

基于脱氧核糖核酸计算模型的概念创新设计

吴海龙¹, 苏 强¹, 谭丽娜¹, 贾 赛²

(1. 同济大学 经济与管理学院, 上海 201804; 2. 西安武警工程大学 军事基础教育学院, 西安 710000)

摘要: 采用脱氧核糖核酸(DNA)计算方法解决了概念模型筛选问题。首先,引入骨骼结构表示概念模型,构建了骨骼结构及设计了偏转角度和骨骼弧度DNA分子编码方式。其次,在提出了DNA计算的算法流程后用计算机模拟实验的方法验证了算法流程的可行性。然后,依据提取出的DNA分子片段长度并按照K-means方法把每种组件划分为了5类。再次,采用信息熵计算方法,计算出了每一个评价参数权重之后,建立了概念模型的评价函数。最后,实验结果表明:合理的评价函数可以描述出概念模型的结构特点,进而根据其评价函数可以快速有效地找到符合需求的设计方案。此外,由于相似的概念模型其评分很接近,可以通过比较相近评分的概念模型提取出共有的设计组件,从而为后续设计开发提供素材。

关键词: 脱氧核糖核酸计算; 创新设计; 骨骼结构

中图分类号: TP319

文献标志码: A

Conceptual Innovative Design Based on Deoxyribonucleic Acid Computing Model

WU Hailong¹, SU Qiang¹, TAN Lina¹, JIA Sai²

(1. School of Economics and Management, Tongji University, Shanghai, 201804 China; 2. School of basic military education, Engineering University of Chinese Armed Police Force, Xi'an 710000, China)

Abstract: In this paper, conceptual model's screening problems are solved via using deoxyribonucleic acid (DNA) computing. Firstly, it analyzes skeleton structure to represent conceptual model. And then it proposes the deflection angle and skeleton curve DNA encoding method. Secondly, it designs the DNA computing algorithm process, whose feasibility is verified by the computer simulation. According to different lengths, DNA fragments are divided into 5 categories through K-means method. Thirdly, the evaluation parameter weights are determined by using information entropy calculating method and the evaluation

function of conceptual model is established. Finally, the experimental results show that a reasonable evaluation function can basically describe the structural features of conceptual model and it could help designers find solutions effectively and efficiently. Furthermore, some common components can be extracted by comparing their scores due to the conceptual models, because similar scores of models have a high similarly and thus common components can provide raw materials for further design in the future.

Key words: Deoxyribonucleic acid computing; innovative design; skeleton structure

概念设计是产品创新的核心阶段,其设计的理论和方法反映了设计人员不同的思维规律^[1]。创新设计系统主要通过组织、配置和推送各类资源来支持创新设计,现有研究主要集中在三方面:①提供创新方法和理论^[2];②提供计算机辅助设计等可视化建模工具,增加设计想法的可视化程度,减少设计认知负荷^[3];③提供与设计相关的知识库和信息源,用于扩展问题空间和方案空间,扩大创新设计的激励范畴^[4]。其中结构变异是一种更新知识库和信息源主要的设计方法^[5]。结构变异设计方法是通过设计者从一个已知的可行结构方案出发,通过变异设计,得到大量的可行方案。结构变异设计的目的是寻找符合设计要求的、独立的结构设计方案,以便能够通过参数设计得到优化的结构解^[6]。通过变异设计得到的独立的设计方案数量越多,覆盖的范围越广泛,通过参数设计得到全局最优解的可能性就越大。但是随着结构设计的范围的增大,如何筛选可行的设计方案成为困扰设计者的主要问题。文献[7]提出采用遗传算法优化候选种群,但是二进制编码结构不能准确描述概念设计的进化方向。文献[8]提出构建模型库的方式筛选可行方案,但是该方案模型库却

需要预先设定,设计的思路会有受到一定限制.

随着概念模型组件的增加,概念模型的理论数量呈指数增长,概念模型的筛选问题已成为一个非确定多项式问题(non-deterministic polynomial, NP)^[7-8].因此,一般的优化算法难以满足概念设计模型优化的需求.脱氧核糖核酸(Deoxyribonucleic acid, DNA)计算自问世以来以运算速度快,高度并行性,信息储存量巨大等特点获得了许多学者的关注.自 Adleman^[9]提出了利用DNA分子计算解决7个节点的哈密尔顿路径问题后,DNA计算迅速成为活跃的研究领域.而后 Lipton^[10]很快地提出了基于DNA模型的DNA算法,成为求解NP完全问题的有效算法.但是由于DNA编码困难成为制约DNA计算快速发展的主要障碍.近年来,Jiao和Zhong等^[11]采用高光谱遥感技术大大提高了识别DNA序列的效率和准确度.Chaves-González等^[12]采用多目标优化的方式优化DNA分子序列,采用这种方式编码的分子序列大幅度提高了计算的稳定性,为后续研究提供了有效的技术支撑.因此,在现有技术条件下,一些学者采用DNA计算并成功解决了许多NP问题,如最小二分法问题^[13]、最小生成树问题等^[14].此外,利用DNA计算可以解决一部分数据挖掘问题,比如网格聚类^[15]、空间聚类等^[16].由于创新设计中概念模型的选择是一个NP问题,因此将DNA计算用于解决概念模型筛选问题既具有理论的可行性,又具有现实的必要性.

1 计算模型

1.1 DNA计算原理

DNA是由2条极长的脱氧核苷酸链组成的.每一分子的脱氧核苷酸是由一分子的磷酸、一分子脱氧核糖和一分子的含氮碱基组成的,4种含氮碱基两两配对(A=T,C=G).DNA计算基于双螺旋碱基互补配对原则,将问题转化为相应的编码链,通过一系列的生化操作,提取最终反应生成物得出问题解.

1.2 概念设计模型

产品的概念设计模型有多种表达方式,文献[17]提出的骨骼结构表示设计模型.针对一个设计对象,骨骼是多条首尾相连连接的线段所组成可以在一定程度上表征设计对象外观特征的线段集合.文献[7]提出增加骨骼弧度用于描述骨骼结构,这种改进的骨骼结构能更加准确地描述模型的变化.因此,本文采用改进的骨骼模型描述概念设计模型.

定义1 一个具有n段线段长的骨骼S,它可以看作为一个三元组(α, d, w),即三个向量的集合,其中 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{n-1})$, $d = (d_1, d_2, \dots, d_{n-1})$, $w = (w_1, w_2, \dots, w_{n-1})$.其中, $\alpha_i \in \{\alpha: |\alpha| < \beta\}$ 是角度的定义域, $w_i \in \{w: |w| < \delta\}$ 是骨骼的弧度定义域.

如图1所示: α_i 表示点*i*和点*i*-1连成的线段与点*i*-1和点*i*-2之间延长线的角度. d_i 表示第*i*条线段的长度. w_i 表示骨骼*i*的弧度.为了阐明定义的方便,下面举例说明.首先随机生成点*p₀*,以*d₁*长度生成点*p₁*,然后以*d₂*为长度,*p₀p₁*的反向延长线为轴,左右摆动,根据偏转角度 α_2 确定点*p₂*的位置.

采用圆上任意两点之间的弧型来表示一段骨骼,圆弧的两个端点就代表了骨骼的两个端点.那么骨骼的弧度就可以采用这段圆弧的弧度来表示.弧度 $w_i \in \{w: 0 < w < 2\pi\}$.规定顺时针方向为正,逆时针方向为负. w_i 线段的弧度可以向上凸,也可以向下凹.同样规定两点之间弧线在两点连接的直线之上的弧度为正,在连线之下弧度为负.如图2所示,点*p₁*、*p₂*之间是一段骨骼,*v₀*是连线的中点,*v₁*、*v₂*、*v₃*分别是三条骨骼的中心点.显然, $w_1 > w_2 > w_0 > w_3$,所以*v₁*、*v₂*的弧度为正,*v₃*的弧度为负.以此类推,直到所有的点位置都确定.任意一个产品概念模型,都可以表述为若干部件的组合,骨骼结构适用于描述线条形产品的设计,其中的每一段骨骼表示一个部件.

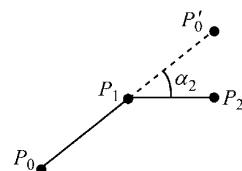


图1 偏转角度

Fig.1 Deflection angle

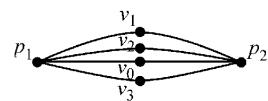


图2 骨骼弧度

Fig.2 Skeleton curve

1.3 评价函数

一个优秀的设计往往是由许多部分组成的.在各个部件的组合过程中,单个最优秀的个体组合的效果往往不能达到最优秀的效果,解决这个问题传统的办法是依靠设计人员的经验.采用这种办法效率很低,特别是针对组成部件数量特别多,组合

方法特别复杂的案例,传统的设计方法就无能为力。例如:假设产品的部件数量 n ,每个部件 i 拥有的属性(或参数) m_i ,最大有可能产生 $\prod_{i=1}^n m_i$ 种选择。对于设计者来说,采用人工方式一一验证每一种实验方式成了一个不可能完成的任务。文献[18]提出采用某些评价函数的方法辅助设计人员决策。为了描述设计产品,文献[7]提出根据目标函数找出一个适应度函数用于描述设计的产品。将概念设计模型中主要影响因素纳入评价函数中,函数的结果可以描述整体造型的特点。由于任意一条骨骼结构都可以用 n 个偏转角度和 m 段骨骼表示,因此,该模型的评价函数应该包括偏转角度和骨骼弧度两个变量。故评价函数如式(1)所示:

$$f = w_1\alpha_1 + w_2\alpha_2 + w_3\alpha_3 + w_4\alpha_4 + w_5\Omega_1 + w_6\Omega_2 + w_7\Omega_3 + w_8\Omega_4 \quad (1)$$

式中: w_i 表示因素 i 的评价权重,且 $\sum_{i=1}^8 w_i = 1$; α_i 表示偏转角度 i 标准得分; Ω_i 表示骨骼弧度 i 的标

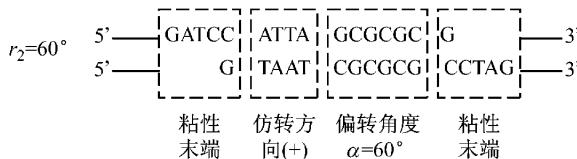


图 3 偏转角度编码结构图

Fig. 3 Deflection angle encoding structure

2.2 算法流程

步骤 1 利用生物技术手段合成含有 n 种内切酶切割位点的闭环 DNA 分子,分子长度为 200 bp。并利用 PCR(polymerase chain reaction) 技术进行扩增,放入试管 a 中。

步骤 2 根据第 2.1 节中所述编码方式,合成代表不同偏转角度和骨骼弧度的线性 DNA 分子片段。利用聚合酶链式反应技术扩增片段并将 r_i 内切酶分位点所表示所有可能的角度(或弧度)的 DNA 片段放入第 i 号试管中。令 $n=1$ 。

步骤 3 将对应的第 n 号限制性内切酶和试管 a 中的溶液混合均匀放入试管 b 中,在一定条件下将闭环 DNA 分子在第 n 号内切酶分位点,打开变成线性 DNA 分子。将第 n 号试管溶液倒入试管 b 中混合,加入 T_4 DNA 连接酶,使 DNA 片段与打开的闭环 DNA 分子发生杂交反应。去掉两边的保护基,再用高效液相层析法除去杂质,在 T_4 DNA 连接酶作用下,使所有打开的闭环 DNA 重新关闭,成为闭环 DNA,再利用高效液相层析法除去杂质。

准得分。

2 DNA 计算操作流程

2.1 DNA 编码

实验中 DNA 编码采用可变长度的编码方式。在满足生物反应要求基础上,采用不固定编码长度的编码方式,通过编码链长短来表示参数的大小变化。

随着偏转角度绝对值的增大,编码链的长度依次增加。设定长度每增加 1 bp,偏转角度增大 10° 。偏转角度 $\alpha \in [-180^\circ, 180^\circ]$,则偏转角度片段长度 $L_\alpha \in [9 \text{ bp}, 27 \text{ bp}]$ 。如 r_2 位置表示偏转角度 $\alpha=60^\circ$ 时的编码(图 3):

随着弧度逐渐增加,编码链的长度也依次增加。设定骨骼弧度每增加 1 bp,骨骼弧度增加 $1/12\pi$ 。骨骼弧度 $\Omega \in [0, 2\pi]$,则骨骼弧度片段长度 $L_\Omega \in [5 \text{ bp}, 29 \text{ bp}]$ 。如 r_3 位置表示骨骼弧度 $\Omega=2/3\pi$ 时的编码(图 4):

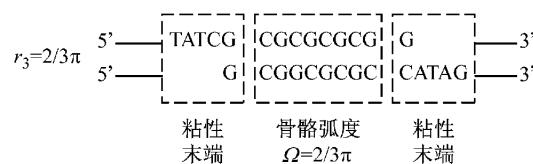


图 4 骨骼弧度编码结构图

Fig. 4 Skeleton curve encoding structure

步骤 4 清空试管 a ,将含有新生成的闭环 DNA 分子溶液倒入试管 a 中,清空试管 b 。如果 $n < 8, n=n+1$,转入第 3 步;否则,进入第 5 步。

步骤 5 加入第 0 号限制性内切酶,将闭环 DNA 分子打开变成线性 DNA 分子,除去相应的杂质。加热将双链 DNA 解链,加入所需要骨骼弧度数所对应的编码链的补链,退火反应,生成 DNA 分子双链,去除相应杂质。

步骤 6 利用凝胶电泳实验测量试管 a 中 DNA 分子长度,根据适应度函数值,选取适合长度的 DNA 分子,并测量其分子序列。

步骤 7 将测量出的 DNA 分子序列转化为相对应的骨骼弧度和偏转角度,既实验的结果。

3 设计实例

3.1 吊灯设计实例

现实生活中有很多物体的主要特征可以用骨骼结构表示,如:植物的枝干,相关的艺术图形,室内部

分的吊灯造型。如图 5 展现的是一个精美的吊灯骨架抽象,其中圆圈代表灯泡,曲线代表支架。

去掉多余的装饰品以及不相干的物件,通过手工方式可以得到吊灯轮廓图。图 6 是作者通过手工操作抽取的吊灯的骨骼结构,可以看出骨骼结构的抽象图并没有完全与原图像一一对应。骨骼的结构代表了整个吊灯框架,对骨骼结构的设计也就代表了对整个吊灯的设计。根据提取出来的骨骼,确定吊灯的每一条边由 5 个关键的点(包括起始点,起始点不进行编码),4 段骨骼组成。根据图 6,把整个吊灯抽象地表示成由 5 条相同的骨骼的线段。由于 5 条骨骼空间结构是相同的,可以只用拿出一条边在二维平面内进行编码,通过相应的技术手段就可以实现对所有骨骼的编辑。

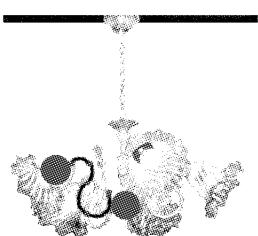


图 5 骨骼抽取

Fig.5 Abstracted skeleton

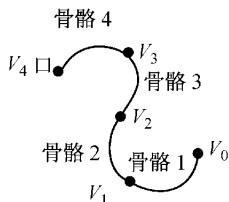


图 6 骨骼结构

Fig.6 Skeleton structure

3.2 计算机仿真模拟试验

由于客观实验条件所限,本文采用计算机模拟 DNA 计算的生化反应过程。由于本文中所用的 DNA 编码方法和计算模型是基于已有的 DNA 计算方法和模型,本文中的计算和操作的正确性依赖于已有的 DNA 计算方法和模型的正确性。这些现有的 DNA 计算方法和模型已经被各领域的专家和学者充分验证并使用^[9-18],因此本文的方法和模型的正确性也是值得信赖的。实验使用的计算机 CPU Intel (R) Core(TM)2 Duo CPU E7400,主频 2.8GHz, 2.0GB 内存。实验环境是基于 Microsoft XP SP3 操作系统,Microsoft Visual studio 2005 平台编译而成,其中实体的造型是利用 ACIS—HOOPS 类库实现。

尽管电子计算机不具有 DNA 计算超高速的计算速度和海量的并行性,但是仍然可以借助电子计算机模拟和验证 DNA 计算的实验可行性。实验随机产生 1 000 000 条 DNA 编码,初始闭环 DNA 长度 200 bp,每个骨骼弧度单链 DNA 片段为 5~29 bp,每个偏转角度单链 DNA 片段为 9~27 bp。故最终结果 DNA 编码链长度为 256~424 bp 之间。

3.3 参数标准化

案例中包括偏转角度和骨骼弧度两类评价指标,每类指标分别有 4 个参数,共计 8 个评价参数。根据角度和弧度的大小,本文将每一个评价参数划分 5 类,从小到大依次为“非常平缓”、“较平缓”、“一般”、“较弯曲”和“非常弯曲”。设计人员可以根据设计需求,分别给每一个类别一个基准分。假如在该案例中,希望得到概念模型“较弯曲”的造型,则可以依次设定其基准分为 $S=\{1, 3, 5, 7, 5\}$ 。分别对每一个指标采用 K-Means 方法聚类,得到每一个指标的聚类中心(C)。见表 1~2。

表 1 偏转角度聚类中心

Tab.1 Cluster centers of deflection angle

	非常平缓	较平缓	一般	较弯曲	非常弯曲
α_1	9	11	15	19	24
α_2	10	14	18	21	26
α_3	11	15	19	23	26
α_4	10	13	18	23	26

表 2 骨骼弧度聚类中心

Tab.2 Cluster centers of skeleton curve

	非常平缓	较平缓	一般	较弯曲	非常弯曲
Ω_1	5	7	11	17	25
Ω_2	6	11	17	22	27
Ω_3	6	11	18	24	28
Ω_4	6	11	16	22	27

根据每一个点的隶属度及其类给定的基准分,可以得到每一个参数的标准分。如图 7 所示,生成某一个概念设计模型中 $\Omega_1=10$,则其隶属度 $Q_{\Omega_2}=\{0, 0.25, 0.75, 0, 0\}$ 。进而可以计算出其标准得分, $S_{\Omega_2}=S \cdot Q_{\Omega_2}^T=\{1, 3, 5, 7, 5\} \cdot \{0, 0.25, 0.75, 0, 0\}^T=$

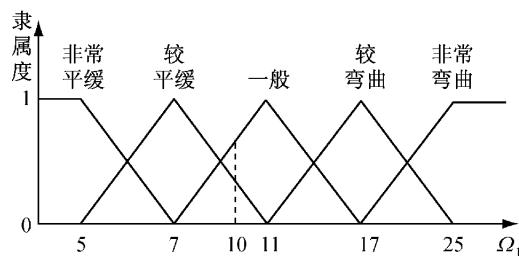


图 7 隶属度函数

Fig.7 Membership function

9/2. 同理, 可以得到每一个骨骼弧度或偏转角度的标准分.

3.4 参数权重

采用香农的信息熵理论确定个评价参数的权重. 根据香农的信息熵理论^[19], 概念设计模型中, 如果某一个参数为区分与其他参数提供的有效信息越多, 则就给该参数赋予较高的权重, 反之, 如则赋予较低的权重. 采用模糊评价方法可以得到每一个偏转角度和骨骼弧度的评价权重, 见表 3.

表 3 评价参数权重表

Tab. 3 Weights of valuation parameters

偏转角度				骨骼弧度				
α_1	α_2	α_3	α_4	Ω_1	Ω_2	Ω_3	Ω_4	
权重	0.078	0.122	0.129	0.128	0.101	0.127	0.175	0.140

3.5 实验结果

根据表 3 和式(1), 可以计算出每一个概念设计模型的评价得分. 开发者可以根据预先设定的需求, 选择出符合设计意愿的概念设计模型. 图 8 分别给出了不同评分阶段下的吊灯概念设计模型. 可以发现:

(1) 随着评分的升高, 概念模型越来越符合预先设定的需求. 图 8 中分别展示了“较弯曲”(图 8a)、“非常弯曲”(图 8b)、“一般”(图 8c)、“较平缓”(图 8d)和“非常平缓”(图 8e)5 种设计结构. 其中“较弯曲”设计评分最高, “一般”和“非常弯曲”评分次之, “较平缓”再次之, “非常平缓”设计评分最低.

(2) 通过设定评分的阈值, 就可以挑选出设计“较弯曲”的设计模型. 评分相近的设计模型可能表现出具有类似的设计结构. 比如, 在本案例中评分在 6.24~6.26 之间的吊灯设计模型多数具有类似底座设计, 评分在 4.63~4.65 之间的吊灯设计模型会具有闭包结构等.

以上两点发现表明, 设计者既可以通过评分的高低筛选出符合预先设定需求的设计结构, 也可以通过选取特定评分的模型, 提取出共有的设计结构, 用于设计后续阶段使用.

4 结语

将 DNA 计算引入创新设计中, 有助于解决当前创新设计中存在的两个主要问题. 首先借助 DNA 计算特性, 可以提高设计者寻找概念模型的效率. 根据设计的需求, 设定合适的基准分. 通过骨骼结构描述而设定的评价函数可以帮助设计者准确而有效地描

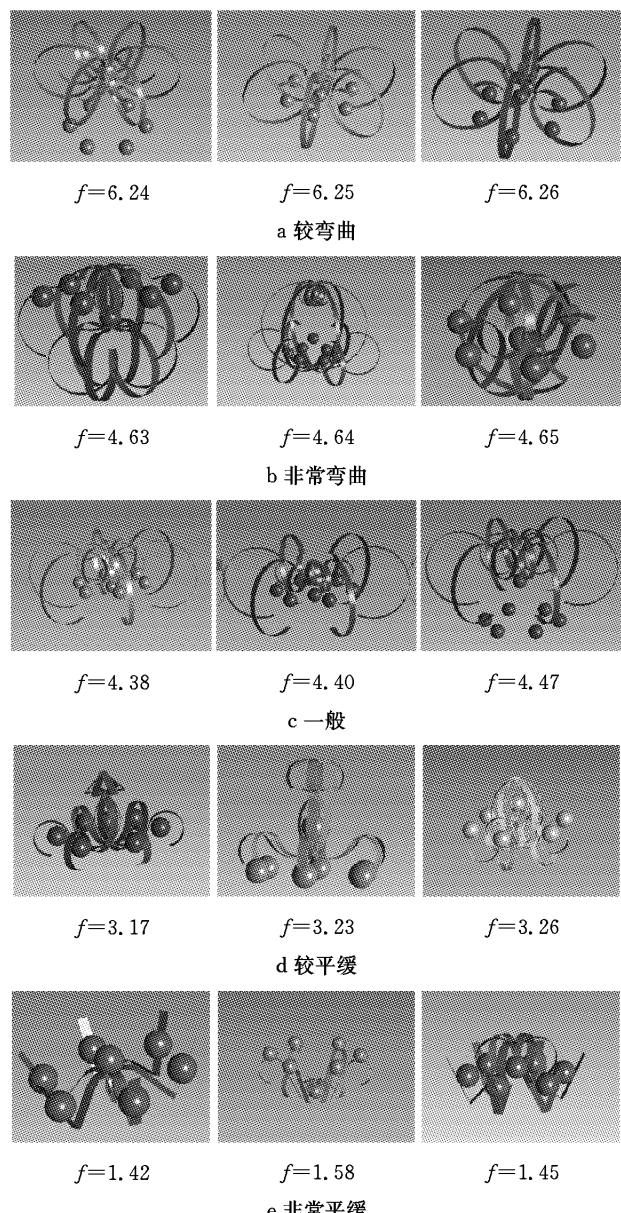


图 8 吊灯概念设计模型

Fig. 8 Lamp conceptual design model

述概念设计模型. 可以有针对性地搜寻目标模型, 大大减少了搜寻空间, 从而提高了概念模型的搜寻效率. 其次, 实验中发现设计相近的概念模型, 评价分数也十分接近. 设计者可以通过比较相近评分结构, 提取出共有的组件. 经过整理之后的设计组件可以大大丰富设计的案例库, 为后续的设计工作提供有效帮助. 当然, DNA 计算为寻找合适的设计模型提供了一种有效的计算方法, 其仍具有和其他算法进一步组合开发的潜力. 将 DNA 计算与其他算法结合是未来 DNA 计算解决创新设计问题的一个研究方向.

参考文献:

- [1] Faerber S J, Carbon C C. Jump on the innovator's train: cognitive principles for creating appreciation in innovative product designs [J]. Research in Engineering Design, 2013, 24(3): 313.
- [2] Komoto H, Tomiyama T. A framework for computer aided conceptual design and its application to system architecting of mechatronics products [J]. Computer-Aided Design, 2012, 44(10): 931.
- [3] Koutsabasis P, Vosinakis S, Malisova K, et al. On the value of Virtual Worlds for collaborative design [J]. Design Studies, 2012, 33(4): 357.
- [4] Goel A K, Vattam S, Wiltgen B, et al. Cognitive, collaborative, conceptual and creative-four characteristics of the next generation of knowledge-based CAD systems: a studying biologically inspired design [J]. Computer-Aided Design, 2012, 40(10): 879.
- [5] Tseng K C, ElGanzoury W. An intelligent system based on concurrent engineering for innovative product design at the conceptual design stage [J]. The International Journal of Advanced Manufacturing Technology, 2012, 63(5-8): 421.
- [6] Kremer G O, Chiu M C, Lin C Y, et al. Application of axiomatic design, TRIZ, and mixed integer programming to develop innovative designs: a locomotive ballast arrangement case study [J]. The International Journal of Advanced Manufacturing Technology, 2012, 61(5-8): 827.
- [7] 吴海龙, 刘希玉, 向来生. 基于遗传算法的骨骼结构模型概念创新设计[J]. 计算机应用, 2011, 31(4): 1155.
WU Hailong, LIU Xiyu, XIANG Laisheng. Skeleton structure of conception innovative design based on genetic algorithm [J]. Journal of Computer Applications, 2011, 31(4): 1155.
- [8] 张凯, 刘希玉. DNA 计算在产品创新设计中的应用[J]. 计算机工程, 2011, 37(22): 219.
ZHANG Kai, LIU Xiyu. Application of DNA Computation in Product Creative Design [J]. Computer Engineering, 2011, 37(22): 219.
- [9] Adleman L M. Molecular computation of solutions to combinatorial problems [J]. Science, 1994, 266 (5187): 1021.
- [10] Lipton R J. DNA solution of hard computational problems [J]. Science, 1995, 268(5210): 542.
- [11] Jiao H, Zhong Y, Zhang L. Artificial DNA computing-based spectral encoding and matching algorithm for hyperspectral remote sensing data [J]. IEEE Transactions on Geoscience and Remote Sensing, 2012, 50(10): 4085.
- [12] Chaves-González J M, Vega-Rodríguez M A. DNA strand generation for DNA computing by using a multi-objective differential evolution algorithm [J]. Biosystems, 2014, 116: 49.
- [13] Liu X, Yang X, Li S, et al. Solving the minimum bisection problem using a biologically inspired computational model [J]. Theoretical Computer Science, 2010, 411(6): 888.
- [14] Wang Z, Huang D, Meng H, et al. A new fast algorithm for solving the minimum spanning tree problem based on DNA molecules computation [J]. Biosystems, 2013, 114(1): 1.
- [15] Zhang H, Liu X. A CLIQUE algorithm using DNA computing techniques based on closed-circle DNA sequences [J]. Biosystems, 2011, 105(1): 73.
- [16] Liu X Y, Xue J. Spatial cluster analysis by the bin-packing problem and dna computing technique [J]. Discrete Dynamics in Nature and Society, 2013, 35(10): 845.
- [17] 郑自然. 基于智能计算的计算机辅助建模方法研究[D]. 济南: 山东师范大学, 2010.
ZHENG Ziran. Research on computer-aided modeling technique based on intelligent computing [D]. Jinan: Shandong Normal University, 2010.
- [18] 刘弘, 刘希玉. 支持外观造型创新设计的进化计算方法[J]. 计算机辅助设计与图形学学报, 2006, 18(1): 101.
LIU Hong, LIU Xiyu. An evolutionary computing approach for supporting creative configuration design [J]. Journal of Computer-Aided Design & Computer Graphics, 2006, 18(1): 101.
- [19] Su Q, Lai S J, Liu L. Geometric computation based assembly sequencing and evaluating in terms of assembly angle, direction, reorientation, and stability [J]. Computer-Aided Design, 2009, 41(7): 479.