

第 3 章

目 标

第 一个维度, **目标(objective)**, 是要生成的音乐内容的性质。

3.1 属性

我们从一个目标的主要五方面的属性, 来考虑这一问题:

(1) 类型(type): 生成内容的音乐本身的性质。例如, 生成旋律、复调及伴奏音乐?

(2) 目标(destination): 使用(处理)拟生成内容的实体。例如, 是生成供音乐家演奏用的乐谱? 还是供软件或音频系统播放的文件?

(3) 应用(use): 目标实体处理生成内容的方式。例如, 是播放生成的音频文件? 或演奏生成的乐谱?

(4) 模式(mode): 对生成音乐的控制方式, 即是否有交互性的人工干预(interaction)? 或完全自动生成没有任何干预(automation)?

(5) 风格(style): 生成音乐的风格。例如, 是生成具有约翰·塞巴斯蒂安·巴赫复调风格的合唱歌曲? 还是生成具有沃尔夫冈·阿马迪乌斯·莫扎特特色的奏鸣曲? 还是生成科尔·波特的歌曲或韦恩的短音乐? 事实上, 生成什么样的音乐风格, 与选择什么样的音乐样例(语料库)作为训练样本有关。

3.1.1 音乐类型

主要的音乐类型如下。

(1) **单音旋律 (single-voice monophonic melody)**, 简称**旋律(melody)**。它是由一种乐器(译者注: 钢琴、吉他等复调乐曲除外)演奏的或歌唱的音符序列, 在同一时刻最多出现一个音符。比如由单音乐器(如长笛)演奏的音乐^①。

^① 虽然有非标准的技术来产生多个音符, 但是最简单的还是边唱边弹。也有其他非标准的复音(diphonic)技术。

(2) **单音复调 (single-voice polyphony)**, 也称 **单轨复调 (single-track polyphony)**, 简称 **复调 (polyphony)**。它是一种乐器的音符序列, 可以同时演奏多个声部音符。例如, 由钢琴或吉他等复调乐器演奏的音乐。

(3) **多声部复调 (multivoice polyphony)**, 也称 **多轨复调 (multitrack polyphony)**, 简称 **多声部 (multivoice)** 或 **多轨部 (multitrack)**。它是一组多重声音/多音轨组成的曲调, 是一种以上的声音或乐器发出的乐音。例如, 由女高音、女低音、男高音和男低音组成的合唱, 或由钢琴、贝斯和鼓组成的爵士三重奏。

(4) 对给定主旋律的**伴奏 (accompaniment)**, 例如:

① **对位 (counterpoint)**: 由一个或多个旋律(声音)组成(译者注: 对位法是在音乐创作中使两条或者更多条相互独立的旋律同时发声并且彼此融洽的技术, 在巴洛克时期的复调音乐中得到了广泛的应用)。

② **和弦进行 (chord progression)**: 提供一些相关的**和声 (harmony)**。

(5) **将旋律 (melody) 与和弦进行 (chord progression) 相结合**。比如在爵士乐中常见的**领谱 (lead sheet)**^①, 它可能包括**歌词 (lyrics)**^②。

值得注意的是, 本节所述的**类型 (type)**, 在目标的五个属性中, 是最为重要的, 因为它抓住了音乐内容生成目标的本质。在本书中, 为了方便, 经常会根据**类型 (type)**, 例如, 将旋律作为一个简化的问题类型) 来确定生成目标。当考虑用户与内容生成过程的**交互 (interaction)**时, 下面的三方面(即**目标 (destination)**、**应用 (use)**、**模式 (mode)**), 才变得更重要。

3.1.2 目标与应用

目标与应用示例如下。

(1) **音频系统 (audio system)**: **播放 (play)**生成的音频文件。

(2) **音序器软件 (sequencer software)**: **处理 (process)**生成的 MIDI 文件。

(3) **人 (humans)**: **弹奏并诠释 (interpret)**生成的乐谱音乐。

3.1.3 生成模式

对于音乐的生成而言, 主要有以下两种生成模式。

(1) **自主 (autonomous)**与**自动式 (automated)**生成: 无须人工干预。

(2) (在某种程度上具有)人机**交互 (interactive)**生成: 有控制界面, 用户在生成过程中可以进行交互式的控制。

由于用于音乐生成的深度学习是最近才出现的, 而且基本的神经网络技术是非交互的, 因此我们分析过的大多数系统的互动性还不是很强^③。正如在 6.11.4 节将要介绍的

① 在图 4-13 中, 展示了一个五线谱领谱的例子。

② 注意, 歌词也是可以生成的。虽然这个目标超出了本书范围, 但我们将在后面的 4.7.3 节中看到, 在某些系统中, 音乐被编码为文本, 因此相关类似的技术也可以生成歌词。

③ 在 6.15 节将介绍一些具有人机交互的系统例子。

作为先驱的 FlowComposer 原型系统^[152]那样,为音乐家设计(用于作曲、分析、和声分析、编曲、制作、混音等)完全交互式的支持系统,已经成为在设计中需要考虑的一个重要问题。

3.1.4 音乐风格

如前所述,选择什么样的音乐数据集进行模型训练,决定了最后生成什么风格的音乐,4.12 节中将进一步讨论这个问题。我们会发现,训练数据集的风格,尤其是其**一致性 (coherence)**、**覆盖范围 (coverage)**(相对于**稀疏性 (sparsity)**而言)和**范围 scope**(专业化的 VS. 大众化的)等,对于生成好的音乐内容来说,是重要的基础。