

公立はこだて未来大学
2024年度
システム情報科学実習
グループ報告書
**Future University Hakodate 2024 System Information Science Practice
Group Report**

プロジェクト名
クリエイティブAI
**Project Name
Creative AI**

グループ名
音響班
**Group Name
Audio Group**

プロジェクト番号/Project No.
7

グループリーダー/Group Leader
1022161 富永鈴 Rey Tominaga
グループメンバー/Group Member
1022187 早川暉人 Akito Hayakawa
1022228 横沢永遠 Towa Yokosawa

指導教員
村井源 迎山和司 中田隆行
**Advisor
Hajime Murai Kazushi Mukaiyama Takayuki Nakata**

提出日
2025年1月 日
**Date of Submission
Generally. ,2025**

概要

クリエイティブAIプロジェクトでは、人工知能を用いた創造性の再現を目指して、ゲームを制作し、人工知能の創作方面における有用性を検討する。創作するものをゲームとした理由は、シナリオや音楽、視覚表現、ゲームシステムなど多様な創造性を表現することができるからである。人工知能を用いた創作の面白さは、未だに定義されていない領域であり、これを解明していくことが最終的な目的である。

このプロジェクトでは昨年までは4つの班に分かれて進行してきたが、今年度は人数の関係から、5つの班に分かれて進行することになった。ゲームシナリオや細かい世界観、伏線の作成を行う伏線班、登場人物の言動の出力を担当するキャラクター班、マップチップや画像の作成を行う視覚班、ゲームシステムの構築や、BGM やSE を担当する音響班、各班の成果物の統合などを担当するシステム班である。

伏線班では、伏線の自動生成に向け伏線が示す情報についての分析を行った。分析ではクラスタリングと因子分析の2つの手法を用いた。クラスタリングとは、多数のデータを複数のクラスターに分割する作業のことである。因子分析とは多変数解析のひとつであり、多数の変数で表現されるデータをより少数の因子に置き換えて表現する手法である。結果について、クラスタリングでは伏線のデータを3つのカテ

ゴリーに分類することができた。各カテゴリーの中身について細かく見ていくと、カテゴリー1には、今後の展開を示す伏線、カテゴリー2には奇妙な出来事などと一緒に張られた伏線が多く含まれていた。また、カテゴリー3にはカテゴリー1やカテゴリー2に含まれなかった余りのデータがまとめられていた。次に、因子分析の結果では因子寄与率の大きい上位六つの因子について分析・考察を行った。それぞれの因子は恐ろしい状況から犠牲者が発生することを示唆する伏線の因子や異常なことや不自然なことが提示される伏線の因子、物語終盤に明かされる意外な情報に関わる伏線の因子、現状や舞台、今後の展開を示唆する伏線の因子、新たな人物の登場に関わる伏線の因子、物語序盤の不可思議な状況に提示される伏線の因子であると解釈した。

キャラクター班では、ホラーゲームにおけるキャラクターの反応パターンを分析し、大規模言語モデルを用いて自動生成を行った。まず、ホラーゲームを対象として、キャラクターが恐怖を感じた際に起こす反応のデータを集め、クラスタリングを用いて分類した。分類結果を4つのクラスターに分け、全体の要素平均値と比較することで他のクラスターとの差異を定量的に評価した。次に、各クラスターを解釈し、それぞれ「ノーマルモデル」「状況変化モデル」「敵対者モデル」「ものや現象モデル」とした。以上の結果を確率表にまとめ、システムの作成を行った。このシステムは、キャラクターの反応を使用者が入力した情報と分析によって作成した確率表を用いてランダムに決定し、そのセリフと行動をOpenAIで生成してテキストファイルに書き出すものである。本システムを用いてキャラクターの反応を生成することで、シナリオ制作者の負担が大幅に軽減される可能性が示された。この成果は、AIを活用したキャラクター生成の実用性と可能性を示すものであり、今後さらなる改良を加えることで、より精度の高いセリフと行動の自動生成が期待される。

視覚班ではStable Diffusion v1-5モデルをLoRAでカスタマイズし、ホラーゲーム『青鬼』シリーズのマップを学習データとした。学習元と出力画像を比較するため、生成されたマップをt検定で評価した。生成画像は曖昧な箇所があるものの、元データの再現性は高いと評価された。ただし、アイテム部屋の配置に関しては弱点が見られた。また、本ゲームにおけるアイテム画像、メニューUI、マップチップなどの視覚表現は、Stable Diffusion、Blender、Figmaなどのツールを用いて行なった。生成画像は曖昧な箇所があるものの、元データの再現性は高いと評価された。ただし、アイテム部屋の配置に関しては弱点が見られた。また、Blenderを視覚表現に用いることで、ゲームの世界観を緻密に表現することができた。

音響班では、音楽の感情表現を深く研究し、AIを用いたメロディ生成に取り組んだ。具体的には、ラッセルの円環モデル(図1)を基に、感情(例: 喜び, 悲しみ, 興奮など)と音楽の特徴量(テンポ, 音程, ダイナミクスなど)との関係性をホラー映画のサウンドトラックの評価を行うことで分析した。この結果を活用することで、生成される音楽が特定の感情を効果的に伝えるよう設計した。技術面では、Pythonを主なプログラミング言語として使用した。また、音楽データの処理や解析には、Midoライブラリを利用してMIDIデータを操作し、感情に適したメロディの自動生成AIを制作した。メロディの自動生成AIにはLSTMモデル(図2)を利用した。その結果、特定の感情を意図的に喚起する音楽を自動生成するシステムを完成させた。しかし、複数のパターンがある感情に関しては、その感情を喚起させるメロディが生成されにくかった。そのため、AIが学習を行う際の重きを変える必要があることが分かった。また、怖い曲は楽曲の特徴が一貫しているのに対し、楽しい曲は楽曲の特徴が発散しているのに加え、より音楽理論に乗っ取ったきれいなメロディが求められるため、怖い曲のメロディを生成するよりも楽しい曲のメロディを生成する方が難しいことが分かった。

システム班では、他班の制作物の統合のために、スクリプトエンジンというシステムを開発した。このシステムは他班にJSONファイル形式でデータを記述してもらい、ゲームに反映させるというものである。このシステムは3つの部分に分かれている。一つ目はマップエンジンで、マップデータを記述することでゲーム上の各マップの見た目と、マップの地点の性質を指定することができ、主に視覚班との連携に使用した。二つ目はオブジェクトエンジンで、マップの位置とフラグ条件、イベントの種類を記述することで、

三つ目は会話エンジンで、会話の内容や、会話中に発生するイベント、会話画面の画像の指定を記述することで、会話データと会話画面の編集ができる。主に伏線班との連携に使用した。これらのエンジンを使用することで、JSONファイルを変更することで素材の差し替えができるため、統合のためにシステム班のメンバーに依頼する時間の削減につながった。

全体では、ゲーム内には未だ不足が残る箇所がありつつも、軸となるシナリオや付随するミニゲームの大枠は完成した。また現状として、各班はそれぞれ目標としていた人工知能の活用もその大部分を達成することができた。今後はより操作性や遊戯性を追求したゲーム部分の制作を進めるとともに、人工知能活用の面においても更にデータを収集して結果の精度を上げることや、システムの構造をより適切なものに変更するなどの改善を行っていく。

(* 文責: 富永鈴)

Abstract

The goal of the Creative AI project was to reproduce human creativity using artificial intelligence. To achieve this goal, we decided to create a computer game because a high level of creativity is required to produce various aspects of a game, including scenarios, music, visuals, and game systems. The intellectual appeal of creating with artificial intelligence has not yet been investigated, and we aim to uncover this question.

In this project, the work was divided into four groups until last year. However, due to the number of participants this year, the project has been divided into five groups: the Foreshadowing Group, responsible for creating the game's scenario, detailed world-building, and foreshadowing; the Character Group, responsible for generating the actions and dialogues of characters; the Visual Group, responsible for creating map chips and images; the Audio Group, responsible for constructing the game system, BGM, and sound effects; and the System Group, responsible for integrating the outputs of each group.

The Foreshadowing Group analyzed the information conveyed by foreshadowing in preparation for automated foreshadowing generation. The analysis used two methods: clustering and factor analysis. Clustering involves dividing a large dataset into multiple clusters, while factor analysis is a multivariate analysis method that represents data expressed by many variables using fewer factors. The clustering results classified the foreshadowing data into three categories. Upon closer examination, Category 1 included foreshadowing that hinted at future developments, Category 2 included foreshadowing paired with strange events, and Category 3 contained residual data not classified in the other two categories. Factor analysis focused on the top six factors with high factor contributions, interpreting them as follows: foreshadowing suggesting horrific situations leading to victims, foreshadowing presenting abnormal or unnatural events, foreshadowing related to surprising revelations near the story's end, foreshadowing indicating current or future developments, foreshadowing involving the appearance of new characters, and foreshadowing tied to mysterious situations at the story's beginning.

The Character Group analyzed reaction patterns of characters in horror games and automated their generation using large language models. Initially, data on character reactions to fear were collected from horror games and classified using clustering. The results identified four clusters, which were quantitatively evaluated against overall averages. Each cluster was interpreted as the "Normal Model," "Situation Change Model," "Antagonist Model," and "Object/Phenomenon Model." These results were summarized in a probability table, forming the basis of a system that determines character reactions randomly based on user-input information and the analyzed probability table. The system generates dialogues and actions using OpenAI, outputting them as

text files. By automating character reactions, the system significantly reduces the workload of scenario creators. This achievement demonstrates the practicality and potential of AI-assisted character generation, with further improvements expected to enhance the accuracy of generated dialogues and actions.

The visual Group customized the Stable Diffusion v1-5 model using LoRA, with maps from the horror game series *Ao Oni* as the training data. To compare the training data with the generated images, a t-test was used for evaluation. While the generated maps had some ambiguous areas, they were highly rated for their reproducibility of the original data. However, weaknesses were observed in the arrangement of item rooms. Additionally, the visual elements of the game, such as item images, menu UI, and map chips, were created using tools like Stable Diffusion, Blender, and Figma. By utilizing Blender for visual representation, the game's world was expressed with intricate detail, enhancing its immersive quality.

The Audio Group conducted in-depth research on emotional expressions in music and worked on melody generation using AI. Based on Russell's Circumplex Model (Figure 1), they analyzed the relationship between emotions (e.g., joy, sadness, excitement) and musical features (e.g., tempo, pitch, dynamics) by evaluating soundtracks from horror movies. Using these results, they designed music to effectively convey specific emotions. Technically, Python was the primary programming language, with the Mido library used for MIDI data processing and analysis. An AI system for emotional melody generation was created using an LSTM model (Figure 2). The system successfully generated music that evoked specific emotions, though it struggled with emotions having multiple patterns. Adjustments in the AI's training emphasis were identified as necessary. Furthermore, it was found that generating happy melodies was more challenging than scary ones due to the need for adherence to music theory and diverse characteristics in joyful music, unlike the consistent features of scary music.

The System Group developed a script engine to integrate the outputs of other groups. This system allowed other groups to describe data in JSON format, which was then reflected in the game. The system comprised three parts. The Map Engine allowed the description of map data, specifying the appearance and properties of each location, mainly used in collaboration with the Visual Group. The Object Engine enabled the description of map positions, flag conditions, and event types. The Conversation Engine allowed the description of conversation content, events during conversations, and conversation screen images, mainly used in collaboration with the Foreshadowing Group. These engines allowed for material replacement by modifying JSON files, significantly reducing the time required for integration tasks by the System Group members.

Until last year, this project had proceeded in four groups, but this year, due to the number of people involved, the project will proceed in five groups. The groups consist of the Foreshadowing Group, which is in charge of creating game scenarios, detailed worldviews, and foreshadowing; the Character Group, which is in charge of outputting characters' words and actions; the Visual Group, which creates map chips and images; the Music Group, which creates background music and sound effects; and the System Group, which builds game systems and integrates the work of the various groups.

(* responsibility for wording of article: Towa Yokosawa)

目次

第1章 はじめに

- 1.1 音響班とは
- 1.2 班の概要
- 1.3 AIの活用について
- 第2章 関連研究
 - 2.1 関連学習
 - 2.2 ラッセルの円環モデル
- 第3章 プロジェクト学習の目標
 - 3.1 目標
- 第4章 目的を達成するための手法、手段
 - 4.1 感情の種類決め
 - 4.2 音楽特徴量
 - 4.3 音楽要素12項目
 - 4.4 Praat
 - 4.4.1 Praatとは
 - 4.4.2 利点
 - 4.4.3 欠点
 - 4.4.4 分析可能項目
 - 4.5 DAW(Digital Audio Workstation)
 - 4.5.1 DAWとは
 - 4.5.1.1 主な機能
 - 4.5.1.1.1 StudioONE
 - 4.5.1.1.2 Logic Pro
 - 4.5.1.2 使用例
 - 4.5.1.3 魅力と活用の幅
 - 4.5.2 使用例
 - 4.5.3 魅力と活用の幅
 - 4.6 映画音楽の評定
 - 4.6.1 映画「IT」の音楽評定
 - 4.6.2 映画「口裂け女」の音楽評定
 - 4.7 MIDI
 - 4.7.1 MIDIファイルとは
 - 4.7.1.1 特徴
 - 4.7.1.2 用途
 - 4.7.1.3 仕組み
 - 4.7.1.4 収集と作成
 - 4.8 BGM
 - 4.8.1 ゲームBGMとは
 - 4.8.2 AIにBGMを作成させるシステム
 - 4.8.3 AIBGM
 - 4.9 ゲームの効果音とは
 - 4.9.1 SEとは
 - 4.9.2 SE作成プロセス
 - 4.9.2.1 独自音源
 - 4.9.2.2 外部音源
 - 4.10 システム
 - 4.10.1 開発環境
 - 4.10.2 ライブラリ
 - 4.10.2.1 os
 - 4.10.2.2 pretty_midi
 - 4.10.2.3 csv
 - 4.10.2.4 pandas as pd
 - 4.10.2.5 numpy as np

- 4.10.2.6 tensorflow as if
 - 4.10.2.7 sklearn.metrics import accuracy_score
 - 4.10.2.8 tensorflow.keras.models import Sequential
 - 4.10.2.9 tensorflow.keras.optimizers import Adam
 - 4.10.2.10 tensorflow.keras.layers import LSTM, Dense Unput, Masking
 - 4.10.2.11 tensorflow.keras.losses import MeanSquaredError
 - 4.10.2.12 sklearn.preprocessing import SrandardScaler
 - 4.10.2.13 mido import MidiFile, MidiTrack, Message
 - 4.10.2.14 tensorflow.keras.models import load_model
 - 4.10.3 ニューラルネットワークの基礎
 - 4.10.4 学習モデル
 - 4.10.5 LSTM(Long Short-Term Memory)
 - 4.10.5.1 LSTMとは
 - 4.10.5.2 構造
 - 4.10.5.3 利点
 - 4.10.5.4 応用と活用
 - 4.10.5.5 試用部分
- 第5章 結果
- 5.1 人力BGM作成
 - 5.1.1 悲しい&不穏&若干恐怖
 - 5.1.2 謎&困惑&少しの恐怖
 - 5.1.3 歓喜
 - 5.1.4 オルゴールバージョンの歓喜&悲しい
 - 5.1.5 恐怖&淡々
 - 5.1.6 恐怖&激しい&蹂躪
 - 5.1.7 タイトル曲
 - 5.2 AIBGM作成
 - 5.2.1 恐怖&グロテスク&激しい&蹂躪
 - 5.2.2 葛藤&恐怖&焦燥
 - 5.3 成果
- 第6章 考察
- 第7章 参考文献

第1章 はじめに

1.1 音響班とは

音響班では昨年度のプロジェクトを参考にしつつ、先輩方の指導を受けながら環境構築を進めた。昨年度は3つの楽器に絞り、AIを活用して場面ごとにフレーズを組み合わせる手法を採用していた。今年度はその手法を踏まえながら、音楽要素の分析方法やシステムの制作手法を学び、AIによる作曲の基盤となるメロディ生成を目指して、最適なモデルの開発に取り組んだ。

(* 文責: 富永鈴)

1.2 班の概要

ゲームのプレイヤーにより高い没入感を与えることを目的に、BGMと効果音を作成する。BGMでは、場面の感情に適したループ音源を使用する。効果音はフリー音源を使用し、自分たちで編集する。

(* 文責: 富永鈴)

1.3 AIの活用について

楽曲制作のすべてをAIに任せることは、感情の種類ごとに旋律の特徴をAIに学習させるために必要なデータ量は膨大となることが判明したため、楽曲制作のすべてをAIに任せることは難しいと判断した。そのため、楽曲のもととなるメロディの生成に特化したAIを作成し楽曲生成を行った。

(* 文責: 富永鈴)

第2章 関連研究

2.1 関連学習

以下は,MeinelとBullerjahn(2022)の要約である.この論文では,ホラー映画における音楽の配置が心理生理学的ストレス反応に与える影響を調査している.

目的:

この研究は,ホラー映画のシーンにおける音楽の配置が観客の心理生理学的反応(皮膚伝導率や心拍数)に与える影響を調べることを目的としている.

方法:

研究では,映画「REC」の2分20秒のシーンを39名の参加者に視聴させ,以下の3つの条件を比較した.

1. 同期条件: 音楽が視覚的ショックと一致して再生.
2. 非同期条件: 音楽が視覚的ショックの前に再生.
3. コントロール条件: 音楽なし.

結果:

- ・音楽がある条件(同期・非同期)では,皮膚伝導率が有意に高くなることが確認された.
- ・非同期条件では,音楽がショックの直前にストレスレベルを上昇させる効果が観察された.

結論:

音楽の配置はホラー映画視聴中の視覚的ショックを強調し,ストレス反応を増幅する.特に,非同期的な音楽配置が,ショックが訪れる前から視聴者を警戒させる効果を持つことが明らかになった.

(* 文責: 横沢永遠)

2.2 ラッセルの円環モデル

1980年に心理学者ジェームズ・A.ラッセルによって提唱された.ラッセルの円環モデル(Russell, 1980)は,感情の構造を説明するための理論的枠組みである.このモデルは,感情が2次元の空間上で分布することを示し,これにより異なる感情の関係性を視覚的に理解しやすくしている.・基本構造

ラッセルの意円環モデルは,円形の2次元空間で感情を配置している.この2次元空間は,以下の2つの軸から成り立っている.

1. 気分価:

水平方向の軸で,感情の快楽的な側面を示す.右側が「快」(例: 幸福,喜び)で,左側が「不快」(例: 悲しみ,怒り)を表す.

2. 覚醒度:

垂直方向の軸で,感情の覚醒状態を示す.上側が「覚醒」(例: 興奮,緊張)で下側が「非覚醒」(例: 落ち着き,リラックス)を表す.

・感情の配置

円環モデル内では,これらの軸に基づいて感情が配置される.例えば,「幸福」(happiness)は気分価が「快」で覚醒度が高く,「リラックス」(relaxation)は気分価が「快」で覚醒度が低い.同様に,「怒り」(anger)は気分価が「不快」で覚醒度が高く,「悲しみ」(sadness)は気分価が「不快」で覚醒度が低い.

(* 文責: 横沢永遠)

第3章 プロジェクト学習の目標

3.1 目標

AIを活用して,楽曲の基盤となるメロディの生成を目指す.

(* 文責: 早川暉人)

第4章 目的を達成するための手法、手段

4.1 感情の種類決め

上記したラッセルの円環モデルに基づいて,音楽から受け取る感情を評価するために,特定の感情カテゴリを設定し,音楽特徴量などのデータを収集することにした.今回は以下の4つの感情カテゴリを選定した.

1. 悲しい(sad)
2. 怒り(anger)
3. 楽しい(happy)
4. 安心(peace of mind)

なお、安心(deep emotion)については、ラッセルの円環モデルに直接含まれる感情カテゴリではない。しかし、本研究において、音楽が聴衆に与える不快情緒的体験や、感銘を受けるような情動を適切に評価するために、このカテゴリを独自に設ける必要があると考えた。安心は、単なる高次のポジティブな感情という枠を超えた、個々の音楽体験における重要な要素であり、本研究の目的に沿ったデータ収集や分析を行う上で欠かせないものである。

これらの感情カテゴリを基に、音楽から受け取る感情ごとの評定や、音楽特徴量のデータ収集を進めていく。これにより、感情と音楽特徴量の関係を明確にし、感情表現のメカニズムを解明することを目指した。

(* 文責: 横沢永遠)

4.2 音楽特徴量

音楽特徴量とは、音楽の特性や性質を数値やカテゴリデータで表現するものであり、音楽データの分析、分類、検索、推薦システム、さらには自動作曲や感情分析など多岐にわたる分野で活用されている。以下に、音楽特徴量をカテゴリごとに詳しく説明する。

1. 時間的特徴量(Temporal Features)

時間的特徴量は、楽曲のテンポやリズム、音の発生感覚など、時間に関連する側面を表現する。

・テンポ(BPM: Beats Per Minute)

楽曲の速さを表す特徴量で、ダンスミュージックでは高速なテンポ、バラードでは低速なテンポが一般的である。

・リズムパターン

拍子やビートの強弱パターンを示す特徴量で、楽曲のノリやダンス適性を評価する際に重要だ。小節単位のリズムパターン分析や、サブビート構造の抽出が含まれる。

・オンセット(音の開始タイミング)

音が発生する瞬間を検出する特徴量で、リズムの詳細解析や演奏表現の研究に利用される。

2. 音高特徴量(Pitch Features)

音高特徴量は、音の高さやその変化に関する情報を表す。

・メロディ

音高の連続的な変化を捉え、主旋律の構造や動きを分析するための基礎データとなる。

・音程(インターバル)

連続する音の高さの差を示し、旋律の方向性や感情的なニュアンスを把握する。

・キー(Key)

曲の調性を示し、メジャーキー(明るい印象)やマイナーキー(暗い印象)を区別することで、楽曲の感情分析やジャンル特定に役立つ。

3. 音色特徴量 (Timbre Features)

音色は、音の「質感」や「個性」を示す要素であり、異なる楽器や音源を区別する際に重要である。

・スペクトル特徴量

音の周波数成分を分析し、高音域・低音域のバランスや倍音構造を表す。

例: フーリエ変換やケプストラム分析に基づく情報。

・MFCC (Mel-Frequency Cepstral Coefficients)

音色を圧縮してコンパクトに表現する特徴量で、音声人認識や音楽分野で広く利用されている。

・スペクトル平坦性

音がどれだけ「純音」に近い(例: フルート)や「雑音」に近い(例: ドラム)を示す指標である。

4. ダイナミクス特徴量 (Dynamics Features)

ダイナミクスは音の強弱や変化を表現する特徴量で、楽曲の演奏表現を分析するのに欠かせない。

・音量 (ラウドネス)

音の強さを示し、楽曲の感情的な起伏や、聴覚心理モデルの研究に使用されている。

・振幅包絡 (エンベロープ)

音波系の時間的な強度変動を示すもので、波形の各頂点を結ぶ比較的ゆっくりとして変動部分だ。これにより、音の強度の変化を表現できる。音波系は通常、早く振動する時間微細構造(搬送波)と、この振幅包絡の成分が重ね合わさったものと考えられる。

5. ハーモニー特徴量 (Harmony Features)

ハーモニー特徴量は、同時に鳴る音同士の関係性や、楽曲全体の和音構造を表す。

・コード(和音)

和音の種類(例: メジャーコード、マイナーコード、セブンスコード)を特定し、楽曲の感情的トーンやジャンル特性を分析する。

・和音進行

楽曲内での和音の変化パターンを捉え楽曲の構造や展開を理解するための手がかりとなる。

6. 形式特徴量(Structural Features)

形式特徴量は、楽曲全体の構造や繰り返しパターンを分析する際に使用される。

・セクション(Section)

楽曲の構成要素(例: イントロ、バース、コーラス)を特定し、楽曲のドラマ性や記憶しやすさを評価する。

・リピート(繰り返し)

楽曲内でのフレーズやモチーフの繰り返しを分析し、リスナーの印象形成や楽曲記憶に関する研究に活用される。

7. 応用例

これらの特徴量は、以下のような応用に活用されている。

・音楽分類とジャンル判別

特徴量を基に、楽曲をポップス、ジャズ、クラシックなどのジャンルに分類する。

・楽曲推薦システム

ユーザの好みにあった楽曲を自動で選定するアルゴリズムに利用される。

・作曲支援

特徴量から得られるパターンを基に、新しい楽曲やメロディを生成する。

・感情分析

音楽のテンポ、キー、ハーモニーなどを分析して、楽曲が引き起こす感情(悲しみ、喜び、怒りなど)を予測する。

本研究では作曲支援と感情分析に活用した。

(* 文責: 横沢永遠)

4.3 音楽要素12項目

映画音楽の評価を目的として、音楽要素の12項目を使用し、さらに心理要素を主観的に評定することを決定した。この取り組みは、P.N.ジュスリン他(2008)で実際に使用されている分離した音楽的要素から12項目を抜粋し、それらを基にしている。以下に映画音楽の評定で使用した音楽要素12項目と詳しい説明を記述する。

・音楽要素の12項目

1. 振幅要素(Amplitude Element):

音の強さや大きさを示す。音量の変化や強弱の表現が含まれる。

2. アーティキュレーション(Articulation):

音と音のつながり方や音の出し方のニュアンスを示す。代表的なアーティキュレーションにはスタッカート(短く切る)、レガート(滑らかに繋げる)、マルカート(各音を強調)、テヌート(音を引きのばす)、アクセント(特定の音を強く演奏)、ポルタート(スタッカートとレガートの間)、スフォルツァンド(瞬間的に強く)、フェルマータ(音符や給付を通常よりも長く保持)、グリッサンド(特定の音を早く繰り返す)、トリル(主要な音と隣接する音を交互に繰り返す)、モルデント(主要な音とそのすぐ下または上の音を素早く交互に演奏)などがある。

3. ハーモニー(Harmony):

同時になる音の組み合わせ,和音の構成や進行を示す.和音の種類やその響き,単純協和(短音),複雑,不協和音などを使用した.

4. 音の大きさの変動(Dynamics):

音量の変化や強弱の変化を示す.代表的なものには,クレッシェンド(次第に大きく)やデクレッシェンド(次第に小さく),スフォルツァンド(特定の音を突然強調),フォルテピアノ(音を強く演奏し,その直後にすぐ弱く)などがある.

5. 旋律(Melody):

音の連続によって形成される楽曲の主な音の流れ.メロディラインの構造やその動きが重要.今回は順次と跳躍に振り分けた.

・順次(Stepwise motion):

音階上で隣り合う音同士の動きを指す.例えば、「ド」から「レ」,あるいは「ミ」から「ファ」への音程変化が順次に該当する.この動きは滑らかで自然な印象を与えやすく,メロディに安定感や親しみやすさをもたらすことが多い.順次進行はバロックや古典派音楽をはじめ,幅広いジャンルで一般的に用いられている.

・跳躍(Leaping motion):

音階上で隣合わない音への動きを指す.例えば、「ド」から「ミ」や「ソ」,あるいは「ラ」から「ド」への移動が跳躍である.跳躍進行はダイナミックでまだ津印象を与え,しばしば感情の高まりや劇的な表現を伴う.跳躍を多用した旋律は,エネルギーで個性的なメロディラインを作り出すことができる.

6. 調(Mode):

音階やスケールの種類を示す.代表的なものとして,ハ長調(すべての音が白鍵),ト長調(F#が含まれる),ニ長調(F#とC#が含まれる),イ長調(C#とF#とG#が含まれる),ホ長調(F#とG#とC#とD#が含まれる),ヘ長調(B♭が含まれる),変ロ長調(B♭とE♭が含まれる),変ホ長調(E♭とA♭とB♭が含まれる),イ短調(#と♭が含まれない自然短音階),ニ短調(B♭が含まれる)がある.今回は短調と長調に振り分けた.

7. 音域(Range):

楽曲で使用される音の最低音から最高音までの範囲.

8. 調性(Tonality):

楽曲の中心となる音や和音を示す.今回は調性が明確なもの,無調性に振り分けた.

9. リズム(Rhythm):

音の時間的な配置やパターンを指す.今回は一定化不規則化で分類した.

10. テンポ(Tempo):

楽曲の速さを指す.どのように変化しているのかを評定した.

11. 音色(Timbre):

音の質感や特性を示す.楽器の種類や演奏方法について評定した.

12. 展開(Development):

楽曲の構造や展開の仕方を示す.テーマの反復や変奏,クライマックスの配置を評定した.

(* 文責: 横沢永遠)

4.4 Praat

4.4.1 Praatとは

Praatは,音声を分析,変換,合成することができる無料のソフトウェアだ.アラン田のアムステルダム大学のPaul Boersma氏とDavid Weenink氏によって開発され,音声学や言語学の研究分野で広く利用されている.

(* 文責: 冨永鈴)

4.4.2 利点

1. 手軽に入手可能で無償

Praatは無料でダウンロード可能で,初心者でも手軽に使用できる.

2. 短時間で再合成音声や実験の準備が可能

音声の再合成や加工に必要な準備が比較的容易で、効率的に作業を進めることができる。

3. 視覚的音声分析に最適

ほかのフリー音声分析ソフトと比較し、音声の視覚化機能に優れており、語学的音声分析に最適である。

(* 文責: 富永鈴)

4.4.3 欠点

1. 再合成音声の自然性に限界がある

再合成した音声が完全に自然な音質になるとは限らず、特に細かいニュアンスを必要とする場面では不十分な場合がある。

(* 文責: 富永鈴)

4.4.4 分析可能項目

1. スペクトログラム(Spectrogram/Spectrum)

音の分析結果を視覚的に示すパターンで、以下の情報を含む。

- ・時間(Time): 横軸で表示
 - ・周波数(Frequency): 縦軸で表示
 - ・音の強さ(Intensity): グレースケールの濃淡で表現
- 典型的なスペクトログラムは、時間、周波数、音の強さを3次元的に表現するグラフである。

2. ピッチ(Pitch)

音の高低(基本周波数)を示し、青い曲線でその推移を視覚化する。

3. 音の強度(Intensity)

音声波形の強さのみを表示。特に摩擦音や子音を判断する必要がない場合に役立つ。

4. フォルマント(Formant)

声道の共鳴特性を表すもので、以下のびょうそで特定される

- ・フォルマント周波数(Formant Frequency): 中心周波数
- ・帯域幅(Bandwidth): 共鳴の範囲

5. 声帯振動(Pulses)

音響音声学における1回の声帯振動を示す。音声の周期性を把握する際に重要である。

6. 波形(Waveform)

音声信号の振幅を時間の関数として表すグラフで、Praatではデフォルトで表示される。これにより、音声データの基本構造を簡単に確認できる。

(* 文責: 富永鈴)

4.5 DAW(Digital Audio Workstation)

4.5.1 DAWとは

DAWとは、音楽制作を行うためのソフトウェアで、デジタル形式での録音、編集、ミキシングを可能にするツールである。DAWは音楽制作プロセス全体をデジタル環境で完結させることができ、現代の音楽制作において欠かせない存在となっている。

(* 文責: 横沢永遠)

4.5.1.1 主な機能

DAWは以下のような多彩な機能を備えている

1. 録音

マイクや楽器を接続し、デジタル形式で高品質な音声データを録音した。

2. 編集

音声データやMIDIデータをコア核編集することが可能で、タイミングの調整や音の切り貼りなどを行った。

3. ミキシング

複数の音源を重ね、音量バランスや音像定位(パンニング)を調整して、楽曲全体の仕上がりを整えた。

4. エフェクトの追加

リバーブ、ディレイ、イコライザーなどのエフェクトを適用し、音楽の質感や深みを向上させた。

5. 楽器シミュレーション

ソフトウェアシンセサイザーやサンプラーを使用して、ピアノ、ギター、ドラムなど多彩な楽器音を生成した。

6. 書き出し(エクスポート)

完成した楽曲を様々なフォーマット(MP3,WAVなど)で出力し、ゲーム制作に対応させた。

(* 文責:横沢永遠)

4.5.1.1.1 StudioONE

Studio Oneは、PreSonus社が開発・提供しているDAW(Digital Audio Workstation)ソフトウェアで、音楽生成やオーディオ編集のプロセスを統合的にサポートするツールである。本研究では音楽生成の際に活用した。以下に主な特徴と機能について説明する。

1. 直観的な操作性

特徴的なポイントの一つは、洗練されたユーザーインターフェースだ。画面上で視覚的にわかりやすく整理された構成により、初心者でもスムーズに音楽制作を始めることができる。ドラッグ&ドロップ操作を多用した設計により、音声ファイルやエフェクト、インストゥルメントトラックの追加が直観的に行える。

2. 統合されたワークフロー

音楽制作の各工程(作曲、編曲、録音、ミキシング、マスタリング)を一つのソフトウェア内で完結できるよう設計されている。特にマスタリング機能がほかのDAWと一線を画しており、作曲データとマスタリングデータのリンク機能により、変更をリアルタイムで繁栄させることが可能だ。この一貫したワークフローにより、作業効率が大幅に向上する。

3. 多彩なプラグインとサウンドライブラリ

高品質なエフェクトプラグインや音源が標準で搭載されている。例えば、「ProEQ」「Compressor」「Limiter」といったプロフェッショナル向けのエフェクトが容易されており、録音したオーディオデータを精密に編集することが可能。また、数1,000種類サンプル音源やプリセットが含まれたサウンドライブラリも提供されており、ジャンルを問わず活用できる。

4. MIDI編集とソフトウェアインストゥルメント

高度なMIDI編集機能を備えており、ピアノロールエディタでの詳細な編集が可能である。また、付属のインストゥルメントを用いることで、電子音楽やオーケストラ楽曲など、多様なサウンドデザインが行える。

5. スクリプトとカスタマイズ

ユーザー自身がスクリプトを作成することで機能を拡張することが可能である。また、ショートカットキーのカスタマイズにも対応しており、ユーザーの作業スタイルに合わせた環境を構築できた。

(* 文責:横沢永遠)

4.5.1.1.2 Logic Pro

Logic Pro(ロジックプロ)は、Apple社が開発しているプロフェッショナル向けのDAW(Digital Audio Workstation)ソフトウェアである。主に音楽制作やオーディオ編集のために設計されており、その洗練されたユーザーインターフェースと豊富な機能から、作曲家、プロデューサー、エンジニアをはじめ、多くの音楽業界関係者に指示されている。今回は、音楽生成と効果音の編集に使用した。以下に主な機能と特徴について記載する。

1. 直観的で美しいデザイン

Apple製品の特長であるシンプルかつ美しいインターフェースを採用している。Mac専用として最適化されているため、安定性が高く、macOSとの連携がスムーズだ。初心者でも操作に迷いにくい設計になっている一方で、高度な機能も簡単に利用できるように配置されている。

2. 統合された作曲と編集機能

アイデアを簡単に形にできる作曲機能が充実している。特に以下のような機能が減収作業を効率化する

- Live Loops: ループベースの音楽生成に適した機能で、リアルタイムでアイデアを試しながら作曲可能。
- MIDI編集: ピアノロールやステップエディタを使用した詳細な編集。
- アルペジエーターやコードトリガー: MIDIトラックで複雑な音楽的パターンを簡単に生成

3. 膨大な音源とサウンドライブラリ

標準で膨大な音源ライブラリを提供している。

- ・サウンドエフェクト: リバーブ,ディレイ,EQ,コンプレッサーなど,高品質なエフェクトが多数搭載
- ・ループライブラリ: 数1,000種類のApple Loopsが付属し,ジャンルを問わず即座に利用可能。

4. Drummer機能

特有の機能である「Drummer」は,リアルなドラムパターンを生成するためのツールだ。AIがドルむパターンを生成し,ジャンルやスタイルに応じてカスタマイズが可能で,作曲の初期段階でのアイデアを効率的に広げることができる。

5. 強力なミキシングとマスタリングツール

プロフェッショナルなミキシングとマスタリングツールを内蔵しており,最終的な音質調整やアルバム制作まで対応可能である。以下の機能が時に強力である。

- ・Smart Tempo: 異なるテンポのトラックを自動で同期。
- ・Flex TimeとFlex Pitch: オーディオトラックのタイミングやピッチを簡単に編集。
- ・マスタリング機能: 標準装備のLimiterやMultiPressorを活用し,商業レベルの音質を実現。

(* 文責: 富永鈴)

4.5.1.2 使用例

本研究では,以下のDAWソフトウェアを使用して楽曲や効果音を制作した。

1. Logic Pro

Appleが提供するプロフェッショナル向けのDAWで,直観的な操作性と高度な機能を兼ね備えている。

2. Studio One Professional

豊富なプラグインと柔軟な編集機能を持つ,音楽制作の幅を広げるDAWである。

3. Domino

シンプルなインターフェースで初心者でも扱いやすいMIDI編集ソフトである。

以上の3つのDAWを駆使して,以下のような作業を行った。

・MIDIファイル化

メロディや和音をMIDI形式で入力・編集し,楽曲の基盤を構築した。

・AI生成メロディを基にしたBGM制作

AIが生成したメロディを基に,楽曲の編曲や最終的な仕上げを行った。

・サウンドエフェクト(SE)の編集

効果音のデザインや加工にもDAWを活用した。波形編集やリバーブ・ディレイの適用により,音響表現を強化した。

(* 文責: 富永鈴)

4.5.1.3 魅力と活用の幅

DAWは,楽曲制作だけでなく,効果音(SE)の制作,ポッドキャストや映画音響の編集など,幅広い音響関連作業に対応可能であった。音楽や音響表現を追求するあらゆるクリエイターにとって,DAWは必須のツールといえる。

音響班は実践を通じて,DAWの多機能性と可能性を活かした高品質な作品制作を実現した。

(* 文責: 富永鈴)

4.6 映画音楽の評定

4.6.1 映画「IT」の音楽評定

上記した音楽要素12項目に基づいて映画「IT」の映画音楽について評定を行った。以下に実際の評定化結果を示す画像を掲載する。

時間	場面	1.振幅要素 (強弱)	2.アーティキュ レーション	3.ハーモニー
----	----	----------------	-------------------	---------

0:00	オープニング	丸い	レガート	単純調和(単音)
0:43 46:00	導入	丸い	レガート	複雑,伴奏あり
4:30	神のボート づくり	丸い	レガート	複雑,伴奏あり
5:13 24:00 35:00	本編の始まり, ボート流し	鋭い	レガート	複雑
9:00	腕食す場面	鋭い	レガート	なし
9:30 26:00	タイトル場面	丸い	レガート	単純協和(単音)
16:20 39:00 45:00	恋の芽生え	丸い	レガート	単純調和
18:20	現実直視	丸い	レガート	単純調和
22:43	フルートの女 性	鋭い	レガート	単純調和
26:25	風船と卵	丸い	レガート	単純調和
27:50	首無エッグ ボーイ	鋭い	スタッカート	単純調和
41:40	女の子髪切り	丸い	レガート	単純調和
43:00	水遊び	丸い	レガート	単純調和
49:55	エディ	鋭い	レガート	複雑/不協和音
53:00	女の子襲う	鋭い	スタッカート	複雑/不協和音
時間	場面	1.振幅要素 (強弱)	2.アーティキュ レーション	3.ハーモニー
1:20:25	エディ落下	丸い	レガート	複雑/不協和音
1:36:48	お父さん暴走	鋭い	レガート	複雑/不協和音
1:39:14	ベバリ 連れ去られた	丸い	レガート	複雑
1:44:36	浮かぶ子供達	丸い	レガート	複雑
1:47:27	マイク	鋭い	レガート	複雑/不協和音
1:52:15	ジョージを 見つける	鋭い	レガート	単音
1:57:30	戦闘	鋭い	レガート	複雑

2:02:52	皆が降りてくる	丸い	レガート	複雑協和
2:06:03	血	丸い	レガート	協和音

4.音の大きさの変動	5.旋律(メロディ)	6.調	7.音域	8.調性	9.リズム	10.テンポ
急速変動		短調	高音	最後わざと音外し	不規則	遅い
急速変動		短調	高音	調性的	不規則	遅い
急速変動		長調	高音	調性的	規則的	遅い
急速変動		長調	高音	調性的	規則的	ミドル
急速変動		短調	低高音	調性的	一定	遅い
急速変動		短調	高音	半音階的	一定	遅い
クレッシェンド		長調	高音	調性的	一定	遅い
クレッシェンド		長調	高音	調性的	一定	遅い
クレッシェンド		短調	高音	フルート: 調性的 声: 無調的	不規則	遅い
一定		長調	高音	最後ずれ	一定	遅い
4.音の大きさの変動	5.旋律(メロディ)	6.調	7.音域	8.調性	9.リズム	10.テンポ
一定(大きい)		短調	低音	調性的	規則的	早い
一定		短調	高音	調性的	一定	遅い
クレッシェンド		長調	高音	調性的	規則的	遅い
一定(大きい)		短調	低音	なし	規則的	早い
急速変動		短調	低音	調性的	不規則	ない
デクレッシェンド		短調	高音	無調的	一定	ミドル
クレッシェンド		長調	高音	調性的	不規則	ない
変化し続けている		長調	高音	調性的	一定	遅い

クレッシェンド		短調	高音	調性的	一定	遅い
クレッシェンド		長調	高音	調性的	一定	徐々に早く
クレッシェンド		長調	高音	調性的	一定	早い
クレッシェンド		長調	高音	調性的	一定	早い
クレッシェンド		長調	高音	調性的	一定	遅い
一定		長調	高音	調性的	一定	遅い

11.音色	12.展開	★心理要素
子供らしい高い声	伴奏,繰り返し	不安,恐れ,不快,恐怖
ピアノ	オープニング続き (キー違い)	穏やかだけど少し 不安,意味ありげ
ピアノ,シンセ	場面変化,展開	穏やか,安心感
ストリングス	場面に合わせたブレイク	ワクワク,前向き
ストリングス,ホーン,破裂音	繰り返し	恐怖,衝撃
子供らしい高音,ヴァイオリン,サウンドエフェクト	オープニング曲のリミックス	不安
オルガン	広がっていく	希望,誕生
ストリングス,ハーブ	途中で音数と音量が上がる	悲しい
11.音色	12.展開	★心理要素
フルート,ストリングス,子供の声	場面合わせ	恐怖,衝撃
オルゴール,ストリングス	場面に合わせて展開	不安定
ストリングス	繰り返し	焦燥感,緊張
ピアノ,ストリングス	繰り返し	悲しみ
ピアノ,コーラスっぽいストリングス	広がっていく	悲しい,少し希望
ドラム,ノイズ(ベース)	繰り返し,スライドで下がる	焦燥感,衝撃
ヴァイオリン,低いホーン	ヴァイオリンの高音ピッチ 上げ	恐怖
アコーディオン(ピッチずれ)	1小節のみ	恐怖,不安定,絶望的

ヴァイオリン	歪み	恐怖
ストリングス,ハープ	場面に合わせて展開	高揚,覚悟
ストリングス,コーラス,シンバル	場面に合わせて展開	絶望
ストリングス,ノイズ	場面に合わせて展開	緊張
ストリングス	繰り返し	緊迫
ストリングス,金属音,フルート	徐々に盛り上がる	興奮
ハープ,ストリングス	場面によって展開	安心
フルート,ピアノ	繰り返しながら展開	決意

以下に,各音楽要素12項目からわかる特徴を記載する.

1. 振幅要素

- ・丸い:
多くの場面で使用されている.穏やかな場面やリラックスした雰囲気表現の際に使用されていた.
- ・鋭い:
緊張感や驚きを表現する場面で使用されていた.

2. アーティキュレーション

- ・レガート:
ほとんどの場面で使用されており,滑らかで連続的な音の流れを示していた.
- ・スタッカート:
一部の場面で使用されており,断続的で鋭い音の流れを表現していた.

3. ハーモニー

- ・単純協和:
穏やかな場面や感情的なシーンで多く使用されていた.
- ・複雑/不協和音:
緊張感や不安を高めるために使用されていた.

4. 音の大きさの変動

- ・急速変動:
劇的な変化を伴う場面で頻繁に使用されていた.
- ・クレッシェンド/デクレッシェンド:
感情の高まりや沈静化を示す際に使用されていた.

5. メロディ

- ・順次:
いくつかの場面で見られ,穏やかな動きを示していた.
- ・跳躍:
劇的な変化を伴う場面で頻繁に使用されていた.

6. 調

- ・短調:
悲しみや不安を表現する場面で頻繁に使用されていた.
- ・長調:
前向きな感情や希望を表現する場面で使用されていた.

7. 音域

- ・高音:

多くの場面で使用されており,緊張感や不安を高める効果があった.

・低音:

特定のシーンで使用され,恐怖感を表現していた.

8. 調性

・調性的:

多くの場面で使用されており,一貫した音楽の流れを示していた.

・無調性的:

緊張感や予測不能な状況を表現するために使用されていた.

9. リズム

・一定:

多くの場面で使用されており,不安定感や規則性を示していた.

・不規則:

緊張感や予測不能な状況を表現するために使用されていた.

10. テンポ

・遅い:

多くの場面で使用されており,不安定感や規則性を示していた.

・早い/ミドルテンポ:

緊張感や動きを表現するために使用されていた.

11. 音色

・ストリング:

多くの場面で使用され,感情の深さやドラマチックな効果を出していた.

・ピアノ/フルート:

特定のシーンで使用され,特有の雰囲気や感情を表現していた.

12. 展開

・繰り返し:

一貫性や安定感を示す際に使用されていた.

・場面に合わせた展開:

場面の変化に応じて音楽が展開し,物語の進行をサポートしていた.

(* 文責: 横沢永遠)

4.6.2 映画「口裂け女」の音楽評定

上記した音楽要素12項目に基づいて映画「口裂け女」の映画音楽について評定を行った.以下に実際の評定結果を示す画像を掲載する.

時間	場面	1.振幅要素(強弱)	2.アーティキュレーション	3.ハーモニー
00:00	オープニング	丸い	レガート	複雑/不協和音
1:00	導入	丸い	レガート	単音
5:26	放課後	丸い	レガート	不協和音
16:00	美香ちゃん	丸い	レガート	複雑/不協和音
17:27	口裂け女登場	丸い	レガート	不協和音
23:00	ニュース	丸い	レガート	不協和音
27:00	先生帰り道	丸い	レガート	不協和音

29:00	男の子連れ去られる	鋭い	レガート	不協和音
47:30	お母さん逃亡	丸い	レガート	単音
58:00	赤い屋根の家	丸い	レガート	なし

4.音の大きさの変動	5.旋律	6.調	7.音域	8.調性	9.リズム
一定	順次	短調	全部	調性的	規則的
一定	跳躍	短調	低音	調性的	規則的
デクレッシェンド	跳躍	短調	高音	無調的	徐々に遅く
一定	跳躍	短調	高音	無調的	一定
一定	跳躍	短調	高音	無調的	規則的
一定	跳躍	長調	高音	調性的	規則的
強弱音	跳躍	短調	高音	無調的	規則的
一定	順次	短調	低音,高音	無調的	不規則
一定	跳躍	短調	低音	調性的	規則的
一定	一定	なし	高音	なし	規則的

11.音色	12.展開	★心理要素
ベース,シンセ,ヴァイオリン,風(ノイズ),トライアングル	繰り返し	不穏
オルガン,ヴァイオリン	繰り返し	不穏
ピアノ,ヴァイオリン	繰り返し	不安
ピアノ	繰り返し	同様
ピアノ,ヴァイオリン,シンバル	場面に合わせて展開	衝撃
ピアノ	繰り返し	恐怖,寂しい
ピアノ	場面に合わせて展開	疑惑
ノイズ,ベース,オルガン	場面に合わせて展開	恐怖
オルガン	繰り返し	恐怖

反響音,トンネル	繰り返し	恐怖
----------	------	----

以下に,各音楽要素12項目から分かる特徴を記載する.

1. 振幅要素

- ・丸い:
多くの場面で使用されており,穏やかな印象を与えていた.
- ・デクレッシェンド:
放課後の場面で使用され,徐々に静かになる効果を出していた.
- ・強弱音:
緊張感を高めていた.

2. アーティキュレーション

- ・レガード:
多くの場面で使用され,滑らかで連続した音の流れを作っていた.

3. ハーモニー

- ・不協和音:
劇的な変化や緊張感を伴う場面で頻繁に使用されていた.
- ・複雑/不協和音:
導入や,口裂け女登場の場面で使用され,不安定さを強調.
- ・単音:
単純さと静けさを示した場面と,急激な変化を表現する場面に使用されていた.
- ・なし:
静寂を表していた.

4. 音の大きさの変動

- ・一定:
多くの場面で使用され,安定感を示していた.
- ・強弱音:
ダイナミックな緊張感を演出する場面に使用されていた.

5. メロディ

- ・跳躍:
劇的な変化を伴う場面で頻繁に使用されていた.
- ・順次:
いくつかの場面で見られ,穏やかな動きを示していた.

6. 調

- ・短調:
多くの場面で使用され,悲しみや不安を強調していた.
- ・長調:
ニュースの場面で使用され,希望や安定感を示していた.

7. 音域

- ・高音:
多くの場面で使用され,緊張感や不安を強調していた.
- ・低音:
導入や男の連れ去られる場面で使用され,恐怖や不安を示していた.
- ・高低音:
広い音域をカバーし複雑な感情を表現していた.

8. 調性

- ・無調性的:
多くの場面で使用され,不安定さや近淳を示していた.
- ・調性的:
安定感や秩序を示していた.

9. リズム

・規則的:

多くの場面で使用され,安定感を示していた.

・不規則:

男の子が連れ去られる場面で使用され,緊張感や予測不可能性を示していた.

10. テンポ

・遅い:

すべての場面で使用され,静けさや落ち着きを強調していた.

11. 音色

・オルガン/ヴァイオリン:

導入の場面で使用され,荘厳な雰囲気を示していた.

・ピアノ/ヴァイオリン:

放課後の場面で使用され,2つの音色が重なり柔らかな音色に聞こえた.

・ピアノ:

口裂け女登場の場面で使用され,主にメロディとして使用されていた.

・ピアノ/ヴァイオリン/シンバル:

緊張感を高めていた.

・オルガン:

悲しみや不安を示していた.

・反響音(トンネル):

神秘的な雰囲気を作っていた.

12. 展開

・繰り返し:

多くの場面で使用され,一定のリズムやメロディを強調している.

・場面に合わせて展開:

感情や緊張感の変化を反映していた.

(* 文責: 横沢永遠)

4.7 MIDI

4.7.1 MIDIファイルとは

MIDIファイルとは,Musicak Instrument Digital Interfaceという音楽情報をデジタルでやり取るするための規格を利用して作成されたファイル形式のことである.MIDIは,1980年代に複数の電子楽器メーカーによって策定され,現在も音楽制作や演奏で広く利用されている標準規格だ.MIDIファイルは,実際の音声データを含まず,代わりに楽譜に基づいた演奏情報を記録する.

(* 文責: 冨永鈴)

4.7.1.1 特徴

1. 音声データを含まない軽量なファイル形式

MIDIファイルは,楽曲の音符(ピッチやタイミング),テンポ,強弱,使用する楽器(音色)など,音楽の演奏に必要な情報を数値データとして記録する.このため,MIDIファイルのサイズは非常に小さく,通常の音声ファイル(MP3やWAVなど)に比べてストレージ容量をほとんど消費しない.

2. 柔軟性と編集のしやすさ

MIDIファイルは,録音された音声ではなく,楽譜のような情報を記録しているため,あとからテンポや音色,音量,音符のタイミングを自由に変更することが可能だ.これにより,制作中の楽曲の調整やアレンジが簡単に行える.

3. 再生環境に依存した音色

MIDIファイル自体は音声データを含まないため,再生する環境や使用するソフトウェア/ハードウェアによって音色が変わる特徴がある.例えば,同じMIDIファイルでも,使用するシンセサイザーや音源モジュールによって異なるサウンドが再生される.

(* 文責: 横沢永遠)

4.7.1.2 用途

1. 音楽制作

MIDIファイルはDAWや電子楽器で読み込んで編集や再生が可能だ。作曲者は、メロディや伴奏のアイデアをおMIDIデータとして保存し、あとでアレンジや詳細な編集を行うことができる。

2. 演奏の自動化

電子ピアノやキーボードなどの楽器では、MIDIファイルを使用して自動演奏が可能。これにより、練習やパフォーマンスで役立つバックトラックを簡単に作成できる。

3. ゲームやデジタルコンテンツ

BGMや効果音の制作にも使用される。小さなファイルサイズで多様なサウンドを生成できる為、ストレージやメモリが限られた環境で特に有用だ。

(* 文責: 横沢永遠)

4.7.1.3 仕組み

1. 基本構造

MIDIファイルは、以下のような情報をデータとして含んでいる

- 音符のデータ: どの音をどのタイミングで、どのくらいの長さで演奏するかを指定
- テンポ情報: 楽曲の速度を制御
- 音量と強弱: 各音符の大きさやアクセントのつけ方を記録
- 楽器の指定: 使用する音色や楽器を設定(例: ピアノ, ギター, ドラム)
- コントロール情報: ピッチベンドやモジュレーションなど、演奏表現を制御

2. 標準MIDIファイル(SMF)

MIDIファイルには、楽曲全体の情報を記録する「標準MIDIファイル(SMF)」形式があり、ほとんどのDAWや電子楽器がこの形式に対応している。SMFには以下の3種類がある。

- フォーマット0: すべての演奏データが1つのトラックにまとめられた形式
- フォーマット1: 複数のトラックに分かれた形式(パートごとに独立)
- フォーマット2: 各トラックが独立して保存される形式(あまり一般的ではない)

(* 文責: 横沢永遠)

4.7.1.4 収集と作成

AIによるメロディの分析を実施するにあたり、WAVやMP3といった音声データ形式では適切な解析が困難であることが判明した。これは、これらの音源には主要なメロディに加えて様々な不随音が含まれており、正確に音階を抽出及び処理することが難しいためである。この問題を解決するために、収集した音源をMIDIファイル形式へ変換するプロセスを採用した。

音源の変換には、DAWソフトウェアであるLogic Proを使用した。音源は、どんが配信サイトYouTubeから収集し、以下の条件を基に選定を行った。

1. フリー音源
2. ゲームの背景音楽(BGM)
3. 映画のサウンドトラック

特に主要メロディがループする特徴を持つ音源に注目し、収集した音源を基に変換を実施した。変換の際、ループ部分は統一して2週分とし、音符ごとの強弱(ベロシティ)は一律に設定した。これにより、音源の均質性を保ちながらデータセットを作成することが可能となった。また、完成したMIDIファイルは、収集車の主観に基づいて以下の4つの感情カテゴリに分類した。

- 恐怖
- 楽しい
- 悲しい
- 安心

この感情分類は、AIによる解析及び学習時の参照データとして活用することを目的として行われた。これにより、音楽の感情的特徴に着目した行動な分析を可能にするデータ基盤が整備された。

このプロセスは、MIDI形式の特性を最大限に活用し、感情解析に適したデータセットの構築に貢献したものである。

(* 文責: 富永鈴)

4.8 BGM

4.8.1 ゲームBGMとは

ゲームのBGM(バックグラウンドミュージック)は、ゲーム体験を構成する重要な要素の一つだ。BGMは、ゲームの雰囲気やストーリー展開を強調し、プレイヤーの没入感を高める効果がある。また、感情の喚起や操作のガイドとしても機能し、プレイヤーに視覚情報だけでは伝わりにくいメッセージを補足する。以下に、BGMがもたらす具体的な効果や役割について詳しく説明する。

1. 雰囲気の形成

ゲームのジャンルやテーマに応じて、BGMは全体の雰囲気を形作る。例えば、ホラーゲームでは緊張感や不安感を高めるために低音域の持続音や不協和音が多用される一方、ファンタジーPG では壮大なオーケストラ音楽が使用されることが多い。これにより、プレイヤーはそのゲームの世界観に浸りやすくなる。

2. 感情の喚起とストーリーテリング

BGMはプレイヤーの感情を直接的に喚起する。特定の音楽が流れることで、キャラクターの悲しみ、喜び、怒り恐怖などの感情がより鮮明に伝わる。また、ゲームのストーリー振興に合わせた、BGMの変化は、プレイヤーの感情的な反応を導く。

3. 没入感の向上

BGMはゲームプレイ中の現実逃避感を高める役割を果たす。例えば、リアルな環境音やテーマにあった音楽が絶え間なく流れることで、プレイヤーは仮想世界に深く没入できる。

4. テンポと緊張感の調整

BGMのテンポやリズムは、プレイヤーの行動や心理状態を誘導する。例えば、戦闘シーンでは早いテンポの音楽が使用され、緊張感や興奮感を高める。一方で、探索パートでは穏やかなテンポの音楽が流れ、リラックスできる雰囲気を作り出す。

5. 記憶と感情の定着

ゲームのBGMはプレイヤーの記憶と強く結びつく。特定の曲が流れると、そのゲームで体験した出来事や蘇ることがある。この効果は、プレイヤーがそのゲームに再び戻ってきたい要因の一つだ。

6. プレイヤーへの操作ガイド

BGMは、ゲーム内の状況やプレイヤーの次の行動を音で示す場合もある。例えば、BGMが静かになったり変化したりすると、重要なイベントや数の出現を予感させる。

7. プレイヤー間の共有体験の形成

BGMはゲームコミュニティ内で共有される文化的なシンボルとしても機能する。特定の曲や音楽を聴くだけで、プレイヤー同士がそのゲームに関する思い出や感情を共有することができる。

8. 音楽ジャンルや音響技術の実験の場

ゲームBGMは新しい音楽ジャンルや技術を試す場ともなる。例えば、ゲームのインタラクティブ性を活かしたダイナミックBGM(プレイヤーの行動に応じてリアルタイムで変化する音楽)は、音楽の新しい表現方法を開拓する。

まとめとして、ゲームBGMは、単なる音楽以上の存在であり、ゲームの魅力を支える重要な要素だ。適切に設計されたBGMは、ゲームの面白さや安心感を何倍にも高める力を持っている。制作側としてもプレイヤーとしても、BGMの効果を意識することで、より深いゲーム体験を追求できる。

(* 文責: 横沢永遠)

4.8.2 AIにBGMを作成させるシステム

まず、学習データとして音高の順列データを、集めたMidiファイルから収集した。その後、集めた学習データからLSTMモデルを用いてAIに音高の順列を学習させた。学習させたのちに音が10個であるメロ

ディの元を人力で作成した。AIのメロディ自動生成でそのメロディの元を入力し、出力としてメロディの続きが作られるようにシステムを開発した。また、感情ごとにモデルの重みの計算方法を変え、それぞれの感情に合わせたメロディ生成を行えるよう作成した。

(* 文責: 早川暉人)

4.8.3 AIBGM

4.9 ゲームの効果音とは

4.9.1 SE(Sound Effects)とは

ゲームの効果音は、ゲームプレイをよりリアルで直観的なものにするために欠かせない要素だ。効果音は、キャラクターの動き、環境の変化、プレイヤーのアクションに連動して発生する音を目指し、ゲーム全体の演出やフィードバックを支える役割を担っている。以下に、効果音の役割やもたらす効果について説明する。

1. リアリティの向上

効果音は、ゲームの世界をよりリアルに感じさせるために重要だ。プレイヤーがゲーム内の物体やキャラクターとインタラクションするたびに、視覚情報を補完する音が鳴ることで、実際にその行動が行われているかのような感覚を生み出す。

2. 操作のフィードバック

効果音は、プレイヤーが起こった操作に対して即座に反応を返す役割を果たす。これにより、プレイヤーは自分の操作が正しく認識されたことを確認でき、ゲーム体験がより直観的になる。

・ボタン操作の音

メニュー画面で選択したときの「ピッ」という音や、ボタンを押した瞬間になる音は、プレイヤーに正しい操作を行ったという安心感を与えた。

・攻撃やスキル使用の音

敵に攻撃がヒットした際の「ズバツ」という効果音は、攻撃の手ごたえを感じさせ、ゲームの爽快感を向上させる。

3. 感情の喚起

効果音は、プレイヤーの感情を操作する重要な手段でもある。例えば、緊張感を高める音や、驚きを与える音は、視覚的な演出と組み合わせることで感情のインパクトを増幅させる。

・驚きの演出

急に鳴り響く不協和音や高音の効果音は、ホラーゲームなどでプレイヤーを驚かせるのに効果的である。

・成功感の強調

ミニゲームをクリアした瞬間やステージクリア時の「キラキラ」とした効果音は、達成感を増幅させる。

4. 状況の伝達

効果音は、プレイヤーにゲーム内の状況を音で伝える重要な情報源である。視覚的に捉えにくい変化やイベントを音で補完することで、プレイヤーの理解を助ける。

・敵の接近を知らせる音

敵が近づくと低音の効果音がだんだん大きくなることで、プレイヤーに危険を警告する。

・環境の変化

強風が吹く音は、ゲーム内の天候や時間の変化を伝える。

5. インタラクション演出

プレイヤーがゲーム内のオブジェクトとやり取りする際、そのアクションに連動した効果音が鳴ることで、ゲーム体験がより豊かになる。

・物体の移動や破壊音

荷物を探る「ガサガサ」という音や、ガラスが壊れる「パリン」という音は、物理的なリアリティを強調する。

・武器やアイテムの使用音

ナイフを使用した音は、武器の種類や威力を音で表現し、プレイヤーのアクションに重みを加える。

6. UI(ユーザーインターフェース)の補強

ゲームのUIにも効果音は重要だ。選択や操作ミス音を音で知らせることで、プレイヤーがスムーズにゲームを進行できるようになる。

・選択音

ポジティブな選択には明るい音、ネガティブな選択には低音が使用されるなど、音が操作の結果を直観的に示す。

・エラー音

無効な操作を試みたいときに短い「ブツ」という音が鳴ることで、何かが間違っていることを即座に理解させる。

7. ブランドや作品の個性の強調

特定の効果音は、ゲームのアイデンティティを強化させる。これにより、プレイヤーは音を聞いただけでそのゲームを思い出すことができる。

8. 没入感の向上

資格と聴覚の情報を統一することで、プレイヤーはゲームの世界に深く入り込むことができる。例えば、自然な生活音がリアルに再現されることで、架空の世界が現実のように感じられる。

9. プレイヤーの学習と誘導

効果音は、ゲーム内のルールやしくみを自然に学ばせる手段としても利用される。特定の行動が正しいかどうかを効果音で知らせることで、プレイヤーは繰り返しの中で学んでいく。

10. サウンドデザインの進化

効果音の品質は技術の進化とともに向上し、より複雑でリアルな表現が可能になっている。例えば、3Dオーディオ技術を活用した効果音は、音の方向や距離感を感じさせ、立体的な体験を提供する。

まとめとして、ゲームの効果音は、単なる音以上の役割を持ち、リアリティの向上、感情の喚起、フィードバックの提供、没入感の強化など、多岐にわたる効果をもたらす。優れた効果音は、プレイヤーに安心や達成感を与え、ゲーム体験そのものを豊かにする力を持っている。また、ゲームデザインにおける効果音の活用は、プレイヤーに忘れられない体験を提供するための重要な要素だ。

(* 文責: 横沢永遠)

4.9.2 SE作成プロセス

効果音(SE)の作成においては、独自の録音と既存の音素材を組み合わせたアプローチを採用した。以下に詳細を述べる。

(* 文責: 富永鈴)

4.9.2.1 独自音源

特定の効果音、例えば「ざわざわとした騒音」は、自ら音を立てて録音を行った。子の音源を基に、3種類の異なる音を重ね合わせ、必要に応じて編集を施すことで、臨場感やリアリティを追求した効果音を作成した。

(* 文責: 富永鈴)

4.9.2.2 外部音源

既存の効果音については、以下の信頼性の高い音素材提供サイトから適切な音源を引用した

- ・効果音ラボ: 強風、肉を切り裂く音、打撃音、ガラスが割れる音、物体破壊音、足音
- ・Pixabay: 蛇の音
- ・otologic: 骨折音、風船の音
- ・ONJIN: 荷物を探る音

これらの効果音は、対象のシーンに合わせた音響表現を強化するために用いられた。

(* 文責: 横沢永遠)

4.10 システム

4.10.1 開発環境

本研究において、プログラミング言語としてPythonを採用した。Pythonを選択した主な理由は、過去の研究やプロジェクト学習との引継ぎが容易であること、そしてその記述のサンプルからプログラミング初心者にも取得しやすい点である。Pythonは、多数のライブラリやフレームワークを備えた高水準のスクリプト言語であり、科学技術計算、データ解析、機械学習、Web開発など幅広い分野で利用されている。

開発環境には、Anaconda Navigatorから利用可能なSpyder(Scientific Python Development Environment)を用いた。この開発環境を選択した理由は、以下の3点に集約される。

1. Python専用の統合開発環境(IDE)である点:

Spyderは、Pythonを主に使用する化学技術系の開発者向けに設計されており、変数エクスポーラや統合デバッガーなど、効率的なプログラミングを可能にする機能を備えている。

2. 学習及びコミュニケーションの利便性:

昨年度の研究に関与した先輩方から直接指導を受けやすく、加えて班員同時に知識や経験を共有するうえで、Spyderのシンプルなインターフェースが役立った。

3. ライブラリや仮想環境の管理の容易性:

SpyderはAnaconda Navigatorと統合されており、必要なライブラリの追加や仮想環境の構築が直観的な操作で行える。そのため、班員全体で同一の環境を構築し、作業を効率化することが可能であった。

加えて、Anaconda Navigatorの利用により、NumPyやPandas, Matplotlib, Scipyといった主要なライブラリを容易にインストール・管理することができ、Pythonによる開発作業を効率的に進めることができた。これらのライブラリは、数値計算やデータ処理、グラフ描画、信号処理などの機能を提供するため、音響班の作業においても非常に有用であった。

以上のような環境を整備した結果、班員間でスムーズなコミュニケーションと情報共有が実現し、プロジェクト全体の進捗を効率的に進めることが可能となった。Python及びSpyderを採用することは、過去の研究との連携やチーム内での作業効率を向上させるための合理的な選択であったと考えられる。

(* 文責: 早川暉人)

4.10.2 ライブラリ

本プロジェクトで利用した主なライブラリについて以下に説明する。

(* 文責: 横沢永遠)

4.10.2.1 os

osライブラリは、ファイル操作やディレクトリ操作を可能にする標準ライブラリである。Midiファイルの処理において、ディレクトリ内のファイルを動的に探索し読み込むために使用した。これにより、大量のMidiファイルを効率的に処理できる柔軟なファイル管理を実現した。

(* 文責: 早川暉人)

4.10.2.2 pretty_midi

pretty_midiは、MIDIデータの詳細な解析や操作を簡便に行うライブラリである。テンポ、音符情報、楽器の特定など、音楽データに関する高度な解析を行う際に利用した。このライブラリの活用により、音楽構造の理解を深めるとともに、音楽生成におけるカスタマイズ性を向上させた。

(* 文責: 早川暉人)

4.10.2.3 csv

csvライブラリは、構造化データの保存及び読み込みに利用される。MIDIファイルから収集した特徴量データをCSV形式で保存し、ほかの解析ツールやライブラリで再利用可能にした。

(* 文責: 早川暉人)

4.10.2.4 pandas as pd

NumPyは、効率的な数値計算を可能にするライブラリであり、多次元配列の操作や大量データ処理を行う際に活用した。プロジェクトでは、データ前処理や保存プロセスにおいて、NumPy配列を用いて高速なデータ操作を実現した。

(* 文責: 早川暉人)

4.10.2.5 numpy as np

Pandasは、データフレームを利用したデータ操作や解析を容易にするライブラリである。MIDIデータから抽出した特徴量を統一フォーマットで管理し、視覚的かつ効率的なデータ解析を可能にした。また、データの保存形式としてCSVファイルを選択し、永続化を実現した。

(* 文責: 早川暉人)

4.10.2.6 tensorflow as tf

TensorFlowは、ディープラーニングの構築と実行を可能にするライブラリである。本プロジェクトでは、音楽生成モデルを構築するために使用し、LSTMネットワークの実装に特価した部分を中心に活用した。

(* 文責: 早川暉人)

4.10.2.7 sklearn.metrics import accuracy_score

音楽生成モデルの評価において、正確性を測定するために用いた。

(* 文責: 早川暉人)

4.10.2.8 tensorflow.keras.models import Sequential

Sequential APIを用いて、LSTM層や全結合層を組み合わせた音楽生成モデルを構築した。柔軟なネットワーク設計が可能で、複雑なデータ構造にも対応できた。

(* 文責: 早川暉人)

4.10.2.9 tensorflow.keras.optimizers import Adam

Adamアダプティマイザを用いてモデルの最適化を行い、収束速度の向上を図った。

(* 文責: 早川暉人)

4.10.2.10 tensorflow.keras.layers import LSTM, Dense Unput, Masking

LSTM層は、時系列データであるMIDI情報の学習に用い、Dense層は出力結果の調整に利用した。また、Masking層を用いて可変長のシーケンスデータを効率的に処理した。

(* 文責: 早川暉人)

4.10.2.11 tensorflow.keras.losses import MeanSquaredError

音楽生成モデルの損失関数として利用し、生成データと目標データの差異を最小化するために用いた。

(* 文責: 早川暉人)

4.10.2.12 sklearn.preprocessing import StandardScaler

MIDIデータの標準化処理に利用し、モデルの学習を安定化させた。

(* 文責: 早川暉人)

4.10.2.13 mido import MidiFile, MidiTrack, Message

Midoライブラリを活用し、MIDIファイルの生成や解析を実施。特に、MidiTrackやMessageオブジェクトを操作し、テンポやノートのカスタマイズを行った。

(* 文責: 早川暉人)

4.10.2.14 tensorflow.keras.models import load_model

事前に学習済みの音楽生成モデルをロードし、効率的に推論を行った。

(* 文責: 早川暉人)

4.10.3 ニューラルネットワークの基礎

ニューラルネットワークは、人間の脳を模倣したモデルであり、神経細胞(ニューロン)が電気信号を伝える仕組みをシミュレーションしている。ニューロンは複数の入力を受け取り、その重み付き和を計算し、活性化関数を通じて発火(出力)するかを決定する。

代表的な活性化関数には以下のものがある:

- ・シグモイド関数: 値を0~1に収束させる。
- ・ReLU関数: 計算が単純で高速、大規模データに適している。
- ・Softmax関数: 出力を確立として扱い、分類問題で活用される。

(* 文責: 横沢永遠)

4.10.4 学習モデル

学習モデルは、特定のタスクを実行するためにデータからパターンや規則を学ぶ計算機アルゴリズムのことを指す。これらにより、未知のデータに対して予測や分類を行うことができる。まず、モデルを訓練するためのデータを収集した。データから有用な音楽直腸量を抽出し、モデルの性能を向上させることができた。

(* 文責: 早川暉人)

4.10.5 LSTM(Long Short-Term Memory)

4.10.5.1 LSTMとは

LSTMは、リカレントニューラルネットワーク(RNN)の一種であり、時系列データのように順序があるデータの学習に特化した構造を持つモデルである。RNNは、過去の情報を保持して次の出力に反映できるため、自然言語処理や音声認識、音楽生成などの分野で活用されている。しかし、通常のRNNには「勾配消失問題」という問題があり、長いシーケンスを学習する際に過去の情報が徐々に失われてしまう。この問題を解決するために考案されたのがLSTMだ。

(* 文責: 横沢永遠)

4.10.5.2 構造

LSTMは、標準的なRNNを改良したもので、「セル状態(Cell State)」と「ゲート(Gate)」と呼ばれる独自の構造を持っている。この仕組みにより、重要な情報を長期的に保持し、不必要な情報を削除することが可能である。

1. セル状態(Cell State)

セル状態は、時系列データを処理する際の「記憶」の役割を果たす。時系列全体にわたり重要な情報を流すバックボーンとして機能し、必要に応じてゲートを通じて情報を調整する。

2. ゲートの種類と役割

LSTMには以下の3つのゲートがあり、それぞれ異なる役割を持つ。

・入力ゲート(Input Gate)

新しい情報をどれだけ受け入れるかを決定する。現在の入力と過去の状態を基に、セル状態に追加する情報を調整する。

・忘却ゲート(Forget Gate)

過去の情報をどれだけ忘れるかを制御する。これにより、長期的に不必要な情報を効率的に削除できる。

・出力ゲート(Output Gate)

セル状態に保持している情報から、どれだけ出力するかを調整する。このゲートにより、次のニューロンに伝える情報を適切に選別する。

(* 文責: 横沢永遠)

4.10.5.3 利点

LSTMは以下のような特徴的な利点を持ち,RNNに比べてより強力な性能を発揮する:

- ・長期的依存関係への対応:
勾配消失問題を起きにくいいため,長い文章や音楽のような長期的な依存関係を持つデータを処理するのに適している.

- ・効率的な情報保持:

忘却ゲートと入力ゲートにより,不必要な情報を削除しながら重要な情報を長期間保持できる.

(* 文責: 横沢永遠)

4.10.5.4 応用と活用

LSTMは,時系列データを扱う様々な分野で活用されている.以下に例を3つ挙げる.

- ・自然言語処理: 文脈を保持しながら文章を生成する.
- ・音声認識: 連続する音データから文字列を生成する.
- ・音楽生成: 時間的な関係性を保ちながら新しいメロディを創造する.

(* 文責: 早川暉人)

4.10.5.5 試用部分

LSTMは今回の学習モデルとして利用した.LSTMモデルでは音高の順列データを入力データとして利用し,その一瞬一瞬の音から次の音を予測できるように学習させた.また,それぞれの感情ごとに分け,LSTMを学習させた.LSTMの学習で忘却度合いを設定するために,独自のloss関数を設定した.ここでは,伴奏と最後の伴奏をより学習するように重みを増やしたloss関数を作成し使っている.

(* 文責: 早川暉人)

第5章 結果

スケジューリングの関係で,すべて人力で制作するものとAIが生成したメロディを使用して制作するものに分けて作曲した.楽曲の雰囲気がゲームにマッチするように,伏線班と密にコミュニケーションを取るよう心掛けた.完成形の前にデモ音源を作り,曲調の確認を行い修正,改善を行った.ここでは完成したそれぞれの曲を紹介する.

(* 文責: 冨永鈴)

5.1 BGM作成

5.1.1 悲しい&不穏&若干恐怖

目覚め: 異世界で目を覚ますシーンで流れる.悲しい,不穏,若干の恐怖という要素を基に,新しい世界で目を覚ましたことを表現するため,広がりのある展開を意識した.

(* 文責: 冨永鈴)

5.1.2 謎&困惑&少しの恐怖

アザミのテーマ: アザミとの会話のシーンで流れる. 謎,困惑,少しの恐怖という要素を基に,あえて不協和音を使用し不穏な感じを表現した.

(* 文責: 冨永鈴)

5.1.3 歓喜

薬の完成: アザミの好感度が高い状態で薬が完成したシーンで使用. 歓喜を表現できるよう,Cメジャースケールの4つのメロディを重ね,不協和音のないきれいな印象の作曲を心掛けた.

(* 文責: 横沢永遠)

5.1.4 オルゴールバージョンの歓喜&悲しい

薬の完成オルゴールバージョン: レイとアザミが静観するシーンで使用. 歓喜の中に少しの悲しさが含まれる演出をするため,音数を減らしさみしさが残る作曲を行った.

(* 文責: 横沢永遠)

5.1.5 恐怖&淡々

異世界の探索: マップ上を探索するシーンで使用. 淡々とした恐怖やグロテスクという要素を基に, 本作のテーマである変身というキーワードに沿った変化のある展開を意識した.

(* 文責: 富永鈴)

5.1.6 恐怖&激しい&蹂躪

終わりなき浸食: 主人公の狂気があふれ出て, 周囲のものを破壊するシーンで使用. 恐怖, 激しい, 蹂躪という要素を基に攻撃的な音選びを行い, 主人公の変化を表現するためにリズムや音数が変化する展開を意識した.

(* 文責: 富永鈴)

5.1.7 タイトル曲

還道: タイトルで流れる. 悲しい感情を表現できるメロディを意識し, あえて音数を減らしたシンプルな楽曲にした.

(* 文責: 横沢永遠)

5.2 AIBGM作成

5.2.1 恐怖&グロテスク&激しい&蹂躪

歪んだ認識: レイがバケモノになったヒカルの幻覚を見るシーンで使用. AIが生成したメロディは区切り方やのびし方を整えてから使用した. 富永が恐怖, グロテスク, 激しい, 蹂躪という要素を基に, 幻覚を見る際の気持ちの悪い感覚を表現できるような編曲を行った.

(* 文責: 富永鈴)

5.2.2 葛藤&恐怖&焦燥

テッセンのテーマ: テッセンというバケモノと会話をするシーンで使用. AIが生成したメロディは区切り方を整え, 繰り返しの回数を変えてから使用した. 葛藤, 恐怖, 焦りという要素を基に, 一曲の中で拍子が変わり, 違和感や不思議な感覚を与えられる変拍子を使い, 焦燥感のある編曲を行った.

(* 文責: 富永鈴)

5.3 成果

AIが生成したメロディを使うと一貫して, 今までにない技術や引き出しを使うことになり, 全く新しい曲を作ることができた. 怖い曲は急激な音高の変化や調からの逸脱など, 曲調が一貫しているためAIが生成したメロディとの相性がとても良いと感じた. 楽しい曲は怖い曲に比べてより音楽理論に乗っ取った曲にする必要があるのに加え, 曲の特徴やジャンルが様々で, 一概に定義するのが難しいといえるため, AIによるメロディの生成は難しく相性が悪いと感じた.

(* 文責: 富永鈴)

第6章 考察

メロディの生成に特化した学習モデルは怖いメロディを予測しやすい傾向にあると考えられる. 楽しいと怖いという感情を呼び起こす曲の形をそれぞれで作成し, それを元にして楽曲生成をAIに行かせた結果, AIにおける怖いという感情に沿った楽曲生成は人が聞いて怖いという感情を引き起こすことができた. 一方で楽しいという感情に沿った場合, 実際に聞いた際に楽しいという感情を生み出すことはなかった. これについて考えられることとして, 怖いという感情を呼び起こす曲はほとんどの曲の形が同じであるため, 予測がしやすかったと考えられる. 一方で楽しいという感情を呼び起こす曲の形は様々であるため, 学習モデルが曲の形を予測しづらいことが原因で生成が難しくなったと考えられる. そのため, AIに楽しい曲を生成させるにはデータ量を増やし, 曲の形を細かく学習できる必要がある. また, 最初の伴奏や最後の伴奏など, 同じ感情を呼び起こす曲にも始まり方や終わり方などに特徴がある. そのため, 曲ごとの特徴となる部分の音高に注目して学習モデルの重きを調整することが必要であると考えられる. その他にも, 今回は音高という要素に絞ってデータの学習を行ったため, 学習モデルの精度を上げるために他の音楽要

素として音の長さや音の大きさも取り入れるべきであると考えられる。それにより、より曲の形を細かく認識でき、それぞれの曲の違いをはっきりさせることで予測がより正確になるからである。

(* 文責: 早川暉人)

第7章 参考文献

- Boersma, Paul & Weenink, David (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4.26, retrieved 8 January 2025 from <http://www.praat.org/>
- Furukawa S (2016). Processing of temporal information in the auditory system. *Audiology Japan*, 59, 615–622.
- Meinel, L. S., & Bullerjahn, C. (2022). More horror due to specific music placement? Effects of film music on psychophysiological responses to a horror film. *Psychology of Music*, 50(6), 1837-1852.
- P.N.ジュスリン, & J.A.スロボダ (編).大串健吾,星野悦子,&山田真司(監訳). (2008). 音楽と感情の心理学. 誠信書房
- Russell, J. A. (1980). "A circumplex model of affect." *Journal of Personality and Social Psychology* 39: 1161-1178.

全学計算機システム

[https://www.u.tsukuba.ac.jp/~hirai.akiyo.ft/meeting14/SLAA\(praat_HO\)\(140417\).pdf](https://www.u.tsukuba.ac.jp/~hirai.akiyo.ft/meeting14/SLAA(praat_HO)(140417).pdf)