International Journal of Advances in Image Processing Techniques– IJIPT
Volume 2 : Issue 2 [ISSN : 2372-3998]
Publication Date : 30 October, 2015

Skin Cancer Images Thresholding Based on Gamma Distribution

Ali El-Zaart
Department of Mathematics and Computer Science, Beirut Arab University, Beirut, Lebanon

***Abstract*—Skin cancers are the most common form of cancers in humans [An accurate segmentation of skin images can help the diagnosis to define well the region of the cancer. Image processing techniques has been used for processing the skin cancer images. Image thresholding is an important concept, both in the area of objects segmentation and recognition. This paper proposes a new thresholding method for skin cancer images based on between class variance with a mixture of Gamma distributions. The proposed method estimates the threshold iteratively in order to decrease the running time of the algorithm. The experimental results showed that the new method thresholded well the skin cancer images.**

Keywords—*skin images; image thresholding; between class variance; Gamma distribution.*

I. Introduction

Image segmentation is an important step in image analysis, pattern recognition, and computer vision. The segmentation is used to detect an object in many fields. In skin images, the segmentation can detect the cancer regions. The present method is encountered this problem by using Gamma distribution to present data on images that have more than one class. Gamma distribution is a statistical model that is used to analyze distribution of gray levels in image histogram, which is used in different thresholding methods. In this paper, it is used since its capability to provide symmetric and non-symmetric histograms. Threshold techniques can be classified mainly into six groups according to the information they are exploiting [8, 9, 10, 11, 12, 13]. They are: 1) Histogram-based Methods. 2) Clustering-based Methods. 3) Object attribute-based Methods. 4) Spatial Methods. 5) Local Methods 6) Entropy-based Methods. However, this work's goal is improving Otsu's method [6, 7], so it can estimate the optimal threshold values of images based on Gamma distribution, by using iteratively algorithm. Otsu's method is weak when it comes to dealing with low contrast images or where the object is small as in [3]. In general, Otsu's method produces a threshold value that maximizes between-class variances. This work focuses on improving Otsu's thresholding method to generate threshold value automatically for a grayscale images based on Gamma distribution. Symmetric and asymmetric histogram distribution of the intensity values can be represented by using Gamma distribution [5]. Gamma distribution is extensively more than Gaussian distribution, which only represents symmetric histogram distribution of the intensity values.

This paper is organized as follows: section II presents the relates works. Section III, demonstrates the proposed method which is based on Gamma distribution. Section IV, showed the experimental results and finally, in section V presented the conclusion.

II. Related Works

Between Classes variance method [1] is one of the most popular methods for its simplicity and efficiency [4]. It is based in thresholding technique. It depends on selecting the optimal threshold value that maximizes the between-class variance of resulting object anod background classes. The search for the optimal threshold done sequentially until finding a value that makes variance between two classes or more maximum. In this section, will demonstrate Otsu's method and its development which made it fast in computational. If the image in the two dimensions is represented by the function f (x, y) and the values of it in gray-level have the range between [0...L], where L=255. And let N represent the whole number of pixels in the image. The number of pixels with gray-level i is h(i), i= [0,1,..., L] that represent the histogram of the image. For the simplest, normalize of the histogram was computed which represents the probability of occurrence of gray-level i as the follows:

$$p(i) = \frac{h(i)}{N} \qquad p(i) \geq 0 \qquad \sum_{i=0}^{L-1} p(i) = 1$$

The total average or mean value of the image computed as:

$$\mu_T = \sum_{i=0}^{L-1} i p(i)$$

For the single thresholding, image pixels will be divided into two classes C1{0,1,..., t}, and C2{t,,..., L-1}, where t is the threshold value. Usually, the resulted image from this method corresponds to separate the object (C2 that represents the class of bright pixels) from the background (C1 that represents the class of dark pixels). The probability of these two classes is:

$$p_1(t) = \sum_{i=0}^{t-1} p(i) \text{ and } p_2(t) = \sum_{i=t}^{L} p(i)$$

And the mean for them are:

$$\mu_1(t) = \sum_{i=0}^{t} ih(i)/p_1(t) \text{ and } \mu_2(t) = \sum_{i=t}^{L-1} ih(i)/p_2(t)$$

Then by using discriminate analysis [1], Otsu proves that optimal threshold (t*) can be obtained by maximizing the between-class variance as:

$$t^* = \text{Arg}\ \max_{0 \leq t \leq 255}\{\sigma_B^2(t)\} \qquad (1)$$

International Journal of Advances in Image Processing Techniques– IJIPT
Volume 2 : Issue 2 [ISSN : 2372-3998]
Publication Date : 30 October, 2015

Where the between-class variance ($\sigma_B^2(t)$) is defined as:

$$\sigma_B^2(t) = p_1(t)(\mu_1(t) - \mu_T)^2 + p_2(t)(\mu_2(t) - \mu_T)^2 \quad (2)$$

Many papers have been published which study the improvement of this method [7, 21]. In [2], the author improves Otsu's method to make it more efficient in computational side. He used the derivation computation to find the optimal threshold value in iteratively manner rather than searching in sequential that decrease the loop, and then the time of CPU needed for the computational process. He concluded to the following formula for two classes:

$$t = \frac{\mu_1 + \mu_2}{2}$$

This formula has convergence property [2] (in each iteration the threshold value converges to be optimal). In [7], the authors also developed algorithm to compute multilevel threshold values faster, using recursive algorithm with look-up table that shortens the huge number of mathematical operations needed in this situation. However, this method is only applied when the histogram is symmetric but for asymmetric histogram it is better to use Gamma distribution as described in the next section.

III. Proposed Method

The proposed method is based on the classical method of between class variance developed by Otsu [1]. We model the image histogram by a mixture of distributions. Histogram represents statistical information for image pixels. It describes pixels intensity distribution in an image by graphing the number of pixels intensity at each gray level or color intensity level. The histogram can be symmetric and non-symmetric, in our proposed method we used Gamma distribution in order to approximate the skin cancer images histogram by a mixture of Gamma distributions. The Gamma function defines as [5]:

$$f(x, \mu, N) = \frac{2q}{\mu}\frac{N^N}{\Gamma(N)}\left[\frac{qx}{\mu}\right]^{2N-1} e^{-N(\frac{qx}{\mu})^2} \quad (3)$$

$q = \Gamma(N + 0.5)/\sqrt{N}\Gamma(N)$, x is the intensity of the pixel, μ is the mean value of the distribution and and N is the shape of distribution. In our method, Gamma distribution used to estimate the mean values of the image modes and then find the optimal threshold value. Figure 1, shows the Gamma distribution for one mode with different shape parameter N and same value of mean μ.

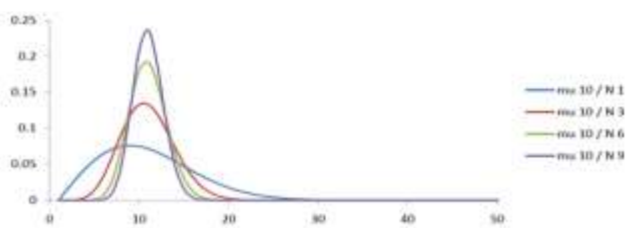


Figure 1: Gamma Distribution with the same mean and different shape parameter N.

This Section will demonstrate our proposal method which is based on Otsu's method using Gamma distribution. We improved our method and gave it the ability to deal with small objects.

Hui-Fuang suggested adding a weight with Otsu's method for detecting small object [3]. The author observes when select the threshold value to be on the valley of the two peaks (from the histogram), the probability of occurrence at the threshold value (pt) has to be small. Based on this observation he proposed to add the weight to Otsu's method which will improve it for selecting threshold values. However, this method called "valley-emphasis". Objective to select a threshold value that has small probability of occurrence (valley in the gray-level histogram), and it also maximizes the between class variance, as in the Otsu method [3]. By applying the weight (1-pt) to Otsu's method is the key of the valley-emphasis method. The smaller the (p(t)) value, the larger the weight (1-pt) will be. This weight ensures that the result threshold value will always be located in at the valley or bottom rim of the gray-level distribution [3]. As Otsu's method, the valley-emphasis method also attempts to maximize the between-class variance of the histogram. Based on Hui-Fuang's work we proposed our method for bimodal thresholding as the following:

$$\eta(t) = [p_1(t)(\mu_1(t) - \mu_T)^2 + p_2(t)(\mu_2(t) - \mu_T)^2]\big(1 - p(t)\big) (4)$$

Where the probability of occurrence of gray-level t is defined as:

$$p(t) = \frac{h(t)}{n} \; ; \qquad p'(t) = \frac{h(t) - h(t-1)}{n}$$

n is the total number of pixels in a given image. $p'(t)$ is the first derivative of the probability.

We consider h(i), i = 0 … 255 be the bimodal histogram of the original image. We assume that the histogram of an image can be seen as a combination of Gamma distributions. The mean values of each mode can be estimated using Gamma distribution [5]. We used Gamma distribution because it has the ability to represent both symmetric and non-symmetric mode rather than the limited Gaussian distribution that describes only the symmetric mode better. However, the mean values will be as the following:

$$\mu_1(t) = \frac{\sqrt{\sum_{i=0}^{t-1} h(i) i^2 q^2}}{\sqrt{p(t)}} \qquad (5)$$

$$\mu_2(t) = \frac{\sqrt{\sum_{i=t}^{L} h(i) i^2 q^2}}{\sqrt{p_2(t)}} \qquad (6)$$

Where μ_T the total mean of image, μ_1 the mean value of the first class and μ_2 the mean of the second class. h(i), i = 0 … 255 is the histogram of image. Here, we aim to divide the histogram in two classes $C_1\{0,1,..., t\}$, and $C_2\{t+1,..., 255\}$. Where t is the threshold value. Consequently we can define the thresholded image as:

$$g(x, y) = \begin{cases} 0 & f(x, y) \le t \\ 255 & f(x, y) > t \end{cases} \qquad (7)$$

Where f(x,y), is the original image, and g(x,y) is the thresholded image result.

Therefore, we calculate the first derivative of Eq. (4) then we set it to zero [2] to get the optimal threshold value as:

$$[p_1(t)(\mu_1(t)-\mu_T)^2 + p_2(t)(\mu_2(t)-\mu_T)^2]'(1-p(t)) + [p_1(t)(\mu_1(t)-\mu_T)^2 + p_2(t)(\mu_2(t)-\mu_T)^2](p(t)-1)' = 0$$

The optimal threshold will be [13]:

$$t = +\sqrt{\frac{-R(t)}{S(t)}}; \qquad (8)$$

Where

$$R(t) = h(t)(n-h(t))\mu_T(\mu_2(t)-\mu_1(t)) - [h(t) - h(t-1)][p_1(t)(\mu_1(t)-\mu_T)^2 + p(t)(\mu_2(t)-\mu_T)^2] (9)$$

and

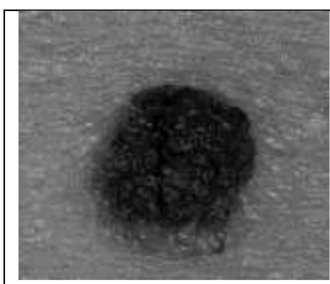
$$S(t) = h(t)q^2(n-h(t))[\frac{1}{\mu_2(t)} - \frac{1}{\mu_1(t)}]$$

IV. Experimental Results

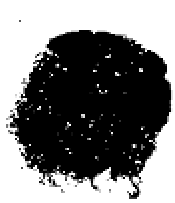
In this section, we applied the proposed thresholding method on four skin images. The first column in figure 2 represents the original skin images and the second column represents the thresholded images with the estimated threshold. We noticed that our proposed method segmented well the skin images.

V. Conclusion and Future Works

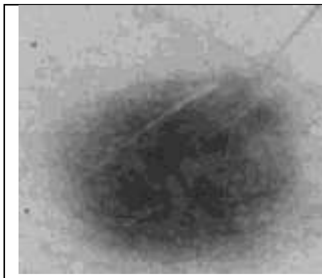
This paper presented a new thresholding method based on between class variance with a mixture of Gamma distributions. Gamma distribution can model data where the histogram is skewed to right or symmetric. We developed a new formula for image thresholding and we estimated the threshold iteratively that make the running time more quickly than the sequential search on the optimal threshold. The experimentation showed that our proposed method gave good in the image thresholding. As future work, we use the segmented image in order to extract features and classify the region as skin cancer or not.



Original Image

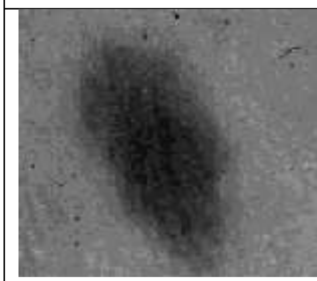


Result Image (t=63)

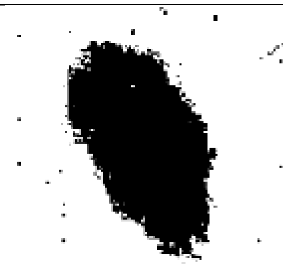


Original Image

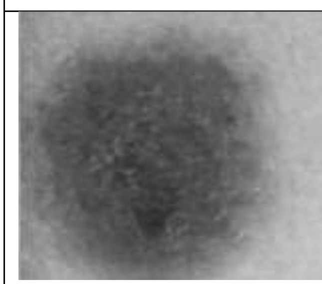
Result Image (t=116)



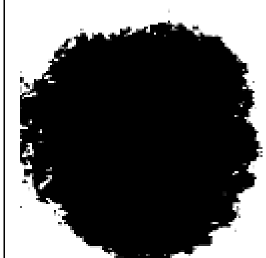
Original Image



Result Image (t=81)



Original Image



Result Image (t=102)

Figure 2: Original skin cancer images with thresholded images

References

[1] N. Otsu, “A threshold selection method from gray-level histogram,” IEEE Transactions on Systems, MAN, and Cybernetics, Vol. SMC-9, No. 1, 1979, pp. 62-66.

[2] K. Chin Lin, “Fast thresholding computation by searching for zero derivatives of image Between-Class Variance,” the 27th Annual Conference of the IEEE Industrial Electronics Society, 2001, pp. 393-397.

[3] H. Fuang Ng, "Automatic Thresholding for Defect Detection," Pattern Recognition Letters, Vol. 27, 2006, pp.1644-1649.

[4] Z. Hou, Q. Hu and W.L. Nowinski, “On minimum variance thresholding,” Pattern Recognition Letters, Vol. 27, 2006, pp. 1732-1743.

[5] A. El Zaart, D. Ziou, S. Wang and Q. Jiang, " Segmentation of SAR images,” Pattern Recognition Journal, Vol. 35, No. 3, 2002, pp. 713-724.

[6] J. C. Russ, “The image processing handbook,” CRC Press, Boca Raton, NW, USA, 5th edition, 2007.

[7] P. Liao, T. Chen, and P. Chung, “A Fast Algorithm for Multilevel Thresholding,” Journal of Information Science and Engineering, Vol. 17, 2001, pp. 713-727.

[8] T. Logeswari ,M. Karnan (2010). An improved implementation of brain tumor detection using segmentation based on soft computing. J. Cancer Res. Exp. Oncol., 2(1): 006-014.

[9] K. Mirzaei, H. Motameni, R. Enayatifar (2011). New method for edge detection and de noising via fuzzy cellular automata. Int. J. Phy. Sci. 6(13): 3175-3180.

[10] P. Sathya, R. Kayalvizhi(2010). PSO-Based TsallisThresholdingSelection Procedure for Image Segmentation. Int. J. Comput. Appl.5(4): 39-46.

[11] O. Sojodishijani, V .Rostam, A. Ramli (2010). A video-rate color image segmentation using adaptive and statistical membership function. Sci. Res. Essays, 5(24): 3914-3925.

[12] L. Yu (2008). Vehicle extractions using histograms and genetic algorithm based fuzzy image segmentation from high resolution UAV aerial imagery. The International Archives of the Photogrammetric, Remote Sens. Spat.

[13] Ali Ghosn, Ali El-Zaart, Eidah Assidan "Mammogram Images Thresholding Based on Between-Class Variance Using a Mixture of Gamma Distributions", 2nd IEEE International conference on Advances In computational Tools For Engineering applications (ACTEA2012). Notre Dame University , Lebanon. December 12-15, 2012.

About Author :

Ali El-Zaart received a M.Sc. degree in computer science from the university of Sherbrooke, Sherbrooke, Canada in 1996, and Ph.D. degree in computer science from the university of Sherbrooke, Sherbrooke, Canada in 2001. His research domain interests include medical image processing and analysis, pattern recognition, computer vision, remote sensing, and biometrics. From 2001 to 2011 he was an assistant and then associate professor in computer science at the King Saud University. Since 2011, he is an associate professor at the Beirut Arab University.