

ABCI 3.0開発加速利用 (2025年度) 成果概要 (公開用)

課題名：マルチモーダル実況生成とその評価

実施時期：2025年4月1日から2026年3月31日

所属機関名：産業技術総合研究所人工知能研究センター

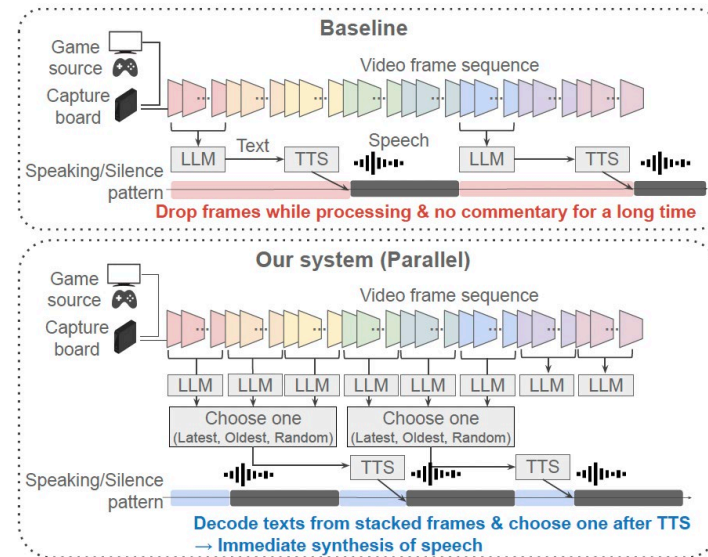
代表者氏名：石垣 達也

成果概要：

内閣府BRIDGEプロジェクトの一環として、2023年度から3年間、リアルタイム言語生成・リアルタイム実況生成に関する研究開発を実施した。ABCIは、ゲーム映像・構造化時系列データ・音声を用いた深層学習モデルおよび大規模言語モデルの学習・推論・評価に利用した。主な成果として、レーシングゲームにおけるリアルタイム実況生成システム、マルチモーダルLLMを用いたゲーム映像実況生成手法、実況音声合成に関する研究成果を得た。関連成果はIEEE CoG2024、LREC 2026、IJCAI2026、IEEE Access等に採択された。

成果のポイント：

本研究では、スポーツ・eスポーツ配信において視聴者の理解や楽しさを高めるため、映像・走行データ・テキスト・音声を統合したリアルタイム音声実況生成技術を開発した。レーシングゲームおよびゲームタイトルを対象に、速度、ステアリング操作、映像情報などをリアルタイムに取得し、試合状況に応じた実況文を生成するシステムを構築した。さらに、マルチモーダルLLMを用いてゲーム映像から実況を生成する手法や、実況音声合成により視聴体験を向上させる手法を検討した。ABCIを用いることで、大規模な深層学習モデルの学習・推論実験、多数条件での比較評価、マルチモーダルデータを用いたモデル開発を効率的に実施できた。これらの成果は、将来的にスポーツ配信、eスポーツ配信、ライブ配信プラットフォームにおける自動実況・多言語実況・個人化実況への応用が期待される。



成果についてより詳細な情報を提供しているWebページ、発表論文などの情報：

- ・ Ryosuke Matsushita et al. “Measuring Time Delay Tolerance in Third-Person Live Commentary for Game Recordings.” IEEE Conference on Games (CoG 2025).
- ・ Anum Afzal et al. “Real-Time Generation of Game Video Commentary with Multimodal LLMs: Pause-Aware Decoding Approaches.” LREC 2026.など