

# 連続撮影されたSPM画像による 相互補完的なノイズ除去

安江 楓真<sup>1,a)</sup> 近藤 佑樹<sup>1,b)</sup> 新庄 康太<sup>1,c)</sup> 秋田 和俊<sup>1,d)</sup> 三星 響<sup>1,e)</sup> 吉村 雅満<sup>1,f)</sup>  
浮田 宗伯<sup>1,g)</sup>

**概要：**走査型プローブ顕微鏡 (Scanning Probe Microscopy; SPM) は様々な分野で欠かせない分析装置の1つである。しかし様々な要因から測定画像にノイズが発生してしまうというデメリットがある。ノイズが発生した場合にはノイズが少ない画像が得られるまで撮影を続ける必要がある。この問題に対して、SPMのパラメータを調整するというハードウェア的な解決策も有るが、最適なパラメータが試料によって異なることから難易度が高く、時間もかかってしまう。そこで、深層学習を利用しソフトウェア的なアプローチを行う。本研究ではノイズが入った複数枚のSPM画像を用いてノイズ除去を行う手法を提案する。この手法は従来の1枚の画像からのノイズ除去手法と異なり、大きなノイズにも対応できるというメリットを持つ。提案手法に合ったSPM画像のデータセットは公開されていなかったため、本研究ではSPM画像のデータセット作成を行い、学習ペア画像選択のためにSPM画像のノイズ定量化手法を提案する。提案手法により生成された画像は定量的・定性的にノイズが減少していると評価された。

**キーワード：**走査型プローブ顕微鏡, 原子間力顕微鏡, ノイズ除去, 深層学習

## 1. はじめに

走査型プローブ顕微鏡 (Scanning Probe Microscopy; SPM) は、探針により試料表面を走査して表面形状を得る顕微鏡の総称である。SPMには走査型トンネル顕微鏡 (Scanning Tunneling Microscopy; STM)[1] や原子間力顕微鏡 (Atomic Force Microscopy; AFM)[2] などの種類がある。AFMやSTMは試料に物理的に接触せず測定を行うことから、非破壊で試料表面の形態を測定できるというメリットがある。また、試料の高さ方向に原子スケールの分解能を持っている[3]。このような特徴からSPMは製造分野や物質分野などで広く利用されている。

しかしSPMでは、測定結果にノイズが発生することがある。ノイズの原因の1つに、走査するSPMの探針と試料の誤った接触がある。図1(a)と(b)は同じ試料・同じ箇所のカボンナノチューブ (Carbon Nanotube; CNT) を撮影

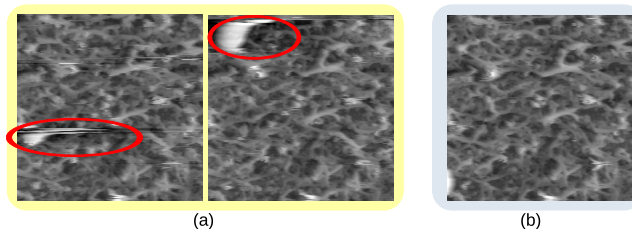


図1 (a) ノイズが多いSPM画像, (b) ノイズが少ないSPM画像。

した3枚のAFM画像である。図1(a)では図1(b)と比較すると画像赤丸部分にノイズが発生していることが確認できる。ノイズは横に引き伸ばされたように発生している。これはSPMが横方向の走査によって像を得ていることに由来する。ノイズが多く発生した場合は、ノイズが少ない画像が得られるまで再撮影を行う必要がある。この問題に対し、SPMのフィードバック感度であるGainや、針の高さを調節するSetpointを変更するといったハードウェア面での解決策もあるが、これらは試料によっても最適な値が異なるため経験と知識が必要となり、難易度が高い。

そこで本研究では、ソフトウェアによるアプローチでSPM画像のノイズ除去を実現する。Computer Vision(CV)分野では、現在の深層学習による画像処理の基礎となる畳み込みニューラルネットワーク (Convolutional Neural Network;

<sup>1</sup> 豊田工業大学  
Toyota Technological Institute  
a) sd22434@toyota-ti.ac.jp  
b) yuki.kondo.ab@gmail.com  
c) sd19049@toyota-ti.ac.jp  
d) sd21501@toyota-ti.ac.jp  
e) sd22430@toyota-ti.ac.jp  
f) yoshi@toyota-ti.ac.jp  
g) ukita@toyota-ti.ac.jp

CNN)の登場によりノイズ除去の手法が確立されており、CNNを利用したノイズ除去手法は多数存在している。しかし、図1(a)赤丸部分のような大きなノイズを除去することは難しい。なぜなら、1枚の画像入力のみではノイズ領域を補完するための情報が不十分だからである。

そこで、大きなノイズに対応できるノイズ除去手法として、本研究では複数枚のSPM画像を利用したノイズ除去を提案する。これには、SPMで同じ箇所を複数回撮影することでノイズの位置が変化することを利用する。

例えば図1(a)では、2枚の画像でノイズの位置が異なっている。提案手法ではこのようにノイズの位置が異なる画像をネットワークに入力する。ネットワークは画像のノイズ箇所を相互に補完することで、ノイズ除去を行った画像を出力する。そのためこの手法では、従来手法では難しい課題である大きなノイズ部分でも、他の画像で補完することでボケを抑制しつつノイズ除去できる。ここで入力する複数枚の画像はノイズが入っていても良いため、スクリプトにより連続撮影をすることで入手できる。そのため全体がきれいな画像を撮影する必要がなくなり、ハードウェア的な調整の難易度と手間を軽減することができる。

また、本研究ではSPM画像データセットの作成を行い、SPM画像のノイズ定量化手法についても提案する。

## 2. 関連研究

### 2.1 SPM画像のノイズ除去手法

一般的な画像のノイズ除去は多数研究されているが[4],[5],[6],[7]、SPM画像には一般的な画像ノイズと異なる特殊な性質がある。よって、この特性を考慮したノイズ除去が望まれる。

Zhaoらは、2021年に機械学習を用いた透過型電子顕微鏡法(Scanning Transmission Electron Microscope; STEM)のノイズ除去手法を発表した[8]。STEMではノイズの少ないスロー撮影と、ノイズが大きい高速撮影が行える。この研究ではU-Net[9]を用いて高速撮影による画像のノイズ除去を行うことで、結果として高速撮影でもスロー撮影と遜色ないレベルまでノイズが減少させることに成功している。学習はSTEMの高速撮影によるノイズが多い画像1枚をU-netの入力とし、50枚の高速撮影の平均値画像を教師データとしている。

しかしこの手法やSPM画像に対する従来のノイズ除去[10]では、予測による補完のため正確性に欠けるという問題点がある。

### 2.2 複数画像を用いたノイズ除去手法

複数画像を用いたノイズ除去手法にKPN(Kernel Prediction Network)がある[11]。ノイズ除去学習時の教師画像にはノイズのない画像を使用し、この画像にガウシアンノイ

ズの付与と移動を行って実画像を再現した合成画像複数枚を入力画像としている。訓練時は合成画像を用いているが、テスト時には実際にカメラで連続撮影された複数枚の画像を用いてもノイズ除去が成功したと示されている。

しかしSPMではノイズのない教師画像を得るコストが大きく、データセットとして大量に得ることが難しいため、この手法を適用することが難しい。また、ノイズがない画像が得られたとしても、SPM画像のノイズは振動や探針の状態、走査によるノイズ、熱ドリフトなど様々な要素が関わってくるため実際のノイズを再現するのが難しい。これらの理由から、本研究では入力画像を生成する方法は採用しなかった。

## 3. 提案手法

### 3.1 複数画像を用いたノイズ除去

従来のノイズ除去手法は、大きなノイズを除去し正確に復元することが難しかった。そこで本研究では複数枚の画像を用いることで、大きなノイズに対しても正確なノイズ除去を提案する。

ノイズ除去の概要を図2に示す。まずSPMにより同じ箇所を $n$ 回撮影した画像を用意する。これを1組の画像と呼ぶ。これらの画像のうち、1枚を教師画像とし、残りの画像のうち $m$ 枚( $1 \leq m \leq n-1$ )を学習画像としてネットワークに入力することで学習を行う。この時入力する枚数は実験を行って比較する(4.6.3節)。

提案手法による学習には同じ箇所を複数回撮影した画像が必要だが、SPMによって同じ箇所を複数枚撮影した公開データセットは存在していない。そこで本研究では実際にSPMで試料を撮影したデータセットを作成する。詳しい撮影条件や使用した試料については4.1節で説明する。今回撮影した画像は1組の画像の中でも移動や変形が発生しているため、位置合わせを行う。位置合わせについては3.2節で説明する。

次に、ネットワークにノイズ除去を学習させるためには教師画像はノイズが少ない画像でなければならない。なぜならば、教師画像としてノイズが多い画像を使用した場合には、ネットワークはノイズの多い画像を出力するように学習してしまうからである。しかしSPMではノイズが全く含まれない画像を得ることが出来ないため、提案手法では1組の画像のうち最もノイズが少ない画像を教師画像とする。大量の画像が含まれるデータセットに対しノイズが少ない教師画像を人力でラベリングすることは現実的ではないため、本研究では前処理として画像のノイズを定量化し、教師画像を選択する手法を提案する。SPM画像のノイズ定量化については3.3節で述べる。

SPMは撮影に時間がかかるため、今回用意できたデータセットは画像枚数が十分ではない。そこでデータ拡張を行

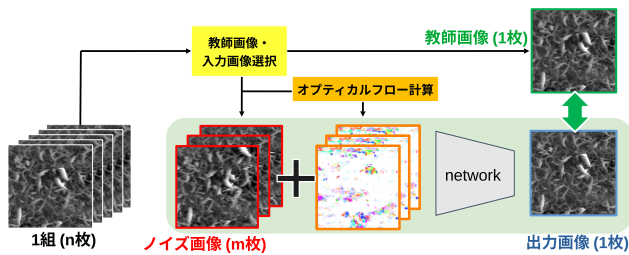


図2 提案手法によるネットワーク学習の概略.

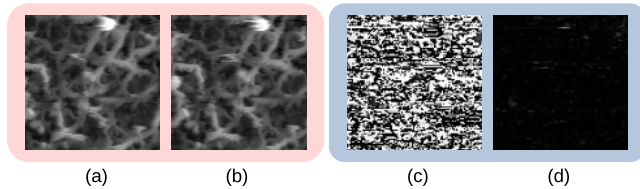


図3 オプティカルフローにより画像の位置合わせを行った例. (a) 入力画像1, (b) 入力画像2, (c) aとbの差分, (d) オプティカルフローによる位置合わせ後の差分.

う必要がある. データ拡張については3.4節で述べる.

詳細なネットワークの構造については4.2節で述べる.

### 3.2 位置合わせ

今回撮影したデータセットはノイズがあるだけでなく, 画像内で熱ドリフトによる [12] 非線形な変形や移動がある. そこで, 位置の基準となる画像を1枚決め, 他の画像をそこに合わせて変形する. これをワーピングと呼び, 提案手法では入力画像とともにオプティカルフローを入力することでノイズ除去ネットワークにワーピングを学習させて行う.

オプティカルフローにより画像のワーピングを行う例が図3である. 図3(a)(b)は2枚の入力画像で, 同じ箇所を連続で撮影した画像である. この画像ではほぼ同じように見えるが, 実際にはわずかに移動や変形がある. 差分を示したのが図3(c)で, (a)(b)間の差分の大きさを表している. 白い箇所は差が大きく, 黒い箇所は差が小さい. 図3(c)は全体に白が含まれる画像になっているため, 図3(a)(b)は位置ズレや変形があり位置が一致していないものと考えられる. (d)はオプティカルフローによって位置合わせをした画像2枚の差分である. 全体が黒っぽいことから, 位置合わせ後の画像は差が少ない. このことから, オプティカルフローを利用することで位置合わせが可能だと分かる.

以上のようなワーピングは入力前に前処理として行うこともできる. そこで, 暗的にワーピングを行う場合と, 事前にワーピングを行ってから入力する場合4.6.2節で比較する.

### 3.3 画像微分を用いたノイズの定量化

SPMは探針の走査方向に引き伸ばされるようなノイズがあることを利用し, 画像微分を用いたノイズの定量化手法

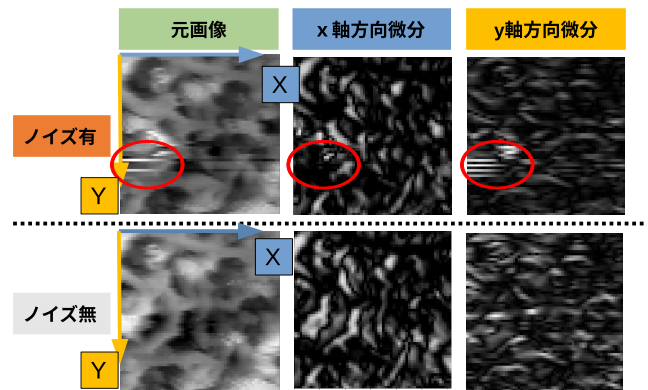


図4 ノイズの有無による画像微分の大きさの変化.

を提案する. 画像微分とは, 微分フィルタを画像の各画素に適用し, 結果として画像を得る処理である. 図4に画像微分を行った例を示す. 画像微分した結果はマイナスを含むため絶対値を取り, 画素の最大値が255になるよう正規化することで可視化している. 各軸方向の画素値の勾配が大きい箇所は, 微分結果の画像は明るく示される. 対して画素値の勾配が小さい箇所は微分の絶対値が小さくなり, 微分結果が暗く示される.

まずノイズが多い画像の微分を確認する. 図4左上にノイズが入った元画像, 図4上段中央と右に画像微分を行った結果を示している. 図4左上の赤丸部分に注目すると  $x$  軸方向に同じ画素値が続いている. そのため微分の値は小さく,  $x$  軸方向の画像微分の結果(図4上段中)では暗く表れている. 一方で図4左上の赤丸部分  $y$  軸方向に注目すると画素値の変化が大きい. そのため微分の値は大きく,  $y$  軸方向の微分結果(図4右上)では明るい模様が現れている. これらより, ノイズが多い画像では,  $x$  軸方向の微分値の総和と  $y$  軸方向の微分値の総和を比べた場合に差が大きいと考えられる.

一方で, ノイズが少ない画像を図4下段に示している. 各軸方向の微分結果は似ていることから, ノイズが少ない場合には  $x$  軸方向の微分値の総和と  $y$  軸方向の微分値の総和を比べた場合に, 差が小さいと考えられる.

以上より, 画像微分によるノイズ量  $Q_{diff}$  を以下のように定義する.

$$Q_{diff} = \left| \sum_i \sum_j |d_{x,i,j}| - \sum_i \sum_j |d_{y,i,j}| \right| \quad (1)$$

$|d_{x,i,j}|$  と  $|d_{y,i,j}|$  はそれぞれ座標  $(i,j)$  の画素に対して  $x$  軸,  $y$  軸方向に画像微分を適用して得られた画素値である. この手法によりノイズ量を定量化し, ある1組の画像を  $Q_{diff}$  によって並び替えると図5のようになる. この例では, 人間が見て定性的にノイズが多い画像ほど  $Q_{diff}$  も高くなっていることがわかる.

提案手法では, 図5において, 最もノイズが少ない画像である緑枠の画像を図2での緑枠の教師画像として利用す

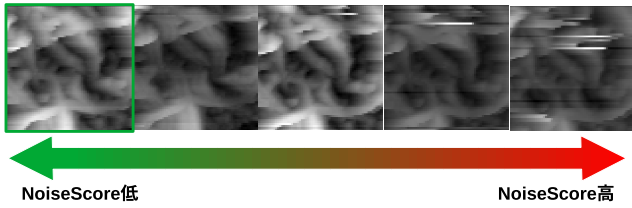


図5  $Q_{diff}$  によって1組の画像をソートした例.

る。また、教師画像以外の画像のうち複数枚を選択してノイズ画像としてネットワークに入力する。

以上より、教師画像として最もノイズの少ない画像を選択出来たことで、ネットワークにノイズ除去を学習させることができる。

### 3.4 データ拡張

一般的にネットワークの学習をするためには大量のデータが必要になるが、SPMは撮影に時間がかかるため、大量の画像を用意するのが困難である。実際に、今回のデータセットは全部で162組しかない。そのためそのまま学習すると、ネットワークがデータセットに対して過剰に適合してしまう「過学習」に陥る可能性がある。

そこで、データ拡張(Data Augmentation)を行う。データ拡張とは、データセットの画像に様々な条件の変換を加えて画像を増やす手法である。本研究では画像のランダムな部分を切り抜くランダムクロップ、画像の上下左右反転、ガンマ補正を行った。ガンマ補正とは、画素値を以下の式に従って変更し明暗を調節する処理である。

$$x_{gamma} = 255 \times \left( \frac{x}{255} \right)^\gamma \quad (2)$$

$x$ は元画像、 $x_{gamma}$ は補正後の画像、 $\gamma$ はパラメータである。実験では $0.8 \leq \gamma \leq 1.1$ でランダムに決定した。また、上下・左右反転はそれぞれ独立に $\frac{1}{2}$ の確率で行った。

ネットワークに入力する画像サイズを学習中に変更することは出来ないため、ランダムクロップのサイズは今回の実験では一定に $64 \times 64$ とした。また、ランダムクロップは1組の画像に対して同じ箇所で行い、切り抜いたあとの組に対してノイズの定量化による教師画像と入力画像の選択をする。

本研究ではランダムクロップに過学習防止以外のメリットもある。SPMのデータセットでは図1で示したような局所的なノイズが多い。そのため一部分を切り抜くことにより局所的にノイズが少ないきれいな領域を選択しやすくなるため、教師画像にノイズが少なくなることが期待でき、ノイズ除去を学習しやすいと考えられる。

## 4. 実験

### 4.1 データセット

本研究のテーマに沿った公開データセットは存在していないため、本研究では複数画像によるノイズ除去のための

SPM データセットを作成した。

撮影した試料の条件は以下である。

- 基板：SiO<sub>2</sub>(300nm)/Si
- サイズ：5 × 5mm

CNT 試料の撮影は AFM で行なった。条件は以下である。

- 装置：OXFORD INSTRUMENTS Asylum Research 社製 MFP-3D Origin
- 一箇所あたり 10 枚撮影し 1 組の画像とする
- 撮影範囲：2 × 2 μm
- 解像度：256 × 256 ピクセル
- データセット 1：学習用画像 143 組 + テスト用画像 19 組
- データセット 2：学習用画像 76 組 + テスト用画像 19 組

撮影された画像は、画像全体でノイズが多すぎる組を含む。そのため1組の画像全てでノイズが多い組を目視で判断し取り除いたデータセットをデータセット2とした。また、それを行わず撮影された学習用画像を全てを含むデータセットをデータセット1としている。4.6.4節では2つのデータセットでそれぞれで学習し、比較を行う。

### 4.2 ネットワークの詳細

提案手法の実装には、映像や連続時系列画像の超解像 [13], [14], [15], [16] のために提案されている RBPN(Recurrent Back-Projection Network for Video Super-Resolution) を利用する [17]。RBPN は DBPN[18], [19] という画像超解像ネットワークを動画超解像タスクに拡張したネットワークで、動画のある1フレームと前後の数フレームでの物体の写り方の変化を利用して元動画より高い解像度の動画を出力する。RBPNの特徴は、回帰構造をもちフレームを1枚ずつ入力して処理することである。

本研究では画像超解像はせず、複数枚入力できるという特徴を利用しノイズ画像複数枚をネットワークに入力する。RBPNは超解像タスクのための構造であるため、そのまま利用すると解像度 $(x, y)$ の画像は $(2x, 2y)$ などに拡大される。そのため最終層で元の解像度 $(x, y)$ にダウンスケーリングした画像を最終的な出力とした。

3.2節でも説明したように、AFMで撮影された画像は、熱の影響により非線形な変形や移動があるため、オプティカルフローを画像とともに入力することで、ネットワークに画像の位置合わせを暗的に学習させることを目指す。RBPNを用いたノイズ除去器の概略を図6に示す。ネットワークに入力するノイズ画像を $m$ 枚とすると、1枚目から $m$ 枚目の入力画像は $I_1$ から $I_m$ で表されている。ネットワークに最初に入力される画像 $I_1$ はベースフレームと呼ぶ。ベースフレームは図6では最も左に $I_1$ で示されている画像である。

2枚目の画像 $I_2$ は、ベースフレーム $I_1$ と入力画像 $I_2$ の



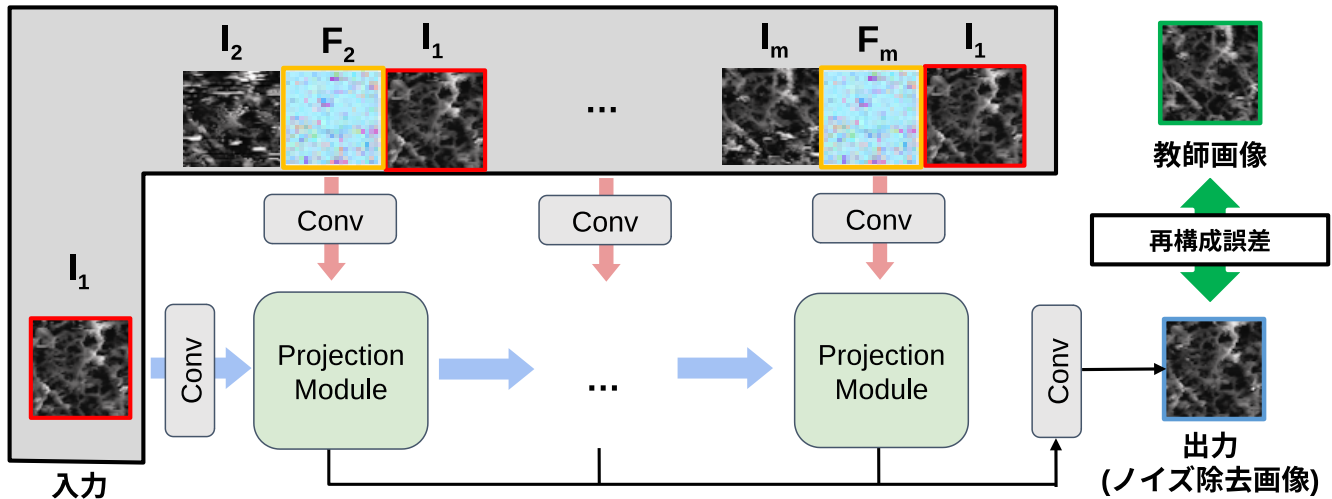


図6 RBPNを用いたノイズ除去器の概略.  $m$  枚目の入力画像とオプティカルフローがそれぞれ  $I_m, F_m$  で示されている.

間のオプティカルフローと、ベースフレーム  $I_1$  をチャンネル方向に結合して入力する。オプティカルフローを入力するのは、3.2 節でも述べたとおり変形や移動に対してロバストな学習をするためである。ベースフレームに対するオプティカルフローであるため、ベースフレームも入力することでネットワークが解釈しやすくなると考えられる。これ以降  $m$  枚目は同様に、 $I_m, F_m, I_1$  が入力される。

### 4.3 学習条件

本研究で作成したデータセットは画像枚数が少ないため、ファインチューニング (Fine Tuning) を行う。ファインチューニングとは、別のデータセットで学習された学習済みモデルを再度新たなデータセットで学習してモデルの重みを微調整する手法で、目的のデータセットが少ない場合に特に有効である。

今回の場合では SPM 画像のデータセットが少ないため、風景や人物などの動画からなる一般データセットにより学習されたモデルを使用する [17]。使用する学習済みモデルは超解像タスクのためのモデルであるが、ノイズ除去のタスクとして SPM 画像で再度学習することで、SPM 画像に特化した学習を行うことができた。後の節で述べる実験では、特に示さない限り学習済みモデルを利用する。

その他、学習時に変更できるパラメータのうち、本研究の実験で変更しないパラメータを下記に示す。

- モデルの最適化 (Optimizer): Adam(Adaptive moment estimation)[20]
- 損失関数 (loss): MAE(Mean Absolute Error)
- 学習率 (learning rate):  $10^{-4}$
- ミニバッチサイズ: 12

MAE は、2 枚の画像でピクセルごとの画素値の差の総和であり次のように定義される

$$L_{MAE} = \frac{1}{N} \sum_{n=1}^N |I_{target,n} - I_{output,n}| \quad (3)$$

$L_{MAE}$  は再構成誤差、 $I_{target,n}$  は  $n$  枚目の教師画像の画素値、 $I_{output,n}$  は出力画像の画素値である。 $N$  はミニバッチサイズである。

### 4.4 ノイズ除去の評価指標

SPM で撮影された画像においてノイズ除去の真値は不明である。そこでテストデータで評価する際も、教師画像と同じようにノイズが少ない画像をデータセットから選択しておき、残る画像で行ったノイズ除去の出力画像と比較することによって定性評価・定量評価を行う。

ノイズ除去は 3.3 節で提案したノイズ定量化手法の式 1 による  $Q$  で評価する。この値は小さいほどノイズが少ないことを表す。評価はテスト画像 19 組全てで行い、その平均値を評価結果とする。

画像全体の再現度を、SSIM (Structural similarity index measure) で評価する [21]。なぜなら SPM の画像は教師画像と出力画像で位置ズレがあるため、画像処理分野の一般的な評価指標である MSE (Mean Squared Error) や、PSNR (Peak Signal-to-Noise Ratio) を利用することが出来ないからである。SSIM は画像のピクセルごとの一致ではなく、画像の輝度やコントラストが画像内でどのような相関を持っているかを広域的に評価する手法で、画素ごとの比較を行わないため画像の位置ずれに影響されづらい。SSIM は 1 に近づくほど 2 枚の画像の構造が似ていることを表す。そのため教師画像と出力画像の比較において、教師画像にノイズが多く含まれている場合にはノイズ除去がうまく出来ても SSIM が低い値になってしまう。実際、テストデータセットでは教師画像であってもノイズが含まれている画像がほとんどである。そこで、テストデータの中で、教師画

像にはほぼノイズが含まれない1組の画像のみをSSIMの評価に使用することとする。1組の画像だけで正確に評価するために、1組の画像から複数通りのSSIMを計算し、その平均を最終的なSSIMとする。例えばノイズ除去の入力画像が7枚であれば、10枚1組の画像のうち1枚を教師画像とし、残った9枚のうち7枚を入力画像とするので入力画像は ${}^9C_7 = 36$ 通り選択することができる。つまり36通りのSSIMの平均を最終的なSSIMとする。

ノイズ除去を行った画像のボケを、ラプラシアンフィルタによる手法で評価する。このフィルタは主にエッジ検出のために用いられ、エッジがはっきりしている画像では畳み込んだ値が大きくなり、ボケた画像ではエッジが失われるため値が小さくなる。つまり画像の鮮明さを表す値となる。この指標は、SSIMと同様、ノイズが多く含まれている場合には正しく評価できない。なぜなら、ノイズ部分では画像値の変化が大きく、ラプラシアンフィルタの値が大きくなるからである。そのため、教師画像とノイズ除去結果の鮮明さを比較するためにはノイズが少ない教師画像のみを評価対象とする必要がある。そこで、鮮明さについてもSSIMと同様に1組の画像のみを用い、複数通りの鮮明さを計算しその平均を最終的な鮮明さとする。

#### 4.5 ノイズ定量化手法の評価

3.3節で提案した画像微分によるノイズの定量化手法を評価する。評価のため、データセット2から31組の画像を使用した。提案手法と目視でそれぞれノイズが最も少ない画像を選択し、一致率を調べる。

データセットには見た目がほぼ同じ画像が多数含まれていたため、目視でノイズが少ない画像を選択する際には同程度きれいな画像を任意の枚数選択可とした。図7に選択する画像の例を示した。緑枠で囲ったのは目視で選択した画像、赤は選択されなかった画像である。また、星をつけた画像は提案手法で選択した画像である。

この方法による一致率は80.6%だった。つまり提案手法による選択は概ね人間の選択と一致するといえる。しかしまだ改善の余地があり、良い学習データを作成するためにはノイズ定量化と教師画像選択の正確性をさらに高める必要がある。

#### 4.6 ノイズ除去

##### 4.6.1 ベースとなるノイズ除去結果

この節では提案手法によるノイズ除去の実験結果について述べる。データセットは4.1節で示したデータセット1を使う。1組中の画像から最もノイズが少ない画像を教師画像とし、残る画像の中から7枚をランダムに選択する。選択した画像の中でもっともノイズが少ない画像をベースフレームとし、残りはランダムに並び替えてネットワーク

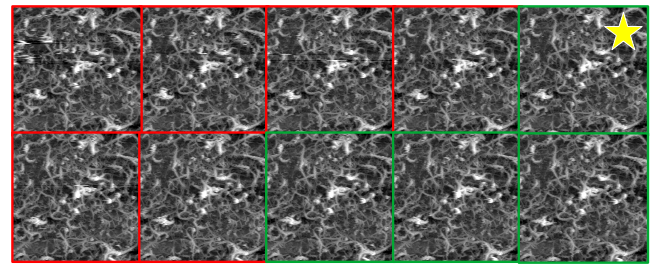


図7 似ている画像の例。緑で囲った画像は、目視でノイズが少ないと判断した画像。赤で囲った画像はそれ以外の画像である。星がついた画像は、提案手法により選択された画像。

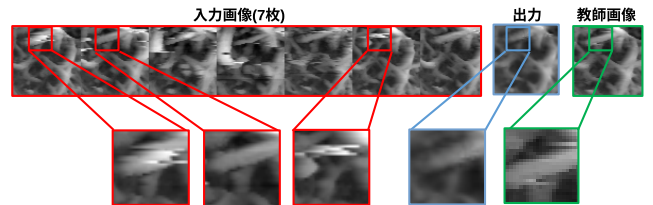


図8 ベースとなるノイズ除去の結果。

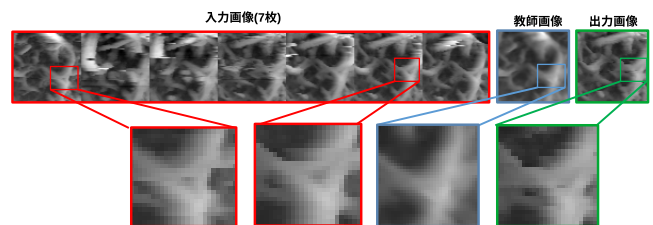


図9 ノイズ除去の結果としてボケが発生した例。

に入力する。

上記の条件で訓練したモデルでテストした結果の中から、ノイズ除去に成功している例を図8に示した。上段のうち赤枠に示されているのが入力画像、青枠に示されているのが出力画像、緑枠に示されているのが教師画像である。また、下段にはそれぞれの画像で右下を拡大して示している。入力画像ではノイズが多いが、出力画像では教師画像と同じようにノイズが少ないことが確認できる。しかし、この出力画像は教師画像や入力画像に比べて若干ボケており、輪郭が不明瞭になってしまっている。

次に、ノイズ除去が失敗している例を図9に示す。図9では教師画像や入力画像に比べて出力画像は大きくボケてしまっている。この例では入力画像間での構造の変化が大きいため、ネットワークが位置合わせできていないと考えられる。また教師画像の移動は入力画像からは予測できないことから、ボケた画像を出力することにより損失を最小限に留めるような学習が行われこのような結果が得られたと考える。

表1に定量評価の結果が示されている。評価指標は4.4節で示したものである。それぞれの列で最も良い結果は赤字で、最も悪い結果は青字で示されている。

表1より、ベース結果のノイズは教師画像に比べて半減

表 1 実験によって得られた出力画像のノイズ量・SSIM・鮮明さの定量評価。列名の横の↑は高いほど良い値であること、↓は低いほど良い値であることを示す。赤字はその列で最も良い結果、青字はその列で最も悪い結果を示している。

	ノイズ量 $Q$ ↓	SSIM ↑	鮮明さ ↑
教師画像	0.0328	-	936
ベース結果 (4.6.1 節)	0.0157	0.474	266
位置合わせ (4.6.2 節)	<b>0.0221</b>	<b>0.799</b>	<b>300</b>
入力枚数 9 枚 (4.6.3 節)	<b>0.0121</b>	0.584	<b>197</b>
入力枚数 4 枚 (4.6.3 節)	0.0175	<b>0.305</b>	258
入力枚数 1 枚 (4.6.3 節)	0.0192	0.338	272
データセット 2 (4.6.4 節)	0.0175	0.590	231

しているが、鮮明さは下がっていることが定量的に確認できる。SSIM は後の節で説明する実験と比べると低い値であり、教師画像を再現できていないといえる。

#### 4.6.2 位置合わせによるノイズ除去結果の変化

提案手法ではオプティカルフローをネットワークに入力し、ワーピングは行わなかったが、ワーピングを行ってから入力した場合にどのような結果となるか実験する。ワーピングは、基準となるベースフレームに他の入力画像と教師画像を合わせるように行う。テスト時は、入力画像の位置合わせのみ行う。

位置合わせした画像によりノイズ除去の学習を行い、テストした結果を図 10 に示した。上段は位置合わせせずに学習した結果で、下段は位置合わせした画像を入力として学習し、位置合わせした画像によってテストした結果である。左側には画像のボケが改善した例、右側には改善しない例を示している。成功例の赤丸部分に注目すると位置合わせ後ではボケが少ないことが分かる。しかし、右の失敗例の黄丸部分に注目すると、位置合わせしていない元画像ではノイズが少ないが、位置合わせ後ではノイズが大きく残ってしまっている。表 1 の定量的評価を確認すると、SSIM と鮮明さは最も良い値であるものの、ノイズ量はベースの結果に比べて増加しており、定性評価と一致している。実際に SPM でノイズ除去を行う状況と考えたときには、ボケの改善よりもロバストなノイズ除去が重要であることから、明示的にワーピングしない手法に優位性があると考えられる。

ワーピングによってノイズが増えてしまったのは、位置合わせに原因があると考えられる。位置合わせするにはオプティカルフローを正しく計算する必要があるが、ベースフレームにノイズがあるとオプティカルフローがうまく推定できない部分が生まれてしまう。推定が失敗したオプティカルフローで学習時の教師画像を位置合わせすると、教師画像にノイズが移ってしまい教師画像のノイズが増えてしまう。そのため、ノイズを除去するような学習ができず、ノイズが除去されなかったのだと考えられる。

#### 4.6.3 入力画像枚数に対する出力画像の変化

4.6.1 節では 7 枚の入力によるノイズ除去を行ったが、入

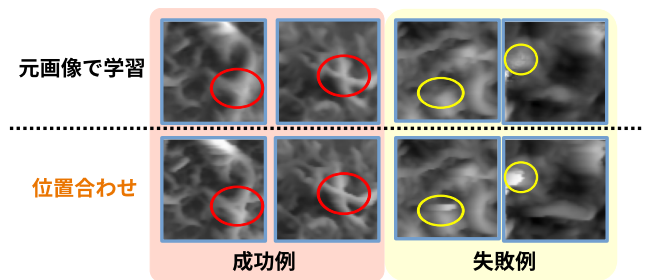


図 10 オプティカルフローによる位置合わせ後にノイズ除去を行った場合の成功例と失敗例。成功例ではボケが比較しやすい部分を赤丸で示している。失敗例では位置合わせによりノイズ除去に失敗した部分を黄丸で示している。

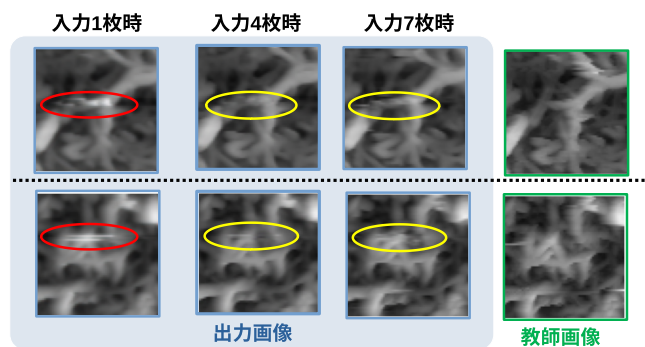


図 11 入力枚数を 1, 4, 7 枚と変化させた場合の出力画像と、教師画像の例。

力枚数を増加させるメリットを得られているか確認する。4.6.1 節と同様の条件で、入力枚数のみを変更して学習した時の出力の変化を図 11 に示した。ただし、入力枚数が 1 枚のときのみ学習済みモデルを利用していない。

入力枚数が 1 枚の時の結果が図 11 左から 1 番目の列に示されている。赤丸で囲った部分には白いノイズが入っており、ノイズ除去できていない。入力枚数を増やした場合の結果は図 11 左から 2 番目、3 番目の列に示されている。入力画像が 1 枚の時ノイズがあった部分を黄丸でそれぞれ示してあり、入力画像が 1 枚の時と比較してノイズが軽減されていることが分かる。4 枚と 7 枚の間では差があまりないが、複数枚入力によって 1 枚入力よりもノイズ除去能力が向上することが確認できる。表 1 の定量評価の結果を見ても、入力画像 1 枚より 4 枚、入力画像 4 枚より、7 枚入力であるベース結果のノイズ量が少ないことが確認できる。画像が 9 枚の際にはこれらの変化の中でノイズが最も少なく SSIM も良い結果だが、鮮明さは低くなっている。原因として、入力される情報が増えたがネットワークがうまく位置合わせできなかったために画像がボケてしまったと考えられる。

#### 4.6.4 2つのデータセットの比較

ノイズの多い画像を含むデータセット 1、ノイズが多い画像を取り除いたデータセット 2 でそれぞれ学習した結果を図 12 に示した。上段はデータセット 1 で学習した結果、



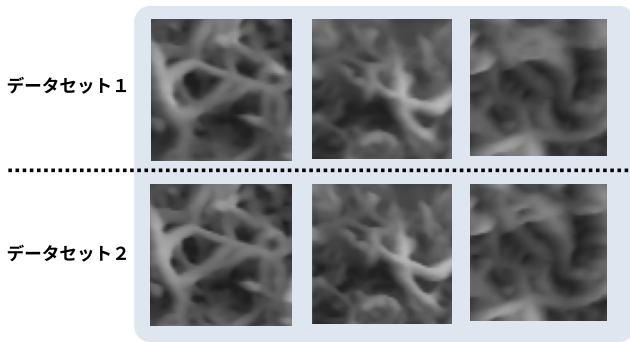


図 12 データセット 1・2 でそれぞれ学習し、ノイズ除去を行った例。

下段はデータセット 2 で学習した結果である。データセット 1 のノイズ除去はベースの結果と同じモデルである。比較しても定性的にはノイズの量・ボケには顕著な差は生じていない。表 1 の定量評価の結果を見てもデータセット 2 はベースとなる結果と比べてノイズ量はほぼ同じで、SSIM は改善しているが鮮明さは同等である。これらよりノイズが多いデータセットはあまり悪影響を与えておらず、質よりも量を重視してデータセットを収集することが効果的である可能性が考えられる。一方で、データセット 2 は 76 組しかないためデータ量が不十分であった可能性もあり、データセットを拡張してさらなる検証を行う必要がある。

## 5. まとめと今後の展望

本研究では、SPM 画像の撮影の手間を軽減するために SPM 画像のノイズ除去手法を考察した。従来のノイズ除去手法ではボケが発生したり、そもそも大きなノイズを除去することが難しかった。そこで、複数枚の画像を利用することで、大きなノイズでも除去できる手法を提案した。また、従来手法のように 1 枚の教師画像からノイズの入った入力画像を生成することが難しいため、実際の SPM 画像からなるデータセットを作成し、SPM 画像のノイズを定量化する手法を提案した。

RBPN を用いて提案手法によるノイズ除去ネットワークを実装し、作成したデータセットによって実験した結果、ノイズ除去を行うことができたが出力画像は入力画像に比べてボケが増加した。そこで、改善するために条件を変更してさまざまな実験を行い、結果を考察した。今後の研究では、実験した複数の条件を組み合わせるより正確かつ鮮明なノイズ除去を行う。また、画像のボケを減らすために位置合わせを工夫したり、教師画像でノイズがない部分のみ損失を計算するなどして損失関数を変更することが考えられる。

本研究は科研費（19K12129, 22H3618）の助成を受けたものである。

## 参考文献

- [1] Gerd Binnig and Heinrich Rohrer. Scanning tunneling microscopy—from birth to adolescence. *Rev. Mod. Phys.*, 59:615–625, 1987.
- [2] G. Binnig, C. F. Quate, and Ch. Gerber. Atomic force microscope. *Phys. Rev. Lett.*, 56:930–933, 1986.
- [3] Foundation for Promotion of Material Science and Technology of Japan. [afm] 原子間力顕微鏡法. <https://www.mst.or.jp/method/tabid/156/Default.aspx>, 2015. (Accessed on 4/1/2021).
- [4] Abdelrahman Abdelhamed et al. NTIRE 2019 challenge on real image denoising: Methods and results. In *CVPR Workshop*, 2019.
- [5] Andrey Ignatov et al. NTIRE 2019 challenge on image enhancement: Methods and results. In *CVPR Workshop*, 2019.
- [6] Seungjun Nah et al. NTIRE 2020 challenge on image and video deblurring. In *CVPR Workshop*, 2020.
- [7] Seungjun Nah, Sanghyun Son, Suyoung Lee, Radu Timofte, and Kyoung Mu Lee. NTIRE 2021 challenge on image deblurring. In *CVPR Workshop*, 2021.
- [8] Yifang Zhao, Suguru Koike, Rikuto Nakama, Shiro Ihara, Masatoshi Mitsuhashi, Mitsuhiro Murayama, Satoshi Hata, and Hikaru Saito. Five-second stem dislocation tomography for 300 nm thick specimen assisted by deep-learning-based noise filtering. *Scientific Reports*, 11:2045–2322, 2021.
- [9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *MICCAI*, 2015.
- [10] Yufan Luo and Sean B. Andersson. Image reconstruction for sub-sampled atomic force microscopy images using deep neural networks. *Micron*, 130:102814, 2020.
- [11] Ben Mildenhall, Jonathan T Barron, Jiawen Chen, Dillon Sharlet, Ren Ng, and Robert Carroll. Burst denoising with kernel prediction networks. In *CVPR*, 2018.
- [12] Semih Sevim, Selin Tolunay, and Hamdi Torun. Micromachined sample stages to reduce thermal drift in atomic force microscopy. *Microsystem Technologies*, 21:1559–1566.
- [13] Goutam Bhat et al. NTIRE 2021 challenge on burst super-resolution: Methods and results. In *CVPR Workshop*, 2021.
- [14] Dario Fuoli et al. AIM 2020 challenge on video extreme super-resolution: Methods and results. In Adrien Bartoli and Andrea Fusiello, editors, *ECCV Workshop*, 2020.
- [15] Seungjun Nah et al. NTIRE 2019 challenge on video deblurring: Methods and results. In *CVPR Workshop*, 2019.
- [16] Seungjun Nah et al. NTIRE 2019 challenge on video super-resolution: Methods and results. In *CVPR Workshop*, 2019.
- [17] Muhammad Haris, Greg Shakhnarovich, and Norimichi Ukita. Recurrent back-projection network for video super-resolution. In *CVPR*, 2019.
- [18] Muhammad Haris, Greg Shakhnarovich, and Norimichi Ukita. Deep back-projection networks for super-resolution. In *CVPR*, 2018.
- [19] Muhammad Haris, Greg Shakhnarovich, and Norimichi Ukita. Deep back-projectinetworks for single image super-resolution. *IEEE Trans. Pattern Anal. Mach. Intell.*, 43(12):4323–4337, 2021.
- [20] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. <https://arxiv.org/abs/1412.6980>, 2017. (Accessed on 5/1/2021).
- [21] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13:600–612, 2004.