

論文 / 著書情報
Article / Book Information

題目(和文)	参照画像とテキストを用いたスケッチの自動着色法
Title(English)	Automatic Sketch Colorization using Reference Image and Text
著者(和文)	YANDingkun
Author(English)	Dingkun Yan
出典(和文)	学位:博士(工学), 学位授与機関:東京科学大学, 報告番号:甲第477号, 授与年月日:2025年9月22日, 学位の種別:課程博士, 審査員:齋藤 豪,小池 英樹,篠田 浩一,下坂 正倫,井上 中順
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Institute of Science Tokyo, Report number:甲第477号, Conferred date:2025/9/22, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名		Yan DingKun	
論文審査 審査員		氏 名	職 名		氏 名	職 名
	主査	齋藤 豪	准教授	審査員	井上中順	准教授
	審査員	小池英樹	教授			
		篠田浩一	教授			
		下坂正倫	准教授			

論文審査の要旨(2000 字程度)

「Automatic sketch colorization using reference images」という題目の本論文は、線画に対して、カラーイラスト画像を参照して自動彩色することを基本とし、さらにユーザの微修正指示も反映できる機械学習を用いた新たな手法について述べており、全 7 章からなる。各章の概要は以下の通りである。

第 1 章では、本論文の背景および目的について述べている。イラストの制作は手作業が多いが、その中で線画の彩色作業の自動化が望まることから、機械学習の技術を用いた解決が本研究の動機となっていると述べた上で、従来の機械学習による自動彩色法の中でユーザ入力が最も少ない参照カラーイラストを用いた手法に本論文では焦点を当てるとしている。

第 2 章では、機械学習の初期の多層パーセプトロンや畳み込みニューラルネットワークの手法から、現代のトランスフォーマー、変分オートエンコーダ、敵対的生成ネットワーク、拡散モデルに至るまでの進化を辿った後、機械学習を用いたスケッチの自動カラー化における最先端技術を概観している。

第 3 章では、本研究での学習用データには線画像、参照カラーイラスト画像、正解彩色画像の組が必要であり、その組には線画像と参照カラーイラスト像間での意味的同一性と汎化に必要な差異の多様性が求められることを説明し、学習用データセットの構築とその際のデータ拡張法について述べている。また線画では線の情報が重要であるがその情報を効果的に学習器に入力するための前処理として線画の白黒反転についても記載している。加えて、自動彩色の定性評価の指針と定量評価法について述べている。

第 4 章では、本論文での提案手法の 1 つ目である敵対的生成ネットワークアーキテクチャを基本とする線画の参照カラーイラストに基づいた彩色とその彩色結果に対する単語による部分的な色の微修正を行う手法について述べている。本手法は、参照カラーイラストを事前学習させた特徴量検出器からの特徴ベクトルに基づいた自動彩色を行うオートエンコーダ部分の学習を第一段階で行った後、単語による彩色の部分変更を可能とするための第二段階の追加ネットワークの学習を行うことで、線画に対して参照カラーイラストに基づいて比較的安定した彩色を行いつつ単語とスカラー値を与えることで追加の部分変更とその変更の程度を指定する機能をユーザに提供することができるという特徴を持つ。従来法と比較した出力結果のユーザテストによる定性評価では優れた出力であることを示している。また FID 法による定量評価では、従来法と比較して 20 ポイント以上の優位性を示している。しかしながら、さらなる高品質な結果を得るためには、敵対的生成ネットワークのアーキテクチャでは困難であるとしている。

第5章では、第4章の手法を上回るために本論文での提案手法の2つ目である ColorizeDiffusion と呼ぶ潜在空間の拡散モデルによる線画の参照カラーイラストに基づいた彩色手法について述べている。参照カラーイラストの CLIP エンコーダによる出力を潜在拡散モデルの主要部に挿入する際に加えるノイズの程度をノイズ除去ステップ数によって変化させることによって、線画像の特徴と参照カラーイラストの特徴の影響を調整して潜在空間内での彩色画像を表す点の軌跡を期待する線画像の彩色画像へと導く手法である。また、参照カラーイラストの背景にマスクを施して注意を限定するアテンション機構についての導入法についても述べている。従来法や第4章の手法との出力結果の比較において、ユーザテストによる定性評価と FID 法による定量評価共に優れていることを示している。

第6章では、機械学習による自動彩色の手法において、参照画像を特徴量化するエンコーダを全体の学習時に合わせて学習する場合と、事前学習させておき、全体の学習では固定化しておく場合についての議論を行い、本論文での提案した手法が採用している後者の手法の優位性について述べている。

第7章では、本論文で提案した手法についてまとめ、今後の課題と応用の方向性について述べている。

以上、本論文は、対象の線画に対して、カラーイラスト画像を参照して彩色する課題に対して2つの手法を提案している。本課題は対象と参照が共に画像であるというシングルモーダルな枠組みで、参照の情報が対象の情報量より多いなかで、対象の情報が最終結果に安定して表出されることを必須とする課題であると言える。本論文で述べられている研究はそれに対する機械学習を適用する際の有効な手法の提案と見なすことができ、その研究成果は画像、映像表現における応用において非常に有用であり、工学上、工業上、寄与するところが多い。よって、本論文を博士(工学)の学位論文として価値あるものと認める。

注意:「論文審査の要旨及び審査員」は、東京科学大学リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。