

研究報告書

「深層学習による高品質デジタル リマスター技術の創出」

研究期間：2018年4月～2020年3月
研究者番号：50202
研究者：飯塚 里志

1. 研究のねらい

本研究では、戦前の記録映像や白黒映画など古い動画像コンテンツを対象とし、経年劣化の修復や色彩の復元、さらには極端な低画質画像の高画質化や色調補正まで含めた、包括的なデジタルリマスター技術の創出に取り組む。多くの歴史的な映像は色情報のないグレースケールデータであり、フィルムの経年劣化によって画質も極端に劣化している。このような映像データのデジタルリマスターには、フィルムノイズやブロックノイズの除去や明度およびコントラストの補正、超解像、色彩復元など複雑な画像処理が必要であり、現在は専門技術をもつ熟練者が膨大な時間をかけて行っている。本研究では、この複雑なリマスター処理を計算機上で再現できる新たなニューラルネットワークモデルを構築し、そのモデルを効果的に学習できるフレームワークの確立を目指す。このためには、経年劣化をシミュレーションして教師データを構築するフレームワークの構築、および複雑な修復・補正処理および色彩復元を同時に行うことのできる大規模ネットワークモデルを構築する必要がある。この技術が確立できれば、実世界の様々な映像コンテンツを誰でも容易に高品質化できるようになり、これまで「過去の出来事」として認識されていた古い劣化映像をより身近なものとして蘇らせ、人々の新たな歴史認識や価値観の創造につながることを期待できる。さらに本技術は、教師データを用意することが難しく複雑な変換を必要とする動画像合成・編集技術に応用することが可能であり、画像中の物体を別の画像に自然に合成する技術や、手書きイラストからの自然画像を自動生成する技術など、新たなコンテンツ制作の支援技術の創出につながることを期待できる。特に本研究は実世界の画像データを扱うことを目標としているため、より実用的なアプリケーションに向けた基盤技術として実社会のコンテンツ産業に貢献できると考えている。

2. 研究成果

(1) 概要

古い映像の包括的な高品質デジタルリマスターを実現するため、動画の劣化修復を行えるニューラルネットワークモデルの構築、参照画像を利用した高精度の動画着色ネットワークモデルの構築、および劣化シミュレーションによる教師データセットの構築を実現した。また、それらのネットワークをまとめて単一のリマスターネットワークとして end-to-end で学習できるフレームワークを確立した。提案手法では、時空間畳み込み演算および temporal source-reference attention 機構から成るリマスターネットワークを構築することで、入力動画の様々な劣化を自動検出・修復し、さらに参照画像を用いてユーザの意図するカラー化を可能にしている。時空間畳み込み演算を用いることで、複数のフレーム間情報を計算できるため、高精度の修復およびフレーム間で一貫性を保持した着色が可能となる(図 1)。さらに、本

研究で新たに提案する temporal source-reference attention は、時系列特徴マップの行列計算によって入力動画と参照画像の類似度を計算し、任意の枚数の参照画像の色彩情報を大域的かつ同時に入力動画に反映させることができる。これにより、これまでの主流である再帰計算によるアプローチの様々な制約を解消し、少数の参照画像で長時間の動画を効果的に着色することができる。様々な検証実験を通じて、提案手法はこれまでのアプローチに対して顕著な精度向上を実現することを確認した。

この研究成果はコンピュータグラフィクス分野の最難関国際会議の一つである SIGGRAPH ASIA 2019 などに採択された[1,2,3]。また社会応用としても、実際に本技術を用いて 1959 年の伊勢湾台風の白黒映像のカラーリマスターを行った。このカラー化された映像は複数のテレビ番組で放送され[6,7]、多くの視聴者に過去の映像をより鮮明に伝える役割を果たし、大きな反響を呼んだ。この映像を特集した番組「伊勢湾台風 60 年 色と記憶」は第 61 回科学技術映像祭において文部科学大臣賞を受賞した。さらに、本技術は最先端の AI 技術として 2019 年 11 月より日本科学未来館の常設展示に採用されている。



図 1: 提案手法によるリマスター結果。6 枚の参照画像のみを用いて 700 フレームのリマスターを行っている。右に示されている画像が使用された参照画像である。左の 1 段目が入力動画であり、2 段目が劣化修復や画質補正の結果、3 段目がカラー化された結果である。提案手法は自動で劣化の修復から超解像、コントラスト補正などを行い、さらにユーザが用意した参照画像に合わせて色彩を復元することができる。

(2) 詳細

研究課題 A「時系列情報を考慮した動画の複合的劣化修復手法の確立」

古い映像の多様かつ複合的な劣化を自動修復するため、時系列情報を考慮して修復処理を行うことのできる畳み込みニューラルネットワークモデルを考案した(図 2 左)。この修復・補正ネットワークは、グレイスケールの入力フレームの様々なノイズや汚れを取り除き、さらにコントラスト補正や超解像まで行う。このモデルはすべての層が時空間畳み込み層で構成されるエンコーダ・デコーダ型のネットワークであり、入力と出力の間にスキップ接続をもつ。修復処理においては入力と出力でほとんど変化しない領域もあるため、入力と出力の間にスキップ接続を持たせることで、ネットワークは入力画像の変化箇所のみを学習するようになり、効果的な学習が可能となる。また、ノイズはフレーム間で変化することが多いため、時空間畳み

込みによって複数のフレーム情報を計算に取り入れることで、どこがノイズかを判断しやすくなり、精度を向上させることができる。このネットワークは既存手法と比較実験を行い、最も高精度であることを確認した。以上より、研究目的の一つである、古い映像の多様な劣化を自動かつ高精度で取り除くことのできるネットワークモデルの構築を達成した。

研究課題 B「ユーザが制御可能な動画の高精度カラー化手法の確立」

グレイスケール動画の高精度カラー化のため、本研究では参照カラー画像を利用した source-reference ネットワークを考案した。このネットワークは、修復・補正ネットワークの出力および任意の枚数の参照カラー画像を入力とし、参照画像の色彩を反映したカラー化を行う(図 2)。これにより、ユーザは少数の参照画像によって大まかな色彩を指定するだけで、動画全体のカラー化を行うことが可能となる。このネットワークでは、空間畳み込み演算や時空間畳み込み演算、2 種類の attention 機構を組み込むことで、時系列方向まで含めた大域的な情報を考慮して色度マップを計算できる。特に、本研究で考案した temporal source-reference attention は、ネットワーク内部の特徴マップ上で入力フレームと類似した箇所を全参照画像から計算することで、そのカラー情報を高精度で出力に反映することができる。従来の動画着色手法は再帰計算によってフレームを順次着色していくアプローチを取っているが、これは一度着色に失敗するとその後のフレームの着色結果が急速に悪化する上、シーンが切り替わるとうまくいかない。これに対し、提案手法は任意の枚数の参照画像全体から必要な情報を抽出し、入力動画全体に効果的に反映させることができる。これにより、長時間の動画着色を少数の参照画像で効率的に行うことができる。様々な種類の動画を用いて検証実験を行い、提案手法は既存手法を大きく上回る精度を達成することを確認した。

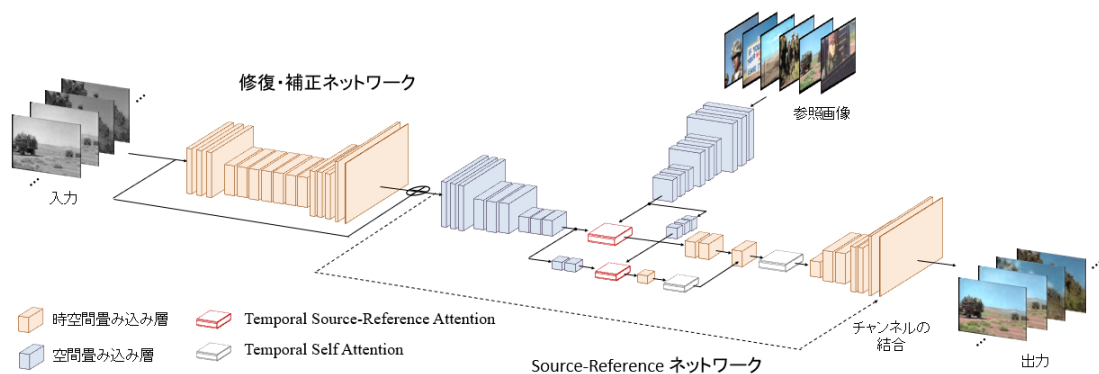


図 2: 提案モデル全体の概要図。提案モデルは修復・補正ネットワークと source-reference ネットワークから構成される。グレイスケール動画のフレーム群が入力されると、まず時空間畳み込み層で構成された修復・補正ネットワークによってフレームの修復処理が行われる。その後、修復されたフレームおよび任意の枚数の参照画像が source-reference ネットワークに入力され、色度マップが出力される。最後に、修復・補正ネットワークによって修復されたフレームを輝度チャンネルとして色度マップと結合することで、リマスターされたカラー動画が出力される。

研究課題 C「学習データセットの構築」

上記のネットワークモデルを学習させるため、本研究では高品質の動画を劣化シミュレーションによって昔の映像のように劣化させることで、劣化動画とその高品質版のペアを生成するアルゴリズムを構築した。まず、公開されている大規模動画データセットから、本タスクに適した動画約 1500 本(1000 万フレーム)を手動で収集した。ここで、本タスクに適した動画とは、動きのあるシーンや高画質で自然な色のシーンを撮影した動画であり、逆に適していない動画は白黒動画やテレビゲーム動画など不自然な色彩の動画、および固定カメラでほとんど動きのない動画などである。

この動画に対し、example ベースとアルゴリズムベースを組み合わせた劣化シミュレーションおよびデータ拡張により、多様な教師データを生成した。Example ベースでは、フラクタルノイズや砂状ノイズ、スクラッチ、ダストノイズなど、実際の映像に現れる様々なノイズデータをあらかじめ用意し、これにスクーリングや回転などのデータ拡張を施しながらランダムに組み合わせ、多種多様な劣化を生成する。アルゴリズムベースでは、圧縮ノイズやガウシアンノイズ、低解像度化、コントラストや明度の低下などによって画質を劣化させる。これらを組み合わせ、多様な劣化を自動生成し、様々な動画で効果的な劣化修復を学習できるようにした。敵対的学習を用いた教師なし学習なども検討したが、実際、提案手法のように人工的に教師ありデータを生成した方が顕著に精度が向上するため、今回のタスクではこのようなアプローチを採用した。

【研究目的の達成状況】

時空間畳み込み演算や新たな attention 機構を組み込んだニューラルネットワークモデルの構築や、劣化シミュレーションによる教師データの生成により、本研究の目標であった、包括的なデジタルリマスター技術の創出を実現した。この研究成果は CG 分野の最難関国際会議にも採択されたことから、学術的に重要な意義があったと考えられる。さらに、もう一つの目標であった社会応用についても、テレビ局と連携し本研究成果を用いて実際の白黒映像をカラーリマスターし、テレビ放送を通じて視聴者から大きな反響を得ることができた。このことから、実用的かつ社会的にもインパクトのある技術を開発できたと思われる。以上のことから、まだ技術的に改善の余地はあるものの、学術的・実用的両面でインパクトのある技術を開発でき、当初の研究目標を達成できたと考えている。

3. 今後の展開

本研究は、これまでは「過去の出来事」として認識されていた白黒映像を身近なものとして蘇らせ、人々の新たな歴史認識や価値観の創造につながるものが一つの目標である。実際に、本研究の成果によって過去の記録映像のカラーリマスターに成功し、視聴者から「昔の出来事が非常に身近に感じられた」、「衝撃を受けた」など大きな反響があった。今後は、参照画像として使用できる適切な画像をウェブやデータベースから自動抽出できるようなアルゴリズムを確立し、ユーザの負担なしに高品質デジタルリマスターを行えるようにしたい。これが実現できれば、より誰でも簡単に過去の映像を高品質化できるようになり、大量の映像のアーカイブをもつ出版業界やテレビ業界に新たなコンテンツ活用の可能性を示せると考えている。さらに学術的にも、参照ベースの様々な動画像処理タスクに対して、幅広い応用が期待できる。

4. 自己評価

・研究目的の達成状況

上記の通り、目標としていた過去の映像の多様な劣化修復・補正から色彩の復元まで含めた、包括的なデジタルリマスター技術の創出を実現できたため、研究目標は達成できたと考えている。今後はさらに手法を改良し、動きの激しいシーンでも良好にカラー化できるようにしたり、参照カラー画像をデータベースなどから自動で取得できるようにしたりするなど、より実用的な手法の確立を目指したい。

・研究の進め方(研究実施体制及び研究費執行状況)

基本的には計画通りに進めることができたと思われる。研究費もそのほとんどをGPUサーバの購入に充てることで、複雑で大規模な提案ネットワークモデルを効率的に学習することができ、その結果研究期間内に一定の目標を達成できたと考えている。

・研究成果技術及び学術・産業・社会・文化への波及効果

研究成果は最難関国際会議の一つである SIGGRAPH Asia に採択されたことから、学術的にインパクトがあったと思われる。さらに、実際に研究成果を用いて過去の悲惨な災害の記録映像をカラーリマスターし、テレビ番組を通じて現代に伝えることで、多くの人々の関心を得ることができた。この映像は第 61 回科学技術映像祭において文部科学大臣賞を受賞したことも含め、産業的・社会的にも大きなインパクトがあったと考えている。さらに、この研究成果は日本科学未来館の常設展示として採用されているため、今後さらに一般に波及していくことが期待される。

・研究課題の独創性・挑戦性

極端に劣化したグレイスケールの動画データを対象とし、その劣化箇所の修復から色彩情報の復元まで行う手法はこれまでに存在しない。動画像のノイズ除去や画質改善の技術はこれまでも提案されているが、これらは単純なガウシアンノイズやぼけの除去を対象としていることがほとんどであり、多様なパターンで劣化した実際の古い動画ではうまくいかない。また動画の着色では、従来の再帰型アプローチでは長期の着色が困難な上、劣化した映像では着色精度が極端に低下してしまう。これに対し本研究では、修復から色彩復元まで単一のフレームワークで行える深層学習ベースの手法を確立することで、これまで困難であった高精度のデジタルリマスターを実現することが、独創的かつ挑戦的な点である。

5. 主な研究成果リスト

(1)論文(原著論文)発表

1. Satoshi Iizuka and Edgar Simo-Serra. "DeepRemaster: Temporal Source-Reference Attention Networks for Comprehensive Video Enhancement", ACM Transactions on Graphics (Proc. of SIGGRAPH Asia 2019), 38, 9, 1-13, 2019.
2. Kazuma Sasaki, Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. "Learning to Restore Deteriorated Line Drawing", The Visual Computer (Proc. of Computer Graphics International 2018), 34, 6-8, 1077-1085, 2018.

(2)特許出願

研究期間累積件数:0件

(3)その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

3. 飯塚里志, シモセラエドガー. "DeepRemaster: Temporal Source-Reference Attentionを用いた動画のデジタルリマスター", Visual Computing 2019, 2019-6. (査読付き国内会議, VC賞(最優秀研究発表賞)を受賞)
4. 飯塚里志, "ディープラーニングによる対話的な画像処理", パターン認識・メディア理解研究会(PRMU), 2018/10/19. (招待講演)
5. 飯塚里志, "深層学習と画像編集", 可視化情報シンポジウム 2018, 2018/9/15. (招待講演)
6. 「伊勢湾台風 60 年, 色が呼びさます記憶」, CBC テレビ, 2019/9/26. (テレビ報道, そのほか TBS テレビ「NEWS23」「チャント!」「ゴゴスマ!」などで放送)
7. 「伊勢湾台風の被災映像、AIで初のカラー化」, 中日新聞第一面, 2019/9/26.