



LLM を用いた会話の感情判別に連動させたアニメ的演出による 2 次元 AI キャラクター『こはね』の創作

Creation of the 2D AI Character 'Kohane' with Anime-Style Effects Linked to Emotion Detection in Conversations using LLM

原田海夢¹⁾, 児玉幸子¹⁾

Kaimu HARADA, Sachiko KODAMA

1) 電気通信大学 情報理工学研究所 (〒182-8585 東京都調布市調布ヶ丘 1-5-1, h2430106@gl.cc.ucc.ac.jp, kodama@inf.ucc.ac.jp)

概要: 本研究では, 2 次元キャラクターに LLM を用いてアニメ的演出を連動させる手法を構築し, 『こはね』を創作した. アニメキャラクターは人間との音声またはテキスト入力を用いる会話における感情の状態に応じて漫符や背景, 構図などの演出と共に表示される. こはねの発話と心内文として表示される思考の各テキストに会話の感情に関する LLM の判別結果をアノテーションし, リアルタイムに 174 通りのアニメーション表現と連動させた.

キーワード: AI キャラクター, イラスト, アニメーション, コミュニケーション, LLM

1. はじめに

多くの人が夢見てきたアニメキャラクターとの対話は, 近年の大規模言語モデル (LLM: Large Language Models) の進化により現実味を帯びてきている. 実際, エンターテインメントのための AI キャラクターが多く登場してきている[1][2]. しかし, 現在開発されている多くの AI キャラクターでは, 写実的な表現で人間らしく自然に振る舞わせることが主な目的となっており, アニメキャラクターとの対話が迫られているとは言い難い.

アニメらしさを再現するためには, キャラクターの感情や状況に応じた表現技法, 例えば漫符や背景, 構図などの演出が重要である. アニメでは, キャラクターが怒っているときには頭に怒りマークを表示し, 背景に炎のエフェクトを加え, 顔が大きく映るような構図を用いることで, 視聴者にその感情を強く伝えている. また, キャラクターが喜んでいるときにはジャンプをする動きや, 感情が高まっているときには集中線を画面に加えるなど, 視覚的に豊かな表現が多用される. 加えて, キャラクターの心情を内面的に表現する「心内文」も, 視聴者の共感を引き出す重要な要素となっている.

そこで本研究では, AI キャラクターの感情がアニメらしく自然に伝わる対話を実現するために, LLM を用いてキャラクターの発話内容と心内文にアニメ的演出を連動させる手法を提案する. 具体的には, キャラクターの各会話

内容の感情を LLM で分析し, その結果に基づいてアノテーションする. そして, このアノテーションに基づいて, リアルタイムでキャラクターの喜びや怒りなどの感情を演出するアニメーション表現と連動させる.

本手法を検証するために, 2 次元 AI キャラクター『こはね』を作成し, こはねの発話内容と心内文に応じたアニメ的演出の自動生成システムを開発した.

2. 関連研究

石井らは, 既存のアニメキャラクターと対話することができるテキスト対話エージェントシステムの開発を行った[3]. このシステムでは, そのキャラクターらしさを再現するために固有モーションに注目し, 発話内容からルールベース方式で導入している.

Hasan らは, LLM を用いて感情を反映した 3 次元仮想エージェントと対話することができるアプリケーションを開発した[4]. 仮想エージェントの感情を LLM によって 7 つ (中立, 喜び, 悲しみ, 恐れ, 怒り, 驚き, 嫌悪) に分類し, 表情や音声に対応するものに制御している.

本研究では, これらの研究を基に LLM でキャラクターの発話内容と心内文から, 感情, モーション, エフェクト, リアクションに関する情報を判別し, それらに対応する演出とともにキャラクターのアニメーションを表示する. 本手法では, 既存の研究では困難であった①基本感情以

外の複雑な感情の判別, ②複数の判別結果の組み合わせ, ③会話以外の要素の判別を可能とした. また, In-Context Learning で LLM を活用するため, 拡張がしやすくなっている.

3. 『こはね』の創作

3.1 作品の概要

本作品では, キーボード入力またはマイクからの音声入力を通じてこはねと会話を行うことができる. また, 画面横のボタンを押すことで, 会話以外のコミュニケーションもとることが可能である.

Unity 上で開発をし, こはねはLive2Dモデルで作成した. 音声入力は Python 上で行い, 音声認識には Whisper API[5]を使用した. LLM には GPT-4o[6]を使用し, こはねの音声には VOICEVOX 春日部つむぎ[7]を使用した.

こはねの性格は天真爛漫でポンコツ, 好きなものはドーナツ, 固有モーションはウインクと設定した[8].



図 1 こはねとの会話画面

3.2 テキスト生成

キャラクターの会話内容と演出を連動させるため, 各会話内容に感情判別情報を付与したテキストを生成する必要がある. しかし, このテキストを直接 LLM で生成することは困難である. そこで, 本研究では会話生成モジュールとアノテーションモジュールの 2 ステップで生成する手法を提案する. このアプローチにより, 生成されたテキストに対してより詳細な感情判別情報を付与することが可能となり, 表 1 に示す wink や aseri などといった複雑な感情の判別, 「ウインクをする」や「押し黙る」などなどといった発話以外に関する判別を実現した.

3.2.1 会話生成モジュール

会話生成モジュールでは, キャラクターの発話内容と心内文を生成する. LLM を使用し, キャラクターのプロンプトを In-Context Learning させることで, キャラクターらしい自然な回答を生成する. また, このプロセスでは発話内容, 心内文, その他のことを区別するための制御も行う. 具体的には, 心内文は () で囲んで生成し, 会話以外のことは [] で囲んで生成するように制御している. このようにすることで, 会話の中での各要素が明確に識別可能となり, 後続のアノテーション作業が容易になる.

表 1 感情判別情報一覧

state	mindstate	reaction	effect	motion
neutral wink/WINK smile/SMILE anticipation /ANTICIPATION Laugh/LAUGH Odoroki /ODOROKI angry/ANGRY aseri/ASERI proud/PROUD flush/FLUSH	neutral thinking happy anticipation angry aseri	wink flush realize silent	focus	bounce swing

表 2 テキスト生成例

入力	こはねは何が好き？
会話生成モジュール	(ふふん, こはねの好きなこと, いっぱい知ってほしいな!)えっとね, こはねは甘いものが大好きなんだよ!特にドーナツにはこだわりがあるの.でも, それだけじゃなくて, ねこも大好きなの!見つけたら絶対に話しかけちゃうもんね.そうそう, 褒められるのもすごく嬉しいんだよ!褒められるとすぐに自慢しちゃうかも.[ウインクをする]君は何が好きなの?教えて教えて!
アノテーションモジュール	(([mindstate]anticipation[mind]ふふん, こはねの好きなこと, いっぱい知ってほしいな!), ([state]smile[text]えっとね, こはねは甘いものが大好きなんだよ!特にドーナツにはこだわりがあるの.), ([state]smile[text]でも, それだけじゃなくて, ねこも大好きなの!見つけたら絶対に話しかけちゃうもんね.), ([reaction]wink), ([state]WINK[text]そうそう, 褒められるのもすごく嬉しいんだよ!褒められるとすぐに自慢しちゃうかも.), ([state]WINK[text]君は何が好きなの?教えて教えて!))
入力	こはねが好きだよ
会話生成モジュール	(きゃー!すごく嬉しい!こはねも君のこと大好きだよ!)ほんとに?えへへ, 嬉しいよ![恥ずかしくて赤面になる]こはねも君のこと大好きだよ!一緒にたくさん遊んで, いっぱいお話ししようね!
アノテーションモジュール	(([mindstate]happy[mind]きゃー!すごく嬉しい!こはねも君のこと大好きだよ!), ([state]smile[text]ほんとに?えへへ, 嬉しいよ!), ([reaction]flush), ([state]FLUSH[text]こはねも君のこと大好きだよ!一緒にたくさん遊んで, いっぱいお話ししようね!))

3.2.2 アノテーションモジュール

アノテーションモジュールでは, 会話モジュールで生成された各会話内容に感情判別情報を付与する. このプロセスでは, まずプロンプトの制御により, 会話内容を小会話ごとに () で囲む. その後, 各会話内容に対して感情判別を行い, 感情判別情報を追加する. 感情判別は, 発話時の状態を表す[state], 思考時の状態を表す[mindstate], リアクションを表す[reaction], アニメ的エフェクトを表す[effect], 全身の動きを表す[motion]に対して行う. 各項目の感情判別情報一覧は表 1 に示す. 各小会話の中では必ず[state], [mindstate], [reaction]のいずれか1つの項目を追加する. [state], [mindstate], [reaction] は重複することができず,

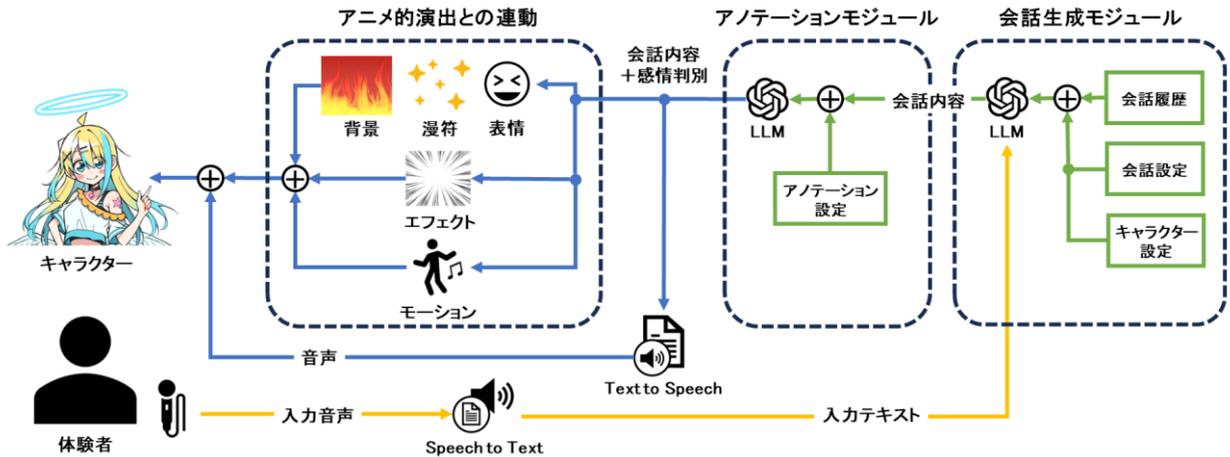


図2 システム図

各小会話において一意の状態を表す。[effect], [motion]は他の項目と重複して追加することができ, [state], [mindstate], [reaction]の後に必要に応じて追加する。[state]においては、通常は全て小文字の感情判別情報を追加し、感情が高まっている場合は全て大文字の感情判別情報を追加する。この方法により、感情の強弱を区別している。したがって、感情判別情報の組み合わせは表1から(19+6+4)×2×3の174通り存在し、複雑な感情表現が可能となる。発話内容は[text]の項目、心内文は[mind]の項目に追加する。表2にテキストの生成例を示す。

3.3 アニメ的演出との連動

3.2で生成したテキストを演出と連動させる。具体的には、[state], [mindstate], [reaction]での感情判別情報に応じて表情の変化、漫符の表示、背景演出の表示を行う。表情はLive2Dパラメータを制御することで作り出す。漫符はLive2Dアニメーションを再生し、背景演出は背景画像を表示する。また、[effect], [motion]の感情判別にに応じてエフェクトとモーションを追加する。エフェクトはエフェクト画像の表示とカメラの調整などを行い、モーションはLive2Dアニメーションを再生し、アイドルリングモーションと重ね合わせる。演出との連動例を図3に示す。

4. 実験

4.1 実験手法

本研究で提案したアニメ演出との連動がAIキャラクターの印象にどのような影響を与えるのかを調査した。被験者にアニメ演出が連動している場合としていない場合のこはねと対話をしてもらい、対話後にアンケートを行った。被験者の半数はアニメ演出が連動なし→連動ありの順、残りの半数は連動あり→連動なしの順で対話を行った。アンケートは7段階のリッカート尺度で回答してもらい、数字が小さいほど「当てはまらない」、7に近づくほど「当てはまる」とした。

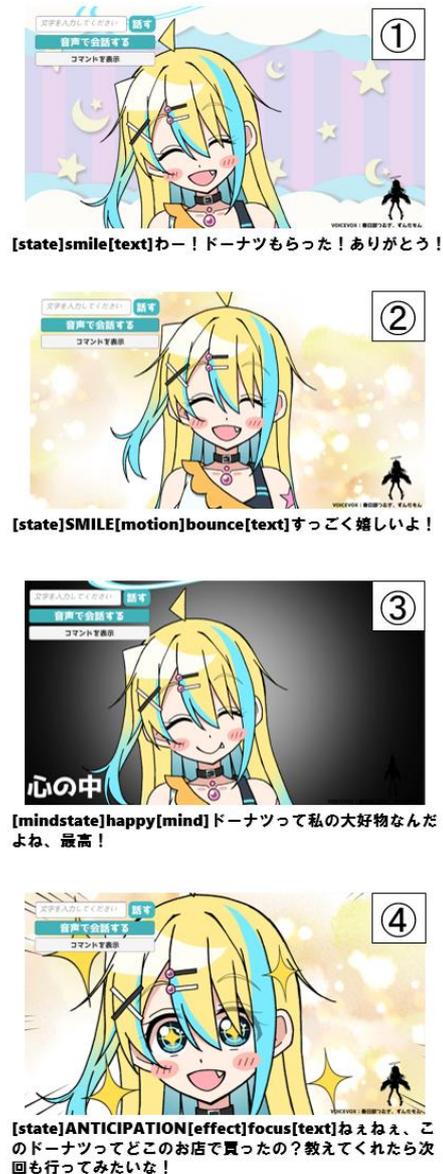


図3 アニメ演出との連動例
①～④は体験者がこはねにドーナツを上げた直後の発話と心内文の順番を示している

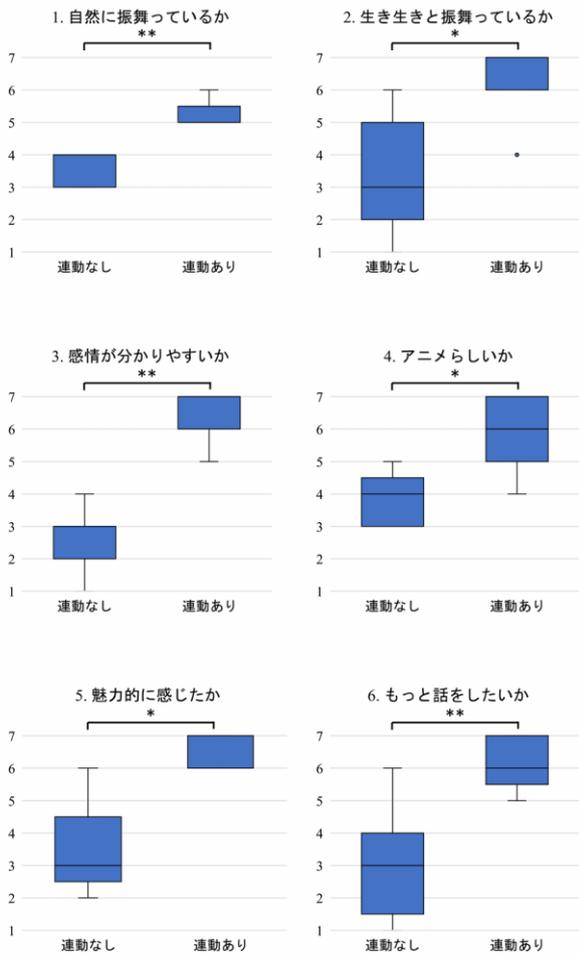
4.2 実験結果と考察

大学生9名を対象として行った。表3にアンケート項目、図4にアンケート結果の箱ひげ図を示す。対応のある2群のデータ間分析であり、サンプル数9とデータに正規性も見られないため、Wilcoxonの符号付順位和検定による分析を行った。p値が0.05未満の設問を有意差ありとしたところ、全ての設問で有意差が見られた。特に設問1, 3, 6ではより強力な有意差が見られた。

この結果から、本研究で提案したアニメ演出との連動手法によってAIキャラクターの自然さとアニメらしさが向上することがわかった。また、これによりキャラクターの魅力をもっと引き出せることも分かった。

表3 アンケート項目

	設問
1	自然に振舞っているか
2	生き生きと振舞っているか
3	感情が分かりやすいか
4	アニメらしいか
5	魅力的に感じたか
6	もっと話をしたいか



*p<.05, **p<.01

図4 Wilcoxonの符号付順位和検定による分析結果

5. おわりに

本研究ではアニメキャラクターらしい感情表現の演出を伴うAIキャラクターと自然に会話を行うこと目的とし、LLMを用いてキャラクター会話の中の感情を分析し、感情に合わせたアニメ的演出を連動させる手法を提案した。この手法では既存研究で困難であった①基本感情以外の複雑な感情の判別、②複数の判別結果の組み合わせ、③会話以外の要素の判別を可能とした。

研究では本提案システムを開発し、2次元AIキャラクターこはねを創作した。こはねを用いた実験の結果、提案手法によるアニメ演出との連動によってAIキャラクターの自然さとアニメらしさが向上し、キャラクターの魅力をもっと引き出せることが確認された。このことから、提案手法がアニメキャラクターとの対話の実現につながることを期待される。今後はより多くのアニメ演出との連動や、キャラクターの自律した行動制御を行うためのAIエージェントの実装を行い、生きているかのように感じられるアニメキャラクターを創り、彼らとの対話を実現することを目指す。

参考文献

- [1] Gatebox 株式会社, Gatebox, <https://www.gatebox.ai/>
- [2] 株式会社 Pictoria, 紡ネン, <https://tsumuginen.com/>
- [3] Ryo Ishii, Ryuichiro Higashinaka, Koh Mitsuda, Taichi Katayama, Masahiro Mizukami, Junji Tomita, Hidetoshi Kawabata, Emi Yamaguchi, Noritake Adachi, and Yushi Aono, “Methods of Efficiently Constructing Text-dialogue-agent System using Existing Anime Character”, Journal of Information Processing, Vol. 29, pp. 30–44, Jan. 2021.
- [4] Masum Hasan, Cengiz Ozel, Sammy Potter, and Ehsan Hoque, “SAPIEN: Affective Virtual Agents Powered by Large Language Models”, 2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos, 2023.
- [5] OpenAI, Whisper, <https://openai.com/index/whisper/>
- [6] OpenAI, GPT-4o, <https://openai.com/index/hello-gpt-4o/>
- [7] Hiroshiba Kazuyuki, VOICEVOX 春日部つむぎ, <https://voicevox.hiroshiba.jp/>
- [8] 原田海夢, “LLMを用いた個性的なAIキャラクター『こはね』のデザインに関する研究”, 令和5年度電気通信大学卒業論文