

医療面接教育のための仮想模擬患者を用いた 没入型音声対話システム

Immersive Spoken Dialogue System with Virtual Simulated Patient for Medical Interviewing Education

吉田 拓実^{1*} 上乃 聖¹ 李 晃伸¹
Takumi Yoshida¹ Sei Ueno¹ Akinobu Lee¹

¹ 名古屋工業大学

¹ Nagoya Institute of Technology

Abstract: 医学生に対する医療面接教育が重要視されているが、教育に用いられる模擬患者の数は不足している。本研究では音声対話技術により模擬患者とのやりとりを仮想で再現する医療面接シミュレータの構築をめざす。没入型 VR デバイスと高精細な模擬患者アバターを組み合わせることで、現実感の高い医療面接空間を再現する。また、マルチモーダル LLM を活用し仮想模擬患者との医療面接を実現する。実験として医療従事者によるシステム評価を行った結果、没入感や応答の柔軟さ、対話のテンポに肯定的な意見が見られたが、応答の長さや相槌の工夫、音声の単調さといった課題が明らかになった。

1 はじめに

医療において医療面接と呼ばれる患者とのコミュニケーションが重要視されている。近年、医学生の能力を測る客観的能力試験 (OSCE) にも医療面接は導入されており、医学生に対する医療面接教育の重要性は増している。この教育において模擬患者が重要な役割を担っている。模擬患者は教育を目的に患者役を演じる人のことを指し、医療面接における対話内容を再現するとともに、声色や身振り手振りといった非言語的な情報も再現する。

しかしながら全国で模擬患者の不足や高齢化 [1] が課題となっている。特に近年の医師の働き方改革により、模擬患者を養成する時間が確保できない医師が増加しており、この状況が模擬患者の不足をより深刻化させる恐れがある。また、模擬患者に知識や技術の格差がある [2] ことで模擬患者間に演技能力のばらつきが生じることがあり、医学生の練習に悪影響を及ぼす可能性が指摘されている。

これらの問題を解決するため、仮想模擬患者 (Virtual Simulated Patient : VSP) と呼ばれる、模擬患者を仮想空間で再現するシステムに関する研究が行われている。これにより、模擬患者が不足している場所でもシステムを利用することで模擬患者を用いた教育が実施できると期待されている。

この仮想模擬患者に関して、文章を用いてやりとりする対話システムとしての形は多く研究されているが、音声や 3D キャラクタによるジェスチャなどを用いるものについての研究は数が少なく、練習で利用できるほど実際の医療面接に近いリアルさがあるとは言えない。また仮想空間から得られる没入感が、学習に効果的な影響を与えることが示唆されている [3]。

そこで、これらを踏まえた医療面接教育で基盤として利用可能な仮想模擬患者システムの研究が始まっている。これは、従来の仮想模擬患者システムより実際の医療面接環境に近い状態で、質の高い練習を実施することを目標としたものである。本研究では音声対話技術による模擬患者とのやりとりを VR デバイスを用いて仮想で再現する医療面接システムの構築をめざす。そのために、没入型 VR デバイスと高精細な模擬患者エージェントを組み合わせ、没入感の高い医療面接環境の再現を行う。

2 関連研究

仮想模擬患者の研究の1つとして、オハイオ州立大学の研究がある。この研究では「ChatScript」という対話管理システムと図1に示す3Dキャラクターを用いて、iPadやディスプレイの画面上に模擬患者を再現している [4]。また、図2に示す広島大学が開発した「VR OSCE」 [5] というシステムがあり、こちらはVR上で

*連絡先: 名古屋工業大学
愛知県名古屋市昭和区御器所町
E-mail: t.yoshida.479@stn.nitech.ac.jp



図 1: オハイオ州立大学で作成された仮想模擬患者



図 4: システム画面の様子



図 2: VR OSCE

OSCEにおける技能を練習できるシステムである。この技能の中には医療面接も含まれ、ユーザは画面に表示された選択肢から適切な応答を選択して会話を行うことで医療面接対話を再現している。また、近年ではLLMを活用したVSPの研究が登場した。そのうちの1つが図3に示すSynthetic Patientである[6]。この研究では、LLMとしてGPT-4を用いた音声対話システムを構築し、患者の外見をText to Imageにより作成している。また、作成した患者の画像を用いて、動画を作成することで、画面内の患者役と対話を可能にする。



図 3: Synthetic Patient

3 仮想模擬患者を用いた没入型音声対話システム

3.1 システム概要

本研究では、没入型VRデバイスであるMeta Questと高精細な模擬患者アバターを組み合わせることで現実感の高い医療面接空間を再現することに加え、音声対話技術による高速で柔軟な医療面接対話の再現を行う。図4に示すような医療面接空間と模擬患者モデルをMeta Questの画面で表示しており、さらに環境音の導入や仮想の手の表示によって没入感を高めている。

模擬患者の3Dモデル、および診察室の3Dモデルは実際の模擬患者や藤田医科大学の診察室の撮影資料をもとに作成されたものである。また、模擬患者モデルについては、リップシンクによる口の動きの再現や、入退室の動作、表情変化を行うことによって従来のシステムより人間らしい表現を可能にしている。本研究においてはすべての症状の動作を再現することが困難であったため、腹痛に限定した再現を行っている。

医療面接対話の再現に関して、本研究ではOpenAIが提供する低遅延マルチモーダル対話用のAPIである「Realtime API」を用いている。低遅延であるため従来のシステムよりも高速な応答が可能になることで、医療面接における対話のテンポの再現を行っている。また、応答内容の制御に関しては、3.2節で詳細を述べる。

3.2 対話内容の制御

対話内容の制御はRealtime APIで使用するInstructionを用いて行った。Instructionは、[7]に示された医療面接の流れをもとに作成した。これは医療面接には決まった流れが存在しているため、その流れに沿った対話を実現する必要があるからである。具体的な流れ

としては、まず患者の入室から着席・自己紹介に続いて症状を伝達する。その後、患者の背景や解釈モデルについて伝達する。そして、睡眠やアレルギーなどのシステムレビューを確認したのち、伝えた内容が間違っていないか確認を行って患者は退出する。

また、患者にあった音声の再現および、模擬患者に渡される患者シナリオに沿った対話の再現のための Instruction も作成した。音声に関しては本研究で用いる 3D モデルに合わせて、「老人のようなしゃがれた音声」であることや音声の速度、音声の高さ、そしてどのような感情を表出するかについて指示を行った。患者シナリオは患者の氏名や主訴、現病歴や解釈モデルなどの内容を記載しているものであり、実際の模擬患者が用いるものをそのまま LLM にも Instruction として与えている。なお、解釈モデルは患者の病状の対する解釈を示したもので、病気の原因や治療の意思を示している。

3.3 システム構成

構築したシステムの構成を図 5 に示す。システムの流

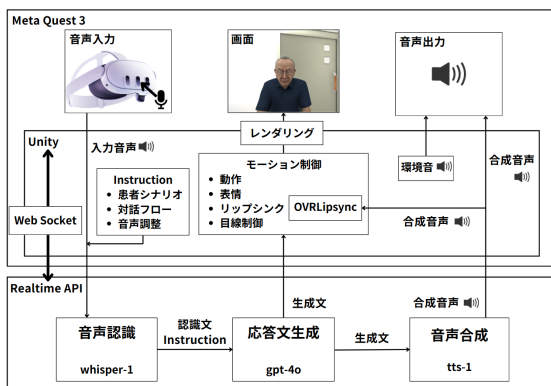


図 5: 全体構成

れは以下のとおりである。まず、ユーザが VR ヘッドセットを被った状態で発話することで、Meta Quest 3 のマイクに音声が入力される。入力した音声は WebSocket 通信により Realtime API へ送信される。API は事前に渡されたシナリオ情報および対話の流れの情報が設定された Instructions に基づいて応答内容を決定する。この際に応答は文章と音声の両方で受け取る。受け取った音声情報は Unity の関数により再生されると同時に、OVR Lipsync へ送信される。これにより、仮想模擬患者モデルは音声に合わせてリップシンクを行うことができる。一方で受け取った文章情報については、文章中にキーワードが含まれる場合、モーションの再生のオンオフや、表情の切り替えを行うようにしている。こ

れに加え、自動瞬きと視線追従を別で動かすことで、仮想模擬患者の動作を制御している。

4 評価実験

4.1 実験タスク・条件

医学生に対して医療面談の指導を行っている藤田医科大学の医療従事者 3 名に対して、構築したシステムを用いた 10 分程度の医療面接対話を実施した。実験場所は藤田医科大学の人通りの少ない静かな部屋である。なお、実験では男性仮想模擬患者システムを利用した。図 6 に実験風景を示す。



図 6: 実験風景

4.2 評価方法

医療面接対話を行ったのち、医療従事者にインタビューを実施した。表 1 に示すように評価項目は大きく分けて、環境や没入感に関するもの、対話の進行に関するもの、表情・動作に関するもの、音声に関するものである。

4.3 実験結果・考察

模擬患者の外見や診察室の空間のリアリティは十分高く、医療面接空間として没入するうえで十分であることが分かった。また、3D 模擬患者の素顔は実際の人間のようで十分にリアルであり、医療面接で医師が感じる緊張感についても十分な再現ができていたことが確認できた。

対話の進行に関しては十分高速な応答ができていたほか、患者シナリオに乗せていない事柄に対しても柔軟な対応ができていた。しかしながら、「ええ」、「そうですね」といった相槌をすることでシステムが反

表 1: 評価実験のインタビュー項目

番号	項目
i-1	医療面接における緊張感ほどの程度感じられたか
i-2	没入した空間は医療面接を練習する環境として十分か
ii-1	対話の流れはどの程度医療面接に基づいていたか
ii-2	対話のテンポ（リズム）は適切であったか
ii-3	症状の表出回数は適切であったか
iii-1	仮想模擬患者の表情はどの程度対話に合っていたか
iii-2	仮想模擬患者の視線に違和感があったか
iii-3	仮想模擬患者の動作はどの程度対話に合ったか
iv-1	仮想模擬患者の音声は表示されているモデルに対して合っていたか
iv-2	発話中の音声の速度はどうだったか
iv-3	マイクで話し終わってからシステムの音声に戻るまでの時間はどのようだったか

応してしまい対話の進行が崩れることがあった。これは、システムが常に評価者の音声を受け取る設定にしていたことが原因である。さらに、1回の発話が長すぎるという評価が目立った。これは、現在の質問と関係していないことまで補足してしまうことや、質問を復唱して応答することで応答が長くなってしまったことが原因である。評価者によると、「医者が患者から言葉を引き出すことも医療面接における重要な能力である」ため、応答は必要な情報を含んだ範囲でできるだけ短くすることが望ましいという結果になった。

また、表情・視線や動作に関しては本研究の範囲では十分に用意することができていないため、対応不足という評価になった。表情に関しては考えているときに表出するしかめつらを用意すること、視線は目を伏せるタイミングと目を合わせるタイミングを適度に調整すること、そして動作に関しては、考えるために顎に手を当てる動作や、患者側の相槌動作、下を向く動作を用意することが求められる。

最後に音声に関しては仮想模擬患者のモデルに対する音声が入っている評価がされた一方で、音声の単調さや抑揚の無さという問題が明らかになった。これは、評価者によると「音声は症状を判断するために重要なものであるため、音声が単調であると正しく情報収集ができないゆえ気になった」からであった。

5 むすび

本研究では、音声対話技術により模擬患者のやりとりを仮想で再現する医療面接システムの構築をめざし、没入型 VR デバイスと高精細な模擬患者エージェントを組み合わせることで、現実感の高い医療面接空間を再現した。評価実験の結果、外見や診察空間のリアリティに関しては十分に高い評価を得られ、十分な没入感や緊張感を獲得することができた。また、発話に関して、シナリオにない情報を対話の流れに沿って補完

できたり、十分に高速な応答ができていたことが分かった。しかしながら、応答の長さや、相槌についての考慮不足、患者の音声の単調さや抑揚の無さといった問題が明らかになった。

今後は実際の医療面接対話のデータを入手し、発話の長さの改善や相槌の際の対話制御を行うほか、愁訴感情をリアルに再現する模擬患者音声の作成による音声の課題の解決が考えられる。さらには、実際の医学生による大規模な評価実験や、医療面接の自動評価を行うシステムとの統合も見据えてシステムの改善を重ねることが考えられる。

謝辞

本研究は JSPS 科研費 24H00170 「AI による対話技術を活用した模擬患者アバターでの仮想空間医療面接教育基盤の創生」によるプロジェクトのもと行われました。本研究を進めるにあたって、全面的な協力をいただいた藤田医科大学の医師の石原慎様、荒川敏様、寺澤晃彦様をはじめとする関係者の方々に深く感謝申し上げます。

参考文献

- [1] 医道審議会医師分科会医学生共用試験部会.: 令和7年度以降の共用試験に関する意見, 2024
- [2] 中村もとゑ, 山崎歩, 渡邊聡美, 鈴木香苗, 眞崎直子.: 看護系大学における模擬患者の養成および活用の現状と課題, 日本赤十字広島看護大学紀要, Vol.16, pp.29-38, 2016
- [3] Kleinert Robert, Wahba Roger, Chang De-Hua, Plum Patrick, Hölscher Arnulf H, Stippel Dirk L.: 3D Immersive Patient Simulators and Their Impact on Learning Success: A Thematic Review, *J Med Internet Res*, Vol.17, No.4, e91, 2015
- [4] Maicher Kellen, Danforth Douglas, Price Alan, Zimmerman Laura, Wilcox Bruce, Liston Beth, Cronau Holly, Belknap Laurie, Ledford Cynthia, Way David, Post Doug, Macerollo Allison, Rizer Milisa.: Developing a Conversational Virtual Standardized Patient to Enable Students to Practice History-Taking Skills, *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, Vol.12, No.2, pp.124-131, 2017

- [5] 佐伯 勇, 服部 稔, 栗原 将, 児島 正人, 本田 有紀子, 蓮沼 直子, 高橋 信也, 粟井 和夫, 檜山 英三.: VR システムを利用した OSCE 教材「VR OSCE」の開発, 日本シミュレーション医療教育学会雑誌, Vol.11, pp.108-111, 2023
- [6] Chu Simon N, Goodell Alex J.: Synthetic Patients: Simulating Difficult Conversations with Multimodal Generative AI for Medical Education, *arXiv preprint arXiv:2405.19941*, 2024
- [7] 伴信太郎, 鈴木富雄, 青松棟吉, 西城拓也, 阿部恵子, 桑島愛.: よくわかる医療面接と模擬患者, 名古屋大学出版会, 2011

付録

本研究で使用した 3D モデル



図 7: 模擬患者の 3D モデル

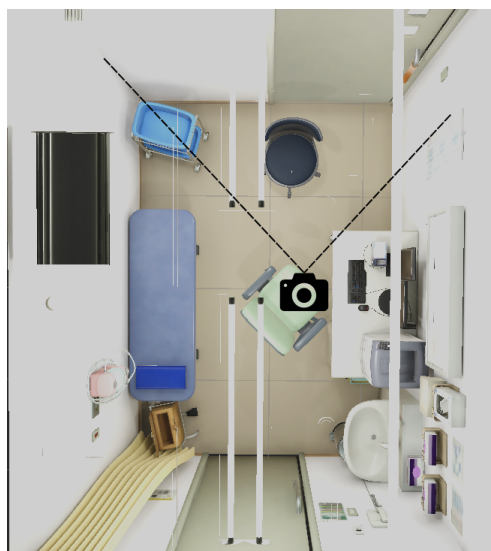


図 8: 診察室の 3D モデル