PAPER • OPEN ACCESS

Segmentation of Irrelevant Regions Using Color Thresholding Method: Application in Breast Histopathology Images

View the article online for updates and enhancements.

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

Segmentation of Irrelevant Regions Using Color Thresholding Method: Application in Breast Histopathology Images

Tan Xiao Jian[1], Mustafa Nazahah[1*], Mashor Mohd Yusoff[1], Ab Rahman Khairul Shakir[2]

[1]University Malaysia Perlis, 02600 Arau, Perlis, Malaysia.
[2]Hospital Tuanku Fauziah, Jalan Tun Abdul Razak, 01000 Kangar, Perlis, Malaysia.

*nazahah@unimap.edu.my

Abstract. Segmentation of irrelevant regions in the breast histopathology image is essentially performed in preliminary or post processing stages. This study presents a color thresholding method to segment the irrelevant regions, specifically blood cells and hemorrhage in the RGB breast histopathology images. The conventional gray level global thresholding method was extended to color images by modifying the thresholding algorithm. The color thresholding method was performed by decomposing the RGB input image into coherent regions based on similarity in color feature. Area Overlap Measure (AOM) and Combined Equal Importance (CEI) are two statistical metrics used to evaluate the performance of the color thresholding method. The obtained AOM and CEI for the overall dataset in blood cells are 0.92 and 0.89, respectively. The color thresholding method is found to be effective and is able to produce plausible thresholding results in segmentation of irrelevant regions in RGB breast histopathology images.

1. Introduction

Image segmentation has a paramount role in meaningful analysis and interpretation of acquired image. It is a fundamental step in many video, image, graphical and computer vision related applications where the subsequence steps are strongly rely on the quality of the segmentation [1]. The main purpose of image segmentation is to decompose the input image into spatially coherent regions sharing similar attribute and homogeneous properties.

Over the last decades, many segmentation techniques have been proposed to address the different segmentation problems. However, a universal segmentation approach that is suitable for all applications (in monochrome and color images) is yet to exist [2]. Segmentation can be performed through bi-level thresholding where the input image is decomposed into two homogeneous regions. The thresholding value can be obtained based on the textural feature, intensity value and relative edges [1]. Gray level global thresholding [3, 4], fuzzy analysis based segmentation [5, 6] and landscape analysis [7, 8] are few typical segmentation approaches. Among these approaches, gray level global thresholding is the most commonly used segmentation technique with many related algorithms have been proposed in recent years [1, 9, 10]. Gray level global thresholding approach can be extended to color images by modifying the

thresholding algorithm. The color thresholding defines the regions of interest based on the similarity in color feature (i.e., red (R), green (G) and blue (B) channels).

The main objective of this study is to segment blood cells (Figure 1 (a)) and hemorrhage (Figure 1 (b)) from the background areas on the RGB breast histopathology images. Blood cells and hemorrhage are irrelevant components in the breast histopathology image. Removal of these components is normally performed at the preliminary or post processing stage to ensure the smoothness in the relevant regions (i.e., tumor regions) segmentation. In this study, the gray level global thresholding method was modified for color image segmentation. A set of comprehensive empirical study was performed to select the optimal thresholding values. The organization of this paper is as follows: Section 2, methodology, Section 3, experimental results and Section 4, conclusion.

(a) Blood cells   (b) Hemorrhage

Figure 1. Examples of breast histopathology images with blood cells and hemorrhage.

2. Methodology

Figure 2 shows the block diagram of the proposed method. The proposed method consists of three main steps: color normalization, empirical analysis and color thresholding. The details of each sub-step are described in the following subsections.
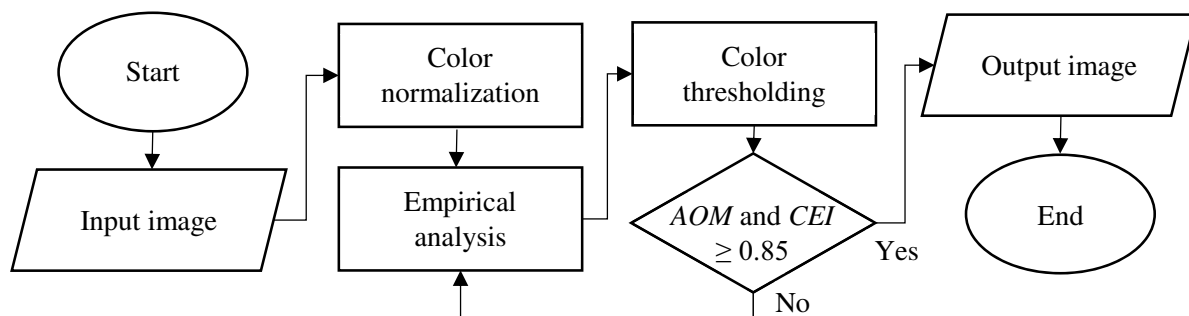
Figure 2. Block diagram of the proposed method.

2.1 Color Normalization

Color normalization step aims to standardize and reduce color inconsistency in the RGB input histopathology images. This step was done using histogram matching algorithm [11]. For this purpose, one breast carcinoma histopathology image was selected as a reference image. The reference image was selected based on the following criterions: (1) the reference image should be chosen from the dominant region and captured under 10x magnification and (2) the reference image should have the common image contents such as tumor regions, non-tumor regions, blood cells, hemorrhage and background to reduce unwanted bias during histogram matching. The R, G and B channels from input image were matched to the R, G and B channels of the reference image.

2.2 Empirical Analysis and Color Thresholding

The conventional gray level global thresholding method [12] in monochrome image is performed by selecting a threshold value from a range of intensity values (i.e., 0 to 255). The equation of the conventional gray level global thresholding method is given in Equation (1), where T is the threshold value, $f(x,y)$ is the original input pixel value and $g(x,y)$ is the output pixel value.

$$g(x,y) = \begin{cases} 0, & f(x,y) < T \\ 1, & f(x,y) \geq T \end{cases} \tag{1}$$

As the RGB input histopathology consists of three color channels, the thresholding process was performed once at a time for each color channel. The Boolean AND operator was used to combine the thresholding process into one rule as given in Equations (2) and (3). To find the suitable values for the thresholding purpose, a set of comprehensive empirical analysis was performed. The study was conducted on five different breast histopathology slides. Three histopathology images were captured from each slide at the dominant area, under 10x magnification. The 15 histopathology images were first pre-processed using the method described in Section 2.1. For each image, the threshold value of R, G, and B were manually adjusted until all the blood cells and hemorrhage were segmented from the background. Table 1 shows the results of the empirical study performed on the 15 histopathology images. The mean value in each color channel was selected as the thresholding value. The blood thresholding values in R (R_{blood}), G (G_{blood}), and B (B_{blood}) are 139, 100, and 130, respectively. The same procedure was applied to the hemorrhage. The hemorrhage has brighter areas (i.e., in terms of intensity values) compared to the blood cells. The hemorrhage threshold values for R ($R_{hemorrhage}$), G ($G_{hemorrhage}$), and B ($B_{hemorrhage}$) are 200, 171, and 171, respectively.

$$g_{blood}(x,y) = \begin{cases} 1, & R(x,y) < R_{blood} \ AND \ G(x,y) > G_{blood} \ AND \\ & \qquad B(x,y) > B_{blood} \\ 0, & otherwise \end{cases} \tag{2}$$

$$g_{hemorrhage}(x,y) = \begin{cases} 1, & R(x,y) < R_{hemorrhage} \ AND \ G(x,y) > G_{hemorrhage} \ AND \\ & \qquad B(x,y) > B_{hemorrhage} \\ 0, & otherwise \end{cases} \tag{3}$$

Table 1. R, G, and B values obtained from the empirical study performed on the 15 breast histopathological images.

| Threshold values | Irrelevant objects | | | | | |
| | Blood cells | | | Hemorrhage | | |
	R	G	B	R	G	B
Mean	139	100	130	200	171	171
Maximum	143	112	139	207	179	174
Minimum	128	96	124	192	165	167

2.3 Evaluation Metric

Two statistical metrics were used: Area Overlap Measure (AOM) and Combined Equal Importance (CEI). AOM is referred as a ratio of the intersection to the union of the two areas to be compared. This metric is commonly used to evaluate the performance of the segmentation algorithm. CEI is an equation that combined AOM, over-segmentation, and under-segmentation to provide them an equal importance. The

equations of AOM and CEI are given in Equations (4) and (5) [13], where A is to the result obtained from the proposed procedure and B is the ground truth.

$$AOM = \frac{area|A \cap B|}{area|A \cup B|} \tag{4}$$

$$CEI = \frac{AOM + \left(1 - \frac{area|B| - area|A \cap B|}{area|B|}\right) + \left(1 - \frac{area|A| - area|A \cap B|}{area|A|}\right)}{3} \tag{5}$$
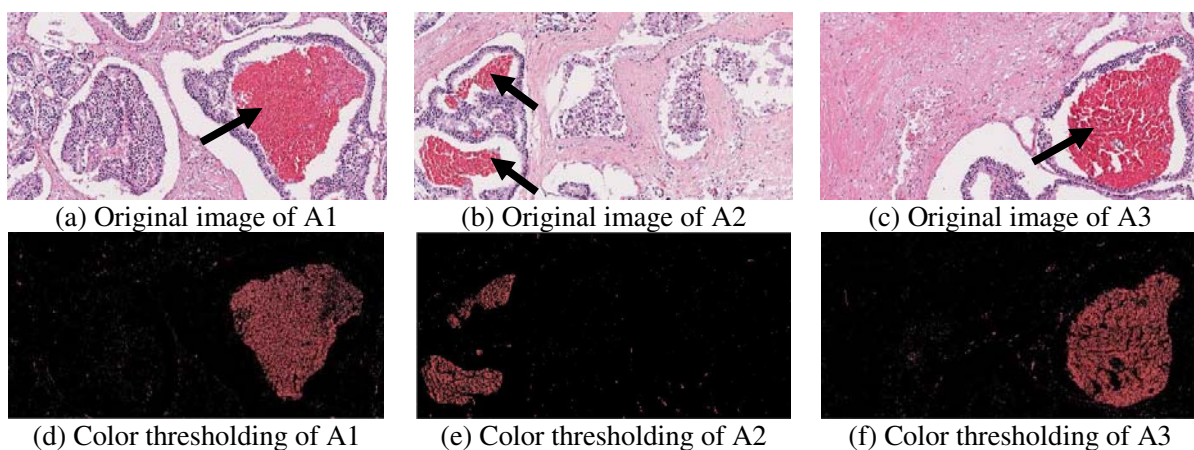
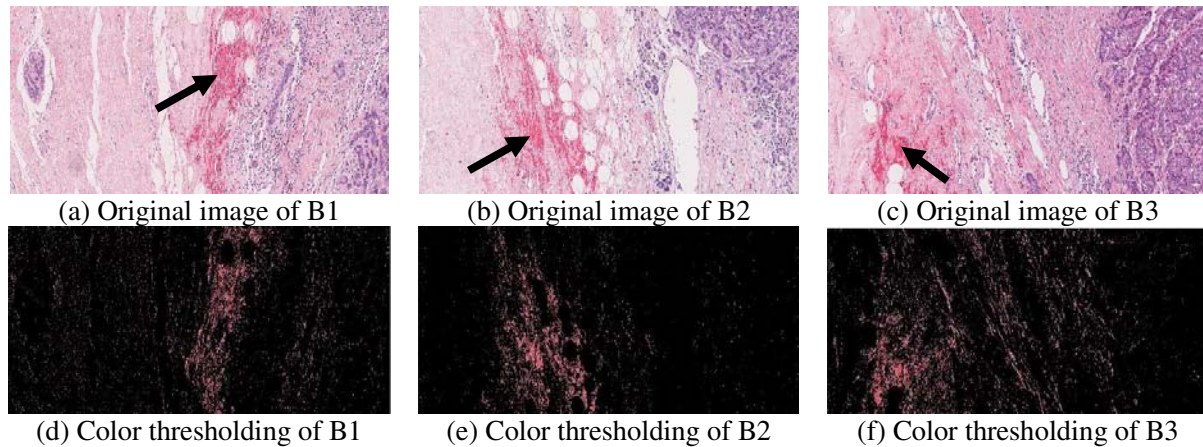## 3. Experimental Results

3.1 Dataset

The histopathology slides of breast carcinoma were locally collected, in Malaysia. The breast histopathology slides were obtained from the Pathology Department, Hospital Tuanku Fauziah, Kangar, Perlis, Malaysia. A total of five breast histopathology slides were used in this study. These slides were prepared under a standard staining procedure aligned with the guideline of MOH, Malaysia, as in [14], using the Hematoxylin and Eosin (H&E) stain. An Aperio CS2 WSI scanner was used to convert the slides into digital form (i.e., digital histopathology slides). The input histopathology images were captured using the Aperio ImageScope at 10x magnification. The captured histopathology image is an 8-bit RGB color with dimensions of 614x1240 pixels (size of pixel: 0.2521 µm per pixel) and is in tiff file format. For the purpose of evaluation, a total of 50 histopathology images (i.e., 25 images with blood cells and 25 images with hemorrhage) were captured, such that 10 images were captured from each slide which corresponding to the different dominant regions. For evaluation purpose, ground truth images for the 50 histopathology images were annotated by an experienced histopathologist.

3.2 Results and Discussion

Three sample images were selected from each dataset, namely A1, A2 and A3 for blood cells images; and B1, B2 and B3 for hemorrhage images. Figures 3 and 4 show the outputs of color thresholding for the blood cells and hemorrhage images, respectively. In each figure, images (a) to (c) show the original RGB input images and images (d) to (f) show the outputs of color thresholding for the respective images.

(a) Original image of A1   (b) Original image of A2   (c) Original image of A3

(d) Color thresholding of A1   (e) Color thresholding of A2   (f) Color thresholding of A3

Figure 3. Blood cells images.

| (a) Original image of B1 | (b) Original image of B2 | (c) Original image of B3 |

| (d) Color thresholding of B1 | (e) Color thresholding of B2 | (f) Color thresholding of B3 |

Figure 4. Hemorrhage images.

Table 2 shows the obtained AOM and CEI for the blood cells images. The hemorrhage is complex in structure and is humanly impossible to manually delimit each region in pixel-based accurately. Therefore, for hemorrhage, the performance of color thresholding is solely depends on human intervention. Referring to Figures 3 and 4, based on the human intervention, the color thresholding method is found to be effective and is able to produce plausible segmentation outputs. The color thresholding method is fast in computation as the thresholding algorithm do not involves with complex mathematical equation. The color thresholding method could be used in the preliminary or post processing stages to remove the irrelevant regions (e.g., blood cells and hemorrhage) from the breast histopathology images.

Table 2. The obtained AOM and CEI for blood cells images.

Image	AOM	CEI
A1	0.94	0.90
A2	0.89	0.86
A3	0.91	0.89
Average for 25 images	0.92	0.89

4. Conclusion

In this study, the conventional gray level global thresholding method was extended to color image segmentation. The color thresholding method was used to segment blood cells and hemorrhage from the RGB breast histopathology images. The thresholding values for each color channel were selected based on the comprehensive empirical analysis done on the respective components. Qualitative and quantitative (blood cells images only) evaluations were performed to evaluate the performance of the color thresholding method. The obtained AOM and CEI for the overall blood cells dataset are 0.92 and 0.89, respectively. This study concludes that the color thresholding method is effective in the segmentation of blood cells and hemorrhage by decomposing the input RGB images into coherent regions based on the similarity in color feature.

References

[1]   Sarkar S, Das S and Chaudhuri SS 2015 A Multilevel Color Image Thresholding Scheme based on Minimum Cross Entropy and Differential Evolution. *Pattern Recognit Lett*;54:27–35. doi:10.1016/j.patrec.2014.11.009.

[2]   Harrabi R and Braiek E Ben 2012 Color image segmentation using multi-level thresholding approach and data fusion techniques : application in the breast cancer cells images. *J Image Video Process*;11:1–11.

[3]   Jian TX, Mustafa N, Mashor MY and Shakir K 2018 Segmentation Based Classification for Mitotic Cells Detection on Breast Histopathological Images. *J Telecommun Electron Comput Eng*;10:2–5.

[4]   Cheremkhin PA and Kurbatova EA 2019 Comparative appraisal of global and local thresholding methods for binarisation of off-axis digital holograms. *Opt Lasers Eng*;115:119–30. doi:10.1016/j.optlaseng.2018.11.019.

[5]   Jian TX, Mustafa N, Mashor MY and Shakir K 2018 An Improved Initialization Based Histogram of K-Mean Clustering Algorithm for Hyperchromatic Nucleus Segmentation in Breast Carcinoma Histopathological Images. *Lect Notes Electr Eng* 2018:1–6.

[6]   Alzu S, Shehab M, Al-Ayyoub M, Jararweh Y and Gupta B 2018 Parallel implementation for 3D medical volume fuzzy segmentation. *Pattern Recognit Lett*;0:1–7. doi:10.1016/j.patrec.2018.07.026.

[7]   Tan XJ, Mustafa N, Rahman KSA, Mashor MY and Ang WC 2018 Simple Landscapes Analysis for Relevant Regions Detection in Breast Carcinoma Histopathological Images. *Int. Conf. Comput. Approach Smart Syst. Des. Appl., IEEE*; 2018, p. 1–5. doi:10.1109/ICASSDA.2018.8477610.

[8]   Klonowski W, Stepien R and Stepien P 2010 Simple fractal method of assessment of histological images for application in medical diagnostics. *Nonlinear Biomed Phys*;4:7. doi:10.1186/1753-4631-4-7.

[9]   Jian TX, Mustafa N, Mashor MY and Shakir K 2018 Hyperchromatic Nucleus Segmentation on Breast Histopathological Images for Mitosis Detection. *J Telecommun Electron Comput Eng*;10:27–30.

[10] N SriMadhava Raja and V. Rajinikanth 2014 Otsu Based Optimal Multilevel Image Thresholding Using Firefly Algorithm. *Modelling and Simulation in Engineering*;2014:1. doi:http://dx.doi.org/10.1155/2014/794574.

[11] Arai K, Kadoya N, Kato T, Endo H, Komori S, Abe Y and et al 2017 Feasibility of CBCT-based proton dose calculation using a histogram-matching algorithm in proton beam therapy. *Phys Medica*;33:68–76. doi:10.1016/j.ejmp.2016.12.006.

[12] Kulkarni N 2015 Color Thresholding Method for Image Segmentation of Natural Images. *Int J Image, Graph Signal Process*;1:28–34. doi:10.5815/ijigsp.2012.01.04.

[13] Abbas Q, Celebi ME and Garcia IF 2013 Breast mass segmentation using region-based and edge-based methods in a 4-stage multiscale system. *Biomed Signal Process Control*;8:204–14. doi:10.1016/j.bspc.2012.08.003.

[14] Guideline CP, editor 2010 Management of Breast Cancer. 2nd ed. *Ministry of Health Malaysia.*