

文章编号: 0427-7104(2017)02-0167-08

使用数学模型选择的方法对古典钢琴演奏中乐句节奏类型的决定因素进行建模

李圣辰

(北京邮电大学 信息光子学与光通信研究院, 北京 100876)

摘要: 在古典钢琴演奏中, 对乐句中节奏线条进行分类是一种常用做法。很多音乐学家在音乐学研究中对于乐句中节奏线条类型的选择因素做出了不同的假设。这些假设通常关注两大类因素: 乐句本身的属性和节奏变化的时序关系。使用乐句位置和前序乐句中节奏线条的类型作为对于乐句节奏线条类型决定因素的抽象成因,⁴ 一个贝叶斯图模型假设被用来进行对比测试, 从而揭示古典钢琴演奏中对于乐句的节奏线条类型的决定因素。根据相对熵、交叉熵和作者自行拟定的交叉熵比在内的 3 个模型表现衡量标准, 联合模型, 即考虑单一乐句节奏型的选取与前序乐句中节奏型和乐句位置均相关的候选贝叶斯图模型, 能够较为准确地折射出在古典钢琴演奏过程中, 为乐句选取节奏型的影响因素。

关键词: 节奏型; 模型选择; 钢琴演奏; 节奏型影响因素; 贝叶斯图模型

中图分类号: TP3-05

文献标志码: A

DOI:10.15943/j.cnki.fdxb-jns.2017.02.004

在古典钢琴演奏中, 乐句中节奏线条的变化可以被分为若干类^[1-2]。为了方便描述, 我们将每类乐句中节奏线条变化的平均值称为节奏型^[3]。在实际钢琴演奏当中, 在不同乐句中选用不同的节奏型是构成演奏过程中音乐表现力的一种方式。演奏者在演奏过程中对于某一乐句中节奏型的选择受很多不同因素的影响, 演奏者在演奏过程中对于节奏型的选择成为音乐学研究的课题之一。通过了解演奏者如何设计和选择节奏型, 可以更好地了解演奏中表现力的形成机制, 为进行音乐表现力客观评价和生成音乐表现力打下基础。

由于影响演奏中节奏变化的因素众多, 现特将演奏者在演奏过程中对于乐句节奏型的选择因素抽象为两种概括性因素: 乐句本身的属性和前序乐句中采用节奏型, 并使用不同的假设建立不同的贝叶斯图模型, 然后通过数学模型选择的方法来对两种因素对节奏型选取的影响进行测试。

本文选取的两种节奏型决定性因素是基于众多已有音乐学研究做出的。在已有的音乐学研究成果中, 有关于节奏型决定因素, 或就某一乐句中节奏变化的研究不在少数。Widmer 等^[4-5]就计算机如何合成有表现力的节奏变化进行了讨论。对于节奏表现力的合成, 一般使用基于经验数据的机器学习方法。计算机根据钢琴演奏者已有的录音, 将乐谱与相应的节奏变化相关联。当计算机要为一个未曾学习过的乐句合成有表现力的节奏变化时, 会找出与被合成乐句相似度最高的已经学习过的乐句, 然后利用已学得的经验为待合成乐句合成具有表现力的节奏变化。因此, 根据已有文献, 乐句本身的各种特性可以对节奏型的选择产生影响。

除去乐句本身的特性, 在文献[4]中, 作者还指出了有关时序特征对于生成乐句节奏表现力的潜在影响。实际上, 计算机音乐学的经典理论之一就是使用二次函数曲线来拟合音乐演奏中节奏的变化^[6]。这种基于数学拟合的分析方法正是基于节奏变化中的时序特性。在传统的基于规则的计算机音乐分析系统 KTH 中^[7], 乐句中节奏变化的影响因素也同时包含了乐句本身的属性(如调性)和演奏时节奏时序序列的特性。

由于在一首乐曲中, 同一乐曲的各个属性一定相同, 我们使用在乐曲中的乐句位置来代表乐句自身

收稿日期: 2017-01-06

作者简介: 李圣辰(1988—), 男, 博士, 讲师, E-mail: shengchen.li@bupt.edu.cn.

属性这一概念,并用被分析乐句的前序乐句所选择的节奏型来代表演奏时节奏的时序序列特性。尽管这两个高度抽象的概念不足以就乐句中节奏型的选取做出详尽的解读,但是通过对两种因素的对比和测试,我们仍然可以对这两种因素对乐句节奏型的影响关系进行简略地了解。

在本文中,我们使用贝叶斯图模型对两种抽象概念对乐句节奏型选取的影响做出不同的假设。贝叶斯图模型,或称概率图模型,是一种对随机变量进行分析,并对随机变量间的关系进行描述的方法^[8]。我们根据乐句的位置和乐句前序节奏型对于当前乐句节奏型的选取做出不同的假设,并根据相应假设构建了 4 个相应的贝叶斯图模型,即根据 1) 乐句节奏型选取与乐句位置和前序乐句节奏型无关;2) 乐句节奏型选取仅与乐句位置相关;3) 乐句节奏型选取仅与前序乐句节奏型相关;4) 乐句节奏型选取仅与乐句位置和前序乐句节奏型二者都相关等 4 种假设并构建相应的贝叶斯图模型。我们将测试各贝叶斯图模型,并根据它们的表现来评估 4 个假设的可行性。

为了测试 4 个候选贝叶斯图模型的表现,我们选用信息论中的相对熵和交叉熵作为衡量候选贝叶斯图模型(以下简称候选模型)的衡量标准。在信息论中,交叉熵和相对熵均被用于对两个数据的信息差异进行测量。我们将数据集按照一定比例分为两部分:训练集和测试集。我们利用候选模型求得在不同乐句位置节奏型序列的分布情况,并比较两个数据集中分布的差异。如果在某个特定的候选模型下,根据训练集得到的模型能够在验证集中获得较大模型似然概率,则说明该模型性能良好。除了经典的相对熵和交叉熵,我们还使用相对熵与测试集信息熵之比作为衡量候选贝叶斯图模型的衡量标准。

本文各个测试的数据库有两个来源:一个是公开的肖邦玛祖卡数据库,一个是文献[1]中所使用的伊斯拉美数据库。玛祖卡数据库由英国录音历史与分析中心(Centre for the History and Analysis of Recorded Music, CHARM)研究组公布,并被诸多相关研究^[2,9]所引用。该数据库含有 5 首肖邦玛祖卡的数据,但为定义乐句节奏型,本文参照文献[1]只选取了其中两首(Op. 24/2 和 Op. 30/2)。数据库中均包含 3 首乐曲的多个演奏版本,其中包括 25 个伊斯拉美版本,64 个玛祖卡 Op. 24/2 版本和 34 个玛祖卡 Op. 30/2 版本。

1 乐句节奏型的定义

在本部分,我们将就乐句节奏型的定义过程给出介绍。乐句节奏型的详细定义和讨论可以参见文献[1]。在本文所使用的数据库中,有关节奏变化的信息以演奏过程中每个拍节发生时刻的方式被记录下来。在本文中,我们取文献[10]中关于节奏的定义,即在单位时间内,音乐在乐谱上的单位所演奏的速率。根据该定义,如果一首乐曲共有 $t+1$ 拍,且每拍发生的时刻记为 $\{t_1, t_2, \dots, t_{n+1}\}$,其中 t_{n+1} 为乐曲结束时刻。在 t_i 时刻,节奏 τ_i 可被表示为 $\tau_i = \frac{1}{t_{i+1} - t_i}$ 。因此,我们可以得到一首曲子在各个拍节上的节奏变化为 $\{\tau_1, \tau_2, \dots, \tau_n\}$ 。由于在本文选取的数据库中,每首乐曲的乐句在曲目中均为等长,如在乐曲中每个乐句有 w 拍,则我们可以用向量 $\mathbf{T}_i = \{\tau_1, \tau_{i+1}, \dots, \tau_{i+w}\}$ 来表示一个乐句中的节奏变化。使用期望最大化(expectation maximization)方法,我们得到数据库中各乐句内部的节奏变化 \mathbf{T}_i 应当符合具有 A 个成分的高斯混合分布,即:

$$p(\mathbf{T}_i) = \sum_{a=1}^A \pi_a N(\tau_i | \mu_a, \Sigma_a). \quad (1)$$

我们认为每一个高斯成分对应一个节奏型的定义,且每个节奏型的形状可以由相应高斯成分的平均数(μ_a)表示。对于一个乐句 i ,其所属节奏型可以由乐句内节奏变化 \mathbf{T}_i 在各个高斯成分中出现的概率所决定,即乐句 i 所属的节奏型 \mathbf{T}_i^* 为:

$$\mathbf{T}_i^* = \arg \max_a \pi_a N(\tau_i | \mu_a, \Sigma_a). \quad (2)$$

根据文献[1]中的方法,使用高斯混合模型对于乐句内节奏的分布进行建模,并取在交叉验证实验中验证集模型似然概率最大的高斯混合模型来定义节奏型,即为肖邦玛祖卡 Op. 24/2 定义 8 个节奏型,为肖邦玛祖卡 Op. 30/2 定义 4 个节奏型,为伊斯拉美定义 2 个节奏型。如果我们赋予每个节奏型以相应的

颜色，使用色块代表数据库中每次演奏中的每个乐句所采用的节奏型，我们可以得到节奏变换图^[11]。

下面我们以肖邦玛祖卡 Op. 24/2 为例，展示一幅节奏变换图（图 1）。在节奏变换图中，每一行代表一位演奏家的演奏，每一列代表一个乐句，每种色块代表一种节奏型。两个节奏型的节奏变化线条越相近，则其表示颜色亦越接近。通过观察图 1，我们发现在本文中提出的乐句位置和节奏型时序序列的两种潜在乐句节奏型决定因素有其合理性。观察每一列，不同乐句的节奏型采用分布有较大差异。如果我们将每一列的节奏型分布进行统计，可以得到不同乐句中节奏型采用情况的分布。在图 2 中，我们以肖邦玛祖卡 Op. 24/2 为例，展示不同乐句的节奏型采用情况。在图 2 中，每一列代表一个乐句。此外，被采用的节奏型比例按照节奏型的相似程度相排列，其中节奏型的配色方案与图 1 中节奏型的配色方案相同。在图 3（见第 170 页）中，我们统计了当某个特定节奏型出现后在下一乐句各个节奏型的分布情况。通过观察图 3，我们可以发现采用不同节奏型的乐句，后续乐句的节奏型采用分布各不相同。比如，在采用节奏型 8 的乐句，后续乐句的节奏型一般为 8；而采用节奏型 5 的乐句，后续乐句选用节奏型 2 的情况较多。

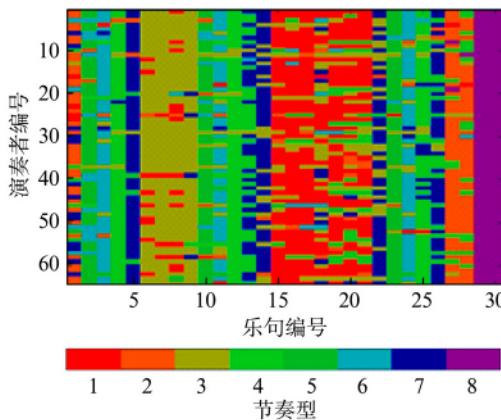


图 1 节奏变换图(肖邦玛祖卡 Op. 24/2)

Fig. 1 Tempo variegation map
(Chopin Mazurka Op. 24/2)

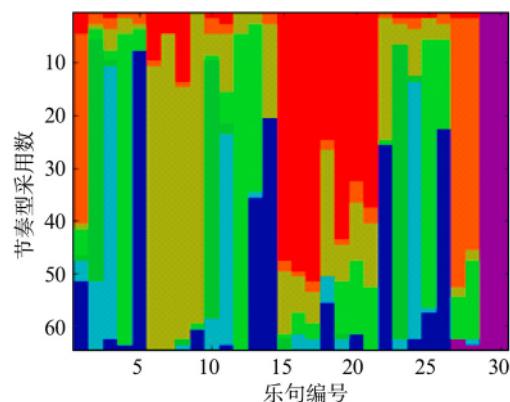


图 2 每个乐句使用的节奏型统计(肖邦玛祖卡 Op. 24/2)

Fig. 2 The expressive timing cluster used for each phrase in Chopin Mazurka Op. 24/2

至此，我们完成了定义节奏型的工作，亦通过观察节奏型的分布再次验证了本文所拟假设的合理性。下面，我们将对各个假设所对应的贝叶斯图模型进行说明。

2 候选贝叶斯图模型

根据对于节奏变换图的观察和对已有研究成果的参考，我们拟定了 4 个候选贝叶斯模型：1) 独立模型 (Independent Model, IM)：乐句节奏型的选取与乐句位置和前序乐句节奏型无关；2) 位置模型 (Positional Model, PM)：乐句节奏型的选取仅与乐句位置相关；3) 时序模型 (Sequential Model, SM)：乐句节奏型的选取仅与前序乐句节奏型相关；4) 联合模型 (Joint Model, JM)：乐句节奏型的选取与乐句位置和前序乐句节奏型均相关。4 个候选模型所对应的贝叶斯图模型分别表示为图 4(a), 图 4(b), 图 4(c) 和图 4(d)（见第 170 页）。在图 4 中，每个椭圆形中的变量被看作是一个随机变量，箭头表示随机变量间的依赖关系，即如果 $A \rightarrow B$ ，则 B 事件的概率分布受 A 事件的概率分布影响。对于被分析的乐句，其节奏型被记为节奏型 2；被分析乐句前序乐句所采用的节奏型被记为节奏型 1。

按照贝叶斯图模型的定义，以上 4 个候选贝叶斯图模型均可以表达为乐句位置和乐句及其前序乐句所采用节奏型的联合概率分布，即 $p(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*)$ 。由于候选贝叶斯图模型中所假设的随机变量间的关系不同，所以我们可以用不同的方法求得 $p(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*)$ ，即根据独立模型、位置模型和时序模型，联合概率分布可以写为：

$$p_{\text{IM}}(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*) = p(\text{position}) \times p(\mathbf{T}_{i-1}^*) \times p(\mathbf{T}_i^*), \quad (3)$$

$$p_{\text{PM}}(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*) = p(\text{position}, \mathbf{T}_i^*) \times p(\mathbf{T}_{i-1}^*), \quad (4)$$

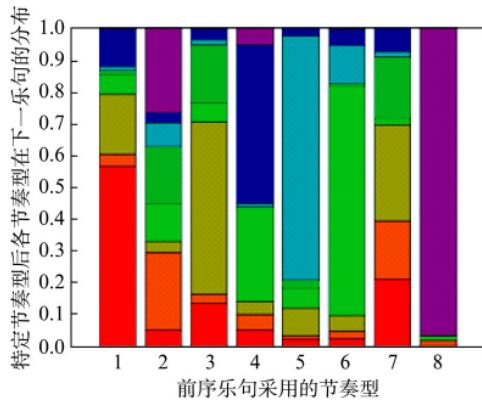


图 3 每种节奏型被选用后后续乐句采用乐句型的分布

Fig. 3 The distribution of expressive timing cluster after a certain expressive timing cluster is engaged in a phrase

$$p_{SM}(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*) = p(\text{position}) \times p(\mathbf{T}_{i-1}^*, \mathbf{T}_i^*). \quad (5)$$

由于联合模型中 3 个随机变量均有关联,故联合分布 $p(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*)$ 不能展开,只能通过计数方法求得. 特别需要强调的是,在有的乐句中,有些特定的节奏型序列未曾出现,导致联合概率分布上出现零点. 这些零点在验证模型表现时会导致“零概率”问题的出现,即概率为零的事件熵为无限大. 为了避免“零概率问题”,我们采用贝叶斯概率估计的方法^[8],为这些零概率点增加一个极小的概率值. 例如,如果在 X 个样本中有 x_1 个样本的值为 1,则 $X=1$ 的概率 $p(X=1)$ 被估计为

$$p(X=1) = \frac{x_1 + \frac{1}{X}}{X+1}. \quad (6)$$

3 模型评估方法

在这一部分,我们将对候选贝叶斯图模型的评估方法作出说明并着重阐述信息论中熵等的定义同交叉验证之间的关系. 传统的数学模型评估方法如赤池信息量准则(Akaike Information Criterion, AIC)、贝叶斯信息量准则(Bayesian Information Criterion, BIC)等模型选择参数^[11]要求候选模型应当处于同一模型空间内,因此扩展性受到制约. 本文所述的 4 个候选贝叶斯图模型均对 $p(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*)$ 作出了描述,因此我们采用几个有关信息论中熵的定义并利用其与交叉验证中的等效性来对候选贝叶斯图模型进行评估.

在交叉验证中,我们可以利用测试集模型似然来衡量候选模型的表现,即将包括 N 个数据点的数据集分为测试集和训练集,如训练集 Q 的联合概率分布记为 $q(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*)$,测试集 P 的联合概率分布记为 $p(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*)$,那么对数模型似然(logL)的平均数,即根据训练集训练产生的模型 M 观测到测试集 P 概率的对数可以写为:

$$\begin{aligned} \text{logL} &= \frac{1}{N} \sum N \cdot p(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*) \cdot \log q(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*) = \\ &\quad \sum p(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*) \cdot \log q(\text{position}, \mathbf{T}_{i-1}^*, \mathbf{T}_i^*). \end{aligned} \quad (7)$$

式(7)中,对数所选取的底数不影响本文实验中的各结论. 为了保持相关变量的物理意义,本文中各数据结果取底数为 2.

在信息论中,交叉熵(H_c)的定义为:

$$H_c = - \sum p \log q. \quad (8)$$

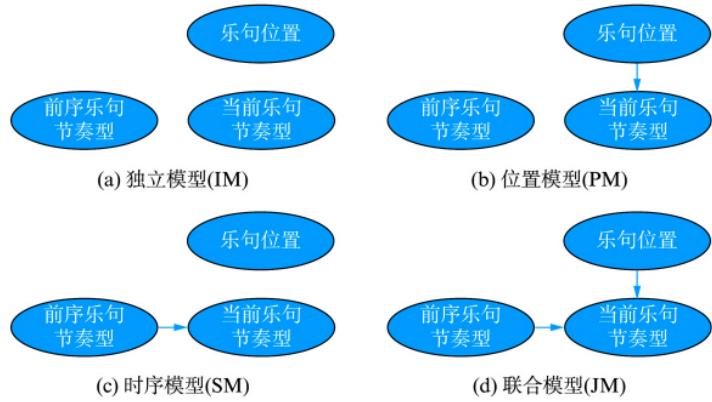


图 4 候选贝叶斯图模型中变量关系的假设

Fig. 4 The assumptions between variables in the candidate Bayesian graphical models

由此可以看出交叉熵和交叉验证中的模型似然具有等效关系,故我们可以利用 P 和 Q 的交叉熵来衡量候选模型的表现。由于在不同的测试中,测试集 P 的复杂度并不一致,而 P 的复杂度在信息论中通过求取 P 的信息熵来进行衡量(H_P),因此可以使用交叉熵与测试集 Q 的信息熵之差作为衡量模型表现的一种方式,而这种方式在数学上与信息论中相对熵(H_r)的定义等效,即:

$$H_r = -\sum p \log \frac{q}{p} = -\sum (p \log q - p \log p) = H_c - H_P. \quad (9)$$

相对熵可以通过求取交叉熵与测试集 Q 的信息熵来避免由于测试集 Q 的复杂程度差别导致的模型评估误差。同理,我们亦可使用交叉熵与测试集 Q 的信息熵之比来衡量候选模型的表现,以避免由于测试集 Q 的复杂程度所带来的模型评估误差,即交叉熵比(Δ):

$$\Delta = \frac{H_c}{H_P}. \quad (10)$$

我们将此评估参数称为交叉熵比,以便于下文中的表示。对于我们所选择的交叉熵、相对熵和交叉熵比等 3 个模型选择参数,我们将在后文中就其各自特点进行讨论。

4 模型评估结果

本部分将分析使用交叉熵、相对熵和交叉熵比对候选模型进行测试的结果。在展示具体结果前,我们首先就本部分实验的设计进行说明。由于本实验所采用的模型选择参数(交叉熵、相对熵和交叉熵比)均与交叉验证中的模型似然在数学上等效,因此,我们在测试集和训练集的划分过程中采用交叉验证中常用的五折分析法。我们首先将数据划分为 5 部分,每次选取一部分作为测试集,剩下的 4 部分作为训练集,训练得到模型,并求出根据该模型观察到验证集数据出现的似然概率。然后重复该过程 4 次,每次选用不同的部分作为测试集。最后,将 5 次得到的验证集似然概率进行平均,以得到一次五折实验的结果。

为避免五折实验过程中划分五折带来的随机性影响,每首乐曲均进行了 100 次五折实验,这 100 次五折实验的验证集平均模型似然概率如表 1 所示。

表 1 模型评估结果
Tab. 1 Model evaluation results

曲目	模型	H_c	H_r	Δ
伊斯拉美	IM	7.17	0.97	0.16
	PM	7.64	1.45	0.23
	SM	7.06	0.86	0.14
	JM	6.80	0.61	0.10
肖邦玛祖卡 Op. 24/2	IM	10.63	4.22	0.66
	PM	13.31	6.90	1.08
	SM	9.69	3.24	0.51
	JM	7.74	1.36	0.20
肖邦玛祖卡 Op. 30/2	IM	5.80	1.69	0.41
	PM	6.60	2.49	0.60
	SM	5.60	1.49	0.36
	JM	4.92	0.81	0.19

注: 加黑的数字为同一曲目表现最好的模型。

在表 1 中,我们展示了 4 个候选模型在我们选择的 3 首乐曲中的表现。需要注意的是,表中采用的相

对熵、交叉熵和交叉熵比均使用负逻辑,故数值越小表示模型表现越好。根据结果,在 3 首被选取的乐曲中,我们选取的 3 个模型选择参数均给出了相同的评估结果,即:联合模型优于时序模型,时序模型优于独立模型,独立模型优于位置模型。

在评估模型对数据的契合程度时,除了模型能够达到的最优效果,数据鲁棒性也是重要的参考指标之一。所谓数据鲁棒性,是指当用于训练模型的数据非常有限时,所得到模型的表现。为了测试鲁棒性,我们更改数据库中训练集和测试集的比例,并观察各个参考模型的表现。在表 2 中,我们展示了在鲁棒性测试实验中,各个候选模型表现最差与最好的交叉熵比之比,故较小的数值表示较好的数据鲁棒性。

表 2 候选贝叶斯图模型的数据鲁棒性
Tab. 2 The data-size robustness of the candidate Bayesian graphical models

曲目	最少可用训练 数据点个数	数据鲁棒性			
		IM	PM	SM	JM
伊斯拉美	39	1.20	1.91	1.21	4.86
肖邦玛祖卡 Op. 24/2	29	1.09	1.02	1.37	3.33
肖邦玛祖卡 Op. 30/2	7	1.20	1.09	1.48	3.31

从表 2 中我们可以看到,各个候选模型的数据鲁棒性与候选模型的复杂度成反比,即越复杂的模型(参数越多的模型),其数据鲁棒性越好。在候选模型中,鲁棒性最好的模型是独立模型,其他候选模型的排位是位置模型,时序模型和联合模型。应当指出的是,候选模型仅有优秀的数据鲁棒性是不够的,表 2 中展示的数据鲁棒性结果仍然需要参考表 1 中候选模型对于实际数据分布的契合程度。故总体上来说,独立模型和位置模型的高数据鲁棒性仍然不足以弥补其在模型评估结果中的糟糕表现。

需要指出的是,对于伊斯拉美这首乐曲,由于其单个演奏所含的可用训练数据点较多,时序模型的数据鲁棒性要强于位置模型的鲁棒性。综合候选模型的表现,可以发现时序模型的鲁棒性虽然没有独立模型和位置模型强,但是根据其优良的模型表现,我们认为时序模型在数据极其有限时可以被用来作为乐句节奏型预测的贝叶斯图模型以替代复杂度较高的联合模型。

根据以上结果,我们可以看出,前序乐句节奏型和乐句位置对于单一乐句节奏型的选择均具有重要的影响。对于仅仅考虑前序乐句节奏型的时序模型来说,其对于乐句节奏型选取的预测结果仍然较为准确。而乐句位置的信息仅在与节奏型时序信息综合的基础上,可以取得更加准确的效果。需要注意的是,如果仅仅考虑乐句位置,相关候选模型的表现甚至低于基于普通统计规律的独立模型。因此,对于仅仅利用乐句位置信息来探究乐句节奏型选取的计算音乐学方法来说,其结果的可靠性可能需要被重新考虑。

我们在实验过程中发现不同的模型选择参数对数据集随机性的鲁棒性也各不相同。理论上,随着训练集所占数据库的比例由小到大,模型选择参数应当呈现一条 U 形曲线,即:在模型训练集所占数据库比例较小时,模型表现随训练集比例增加而提高;在模型训练集所占数据库比例超过一定数值时,由于测试集较小,不能代表实际数据的分布,导致模型表现随训练集比例增高而降低。实际实验中进行了 100 次五折实验,然而 100 次五折数据划分只占所有五折划分可能的一小部分,故测试集组成的随机特性仍然不会完全消除,故实验结果不会呈现理想的 U 形曲线。

我们将选取的模型选择参数随训练集所占数据库比例的变化情况画成一条曲线,根据曲线一阶变化的过零率来判断模型选择参数随训练集所占数据库比例变化的单调性。理论上,模型选择参数随数据库比例变化的曲线仅有一次单调性改变,故选择参数随训练集所占数据库比例的一阶变化过零率应当很小。在本文中选取的 3 个模型选择参数随训练集所占数据库比例的一阶变化过零率如表 3 所示。在表 3 中,较高的过零率被以斜体显示,意味着此时该模型选择参数对数据的随机性最为敏感。根据表 3 中的数据,交叉熵对于数据随机性最为敏感,相对熵次之,相对熵对于数据随机性最不敏感。因此,在衡量两个模型的差异时,应当首先选用相对熵或交叉熵比,避免使用交叉熵。

表3 模型选择参数受数据集随机性影响的比较
Tab. 3 The impact of dataset randomness to model selection criteria

曲目	模型	H_c	H_r	Δ
伊斯兰	IM	0.59	0.14	0.14
	PM	0.32	0.14	0.14
	SM	0.68	0.23	0.14
	JM	0.14	0.14	0.23
肖邦玛祖卡 Op. 24/2	IM	0.66	0.54	0.54
	PM	0.70	0.61	0.64
	SM	0.44	0.34	0.38
	JM	0.41	0.41	0.44
肖邦玛祖卡 Op. 30/2	IM	0.48	0.29	0.29
	PM	0.48	0.29	0.35
	SM	0.55	0.48	0.35
	JM	0.42	0.42	0.42

注：数据为数据集比例与模型表现的曲线的一次差值的过零率，数据越小表示模型选择参数受数据随机性的影响越小。

5 结语

本文调查了古典钢琴演奏中每个乐句节奏型选取的潜在影响因素。实验结果表明，古典钢琴演奏中每个乐句节奏型的选取受节奏型序列的影响，而乐句本身属性对于乐句节奏型的影响应当叠加在节奏序列的影响之上。如果作为一个独立的影响因素，乐句本身的属性对于演奏中乐句节奏型的选取较为有限。

本文一共测试了4种候选贝叶斯图模型，每种模型假设节奏型时序序列和乐句位置对于单一乐句节奏型造成不同的影响。通过比较相对熵、交叉熵和本文新列举的交叉熵比等3项模型选择参数，我们发现考虑节奏型时序序列和乐句位置均对乐句节奏型造成影响的联合模型较其他候选模型的表现更好。在仅有有限数据进行模型训练时，仅考虑前序节奏型的时序模型由于具有更低的模型复杂度而具有更强的模型数据鲁棒性。比较本文所采用的模型选择参数：交叉熵、相对熵和交叉熵比，我们发现交叉熵比受训练集随机性的影响较小，故在使用贝叶斯图模型进行评估时，可以考虑使用相对熵与交叉熵比作为模型选择的依据。

本文利用了3个数据库对于决定乐句节奏型的因素进行了探索性的研究，但由于本文所采用的节奏型定义方法是根据文献[1]所提出的，受乐曲中乐句长度的限制，故无法被大规模应用于大型数据集上。亦鉴于此，文中的各项结论的适用范围仍有待进一步探索。此外，如大规模数据集得以应用，则文中所抽象的乐句位置和前序乐句节奏型的概念亦可展开。以乐句位置为例，其笼统地涵盖了包含调式、旋律等诸多音乐元素，故需大量数据对其所涵盖元素对乐句节奏型选择造成的影响一一加以甄别。

致谢：本文系作者在英国伦敦大学玛丽女王学院攻读博士学位期间的研究成果，英文版的详细内容可参见博士论文(<http://www.eecs.qmul.ac.uk/~simond/phd/ShengchenLi-PhD-Thesis.pdf>)。本文工作得到了马克·普拉博利教授，西蒙·迪克森博士和道恩·布莱克博士的指导，在此特表感谢。

参考文献：

- [1] LI S, BLACK D, PLUMBLEY M. Model analysis for intra-phrase tempo variations in classical piano performances [C/OL]// Proceedings of Computer Music Multidisciplinary Research. <http://cmr.soc.plymouth.ac.uk/cmmr2015/proceedings.pdf>.
- [2] SPIRO N, GOLD N, RINK J. The form of performance: Analyzing pattern distribution in select

- recordings of Chopin's Mazurka op. 24 no. 2 [J]. *Musicae Scientiae*, 2010, **14**(2): 23-55.
- [3] LI S, BLACK D, CHEW E, et al. Evidence that phrase-level tempo variation may be represented using a limited dictionary [C/OL]//Proceedings of International Conference on Music Perception and Cognition, 2014: 405-411. http://vbn.aau.dk/files/204037394/icmpc_apscom_2014_Proceedings_.pdf.
- [4] WIDMER G, FLOSSMANN S, GRACHTEN M. YQX plays chopin [J]. *AI Magazine*, 2010, **31**(3): 23-34.
- [5] TOBUDIC A, WIDMER G. Relational IBL in music with a new structural similarity measure [C]// Proceedings of the 13th International Conference on Inductive Logic Programming (ILP'03). Berlin Heidelberg: Springer, 2003: 365-382.
- [6] TODD N. The dynamics of dynamics: A model of musical expression [J]. *Journal of Acoustical Society of America*, 1992, **91**: 3540-3550.
- [7] FRIBERG A, BRESIN R. Overview of the KTH rule system for musical performance [J]. *Advances in Cognitive Psychology*, 2006, **2**(2-3): 145-161.
- [8] KOLLER D, FRIEDMAN N. Probabilistic graphical models: Principles and techniques [M]. Cambridge, Massachusetts: MIT Press, 2009: 733.
- [9] SAPP C. Hybrid numeric/rank similarity metrics for musical performance analysis [C]//Proceedings of International Conference on Music Information Retrieval. Philadelphia, USA: Lulu com, 2008: 501-506.
- [10] DIXON S. Automatic extraction of tempo and beat from expressive performances [J]. *Journal of New Music Research*, 2001, **30**(1): 39-58.
- [11] CLAESKENS G, HJORT N. Model selection and model averaging [M]. Cambridge: Cambridge University Press, 2008: 1.

Modelling the Decisive Factor of Expressive Timing Clusters for a Phrase in Classical Piano Performance with Mathematical Model Selection Tests

LI Shengchen

(Institute of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: In performed classical piano music, the clustering of expressive timing within a phrase is a common methodology. Many musicologists studied how the expressive timing should be varied within a phrase with different hypothesis. There are two major factors considered to contribute to expressiveness: The phrase itself and the sequence of expressive timing. Using the cluster of expressive timing used in a phrase and the previous phrase and the location of a phrase, four candidate Bayesian models are built to assert different dependencies of the use of expressive timing cluster. With relative entropy, cross entropy and a novel measure cross entropy ratio, the expressive timing cluster used in the previous phrase is more decisive than the location of phrase, which will only affect the expressive timing cluster of a phrase with the expressive timing cluster in the previous phrase jointly.

Keywords: cluster of expressive timing; model selection; piano performance; effects on cluster of expressive timing; Bayesian graphical models