

客観的感情尤度と感情強度に基づく主観的体験感情の推定

Felt-emotion estimating method on the basis of objective emotion probability and intensity

齋藤 晶太
Shota Saito

広島市立大学大学院 言語音声メディア工学研究室

Language and Speech Research Laboratory, Graduate School of Information Sciences, Hiroshima City University

概要 本研究では、被験者が体験した感情に近い“主観的評価感情”ラベルを付与した音声データベースによって構築した機械学習器による感情推定手法を提案する。しかし主観的評価ラベルを付与した大規模データベースは公開されていないため、異なる感情ラベルを付与した大規模データベースに対して主観的評価感情ラベルを推定する手法を適用することでデータ数の不足を補う。主観的評価感情をより正しく推定するため、本稿では客観的感情尤度と感情強度のラベルを付与されたデータベースを用いる。

1 はじめに

現在、コンピュータ上で感情や感性を扱う研究が盛んにおこなわれている。解析対象が“感情”という抽象的なものであるため、感情推定処理には人手で判別ルールを作成しなくてもよい機械学習が使われることが多い。教師付き機械学習を行うためには学習するデータに正解ラベルを付与する必要があるが、感情推定研究においては、同じ“感情”の名のもとに、様々な感情ラベルが混在している。図1に感情推定研究で用いられている感情ラベルの種類を示す。

このうち多くの感情推定システムで本来目指しているのは、「観測対象である人物が現在体験している感情」の推定である。しかし、体験している感情そのものをリアルタイムに取得することは非常に困難であるため、体験している感情の代わりとなる感情ラベルを用いて研究を行っている。以下に**体験した感情**の代用となっている感情について説明する。

情動喚起刺激: 対象となる感情を人為的に喚起して、その時の被験者の状態を計測する。IAPS や IADS のような情動喚起刺激データベースでは、各刺激を提示することによって多くの人が喚起された感情がラベリングされている。実験に際して被験者がどのような感情を喚起したかの評価を行う必要は無いが、必ず全ての被験者がこの刺激によって同じ感情を喚起しているという保証は無い

自覚している感情: 被験者にある感情が生じた時、感情生起後に被験者自身によって評価された感情。ただし、計測実験と同時に評価作業を行うと、「評価する」という脳内プロセスが生理情報や次の感情生起傾向に影響を与える恐れがある。一方、実験後の評価だと被験者自身の記憶も薄れているため、他者による評価と同様に精度が低下する恐れがある

表出しようとする感情: 被験者に“ある感情状態と思わ

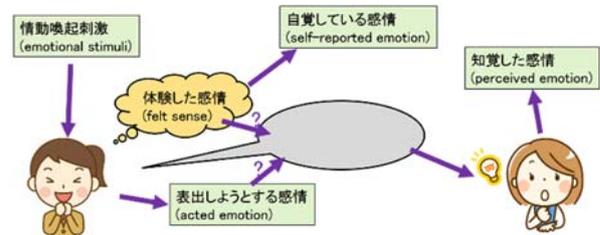


図1: 感情推定研究で用いられている感情ラベル

れるように” action をしてもらい、相手に伝えるために演技っぽくなり、特徴を大げさにしたりシンボリックな action やゼスチャーをする傾向があると言われていた。ただ、Stanislavski System のように「被験者自身により対象感情をその場で喚起」した場合、技術的な演技とどう異なるのかについては現在分析実験中である

知覚した感情: 被験者自身の感情ではなく、被験者を観測している他者が推定した感情。被験者が体験した感情から、表出プロセス、(他者による)知覚プロセス、認知プロセスを経由するため推定精度が下がる恐れがあるが、複数の評価者を用いることができるため評価基準の揺らぎの影響は軽減できる。また、データ収集作業と完全に切り離して評価作業を行うことができる。

現在公開されている感情音声データベースでは、OGVC の自発音声や UUDB など多くのデータベースのように知覚感情に基づいてラベリングされているか、あるいは OGVC の演技音声や広島市大データベース[1]の演技音声のように表出しようとする感情に基づいてラベリングしている。一方、脳科学や心理学の分野では情動喚起刺激の感情ラベルに基づいた実験が多い。しかし、自覚している感情に基づくラベリングをしたデータベースは現在ほとんど公開されていない。

そこで本研究では、小規模の自覚感情ラベル付きデータベースを構築し、それを用いた話者感情推定手法を提案する。さらに、自覚感情ラベル付きデータベースのデータ数の少なさを補うために、他の大規模データベースのラベルを付け替える手法[2]を用いて、自覚感情ラベル付きデータを増やす。ラベル付け替え処理に際しては、音響的特徴量だけでなく、客観的評価による感情尤度と感情強度も素性に用いる。

2 主観的評価感情を推定する感情推定手法

本論文で提案する感情推定手法の処理の流れを図2に示す。提案手法では、入力音声の音響的特徴量から発話者の主観的感情を機械学習器 (SVM1) によって推定する。SVM1 は、被験者自身の評価による感情ラベルを

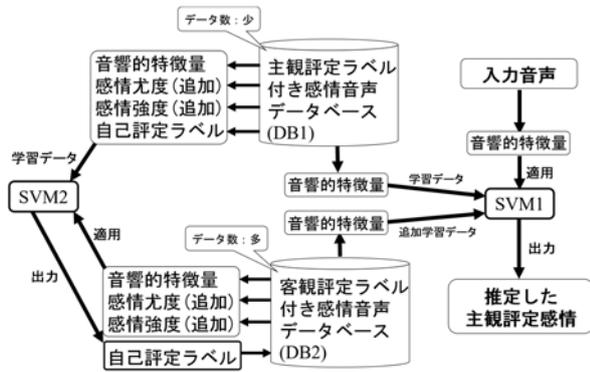


図 2：提案手法の処理の流れ

付与されたデータベース (DB1) を用いて、発話音声の音響的特徴量を入力、感情ラベルを出力とするように学習される。しかし DB1 は我々が構築した小規模データベースであるため、データ数は少ない。そこで DB1 を使って異なる大規模感情音声データベース (DB2) の音声データに対して主観的感情ラベルを推定し付与することで、DB1 のデータ数不足を補う。

3 節では主観的感情をラベリングした感情音声データベースの構築方法について、4 節では異なるラベルが付与されたデータに対して主観的感情ラベルを推定する手法について説明する。

3 主観評定による感情音声データベース

以下に主観評定感情ラベル付きデータベースの構築方法を示す。

- (1) 2 人 1 組の被験者の片方にビデオカメラと高指向性マイクをセットする
- (2) (1)の状態 でゲームソフトをプレイしてもらい、プレイ中の表情と発話音声を収録する
- (3) 1 人目の収録が終了したらもう 1 人の被験者に(1)～(2)の手順を同様に行う
- (4) 両者の収録終了直後に、収録風景の動画を見ながら感情が喚起した時点の発話とその時の生起感情を自己評定により回答してもらう

生起感情の事後評定は、データ収集直後に収録した動画を提示し、感情が生起していると被験者自身が感じた発話に対して、各生起感情の尤度を[0,1]の実数値で

回答する。本研究では、怒り、嫌悪、恐れ、幸福、悲しみ、驚きの 6 種類の感情を対象としている。本収録作業により、合計 54 件の感情発話音声データが収集できた。

4 主観的感情ラベルの推定手法

推定対象データベースの感情ラベルを利用した主観的感情ラベル推定手法[2]の流れを図 3 に示す。図中の太い矢印は大規模データベースで用いられている感情ラベルを表す。本研究では[3]のデータベースを用いるため、感情ラベルは客観的に評定された感情尤度と感情強度である。被験者は各音声データについて 8 感情 (angry, happiness, sadness, dislike, surprise, interest, anxiety) それぞれの感情尤度 (0.0~1.0) と感情強度 (0.0~3.0) を回答し、回答の平均値に基づいてラベルを付与している。

まず学習パートで主観的感情ラベル付きデータベース (DB1) の各データに客観的感情尤度と感情強度のラベルを追加し、追加した感情ラベルと音声データの音響的特徴量から主観的感情ラベルを推定する分類器 (SVM2) を構築する。そして大規模感情音声データベース (DB2) の客観的感情尤度と感情強度のラベル及び音声データの音響的特徴量を SVM2 に適用することで、主観的感情ラベルを推定する。そして DB2 に主観的感情ラベルを付与したものを追加学習データとして図 2 の SVM1 の学習に用いることで、SVM1 の学習データ数を増やすことができる。

参考文献

- [1] 目良他, 演技感情と推定感情のタグを付与した感情音声コーパスの構築, 日本音響学会 2017 年春季研究発表会講演論文集, pp.1471-1474 (2017)
- [2] 長岡他, 感情音声コーパス共通化のための新たな感情ラベル推定における既存感情ラベル併用の効果, 日本音響学会誌, Vol.73, No.11, pp.682-693 (2017)
- [3] 齋藤他, 音声に明確に表出しない感情の認識のための感情音声データベースの構築, 日本音響学会 2017 年秋季研究発表会講演論文集, pp.93-96 (2017)

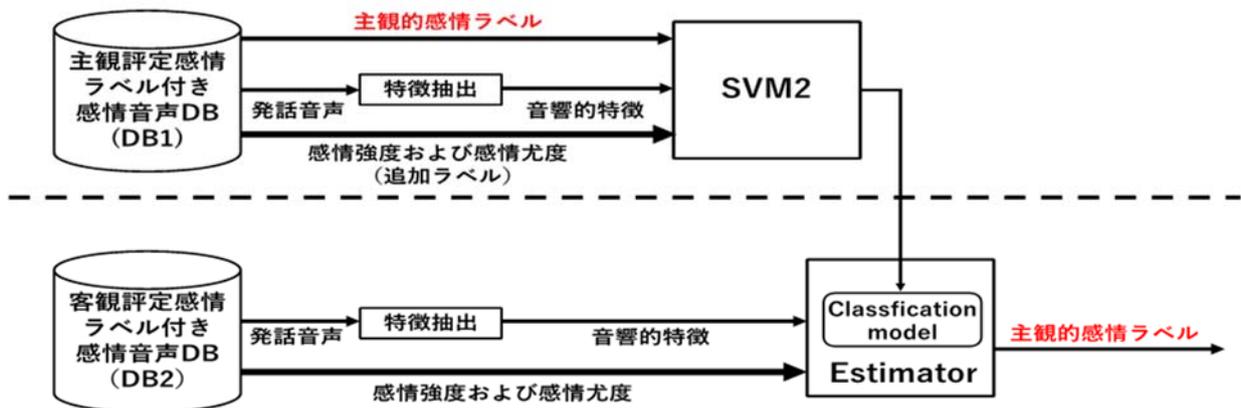


図 3：推定対象データベースの客観的感情ラベルを利用した主観的感情ラベル推定の流れ