

## 【研究区分：先端的研究】

研究テーマ：AIによる認知症患者意思理解・支援に関する技術の提案	
研究代表者：地域創生学部 地域創生学科 地域産業コース 准教授 陳金輝	連絡先：kinki@pu-hiroshima.ac.jp
共同研究者：神戸大学都市安全研究センター・ 同大大学院システム情報学研究科（兼務）教授 滝口哲也	
<b>【研究概要】</b> 認知症又は知的障がい者はパターンが多様性であり、AI技術による障がい者行動・意図の理解に困難がしつつある。そのため一つのモダリティだけではなく、複数のモダリティを活用したユニバーサルコミュニケーション支援が必要である。本研究は、機械学習を軸にした複数モダリティの統合による障がい者意図の理解、さらに患者に支援できるような技術の実現を目指す。これらの課題に対処していくため、今年度は障がい者の位置推定、ルージュテストの研究に取り組み、主として評価高い国際ジャーナルに研究成果が掲載されている。	

### 【研究内容・成果】

#### 1. 背景、目的及び令和2年度実施内容

人間認知は知的な活動であり感情、意図など主観的知覚に関わり、それらは、特に、障がい者認知関連、そのメカニズムは十分に解明されたわけではなく、このことは他分野の関連研究においても同様である。

人工知能の分野において、アフェクティブ・コンピューティングという、人間の感情、意図、行動などに関係するコンピューティングの研究がある。現状では、これらのほとんどは表情解析や情緒計算といった単一の個別課題についてのみしか検討されていない。即ち、それらの研究はほぼ例外なく単純な認識や検出に関連する課題について、各自の手法を提案したうえでモデルが組み立てられている。たとえば、典型的な「FACS」法は感情表出のときに現れる、顔面上の特定の部位における動作のシステムであり、表情画像の(特徴)記述や識別に広く用いられている(金出武雄ら, 2010)。我々も、画像の「空間変形に対する不変な(ディープ)」特徴記述という視点から、畳み込みニューラルネットワーク(ディープラーニング技術の1つ)に基づく感情など生体データの解析を行ってきた(参考文献1)。

しかし、認知症患者又は知的障がいの人々は一人一人その障がい特性が多様であり、コミュニケーションが困難な場合がある。そのため複数のモダリティ(音、映像など)モデルの協働によるユニバーサルコミュニケーション支援が必須、不可欠である。本研究は機械学習を用いた視聴覚モデルによる認知症行動の理解、及びそれをコミュニケーションへ応用との統合制御について、実証実験まで行うものである。令和2年度は a)障がい者行動を解析するための位置推定を行った。また、b)画像認識技術を用いたルージュテストによる軽度認知機能障害に対する早期診断の支援研究を一部実施した。今回は a)を主として報告する。

#### 2. 提案手法の概要

障がい者行動を解析するため、その位置情報が不可欠である。本研究では、部屋の音響伝達特性を識別することで、音源(=障がい者)の位置を推定する手法を提案する。口又は動作から発せられた直後のクリーン音声は、壁からの反射音(残響)やスペクトルの減衰などの、部屋音響伝達特性が畳み込まれた残響信号としてマイクに観測される。この音響伝達特性は、音源の位置に依存して変化することが知られている。この手法では、この音響伝達特性が音源位置に依存する点に着目し、音響伝達特性をあらかじめ音源位置毎に学習しておき、評価音声についても、その音響伝達特性を識別することで、音源位置を単一マイクで推定する手法を提案する。

さらに、より正確に音響伝達特性を推定するために、距離情報を音声モデルで、方位情報を、画像認識モデルでそれぞれ音声推定する。複数のモデルを用いた方がクリーン音声をより詳細にモデル化でき、それにより音響伝達特性の推定も正確になると考えられる。

距離情報について、HMM (Hidden Markov Model) は複数の状態からなる状態遷移モデルであり、音声の時間的変化を表現することができるため、クリーン音声を HMM でモデル化して推定する。一方、方位情報については、

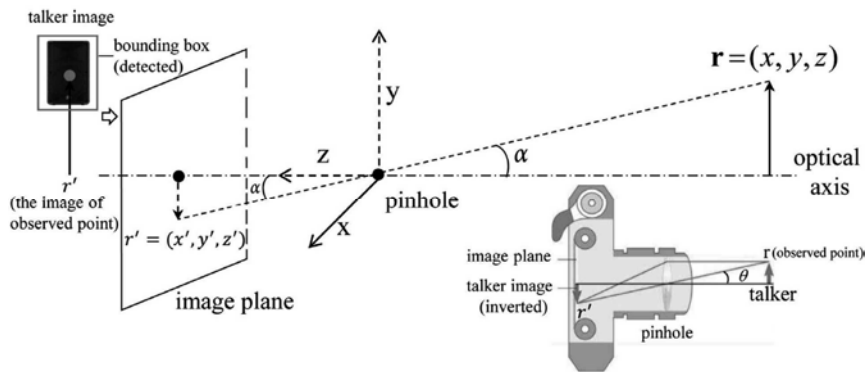


図 1. 凸レンズイメージングによるオブジェクト方位の推定

図 1 (成果論文 1) とも掲載) に示すよう、凸レンズイメージング法による推定する。過去研究成果(参考文献 1)の改善手法であり、方位を安定的計算するため、フーリエ解析を用いて局所的極座標化する計算より、回転・反転されても不変性を持つ新たな polar-HOG 特徴を提案した。

### 3. 評価・考察

音響伝達特性の特徴量ベクトルである MFCC (Mel-Frequency Cepstrum Coefficients) の各次元の中には、音源位置の識別に有効な次元とそうでない次元があり、かつその次元は音源位置によって異なるという考えに基づいて、Multiple Kernel Learning (MKL) を用いた音源位置毎の次元重みを学習させる手法を提案する。MKL は本来、様々なカーネル関数を統合することによって認識器の性能を強化するために用いられてきた手法であるが、本手法では MKL を MFCC の各次元の重みを求める目的で使用。また、MKL は Support Vector Machine (SVM) 系識別器であり、polar-HOG を組み合わせた識別枠組みによる安定的な方位を推定でき、その有効性が実験で確認できた。音源位置の識別実験において、提案手法と複数のマイクロホンを用いた音源位置・方向推定の代表的な Cross-power Spectrum Phase (CSP) など手法を比較し、提案法の有効性を示した。また、実験結果の詳細データを、AI 領域トップジャーナルの 1 つである Pattern Recognition に発表した (成果論文 1)。

なお、今回は学習データ量の制限でディープラーニングの手法が検討できなかった。

### 4. まとめと今後の予定

部屋の音響伝達特性が音源の位置に依存する点に着目し、音響伝達特性を識別することと凸レンズイメージングによる方位を推定することで、認知症障がい者の位置を推定する手法を提案した。提案法は、複数組のマイクロホンアレーや反射板も用意する必要がなく、マイク一つのみで実装することができる。

一方、この手法は、強雑音環境下や、別の話者など、学習時と評価時の環境が異なる場合は精度に影響が与えられ、今後、よりロバストな手法、少量データにも対応可能な zero-shot など深層学習法を導入する予定である。また、現在取り込んでおり、画像認識技術を用いたルージュテストによる軽度認知機能障害に対する早期診断支援技術の研究開発についても今後引き続き推進していく。

研究成果(下線：本研究プロジェクトの構成員，\*付：研究代表者)

1. J. Chen\*, R. Takashima, X. Guo, Z. Zhang, X. Xu, T. Takiguchi, E. R. Hancock: “Multimodal Fusion for Indoor Sound Source Localization”, Pattern Recognition, 2021. (forthcoming)

#### 参考文献

1. J. Chen\*, Z. Luo, Z. Zhang, F. Huang, Z. Ye, T. Takiguchi, E. R. Hancock: “Polar Transformation on Image Features for Orientation-Invariant Representations”, IEEE Trans. Multimedia Vol. 21(2) PP.300 - 313, 2019.