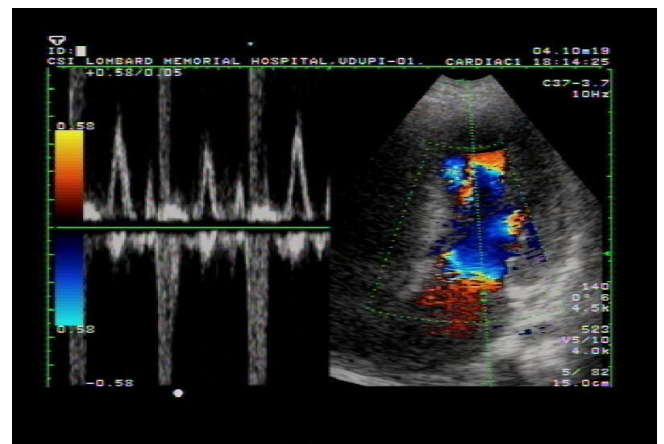


Fig-1: Example of Doppler Echocardiography

Continuous Wave Doppler: It is one of the ancient and easier of two types. Figure depicts Continuous Wave Doppler (CWD) representing Aortic Stenosis. The Continuous Wave Doppler generates and receives ultrasound waves continuously. Here two crystals are used one is to send and the other is to receive the echoes. Transmitter provides continuous sinusoidal wave. The receiver detects the shift. The analog recorder or spectral analyzer creates and records the sound, which is audible. Relative importance is assigned to the spectral analysis,

which keeps in isolation the system into individual components. Involving high sensitivity to low velocities and detection of high velocity without aliasing is the main advantage. At the same time, it has some disadvantages. Continuous Wave Doppler finds it difficult to distinguish between sending and receiving signals or external echoes, as also it does not present an exact image like Pulsed Wave Doppler.

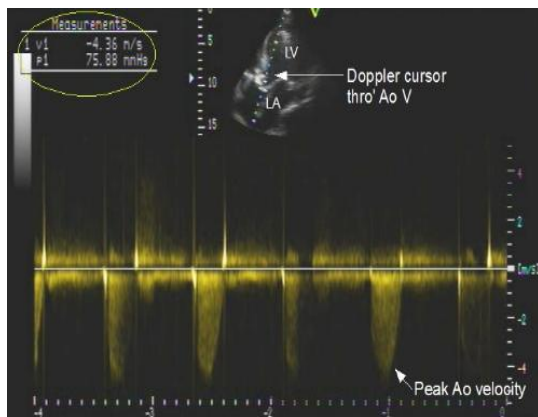


Fig-2: Continuous Wave Doppler imaging in Aortic Stenosis

One of the main advantages of CW Doppler system is in precisely evaluating high blood velocities with competence. It is certain that the highest velocities in valvular and congenital heart diseases are recorded with precision. The velocities, which exceed 1.5 m/s, are often seen in such disorders, but to recognize the full abnormal flow profile the precision with velocity measurement is of great importance. One main drawback of CW Doppler is its lack of sensitivity or depth discrimination. To permit selective placing of a given Doppler sample volume in space, there is no facility for imaging or range gating because CW Doppler is constantly transmitting and receiving from two separate transducer heads (crystals). Normally true Continuous Wave Doppler is functionally a stand-alone technique whether or not the capacity is outfitted within a 2D imaging transducer. It is because the output from a CW analysis contains Doppler shift data from every red cell reflecting ultrasound back to the transducer along the course of the ultrasound beam.

Pulsed Wave Doppler: Like the M-mode transducer the Pulsed Doppler system also uses a transducer, which simultaneously transmits and receives ultrasound. The sample volume refers to the main help of Pulsed Doppler with its main ability in providing Doppler shift data from a small segment along the ultrasound beam. Operator controls the location of the sample volume. For a given time X, an ultrasound pulse is sent into the tissues, which travel until it reaches blood cell, and reflected back. Over the same time interval it will then return to the transducer, with a shifted frequency. The total time to

transmit and receive is 2X. The distance to sample volume is ultrasound speed divided by round-trip travel time because the speed of ultrasound in the tissues is constant. This procedure is carried on frequently each second through many transmit-receive cycles.



Fig-3: Pulsed Wave Doppler Imaging

“Consequently this range gating is relying on a timing mechanism that just samples the returning Doppler shift information from a specific region. This is standardized that operator selects a particular place for the sample volume; the range gate is permit only Doppler shift information in that region, which is displaced as output. Every other returning ultrasound data is fundamentally not noticed. Other major assistance of Pulsed Wave Doppler (PWD) is the fact that a few imaging should carry on alternately with Doppler. Hence the sample volume may be displayed on the actual two dimensional for guidance. PW Doppler’s ability is plausible in mixtures of imaging from mechanical or phased array imaging system. Usually it is dirigible through the two-dimensional field of view, even though not all systems have this capability.

2. Related work

Li et al. (2012)

Author has suggested new approach to lessen over-segmentation using both pre and post processing for watershed segmentation. A new watershed segmentation process that combines pre-processing of the image and post-processing of image objects to produce the final segmentation results. In the initial stage of the watershed transform, this not only generates a gradient image from the original image. It also presents the texture gradient. The texture gradient can be obtained by a gray-level co-occurrence matrix.

Ghoshale et al. (2013)

Author has described the several edge sharpening filters and to discover the effect on the output image using watershed algorithm. A spatial sharpening filter on the performance of the segmented images and mathematical

morphology plays a very important role. Morphology is the special type of filtering and structuring elements used for shape smoothing and elimination of small holes.

Rahman et al. (2013)

Author present, a novel image segmentation process based on adaptive threshold and masking operation with watershed algorithm. Whose objective is to overcome over-segmentation problem of the traditional watershed algorithm.

Kailash Sinha et al. (2014)

This paper presents a comparative study of three segmentation methods implemented for tumor detection. The methods include k-means clustering with watershed segmentation algorithm, optimized k-means clustering with genetic algorithm and optimized c-means clustering with genetic algorithm. Traditional k-means algorithm is sensitive to the initial cluster centers. Genetic c-means and k-means clustering techniques are used to detect tumor in MRI of brain images. At the end of process the tumor is extracted from the MR image and its exact position and the shape are determined.

Arti Taneja et al. (2015)

This paper analyses the performance of algorithms for image segmentation in detail. Intensity and texture based image segmentation is the two levels of the level set method. The combination of both intensity and texture based image segmentation provides better results than the traditional methods. The detailed survey of segmentation techniques provides the requirement of the suitable enhancement method that supports both intensity and texture based segmentation for better results.

Kaushik Subramanian et al. (2006)

This paper present an overview of feature extraction method for off line recognition segmented character, in selection of a feature extraction method it is most likely important factor in achieving high recognition performance in signature recognition system. Kaushik described a novel approach for feature learning and information security by unifying image processing techniques and image encryption. Enhanced images are processed by edge detectors followed by segmentation using the Radon transform. These characteristics are then transformed to a multi-dimensional space using the Gabor transform to obtain the required feature vector. The uniqueness and the preservation of the image features are found to be maximum using this form of representation, this acquired vector is transmitted by using matrix-based encryption techniques.

Dennis, C.R. and Weisberg S. et al. (2007)

Feature extraction method are designed for different representations of the character such as solid binary character, character contours, skeleton or grey level sub-

images of each individual character. The feature extraction methods are discussed in term of invariance properties, reconstruct ability and expected distortions as variability of the character.

Mrozek et al. (2000)

Author discussed the problem of extracting linear features from a noisy image and counting the number of branching points and this may be successfully solved by homological methods applied directly to the image without the need of skeletonization and the analysis of the resulting graph. The superimposition of a mask set over the original image and works even when the homology of the feature is trivial and in arbitrary dimension saves the inherent limitation.

Nikolay et al. (2004)

This paper describes dimensionality reduction as a widespread pre-processing step in high-dimensional data analysis, visualization and modelling, and one of the simplest ways to reduce dimensionality is by feature extraction. Feature Extraction (FE) is a more general method in which one tries to develop a transformation of the input space onto the low-dimensional subspace that preserves most of the relevant information.

Agarwal S. et al. (2002)

The form of information learning is required in an accurate recognition and classification of the image. This process has made great progress using Bayesian classifiers and decision trees. But if the image is covered by noise, and patterns are confused with each other or incomplete, image based modeling is very difficult to some of these noise and feature extraction problems, we need a new feature learning algorithm.

Steven et al. (2007)

Author describe the implementation and performance of a Genetic Algorithm (GA) which generates image feature extraction algorithms for remote sensing applications, these describe as the basis set of primitive image operators and present of chromosomal representation of a complete algorithm, This application has been geospatial feature extraction using publicly available multi-spectral aerial-photography data set.

Srinivas C. C. et al. (2004)

Srinivas address two issues that are fundamental to the analysis of naturally-occurring datasets i.e. how to extract features that arise at multiple-scales and how to cluster items in a dataset. Srinivas also present two spectral methods they are; *Sparse Principal Component Analysis S-PCA* and *EigenCuts*. In *Sparse Principal Component Analysis (S-PCA)* this based on the discovery that natural images exhibit structure in a low dimensional subspace in a local, scale-dependent form and a framework for learning a linear, orthonormal basis representation for structure

intrinsic to a given dataset. It is also motivated by the observation that PCA does not typically recover such representations, due to its single minded pursuit of variance. *Eigen Cuts* this is an algorithm for clustering items in a dataset using their pair-wise similarities, and also a clustering algorithm for finding stable clusters in a dataset. In the eigen cuts, the algorithm performs clustering by removing these identified bottlenecks in an iterative fashion.

3. RESEARCH METHODOLOGY

Algorithm 1: Existing K-Means Clustering

(2) k - number of Clusters
 (3) D - data set containing n objects
 (4) C_j - clusters containing subset of n objects, where $j \in \{1, 2, \dots, k\}$ Randomly choose k objects from D as the initial clusters

Repeat

for each object il (where, $l \in \{1, 2, \dots, n\}$) do

(re)assign il to C_j using Euclidean distance calculation (similar objects) for each Cluster C_j do

Update the cluster means (or centroids) considering all objects currently in the cluster

Until no more reassigning

end K_MeansCluster.

It then computes the new mean for each cluster. This process iterates until the criterion function converges. The K-Means procedure is summarized in Figure 3.12. Typically, the square-error criterion is used to terminate the algorithm and is defined in equation 3.32.

$$E = \sum_{i=1}^K \sum_{p \in C_i} |p - m_i|^2 \quad (3.32)$$

where, E is the sum of the square error for all objects in the data set; p is the given object; m_i is the mean of cluster C_i . In other words, for each object in each cluster, the distance from the object to its cluster center is squared, and the distances are added.

Algorithm 2: Proposed K-Means Clustering

(5) k - number of Clusters
 (6) D - data set containing n objects
 (7) C_j - clusters containing subset of n objects, where $j \in \{1, 2, \dots, k\}$ Randomly choose k objects from D as the initial clusters

Repeat

Parfor each object il (where, $l \in \{1, 2, \dots, n\}$) **do**

(re)assign il to C_j using Euclidean distance calculation (similar objects) **parfor each** Cluster C_j **do**

Update the cluster means (or centroids) considering all objects currently in the cluster

Until no more reassigning

end K_MeansCluster.

Algorithm 3: SNGFilter(I)

```
// Input: Segmented binary image
// Output: Filtered Image, F
// pixels[r, c] ← I: pixel data with r rows and c columns
1 // Initialization
2 WHITE ← 200 // Threshold for white pixels
3 BLACK ← 25 // Threshold for black pixels
4 d ← 6; // Length of island to be filtered
1. count ← 1
2. flag ← FALSE
3. for i ← 1 to r do
4. j ← 0
5. while j < c - 1 do
6. current ← pixels[i, j]
7. if current ≥ WHITE
8. next ← pixels[i, j+1]
9. if (i == 0) up = WHITE else up = pixels[i - 1, j]
10. pos ← j + 1
11. while next ≥ BLACK and up ≥ WHITE and count ≥ d and j ≥ c-1 do
12. next = pixels[i, j + 1]
13. if (i == 0) up = WHITE else up = pixels[i - 1, j]
14. count ← count + 1
15. flag ← TRUE
16. if flag = TRUE and pixels[i, j] ≥ WHITE
17. for k ← 1 to count do
18. pixels[i, pos] ← 255
19. pos ← pos+1
20. j ← j + count // search next block
21. count ← 1 // reset count and flag
22. flag ← FALSE
23. F ← pixels
return F
```

Algorithm 4: MGAC

- Input: Echo Image, I
- Output: Contour Image, C

```
1 Initialize  $\alpha, \lambda, \mu, \epsilon, g$  2  $T \leftarrow 800$ 
4  $roi \leftarrow$  get the initial ROI // initial contour
•  $u \leftarrow$  GeometricActiveContour ( $roi, g, \alpha, \lambda, \mu, \epsilon$ ) // initial level set computation
•  $area \leftarrow$  Area( $u$ )
•  $prev \leftarrow temp \leftarrow u$ 
•  $prevArea = area$ 
• for  $i \leftarrow 1$  to 100 do
•  $C \leftarrow$  GeometricActiveContour ( $prev, g, \alpha, \lambda, \mu, \epsilon$ )
•  $currentArea \leftarrow$  Area( $C$ )
• if Abs( $currentArea - prevArea$ ) >  $T$ 
•  $\alpha = 0.5$  // slow down
• else
•  $\alpha = 1.5$  // normal speed
•  $prev = C$ 
•  $prevArea = currentArea$ 
• return  $C$ 
end GeometricActiveContour
```

4. RESULTS AND DISCUSSION

Essentially, three different types of experiments are carried out. The first one compares running times of the conventional K-Means, FK-Means, and Proposed K-Means algorithms by choosing the number of clusters (k) as 3, and later to 5, data size (n) as 100,000, number of dimensions (d) as 1, and number of iterations (Q) as 4 being executed on two different machines M1 and M2. The second type is by varying the data size from 10,000 pixels to 100,000 pixels with $k = 3$, $d = 1$, and $Q = 4$ against the proposed K-Means, FK-Means, and Conventional K-Means implementations. For consistency, all three algorithms are executed on the same dataset with the same initial placement of centers. Here execution time of the algorithm depends on the number of iterations, type of image, and pixels distribution. The experiments were repeated several times to check for consistent results.

Figure 4 shows ten 2D echo images of size 400×250 (100,000 pixels) and the pixel data is one dimension (grayscale intensity values) being used for clustering, the number of clusters = 3. This figure also shows the segmented images along with the original images after 4 iterations.

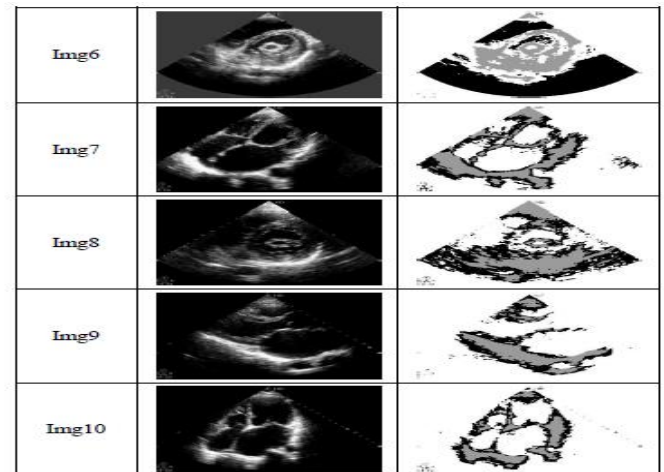
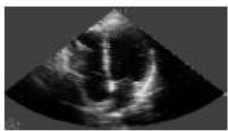





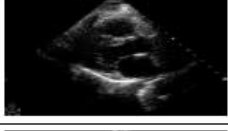





Fig-4: Two-Dimensional original and segmented echo images in different views (400x250)

Image Id	Input Image	Segmented Image
Img1		
Img2		
Img3		
Img4		
Img5		

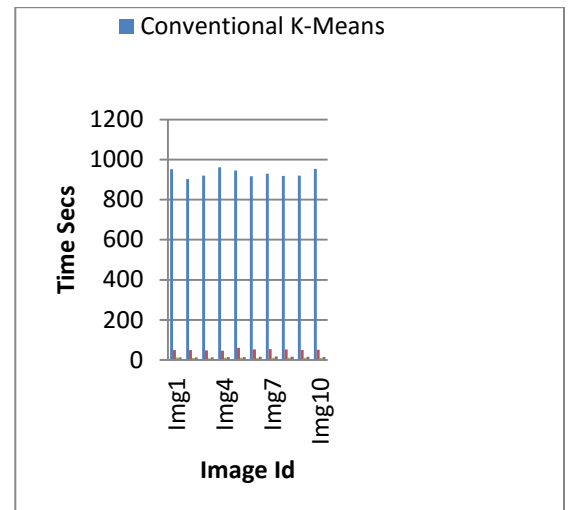


Fig-5: shows a graphical representation showing three K-Means algorithms with $k = 3$ and Parallel K-Means with $k = 5$.

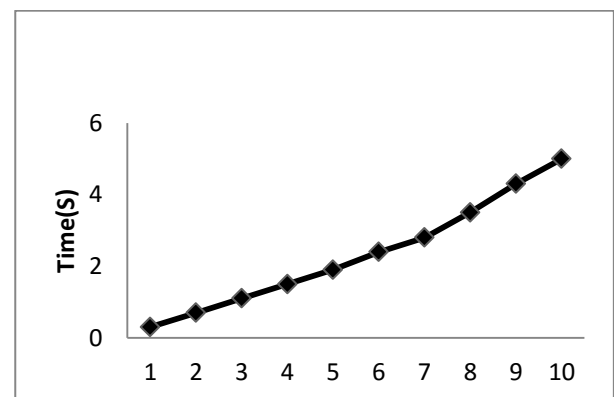


Fig-6: Plot of running time with increase of the number of iterations (Q) from 1 to 10 and $n = 100,000$ on machine M1.

5. CONCLUSIONS

In this research we designed framework to perform the accurate medical decisions using echo images automatically. The parallel K-means clustering algorithm is proposed and it is evaluated with existing methods in terms of speed efficiency. Due to the parallel computing proposed segmentation method minimized the time required for ROI extraction. After segmentation, novel post-processing algorithm designed. This algorithm was required to elimination the noisy part in segmentation regions to improve the prediction accuracies. This automated tool further designed the modified active counter based on ROI boundary tracing algorithm. This method was also evaluated in this thesis with respect to accuracy and speed of ROI boundary tracing. Finally the proposed classifier designed is used with different features extracted using geometric features, histogram features and statistical features. These features are used for color Doppler echo images analysis. The accuracy of proposed approach is showing approximate 97 % which is far better than present weka classifier tool.

REFERENCES

- [1] Barry E. Bleske, et. al., "Cardiovascular Risk Among University Students from Developed and Developing Nations", *The Open Cardiovascular Medicine Journal*, 5, 117-122, 2011.
- [2] G P Bezante, et al., "Left ventricular myocardial mass determination by contrast enhanced color Doppler compared with magnetic resonance imaging", *Heart Journal*, pp. 38-43, Feb 2005.
- [3] K. Bommanna Raja, M. Madheswaran, "Texture pattern analysis of kidney tissues for disorder identification and classification using dominant Gabor wavelet", *Machine Vision and Applications*, Vol. 21, pp. 287-300, Springer-Verlag, 2010.
- [4] E Boonchieng, W Boonchieng, and R Kanjanavanit, "Edge-detection and Segmentation Methods for Two-Dimensional Echocardiograms", *Computers in Cardiology*, 31, pp. 541-544, IEEE, 11 July 2005.
- [5] Bell. S. Brockhausen, P. "Discovery of Data Dependencies in Relational Databases", *Machine Learning and Knowledge Discovery in Databases*, The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases - ECML, March 1995.
- [6] Carlos Ordonez, "Integrating K-Means clustering with a Relational DBMS using SQL," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, No. 2, pp. 188-201, February 2006.
- [7] Carlos Ordonez, "Bayesian Classifiers Programmed in SQL," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, No. 1, pp. 139-144, January 2010.
- [8] Chotirat Ratanamahatana, "Scaling up the Naive Bayesian Classifier: Using Decision Trees for Feature Selection", *Proceedings of Workshop on Data Cleaning and Preprocessing (DCAP 2002)*, at IEEE International Conference on Data Mining (ICDM 2002), Maebashi, Japan. Dec 2002.
- [9] Christopher A. Troianos, MD, and Steven Konstadt, MD, "Evaluation of Mitral Regurgitation", *Seminars in Cardiothoracic and Vascular Anesthesia*, Vol 10, No 1 (March), 2006: pp 67-71.
- [10] Chunming Li, Chenyang Xu, Changfeng Gui, and Martin D. Fox, "Level Set Evolution Without Re-initialization: A New Variational Formulation", *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Jun 20-26, 2005.
- [11] C. L. Curotto, "Implementing Data Mining Algorithms with Microsoft SQL Server", *Wit Press*, UK, 2005.
- [12] Daniela Xhemali, "Naïve Bayes vs. Decision Trees vs. Neural Networks in the Classification of Web Pages", *International Journal of Computer Science Issues (IJCSI)*, Vol. 4, No. 1, pp. 16-23, 2009.
- [13] C. J. Date, *An Introduction to Database Systems*, Pearson Education 7th ed., 2000.
- [14] Dong in, Jia Pan, et al, "Medical Image Categorization based on Gaussian Mixture Model", *IEEE Computer Society, Intl. Conf. on Biomedical Engineering and Informatics*, 27-30, May 2008.
- [15] Donka Angelova • Lyudmila Mihaylova, "Contour segmentation in 2D ultrasound medical images with particle filtering", *Machine Vision and Applications* (2011) 22:551-561, Springer-Verlag, 2011, DOI 10.1007/s00138-010-0261-4.