

拡散モデル正則化を駆使した高精度レンズレスイメージング

大阪大学大学院基礎工学研究科 准教授 中村 友哉 *

1. はじめに

レンズレスイメージングは、従来のレンズを符号化開口や位相マスクのような薄型光学素子に置き換え、入射光をセンサ上に多重化して取得する撮像方式である。計測後の画像再構成処理を組み合わせることで、レンズレス光学系でありながらイメージングを実現する。概念図を Fig.1 に、実現例を Fig.2 に示す。従来型の結像光学系を必要としないため、この手法はカメラの薄型化、軽量化、および広視野化に有望であり、ウェアラブルデバイス、IoT センサ、医療用内視鏡など、小型化や低コスト化が求められる応用において特に期待されている。しかし、得られるセンサ画像は、シーン情報が空間的に絡み合った符号化計測画像であり、シーン再構成は極めて悪条件な逆問題となる。

この問題に対する再構成手法は、大きく三つに分類できる。第一に、ウィナーフィルタや交互方向乗数法(ADMM)アルゴリズム^[1] のような物理モデルに基づく手法は、強い計測整合性を達成するが、知覚品質の面では限定的な再構成結果となる。第二に、U-Net や FlatNet^[2] のような教師あり学習手法は、深層学習の変換自由度を利用することで高速かつ高品質な再構成を可能にするが、物理整合性を明示的に保証するものではない。また、これらの手法は順モデルとは異なるレンズ付きカメラで撮影した画像を教師データとして用いることが多く、光学系パラメータに起因するバイアスが再構成に影響し得る。近年は、第三のアプローチとして、拡散モデルの援用が注目されている。拡散モデルはノイズから段階的に画像を生成する生成モデルの一種であり、自然画像の分布を表現する強

力な事前情報として、さまざまなコンピュータシミュレーションの逆問題に応用されている^[3]。

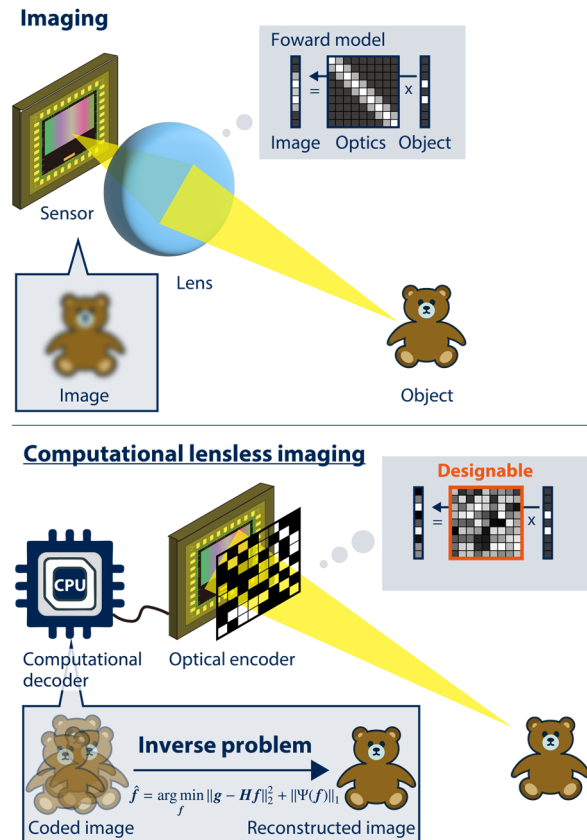


Fig. 1 レンズレスイメージングの概念図

特にレンズレスイメージングにおいては、値域-零空間分解を活用する DDNM+^[4]アルゴリズムが、高い知覚的リアリズムと物理的忠実性の両方を達成できることを示している。しかし、これらの拡散モデルに基づく手法は反復的なサンプリングを必要とするため、リアルタイム推論は実用的ではない。したがって、既存の手法には、計測整合性、知覚品質、推論速度の間に根本的なトレードオフが存在する。

* (助成採択時) 大阪大学産業科学研究所 准教授

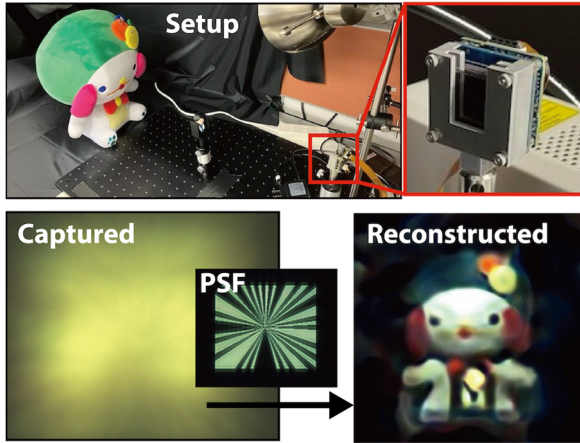


Fig. 2 レンズレスイメージングの実験例

本稿では、値域-零空間分解に基づく DDNM+ の構造化された推論を単一パスのフィードフォワードネットワークへ蒸留する Null-Space Diffusion Distillation(NSDD)を用いたレンズレス再構成アルゴリズム^[5]について、本助成を受けて我々が設計した原理と実験結果を紹介する。

2. 提案手法

レンズレスイメージングシステムにおける画像再構成問題は、次の線形計測物理モデルの逆問題として定式化される。

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \quad \text{Eq.(1)}$$

ここで、 $\mathbf{x}$ は未知のシーンを表現するベクトル、 $\mathbf{y}$ は観測された計測画像ベクトル、 $\mathbf{A}$ は光学系の作用を表現する行列、 $\mathbf{n}$ はノイズベクトルである。DDNM+は、各反復的な拡散サンプリングステップにおける再構成を、計測と整合する値域空間成分と、計測のみからは決定できない零空間成分に分離する。前者は観測に従うよう制約され、後者は拡散モデルによって補完される。これにより、高い知覚的品質と物理的忠実性の両方が達成されるが、およそ 1,000 回程度の反復を必要とするため、高速性が求められる応用に対する実用は困難である。

NSDD は、DDNM+による再構成を教師信号と

して事前計算し、その再構成挙動を単一パスのフィードフォワードネットワークへ蒸留する。具体的には、まず各学習用計測画像に対して DDNM+を実行し、復元画像を教師データとして保存する。次に、DDNM+で使用した拡散モデルと同一の U-Net の重みで生徒ネットワークを初期化し、教師の復元画像との平均二乗誤差を損失関数として最適化する。生徒モデルは、観測された計測画像 $\mathbf{y}$ と擬似逆再構成 $\mathbf{A}^\dagger\mathbf{y}$ の両方を入力とし、最終的な再構成画像 $\hat{\mathbf{x}}_s$ を次のように出力する。

$$\hat{\mathbf{x}}_s = \mathbf{A}^\dagger\mathbf{y} + \hat{\mathbf{r}}(\mathbf{y}, \mathbf{A}^\dagger\mathbf{y}) \quad \text{Eq.(2)}$$

ここで、 $\mathbf{A}^\dagger\mathbf{y}$ はウィナーデコンボリューションによって得られる値域空間成分であり、 $\hat{\mathbf{r}}$ はネットワークによって予測される零空間残差補正である。生徒ネットワークが予測するのは画像そのものではなく、 $\mathbf{A}^\dagger\mathbf{y}$ からの残差のみである。 $\mathbf{A}^\dagger\mathbf{y}$ は計測に整合する成分を既に含んでいるため、ネットワークは計測からは決まらない零空間の情報のみを学習すればよい。このように、蒸留において値域-零空間分解の構造を保持することで、NSDD は反復的サンプリングなしに、物理整合性と知覚品質の両方を維持した高速再構成を達成する。

NSDD の学習は二段階で行われる。第一段階では、21,000 枚のレンズレス計測画像それぞれに対して DDNM+を実行し、教師復元画像を生成する。この処理には単一 GPU で約 52 時間を要する。第二段階では、生成された教師データを用いて生徒ネットワークを約 25 時間で学習する。これらの計算コストは一度きりであり、学習完了後は 1 枚あたり約 0.034 秒での高速推論が可能となる。

3. 実験結果

実験には二つのレンズレスイメージングデータセットを用いた。Lensless-FFHQ は、振幅マスク型レンズレスカメラにより 21,000 枚の顔画像を撮

影したデータセットであり、PhlatCam は異なる位相マスク型レンズレス光学系により取得された多様なシーンのデータセットである。評価指標には、構造的類似度 (SSIM)、知覚品質 (LPIPS)、無参照知覚品質 (CLIP-IQA)、および計測整合性誤差 ($\|y - Ax\|$) を用いた。従来手法、教師あり手法、および拡散モデルに基づく従来の再構成手法の比較を行った。結果、NSDD は、DDNM 教師に近い知覚品質と計測整合性を達成しつつ、推論時間を約 600 倍短縮した。例えば Lensless-FFHQ データセットにおいて、NSDD の SSIM は 0.4984、LPIPS は 0.3807 であり、DDNM+ の SSIM 0.4830、LPIPS 0.3717 と比較して同等の水準を維持している。一方、教師あり手法である FlatNet は SSIM 0.5042 と高いものの計測整合性が大幅に劣り、物理モデルに基づくウィナーフィルタは計測整合性に優れるが SSIM が 0.5432 にとどまり知覚品質は低い。その他網羅的な定量評価結果やその実験条件の詳細は文献^[5]を参照されたい。

さらに、蒸留における値域-零空間分解の有効性を検証するため、全画像を直接蒸留する直截的な手法との比較実験を行った。Lensless-FFHQ において、全画像蒸留の SSIM は 0.4719、LPIPS は 0.4071 であったのに対し、NSDD は SSIM 0.4984、LPIPS 0.3807 と、知覚品質・構造類似度のいずれにおいても優れた結果を示した。この結果は、値域空間成分を解析的に固定し、ネットワークの学習を零空間の補完に集中させることが、蒸留の効率と精度の両方を向上させることを示している。



Fig. 3 OOD 条件下でのレンズレス撮像結果

実シーンへの汎化性能を評価するため、Fig.3 に示すように、レンズレスカメラを試作し、それを用いて実環境の物体のレンズレス計測を行った。再構成用のモデルは、モニタ表示された顔画像のみ

で学習した。この学習条件に対して、撮像実験系は一般拡散物体を用いたため、学習と撮像条件に乖離が発生している（OOD: Out-Of-Distribution）。教師あり手法はこの OOD 条件下の実験において再構成に失敗した一方、DDNM+および NSDD はいずれも、実在する顔に似た物体および非顔物体に対して高品質な再構成を生成した。この結果は、実用化において重要な意味を持つ。レンズレスカメラを任意のシーンで使用するためには、学習時に想定していない被写体に対しても安定して高速に動作することが不可欠であり、NSDD がその要件を満たすことが示された。これらの結果は、値域-零空間分解が実用的なレンズレスイメージングに対するロバストな枠組みを提供し、そのロバスト性が提案する蒸留を通じても保持されることを示唆している。

4. まとめと今後の課題

本稿では、拡散モデルを用いたレンズレスイメージングの再構成手法に関する研究成果^[5]について述べた。レンズレス画像再構成には計測整合性、知覚品質、推論速度の間に根本的なトレードオフが存在し、従来手法はこれらを同時に満たすことが困難であった。我々の提案した NSDD 法は、DDNM+法の値域-零空間分解に基づく構造化された推論過程を単一パスのネットワークへ蒸留することで、物理整合性と知覚品質を維持しながら約 600 倍の高速化を実現した。さらに、学習時に想定していない実物体に対しても高速かつ安定した復元を示し、実用的なレンズレスイメージングへの適用可能性を確認した。

一方で、本手法にはいくつかの課題が残る。第一に、提案手法の性能は計測物理モデルおよび擬似逆変換 $\mathbf{A}^\dagger$ の精度に依存しており、現状ではシフト不変な点像分布関数(PSF: Point Spread Function)を仮定している。実際のレンズレスカメラでは、視

野周辺部で PSF が空間的に変化するため、空間変化 PSF への対応は実用化に向けた重要な課題である。第二に、本研究では二次元シーンの復元のみを扱ったが、三次元シーンへの拡張や、マルチスペクトルイメージングへの拡張も有望な方向である。第三に、蒸留の教師モデルである DDNM+自体の性能向上、例えばより高精度な計測物理モデルの導入やノイズモデルの改良も、NSDD の復元品質をさらに改善し得る。拡散モデルに基づくレンズレスイメージングは、薄型・軽量な次世代撮像システムの実応用に向けた有望な研究方向であると考ええる。

5. 参考文献

- [1] N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, R. Ng, and L. Waller; DiffuserCam: lensless single-exposure 3D imaging, *Optica*, 5 - 1, (2018), pp. 1–9.
- [2] S. S. Khan, V. Sundar, V. Boominathan, A. Veeraraghavan, and K. Mitra; FlatNet: Towards Photorealistic Scene Reconstruction From Lensless Measurements, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44 - 4, (2022), pp. 1934–1948.
- [3] H. Kawachi, J. Neto, Y. Yagi, H. Nagahara, and T. Nakamura, “Single-Image Depth from Defocus with Coded Aperture and Diffusion Posterior Sampling,” *IEEE Transactions on Computational Imaging*, 12, (2026), pp. 1034–1045.
- [4] Y. Wang, J. Yu, and J. Zhang; Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model, *The International Conference on Learning Representations (ICLR)*, (2023).
- [5] J. Neto, H. Kawachi, Y. Yagi, and T. Nakamura; Null-Space Diffusion Distillation Unlocks

Speed, Fidelity and Realism in Lensless
Imaging, IEEE International Conference on
Computational Photography (ICCP), (2026).

6. 謝辞

本研究は島津科学技術振興財団の助成を受けて
実施された。