Colour Based Segmentation of Red Blood Cells using K-means and Image Morphological Operations.

Faten Abushmmala[#1], Mohammed Alhanjouri [#2]

[#1#2]*Computer Department, Islamic University*
[1]faten.fffa@gmail.com
[2]malhanjouri@iugaza.edu.ps

Abstract

This paper is a practical application of colour based segmentation on medical imaging. This application is Red Blood Cells (RBC) images to obtained specific results, that accomplished using digital image morphological operations and K-means algorithm. This paper main goal not the segmentation process itself but its modified to isolate the Red Blood Cells alone. Images of Red Blood Cells films were captured and we obtained 100 images. Theses images used to test our methodology. After using k-means for the segmentation process; various image morphological operations were applied to reach the desired goal as intended. The segmentation using K-means applied in two spaces RGB and LAB space. Better outcome used to apply image morphological operations to enhance the separated Red Blood Cells. In RGB Space the success rate were 98.2% and in LAB space 88.9%.

Keywords: RBC segmentation, K-means segmentation, image processing.

I. INTRODUCTION

Segmentation of medical images is extremely challenging due to poor image contrast that result in missing or diffuse boundaries. Consequently, this task involves incorporating as much prior information as possible (e.g., texture, shape, and spatial location of organs) into a single framework. Color based segmentation using K-means were suggested in [1]. Several automated methods have been developed to process the acquired images and identify features of interest [2], including intensity-based methods, region-growing methods and deformable contour models. Intensity-based methods identify local features such as edges and texture in order to extract regions of interest. However, due to the low contrast information in medical images, an effective segmentation often requires extraction of a combination of features such as shape and texture or pixel intensity and shape. In this paper the color of red blood cells was our way to construct the segmentation as best as we can. For that this paper investigate the color space to compare which will give better results for the Red Blood Cells (RBC). Image spaces discussed in this paper are LAB space and RGB space.

A. LAB Space

LAB* color space** in fig 1 is a color-opponent space with dimension ***L for lightness and ***A*** and ***B*** for the color-opponent dimensions, a space which can be computed via simple formulas from the *XYZ* space, but is more perceptually uniform than *XYZ* [3]. *Perceptually uniform* means that a change of the same amount in a color value should produce a

change of about the same visual importance. When storing colors in limited precision values.

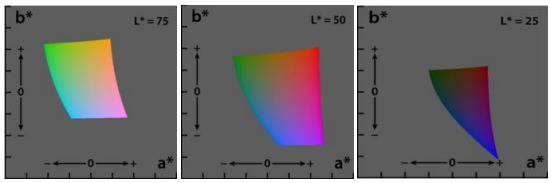

Fig1. LAB Space

Unlike the RGB, LAB color is designed to approximate human vision. It aspires to perceptual uniformity, and its L component closely matches human perception of lightness. It can thus be used to make accurate color balance corrections by modifying output curves in the A and B components, or to adjust the lightness contrast using the L component which model the output of physical devices rather than human visual perception, these transformations can only be done with the help of appropriate blend modes in the editing application.

B. RGB Space.

The **RGB color model** in fig 2 is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green, and blue. The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human perception of colors.

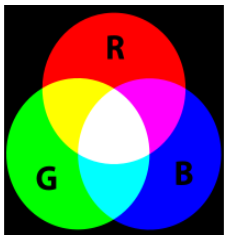

Fig2. RGB space.

In RGB space the features are r, g, b, x, y, z where r, g and b values along with coordination play strong part in clusters determination, but in LAB space the location (x ,y) does not matter.

C. K-means

The K-means algorithm, first developed four decades ago [4], is one of the most popular centre-based algorithms that attempts to find K clusters which minimize the mean squared quantization error , MSQE. The algorithm tries to locate K prototypes (centroids) throughout a data set in such a way that the K prototypes in some way best represent the data. A summarization of the K-Means algorithm through the following steps [5]:

1. Initialization

a) Define the number of prototypes (K).

b) Designate a prototype (a vector quantity that is of the same dimensionality as the data) for each cluster.

2. Assign each data point to the closest prototype. That data point is now a member of the class identified by that prototype.

3. Calculate the new position for each prototype (by calculating the mean of all the members of that class).

4. Observe the new prototypes' positions. If these values have not significantly changed over a certain

number of iterations, exit the algorithm. If they have, go back to step 2.

The main problem of the K-means algorithm [5] is its dependency on the prototypes' initialization. If the initial prototypes are not chosen carefully the computation will run the chance of converging to a local minimum rather than the global minimum solution. Thus initializing prototypes appropriately can have a big effect on K-Means. The performance function for K-Means may be written as

$$J_{Km} = \sum_{i=1}^{N} min \ \|x_i - m_j\|^2 \qquad 1$$

D. Image morphological Operations

Mathematical Morphology is one of the most productive areas in image processing. The motivation comes from the collection of structural information about the image domain. The content of mathematical morphology is completely based on set theory. By using set operations there are many useful operators defined in mathematical morphology. For instance erosion, dilation, opening and closing are these kind of operations which are beneficial when dealing with the numerous image processing problems. Sets in mathematical morphology represent objects in an image. For example when we talk about all the black or white pixels as a set in a binary image we completely mean a morphological description of the image. In a binary image, the sets are the members of the 2-D image domain with their integer elements. Each element in the image is represented by a 2-D tuple whose elements are x and y coordinate. These coordinates show the resultant function of a light perceptive function from a digital sensor. Gray scale images can be represented in ZxZxZ domain too. ZxZxZ is a 3-D domain that includes X and Y spatial coordinate with a 3rd dimension Z coordinate (X=width, Y=length and Z=height). Sets in higher dimensional spaces can contain other characteristics of image such as color and time varying components. Mathematical morphological operations are based on simply transforming an input image with a specific structural element

Some of the morphological operations [6] defined on binary images are translation, reflection, complement, difference, dilation, erosion , opening, closing, hit-or-miss, transform, boundry extraction, region filling, connected components, convex hull, thinning, thickening, skeletons and prunning.. In his paper opening and closing [6] op:

Dilation is one of the operators in the area of mathematical morphology, the other being erosion. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels . Thus areas of foreground pixels grow in size while holes within those regions become smaller. In dilation we increase the white pixel in the image making, it look broader. Every background pixel that is touching an object pixel is changed into an object pixel.

1. OPENING

In morphology, opening is the dilation of the erosion of a set A by a structuring element B:

$$A \circ B = (A \ominus B) \oplus B, \qquad 2$$

where and denote erosion and dilation, respectively. Together with closing, the opening serves in computer vision and image processing as a basic workhorse of morphological noise removal. Opening removes small objects from the foreground (usually

taken as the dark pixels) of an image, placing them in the background, while closing removes small holes in the foreground, changing small islands of background into foreground. These techniques can also be used to find specific shapes in an image. Opening can be used to find things into which a specific structuring element can fit.

2. CLOSING

In mathematical morphology, the **closing** of a set (binary image) A by a structuring element B is the erosion of the dilation of that set,

$$A \cdot B = (A \oplus B) \ominus B, \qquad 3$$

where and denote the dilation and erosion, respectively. In image processing, closing is, together with opening, the basic workhorse of morphological noise removal. Opening removes small objects, while closing removes small holes.

II. METHODOLOGY

We use color based segmentation to separate red blood cells from other cells and from the background. On our study we have 100 images of blood films, these images prepared using Wight stains. K-means used to cluster the blood films images into 3 cluster, first cluster is red blood cells, second cluster is the background and third cluster is for other blood cells, repeating clustering three times to avoid local minima. The result from clustering the blood films images in RGB space (fig 3) ; each color represents different cluster alone, nuclei of cells and the background in one cluster not quite the wanted result. Each colors in fig 3 (a) represents a different cluster.

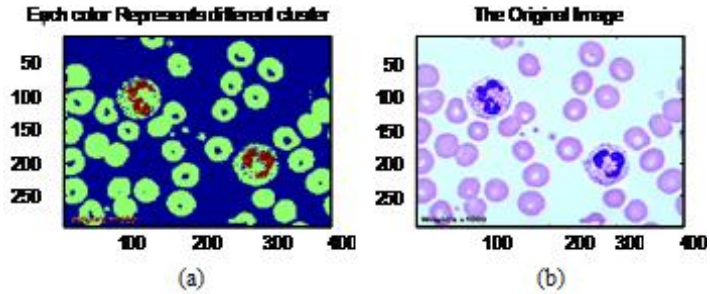

Fig. 3.Clustering using K-means in RGB space, Fig. 3. (a) the clustering result, Fig. 3.(b) the original image.

The features chosen to be fed into K-means (8 features) are:

R: red value for each pixel.

B: blue value for each pixel.

G: green value for each pixel.

x: x- coordinate for each pixels.

y: y- coordinate for each pixels

I: intensity value for each pixel.

S: saturation value for each pixel.

H: hue value for each pixel.

Using these equations:

$$I = \frac{1}{3}(R + G + B), \qquad 3$$

$S = 1 - \left(\frac{3}{(R+G+B)}\right) \times a$, Where **a** is the minimum of R, G and B

$$H = \cos^{-1}\left(\frac{0.5\times((R-G)+(R-B))}{((R-G)^2+(R-B)\times(G-B)^{0.5})}\right) \qquad 4$$

Using same clustering techniques (K-means) and the same features for segmenting/clustering, but in the LAB space, each cluster in a separate image alone is shown in fig 4..

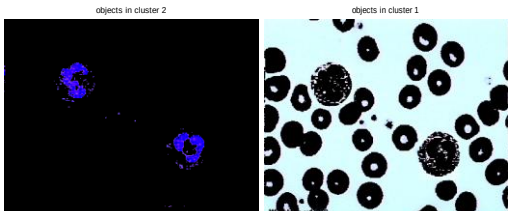

Fig. 4.Clustering of foreground and using K-means in LAB space.

In fig 4, the white cells (large cells) still segmented with the red blood cells, this means that post processing operation required; the result form the clustering in the RGB space is much better, since colors manipulation operation are more flexible. In segmenting (clustering) in the LAB space the image entirely fed into K-means; only pink cells must be in one cluster along with their centers (nuclei).

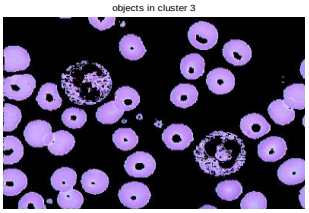

Fig. 5. Third cluster red cells.

After segmenting/clustering the blood films images using K-means into 3 clusters, first cluster is Pink objects (red blood cells), second cluster is the background and third cluster is for the other cells (in RGB space), repeating the clustering process three times to avoid local minima. Converting fig 5 into binary image is presented in fig 6.

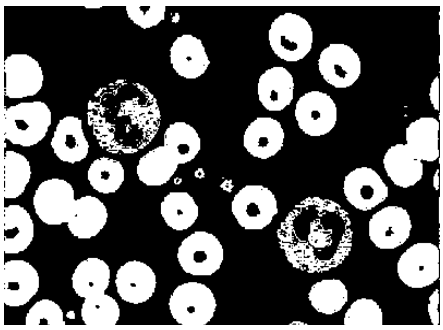

Fig.6. . Binary image of figure 3

For the binary image, some morphological operations performed to clean the image from unwanted parts (such as the fractions of white cells). First: opening operation with disk of five pixels radius is done (notice here that the images of the blood films are of the same size).

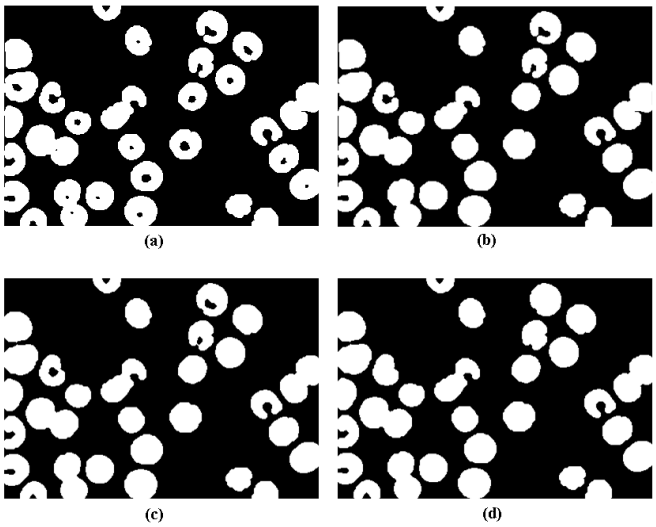

Fig. 7. Some morphological operations: (a) Binary image after opening morphological operation (R=5 pixels). (b) Image filling morphological operation (holes). (c) Opening morphological operation (R=2 pixels). (d) Closing morphological operation (holes).

The morphological operations that used in fig 6 are necessary to separate the red blood cells alone, these red blood cells is cleaned and ready to be used.

Previous phase results gave each red blood cell alone as, but the produced binary image lost a lot of it is properties due to enhancement and morphological operations usage. To solve this problem, we located the centroid for each object in the binary image in hand and convert the **original image (the first blood films images)** to binary image, and use the centroids as flags to locate the wanted objects as shown in fig 8.

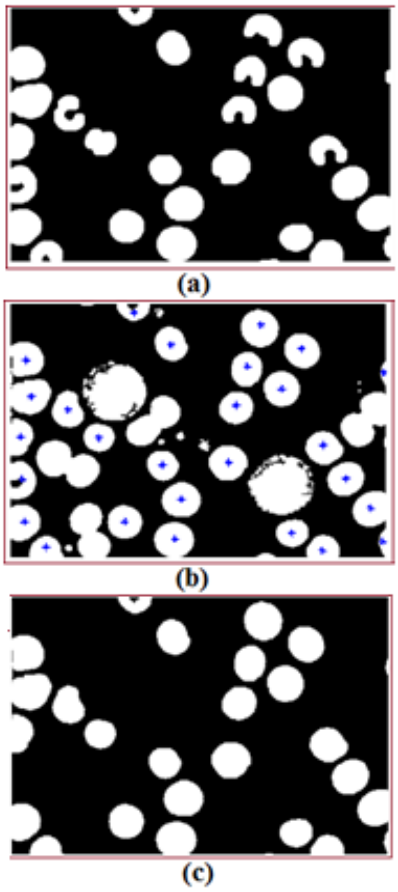

Fig. 8. Cell Segmentation (a) The binary image (non-overlapped cells). (b) The centroids of cells in binary image as flag in the original binary image. (c) The resultant image after flagging operation

This can be summaries in the following steps:

- Resize images to one size (Average size) among our images was 1936 pixels.
- Clustering the image based on color. (Features were: red, green, blue, the coordination, saturation, intensity and hue), we are only concerned with the pink objects (the red blood cells).
- Filtering the pink cluster with smoothing filter then convert it into binary image.
- Applying opening morphological operation (R=5 pixels) on the binary image.
- Applying image filling morphological operation (holes) on the result from the third step.
- Applying opening morphological operation (R=2 pixels) on the result from the fifth step.
- Applying closing morphological operation (holes) on the result from the sixth step.
- The object exist in result of sixth step is the only object we concerned in, the centroids of each object in the binary image from the sixth step made as flags to the binary image of the original image, mapping the original image by centroids (flags) to locate object needed with all their details.

III. RESULT

As we mentioned before, we have 100 red blood cell images. Each image contains max 10 cells of RBC. Our methodology applied on each image at a time and a ratio of the number of the isolated cells to the actual number of cells.

$$Ratio = \frac{I_{RBC}}{A_{RBC}} \qquad 5$$

Then an average taken for all these ratios, all that done on both spaces (RGB and LAB). the average ratio for RGB space is 98.2% and for LAB space is 88.9%. this result considered fairly good. there is a big chance if other types of stains used in making the blood films the result will be different, lighting also considered huge factor may influence the results. For these reasons all these image captured together in the same time and with the same camera and microscope.

IV. CONCLUSION

In this paper we used red blood cells images to apply a segmentation process to isolate the RBC's from the images, we applied K-means on the images using distinctive features in two spaces, from the results the RGB space gave more appropriate results on all samples we have. Be aware that these images of blood films captured under microscope and the stained used to prepare the blood films are Wight stain to obtained such range of colors. The success rate of isolating the wanted red blood cells in RGB space was 98.2% while LAB space was 88.9%..

REFERENCES

[1] Faten F. AbuShmmala and Wesam M. Ashour, "Color Based Image Segmentation using Different Versions of K-Means in two Spaces", Global Advanced Research Journal of Engineering, Technology and Innovation (GARJETI), 1(1):030-041, 2013. ISSN: 2315-5124**.**

[2] D. L. Pham, C. Xu, J. L, Prince, "Survey of current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, Vol. 2, pp. 315-337, 2000.

[3] J.S.Marques, N.P.Blanca and P.Pina, "Pattern Recognition and Image Analysis" Singh, S.; Singh, M.; Apte, C.; Perner, P. Estoril, Portugal, (2005), ISBN: 978-3-540-28833-6.

[4] 6 J. MacQueen. "Some methods for classification and analysis of multivariate observations". Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, pages 281–297, 1967

[5] W. Barbakh, Ying Wu, and Colin Fyfe. "Non-standard parameter adaptation for exploratory data analysis" (2009), University of the west of Scotland, Scotland ISBN: 978-3-642-04004-7

[6] R. C. Gonzalez and R. E. Woods " Digital Image Processing", Prentice Hall, Upper Saddle River, New Jersey 07458.2002, ISBN: 0-201-18075-8.