2章 漫画・線画の着色技術

2-1 漫画の半自動着色手法 Comicolorization

古澤 知英†

キーワード：自動彩色，漫画，ディープラーニング

1. まえがき

漫画は世界中で人気を博している。毎年数多くの漫画が生まれるだけでなく、過去の漫画を再出版されることや、デジタル漫画化する作品も多く存在している。漫画の表現も国によりさまざまである。アジアでは、漫画家の少人数体制で白黒漫画を制作することが主流となっている。一方で、欧米ではカラーリストと呼ばれる着色を専門とする人を含めたチームで1つの漫画を制作することが多く、カラー漫画が普及している。このような背景から、アジアで制作された白黒漫画を着色し、再出版することでカラー漫画を主流とする欧米のファンやデジタル漫画のファンの獲得をしている事例も増えている。

しかし、カラー漫画の需要に伴わず日本などでは漫画制作のための人手不足に悩まされている上に、白黒漫画の着色には多大な労力が必要とされる。漫画は、通常1巻100ページにも及ぶが、漫画の着色作業はいまだ手作業で行われていることが多い。そこで、漫画の特徴を考慮し、漫画に特化した着色作業の自動化技術が求められている。

本稿にて紹介する提案手法、Comicolorizationでは、白黒漫画と参照画像を入力として、多層畳み込みニューラルネットワークによりカラー漫画を半自動生成することができる。本稿では、白黒漫画の自動着色に関する関連研究の紹介、およびわれわれの提案手法、Comicolorizationによる白黒漫画の半自動着色手法を紹介する。

2. 関連研究

白黒漫画の着色作業の効率化の究極の目標は、白黒漫画の入力のみで、end-to-endでカラー漫画を生成することである。そのような目標を白黒漫画ではなく、自然画像に対して行うことを可能にした研究が報告された¹⁾。飯塚らの手法では、多層畳み込みニューラルネットワークによりグ

(a) ベースライン1 (b) ベースライン2 (c) 提案手法

図1 従来手法の比較

(a) 自然画像の自動着色手法による結果¹⁾、(b) (a) 手法に adversarial loss²⁾を追加した手法による着色結果、(c) 提案手法による着色結果。(利用したイラストは、亘井によるもの)

レースケール自然画像から全自動で色の復元を行うことができる。しかし、飯塚らの手法を用いて白黒漫画を自動着色すると、鮮やかでない色で着色されることが多い(図1(a))。図1(a)のどの画像も淡い色で髪髪・服が着色されていることがわかる。

Generative adversarial network²⁾を用いることで、全自動のまま、鮮やかな色で着色を行う方法も報告された³⁾。図1(b)は、飯塚らのネットワークモデルに対して、adversarial loss²⁾を追加した着色結果を示している。頭髪の色に注目すると、2列目では黄色で、3列目では緑色であり、いずれも鮮やかな色で着色されている。しかし、この着色結果は、漫画の着色結果としては問題がある。漫画では、複数コマで登場する同一キャラクターは同じ色で塗られているべきである。

図1(a)、図1(b)で漫画の着色において適当でない色で着

†株式会社ドワンゴ

"Colorization Techniques for Manga and Line Drawings: Comicolorization: Semi-Automatic Manga Colorization" by Chie Furusawa (DWANGO Co., Ltd., Tokyo)

色される原因は、いずれも漫画で使用される色の多様性により生じていると考えられる。自然画像の被写体の多くは自然物である一方、漫画の多くは創作物により構成される。自然物は、空や海は青、森は緑、人間の肌はベージュなど色が一意に定まりやすい。しかし、漫画の創作物は、目、髪、背景など、同じオブジェクトでも作品やキャラクターによって使われる色が異なる上に鮮やかな色を使うことが多い。さらに、色彩設定という行程があるように、制作者があらかじめ色使いに関して強い意図を持っている場合が多い。

そこで、全自動ではないが、多層畳み込みニューラルネットワークに追加の色指定に関する情報を付与することで、制作者の意図を補足的に教え、スケッチの半自動着色を行う研究も報告された^{4) 5)}。追加の色指定の情報は、カラーズクリブルやカラードットと呼ばれ、ユーザインタフェースを通して画像上に直接色の線やドットを明示する方法で与えることができる。

本手法では、参照画像から抽出した色彩特徴量を多層畳み込みニューラルネットワークに入力することで、画像入力ごとにカラーストロークを入れることなく、着色の意図に関する追加情報を与える。色彩特徴量により、各コマは鮮やかな色を用い、同一キャラは同一の色使いで着色がされる。

また、本手法では、自動コマ分割やレイアウト復元も導入することで、イラストのような漫画中の1コマだけでなく、漫画1作品(複数コマが含まれる漫画ページ)をend-to-endで着色できる。さらに、よりユーザの意図に近いカラー漫画を最終的に得ることができるようにするため、ユーザインタラクティブな着色結果の修正の機能も備えている。以降、本システムの概要および各工程詳細、結果および応用事例について記す。

3. システム概要

手法概要を図2に示す。本手法では、主に4つの工程に

分かれる。コマ分割・半自動着色・着色結果修正・ページレイアウト復元である。以下の項では各工程詳細を述べる。

3.1 自動コマ分割

多くの先行研究では、漫画でいう1コマ分の画像にあたる単一のイラストを入力画像とし、自動着色することが目標となっていた。しかし、われわれは、漫画1作品を効率的に着色することに焦点を当てるため、入力の漫画ページをコマに分割することから始める。

本システムでは、入力の漫画を二値化した後、コマ分割を自動で行う。トレーニング時の入力画像に合わせるため、二値化を用いている。また、コマ認識および分割のアルゴリズムは、石井らの手法⁶⁾を元に行っている。石井らの手法では、画像中の各行および列の黒ピクセルの勾配に注目し、コマの枠線を認識している。

3.2 参照画像を用いた半自動着色

コマに分割された各画像は、多層畳み込みニューラルネットワークに入力され、入力時に指定された参照画像の情報をを用いて自動着色される。本手法で用いられるネットワークは飯塚らによる手法¹⁾に基づき構築した。

飯塚らの手法は、4つのネットワーク(低レベル特徴ネットワーク、中レベル特徴ネットワーク、大域特徴ネットワーク、着色ネットワーク)から構成されており、画像全体から抽出した大域特徴ベクトルと、より小さな画像領域から抽出した局所特徴ベクトルを結合層で1つのベクトルに統合し、着色ネットワークに入力する。また、大域特徴ベクトルを入力として、画像に対応するラベルを推測するクラス分類タスクを行う。

本手法では、飯塚らの手法を漫画の着色に適応するために、主に3つの改良を行った。1つ目は、参照画像から抽出した色彩特徴量を用いることである。2つ目は、キャラクターおよび作品のラベルを予測するクラス分類タスクを行うことである。3つ目は、adversarial lossを損失関数に加えることである。

図2 手法概要

(利用したイラストは、衿沢すーたろー、巨井、らぐほのえりかによるもの)

- (1) 色彩特徴量：入力した参照画像は、パレットと呼ばれる色彩特徴量を求めるために用いられる。パレットは、参照画像の 256^3 次元の色ヒストグラムから作られる。次元の圧縮のため 256^3 次元から 6^3 次元に量子化する。そのうち、イラストごとの各領域と色の違い（顔がアップの画像と引きの画像では肌の面積が異なるなど）の頑健性を保持するために各色の有無だけを示す様、各色が1か0で表現されたベクトルにする。パレットは、特徴ベクトルと同様に結合層で統合したあと、着色ネットワークに入力する。飯塚らの手法のネットワークモデルの結合層と同様に 28×28 にブロードキャストして局所特徴ベクトルとした。
- (2) クラス分類：飯塚らの手法では、ラベル予測を取り入れることで、各シーンのラベルから着色にふさわしい色をネットワークが選ぶことができた。そこで、われわれは自然画像のシーンではなく、漫画の作品名またはキャラクタ名に対するラベルを用いて、同様の効果を図る。
- (3) Adversarial loss：2節で述べたように、adversarial lossを目的関数に含めることで、自動着色においてより鮮明な色を用いて着色を行うことができる報告がある^{2)~4)7)}。われわれの手法でも同様に、ネットワークにより生成された着色後画像とground-truthの画像の判別機をネットワークに組込む。判別機の損失の算出にはシグモイド・クロスエントロピーを用いる。

3.3 着色結果修正

本手法では、漫画ページおよび参照画像の入力をするこのみで着色結果を得ることができるが、よりユーザ意図通りの着色結果を得るために2つのユーザインタラクティブな修正機能を取り入れている。2つの機能はそれぞれ、大域的な修正と局所的な修正に適しており、ユーザの意図に合わせて1つのみの機能を使うことも、組合せて用いることもできる。大域的な修正にはヒストグラム、局所的な修正にはカラードットと呼ばれるテクニックを用いる。

- (1) ヒストグラム：大域的な修正を行う場合、ネットワークに入力する色彩特徴をパレットの代わりにヒストグラムを用いる。ヒストグラムは、各色の割合を表現しているため、各色の量を調整することができる。パレットの色選びが好ましくなかった場合、ヒストグラムを用いることで、大域的にふさわしい色使いの着色結果を得ることができる。
- (2) カラードット：局所的な修正を行うためには、カラードットを付与する。カラードットでは、パレットやヒストグラムなどの色の組合せの指定ではなく、画像中の特定の箇所の色を指定することができる。得られた着色結果上にユーザインタフェース上で指

定の色のドットを、色を付与したい領域上に直接付与することによって、ドットの周りが自動的にユーザの選んだ色で着色される。トレーニング時に、ネットワークは擬似的なカラードットを与えているため、領域をはみ出しにくく、かつ、色彩特徴量とともに色付けのサポートをすることができる。

3.4 ページレイアウト復元

着色を終えたコマ画像は、入力した元のページのコマの構成に戻される。元のレイアウトに戻すためには、コマのサイズと位置を復元する必要がある。復元前に見た目の質の担保および向上のために2つの事柄を行う。

1つは、着色結果画像中で用いられている白および黒に近い色を各々 $(R, G, B) = (255, 255, 255)$, $(0, 0, 0)$ の白・黒にする。われわれの半自動着色ネットワークでは、量子化された色彩特徴を用いるので、真白・真黒で塗られない場合がある。漫画本来の白い背景と黒い線の見た目に近づけるための改善として、このような色の微調整を自動で行う。

また、ネットワークの出力が、元のコマのサイズより小さい場合、レイアウト復元のために単純な拡大を行ってしまうと、吹き出し中の文字や輪郭線、コマの枠線などのジャギーが目立ち、見た目の品質を損なう。そこで、超解像度合成によるコマのリサイズを行う。これらの処理を通して、入力時のサイズの着色後コマ画像を得る。

自動コマ分割時に保存したレイアウト情報に基づき、配置を復元する。サイズおよび位置の復元が終わったページに対し、入力漫画ページの輪郭線のみを漫画ページを重ね、最終的な着色後ページを得る。漫画は、塗りつぶしがあることが多く、最後の処理で輪郭線を重ねるためには塗りつぶしのない輪郭線のみを得る必要がある。輪郭線は、Erosionをかけた入力漫画画像と二値化をかけた入力漫画画像の差分を求めることで、輪郭線漫画画像を得られる。

4. 着色結果

パレットを用いて半自動着色された漫画ページを図3に示す。左列は、着色に用いた参照画像で、同列側の漫画ページが着色後漫画の一部のページである。図3では、パレットを用いたため、ユーザの手作業による修正は行っていない。各漫画ページが、左列の参照画像と同様の色合いで各コマが着色された。また、同じキャラクタが同じ色合いで着色されていることも確認できた。さらに、着色後、元のレイアウトにふさわしく復元されていることも確認できた。各参照画像には、nico-opendata^{*1}の画像を用い、各漫画は、Manga109⁸⁾の漫画ページを用いた。

さらに、コマの着色結果を図4に示す。図4では、同一のキャラクタの複数コマが参照画像と同様な鮮やかな色合いで着色されたことを示している。図4(c)では、一度、

*1 <https://nico-opendata.jp/ja/index.html>

図3 ページ着色結果

左列が同列右側の漫画ページの着色の際に用いた参照画像(利用したイラストは、亘井、ボコ太、はじめ、トキア、マッハダイ、らくほのえりか、ライによるもの。利用した漫画ページは、Manga109⁸⁾データセットから)。

図4 コマ着色結果

(利用したイラストは、村雨シア、杵沢すーたろーによるもの)

本手法により着色した結果を参照画像として、他のコマを着色することもできたことを示している。したがって、参照画像として利用できるカラー画像がなかったとしても、最初に1枚だけ本手法で着色した後、着色後画像を他のコマに対する参照画像として用いるというシステムの使用の

仕方も考えられる。また、図4(c)の着色結果の上から2列目と4列目では、修正機能を用いた。上から2列目では、ヒストグラムを調整することで、服へしみ出したキャラクタ髪色の黄色が減少していることが確認できた。また、上から4列目では、肌色、青色、黄色のカラードットをそれぞれ額、襟縁、髪の毛に付与することで、特定の領域の色の指定ができた。本手法に関するコードは、GithubのDwango Media Villageオーガニゼーション^{*2}にて公開している。

5. トレーニング

学習済み自動着色モデルおよび超解像度モデルを使用している。本節では、自動着色モデルと超解像度モデルの詳細を記述する。

自動着色のトレーニングデータセットは、ニコニコ静画^{*3}に投稿されたカラーイラストを用いた。ニコニコ静画では、イラスト共にキャラクタ名やタグを付与できるサービスになっている。そこで、クラス分類に用いるラベルは、イラストに付与されたタグ情報428ラベルを用いた。

さらに、トレーニングデータセットには漫画ではなくイラストを用いた。データセットとして使用するのに十分な量の漫画がなく、漫画にはキャラクタが登場するコマが多く含まれているので、イラストを用いて学習した。

イラストは、カラーイラストに対して大津の二値化⁹⁾を

*2 <https://github.com/DwangoMediaVillage/Comicolorization>

*3 <http://seiga.nicovideo.jp/>

図5 自動着色ネットワークモデル
(利用したイラストは、衿沢すーたろー、らぐほのえりかによるもの)

かけることで、元のカラー画像を ground-truth とし、二値化後の画像を入力画像とすることでトレーニングデータペアとした。漫画は、塗りつぶし等の線画イラストとは異なる表現を含んでいるため、本手法では、入力画像の生成に二値化を用いた。

また、カラーイラストの中で白黒漫画の自動着色に適する画像のみを学習に用いるため、2つの観点で画像を選んだ。1つは、漫画には鮮やかな色が用いられることが多かったため、自動着色による結果でも鮮やかな着色を学習できるように、彩度が低いカラーイラストを除外した。彩度の指標にはHSV色空間のSの値を用いた。さらに、漫画にはキャラクターが登場するため、自動着色する際にキャラクターをふさわしく塗れるよう、キャラクター画像を用いた。イラスト中でキャラクター顔認識^{*4}された画像をキャラクター画像とした。これらの条件により、16万枚のイラストをトレーニングデータセットに用いた。

各イラストは、自動着色ネットワークに入力する際に256×256ピクセルにリサイズしたのち、224×224ピクセルをランダムクロップした。また、50%は左右を反転させた画像もデータセットとして使用することで、ロバスト性を強化した。

ネットワークモデルは図5のようにになっている。システムには、バッチサイズ30で55万イテレーションの学習を行ったモデルを使用した。各層の活性化関数には最終層以外ReLU関数を用い、最終層のみシグモイド関数を用いて学習した。また、バッチ正規化層を各畳み込み層の後ろに挿入した。モデルのパラメタは、低レベル特徴ネットワーク、中レベル特徴ネットワーク、大域特徴ネットワークは飯塚らの手法と同様の値を使用し、着色ネットワークおよび判別機は表1のパラメタを用いた。

モデルの最適化の際の損失には、Lab色空間での平均二乗誤差、Adversarial Lossおよびクラス分類誤差を用いた。

表1 着色ネットワークおよび判別機のパラメタ
(a)低レベル特徴ネットワーク、(b)中レベル特徴ネットワーク、
(c)大域特徴ネットワークは飯塚らの手法と同様の値を使用した。

(d) 着色ネットワーク			
タイプ	カーネル	ストライド	出力
Fusion	—	—	256
Conv.	3×3	1×1	128
Upsample	—	—	128
Conv.	3×3	1×1	64
Conv.	3×3	1×1	64
Upsample	—	—	64
Conv.	3×3	1×1	32
Conv.	3×3	1×1	3
(e) 判別機			
タイプ	カーネル	ストライド	出力
Conv.	4×4	2×2	644
Conv.	4×4	2×2	128
Conv.	4×4	2×2	256
Conv.	4×4	2×2	512

各損失を1：1：0.003の重みづけをした損失を用いた。最適化手法はAdam¹⁰⁾を用いた。

修正の際に、付与されたカラードットが着色のために適切に色が広がるようにするため、学習時にも入力白黒画像に擬似的なカラードットを付与した。Ground-truthの画像からランダムに選ばれた0～15点の1ピクセルを擬似的なカラードットとして入力白黒画像の同じ位置に付与した。

超解像度合成モデルは、3.4項にて説明したレイアウト復元の際に用いた。超解像度合成モデルの構成は、Ronnebergerら¹¹⁾に倣った。なお、着色後に超解像度合成により高解像度な着色画像を生成するという発想は、Paints Chainer^{*5}に基づいている。これにより、224×224ピクセルの画像から448×448ピクセルの高解像度の画像合成を得た。

*4 https://github.com/nagadomi/lbpcascade_animeface

*5 <https://github.com/pfnet/PaintsChainer>

図6 ハードウェアへの応用例

Jetsonおよびプロジェクタと組合せることで、実世界のイラストへの着色も可能になった。

図7 コントローラの様子

6. プロジェクションを用いた応用例

応用例として、ComicolorizationとJetsonとプロジェクタを組合せたハードウェアを用いて実世界のイラストにも着色結果を投影できるようにした。使用中の様子および投影結果を図6に示す。Jetsonとプロジェクタを組合せたコントローラを図7に示す。Jetsonは、NVIDIAが開発しているGPU計算が可能な小型のハードデバイスで、今回は、Jetson TX1を用いた。Jetsonに備え付けられたカメラを用いて撮影した実世界のイラストや漫画をComicolorizationに入力し、Jetsonに備わったGPUを用いて着色結果を計算し、プロジェクタで実世界イラスト上に着色結果を投影することで、実世界でも線画や漫画を自動着色する体験を実現した。コントローラは持ち運べるため、実世界のさまざまなものに投影できた*6。

7. むすび

本稿では、白黒漫画の半自動着色手法について紹介した。本手法では、漫画の特徴に焦点を当て、1作品を着色する手法を紹介した。参照画像を用いることで、コマ間で同キャラクターが同じ色合いかつ鮮やかな色で着色できた。また、応用として実世界のイラストに対し、自動着色結果を投影した。

*6 <http://www.nicovideo.jp/watch/1514196011>

今後の課題として、参照画像のキャラクタに対する自動割当てや、複数キャラクタが登場するコマでもふさわしく着色するための改善を行いたい。

Comicolorizationを含むさまざまな技術を通して、漫画やイラストに関する専門家しかできなかったタスクを専門的な知識を持たない人でも行えるようになる技術が発達することは胸の躍ることではないだろうか。本稿から漫画やイラストに関する研究に興味を持ち、さらにnico-opendataや弊社のサービスを活用してくださる方が増えるとさらに嬉しく思う。最後に、本稿の研究や執筆にあたって、お世話になったすべてのイラスト制作者、山田祐雅氏、松井勇佑氏、大元司氏、そして共著者の廣芝和之氏、大垣慶介氏、小田桐優理氏には重ねて感謝を申し上げたい。

(2018年1月30日受付)

〔文 献〕

- 1) S.Iizuka, E. Simo-Serra and H. Ishikawa: "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 35.4 (2016)
- 2) I.J. Goodfellow, J.P.-Abadie, M. Mirza, Courville and Y. Bengio: "Generative adversarial networks", Neural Information Processing Systems (2014)
- 3) P. Isola, J.-Y. Zhu, T. Zhou and A.A. Efros: "Image-to-Image Translation with Conditional Adversarial Networks", IEEE Computer Vision and Pattern Recognition (2017)
- 4) P. Sangkloy, J. Lu, C. Fang, F. Yu and J. Hays: "Scribbler: Controlling Deep Image Synthesis with Sketch and Color", IEEE Computer Vision and Pattern Recognition (2017)
- 5) R. Zhang, J.-Y. Zhu, P. Isola, X. Geng, A.S. Lin, T. Yu and A.A. Efros: "Real-Time User-Guided Image Colorization with Learned Deep Priors", ACM Transactions on Graphics, 9, 119:1-119:11 (2017)
- 6) D. Ishii, K. Kawamura and H. Watanabe: "A Study on Control Parameters of Frame Decomposition Method for Comic Images", Institute of Electronics, Information and Communication Engineers Tech. Rep., 109, 187-192 (2009)
- 7) A. Radford, L. Metz and S. Chintala: "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", International Conference on Learning Representations (2017)
- 8) Y. Matsui, K. Ito, Y. Aramaki, T. Yamasaki and K. Aizawa: "Sketch-based Manga Retrieval using Manga109 Dataset", Multimedia Tools and Applications, Springer (2016)
- 9) N. Otsu: "A Threshold Selection Method from Gray-Level Histograms", IEEE Transactions on Systems, Man and Cybernetics, 9, 62-66 (1979)
- 10) D.P. Kingma and J. Ba: "Adam: A Method for Stochastic Optimization", International Conference on Learning Representations (2015)
- 11) O. Ronneberger, P. Fischer and T. Brox: "U-Net: Convolutional Networks for Biomedical Image Segmentation", Medical Image Computing and Computer-Assisted Intervention, Springer (2015)

古澤 知英 2016年、早稲田大学大学院先進理工学研究科修了。同年、(株)ドワンゴ入社。現在、Dwango Media Villageという機械学習・画像処理・Computer Graphicsの研究とその応用を行うグループに所属。