Segmentation of Blast using Vector Quantization Technique

H.B.Kekre,Ph.D
Senior Professor, MPSTME, NMIMS
University, Vile-Parle(w), Mumbai-400056, India

Archana B. Patankar,Ph.DAsst. ProfessorComputer departmentTSEC,Bandra(W)
Mumbai – 50, India

Hema Ramesh Galiyal
ME Student
Computer department
TSEC, Bandra (W),
Mumbai – 50, India

ABSTRACT
Image enhancement plays an important role in computer vision and image processing. Leukemia is a malignant disease (cancer) seen in people of any age groups either in children or adults aged over 50 years. It is characterized by the uncontrolled accumulation of immature white blood cells. Further the noises and blurriness effect often lead to false diagnosis of leukemia .The recognition of acute leukemia blood cell based on color image is one of the most challenging tasks in image processing. Also, the conventional method of manual counting using a microscope is a time consuming, produces errors and put an intolerable amount of stress to technicians. As a solution to this problem, this paper proposed vector quantization technique for segmentation of blast in acute leukemia images. This method is applied on 115 microscopic images and succeeds with specificity of 90% and sensitivity of 60% to detect abnormal white blood cells (blast). Images used are availabe at www.dti.unimi.it/fscotti/all misp.mui.ac.ir/data/microscopic-image-data.html

General Terms
Algorithms

Keywords
Leukemia, Segmentation, VQ, LBG, KPE.

1. INTRODUCTION
Cancer is the uncontrolled growth of abnormal cells in the body. Cancerous cells are also called malignant cells. There are various types of cancers one of which is Leukemia. The term Leukemia, which is derived from Greek, literally means white blood. It is characterized by abundance of abnormal white blood cells in the body. It is a cancer that starts in the tissue that forms blood. To understand cancer, it helps to know how normal blood cells form. Most blood cells develop from cells in the bone marrow called stem cells. Bone marrow is the soft material in the center of most bones. Stem cells mature into different kinds of blood cells. Each kind has a special job, White blood cells, red blood cells, and platelets are made from stem cells as the body needs them. When cells grow old or get damaged, they die, and new cells take their place. In a person with leukemia, the bone marrow makes abnormal white blood cells. The abnormal cells are leukemia cells.
Unlike normal blood cells, leukemia cells don't die when they should. They may crowd out normal white blood cells, red blood cells, and platelets. This makes it hard for normal blood cells to do their work. So, Leukemia should be detected at early stage to start with treatment as soon as possible .The types of leukemia can be grouped based on how quickly the disease develops and gets worse. Leukemia is either chronic (which usually gets worse slowly) or acute (which usually gets worse quickly).
Leukemia is divided into two categories: Mylogenous or Lymphocytic, each of which can be acute or chronic. Acute Mylogenous Leukemia (AML) is a cancer of the myeloid line of blood cells, characterized by the rapid growth of abnormal white blood cells that accumulate in the bone marrow and interfere with the production of normal blood cells. Chronic Mylogenous Leukemia (CML), also known as chronic granulocytic Leukemia (CGL), is a cancer of the white blood cells [1, 2]. It is a form of leukemia characterized by the increased and unregulated growth of predominantly myeloid cells in the bone marrow and the accumulation of these cells in the blood. Acute Lymphoblastic Leukemia (ALL) is a form of leukemia, or cancer of the white blood cells characterized by excess lymphoblast. Malignant, immature white blood cells continuously multiply and are overproduced in the bone marrow. Chronic Lymphocytic Leukemia (CLL) causes a slow increase in white blood cells called B lymphocytes, or B cells. Currently, the microscopic investigation of blood cells is performed manually by Hematologists through visual identification under the microscope [3]. However, the manual recognition method requires a lot of time and effort. This method is therefore in appropriate to be utilized in large hospitals. Several algorithms and techniques have been developed for blood cells recognition. Image enhancement at the pre-processing stage becomes the most important process for a successful feature extraction and diagnosis of Leukemia. Contrast of the acute leukemia image is one of the factors that may influence the accuracy of interpretation by Hematologists [4-5].
The work we have done is to propose a segmentation process which segments abnormal white blood cells (blast), present in the acute leukemia image using vector quantization techniques.

1.1 Vector Quantization
Vector Quantization (VQ) [6, 7] is an efficient technique for data compression. VQ has been very popular in a variety of research fields such as speech recognition and face detection [8]. Vector quantization is a technique in which clustering is performed where clustering means grouping common region together. VQ is a technique in which a codebook is generated for each image. A codebook is a representation of the entire image containing a definite pixel pattern which is computed according to a specific VQ algorithm. In this algorithm the image is divided into equal size parts (blocks) that represent training vector. As we know formation of training vector is

first step for cluster formation [9-10]. Vector Quantization VQ can be represented as a mapping function that maps k-dimensional vector space to a finite set CB = {C1, C2, CN }. The set CB is called codebook consisting of N number of code vectors and each code vector Ci = {ci1, ci2, cik} is of dimension k [11-14]. The key to VQ is the good codebook. Codebook can be generated in spatial/transform domain by clustering algorithms. The method most commonly used to generate codebook is the Linde-Buzo-Gray (LBG) algorithm which is also called as Generalized Lloyd Algorithm (GLA).

2. ALGORITHM FOR SEGMENTATION

It is difficult, however, to compare the effectiveness of these methods because each gave appropriate results and the results varied between training and testing.

2.1 Proposed Algorithm

In this proposed algorithm color plane of original image is used. For further image segmentation blue and green plane of image are used as an input image.

The proposed technique has three steps to follow:

1. Pre processing.
2. Region forming using vector quantization technique.
3. Post Processing.

2.1.1 Pre processing

For complete color image we will separate three color planes red, green and blue. As the color of white blood cell is purple we have used collectively green and blue plane.

2.1.2 Vector Quantization

Vector quantization is a classical quantization technique from signal processing which allows the modeling of probability density functions by the distribution of prototype vectors. It was originally used for data compression. It works by dividing a large set of points (vectors) into groups having approximately the same number of points closest to them. Each group is represented by its centroid point, as in k-means and some other clustering algorithms [15].

2.1.2.1 Linde Buzo and Gray Algorithm (LBG)

In this algorithm, from the initial feature vector (training set) of an image, centroid is computed as a first code vector. By adding constant error to the code vector two code vectors V1 and V2 are generated [16]. Euclidean distance for all training vectors is computed with code vectors V1 and V2. Based on the smallest distance, training vectors are clustered into either V1 or V2. This process is repeated till the codebook of desired size is obtained. In Figure1 [17] two vectors v1 & v2 are generated by adding constant error to the codevector.

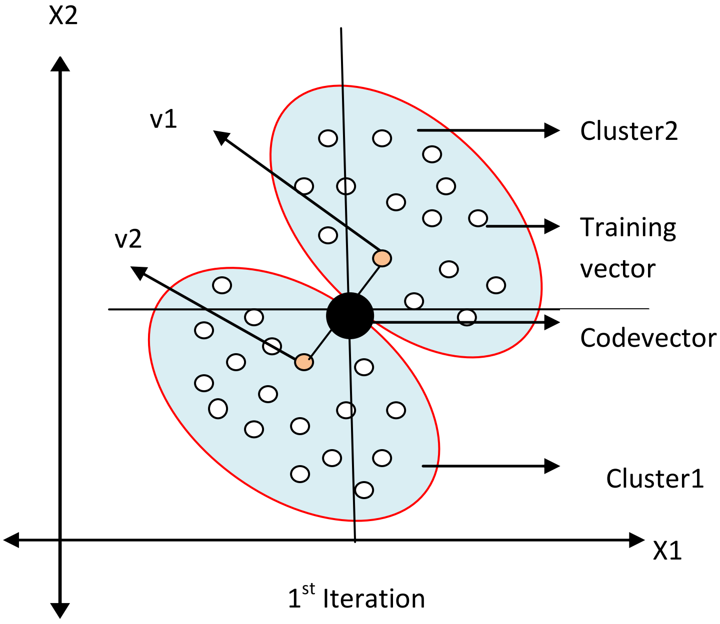


Figure 1: LBG for 2 dimensional case

2.1.2.2 Kekre's Proportionate Error (KPE) Algorithm

In this codebook generation algorithm, proportionate error is added to the centroid. The proportionate error is decided by magnitude of coordinates of the centroid. While adding proportionate error a safe guard is also introduced so that neither v1 nor v2 go beyond the training vector space eliminating the disadvantage of the LBG.Remaining procedure is same as in LBG algorithm mentioned above. In both the algorithms, codebook of desired size for trainee and test images are generated. Figure 2 shows the cluster elongation after adding proportionate error [18-19].

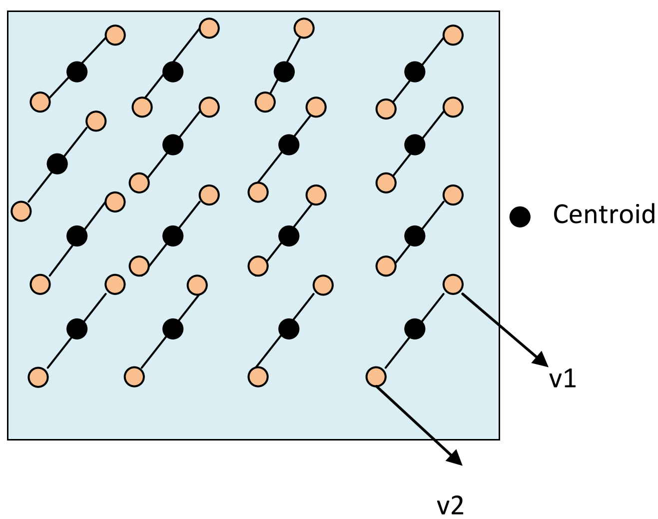


Figure 2: Orientations of the line joining two vector v1 and v2 after addition of proportionate error to the centroid

2.1.3 Post Processing

If some unwanted regions are present in the segmented image then post processing is required to achieve only blast in the final image. For that morphological operations and median filters are used.

3. RESULTS

In this section, the final results of blast detection are presented. This algorithm was tested on 115 microscopic images of size 256x256 captured with a simple light microscope with three ocular lenses and an analog video CCD which is coupled to a Pinnacle to digitize the captured images. Image database is availabe at
www.dti.unimi.it/fscotti/all.
misp.mui.ac.ir/data/microscopic-image-data.html
In this study, we successfully designed a system that can segment the blast (abnormal white blood cell) and evaluated the accuracy of the proposed methods. The specificity and sensitivity of this method is calculated using the following formula: [20, 21].

$$Sensitivity = TP / (TP+FN) \qquad (1)$$
$$Specificity = TN / (TN+FP) \qquad (2)$$

Where TP, FP, FN, and TN stand for true positive results, false positive results, false negative results, and true negative results.True Positives (*TP*) is the number of images correctly classified as positive by the test; True Negatives (*TN*) is the number of images correctly classified as negative by the test; False Positive (*FP*) is the number of images classified as positive by the test, but they are not; False Negative (*FN*) is the number of images classified as negative by the test, but they are not [22].

Table 4.1 Segmentation performance of clustering algorithm for AML type

Algorithm	Sensitivity	Specificity
LBG	60%	90%
KPE	67%	92%

Table 4.2 Segmentation performance of clustering algorithm for A LL type

Algorithm	Sensitivity	Specificity
LBG	55%	90%
KPE	60%	92%

Table 4.3 Different Vector quantization Algorithms applied on AML images

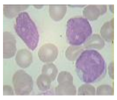
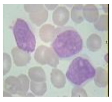
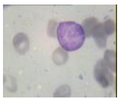
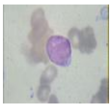
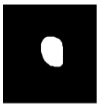
Original AML Image	Image after applying LBG	Image after applying KPE
 FIG 1.1	 FIG 1.2	 FIG 1.3
 FIG 2.1	 FIG 2.2	 FIG 2.3

Table 4.4 Different Vector quantization Algorithms applied on ALL images

Original ALL Image	Image after applying LBG	Image after applying KPE
 FIG 3.1	 FIG 3.2	 FIG 3.3
 FIG 4.1	 FIG 4.2	 FIG 4.3

4. CONCLUSION

In this paper vector quantization is used which is commonly used for data compression. Basically vector quantization is a clustering algorithm and can be used for texture analysis. Here the results of proposed algorithm using LBG and KPE for blast detection in acute leukemia images are displayed. The results are compared with well known public dataset of blood samples, which is specifically designed for the evaluation and comparison of the performances of algorithms for segmentation and image classification.

5. ACKNOWLEDGEMENT

The authors would like to record their appreciation to Dr. Fabio Scotti Assistant Professor Università degli Studi di Milano Dipartimento di Informatica Via Bramante 65 - 26013 Crema - Italy for providing public and free dataset of microscopic images of blood samples.

6. REFERENCES

[1] N.R.Mokhtar, 1nor Hazlyna Harun, 1m.Y.Mashor, 2h.Roseline, 1nazahah Mustafa, 1r.Adollah , 1h. Adilah, 1n.F.Mohd Nasir, "Image Enhancement Techniques Using Local Global, Bright, Dark And Partial Contrast Stretching For Acute Leukemia Images", The World Congress On Engineering 2009 Vol I, pp.1-6.

[2] Aimi Salihah, A.N., M.Y.Mashor, Nor Hazlyna Harun "Colour Image Enhancement Techniques for Acute Leukemia Blood Cell Morphological Features", IEEE 2010 pp.3677-3682.

[3] R. Ravindraiah, Dr.M.n. Giri Prasad ,M.Venkata Srinu **"**Qualitative Evaluation Of Enhancement Methods For Analysis Of Acute Leukemia Images" International Journal of Engineering Science and Technology (IJEST),pp.6447-6453.

[4] Aimi Salihah, A.N., M.Y.Mashor, Azian Azamimi Abdullah ,"Improving Blast Segmentation of Acute Myelogenous Leukemia (AML) Images Using Bright Stretching Technique" , Proceedings of the International Postgraduate Conference on Engineering (IPCE 2010) 16 - 17 October 2010, Perlis, Malaysia, pp.16-21.

[5] N.H.Abd Halim*1, M.Y.Mashor*2, A.S.Abdul Nasir*3,N.R.Mokhtar*4, H.Rosline#5," Nucleus Segmentation Technique for Acute Leukemia", 2011 IEEE 7th International Colloquium on Signal Processing and its Applications,pp.192-197.

[6] Jaspreet Kaur, Amita Choudhary," Comparison of Several Contrast Stretching Techniques on Acute Leukemia Images", International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 1, July 2012,pp.332-335.

[7] Rohini Patil, Prof. Mandar Sohani, Prof.Sachin Bojewar," Acute Leukemia blast counting using RGB, HIS color spaces", pp.1-6.

[8] Aimi Salihah, A.N. #1, M.Y.Mashor#2, Nor Hazlyna Harun#3, Azian Azamimi Abdullah#4, H.Rosline*," Improving Colour Image Segmentation on Acute Myelogenous Leukaemia Images Using Contrast Enhancement Techniques", 2010 IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES 2010), Kuala Lumpur, Malaysia, 30th November - 2nd December 2010.pp 246-251

[9] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," IEEE Transactions on Communication., vol. COM-28, pp. 85-94, Jan., 1980.

[10] R. M. Gray, "Vector quantization," *IEEE ASSP Magazine*, vol. 1, pp. 4-29, 1984.

[11] J. Zheng and M. Hu, "An anomaly intrusion detection system based on vector quantization," IEICE Trans. Inf. Syst., vol. E89-D, no. 1, pp.: 201–210, Jan. 2006

[12] 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: http://www.waset.org/ijcs.pp.: 357-362, 23-24 Jan 2009, Fr. CRCE, Mumbai. Available on ACM portal.

[13] H. B. Kekre, Tanuja K. Sarode, Bhakti Raul, "Color Image Segmentation using Vector Quantization Techniques Based on Energy Ordering Concept" International Journal of Computing Science and Communication Technologies (IJCSCT) Volume 1, Issue 2, pp: 164-171, January 2009.

[14] H. B. Kekre, Tanuja K. Sarode, "Speech Data Compression using Vector Quantization", WASET International Journal of Computer and Information Science and Engineering (IJCISE), vol. 2, No. 4, pp.: 251-254, Fall 2008.

[15] H. B. Kekre, Tanuja K. Sarode, "Multilevel Vector Quantization Method for Codebook Generation", International Journal of Engineering Research and Industrial Applications (IJERIA), Volume 2, No. V, 2009, ISSN 0974- 1518, pp.: 217-231..

[16] Dr.H.B.Kekre1, Saylee M. Gharge2 and Tanuja K. Sarode3," Tumor Demarcation in Mammography Images using LBG on Probability Image", International Journal of Computer Applications (0975 – 8887) Volume 3 – No.8, June 2010,pp 47-53.

[17] Dr. H.B.Kekre, Sudeep D. Thepade, Tanuja K. Sarode and Vashali Suryawanshi," Image Retrieval using Texture Features extracted from GLCM, LBG and KPE", International Journal of Computer Theory and Engineering, Vol. 2, No. 5, October, 2010 1793-8201.

[18] Dr.H.B.Kekre, Tanuja K. Sarode," New Clustering Algorithm for Vector Quantization using Rotation of Error Vector", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 3, 2010 ,pp 159-165.

[19] Dr.H.B.Kekre1, Saylee M. Gharge2 and Tanuja K. Sarode3," Image Segmentation of Mammographic Images Using Kekre'S Proportionate Error Technique on Probability Images", International Journal of Computer and Electrical Engineering, Vol.2, No.6, December, 2010,1793-8163.

[20] Ruggero Donida Labati, Vincenzo Piuri, Fabio Scotti," All-Idb: The Acute Lymphoblastic Leukemia Image Database For Image Processing", Università degli Studi di Milano, Department of Information Technology, via Bramante 65, 26013 Crema, Italy,pp 2089-2092.

[21] Ramin Soltanzadeh,1 Hossein Rabbani,1 and Ardeshir Talebi2," Extraction of Nucleolus Candidate Zone in White Blood Cells of Peripheral Blood Smear Images Using Curvelet Transform " Volume 2012, Article ID 574184, 12,pp 1-12.

[22] Abdul Nasir, A. S.1, Mustafa, N.2, Mohd Nasir, N. F.3" Application of Thresholding Technique in Determining Ratio of Blood Cells for Leukemia detection", 11 – 13 October 2009, Batu Ferringhi, Penang, MALAYSIA,pp.2A1-2A6.