AFFINE SUBSPACE MODELS AND CLUSTERING FOR PATCH-BASED IMAGE DENOISING

Tharindu Wickremasinghe
University of Moratuwa, Colombo, Sri Lanka
Purdue University, West Lafayette, IN, USA

Marco F. Duarte
University of Massachusetts Amherst, USA

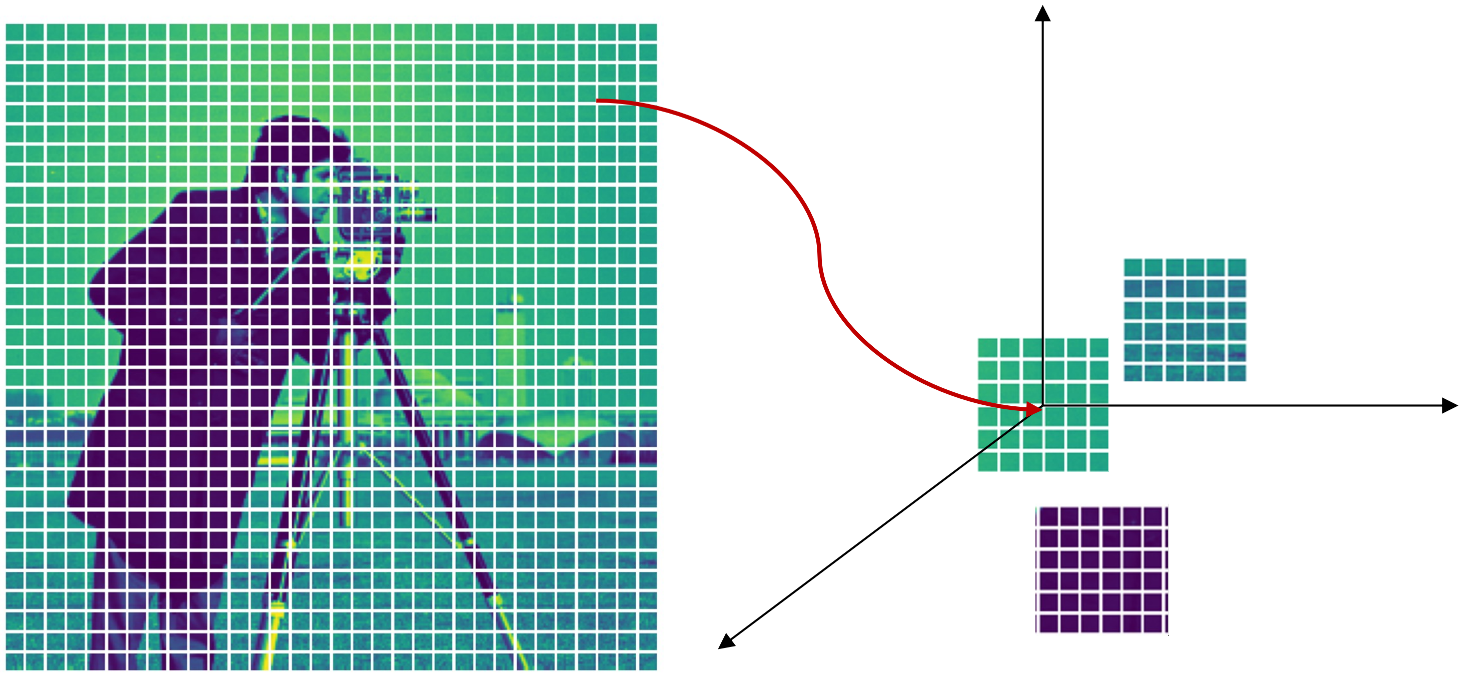
Denoising and Geometric Structure

$$Y = X + \eta$$

- Denoising is ubiquitous in many signal processing applications; image denoising has resurged as a powerful tool in image modeling (e.g., regularization, diffusion-based synthesis)
- Our focus is patch-based denoising methods: look beyond the immediate pixel neighborhood for denoising information
- identify similar image patches (clusters), and compute a weighted average to denoise, e.g., Non-Local Means (NLM)
- It turns out that the set of image patches is being modeled as a **union of subspaces**

A. Buades, B. Coll and J.-M. Morel, "A non-local algorithm for image denoising," IEEE Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 2005

Linear Subspaces and Non-Local Means (NLM)



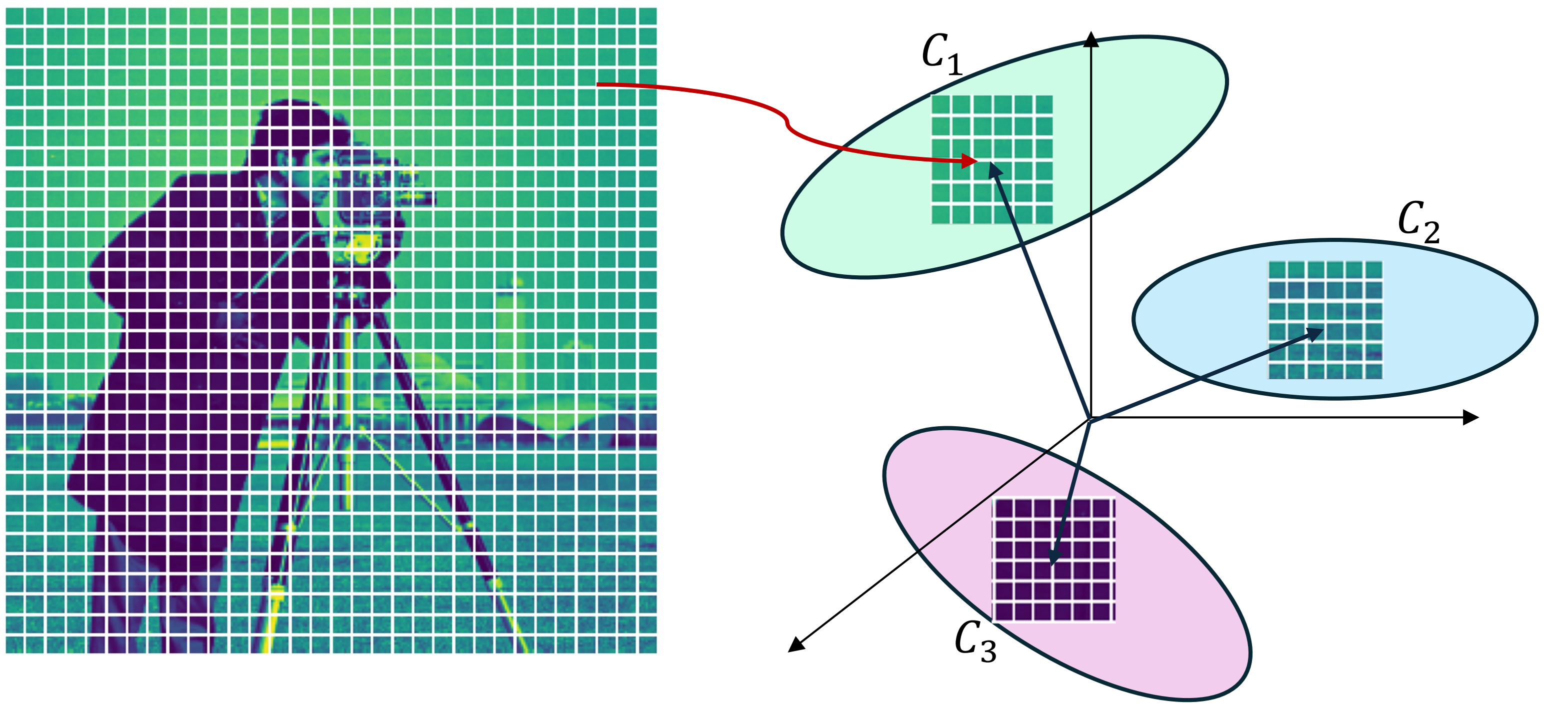
- A weighted average considers all patches x_i:

$$\hat{x} = \sum_i w_i x_i$$

- The weight w_i is a measure of the similarity of other patches to the current patch
- This approach for patch estimation assumes that similar patches (cluster) are contained in a single subspace
- NLM thus models the patches of an image as a **Union of Subspaces**

A. Buades, B. Coll and J.-M. Morel, "A non-local algorithm for image denoising," IEEE Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 2005

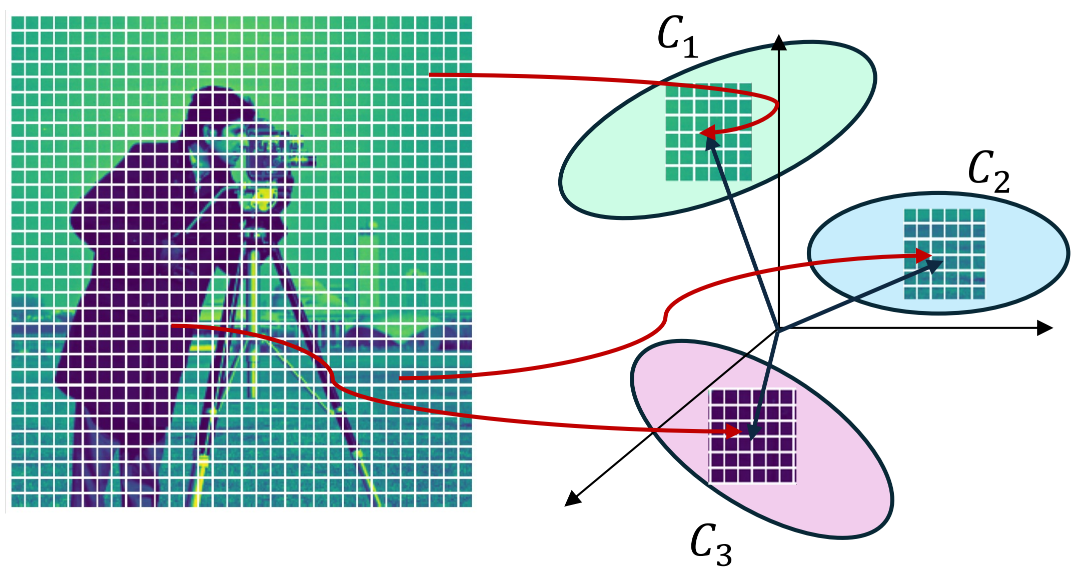
The Benefits of Affine Subspaces for Patch Models



$$\hat{x} = \sum_i w_i x_i$$

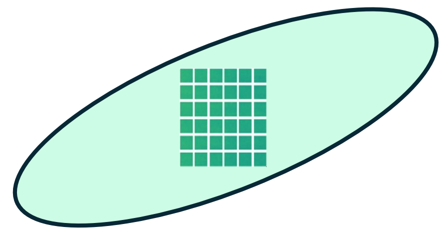
- Images are non-negative
- Each subspace has a non-zero mean
- This structure is better captured by an affine subspace model: $\sum_i w_i = 1$
- We review several algorithmic approaches to solve the noisy affine subspace clustering problem
- We also present a simple denoising algorithm that relies on the affine subspace clustering model using least squares projection

Methodology



① **Subspace Clustering:**

1. Basis Pursuit Denoising (BPDN)
2. Non-Negative Constrained Lasso (NNCL)
3. Non-Negative Lasso (NNL)



② **Cluster Encoding:**

1. Mean vector (affine)
2. PCA components (subspace)

③ **Cluster Denoising:**

ML estimate assuming AWGN:

Patch (Affine) Subspace Projection (PASP)

Robust Subspace Clustering (RSSC)

Algorithm 1: Robust Subspace Clustering Procedure
Input: A data set arranged as columns of a matrix $X \in \mathbb{R}^{n \times N}$

1. For each $i \in \{1, \dots, N\}$, produce a sparse coefficient sequence $\{\hat{w}_i\}$ by regressing the i^{th} column vector x_i onto the other columns of X. Collect these as columns of a matrix B.

2. Form the similarity graph G with nodes representing the N data points and edge weights given by
$$W_{ij} = |B_{ij}| + |B_{ji}|$$

3. Sort the eigenvalues $\delta_1 \geq \delta_1 \geq \cdots \geq \delta_N$ of the normalized Laplacian of G In descending order, and set
$$\hat{L} = \arg\max(\delta_i - \delta_{i+1}),\, i = 1, \dots, N-1$$

4. Apply a spectral clustering technique to the similarity graph using $\hat{L}$ as the estimated number of clusters to obtain the partition $X_1, \dots, X_L$

M. Soltanolkotabi, E. Elhamifar, and E. J. Candès, “Robust subspace clustering,” The Annals of Statistics, vol. 42, no. 2, pp. 669 – 699, 2014

Robust Affine Subspace Clustering (RASSC)

Algorithm 2: Robust Affine Subspace Clustering Procedure

Input: A data set arranged as columns of a matrix $X \in \mathbb{R}^{n \times N}$

1. For each $i \in \{1, \dots, N\}$, produce a sparse coefficient sequence $\{\hat{w}_i\}$ by regressing the i^{th} column vector x_i onto the other columns of X. Collect these as columns of a matrix B.

Set sequences to enforce affineness

i) Basis Pursuit Denoising (BPDN):

$$\hat{w}_i = \arg\min \|w\|_1 \; s.t. \; \|Xw - x_i\|_2 \le \sigma$$

ii) Non-Negative Constrained Lasso (NNCL):

$$\hat{w}_i = \arg\min \|Xw - x_i\|_2 \; s.t. \; \|w\|_1 \le \tau, w \ge 0$$

iii) Non-Negative Lasso (NNL):

$$\hat{w}_i = \arg\min \|Xw - w_i\|_2 + \alpha \|w\|_1 \; s.t. \; w \ge 0$$

Other names for (ii-iii): positive lasso, non-negative garrotte, …

M. Yuan and Y. Lin. "On the non-negative garrotte estimator." J. Royal Stat. Soc. B: Statistical Methodology 69.2 (2007): 143-161.

B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least Angle Regression," Annals of Statistics, 32 (2004): 407–499.

A Comparison of Regression Approaches

i) Basis Pursuit Denoising (BPDN):

$$\hat{w}_i = \arg\min \|w\|_1 \; s.t. \|Xw - x_i\|_2 \leq \sigma$$

ii) Non-Negative Constrained Lasso (NNCL):

$$\hat{w}_i = \arg\min \|Xw - x_i\|_2 \; s.t. \|w\|_1 \leq \tau, w \geq 0$$

iii) Non-Negative Lasso (NNL):

$$\hat{w}_i = \arg\min \|Xw - w_i\|_2 + \alpha\|w\|_1 \; s.t. \; w \geq 0$$

	BPDN	NNCL	NNL
How is affineness ($\sum w_i = 1$) enforced?	Not enforced	Directly: Achieve $\sum\lvert w_i\rvert = \sum w_i = 1$ at converging solution by setting $\tau = 1$	Indirectly: try to tune α so that $\sum\lvert w_i\rvert = \sum w_i = 1$; we still enforce $w_i \geq 0$
Do we achieve $\sum w_i = 1$? Need to tune parameters for it?	Not achieving $\sum\lvert w_i\rvert = 1$ Will find β such that $\|Xw - x_i\|_2 \sim \sigma$	$\sum w_i = 1$ for **most** patches. $\sum\lvert w_i\rvert = 1$ for **all** patches	Very hard to tune α with noisy images. (Need to change parameters with image and noise level) Not achieving $\sum\lvert w_i\rvert = 1$ for most patches

Robust Affine Subspace Clustering (RASSC)

Algorithm 2: Robust Affine Subspace Clustering Procedure
Input: A data set arranged as columns of a matrix $X \in \mathbb{R}^{n \times N}$

1. For each $i \in \{1, \dots, N\}$, produce a sparse coefficient sequence $\{\widehat{w}_i\}$ by regressing the i^{th} column vector x_i onto the other columns of X with $\sum \widehat{w}_i = 1$. Collect these as columns of a matrix B.

NNCL is better to enforce affineness

2. Form the similarity graph G with nodes representing the N data points and edge weights given by
$$W_{ij} = |B_{ij}| + |B_{ji}|$$

3. Sort the eigenvalues $\delta_1 \geq \delta_1 \geq \cdots \geq \delta_N$ of the normalized Laplacian of G In descending order, and set
$$\hat{L} = \arg\max(\delta_i - \delta_{i+1}), i = 1, \dots, N - 1$$

4. Apply a spectral clustering technique to the similarity graph using $\hat{L}$ as the estimated number of clusters to obtain the partition $X_1, \dots, X_L$

Robust Affine Subspace Clustering (RASSC)

Algorithm 2: Robust Affine Subspace Clustering Procedure
Input: A data set arranged as columns of a matrix $X \in \mathbb{R}^{n \times N}$

1. For each $i \in \{1, \ldots, N\}$, produce a sparse coefficient sequence $\{\widehat{w}_i\}$ by regressing the i^{th} column vector x_i onto the other columns of X with $\sum \widehat{w}_i = 1$. Collect these as columns of a matrix B.

NNCL is better to enforce affineness

2. Form the similarity graph G with nodes representing the N data points and edge weights given by
$$W_{ij} = |B_{ij}| + |B_{ji}|$$

3. Sort the eigenvalues $\delta_1 \geq \delta_1 \geq \cdots \geq \delta_N$ of the normalized Laplacian of G In descending order, and set
$$\hat{L} = \arg\max(\delta_i - \delta_{i+1}), i = 1, \ldots, N-1$$

4. Apply a spectral clustering technique to the similarity graph using $\hat{L}$ as the estimated number of clusters to obtain the partition $X_1, \ldots, X_L$

Robust Affine Subspace Clustering (RASSC)

Algorithm 2: Robust Affine Subspace Clustering Procedure
Input: A data set arranged as columns of a matrix $X \in \mathbb{R}^{n \times N}$

1. For each $i \in \{1, \dots, N\}$, produce a sparse coefficient sequence $\{\widehat{w}_i\}$ by regressing the i^{th} column vector x_i onto the other columns of X with $\sum \widehat{w}_i = 1$. Collect these as columns of a matrix B.

NNCL is better to enforce affineness

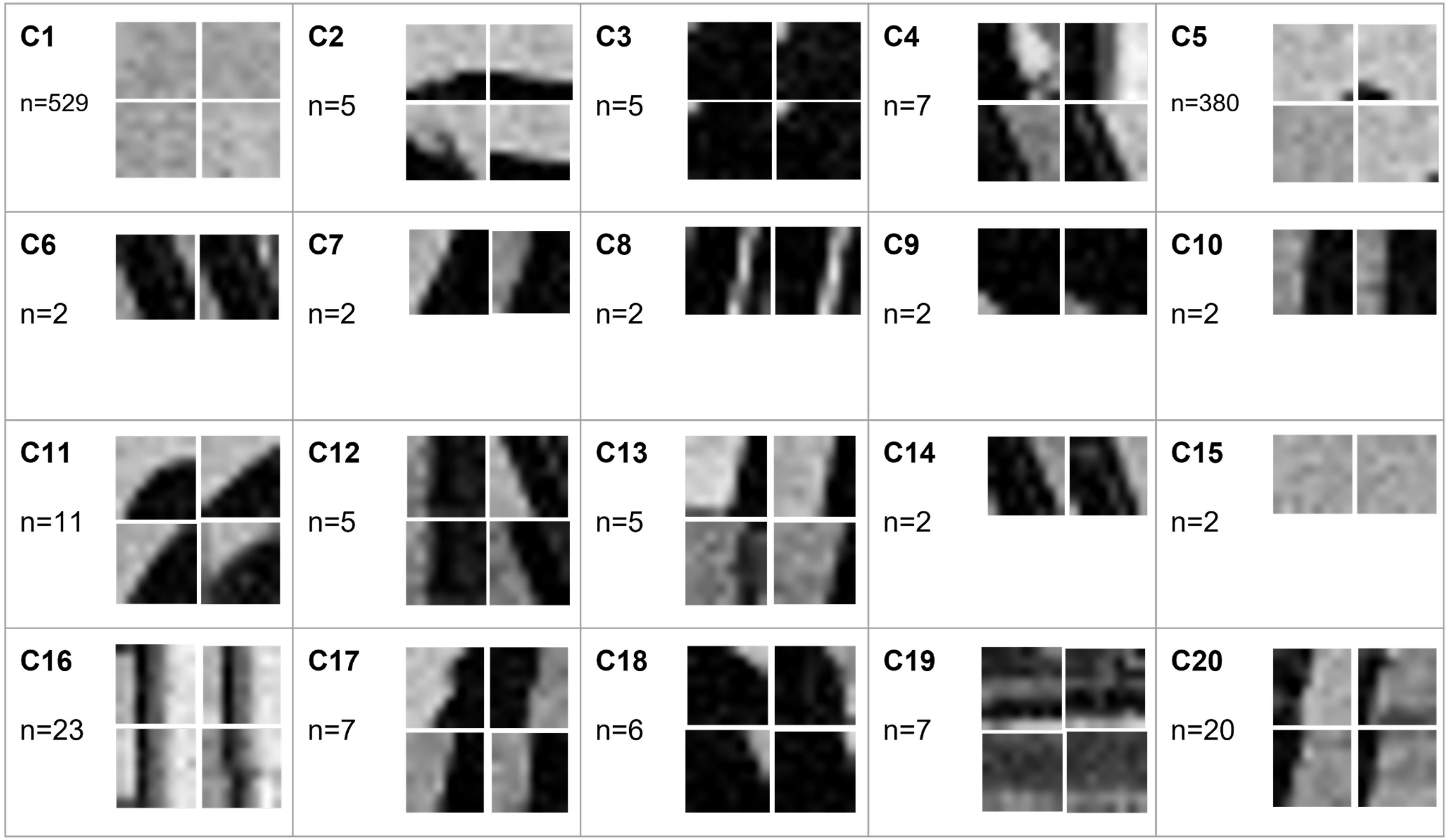
2. Form the similarity graph G with nodes representing the N data points and edge weights given by
$$W_{ij} = |B_{ij}| + |B_{ji}|$$

3. Sort the eigenvalues $\delta_1 \geq \delta_1 \geq \cdots \geq \delta_N$ of the normalized Laplacian of G In descending order, and set
$$\hat{L} = \arg\max(\delta_i - \delta_{i+1}),\, i = 1, \dots, N-1$$

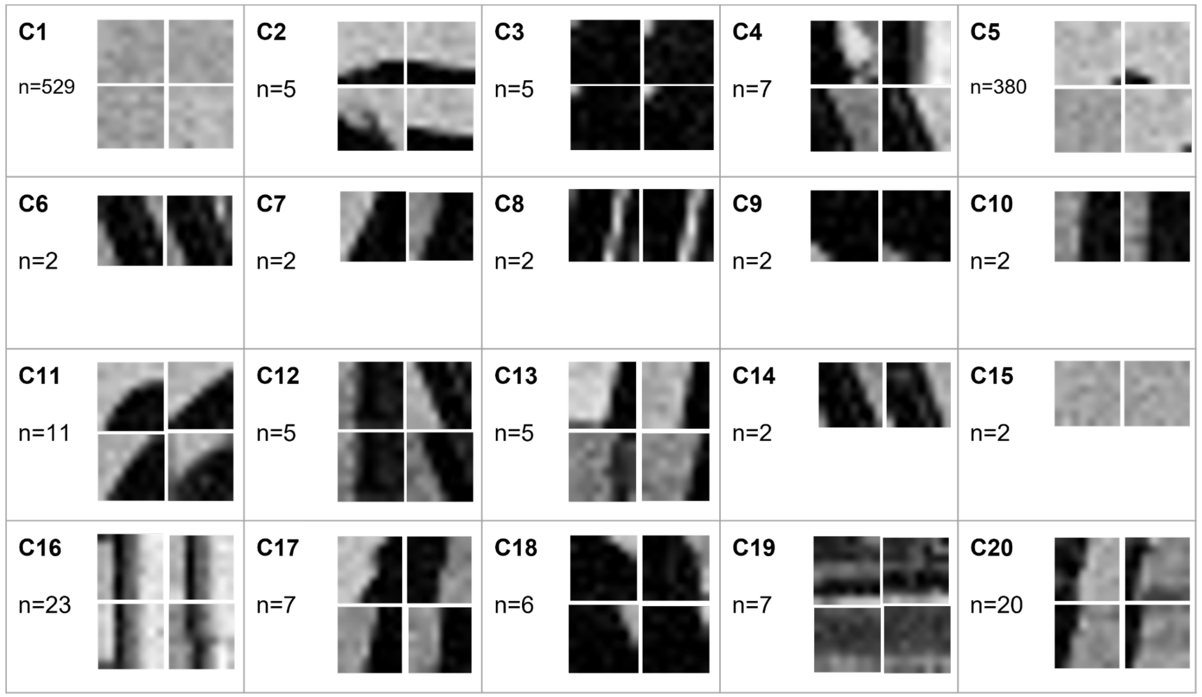
A fixed number of clusters can be a hyperparameter

4. Apply a spectral clustering technique to the similarity graph using $\hat{L}$ as the estimated r
to obtain the partition $X_1, \dots, X_L$

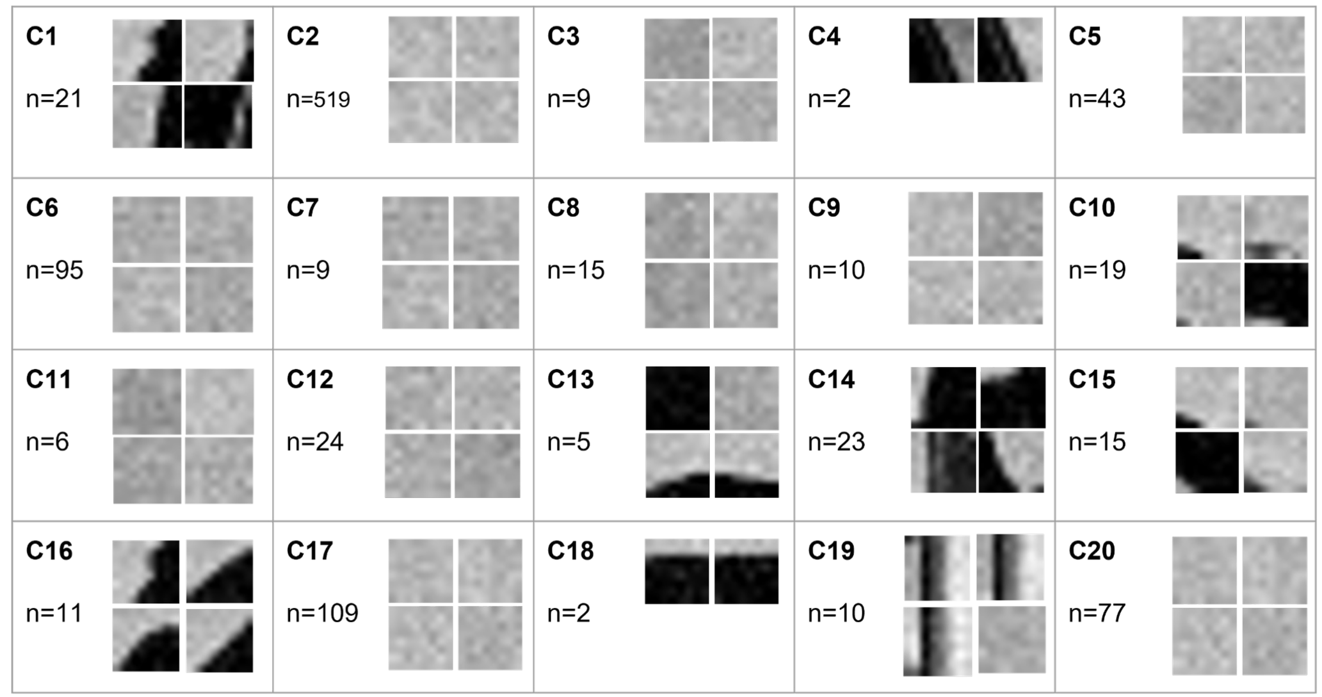
Robust Affine Subspace Clustering (RASSC) with NNCL



Robust Affine Subspace Clustering (RASSC) with NNCL

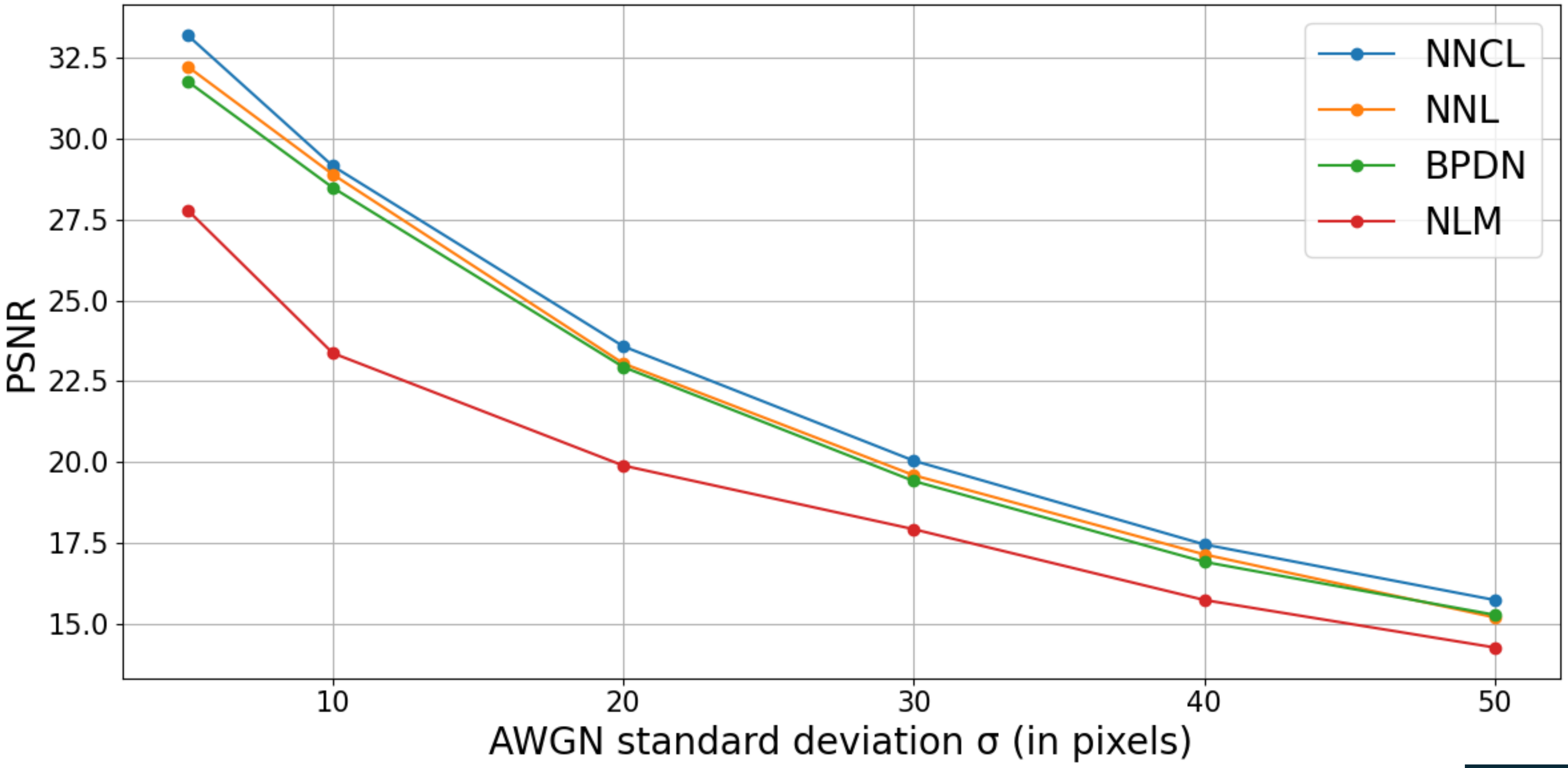


Robust Subspace Clustering (RSSSC) with BPDN



RASSC vs. Baseline: Non-Local Means (NLM)

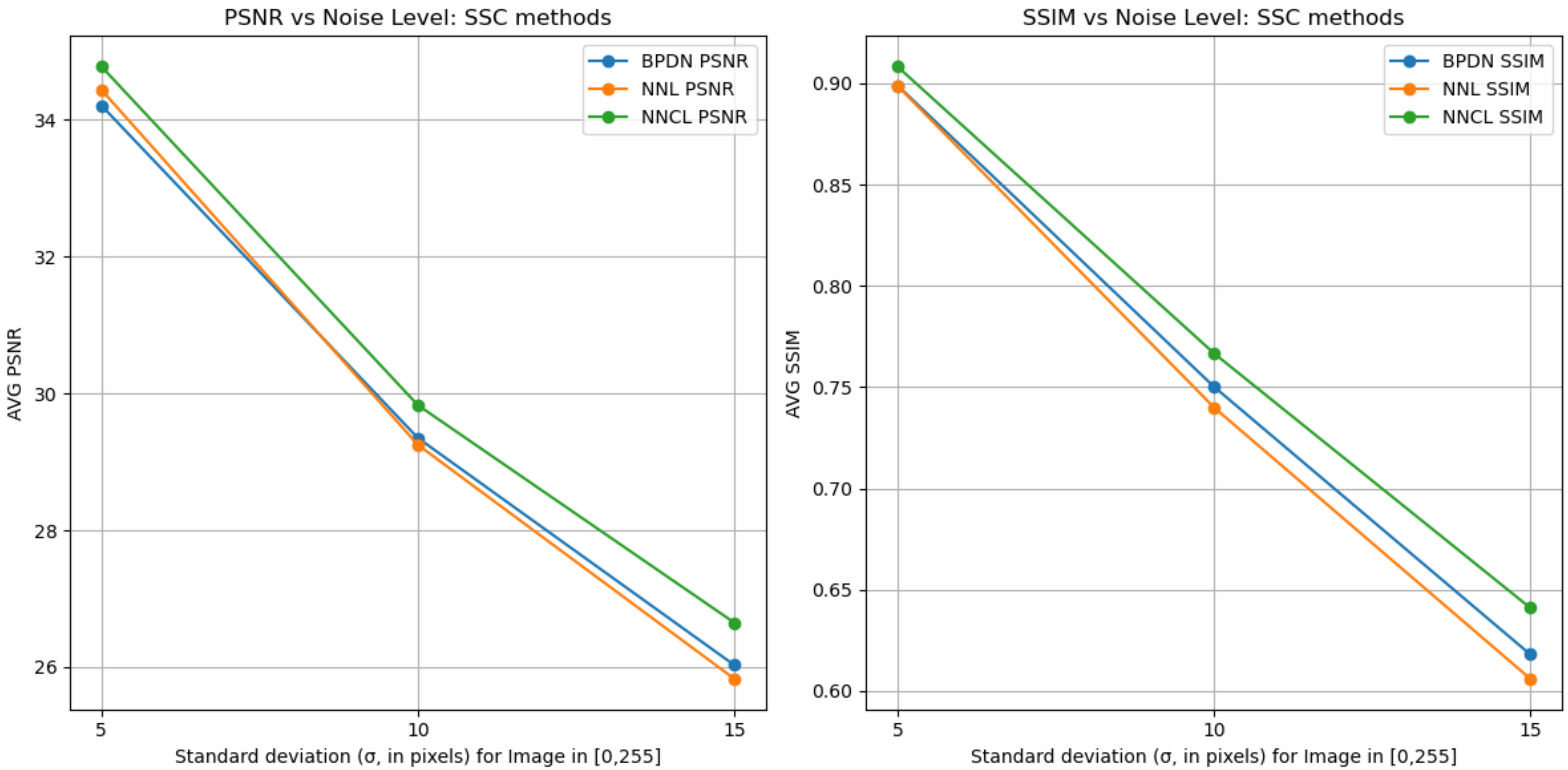
Denoising of clusters with Patch Affine Subspace Projection (PASP) in each cluster, compared to Denoising with NLM from scikit-learn package



Robust Affine SSC (NNCL/NNL) is better as expected

Image Data: https://github.com/antimattercorrade/Image_Denoising

RASSC works best when affine constraints are enforced



NNC LASSO is better as expected

Image Data: https://github.com/antimattercorrade/Image_Denoising

Conclusions and Future Work

1. In patch-based image processing literature, affine models should be explored

2. Experimental evidence matches with our intuition:
   there are benefits for modelling the set of image patches as union of affine subspaces

3. Applying our model as an augmentation to existing patch-based denoising methods
   improves performance in the downstream task of denoising

We thank the IEEE SPS and the ME-UYR program (now SigMA)
for sponsoring this research collaboration

IEEE Signal Processing Society