

画像処理AIによるリアルタイムMRI動画のアップスケールの試み

藤木祐輔（甲南大）、北村達也（甲南大）

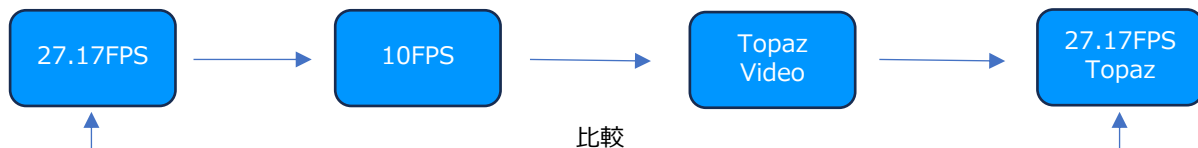
1. 背景と目的

- リアルタイムMRIは音声生成過程の研究に有用
- しかし、現行技術では動画のフレームレートは**数十FPS**に制限
- 画像処理AIでフレームを正確に補間できるか検証し、現在の計測技術では把握できない時間分解能で調音運動を観測できないか？

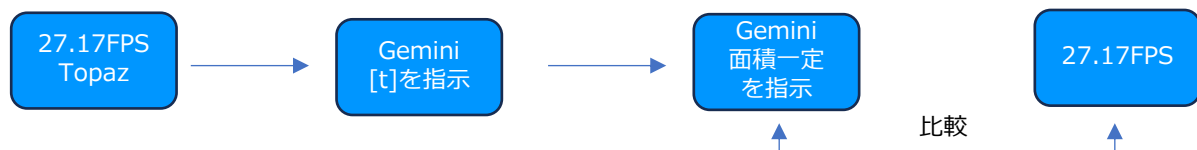


2. 研究方法

- 歯茎破裂音の前後フレームから歯茎破裂音のフレームを生成できるか調査する
- STEP 1: フレームレート (FPS) の調整
rtMRI動画のFPSを一度下げ、再びTopaz Videoで向上、この2つの動画を比較



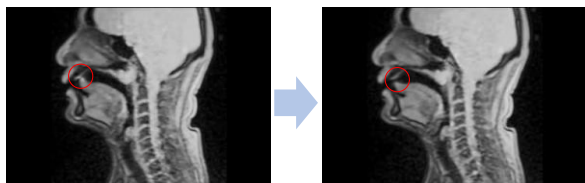
- STEP 2: 生成AI(Gemini)への指示と出力
発話時に舌尖と歯茎が接触する([t]とする)ことを指示。
Geminiの出力を安定化させるために、二次元画像上の舌の面積を一定にするように指示
元動画で舌尖と歯茎が接触したフレームとGeminiで作成した[t]の状態のフレームを比較



3. 実験結果

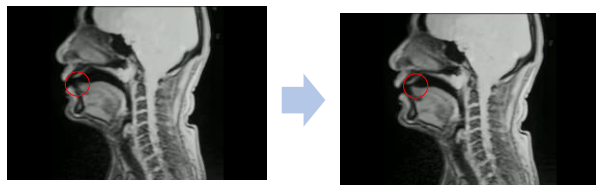
- 1.Topaz Videoによるフレーム補間:rtMRI動画のフレームをTopaz Videoで補間したが、元のrtMRIに存在しない中間フレームの生成は**困難**であった

元動画から切り取った2フレーム



舌尖と歯茎が接触するフレームが存在している。

Topaz Videoで再びFPSを向上させた動画から切り取った2フレーム



舌尖と歯茎が接触するフレームは生成されない。

- 2.Geminiによる画像生成と特徴再現:
舌尖と歯茎の接触を指示条件として Geminiに適用
生成された画像は、**元動画の特徴** (舌の形状、接触の有無など)を一部 **再現可能**であった

元動画で舌尖と歯茎が接しているフレームとGeminiで作成した、舌尖と歯茎が接触するよう指示した画像



比較



4. 結論と今後の展望

- 本研究は、画像生成AI (Gemini)がrtMRI発話動画のフレーム生成において有効である可能性を示した
- しかし現在の画像生成AIは、生物学的な筋肉の物理シミュレーションを行って舌形状を計算したわけではなく、あくまで前後の画像フレームのピクセル情報と、学習した[t]の形状パターンを組み合わせで生成したもの
- 今後の課題として、複数の話者や異なる発話タスクに適用しAI生成動画の有用性を検証する必要がある