

音楽感情認識に基づく聴覚障害者向け 音楽視覚提示システムの構築



国立大学法人

筑波技術大学

産業技術学部
(大川ゼミ)

武田 晴彦、 大川 学

1 研究背景

音楽感情認識 (MER)

- **音楽**
文化や言語を超えて人類共通の芸術表現が可能、
聴き手に豊かな感情を提供する
- **聴覚障害者での音楽**
音楽は聴覚を前提とした表現であり、
聴覚障害者にはその感情的側面の享受が困難
- **音楽の感情表現**
エンターテインメント分野（映画・ゲームなど）や、
感情の可視化・分析などの技術応用に活用されている

2 研究目的

【課題】

- 音楽は言語化困難、字幕では「♪」記号表示のみ
⇒「楽しい曲」「悲しい曲」？ 情景伝達も困難
- ピアノなどの楽器演奏では歌手表情・歌声情報が
欠落するため、さらに感情把握が困難

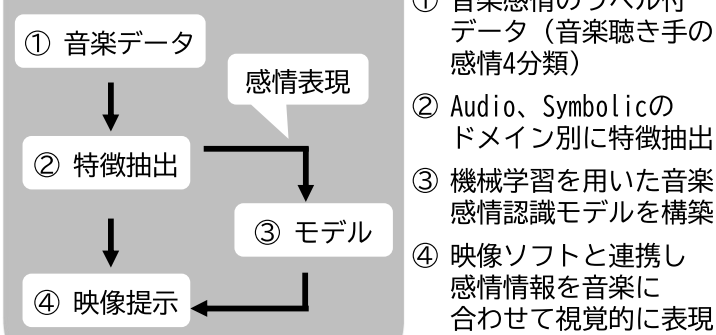


【対策】

- 音楽を「聴いて楽しむ」から「見て楽しむ」へ
⇒ 聴覚障害者の新たな鑑賞体験の創出を目指す
- 音楽を「聴いて作る」から「見て作る」へ
⇒ 聴覚障害者の新たな音楽創作支援を目指す

3 提案手法

【システム概要】



ドメイン	内容	特徴	モデル
Audio	音響情報	MFCC(メル周波数ケプストラム係数)	CNNベース
Symbolic	楽譜的記号情報(演奏構造)	音符長、ベロシティ、拍密度、キーなど	Transformerベース

4 実験方法

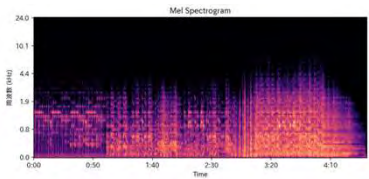
- データセット：EMOPIA [1]
 - 感情ラベル付きのポップピアノ音楽データセット
 - 1,087の音楽クリップ（アニメ、ポップ、映画サウンドトラック、個人制作楽曲など）
 - 感情ラベルは、ラッセルの円環モデルに基づき、Valence（快・不快）とArousal（覚醒度）の4象限分類
- 構築した機械学習モデルからArousalとValenceを抽出、音楽の流れに合わせて音楽感情をリアルタイムに可視化

[1] H. Hung et al., "EMOPIA: A Multi-Modal Pop Piano Dataset for Emotion Recognition and Emotion-Based Music Generation," in Proc. ISMIR, 2021.

象限	感情カテゴリ	Valence (快・不快)	Arousal (覚醒度)
Q1	Positive-Energy	高(ポジティブ)	高(活力あり)
Q2	Negative-Energy	低(ネガティブ)	高(緊張・怒り)
Q3	Negative-Calm	低(ネガティブ)	低(沈静)
Q4	Positive-Calm	高(ポジティブ)	低(穏やか)

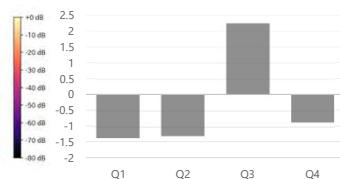
5 実験結果

音楽特徴量の可視化（例）



サンプル音楽のメルスペクトログラム

感情分類の結果（例）



サンプル音楽の4象限分類

6 まとめ

- 音楽感情認識に基づく聴覚障害者向け視覚提示の構築を目指し、現在までに提案手法のプロトタイプを作成
- 今後の展望：
 - 動的な視覚表現への拡張
 - ユーザ評価の実施による有効性検証
 - 音楽作成支援への応用可能性の検討

【謝辞】

- 本研究発表の一部は、JSPS 科研費 JP 25K15107の助成によります。
- 本研究にご協力いただいた関係者の皆様に、心より感謝申し上げます。

【問い合わせ先】

実施責任者

■ 筑波技術大学 産業技術学部産業情報学科 准教授 大川 学

■ Email: mnb_okawa7[at]a.tsukuba-tech.ac.jp (※ [at]は@に置換えてください。)