

# 知っておきたいキーワード

## NeRF: Neural Radiance Fields (ニューラルラディアンسフィールド)

長沼宏昌†

†株式会社ソニー・インタラクティブエンタテインメント

"NeRF: Neural Radiance Fields" by Hiromasa Naganuma (Sony Interactive Entertainment Japan, Tokyo)

キーワード: ニューラルラディアンスフィールド, 3D再構築, ディープラーニング, 写真的レンダリング

### まえがき

近年、コンピュータビジョンとAIの技術は急速に進化しています。特に、3次元空間の再構成技術の発展は目覚ましいものがあります。その基幹技術として、NeRF (Neural Radiance Fields)<sup>1)</sup>は特に注目された技術です。

発音はナーフと読むようです。NeRFの革新的なアプローチは、3次元空間の視覚情報を従来手法に比べて高い精度と再現性で実現しました。

NeRFが登場して以来、さまざまな派生技術が次々と生み出されています。それらの技術は3Dモデル生成、仮想現実、映画やゲーム制作など、幅広い応

用分野において活用されています。

本記事では、NeRFの基本的な概念とそれを構成する技術について解説を行います。私たちの周りには多くのコンテンツがどのように作られているのか、その裏側にある技術と工夫に触れていただければ幸いです。

### NeRFの基本概念

NeRFは2020年に開催されたECCV (European Conference on Computer Vision)において、Best Paper Honorable Mentionを受賞した革新的な技術です。この技術は、複数の視点からの画像を基に、新しい視点の画像を生成することを可能にします。任意の視点の画像を生成する技術としては、既存手法として360°撮影技術やフォトグラメトリといった要素技術がすでにありました。

360°撮影技術は全方位を撮影し、合成することで360°視点の映像を実現します。しかし、この方法では撮影位置からの視点の移動は不可能であり、視点の自由度に限界がありました。

フォトグラメトリは、複数の写真を用いて物体や環境の3Dモデルを生成する技術です。対象の立体的な再現が可能になりますが、多くの画像データと高度な計算が必要であり、再現精度に限界がありました。

NeRFはフォトグラメトリと同様に、さまざまな角度から撮影した写真

を用いて自由視点映像を生成しますが、その過程でニューラルネットワークを利用します。これにより、従来の技術では再現が難しかった光の反射や透けて見える景色のような複雑な情報を含む環境下でも、高精細な画像生成が可能になります。

NeRFは、視点の自由度や再現精度において従来技術を大きく上回り、3次元空間内の視覚情報を高精度に再生するための新たなアプローチを提供しました。

### NeRFの仕組み

NeRFを理解するために、まず本論文を紹介しているWebサイト<sup>2)</sup>をご覧くださいと思います。ここではNeRFを使用した多数の事例が動画として紹介されています。

NeRFは、撮影したシーンの新しい視点の画像を合成する手法であり、複数の視点から得られた画像を基に、連続的なボリュメトリックシーン関数を最適化します。この技術により、視点の変更に伴ってリアルタイムに新しい視点の画像を生成することができます。

図1に論文から引用した図を掲載しました。NeRFは以下のステップで構成されています。

- (a) 5D入力 (位置+方向)  
NeRFの入力は5次元ベクトルです。3次元空間の位置  $(x, y, z)$  とその点からの視線方向  $(\theta, \phi)$  で構成されます。
- (b) 出力 (色+密度)  
5次元ベクトルは、ニューラルネットワーク  $F_{\theta}$  に入力されます。このネットワークは、与えられた位置と方向に基づいて、その点の放射輝度 (RGB) と体積密度 ( $\sigma$ ) を出力します。放射輝度とはその点からどれだけの光が放射されるかを示し、体積密度はその点の物質の密度を示します。
- (c) ボリュームレンダリング  
シーン内の各点から視点までの光の経路に沿って、色と密度の値を集積します。これは、光線がシーンを通過する際にどのように色が変化し、どのよ

うに吸収されるかをシミュレートしています。このプロセスにより、最終的な画像が生成されます。

- (d) レンダリングロス  
生成された画像と実際の撮影画像との誤差を計算します。この誤差を最小化するためにニューラルネットワークのパラメータを最適化します。これにより、ネットワークがより正確にシーンを再現するように学習を進めます。
- 以上がNeRFの基本的な処理になります。3次元位置と視線方向を入力として、ニューラルネットワークが色と密度を出力し、それを基にレンダリングを行い、最終的に撮影画像との誤差を最小化することで、高精細でリアルな新しい視点の画像を生成することが可能となります。

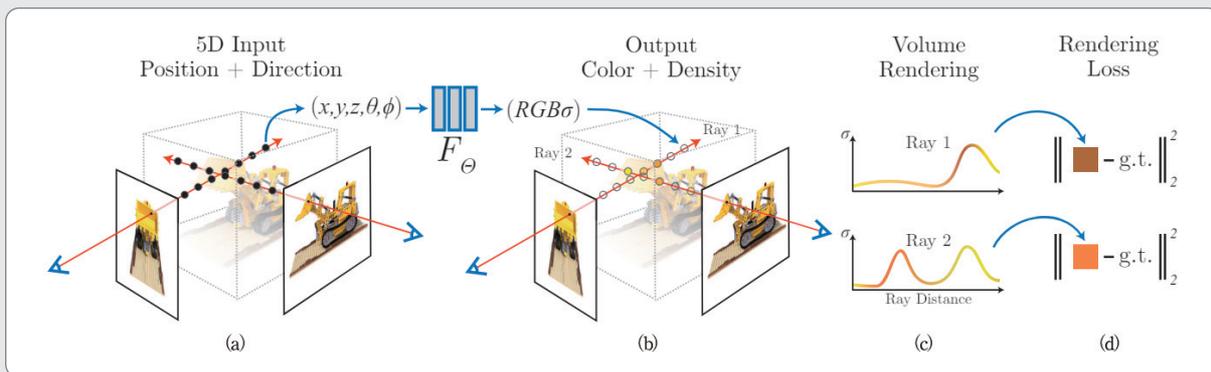


図1 NeRFの仕組み (文献<sup>1)</sup>より引用)

### NeRFの実例

NeRFの使用例として、nerfstudio<sup>3)</sup>を紹介します。nerfstudioは、NeRFを簡単に実装し、さまざまなユースケースで利用できるツールセットです。

#### (1) 特徴

NeRFの入力は5次元ベクトルのため、撮影した画像だけでは情報が足りません。そのために前処理が必要となります。nerfstudioは、このデータの前処理や訓練、レンダリング処理などを別々に行うことができ、それぞれの結果の確認も容易です。

#### (2) 環境構築

nerfstudioではさまざまな環境構築方法が提示されています。この中でも

一番容易だと思われるのが、Dockerイメージを使用する方法です。注意点として、GPUとCUDAのバージョンについて推奨されている記載があることです。

#### (3) 撮影したデータを使用する

nerfstudioでは撮影したデータから5次元ベクトルを生成することもできます。文献<sup>4)</sup>にさまざまな入力データに対応した変換手法について記載があります。

#### (4) COLMAPによるカメラポーズ推定

5次元ベクトルを生成するために、nerfstudioではCOLMAPというSfM (Structure from Motion) による解析ツールを用いてカメラポーズを推定し

ています。SfMはコンピュータビジョンとフォトグラメトリの分野で長い歴史があり現在でも重要な基幹技術です。

#### (5) NeRFのトレーニングとレンダリング

COLMAPの出力を基にNeRFモデルをトレーニングします。viewerで学習結果を確認し、レンダリングしたいカメラ情報を決めます。最後にレンダリングを行って完成です。

nerfstudioは論文<sup>5)</sup>もあります。nerfstudioはNeRF系論文のためのフレームワークとして利用することが容易なツールセットです。個人のローカル環境でこのような素晴らしい画像を生成できることに、時代の進歩を感じます。

## むすび

今回は実例としてnerfstudioを紹介しましたが、自由視点映像生成をもう少し手軽に試したい方にはLuma AI<sup>6)</sup>をお勧めします。Luma AIは、スマートフォンで動画を撮影するだけで新しい視点の画像を生成することが可能なサービスです。出力結果は図2に掲載しました。

2020年にNeRFが登場して以来、さまざまな派生技術が開発されていま

す。Luma AIはこれらの派生技術も取り入れています。以下に、NeRF以降の代表的な派生技術をいくつか紹介します。

(1) 3D Gaussian Splatting<sup>7)</sup>

3次元空間の表現方法がNeRFと異なり、レンダリングが高速です。非常に注目されている技術なので、詳細はWebサイト<sup>8)</sup>を参照すると良いでしょう。

(2) NeRF-W<sup>9)</sup>

NeRF in the Wild<sup>10)</sup>は天候や多様な

照明条件の変化に対応した技術です。自然環境での精度が向上しています。

(3) Instant-ngp<sup>11)</sup>

GPUベンダーの研究者によって発表されました。Multiresolution Hash Encoding<sup>12)</sup>という技術を使用して学習時間を短縮しています。

この他にも多数の技術が日々開発され、さまざまな分野での応用が期待されています。今後もNeRFはますます多くの分野で利用されるでしょう。

(2024年6月4日受付)



図2 Luma AIの出力例(左撮影画像, 右レンダリング画像)

## 参考文献

- 1) B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi and R. Ng: "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis", in European Conference on Computer Vision (ECCV) (2020)
- 2) B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi and R. Ng: "NeRF" (2020), <https://www.matthewtancik.com/nerf/>
- 3) NeRF Studio Documentation, <https://docs.nerf.studio/>
- 4) NeRF Studio Custom Dataset, [https://docs.nerf.studio/quickstart/custom\\_dataset.html](https://docs.nerf.studio/quickstart/custom_dataset.html)
- 5) M. Tancik, P.P. Srinivasan, B. Mildenhall, J.T. Barron, R. Ramamoorthi and R. Ng: "Nerfstudio: a Modular Framework for Neural Radiance Field Development", in Proceedings of the SIGGRAPH Conference (2023)
- 6) Luma Labs, <https://lumalabs.ai/>
- 7) B. Kerbl, J. Thies, T. Müller, A. Dai, M. Hälli and A. Kaplanyan: "3D Gaussian Splatting for Real-Time Radiance Field Rendering", in Proceedings of the SIGGRAPH Conference (2023)
- 8) 3D Gaussian Splatting, <https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/>
- 9) R.M.-Brualla, N. Radwan, M.S.M. Sajjadi, J.T. Barron, A. Dosovitskiy and D. Duckworth: "NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections", in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021)
- 10) NeRF in the Wild, <https://nerf-w.github.io/>
- 11) T. Müller, A. Evans, C. Schied and A. Keller: "Instant Neural Graphics Primitives with a Multiresolution Hash Encoding", in Proceedings of the ACM SIGGRAPH Conference (2022)
- 12) Instant Neural Graphics Primitives, <https://nvlabs.github.io/instant-ngp/>



ながさわ ひろまさ  
長沼 宏昌

2000年、ソニー(株)入社。PC事業において録画基板の設計開発に従事。その後、Blu-ray事業にて高画質化エンジンの開発を経て、研究開発部門にて映像技術の開発に携わる。現在、(株)ソニー・インタラクティブエンタテインメントにて、センシング技術の開発に従事。