

基于改进蜂群繁殖算法的 Tsallis 熵阈值图像分割

马巧梅, 李婧婧, 康 琚

(中北大学 软件学院, 山西 太原 030051)

摘 要: 针对传统蜂群繁殖算法存在的不足, 提出了一种改进蜂群繁殖算法. 首先, 将传统蜂群繁殖算法中自定义的繁殖概率和突变概率改进为自适应的, 以提高算法的快速收敛性和全局寻优能力; 其次, 引入精英保留策略, 以缩短寻优时间; 最后, 结合 Tsallis 熵的非广延性, 采用 Tsallis 熵作为阈值图像分割的适应函数, 利用改进后的蜂群繁殖算法实现多阈值图像分割, 避免图像分割过程中非可加信息被忽略, 进而提高图像的分割精度. 为了验证所提出算法的可行性, 对该算法进行仿真, 并与多种算法进行对比研究. 实验结果表明, 该算法在寻优能力、快速收敛能力、图像分割精度、图像分割速度等方面都有进一步的提高.

关键词: 图像分割; Tsallis 熵; 多阈值; 蜂群繁殖算法; 分割速度

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-7180(2018)04-0032-05

Tsallis Entropy Threshold Image Segmentation Algorithm Based on Improved Honey Bee Mating Algorithm

MA Qiao-mei, LI Jing-jing, KANG Jun

(Software School, North University of China, Taiyuan 030051, China)

Abstract: Aiming at the shortcomings of the traditional honey bee mating algorithm, an improved honey bee mating algorithm was proposed. Firstly, the custom breeding probability and mutation probability in the traditional honey bee mating algorithm were improved to be adaptive to improve the fast convergence and global optimization ability of the algorithm; Secondly, the elite retention strategy was introduced to shorten the optimization time; Finally, with the non-extensibility of Tsallis entropy, using the Tsallis entropy as the fitness function of threshold image segmentation, and the improved honey bee mating algorithm was used to achieve multi-threshold image segmentation to avoid the non-add-on information of the image segmentation process being ignored, thereby improving the accuracy of image segmentation. In order to verify the feasibility of the proposed algorithm, the algorithm was simulated, and compared with some algorithms. The experimental results showed that the algorithm had further improved the optimization ability, fast convergence ability, image segmentation accuracy, image segmentation speed, and so on.

Key words: image segmentation; Tsallis entropy; multi-threshold; honey bee mating algorithm; segmentation speed

1 引言

图像分割是指根据图像多方面的特征将图像划分为互不相交的区域并获得图像中感兴趣的目标区域. 目前对图像分割技术的改进措施主要集中在以下三个方面: 对分割技术的改进、针对不同领域的图

像特征不同而提出不同的改进方法、图像分割方法与智能算法相结合^[1]. 本文提出的算法主要是与传统蜂群繁殖算法(HBMA)相结合, 对传统蜂群繁殖算法提出进一步的改进, 并应用于多阈值图像分割中.

目前, 对传统蜂群繁殖算法的改进及应用主要

有: Horng M H 提出了基于蜂群繁殖优化算法的最大熵阈值分割方法^[2] (MEHBMOT); Jiang Y 提出了合作蜂群繁殖优化算法(CHBMA)在多阈值图像中的应用^[3], 随后提出了基于贝叶斯定理的蜂群繁殖算法(BHBMA)^[4]; Marinakis 用蜂群繁殖优化算法解决旅行商问题^[5]; 等等. 但蜂群繁殖算法仍存在以下问题: 第一, 算法中的繁殖概率和突变概率为自定义的, 降低了算法的收敛性和全局搜索能力; 第二, 算法中的每一次迭代过程都是一次重新寻优过程, 未考虑先前迭代过程所求得的最佳分割阈值, 增加了迭代次数和计算时间; 第三, 图像分割过程中, 算法中适应函数的设定未考虑图像中非可加信息易被忽略这一缺陷, 从而降低了图像分割精度. 针对上述问题, 本文引入了自适应的繁殖概率和突变概率, 并结合精英保留策略和 Tsallis 熵的优点提出一种基于改进蜂群繁殖算法的 Tsallis 熵阈值图像分割算法(THBMO).

2 HBMA 算法

2.1 HBMA 算法的一般流程

HBMA 算法最早由 Abbass 于 2001 年提出的^[6]. 算法中采用退火算法来表示蜂王与雄蜂交配成功的概率, 如式(1)所示:

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

式中, Δf 表示蜂王 Q 与雄蜂 D 适应值的差的绝对值, $S(t)$ 表示在时间 t 时蜂王 Q 的飞行速度. 在每一次交配完成之后, 蜂王的速度和能量都会依据式(2)减少, α 表示减小因子, 其取值范围为($\alpha \in [0, 1]$).

$$S(t+1) = \alpha \times S(t) \quad (2)$$

HBMA 算法的一般流程如下:

- (1) 算法从配对飞行开始, 蜂王(最优解)按一定概率从蜂群中选择一组雄蜂(即精子), 形成受精囊.
- (2) 随机从受精囊中选取一组雄蜂与蜂王交配, 生成幼蜂(备选解).
- (3) 对生成的幼蜂的基因型进行突变处理, 生成一组新的幼蜂.
- (4) 比较幼蜂与蜂王的适应值的大小, 若存在适应值大于蜂王的幼蜂, 则替换蜂王(更新操作).
- (5) 检查当前的蜂群规模与蜂王的受精囊容量是否达到预定值, 若达到预先定义的值, 则终止算法, 否则, 循环上述步骤.

(5) 检查当前的蜂群规模与蜂王的受精囊容量是否达到预定值, 若达到预先定义的值, 则终止算法, 否则, 循环上述步骤.

2.2 对 HBMA 算法的改进

2.2.1 繁殖概率和突变概率的改进

HBMA 算法的步骤(2)和步骤(3)中只是简单

定义了固定值繁殖概率 P_c 和突变概率 P_m , 实际上 P_c 和 P_m 对算法的全局搜索能力和快速收敛能力起至关重要的作用^[7]. P_c 值越大, 新个体产生的速度越大, 未能体现幼蜂对雄蜂和蜂王基因型的继承, P_c 值越小, 新个体产生的速度越慢, 从而使得搜索速度变得缓慢; P_m 值越大, 幼蜂发生突变的几率越大, 使得搜索过程变成纯粹的随机搜索, P_m 值越小, 幼蜂发生突变的几率越小, 不容易产生新的个体^[8]. 故本文提出了一种自适应的 P_c 值、 P_m 值的定义方法, 用以提高算法中个体更新能力和最优个体的搜索能力. 对 P_c 和 P_m 的改进方法如式(3)和(4)所示.

其中, $f_{\max}$ 表示当前蜂群个体中的最大适应值, f_{avg} 表示当前蜂群中所有个体的平均适应值, f 表示当前突变个体的适应值, f' 表示进行繁殖过程的雄蜂和蜂王中较大的适应值. 文献[9]中对 P_{cmax} 、 P_{cmin} 、 P_{mmax} 、 P_{mmin} 进行了讨论, 这里采用该文献结果, 令 $P_{\text{cmax}} = 0.9$, $P_{\text{cmin}} = 0.6$, $P_{\text{mmax}} = 0.1$, $P_{\text{mmin}} = 0.01$.

$$P_c = \begin{cases} P_{\text{cmax}} - \frac{(P_{\text{cmax}} - P_{\text{cmin}})(f' - f_{\text{avg}})}{f_{\max} - f_{\text{avg}}}, & f' \geq f_{\text{avg}} \\ P_{\text{cmax}}, & f' < f_{\text{avg}} \end{cases} \quad (3)$$

$$P_m = \begin{cases} P_{\text{mmax}} - \frac{(P_{\text{mmax}} - P_{\text{mmin}})(f - f_{\text{avg}})}{f_{\max} - f_{\text{avg}}}, & f \geq f_{\text{avg}} \\ P_{\text{mmax}}, & f < f_{\text{avg}} \end{cases} \quad (4)$$

2.2.2 精英保留策略

HBMA 算法第(4)步中幼蜂对蜂王的替换是进行整体的更换, 而未将幼蜂个体和蜂王个体的适应值进行比较, 从而增加了寻优过程的负担. 本文应用精英保留策略, 对幼蜂群体和蜂王群体中的个体进行适应值的比较, 保留蜂王中较优的个体, 并用幼蜂中的较优个体替换蜂王中的其余个体, 从而保证了当前最优个体不被替换、突变, 理论上提高了算法的寻优能力, 缩短了寻优时间^[9].

3 Tsallis 熵多阈值分割原理

Tsallis 上多阈值分割原理如下: 被分割的原图像 I 有 L 个灰度级, 即这些灰度级在范围 $\{0, 1, 2, \dots, L\}$ 内. 然后定义 $P_i = h(i)/N$, 且 i 在范围 $[0, L]$ 内, 其中 $h(i)$ 表示图像 I 中灰度级为 i 的像素数量,

N 表示图像 I 中像素的总数. 假设待分割图像 I 选取 c 个阈值, 阈值分别为 $[t_1, t_2, \dots, t_c]$, 则适应值标准为使下述的目标函数最大化^[10]:

$$f([t_1, t_2, \dots, t_c]) = H_0 + H_1 + H_2 + \dots + H_c \quad (5)$$

$$\begin{aligned} w_0 &= \sum_{i=0}^{t_1-1} p_i, H_0 = \frac{1 - \sum_{i=0}^{t_1-1} \left(\frac{p_i}{w_0}\right)^q}{q-1} \\ w_1 &= \sum_{i=t_1}^{t_2-1} p_i, H_1 = \frac{1 - \sum_{i=t_1}^{t_2-1} \left(\frac{p_i}{w_1}\right)^q}{q-1} \\ w_2 &= \sum_{i=t_2}^{t_3-1} p_i, H_2 = \frac{1 - \sum_{i=t_2}^{t_3-1} \left(\frac{p_i}{w_2}\right)^q}{q-1}, \dots \\ w_c &= \sum_{i=t_c}^{L-1} p_i, H_c = \frac{1 - \sum_{i=t_c}^{L-1} \left(\frac{p_i}{w_c}\right)^q}{q-1} \end{aligned} \quad (6)$$

4 THBMO 算法原理

THBMO 算法流程图如图 1 所示.

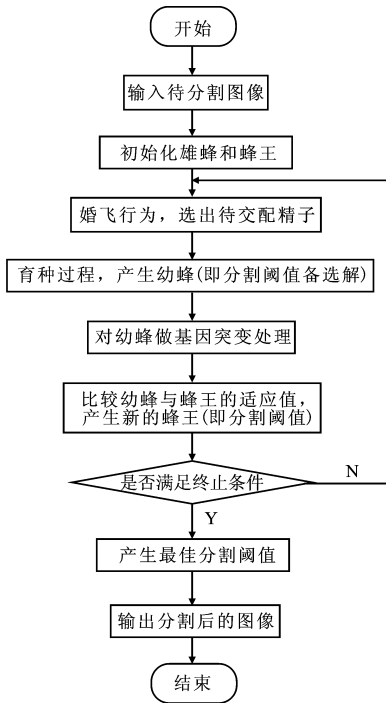


图 1 算法流程图

详细步骤如下:

(1) 初始化雄蜂和蜂王

初始化 m 组雄蜂, 且每一组的雄蜂数量为 c , 由矩阵 D_{mc} 表示. 其中 d_i^j 表示第 i 组雄蜂中第 j 个雄蜂, 且其取值范围为 $[0, 1, \dots, L]$, 且 $d_i^j < d_i^{j+1}$, 如式

(7) 所示. 在所有的雄蜂组 D_h 中选择适应最大的一组作为蜂王 Q , 即最佳分割阈值. 蜂王的表示方法式 (8) 所示.

$$D = [D_1, D_2, \dots, D_m], \quad (7)$$

$$D_i = (d_i^1, d_i^2, \dots, d_i^c)$$

$$Q = [X_{best}^1, X_{best}^2, \dots, X_{best}^c] \quad (8)$$

初始化过程即将 $[0, 255]$ 像素值分为 m 组, 每组 c 个像素值, 且从小到大排列, 选取一组适应值最大的雄蜂作为初始的最佳阈值.

(2) 婚飞行为

使用模拟退火算法来选择最佳雄蜂与蜂王 Q 进行交配, 即被选取的雄蜂 D_k 的模拟退火函数值 $P(Q, D)$ 最大. 婚飞行为之后, 蜂王的速度和能量将依据公式 (2) 降低. 此后婚飞行为将继续, 直到蜂王精囊中的精子数达到精囊容量 $nsperm$, 且精囊容量 $nsperm$ 的值是预先设定的, 其数值必须小于雄蜂的数量. 被选择的精子的表示方法如式 (9), 其中 Sp_i 表示蜂王精囊中的第 i 个个体.

$$Sp_i = (X_i^1, X_i^2, \dots, X_i^c) \quad (9)$$

婚飞过程即从步骤 1 生成的 m 组像素值中依据退火函数值从大到小的顺序选取 $nsperm$ 组作为备选分割阈值.

(3) 育种过程

幼蜂产生于蜂王和蜂王精囊中个体间的交配行为. 在育种过程中, 如果精囊中的第 j 个个体相应的随机数 R_i 小于当前繁殖率 P_c , 则 Sp_j 将被选中与蜂王交配. 蜂王和雄蜂的基因型将通过下式传递给第 j 个幼蜂个体 $Brood_j$, 参数 β 是每一次育种过程中随机生成的, 且其取值范围为 $[0, 1]$.

$$Brood_j = Q \pm \beta \times (Sp_j - Q) \quad (10)$$

幼蜂可表示为:

$$Brood_j = (Brood_j^1, Brood_j^2, \dots, Brood_j^c) \quad (11)$$

育种过程即对最佳分割阈值的扰动过程, 幼蜂表示扰动结束后生成的 $nsperm$ 组备选分割阈值. 此步骤中的繁殖率 P_c 是依据公式 (3) 计算得出的.

(4) 基因突变

幼蜂的基因型将通过基因突变得以提高, 具体步骤如下:

① 每一个幼蜂将对应一个随机生成数 R_i .

② 如果 R_i 小于当前突变率 P_m , 则第 i 个幼蜂将根据下式进行突变, 其中 $\delta \in [0, 1]$, $0 < \epsilon < 1$

$$Brood_i^k = Brood_i^k \pm (\delta + \epsilon) \times Brood_i^k \quad (12)$$

③ 选取适应值最大的幼蜂 $Brood_j$ 作为蜂王的候选.

④ 如果 $Brood_j$ 的适应值大于蜂王的适应值,则用 $Brood_j$ 中较优个体替换蜂王中较差个体,保留目前蜂王中的优秀个体,即精英保留策略。

基因突变即对上一步中生成的 $nsperm$ 组分割阈值的进一步优化,以避免局部最优。此步骤中的繁殖率 P_m 是依据公式(4) 计算得出的。

(5)检查终止条件

如果满足算法的终止条件(包括精囊容量和蜂群规模),则该迭代完成。否则将丢弃先前的实验结果(即生成的幼蜂集合),转到步骤 2,直到迭代完成。

5 实验研究与分析

本文采用 MATLAB2014 编码,并在 window7 系统中使用具有 2.30 GHz CPU 和 2 GB RAM 的个人计算机实现仿真。实验采用了不同种类的图像来证明本文算法的优越性,下面仅列出了图像“lena”的实验结果。

5.1 q 值的设置

图 2 中列出了当分割阈值数为 2 时,图像“lena”、在不同 q 取值时的分割结果。其中, x 轴、 y 轴表示图像分割时阈值的取值, z 轴表示不同的 q 值,游标表示图像达到最佳分割阈值时所对应的 x 、 y 、 z 的取值。

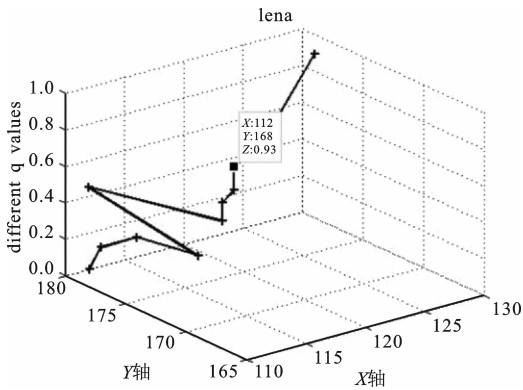


图 2 不同 q 取值的分割结果

从图 2 可以看出,当图像的分割效果最佳时, q 取值为 0.93。大量实验证明 q 的最佳取值范围为 (0.8,1),故本文仿真实验中将 q 统一赋值为 0.98。

5.2 算法中参数的设置

对算法仿真所需的参数进行设定,如表 1 所示。其中前八项参数的设置参考自文献[2],四项参数的设置已在本文 2.2 小节中进行了说明。

5.3 $C=4$ 分割结果与对比

图 3 列出当阈值个数为 4 时,本文算法 THB-

MO、BHBMA、MEHBMOT、CHBMA、HBMA 对图像“lena”的分割结果。

表 1 算法中涉及到的参数

参数	描述	取值
	蜂王的数量	1
m	雄蜂的数量	150
C	阈值的个数	1,2,3,4
L	图像非灰度	255
α	飞行速度减小因子	0.98
$nsperm$	精囊的容量	50
$S(0)$	初始飞行速度	1.0
ϵ	突变因子	0.5
P_{cmax}	最大繁殖概率	0.9
P_{cmin}	最小繁殖概率	0.6
P_{mmax}	最大突变概率	0.1
P_{mmin}	最小突变概率	0.01



图 3 $C=4$ 时的图像“lena”分割结果比

从图 3 可以看出,本文算法 THBMO 对“lena”图像中人物阴影部分的发丝、发尾分割的更佳明显,且人物的面部特征保留完好,分割效果更佳。

为了进一步证明该算法的优越性,比较不同分割方法的分割阈值、计算时间和峰值信噪比 (PSNR)。PSNR 是通过均方差 MSE 定义的,峰值信噪比越大,表示图像的分割效果越好。PSNR 的计算方法如下:

$$PSNR = 20 \times \log_{10} \left(\frac{255}{\sqrt{MSE(x,y)}} \right) \quad (13)$$

$$MSE(x,y) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x(i,j) - y(x,j))^2 \quad (14)$$

表 2 列出不同分割方法对图像“lena”在阈值 C

=4 时的结,表中加粗数据为计算时间和 PSNR 结果较优的一方.

表 2 C=4 时阈值、计算时间、PSNR 的比较

图像	方法	分割阈值	计算时间/s	PSNR/dB
lena	THBMO	(74,112,144,179)	1.14	21.98
	BHBMA	(56,100,144,180)	1.11	21.04
	MEHBMOT	(56,100,144,182)	1.20	19.71
	CHBMA	(56,100,144,169)	1.17	19.89
	HBMA	(56,100,149,208)	1.20	19.71

从表 2 可以看出,本文算法具有更好的分割效果,在计算时间上也较优.

5.4 收敛能力和寻优能力

为了提高算法的全局搜索能力和快速收敛性, THBMO 算法提出繁殖概率和突变概率的自适应定义方法,下面将通过图 4 来说明该算法的优越性. 图 4 表示图像“lena”在分割阈值数为 4 时,不同算法的寻优迭代次数. 其中,横坐标表示迭代次数,纵坐标表示适应值.

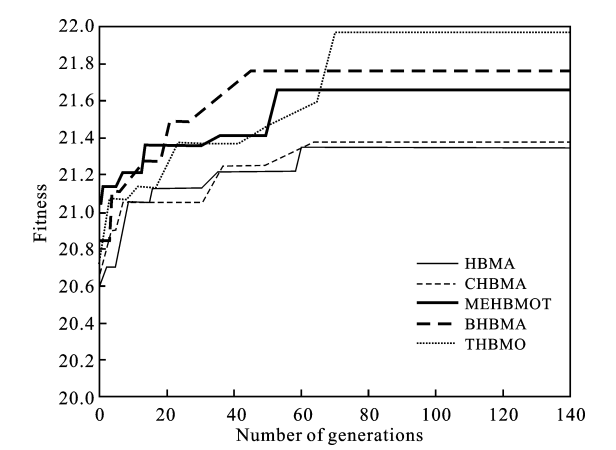


图 4 C=4 时图像“lena”收敛性与适应值的比较

从图 4 可以看出:当算法收敛时, THBMO 算法最优解的适应值大于其余四种算法的适应值,且迭代次数约为 80,故 THBMO 算法具有较好的寻优能力和收敛能力.

6 结束语

针对传统蜂群繁殖算法中现阶段存在的问题,本文引入了自适应的繁殖概率和突变概率以及精英保留策略,以提高算法的全局搜索能力和快速收敛能力,并结合 Tsallis 熵的非广延性,避免了图像分割过程中非可加信息被忽略,从而提高分割精度. 大量实验证明该算法在全局搜索和快速收敛能力、图像的分割精度和分割速度等方面都有了进一步的提升,且分割后的图像特征保留完好,具有较好的实用价值.

参考文献:

[1] 肖林云,陈秀宏,林喜兰. 邻域信息熵的核模糊 C 均值聚类图像分割算法[J]. 微电子学与计算机, 2016, 33(11):166-170.

[2] Horng M H. A multilevel image thresholding using the honey bee mating optimization [J]. Applied Mathematics & Computation, 2010, 215(9):3302-3310.

[3] Jiang Y, Yeh W C, Hao Z, et al. A cooperative honey bee mating algorithm and its application in multi-threshold image segmentation [J]. Information Sciences, 2016, 369(1):171-183.

[4] Jiang Y, Tsai P, Yeh W C, et al. A honey-bee-mating based algorithm for multilevel image segmentation using Bayesian theorem [J]. Applied Soft Computing, 2017, 52 (C) :1181-1190.

[5] Marinakis Y, Marinaki M. A hybrid honey bees mating optimization algorithm for the probabilistic traveling salesman problem[C]// Eleventh Conference on Congress on Evolutionary Computation. Trondheim, Norway, IEEE Press, 2009:1762 - 1769.

[6] Abbass H A. MBO: marriage in honey bees optimization-a Haplometrosis polygynous swarming approach [C]// Evolutionary Computation, 2001. Proceedings of the 2001 Congress on. [s. l.], IEEE, 2015: 207-214.

[7] Srinivas M, Patnaik L M. Adaptive probabilities of crossover and mutation in genetic algorithms [J]. IEEE Transactions on Systems Man & Cybernetics, 1994, 24(4):656-667.

[8] Jiang J, Jiang T, Zhai S. A novel recognition system for human activity based on wavelet packet and support vector machine optimized by improved adaptive genetic algorithm [J]. Physical Communication, 2014, 13(c): 211-220.

[9] Ren Z W, San Y. Improved adaptive genetic algorithm and its application research in parameter identification [J]. Journal of System Simulation, 2006, 18(1):41-40.

[10] 邸秋艳. 基于 Tsallis 熵的阈值图像分割方法研究[D]. 河北,秦皇岛:燕山大学,2009:19-21.

作者简介:

马巧梅 女,(1969-),博士,副教授. 研究方向为图形图像处理、信息安全技术.

李婧婧(通讯作者) 女,(1992-),硕士研究生. 研究方向为图形图像处理. E-mail:814331631@qq. com.

康 珺 女,(1978-),博士,副教授. 研究方向为人工智能、系统建模.