

何委徽,王家林,胡龙胜. 实数编码多种群遗传算法的改进及应用. 地球物理学报, 2009, **52**(10):2644~2651, DOI:10. 3969/j. issn. 0001-5733. 2009. 10. 025

He W H, Wang J L, Hu L S. The improvement and application of real-coded multiple-population genetic algorithm. *Chinese J. Geophys.* (in Chinese), 2009, **52**(10):2644~2651, DOI:10. 3969/j. issn. 0001-5733. 2009. 10. 025

实数编码多种群遗传算法的改进及应用

何委徽,王家林,胡龙胜

同济大学海洋地质国家重点实验室,上海 200092

摘 要 本文首先从分析实数编码的 GA 理论出发,提出实数编码条件下不需要继续沿用二进制编码条件下常用的传统单点、两点及多点交叉算子,基因选择可采用基因位独立(locus independent)的方式进行. 据此研制了相应的交叉算子,并给出了一种实现方式,试验计算验证了该算子的有效性. 其次,针对地球物理反演问题的复杂性以及遗传算法参数设置对问题的敏感性,将算法性能调整的灵活性作为方法选择的重要标准,采用了多种群遗传算法作为反演方法,设计了两个种群级的改进策略,并验证了其效能. 最后,将改进的实数编码多种群遗传算法应用于重力模型反演,完成了梯形变密度体组合模型和 2.5 度体多边形棱柱组合模型的反演试算,取得了预期效果,并成功地应用于新疆地区实际资料的处理与解释.

关键词 实数编码,多种群遗传算法,重力,反演

DOI:10. 3969/j. issn. 0001-5733. 2009. 10. 025

中图分类号 P631

收稿日期 2009-05-20,2009-09-07 收修定稿

The improvement and application of real-coded multiple-population genetic algorithm

HE Wei-Hui, WANG Jia-Lin, HU Long-Sheng

State Key Laboratory of Marine Geology, Tongji University, Shanghai 200092, China

Abstract This paper first examines the theory of Real-Coded Genetic Algorithm(RCGA) and put forward that locus independent selection strategy should be adopted when performing RCGA crossover operation. According to this strategy a new crossover operator is designed and tested. Secondly, the paper pays great emphasis on the scalability of the GA algorithm when selecting GA strategies, thus the Multiple Population Genetic Algorithm(MPGA) is selected as the framework of the algorithm and two population-level strategies are designed and tested. Combining the new crossover operator and these two strategies together, the paper at last comes to the improved MPGA and it is then applied to the multi-layer gravity inversion of both experimental models and field data of Xinjiang area, China.

Keywords Real coded, Multiple population Genetic Algorithm, Gravity, Inversion

基金项目 国家重点基础研究发展规划(973 计划)项目“南海大陆边缘动力学及油气资源潜力”中的子课题“南海东北部中生代的断裂特征和盆地分析”(2007CB411706-02);国家高技术研究发展计划(863 计划)课题“海区残留盆地油气资源综合地球物理勘探技术”的子课题“断裂与火成岩分布的重磁提取技术”(2006AA09Z359)资助.

作者简介 何委徽,男,1979 年生,现为同济大学海洋与地球科学学院在读博士生. 主要研究方向为综合地球物理解释、地球物理联合反演、非线性反演. E-mail:moofen2006@163.com

1 引言

遗传算法(Genetic Algorithm,简称GA)是以选择、交叉、变异三个简单遗传操作为特征的模拟生物进化过程的计算方法,整个计算过程只依赖个体(即编码后的模型参数全体)的适应度值(在反演中是目标函数值),GA具有实现简单、适用性强的特点,获得了极为广泛的应用.选择算子引导搜索向可能的最优区域前进,“开采”搜索空间中具有适应度值较高的个体,会显著降低多样性,一般将倾向于“开采”的策略称为具有较大的选择压,反之称为具有较小的选择压,选择压太大会使种群多样性迅速消失而陷入早熟,选择压太低又会导致收敛缓慢.而交叉和变异算子的作用则是“勘探”搜索空间以寻找那些可能的最优区域,它们以保持种群的多样性为主要目的.“开采”与“勘探”的平衡实质上是种群多样性和演化速度间的平衡.

重力异常反演是一个非线性多参数最优化问题,由于重力反演多解性的影响,使同时反演模型几何参数(埋深规模尺寸)和物性参数(密度)的问题变得困难,尤其是反演多层密度界面时更是如此.用重力异常反演多层密度界面大致有两条途径^[1]:一是通过场分离后逐一对单层界面反演的方法,但场分离问题本身的困难对该方法的使用产生了一定的限制;另一个办法是直接求解:如可用于定性和半定量解释的重力归一化总梯度法、以直角侧边梯形组合体为模型的正则化非线性反演法等.遗传算法作为一种全局寻优的随机搜索方法十多年来也逐渐地被应用于重力界面的直接反演.刘云峰等^[2]把二维连续密度界面离散化,应用遗传算法进行反演,取得了较好结果.石琳柯等^[3]根据模型分解的思想,以遗传算法为基础,采用变密多层地壳密度模型,结合多种地球物理资料,计算出了郑州—临汾剖面的地壳密度结构.陈军等^[4]提出了采用多段自适应交换策略的改进遗传算法,并应用于多层密度界面的反演,提高了反演速度.陈超等^[5]通过混合编码来改善GA的计算性能也取得了很好的效果.

由于遗传算法的性能对遗传算子、参数设置以及待求问题本身都很敏感,所以一般都要对Holland最初提出的标准遗传算法进行改进后再应用于特定的问题.这些改进可以分为三个方面:一是编码;二是微观遗传策略,主要是指选择算子、交叉算子和变异算子的设计以及参数选择;三是宏观遗

传策略,如稳态遗传算法(Steady-state GA),多种群算法等.前人工作中应用的主要是二进制编码的遗传算法(Binary-Coded Genetic Algorithm,以下简称BCGA),改进则主要集中在编码方式和算子设计上.

本文的思路有二:一是采用实数编码的遗传算法(Real-Coded Genetic Algorithm,以下简称RCGA),这是因为RCGA相比BCGA具有编码更接近问题空间、编码长度更小、计算速度更快的明显优点^[6];二是将GA算法调整的灵活度作为算法设计和改进的重要标准,从而选择了多种群遗传算法作为框架设计改进策略.

2 理论研究

遗传算法的基础理论研究以收敛性分析为主,目前可以分为随机理论模型和进化动力学理论.前者以马尔可夫链为工具描述和研究遗传算法,其高度的抽象性使它在具有广泛适用性的同时也失去了对GA具体应用的指导意义.Holland提出的模式定理和隐含并行性原理以及积木块假设(Building block hypothesis)构成了遗传算法进化动力学的基本定理,对遗传算法的理论和实践均具有重要指导意义.模式定理说明了超过种群平均适应度值的模式将在遗传算子的作用下存活并重新联结从而形成最终解,这个过程也可以被形象地看成是一个“搭积木”的过程.但模式定理和积木块假设能直接作为RCGA算法设计的指导吗?又如何分析和理解RCGA演化的过程呢?国内的文献一直没有提及相关的内容,而这对于实数编码的遗传算法应用又是至关重要的,因此本文首先介绍有关模式定理在RCGA过程分析中推广和发展的相关理论.

二进制编码表可以直接扩展到任意的整数编码

表:
$$c = \prod_{i=1}^n (a_i, \beta_i),$$
其中 $a_i, \beta_i \in \{0, \dots, L_i\}$ 且 $a_i \leq a_i \leq \beta_i \leq b_i, [a_i, b_i]$ 是第 i 个参数的变化区间.由于实际问题总是有精度要求的,所以当编码表足够大时,整数编码表也就变成了实际使用的实数编码表.因此Eshelman^[7]认为RCGA其实应该叫做“整数编码的遗传算法(Integer Coded GA, ICGA)”,而ICGA中的宽区间就对应着BCGA中的低阶模式,随着演化的进行算法将逐步地自动聚焦于短的区域,这就对应着BCGA中的高阶模式.由于这一过程是自动进行的,所以计算中没有必要人为地将实

数映射为整数. 因此本文采用了直接的实数编码, 编码空间就是问题的模型空间.

作为模式定理在 RCGA 下的替代, Goldberg^[8] 的拟编码表理论力图找出演化中的有效积木块, 并称之为拟字符(virtual character), 拟字符的集合就构成了拟编码表(virtual alphabet). 首先他观察到遗传算法在演化初期选择算子主导了演化, 使得超过平均适应值水平的个体在很短的时间内就会充满种群. 这个时间称为接管时间: $t = \log_2 \log_2 n_p$, t 是演化代数, n_p 是种群规模. 依据此式, 即使是规模达到 10^9 种群接管时间大约也只有 5 代, 一般规模种群的接管时间在 3~4 代. 因此对初期最优个体生成有贡献的显然主要是这部分个体, 令 $\lambda = \bar{\Phi}$, 可以定义遗传算法的选择集为:

$$L(\Phi, \lambda) = \{x : x \in \Omega \text{ 和 } \Phi(x) \geq \lambda\},$$

进一步可在选择集中依照个体间的关联性划分出多个子集, 将每个极大的关联子集称为选择集的一个特征. 定义 $f(x)$, $x \in \Omega$ 为向量的概率密度函数, 则由索引集合 I 定义的 n 个实数变量的边缘密度函数可写为:

$$f_I(x_I) = \int_{\Omega_r} f(x) d\Omega_r, I' = \{1, \dots, n\} - I,$$

据此就可以定义 I 中自由变量 $x_i, i \in I$ 的“平均切片”:

$$\bar{\Phi}_I(x_I) = \int_{\Omega_r} f_r(x_r) \Phi(x) d\Omega_r$$

式中 $d\Omega_r = \prod_{j \in I'} dx_j$. 当 I 中只有一个变量时, 就得到了单变量的平均切片 $\bar{\Phi}_i$. 这样就可以定义拟字符为 $\bar{\Phi}_i$ 上一个最优的吸引盆与一个特征的非空交集, 而 $\bar{\Phi}_i$ 上所有拟字符的集合就构成了该维的拟编码表.

3 基于基因位独立的交叉算子

3.1 算子的提出

传统上 RCGA 交叉操作的基因选择方式直接沿用了 BCGA 的选择方式, 对实数编码的连续优化问题并不很适合. 本文就此提出了基于基因位独立(locus independent)的交叉算子. 如图 1 所示, 图 1 (a、b)对比了传统的单点和两点交叉的基因位选择方式的不同. 为弥补基因位独立的选择方式在交叉基因数量上的损失, 在交叉中需要提供一个规模因子 α , 设基因长度为 l , 则每个个体实际交叉的基因数量是 $\text{round}(\alpha \cdot l)$, $\alpha \in [0, 1]$. α 在演化过程中可以是固定的预设值, 也可以是一个与演化代数有关

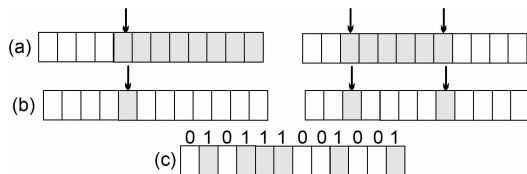


图 1 基于基因位独立的交叉算子

(a) 传统单点和两点交叉算子选择方式; (b) 基因位独立的单点和两点交叉选择方式; (c) 基因位独立选择方式的一种实现.

Fig. 1 The locus independent crossover operator

(a) The traditional one point and two point crossover strategy; (b) The locus independent one point and two point crossover strategy; (c) A concrete implementation of locus independent crossover strategy.

的因子, 从而建立起交叉强度可控的自适应交叉算子, 图 1c 给出了基因位选择的一种实现方式: 每次交叉时, 产生 1 个长度为 l 的随机二进制序列, 其中产生 1 的概率为 α , 产生 0 的概率为 $1-\alpha$, 然后对值为 1 的基因位进行交叉操作. 因为这里提出的只是基因位选择的方式, 而现有的对交叉算子的改进方式都是集中在基因位选择以后的处理上, 所以该方式可以与任何已有的算子改进相结合, 比如本文采用的交叉算子就是基于基因位独立的算术交叉算子(下称 needle 算子).

基因位独立的交叉选择方式与传统方式的根本不同在于基因位独立方式不考虑被选中基因在位置上的连续性. 那么这样做是否合理呢? 本文提供如下分析:

首先在 BCGA 的模式理论框架中考查这个问题. Holland^[9] 在其早期的研究中就曾指出, 交叉算子的重组作用是双重的: 一方面重新联结已有的低阶有效模式而成为高阶模式, 称为模式保存能力; 另一方面交叉算子还具有对有效基因位进行重组而产生新的低阶有效模式的能力, 称为模式重组能力. 前者要求交叉算子的基因位选择尽可能相互关联, 而后者则无此要求. 从计算效果来看, 即便是在 BCGA 条件下对高阶模式破坏性更大的一致交叉算子的效果也要好于对模式保护能力最强的传统单点交叉算子^[10], 而且交叉算子不可能同时具备强大模式保存能力和重组能力^[11].

再从整数编码的观点来考查, RCGA 确实存在着与 BCGA 完全类似的模式概念, 但随着编码精度的提高, 在一般种群规模下(30~1000)产生有效的高维模式成为一个极小概率事件. Goldberg 的拟编码表理论将 RCGA 实际使用的有效模式数量降低

了很多,即使如此在 RCGA 下的传统意义的高维特征也还是一个极小概率事件:不失一般性,考虑一个 $[0,1]^n \rightarrow R$ 问题,设特征在该维上所占的长度为 δ ,则 δ 也可以看成是特征在编码区间上的长度比率,则一个 2 变量的模式就是一个 δ^2 的面积, k 个变量的模式就在 k 维超立方体上占有 δ^k 的体积,如果种群规模刚好能够保证在种群产生一个拟字符 $O(1)$,或者说该拟字符在种群产生的概率为 $\delta \approx 1/n_p$,则种群中具有所有 2 维特征的概率是 $O(1/n_p)$,更高阶特征的概率更低。

综上所述,从二进制编码的观点来看,RCGA 的交叉算子的高维模式生存能力是不重要的,模式重组能力才是最重要的。如果再考虑实数编码意义下的模式理论,则可以认为考虑高维模式是完全没有必要的,当采用基因位连续的交叉算子时反而可能缩小交叉算子的搜索空间。所以,基因位独立的选

择方式更适合 RCGA 交叉算子的选择方式。

3.2 基因位独立的交叉算子的测试

本文使用测试函数 F_1 和 F_2 对算子或算法的性能进行测试。 F_1 表达式为: $f(x) = \sum_{i=1}^n x_i^2$, $-5.12 \leq x_i \leq 5.12, i = 1, \dots, n$, 当 $x^* = (0, \dots, 0)$, 取得惟一全局极小值点 $f(x^*) = 0$, 图 2a 是其 $n=2$ 时的图像。

F_2 表达式为: $f(x) = 10n + \sum_{i=1}^n (x_i^2 - 10\cos(2\pi x_i))$, 式中 $-5.12 \leq x_i \leq 5.12, i = 1, \dots, n$, 有多个局部极值,当 $x^* = (0, \dots, 0)$ 时,取得全局最小值 $f(x^*) = 0$, 图 2b 是其 $n=2$ 时的图像。本文的测试计算中 F_1 和 F_2 的自变量个数均为 $n=10$ 。适应度的尺度变换函数为 $\text{Fit}(f(x)) = 1/(1+f(x))$, 将原始函数值映射为 $[0,1]$ 上的适应度值。

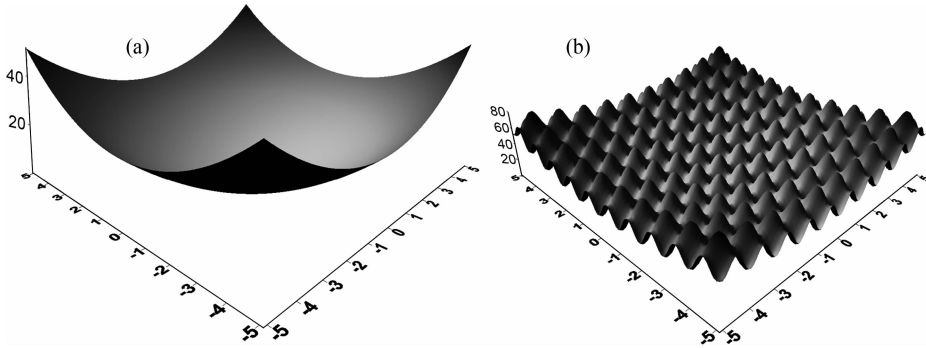


图 2 测试函数

(a) 测试函数 F_1 ; (b) 测试函数 F_2 ; 纵横坐标分别表示测试函数第一维和第二维变量值。

Fig. 2 The trial functions

(a) Trial function F_1 ; (b) Trial function F_2 .

对交叉算子的测试计算采用实数编码的单种群遗传算法,主要参数设置为:种群规模 80,选择算子采用轮转赌盘策略,交叉率为 0.7,变异算子为动态变异算子,变异率为 0.06。当 $|\text{Fit}(f(x)) - 1| < 1 \times 10^{-5}$ 时,认为计算完全收敛,表 1 给出了算法收敛时的演化代数,未收敛的给出了停止代数时的适应度值,计算结果均为 20 次计算的平均值。

对表 1 的试验结果分析:(1)标准的单点和两点算子性能远远落后于采用线性交叉的其他算子;(2)两点的线性交叉算子的性能与采用基因位独立选择方式的 needle 算子的性能基本相当,后者略好于前者,说明 RCGA 下的高阶模式的破坏对算法性能没有影响,而两者的性能相当则是因为两点线性交叉算子本身也是高维模式保存能力较差的算子;(3)从试验编号 4~7 可以看到,对单峰函数 F_1 的勘探力度越大其收敛速度越快,而对复杂的多峰函数 F_2 ,

遗传算子的勘探力度存在最优值,这里当 needle 算子的交叉规模达到 0.2 时效果最好。(4)除性能稍强外,needle 算子相比两点线性交叉算子,一个突出的优点是“勘探力度”可以灵活调整,结合交叉率可以提供对交叉算子性能的灵活控制。

表 1 基因位独立的交叉算子与传统算子的性能比较

Table 1 Comparison of locus independent crossover and traditional crossovers

编号	算子及参数	收敛结果	
		F_1	F_2
1	单点	4970	5000(0.07176)
2	两点	4930	5000(0.05023)
3	两点线性交叉	105	325
4	needle($a=0.2$)	150	350
5	needle($a=0.4$)	115	430
6	needle($a=0.6$)	100	1225
7	needle($a=0.8$)	75	805

4 改进的实数编码多种群遗传算法

提高 GA 算法性能的一个重要方向就是并行遗传算法(parallel GA, PGA),可划分为三种基本类型:全局单种群主从模式、全局单种群细粒度模型 PGA、和多种群或粗粒度模型 PGA(multiple population GA,MPGA).

多样性是 GA 能够搜索到全局最优解的基本条件.有效控制种群的多样性是提高 GA 性能的一个重要途径.单种群演化较难控制群体多样性与选择压力之间的有效平衡,这时算法或出现早熟或频繁在非有效区域进行搜索而导致搜索效率过低.MPGA 通过对不同的子种群配置不同的初始化值和参数而有效地控制子种群的独立演化特性,同时还通过种群间的个体迁移来保证子种群间的融合,从而实现算法在整体上的性能提升.

据此,本文设计了如图 3 所示的采用环形拓扑的 MPGA.算法共有 7 个子种群,0~5 种群构成一个均衡环,种群 6 是一个精英池,单向地从其他 6 个种群接受最优个体,它不影响环结构的搜索性能.算法主要设计了两个种群级的宏观改进策略.策略一:在种群环上同时选择了选择压很大的轮转赌盘(Roulette)选择算子和选择压相对较小的线性排名(Rank-based)选择算子,如图 3 所示,种群 0,2,4 采用 Rank 选择算子,1,3,5 采用 Roulette 选择算子.策略二:在由当前最优解构成精英池中使用“开采能力强勘探力度小”的算子组合进一步“开采”,以期获得目前条件下的更优解.

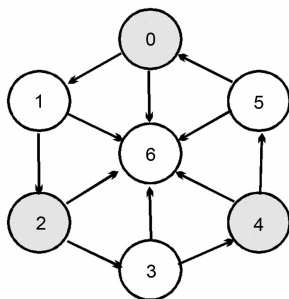


图 3 本文设计的多种群遗传算法拓扑结构

Fig. 3 The designed structure of MPGA

下面仍使用测试函数 F_1 和 F_2 对多种群遗传算法进行测试.MPGA1 是没有精英池(种群 6)的环结构,子种群规模为 80,交叉算子采用规模因子为 $\alpha=0.6$ 的 needle 算子,交叉率为 0.6,变异算子为动态变异算子,变异率为 0.06,迁移率 0.05,迁移间隔

为 1.MPGA2 则是本文设计的带精英池的环结构,其精英池(种群 6)的选择算子为线性排名的选择算子,交叉算子为 $\alpha=0.4$ 的 needle 算子,变异算子为动态变异算子,变异率为 0.06,其他种群(0~5)的参数与 MPGA1 完全相同.简单地说,MPGA1 采用了前述策略一,MPGA2 同时采用了策略一和策略二,主要试验目的是测试两种选择算子的不同组合对于收敛速度的影响,结果均为 20 次计算的平均结果.

对表 2 的结果分析可知:(1)当在部分子种群中使用线性排名选择算子时,MPGA1 和 MPGA2 的性能均得到了明显改善,证明策略一是有效的.(2)同条件下 MPGA2 的收敛速度要好于 MPGA1 的,证明了策略二的有效性;(3) F_1 的收敛速度随采用 Rank 种群的数量上升而增大, F_2 则存在最优配置结构,当 Rank 种群数量为 3 个并采用图 3 所示的结构时 MPGA1 和 MPGA2 均出现最快收敛速度.

采用如图 3 所示拓扑结构的 MPGA2 即为本文最终确定的多种群算法设计,后文简记为 MPGA0,除特别给出的参数外,其余参数均与本试验所采用的参数相同.

表 2 种群策略的性能测试

Table 2 Trials of population-level strategies

Rank 数目	收敛代数			
	MPGA1		MPGA2	
	F_1	F_2	F_1	F_2
0	96	222	90	216
1	87	177	72	141
2	69	186	51	135
3	66	129	72	108
4	69	144	63	132
5	66	144	63	132
6	54	192	60	135

5 重力资料的反演研究

在充分利用其他地质地球物理资料的情况下,可以将重力反演问题描述为一个对目标函数的带约束的优化问题.考虑到遗传算法的适应度值函数应为正值,且为防止适应度函数造成的超级个体的出现,对计算结果进行了归一化处理,可将适应度函数写成下面的形式:

$$\phi = \frac{1}{M} \sum_{i=1}^M \left(1 - \frac{g_i^{\text{cal}}(c_i)}{g_i^{\text{obs}}} \right)^2, c_i \in [a_i, b_i]$$

式中 M 为染色体长度(染色体由模型参数构成), g_i^{cal} 为计算重力值, g_i^{obs} 为实际观测值或理论值, c_i 为基因值. 令模型地质单元数目为 n , 各地质单元包含的参数数目为 m_1, m_2, \dots, m_n , 则染色体长度为:

$M = \sum_{i=1}^n m_i$. 参数约束范围在模型计算中取 $|b_i - a_i| = 0.2c_i^{\text{theory}}$, 在实际计算中根据已知的地质地球物理资料确定. 反演均采用指定最大演化代数的方式作为结束条件.

5.1 梯形变密度体组合模型试验

正演采用如图 4 所示的梯形变密度组合模型^[1], 试验中变密度梯形的个数 $n=3$. 每个梯形体参数为 $m_1=m_2=m_3=8$, 分别为梯形 4 个顶点的横坐标、顶底的深度值和顶底的密度值. 该模型的密度在梯形顶底深度间纵向线性变化, 通过块体组合可模拟地质体的纵横向变化.

作为对比, 计算分别采用了单种群遗传算法和 MPGA0, 其中单种群遗传算法的参数为: 个体长度为 24, 种群规模为 480, Roulette 选择算子, $\alpha=0.6$ 的 needle 交叉算子, 交叉率为 0.7, 动态变异算子, 变异率 0.06.

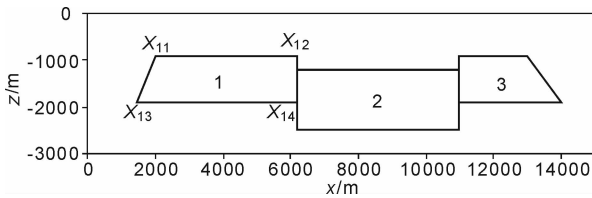


图 4 梯形变密度体组合模型

Fig. 4 Combined trapezoidal model with density variation

表 3 的计算结果为 10 次计算中选取的最好结果. 可以看到:MPGA0 的平均相对误差为 0.03526, 最大相对误差 0.07309, 最小相对误差为 0.00207. 如果采用传统的单种群遗传算法的平均相对误差 0.098477, 最大相对误差为 0.448864, 最小相对误差 0.00229. 前者的结果明显好于后者.

5.2 2.5 度多边形棱柱体的界面模型试验

试验计算中正演采用 2.5 度体的任意多边形的棱柱体模型^[1]. 界面(地质单元)数目为 $n=3$, 界面参数前 m_i-1 个为该界面的控制点深度值(单位: km), 最后一个参数 m_i 为该界面的密度值(单位: 0.2 g/cm^3), 数目分别为 $m_1=8, m_2=9, m_3=9$, 染色体长度 $M=26$ (即为 23 个深度点和 3 个密度值).

地面观测点 70 个, 点距 0.5 km, 每层密度均

表 3 梯形变密度体组合模型的反演结果

Table 3 Inversion of the combined trapezoidal model with density variation

参数名称	理论值	单种群		MPGA0	
		反演结果	相对误差	反演结果	相对误差
X ₁₁	2000	1853.49	0.073255	2029.81	0.01491
X ₁₂	6200	6398.18	0.031965	5908.86	0.04696
X ₁₃	1440	1575.86	0.094347	1485.86	0.03185
X ₁₄	6200	6309.99	0.01774	6247.42	0.00765
H ₁₅	900	968.667	0.076297	953.092	0.05899
H ₁₆	1900	2056.05	0.082132	2038.87	0.07309
S ₁₇	2.74	2.86828	0.046818	2.76529	0.00923
S ₁₈	2.81	2.61877	0.068053	2.79605	0.00496
X ₂₁	6200	6944.57	0.120092	6483.55	0.04573
X ₂₂	11000	15937.5	0.448864	10392.5	0.05523
X ₂₃	6200	4787.54	0.227816	5922.79	0.04471
X ₂₄	11000	10862.4	0.012509	10566.3	0.03943
H ₂₅	1200	1355.53	0.129608	1246.31	0.03859
H ₂₆	2500	2316.61	0.073356	2409.28	0.03629
S ₂₇	2.79	2.95483	0.059079	2.89358	0.03713
S ₂₈	2.86	2.78477	0.026304	2.72373	0.04765
X ₃₁	11000	11428.4	0.038945	11241	0.02191
X ₃₂	13000	11890.2	0.085369	13490.6	0.03774
X ₃₃	11000	12513.9	0.137627	11051.9	0.00472
X ₃₄	14000	13876.1	0.00885	14029	0.00207
H ₃₅	900	762.139	0.153179	948.993	0.05444
H ₃₆	1900	2269.83	0.194647	1988.7	0.04668
S ₃₇	2.74	2.3172	0.154307	2.68091	0.02157
S ₃₈	2.86	2.86655	0.00229	2.67495	0.0647
平均相对误差:			0.098477		0.03526

注:表中水平距离 X 和深度值 H 单位为 m, 密度 S 单位为 g/cm^3 .

一, 层间密度差均为 0.2 g/cm^3 . 反演算法采用 MPGA0, 子种群规模为 300.

对三个界面分别正演后使用 MPGA0 进行反演. 反演结果的平均相对误差为 0.027, 最大相对误差 0.073, 最小相对误差 0.0014. 图 5 中实线为理论计算值, 虚线为反演结果, 除个别点以外, 反演界面与理论值基本吻合.

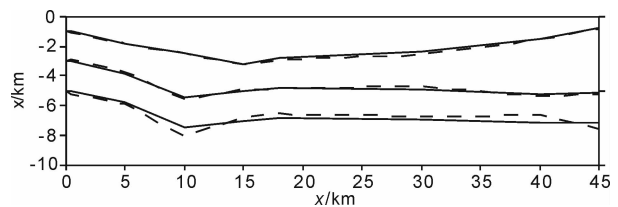


图 5 界面模型的反演

Fig. 5 Inversion of the interfaces model

5.3 实际资料反演

本文选取了新疆准噶尔盆地的某测线进行实际处理解释. 测线为西南东北走向, 全长 184 km, 位于盆地南部重力异常低值异常区. 对观测数据进行了预处理后, 本文选取其中 90 km (10~100 km) 进行反演. 根据已有地质地震资料, 确定地层分别为白垩系、上侏罗统、中下侏罗统、三叠系、二叠系和石炭系, 各层密度分别为: 2.4、2.48、2.54、2.60、2.61、2.68(g/cm³). 但是地震资料上基底的二叠系和石炭系的分界面并不清楚, 只能得到上面 4 个分界面, 如图 6a 所示. 本文尝试采用 MPGA0 反演包括二叠系和石炭系分界面在内的 5 个地层分界面. 首先用地震资料解释结果作为先验信息给出界面的搜索范围, 然后利用 MPGA0 进行反演. 参数为: 地面观测点 41 个, 点距 2 km. MPGA0 的交叉率为 0.85, 变异率为 0.01, 种群大小为 300.

重力异常曲线拟合如图 6b 所示, 其中实线为计算值, 虚线为观测值. 图 6c 中的实线为重力反演结果, 作为比较, 以虚线绘出了地震解释所揭示的界面. 可以看到上部的 4 个界面重力反演结果与地震解释结果基本吻合, 最大相对误差为 0.067, 最小相对误差为 0.0005, 平均相对误差 0.018, 个别地方的较大误差可能是与边界效应有关, 结果验证了方法的有效性. 同时, 反演还揭示了石炭系的顶界面的分布, 弥补了地震解释的不足.

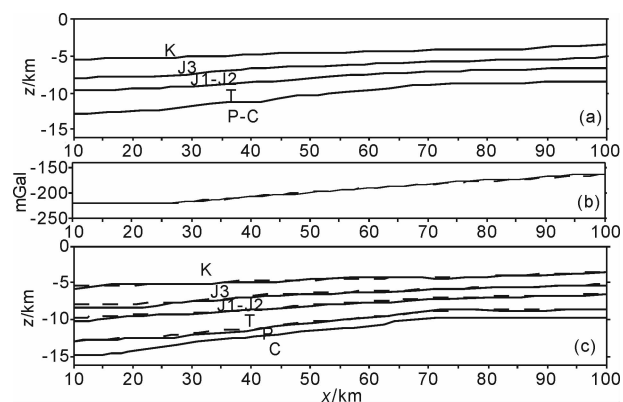


图 6 实际资料的反演

(a)地震数据分层;(b)数据拟合;(c)模型拟合.

Fig. 6 Inversion of the field data

(a)Layers of seismic data; (b)Data fitting; (c)Model fitting.

6 结论

本文在分析实数编码 GA 理论的基础上提出了基因位独立的交叉算子基因选择方式, 研制了相应

的交叉算子, 并给出了一种实现方式, 从而在微观策略层面对交叉算子实现了改进.

基于对地球物理反演问题的复杂性和 GA 对问题的敏感性的认识, 本文将算法性能调整的灵活性作为方法选择的重要标准, 从宏观策略层面设计了两个种群级的改进策略.

最后, 将改进的实数编码 MPGA 应用于重力模型反演, 完成了梯形变密度组合模型和 2.5 度体多边形棱柱组合模型的反演试算, 取得了预期效果, 并应用于新疆地区实际资料的处理与解释, 取得了对该测线石炭系地层顶界面的进一步认识.

结果表明改进后的 MPGA 具有较好的性能, 能够满足多层密度界面重力反演的需要.

参考文献(References)

- [1] 王家林, 王一新, 万明浩. 石油重磁解释. 北京: 石油工业出版社, 1991
Wang J L, Wang Y X, Wan M H. Petroleum Gravity and Magnetic Data Interpretation(in Chinese). Beijing: Petroleum Industry Press, 1991
- [2] 刘云峰, 沈晓华. 二维密度界面的遗传算法反演. 物探化探计算技术, 1997, 19(2):138~142
Liu Y F, Shen X H. Nonlinear inversion of gravity anomalies caused by 2D surface of geologic structures genetic algorithms. *Computing Techniques for Geophysical and Geochemical Exploration*(in Chinese), 1997, 19(2):138~142
- [3] 石琳珂, 孙铭心, 王广国等. 地球物理遗传反演方法. 北京: 地震出版社, 2000
Shi L K, Sun M X, Wang G G, et al. Genetic Inverse Methods of Geophysical Problems (in Chinese). Beijing: Seismic Press, 2000
- [4] 陈军, 王家林. 应用改进的遗传算法反演多层密度界面. 地球科学, 2001, 25(6):651~655
Chen J, Wang J L. Application of improved genetic algorithm to inversion of multi-layer density interface. *Earth Science*(in Chinese), 2001, 25(6):651~655
- [5] 陈超, 刘江平, 余丰. 求解位场反演问题的混合编码遗传算法. 地球物理学报, 2004, 47(1):119~126
Chen C, Liu J P, Yu F. The inversion of gravity data by using hybrid encoding genetic algorithm. *Chinese J. Geophys.* (in Chinese), 2004, 47(1):119~126
- [6] Z. 米凯利维茨. 演化程序——遗传算法和数据编码的结合. 北京: 科学出版社, 2000
Michalewicz Z. Genetic Algorithms + Data Structures = Evolution Programs. Beijing: Science Press, 2000
- [7] Eshelman L J, Schaffer J D. Real-Coded Genetic Algorithms and Interval-Schemata. In: *Foundation of Genetic Algorithms 2*. San Mateo: Morgan Kaufmann Publisher, 1993. 187~202
- [8] Goldberg D E. Real-coded Genetic Algorithms, *Virtual*

- Alphabets, and Blocking. *Complex Systems*, 1991, 5:153~171
- [9] Holland J H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, 2nd edition. Cambridge: MIT Press, 1992
- [10] Syswerda G. Uniform crossover in genetic algorithms. In: *Proceedings of the Third International Conference on Genetic Algorithm(ICGA 3)*, San Mateo: Morgan Kaufmann Publisher, 1989. 2~9
- [11] 李敏强, 寇纪淞, 林丹等. *遗传算法的基本理论与应用*. 北京: 科学出版社, 2002
- Li M Q, Kou J S, Lin D, et al. *Foundamental Theory and Application of Genetic Algorithms (in Chinese)*. Beijing: Science Press, 2002
- (本文编辑 胡素芳)

欢迎订阅《地球物理学报》

《地球物理学报》创刊于 1948 年,是中国科学院地质与地球物理研究所、中国地球物理学会联合主办的有关地球物理科学的综合性学术刊物. 主要刊载固体地球物理、应用地球物理、地磁和空间物理、大气和海洋地球物理,以及与地球物理密切相关的交叉学科研究成果的高质量论文. SCI 等国际著名检索系统收录,国内外公开发行. 2010 年《地球物理学报》为月刊,每期定价 70 元,全年定价 840 元(含邮资). 国内邮发代号 2-571. 也可直接从编辑部订购,保证快捷准时.

订阅地址:北京 9825 信箱《地球物理学报》编辑部 邮编:100029

电话:010-82998105 传真:010-82998107

电子邮件:prog@mail.igcas.ac.cn xtq@mail.igcas.ac.cn

联系人:刘少华 肖台琴

开户行:中国农业银行北京健德支行

银行帐号:190901040000456

收款单位:中国科学院地质与地球物理研究所

(务必在注释行写上:购《地球物理学报》款,同时写上您的姓名和联系地址及发票单位)