基于小波阈值收缩消噪的浮点数编码遗传算法研究

崔明义
CUI Ming-yi

河南财经学院 信息学院,郑州 450002
Information School,Henan University of Finance & Economics,Zhengzhou 450002,China
E-mail:mycui369@yahoo.com.cn

CUI Ming-yi.Research on floating point representation genetic algorithm based on wavelet threshold shrinkage denoising. Computer Engineering and Applications,2009,45(34):38-40.

Abstract: Floating Point Representation (FPR) is of the strongpoint of high precision and facilitating search on high-dimension space.It is superior to other representation in function optimization and restriction optimization.But,the noise in run environment has brought about with Floating Point Representation Genetic Algorithm (FPRGA).It has often neglected by researchers.Simple FPRGA uses bounded random mutation.It cannot avoid the noise to influence on the algorithm performance.This paper presents the Floating point representation Genetic Algorithm based on Wavelet threshold Shrinkage Denoising (FGAWSD).A filter is structured.Mutation operation is replaced with different thresholds denoising.The experiments are done.The result of the research and the experiments indicates that the method is reliable in theory,is feasible in technique.The precision of optimal solution of the algorithm can be enhanced by selecting proper threshold.The method is of high stability.

Key words: genetic algorithm;wavelet threshold;shrinkage denoising;mutation

摘 要:浮点数编码具有精度高、便于高维大空间搜索的优点,在函数优化和约束优化领域明显有效于其他编码。但浮点数编码遗传算法在运行环境中产生的噪音对算法性能的影响并未引起人们的重视。传统的浮点数编码遗传算法采用的是有界随机变异,不能消除噪音对算法性能的影响。提出了基于小波阈值收缩消噪的浮点数编码遗传算法,建立滤波器,采用不同的阈值消噪取代变异操作,并进行了实验。该研究和实验结果表明,这种方法理论上是可靠的,方法上是可行的,选择适当的阈值,可明显提高算法的全局最优解精度,具有较高的稳定性。

关键词:遗传算法;小波阈值;收缩消噪;变异
DOI:10.3778/j.issn.1002-8331.2009.34.012 **文章编号**:1002-8331(2009)34-0038-03 **文献标识码**:A **中图分类号**:TP301.6

1 引言

近几年,浮点数编码遗传算法(Floating Point Representation Genetic Algorithm,FPRGA)研究正在引起人们的重视。研究表明,浮点数编码具有精度高、便于高维大空间搜索的优点,在函数优化和约束优化领域明显有效于其他编码,这已被得到了广泛地验证[1-4]。在 FPRGA 研究中,噪音问题也越来越引起人们的关注。但由于 FPRGA 运行过程中噪音产生的机理和其对算法性能影响的复杂性,许多研究仅仅徘徊在对算法含噪解的处理上。Tsutsui 和 Ghosh[5-6]研究了遗传算法(Genetic Algorithm,GA)鲁棒解搜索模式,在性能上对含噪的鲁棒解搜索进行了分析。Kenneth[7]使用局部搜索元启发方法,找到 GA 的含噪鲁棒最优解。Kenneth 研究表明,当噪音服从标准正态分布时,其标准方差 σ 对算法性能的影响是十分敏感的。参数 σ 的稍微增大,就很难保证鲁棒解的最优性。显然,求 GA 的鲁棒解并不是解决 GA 噪音的最好方法,其应用的局限性也十分明显。

噪音来源于 FPRGA 的实际应用环境,在 FPRGA 运行过程中反映在个体编码和适应度评价过程中。要尽可能地使噪音对 FPRGA 性能的影响最小,较可靠的方法就是算法运行过程中的消噪。而小波在消除信号和图像噪音方面已有许多成功的应用,显示了小波在消噪领域的独特优势。将小波用于 FPRGA 的消噪却鲜有研究者涉及。2006 年,MA Qi-ming[8]等将基于小波变换阈值收缩(Wavelet Transform Threshold Shrinkage,WTS)平移不变阈值收缩(Translation-Invariant Threshold Shrinkage,TIS)的 GA 用于含噪函数的消噪。2007 年,S.H.Ling[9]等提出了一种具有新遗传操作的实数编码遗传算法(Real-Coded Genetic Algorithm,RCGA),使用平均有界交叉和小波变异实现遗传进化。这些成果,尽管都将小波用于 GA 的研究,但前者是将基于小波阈值收缩的 GA 用于函数消噪,后者是将小波变异用于 RCGA。均没有涉及将小波用于 FPRGA 的消噪研究。提出将小波收缩消噪引入 FPRGA(Floating Point Representation Genetic Algorithm Based on Wavelet Threshold Shrinkage Denoising,FGAWSD)的算法,通

基金项目:河南省基础与前沿技术研究计划(No.082300410100)。
作者简介:崔明义(1958-),男,教授,博士,研究方向:计算智能、软件和网络安全。
收稿日期:2008-11-03 **修回日期**:2008-12-08

过变异消除 FPRGA 运行环境中产生的噪音，并编程予以实现。研究和实验结果表明，该算法可显著地提高 FPRGA 的性能，理论上是可靠的，方法上是可行的。

2　FPRGA 的噪音特性和小波的性质

2.1　FPRGA 的噪音特性

定义 1　假设 FPRGA 在遗传运算环境中产生的噪音是一个随机过程 $M(t)$，对于每个 $t_i,t_j(t_i\neq t_j)$，如果 $M(t)$ 的取值 $M(t_i)$ 和 $M(t_j)$ 都不相关，即自协方差

$$C_{xx}(t_i,t_j)=0\quad t_i\neq t_j \tag{1}$$

则称随机过程 $M(t)$ 为白噪音。

这种白噪音污染了浮点数编码。设 $F^{(0)}$ 为不含噪音的浮点数编码，平稳且遍历，而遗传过程中得到的浮点数编码为 X，则

$$X=F^{(0)}+n^{(0)} \tag{2}$$

$n^{(0)}$ 为加性白噪音，且与 $F^{(0)}$ 不相关。假设 $n^{(0)}$ 为高斯白噪音，则它的概率密度函数满足正态分布统计特性。

2.2　小波的性质

设连续时间函数 $\psi(x)$ 为小波，则该函数满足以下性质：

$$(1)\int_{-\infty}^{\infty}\psi(x)\mathrm{d}x=0 \tag{3}$$

$$(2)\int_{-\infty}^{\infty}|\psi(x)|^2\mathrm{d}x<\infty \tag{4}$$

性质（1）说明 $\psi(x)$ 的总正动量和总负动量相等；性质（2）说明 $\psi(x)$ 的大多数能量都限制在有限的时空区域内。小波的这些性质与高斯白噪音的一些正态分布统计特性相吻合。

3　小波阈值收缩消噪

3.1　小波变换

设 $X(t)$ 为被噪音污染的浮点数编码，$\psi(t)$ 为小波基函数。假设 $\hat{\psi}(\omega)$ 是 $\psi(t)$ 的傅里叶变换，则以下方程成立

$$C_w=\int_0^{\infty}\frac{|\hat{\psi}(\omega)|^2}{\omega}\mathrm{d}\omega<+\infty \tag{5}$$

$X(t)$ 的小波变换为

$$WX(a,b)=\int_{-\infty}^{+\infty}X(t)\frac{1}{\sqrt{|a|}}\psi^*\left(\frac{t-b}{a}\right),a,b\in R,a\neq 0 \tag{6}$$

这里 a 是尺度因子，b 是平移因子。ψ^* 是共轭的。参数 a、b 在一个连续的实数集上变化，它们决定了小波变换的质量。在实际应用中，常将 a、b 离散化，可分别取为序列 $(2^j)_{j\in Z}$ 和 $(k2^j)_{k,j\in Z}$。则离散后的小波变换为

$$WX(k,j)=2^{j/2}\int_{-\infty}^{+\infty}X(t)\psi^*(2^jt-k)\mathrm{d}t \tag{7}$$

式（7）可通过快速塔式算法来实现。

3.2　小波阈值收缩消噪方法

小波变换阈值收缩消噪依赖于小波域中浮点数编码和噪音的统计性质。其实现包含小波变换域的收缩。可由三步实现：线性小波变换、非线性收缩消噪、线性逆小波变换。

设第 i 代被噪音 n_i 污染后的编码为 X_i，无噪编码为 f_i，则

$$X_i=f_i+n_i,i=1,2,\cdots,N \tag{8}$$

式中 n_i 相互独立，服从高斯正态分布 $N(0,\sigma)$。假设在 GA 运行过程中用一个操作 Th 取代变异操作，而 Th 能够使 X_i 具有最小均方差（MSE）

$$MSE=\frac{1}{N}\|\hat{\boldsymbol{f}}-\boldsymbol{f}\|^2=\frac{1}{N}\sum_{i=1}^{N}(\hat{f}_i-f_i)^2 \tag{9}$$

这里，$\boldsymbol{f}=(f_1,f_2,\cdots,f_N)^{\mathrm{T}}$，$\hat{\boldsymbol{f}}=Th(X)$，则可实现 X 的消噪。式（9）需满足一个条件：$\hat{\boldsymbol{f}}$ 至少与 $\boldsymbol{f}$ 有同样的平滑度。

在小波域中，几个大幅小波系数控制着编码能量，而噪音能量正态分布在所有小幅小波系数中，特别在细节部分。式（8）可写成

$$X=\boldsymbol{f}+n \tag{10}$$

这里，X 是被噪音污染的编码，$\boldsymbol{f}$ 是无噪编码，n 是遗传环境中产生的噪音。设 $w(\cdot)$ 和 $w^{-1}(\cdot)$ 分别为小波变换和逆小波变换，$D(\cdot,\lambda)$ 为用硬阈值或软阈值的消噪操作。则小波阈值收缩消噪可依据下式实现

$$y=w(X) \tag{11}$$

$$z=D(y,\lambda) \tag{12}$$

$$\hat{\boldsymbol{f}}=w^{-1}(z) \tag{13}$$

在小波阈值收缩消噪中，阈值的选择对消噪起着十分重要的作用，直接影响着消噪效果和算法性能。

3.3　小波阈值选择

可将含噪编码 X 分解成一个正交小波系列 $X=\sum_{\lambda\in\Lambda^J}\bar{X}_\lambda\psi_\lambda$。这里复指数 $\lambda=(j,i)$ 表示规模 j 和小波的位置 i。对应的指数集为

$$\Lambda^J=\{\lambda=(j,i),j=0,1,\cdots,J-1,i=0,1,\cdots,2^j-1\}$$

通过对小波系数 $\bar{X}_\lambda$ 的阈值化，变异相应编码。为此，定义非线性算子：

$$S_T:X\mapsto S_T(X)=\sum_{\lambda}\rho_T(\bar{X}_\lambda)\psi_\lambda \tag{14}$$

其阈值函数或为硬阈值

$$\rho_T(\alpha)=\begin{cases}\alpha & \text{if } |\alpha|>T\\ 0 & \text{if } |\alpha|\leqslant T\end{cases}$$

或为软阈值

$$\rho_T(\alpha)=\begin{cases}sign(\alpha)(|\alpha|-T) & \text{if } |\alpha|>T\\ 0 & \text{if } |\alpha|\leqslant T\end{cases}$$

这里，T 为阈值。用 Λ_T 表示通过阈值函数 ρ_T 选择的小波系数 $\bar{X}$ 的指数子集，即 $\Lambda_T=\{\lambda\in\Lambda^J,|\bar{X}_\lambda|>T\}\subset\Lambda^J$。那么，编码 $\boldsymbol{f}$ 和它的估计值 $\hat{\boldsymbol{f}}$ 的相对二次误差可定义为

$$\varepsilon(T)=\frac{\|\boldsymbol{f}-\hat{\boldsymbol{f}}\|^2}{\|\boldsymbol{f}\|^2} \tag{15}$$

其下界 $\min_T\ \varepsilon(T)$ 在所有编码 $\boldsymbol{f}\in\mathcal{H}$（$\mathcal{H}$ 是广义的函数空间，包括 Hölder 和 Besov 空间）的极大极小误差附近。也表明阈值

$$T_D=\sigma_{n'}(2\ln N')^{1/2} \tag{16}$$

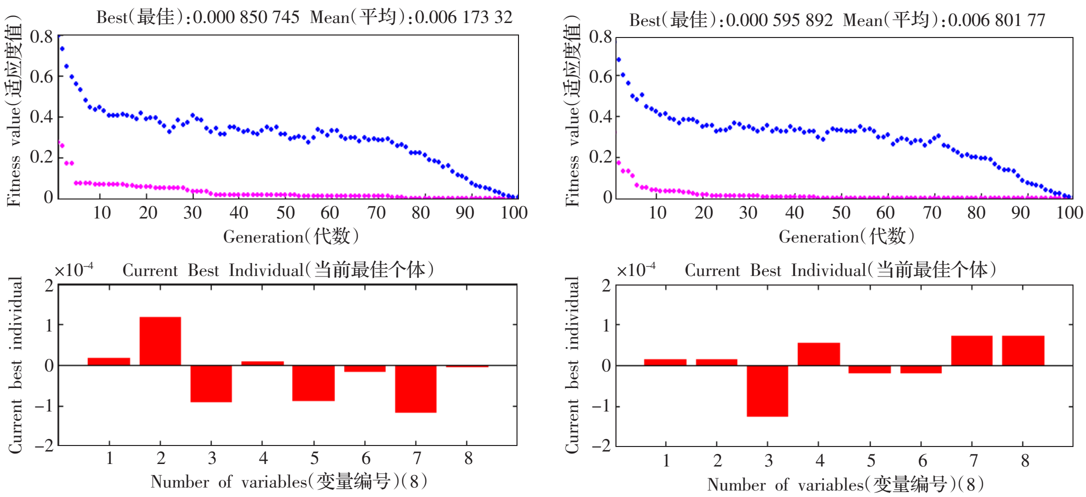
的误差 $\varepsilon(T_D)$ 在 $\varepsilon(T)$ 的最小值附近。由于 T_D 只取决于噪音的方差，与使误差 $\varepsilon(T)$ 最小的 $T_{\min}$ 值比较，T_D 也被称为通用阈值。然而，在许多应用中，$\sigma_{n'}$ 是未知的，必须从有效含噪编码 X 中予以估计。

为了估计噪音，仅研究残余部分，而不考虑 $S_T(X)$。即

$$S_T^c(X)=(Id-S_T)(X)=X-S_T(X)=\sum_{\lambda\in\Lambda^J}\rho_T^c(\bar{X}_\lambda)\psi_\lambda=\sum_{\lambda\in\Lambda_T^c}\bar{X}_\lambda\psi_\lambda \tag{17}$$

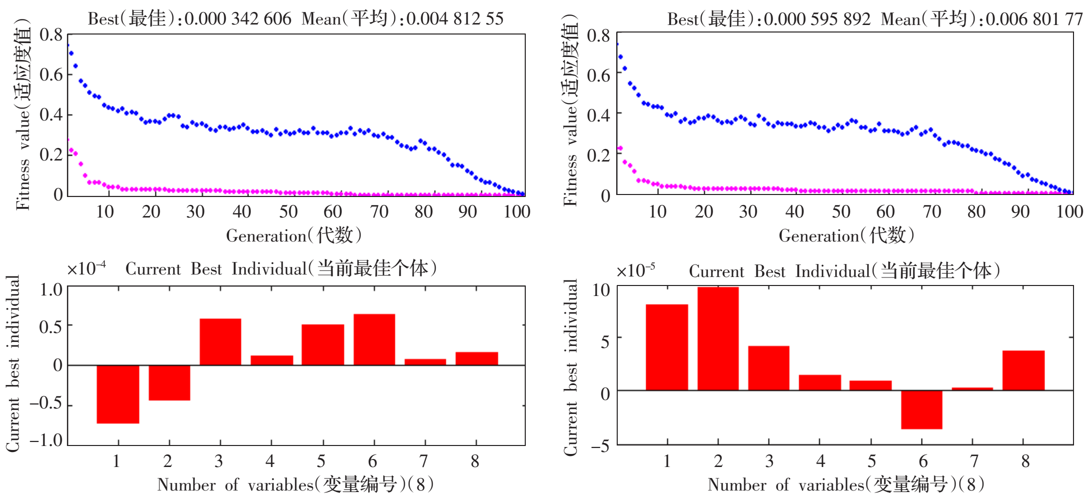
这里，Id 表示恒等元。补码算子 S_T^c 使用补阈值函数 $\rho_T^c=Id-\rho_T$，并定义补指数集为 $\Lambda_T^c=\Lambda^J\backslash\Lambda_T$。残余 $S_{T_D}^c(X)$ 是高斯白噪音 n 一个近似最优估计，其相对误差为

$$\varepsilon'(T)=\frac{\|X-S_T(X)-n\|^2}{\|n\|^2}=\frac{\|f+n-S_T(X)-n\|^2}{\|n\|^2}=\frac{\|f\|^2}{\|n\|^2}\varepsilon(T) \tag{18}$$



（a）FPRGA （b）FGAWSD（T=0.1）

图 1 两种方法的比较



（a）FGAWSD（T=0.5） （b）FGAWSD（T=1.0）

图 2 不同阈值的比较

4 应用实例

为了验证上述方法的可靠性和可行性，取著名的 F8 函数（Rastrigin's function）

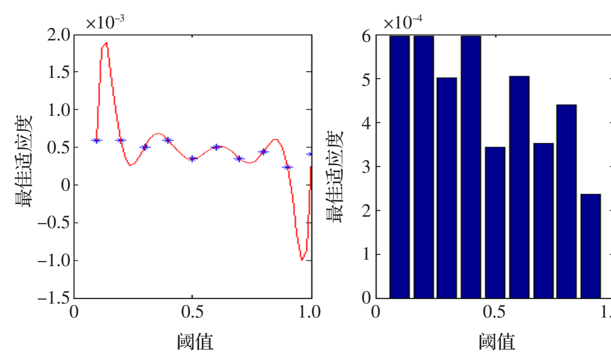
$$\min\ f(x)=10\times n+\sum_{i=1}^{n}[x_i^2-10\times(2\pi x_i)] \tag{19}$$

s.t. $x_i\in[-5.12,5.12]$

用 MATLAB 编程。群体规模 M=80，进化代数 G=100，在相同样本和参数下，分别对浮点数编码遗传算法（FPRGA）和该文的方法（FGAWSD）进行了比较实验。同时，也取不同的阈值对该方法进行了实验。程序运行结果如图 1~图 3 所示，实验数据如表 1 所示。实验结果表明：

（1）在最佳适应度、平均适应度和最优解等性能指标上，该方法均优于传统的浮点数编码遗传算法。说明该方法的性能是稳定可靠的，技术是可行的。

（2）在[0.1，1.0]的范围内取不同的阈值，在[0.1，0.5]区间性能比较稳定，在[0.6，1.0]区间波动稍大，而且当阈值 T=0.5 时性能较好。所以，阈值取 0.5 较为合适。



（a）最佳适应度拟合曲线 （b）不同阈值的最佳适应度

图 3 不同阈值的最佳适应度拟合曲线和柱状图

5 结论

浮点数编码具有精度高、便于高维大空间搜索的优点，在函数优化和约束优化领域明显有效于其他编码。但 FPRGA 在运行环境中产生的噪音对算法性能的影响并未引起人们的重视。传统的 FPRGA 变异采用的是有界随机的方法，这种变异方法不能消除算法运行环境中产生的噪音，因而也不能减少噪音对算法性能的影响。根据小波理论，建立小波滤波器，对浮点数编码进行阈值收缩消噪，取代变异操作。研究和实验结果表明，选择适当的阈值，FGAWSD 不仅在搜索全局最优解和提高搜

（下转 84 页）