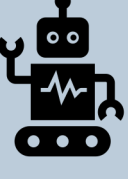


人間の感性に迫る！

AIがつくる豊かなピアノ演奏

Jingjing Tang[#], Xin Wang^b, Zhe Zhang^b, 山岸 順一^b, Geraint Wiggins[#], George Fazekas[#]
Queen Mary University of London
b 山岸研究室

どんな内容？

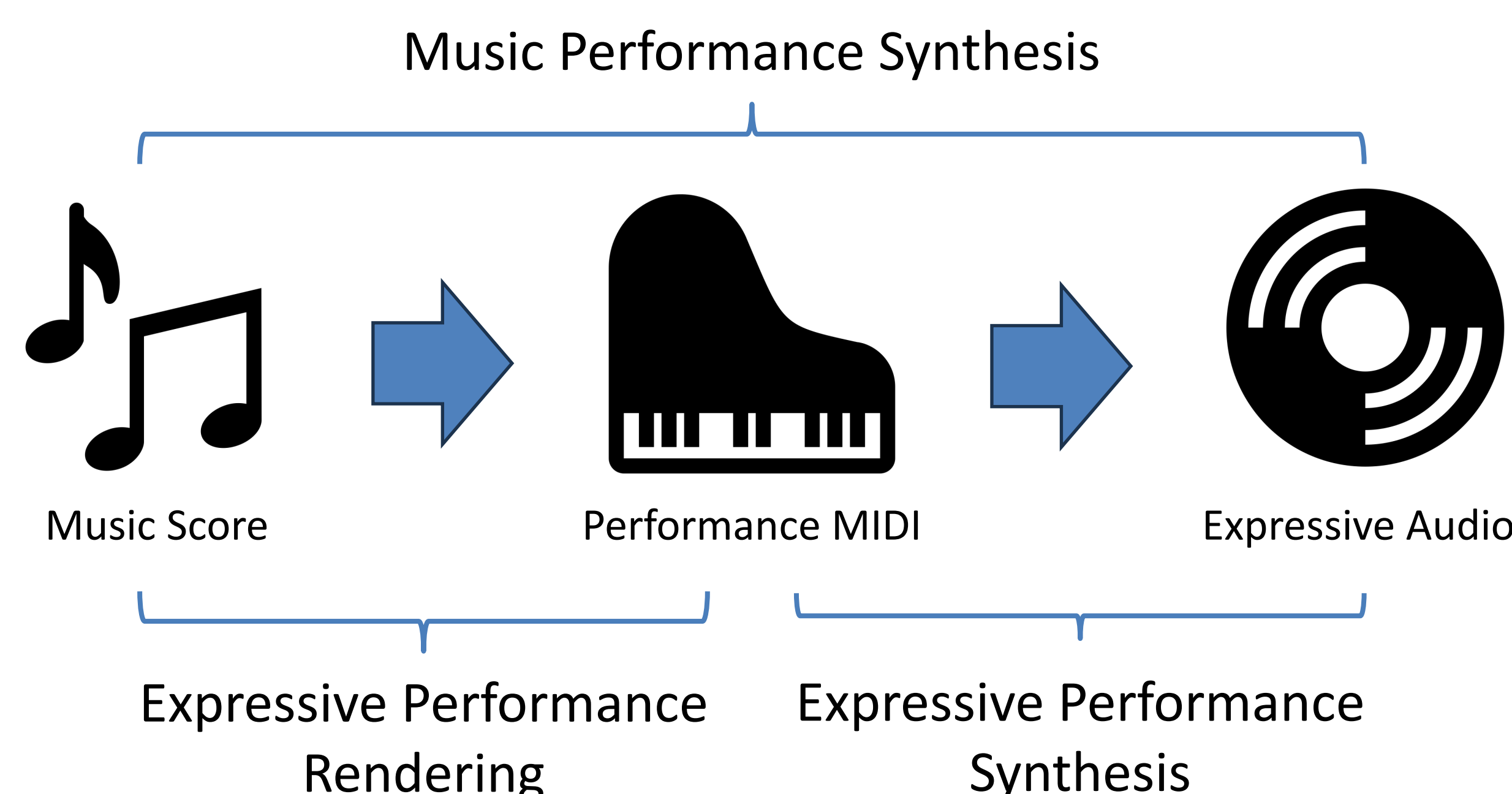
- 言語モデルを活用したピアノ演奏生成AI『MIDI-VALLE』の開発。
- MIDIと音声をトークン化し、音楽情報を言語的に扱うことで多様な表現を学習。
- AIピアノ演奏サンプルを聞いてみよう～

何を検討している？

- 言語モデルを音楽表現に適用する方法の開発。
- 多様な音楽ジャンルや録音環境への汎化能力。
- 人間らしい表現力の再現性向上。

背景・目的

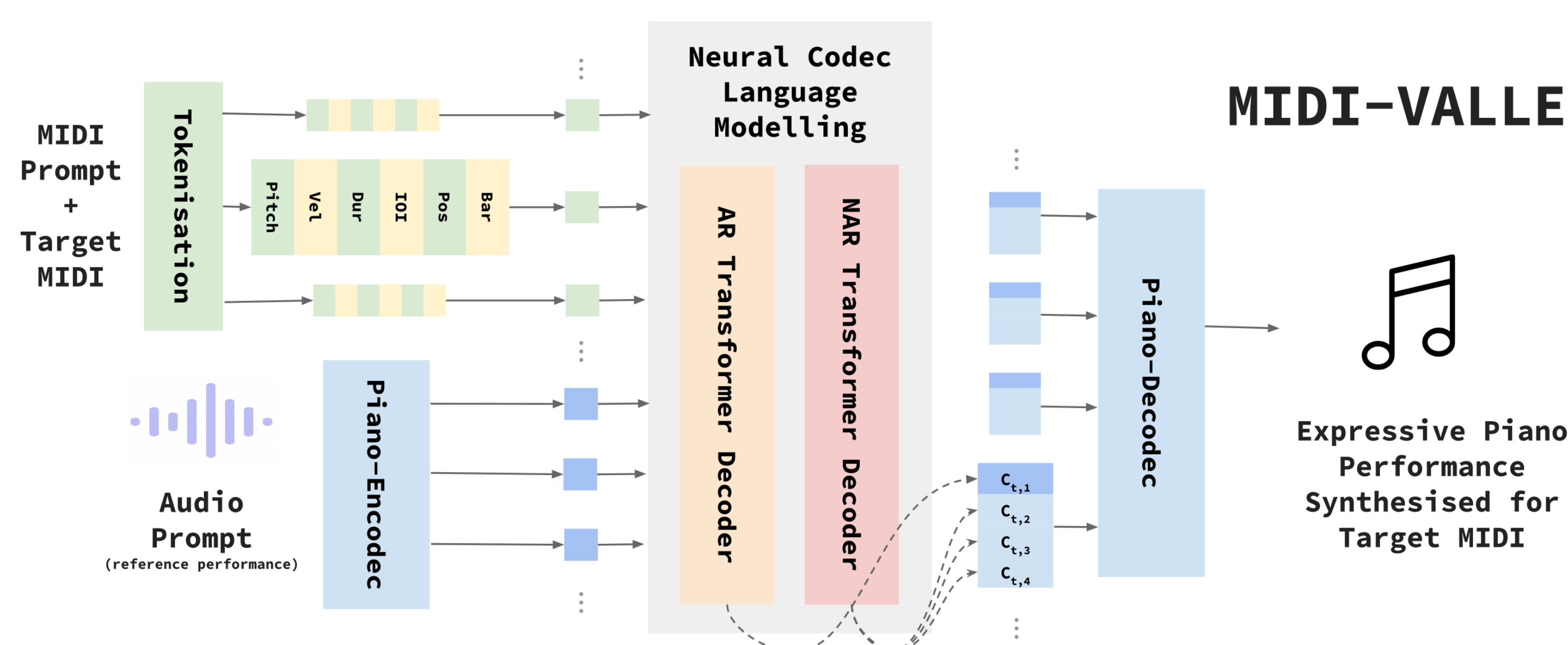
- ピアノ演奏は、楽譜の音符だけでなく、強弱やテンポの揺れ、音色や響きで聴き手を引き込む。
- 従来のAIは特定のジャンルや録音環境に限られ、多様な条件下で豊かな表現を再現できない。
- 本研究では、さまざまなスタイルや環境でも人間らしい演奏を生成するAI技術を目指す。



研究内容（方法・結果・結論）

方法

- MIDIデータと音声データをトークン（言語的単位）としてエンコード。
- 幅広いジャンルの音楽と多様な録音環境で収録されたデータを用いて学習。
- 短い参照音源をもとに環境特性に適応可能。

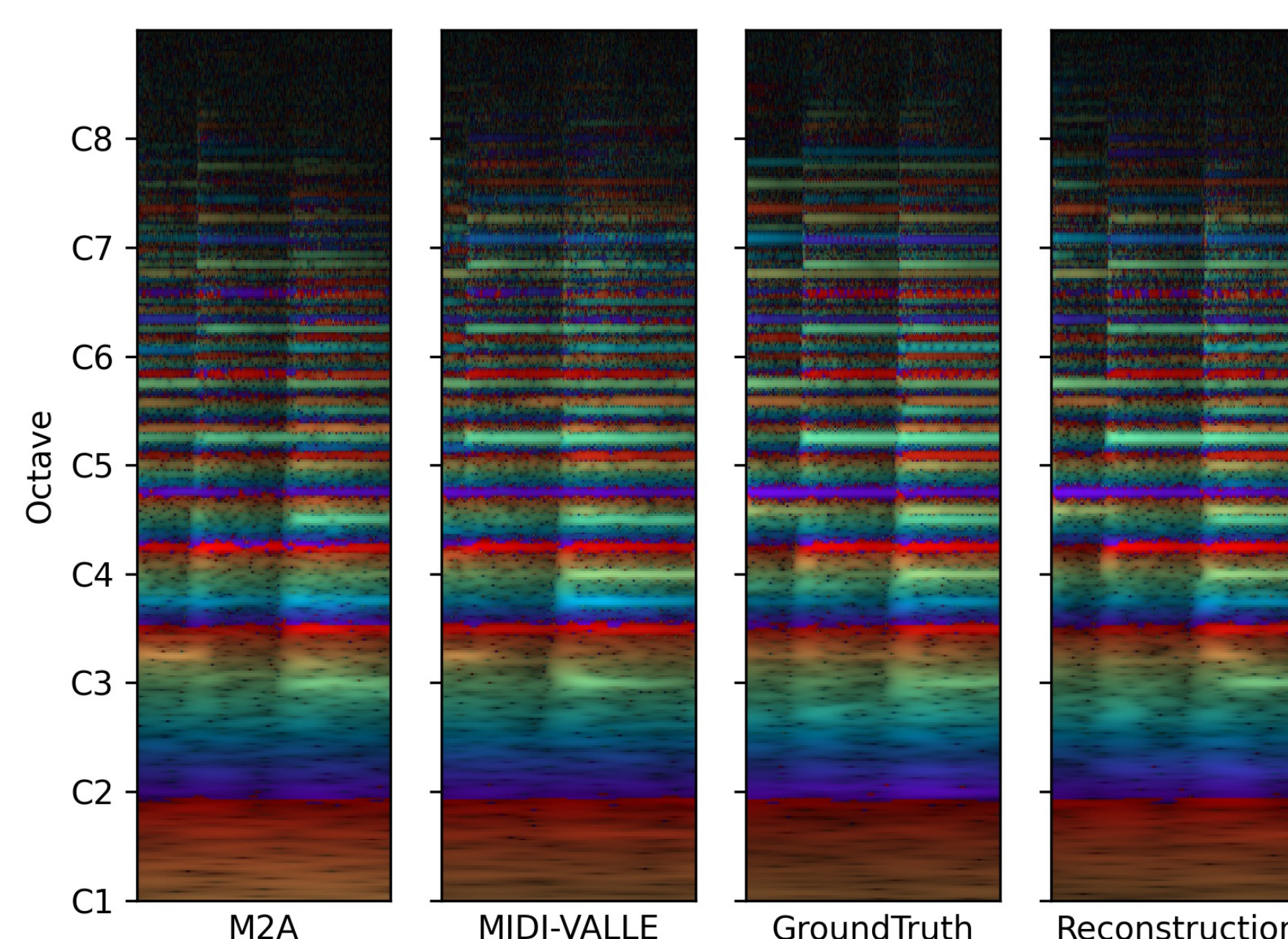


利用したデータセット

- ATEPP：多様なクラシックピアノ演奏データを含む。
- Maestro：国際ピアノコンクールの演奏を収録した高品質なデータ。
- Pijama：ジャズのピアノソロをライブおよびスタジオ環境で収録したデータ。

結果

- 主観評価（リスニングテスト）において従来のAIモデルを大きく上回る表現力を達成。
- Fréchet Audio Distance (FAD)において従来モデル比75%以上の性能向上を実現。



結論

言語モデルの応用により、クラシックからポップまで幅広いジャンルで人間並みの表現力を再現。

現在取り組んでいる事

- 幅広い音楽ジャンルへの汎化能力のさらなる改善。
- MIDIと音声間の精度と整合性を高めるための技術開発。
- 新しい音楽スタイルへの対応力向上を目指したモデルの拡張。