

摸石头过河算法与分布估计混合算法

高 尚¹ 曹存根²

(江苏科技大学计算机科学与工程学院 镇江 212003)¹

(中国科学院计算技术研究所智能信息处理重点实验室 北京 100190)²

摘 要 依据摸石头过河算法与分布估计算法的优点,提出了一种混合算法。该算法以一个解为起点,向该起点附近邻域随机搜索若干个解,找出这些解中最好的一个解;并挑选部分优秀个体的中心与最好解进行交叉操作,以此解作为下次迭代的结果,然后以此点为起点,再向附近邻域随机搜索若干个解,以此类推。对几个经典测试函数进行实验的结果表明,利用摸石头过河与分布估计算法能够极大地提高收敛速度和精度。

关键词 随机优化算法,连续空间优化,摸石头过河,分布估计算法

中图分类号 TP18 文献标识码 A DOI 10.11896/j.issn.1002-137X.2016.1.019

Solving Continuous Optimization Problem by Hybrid Wading across Stream Algorithm-Estimation Distribution Algorithm

GAO Shang¹ CAO Cun-gen²

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)¹

(Key Laboratory of Intelligent Information Processing Institute of Computing Technology,

Chinese Academy of Sciences, Beijing 100190, China)²

Abstract According to advantage of wading across the stream and estimation distribution algorithm, a hybrid wading across stream algorithm-estimation distribution algorithm (Hybrid WSA-EDA) was put forward. The Hybrid WSA-EDA acts a solution as a start point, then searches several random solutions near the start point, and finds out the best solution of these solutions. The center of some selection of good individual is crossed with the best solution and this solution is taken as the next start point, and then several random solutions near this start point are searched, and so on. For solving continuous optimization problem, the improved wading across stream algorithm shrink the search space gradually. The experiment results of some classic benchmark functions show that the Hybrid WSA-EDA extraordinarily improves the convergence velocity and precision.

Keywords Random optimization algorithm, Continuous space optimization, Wading across stream algorithm, Estimation distribution algorithm

1 引言

对于非线性规划,目前还没有适合各种问题的一种统一解法,各种方法都有自己特定的适用范围。对于传统的解析法,其要求目标函数与约束函数具有连续性并且其导数存在。但在某些实际问题中,由于目标函数很复杂,有时甚至无法写出其表达式,当然更无法求其导数,因此解析法就不再适用了。目前采用智能优化算法进行求解比较流行,基于随机搜索技术的进化算法^[1]、模拟退火算法^[2,3]、蚁群优化算法^[4]、粒子群优化算法^[5-9]和混沌优化算法等优化算法^[10,11]在解决全局最优问题中得到广泛应用。但这些智能算法都存在某些缺陷,如需要增加一些额外的运算,进化算法中需要变异和选择操作等,这导致计算复杂性高、实时性差、搜索精度差、易陷

于局部最优解。本文依据摸石头过河算法,结合分布估计算法,提出一种混合分布估计算法,来解决连续优化问题。与其他优化法相比,该算法没有额外计算,降低了计算复杂度。

2 摸石头过河算法的思想

摸石头过河算法(Wading across Stream Algorithm, WSA)的思想来源于“摸石头过河”的思想^[12]:摸到一个“石头”后,向该“石头”周围摸索其它石头,当再摸到一个“石头”后,再向该“石头”周围摸索其它石头,以此类推进行搜索。摸石头过河算法以一个解为起点,向该起点附近邻域随机搜索若干个解,找出这些解中的最好的一个解,以此解作为第 2 次迭代的结果。然后以此点为起点,再向附近邻域随机搜索若干个解,找出这些解中的最好的一个解,以此解作为第 3 次迭

到稿日期:2015-03-20 返修日期:2015-05-01 本文受人工智能四川省重点实验室开放基金(2012RYJ04),中科院智能信息处理重点实验室开放课题(IIP2013-1)资助。

高 尚(1972—),男,博士,教授,主要研究方向为智能计算, E-mail: gaoshang@sohu.com; 曹存根(1964—),男,博士,研究员,主要研究方向为大规模知识获取和共享、计算机辅助教学、情感计算等。

代的结果。后面的步骤以此类推,直到达到最大迭代次数或其它停止条件为止。其迭代过程如图1所示。

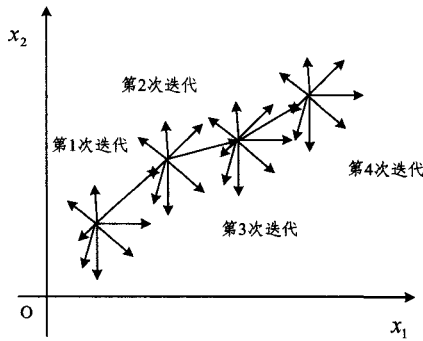


图1 “摸石头过河”算法迭代过程

摸石头过河算法与模拟退火算法有点类似,但效果比模拟退火算法好,并且算法比模拟退火算法简单。模拟退火算法是从1个解搜索下一个解,而摸石头过河算法是从周围若干个解中摸索比较好的解作为下一个解,利用了“群”优势。模拟退火算法根据 Metropolis 准则来接受或舍弃下一个解,而摸石头过河算法直接选取周围若干个解中比较好的解为下一个解,操作简单。

3 分布估计算法

分布估计算法(Estimation Distribution Algorithm, EDA)的概念最初由 Muhliebe H 和 Paass G 在 1996 年提出,分布估计算法是在遗传算法的基础上发展起来的一种全新的进化模式。在传统的遗传算法中,用种群表示优化问题的一组候选解,种群中的每个个体都有相应的适应值(目标值),然后根据适应值大小进行选择、交叉和变异等模拟自然进化的操作,反复进行,对问题进行求解;而在分布估计算法中,没有传统的交叉、变异等遗传操作,取而代之的是概率模型的学习和采样,分布估计算法通过一个概率模型表示候选解在空间的分布,采用统计学习手段从群体宏观的角度建立一个描述解分布的概率模型,然后对概率模型随机采样产生新的种群,如此反复进行,实现种群的进化,直到终止条件。根据概率模型的复杂程度以及不同的采样方法,分布估计算法发展了很多不同的具体实现方法,但是它们都可以归纳为下面两个主要步骤:首先建立描述解空间的概率模型,通过对种群的评估,选择优秀的个体集合,然后采用概率统计等手段构造一个描述当前解集的概率模型;其次,由概率模型随机采样产生新的种群,一般采用蒙特卡罗方法,对概率模型采样得到新的种群。遗传算法中的交叉和变异会破坏已经进化好的个体,分布估计算法用建立概率模型和采样样本两大操作取代了遗传算法中的交叉算子和变异算子,以一种带有“全局操控”性的操作模式解决了遗传算法中存在的缺陷。而且分布估计算法不需要太多的参数设置,编程比遗传算法简单。

分布估计算法的基本步骤如下:

步骤1 在搜索空间内按均匀分布随机产生 N 个点,组成初始群体。

步骤2 根据适应值评价函数计算群体中的各点的适应值,保留最好解。

步骤3 根据适应值,运用一定的选择策略选出适应值

较好的 m 个个体组成优势群体。

步骤4 估计优势群体的概率分布模型。

步骤5 根据估计的概率模型进行采样,产生一些新个体。

步骤6 若满足某种停止准则,则算法结束,群体中的最好个体就是优化的结果;否则算法转到步骤2继续执行。

4 解连续性优化问题的摸石头过河算法与分布估计混合算法

4.1 基本摸石头过河算法

连续空间优化问题可表示为:

$$\begin{aligned} \min & f(x_1, x_2, \dots, x_n) \\ \text{s. t. } & a_i \leq x_i \leq b_i (i=1, 2, \dots, n) \end{aligned} \quad (1)$$

解连续空间优化问题的摸石头过河算法如下:

步骤1 设置算法参数:搜索解的个数 m , 邻域半径 L , 迭代次数 $n_{\max}$ 。

步骤2 随机产生一个解 $(x_1^*, x_2^*, \dots, x_n^*)$, 计算其目标值 f^* , $k=0$ 。

步骤3 在解 $(x_1^*, x_2^*, \dots, x_n^*)$ 邻域半径 L 内产生 m 个邻域解:

$$X' = \begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1n} \\ x'_{21} & x'_{22} & \dots & x'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{m1} & x'_{m2} & \dots & x'_{mn} \end{bmatrix}$$

其中 $x'_{ij} = x_j^* + L \cdot r_{ij}$, r_{ij} 为在区间 $[-1, 1]$ 内均匀分布的随机数。

步骤4 计算这 m 个解的目标函数值 $f'_1, f'_2, \dots, f'_m$, 再找出最好解 $(x_1^*, x_2^*, \dots, x_n^*)$ 及其目标值 f^* , 保留最好解, $k=k+1$ 。

步骤5 若 $k > n_{\max}$, 则算法结束, 输出保留的最好解; 否则执行步骤3。

该摸石头过河算法的时间复杂性估算如下: 因为计算目标函数值的操作花时间最多, 所以时间复杂性大约为 $O(m \cdot n_{\max})$ 。

4.2 摸石头过河与分布估计混合算法

解连续空间优化问题的摸石头过河与分布估计混合 (Hybrid WSA-EDA) 算法如下:

步骤1 设置算法参数:搜索解的个数 m , 邻域半径 L , 迭代次数 $n_{\max}$;

步骤2 随机产生一个解 $(x_1^*, x_2^*, \dots, x_n^*)$, 计算其目标值 f^* , $k=0$ 。

步骤3 在解 $(x_1^*, x_2^*, \dots, x_n^*)$ 邻域半径 L 内产生 m 个邻域解:

$$X' = \begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1n} \\ x'_{21} & x'_{22} & \dots & x'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{m1} & x'_{m2} & \dots & x'_{mn} \end{bmatrix}$$

其中 $x'_{ij} = x_j^* + L \cdot r_{ij}$, r_{ij} 为区间 $[-1, 1]$ 内均匀分布的随机数。

步骤4 计算这 m 个解的目标函数值 $f'_1, f'_2, \dots, f'_m$, 再找出最好解 $X^* (x_1^*, x_2^*, \dots, x_n^*)$ 及其目标值 f^* , 保留最好

解,按适应度从高到低的顺序对种群进行排序,并从中选出最优的 N 个个体($N \leq m$),计算这 N 个个体所包含的中心点 $\bar{X}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$,与最好解进行交叉操作,修改最好点($x_1^*, x_2^*, \dots, x_n^*$)的位置, $X^* \leftarrow \alpha X^* + (1-\alpha)\bar{X}$, α 为 $[0, 1]$ 的随机数, $k = k + 1$ 。

步骤 5 若 $k > n_{\max}$,则算法结束,输出保留的最好解;否则执行步骤 3。

4.3 数值仿真与分析

4.3.1 与随机优化算法比较

摸石头过河算法和摸石头过河与分布估计混合算法实质上都计算了 $m \cdot n_{\max}$ 个目标函数值,从中找出最好解。随机优化算法(Random Optimization Algorithm, ROA)的思路是从解空间中直接随机产生 $m \cdot n_{\max}$ 个解,计算这 $m \cdot n_{\max}$ 个目标函数值,从中找出最好解。

为了说明摸石头过河与分布估计混合算法的优势,将其与随机优化算法、摸石头过河算法进行比较。以下面的测试函数为例。

对于测试函数:

$$\min F_0 = \sum_{i=1}^4 [x_i^2 - 10 \cos(2\pi x_i) + 10]$$

$$\text{s. t. } -1 \leq x_i \leq 1, i=1, 2, 3, 4$$

当 $x_1 = x_2 = x_3 = x_4 = 0$ 时, $F_{\min} = 0$ 。

将摸石头过河和摸石头过河与分布估计混合算法参数设置如下: $m=200, n_{\max}=100, N=1000, L=0.1, \alpha=0.8$ 。对于随机优化算法产生的 $m \cdot n_{\max}=20000$ 个解,从中找出最好的解。算法各测试 100 次,统计数据如表 1 所列。 α 为 1 时,混合算法是摸石头过河算法,而 α 为 0 时,混合算法是分布估计算法。

表 1 3 种算法的比较

算法	最好解	最坏解	平均值
随机优化算法(ROA)	0.07185	2.87820	1.53764
摸石头过河算法(WSA)	0.00263	0.05002	0.022111
摸石头过河-分布估计混合算法(HWSA-EDA)	0.000368	0.048462	0.001985

从表 1 可以看出,摸石头过河与分布估计混合算法的效果比摸石头过河算法和随机优化算法效果好很多。

4.3.2 收敛性比较

用以下 4 个经常被国内外学者用来测试优化算法有效性的测试函数^[10]:

$$\min F_1 = \sum_{i=1}^{30} x_i^2, -1 \leq x_i \leq 1, i=1, 2, \dots, 30$$

$$\min F_2 = \sum_{i=1}^{30} |x_i| + \prod_{i=1}^{30} |x_i|, -1 \leq x_i \leq 1, i=1, 2, \dots, 30$$

$$\min F_3 = \max_{1 \leq i \leq 30} |x_i|, -1 \leq x_i \leq 1, i=1, 2, \dots, 30$$

$$\min F_4 = \sum_{i=1}^{30} (\sum_{j=1}^i x_j)^2, -1 \leq x_i \leq 1, i=1, 2, \dots, 30$$

与当前最流行的粒子群优化算法(Particle Swarm Optimization Algorithm, PSO)作比较。摸石头过河算法和摸石头过河与分布估计混合算法参数设置如下: $m=200, n_{\max}=200, N=1000, L=0.1, \alpha=0.8$ 。粒子群算法的参数设置如下:粒子数为 50, $c_0=1, c_1=2, c_2=2, V_{\max}=0.5$ 。3 种算法的迭代过程如图 2 所示。从图 2 可以看出,摸石头过河与分布估计混合算法的收敛速度最快。

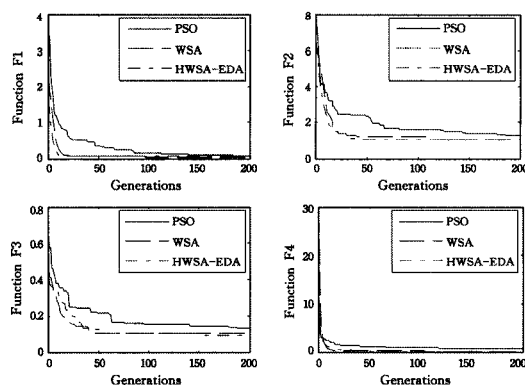


图 2 3 种算法的比较

结束语 依据摸石头过河算法和分布估计算法,提出一种混合优化算法来求解连续空间优化问题,该算法易于设计和实现。测试结果表明,该算法具有搜索速度快、精度高和不易陷入局部极值点的特点,因而具有较好的全局搜索能力,其应用前景非常广泛。该方法具有一定潜力,值得推荐。

参考文献

- [1] Antamoshkin A N, Kazakovtsev L A. Random search algorithm for the p-median problem[J]. Informatica (Slovenia), 2013, 37 (3): 267-278
- [2] Ge Zhen-zhen, Zhou Jun, Lin Peng. The Control-Configured Optimization Method for Hypersonic Vehicles Using the Improved Simulated Annealing Algorithm[J]. Journal of Astronautics, 2013, 34(11): 1427-1433 (in Chinese)
葛振振, 周军, 林鹏. 采用改进模拟退火算法的高速飞行器随控总体优化方法[J]. 宇航学报, 2013, 34(11): 1427-1433
- [3] Li Yuan, Jiang Jiang, Zhang Min-xuan, et al. Automatic Word-Length Determination Tool Based on Simulated Annealing Algorithm[J]. Journal of Shanghai Jiaotong University, 2013, 47 (1): 76-80, 85 (in Chinese)
黎渊, 蒋江, 张民选, 等. 基于模拟退火算法的浮点转定点自动位宽优化工具[J]. 上海交通大学学报, 2013, 47(1): 76-80, 85
- [4] Zhang Qi, Ma Jia-cheng, Xie Wei, et al. Improved Ant Colony Algorithm-Based Path Planning for Mobile Robot[J]. Journal of Northeastern University (Natural Science), 2013, 34(11): 1521-1524 (in Chinese)
张琦, 马家辰, 谢玮, 等. 基于改进蚁群算法的移动机器人路径规划[J]. 东北大学学报(自然科学版), 2013, 34(11): 1521-1524
- [5] Eberhart R C, Kennedy J. A new optimizer using particles swarm theory[C]//Proc Sixth International Symposium on Micro Machine and Human Science. Nagoya, Japan, 1995: 39-43
- [6] Shi Y H, Eberhart R C. A modified particle swarm optimizer [C]//IEEE International Conference on Evolutionary Computation. Anchorage, Alaska, 1998: 69-73
- [7] Kennedy J, Eberhart R. Particle swarm optimization [C]//Proc IEEE International Conference on Neural Networks. Perth, 1995: 1942-1948
- [8] Guo Tong, Lan Ju-long, Li Yu-feng, et al. Adaptive fractional-order Darwinian particle swarm optimization algorithm [J]. Journal of Communications, 2014, 35(4): 130-140 (in Chinese)
郭通, 兰巨龙, 李玉峰, 等. 自适应的分阶达尔文粒子群优化算法[J]. 通信学报, 2014, 35(4): 130-140
- [9] Hu Wang, Yen G G, Zhang Xin. Multiobjective Particle Swarm Optimization Based on Pareto Entropy[J]. Journal of Software,

- [10] Gao Shang, Yang Jing-yu. Swarm intelligence algorithm and its application [M]. Beijing: China Water and Power Press, 2006: 112-117 (in Chinese)
高尚, 杨静宇. 群智能算法及其应用[M]. 北京: 中国水利水电出版社, 2006: 112-117

(上接第 56 页)

从图 8(a)和图 8(b)可以看出,两者存在仿射差异,在曲率图上目标图像的轮廓长度发生变化,并且曲率的值也发生相应的变化,总体趋势具有一致性;如图 8(c)和图 8(d)可以看出,当两个目标存在旋转差异时,轮廓长度没有变化,由于计算方向发生变化,使得曲率值发生变化。

图 9(a1)、(b1)、(c1)、(d1)为实验样本对二配准之后叠加在一起显示的结果,可以方便地看出两个配准的吻合程度。

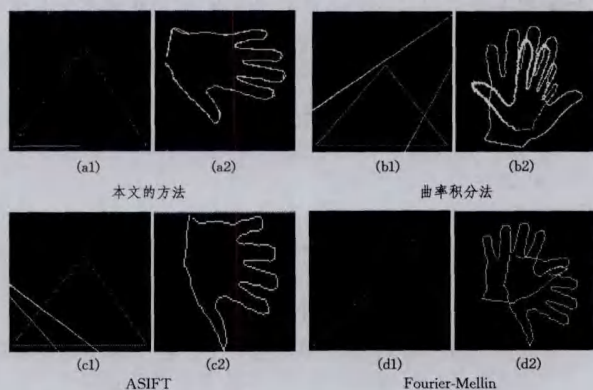


图 9 Kimia 样本配准结果

从上述实验结果的对比中可以看出,本文的方法可以成功匹配。曲率积分法由于在曲率积分图中根据最大互相关找到最佳偏移位置时存在误差,因此导致配准存在误差。ASIFT、Fourier-Mellin 方法由于轮廓线所含信息较少,造成匹配精度下降。从表 1 中算法检测结果可以看出,本文方法在互相关程度上具有较好优势, RMSE 误差较小。

表 1 算法检测结果

算法	实验样本对一		实验样本对二	
	NCC	RMSE	NCC	RMSE
本文的方法	0.706	1.086	0.786	0.693
曲率积分法	0.605	2.820	0.461	1.851
ASIFT	0.004	14.542	0.055	1.414
Fourier-Mellin	0.342	461.06	0.0926	552.291

结束语 本文提出基于曲率的全仿射曲线图像匹配方法,对样本图像根据全仿射模型建立样本库,对样本库中每一个样本通过等间隔偏移进行采样后的子样本建立曲率样本子集,将目标图像与样本子集中的每一个样本进行快速最近邻搜索算法相似性匹配,通过匹配结果找到最优匹配样本。实验结果表明,基于曲率的全仿射曲线图像匹配具有较高的匹配成功率,并且与传统算法相比具有更多的优点。

参考文献

- [1] Turk M A, Pentland A P. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86

- [11] Gao Shang, Yang Jing-yu. Research on Chaos Particle Swarm Optimization Algorithm[J]. Pattern Recognition and Artificial Intelligence, 2006, 19(2): 266-270 (in Chinese)
高尚, 杨静宇. 混沌粒子群优化算法研究[J]. 模式识别与人工智能, 2006, 19(2): 266-270
[12] Gao Shang, Yu Hua-long, Qiu Ling, et al. The wading across stream algorithm[J]. International Journal of Computers and Applications, 2014, 36(4): 127-132

- [2] Chen Q, Defrise M, Deconinck F. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(12): 1156-1168
[3] Li X M, Zhang J F. An Evaluation Method for Image Registration by Machine Learning[J]. Acta Automatica Sinica, 2008, 34(1): 7-13 (in Chinese)
李晓明, 张继福. 一种基于学习的自动图像配准检验方法[J]. 自动化学报, 2008, 34(1): 7-13
[4] Sapiro G, Tannenbaum A. Affine invariant scale-space[J]. International Journal of Computer Vision, 1993, 11(1): 25-44
[5] Mokhtarian F, Mackworth A. Scale-based description and recognition of planar curves and two-dimensional shapes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986(1): 34-43
[6] Mokhtarian F, Abbasi S. Shape similarity retrieval under affine transforms[J]. Pattern Recognition, 2002, 35(1): 31-41
[7] He X C, Yung N H C. Curvature scale space corner detector with adaptive threshold and dynamic region of support [C]// Proceedings of the 17th IEEE International Conference on Pattern Recognition (ICPR'04). Aug 2004, 2: 791-794
[8] Cui M, Femiani J, Hu J, et al. Curve matching for open 2D curves[J]. Pattern Recognition Letters, 2009, 30(1): 1-10
[9] Lim K B, Du T H, Wang Q. Partially occluded object recognition [J]. International Journal of Computer Applications in Technology, 2011, 40(1): 122-131
[10] Zhang Y S, Zou Z R. Automatic registration method for remote sensing images based on improved ORB algorithm [J]. Remote Sensing for Land and Resources, 2013, 25(3): 20-24 (in Chinese)
张云生, 邹峥嵘. 基于改进 ORB 算法的遥感图像自动配准方法[J]. 国土资源遥感, 2013, 25(3): 20-24
[11] Yuan X G, Peng G H, Wang L. GPU-based Real Time Image Registration with Variant SIFT[J]. Computer Science, 2011, 38(3): 300-303 (in Chinese)
袁修国, 彭国华, 王琳. 基于 GPU 的变型 SIFT 算子实时图像配准[J]. 计算机科学, 2011, 38(3): 300-303
[12] Morel J M, Yu G. ASIFT: A new framework for fully affine invariant image comparison [J]. SIAM Journal on Imaging Sciences, 2009, 2(2): 438-469
[13] Muja M, Lowe D G. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration [C]// Proc of the VISAPP International Conference on Computer Vision Theory and Applications, 2009: 331-340
[14] Kimia shape database [DB/OL]. <http://www.lems.Brown.edu/vision/software/index.html>