文章编号:1001 - 9081(2009)11 - 3056 - 04

基于改进非支配遗传算法的 DNA 编码序列优化方法

王延峰[1,2],申永鹏[1],张勋才[1,2],崔光照[1,2]
(1. 郑州轻工业学院 电气信息工程学院,郑州 450002; 2. 河南省信息化电器重点实验室,郑州 450002)
(shyplm@126.com)

摘 要:针对 DNA 计算中的编码序列设计问题,分析了 DNA 编码序列设计的目标和需要满足的约束条件,并建立了相应的数学模型。通过将约束条件引入非支配排序过程,提出了一种改进的 NSGA-Ⅱ算法。实验结果表明,该算法具有良好的收敛特性和种群多样性,能为可控的 DNA 计算提供可靠的编码序列。

关键词:DNA 计算;编码序列;遗传算法;NSGA-Ⅱ
中图分类号: TP18;TP301.5 **文献标志码**:A

DNA codewords design based on improved NSGA-Ⅱ

WANG Yan-feng[1,2], SHEN Yong-peng[1], ZHANG Xun-cai[1,2], CUI Guang-zhao[1,2]
(1. *College of Electrical Information Engineering, Zhengzhou University of Light Industry, Zhengzhou Henan* 450002, *China*;
2. *Henan Key Laboratory of Information-based Electrical Appliances, Zhengzhou Henan* 450002, *China*)

Abstract: Concerning DNA codewords design, the authors set up the mathematical model by analyzing the objectives and the restrictions that should be satisfied. A new codewords design method named the Improved Non-dominated Sorting Genetic Algorithm (INSGA-Ⅱ) was proposed by introducing the constraints to the non-dominated sorting process. The experiments demonstrate that INSGA-Ⅱ has higher convergence speed and better population diversity than those of NSGA, and can provide reliable and effective codewords for the controllable DNA computing.

Key words: DNA computing; coded sequence; Genetic Algorithm (GA); NSGA-Ⅱ

0 引言

DNA 编码序列设计问题自 DNA 计算诞生以来就一直是该领域研究的核心问题之一,编码序列设计的优劣决定了 DNA 计算的精度与可靠性。如何设计具有高质量的 DNA 编码序列已成为 DNA 计算领域的一个研究热点。

DNA 计算中的编码序列设计问题可定义为:系统地将一个算法问题的实例映射为特殊的 DNA 分子序列,这些 DNA 分子序列应能够确保随后所进行的生化反应中不出现任何错误,而且生化反应后产物中须包含有足够多的、稳定可靠的、能被成功提取的原始问题的解[1]。具体而言,DNA 编码序列设计的目的就是通过对 DNA 序列的优化从而最大限度地减小假阳性和假阴性的出现。假阳性是指不完全互补的 DNA 分子在适当的条件下也能够杂交形成双链分子;假阴性是指完全互补的 DNA 分子在反应过程中由于种种原因而没有杂交。DNA 编码序列的设计通常是利用约束条件来筛选序列,使其最大限度地满足要求。序列的设计方法理论上也可以分为阳性设计和阴性设计。阳性设计方法要求最优化亲和性使匹配序列充分杂交,包括 T_m 值约束、GC 含量约束、能量最小限度约束等。阴性设计方法则是最优化序列特异性,包括汉明距离约束、序列最小对称约束和最小自由能约束等。目前主要的研究方法包括文献[2]中提出的模板-映射法、文献[3]中提出的进化算法、文献[4]中提出的最小长度子串评价方法等。

DNA 编码序列设计问题是一种多目标优化问题。非支配遗传算法 NSGA-Ⅱ(Non-dominated Sorting Genetic Algorithm Ⅱ)是一种在 NSGA 算法的基础上通过引入精英策略、密度值估计策略和快速非支配排序策略的多目标遗传算法,它在很大程度上改善了 NSGA 的缺点。但是在目标函数太多的情况下会降低其收敛速度和种群多样性。本文通过将约束条件引入快速非支配排序过程,降低了算法的时间复杂度,提高了收敛速度和种群多样性,并设计出了 INSGA-Ⅱ的具体实现程序,得到了较高质量的 DNA 编码序列,验证了算法的有效性。

1 设计标准

通常,为保证设计出的序列具有稳定的化学性质,同时序列间有足够大的互异性,序列(间)必须满足以下约束条件。

1) 解链温度(T_m)约束。解链温度是决定反应效率的一个重要参数。为了有效降低不完全匹配双链产生的概率,保证生化反应的可靠性,要求参加反应的序列的 T_m 值基本一致。这里采用文献[5]中提出的 T_m 值计算公式:

$$T_m(x_i) = \frac{\Delta H}{\Delta S + R \times \ln(C/4)} - 273.15 \quad (1)$$

其中:R 为摩尔气体常数(1.987 cal / mol · K);C 为核酸浓度;ΔS 和 ΔH 分别表示在一定温度下各碱基之间的熵变与焓变。

2) 组分约束(GC 含量):GC 含量指的是 DNA 序列中碱基 G(鸟嘌呤)和 C(胞嘧啶)个数占所有碱基个数的比例。由于 G(鸟嘌呤)和 C(胞嘧啶)之间有三个氢键,而 A(腺嘌呤)

收稿日期:2009 - 05 - 11;**修回日期**:2009 - 07 - 26。 **基金项目**:国家自然科学基金资助项目(60573190;60773122);河南省基础与前沿技术研究计划资助项目(082300413203);河南省创新型科技人才队伍建设工程资助项目。

作者简介:王延峰(1973 -),男,河南南召人,副教授,博士,主要研究方向:生物计算、智能优化; 申永鹏(1985 -),男,河南内黄人,硕士研究生,主要研究方向:DNA 计算、智能优化; 张勋才(1981 -),男,河南郸城人,博士,主要研究方向:DNA 计算、算法设计与优化; 崔光照(1957 -),男,河南洛宁人,教授,博士,主要研究方向:智能计算、数字系统。

和 T(胸腺嘧啶)之间只有两个氢键,所以 GC 含量对保持序列化学性质的稳定非常重要。GC 含量也会影响 T_m 值的稳定,从而影响不完全匹配双链产生的概率。一般要求 $GC(x) \in [0.4, 0.6]$。

$$GC(x_i) = \frac{\#G + \#C}{|x|} \tag{2}$$

其中:$\#G$、$\#C$ 分别表示序列中碱基 G 和 C 的数量;$|x|$ 表示序列的长度。

3)二级结构约束。当 DNA 计算中出现长度较短(一般为 10 ~ 100 bp)的 DNA 单链序列时,由于不能形成更为复杂的折叠的卷曲,单链在有一定长度的反向互补序列时,就会形成图 1 所示的发夹结构。

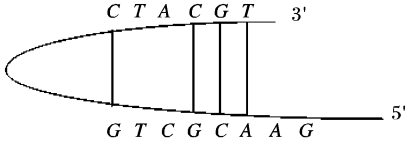

图 1　单链 DNA 形成发夹结构

虽然这种二级结构也能用于 DNA 计算,但是本文仅考虑所有 DNA 序列仅包含一级结构时的情形。所以,为降低生化操作时发夹等二级结构出现的概率,应尽量减小 DNA 单链中反向互补序列的长度。

$$f_{Hp}(X) = \sum_{i=1}^{n} \max_{k=-(n-2),\cdots,n-2} \sum_{j=1}^{|x|} c(x_{i_j}, \overrightarrow{x}_{i_{j+k}}) \tag{3}$$

其中:x_{i_j} 表示序列 x_i 中的第 j 个碱基;$\overrightarrow{x}_{i_{j+k}}$ 表示 x_i 的反向序列中的第 $j+k$ 个碱基;且有:

$$c(a,b) = \begin{cases} 1, & (a,b) = (A,T) \text{ or } (T,A) \text{ or } (G,C) \text{ or } (C,G) \\ 0, & \text{其他} \end{cases} \tag{4}$$

$f_{Hp}(x)$ 计算的是每个 DNA 单链中分别从 3′ 和 5′ 开始能匹配的碱基个数。显然,$f_{Hp}(x)$ 越小,形成二级结构的概率就越低。

4)连续性约束。当 DNA 序列中某一碱基连续出现的次数过多时(比如"$AAAAAA$")会导致 DNA 分子结构的不稳定,容易造成配对错位,使杂交反应不能得到良好的控制,所以必须对序列中连续碱基出现的次数加以控制,以尽量减小连续碱基出现的次数。

$$f_{Con} = \sum_{i=1}^{n} \left[\max_{j=1,\cdots,|x|} \sum_{k=j+1}^{|x|} u(x_{i_j}, x_{i_k}) - Con' + 1 \right]^2 \tag{5}$$

其中,Con' 表示由用户设定所允许的连续碱基出现的最大次数。

$$u(x_{i_j}, x_{i_k}) = \begin{cases} 1, & x_{i_j} = x_{i_k} \\ \text{break}, & \text{其他} \end{cases} \tag{6}$$

其中"break"表示跳出当前步骤的计算。

显然,f_{Con} 越小表示序列中连续碱基出现的次数越小。

5)汉明距离约束。汉明距离约束指的是任意两个序列中相对应位置上的碱基应尽可能不同,以确保序列之间相似度尽可能小,降低交叉杂交反应发生的概率。常见的距离约束还有反向汉明距离、移位汉明距离和补序列汉明距离等[1]。结合算法的时间复杂度,这里采用基本汉明距离和反向汉明距离。

$$f_{Hd}(X) = \sum_{i=1}^{n} f_{Hd}(x_i) \tag{7}$$

$$f_{Hd}(x_i) = \sum_{j=1}^{n} (Hd(x_i, x_j) + Hd(x_i, \overrightarrow{x}_j)) \tag{8}$$

其中 $\overrightarrow{x}_j$ 表示序列 x_j 的反向序列,且有:

$$Hd(x_i, x_j) = \sum_{m=1}^{|x|} d(x_{i_m}, x_{j_m}) \tag{9}$$

$$d(u,v) = \begin{cases} 1, & u = v \\ 0, & \text{其他} \end{cases} \tag{10}$$

2　数学模型

上面共考虑了五个设计准则,由于五个目标函数的多目标优化问题很难达到良好的收敛性和种群多样性,并且具有很高的时间复杂度。分析表明,把一些设计准则作为约束条件会比作为目标函数具有更好的效果。所以,这里把解链温度 T_m 和组分约束 GC 含量作为约束条件,把二级结构、序列的连续性、汉明距离作为目标函数,构造了一个带约束多目标优化模型。

$$\min F(X) = (f_{Hp}(x), f_{Con}(x), f_{Hd}(x)) \tag{11}$$

$$\text{s.t.} \quad \begin{cases} T_m^{\mathrm{L}} \leqslant T_m(x_i) \leqslant T_m^{\mathrm{U}}, & i = 1,2,\cdots,n \\ GC^{\mathrm{L}} \leqslant GC(x_i) \leqslant GC^{\mathrm{U}}, & i = 1,2,\cdots,n \end{cases}$$

其中:T_m^{L} 和 T_m^{U} 分别代表设置的 T_m 值的下限和上限,它们的选取是由特定的实验环境和对 T_m 值精确度的要求所决定的,文献[6]对 T_m 值的计算和选取做了具体描述,这里分别选取 66.5 和 61.5;GC^{L} 和 GC^{U} 分别代表设置的 GC 含量的下限和上限,可以根据具体的实验需求选取 GC 含量的下限和上限,这里分别选取 0.4 和 0.6。

3　算法设计

NSGA-Ⅱ是在 NSGA 算法基础上改进得到的高性能遗传算法[7],它主要采取三个策略来改善算法的性能:1)采用时间复杂度很低的快速排序法检查解的非支配排序等级;2)采用新的子代选取办法保证子代种群的优良;3)采用拥挤距离度量同一支配等级下解在目标空间的分布情况。为了解决上面提出的带约束条件 DNA 编码序列优化问题,这里对 NSGA-Ⅱ的非支配排序过程进行了改进。

在 NSGA-Ⅱ中,如果 p 支配 q,那么必须满足以下两个条件[8]:1)对于所有的目标函数 p 不比 q 差,即 $f_k(p) \leqslant f_k(q)$,$(k = 1,2,3,\cdots,r)$;2)至少存在一个目标函数使 p 比 q 好,即 $\exists l \in \{1,2,3,\cdots,r\}$,使得 $f_l(p) < f_l(q)$,其中,r 为目标函数个数。那么,在构造非支配集前,必须首先通过一个二重循环计算两个向量 $\{n_i\}$ 和 $\{s_i\}$,其中 $i \in Pop$,n_i 记录支配个体 i 的个体数目,s_i 记录被个体 i 支配的个体的集合 s_i,即有:

$$\begin{cases} n_i = |\{k \mid k > i \quad i,k \in Pop\}| \\ s_i = \{j \mid i > j \quad i,j \in Pop\} \end{cases} \tag{12}$$

在比较两个个体的支配关系的时候,首先检查两个个体是否满足约束条件,然后再根据其满足约束条件的情况进行非支配排序。即对于任意两个进行比较的个体 A、B,如果 A、B 均满足约束条件,则非支配解拥挤距离大者入选;如果 A 满足,B 不满足,则 A 入选;如果 A,B 都不满足,则由式(13)计算其 Pen 值,Pen 值小者入选。

$$Pen(p) = \omega_{gc} \cdot range(GC(p), GC^{\mathrm{L}}, GC^{\mathrm{U}}) + \omega_{tm} \cdot range(T_m, T_m^{\mathrm{L}}, T_m^{\mathrm{U}}) \tag{13}$$

其中:ω_{gc} 表示约束条件 GC 含量的权值;ω_{tm} 表示约束条件 T_m 值的权值,本文中均取 0.5。$range$ 函数的表达式如下[9]:

$$range(t,l,u) = \begin{cases} l-t, & t < l \\ t-u, & t > u \\ 0, & \text{其他} \end{cases} \tag{14}$$

上述过程由如下伪代码详细描述：

```
Fast non-dominated sort (P)
for each p ∈ P
  S_P = ∅; n_p = 0; p_rank = 1;
for each q ∈ P
  if (pen(p) = 0, pen(q) = 0) then
    if (p < q) then S_P = S_p ∪ q;
    else if (p > q) then n_p = n_p + 1;
  if (pen(p) = 0, pen(q) ≠ 0) then n_p = n_p + 1;
  if (pen(p) ≠ 0, pen(q) = 0) then S_P = S_p ∪ q;
  if (pen(p) ≠ 0, pen(q) ≠ 0) then
    if (p < q, pen(p) > pen(q)) then S_P = S_p ∪ q;
    if (p > q, pen(p) < pen(q)) then n_p = n_p + 1;
    else
      random (S_P = S_p ∪ q, n_p = n_p + 1);
  if (n_p = 0) then
    P_r = 1, F_1 = F_1 ∪ {p};
  i = 1;
  while F_i ≠ ∅
    F_{i+1} ≠ ∅;
  for each q ∈ S_P
    n_q = n_q - 1; q_rank = q_rank + p_r;
    if (n_q = 0) then
      q_r = i + 1;
      F_{i+1} = F_{i+1} ∪ {q};
      i = i + 1;
```

其中：S_P 表示由 p 支配的所有个体；n_p 表示支配 p 的个体数目；p_r 表示 p 的 Pareto 排序值；p_{rank} 表示 p 的累积 Pareto 排序值；F_i 表示第 i 级非支配集；$pen(p)$ 是计算 p 约束情况的一个函数。

算法的主要时间开销用于计算 n_i 和 s_i，时间复杂度为 $O(rN^2)$，r 为目标函数个数，N 为种群规模。这里，通过减少进化算法的目标函数，从而达到了降低算法时间复杂度的目的。

整个算法的流程如图 2 所示。

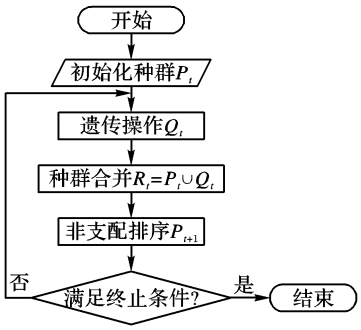

图 2 INSGA-Ⅱ算法流程

采用四进制数字编码基因的每个碱基，具体的编码规则如下：

{ 0 - A, 1 - C, 2 - T, 3 - G }

种群的初始化采用不考虑约束条件的随机初始化。初始化的过程如下：

1）随机产生 m 个均匀分布于[0,3]的整数，作为一个长度为 m 的序列；

2）重复步骤 1）共 n 次，得到由 n 个序列组成的染色体序列集合；

3）重复以上两个步骤 l 次，得到种群规模为 l 的初始种群。

选择操作采用锦标赛算子根据累积排序值和拥挤距离评价个体的优劣，累积排序值小，拥挤距离大的个体参与繁殖。交叉操作随机采用单点交叉、多点交叉和洗牌交叉，充分保证了种群分布的多样性。交叉操作和变异操作分为两个阶段，首先进行染色体之间的交叉和变异，即交换集合中的序列，然后在每个染色体内进行序列之间的交叉和变异操作[10]。算法停止准则一般是经过指定的遗传代数后停止运算。

4 仿真结果与分析

在 Matlab 7.0 编程环境下，使用 INSGA-Ⅱ对编码序列过程进行仿真，运行环境是 Intel Pentium Dual E2104 1.6 GHz，512 MB 内存，Windows XP。种群规模 50，交叉概率 0.7，变异概率 0.001，每个个体中包含 10 条长度为 20 的 DNA 序列。表 1 列出的是从进化 200 代后产生的结果中随机选出的一组 DNA 序列。

表 1 INSGA-Ⅱ编码序列性质表

INSGA-Ⅱ编码序列(5-3)	连续性	二级结构	相似性	T_m	GC
ATATTGATCAGGCCTAGACG	0	0	161	63.5028	0.45
GCGATGATATACTGTGGCGA	0	0	172	64.6527	0.50
CCATAACCGAACCGACTGTA	0	0	152	63.8384	0.55
ATCCATGTCTTGAACAGTGC	0	0	150	64.7618	0.50
GCTCATCAGTGTGCTACTCT	0	3	186	63.2197	0.50
ACCGGGAAGAAGAGCTTTGA	0	0	164	65.9024	0.50
ATCATACTCCGGAGACTACC	0	0	156	62.1413	0.50
ACCGCGTCTACCGAAGAATT	0	3	168	65.6801	0.50
GATGGTTCAGTATACGTGTA	0	3	152	63.8206	0.40
TCAGTAAGAGGACCGTAACG	0	0	169	65.1274	0.50

表 2 文献[11]编码序列

文献[11]编码序列(5-3)	连续性	二级结构	相似性	T_m	GC
CGAGACATCGTGCATATCGT	0	6	170	64.5958	0.5
TATAGCACGAGTGCGCGTAT	0	13	197	66.1329	0.5
GATCTACGATCATGAGAGCG	0	13	191	61.9863	0.5
TCTGTACTGCTGACTCGAGT	0	0	199	64.6786	0.5
CGAGTAGTCACACGATGAGA	0	3	198	63.2197	0.5
AGATGATCAGCAGCGACACT	0	0	196	66.0585	0.5
TGTGCTCGTCTCTGCATACT	0	3	188	65.6801	0.5
AGACGAGTCGTACAGTACAG	0	6	178	62.8642	0.5
ATGTACGTGAGATGCAGCAG	0	0	177	64.8098	0.5
ATCACTACTCGCTCGTCACT	0	0	180	65.0757	0.5

表 3 NACST/Seq. 编码序列[10]

NACST/Seq.编码序列(5-3)	连续性	二级结构	相似性	T_m	GC
GTGACTTGAGGTAGGTAGGA	0	0	158	62.2239	0.5
ATCATACTCCGGAGACTACC	0	0	156	62.1413	0.5
CACGTCCTACTACCTTCAAC	0	0	178	61.9397	0.5
ACACGCGTGCATATAGGCAA	0	3	154	67.2662	0.5
AAGTCTGCACGGATTCCTGA	0	0	163	65.9669	0.5
AGGCCGAAGTTGACGTAAGA	0	0	164	65.9024	0.5
CGACACTTGAAGCACACCTT	0	0	166	65.4594	0.5
TGGCGCTCTACCGTTGAATT	0	0	151	66.8953	0.5
CTAGAAGGATAGGCGATACG	0	3	151	61.0777	0.5
CTTGGTGCGTTCTGTGTACA	0	0	149	65.1612	0.5

图 3 所示为 INSGA-Ⅱ、文献[11]方法、NACST/Seq. 三种编码方法产生的序列的二级结构、相似性适应函数均值以及

T_m 值的标准差,可以看出:INSGA-Ⅱ方法在控制生成序列的二级结构、相似性和 T_m 值方面均有良好的效果,尤其是在控制 T_m 值的一致性方面明显优于其他两种方法。在二级结构和相似性两个方面效果虽然不明显,但总体来说 INSGA-Ⅱ是解决 DNA 编码序列问题的一种有效方法。

目标函数的收敛特性曲线如图 4 所示,从中可以看到三个目标函数在 40 代左右就已经趋于稳定,并且产生了优良的结果,说明 INSGA-Ⅱ在解决 DNA 编码序列优化问题上具有很好的收敛性,并且能够保证种群的多样性。

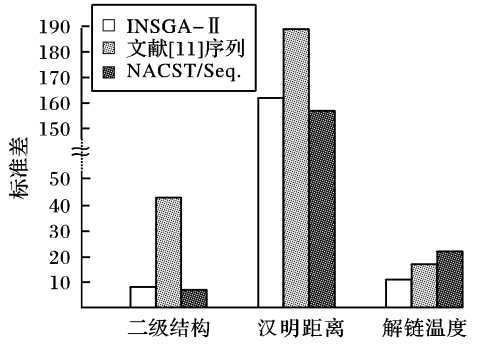

图 3　三种 DNA 编码序列比较结果图

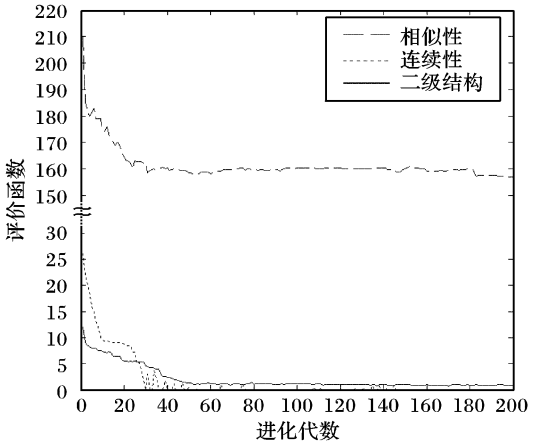

图 4　INSGA-Ⅱ收敛特性曲线

5　结语

一组能够保证可控 DNA 计算生化反应顺利进行的 DNA 序列是 DNA 计算的先决条件,所以如何设计一组满足各种约束条件的 DNA 序列是近几年来 DNA 计算领域的核心内容之一。由于 DNA 编码序列各约束条件之间存在着相互制约,而且不同的生物实验条件和方法对各约束条件的要求也不同,所以根据相关约束条件建立一套通用的 DNA 编码序列设计方案是一件非常困难的事情。本文提出的 INSGA-Ⅱ DNA 编码序列设计方法是在 NSGA-Ⅱ方法的基础上与 DNA 编码设计具体要求相结合设计出的一种方法,具有算法时间复杂度小,通用性强等优点,并通过与文献[11]和 NACST/Seq. 两种方法相比较验证了算法的有效性。进一步的工作应该结合具体的生化操作设计约束性更强,算法更简单的约束条件,并设计出评价 DNA 编码序列质量的通用评价体系。

参考文献:

[1] GARZON M H, DEATON R J. Codeword design and information encoding in DNA ensembles [J]. Natural Computing, 2004, 3(3): 253 – 292.

[2] FRUTOS A G, LIU Q, THIEL A J, *et al.* Demonstration of a word design strategy for DNA computing on surface [J]. Nucleic Acids Research, 1997, 25(23): 4748 – 4757.

[3] DEATON R, MURPHY R C, ROSE J A, *et al.* A DNA based implementation of an evolutionary search for good encodings for DNA computation [C]// Proceedings of 1997 IEEE International Conference on Evolutionary Computation. Washington, DC: IEEE Press, 1997: 267 – 271.

[4] FELDKAMP U, BANZHAF W, RAUHE H. A DNA sequence compiler [EB/OL]. [2009 – 03 – 16]. http://www. cs. mun. ca/ ~ banzhaf/papers/DNASeqComp. pdf.

[5] SANTALUCIA J, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics [C]// Proceedings of the National Academy of Sciences of the United States of America. Washington, DC: National Academy of Sciences of the United States of America, 1998, 95: 1460 – 1465.

[6] 王非,杨欣,郑珩. 核酸熔解温度的研究进展与寡核苷酸设计平台的研制[J]. 生命的化学,2004,24(2):157 – 160.

[7] DEB K, PRATAP A, AGARWAL S, *et al.* A fast and elitist multiobjective genetic algorithm: NSGA-Ⅱ [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182 – 197.

[8] 郑金华. 多目标进化算法及其应用[M]. 北京: 科学出版社, 2007: 1 – 24.

[9] SHIN S-Y, LEE I-H, KIM D, *et al.* Multiobjective evolutionary optimization of DNA sequences for reliable DNA computing [J]. IEEE Transactions on Evolutionary Computation, 2005, 9(2): 143 – 158.

[10] SHIN S Y, KIM D M, LEE I H, *et al.* Evolutionary sequence generation for reliable DNA computing [C]// CEC '02: Proceedings of the 2002 Congress on Evolutionary Computing. Washington, DC: IEEE Press, 2002, 1: 79 – 84.

[11] TAKAHARA A, YOKOMORI T. On the computational power of insertion-deletion systems [J]. Natural Computing, 2003, 2(4): 321 – 336.

(上接第 3039 页)

[3] PARINYA S, WIDHYAKORN A, SOMCHAI J, *et al.* Two-dimensional linear discriminant analysis of principle component vectors for face recognition [J]. Journal Article, 2006, 89(7): 2164 – 2170.

[4] NOUSHATH S, KUMAR G H, SHIVAKUMARA P. $(2D)^2$LDA: An efficient approach for face recognition [J]. Pattern Recognition, 2006, 39(7): 1396 – 1400.

[5] BARTL ETT M S. Face image analysis by unsupervised learning and redundancy reduction [D]. California, CA: University of California, 1998: 27 – 37.

[6] YANG C-H T, LAI S-H, CHANG L-W. Robust face matching under different lighting conditions [C]// ICME '02: 2002 IEEE International Conference on Multimedia and Expo. Washington, DC: IEEE Press, 2002, 2: 149 – 152.

[7] EBRAHIMPOUR-KOMLEH H, CHANDRAN V, SRIDHARAN S. Robustness to expression variations in fractal-based face recognition [C]// Proceeding of the 2001 International Conference on Signal Processing and its Applications. Washington, DC: IEEE Press, 2001: 359 – 362.

[8] 郝广涛,胡步发. 基于 DDGVFSnake 和 Gamma 方法处理人脸光照不均[J]. 计算机应用,2007, 27(4):925 – 928.

[9] CHOI H, CHOI S. Relative gradient learning for independent subspace analysis [C]// Proceeding of the 2006 International Joint Conference on Neural Networks. Washington, DC: IEEE Process, 2006: 3919 – 3924.

[10] LAI S-H, WEI S-D. Reliable image matching based on relative gradients [C]// Proceeding of the 16th International Conference on Pattern Recognition. Washington, DC: IEEE Process, 2002: 802 – 805.