SAND2022-5244C

An Improved Process to Colorize Visualizations of Noisy X-Ray Hyperspectral Computed Tomography Scans of Similar Materials

Joshua M. Clifford, Ben Limpanukorn, Edward S. Jimenez, *Sandia National Laboratories*

***Abstract*— Hyperspectral Computed Tomography (HCT) Data is often visualized using dimension reduction algorithms. However, these methods often fail to adequately differentiate between materials with similar spectral signatures. Previous work showed that a combination of image preprocessing, clustering, and dimension reduction techniques can be used to colorize simulated HCT data and enhance the contrast between similar materials. In this work, we evaluate the efficacy of these existing methods on experimental HCT data and propose new improvements to the robustness of these methods. We introduce an automated channel selection method, compare the Feldkamp, Davis, and Kress filtered back-projection (FBP) algorithm with the maximum-likelihood estimation-maximization (MLEM) algorithm in terms of HCT reconstruction image quality and its effect on different colorization methods. Additionally, we propose adaptations to the colorization process that eliminate the need for a priori knowledge of the number distinct materials for material classification. Our results show that these methods generalize to materials in real-world experimental HCT data for both colorization and classification tasks; both tasks have applications in industry, medicine, and security, wherever rapid visualization and identification is needed.**

Keywords—x-ray computed tomography, hyperspectral imaging, colorization, maximum-likelihood expectation-maximization

I. Introduction

This paper further develops approaches originally presented in [1] for preprocessing and dimension reducing hyperspectral computed tomography (HCT) data to represent all channels as a single colorized image. The previous paper used simulated data and provided several potentially useful pipelines for coloring objects in a HCT scan. This paper not only improves the robustness of those methods in numerous ways and tests them on experimental (non-simulated) data. The experimental data also consists of ceramic cylinders with highly similar compositions and material properties. [2] demonstrated that these similar materials can be uniquely identified using an iterative hierarchical clustering method, assuming a priori knowledge of the number of distinct materials in a scan and roughly the number of pixels associated with each material. Our work aims to also improve upon this method to allow identification of each distinct material without any prior knowledge.

II. Methods

A. Data

The HCT system is described in Jimenez et al. [3] and the experimental scans are described in Gallegos et al. [2]. The HCT system leverages the same geometry as the simulated data in Clifford et al. [1], with a single source detector composed of five customized Multix ME100 modules providing a 640 square pixel array calibrated for 300 keV across 128 channels. The scan contains six ceramic cylinders (made from room-temperature glass mica, high-temperature glass mica, alumina-silicate, alumina-bisque, alumina, and water-resistant zirconia) placed in a circular arrangement. A second scan also placed a steel penny along the circle and a wood block in the middle.

B. Preprocessing

Our first preprocessing step is to compute a mask isolating the objects of interest from the background. The majority-vote thresholding algorithm from [1] needed a lower threshold than 0.5 to not erroneously remove objects due to increased photon starvation. This process was made more robust by determining a mask by combining multi-Otsu thresholding and binary erosion on the full (non-hyperspectral) reconstruction. Next, channels were dropped by comparing masks built on channels to the mask built on the full reconstruction, with channels differing on more than 2% of pixels being dropped (in this case channels 109-128). This automated channel drop method is an improvement over the more manual exploratory data analysis used in [1]. Another preprocessing improvement is a shift from using filtered back-projection (FBP) to reconstruct each channel image to using maximum-likelihood expectation-maximization (MLEM). The final preprocessing step used was to apply a median filter (3x3 pixels) on each channel's image. Note an additional improvement in robustness of the preprocessing pipeline from [1] is agnostic to objects in a scan.

C. Colorization

Uniform Manifold Approximation and Projection (UMAP) is again utilized to colorize objects by transforming the 128-channel HCT data to a 3D RGB color space. It provides high contrast between objects with distinct materials, but it is not as interpretable of a colorization scheme, only indirectly related to objects' hyperspectral signatures. Other coloring methods used are polynomial regression, similar to [1] but fit to each pixel rather than an entire object, and functional principal component analysis (FPCA). Each method had its RGB dimensions [0,1] normalized. In order to color a new scan, the extremes used in the original normalization were used to normalize the new results (outside of the [0,1] interval, pixels within -0.05 or 1.05 reassigned to 0 or 1, while others were not colored).

III. Results

A. FBP vs. MLEM Reconstruction Quality

To quantify the image quality of MLEM and FBP reconstructed scans, we use the signal difference-to-noise ratio

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.

(SDNR) defined by [4], which compares the contrast of the objects against background noise. To allow SDNR calculation, we manually segment the objects and background.

We observe generally higher (better) SDNR values for the MLEM reconstructed materials compared to the same materials reconstructed using FBP. On average, the SDNR for MLEM was 46% higher than the SDNR for FBP. These results align with previous work indicating MLEM reconstructions produce a statistically significant reduction in noise compared to FBP [4]. However, a downside of the MLEM algorithm is that it is much more computationally expensive than the FBP algorithm.

An initial visualization comparing FBP and MLEM, with no preprocessing besides thresholding and erosion, makes apparent the massive improvement from using MLEM. With FBP, objects are not smoothly colored and only zirconia is slightly distinguishable from the other ceramics.

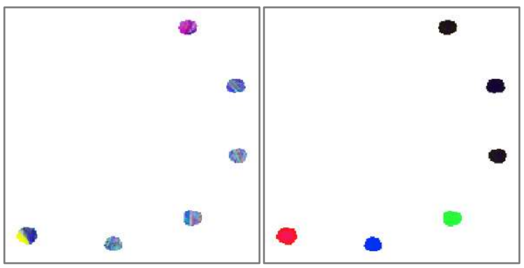

Fig. 1. UMAP coloring comparison: FBP (left) and MLEM (right).Only thresholding and erosion preprocessing.

Since MLEM provides much less noisy images, the remainder of the results will focus on using that reconstruction.

B. Colorization

1) UMAP

While the MLEM UMAP colored image in Fig. 1 shows contrast between some of the objects, the other objects in that image do not have distinct colors assigned. However, after median filtering and removing channels, UMAP produces distinct colors between all of the ceramic objects (Fig. 2).

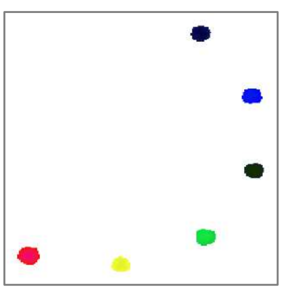

Fig. 2. UMAP coloring after median filtering and channel drop.

2) Polynomial Regression & FPCA

While UMAP performs well in producing distinct colors for each distinct material in the image, the polynomial regression and functional principal components analysis methods do not (Fig. 3). The main driver of the worse performance is the large relative magnitude of difference in hyperspectral signature for the zirconia. If the zirconia object is artificially removed from the image, then these colorization methods are more effective in producing distinct colors for each remaining object.

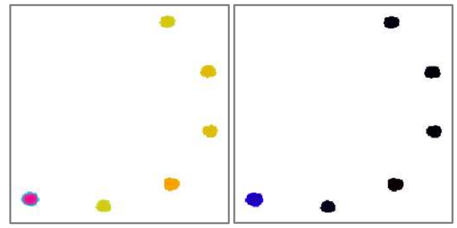

Fig. 3. Coloring comparison after median filtering and channel drop: Polynomial regression (left) and FPCA (right).

C. Introducing New Objects

When the UMAP manifold built on the ceramics-only scan is used to color the scan with the block and penny added, the colors assigned to the ceramic objects ends up being almost identical (Fig. 4).

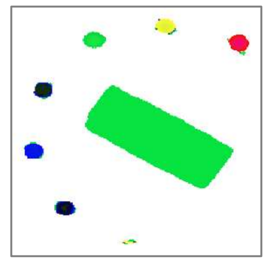

Fig. 4. UMAP coloring trained on original ceramics only data on new data with additional materials (wood block, steel penny).

D. Using Colorization for Material Identification

The UMAP results on the ceramics-only data were used to determine how many objects were in the scan. Calculating an optimum average silhouette width for varying potential number of clusters suggested six clusters, which demonstrates that the colorization method can inform material classification, at the very least as a supplement to iterative hierarchical clustering.

Acknowledgements

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. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

[1] J. Clifford, E. Kemp, B. Limpanukorn, E. S. Jimenez, "A process to colorize and assess visualizations of noisy x-ray computed tomography hyperspectral data of materials with similar spectral signatures," in *Proc. 2021 IEEE Nucl. Sci. Symp. Med. Imaging Conf.*

[2] I. O. Gallegos, et al., "Unsupervised learning methods to perform material identification tasks on spectral computed tomography data," in *Proc. SPIE 10763: Radiation Detectors in Medicine, Industry, and National Security XIX*, 2018, doi: 10.1117/12.2326394

[3] 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," in *Proc. SPIE 10393: Radiation Detectors in Medicine, Industry, and National Security XVIII*, 2017, doi: 10.1117/12.2275850

[4] Gomi, T. (2015) Comparison of Different Reconstruction Algorithms for Decreasing the Exposure Dose during Digital Breast Tomosynthesis: A Phantom Study. *Journal of Biomedical Science and Engineering*, **8**, 471-478. doi: 10.4236/jbise.2015.88044