

# マルチモーダルデータによるアイスブレイク時の場の和み度の推定

胡 献引<sup>†1</sup> 伴 佑樹<sup>†2</sup> 濱村 航明<sup>†3</sup> 鈴木 晴菜<sup>†4</sup> 足利 朋義<sup>†5</sup> 割澤 伸一<sup>†6</sup>  
東京大学 東京大学 ダイキン工業 ダイキン工業 ダイキン工業 東京大学

## 1. はじめに

アイスブレイクとは、本題に入る前に雑談をすることで緊張感が漂う場を和ませるコミュニケーション方法である。和やかな雰囲気の中では話しやすく感じ、より多くの意見を引き出したり新しいアイデアを自由に発表するのに役立つため、アイスブレイクは協創的作業の生産性向上にとって重要な手段である。アイスブレイクの効果を指標化することで自動ファシリテーションや、場を和ませる空間デザインなどの応用創出が考えられる。

先行研究では会話参加者の主観評価を指標として用い、事後で比較してアイスブレイクの設計要素に関する洞察を得ている。一方で、会話を行う空間の最適化に対しては、設計要素が多様かつ高精度であるため、従来手法を用いた実験コストが高い。そのため、会話の中断を必要とする主観的評価ではなく、客観的データを用いて会話を行う最中で評価できる指標が重要になってきた。また、環境設計は複雑であるため、単なる客観的指標の変化を見て設計するのではなく、環境要素に対して参加者がどう感じとるかを知ること設計のヒントをもらえるに大事であると考えられる。

情報センシングと処理技術の発展につき、場の和み度を客観的データから推定して指標化する先行研究が行われた。例えば、稲熊 [1] らは、笑いのタイミングを適切に把握した人間らしい対話ロボットの開発が目的であるため、場が和んだタイミングを話者間の共有笑いの発生タイミングとして定義した。しかし、これらの先行研究における評価指標の定義は、応用先から特徴付けられ、アイスブレイクの本来の目的からの乖離が見られた。

本研究では、アイスブレイクの目的に即し、場の和み度を複数人がいる雑談の場における参加者が感じる話しやすさとして定義する。そして、この主観的指標を、生体情報、視線、姿勢、表情を含めたマルチモーダルデータから、畳み込

みニューラルネットワーク (CNN) を用いて推定することで、会話の中断を必要としない場の和み度の客観的指標作りを試みた。環境要素の例として温度取り上げ、参加者の温熱環境に対する評価を取り、場の和み度との関係を分析した。

## 2. 実験とデータ処理

本実験は 44 名 (男性 22 名, 女性 22 名, 平均年齢 35 ± 8.4 歳) の実験参加者からなる 11 組の 4 人グループでアイスブレイク会話実験計画を行い、推定モデルの学習と評価に必要なデータを収集した (図 1)。各グループは男性と女性が 2 名ずつで構成される。

実験開始後、参加者に実験に関する説明を行い、その後機器を装着してもらった。装着するデバイスとしては、心電センサ、呼吸バンド、皮膚電位センサ (biosignals plus, biosignals Pro) およびメガネ型アイトラッカー (Tobii, Pro Glasses 3) がある。他のデバイスとして、顔動画を撮影するウェブカメラと、姿勢データをセンシングする深度カメラ (Microsoft, Azure Kinect) が整備された。また、対話風景を撮影するために、参加者 4 人を写せる 360 度カメラを使用した。機器装着の完了後、現状の温熱環境に対する評価として、紙アンケートにより温冷感を 5 段階で回答させた。次に各センサーデータを確認した後、約 20 分間のアイスブレイクを行なった。具体的には、ファシリテーターの指示に応じて、4 人が順番に簡単な自己紹介をしてから、お題「どこでもドアがあったら、どこにいきますか?」について発表した。他の参加者はそれについて自由にコメントするように指示した。アイスブレイク終了後、再び温冷感について回答した。場の和み度の真値を取るために、アイスブレイク期間の会話風景動画を 2 分間セグメントに区切ったものを順番に閲覧しながらアンケートに回答するウェブシステムを作成した。この動画振り返りアンケートシステムの準備に時間がかかるため、約 40 分間の休憩時間を設けた。休憩中の参加者はカードゲーム ITO をプレイした。最後に、参加者全員が会話風景の動画を振り返りながら話しやすさについて 5 段階評価した。

全ての場の和み度評価付きの 2 分間におけるマルチモーダルデータに対して以下の処理を行い、9 種類の特徴量を抽

Estimation of Nagomi During Icebreak Through Multimodal Data

<sup>†1</sup> HU XIANYIN, The University of Tokyo

<sup>†2</sup> BAN YUKI, The University of Tokyo

<sup>†3</sup> HAMAMURA KOMEI, DAIKIN Industries, LTD.

<sup>†4</sup> SUZUKI HARUNA, DAIKIN Industries, LTD.

<sup>†5</sup> ASHIKAGA TOMOYOSHI, DAIKIN Industries, LTD.

<sup>†6</sup> WARISAWA SHIN'ICHI, The University of Tokyo



図1 実験風景

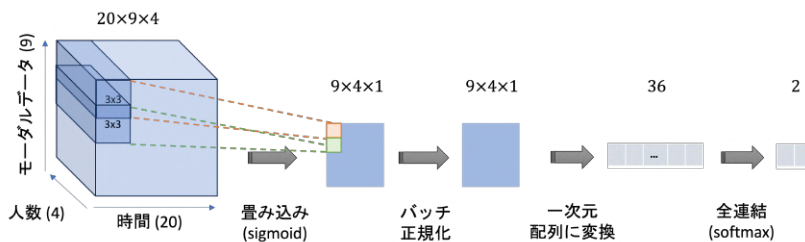


図2 CNNの構造

出した．心電の信号からは R ピークを検出し，R-R 間隔の逆数をとった心拍率の時系列データを利用した．呼吸の信号からは呼吸間隔を算出し，同様に間隔の逆数をとった呼吸率の時系列データを利用した．皮膚電位 (EDA) に関しては，持続的 (Tonic) 成分と一過性的 (Phasic) 成分に分離し利用した．瞳孔径は生データを利用した．視線については，アイトラッカーで録画した視野動画の全フレームに対して既存の深層学習モデル [2] を用いてピクセルワイズのラベリングを行い，視線先座標のラベルをデータとして利用した．ラベルとしては背景，人の顔と人の身体の 3 種類がある．さらに，視線先が人の顔であるフレームに対して，注視する人の相対位置のラベルについても，評価者一名がフレーム画像を見て分類を行なった．顔表情に関しては，既存の感情認識モデル (Bleed AI) を用いて顔動画の全フレームに対して 7 種類の感情確率を算出し，その中の楽しい表情の確率の時系列データを利用した．最後に姿勢に関しては，実験開始初期と比べ身体平面の変化角度を前傾角度として利用した．

推定の目的変数である場の和み度は，4 人の参加者の中，任意の複数の  $n$  人が感じる話やすさの平均値 ( $n = 2, 3, 4$ ) とした．

そして目的変数  $y$  を離散化し， $y > 3$  は場の和み度が高い群とし， $y \leq 3$  は場の和み度が低い群として二値化した．

図 2 で示した通りの 2 分類 CNN モデルを構築した．入力データ  $X$  は，時間，データモダル，参加者の 3 軸から構成したデータである ( $X \in \mathbb{R}^{20 \times 9 \times 4}$ )．メンバー軸においては， $n < 4$  の場合，場の和み度の評価に貢献しない参加者の行動データは全て 0 で埋めた．時間軸においては，異なるモダルのデータを 20 サンプルまでにダウンサイジングして揃った．畳み込み層のフィルターのサイズは  $3 \times 3$ ，移動幅は 2 に設定した．損失関数は cross entropy 関数を使用した．学習アルゴリズムは Adam を用い，学習率は 0.01 に設定した．

### 3. 結果と考察

推定モデルを leave-one-out 交差検証を用いて評価した．具体的には，11 グループ内の 10 グループは学習データとして CNN に学習させ，残りの 1 グループをテストデータと

して推定精度および F1 点数を算出し評価した．構成される学習データにおいて，二群のサンプル数のバランスがもっとも良かった上位の 3 グループである第 2, 5, 10 グループをテストデータとした．結果として，3 回交差検証の平均精度は 0.74，平均 F1 スコアは 0.82 であった．CNN によりマルチモーダルデータを用いて場の和み度をランダムな推定より高い精度で推定できた．学習済みの CNN の出力は，自動ファシリテーションやアイスブレイクを行う空間の設計要素や因子範囲を特定する際の参考情報として有用であると考えられる．また，ウィルコクソンの順位検定を用い，温冷感についてアイスブレイク前後で 4 人平均値を比較した結果，行う後は前より有意に高かった ( $p = 0.0056$ )．場の和み度についても，最初の 2 分間より，最後の 2 分間の真値 ( $n = 4$  のみ) は有意に高かった ( $p = 0.0036$ )．この結果からアイスブレイクを行う後に温冷感と主観和み度がともに上昇したことが示唆された．アイスブレイクに適した温熱環境として，少し暖かく感じさせるような空間が良い可能性があると考えられる．

### 4. おわりに

本研究では，アイスブレイクにおいて，生体情報，視線，顔表情および姿勢からなるマルチモーダルデータをからある程度主観的場の和み度を推定できることを確認した．さらに，アイスブレイクを行なった後場の和み度と温冷感が共に上昇することがわかった．今後は，推定モデルの汎化性を検証することと，フィルターを可視化して場の和み度の推定に有用な客観的特徴を明確にする予定である．

### 参考文献

- [1] 稲熊寛文，井上昂治ほか初対面対話における場の和みのマルチモーダルな分析と検出，第 78 回全国大会講演論文集，Vol. 2016, No. 1, pp. 515–516 (2016)．
- [2] Lin, K. et al.: Cross-domain complementary learning using pose for multi-person part segmentation, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 31, No. 3, pp. 1066–1078 (2020)．