Multilevel Minimum Cross Entropy Threshold Selection based on Honey Bee Mating Optimization

Ming-Huwi Horng
Department of Computer Science and Information Engineering
National PingTung Institute of Commerce
51 Min Sheng E.Road, PingTung 900
TAIWAN
horng@npic.edu.tw

Abstract: - Image entropy thresholding approach has drawn the attentions in image segmatation. The endeavor of this paper is focused on multilevel thresholding using the minimum cross enrtop criterion. In the literature, the particle swarm optimization (PSO) had been applied to conducting the thresold selection. The adopted algorithm used in this paper is the honey bee mating optimization (HBMO). In experimnats, the two different methods are implemented for comparison with the results of segmantation. Compared to the results of PSO, the threshold selection of HBMO is more close to the optimal ones examined by the exhaustive search method. Furthermore, the segmentation result of HBMO is superior to PSO method, but it still slower than ones of PSO.

Key-Words: - Image segmentation, Image thresholding, Cross entropy, Particle swarm optimization, Honey bee mating optimization

1 Introduction

Image thresholding is a basic component of many computer vision systems. While humans can easily differentiable an object from complex background, image thresholding is a difficult task for separate them. The gray-level histogram of an image is usually considered as efficient tools for development of image thresholding algorithms. The main objective is to determine a threshold for bi-level thresholding or several thresholds for multi-level thresholding for image segmentations. Several algorithms of multi-level thresholding have been proposed in literature that included the works of Kapur et al [1], Otsu [2] and fast Ostu's implementation [3]. Among the tremendous amount of image thresholding technniques, entropy-based approaches have drawn many attentions. Yin [4] proposed a new method that adopts the particle swarm optimization to select the thresholds based on the minimun cross-entropy. Madhubanti et al. Uses the hybrid cooperative-comprehensive learning PSO algorithm based on amximun entropy criterion[5].

Besides, over the last decade, modeling the behavior of social insects, such as ants and bees, for the purpose of search and problems solving has been the context of the emerging area of swarm intelligence. Therefore, the honey-bee mating may also be considered as a typical swarm-based approach for searching for the optimal solution in many application domains such as clustering [6], market segmentation [7] and benchmark mathematical problems [8].

Honey-bee mating optimization (HBMO) may also be considered as a typical swarm-based approach for optimization, in which the search algorithm is inspire by the process of mating in real honey-bees. The behavior of honey-bees is the interaction of their (1) genetic potentiality. (2) ecological and physiological environments, and (3) the social conditions of the colony, as well as various prior and ongoing interactions between these three parameters. This paper introduces a new approach for multi-level thresholding adopted the honey bee mating optimization based on the minimun cross-entropy criterion. The rest of the paper is organized as follows. Section 2 introduces the honey bees mating optimization. Section 3 presents the multilevel thresholding using HBMO. Performance evaluation is presented in detail in Section 4 Conclusions are presented in Section 5.

2. Honey bee mating optimization

A honeybee colony typically consists of a single egg-laying long-lived queen, anywhere from zero to several thousands drones. Queens are specialized in egg laying. A colony may contain one queen or more during its life cycle, which named monogynous and/or polygynous. A queen bee may live up to 5 or 6 years, whereas worker bee and drones never live more than 6 months. After the mating processes, the drones die. The drones are the fathers of colony. They are haploid and act as amplify their mother's genomes without altering their genetic composition, expect through the mutation. The drones practically considered as agents that pass one of their mother's gametes and function to enable females to act genetically as males. Worker bees specialized in brood care and sometimes lay eggs. Broods arise either from fertilized (represents queen or worker) or unfertilized (represents drones) eggs.

In the marriage process, the queen(s) mate during their mating flights far from the nest. A mating flight starts with a dance performed by the queen who then starts a mating flight during which the drones follow the queen and mate with her in the air. In each mating, sperm reaches the spermatheca and accumulates there to form the genetic pool of the colony. Each time a queen lays fertilized eggs, she randomly retrieves a mixture of the sperm accumulated in the spermatcheca to fertilize the egg. In practices, the mating flight may considered as a set of transitions in a state-space where the queen moves between the different states in some speed and mates with the drone encountered at each state probability. Furthermore, The queen initialized with some energy content during the flight mating and returns to her nest when the energy is within some threshold from zero to full spermatches.

In the development of the algorithm, the capability of workers is restrained in brood care and thus each worker may be regarded as a heuristic that acts to improve and/or take care of a set of broods. An annealing function is used to describe the probability of a drone (D) successfully mates with the Queen (Q), shows as Eq. (1).

$$P(Q,D) = \exp[-\Delta(f)/S(t)] \tag{1}$$

where $\Delta(f)$ is the absolute difference of the fitness of D and the fitness of Q, and the $S(t)$ is the speed of queen at time *t*. After each transition of mating, the queen's speed and energy decays according to the following equations:

$$S(t+1) = \alpha(t) \times S(t) \tag{2}$$

where α is the decreasing factor ($\alpha \in [0,1]$). Workers adopt some heuristically mechanisms such as crossover or mutation to improve the brood's genotype based on a pre-defined heuristic fitness value. The fitness of the resulting genotype is determined by evaluating the value of the objective function of the brood genotype or its normalized value. The five stages of constructing the HBMO algorithm had been proposed by M. Fathian *et al.* [9] that are used to develop the algorithm for multi-level image thresholding.

3. Proposed approach

The proposed algorithm has two main phases. The first phase involves generating the objective function based on image entropy for later developing the HBMO algorithm. The second phase introduces the HBMO algorithms for multi-level image thresholding based on the M. Fathian's five stages model.

3.1. Cross entropy measure criterion

The cross entropy was developed by Kullback in [9]. Let $F = \{f_1, f_2, ..., f_N\}$ and $G = \{g_1, g_2, ..., g_N\}$ be two probability distributions on the same set. The cross entropy between F and G is defined by

$$D(F,G) = \sum_{i=1}^{N} f_i \log \frac{f_i}{g_i} \tag{3}$$

The minimum cross entropy thresholding (MCET) algorithm selects these thresholds by minimizing the cross entropy between the original image and the resulting image. Let ***I*** be the original image and $h(i)$, $i = 1,2,...,L$, be the corresponding histogram with L being the number of gray levels. Then the resulting image, denoted by I_t, using *t*as the threshold value is constructed by

$$I_t(x,y) = \begin{cases} \mu(1,t) & I(x,y) < t, \\ \mu(t, L+1) & I(x,y) \geq t, \end{cases} \tag{4}$$

where

$$\mu(a,b) = \sum_{i=0}^{b-1} ih(i) \Big/ \sum_{i=a}^{b-1} h(i) \tag{5}$$

The cross entropy is then calculated by

$$D(t) = \sum_{i=1}^{L} ih(i)\log(i) - \sum_{i=1}^{t-1} ih(i)\log(\mu(1,t)) - \sum_{i=t}^{L} ih(i)\log(\mu(t,L+1)). \tag{6}$$

The MCET determines the optimal threshold t^* by minimizing the cross entropy based on Eq. (6).

$$t^* = \arg\min_t \{D(t)\}. \tag{7}$$

Since the first term is constant for a given image, the objective function can be re-written as

$$\eta(t) = -\sum_{i=1}^{t-1} ih(i)\log(\mu(1,t)) - \sum_{i=1}^{L} ih(i)\log(\mu(t,L+1)$$

$$= -(\sum_{i=1}^{t-1} ih(i)) \log(\frac{\sum_{i=1}^{t-1} ih(i)}{\sum_{i=1}^{t-1} h(i)}) - (\sum_{i=t}^{L} ih(i)) \log(\frac{\sum_{i=t}^{L} ih(i)}{\sum_{i=t}^{L} h(i)})$$

$$= -m^1(1,t)\log(\frac{m^1(1,t)}{m^0(1,t)}) - m^1(t,L+1)\log(\frac{m^1(t,L+1)}{m^0(t,L+1)})$$

(8)

where $m^0(a,b) = \sum_{i=a}^{b-1} h(i)$ and $m^1(a,b) = \sum_{i=a}^{b-1} ih(i))$ are the zero-moment and first-moment on partial range of the image histogram.

Assume that it is required to select n thresholds denoted by $t_1, t_2, t_3, ..., t_c$. For the convenience of computation, the two dummy thresholds $t_0 = 1$, $t_{c+} = L + 1$ are used. The objective function then becomes

$$\eta(t_1, t_2, ..., t_n) = -\sum_{i=1}^{c+1} m^1(t_1, t_2, ..., t_i) \log(\frac{m^1(t_{i-1}, t_i)}{m^0(t_{i-1}, t_i)}).$$

(9)

3.2. Image thresholding based on HBMO

In this paper the multi-level image thresholding using the HBMO algorithm is developed based on the popular the five stages approach proposed by M. Fathian [9]. The details of our method include five stages show as follows.

Stage 1: Generate the initial drone sets and queen

Generate *m* drones with gene length *c*, denoted by the matrix D_{mc}.

$$D = [D_1, D_2, ..., D_m]$$
$$D_i = (d_i^1, d_i^2, ...d_i^c) \quad (10)$$

where d_i^j is the *j*-th gene value of *i*-th drone that is restricted into [0,..,L] and the $d_i^j < d_i^{j+1}$ for any j. Among all drones, the drone D_h with the maximum value of $f(D_i)$ for all *i* is assigned as the queen Q. The Queen Q is described X_{best} as shown as the Eq. (11)

$$Q = [X_{best}^1, X_{best}^2, ..., X_{best}^c] \quad (11)$$

Stage 2: Flight matting

The stage2 do the flight matting of queen Q. The best drone D_k with the largest $\mathrm{Prob}(Q, D_i)$ Prob among the drone set D is selected the object of matting for the queen Q. After the flight matting the queen's speed and energy decay is reduced by Eq.(2). The flight matting is continues until the speed $S(t)$ is less than a threshold *d* or the number of sperms of the queen's spermatcheca is less than the threshold n_{sperm}. In general, the values of *d* and n_{sperm} are predefined by users and the n_{sperm} is less than m. The selected sperms, Sperm, are described by Eq. (12).

$$Sperm = [Sp_1, Sp_2, ..., Sp_{n_{sperm}}]$$
$$Sp_i = (X_i^1, X_i^2, ..., X_i^c) \quad (12)$$

where Sp_i is the *i*-th individual in the queen's spermatheca.

Stage 3. Breeding process

In this step, a population of broods is generated based on matting between the queen and the drones stored in the queen's spermatheca. At first the *j*-th individual is selected if the random number R_j is less than a user-defined breeding ratio P_c to breed. The breeding process can transfer the genes of drones and the queen to the *j*-th individual based on the Eq. (13).

$$Brood_j = Q \pm \beta \times (Sp_j - Q) \quad (13)$$

The $Brood_j$ is a brood that generated by the queen and the *j*-th individual of spermatheca of queen. The parameter β is randomly generated in the interval [0, 1]

Stage 4. Brood mutation with the royal jelly by works.

The population of broods is improved by applying the mutation operators as follows:

Step1. For all broods, the random number R_i of *i*-th brood is generated.

Step 2. The *i*-th brood needs mutation if the R_i is more than the predefined mutation ratio P_m based on the Eq. (14). For the mutated brood $Brood_i$, the only one gene X_i^k do the mutation. The *k* is randomly selected in integer interval (1,…,*c*).

$$Brood_i^k = Brood_i^k \pm (\delta + \varepsilon) \times Brood_i^k$$
$$\delta \in [0,1],\ 0 < \varepsilon < 1 \quad (14)$$

The δ is randomly generated and ε is pre-defined.

Step 3. The best brood, $brood_{best}$ with maximum objective function value is selected as the candidate queen.

Step 4. If the objective function of $brood_{best}$ is superior to the queen, the queen replaces with $brood_{best}$.

Stage 5. Check the termination criteria

If the termination criteria satisfied finish the algorithm, else generate new drones set and go to stage 2.

4 Experimental Results

In this section, the proposed algorithm is employed to solve the multilevel image thresholding. The PSO based multilevel threshold selection based on the cross entropy criterion is implemented for later comparison. All programs are designed by using the Matlab package under a personal computer with 2.4GHz CPU, 1G RAM with Microsoft windows XP system. Two popular images, that are Lena and Pepper with image size 256×256, respectively, are used for conducting by using the four methods in experiments. The two images are shown in Fig.1.

(a)Lena original image | b) Pepper original image

Fig. 1. Two experimental images

In the implementation of HBMO algorithm, it needs many parameters that are predefined and randomly generating. Table 1 shows the parameters of HBMO algorithm. In order to consistent comparisons, two performance indicators that are peak signal to noise ratio (PSNR) and uniformity measure, are used. For the sake of completeness we define PSNR, measured in decibel (dB) as

$$PSNR = 20\log_{10}\left(\frac{255}{RMSE}\right) \tag{15}$$

where RMSE is the root mean-squared error, defined as

$$RMSE = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j) - \hat{I}(i,j))^2} \tag{16}$$

Here I and $\hat{I}$ are original and thresholded images of size $M \times N$, respectively.

parameter	explanation	Value
	Number of queens	1
m	Number of drones	150
c	Number of threshold	2,3,4, or 5
L	The grayscale of image	255
α	Speed reduction schema	0.98
n_{sperm}	Capacity of spermatheca	50
S(0)	Speed of queen at first of flight	Randomly $\in [0.5,...,1]$
P_c	The breeding ratio	0.8
P_m	Mutation ratio	0.001
ε	Mutation variation	0.5

Table 1. The parameters used in the HBMO algorithm

For applying the multilevel thresholding on images LENA and PEPPER, we execute the PSO-based with three particles and the HBMO-based algorithms with one queen for 200 iterations. The exhausive search method is also conducted for derived the optimal solutions for comparison. Table 2 shows the selection of thresholds of MCET derived by PSO method, HBMO method and exhaustive search method. The derived MECT thresholds by the three methods are equivalent or very close (for the 2,3,4-threshold) to the optimal thresholds derived from the exhaustive method, however, in the 5-threshold the selected thresholds of PSO-based method are apparently different from ones generating from the two other methods. In compareison with PSO-based methods, the solution of HBMO-based method are more powerful in searching the optimal solutions.

Table 3 displays the consumed computionation times and the correspondind PSNR values for images Lena and Pepper. From the table we found that the computation times of exhaustive search method is exponential, particularly, the needed CPU times for $k \geq 4$ are absolutely unacceptable. The compution times of segmentation using the HBMO-based methods are nearly double compared to the

PSO-based method. Howevre, the needs of computation of this two methods are negligible beacuse the needs is less than one second. .

In general, the segmented images are more informative as the number of the thresholds increases. The PSNR value of the image may reflects to the quality of image. The experimental results of Table 3 reveal that the resulting images of Lena and Pepper using the HBMO-based method are superior to the ones of PSO-based method.

4 Discussion and Conclusion

This paper presents a new multilevel image thresholding scheme based on the honey bee mating optimization (HBMO) algorithm. From the Table 2 and Table 3 we find the two important contributions. One is that the proposed HBMO-based method can more efficient to search the near optimal solutions compared to the exhaustive search method. The other is that the quality of segemantion images using HBMO-based method is superior to the ones of PSO-based method. The experimental result is promising and it encourages further research for applying the HBMO algorithm to comples image processing and pattern recognition.

References:

[1] J. N. Kapur, P.K. Sahoo, and A.K.C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, *Computer Vision Graphics Image Processing*, 29, 1985, pp. 273-285.

[2] N. Ostu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, Cybernetics*, SMC-9, 1979, pp. 62-66.

[3] P.S. Liao, T. S. Chen and P. C. Chung, A fast algorithm for multilevel thresholding, *Journal of Information Science and Engineering*, 17, 2001, pp. 713-727.

[4] P. Y. Yin, Multilevel minimum cross entropy threshold selection based on particle swarm optimization, *Applied Mathematics and Computation,* Vol. 184, 2007, pp. 503-513.

[5] M. Madhubanti and A. Amitava, A hybrid cooperative-comprehensive learning based algorithm for image segmentation using multilevel thresholding", *Expert systems with Application*, 34, 2008, pp. 1341-1350.

[6] F. Mohammad, A. Babak and M. Ali, Application of honey-bee mating optimization algorithm on clustering, *Applied Mathematics and Computations*, 2007, 1502-1513.

[7] B. Amiri and M. Fathian, Integration of self organizing feature maps and honey bee mating optimization algorithm for market segmentation, *Journal of Theoretical and Applied Information Technology*, 2007, pp. 70-86.

[8] D. Karaboga and B. Basturk, On the performance of artifical bee colony algorithm, *Applied Soft Computing*, 2008, pp. 687-697.

[9] S. Kullback, Information Theory and Statistics, Dover, New York, 1968.

[10] G. H. Li, c. K. Lee. Minimum cross entropy thresholding. Pattern Recognition, 26, 1993, pp. 617-625.

Image	k-thresholds	Selected thrsholds		
		HBMO	PSO	Exhastive
Lena	2	53,117	53,117	53,117
	3	46,95,150	46,95,150	46,95,150
	4	40,77,114,160	47,73,111,159	40,77,114,160
	5	30,55,86,120,162	39,73,102,130,170	29,53,83,117,161
Pepper	2	52,125	52,125	52,125
	3	48,107,157	48,107,157	48,107,157
	4	35,75,117,163	37, 76,117,163	35,75,117,163
	5	34,71,104,136,171	35,75,117,163,239	34,71,104,137,172

Table 2. The minimum creoo entropy thresholds for LENA and Pepper derived by the HBMO-based method, PSO-based method and the exhausive search method,.

Image	k-thresholds	Execution time(sec)/PSNR(dB)					
		HBMO		PSO		Exhastive	
Lena	2	0.82	16.05	0.44	16.05	1	16.05
	3	0.83	18.32	0.45	18.32	91	18.32
	4	0.84	20.35	0.46	20.14	6300	20.35
	5	0.87	21.77	0.45	21.57	331302	21.60
Pepper	2	0.81	15.31	0.45	15.31	1	15.31
	3	0.82	17.90	0.45	17.90	92	17.90
	4	0.85	20.22	0.43	20.20	6345	20.22
	5	0.87	21.92	0.46	20.22	336181	21.96

Table 3. The execution time and peak signal and noise ratio