

PAPER • OPEN ACCESS

Tongue Image Segmentation using Hybrid Multilevel Otsu Thresholding and Harmony Search Algorithm

To cite this article: Muhammad Fachrurrozi *et al* 2019 *J. Phys.: Conf. Ser.* **1196** 012072

View the [article online](#) for updates and enhancements.

You may also like

- [Sudlow site II of human serum albumin remains functional after gold nanocluster encapsulation: a fluorescence-based drug binding study of L-Dopa](#)
Ben A Russell, Ashley Garton, Abeer Saeed Alshammari et al.
- [Electrochemical Corrosion Behavior of Orthopedic Biomaterials in Presence of Human Serum Albumin](#)
Shima Karimi and A. M. Alfantazi
- [PEGylated human serum albumin \(HSA\) nanoparticles: preparation, characterization and quantification of the PEGylation extent](#)
E Fahrländer, S Schelhaas, A H Jacobs et al.



Breath Biopsy[®] OMNI[®]

The most advanced, complete solution for global breath biomarker analysis

TRANSFORM YOUR RESEARCH WORKFLOW



Expert Study Design & Management



Robust Breath Collection



Reliable Sample Processing & Analysis



In-depth Data Analysis



Specialist Data Interpretation

Tongue Image Segmentation using Hybrid Multilevel Otsu Thresholding and Harmony Search Algorithm

Muhammad Fachrurrozi¹, Erwin^{2*}, Saparudin³, Nur Rahma Dela⁴, Yenita Mahyudin⁵, and Hardians Kesuma Putra⁶

^{2,4,5} Faculty of Computer Science, Computer Engineering, University of Sriwijaya, Palembang, Indonesia

^{1,3}Departement of Informatics Engineering, University of Sriwijaya, Palembang, Indonesia

⁶Faculty of Medicine, University of Sriwijaya, Palembang, Indonesia

*E-mail: erwin@unsri.ac.id

Abstract. This paper proposes the use of hybrid multilevel thresholding and harmony search algorithm (HSA) methods for segmenting the image of the tongue. The Otsu thresholding algorithm aided by the HSA optimization algorithm is used to look for optimal solution in order to produce the best threshold based on the inserted class. Parameters used there are three levels, namely minimum, medium, and maximum. The dataset used consists of the image of the Hongkong University Polytechnic Biometric research center and the acquisition image. The result shows that the best HSA parameters value is at the medium level with threshold value 5 for benchmark image with image quality of 35.50 dB, maximum level with threshold value 5 in biometric image with image quality of 36.51 dB, and maximum level with threshold value 5 on image acquisition with image quality of 38.02 dB. Furthermore, it is expected to conduct research with a threshold value of more than 5 with maximum HSA parameters so that it can be used better again in the diagnosis of the disease through the tongue.

1. Introduction

The recent development of compute vision is increasing, many applications in various fields such as face recognition [1], [2], [3], multi-object recognition [4], environmental image [5], and biomedical image [6]. In the image, there is the division of an image into different regions or categories that contain each pixel with the same attribute called segmentation. Segmentation is a very important first step in image analysis. If segmentation is done well, then all other stages in image analysis will be simpler [7], [8]. There are three common approaches to segmentation: thresholding is allocated to categories according to the range of values in which a pixel is located, edge-based where pixels are classified as edge or non-edge, and region based by grouping neighboring pixels and having similar values and dividing them into a pixel group that is not different in value. Thresholding is the simplest non-contextual segmentation technique that converts RGB / Grayscale images into binary images that have two regions, one of which contains pixels with input data values smaller than the threshold and the other is larger or above the threshold. In segmentation highly dependent on the threshold image and the selected thresholding method [9], [10].

There are several simple and effective threshold techniques for image segmentation that have been widely used in computer vision and pattern recognition, including otsu, maximum entropy, minimum cross entropy, and histogram. Compared to other threshold segmentation methods, Otsu is the most popular for grayscale image segmentation [11], [12], [13], [14]. The Otsu algorithm can provide color differences in the grayscale color space so that the threshold images can show differences based on the grayscale from 0 to 255 [15], [16]. The Otsu thresholding method involves iterating all possible threshold values and calculating the deployment size for the pixel level on each side of the threshold.

In the process of thresholding required a large calculation, so it takes the optimization process so that computation time is fast. There are several methods of optimization that can be used, one of which is



Harmony Search Algorithm (HSA) [17], [18], [19]. How these method works are that the value of thresholding that has been obtained will be optimized using harmony search algorithm. This algorithm is used to optimize the output of Otsu algorithm because this algorithm is better especially in the case of determining the truss structure design worth 484.85 [20], [21]. Each solution in this algorithm is called Harmony which has a solution archive called harmony memory (HM). At each iteration of harmony search, a new harmony will be generated by considering harmony stored in HM with the probability of harmony memory (HMCR) or using a random element with a 1-HMCR probability.

Image segmentation to be done in this paper is segmentation of tongue image using Otsu and HSA algorithm. This stage is an early stage in the detection of the disease in the image of the tongue which will pass through several other stages.

2. Otsu Multilevel Thresholding

The Otsu algorithm is a global thresholding type first introduced by Nobuyuki Otsu in 1979. The Otsu method itself is widely used because it includes a simple but very effective method which uses only the maximum variance values of different classes as the criteria in image segmentation. Using the L intensity level of the gray image of each RGB color component, the image intensity value can be calculated as equation (1):

$$Ph_i^c = \frac{h_i^c}{N}, \sum_{i=1}^N Ph_i^c = 1, \tag{1}$$

$$c = \begin{cases} 1,2,3, & \text{if RGB,} \\ 1, & \text{if grayscale,} \end{cases}$$

Where:

- Ph_i^c = Distribution probability
- h_i^c = Pixel value corresponding to the intensity level from i until c
- N = the number of pixels in the image
- I = intensity level ($0 \leq i \leq L-1$)
- C = Image component (grayscale or RGB)

The next step of the histogram is normalized in the probability distribution using the following equation (2):

$$\omega_0^c(th) = \sum_{i=1}^{th} Ph_i^c, \omega_1^c(th) = \sum_{i=th+1}^L Ph_i^c. \tag{2}$$

$$C_1 = \frac{Ph_1^c}{\omega_0^c(th)}, \dots, \frac{Ph_{th}^c}{\omega_0^c(th)}, C_2 = \frac{Ph_{th+1}^c}{\omega_1^c(th)}, \dots, \frac{Ph_L^c}{\omega_1^c(th)}$$

Where: $\omega_0(th)$ and $\omega_1(th)$ = Distribution probability from C1 and C2.

Next look for computing average levels and variants between classes using the following equation (3):

$$\mu_0^c = \sum_{i=1}^{th} \frac{iPh_i^c}{\omega_0^c(th)}, \mu_1^c = \sum_{i=th+1}^L \frac{iPh_i^c}{\omega_1^c(th)}, \tag{3}$$

$$\sigma^{2c} = \sigma_1^c + \sigma_2^c, \quad \sigma_1^c = \omega_0^c(\mu_0^c + \mu_T^c)^2, \sigma_2^c = \omega_1^c(\mu_1^c + \mu_T^c)^2,$$

Where:

- μ_0^c and μ_1^c = average rate for class variants 1 and 2
- σ^{2c} = variants between classes, σ_1^c and σ_2^c = class variants 1 and 2

Whereas objective function can be calculated using equation (4) as follows :

$$J(th) = \max(\sigma^{2c}(th)), \quad 0 \leq th_i \leq L - 1, \quad i = 1,2,\dots,k, \tag{4}$$

Where: $th = th_1, th_2, \dots, th_{k-1}$ is a vector that contains multiple thresholds.

3. Harmony Search Algorithm

The HS algorithm (HSA) is one of the optimization algorithms that can be used for position search with maximum function return value. HS has features that distinguish it from simple algorithms and

efficiency searches that have been widely used in areas such as function optimization, mechanical structure design, pipe network optimization, and data classification system optimization. The advantage compared to other optimization methods is that HSA does not require configuration of values based on determinant variables, does not require derivative information, and has several parameters so it is easy to adopt in various optimization problems. There are five parameters in HSA, namely harmony memory (HM), harmony-memory consideration rate (HMCR), pitch adjusting rate (PAR), distance bandwidth (BW), and the number of iteration (NI). The performance of HSAs is strongly influenced by the values given in these parameters.

3.1. Steps

There are 4 steps in the HSA process.

- a) *Initialize Harmony Memory (HM)*. The initial HM consists of a number of randomly generated solutions to the optimization problem under consideration. For n-dimensional problems, HM with HMS size can be represented as equation (5) follows :

$$HM = \begin{bmatrix} x_1^1, & x_2^1, & \dots, & x_n^1 \\ x_1^2, & x_2^2, & \dots, & x_n^2 \\ \vdots & & & \\ x_1^{HMS}, & x_2^{HMS}, & \dots, & x_n^{HMS} \end{bmatrix} \tag{5}$$

Where : $[x_1^i, x_2^i, \dots, x_n^i]$ ($i = 1, 2, \dots, HMS$) is a solution candidate. HMS is usually set between 50 and 100.

- b) *Improvising New Solutions from HM*. Each component of this solution, x_j^i is obtained in the HMCR (harmony memory considering rate). HMCR is defined as the probability of selecting components of HM members, and $1 - HMCR$, therefore, the probability of a result is random. The random results are then mutated in accordance with PAR (pitch adjusting rate). PAR will determine the probability of the candidate from HM.
- c) *Update HM*. New solutions from step 2 are evaluated. If the new harmony vector is better than what is already in HM based on the value of an objective function, new harmony can enter HM, and the worst harmony will not be included in HM.
- d) *Checking criteria for termination*. If the criteria are met (maximum NI), the computation process will be stopped. If not, repeat steps 2 and 3. HMCR and PAR are the two basic parameters in the HS algorithm that control the solution components and affect the convergence speed.

4. Experimental Results

The image being tested is from biometric and acquisition dataset. To maintain compatibility with the same study [22], the number of points used in the threshold test is $th = 2, 3, 4, \text{ and } 5$. Parameter harmony search algorithm is divided into three levels and will be presented in Table 1.

Table 1. Level Parameter Value HSA.

parameter	Minimum	Medium	maximum
HSA			
HM	1	150	300
HMCR	0.7	0.85	0.99
PAR	0.1	0.3	0.5
BW	0001	0:01	0.1
NI	50	1000	2000

PSNR was used in this study as a comparison of the original image quality and image segmentation results. MSE and PSNR is defined in equation (6):

$$MSE = \frac{\sum_{x=1}^B \sum_{y=1}^A (S(x,y) - C(x,y))^2}{AB} \tag{6}$$

$$RMSE = \sqrt{MSE}, \quad PSNR = 10 \times \log_{10} \frac{255^2}{RMSE}$$

Where x and y are the coordinates, A and B are the dimensions, S(x,y) is the original image, and C(x,y) is image segmentation.

4.1. Test Result

The image being tested is a biometric imaginary and acquisition data.

Test Using Biometric Imagery. The biometric image in this study is the image of the tongue that has been used in studies such as image analysis of the tongue [22], [23], [24]. The provided tongue image data is a sample image database of the tongue. There are 12 samples with BMP data format. But for the representation, only 5 test images will be displayed. Testing biometric images shows that the greater the value of the input threshold, the difference in the intensity of the gray image of the object is more visible. Enlargement levels and increased merge boundaries show better results. Testing the biometric image quality using Otsu and HSA will be displayed in the graph. Figure 1 illustrates the use of maximum parameters resulting in the best quality levels among other parameters.

Test Using Acquisition Data. This data acquisition is taken using a camera that has been designed previously. The acquisition results have a JPG image format. The overall size is the same that is 640 x 480. There are 59 acquisition images, but only 5 images are shown in table 4 for representation.

The acquisition image test shown in Table 1 shows a bad PNSR. This is because the image taken by the camera produces images with colors that are less resembling the original object. In addition, shooting in low light conditions results in loss of information that can be found on the tongue such as fungiform papillae and filiform papillae information.

Results of testing the image acquisition using OTSU and HSA uses the HSA value of the parameters. However, at the threshold of 5 with medium HSA parameter value is no better than the maximum value of the HSA parameters. The results of the tests can be seen in Figures 1 and Tables 2.

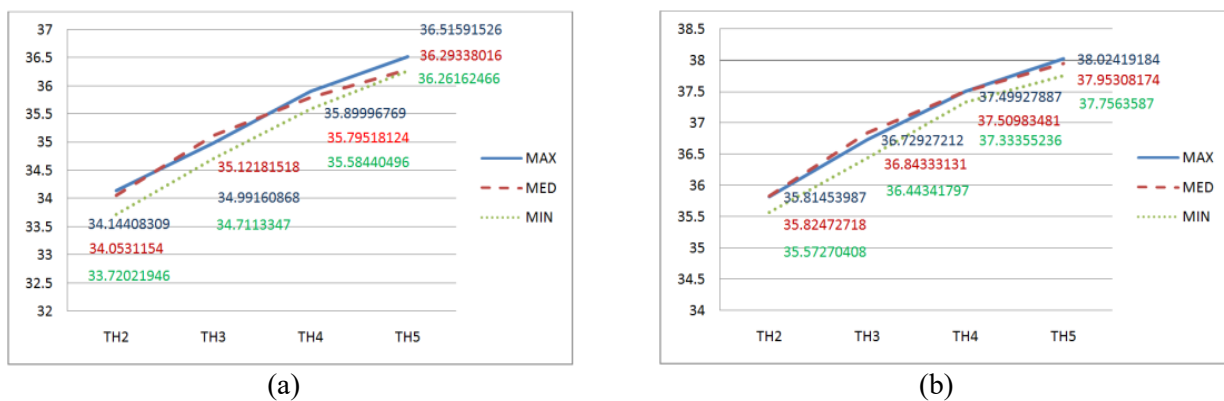
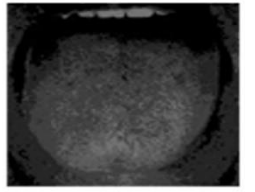
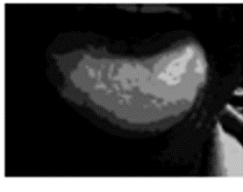
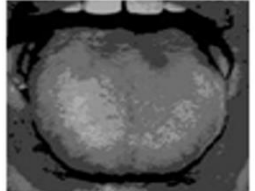
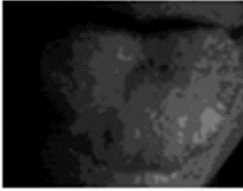
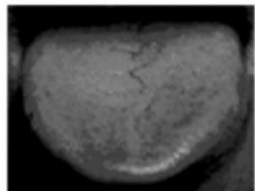
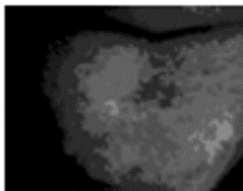
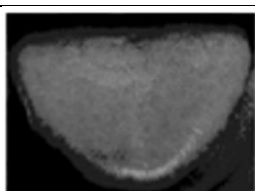
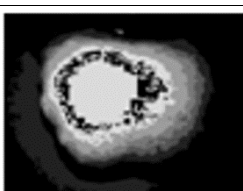
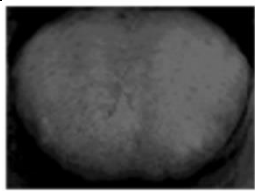
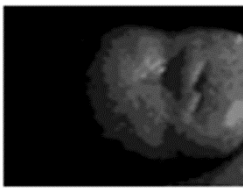


Figure 1. PSNR value a. biometric images, b. acquisition images

5. Conclusion

The experiment uses three levels of HSA parameter values and four levels of threshold values with 12 biometric images in the form of tongues and 59 tongue images derived from acquisitions. The results show that the HSA parameter value is best at a moderate level with a threshold value of 5 for benchmark images with 35.50 dB image quality, maximum level with a threshold value of 5 in biometric images with 36.51 dB image quality, and maximum level with values threshold is 5 in image acquisition with image quality 38.02 dB.

Table 2. Segmentation results

Dataset Biometric			Dataset Biometric		
Image	k	Best Thresholding	Image	k	Best Thresholding
	5	38;64;92;114;242		5	13;46;92;140;192
	5	63;95;111;147;171		5	14;38;59;88;123
	5	37;65;102;124;165		5	29;45;81;99;137
	5	31;82;97;115;147		5	33;85;121;158;219
	5	24;48;66;84;107		5	27;57;83;07;139

6. References

[1] M. Fachrurrozi, Erwin, Saparudin, and Mardiana, “Multi-object face recognition using Content Based Image Retrieval (CBIR),” in *2017 International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2017, pp. 193–197.

[2] M. Fachrurrozi *et al.*, “The grouping of facial images using agglomerative hierarchical clustering to improve the CBIR based face recognition system,” in *2017 International Conference on Data and Software Engineering (ICoDSE)*, 2017, pp. 1–6.

[3] M. Fachrurrozi, S. Masyarif, M. Mardiana, C. Badillah, J. Erlina, and A. Lazuardi, *Real-time Multi-object Face Recognition Using Content Based Image Retrieval (CBIR)*, vol. 8. 2018.

[4] Erwin, M. Fachrurrozi, A. Fiqih, B. R. Saputra, R. Algani, and A. Primanita, “Content based image retrieval for multi-objects fruits recognition using k-means and k-nearest neighbor,” in *2017 International Conference on Data and Software Engineering (ICoDSE)*, 2017, pp. 1–6.

[5] M. Fachrurrozi, Saparudin, and Erwin, “Real Time Monitoring System of Pollution Waste on Musi River Using Support Vector Machine (SVM) Method,” *IOP Conf. Ser. Mater. Sci. Eng.* *197 012057*, pp. 1–8, 2017.

- [6] Erwin, A. Nevriyanto, and D. Purnamasari, "Image enhancement using the image sharpening, contrast enhancement, and Standard Median Filter (Noise Removal) with pixel-based and human visual system-based measurements," in *2017 International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2017, pp. 114–119.
- [7] O. M. Alia *et al.*, "Color tongue image segmentation using Fuzzy Kohonen Networks and Genetic Algorithm," *Proc. SPIE---The Int. Soc. Opt. Eng.*, vol. 3962, no. 1, pp. 176–179, 2010.
- [8] Erwin, Saparudin, and M. Fachrurrozi, "Segmentation and classification models validation area mapping of peat lands as initial value of Fuzzy Kohonen Clustering Network," in *IAES International Conference on Electrical Engineering, Computer Science and Informatics IOP Publishing*, 2017.
- [9] Erwin, Saparudin, A. Nevriyanto, and D. Purnamasari, "Performance Analysis of Comparison between Region Growing , Adaptive Threshold and Watershed Methods for Image Segmentation," in *Proceedings of the International MultiConference of Engineers and Computer Scientists 2018*, 2018, vol. I, pp. 157–163.
- [10] Saparudin, Erwin, and M. Fachrurrozi, "Tongue Segmentation Using Active Contour Model," in *IAES International Conference on Electrical Engineering, Computer Science and Informatics IOP Publishing*, 2017, pp. 1–6.
- [11] D. Oliva, E. Cuevas, G. Pajares, D. Zaldivar, and M. Perez-cisneros, "Multilevel Thresholding Segmentation Based on Harmony Search Optimization," *J. Appl. Math.*, vol. 2013, 2013.
- [12] K. Chen *et al.*, "Multilevel Image Segmentation Based on an Improved Firefly Algorithm," *Int. J. Electr. Comput. Eng.*, vol. 2016, no. 2, pp. 1–12, 2017.
- [13] X. Yang, X. Shen, J. Long, and H. Chen, "An improved median-based Otsu image thresholding Algorithm," vol. 3, pp. 468–473, 2012.
- [14] P. Kalavathi, "Brain Tissue Segmentation in MR Brain Images using Multiple Otsu's Thresholding Technique," no. Iccse, pp. 639–642, 2013.
- [15] R. Singh, P. Agarwal, M. Kashyap, and M. Bhattacharya, "Kapur ' s And Otsu ' s Based Optimal Multilevel Image Thresholding Using Social Spider and Firefly Algorithm," no. v, pp. 2220–2224, 2016.
- [16] F. Kurugollu, B. Sankur, and a. . Harmanci, "Color image segmentation using histogram multithresholding and fusion," *Image Vis. Comput.*, vol. 19, pp. 915–928, 2001.
- [17] M. Mahdavi, M. Fesanghary, and E. Damangir, "An improved harmony search algorithm for solving optimization problems," *Appl. Math. Comput.*, vol. 188, no. 2, pp. 1567–1579, 2007.
- [18] F. Maleki, M. Nooshyar, and G. Z. Fatin, "Breast Cancer Segmentation in Digital Mammograms Based on Harmony Search Optimization," pp. 477–484, 2014.
- [19] B. Alatas, "Chaotic harmony search algorithms," *Appl. Math. Comput.*, 2010.
- [20] B. Pang, D. Zhang, N. Li, and K. Wang, "Computerized tongue diagnosis based on Bayesian networks," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 10, pp. 1803–1810, 2004.
- [21] Z. W. Geem, C.-L. Tseng, and J. C. Williams, *Music-Inspired Harmony Search Algorithm*, vol. 191. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.
- [22] B. Pang, D. Zhang, and K. Wang, "Tongue image analysis for appendicitis diagnosis," *Inf. Sci. (Ny)*, vol. 175, no. 3, pp. 160–176, 2005.
- [23] D. Zhang, H. Zhang, and B. Zhang, "Tongue Color Analysis for Medical Application," *Tongue Image Anal.*, vol. 2013, pp. 207–223, 2017.
- [24] B. Zhang, B. V. K. Vijaya Kumar, and D. Zhang, "Detecting diabetes mellitus and nonproliferative diabetic retinopathy using tongue color, texture, and geometry features," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 2, pp. 491–501, 2014.