

複数の対話音声コーパスにおける感情ラベルの相互推定*

永岡篤, 森大毅 (宇都宮大), 有本泰子 (理研・脳総研)

1 はじめに

パラ言語情報の研究には、音声対話コーパスの利用が必要となる。大部分の研究は単一のコーパスを使用しているが、様々な場面でのパラ言語情報を包括的に研究対象とすることができるため、複数のコーパスの併用が有効である。

コストの大きい人手でのラベリングを避けるため、本研究ではパラ言語情報ラベルを有するコーパスを基にした別のコーパスに対する相互自動ラベリングの検討を行っている。

本稿では、宇都宮大学パラ言語情報研究向け音声対話データベース (UUDB)[1] の感情次元推定モデルを感情評定値付きオンラインゲーム音声チャットコーパス (OGVC)[2] に適用し、それぞれの感情ラベル間の関係性を調査した。さらに、OGVC の感情ラベル推定モデルを UUDB に適用するための予備検討として、OGVC に対する感情カテゴリラベル推定精度を検証した。

2 パラ言語情報相互自動ラベリング

以下で説明する UUDB と OGVC の感情ラベルは異なる。互いの感情ラベルを一方のコーパスに付与することを考える。

UUDB には、パラ言語情報ラベルとして6つの感情次元による評価が付与されている。評価値は1から7まで(4が中立)の数値である。Fig.1に快-不快の感情次元についての例を示す。

OGVC には、Pluchik の立体構造モデルの基本8感情 [3] に加えて、感情が表出していない「平静」と「その他」の10種類の感情ラベルがカテゴリで付与されている。

本研究では、音声の音響特徴量から機械学習によってこれらの感情ラベルを推定するモデルを作成する。その枠組を Fig.2 に示す。[4] では、Fig.2 に示す経路 (b) に先立ち、まず経路 (a) の検証を行いその有効性を確かめた。本稿では、[4] で有効性が確かめられた UUDB の感情次元推定モデルを OGVC に適用する経路 (b) を実施し、推定された感情次元と付与されている感情カテゴリの関係性を調査する。さらに、経路 (d) の実施に

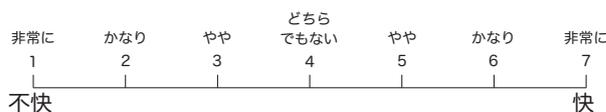


Fig. 1 UUDB における快-不快の感情次元軸

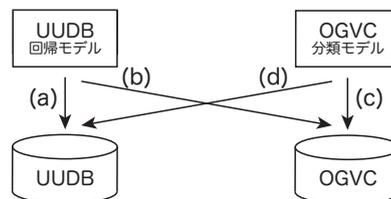


Fig. 2 感情ラベル相互推定の枠組み

先立ち、経路 (c) の精度検証を行う。

3 OGVC に対する感情次元推定実験

UUDB の話者 14 名の全 4840 発話から抽出した、Interspeech 2010 Paralinguistic Challenge ベースラインシステム [5] と同様の 1582 次元の音響特徴量を入力数として、快-不快と覚醒-睡眠の次元ラベルを推定するサポートベクター回帰モデルを学習した。カーネル関数はガウシアンカーネルを用いる。次に、これらのモデルを OGVC に対して適用した。

快-不快と覚醒-睡眠次元の推定値の分布を各感情カテゴリ毎に Fig.3 に示す。Fig.3 を見ると、快-不快では喜びと期待が快寄りに推定され、嫌悪と恐れが不快寄りに推定されていることが分かる。また、覚醒-睡眠では怒りや喜び、驚きが覚醒寄りに推定され、悲しみが睡眠寄りに推定されている。これらは、直感的には妥当な結果と言える。

4 OGVC の感情カテゴリ推定実験

OGVC から学習した感情カテゴリ推定モデルが UUDB に対して (経路 (d)) 有効に機能するためには、少なくとも OGVC に対して (経路 (c)) 有効に機能しなければならない。ここでは予備検討として、OGVC に対する感情カテゴリの推定精度を検証し、相互ラベリング実現のための課題を洗い出す。

学習データは OGVC の感情ラベルが付与されている発話のうち、評価者 3 名による評価が 2 名

* Mutual estimation of emotion label in multiple spoken dialog corpora.

by NAGAOKA, Atsushi, MORI, Hiroki (Utsunomiya University), ARI-MOTO, Yoshiko (RIKEN BSI).

Table 1 感情カテゴリ推定結果の混同行列 [%]

感情	受容	怒り	期待	嫌悪	恐れ	喜び	平静	悲しみ	驚き
受容	21.8	0.7	10.6	6.3	1.0	12.5	34.3	2.3	10.6
怒り	3.8	4.2	13.1	5.5	0.4	20.7	26.2	0.8	25.3
期待	3.5	1.4	20.6	4.7	0.2	25.5	30.9	1.6	11.5
嫌悪	9.0	1.8	9.6	12.5	0.6	12.8	40.6	5.4	7.8
恐れ	7.0	1.4	11.3	4.2	0.7	14.8	28.9	7.7	23.9
喜び	5.9	0.8	11.4	3.0	0.3	41.7	22.5	1.2	13.1
平静	8.6	1.4	11.8	7.9	0.1	12.2	47.6	2.4	8.0
悲しみ	11.1	0.4	7.8	10.3	0.4	9.9	34.2	17.3	8.6
驚き	5.5	2.8	5.9	2.0	0.5	12.5	10.7	1.4	58.7

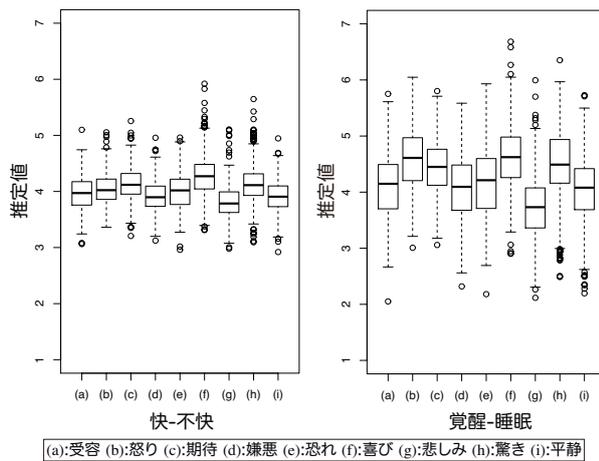


Fig. 3 OGVC の発話に対する快-不快, 覚醒-睡眠次元推定値の分布

以上一致している話者 11 名の 3645 発話である。前節と同様の 1582 次元の音響特徴量を抽出した。

感情カテゴリラベルの分類モデルをサポートベクターマシンによって学習した。カーネル関数にはガウシアンカーネルを用いた。また、カテゴリによって学習データ数が異なることを考慮し、スラック変数のコストパラメータはデータ数が多いほど大きくなるよう設定した [6]。得られた不特定話者のモデルは leave-one-speaker-out 交差検証によって評価した。

得られた分類モデルの全体の正解率は 33.7% であった。混同行列を Table 1 に示す。この表を見ると、平静と誤識別される発話が多い。一方で、驚きの発話は平静に誤識別された数が他の発話と比較して少ない。また、恐れと識別される発話の正解率が低いが、これは学習データ中に恐れラベルが付与された発話が他に比べて少ないことが原因である可能性がある。

期待が喜び、驚きに誤識別されやすく、同様に恐れは驚きに誤識別されやすいが、喜びと驚きが

期待や恐れに誤識別された割合は相対的に小さい。このことから、喜びは期待を、驚きは期待と恐れの一部を包括していると考えられる。しかし、喜びと悲しみや怒りと悲しみのように互いに誤識別されづらい組み合わせも存在していた。よって、喜びと悲しみ、怒りと悲しみは互いに独立した感情であると分かる。これは、感覚的にも矛盾しない結果であった。

5 おわりに

本稿では、パラ言語情報の評価方法の異なる UADB と OGVC の感情次元と感情カテゴリの関係性を調査した。また、OGVC の感情ラベル推定モデルを UADB に適用する前実験として OGVC の感情カテゴリ推定を行った。今後は、OGVC に対して人手ラベリングを実施して得られたパラ言語情報ラベルと本稿で得られた推定結果を比較し有効性を検証する。また、機械学習に用いる特徴量に音響特徴量以外のものを使用することを検討していく。

謝辞 本研究は立石科学技術振興財団研究助成 (課題名: 言語情報と音響情報の統合的利用による感情音声コーパスの大規模化) および JSPS 科研費 (26280100) の支援を受けた。

参考文献

- [1] Mori *et al.*, Speech Communication, Vol. 53, 36-50, 2011.
- [2] 有本泰子, 河津宏美, 音講論 (秋), 385-388, 2013.
- [3] Plutchik, "Emotions: A psychoevolutionary synthesis," Harper & Row, 1980.
- [4] 永岡篤, 森大毅, 音講論 (秋), 389-390, 2014.
- [5] Schuller *et al.*, in Proc. Interspeech 2010, 2794-2797, 2010.
- [6] Rehan *et al.*, RCML 2004, 39-50, 2004.