

演算法作曲之十二小節藍調

Algorithmic Composition of Twelve-Bar Blues Music

指導教授: 溫育璋

學生: 秦敏瑜、何晉維、黃冠霖

國立聯合大學 資訊工程學系

苗栗市南勢里聯大2號

ywwen@nuu.edu.tw

{U1124003, U1124009, U1124034}@o365.nuu.edu.tw

摘要

本研究提出不依賴資料集、完全以樂理規則為基礎的自動作曲系統，強調可控性與可解釋性，用於生成12小節藍調音樂。系統包含兩項核心方法：機率式藍調生成演算法（SBGA）與基因演算法（GA）。此外，本研究整合音樂生成、樂譜渲染與Stable Diffusion 圖像生成，建構為網頁應用程式。實驗結果顯示，在無資料集支援下，兩種演算法皆能生成風格明確且結構合理的藍調旋律，驗證規則式自動作曲的可行性。未來可擴展其規則庫並結合歌詞與歌聲生成，提升系統功能與應用價值。

關鍵詞：基因演算法、12小節藍調、Stable Diffusion

Abstract

This study proposes a fully rule-based automatic composition system that does not rely on any dataset and emphasizes controllability and interpretability, aimed at generating twelve-bar blues music. The system incorporates two core methods: a Stochastic Blues Generation Algorithm (SBGA) and a Genetic Algorithm (GA). In addition, this research integrates music generation, score rendering, and Stable Diffusion-based image generation into a web application. Experimental results show that even without dataset support, both algorithms can produce blues melodies with clear stylistic

features and coherent structure, demonstrating the feasibility of rule-based automatic composition. Future work may expand the rule set and integrate lyric and vocal generation to further enhance system functionality and application value.

Key Word: GA, Twelve-Bars Blues, Stable Diffusion

一、研究動機與目的

隨著人工智慧的發展，現在AI音樂日漸興盛，但現在主流音樂多半依賴大型資料集進行深度學習訓練來模仿市場上著名或高商業價值的音樂，此類性上風格模擬上表現優異，但進而缺乏自我創造性，且使用者也較難理解音樂是如何被規則與結構創造出來的。

本研究的主要目標為在不依賴音樂資料集的前提下，證明音樂文法與樂理知識可以被量化為演算法生成作品。具體而言，本研究目標為建立一套能自動生成十二小節藍調旋律的系統，將音樂理論透過演算法控制音符排列與節奏設計，讓程式輸出具音樂性與即興感的旋律。

因此，樂理規則的建構設計會是這本次專題的關鍵。如何設計出既兼顧音樂的聽感，又可以讓音樂在一定的規則限制下靈活的即興，藉此證明在無資料集的情境下，單純依據樂理知識也能達成自動化作曲的目標。

二、系統架構及介紹

此系統目的在建構一套整合演算法與多模態生成技術的互動式創作平台。在系統設計上，後端以 Python 語言為基礎，運用 FastAPI 架構搭配執行緒（ThreadPool）技術。前端介面則採用 HTML5 與 JavaScript 開發，負責參數傳遞及即時樂譜渲染。在功能方面，系統包含三大引擎：

1. 演算法音樂生成引擎

分別有機率式藍調生成演算法及基因演算法，負責產出結構完整的 MIDI 音訊。

2. 視覺化渲染引擎

整合 MuseScore 4 CLI 工具，將抽象的音訊數據轉譯為 SVG 向量樂譜與 MPOS 座標定位檔，實現網頁的視聽同步。

3. 跨模態影像生成引擎

串接 Stable Diffusion 與 Stable Video Diffusion (SVD) 模型，負責生成封面及動態影片。

三、研究方法

（一）機率式藍調生成演算法

此為以機率為核心的十二小節藍調自動生成演算法，圖一為其主旋律流程圖。首先，程式會建立並初始化一個 PrettyMIDI 物件以作為建立整份 MIDI 檔案的主容器，並為12小節藍調建立獨立的樂器音軌物件，包括主旋律以及其會使用到的伴奏，並根據使用者設定或預設值（調性為 C，拍速為 120 BPM）初始化調性與拍速。旋律生成以拍為基準，起始於藍調音階根音，並透過機率選擇三種走向模式（隨機佔比10%、上行佔比45%、下行佔比45%）決定音高傾向。每一候選音會依與當前被選出的音高之距離、當前小節之和弦相符度與藍調特徵音（b3、b5）加權後，以機率方式選取。

節奏部分則依拍點決定可用節奏型：整拍可出現 Shuffle、三連音、四分音符與

八分音符，比例分別為：40%、30%、10%、20%；非整拍則移除三連音與 Shuffle 並加入二分音符（八分音符：30%、四分音符：35%、二分音符：35%），其比例設計強化藍調質感並維持律動自然。系統亦採用特殊的複製小節機制，在第 3、4、10 小節複製前一小節之音形，再調整音高貼合和弦，以加深音樂的主題感以及複製之旋律與小節和弦的和諧性。

完成十二小節後，所有音符資料將轉為 PrettyMIDI Note 物件並輸出為可播放的 MIDI 檔案，呈現具藍調特色、節奏自然且結構一致的旋律生成結果。

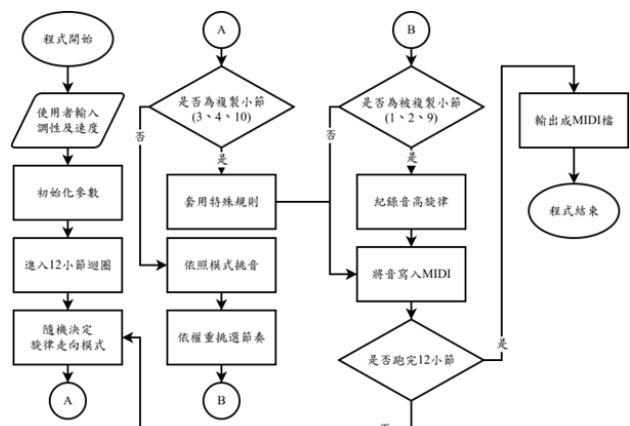


圖1：機率式藍調主旋律流程圖

（二）基因演算法

本方法採用基因演算法（GA）[1][2]，用於自動產生符合藍調風格的主旋律，其流程主要包含使用者輸入（BPM、調性）、初始化族群、適應度評估、交配與突變及生存選擇四個步驟，詳細流程如圖2所示。

1. 旋律編碼與初始化：

每一條染色體代表一條旋律，由基因序列組成。每個基因是旋律的最小單位，形態上分為單音（note）、休止符（rest）、延音（tenuto）彼此各佔用一個時間單位格，以及佔用兩個單位時間格的三連音（triplet）。樂曲長度由預設的藍調和弦進行決定。而初始族群旋律是根據所選的藍調音樂調性的藍調音階範圍內選擇音高，並且在初始化階段引入了多項樂理規則限制，包括：控制連續上下行的步數避免疲

倦感、防止旋律超越音域的折返處理，以及排除邏輯錯誤（如確保三連音的完整性、節奏對齊）。

2. 適應度評估：

適應度函數為衡量旋律品質的核心，本研究採用扣分制，結合人工撰寫的樂理規則和藍調風格特徵來進行評分，主要評估項目在引導旋律具備更好的音樂結構及更具藍調音樂特徵。核心規則包含懲罰不諧和大音程跳用、確保旋律起止於和弦跟音，以及強化正反拍呼應機制（加重和弦音懲罰藍調色彩音在強拍出現的機率）。其他評估規則亦涵蓋樂句結構完整性、音高重複性等，以確保生成之旋律符合基本樂理與聽感。

3. 交配與突變：

交配採用單點交配。為確保結構完整性，切點選擇必須在不破壞三連音結構前提下的位置。突變部分以低機率進行（10%），在排除對三連音的修改下（包含新增、刪除）導入局部隨機性，防止族群過早收斂。

4. 選擇與終止：

採用菁英選擇法（Elitism Selection）。每代將親代與子代合併後選取適應值最高的子集作為下一代族群。演化以固定最大世代數（1000代）作為終止條件，並輸出最終適應值的旋律作為輸出結果。

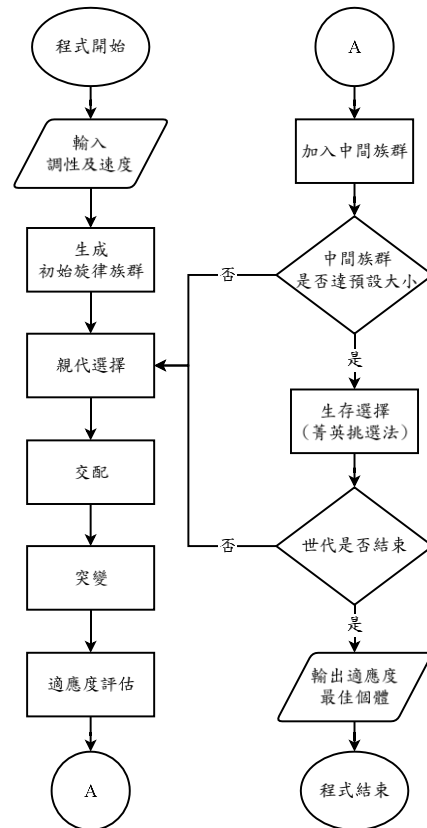


圖2：基因演算法主旋律流程圖

(三) 伴奏

在使用者選擇完音軌要使用哪些後，會進入到生成伴奏流程。伴奏包含有：Chord（全音符和弦）、Shuffle、兩種 Walking Bass（根音、五度音及接近音所組成之「穩定風格版」——使用根音、五度音以及接近音組成，由於幾乎為和弦音，所以會聽起來較為穩定。和由降三度、五度、降七度及接近音組成的「藍調氛圍版」——一、三拍為根音及屬音，二四拍從b3、b7等藍調色彩音中做選擇）及 Drum（採用 Blues drum 風格之鼓組）——根據藍調音樂中的 shuffle 風格設計，使自動生成的旋律配合鼓組後更句音樂性與律動感。

四、實作成果

本節將展示系統實際運作之使用者介面。透過以下畫面，可具體觀察本系統如何將後端的音樂生成演算法、樂譜渲染以及 Stable Diffusion 圖像生成，整合為一個完整且直觀的網頁應用程式。

1. 網站首頁

上半部提供直觀的參數介面，使用者可設定調性、速度並選擇生成引擎；下半部為作品展示區，透過 Stable Video Diffusion 技術將靜態封面轉化為動態影像，提升視覺互動性，如圖2所示。

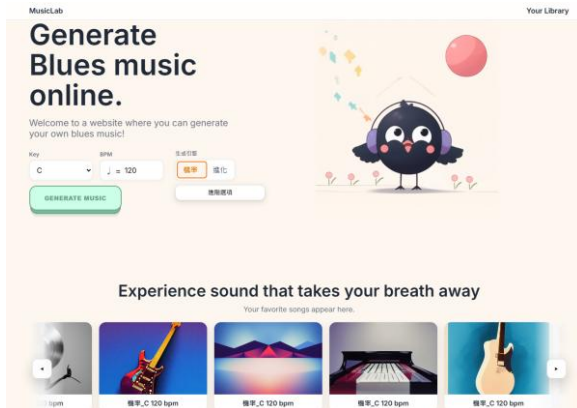


圖2：網站首頁

2. 歌曲管理列表

使用者可在此管理歷史創作，列表左側顯示由 Stable Diffusion 自動繪製的封面，列表右側提供重新命名、刪除歌曲、收藏及下載 MP3 音訊、PDF 樂譜與 MIDI 檔案等功能，如圖3所示。

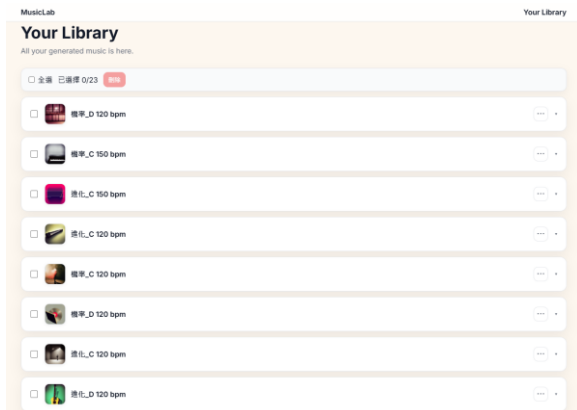


圖3：歌曲管理列表

3. 音樂播放及樂譜視覺化介面

圖中顯示由後端轉檔生成的 SVG 向量五線譜。系統利用 MPOS 座標定位技術，實現歌曲視聽同步效果，如圖4所示。

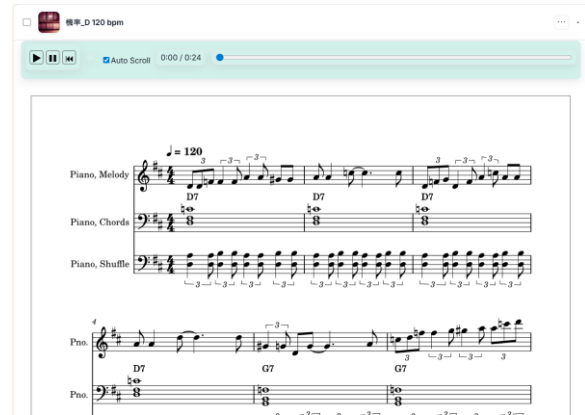


圖4：音樂播放與樂譜視覺化介面

五、結論與未來展望

(一) 結論

本研究成功實作出一套利用音樂文法為核心的藍調自動作曲系統，驗證了規則式演算法作曲在無需依賴大型資料及下生成音樂的可能性。本專案的核心目標為將複雜和隨機性極高的藍調樂理知識量化為演算法規則，透過自製權重機制與基於演化機制等兩種演算法證明只要掌握了音樂的結構與規範，即可實現高度可控的音樂生成。

(二) 未來展望

本研究未來發展方向可分為「加深」與「加廣」兩大主軸，加深可以做到的是在音樂生成後可以結合語言模型（讓使用者輸入文字情緒或主題，由 LLM 輔助生成歌詞）和歌聲合成模型（加入人聲合成技術），完成從詞曲、編曲到演唱的完整歌曲生成流程，使本專案具有更高的商業價值及觀賞性；加廣部分由於在程式實作上已具備良好的模組化設計，未來只需專注於擴展音樂理論庫，即可實現多種曲風變化，也可藉此讓未來想嘗試實作此系統的使用者創作出屬於自己的自定義風格。

六、專題作品影音平台網址

<https://youtu.be/9WtNIRhKdts>

七、參考文獻

- [1] Y.-W. Wen and C.-K. Ting, "Composing Bossa Nova by Evolutionary Computation," in *Proceedings of the

2020 International Joint Conference on
Neural Networks (IJCNN)*, Glasgow,
UK, Jul. 19-24, 2020, doi:
10.1109/IJCNN48605.2020.9207713.

- [2] Y.-W. Wen and C.-K. Ting, “Recent
advances of computational
intelligence techniques for composing
music,” *IEEE Trans. Emerg. Topics
Comput. Intell.*, vol. 7, no. 1, pp. 5–20,
Mar. 2023, doi:
10.1109/TETCI.2022.3221126.