

# 一种基于隐马尔可夫模型的第一类对位生成方法

陈魁, 冯寅

(厦门大学 智能科学系 人工智能研究所 福建 厦门 361005)

**摘要:** 本文简单介绍了目前利用计算机自动生成对位的几种方法。在此基础上提出一种基于隐马尔可夫模型并利用改进的 Viterbi 算法生成第一类对位的方法。在进行相关的实验后表明, 计算机利用这种方法生成的对位结果是有一定的音乐价值的。

**关键词:** 对位法; 第一类对位; 隐马尔可夫模型; 改进的 Viterbi 算法

## 0. 引言

对位法是复调音乐的写作技法, 18 世纪的奥地利音乐家福克斯(J.C.Fux)总结出了“分类法”(Species)。分类法的五种模式: (1)全音对全音(一对一); (2)一个全音对两个二分音符(一对二); (3)一个全音对四个四分音符(一对四); (4)一个全音对两个切分二分音符(一对切分); (5)一个全音对混合节奏(一对混合)。

### 1. 相关研究概况

算法作曲是人工智能学科的一个富有挑战性的研究领域。而对位法作为复调音乐的写作手法, 很早就是算法作曲中一个研究的方向。

基于规则的知识库。早期的自动对位生成系统都是采用的基于约束-满足度规则的技术。比如 Hiller and Isaacson 的 ILIAC, Lewin 的基于独创“Global Rule”规则的系统。以及 Gjerdingen 的遵循一位音乐家观点设计的系统 PRAENESTE。这些早期的成果都是基于遵守给定的对位规则来为给定乐段进行对位的。

进化算法(EA)。进化算法实际上是一种搜索算法。它在所有可能的解中搜寻“最优”的个体。

进化算法中较早被用于自动对位的是遗传算法(GA)。解决对位问题的 GA 以一定数量的符合固定旋律(C.F)的种类, 节奏要求的随机音符群为初始种群, 给定一个适应度函数, 一代一代的进行交叉, 变异, 选择, 最后得到近似最优的结果。然而, 以 GA 来解决对位最困难的是适应度函数的选择, 除了满足基本的旋律规则以及和声规则以外, 如何评价两个都满足规则的乐段哪个更有艺术价值就显得有些困难。

目前较流行的是建立隐马尔可夫模型, 设定状态值, 观察值, 输入一系列的样本数据, 训练参数, 最后根据 Viterbi 算法得到乐段的“最可能序列”。但是隐含状态和观察值的选择不同, 各个隐马尔可夫模型各不相同。隐马尔可夫模型考虑到了前后乐段的相关性, 可以使得乐段之间的结合更加流畅。本文提出了一种以对位结果(Cpt)的音程差为隐含状态的隐马尔可夫模型。然后根据改进的 Viterbi 算法来解决第一类对位的方法。

## 2. 用于自动对位的隐马尔可夫模型的建立

### 2.1 音乐相关知识

音程: 在乐音音序中, 两音之间的音高距离称为音程。

音程的度数: 音程的度数说明了音与音之间音高距离的大小。两音之间包含了几个音(两音本身也算在内)就称几度。

和声音程: 两音同时发响的音程称为和声音程。和声音程的两个音一般纵向记写。

旋律音程: 两音先后发响的音程称为旋律音程。旋律音程一律横向记写, 从左到右读或唱。

这里, 我们以 C.F 来表示“固定旋律”Canto Firmus; 以 Cpt 来表示对位声部 Countpoint。实际中, 我们有时候看到 C.F 在下, Cpt 在上, 有时候看到 C.F 在上, Cpt 在下。然而, 不管他们的位置如何, 他们的遵守的规则基本相同, 但有稍稍的差别。这里, 我们专门研究 C.F 在下方, Cpt 在上方的情况。

### 2.2 用于自动对位的隐马尔可夫模型

这里我们建立一个一阶的 HMM 模型。我们认为 Cpt 乐段的相邻两个音符之间的音程差是与上一对相邻两个音符之间音程差是相关的。而 C.F 乐段的相邻两个音符之间的音程差与 Cpt 乐段的所对应的相邻两个音符之间的音程差有关。

#### 隐含状态的集合

我们把对位结果(Cpt)中相邻两个音符间的音程差看成隐含状态。这里, 我们不考虑音程的大小以及增减。由于在第一类对位中, 一些旋律音程的使用有一定的限制: 不允许使用的比如七度以及八度以外的大跳等。这样我们可以找出所有的可能的状态。禁止出现的状态就不用考虑。所有的状态有: 一度, 向上/向下二度, 向上/向下三度, 向上/向下四度, 向上/向下五度, 向上/向下六度, 向上/向下八度这 13 个。

#### 观察值集合

我们把要进行对位的乐曲片段(C.F)中的相邻两个音符的音程差看成是观察值。理论上来说 C.F 乐段应该遵循和 Cpt 乐段一样的旋律规则, 所以, C.F 乐段的旋律的音程差应该也是 13 个状态, 但是我们为了避免输入的样本中存在一些突破限制的乐段而导致错误, 我们在 13 个观察值基础上加上原来的向上/向下七度音程。实际上, 在观察值矩阵 B 中, 由于从各个状态下观察到的向上/向下七度的概率设置很小, 它并没有太大的影响, 这样一共存在 15 个观察状态。

隐含状态的转移矩阵; 这里我们可以选择随机初始化矩阵 A, 或者是根据基本的音乐知识有轻重的设置各个状态的转移概率, 而后者在数据样本不大, 未必能使得模型达到收敛条件的情况下。可以得到更优的个体。

观察值概率矩阵; 同样可以随机初始化矩阵 B, 或者根据经验值进行赋值。

初始概率分布; 由于采用的是音程差来做为隐含状态, 那么第一个状态应该进行特殊的定义: 我们设定第一个音符与它在调式的五线谱上加一线的位置之间的音程差为初始状态(当然也可以根据情况以别的位置作为参照)。

这样一个 HMM 模型就建立起来了, 但这里只考虑到了 Cpt 旋律线上的相关性。我们知道, 对于复调音乐来说, 不仅本身的旋律要流畅独立, 还必须和 C.F 乐段的对应位置的音符有一定的和声关系, 所以还要遵守一定的和声准则。这些规则我们会加入到选择最优路径的 Viterbi 算法中。

### 2.3 参数学习

由于样本数据的有限, 不可能让模型达到收敛状态。同时在音乐艺术上, 原本就没有最正确的准确答案, 只能说模型的参数会受到输入样本的风格的影响。隐马尔可夫模型参数学习的收敛只能表明目前模型的参数设置趋向于某一种音乐创作的风格。这里, 我们可以选择按人工经验来初始化参数, 然后输入样本, 训练参数, 用于接下来的 Viterbi 算法。

### 2.4 改进的 Viterbi 算法

由于 Cpt 乐段还要遵守一定的和声规则: 与 C.F 的和声音程中要以三, 六度为主。五, 八度少用, 而纯 4 度不能使用等等。我们对经典的 Viterbi 算法进行一些修改, 在 Viterbi 算法执行的

过程中。我们加入一个修正因子。这里可以理解为一个由下一个要到达的隐含状态，从开始到  $t-1$  时刻隐含状态的局部最优隐含状态序列，以及从开始到达  $t$  时刻的观察值序列，这 3 个参数为输入的惩罚函数。

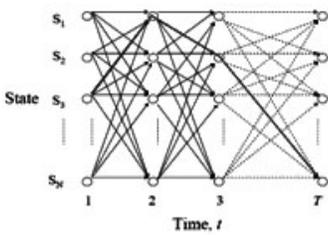


图 1 Viterbi 算法示意图

$\epsilon$  的意义在于:判断 Cpt 中从前一个时刻的各个状态达到目前这个状态的时候是否会导致与 C.F 在和声音程中出现不协和的情况。一旦出现这样的状态,就把从前一个状态出发到达本状态的概率设置为 0。这样做实际是在选择旋律线的过程中加入了和声规则的限制条件,虽然可能改变了原本的“最优”旋律线,但是同时也避免了出现和声不协调的情况。

$\epsilon$  函数的实现: 从上图的普通 Viterbi 算法的示意图我们可以看出,除了时刻以外,其他每个时刻的每个状态都会选择一条局部最优的路径到达自己,最后在通过最后一个状态的选择进行回溯,求得全局最优的序列。如果我们记录下  $t=1$  时刻各个音符到达的位置,进而向前推算,就可以计算出在每个时刻的每个状态的音符位置,再与 C.F 乐段的音符进行比较就能得到音程差。最后根据具体约束条件,得到修正值。除此之外,由于对 Cpt 首尾这两个音符还要有特别的要求,也可以给“更符合条件的音程差”赋予较高的修正,进而让它有更大机会入选。

3. 实验数据以及结果分析

程序使用的状态转移矩阵 A, 观察矩阵 B 以及初始状态, 我们根据基本的乐理知识进行赋值。我们从各种音乐教科书中选择 13 个 C.F 在下方的实例进行训练, 为教科书中的 C.F 在下方的乐段练习题来配对位。

实验结果输出的最可能状态序列为: 4 8 9 2 3 2 3 2 按照原来状态的含义可以得到结果:

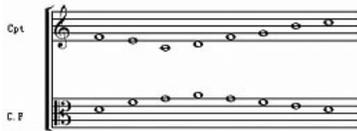


图 2 一个对位实例

对教科书上另一个实例,人工对位与机器对位的比较:

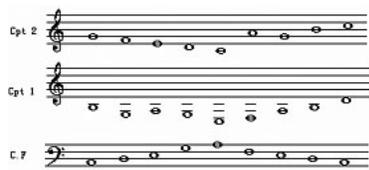


图 3: Cpt 1 为机器生成对位, Cpt 2 为教科书上人工对位

实验结果分析: 对几个不同的乐段进行实验后,可以看出,这种基于马尔可夫链的第一类对位的生成法的确可以生成具有

一定价值的对位结果。但是,其中还存在一些问题等待解决。

从参数的选择上看,在样本数量有限的情况下,初始化的参数对实验结果产生了较大的影响。本人由于对音乐专业知识不是十分熟悉,仅凭个人主观设置的初始参数也必定存在着缺陷,这也一定程度上影响了对位的结果。另一方面上看,如果选择随机初始化初始参数,每次产生的对位结果风格将会迥异。

从参数进化的方面看,样本的选择以及数量也有很大影响。由于样本的数量有限,难以使得参数训练达到一个“收敛”的状态。而从另一角度上看,音乐上的作品本没有对错,最优之分,这里的“收敛”指的是风格上的趋近某一种状态。

而要风格的趋进,最好是采用一个作曲家的作品作为样本。而实验中,实例都来自不同的音乐教科书,实例作者也各不相同,风格也必定不同。这个对实验结果趋向某种风格产生了一定的影响。

添加更多的隐含状态。本文的隐含状态中没有考虑到音程度的大小以及增减,如果把这些因素都考虑进去,这样生成的对位结果应该会更加符合乐理。

所以,最理想的状态是,由一个专业音乐家的经验判断来初始化 HMM 的各个参数,然后,再用这个音乐家的作品来训练参数,这样即使在样本数量不大的情况下,也会相对容易得到一个较理想的结果。

对于对位规则的遵守以及风格上看。对位法规则众多,包括了旋律音程,和声音程,旋律线走势,首尾音符选择,等等,不可能一一形式化且被采用。在人工对位的时候,作曲家们在大体上遵守基本准则的情况下,偶有突破,作品风格带有个人色彩,对位结果自然优秀;利用规则的约束,或是遗传算法产生的机器对位,能遵守事先设定的各个准则,作品风格带有随机色彩;利用隐马尔可夫模型生成的对位,应该来说在旋律线上相关性更好,但是由于对于规则的把握不够,实验中可能会产生一些违反规则的对位结果。在样本实例风格相同以及数量足够的情况下,对位风格趋向输入的样本。

4. 结束语

本文提出了一种利用隐马尔可夫模型生成第一类对位的方法。实验的最终结果未必是完美的,但是希望可以给人带来一丝的灵感和启发。算法作曲作为人工智能中一个富有挑战性的研究领域,还需要我们进行不懈的努力和探索。

参考文献:

1. Mary Farbood, Bernd Schoner. Analysis and synthesis of palestrina-style counterpoint using markov chains Proceedings of International Computer Music Conference. Havana, Cuba, 2001.
2. Soren Tjagvad Madsen. Evolving Palestrinian counterpoint with and EA.
3. 《Visual C++ 数字图象模式识别技术及工程实践》求是科技,张宏林编著,人民邮电出版社 2003
4. 《复调音乐简明教程》林华著 上海音乐学院出版社 2006
5. 《复调音乐基础教程》赵德义 刘永平著 1997
6. 《复调音乐教程》于苏贤著 上海音乐出版社 2001

(上接第 61 页)

致力于对新的加密技术的开发与研究。对称密钥和公开密钥等算法各有各的优缺点,在选择时可以根据自己的安全要求,来选择自己需要的加密算法,也可以将其结合起来使用,以达到更好的加密效果,更好的提高数据的安全性。

参考文献:

1. 信息安全技术基础 李俊宇 编著 冶金工业出版社 2004
2. 计算机网络安全的教程 石志国,薛为民,江俐编著 清华大学出版社 2004
3. 计算机网络安全技术 蔡立军主编 中国水利水电出版社 2002
4. 计算机网络安全技术及应用 邵波主编 电子工业出版社 2005
5. 计算机网络安全技术教程 谢冬青等主编 机械工业出版社 2007