

# Segmentation of Ultrasound Breast Images using Vector Neighborhood with Vector Sequencing on KMCG and augmented KMCG algorithms

Dr.H.B.kekre

Senior professor, Dept. of Computer Engineering  
MPSTME, SVKM's NMIMS University  
Mumbai, India.

Pravin Shrinath

Ph.D. Scholar  
MPSTME, SVKM's NMIMS University  
Mumbai, India.

**Abstract**— B mode ultrasound (US) imaging is popular and important modality to examine the range of clinical problems and also used as complimentary to the mammogram imaging to detect and diagnose the nature breast tumor. To understand the nature (benign or malignant) of the tumor most of the radiologists focus on shape and boundary. Therefore boundary is as important characteristic of the tumor along with the shape. Tracing the contour manually is a time consuming and tedious task. Automated and efficient segmentation method also helps radiologists to understand and observe the volume of a tumor (growth or shrinkage). Inherent artifact present in US images, such as speckle, attenuation and shadows are major hurdles in achieving proper segmentation. Along with these artifacts, inhomogeneous texture present in the region of interest is also a major concern. Most of the algorithms studies in the literature include noise removal technique as a preprocessing step. Here in this paper, we are eliminating this step and directly handling the images with high degree of noise. VQ based clustering technique is proposed for US image segmentation with KMCG and augmented KMCG codebook generation algorithms. Using this algorithm images are divided in to clusters, further these clusters are merged sequentially. A novel technique of sequential cluster merging with vector sequencing has been used. We have also proposed a technique to find out the region of interest from the selected cluster with seed vector acquisition.

Results obtained by our method are compared with our earlier method and Marker Controlled Watershed transform. With the opinion of the expert radiologist, we found that our method gives better results.

**Keywords**— *codebook; seed vector; training set; vector quantization*

## I. INTRODUCTION

B-Mode ultrasound (US) imaging is widely used diagnostic tool to examine the range of clinical problems because of its real-time image availability, non invasive nature and low cost of a scan. It has very low health risk to the patient during examination and image acquisition process relative to the other imaging modality [1, 2]. In the tissue characterization of the breast, US imaging is a complimentary method to the mammogram to distinguish between benign and malignant solid masses [3, 4]. Usually a malignant tumor has different characteristics, such as irregular shape, ill defined margins and heterogeneous echo texture as compare to benign

tumor. Round or oval shape with well defined boundary of a tumor is the most valuable information in the detection as benign and can be used to reduce the number of biopsies performed [5].

Segmentation of US images provides clinically valuable information for radiologists in terms of shape irregularity, boundary definition and quantitative measurement of lesion size. The accurate measurement of a lesion is basically used to monitor the tumor growth and also helps in treatment and planning for surgery [6]. Due to the limitation of acquisition process (dependence on expert radiologist) and technology, detection and measuring the size of tumor manually is difficult and time consuming process. However, with the improvement in the scanning devices (transducers), the inherent artifacts, such as speckle (which decrease signal-to-noise ratio), attenuation, shadows and signal dropout degrades the quality of the US images acquired. Many traditional segmentation algorithms are not suitable for such poor quality images, unless preprocessing steps to remove these artifacts has been used [7]. The artifact such as attenuation, which is causes by the gradual loss in the intensity of the ultrasound waves, generates inhomogeneous intensities within the same tissue type regions (i.e. tumor) and significant overlap between different class tissues (i.e. at the boundary). Blurred boundary and presence of tissue intensity variation in the region of interest are major constraints to achieve accuracy in the automated intensity based segmentation [8]. Some methods are discussed in the literature which handles this issue with the help of multiple images of the same region (sequence of images), but the processing of multiple images together is computationally inefficient [9, 10]. Other methods discussed in the literature for segmentation and classification such as, texture feature, thresholding [11,12], region growing and region merging [13,14], neural network, wavelet and watershed transform [15, 16], clustering [17, 18] etc, are strongly influenced by these inherent artifacts and involves steps to remove them to enhance the quality of images. The most usual artifact in US images is speckle and its degree is depends on human expertise, acquisition process and devices. This phenomenon highly affects the accuracy of segmentation and requires attention, so removal of noise has been extensively studied by the researchers and provides solutions [19-22]. Here, in this paper, we are demarking the boundary of the tumor in high degree noisy and attenuated US images

without involving any preprocessing (i.e. image enhancement) step. We are proposing Vector Quantization (VQ) based clustering with new computationally efficient codebook generation algorithm [23], i.e. Kekre’s Median Codebook Generation (KMCG) and augmented KMCG. Improved method of cluster merging with vector neighborhood is proposed on novel method of sequential merging of clusters [24] and compared with each other.

The other sections of this paper are organized as follows, in section II, vector quantization is discussed with encoding technique and its usability in segmentation. In section III, original VQ based KMCG codebook generation algorithm and augmented KMCG are discussed along with training set formation for horizontal and vertical division of the images. A novel technique of vector neighborhood with vector sequencing is described in section IV along with new technique for closing the holes in the clusters. In section V results are discussed followed by conclusion in section VI.

## II. VECTOR QUANTIZATION

Vector Quantization (VQ) was initially developed and implemented for image compression, with the help of many codebook generation and quantization algorithms [25, 26, 27, 28], but now a days it has been extensively use in other applications, such as image segmentation [29], speech recognition [30], pattern recognition and face detection [31, 32], tumor demarcation in MRI and Mammogram images [33, 34], content based image retrieval [35, 36] etc. In this paper, this method has been used as clustering aid in demarcation of area of interest (cysts and tumor) in breast ultrasound images.

A two dimensional image  $I(X, Y)$  is converted into  $K$  dimensional vector space of size  $M$ ,  $V = \{V_1, V_2, V_3, \dots, V_M\}$  (training set). VQ is used as a mapping function to convert this  $K$  dimensional vector space to finite set  $CB = \{C_1, C_2, C_3, C_4, \dots, C_N\}$ .  $CB$  is a codebook of size  $N$  and each code vector from  $C_1$  to  $C_N$  represents the specific set of vectors of the entire training set of dimensions  $K$  and size  $M$ . The codebook size is much smaller than size of the training set and it can represent entire training set. Here, in this paper, the work has been done in spatial domain and size of the codebook is limited to only eight codevectors, which are further used to forms eight clusters. As discussed in the section III A and B, KMCG and augmented KMCG are used as VQ based clustering algorithms and each cluster represents different regions of the image.

## III. CLUSTER FORMATION USING CODEBOOK GENERATION ALGORITHMS

### A. Kekre’s Median Codebook Generation algorithm (KMCG)

This algorithm was proposed for data compression [37, 38, 39], but this VQ based algorithm, proved its potential and usefulness in various applications, such as segmentation of mammographic images[35], content based image retrieval, face recognition etc. Here, in this paper, this iterative algorithm has been explored for demarcation of tumor from the breast US images.

Initially, Image  $I$ , is divided into  $M$  non-overlapping blocks of size  $2 \times 2$  and these blocks are further converted into

vectors of dimension  $1 \times K$ , ( $K=4$ ) as shown in the Fig. 1(a). Let,  $V$  be the training set, i.e.  $V = \{V_1, V_2, V_3, \dots, V_M\}$ , where  $V_1$  to  $V_M$  are individual vectors of the training set. This entire training set (matrix) of dimension  $M \times 4$  is considered as first cluster and become the input to the KMCG algorithm. To divide this cluster further, during the first iteration of this algorithm, entire training set has been sorted with respect to first value of all vectors (i.e. first column) and obtained the median of this sorted column. This median is consider as first codevector, and then divide training set into two clusters with respect to this codevector ,as shown in Fig. 1(b). (i.e upper part including median is first cluster  $C_1$  and lower part is second cluster  $C_2$ ). In second iteration, clusters  $C_1$  and  $C_2$  are sorted separately with respect to second column (i.e. second value of all the vectors) of the training set and obtain the medians of  $C_1$  and  $C_2$ . Further, cluster  $C_1$  and  $C_2$  are divided into four clusters by using these new medians, as shown in Fig. 1(c). Same procedure has been repeated till we obtained the desired number of clusters. In Fig. 1, cluster formations are shown up to second iteration with  $k = 4$ . After acquiring desired number of clusters, they are merged together sequentially as discussed in section IV.

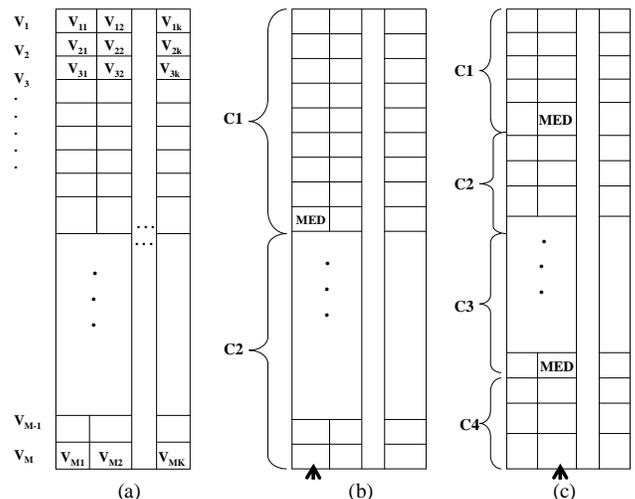


Fig. 1. Clustering Using KMCG Algorithm (A) Entire Training Set Of Size  $M \times K$  ( $K=4$ ) Obtained From Image. (B) Clusters  $C_1$  And  $C_2$  Obtained After First Iteration W.R.T. First Column Shown By Arrow. (C) Four Cluster Obtained From  $C_1$  And  $C_2$  W.R.T. Second Column.

### B. Augmented KMCG

KMCG algorithm has been augmented, to decrease the time require in the clustering process. In this method, vector size is increased to 6 columns, in which last four columns are used to store original gray levels obtained from  $2 \times 2$  blocks of the image. Further, averages of each of these blocks are done separately and stored at the second column in the respective vectors. At the first column, sequence number of the respective vectors has been stored. Eventually, the size of the entire training set used for this method becomes  $M \times 6$ . Therefore, unlike the original KMCG, sorting process is done only once on the second column and accordingly vectors of the entire training set are shuffled by keeping vector intact (i.e. vectors are arranged in the increasing order of their average values stored at second column). To make clusters, during the

process median value is obtained from the second column and entire training set is divided into two clusters. Further, these two clusters are handled separately and obtained median value from each of them from the same column. These median values further divides two clusters into four. This process is repeated until desired number of cluster has been formed. Since sorting process is applied only once on the second column (i.e. column of average gray values of the block), time require for clustering is drastically reduced as compared to original KMCG algorithm.

### C. Training set formation with vector sequencing

Original KMCG and augmented KMCG algorithms are used to form clusters from the training set. Here, in this paper, two separate training sets are formed for each codebook generation algorithms, Fig. 2 shows training set used for augmented KMCG algorithm. First training set is created by dividing image I (X,Y) into 2x2 non overlapping blocks horizontally and sequence number of these blocks are added at the first location of respective vectors as the vector number as shown in Fig. 2, therefore, first column of the training set contains vector sequence number. Similarly, same image is divided into non overlapping blocks vertically and forms the second training set by adding sequence number of blocks at the first column of the vector. Same procedure has been followed to create two training sets for original KMCG algorithm, except calculation of average gray levels. Therefore, size of these training sets becomes  $M \times 5$  and clusters are created according to the method discussed in section III. A.

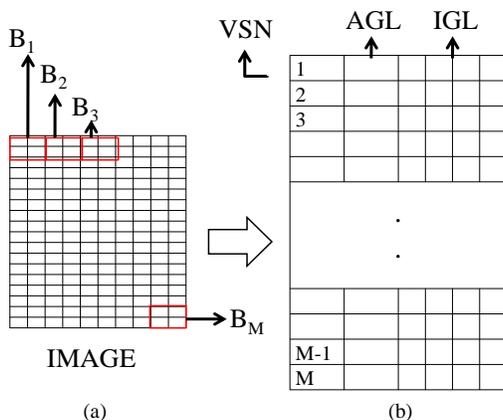


Fig. 2. (a) Original image, divided horizontally into  $M$  nonoverlapping blocks shown by red boxes. i.e.  $B_1, B_2, \dots, B_M$  and stored at Last four column of the respective vector. (b) Training set generated from Image shown in (a), where, first column indicates sequence number of the block used as Vector Sequence Number (VSN), Second column Indicates Average Gray Levels (AGL) calculated from Image Gray Levels (IGL) shown in last four column.

## IV. CLUSTER MERGING AND SEED VECTOR ACQUISITION

### A. Sequential cluster merging

As discussed in section III. C, pair of separate training sets is used for each codebook generation algorithms to make clusters. Initially, first training set (i.e. formed by horizontal division of image) from the pair is divided in to eight clusters using code book generation algorithms, similarly second training set (i.e. formed by vertical division of image) divided

in to another set of eight clusters by same algorithm. So pair of similar clusters is obtained for each algorithm. “Fig. 4”, shows cluster images obtained using KMCG algorithm. “Fig. 5”, shows cluster images obtained using augmented KMCG. These two sets of clusters are handled separately in segmentation process. Further these clusters are merged sequentially one-by-one and forms new sets of merged clusters. As shown in the “Fig. 6”, first cluster is added with second, resultant clusters is then added with third and resultant third cluster added with fourth and so on. Similarly clusters obtained by augmented KMCG are merged and shown in “Fig. 7”.

### B. Seed vector and exploration of clusters with horizontal and vertical sequencing

Images are divided horizontally and vertically into blocks and sequence of these blocks are used as vector sequence number (VSN) and added to the first column of the training set. After making eight clusters for these training sets, first cluster has been sorted by its vector sequence number (i.e. first column) and accordingly vectors are shuffled. Further median is acquired from the first column of the training set and vector which contains this median value is considered as seed vector as shown in “Fig. 3”. In most of the ultrasound images, gray level distribution is inhomogeneous but the pixels with lower gray value are concentrated at region of interest. Therefore after clustering, usually these gray levels are components of the first cluster in the form of vectors. Due to these characteristics of image, mostly seed vector obtained from the first cluster falls in the region of interest Therefore sequence number of the seed vector is used to gather neighboring vectors in the cluster. All the vectors in the cluster are searched with respect to sequence values of seed vector. Searching is done in both, right hand side and left hand side of the seed vector. Since cluster is in sorted order with respect to sequence number, it searches vectors only for consecutive sequence number, if any interlude occurs in between, it stops searching and marked all searched vectors as found and this vector line is considered as seed vector line as shown in “Fig. 3”. Obtained new seed vector by adding sequence number of the first seed vector with number of vectors present horizontally in the row (i.e.  $X/2$ ). To explore other vectors in the region of interest, first this new vector is checked whether it is present or not in the cluster by using equation 1. If this new vector is present in the cluster then second line has been explored, this process is repeated and lines are explored and lower region is grown from the region of interest. Same procedure is repeated to explore the upper part of the original seed vector line, only difference is, instead of addition, subtraction is used to obtained new seed vector (sequence number of the original seed vector is subtracted from number of vectors present horizontally in the row). Furthermore all marked vectors from these clusters are preserved and other vectors are removed. As shown in the “Fig. 3” vectors which are representing noise are removed from the cluster, since they are not the neighbors of the seed vectors and this new cluster is converted to the original size image as shown in the “Fig. 8 (a)”. Similar procedure is followed to explore the cluster with vertical sequencing, except addition and subtraction of value (i.e.  $Y/2$ ) is done with the vector sequence number every time to get the new seed vector. With addition of this value, right

hand side region is explored and with subtraction left hand side of the region is explored. This exploration has been stop similarly when newly generated vector does not present in the cluster. Cluster image for vertical indexing is shown in “Fig. 10 (a)”. Same procedure is applied for second merged cluster and third merged cluster after closing the opening as per discussed in the section IV C and cluster images are formed.

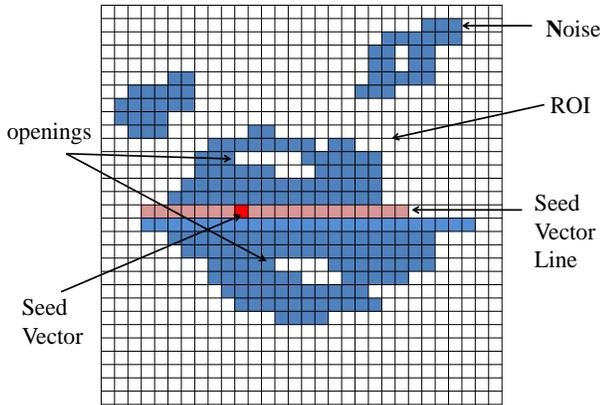


Fig. 3. Image Of First Cluster, Used To Obtain Seed Vector And Generate Seed Vector Line. This Seed Vector Further Used To Acquire Region Of Interest (ROI) By Vector Neighborhood With Vector Sequence Number.

C. Closing the opening of clusters

As discussed in the section IV-A, clusters are merged together one-by-one and sets of eight merged clusters are obtained for both horizontal and vertical sequencing. From this set desired merged cluster has been selected for closing the opening (Break in the sequence number of vector as shown in the “Fig. 3”). These vectors are not present in the cluster but they are part of region of interest). Here in this paper we proposed new technique to close the opening directly on cluster rather on cluster images. Here we select the certain threshold value, which indicates the number of vectors. Then selected cluster has been sorted with respect to its sequence number. This sorted cluster has been traversed sequentially from the first vector, if any consecutive interlude occur with respect to sequence number and it is less than or equal to threshold value then that many number of vector are added to the cluster with zero gray level. Sequence number of first newly added vector is started from sequence number of the vector where the interlude has started plus one. After closing the opening of these clusters, they are converted into images of original size as shown in “Fig. 8(c)” and “Fig. 10(c)”. for horizontal and vertical sequencing respectively.

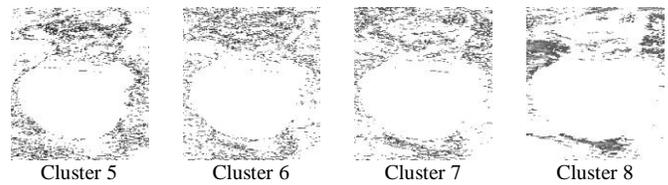
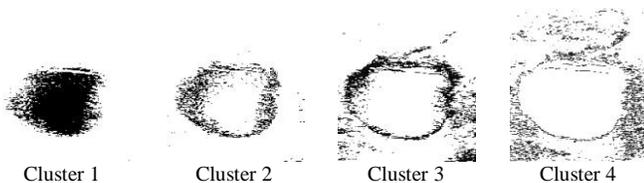


Fig. 4. Eight Cluster Images Obtained Using Original KMCG Algorithm For Dimension K= 4. Cluster Images Obtained For Both Horizontal And Vertical Division Of Image Are Same.

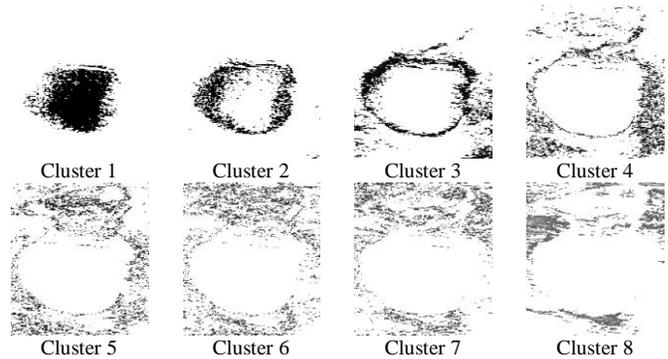


Fig. 5. Eight Cluster Images Obtained Using Augmented KMCG Algorithm For Dimension K= 4. Cluster Images Obtained For Both Horizontal And Vertical Division Of Image Are Same.

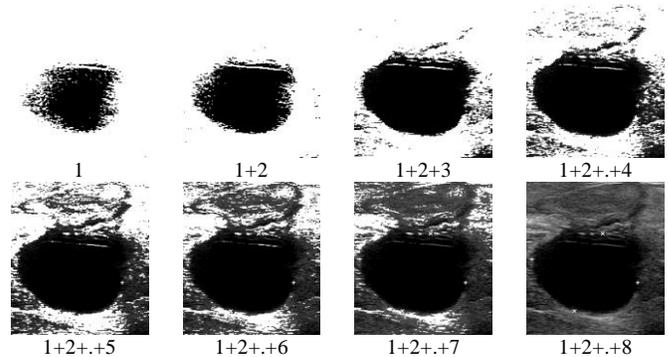


Fig. 6. Eight Sequentially Merged Clusters Images Obtained From Clusters Shown In “Fig. 4” Using KMCG For Both Horizontal And Vertical Division

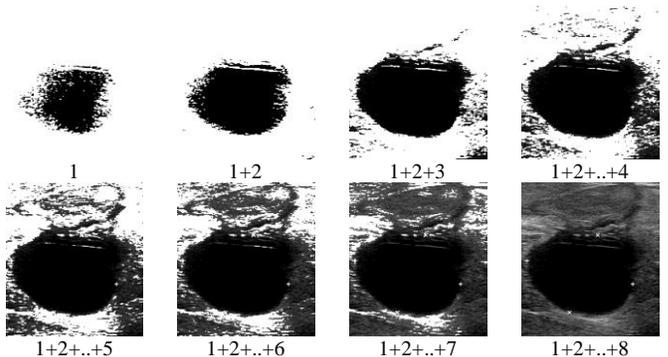


Fig. 7. Eight Sequentially Merged Clusters Images Obtained From Clusters Shown In “Fig. 5” Using Augmented KMCG For Both Horizontal And Vertical Division

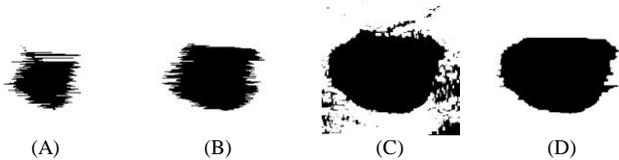


Fig. 8. Images Obtained From The Merged Clusters For **Horizontal Direction** Shown In “Fig. 6” (A) Image Obtained From The First Cluster Generated By Using Seed Vector. (B) Image Obtained From Second Merged Cluster Generated By Using Same Seed Vector Acquired From First Cluster. (C) Image Obtained From Third Merged Cluster After Closing The Holes. (D) Image Obtained From Third Cluster Shown In “Fig. 8(C)” By Using Same Seed Vector

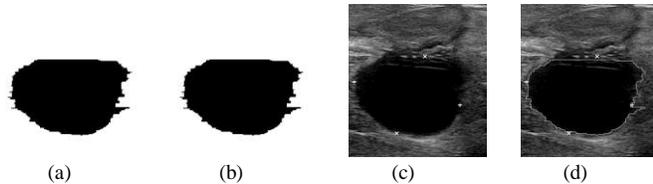


Fig. 9. Resultant Images Obtained By KMCG Algorithm With **Horizontal Division** (A) Image Obtained By Merging Cluster Images Shown In “Fig. 8 (A),(B) And (D)” (B) Segmented Resultant Image (C) Original Breast US Image With Tumor At The Center, Marked By Radiologist (D) Superimposed Image.



Fig. 10. Images Obtained From The Merged Clusters For **Vertical Direction** Shown In “Fig. 6” (A) Image Obtained From The First Cluster Generated By Using Seed Vector. (B) Image Obtained From Second Merged Cluster Generated By Using Same Seed Vector Acquired From First Cluster. (C) Image Obtained From Third Merged Cluster After Closing The Holes. (D) Image Obtained From Third Cluster Shown In “Fig. 10(C)” By Using Same Seed Vector.

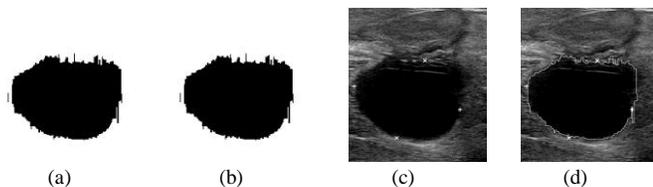


Fig. 11. Resultant Images Obtained By KMCG Algorithm With **Vertical Division** Of Original Image. (A) Image Obtained By Merging Cluster Images Shown In “Fig. 10 (A),(B) And (D)” (B) Segmented Resultant Image (C) Original Breast US Image With Tumor At The Center, Marked By Radiologist (D) Superimposed Image.

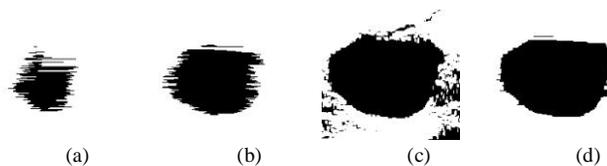


Fig. 12. Images Obtained From The Merged Clusters For **Horizontal Direction** Shown In “Fig. 7”. (A) Image Obtained From The First Cluster Generated By Using Selected Seed Vector. (B) Image Obtained From Second Merged Cluster Generated By Using Same Seed Vector Selected From First

Cluster. (C) Image obtained from third merged cluster after closing the holes. (d) Image obtained from third cluster shown in “Fig. 12.(c)” by selecting same seed vector

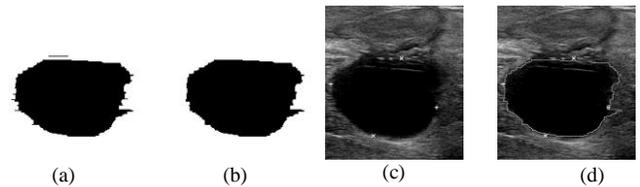


Fig. 13. Resultant Images Obtained By Augmented KMCG Algorithm With **Horizontal Division** Of Original Image. (A) Image Obtained By Merging Cluster Images Shown In “Fig. 12 (A),(B) And (D)” (B) Segmented Resultant Image (C) Original Breast US Image With Tumor At The Center, Marked By Radiologist (D) Superimposed Image

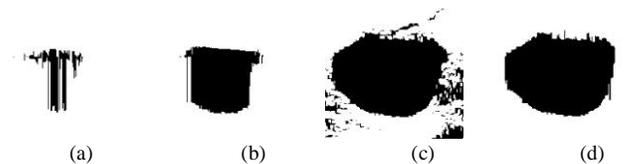


Fig. 14. Images Obtained From The Merged Clusters For **Vertical Direction** Shown In “Fig. 7” (A) Image Obtained From The First Cluster Generated By Selecting Seed Vector. (B) Image Obtained From Second Merged Cluster Generated By Using Same Seed Vector Selected From First Cluster. (C) Image Obtained From Third Merged Cluster After Closing The Holes. (D) Image Obtained From Third Cluster Shown In “Fig. 14.(C)” By Selecting Same Seed Vector.

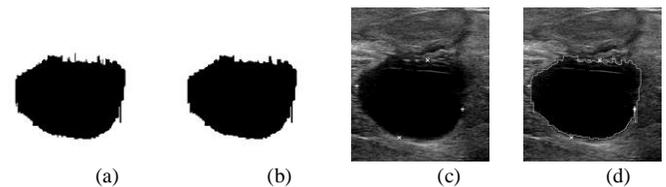


Fig. 15. Resultant Images Obtained By Augmented KMCG Algorithm With **Vertical Division** Of Original Image. (A) Image Obtained By Merging Cluster Images Shown In “Fig. 14 (A),(B) And (D)” (B) Segmented Resultant Image (C) Original Breast US Image With Tumor At The Center, Marked By Radiologist (D) Superimposed Image

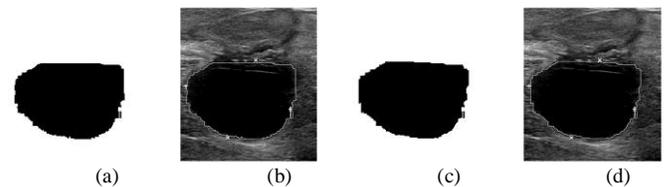


Fig. 16. Segmentation Results Obtained By KMCG Algorithm & Augmented KMCG Algorithm. (A) Image Obtained By Taking INTERSECTION Of Cluster Images Shown In “Fig. 9.(B)” With Cluster Image Shown In “Fig. 11.(B)” (Using KMCG) (B) Segmentation Results Superimposed On Original Image. (C) Image Obtained By Taking INTERSECTION Of Cluster Image Shown In “Fig. 13.(B)” With Cluster Image Shown In “Fig. 15.(B)” (Using Augmented KMCG). (D) Superimposed Image.

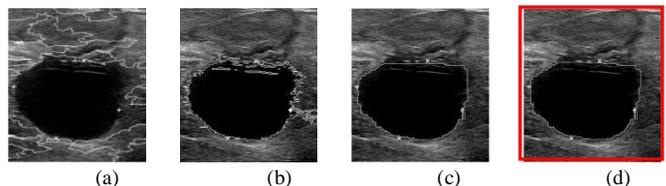


Fig. 17. Results Comparison (A) Results Of Marker Controlled Watershed Algorithm (B) Results By VQ Based Clustering [24] (C) Result By Our Method Shown In “Fig. 16.(B)”. (D) Results By Our Method Shown In “Fig. 16.(D)”.

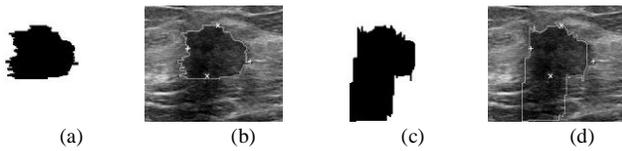


Fig. 18. Resultant Images Obtained By KMCG Algorithm With **Horizontal** And **Vertical Division** Of Original Image. (A) Segmented Image Obtained By Horizontal Division. (B) Superimposed Image (C) Segmented Image Obtained By Vertical Division. (D) Superimposed Image

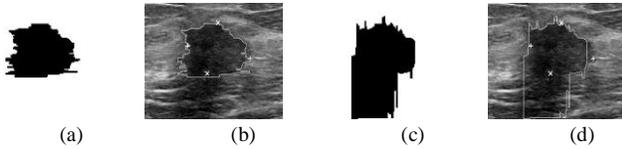


Fig. 19. Resultant Images Obtained By Augmented KMCG Algorithm With **Horizontal** And **Vertical Division** Of Original Image. (A) Segmented Image Obtained By Horizontal Division. (B) Superimposed Image (C) Segmented Image Obtained By Vertical Direction. (D) Superimposed Image

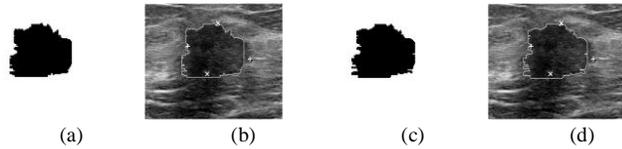


Fig. 20. Resultant Images Obtained By Taking INTERSECTION Of Two Images. (A) Image Obtained By INTERSECTION Of Images Shown In “Fig. 18.(A) And 18 (C)”. (B) Superimposed Image. (C) Image Obtained By INTERSECTION Of Images Shown In “Fig. 19.(A) And 19 (C)”. (D) Superimposed Image

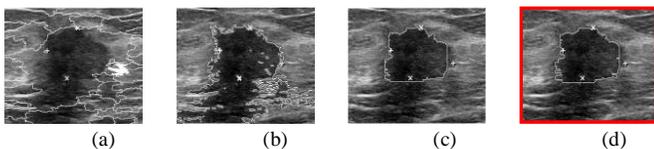


Fig. 21. Results Comparison (A) Results Of Marker Controlled Watershed Transform (B) Results By VQ Based Clustering [24] (C) Result By Our Method Shown In “Fig. 20.(B)”. (D) Results By Our Method Shown In “Fig. 20.(D)”.

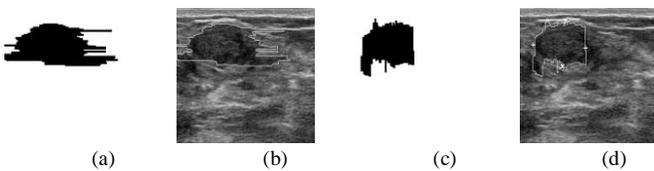


Fig. 22. Resultant Images Obtained By KMCG Algorithm With **Horizontal** And **Vertical Division** Of Original Image. (A) Segmented Image Obtained By Horizontal Division. (B) Superimposed Image (C) Segmented Image Obtained By Vertical Division. (D) Superimposed Image

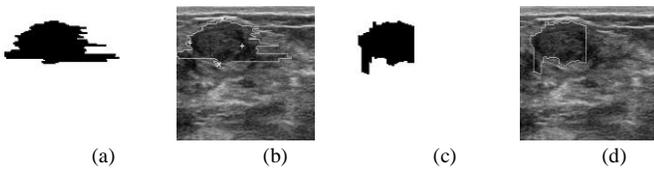


Fig. 23. Resultant Images obtained by augmented KMCG algorithm with **Horizontal** and **Vertical division** of original image. (a) Segmented image obtained by Horizontal division. (b) Superimposed image (c) Segmented image obtained by Vertical division. (d) Superimposed image

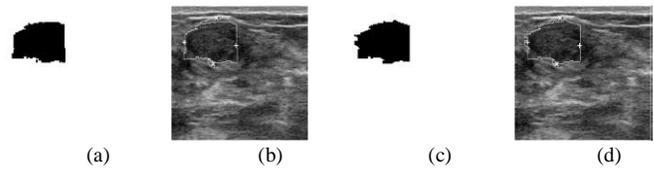


Fig. 24. Resultant Images Obtained By Taking INTERSECTION Of Two Images. (A) Image Obtained By INTERSECTION Of Images Shown In “Fig. 22.(A) And 22 (C)”. (B) Superimposed Image. (C) Image Obtained By INTERSECTION Of Images Shown In “Fig. 23.(A) And 23.(C)”. (D) Superimposed Image

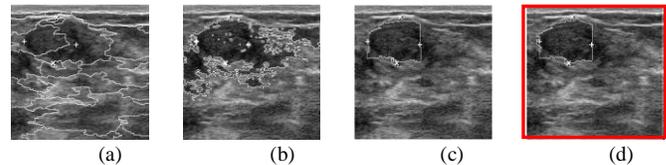


Fig. 25. Results comparison (a) Results of Marker Controlled Watershed transform (b) Results by VQ based clustering [24] (c) Result by our method, shown in “Fig. 24. (b)”. (d) Results by our method, shown in “Fig. 24. (d)”.

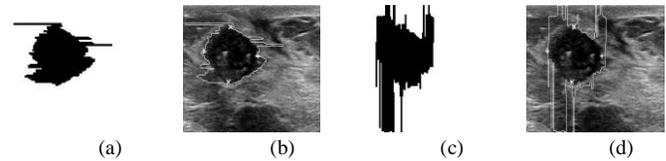


Fig. 26. Resultant Images Obtained By KMCG Algorithm With **Horizontal** And **Vertical Division** Of Original Image. (A) Segmented Image Obtained By Horizontal Division. (B) Superimposed Image (C) Segmented Image Obtained By Vertical Division. (D) Superimposed Image

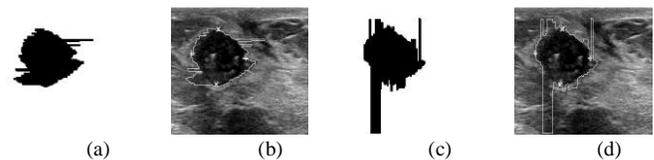


Fig. 27. Resultant Images Obtained By Augmented KMCG Algorithm With **Horizontal** And **Vertical Division** Of Original Image. (A) Segmented Image Obtained By Horizontal Division. (B) Superimposed Image (C) Segmented Image Obtained By Vertical Division. (D) Superimposed Image

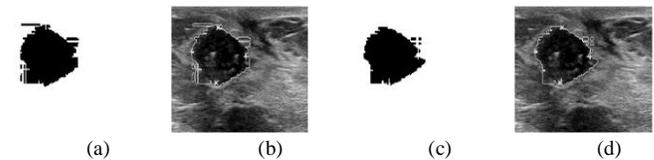


Fig. 28. Resultant Images Obtained By Taking INTERSECTION Of Two Images. (A) Image Obtained By INTERSECTION Of Images Shown In “Fig. 26.(A) And 26 (C)”. (B) Superimposed Image. (C) Image Obtained By INTERSECTION Of Images Shown In “Fig. 27.(A) And 27.(C)”. (D) Superimposed Image



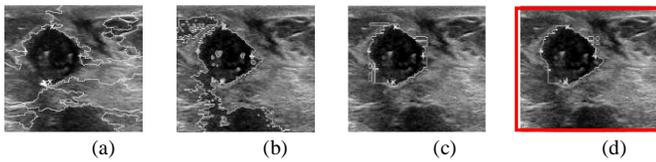


Fig. 29. Results Comparison (A) Results Of Marker Controlled Watershed Transform (B) Results By VQ Based Clustering [24] (C) Result By Our Method, Shown In “Fig. 28.(B)”. (D) Results By Our Method, Shown In “Fig. 28.(D)”.

## V. RESULTS

Here, in this paper, KMCG and augmented KMCG codebook generation algorithms are implemented for clustering process and further improved cluster merging technique is used to get final segmentation results. These methods are tested on real US images, using MATLAB 7.0 and Intel Core2 Duo 2.20GHz processor with 1 GB RAM. Results of four different images are shown.

Using these algorithms, sets of clusters and merged clusters are obtained, for different images. But complete result (results of each and every stage) is displayed only for one image shown in “Fig. 9(c)” and for other original images final segmented results are shown directly. All the results are validated by expert radiologists with visual inspection. Final results are compared with other techniques such as Marker Controlled Watershed transform and our other method [24]. Our results are compared with other method’s results as shown in the “Fig. 17”, “Fig. 21”, “Fig. 25” and “Fig. 29” and best results are shown by red box drawn around the image. Marker controlled watershed transforms gives over segmentation and other method discussed in [24], boundary is not clear around the tumor.

## VI. CONCLUSION

VQ based, KMCG and augmented KMCG algorithms are used for clustering and further it has been used to segment the ultrasound breast images. Here we used training set of size  $M \times 5$  and  $M \times 6$  for KMCG and augmented KMCG respectively. Since augmented KMCG algorithm requires sorting only once, it is computationally efficient than KMCG and on the basis of visual inspection by expert radiologist, it also having better segmentation results. Here in this paper we used vector indexing in the cluster formation process, which further helps to get the seed vector to grow the region of interest. Some images contains inhomogeneous texture within the tumor region, due to this openings may exist. This openings affects the accuracy of the segmentation, therefore we developed a new technique for closing. This technique is implemented directly on the clusters rather than cluster images.

Ultrasound images probably have smooth texture at the region of interest but course texture at the boundary of the normal tissue region and defected tissue region. Therefore tracing the boundary around the area of region is not remains a trivial task. Here in this paper we use US images with strong attenuation and high degree of noise without using any preprocessing task, because during preprocessing step, important information from the image could be lost. Our method focused on the computational efficiency as well as

accuracy of the segmentation. Results are compared with the recently developed marker control watershed transform and our newly developed method [24]. With the help of visual inspection and opinion of the expert radiologists, it is found that our results are improved and accurate

## ACKNOWLEDGMENT

The authors would like to thank Dr. Wrushali More, Dr. Mushtak and Dr. Anita Sable for their valuable guidance and suggestions to understand the nature and shape of the tumor in ultrasound images and validity of the segmentation results

## REFERENCES

- [1] Thomas L Szabo, “Diagnostic Ultrasound Imaging: Inside out”, Elsevier Inc., pp 19-23, Sept 2004.
- [2] S. Michelle Biering, Anne Jones, “Accuracy and Cost Comparison of Ultrasound Versus Alternative Imaging Modalities, Including CT, MR, PET, and Angiography”, *Journal of Diagnostic Medical Sonography*, vol. 25, no 3, pp 138-144, May 2009.
- [3] Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA, “Solid breast nodules: use of sonography to distinguish between benign and malignant lesions”, *Radiology*, pp 123-134, 196(1),1995.
- [4] K. J. W. Taylor, C. Merritt, C. Piccoli, R. Schmidt, G. Rouse, B. Fornage, E. Rubin, D. Georgian-Smith, F. Winsberg, B. Goldberg, and E. Mendelson, “Ultrasound as a complement to mammography and breast examination to characterize breast masses,” *Ultrasound Med. Biol.*, vol. 28, pp. 19–26, Jan. 2002.
- [5] M.A. Dennis, S.H. Parker, A.J. Klaus, A.T. Stavros, T.I. Kaske, S.B. Clark, “Breast biopsy avoidance: the value of normal mammograms and normal sonograms in the setting of a palpable lump”, *Radiology* vol. 219, no 1, pp 168–191. April 2001.
- [6] K. Horsch, M.L. Giger, L.A. Venta, C.J. Vyborny, “Computerized diagnostic of breast lesions on ultrasound”, *Medical Physics*, vol. 29, no.2, pp 157–164, Feb 2002.
- [7] J. Alison Noble, Djamel Boukerroui, “Ultrasound Image Segmentation: A Survey”, *IEEE Transactions on Medical Imaging*, Vol. 25, No. 8, pp 987-1010, Aug 2006
- [8] Guofang Xiao, Michael Brady, J. Alison Noble, Yongyue Zhang, “Segmentation of Ultrasound B-Mode Images With Intensity Inhomogeneity Correction, *IEEE TRANSACTIONS ON MEDICAL IMAGING*, VOL. 21, NO. 1, pp 48-57, JANUARY 2002
- [9] D.Boukerroui, O. Basset, A.Noble, and A. Baskurt, G’erard Gimenez, “A Multiparametric and Multiresolution Segmentation Algorithm of 3-D Ultrasonic Data”, *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, Vol 48 ,No1 pp 64-77, 2002.
- [10] D.Boukerroui, O. Basset, A.Noble, and A. Baskurt, “Segmentation of ultrasound images— multiresolution 2D and 3D algorithm based on global and local statistics”, *Pattern Recognition Letters* Vol 24 , pp 779–790, 2003, Elsevier, Science Direct, 2003.
- [11] H.B.Kekre, Pravin Shrinath, “Tumor Demarcation by using Local Thresholding on Selected Parameters obtained from Co-occurrence Matrix of Ultrasound Image of Breast”, *International Journal of Computer Applications*, Volume 32– No.7, October 2011, Available at: <http://www.ijcaonline.org/archives>
- [12] Dar-Ren Chen, Yu-Len Huang, Sheng-Hsiung Lin “Computer-aided diagnosis with textural features for breast lesions in sonograms” *Computerized Medical Imaging and Graphic*, Elsevier, vol. 35, pp- 220–226, 2011.
- [13] X. Hao, C.J. Bruce, C. Pislaru, J.F. Greenleaf, Segmenting high-frequency intracardiac ultrasound images for myocardium into infarcted, ischemic and normal regions, *IEEE. Trans. Med. Imag.* 20 (12) pp 1373–1383, 2001.
- [14] A. Madabhushi, D.N. Metaxas, Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions, *IEEE Trans. Med. Imag.* 22 (2), pp155–169, 2003.
- [15] D. R. Chen, R. F. Chang, W. J. Kuo, M. C. Chen, and Y. L. Huang, “Diagnosis of breast tumors with sonographic texture analysis

- using wavelet transform and neural networks,” *Ultrasound Med. Biol.*, vol. 28, no. 10, pp. 1301–1310, Oct. 2002.
- [16] Y. L. Huang and D. R. Chen, “Watershed segmentation for breast tumor in 2-D sonography,” *Ultrasound Med. Biol.*, vol. 30, no. 5, pp. 625–632, May 2004.
- [17] Jin-Hua Yu, Yuan-Yuan Wang, Ping Chen, Hui-Ying Xu, “Two-dimensional Fuzzy Clustering for Ultrasound Image Segmentation”, published in the proceeding of IEEE International Conference on Bioinformatics and Biomedical Engineering, pp 599-603,1-4244-1120-3, July 2007.
- [18] Chang Wen Chen, “Image Segmentation via Adaptive k-Mean Clustering and Knowledge-Based Morphological Operations with Biomedical Applications”, *IEEE TRANSACTIONS ON IMAGE PROCESSING*, VOL. 7, NO. 12, PP 1673-1683, DECEMBER 1998
- [19] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, “Statistics of speckle in ultrasound B-scans,” *IEEE Trans. Sonics Ultrasonics*, vol. SU-30, pp. 156–163, Mar. 1983.
- [20] Y.J. Yu, S.T. Acton, “Speckle reducing anisotropic diffusion,” *IEEE Trans on Image Processing* ,vol.11.no.11,pp.1260-1270,2002
- [21] S.Sudha, G.R.Suresh and R.Sukanesh, “Speckle Noise Reduction in Ultrasound Images by Wavelet Thresholding based on Weighted Variance”, *International Journal of Computer Theory and Engineering*, Vol. 1, No. 1, pp 7-12, April 2009
- [22] Peter C. Tay, Christopher D. Garson, Scott T. Acton, John A. Hossack, “Ultrasound Despeckling for Contrast Enhancement”, *IEEE TRANSACTIONS ON IMAGE PROCESSING*, VOL. 19, NO. 7, pp 1847-1860, JULY 2010
- [23] H. B. Kekre, Sudeep D. Thepade, Adib Parkar, “Performance Analysis of Kekre’s Median Fast Search, Kekre’s Centroid Fast Search and Exhaustive Search Used for Colouring a Greyscale Image”, *International Journal of Computer Theory and Engineering*, Vol. 2, No. 4, August, 2010
- [24] H. B. Kekre, Pravin Shrinath, “Tumor Demarcation by VQ based clustering and augmentation with KMCG and KFCG codebook generation algorithms”, 2<sup>nd</sup> world congress of communication and information technology, IEEE conference, Trivendram.. pp 993-998, 978-1-4673-4806-5 IEEE, Oct 2012
- [25] R. M. Gray, “Vector quantization”, *IEEE ASSP Magazine*., pp. 4-29, Apr.1984.
- [26] Pamela C. Cosman, Karen L. Oehler, Eve A. Riskin, and Robert M. Gray, “Using Vector Quantization for Image Processing”, *Proceedings of the IEEE*, pp- 1326-1341, Vol. 81, No. 9, September 1993
- [27] W. H. Equitz, “A New Vector Quantization Clustering Algorithm,” *IEEE Trans. on Acoustics, Speech, Signal Proc.*, pp 1568-1575. Vol-37, No-10, Oct-1989.
- [28] Huang, C. M., Harris R.W., “A comparison of several vector quantization codebook generation approaches”, *IEEE Transactions on Image Processing*, pp 108 – 112, Vol-2, No-1, January 1993.
- [29] H. B. Kekre, Tanuja K. Sarode, Bhakti Raul, “Color Image Segmentation using Kekre’s Algorithm for Vector Quantization *International Journal of Computer Science (IJCS)*, Vol. 3, No. 4, pp. 287-292, Fall 2008. Available at: <http://www.waset.org/ijcs>.
- [30] Chin-Chen Chang, Wen-Chuan Wu, “Fast Planar-Oriented Ripple Search Algorithm for Hyperspace VQ Codebook”, *IEEE Transaction on image processing*, vol 16, no. 6, June 2007.
- [31] Qiu Chen, Kotani, K., Feifei Lee, Ohmi, T., “VQ-based face recognition algorithm using code pattern classification and Self-Organizing Maps”, 9th International Conference on Signal Processing, pp 2059 – 2064, October 2008.
- [32] C. Garcia and G. Tziritas, “Face detection using quantized skin color regions merging and wavelet packet analysis,” *IEEE Trans. Multimedia*, vol. 1, no. 3, pp. 264–277, Sep. 1999.
- [33] H. B. Kekre, Tanuja K. Sarode, Saylee Gharge, “Detection and Demarcation of Tumor using Vector Quantization in MRI images”, *International Journal of Engineering Science and Technology*, Vol.1, Number (2), pp.: 59-66, 2009. Available online at: <http://arxiv.org/ftp/arxiv/papers/1001/1001.4189.pdf>.
- [34] H. B. Kekre, Dr.Tanuja Sarode, Ms.Saylee Gharge, Ms.Kavita Raut, “Detection of Cancer Using Vector Quantization for Segmentation”, Volume 4, No. 9, *International Journal of Computer Applications (0975 – 8887)*, August 2010.
- [35] H. B. Kekre, Ms. Tanuja K. Sarode, Sudeep D. Thepade, “Image Retrieval using Color-Texture Features from DCT on VQ Codevectors obtained by Kekre’s Fast Codebook Generation”, *ICGST-International Journal on Graphics, Vision and Image Processing (GVIP)*, Volume 9, Issue 5, pp.: 1-8, September 2009. Available online at <http://www.icgst.com/gvip/Volume9/Issue5/P1150921752.html>
- [36] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, “Color Texture Feature based Image Retrieval using DCT applied on Kekre’s Median Codebook”, *International Journal on Imaging (IJI)*, Available online at [www.ceser.res.in/iji.html](http://www.ceser.res.in/iji.html)
- [37] H. B. Kekre, Tanuja K. Sarode, “Centroid Based Fast Search Algorithm for Vector Quantization”, *International Journal of Imaging and Robotics (IJIR)*, Volume 1, Number A08, pp. 73-83, Autumn 2008, available: <http://www.ceser.res.in/iji.html>
- [38] H. B. Kekre, Tanuja K. Sarode, “An Efficient Fast Algorithm to Generate Codebook for Vector Quantization”, *First International Conference on Emerging Trends in Engineering and Technology*, IEEE Computer Society, pp 62-67, 978-0-7695-3267-7/08 2008 IEEE
- [39] H. B. Kekre, Tanuja K. Sarode, “Fast Codebook Generation Algorithm for Color Images using Vector Quantization,” *International Journal of Computer Science and Information Technology*, Vol. 1, No. 1, pp: 7-12, Jan 2009.

#### AUTHORS PROFILE

**Dr. H. B. Kekre** has received B.E. (Hons.) in Telecomm. Engineering. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970 He has worked as Faculty of Electrical Engg. and then HOD Computer Science and Engg. at IIT Bombay. For 13 years he was working as a professor and head in the Department of Computer Engg. at Thadomal Shahani Engineering College, Mumbai. Now he is Senior Professor at MPSTME, SVKM’s NMIMS University. He has guided 17 Ph.Ds, more than 100 M.E./M.Tech and several B.E./ B.Tech projects. His areas of interest are Digital Signal processing, Image Processing and Computer Networking. He has more than 450 papers in National / International Conferences and Journals to his credit. He was Senior Member of IEEE. Presently He is Fellow of IETE and Life Member of ISTE. 13 Research Papers published under his guidance have received best paper awards. Recently 5 research scholars have been conferred Ph. D. by NMIMS University. Currently 07 research scholars are pursuing Ph.D. program under his guidance.

**Mr. Pravin Shrinath** has received B.E. (Computer science and Engineering) degree from Amravati University in 2000. He has done Masters in computer Engineering in 2008. Currently pursuing Ph.D. from Mukesh Patel School of Technology Management & Engineering, NMIMS University, Vile Parle (w), Mumbai. He has more than 10 years of teaching experience and currently working as Associate Professor in Computer Engineering Department, MPSTME.