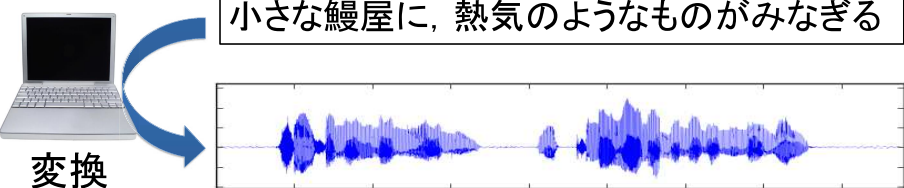


コンピュータがあらゆる声を合成！

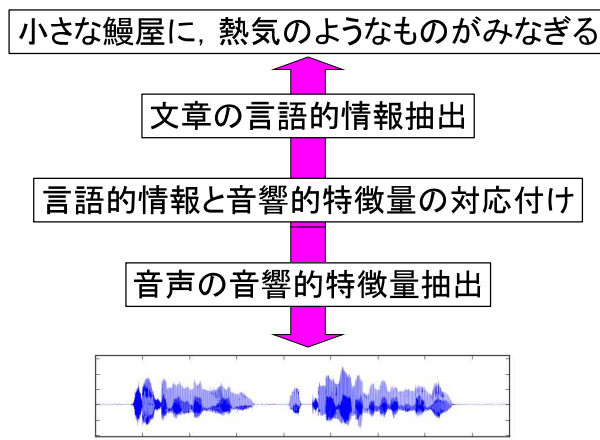
• テキスト音声合成 (TTS: Text-to-speech)



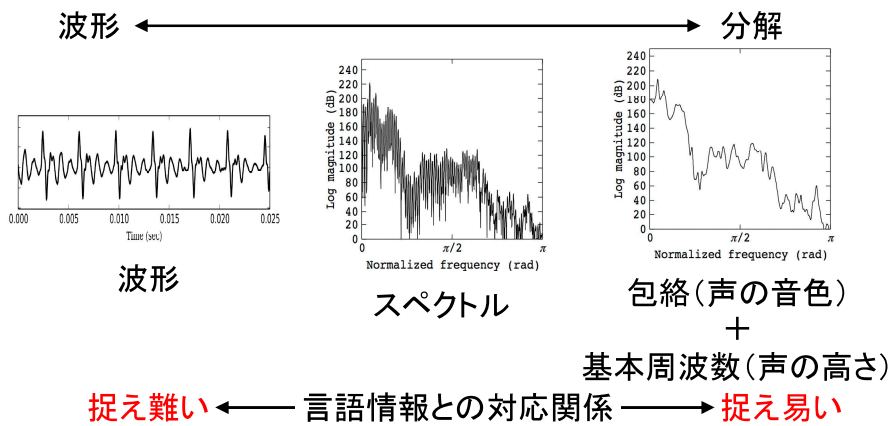
- テキスト音声合成の利用方法は限定的
 - カーナビ, Siri, Pepper, ナレーション合成, 歌声合成
 - 合成困難な声が存在 (例えば, 舞台の声や特殊な歌唱)

現在の問題の解決：性能不足の解消
 未来の価値の創出：表現力の向上

• テキスト音声合成の基礎



• 音声の特徴量

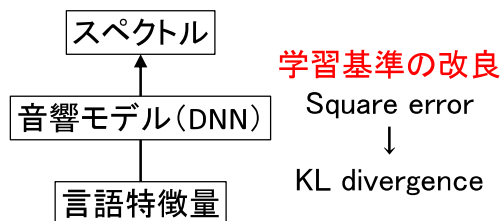


- 従来手法：音声を分析 (分解) して利用
 - 分析が難しい声が存在
 - 分析に失敗すると...
 - 音声合成器に大きな影響
- 音声波形やスペクトルを直接利用
 - 扱いが難しい
 - 機械学習による解決

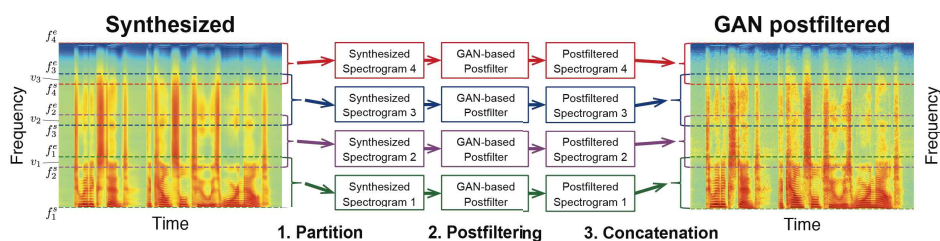
あらゆる音声を高品質に合成可能に
 → TTSを様々な場面で利用可能に

• 研究の詳細

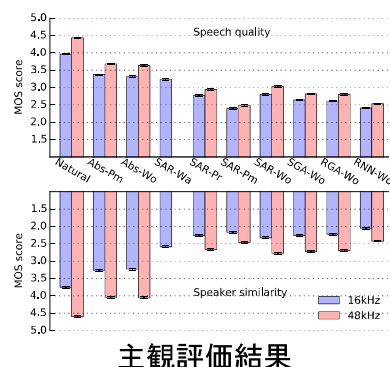
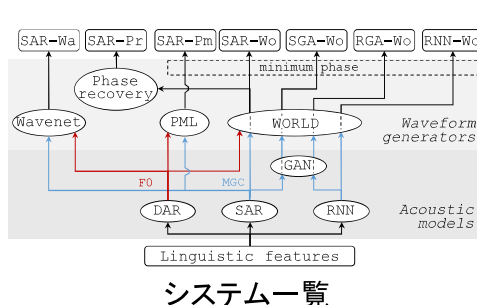
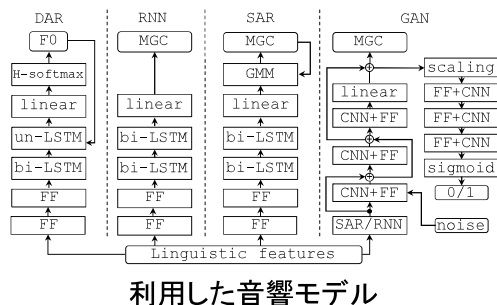
• テキストからスペクトルの予測



• GANポストフィルタの適用



• 音声合成モデルの比較



• 提案手法による合成音声サンプルを用意しています！

主観評価結果