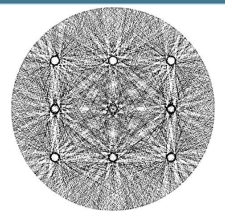
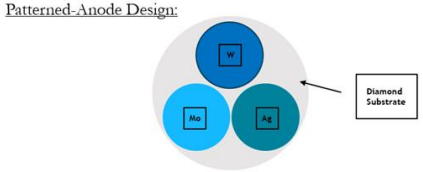
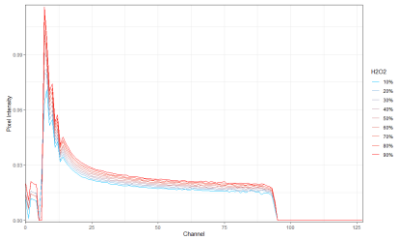
SAND2021-12800C

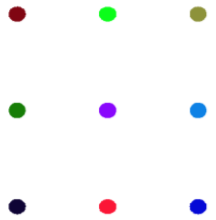
Sandia National Laboratories

A Process to Colorize and Assess Visualizations of Noisy X-Ray Computed Tomography Hyperspectral Data of Materials with Similar Spectral Signatures









Joshua Clifford, Emily Kemp, Ben Limpanukorn, Edward S. Jimenez

2021 IEEE Nuclear Science Symposium and Medical Imaging Conference

Joint NSS-MIC Session 2, October 19

U.S. DEPARTMENT OF ENERGY | National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Background

Past Research

- Denoising hyperspectral data
- Material classification with hyperspectral data
- Coloring hyperspectral data using dimension reduction techniques

Objectives

- Represent hyperspectral computed tomography (HCT) data in a single colorized image
- Maximize contrast between distinct materials with similar compositions (different mixture concentrations of similar materials)
- Maximize smoothness within homogeneous materials

Potential Impact

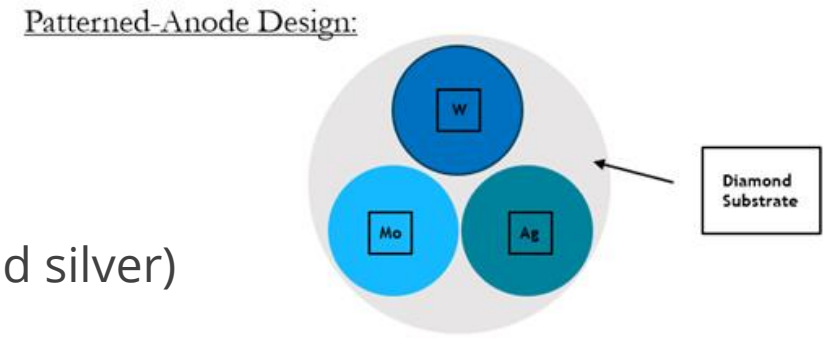
- Rapid human interpretability of complex HCT signatures
- More robust material identification methods
- Industrial, medical, and security-based applications

Methods

Data

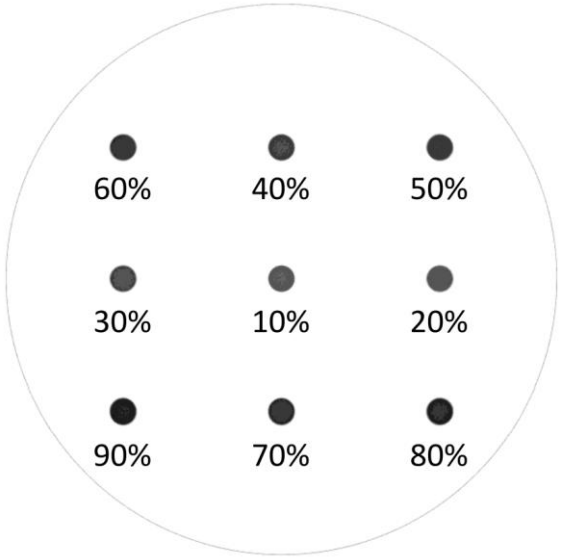
HCT System

- PHITS simulations
- Patterned anode (tungsten, molybdenum, and silver)
- 225kVp electron beam impinging on anode
- Detector channelizing photons into 128 energy channels across 300 keV



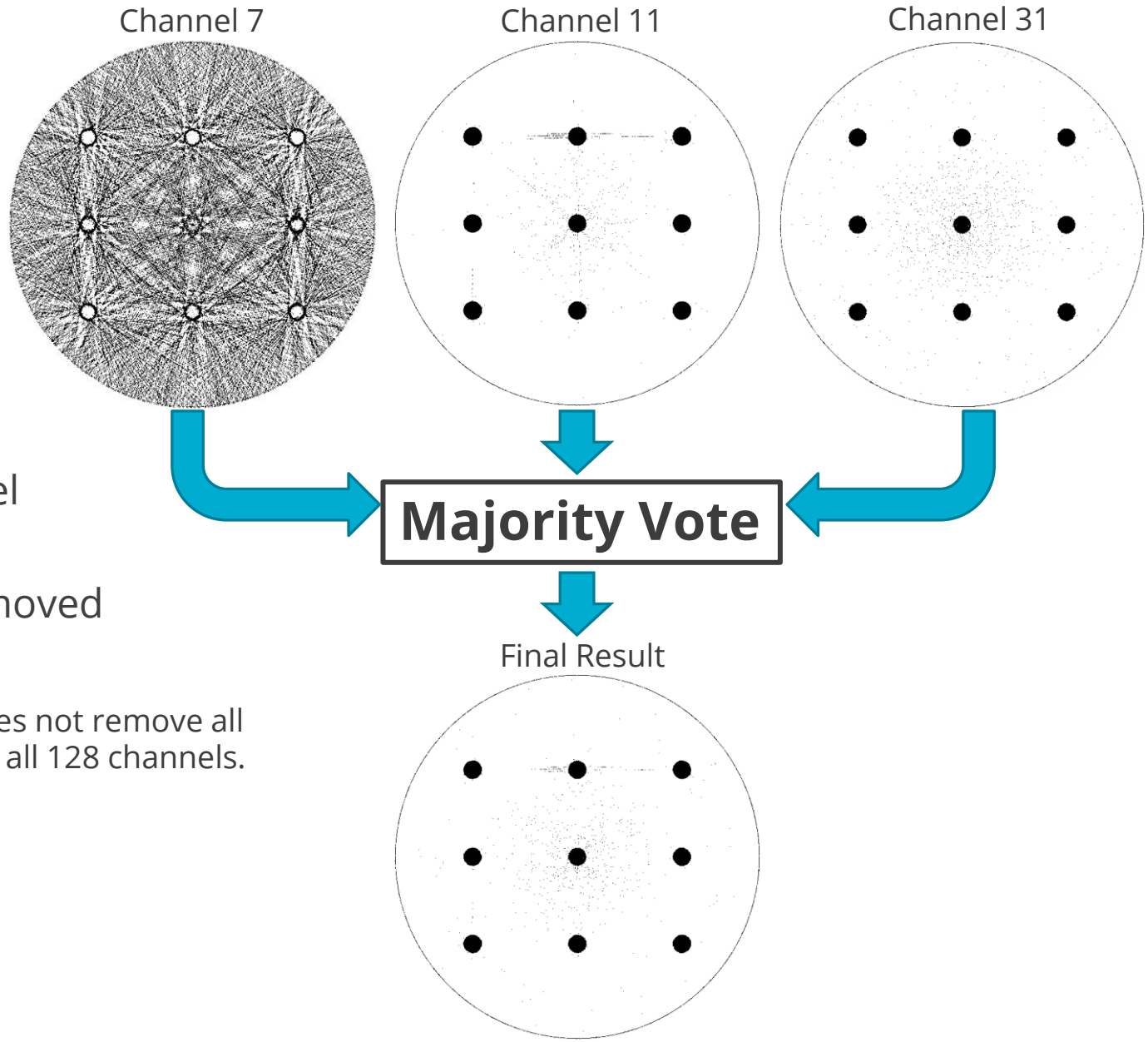
Objects

- Cylindrical
- Mixtures of H_2O and H_2O_2



Reconstructed image of cylinders labeled with their concentrations of H_2O_2 to H_2O

Preprocessing Methods



Novel “majority vote” thresholding

- Multilevel thresholding on each channel
- Each channel result counts as a vote
- Each pixel with < 50% of the vote is removed

Note: In the small 3 channel example to the right, it does not remove all pixels not in objects of interest, but it works well using all 128 channels.

Preprocessing Methods

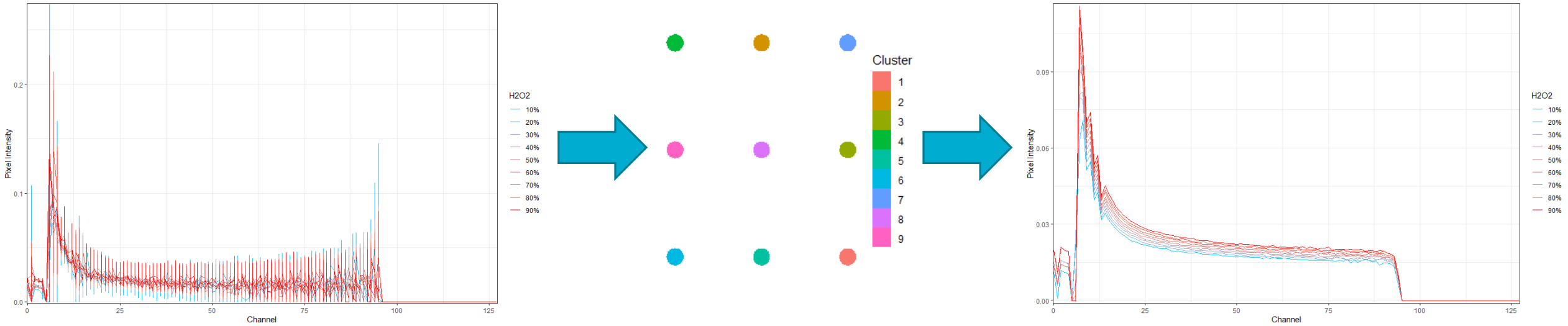
Binary erosion (remove additional remaining artifacts)

Channel filtering (removing uninformative low-energy channels)

Spatial smoothing (box, Gaussian, or median filter)

Clustering (grouping objects by location/channel intensities)
+ object value assignment (channel mean/median or smoothing splines over channels)

Example of Clustering + Object Value Assignment



Colorization Methods

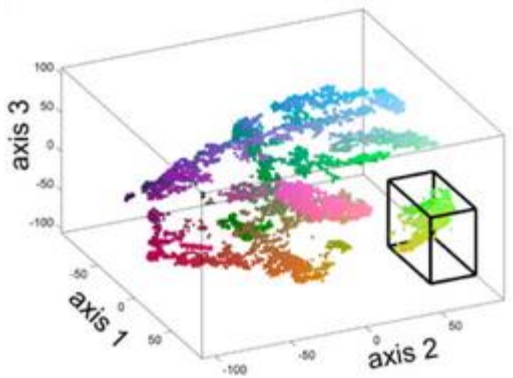
Dimension Reduction

- Uniform Manifold Approximation and Projection (UMAP)
- t-distributed Stochastic Neighbor Embedding (t-SNE)
- Others tested but not presented, did not perform as well
- Represent 128 channels in 3 dimensions for RGB coloring

Example UMAP 2D representation of MNIST digit images from McInnes, Healy, & Melville [12]

Linear Models

- Logistic Regression
  - Predicting concentration from channel intensities
- Quadratic Regression
  - Using model parameters as RGB color dimensions



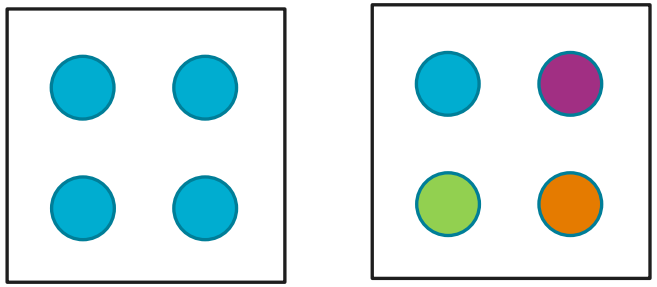
Example t-SNE to RGB translation from Fonville et al. [6]

Colorization Assessment Metrics

After the final colorization is produced, there are several visual attributes that we wished to measure. The metrics to assess these attributes were based on the RGB additive color model.

Contrast

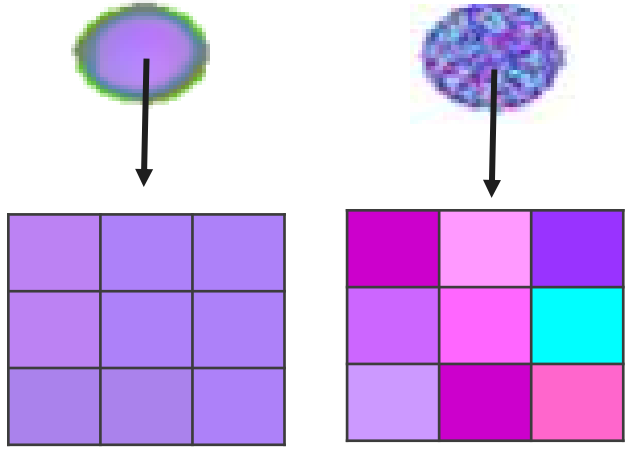
- Color differences between each pair of objects.
- *Think*: The visual/color contrast within an image.



Low Contrast *High Contrast*

Local Color Variance

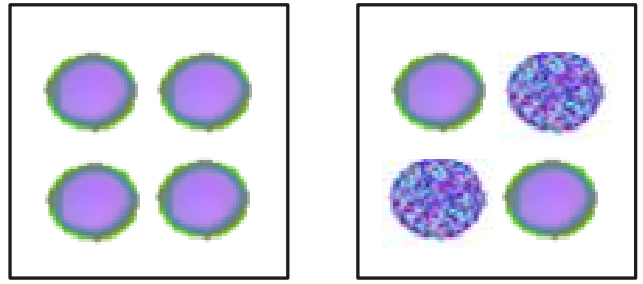
- Color differences between neighboring pixels within objects.
- *Think*: The amount of speckle within an object.



Low LCV *High LCV*

Inconsistency

- Coloring differences within objects.
- *Think*: How similar all the objects are in an image with respect to the smoothness or speckle.



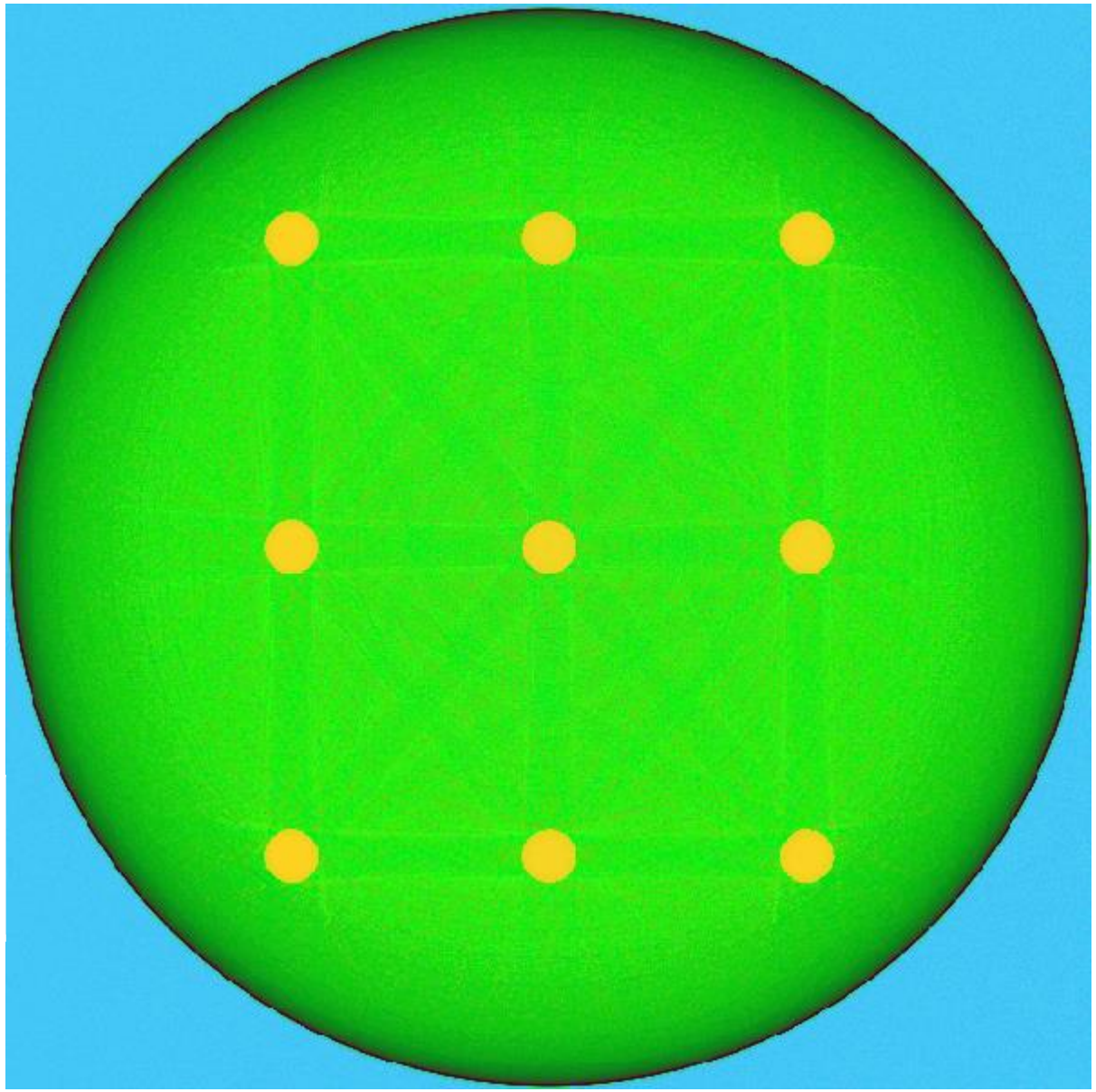
Low Inconsistency *High Inconsistency*

Results

Before Majority-Voting Thresholding + Erosion

Before any preprocessing:

- Many artifacts due to system noise
- No contrast between objects of interest



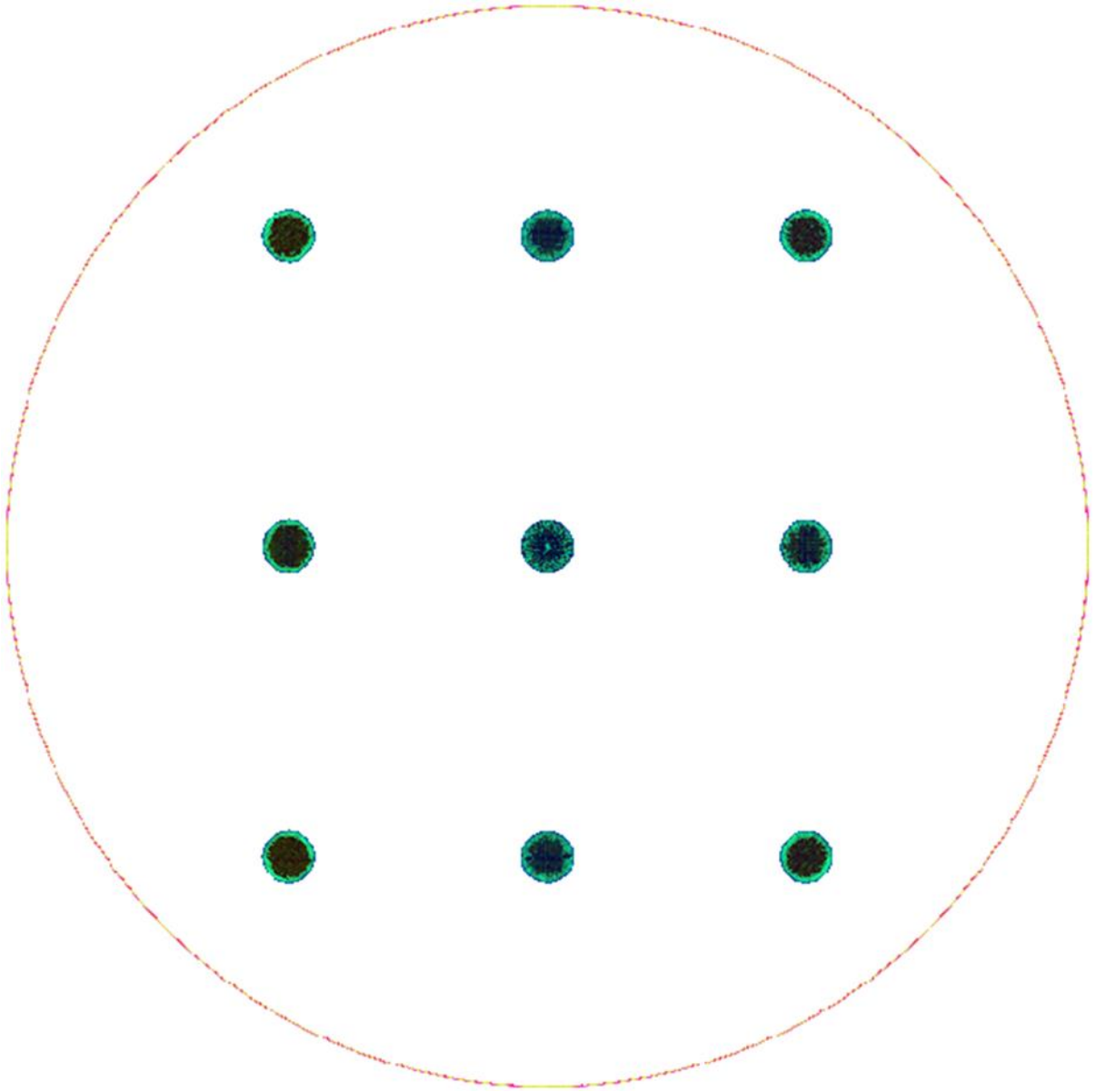
Concentration Key		
60%	40%	50%
30%	10%	20%
90%	70%	80%

Metric Results (only for objects)
Contrast = 0.0082
Local Color Var = 0.0003
Inconsistency = 0.0000

After Majority-Voting Thresholding + Before Erosion

After thresholding:

- Most non-object pixels removed
- Objects have large amounts of local color variance and fairly low contrast
- Remaining non-object pixels are scan boundary artifact due to reconstruction



Concentration Key		
60%	40%	50%
30%	10%	20%
90%	70%	80%

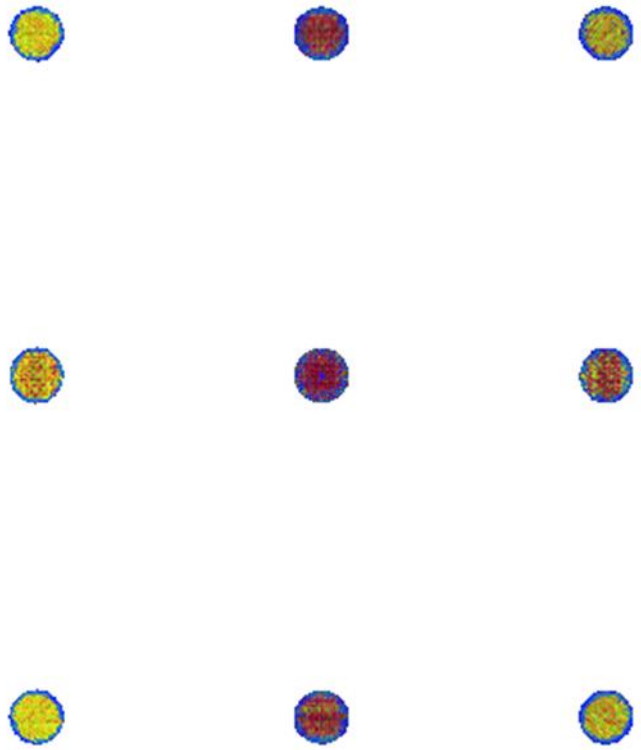
Metric Results (only for objects)
Contrast = 0.1376
Local Color Var = 0.0815
Inconsistency = 0.0001

After Majority-Voting Thresholding + Erosion

Concentration Key		
60%	40%	50%
30%	10%	20%
90%	70%	80%

After thresholding + erosion:

- All non-object pixels removed
- Slightly more contrast between objects and much lower local color variance
- Issue with edge artifact in objects, which was localized to a few low-energy channels



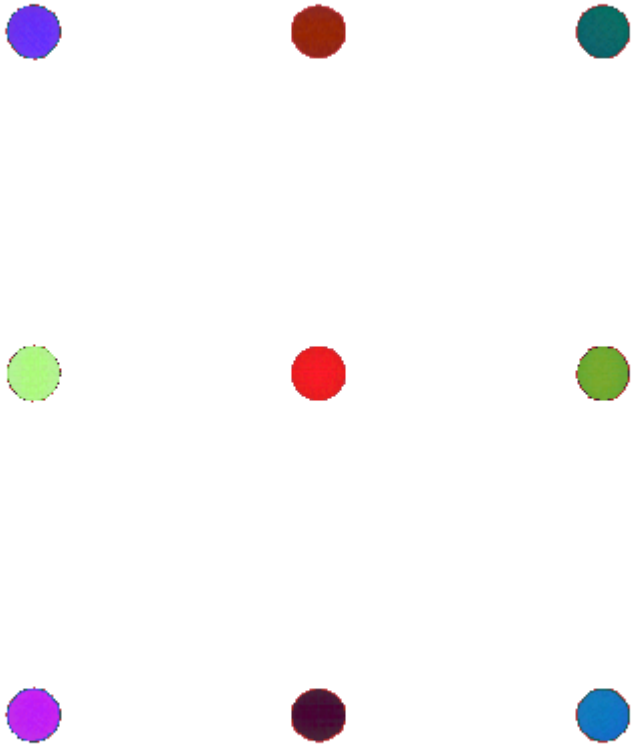
Metric Results
Contrast = 0.1460
Local Color Var = 0.0222
Inconsistency = 0.0001

Edge Artifact Removal

Concentration Key		
60%	40%	50%
30%	10%	20%
90%	70%	80%

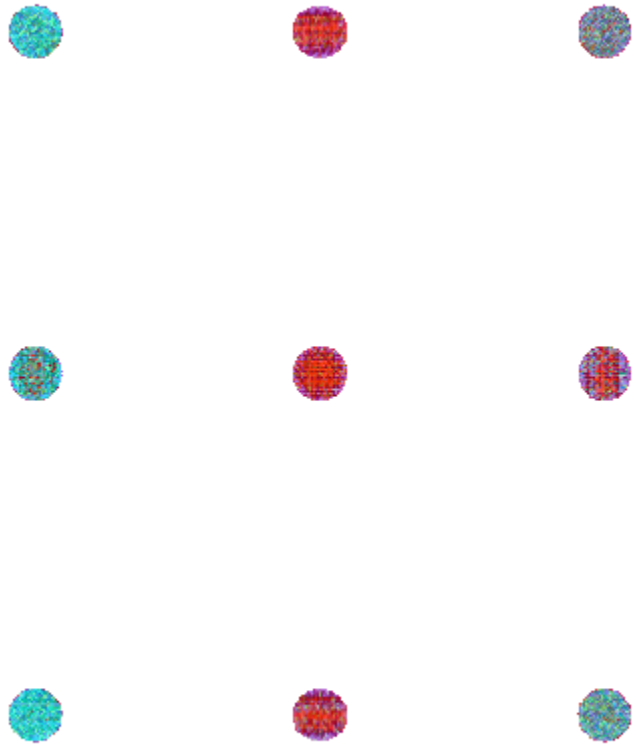
Coloring after removing channels driving object edge artifacts…

Very effective for reference data (high photon count)



Metric Results
Contrast = 0.3961
Local Color Var = 0.0215
Inconsistency = 0.0009

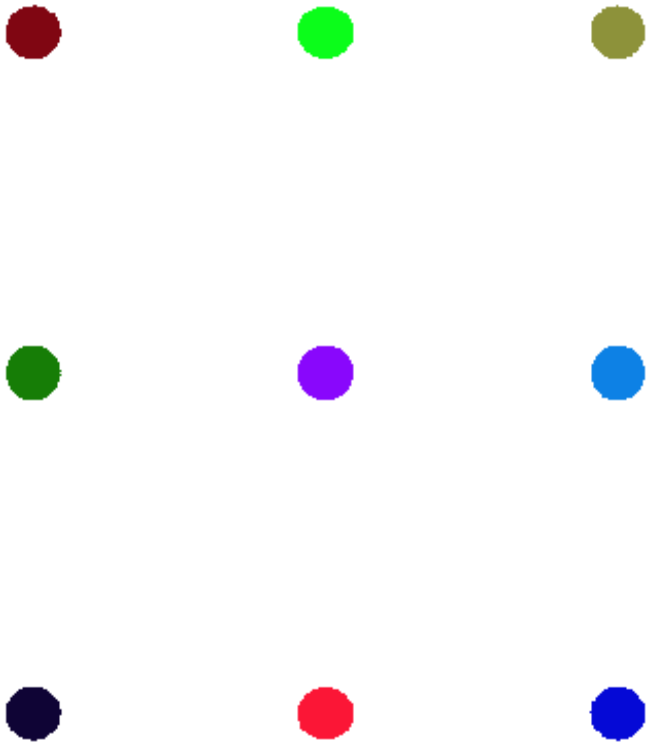
Not so effective for sample data



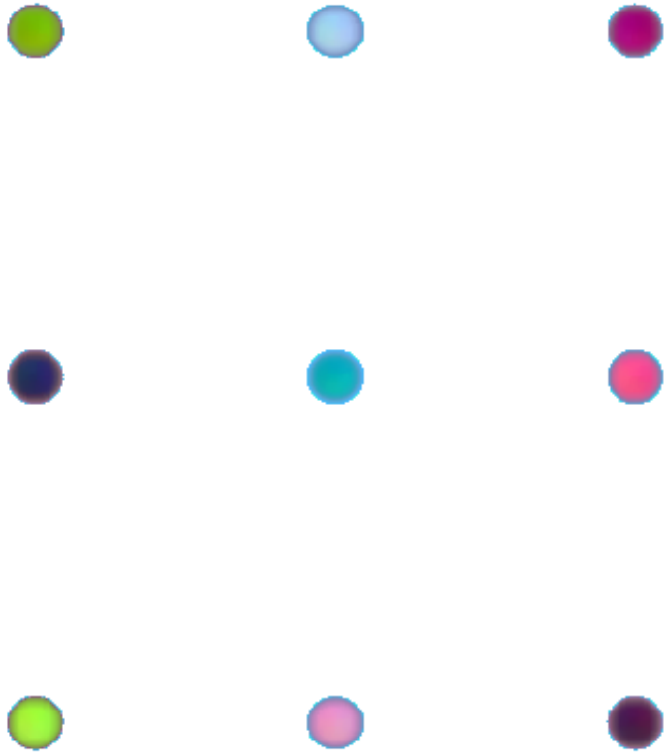
Metric Results
Contrast = 0.2378
Local Color Var = 0.0728
Inconsistency = 0.0007

Median Assignment vs. Median Filtering

Concentration Key		
60%	40%	50%
30%	10%	20%
90%	70%	80%



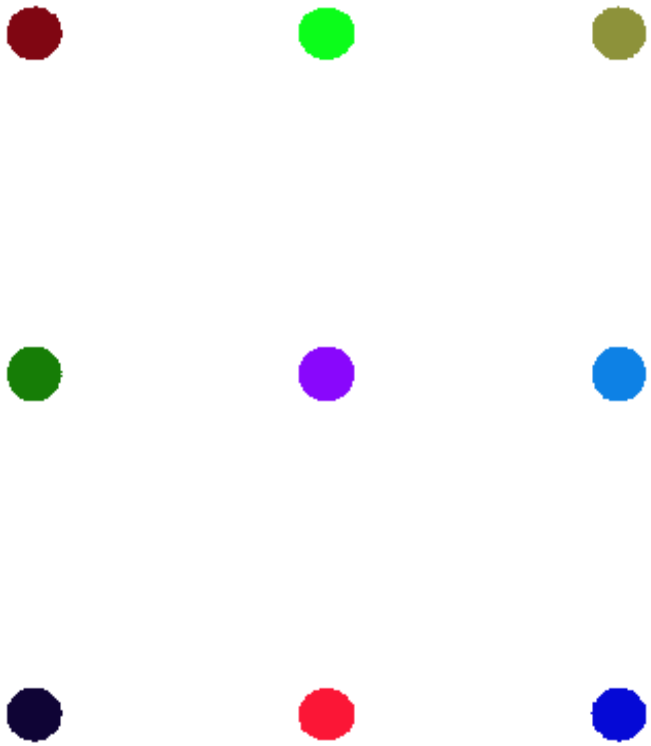
Metric Results
Contrast = 0.4908
Local Color Var = 0.0000
Inconsistency = 0.0000



Metric Results
Contrast = 0.2957
Local Color Var = 0.0118
Inconsistency = 0.0025

Median assignment requires object segmentation, but produces high contrast.
Median filtering provides a useful alternative with simple pixel-level processing.

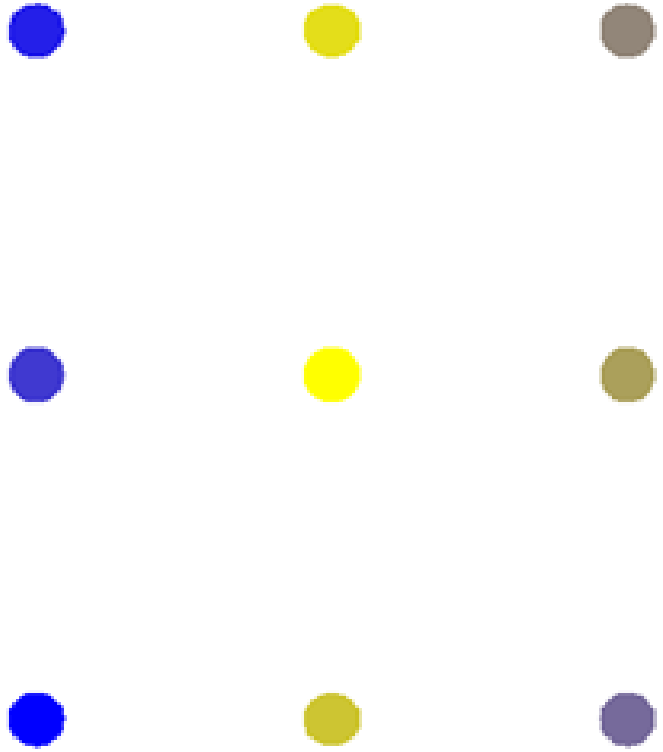
UMAP vs. t-SNE



Concentration Key

60%	40%	50%
30%	10%	20%
90%	70%	80%

Metric Results
Contrast = 0.4908
Local Color Var = 0.0000
Inconsistency = 0.0000



Metric Results
Contrast = 0.4276
Local Color Var = 0.0000
Inconsistency = 0.0000

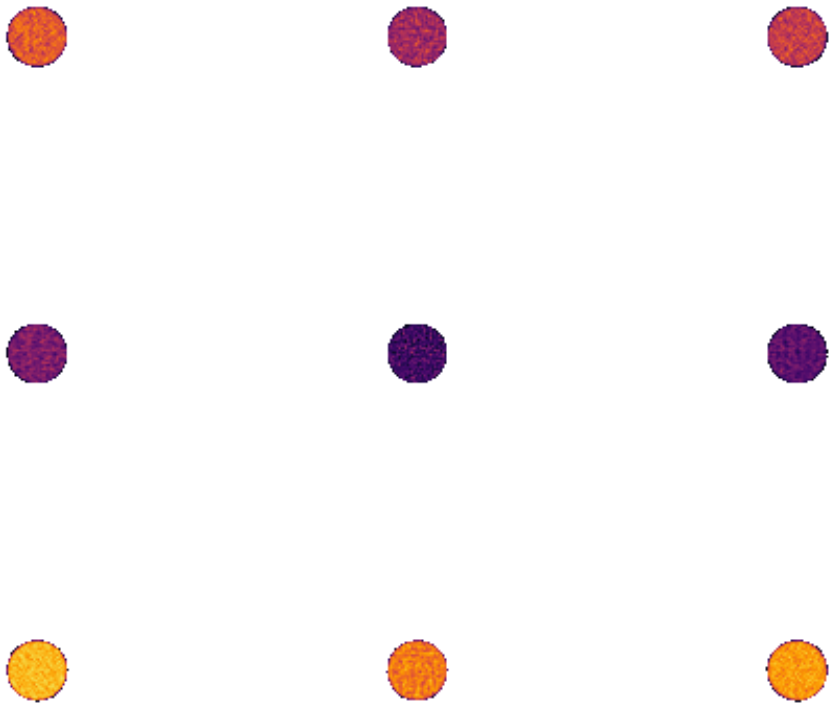
UMAP provides more separation between objects, resulting in better contrast.
However, both methods do not allow easy interpretation of concentrations.

Linear Model Colorization

Concentration Key		
60%	40%	50%
30%	10%	20%
90%	70%	80%

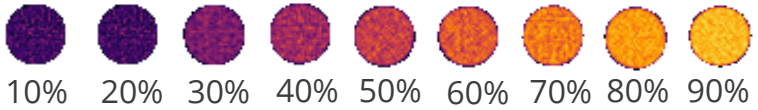
Using linear models retains good contrast, and increases interpretability by coloring concentrations on a meaningful gradient.

Logistic Regression
(on pixels)



Metric Results
Contrast = 0.2651
Local Color Var = 0.0444
Inconsistency = 0.0004

Objects arranged in ascending order by concentration

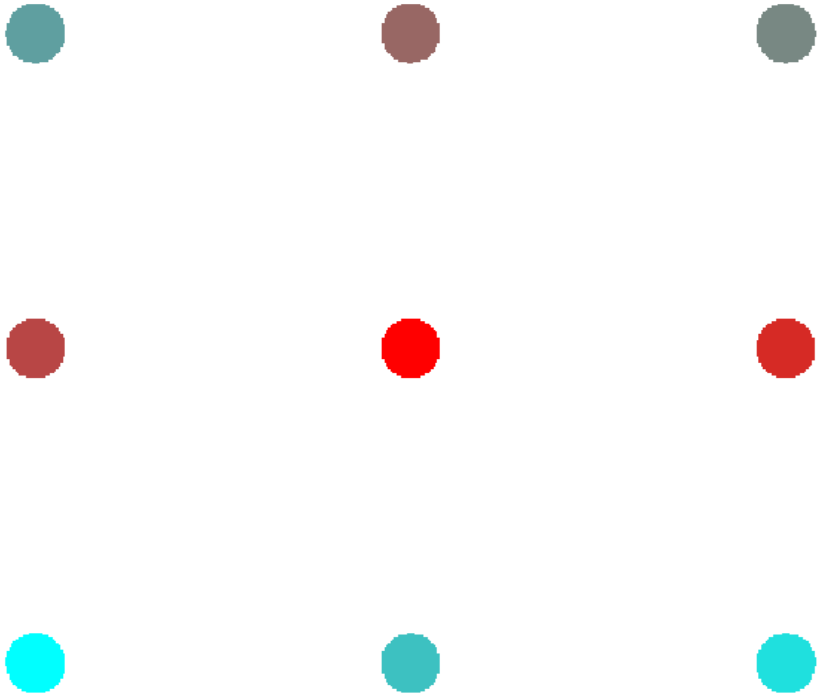


Linear Model Colorization

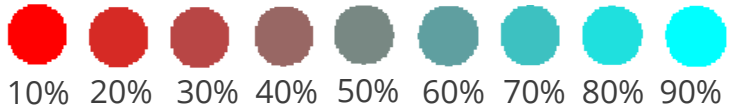
Concentration Key		
60%	40%	50%
30%	10%	20%
90%	70%	80%

Using linear models retains good contrast, and increases interpretability by coloring concentrations on a meaningful gradient.

Quadratic Regression
(median assignment, normalized parameters)



Objects arranged in ascending order by concentration



Metric Results
Contrast = 0.4083
Local Color Var = 0.0000
Inconsistency = 0.0000

Method Comparison

	Preprocessing	None	MLT	• MLT • Erosion	• MLT • Erosion • Channel Removal	• MLT • Erosion • Channel Removal	• MLT • Erosion • Channel Removal • Median Filter	• MLT • Erosion • Channel Removal • Median Assign	• MLT • Erosion • Channel Removal • Median Assign	• MLT • Erosion • Channel Removal • Median Assign
	Colorization	UMAP	UMAP	UMAP	UMAP	Logistic Regression	UMAP	UMAP	t-SNE	Quadratic Regression
Metric Results	Contrast	0.0082	0.1376	0.1460	0.2378	0.2651	0.2957	0.4908	0.4276	0.4083
	Local Color Var	0.0003	0.0815	0.0222	0.0728	0.0444	0.0118	0.0000	0.0000	0.0000
	Inconsistency	0.0000	0.0001	0.0001	0.0007	0.0004	0.0025	0.0000	0.0000	0.0000

Note: For each method, best performing methods highlighted in green and worst performing methods highlighted in red.

Future Directions

Develop methods

- Automate channel selection for filtering
- Improve polynomial regression fit and parameter normalization
- Metrics comparing between different scans of objects

Test robustness across different

- Materials
- Shapes
- Sizes
- Arrangements

Understand HCT data

- Photon counts (understand/calibrate for drift and detector degradation)
- Real-world system (currently under construction)

Thank you! Questions?

References

[1] W. Shi, et al., “Pre-processing visualization of hyperspectral fluorescent data with Spectrally Encoded Enhanced Representations,” Nature Communications, vol. 11, no.726, pp. 1-15, 2020.

[2] A. Bjorgan, and L. L. Randeberg, “Application of smoothing splines for spectroscopic analysis in hyperspectral images,” SPIE BiOS, 2019.

[3] I. O. Gallegos, S. Koundinyan, A. N. Suknot, E. S. Jimenez, K. R. Thompson, and R. N. Goodner, “Unsupervised learning methods to perform material identification tasks on spectral computed tomography data,” Proc. SPIE 10763: Radiation Detectors in Medicine, Industry, and National Security XIX, 2018, doi: 10.1117/12.2326394

[4] H. Yadav, A. Candela, and D. Wettergreen, “A study of unsupervised classification techniques for hyperspectral datasets,” IEEE Geoscience and Remote Sensing Symposium, 2019.

[5] J. M. Murphy, and M. Maggioni, “Unsupervised clustering and active learning of hyperspectral images with nonlinear diffusion,” arXiv:1704.07961v5, 2018.

[6] J. M. Fonville, et al., “Hyperspectral Visualization of Mass Spectrometry Imaging Data,” Anal. Chem., vol. 85, no. 3, pp. 1415–1423, 2013.

[7] S. Ranjan, D. R. Nayak, K. S. Kumar, R. Dash, and B. Majhi, "Hyperspectral image classification: A k-means clustering based approach," 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017, pp. 1-7.

[8] C. Gasser, M. González-Cabrera, M. J. Ayora-Cañada, A. Domínguez-Vidal, and B. Lendl, “Comparing mapping and direct hyperspectral imaging in stand-off Raman spectroscopy for remote material identification,” Journal of Raman Spectroscopy, vol. 50, no. 7, pp. 1034-1043, 2019.

[9] G. Dalton, B. Limpanukorn, E. Kemp, J. Clifford, and E. S. Jimenez, “Monte-Carlo modeling and design of a high-resolution hyperspectral computed tomography system with a multi-material patterned anode for material identification applications,” SPIE Optics+Photonics, 2021.

[10] E. S. Jimenez, K. R. Thompson, A. Stohn, and R. N. Goodner, "Leveraging multi-channel x-ray detector technology to improve quality metrics for industrial and security applications," Proc. SPIE 10393: Radiation Detectors in Medicine, Industry, and National Security XVIII, 2017, doi: 10.1117/12.2275850

[11] T. Sato, et al., “Features of Particle and Heavy Ion Transport code System (PHITS) version 3.02,” J. Nucl. Sci. Technol., vol. 55, pp. 684-690, 2018.

[12] L. McInnes, J. Healy, J. Melville, “UMAP: Uniform Manifold Approximation and Projection for dimension reduction,” arXiv e-prints 1802.03426v3, 2020.

[13] L. van der Maaten, and G. Hinton, “Visualizing data using t-SNE,” Journal of Machine Learning Research, vol. 9, pp. 2579-2605, 2008.