

信頼される AI の基盤技術
2021 年度採択研究代表者

2022 年度
年次報告書

飯塚 里志

筑波大学 システム情報系
准教授

実応用に向けた動画像コンテンツ加工のためのユーザ制御可能な例示ベース深層学習フレーム
ワークの確立

研究成果の概要

ユーザが制御可能な動画コンテンツ加工フレームワークの確立を目標とし、複数のタスクにおいて例示データを活用した深層学習フレームワークを構築した。グレイスケール動画像のカラー化タスクでは、前年度で検証した GAN inversion と呼ばれるアプローチとは別に、拡散モデルをベースとした画像のカラー化について研究に取り組んだ。その結果、汎化性能を向上させるためにバックボーンネットワークの構造や学習手法の改良が必要であるものの、特定のグレイスケール画像に対して良好な結果が得られることが確認できた。ただし、この手法はそのまま使うと計算コストが非常に大きい。今後は動画への応用も踏まえ、より多様なグレイスケール画像で良好なカラー化を行えるモデルの構築や高速化、動画のフレーム間での色の伝播手法の構築などに取り組んでいく予定である。

また上記のタスクとは別に、例示データを考慮した深層学習モデルを活用し、物体の経年劣化シミュレーションやラベルマップからの絵画生成、動画中の人物の表情や顔のポーズを編集できる手法の研究にも取り組んだ。画像中の物体の経年劣化をシミュレーションする研究では、まず錆や汚れなどの経年劣化テクスチャを深層生成モデルによって生成し、そのテクスチャを物体領域に合成することで、多様な物体の劣化を再現することができる。また、絵画生成の研究では、ユーザが山や空、草地などのレイアウトを指定し、それに合わせて写実的な絵画を生成する深層学習フレームワークを提案した。この手法では、新たな合成絵画データセットを作成し、多様な編集操作が可能な生成ニューラルネットワークモデルを構築している。さらに、動画の顔編集の研究では、表情や顔の向きや角度、目の開き具合などを属性ベクトルとして受け取り、それに合わせて人物の顔を生成する深層学習モデルを構築している。これらの研究成果は、国際論文誌および国際会議に投稿中である。

【代表的な原著論文情報】

- 1) Yuantian Huang, Satoshi Iizuka, Kazuhiro Fukui. "Free-view Expressive Talking Head Video Editing," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2023.