

H-017

畳み込みニューラルネットワークを用いたモノクロ画像のカラリゼーション

Colorization of Grayscale Image using Convolution Neural Network

嵐 悠登† 黒木 修隆† 廣瀬 哲也† 沼 昌宏†

Yuto Arashi Nobutaka Kuroki Tetsuya Hirose Masahiro Numa

1. はじめに

モノクロの画像や映像は、過去の資料・作品の中に多く残されている。また、夜間の撮影など、特殊な光源の下ではカラー画像が得られない場合がある。これらの画像をカラーに復元することは、識別能力の向上や過去からの情報創出のためにも需要が見込まれる [1]。現在、モノクロ画像の着色は主に手作業で行われており、絵柄にふさわしい色を作業者の経験や勘によって着色している。これには非常に多くの時間と労力が必要となる。そのため、近年、カラリゼーションと呼ばれるコンピュータ技術を用いてモノクロ画像を自動で着色する研究が盛んに行われている [2],[3]。これらの技術では、復元したい画像に合った色情報を学習させておけば自動で着色されるため、作業の効率化が期待できる。

1.1 従来手法の概要

従来手法[3]では図 1 に示すような 3 層の多層パーセプトロン (MLP) を用いる。入力層と、中間層は 3 ユニット、出力層は 2 ユニットである。モノクロ画像から得た 1 画素の輝度値、およびその周辺 5×5 画素の標準偏差と平均値を入力し、出力として色差信号の Cb, Cr を得る。それらの色差信号を入力輝度値に合成することにより画像の着色を行う。

1.2 問題点

従来手法では入力層と中間層のユニット数が少ないため表現能力が低くなる。しかし、単純に層数やユニット数を増やすと、MLP はユニット間の結合数が膨れ上がり、加重をうまく更新できない問題 (Vanishing gradient problem) が生じる。また、周辺 5×5 画素の情報だけでは、局所領域の形状を考慮できない。その結果、自然な着色が困難になる。

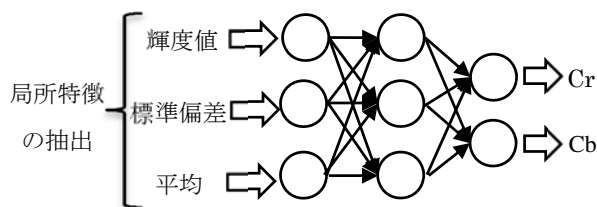


図 1 従来手法のネットワーク構造

2. 提案手法

カラリゼーションにおいて畳み込みニューラルネットワーク (CNN) を用いた手法について提案する。

2.1 提案手法の概要

提案手法では図 2 に示すように 3 層構造の CNN を用いる。CNN を用いることで Vanishing gradient problem を回避しながら特徴マップ数や層の数を増やすことが可能になるので、表現能力の高い着色が期待できる。提案手法の CNN はモノクロ画像を入力とし、Cr, Cb 成分画像を出力する。カラリゼーションの場合は画像認識と異なり、出力の次元数を下げる必要がないため、CNN にプーリング層は設けない。提案手法は学習段階と復元段階の二つの処理に分けられる。

2.2 復元段階

復元段階では、着色したいモノクロ画像を予め学習で構築した CNN に入力することで、色差信号 Cr, Cb の画像を得る。それらを入力画像に合成することでカラー画像を復元する。

入力するモノクロ画像を Y とすると、第 1 層の特徴マップ $F_1(Y)$ は、

$$F_1(Y) = \max(0, Y * W_1 + B_1) \quad (1)$$

となる。ここで、 W_1 はフィルタ、 B_1 はバイアスである。

次に第 2 層の特徴マップ $F_2(Y)$ は、

$$F_2(Y) = \max(0, F_1(Y) * W_2 + B_2) \quad (2)$$

と表される。これらのパラメータの意味は式 (1) と同様である。

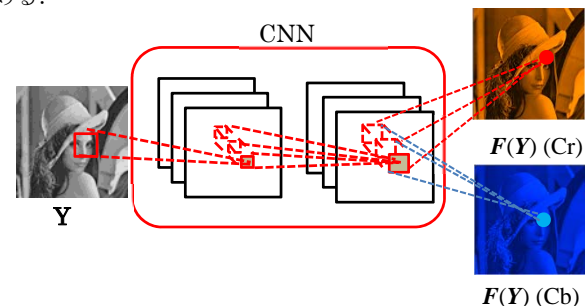


図 2 提案手法の構造

† 神戸大学大学院工学研究科,
Graduate School of Engineering, Kobe University

第 3 層は, $F_2(\mathbf{Y})$ をもとに Cr, Cb 成分の画像を生成する. Cr 成分画像を y_{Cr} , Cb 成分画像を y_{Cb} とすると, 2ch の出力画像 $\mathbf{F}(\mathbf{Y}) = (y_{Cr}, y_{Cb})$ は,

$$\mathbf{F}(\mathbf{Y}) = \mathbf{F}_2(\mathbf{Y}) * \mathbf{W}_3 + \mathbf{B}_3 \quad (3)$$

のように算出される.

2.3 学習段階

学習段階ではカラー画像をモノクロ画像に変換し, 前者を教師信号, 後者を入力 \mathbf{Y} とする. そして, 出力 $\mathbf{F}(\mathbf{Y})$ と教師信号の誤差が小さくなるように誤差逆伝播法を用いて $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$, および $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3$, を決定する.

3. 実験

3.1 実験内容

従来手法と提案手法を用いてモノクロ画像のカラリゼーションを行う. 学習用画像および評価用画像は画像データセット Caltech256 から zebra, fern, face, fireworks の 4 つのカテゴリを用いた. 各カテゴリから 10 枚を評価用画像とし, 学習用画像をそれぞれ, zebra : 76 枚 fern : 92 枚 face : 100 枚 fireworks : 90 枚とする. 評価項目は, 評価用画像 10 枚の PSNR とする. CNN のフィルタサイズは 1 層目から順に 5-5-3, 特徴マップ数は 256-128-2 とした. また, 学習時の学習係数は 0.0001 とし, 200 万回のバックプロパゲーションを行った.

3.2 結果と考察

表 1 に従来手法と提案手法における各カテゴリの PSNR とそれらの平均値を示す. 実験を行った全てのカテゴリで提案手法が従来手法より高い PSNR に達し, 平均で 1.10 [dB] の向上が見られた. これは提案手法の層が深くなったことにより, 表現能力が上がったためだと思われる. 図 3 ~ 6 に出力結果の一部を示す. 従来手法は自然な着色ができず, 全体的に色相のバリエーションが少ない画像となった. それに対し, 提案手法では, 原画像と比べると違いはあるものの, 複数の色で自然な着色が行えた. これは, 特徴マップが増えることで, 多様な色のパターンを学習できたためだと考えられる. 図 3 において従来手法は背景と対象物の色合いが全体的に類似した色で着色されているが, 提案手法では背景と対象物を分離した着色が可能となった. 提案手法では従来手法より入力の参照領域が広く, その形状まで考慮できるためと考えられる. 一方, 図 4 の face は影の部分やエッジ付近に顔と同じ肌色で着色された箇所が見られる. これは, 輝度値成分やテクスチャが顔とそれらの箇所類似していたためだと考えられる.

4. まとめ

本稿では, モノクロ画像に自然な着色を行うことを目的に, 畳み込みニューラルネットワークを用いたカラリゼーション手法を提案した.

実験の結果, PSNR において提案手法は従来手法より 1.10dB 向上した. また, 主観的にも提案手法はモノクロ画像を多様な色合いで着色することを確認した. この結果から, CNN を用いることの有効性が確認できた.

今後の課題として, 物体認識を取り入れたカテゴリの自動分類が挙げられる.

表 1 着色後の PSNR

カテゴリ	PSNR [dB]	
	従来手法	提案手法
zebra	24.36	24.75
fern	19.48	22.46
face	24.89	25.39
fireworks	24.15	24.66
平均	23.22	24.32

参考文献

- [1] 堀内 隆彦, "モノクロ映像のカラリゼーション", 日本色彩学会誌, vol.31, No.4, pp.298-303, 2007.
- [2] T.Welsh et al, "Transferring color to greyscale images", Proc.ACM SIGGRAPH, vol.20, no.30, pp.277-280, 2002.
- [3] 小林 洋平, 白山 晋, "モノクロ画像のカラリゼーションに関する基礎的研究", 映像情報メディア学会誌 Vol.59, No.5, pp.769-775, 2005.



(a) 提案手法

(b) 従来手法

図 3 カラリゼーション結果 (zebra)



(a) 提案手法

(b) 従来手法

図 4 カラリゼーション結果 (face)



(a) 提案手法

(b) 従来手法

図 5 カラリゼーション結果 (fern)



(a) 提案手法

(b) 従来手法

図 6 カラリゼーション結果 (fireworks)