

2025 年 7 月 30 日

報道関係者各位

国立大学法人筑波大学

PTSD 患者のエクスポージャー療法を支援する 自然言語インタラクションによる聴覚 VR 生成システムを開発

自然言語インタラクションにより聴覚 VR を生成するシステムを開発し、PTSD 患者へのエクスポージャー療法への応用可能性を検証しました。臨床専門家の評価で、専門知識がなくても短時間で高品質な音響シナリオを作成でき、治療に十分な有効性と実用性があることが確認されました。

心的外傷後ストレス障害（PTSD）は、事故や災害、暴力などによる強いトラウマ体験が原因で起こる精神障害です。その治療法の一つがエクスポージャー療法で、高い効果があることが科学的に示されています。このエクスポージャー療法では、患者が恐怖の対象となる状況に段階的に触れていくことで、不安や回避反応を和らげていきます。

しかし、患者一人ひとりのトラウマに対応した再現環境の準備には手間がかかります。近年では仮想現実（VR）を用いたエクスポージャー療法も試みられていますが、患者ごとに VR をカスタマイズするにも時間と労力がかかることが課題となっていました。

本研究では、この問題に対処するため、大規模言語モデル（LLM）と音響データセットを活用し、自然言語入力からトラウマの音響を体験できる聴覚 VR を自動生成するシステムを提案しました。ユーザーが言葉でテーマを入力するだけで、それに合った音素材やシナリオが自動で生成されます。

臨床経験のある専門家にこのシステムを評価してもらったところ、使いやすく音の品質も高く、治療現場でも最小限の修正で利用できることが確認されました。特に、実験では、一つの音響 VR の作成にかかった平均時間は約 5 分（最短 50 秒）と、診療現場での活用を想定しても実用的な水準であることが示されました。

今後は、同時に複数の音を扱うような複雑なシーンにおける精度や制御を改善し、実用化に向けてさらなる検証を進めていきます。

研究代表者

筑波大学システム情報系

善甫 啓一 准教授

研究の背景

心的外傷後ストレス障害（PTSD）は、命の危険を伴うような体験により発症する精神障害です。エクスポージャー療法^{注1)}はその治療に高い効果があるとされていますが、実施の困難さや専門人材の不足により、臨床現場での普及は進んでいません。近年では、仮想現実（VR）を活用したエクスポージャー療法が注目されており、中でも視覚に比べて構築コストが低い聴覚 VR^{注2)}の活用が有望視されています。このため本研究では、自然言語入力^{注3)}によって誰でも簡単に治療用の聴覚 VR を生成できるシステムの開発と評価に取り組みました。

研究内容と成果

本研究では、大規模言語モデル（GPT-4）^{注4)}と音響データセットを組み合わせ、自然言語による入力から治療用の音響シナリオを自動生成する聴覚 VR システムを構築しました。ユーザーが「雨の日の交通事故」「夜の繁華街」「飛行機の離陸中」などのテーマを入力すると、システムは関連する音響素材をデータセットから検索し、選定された音の再生順序や空間的配置（座標）を含む 1 分間の音響シナリオを生成します。音響には HRTF（頭部伝達関数）^{注5)}を適用し、聴覚的臨場感を高めた空間音響として統合します。出力前には自然言語による修正指示が可能で、シナリオ内容や音がどの方向から聞こえるか（定位）を段階的に調整できます。

このシステムの臨床応用性を検証する実験では、PTSD や不安障害の臨床経験を持つ医師・心理士 3 人がこのシステムを実際に操作し、三つの異なるテーマごとに 3 回ずつ音響 VR を作成してもらいました。その結果、全ての試行において音響 VR の出力に成功しました。また、合計 9 ケースで出力された音響 VR を、「治療的有効性」「入力内容との一致」「音の明瞭さ」「空間定位」「シナリオの妥当性」「全体的な品質」の 6 項目について 7～10 段階で評価してもらった結果、平均 4.8 点（7 点満点）を記録しました。「飛行機搭乗」シーンなど満点を獲得した例もあり、多くが「軽度の修正で治療に使用可能」と評価されました。操作性に関する SUS スコア^{注6)}も平均 72.7 と高く、一部の試行では即時利用に近い高評価（82.5 点）も得られました（一般的に SUS が 85 点以上であれば“即時実用可能”とされます）。これにより、専門的な音響知識がなくても治療用音響を短時間で作成できる点が高く評価されました。

今後の展開

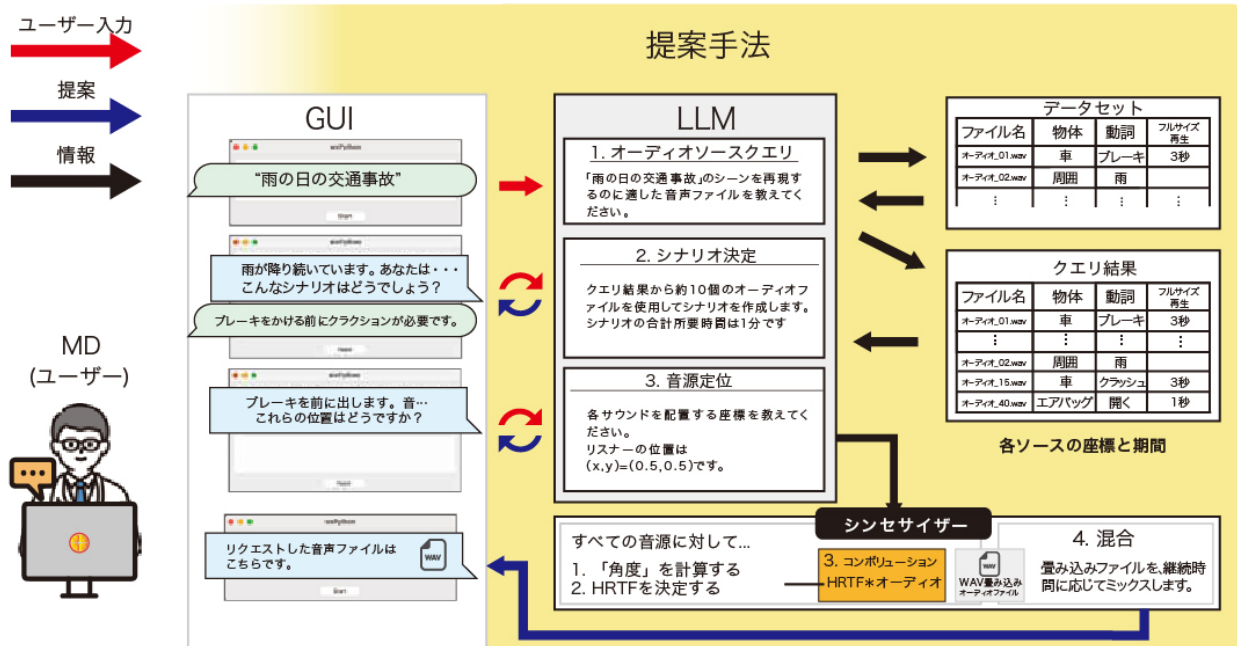
今回の実験では、すべての試行において音響 VR の出力に成功し、多くのケースで「軽微な修正で治療に使用可能」との高評価を得ました。一方で、交通事故など複数の音が短時間に重なる複雑なシナリオでは、音の重なりやタイミングのずれにより、再現度や没入感が低下する傾向が見られました。

そのため今後は、複数の音が同時に再生される場面でも、より自然なタイミングや重なりを再現できるよう、出力調整機能やタイムライン制御の精度向上が課題となります。

また、治療にあたる専門家が自然言語でシナリオを入力し、システムから提示されるテキストベースの構成案だけでは、実際にどのような音響体験が生成されるのかを事前に正確に把握することが難しいという課題も明らかになりました。そこで、今後の改善策として、再生前の試聴機能や音源配置の可視化 GUI の導入も検討しています。さらに、音響修正やシナリオ調整の自由度を高めることで、臨床現場での柔軟な活用が可能になります。

本システムの社会実装に向けては、今後、臨床現場での検証を重ねながら、効果の科学的根拠を蓄積し、実用性や導入可能性を段階的に確認していく予定です。

参考図



図：自然言語入力により音響 VR を自動生成する提案システムの構造

ユーザーが入力した文章に基づき、大規模言語モデルが適切な音響を選び、再生順・音源の位置・持続時間を設計する。最終的に HRTF（頭部伝達関数）により臨場感ある空間音響が合成され、エクスポージャー療法に使用可能な音響 VR として出力される。

用語解説

注1) エクスポージャー療法 (Exposure Therapy)

PTSD の主要な治療法の一つで、恐怖や不安を引き起こす状況に段階的に触れることで、徐々に反応を弱める心理療法。効果的な治療法として推奨され、保険適用もされているが、その実施には専門的な技能と環境が必要で、現場での活用は限定的となっている。

注2) 聴覚 VR (Auditory Virtual Reality)

音の方向や距離を仮想空間内で再現することで、没入感のある聴覚体験を提供する技術。視覚 VR より構築が容易で、PTSD 治療との親和性が高いとされている。

注3) 自然言語入力 (インタラクション)

専門用語やプログラミングを使わずに、人が日常会話のような自然な言葉でコンピューターとやり取りしながら指示を出せる入力方式のこと。

注4) 大規模言語モデル (LLM) GPT-4

膨大な文章を学習し、人間に近い自然な応答や生成ができる AI モデル。GPT-4 は米 OpenAI 社が開発した高性能な大規模言語モデルである。文脈理解・生成力に優れ、シナリオ作成や音響選定にも応用されている。

注5) HRTF (頭部伝達関数)

音が頭・耳・肩などに当たることで生じる音の変化を再現する音響技術。音の方向や距離感を立体的に感じさせるために使用される。

注6) SUS スコア (System Usability Scale)

ソフトウェアやシステムの使いやすさを数値で評価する国際的な指標。スコアは 0～100 点で、標準平均は 68 点とされている。

掲載論文

【題 名】 Development and Evaluation of an Auditory VR Generative System via Natural Language Interaction to Aid Exposure Therapy for PTSD Patients

(PTSD 患者のエクスポージャー療法を支援する自然言語インタラクションによる聴覚 VR 生成システムの開発と評価)

【著者名】 Yuta Yamauchi¹, Keiko Ino², Masanori Sakaguchi³, Keiichi Zempo⁴

¹筑波大学 理工情報生命学術院 システム情報工学研究群 知能機能システム学位プログラム

²国立精神・神経医療研究センター 精神保健研究所

³筑波大学高等研究院 (TIAR)、国際統合睡眠医科学研究機構 (IIIS)

⁴筑波大学 システム情報系 知能機能工学域

【掲載誌】 ACM Transactions on Computing for Healthcare

【掲載日】 2025 年 3 月 19 日

【DOI】 10.1145/3723048

問合わせ先

【研究に関すること】

善甫 啓一 (ぜんぼ けいいち)

筑波大学システム情報系 知能機能工学域 准教授

URL: <http://www.xpercept.aclab.esys.tsukuba.ac.jp>

【取材・報道に関すること】

筑波大学広報局

TEL: 029-853-2040

E-mail: kohositu@un.tsukuba.ac.jp