Leukemia Image Segmentation using K-Means Clustering and HSI Color Image Segmentation

Mashiat Fatma
Department of CSE
Amity University, Noida,
Uttar Pradesh

Jaya Sharma
Department of CSE
Amity University, Noida,
Uttar Pradesh

ABSTRACT

During the unfolding measures that are taken for the purpose of leukemia detection, segmentation of blood cells is a vital step. In this paper two approaches of such segmentation technique is proposed. While one uses K-means clustering, other uses color image based segmentation method. Both the processes segment the image into two regions, blasts & backgrounds. These blasts are our area of interest. The performance measure is based on the comparison of the two proposed techniques tends to find the more suitable approach for correct leukemia image segmentation. The results show that the segmentation based on K-means clustering gives better results preserving important information and removing background noise.

General Terms

Medical Image Processing, Segmentation

Keywords

Leukemia, k-means clustering, blasts, HSI color segmentation

1. INTRODUCTION

One of the leading causes of deaths in the world is leukemia, which in general terms is known as blood cancer. An estimated 350000 people are diagnosed with leukemia every year with a death toll of 257000 annually [6]. Leukemia is the primary cause of deaths among several cancer types in children and young adults under the age of 20 years. It equally affects people of every age. Though the ailment is treatable, the correct and in time diagnosis is very important. To diagnose this ailment, the indicative examination and evaluation of the number and maturity of different cells in blood are carried out. All these diagnostic measures taken involve human and are susceptible to unintentional human errors. Therefore need of an automatic leukemia detection technique arises.

1.1 Image Segmentation

Image segmentation is the process of dividing the image into parts. The process of segmentation involves the severance of areas that hold similar characteristics. In the field of image processing, color plays two important roles. First, color can easily be used for simplifying the process of object identification and object extraction from a scene. Second, human can recognize thousands of color shades and intensities. Various diverse segmentation techniques have been suggested and put forward for blood cells recognition of leukemia images. But there is no single method that is good for all kinds of images. Similarly all the methods are not equally good for a single type of particular image.

1.2 HSI Color Model

The Hue Saturation Intensity model roughly known as HSI color model is very much close to what we call as human ocular recognition. This is why it is better to use HSI rather than RGB because the latter model does not closely relate to human visual perception of color [1]. The HSI color model decouples the intensity component from the color carrying information (hue and saturation) in a color image. Since the hue, saturation and intensity are independent of one another, each of this color space can be processed separately without worrying about the correlation between them.

1.3 K-Means Clustering

This is the most popular unsupervised learning algorithm. K-means clustering is known to be one of the easiest and elementary algorithms that are used to answer and unravel the very popular clustering problems. It conforms to a much uncomplicated procedure in which a given data set is classified abiding some common methods and then is associated to nearest centroid. The process involving K means algorithm is a reiterative technique which outputs k clusters by dividing an input image. The algorithm partitions n objects into k clusters where k < n. The k means clustering was invented in 1956. This algorithm is composed of following steps [2].

a) Select randomly k points into the space represented by pixels that are being clustered.

b) Assign each data to the nearest center.

c) When all data have been assigned, recalculate the new center position.

Repeat step 2 and 3 until the centers are no longer on move. This will result in a group of separated objects from where the criterion for measurement (that is to be minimized) can be calculated.

1.4 Thresholding

Thresholding is one method that can be used for the segmentation purpose. In thresholding-based segmentation, the image histogram can be partitioned into two classes using a single value (called bi-level thresholding) or multiple classes using multiple values(called multi-level thresholding) based on the characteristics of the histogram[3]. In the bi-level thresholding, pixels with intensity values less than the threshold are set as background while others are set as object. In multiple-level thresholding, pixels with the intensity values between two successive thresholds are assigned as a class. Theoretically, the levels of thresholding can be increased infinitely according to the number of objects present in images, but the computation load will also be increased exponentially.
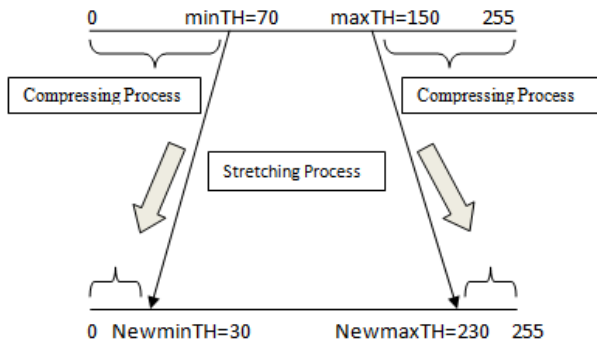
Figure 1: Thresholding **[4]**

The value of 70 and 150 is used as an example of lower Threshold (minTH) and upper threshold (maxTH) values. Threshold value of 30 (NewminTH) and 230 (NewmaxTH) are estimates set as the expected ambit or band of the color level for the output image.

2. METHODOLOGY
In the proposed paper, two approaches of segmentation are used. In first approach the segmentation is based on K-means clustering while the second approach uses thresholding for the purpose of segmentation. These approaches based on the results are then compared and a performance evaluation is carried out.

2.1 Image Acquisition
The first step is to acquire an image of leukemia blood sample. High quality blood samples were acquired from internet.

2.2 Contrast Enhancement Using Linear Contrast Technique
Before performing any major operations on an image, some preprocessing techniques are applied onto the image for enhancing the quality of image and removing the noise and other unnecessary information from the image. Following preprocessing techniques were applied in this process.

This preprocessing is carried out so that the contrast level and brightness level of the image can be enhanced. It is based on the formula given below.

$$outRGB\ (x,y) = 255 * \left[\frac{inRGB(x,y) - aRGB}{bRGB - aRGB}\right] \qquad [2]$$

Where,

inRGB (x, y)   : The original RGB value of the pixel

outRGB (x, y)   : The new RGB value of the pixel

aRGB   : Minimum RGB value
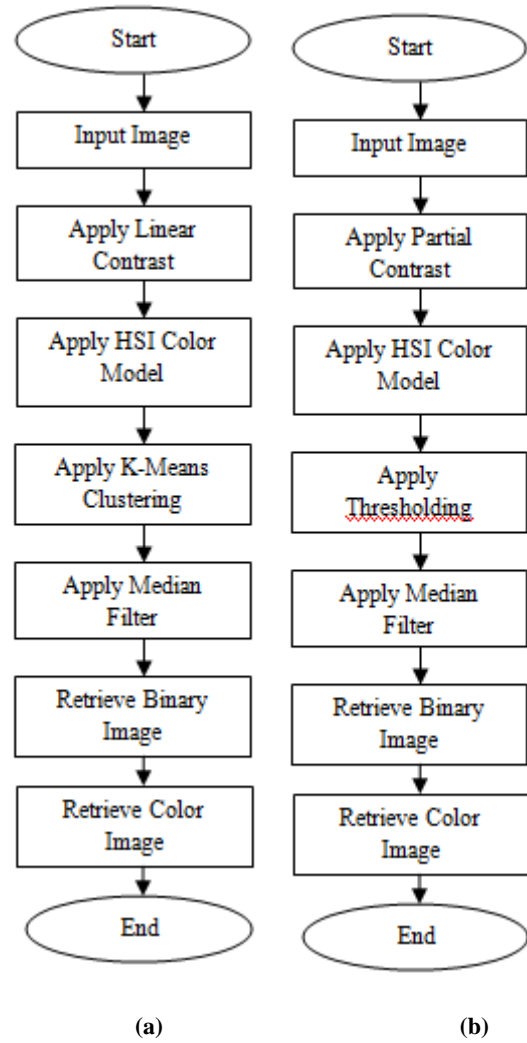
bRGB   : Maximum RGB value

(a)   **(b)**

Figure 2: Segmentation Using
(a) K-means Clustering
(b) Segmentation using Thresholding

2.3 Segmentation Using HSI Color Model & K-Means Clustering
The RGB image is converted to HSI image. Hue and Saturation component of HSI image can be calculated using following formulae

$$Hue = \begin{cases} \theta & if\ B \leq G \\ 360° - \theta & if\ B > G \end{cases} \qquad [3]$$

$$\theta = cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{\frac{1}{2}}} \right\} \qquad [2]$$

Here, the k-means clustering has been used for image segmentation. K-means clustering is based on minimizing the objective function, J as in [5]

$$J = \sum_{i=1}^{n} \sum_{j=0}^{k} \left\| x_i - c_j \right\|^2 \qquad [5]$$

Where n is the number of data, k is the number of cluster, x_i is the i-th sample and c_j is the j-th centre of the cluster.

3. RESULTS & DISCUSSIONS

The following figures show the results obtained after applying the combination of linear contrast technique and color segmentation using HSI color model and K-Means clustering.

(a)

(b)

(c)

(d)

(e)

(f)

(g)
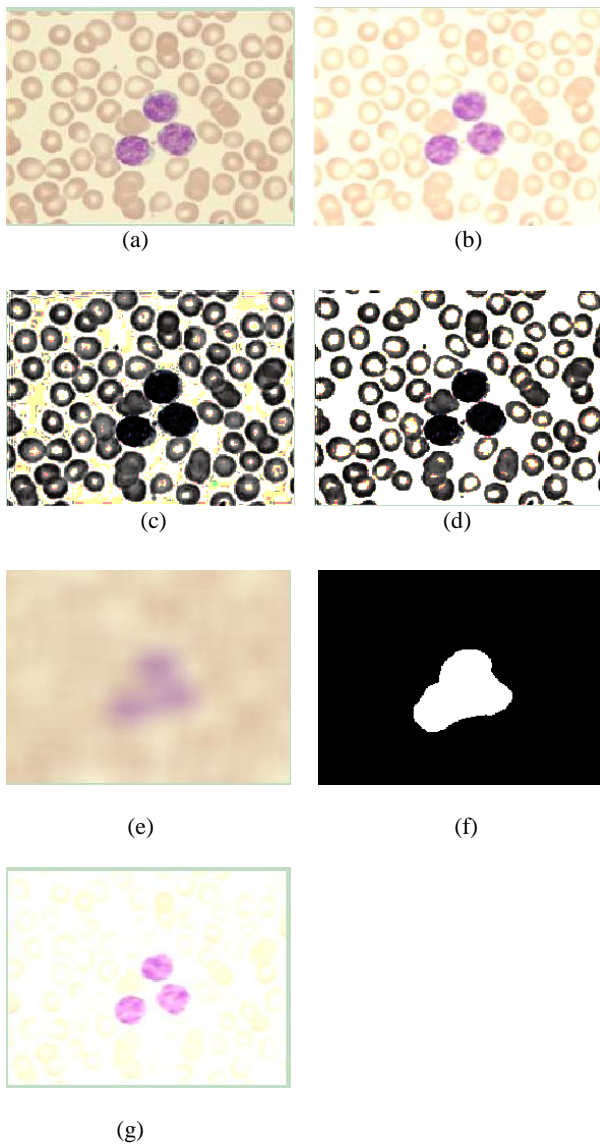
Figure 3: (a) Original image and resultant image after applying (b) Linear Contrast, (c) HSI Color Model, (d) K-Means Clustering, (e) Median Filter, (f) Binary image and (g) Final segmented image

Following figures show the results obtained after applying the combination of linear contrast technique and color segmentation using HSI color model and K-Means clustering.

(a)

(b)

(b)

(d)

(e)

(f)

(g)
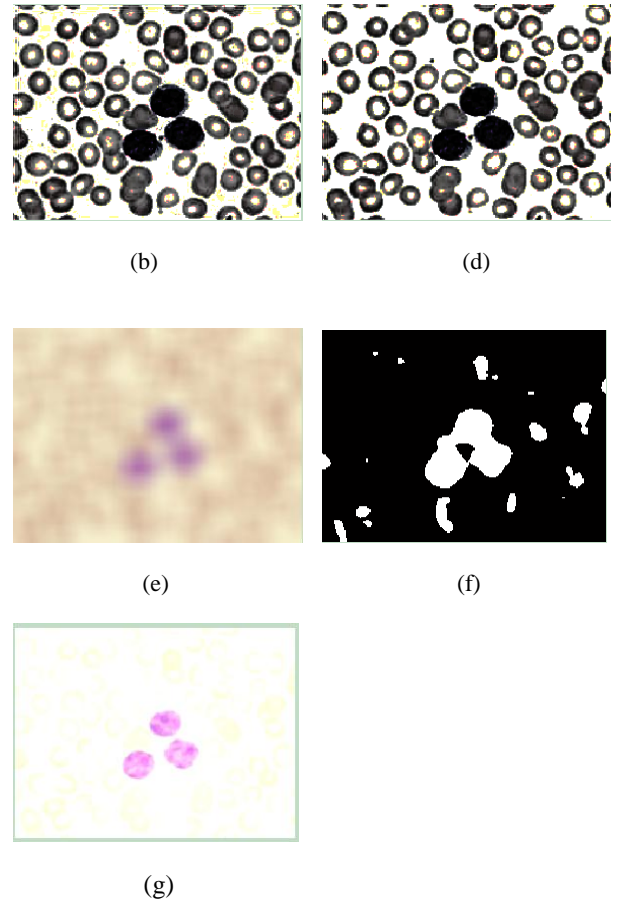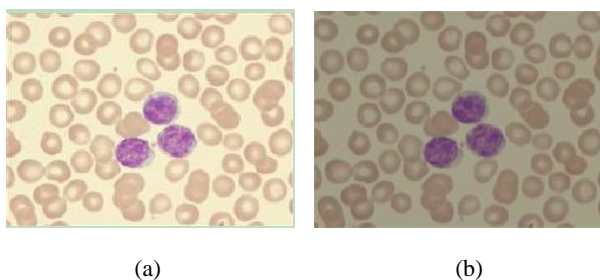
Figure 4: (a) original Image and resultant image after applying (b) Partial Contrast Stretching, (c) HSI Color Model, (d) Thresholding, (e) Median Filter, (f) Binary Image and (g) Final Segmented Image

4. PERFORMANCE COMPARISON TABLE

The table given below gives the comparison of both the algorithms applied on a single image. The pixel values, one in each of the three blasts are taken as input. The corresponding RGB values of the original image, the resultant image after segmentation based on k-means clustering and thresholding respectively are mentioned in the table. From the table we can say that the algorithm that uses k-means clustering gives better results.

Table1: The selected pixel value for segmented images

Values For BLAST 1			
Pixel Value	Original Image (RGB Value)	Using K-Means Clustering(RGB)	Using Thresholding(HSI)
900, 205	146, 82, 158	170,115, 220	224, 124, 236
905,100	159,49,346	190,100,225	235,154,268
Values For BLAST 2			
Pixel Value	Original Image (RGB Value)	Using K-Means Clustering(RGB)	Using Thresholding(HSI)

861, 267	177, 119,167	215,167,209	254, 184, 247
888,295	192,125,182	207,135,169	276,189,258
Values For BLAST 2			
Pixel Value	**Original Image (RGB Value)**	**Using K-Means Clustering(RGB)**	**Using Thresholding(HSI)**
922, 252	204, 149,191	240, 182,243	255,210, 255
939,248	219,158,181	250,179,231	269,218,249

5. CONCLUSION

The results show that both the algorithms i.e. unsupervised segmentation using k-means clustering and segmentation using thresholding were implemented successfully. The significant changes between segmented WBC (White Blood Cell) in and background in leukemia images can easily be seen. The fully segmented WBC (White Blood Cell) is achieved by application of both the algorithms. But better results were obtained when K-Means Clustering was used.

6. FUTURE WORK

The algorithm can be tested in future for further enhancement by application of neural networks for improving the performance. Also other image preprocessing techniques can be applied to improve results. Other segmentation technique can also be tested in order to improvise the segmentation results.

7. REFERENCES

[1] Arthur R. Weeks**,** G. Eric Hague, "Color segmentation in the HSI color space using k-means algorithm", Univ. of Central Florida (USA).

[2] A. S. Abdul Nasir, M. Y. Mashor, H. Rosline, "Unsupervised Color Segmentation of White Blood Cell for Acute Leukemia Images".

[3] R. C. Gonzalez, R .E. Woods, "Digital Image Processing", 3rd Ed, Pearson Prentice Hall, 2008.

[4] E. U. Francis, M. Y. Mashor, R. Hassan, A. A. Abdullah, "Screening of Bone Marrow slide Images for Leukemia using Multilayer Perceptron(MLP)", 2011 IEEE Symposium on Industrial Electronics and Applications (ISIEA2011), September 25-28, 2011, Langkawi, Malaysia.

[5] A. N. A. Salihah, M. Y. Mashor, N .H. Harun, A. A. Abdullah, and H. Rosline, "Improving Colour Image Segmentation On Acute Myelogenous Leukemia Images Using Contrast Enhancement Techniques" in 2010 IEEE EMBS Conference on Biomedical Engineering & Sciences(IECBES 2010), Kuala Lumpur, Malaysia, 2010.

[6] www.asuragen.com/Diagnostics/US/Educational-page.

[7] "Cancer, Leukemia". Retrieved September, 2013, from, http://www.thesurvivorsclub.org/health/surviving-cancer/

[8] "Statistics, Disease Information". Retrieved September, 2013, from, http://www.leukemia-research.org/statistics.

[9] "Leukemia-Topic Overview". Retrieved August, 2013, http://www.webmed.com/cancer/tc/leukemia-topic-overview.