

マルチモーダル情報処理学研究室の紹介

学術研究院環境生命自然科学学域 計算機科学講座

原直

マルチモーダル情報処理学研究室（岡山大学・原研究室）

Multimodal Information Processing Laboratory, Okayama Univ.

Sound processing

音響処理

- ✓ 音声認識
- ✓ 音声合成



Lifelogs

ライフログ

- ✓ 思い出と記録
- ✓ GPS情報の活用



Machine Learning

機械学習

- ✓ 深層学習
- ✓ オンライン学習



Human-Machine Interface

インタフェース

- ✓ 音声対話
- ✓ 可視化と見守り



Multimodal
情報源
マルチモーダル

音声情報処理技術の多くは、**時系列情報に対する機械学習方式**を基盤としており、様々な場面での活用が期待できます。加えて、**人間自身を理解する**ことで技術開発に応用することができるという**学際的な面も持つ技術分野**でもあります。

本研究室では、**音声対話インタフェース**や**生体信号に基づく行動認識**のような、**人と機械の高度なコミュニケーション**を可能とするための技術開発を目指しています。その目標のために、音声情報処理に関する研究を軸として、より**多様な情報をも表現しうる方式**の研究や、**多様な情報源を処理する方式**の研究を進めています。



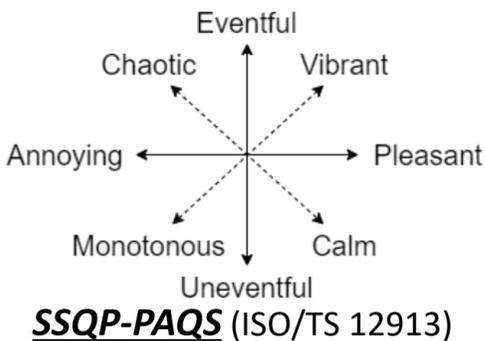
研究に基づく技術基盤の確立

- ◆ 時間信号（短期間）の処理... 音声音響技術を基盤とした技術
- ◆ 時間信号（長期間）の処理... ライフログ技術を基盤とした技術
- ◆ コンセプトドリフトの処置... 異常検知とオンライン学習（模索中）

サウンドスケープ研究への発展

- ✓ LOIS研究会 2022.3
- ✓ Internoise 2023.8

音声 & GPS情報を用いた音環境印象指標の推定



GPS情報の拡張
GPS情報を元に衛星写真を検索



[研究紹介]

環境音からの主観的な騒音マップ生成

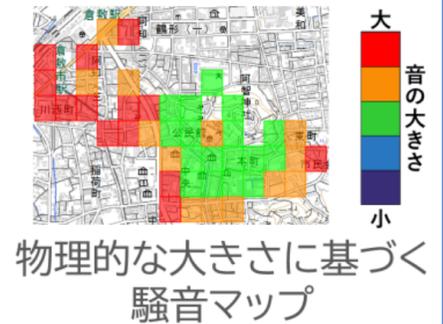
6



■ “音の感じ方”を考慮した騒々しさの推定モデルの構築方法を検討する

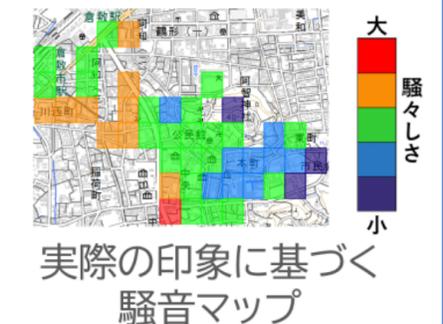
- ◆ 場所の影響 例:「住宅街」と「駅前」
- ◆ 時間帯の影響 例:「朝」と「夕方」
- ◆ 音源の影響 例:「話し声」と「エンジン音」

→ 音の物理的な大きさの可視化ではなく、実際の印象(の推定)を可視化できないか？



■ 推定モデルは、いつまでも使える？

→ 「変化」への適応が必要. でもどうやって？



■ ノイズに適応したらダメなのでは？

→ ノイズ(≒異常データ)かどうかの判定も必要. で、どうやるの？「変化」と「異常」は区別できる？

[研究紹介]

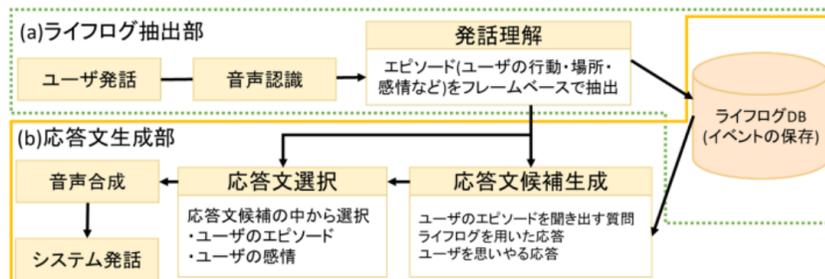
ライフログと成長する音声対話システム

7



- ずっと同じ反応しかしないシステムはつまらない
→ オンライン学習により「成長」させるには？
- 「私のための」音声対話システムに成長してほしい
→ ライフログとして「収集」し、会話に「活用」するには？
- 「私のための」音声対話システムは、「私の変化」にも気づけるかもしれない
→ 見守り, 異常検知, コンセプトドリフト, etc.

- 淡々とした音声で返されてもねえ...
→ 音声合成も重要



共感的対話音声合成研究への発展

- ✓ LOIS研究会 2021.3
- ✓ Interspeech 2023.10

Many-to-one 音声変換と Zero-shot 音声合成を組み合わせることで、ユーザに合わせてられる音声合成

