# 一种基于共同语义空间的个性化音乐生成技术

## 背景简介

音乐生成是让计算机自动创作音乐的技术。音乐创作一直被认为是一种高难度，依赖人类突发灵感的高级思维活动。另一方面，音乐创作又需要遵循严格的规律，如节拍性、强弱性等。这意思着创造音乐是一项既要循规蹈矩，又要寻求新意的艰苦劳动，仅有对音高、节拍等具有敏锐感觉，且思维活跃度极高的少数人能够胜任。幸运的是，这种在严格框架下进行有限创新的工作，计算机具有天然优势，它可以充分保证生成作品的合规性，同时在合规下探索各种可能的创新。让计算机自动生成音乐，可极大减少人类进行音乐创作的工作量，且有望产生挣脱传统思路束缚新颖音乐。即便用机器生成的音乐还不能与人类的音乐家相比，但机器作品可以为人类提供候选或初级作品，使作曲家创作更加容易；同时，计算机生成的音乐还可以为作曲家提供灵感和刺激，激发他们不断创造新的音乐，仿止因长期创作带来的风格惰性和思维困顿，帮助作曲家永褒创作青春。因此，自动音乐生成具有非常广阔的应用前景。

音乐生成一般可分为XX种方法，例如... (补充)

## 问题描述

目前的音乐生成方法无法细致指定生成音乐的风格。要生成不同风格的音乐，或者人为挑选不同风格的训练数据进行与风格相关(style-dependent)的模型训练，或者引入一个表示风格的指示变量，表示想要生成的风格方式。虽然这些是可行的，但这些方法无法对风格进行细致指定。比如，同样是田园风格，反应农耕生活和反应狩猎生活的音乐可能差距较大。再如，如果我们想反应复杂风格，既有田园风格，也有魔幻风格，则传统的方法无法实现。除了风格，音乐还包括意境、情景、叙事等复杂属性。这些复杂属性到相交叉组合，意味着对音乐的描述极为复杂，更别提依这些描述来生成音乐了。

本发明关注音乐生成任务，提出一种新的方法，可通过文字描述来指定生成音乐的风格和属性。例如，通过输入“牛和羊在青青的草地上吃草，旁边有小溪流过”来指定生成音乐所表述的内容，音乐生成器即可生成与这一田园风格相适应的音乐。

## 发明要点

1. 将文字和音乐表达在同一个低维语义空间中，使得文字和音乐在该空间中的距离可计算。这意味着一段音乐（风格、意境等）可用其在该空间中所对应的文字细致刻画。
2. 用歌词与音乐的一一对应关系，训练上述低维语义空间。训练采用sequence-to-sequence的端对端训练方法。该框架中，一句歌词首先通过一个“编码器”成一个在低维语义空间中的表征向量，该向量再通过一个“解码器”产生相应的音乐。训练时，训练目标是生成的音乐与目标音乐的误差最小。
3. 音乐生成时，在编码器段输入若干关键词或句子，解码器即可生成与之相匹配的音乐。由于输入的句子是随意的，意味着这一方法可生成各种风格、各种意境的音乐。

## 发明内容和系统实现方法[补充]

1. 共同语义空间
2. 基于RNN的序列编码-解码器结构
3. 训练与生成

## 方案优势

1. 可指通过文字描述指令算法生成各种风格、意境的音乐。
2. 可通过变换文字描述方法，生成同一风格意境，但旋律不同的各种音乐。
3. 可通过文字描述的发展，生成一段具有相应意境发展结构的音乐，