

ニューラル場による共焦点画像の奥行補間

大橋 智哉^{1,a)} 藤村 友貴¹ 船富 卓哉¹ 向川 康博¹ Nath Sristilekha² 奥田 覚²

概要

本研究ではニューラル場を用いた共焦点画像の奥行補間を提案する。細胞サンプルを断層的に撮影した共焦点画像は、奥行方向の撮影枚数が少ない場合は十分な解像度が得られず、細胞サンプルの観察は困難となる。本研究では、元の共焦点画像を再構成するようにニューラル場の学習を行う。学習後は任意の奥行で画像が生成可能であり奥行の補間が可能である。画像空間での線形補間との比較を行い、ニューラル場を用いた手法の有効性を検証した。

1. はじめに

近年、幹細胞の培養技術が発達し、再生医療などへの応用が期待されている。しかしながら、培養時における細胞の形成過程のメカニズムは依然不明であり、目的とする細胞の作成は困難である。このメカニズムの解明には、対象とする細胞を3次元的に観測し、形成過程における立体構造を理解する必要がある。共焦点顕微鏡は対象とする細胞を断層的に撮影することが可能であり、このようなニーズに利用されている。

共焦点顕微鏡を用いて立体構造を詳細に分析するためには、奥行方向に密な撮影をする必要がある。しかしながら、そのような密な撮影には計測時間や細胞へのダメージを増大させてしまうという問題がある。

図1に示すように、本研究は共焦点画像の奥行補間を目的とする。共焦点画像の奥行補間を行うことで、疎な撮影からでも立体構造の詳細な分析が可能となる。本研究では、近年 Neural radiance field (NeRF) [5] などを用いられているニューラル場を奥行補間に利用する。共焦点画像は対象とする3次元空間を断層的に撮影したものである。そこで、3次元空間中の点の座標を入力とし、撮影した共焦点画像を再構成するようにニューラル場を学習する。学習後は3次元空間中の任意の値を入力可能であり、奥行方向に密な画像を生成することができる。

本研究ではグリッド表現を持つニューラル場の一つである K-planes [3] を用いる。グリッド表現を持つニューラル

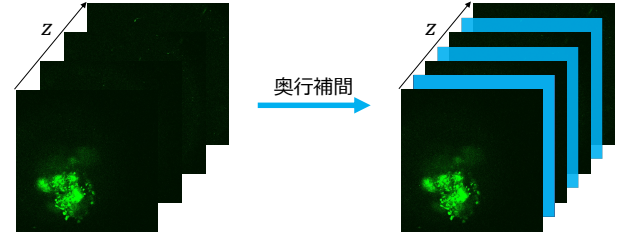


図1 本研究では奥行方向に断層的に撮影された共焦点画像を入力として、各画像間の奥行補間を行う。

場の補間は、特徴空間での線形補間と多層パーセプトロンによる非線形変換に相当する。実験では画像空間における線形補間との比較を行い、ニューラル場を用いた補間の有効性を検証する。

2. 提案手法

2.1 共焦点画像

まず最初に本研究で扱う共焦点画像について説明する。共焦点画像は対象とする空間を断層的に撮影したものであり、 $I(x, y, z) \in \mathbb{R}$ で与えられる。ここで、 (x, y) は各画像のピクセル位置であり、 z は撮影した奥行の位置である。例えば、図1の例は立体組織であるオルガノイド [2] を奥行 z 方向に撮影した共焦点画像である。

2.2 K-planes [3]

本研究ではニューラル場として K-planes [3] を用いる。K-planes はグリッド表現を持つニューラル場 [1], [4], [6], [7] の一つである。NeRF [5] のような多層パーセプトロンのみを用いるものに対して、グリッド上に直接特徴ベクトルを保持することで、学習が高速であるという利点がある。さらに、K-planes では任意の N 次元グリッドを2次元平面の組み合わせで表現することで、メモリコストの削減が可能としている。

K-planes で3次元空間を表現する場合は、互いに直交する3枚の2次元平面 \mathbf{P}_{XY} , \mathbf{P}_{YZ} , \mathbf{P}_{ZX} を用いる。各平面上には M 次元の特徴ベクトルを保持する。例えば、 X, Y を各次元の解像度とすると、 $\mathbf{P}_{XY} \in \mathbb{R}^{X \times Y \times M}$ である。いま、ある3次元点 $\mathbf{q} \in \mathbb{R}^3$ を入力した場合、各平面 $c \in C$, $C = \{XY, YZ, ZX\}$ 上の特徴ベクトル $f(\mathbf{q})_c \in \mathbb{R}^M$ を以

¹ 奈良先端科学技術大学院大学

² 金沢大学

^{a)} ohashi.tomoya.ou3@is.naist.jp

下の式に従って取得する.

$$f(\mathbf{q})_c = \psi(\mathbf{P}_c, \pi_c(\mathbf{q})) \quad (1)$$

ここで, π_c は 3 次元点の 2 次元平面上への射影であり, ψ はバイリニア補間である. その後, アダマール積を用いて特徴ベクトル $f(\mathbf{q}) \in \mathbb{R}^M$ を計算する.

$$f(\mathbf{q}) = \prod_{c \in C} f(\mathbf{q})_c \quad (2)$$

最後に, この特徴ベクトルを浅い多層パーセプトロン g に入力して値 $I^*(\mathbf{q}) \in \mathbb{R}$ を得る.

$$I^*(\mathbf{q}) = g(f(\mathbf{q})) \quad (3)$$

2.3 共焦点画像の学習

提案手法では, 入力として与えた共焦点画像を再構成するように K-planes の学習を行う. 共焦点画像の画像平面の解像度が $X \times Y$, 奥行方向の解像度が Z であるとき, 画像方向と奥行方向に対応するように 3 つの平面 $\mathbf{P}_{XY} \in \mathbb{R}^{X \times Y \times M}$, $\mathbf{P}_{YZ} \in \mathbb{R}^{Y \times Z \times M}$, $\mathbf{P}_{ZX} \in \mathbb{R}^{Z \times X \times M}$ を初期化する. 学習時は $(x, y, z) \in \{1, \dots, X\} \times \{1, \dots, Y\} \times \{1, \dots, Z\}$ をランダムにサンプルし, K-planes に入力して得られた出力に対して再構成誤差を計算する. 具体的には以下を計算する.

$$\mathcal{L}_r = \frac{1}{XYZ} \sum_{x,y,z} (I(x, y, z) - I^*(x, y, z))^2 \quad (4)$$

この誤差を最小化するように, 各平面上の特徴ベクトルと多層パーセプトロンのパラメータを更新する.

K-planes では再構成誤差に加え, 平面上の特徴に対する正則化を加える. 具体的には, 以下のように各平面上での total variation を計算する.

$$\begin{aligned} \mathcal{L}_{TV} = & \frac{1}{XY} \sum_{x,y} (\|\mathbf{P}_{XY}^{x,y} - \mathbf{P}_{XY}^{x-1,y}\|_2^2 + \|\mathbf{P}_{XY}^{x,y} - \mathbf{P}_{XY}^{x,y-1}\|_2^2) \\ & + \frac{1}{YZ} \sum_{y,z} (\|\mathbf{P}_{YZ}^{y,z} - \mathbf{P}_{YZ}^{y,z-1}\|_2^2 + \|\mathbf{P}_{YZ}^{y,z} - \mathbf{P}_{YZ}^{y-1,z}\|_2^2) \\ & + \frac{1}{ZX} \sum_{z,x} (\|\mathbf{P}_{ZX}^{z,x} - \mathbf{P}_{ZX}^{z,x-1}\|_2^2 + \|\mathbf{P}_{ZX}^{z,x} - \mathbf{P}_{ZX}^{z,x-1}\|_2^2) \end{aligned} \quad (5)$$

最終的な誤差関数はこれらの和であり以下となる.

$$\mathcal{L} = \mathcal{L}_r + \lambda \mathcal{L}_{TV} \quad (6)$$

ここで, λ はハイパパラメータであり, 本研究では $\lambda = 0.001$ とした.

2.4 共焦点画像の補間

学習時は対象とする共焦点画像に存在する奥行 $z \in \{1, \dots, Z\}$ のみを入力する. 学習後は任意の奥行 $z \in \mathbb{R}$ を入力することで, もとのデータセットに存在しない奥行の画像を生成することができる. これは実際には, 式 (1) に従って平面上の特徴空間で線形補間を行ったあと, 式 (3) の多層パーセプトロンによる非線形変換で画像を生成することである. 実験では画像空間上での線形補間との比較を行い, K-planes による補間の有効性を議論する.

3. 実験

3.1 実験データ

眼領域の発達に重要なマスター遺伝子 Rax を蛍光標識したマウス胚性幹細胞から網膜オルガノイドが形成される過程を共焦点顕微鏡で撮影した. 実験では培養を始めてから 3.5 日目と 6.5 日目の 2 つの時期に撮影された共焦点画像を用いた. 3.5 日目は分化の初期であり比較的均一な形状である一方で, 6.5 日目では複雑な立体構造をもつ. 各共焦点画像は画像サイズが 512×512 であり, 奥行方向には 15 枚の画像を撮影した. 画像の 1 ピクセルは $2.486 \mu\text{m}$, 奥行方向の間隔は $17.920 \mu\text{m}$ である. 本実験では定量評価のため, $z \in \{1, 3, \dots, 15\}$ の 8 枚で学習を行い, その間の $z \in \{2, 4, \dots, 14\}$ で評価を行った.

3.2 実験結果

共焦点画像の再構成と補間

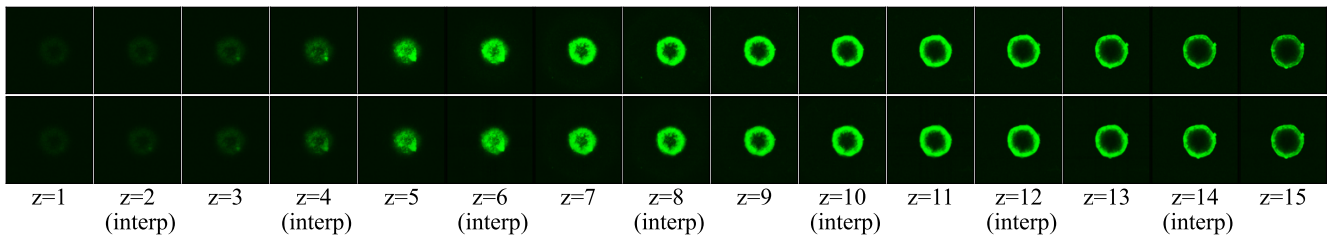
図 2 に共焦点画像の再構成と補間結果を示す. (a) は 3.5 日目, (b) は 6.5 日目の結果である. 上段は撮影した共焦点画像であり, 下段は K-planes で生成した画像である. $z \in \{1, 3, \dots, 15\}$ は学習に使用した画像であるため, 対応する下段の画像は学習画像を再構成したものである. これに対し, $z \in \{2, 4, \dots, 14\}$ は学習に使用しなかった画像であり, 下段は K-planes によって補間した画像である. この結果より, K-planes を用いて学習画像の再構成とその間の補間が可能であることが定性的に確認できる.

画像空間での線形補間との比較

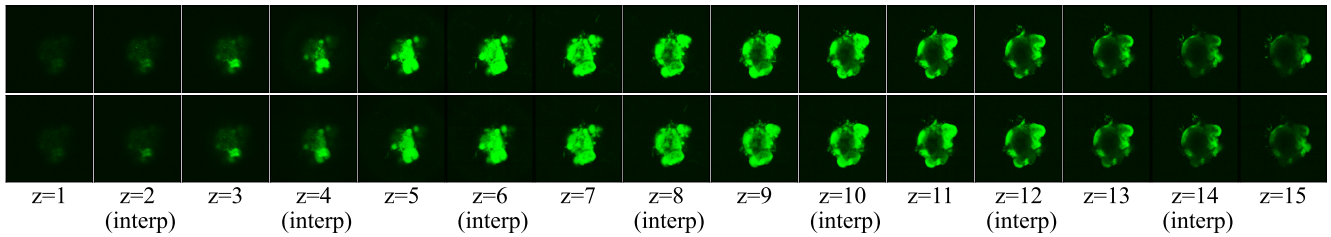
2.4 で述べたように, 提案手法は特徴空間で線形補間をしていることに相当する. この有効性を調査するため, ベースラインとして画像空間での線形補間との比較を行った.

図 3 に画像空間での線形補間との比較例を示す. (a), (b) はそれぞれ 3.5 日目と 6.5 日目の $z = 6$ の補間結果である. 画像空間で線形補間をすると近傍の画像が直接足し合わされるため, (a) の結果に示すように輪郭部分がボケたような画像が生成される. これに対し K-planes は特徴空間で補間を行うため, このようなボケを軽減することが可能である. 一方で K-planes での補間結果には誤差の大きい箇所が点在している.

図 4 にすべての層における補間結果の定量的な比較を示

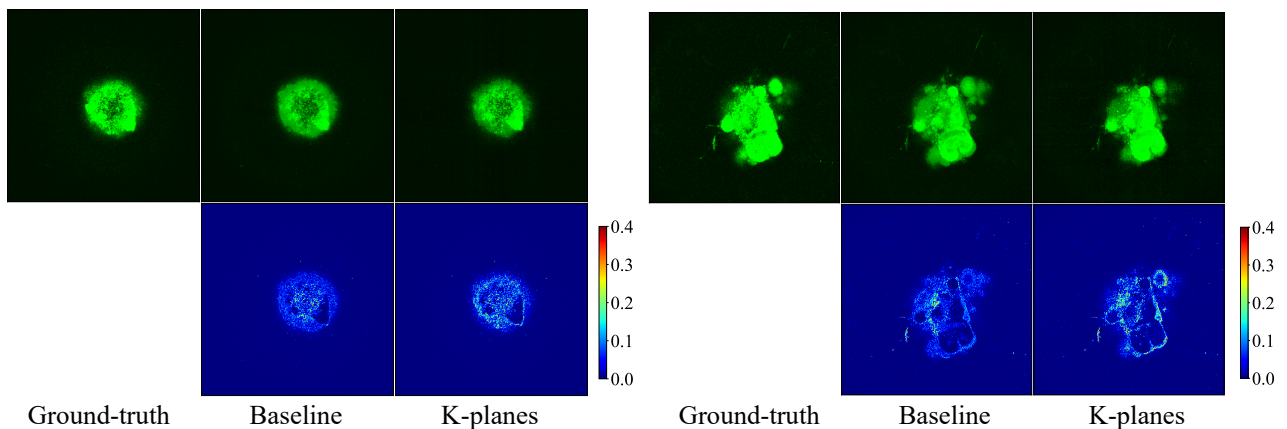


(a) 3.5 日目



(b) 6.5 日目

図 2 (a) 3.5 日目と (b) 6.5 日目の実験結果. 上段は撮影した共焦点画像, 下段は K-planes で生成した画像. $z \in \{1, 3, \dots, 15\}$ は学習に使用した画像であり, $z \in \{2, 4, \dots, 14\}$ は K-planes で補間した画像 (“interp” と表記している).



(a) 3.5 日目

(b) 6.5 日目

図 3 (a) 3.5 日目と (b) 6.5 日目の $z = 6$ の画像について, 画像空間での線形補間 (Baseline) との比較. 下段にはそれぞれ二乗誤差を可視化している.

す. 補間結果の構造的な特徴を評価するため, 評価指標としては SSIM [8] と LPIPS [9] を用いた. ここで, SSIM は大きいほど, LPIPS は小さいほど精度が高いことを表す. (a)(c) に示すように, SSIM の結果では画像空間での線形補間とあまり大きな精度の違いは見られなかった. これに対して (b)(d) の LPIPS を用いた比較では, 奥行が手前の画像ではベースラインの方が精度が高く, 奥の画像では提案手法の方が精度が高いという結果が得られた. 今回用いたデータセットは, 手前より奥で撮影された画像の方が空間的に複雑な構造をもっており, このような画像に対して提案手法は有効であると考えられる. しかしながら, 図 3 でも述べたように提案手法の補間結果には誤差の大きな箇所が点在しており, ベースラインからの大きな精度向上は

みられなかった.

4. まとめ

本研究では, グリッド表現のニューラル場である K-planes を用いた共焦点画像の奥行補間を提案した. 実験結果で示したように, 提案手法は画像空間での線形補間で生じる輪郭部のボケを軽減することが可能であるものの, 誤差の大きな箇所が点在しているため全体的な精度向上はみられなかった. このような誤差を低減するためには, 正則化のパラメータの最適化が必要であると考えられる.

提案手法はニューラル場として K-planes を用いたが, K-planes を NeRF のような新規視点合成に用いる場合は, 3次元空間中の各座標軸に対して等しい解像度でグリッド

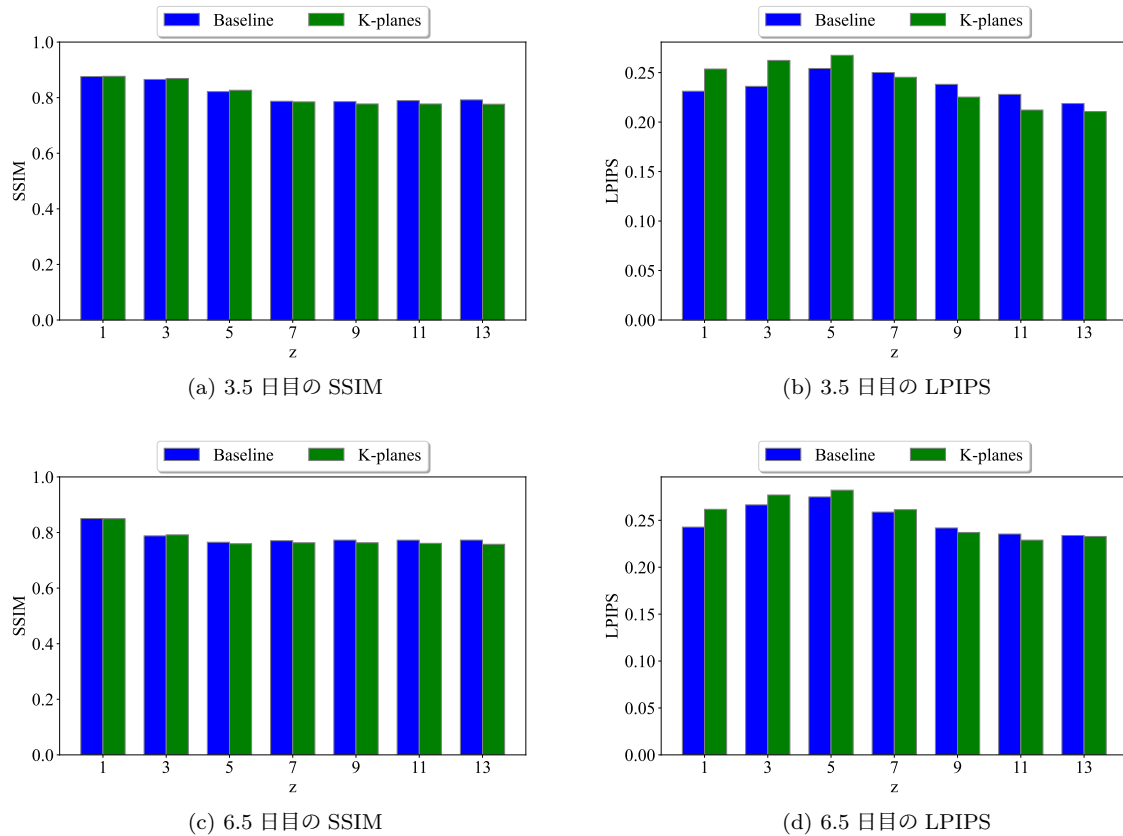


図 4 (a) 3.5 日目と (b) 6.5 日目のそれぞれについて、各層の補間結果の定量評価。ベースラインは画像空間で線形補間を行ったもの。

を定義する。一方で、本研究が対象とする共焦点画像の画像平面と奥行方向を直接グリッドとする場合、画像平面と奥行方向でグリッドの解像度が大きく異なる。今後の課題としては、このようなグリッドの解像度の違いを考慮した、正則化のパラメータの最適化などが挙げられる。

謝辞

本研究は JST CREST JPMJCR23N3 (T.F.), JPMJCR1921 (S.O.), AMED 21bm0704065h0003 (S.O.), JSPS 科研費 JP22K17911 (Y.F.) の助成を受けたものである。

参考文献

[1] Chen, A., Xu, Z., Geiger, A., Yu, J. and Su, H.: TensorRF: Tensorial Radiance Fields, *ECCV* (2022).

[2] Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T. and Sasai, Y.: Self-organizing optic-cup morphogenesis in three-dimensional culture, *Nature*, Vol. 472, No. 7341, pp. 51–56 (2011).

[3] Fridovich-Keil, S., Meanti, G., Warburg, F. R., Recht, B. and Kanazawa, A.: K-Planes: Explicit Radiance Fields in Space, Time, and Appearance, *CVPR*, pp. 12479–12488 (2023).

[4] Fridovich-Keil, S., Yu, A., Tancik, M., Chen, Q., Recht, B. and Kanazawa, A.: Plenoxels: Radiance Fields Without Neural Networks, *CVPR*, pp. 5501–5510 (2022).

[5] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron,

J. T., Ramamoorthi, R. and Ng, R.: NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, *ECCV* (2020).

[6] Müller, T., Evans, A., Schied, C. and Keller, A.: Instant Neural Graphics Primitives with a Multiresolution Hash Encoding, *ACM TOG*, Vol. 41, No. 4, pp. 102:1–102:15 (2022).

[7] Sun, C., Sun, M. and Chen, H.-T.: Direct Voxel Grid Optimization: Super-Fast Convergence for Radiance Fields Reconstruction, *CVPR*, pp. 5459–5469 (2022).

[8] Wang, Z., Bovik, A., Sheikh, H. and Simoncelli, E.: Image quality assessment: from error visibility to structural similarity, *IEEE TIP*, Vol. 13, No. 4, pp. 600–612 (2004).

[9] Zhang, R., Isola, P., Efros, A. A., Shechtman, E. and Wang, O.: The Unreasonable Effectiveness of Deep Features as a Perceptual Metric, *CVPR* (2018).