

# モノクロ画像を入力としたカラー化 NeRF の検討

水上 皓太<sup>†</sup> 出口 大輔<sup>†</sup> 村瀬 洋<sup>†</sup>

<sup>†</sup>名古屋大学

## 1 はじめに

NeRF [1] とは、複数視点で撮影された画像からニューラルネットワークを用いて 3D シーン表現を学習し、学習した 3D シーン表現から 3 次元再構成と新規視点画像生成を行う手法である。NeRF の入力として、例えば特殊なモノクロセンサーや古いモノクロカメラ等で撮影された画像しかない場合などでは、モノクロ画像を入力とするために、得られる 3D シーン表現はモノクロ情報しか持たない。そのため、そのままでは新規視点のカラー画像を生成することはできない。その解決方法の一つとして、NeRF に入力する前のモノクロ画像を従来の画像カラー化手法によりカラー化する方法が考えられる。しかし、従来の画像カラー化手法を視点ごとに独立に適用した場合には、同じ物体であったとしても異なる色が推測される可能性がある。そのため、このような視点による色の一貫性がない画像を用いて NeRF により 3 次元再構成を行うと、視点の移動により物体の色が不自然に変化するシーンが生成される。

本発表では、3 次元的な色の一貫性を保つカラー化 NeRF モデルの構築について検討した内容について報告する。

## 2 提案手法

NeRF は、ニューラルネットワークを用いて 3D シーン表現を学習し、それを用いてボリュームレンダリングすることによって任意

視点の画像を生成する技術である。NeRF は、複数視点画像を入力として 3D 座標と視線方向から色と密度を計算する MLP を学習し、これをボリュームレンダリングに必要な各座標の色と密度の計算に用いている。提案手法では、この MLP を学習する際に、視線方向の変化による色の変化を抑制する損失  $L_{con}$  を加える。以下では、 $L_{con}$  の具体的な計算手順についてのみ示し、MLP の学習の際は従来の損失と  $L_{con}$  の両方を用いる。

まず、ボリュームレンダリングによって各座標の色を計算に用いる光線を  $r_n = (t, d_n)$  とする。ここで、 $t$  はカメラ中心を表す座標であり、 $d_n$  は視線方向を表す。この  $d_n$  に対して、正規乱数を用いてランダムな回転行列  $R$  を生成し、 $d_n$  を回転させた  $d'_n$  を次式により求める。

$$d'_n = R d_n$$

次に、光線  $r_n$  に沿ってサンプリングする座標を  $x_{n1}, \dots, x_{nm}$  とすると、通常の NeRF と

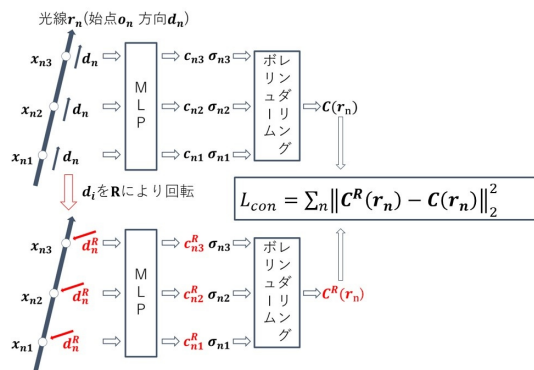


図 1:  $L_{con}$  の計算

A preliminary study on colorized NeRF with grayscale images.

<sup>†</sup> MIZUKAMI Kota, DEGUCHI Daisuke, MURASE Hiroshi, Nagoya University

同じく視線方向  $d_n$  を用いて得られる  $\{c_{nm}, \sigma_{nm}\}$  とランダムに回転させた視線方向  $d'_n$  を用いて得られる  $\{c'_{nm}, \sigma'_{nm}\}$  を次式により得る.

$$\{c_{nm}, \sigma_{nm}\} = MLP(x_{nm}, d_{nm})$$

$$\{c'_{nm}, \sigma'_{nm}\} = MLP(x_{nm}, d'_{nm})$$

ここで, NeRF で用いる MLP では視線方向を変えても得られる密度は変わらないことから  $\sigma_{nm} = \sigma'_{nm}$  とすると, 最終的にボリュームレンダリング (VolR) によって得られる色情報は以下ようになる.

$$C(r_n) = \text{VolR}(c_{n1}, \sigma_{n1}, \dots, c_{nm}, \sigma_{nm})$$

$$C'(r_n) = \text{VolR}(c'_{n1}, \sigma_{n1}, \dots, c'_{nm}, \sigma_{nm})$$

これらを用い, 提案する損失  $L_{con}$  は次式により求められる.

$$L_{con} = \sum_n \|C'(r_n) - C(r_n)\|_2^2$$

### 3 実験

iPhone 13 を用いて物体を撮影した動画から 300 枚程度の異なる視点の画像を切り出し, 実験用データセットに用いた. また, 想定される入力用のモノクロ画像は, 次式により便宜的に生成した.

$$Y = 0.299R + 0.587G + 0.114B$$

モノクロ画像をカラー化する手法には BigColor [2] を用いた. NeRF の実装には nerfacto [3] を用い, 以下の 3 つの手法により得られる画像を比較した.

**ベースライン 1** モノクロ画像をカラー化してから NeRF に入力

**ベースライン 2** NeRF が出力したモノクロ画像をカラー化

**提案手法** モノクロ画像をカラー化してから  $L_{con}$  を追加した NeRF に入力

図 2 に提案手法と比較手法の実験結果を示す. 図 2 からわかるように, 提案手法では視点

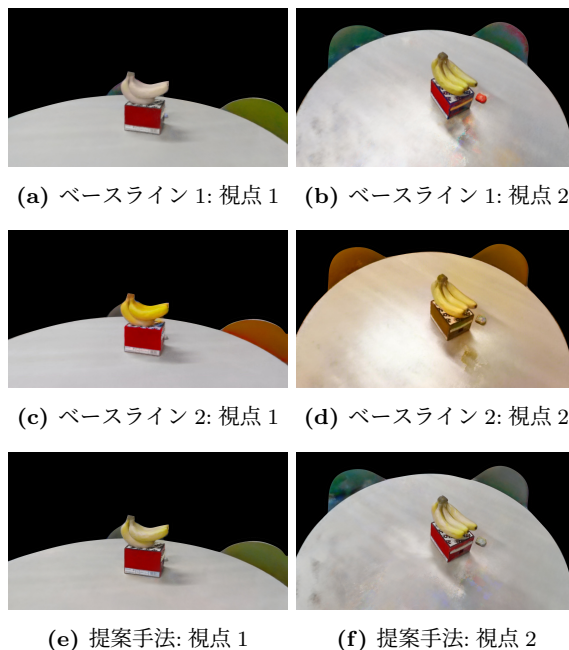


図 2: 実験結果

が移動してもバナナや椅子, 机の色の変化が抑えられていることがわかる.

### 4 むすび

本発表では, 入力として, モノクロ画像しか得られない状況で構築したモノクロ NeRF モデルに対して, 3 次元的な色の一貫性を考慮した損失を取り入れることでカラー化 NeRF モデルを構築する手法を提案した. 今後の課題として, 被験者実験による主観評価, 近赤外線カメラなどのモノクロセンサーへの応用などが挙げられる.

**謝辞** 本研究の一部は JSPS 科研費 23H03474 による.

### 参考文献

- [1] Mildenhall, B., Srinivasan, P.P., Tancik, M., et al. : NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, *Computer Vision - ECCV 2020*, pp.405-421, Nov. 2020.
- [2] Kim, G., Kang, K., Kim, S., et al. : BigColor: Colorization Using a Generative Color Prior for Natural Images, *Computer Vision - ECCV 2022*, pp.350 - 366, Nov. 2022.
- [3] Tancik, M., Weber, E., Ng, E., et al. : Nerfstudio: A modular framework for neural radiance field development, *SIGGRAPH 2023 Conference Proceedings*, pp.1-12, July. 2023.