

効率的なアニメ線画の着色作業にむけた 逐次的な自動着色精度の改善手法

渡邊 大起^{1,a)} 品川 政太郎² 船富 卓哉² 前島 謙宣³ 向川 康博² 中村 哲² 久保 尋之¹

概要

本研究では、アニメ線画の着色作業を効率化することを目的として、着色過程において段階的に増加する着色済み領域の色情報を用い、逐次的に着色精度を改善する手法を提案する。実験では、着色済みの領域の増加に伴い未着色領域における着色精度が改善することが示された。モデルが予測した色ラベルを着色候補として人に提示し、逐次人が確認しながら着色を行う新たなアプローチにより着色作業の効率化が期待できる。

1. はじめに

現在、アニメ産業界では制作者の作業負担の増加が問題視され、負担軽減のために制作工程の効率化が進められている。なかでも、キャラクタの線画を手動で着色する工程は、キャラクタごとに定義される色指定表をもとに、対象の領域の色をひとつずつ確認しながら着色する。これは単純な作業であるが、繰り返し行うため手間がかかる。これまでこの手間を削減する方法として線画と着色済みの画像が一組となるデータを集めて、線画を全自動で着色するモデルを機械学習で構築することが主に行われてきた [3, 5, 8, 11, 18]。しかし、全自動による着色は、すべての領域が正しい色で塗られるという保証がない。そのため商業作品での利用を考えた場合、全自動による着色結果に誤りがないかどうかを手で確認し、修正する作業が新たに必要となる (図 1)。着色結果によっては、この修正作業のコストが無視できないものとなる。なぜなら、多くの場合で全自動着色結果は一見正しく着色できているように見え、細かい着色の誤りを見分けるのは難しく、修正にかかる作業時間の見込みがつかないためである。この問題を解決する方法として我々は従来の自動着色と誤りの修正という着色プロセスを見直し、人が線画を着色することを主として作業の負担軽減と効率化する方法を考える。人手に



図 1 線画から全自動で着色するアプローチ：線画を着色モデルに入力し、モデルが出力した着色結果の誤りを確認し修正する。

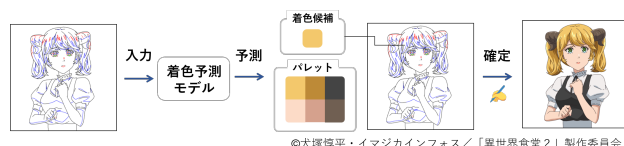


図 2 本研究における着色作業の効率化アプローチ：線画を着色モデルに入力し、モデルが予測した領域における色ラベルを着色候補として人に提示し、人が色の確定と着色を行う。

よる着色作業は、あらかじめ人が目視で色の確認を行った上で着色するため、現状では自動着色よりも正確であるという点で優位である。したがって、各領域で着色される色は、自動着色モデルではなく、人が目視で確認して確定した方が後から修正する作業を必要としないため、最終的に正確かつ効率的であると言える。そこで我々の提案する手法は、人が指定した領域において、自動着色モデルが予測した色と RGB 値を着色候補として人に提示し、選択させる。(図 2)。提示される着色候補が正解である確率が十分高ければ人の確認コストは十分低く済み、提示される色が誤っている場合は、通常の工程通り人がパレットから色を選択すればよい。したがって、全自動で着色された結果から細かい誤りを見つけて修正するよりも、一領域ずつ提示された色を採用するか否かを判断の方が作業者に与える負荷が低く、結果的に効率的であると考えられる。また、最悪作業時間が完全な人手作業をした場合の時間で制約されるため、作業時間の見込みもつきやすく実用上望ましい。

このアプローチの効率性は着色モデルの予測精度に依存する。そこで我々は、着色過程において増加する線画内の着色済み領域に注目する。着色作業を進める上でそれらの領域は既に目視で確認済みであるため正しい色で塗られていることが想定される。したがって、着色済みの領域の色情報をヒントとして、残りの領域の色の推定に有効活用で

¹ 千葉大学

² 奈良先端科学技術大学院大学

³ オー・エル・エム・デジタル・IMAGICA GROUP

a) twatanabe_qlab@chiba-u.jp

できれば、予測精度の改善が期待できる。本論文では、着色過程に応じて増加する着色済みの領域の色情報を用い、逐次的に予測精度を改善する手法を提案する。本手法による学習済みモデルの検証実験で、着色過程に応じて未着色の領域に対する予測精度が改善することを示す。本手法の独創的な点は、機械と人どちらかに依存するアプローチではなく、人と機械が協調的に着色作業を行うアプローチの上で、精度改善と効率性の安定化を図れる点である。システムが着色モデルによって予測された線画の着色候補を人に提示し、人が目視で確認して着色を確定、その着色過程の線画をモデルに再入力して予測された結果を新たな着色候補として確認、着色する。この作業を繰り返す協調的な着色作業を行うことで十分な効率化が期待できる。

2. 関連研究

2.1 全自動による着色手法

機械学習を用いた自動着色に関連する研究として、実写のグレースケール画像への自動着色 [7,14]、漫画・イラスト（線画）の自動着色 [9,15]、アニメ作品におけるキャラクターの線画の自動着色 [8,11,19] など様々なものを着色対象とした研究がある。また、敵対的生成ネットワーク (Generative Adversarial Networks; GAN) [4] やその派生形である条件付き敵対的生成ネットワーク (Conditional GAN) [12] を用いた自動着色手法 [5,6] の研究も取り組まれており、自然な陰影やグラデーションのある着色結果が得られている。アニメ線画における着色は、陰影含めて指定された色で閉領域内を均一に塗る、すなわちべた塗りである点が自然画像や漫画の着色問題と異なる。また画像生成による手法では色にじみなどのアーティファクトが多く生じることから、本研究では Ishii ら [8] と同様にアニメ線画の自動着色の問題をセマンティックセグメンテーションの問題として扱う。

2.2 半自動による着色手法

自動着色の研究の文脈として、完全に自動的な着色を行う手法のほかに、人が着色の補助となるヒントを与え、それをもとに着色を行うような半自動的な着色手法が研究されている。例として、着色する箇所に対してカラーストロークやカラーポイント指定したものを着色のヒントとして使用する手法 [2,17,18] や、人が用意した参照画像をヒントに、その着色スタイルのように着色結果を出力する、所謂スタイル転送のアプローチで着色する手法 [3,16] が挙げられる。完全に自動的な着色による手法とこれらの半自動的な着色手法の違いは、できるだけ希望に近い着色結果を得られるという点である。完全に自動で行う場合、色の塗り間違いや特有の色にじみなどのアーティファクトを修正する時間が多く必要となる可能性がある。しかし、人が事前にヒントを与えることである程度修正時間を軽減する

ことができる [1,5]。本研究は半自動的な着色手法の一つとして位置づけられる。

3. 提案手法

提案手法による協調的な着色作業フローの概要を図 3 に示す。着色作業の大まかな手順は以下の通りである。

- (1) 線画を着色モデルに入力し、閉領域ごとに色を予測
- (2) 人が着色したい線画の閉領域を指定
- (3) 指定された閉領域における着色候補を人に提供
- (4) 着色候補をもとに人が着色（候補に正しい色がなければ、通常通り色指定表をもとに着色）
- (5) 未着色領域があれば、着色された線画をモデルへ再入力し、閉領域ごとに色を予測して手順 (2) に戻る

我々の提案するアプローチには、モデルによって予測された色を着色候補としているため、既存の自動着色モデルを使用できる。しかし、この着色フローの手順に従って部分的に着色済みの線画を入力した場合、既存のモデル学習時に含まれない特徴の線画を入力することから徐々に精度が低下する。本論文では、線画の着色過程において増加する着色済みの領域の色情報をヒントとして、未着色の領域に対する予測精度を更に改善する手法として新たなモデルの学習方法を提案する。既存の自動着色モデルとして Ishii ら [8] のモデルを用い、閉領域内の予測された色ラベルの精度を評価する。Ishii らは、線画への自動着色を画素ごとに色ラベルを予測する、セマンティックセグメンテーションの問題として解いている。本研究も同様に線画における画素ごとに色ラベルを予測する。これまでのモデルの学習には、線画と着色済みの正解画像を組にしたデータセットが用いられていたが、我々は従来のデータセットの線画に対して、データ拡張を施した新たなデータセットを作成して学習に用いる。図 4 にデータセットの作成方法の概要と手順を以下に示す。

- (1) 着色済みの正解画像をもとに閉領域ごとにラベル付け
- (2) ランダムに選択したラベルに対応する線画の閉領域を着色し、線画の全領域に対して着色済み領域の割合が 0%, 20%, 40%, 60%, 80% となるように 5 パターンの着色済み画像を作成

ここで 5 パターンの着色済み画像は、着色過程を模して図 4 の通り着色済み領域を引き継ぐように作成した。上記の手順で作成されたデータセットでモデルの学習を行う。提案手法により学習済みのモデルでは、部分的に着色された線画の入力に対応し、かつ入力される線画における着色済み領域の増加に伴い、段階的に改善された着色予測を出力する。そのため、最初に入力される線画に対する着色精度が低い場合でも、着色作業の補助として効果的に機能し安定した着色作業の効率化を促す。

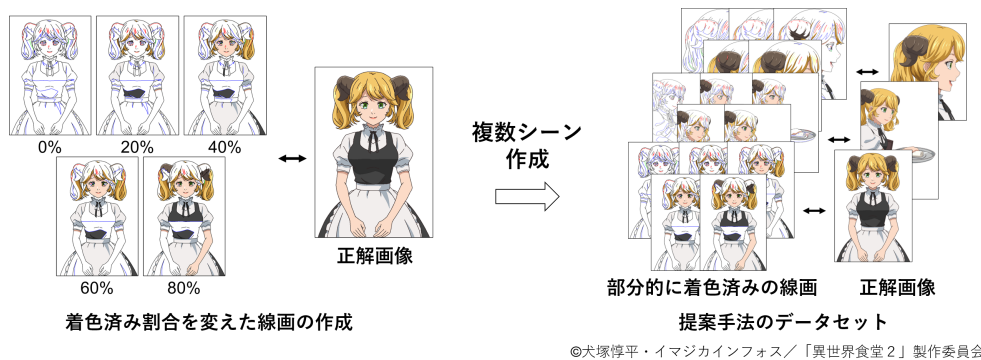
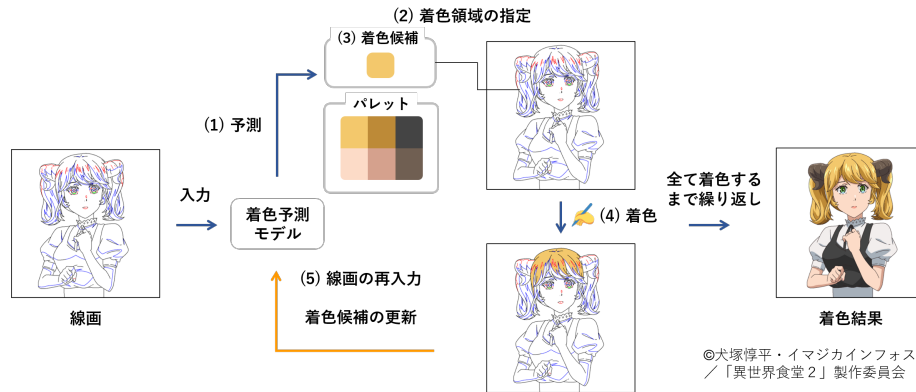


図 4 データセットの作成方法：データ拡張として、正解画像をもとに線画における着色済み領域の割合を 0% から 80% まで段階的に増やした線画を作成し、すべての線画に適用することでデータセットを作成する。

4. 実験

提案手法により、着色済みの領域の色情報をもとに、残りの未着色領域における着色予測の精度改善への有効性を検証する。そのため、着色済みの領域の割合を段階的に増やした線画を入力とし、予測する最上位の色ラベルが正解の色ラベルだった場合および、予測する上位 3 位以内に正解の色ラベルが含まれる場合で、未着色領域の着色予測精度を評価する。提示した着色作業のアプローチ図 3 では、着色候補として予測確率の最上位の色を提示しているが、閉領域内の予測確率上位の何色かを着色候補として提示することも考えられる。選択肢は増加するものの、基本的には最上位の色を確認し、誤っている場合の第 2、第 3 候補として優先的に確認する選択肢として扱うことで効率的な着色作業の補助に用いることができる。検証実験ではこれを踏まえて予測確率最上位の色の精度および、上位 3 位までの色の精度に対して評価を行う。また、同様の設定で学習された従来手法 [8] のモデルとの着色精度を比較する。

4.1 データセットとモデルの学習

提案手法で新たに作成するデータセットと区別するために、従来の線画と着色済みの正解画像の組で構成されるデータセットをベースラインデータセットとする。実験に

におけるベースラインデータセットは、放映済みのアニメシリーズの 1 クールで用いられた線画と着色済み画像の組、計 915 組である。このうち、728 組を Train、94 組を Validation、93 組を Test に分割して使用する。提案手法に基づいてベースラインデータセットを拡張し、0% から 80% の割合で部分的に着色された線画 5 枚と正解画像を 1 組にした部分着色済みデータセットを作成した。部分着色済みデータセットの構成は Train が 3,640 組、Validation が 470 組、Test が 465 組である。これら 2 つのデータセットを用いてモデルを学習し、比較を行う。ネットワークのアーキテクチャは Ishii ら [8] と同様に U-Net [13] を用いる。モデル学習時のパラメータは従来手法、提案手法ともに学習率を $1e-5$ 、エポック数を 1,000、最適化手法には Adam [10] を使用し、モーメント用の係数はそれぞれ標準の設定値 ($\beta_1 = 0.9, \beta_2 = 0.999$) とした。

4.2 評価指標

着色は閉領域単位で行われるため、着色精度の評価指標には閉領域ごとの精度を表す Region-wise accuracy を用い、全閉領域の平均値を比較する。閉領域内で色ラベル i とラベル付けされた画素を、色ラベル j と予測した画素の総数を N_{ij} としたとき、Region-wise accuracy は以下の式で表される。

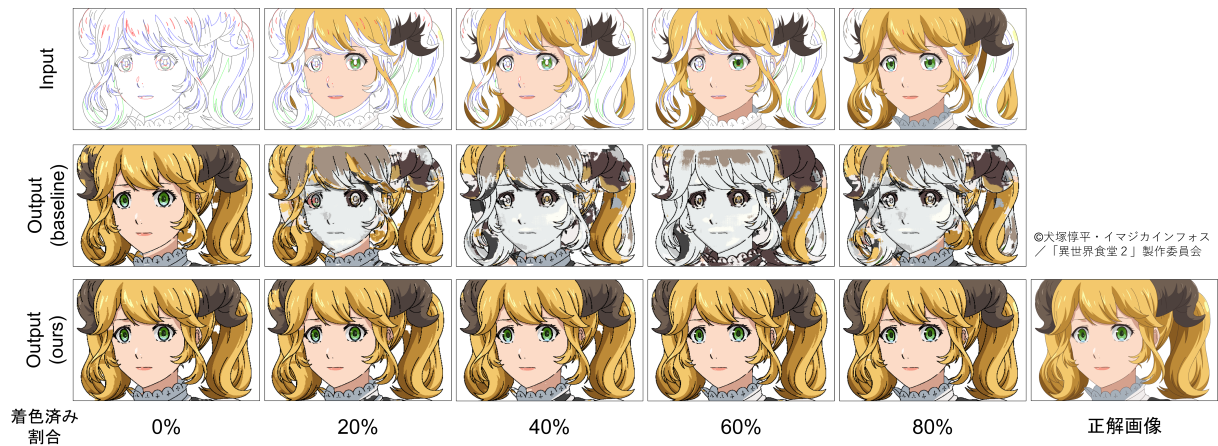


図 5 モデルへの入力画像（部分着色済みの線画）と着色予測結果の例：提案手法の着色結果では、キャラクターのツノにおける着色結果が改善している。従来手法の着色結果では、部分的に着色された線画は学習データに含まれないことから着色ミスが多く、未着色の領域においても改善はしていない。

$$(\text{Region-wise accuracy}) = \frac{\sum_{i=1}^K N_{ii}}{\sum_{i=1}^K \sum_{j=1, j \neq i}^K N_{ij}} \quad (1)$$

なお、色ラベルの予測確率上位 3 位までの色に対する閉領域内の精度は、画素ごとに予測された 3 つの色ラベルの中に正解の色ラベルがあれば、その画素における着色予測は正解と見なして算出する。また、着色精度は着色する割合を 80% まで増加させた線画における着色済み領域を除いた未着色の領域に対して比較評価を行う。

4.3 結果と考察

着色済みの領域の割合を 0% から 80% まで段階的に増やした線画を学習済みのモデルへ入力し、色ラベルの予測確率最上位の色および上位 3 位までの色の着色精度を算出した。提案手法を Ours、従来手法を Baseline として、着色精度の変化を表 1 及び表 2 に示す。提案手法の着色精度に注目すると、着色済みの領域の割合が増加するにつれて、未着色の領域の予測精度が改善していることがわかる。色ラベルの予測確率最上位の色で予測された色ラベルを着色に利用する従来の自動着色の方法に則って作成した着色予測結果の例を、入力した着色画像および正解画像とともに図 5 に示す。提案手法の着色結果では、入力時に未着色であるキャラクターのツノに注目すると、視覚的に正しく着色された画素が増加している様子がわかる。Ishii らの従来手法を図 3 のアプローチに応用し、モデルに対して着色済み画像を入力した場合、その他の領域に対する予測精度が低下していることがわかる。この原因としては従来手法には部分着色済みの線画は学習データに含まれておらず、学習データとは特徴的にかけ離れた線画がモデルに入力されたためと考えられる。また、前述した従来手法の着色精度の低下は主に着色済みの領域の周辺で生じていることから従来の学習データとは特徴的に異なることが確認できる。検

表 1 着色済み割合と各モデルの予測最上位の色の精度

	0%	20%	40%	60%	80%
Baseline	0.6286	0.5418	0.4186	0.2622	0.1308
Ours	0.6934	0.7190	0.7434	0.7682	0.7824

表 2 着色済み割合と各モデルの予測上位 3 位までの色の精度

	0%	20%	40%	60%	80%
Baseline	0.9024	0.8172	0.6668	0.4613	0.2806
Ours	0.9383	0.9387	0.9395	0.9393	0.9397

証実験では学習時のエポック数を 1,000 として設定したが、デフォルトデータセットの Train は 728 組に対して部分着色済みデータセットの Train は 3,640 組であることから、学習時間はそれぞれ約 30 時間と約 150 時間となっており、提案手法では学習時間が増加することが課題である。

5. まとめと今後の展望

着色作業を効率化し、作業者の負担を軽減する新たなアプローチとして、モデルが予測した色ラベルを着色の候補として提供し、人が確認および着色行うアプローチを提案した。本論文では、そのうえで課題となる色ラベルの予測精度を、逐次的に更新可能なモデルの学習方法を提案した。提案手法によって学習されたモデルの検証実験の結果から、着色済み領域の増加に伴い着色予測精度が改善することが分かった。これにより、着色作業が進むにつれて着色予測の精度が上がり、より効率的な着色が可能になるといえる。今後の展望として、より精度の改善が可能なモデル学習のためのデータセット作成方法や、学習時間を抑えつつ同程度の精度を実現可能な学習方法について検討する。また、制作現場で試験的に使用してもらい、効率の評価や、ヒアリングから今後の制作効率化に向けた手法を模索する。

謝辞 本研究は JST 創発的研究支援事業 (JPMJFR206I) 助成を受けたものです。

参考文献

- [1] Cao, R., Mo, H. and Gao, C.: Line Art Colorization Based on Explicit Region Segmentation, *Computer Graphics Forum*, Vol. 40, No. 7, Wiley Online Library, pp. 1–10 (2021).
- [2] Ci, Y., Ma, X., Wang, Z., Li, H. and Luo, Z.: User-guided deep anime line art colorization with conditional adversarial networks, *Proceedings of the 26th ACM international conference on Multimedia*, pp. 1536–1544 (2018).
- [3] Furusawa, C., Hiroshiba, K., Ogaki, K. and Odagiri, Y.: Comicolorization: semi-automatic manga colorization, *SIGGRAPH Asia 2017 Technical Briefs*, pp. 1–4 (2017).
- [4] Goodfellow Ian, J., Jean, P.-A., Mehdi, M., Bing, X., David, W.-F., Sherjil, O. and Courville Aaron, C.: Generative adversarial nets, *Proceedings of the 27th international conference on neural information processing systems*, Vol. 2, pp. 2672–2680 (2014).
- [5] Hati, Y., Jouet, G., Rousseaux, F. and Duhart, C.: Paintstorch: a user-guided anime line art colorization tool with double generator conditional adversarial network, *European conference on visual media production*, pp. 1–10 (2019).
- [6] Hensman, P. and Aizawa, K.: cGAN-based manga colorization using a single training image, *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Vol. 3, IEEE, pp. 72–77 (2017).
- [7] Iizuka, S., Simo-Serra, E. and Ishikawa, H.: Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification, *ACM Transactions on Graphics (TOG)*, Vol. 35, No. 4, pp. 1–11 (2016).
- [8] Ishii et al., D.: Confidence-aware Practical Anime-style Colorization, *Special Interest Group on Computer Graphics and Interactive Techniques Conference Talks*, pp. 1–2 (2020).
- [9] Kim, H., Jhoo, H. Y., Park, E. and Yoo, S.: Tag2pix: Line art colorization using text tag with secant and changing loss, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9056–9065 (2019).
- [10] Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014).
- [11] Maejima, A., Kubo, H., Shinagawa, S., Funatomi, T., Yotsukura, T., Nakamura, S. and Mukaigawa, Y.: Anime character colorization using few-shot learning, *SIGGRAPH Asia 2021 Technical Communications*, pp. 1–4 (2021).
- [12] Mirza, M. and Osindero, S.: Conditional generative adversarial nets, *arXiv preprint arXiv:1411.1784* (2014).
- [13] Ronneberger et al., O.: U-net: Convolutional networks for biomedical image segmentation, *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp. 234–241 (2015).
- [14] Su, J.-W., Chu, H.-K. and Huang, J.-B.: Instance-aware image colorization, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7968–7977 (2020).
- [15] Xie, M., Li, C., Liu, X. and Wong, T.-T.: Manga filling style conversion with screentone variational autoencoder, *ACM Transactions on Graphics (TOG)*, Vol. 39, No. 6, pp. 1–15 (2020).
- [16] Zhang, L., Ji, Y., Lin, X. and Liu, C.: Style transfer for anime sketches with enhanced residual u-net and auxiliary classifier gan, *2017 4th IAPR Asian conference on pattern recognition (ACPR)*, IEEE, pp. 506–511 (2017).
- [17] Zhang, L., Li, C., Wong, T.-T., Ji, Y. and Liu, C.: Two-stage sketch colorization, *ACM Transactions on Graphics (TOG)*, Vol. 37, No. 6, pp. 1–14 (2018).
- [18] Zhang, R., Zhu, J.-Y., Isola, P., Geng, X., Lin, A. S., Yu, T. and Efros, A. A.: Real-time user-guided image colorization with learned deep priors, *arXiv preprint arXiv:1705.02999* (2017).
- [19] 品川政太郎, 久保尋之, 石井大地, 前島謙宣, 船富卓哉, 中村哲, 向川康博: セマンティックセグメンテーションに基づくアニメ作品の自動彩色, *Visual Computing* (2020).