

博士論文要旨

論文題名：多様な感情表現の理解に向けた 音声感情認識の高度化

立命館大学大学院情報理工学研究科
情報理工学専攻博士課程後期課程

ナガセ リョウタロウ
永瀬 亮太郎

本論文では音声伝える感情を推定する音声感情認識の研究に取り組む。この技術は、暮らしやビジネス、医療・福祉の幅広いサービスや商品に応用されている。

音声感情認識は多くの場合機械学習で実現される。特に、深層学習に基づく音声感情認識の研究は盛んに取り組まれており、認識率は年々向上している。これまでの研究では、音響情報のみを入力とし発話全体から喜怒哀楽などの代表的な感情を推定する手法が検討されてきた。そのため、音響情報のみでは誤認識しやすい感情や時間と共に変化する感情、代表的な感情のみでは示せない感情の認識が困難である。これらの感情を認識できるようにするため、本論文では「音響情報と言語情報を併用した音声感情認識」「感情ラベル列を用いた音声感情認識」「感情キャプションを活用した音声感情認識」の手法の検討を行った。

「音響情報と言語情報を併用した音声感情認識」の研究では、声高や周波数スペクトルなどの音響情報と発話した内容などの言語情報を組み合わせた手法を検討し、音声感情認識に効果的な手法を同一条件下で比較した。この実験を通して、認識性能の改善に効果的な音響・言語情報の組合せを明らかにした。また、言語情報として音声認識結果を利用し、2つの情報の統合手法にどのような影響を与えるか明らかにした。

「感情ラベル列を用いた音声感情認識」の研究では、音素情報を利用し感情ラベルを並べた系列を予測するようにニューラルネットワークを学習する新しい手法を提案した。この手法によって、音素情報を活用することが細かな感情の認識に有効であることを明らかにした。また、発話内で感情が変化する音声を用意し、提案手法が時々刻々と変化する感情の認識に有効であることを示した。

「感情キャプションを活用した音声感情認識」の研究では、予測する感情形式として感情の説明文（感情キャプション）を利用し、音声から感情キャプションを推定する手法や分類クラスを推論時に感情キャプションで定義できる手法を提案した。この手法によって、代表的な感情では示せない感情をテキストとして認識できることを示した。また、感情に関連するクラス（e.g. 購買意欲の有無）などもある程度予測できることを示した。

以上の研究を通して、既存手法が認識していた感情の範囲を拡張した高度な音声感情認識の実現を目指した。