IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 2, Apr-May, 2014
ISSN: 2320 – 8791 (Impact Factor: 1.479)
www.ijreat.org

An Approach for Balancing Intensity and Color Discrepancy of Histopathology Images for Accurate Clustering of Tissue Cells for Significant Analysis of Bone Cancer

Vandana B.S[1], Antony P.J[2], Sathyavathi R Alva[3]

[1, 2]Department of Computer Science and Engineering, K V G College of Engineering, Sullia, VTU, India

[3]Department of Pathology, KVG Medical College, Sullia, RGUHS, India

Abstract

the method presents an algorithm for segmenting cells in a microscopic image of bone cancer stained with H&E. Unlike other biopsy, bone tissue samples usually contain different types of cells these include normal cells, blood cells, cancer and fracture cells. The procedure for the approach consists of detecting and classifying cells based on the properties studied .Noise reduction is one of the relevant approach that are used to eliminate the stain artifacts. Median filter is one of the preprocessing steps that is adapted for RGB component to reduce noise, but gives the blur effect. In order to enhance the contrast, without affect to the color information algorithm simply call for converting the image to HIS and the approach endure with histogram equalization, processing only intensity component. To give improved input for the segmentation, algorithm encompasses color normalization of the image that deals with the ambiguous color occurred because of satin artifacts and poor illumination. *K*-means clustering algorithm is applied to segment RGB information. In this method intensity and color variations of the image is work out without accord of computational speed, and the result shows that cancer cells in the background with heterogeneous intensities and colors are properly segmented from the chondrosarcoma image dataset.

Keywords— *bone cancer, Chondrosarcoma, Color - normalization, segmentation*.

1. Introduction

Histopathology is the microscopic examination of biopsy to study the tissue structure, and related disease by the pathologist. A biopsy is the process in which a tissue sample is removed from a part of the body and examined under the microscope to identify for abnormal cells. The slides are stained with standard techniques such as haematoxylin and eosin to demonstrate the bone structure and cellular components. Digital histopathology is the recent development, which is the practice of converting glass slides into digital slides. These slides provide an exhaustive site of disease and its effect on tissues, since the preparation process preserves the underlying tissue architecture. Advances in digital histopathology and image analysis show a great amount of interest in designing efficient automated tool for extracting features for cancer diagnosis. A bone cancer is a neoplastic growth of tissue in bone. Bones and cartilages are types of connective tissue in the body. Bone tissue is made up of osteoblast, osteocytes and osteoclasts.Cartilage comprises of chondroblasts, chondrocytes.Abnormal growths found in the bone can be either noncancerous or cancerous. Common types of primary bone cancer include the Osteosarcoma, Chondrosarcoma, Ewing sarcoma. Osteosarcoma, which arises from osteoid tissue in the bone, Chondrosarcoma, which begins in cartilaginous tissue. The Ewing Sarcoma that arises from elements of primitive nerve tissue in the bone or soft tissue. The study of histocytological feature of bone cancer is necessary to design efficient automated system.

1.1 Related work

A good quality images are prerequisites for any image analysis algorithm. Noise reduction and color normalization are the crucial steps, which improves the histopathological image features against the staining process. Various methods have been presented for the preprocessing, which includes Median filter, thresholding, stain space normalization. Median filter and thresholding technique are used to remove stains artifacts from the image [1],[2],[3].

Stain variance in the histopathology creates inconsistency to perform quantitative analysis. Macenko and Niethammer presented Color deconvolution based normalization overcoming many of the known inconsistencies in the staining

process, thereby bringing slides that were processed or stored under very different conditions in a common, normalized space to enable improved quantitative analysis [4],[5].

Numerous works has been conducted on segmentation of various structures in histopathology images of cancer using methodologies such as adaptive thresholding, fuzzy c-means clustering, Bayesian classifier .It is also studied that to segment cells morphological operations are suitable for uniform background image, and thresholding is preferable for uneven background image [6],[7],[8]. Shivang showed that a Bayesian classifier is used to detect regions of interest by utilizing low-level image features to find the nuclei, cytoplasm of the tissue and by using this method; they presented automated detection and segmentation of prostate gland regions [9].

Anil Z Chitade, Dr. S.K. Katiyar presented a novel mage segmentation algorithm based on color feature by k-means clustering algorithm for satellite dataset. They showed that it is possible to reduce computational cost avoiding feature calculation for every pixel in the image [10].

2. Proposed work

It is imperative to study the basic bone histology, different bone forming cells, cartilage features, ostoid production, metabolism activity, and some of the common disorder to distinguish normal and cancerous tissue. The proposed work includes designing efficient segmentation algorithm, which includes preprocessing method, improves the image quality through median filter and color normalization. Segmentation to extract all the features of different sarcomas by using *k* means clustering algorithm and classification of these features to define the malignancy. A tool that allows perfect analysis and generates greater confidence in results augments the role of the pathologist. An overview of the proposed approach is given in an algorithm

Input: Microscopic image of bone –Osteosarcoma
Output: Segmented image
Method:
Step 1:Read Image
Step 2:Apply median filter for RGB components
for i = 1:3
Irgb(:,:,ii) = medfilt2(Image(:,:,i),[3 3]);
Step 3:Convert RGB Image into HIS
function rgb_to _hsi

$H=\cos^{-1}\left(\frac{0.5(r-g)+(r-b)}{\sqrt{(r-g)^2+(r-b)\times(g-b)}}\right)$

$I=0.3\times(r+g+b)$

$S=1-\frac{\min(r+g+b)}{I}$

Step 4:Apply histogram equalization to intensity component(I) of HSI image
IE=histeq(I)
Step 5:Convert HIS to RGB
Function his_to_rgb

If $0^\circ \le H < 120^\circ$ *then* $B = I(1-S)$,

$$R = I\left[1+\frac{S\cos H}{\cos(60^\circ - H)}\right], G = 3I-(R+B)$$

If $240^\circ \le H < 360^\circ$ *then* $H = H-240^\circ$, $G = I(1-S)$,

$$B = I\left[1+\frac{S\cos H}{\cos(60^\circ - H)}\right], R = 3I-(G+B)$$

If $120^\circ \le H < 240^\circ$ *then* $H = H-120^\circ$, $R = I(1-S)$,

$$G = I\left[1+\frac{S\cos H}{\cos(60^\circ - H)}\right], B = 3I-(R+G)$$

Step 6:Apply the color normalization to remove any intensity variations
For every pixel in the image
Normalized Red=$\frac{r}{\sqrt{r^2+g^2+b^2}}$,
NormalizedGreen=$\frac{g}{\sqrt{r^2+g^2+b^2}}$,NormalizedBlue= $\frac{b}{\sqrt{r^2+g^2+b^2}}$
normalized _RGB Image=cat3(NormalizedRed,NormalizedGreen,NormalizedBlue)
Step 7:Color based segmentation using K-means clustering
Classify the Colors Using K-Means Clustering
Label Every Pixel
Segment the blue Nuclei
Step 8: Calculate cytoplasm nucleus ratio.

2.1 Preprocessing

The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing. Median filter is very much suitable for histopathology image, to remove noise because of staining process. Main advantage of median filter is, it preserves edges while removing noise. Edges are great important here to detect tissue cells. The algorithm deals with color image, it is required to apply median filter to three-color channel- red, green, blue. Median filtering is one kind of smoothing technique that gives blur effect. In order to enhance the contrast, the proposed algorithm uses histogram equalization method. If histogram equalization

technique is applied on histopathogical bone cancer image, the color component appears somewhat different from the original image. Therefore, algorithm simply calls converting the image to HIS format then histogram equalization technique is applied on intensity component of the image, and converting back to RGB for display. The color component is not affected. The result is presented in the fig 1.

2.2 Converting RGB Values to HSI Values

Suppose R, G, and B are the red, green, and blue values of a color. The HSI intensity is given by the equation

$$I = \frac{(R+G+B)}{3} \quad (2.2\text{-}1)$$

now, let m be the minimum value among R, G, and B. The HSI saturation value of a color is given by the equation

$$S = 1 - \frac{M}{I} \quad \text{if } I > 0, \quad (2.2\text{-}2)$$

Alternatively,

$$S = 0 \text{ if } I = \text{zero}.$$

To convert a color's overall hue, H, to an angle measure, uses the following equations:

$$H = \cos^{-1}[(R - \tfrac{1}{2}G - \tfrac{1}{2}B)/\sqrt{R^2 + G^2 + B^2 - RG - RB - GB}] \quad \text{if } G \geq B, \quad (2.2\text{-}3)$$

Alternatively,

$$H = 360 - \cos^{-1}[(R - \tfrac{1}{2}G - \tfrac{1}{2}B)/\sqrt{R^2 + G^2 + B^2 - RG - RB - GB}] \quad \text{if } B > G,$$

where the inverse cosine output is in degrees.

2.3 Converting colors from HIS to RGB

RG sector: $0° \leq H \leq 120°$

$$R = I\left[1 + \frac{S \sin H}{COS(60°-H)}\right] \quad (2.3\text{-}1)$$

$$G = 3I - (R + G) \quad (2.3\text{-}2)$$

$$B = I(1 - S) \quad (2.3\text{-}4)$$

GB sector: $120° \leq H \leq 240°$

$$H = H - 120° \quad (2.3\text{-}5)$$

$$R = I(1 - S) \quad (2.3\text{-}6)$$

$$G = I\left[1 + \frac{S \sin H}{COS(60°-H)}\right] \quad (2.3\text{-}7)$$

$$B = 3I - (R + G \quad (2.3\text{-}8)$$

BR sector: $240° \leq H \leq 360°$

$$H = H - 240° \quad (2.3\text{-}9)$$

$$R = 3I - (G + B) \quad (2.3\text{-}10)$$

$$G = I(1 - S) \quad (2.3\text{-}11)$$

$$B = I\left[1 + \frac{S \sin H}{COS(60°-H)}\right] \quad (2.3\text{-}12)$$

2.4 Color Normalization

Color normalization is concerned with artificial color vision and object recognition. In general, the distribution of color values of histopathology image depends on the staining process, different lighting conditions and microscope feature. Color also very important factor and input for segmentation algorithm .Color normalization allows for object recognition techniques based on color, to compensate for these variations. Algorithm that works for every pixel in the image is given by

$$\text{Normalized Red} = \frac{r}{\sqrt{r^2+g^2+b^2}}, \quad (2.4.1)$$

$$\text{Normalized Green} = \frac{g}{\sqrt{r^2+g^2+b^2}}, \quad (2.4.2)$$

$$\text{Normalized Blue} = \frac{b}{\sqrt{r^2+g^2+b^2}} \quad (2.4.3)$$

Normalized RGB Image is obtained by concatenation of these three normalized color components. The result of algorithm demonstrated in the fig 2.

2.5 Segmentation

Segmentation is one of the most difficult tasks in image processing. The process is subdivides an image into its constituent parts or objects. The approach here is to partition an image based on the *k*-means clustering algorithm based on color information. *K-means* clustering aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into clusters. Given a set of observations ($\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$), where each observation is a *d*-dimensional real vector, *k*-means clustering aims to partition the *n* observations into *k* sets ($k \leq n$) $\mathbf{S} = \{S_1, S_2, \ldots, S_k\}$ so as to minimize the within-cluster sum of squares .

Given an initial set of *k* means $p_1^{(1)}, \ldots, p_k^{(1)}$,Image with color information.

Step1: Assign each observation to the cluster whose mean yields the least within-cluster sum of squares. Difference between two colors using the Euclidean distance metric.

$$s_i^{(t)} = \{x_p : \|x_p - p_i^{(t)}\|^2 \leq \|x_p - p_j^{(t)}\|^2 \ \forall\, 1 \leq j \leq k\}, \quad (2.5.1)$$

Where each x_j is assigned to exactly ones$^{(t)}$,

Step2: Calculate the new means to be the centroids of the observations in the new clusters.

$$p_i^{(t+1)} = \frac{1}{|s_i^{(t)}|} \sum_{x_{j \in s_i^{(t)}}} x_j \quad (2.5.2)$$

The cluster center value, which contains the mean value for each cluster, to detect blue color nuclei, shown in the fig 3.

Finally calculate the cytoplasm nucleus ratio; malignancy is defined if it is 1:2 ratios, resulted in the fig 4.

3. Result discussion

Advances in digital histopathology and image analysis show a great amount of interest in designing efficient automated tool

for extracting features for cancer diagnosis. There are unique challenges in proposed work due to the variation in bone cancer, complex tissue structure, and their relationship. Each cancer feature set are different from one another, particularly more frequent bone cancer like Osteosarcoma, Chondrosarcoma, and Ewing sarcoma. Data set particularly for this work is Chondrosarcoma. The algorithm successfully segments the tissue cells and gives greater confidence for further step. Good quality images are prerequisites for the any segmentation algorithm. Keeping that in mind proposed work includes preprocessing step that has noise removal, and color normalization steps. The result shows that median filter very well works for histopathology bone image helps in elimination of noise created due to stain artifacts, but produces blur effect. To overcome that algorithm calls contrast enhancement technique. Color is the major feature in histopathology image of bone cancer, preprocessing method includes color normalization that gives better input for color based K-means clustering algorithm. Result shows that 95 % cells are successfully segmented and provided for further analysis.

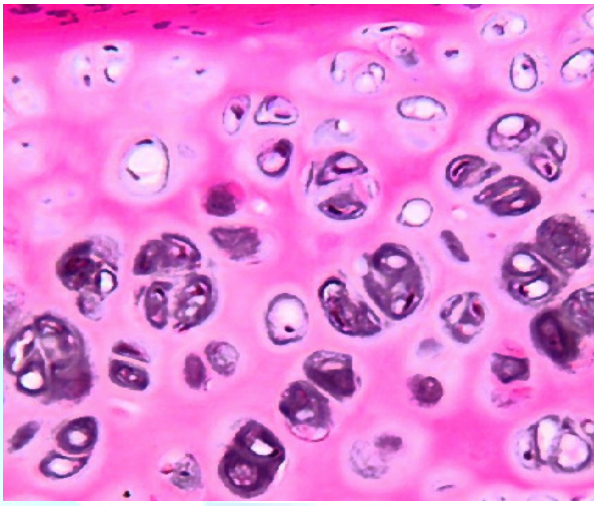

Fig 2 Color Normalization

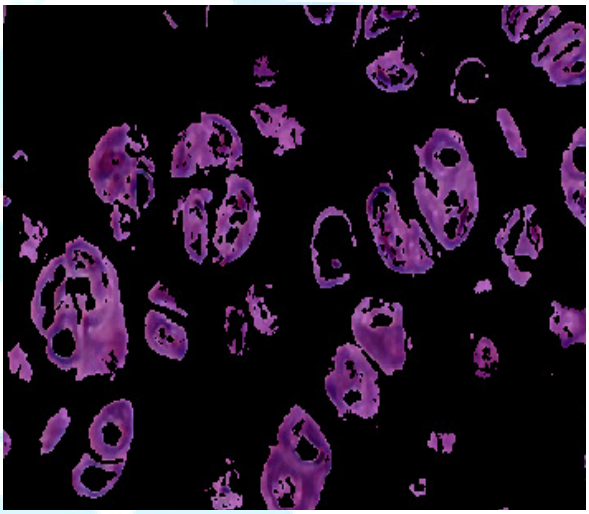

fig 3 result of k-means clustering algorithm

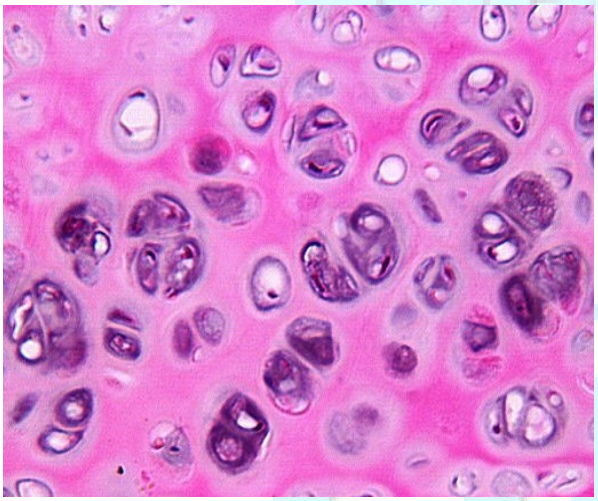

fig 1 Original Image

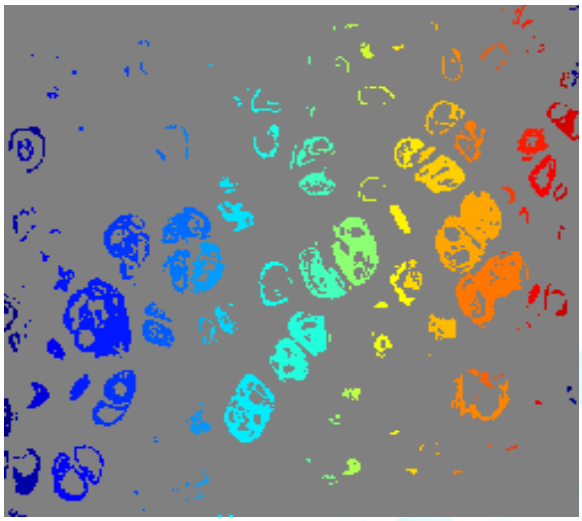

fig 4 labeled regions

4. Conclusion

Identification of various magnitude features of malignant bone cells to be explored in the forth-coming research work towards automation for accuracy. Adjusting the colors of histology slides is not an easy task as we are faced with a great variation not only in staining quality but also in heterogeneous staining of different tissues, and from a great variation of color temperature from the microscope light source. Poor contrasting of the raw images that come from settings of the various microscope cameras on the market further aggravates this. Nevertheless, we do have to normalize these images for to reduce the color variations so that pathologist gets relief from the physical obstacle. We intend to differentiate all the tissue structures from histology bone image so that enhances the reliability in the process.

Multilevel segmentation needed to extract all the features of different sarcomas.

References

[1] Bieniecki, Wojciech. "Oversegmentation avoidance in watershed-based algorithms for color images." Modern Problems of Radio Engineering, Telecommunications and Computer Science, 2004. Proceedings of the International Conference. IEEE, 2004.

[2] Macenko, Marc, et al. "A method for normalizing histology slides for quantitative analysis." Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on. IEEE, 2009.

[3] Chinta, Rambabu, and Martin Wasser. "Three-dimensional segmentation of nuclei and mitotic chromosomes for the study of cell divisions in live Drosophila embryos." *Cytometry Part A* 81.1 (2012): 52-64

[4] Korde, Vrushali R., et al. "Automatic segmentation of cell nuclei in bladder and skin tissue for karyometric analysis." European Conference on Biomedical Optics. Optical Society of America, 2007.

[5] Lopez, Clara M. Mosquera, and Sos Agaian. "Iterative local color normalization using fuzzy image clustering." *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2013.

[6] Phukpattaranont, P., and P. Boonyaphiphat. "Computer-aided analysis of nuclear stained breast cancer cell images." *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on*. Vol. 1. IEEE, 2008.

[7] Zhuang, Hualiang, Kay-Soon Low, and Wei-Yun Yau. "Multichannel Pulse-Coupled-Neural-Network-Based Color Image Segmentation for Object Detection." *Industrial Electronics, IEEE Transactions on* 59.8 (2012): 3299-3308.

[8] Phukpattaranont, Pornchai, and Pleumjit Boonyaphiphat. "An automatic cell counting method for a microscopic tissue image from breast cancer." 3rd Kuala Lumpur International Conference on Biomedical Engineering 2006. Springer Berlin Heidelberg, 2007.

[9] Naik, Shivang, et al. "Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information." *MIAAB Workshop*. 2007.

[10] Chitade, Anil Z., and S. K. Katiyar. "Colour based image segmentation using k-means clustering." *International Journal of Engineering Science and Technology* 2.10 (2010): 5319-5325.