

运用遗传算法进行智能音乐作曲研究

黄澄宇, 汪镭, 康琦, 吴启迪

(同济大学电子与信息工程学院, 上海 201804)

5 **摘要:** 该论文采用遗传算法进行算法作曲, 主要利用遗传算法的全局最优性, 并且相对简单有效的特点。利用简洁的编码方式对基因进行编码, 有效地表达了乐曲的主要信息, 方便了适应度函数的计算和相关遗传操作的进行。针对音乐质量评估问题, 提出了将相关音乐知识的规则和人机交互相结合的模式, 共同对所创作音乐进行合理评价。不仅提高了评价的准确性, 而且大大降低了人的工作量, 产生了良好的音乐工程效果。

10 **关键词:** 遗传算法; 算法作曲; 适应度函数。

中图分类号: TP391

Algorithm Composition By Genetic Algorithm

15 HUANG Chengyu, WANG Lei, KANG Qi, WU Qidi

(Electronic Science and Technology School, Tongji University, Shanghai 201804)

Abstract: Genetic Algorithm is one of global optimality algorithms and it's relatively simple and effective, so we use it to do algorithmic composition. In this paper, encoding genes encode effectively to express the main information of the music, and to facilitate the calculation of the fitness function and the associated genetic manipulation. According to the basic music knowledge and rules, fitness function is designed. At the same time, this paper also use interaction between computer and human to evaluate the music. This strategy not only improves the accuracy of the evaluation, but also greatly reduce the human workload. So this system can produce good result.

20 **Key words:** Genetic Algorithm; Algorithm Composition, Fitness Function.

0 引言

算法作曲时试图应用某个形式化的过程, 使人利用计算机进行音乐创作时介入程度达到最小的研究^[1]。对于算法作曲的研究, 有了一定的发展, 不同研究者利用不同的方法进行算法作曲, 主要算法有: Markov 链, 随机过程, 知识库系统, 音乐文法, 人工神经网络以及遗传算法。当今, 对算法作曲的研究主要面临四个问题, 分别是: 音乐知识的表达问题, 创造性

30 和人机交互问题, 音乐创作风格问题和系统生成的作品质量问题。

遗传算法(genetic algorithms)是一种使用适应函数(fitness function)来演化候选者(染色体)的全局优化算法。在使用遗传算法进行音乐创作的工作中, 主要是构造适应函数, 以此来评估及选择系统生成的旋律问题。作曲系统可以通过建立适应函数的计算模型来演化旋律(染色体)或连接旋律的和声结构。由于在自动作曲过程中, 适应函数很难准确定义, 研究者提出交互式遗传算法。主要包括: Biles 构造的“GenJam”交互式即兴演奏系统, Unemi 用模拟繁殖的方法开发的一个音乐工具 SBEAT 等^[2]。

40 本文提出一种相对简洁方便的编码方式, 采用一些综合基本乐理知识和相关规则建立的

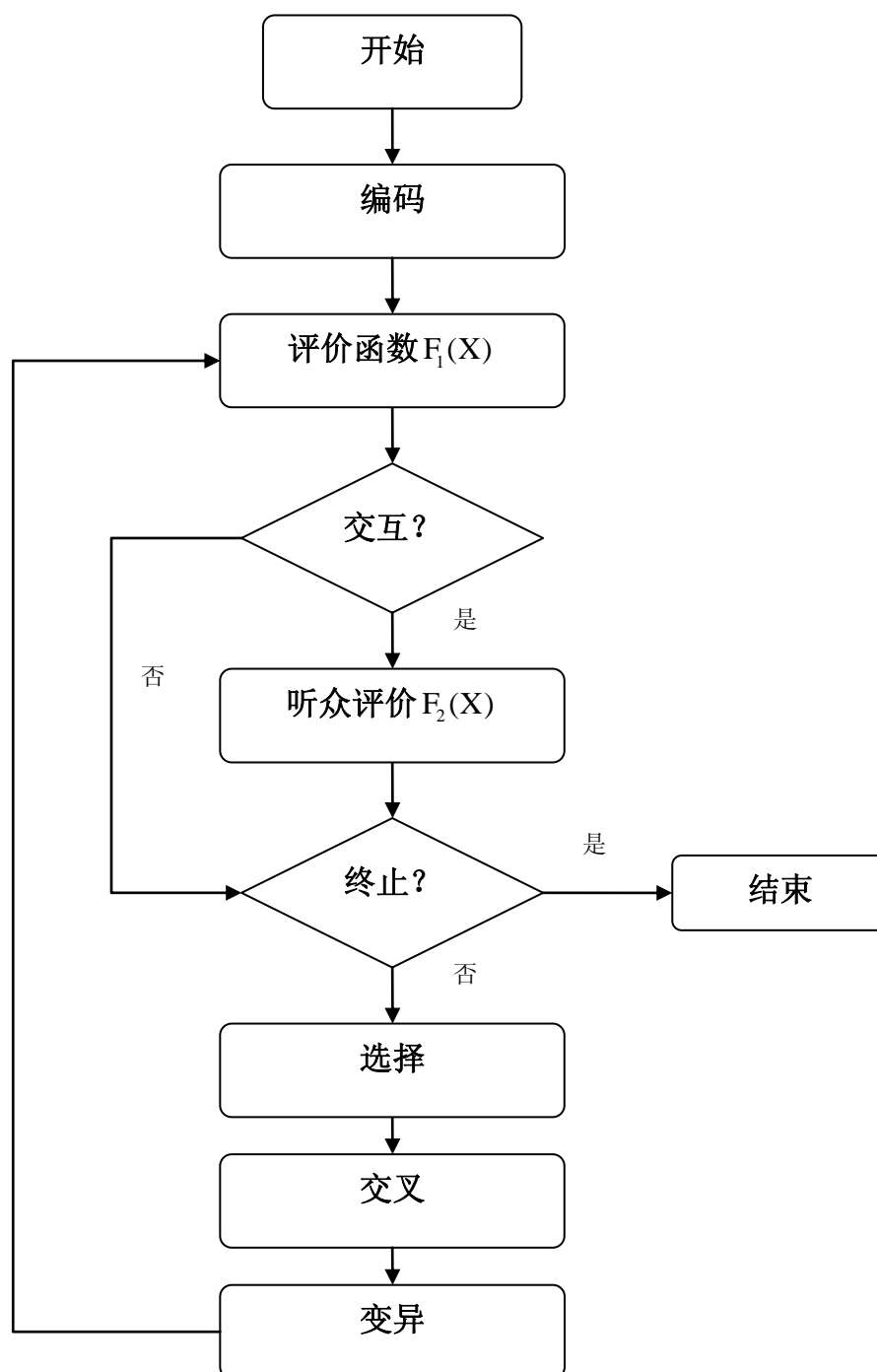
基金项目: 教育部博士点基金(20100072110038), 国家自然科学基金项目(70871091, 61075064, 61034004, 61005090), 教育部新世纪人才计划项目(NECT-10-0633)上海市信息化专项基金项目。

作者简介: 黄澄宇(1988-), 男, 硕士研究生, 主要研究方向: 智能控制

通信联系人: 汪 镭(1970-) 男, 同济大学电子与信息工程学院, 教授, 博士生导师, 研究方向: 智能控制, 智能计算, CIMS 和系统工程方面的研究, 上海, 200092. E-mail: wanglei_tj@126.com

应适度函数的同时，同时辅助人的评价来创作出好的音乐作品，从而提高了系统的性能。

1 利用遗传算法创作乐曲研究



图一 遗传算法流程图

Fig. 1 Flow cart of Genetic Algorithm

遗传算法是一种建立在自然选择和遗传变异等生物进化特征基础之上的全局最优的算法，具有良好的普适性和可规划性，在形式上简单明了，与其他算法综合使用很方便。一段优秀乐曲的产生主要依靠音乐家的“天赋”和“努力”，针对前者，其中的奥妙是难以完全理解的；而对于后者，遗传算法就可以深刻模拟音乐家“努力”创作音乐的过程。如遗传操作中的选择操作就可以模拟音乐家根据自己的知识和经验选择用于后续创作的种子；交叉操

作可以模拟音乐家综合不同乐曲的长处，来产出新的更好的乐曲；突变操作则可以模拟音乐家在创作乐曲过程中灵感^[3]。

1.1 编码

1.1.1 音级的编码

音级就是乐音体系中的各音。音级有基本音级和变化音级两种。乐音体系中，七个具有独立名称的音级叫做基本音级。基本音级的名称是用字母和唱名两种方式来标记的。两个相邻的具有同样名称的音叫做八度。升高或降低基本音级而得来的音，叫做变化音级^[4]。

我们将低音 1 编码为 11，中音 1 编码为 12，高音 1 编码为 13，随着音级从低到高依次增加 10。为了简单起见，本文中通过遗传算法产生的乐曲曲调都为 C 调。下表表示部分音级的编码，偏移半音数是指偏移中音 1 的半音个数。

表 1 部分音级编码表

Tab.1 Parts of Scale degree encoding

简谱音级	1	2	3	4	5	6	7
音级编码	11	22	32	42	52	62	72
偏移半音	0	2	4	5	7	9	11

1.1.2 时值的编码

音符的时值有十六分之一，八分之一，八分之一等多种类型。我们将以十六分音符为标准，记为 0，其他音符的编码是自身音符时值与十六分音符时值的比值加 1^[5]。如下表所示

表 2 时值编码表

Tab.2 Duration encoding

拍数	0.125	0.25	0.375	0.5	0.75	1	1.5	2	3	4
编码	0	1	2	3	4	5	6	7	8	9

1.1.3 音符的编码

采用三位十进制编码方式，前两位表示音级，第三位表示时值。

例如下面《得民心者得天下》中一小节的简谱：

4 3 3 6

则由前面的论述可知，编码为：423,320,320,615

2 适应度函数

智能音乐研究遇到的一个大问题就是生成作品的质量评估的问题。而使用遗传算法进行创作的时候，这个问题变的更加突出。因为每次迭代之后的种群都将面临选择或淘汰的问题。这个过程，适应度函数发挥了重大的作用^[4]。

2.1 概述

适应度函数的定义分为两部分，一部分是基于音乐原理和规则的定义，另一部分是利用听众的评价作为适应度函数，对创作出的乐曲进行评价。而且，随着迭代次数的增加，听众喜爱程度将占据更为重要的比重。即：

$$F(X) = F_1(X) + \sqrt{n} \cdot F_2(X)$$

其中 $F(X)$ 是总体的适应度函数， $F_1(X)$ 为基于音乐原理和规则定义的适应度函数，

$F_2(X)$ 为听众对音乐的评价, n 为迭代次数, X 为个体即染色体。

85 下面将详细介绍这两个适应度函数的构成。

2.2 基于乐理知识的评价。

基于乐理知识的评价标准, 我们将分为十个客观方面 ($i=1\sim 10$)。即:

$$F_1(x) = \frac{1}{10} \sum_{i=1}^{10} g_i(x)$$

1. 两个相邻的音符之前的音程差如果大于八度音阶的话将被记为 0, 反之则记为 1,

90 $g_1(x)$ 则是每个个体的平均值。即: 若一个个体有 a 组相邻音符的音程差大于八度音阶, 剩余的 b 组小于等于八度音阶。则:

$$g_1(x) = \frac{a*0+b*1}{a+b}$$

2. 如果音高是“F”或者“B”的话, 那么 $g_2(x)$ 将被赋予一个很小的值。因为大多数音乐更经常采用像“C,D,E,G,A”的音调。这个函数表明对某种音乐风格的喜好程度, 根据系

95 统的输入参数, 系统将就相应的减少那些不受人喜爱的音调的出现概率。

3. 时值是指时间的长短或者一个特定长度的音程。由于两个连续音符剧烈的时值变化将会给听众烦躁不安, 因此, 如果出现超过 4 的时值变化, 我们就将此记为 0, 反之则为 1。如同 $g_1(x)$ 一样, $g_3(x)$ 也表示一个个体的平均值。所以说这个参数是为了惩罚那些连续音符之间有剧烈变化的音符。

100 4. 如果一个个体中有切分音, 则 $g_4(x)$ 为 0, 反之则为 1。这个参数的目的不是禁止切分音的出现, 仅仅是为了降低其出现的概率。

5. 对于音乐来说, 音调一直上升或者一直下降都不会是好的音乐。所以, 我们用 $g_5(x)$ 来评价一个乐曲的整体轮廓。定义如下:

$$g_5(x) = \begin{cases} 1 & \frac{1}{n-1} \left| \sum_{i=1}^{n-1} \text{interval}_i \right| \leq 1 \\ 0 & \text{else} \end{cases}$$

105 6. 如果一个乐曲持续在一个音高或时值上, 听众将会很快烦躁不安的。因此, 我们用 $g_6(x)$ 来阻止这种情况的发生。如果一个乐曲中, 超过四个音符在同一音高和时值上连续出现, 那么这个参数将被记为 0, 反之, 则为 1。

7. 一段乐曲不能够变化的太剧烈, 用 $g_7(x)$ 评估这种情况。定义如下:

$$g_7(x) = \begin{cases} 1 & \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (|\text{interval}_i| - |\overline{\text{interval}}|)^2} \leq 1 \\ 0 & \text{else} \end{cases}$$

110 8. 众所周知, 一般情况下, 合唱的平均音高要比独唱要高。用 $g_8(x)$ 评估这种情况,

$$g_8(x) = \begin{cases} 1 & \frac{1}{n} \sum \text{pitch}_{ch} > \frac{1}{m} \sum \text{pitch}_v \\ 0 & \text{else} \end{cases}$$

9.另外，对于合唱而言，她的音高的升降变化要相对快一些。

$$g_9(x) = \begin{cases} 1 & \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (|\text{interval}_{ch,i}| - |\text{interval}_{ch}|)^2} > \sqrt{\frac{1}{m-1} \sum_{i=1}^{m-1} (|\text{interval}_{v,i}| - |\text{interval}_v|)^2} \\ 0 & \text{else} \end{cases}$$

115 10.通常情况下，一段乐曲将会以一个长音来结尾。因此，我们用 $g_{10}(x)$ 来描述乐曲是否以超过一拍的长音来结尾，若是的话则为 1，反之为 0。

2.3 交互式系统

单纯的乐理知识和相关规则的适应度函数完全满足听众娱乐的需求，在由系统产生适应度函数之后，引入听众的因素，成为一个交互式的系统，由每一个听众对乐曲进行打分。

120 系统会产生大量的乐曲，听众的时间总是有限的，不可能鉴赏所有的音乐，应用相关乐理知识对大量乐曲进行初选也是必须和必要的。听众的打分是在 0.1~1 分的区间。另外，随着迭代的进行，听众打分所占的权重将会越来越高，从而更加有利于该方法的收敛速度。

2.4 遗传操作

1. 复制对象的选择。

125 根据上述适应度函数的公式，计算每代个体的适应度。为了减小系统压力，根据适应度函数，选择其中的一些个体来进行下一代的复制和变异。选择的策略是：

若第 p 代 n 个乐曲的第 i 个个体 x_i 的适应度函数为 $F(x_i)$ ，若 $F(x_i) > \frac{1}{n} \sum_{i=1}^n F(x_i)$ ，则

第 i 个个体将进行后续的交叉和突变的操作，反之，该个体被舍弃。

2. 交叉规则。

130 交叉操作可以分为音符之间的交叉和小节之间的交叉，先进行音符之间的交叉，再进行小节之间的交叉。

音符之间交叉：编码是由三位数字组成，包含了音级和时值信息。因此，交叉操作是会发生音级或者是时值的交叉变换。而交换的方式有三种，分别是：音级的交叉，时值的交叉和二者同时交叉。对于音级的交叉，只会发生在前两位。对于时值的交叉，只会发生在最后一位。而如果二者同时发生，则整体交叉。对于交叉的概率，一般是取 0.25~0.75，本文对于不同的交叉方式，选取不同的交叉率。其中，音级的交叉率是 0.6，时值和二者同时交叉的交叉率是 0.5。

小节之间的交叉：设定小节的交叉率是 0.6。

3. 变异规则。

140 一般情况下，变异发生的概率较小，因此，本文选择的突变率是 0.02。

对于变异来说，都发生在音符上，而哪个音符突变，是随机发生的。对于变异，可以进行音级变异，可以进行时值变异，也可以将二者同时变异。在此不考虑相邻两音符的音程一般不超过 8 度的约束规则，这个规则在适应度函数中 $g_1(x)$ 进行定义，可以防止发生这种情况的个体进入到下一代的遗传操作中。

145 4. 时值的修正。

在遗传操作的过程中，发生了音符的交叉和变异的情况。此时，有些小节内时值已经不

符合拍号的要求。一般情况下，一段乐曲的拍号都是一定的，这就要求每个小节中各音符的时值之和是固定的。但是在交叉和变异的过程中，都伴随着音符时值的改变。这样就可能出现小节内时值和不符合要求的情况，因此，就要对这种情况进行修正。相应方案为：

- 150 (1) 计算各小节内音符时值之和，即编码中第三位数字。
- (2) 若每小节规定的总拍数为 PH ，而新一代的个体总拍数是 PH' 。比较二者大小。若相等，则不需要修正。
- a. 若 $PH > PH'$ ，则需要增加时值，增加的音符为小节中最后一个音符，增加的时值为 $PH - PH'$ 。
- 155 b. 若 $PH < PH'$ ，则需要减少时值。设 a 为需要减少的时值，则 $a = PH' - PH$ 。若最后一个音符的时值 $b > a$ ，则将该音符的时值变为 $b - a$ ；反之，则将该音符删除，再比较上一个音符，直到该小节满足时值要求。

3 实验

3.1 确定相关参数：

- 160 调号：C
- 拍号：4/4
- 小节数：3

3.2 进行遗传算法进行作曲。

系统根据输入的参数随机生成初始种群。

- 165 相关遗传操作按照前章论述的方法和相应概率进行。

首先有系统根据乐理知识和规则确定的适应度函数进行评价。根据研究发现，当采用这种方法进行了 1000~2000 代之后，系统的平均适应度函数趋于稳定，再进行迭代下去对乐曲的优化意义不大，并且此时的子代数量也相对大大减少。因此，引入人的鉴赏作为评价因素，对乐曲进行评价，经过十次迭代之后，产生乐曲片段。



图 2 乐曲片段
Fig.2 Piece of music

175 4 总结

本文采用遗传算法进行算法作曲研究。

- 首先根据音乐的基本特点，采用简洁的编码方式，准确的表述了音级和时值两个音乐特征，为后续适应度函数的计算以及相关基因操作提供了极大的方便。提出了基于基本乐理知识的适应度函数，为后续进行人机交互评价打下了坚实的基础并且大大降低了人的工作量，
- 180 达到了良好的效果。将遗传算法成功的应用到算法作曲当中，虽然还远不能作曲家创作的乐

曲相提并论。但是把该方法应用到小片段当中已经取得了良好的效果。

[参考文献] (References)

- [1] Gartland-Jones A, Copley P. The suitability of genetic algorithms for musical composition[J]. Contemporary Music Review, 2003, 22(3): 43-55.
- 185 [2] Papadopoulos G, Wiggins G. A genetic algorithm for the generation of jazz melodies[J]. Proceedings of STEP, 1998, 98.
- [3] Todd P M, Werner G M. Frankensteinian methods for evolutionary music[J]. Musical networks: parallel distributed perception and performance, 1999: 313-340.
- 190 [4] Mehrabian A. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament[J]. Current Psychology, 1996, 14(4): 261-292.
- [5] Maeda Y, Miyashita S. Chaotic Interactive Sound Generation System Using Interactive Genetic Algorithm[J]. Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, 2009, 21(5): 768-781.