

感情分析と感情適応型音声合成を統合した AI カウンセリングシステム

キム チャンソプ¹ 松吉 俊¹

¹ 東京工科大学 メディア学部

m0a25111ec@edu.teu.ac.jp, matsuyoshis@stf.teu.ac.jp

概要

本研究では、感情分析機能と感情適応型音声合成を統合した AI カウンセリングシステムの対話特性を評価した。提案システムは、最大3つの感情同時検出と、動的音響パラメータ変調を組み合わせて、ユーザーの複雑な感情状態に応じた音声表現を実現する。カウンセラーの回答をタメ口モードに設定することで、親近感を抱かせるようにした。同様の入力プロンプトを用いた ChatGPT と比較実験を行った。実験の結果、提案システムが優位性を示し、特にタメ口モードによる応答が共感性と親密性において高い評価を得た。本研究は、LLM ベースの複合感情認識と音響パラメータ変調を組み合わせた共感型対話システムの有効性を実証し、カウンセリング領域におけるマルチモーダル AI システムの設計方針を提供する。

1 はじめに

近年、精神健康支援の需要が世界的に増加する一方で、専門カウンセラーの不足が深刻な社会問題となっている。このような状況において、大規模言語モデル (LLM) を活用した AI カウンセリングシステムは、24時間アクセス可能な初期支援ツールとして注目されている。特に、ユーザーの感情状態を適切に認識し、共感的な応答を生成する技術は、効果的なカウンセリング対話の実現に不可欠である。

従来の AI カウンセリングシステムは、ChatGPT 等の汎用 LLM を用いることで高度な応答生成を実現している。しかしながら、テキストのみの対話では、声のトーンや話し方といった非言語情報を通じた感情伝達が欠如しており、対面カウンセリングで重要とされる「共感の表現」が限定的である。

そこで、本研究では、感情分析と感情適応型音声合成を統合した AI カウンセリングシステムを提案する。この論文では、このシステムの詳細とユーザー評価実

験の結果について報告する。

2 関連研究

井手ら [3] は、BART を用いた応答生成と感情認識のマルチタスク学習により、発話の感情を考慮した対話応答を生成する手法を提案した。宮川ら [4] は、発話履歴を考慮しつつシステム発話に対するユーザーの感情を推定する手法を提案した。これらの研究が感情推定に焦点を当てているのに対し、本研究は感情に応じた音声表現の生成という応用的側面に重点を置いている。

市村ら [2] は、感情タイプと単語評価極性との関係を分析し、ロジスティック回帰分析により感情タイプから単語極性を予測し、離散値による感情表現と連続値による感情表現の互換性を実証した。

田辺ら [6] は、物語の登場人物の感情推定において、Plutchik の 8 基本感情をマルチラベル分類問題として扱い、コモンセンス知識グラフ ATOMIC を利用することで推定精度を向上させた。本研究でも感情分析はマルチラベル分類問題として扱う。

頼ら [5] は、音声とテキストというマルチモーダル情報を利用してゲームキャラクターの感情を推定する手法を提案した。頼らがシステム入力段階でマルチモーダル統合を行っているのに対し、本研究では感情分析結果に基づいて出力段階で音響パラメータを動的に制御する。

3 提案システム

提案する AI カウンセリングシステムの主要コンポーネントは、感情分析モジュールと音声合成モジュールである。提案システムとの対話例を図 1 に示す。

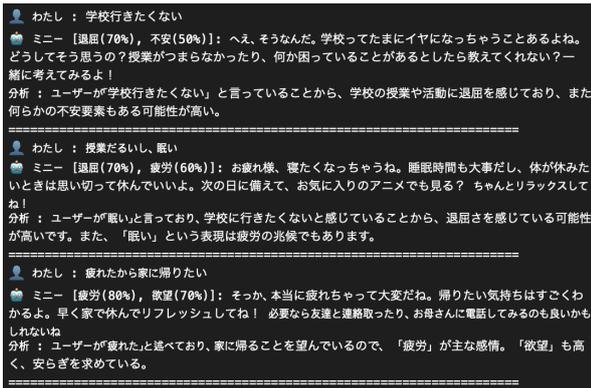


図1 提案システムとの対話例。システム（「ミニー」）の発話は、合成音声として再生される。

3.1 感情分析モジュール

Ollama¹⁾フレームワーク上で動作する Phi-4 モデル [1] にプロンプトを与えることで、感情分析を実現する。与えたプロンプトは以下の3つの特徴を持つ。

- 感情の種類、その感情を検出した理由、カウンセラーとしてのアドバイスの3つを組として出力させる
- 最大3つの感情を同時に検出し、各感情に0~100の強度値を付与する。これは、例えば、「疲労(70%) + 不安(30%)」のように、複雑な感情状態を効率的に表示するためである。
- タメ口での回答を要求する。これは、ユーザの心理的負担を消し、ユーザの自己開示を促進するためであり、既存の丁寧語ベースのカウンセリングシステムとの差別化要因でもある。

本研究では、心理学の感情理論と日本語の感情表現の特性を考慮して、独自に次の14種類の感情を定義した。

悲しみ、怒り、喜び、不安、疲労、平穏、嫌悪、
 小心、恥、当惑、退屈、羨望、欲望、恐怖

3.2 音声合成モジュール

音声合成エンジンとして、オープンソースの日本語音声合成ソフトウェアである VOICEVOX²⁾を採用した。前節の分析結果を使用して、システム発話において以下の4つのパラメータを調整する。

速度、ピッチ、抑揚、音量

各感情に対する音響パラメータの調整規則を表1に示す。これらの規則は、感情と音声特徴に関する心理音響学的知見と、予備実験における評価に基づいて独自

1) <https://ollama.com>

2) <https://voicevox.hiroshima.jp>

に設定した。0~100の感情強度値をパーセンテージに変換し、調整パラメータの重みとして用いる。複数の感情が検出された場合、各感情による調整結果を単純に加算する。この設計により、「悲しみ(70%) + 不安(30%)」のような複合感情に対して、両方の感情特性が適切な比率で反映された音声表現が生成される。

4 評価実験

提案システムを実際に被験者に利用してもらい、評価実験を実施した。

4.1 実験設定

比較対象のベースラインとして、ChatGPT (GPT-4)³⁾に同様のプロンプトを与えたものを用いた。なお、ベースラインはテキスト対話とし、音声合成は利用しない。

被験者は7名であり、提案システムとベースラインの使用順序はカウンターバランスを取った。複雑な感情が含まれる発話をスムーズに入力してもらうために、被験者に以下の3つの状況を提示し、任意に1つの状況を選んでもらった。

- 状況1: 3年付き合った彼/彼女と別れた後、復縁を望んだが、その人には既に新しい恋人がいた。怒りと悲しみで人生の虚しさを感じている。
- 状況2: 授業の後にスーパーに買い物に行く決めていたのに、すっかり忘れて帰宅してしまった。玄関の前でこの事実気づき、怒りと悲しみを感じている。
- 状況3: 海外で働いている私に、日本の親友から結婚式の招待状が届いた。ところが、その日、絶対に外せない重要な取引先とのミーティングがある。そのため、心配と悩み、疲労を感じている。

評価は主に5段階リッカート尺度によるアンケート調査により行った。

4.2 実験結果

感情分析とその理由に関する評価を表2に示す。両システムとも否定的評価は皆無であった。この表から、提案システムのほうがより強い肯定的評価を得たことが分かる。

提案システムの音声合成機能に関する評価を表3に示す。すべての被験者が、出力音声の音響パラメータが変化していたことに気づいたことが分かった。これ

3) <https://chatgpt.com/>

表1 感情タイプと音響パラメータの対応

感情タイプ	速度	ピッチ	抑揚	音量	設計意図
悲しみ	-0.2	-0.05	-0.4	0	ゆっくり、低く、平坦に
喜び	+0.1	+0.05	+0.2	0	やや早く、高く、生き生きと
怒り	+0.2	-0.03	+0.2	+0.3	早く、低く、大きく
不安	+0.1	+0.04	0	0	早口で上擦る
疲労	-0.25	0	-0.3	-0.1	とてもゆっくり、平坦、小さく
平穏	+0.1	+0.05	+0.2	0	落ち着いて明るく
嫌悪	-0.15	-0.04	-0.2	0	低く、やや遅く
小心	+0.05	+0.02	0	-0.4	控えめに、小さく
恥	-0.1	+0.02	0	-0.2	遠慮がちに
当惑	-0.1	0	+0.3	0	戸惑いの抑揚
退屈	-0.1	0	-0.5	0	棒読み気味
羨望	0	+0.03	+0.2	0	やや明るく
欲望	-0.3	-0.06	0	-0.1	ねっとり
恐怖	+0.25	+0.08	+0.5	0	パニック、震える

注：数値は基準値からの変位量。正の値は増加、負の値は減少を示す。

表2 感情分析精度の評価

評価項目	システム	非常に	やや
		そう思う	そう思う
感情の適切な把握	提案システム	5 (71.4%)	2 (28.6%)
	ベースライン	2 (28.6%)	5 (71.4%)
分析理由の納得性	提案システム	3 (42.9%)	4 (57.1%)
	ベースライン	2 (28.6%)	5 (71.4%)

は、設計した感情適応型音響パラメータ制御が、知覚可能なレベルで機能していることを実証している。感情との適合性については、「ややそう思う」が71.4%と多数を占めたものの、「あまり思わない」が28.6%存在した。この結果は、その適合性には改善の余地があることを示唆している。

音声が自然に感じた感情のアンケート結果によると、「悲しみ」が最も自然と評価され（85.7%）、次いで「疲労」（57.1%）、「不安」（42.9%）、「当惑」（28.6%）となった。この結果は、音響パラメータ調整規則において、特にネガティブな感情の表現が効果的であったことを示唆している。

タメ口モードによる共感性・親密性の評価を表4に示す。両システムでタメ口モードに関して約86%（6/7名）が肯定的評価をしており、従来の丁寧語ベースのカウンセリングシステムと異なるアプローチの妥当性が支持されたことが分かる。

カウンセリング効果に関する評価を表5に示す。適切なカウンセリングであったかについて、「非常にそ

う思う」では提案システムがやや優位であるが、ベースラインのほうが一貫した肯定評価を得ている。

人間カウンセラーの代替可能性についてもアンケートを取った。肯定的評価は両システムとも42.9%（3/7名）にとどまり、約4割が否定的評価を示した。これは、AIカウンセリングシステムの現状の限界を示すと同時に、技術的に高度なシステムであっても、人間のカウンセラーが持つ専門性や信頼性を完全には代替できないことを示唆している。

5 おわりに

本研究では、テキストベースの大規模言語モデル（LLM）と音声合成技術を統合し、感情認識に基づく動的な音響パラメータ制御を取り入れたAIカウンセリングシステムを提案した。提案システムは、ユーザの発話から複合的な感情を抽出し、その感情タイプと強度に応じて音声のピッチ、速度、抑揚をリアルタイムに適応させることで、より人間らしい共感的な対話の実現を目指した。

評価実験の結果、提案システムは、比較対象としたテキストベースのシステムに対し、感情分析の精度およびカウンセリングの適切さにおいて高い評価を獲得した。特に「悲しみ」などのネガティブな感情表現においては、設計した音響パラメータ制御が自然な発話として受容され、ユーザの共感性を高める効果が確認された。また、親密さを演出する「タメ口」スタイルの導入も、カウンセリング体験の向上に寄与すること

表3 音声合成機能の評価

評価項目	非常にそう思う	ややそう思う	どちらとも言えない	あまり思わない	全く思わない
音声変化に気づいた	4 (57.1%)	3 (42.9%)	0 (0%)	0 (0%)	0 (0%)
感情との適合性	0 (0%)	5 (71.4%)	0 (0%)	2 (28.6%)	0 (0%)

表4 タメ口モードの評価

システム	非常にそう思う	ややそう思う	どちらとも言えない	あまり思わない	全く思わない
提案システム	5 (71.4%)	1 (14.3%)	0 (0%)	1 (14.3%)	0 (0%)
ベースライン	5 (71.4%)	1 (14.3%)	0 (0%)	1 (14.3%)	0 (0%)

表5 カウンセリング効果の評価

システム	非常にそう思う	ややそう思う	どちらとも言えない	あまり思わない	全く思わない
提案システム	3 (42.9%)	2 (28.6%)	1 (14.3%)	1 (14.3%)	0 (0%)
ベースライン	2 (28.6%)	5 (71.4%)	0 (0%)	0 (0%)	0 (0%)

が示唆された。一方で、音声と感情の適合性については評価が分かれる結果となり、一部の感情カテゴリにおいては音響変化がユーザの期待と乖離するケースも見られた。

今後の課題として、より微細な感情ニュアンス（複合感情）に対応したパラメータ調整規則の精緻化、およびシステム全体のレイテンシ短縮が挙げられる。また、長期的な対話における文脈理解や、ユーザ個人の選好に合わせたパーソナライズ機能の実装も検討が必要である。これらを通じて、実カウンセラーの代替となり得る、より高度な信頼性と親密性を備えた AI カウンセリングシステムの構築を目指す。

参考文献

- [1] Marah Abdin, Jyoti Aneja, Harkirat Behl, Sébastien Bubeck, Ronen Eldan, Suriya Gunasekar, Michael Harrison, Russell J. Hewett, Mojan Javaheripi, Piero Kauffmann, James R. Lee, Yin Tat Lee, Yuanzhi Li, Weishung Liu, Caio C. T. Mendes, Anh Nguyen, Eric Price, Gustavo de Rosa, Olli Saarikivi, Adil Salim, Shital Shah, Xin Wang, Rachel Ward, Yue Wu, Dingli Yu, Cyril Zhang, and Yi Zhang. Phi-4 technical report, 2024.
- [2] 市村真衣, 久野雅樹. 感情カテゴリを考慮した単語極性の推定. 言語処理学会 第 29 回年次大会 発表論文集, 2023.
- [3] 井手竜也, 河原大輔. 生成と分類のマルチタスク学習による感情が考慮された対話応答生成. 言語処理学会 第 27 回年次大会 発表論文集, 2021.
- [4] 宮川由衣, 加藤大貴, 松平茅隼, 平山高嗣, 駒水孝裕, 井手一郎. 発話の交換を考慮した対話システムにおけるユーザ感情推定手法の検討. 言語処理学会 第 29 回年次大会 発表論文集, 2023.
- [5] 頼展韜, 高橋誠史. Rina: マルチモーダル情報を利用したキャラクターの感情推定. 言語処理学会 第 28 回年次大会 発表論文集, 2022.
- [6] 田辺ひかり, 小川哲司, 小林哲則, 林良彦. コモンセンス知識を利用した物語中の登場人物の感情推定. 言語処理学会 第 27 回年次大会 発表論文集, 2021.