WBS # – Project Title: 24.1.3.1—NDA Tech: Multimodal Data Fusion with 3D Gamma-Ray Imaging for Safeguards

HQ Team Lead and PM: Arden Dougan

Summary Statement of Work:

Improve the quantitative results obtained from 3D gamma-ray imagers for use in safeguards inspections. In FY23 the work focused on continuing to improve the numerical results obtained with gamma-ray imagers and explore approaches to project such data onto 3D structures mapped with Self Localization and Mapping (SLAM) hardware and software.

Report Title: 2023 Annual Report: Multimodal Data Fusion with 3D Gamma-Ray Imaging for Safeguards

Names of Authors and Affiliations:

Oak Ridge National Laboratory: Klaus Ziock, Jake Daughhetee, Alex Laminack, Kyle Schmitt

Lawrence Berkeley National Laboratory: Ren Cooper

University of California Berkeley: Kalie Knecht

H3D Inc.: Willie Kaye, David Goodman

Major Highlights:

Greater detail on the following highlights can be found in the attached report:

- Generated 3D coded-aperture images using MLEM iterative image reconstruction.
- Demonstrated that the discrepancies between detector calibration and simulated data are due to detector imperfections remapping event locations
- Collected high-resolution data to allow calibrating individual detectors
- Presented a paper on the project at the 2022 IAEA Safeguards Symposium.
- With related funding, presented aspects of the work at the Annual 2023 INMM/ESARDA meeting.
- Developed a general-purpose prescription to remap MLEM basis sets to voxelized reconstruction spaces without needing to fully regenerate those basis sets.
- Have worked with H3D to facilitate next year's work on implementing the project results in the H3D commercial software base.

Progress: Please see attached report.

Publications: K.P. Ziock, et al., “Nuclear Materials Safeguards with Scene-Data-Fusion-Enabled Gamma-ray Imaging,” Proceedings of the 2022 IAEA Safeguards Symposium: Reflection on the Past and Anticipating the Future,” Vienna, Austria, Oct. 2022.

Related:

K.P. Ziock, et al., “Efficiency Calibration of the H3D H420 Gamma-Ray Imager,” Joint Annual INMM & ESARDA Meeting, Vienna, Austria, May 2023.

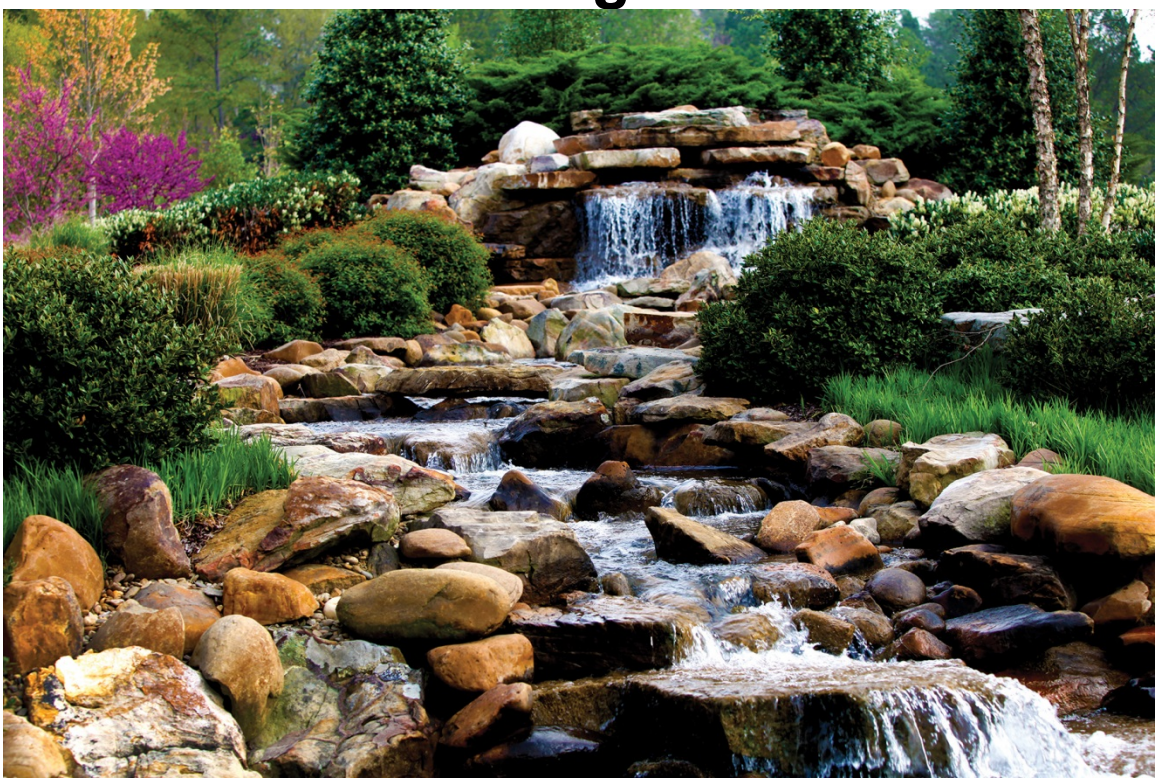
A. Laminack, et al., “3D Source Reconstruction Using Coded-Aperture Gamma-Ray Imaging,” Joint Annual INMM & ESARDA Meeting, Vienna, Austria, May 2023.

K. Knecht, et al., “Scene-informed optimal measurement positions for quantitative safeguards measurements,” Proc. INMM & ESARDA Joint Annual Meeting, Vienna, Austria, 2023.

References: Please see attached report.

ORNL/SPR-2023/3216

2023 Annual Report: Multimodal Data Fusion with 3D Gamma-Ray Imaging for Safeguards



K.P. Ziock
R. Cooper
A. Laminack
J. Daughhetee
F. Gonzalez
D. Goodman
W. Kaye
K. Knecht
K. Schmitt

October 2023

OAK RIDGE National Laboratory

ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via OSTI.GOV.

Website www.osti.gov

Reports produced before January 1, 1996, may be purchased by members of the public from the following source:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone 703-605-6000 (1-800-553-6847)
TDD 703-487-4639
Fax 703-605-6900
E-mail info@ntis.gov
Website http://classic.ntis.gov/

Reports are available to US Department of Energy (DOE) employees, DOE contractors, Energy Technology Data Exchange representatives, and International Nuclear Information System representatives from the following source:

Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831
Telephone 865-576-8401
Fax 865-576-5728
E-mail reports@osti.gov
Website https://www.osti.gov/

ORNL/SPR-2023/3216

Physics Division

2023 Annual Report: Multimodal Data Fusion with 3D Gamma-Ray Imaging for Safeguards

Oak Ridge National Laboratory:
K.P. Ziock, F. Gonzalez, J. Daughhetee, A. Laminack, K. Schmitt

Lawrence Berkeley National Laboratory:
R. Cooper

University of California, Berkeley
K. Knecht

H3D, Inc.
D. Goodman, W. Kaye

September, 2023

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

Contents

Figures

1. INTRODUCTION

1.1 PROJECT GOAL

As a visualization tool to locate radioactive materials, gamma-ray imaging represents a mature technology with several well-developed commercial instruments [1, 2] available to the interested user. However, for use in international safeguards, one is concerned not just with the location of materials, but also the types (isotopes) and quantities of materials that are present. With the excellent semiconductor-detector-based energy resolution available from the referenced systems, the presence of specific isotopes can be readily determined and localized to specific regions of the images. However, accurate isotopics and quantitation require that results from the instruments are elevated to a new level of precision, one that meets the standards for non-destructive analysis. We have successfully demonstrated such performance from one of the instruments (PHDS, HPGe-based coded-aperture imager) by performing a pixel-by-pixel sensitivity calibration of the imager that allows users to facilely integrate the observed flux across an image and determine the amount of material that is present [3].

The goal of this project is to bring a comparable level of calibration to the H3D, Inc. H420, and extend the performance to both the coded-aperture and Compton imaging modalities. However, to achieve accurate results, it is important to know where material is located with respect to the detector so that one can account for the change in observed flux with distance, and to include attenuation by overlying materials. If the material's location is known *a priori*, such information can be added by hand, but frequently this is not the case, and this significantly hampers automated use, while degrading the accuracy of results. To counter this, a second major goal of the project is to utilize contextual sensors that are used to generate a three-dimensional (3D) map of the measurement location. Based on such a map, one can collect data from a few locations, and use such data to create a 3D source distribution map. Observed emission regions can be selected from the maps and integrated to determine the amount of material present. To distribute these advances to interested users (notably, the IAEA), the algorithms required to achieve these results will be integrated into the commercial user software base available from H3D, Inc.

1.2 PROJECT EVOLUTION

The project originally included researchers from Oak Ridge National Laboratory (ORNL) and Lawrence Berkeley National Laboratory (LBNL) as the primary partners, with ORNL serving as the Principal Investigator (PI) organization. LBNL brought expertise in Compton imaging, contextual sensors and related simultaneous localization and mapping (SLAM) techniques. ORNL brought experience in coded-aperture imaging, detector calibration, and quantitative analysis. As the manufacturer of the H420 Imager, H3D, Inc. formed a third major partner, providing expertise on the device as well as the means of final software distribution. Sandia National Laboratory (SNL) contributed as a minor partner, exploring the use of their directional unfolding source term (DUST) algorithm as an alternate means of obtaining quantitative results, particularly from the Compton data. Through LBNL there has also been participation by a graduate student from the University of California at Berkeley (UCB) who is developing algorithms to guide users on optimizing imager placement for long-dwell measurements.

Near the close of the prior year's effort, the LBNL personnel associated with the project pursued opportunities beyond LBNL and it became clear that, while the laboratory retained significant expertise in contextual sensing, no personnel working on the project would continue into the fiscal year covered by this document. LBNL management sought to resolve this issue by appointing an interim principle investigator (PI), and stepping back from the effort for the first half of the reporting period with plans to hire new staff who would support the project after that hiatus. At the same time, ORNL planned to take advantage of the delay by focusing on problems with detector calibration.

While the calibration issue appears to be resolved, the LBNL staffing problem persisted throughout the year, with no clear path forward identified by year's end. Based on this, Headquarters will terminate LBNL participation at the end of the current FY. Because of this, the Compton portion of the project will also end, and since Compton imaging is the primary consumer of the SNL DUST algorithm, that work is also ending.

To achieve its goals of quantitative 3D gamma-ray imaging, the project does require contextual information as previously supplied by LBNL hardware and software. Fortunately, over the life of the project, H3D, Inc. has been working to develop an in-house contextual information capability that should serve the project goals. Barring that, they are also familiar with a commercial solution from Gamma Reality, Inc. that should be suitable for the project needs. While it may seem that changing the contextual information supplier late in the project would further delay attaining our final goal, in fact, all of the contextual information solutions are based on similar hardware (LIDAR, depth cameras, inertial sensors, and possibly GPS), with SLAM software based on the Google, open-source Cartographer platform [4]. Hence, relying on an in-house solution from H3D, Inc. should ease integration, since they are tasked with integrating any advances from the project into their software base, and by relying on their internal solution, that task is already well advanced. It remains only to integrate the quantitative advances from improved calibration and its application to coded-aperture imaging with the rest of the extant software, and that is the primary goal of the first half of FY 2024.

1.3 PRIOR DEVELOPMENTS

The project commenced in the second half of FY2020 with funding to allow procuring H420 imagers and the hardware components of localization and mapping (LAMP) modules, the LBNL contextual imaging platform. With work planned at both institutions, two sets of imagers and hardware were purchased, one for each laboratory. This allowed for independent development while also providing multiple hardware samples to allow for validating that results obtained are hardware independent. In fact, since the start of the project, ORNL has become a partner in the Strategic Laboratory Assessment (SLA) that is developing advanced techniques for the Surplus Plutonium Disposition (SPD) project that will down blend and dispose of surplus cold-war plutonium. As a part of that project, ORNL is exploring the use of H420 imagers to monitor glovebox holdup, and that work is closely related to this effort. Participation in SLA provides not only additional H420 imagers, but also additional resources to further algorithmic development.

In the following summary, the progress is described by the different intellectual threads associated with the project. The narrative is meant only as a general overview. More detailed information can be found in the prior mid-year and year-end reports.

1.3.1 Contextual Information

In FY2021, the first full year of the project, LBNL designed and assembled the LAMP 2 contextual platforms using the hardware purchased in FY 2020. The units were ready for use by the middle of the year with one module delivered to ORNL.

A primary goal of this project is to transition from the previous LAMP measurement paradigm of continual system motion, to one better adapted to IAEA needs and good quantitation. Specifically, we are targeting a concept of operations (CONOPS) that includes two measurement phases: a short initial area survey to generate a contextual map of the measurement location, and a second phase comprised of one or more long-dwell measurements from a few fixed locations. This shift is necessitated by the change in emphasis from finding radiation sources to accurately determining the amount of radioactive material that is present. Such quantitation requires sufficient measurement times to minimize uncertainties due to

counting statistics. Collecting long-dwell measurements from a few fixed locations (albeit enough to allow generating 3D source maps) also allows inspectors using the devices to attend to other duties during the integrations. Additional reasons to use fixed locations are that the contextual sensors and associated data processing represent significant power and information usage, requiring large batteries and data storage. Minimizing the system motion during data collection also reduces motion blur in the resulting images, something of more concern to higher-resolution, coded-aperture images than to the Compton imaging generally used with the contextual systems.

Discussions within the project that lead to this CONOPS lead to questions on IAEA needs for the imagers. Unfortunately, those queries, posed early in calender-year 2022, were not resolved until late in that year when discussions were held at the IAEA during the Holdup Symposium hosted by that organization. Guidance obtained indicated that fixed measurements were the appropriate approach and that expectations on quantitation were low, with the first-order goal of simply validating that spectra from different objects in the field of view were correctly associated with that object, a requirement easily met by the coded-aperture images.

In addition to modifying the SLAM software to the revised CONOPS, LBNL also worked to improve the functionality/reliability of the LAMP 2 software. Specifically, they moved the various software components to a Docker configuration that streamlined module communications and improved error recovery, allowing individual modules to restart if errors occurred, without having to restart the full system.

There was also considerable work performed to see if one could use photogrammetry to replace the light detection and ranging (LIDAR) instruments that form a major component of the SLAM approach. LIDAR units are expensive, power hungry, and relatively large so their elimination from the system would represent a significant improvement in some scenarios. Photogrammetry achieves the same contextual results by analyzing the images from cameras and depth cameras. Unfortunately, while the results were promising, no final solution was achieved before the LBNL staffing issues became problematic.

A final development of the contextual system development is being performed by a graduate student at UC Berkeley. The research uses the contextual information to find optimum locations and imager pointing (or imager pose) from which long-dwell measurements should be collected [5]. To achieve this the scene is broken into voxels of potential source locations and allowed measurement locations. From each potential location, the sensitivity to all of the scene voxels is computed and used to find those locations that simultaneously minimize the variance in sensitivity between voxels while also maximizing the overall sensitivity. This has been implemented and tested using both a brute force approach to validate the overall concept, and a Genetic algorithm that is computationally more efficient and suitable for implementation in a fielded system [6].

1.3.2 Compton Imaging

Compton imaging generates an image by using the muti-site gamma-ray data from the instrument. Specifically, the locations and energy depositions of scatter and subsequent photoelectric capture events from a single gamma-ray can be used to localize the direction of incidence of the gamma-ray to lie on a cone of possible directions [7]. Source locations are identified (and images created) because the cones from gamma-rays that originate from the same location (i.e., a single source) intersect at the source's location. Practically, the angular resolution of this type of image is ~ 20°, with the rings (cone intersections with a sphere at a specific distance) from a point source contributing background across the overall image. This will cause spectral cross contamination if multiple sources are in the field of view. Compton images have the advantage that rings that do not overlap with a specific image location do not add to the background at that location, allowing one to perform background rejection. Compton imaging

works best at higher gamma-ray energies where the probability of simple photoelectric absorption is small. Given that the H420 detectors inherently provide full 3D event localization, Compton imaging is a natural means of generating images and is the imaging modality most developed by H3D, Inc.

To obtain quantitative information on source intensities, LBNL worked with the list-mode Maximum Likelihood Expectation Minimization (lm-MLEM) algorithm [8]. They were able to show that quantitative images of point sources could be obtained [9]. In addition, they explored the impact of the distributed nature of this imaging technique on both images and spectra. Specifically, it is known that higher-energy emissions from other sources in the field of view can add undesirable structure to the images made with lower-energy spectral lines. The problem comes because multi-site Compton-continuum events underlie the lower-energy photopeak, and when projected during image generation, create spatial structure in the image. Such artifacts can be reduced by processing the images starting from the highest-energy lines and working to progressively lower energies by subtracting the artifacts based on the higher-energy data and known Compton-continuum efficiencies [10], however such an analysis is computationally burdensome. Instead, a technique was developed that used the continuum events in a spectral region of interest just above the energy of the lower-energy line of interest. Images were made with those events and included in the lm-MLEM analysis using the precept that the Compton-continuum is relatively flat and that the imaging properties will change minimally with small shifts in energy. They demonstrated that this computationally simpler approach achieved good results by comparing the compensated low-energy images with and without the presence of a higher-energy source, finding good agreement between the two.

An additional Compton imaging advance was made using 3D contextual data to correct for the angular blur in the Compton images. Specifically, in lm-MLEM images, the flux from a source in the near field extended to neighboring voxels that corresponded to a distant wall. When the intensity from those locations was corrected for the apparent greater distance, the overall integrated flux from the image was over an order of magnitude higher than the strength of the single near-field source in the scene. This was largely corrected using the precept that the flux from angularly adjacent voxels that were at significantly different distances from the imager were unlikely to originate from source material at both locations. A regularization term was developed to penalize such situations with the flux collecting to the stronger location, in this case the closer point source [9].

1.3.3 DUST

Sandia's DUST algorithm was originally developed with funding from NA-22 for use with their Gamma Detector Response and Analysis Software (GADRAS) code [11]. It uses the known Compton imaging response function to collect the distributed background from non-overlapping Compton rings, into regions of the image where more overlap occurs, e.g. those locations identified with source. Over the course of the project the algorithm has been developed from a research branch of the GADRAS code to a mature integrated feature. In addition, to make the algorithm function correctly, the H420 calibration files included in GADRAS were updated to improve the overall performance across the GADRAS suite of capabilities. As part of the development, SNL provided a general-purpose tool to allow LBNL (and other users) to incorporate DUST into their code base. Unfortunately, that tool was still being debugged when the LBNL team member working on the project departed.

1.3.4 System Calibration

In order to obtain quantitative images from the H420 imager, we are applying an approach demonstrated to give quantitative determinations of the amount of ^{235}U present in distributed sources from coded-aperture images that are good to 3% [3]. The approach relies on careful mapping of an imager's response function across its spatial and spectral range. The maps are created at regular energy intervals and across

the field of view, with interval spacings small enough that interpolation techniques can be applied to obtain the response at specific energies and locations. However, because radiation sources do not exist with sufficiently fine and regular energy spacings, we use GEANT-based [12] Monte Carlo simulations to generate the calibration "data." The simulation codes are themselves validated against laboratory data collected at several different source energies. Unfortunately, in applying this technique to the H420 imager, we found that sensitivity maps generated at the same energies from laboratory and simulation data were not in good agreement, with the shapes of the 2D sensitivity maps differing substantially. To understand this significant resources were used to address the discrepancies.

As part of the calibration effort, improvements were made to the coded-aperture mask design used by the H420 system. First, the phase of the mask pattern was altered to remove high-frequency spatial changes in the imager response that occurred because the original mask phase has a single open and closed row and column through the center of the mask [13]. This interacted unfavorably with the vertical and horizontal gaps through the center of the detector plane due to the 4-crystal design of the detector. At the same time, we doubled the thickness of the mask from 1.5 mm to 3 mm to improve the response of the coded-aperture imaging at the important lines at 375 keV and 414 keV from ^{239}Pu. While the changes improved the overall high-energy imaging efficiency and the system spatial response at the center of the field of view, it did not fix the discrepancy between the shapes of the experimental and simulated response maps.

In general, we expected the divergence between the experimental and simulated response functions to be due to incorrect blurring of the gamma-ray event locations in the simulations. However, we found that the experimental shapes of the system sensitivity maps between the different mask designs and from several imagers to be similar, so it was hoped that a single blur function could be developed for use with all imagers. However, simple blur functions applied uniformly across the detector area did not fix the problem. Attempts to use a pinhole or slit scans to map the response of the detector as a function of energy met with additional complications from the detector (see 1.3.5 below). In the end, some slit data were successfully collected and used to generate generic spatially dependent blur functions (e.g. crystal edge versus crystal center) but these did not fix the discrepancies.

1.3.5 H420 Instrument Performance/Issues

The H420 represents a complex instrument that has evolved commercially to fulfill the needs of its end users. However, this project requires a greater degree of accuracy from the instrument while also using it in new ways. Invariably this led to issues of operability, complicating automated data collection, and data-handling practices that, while sufficient for current users, were incommensurate with our goal of accurate quantitation. Some of the more challenging issues included: infrequent but numerous restarts of the H420 imager during automated data acquisitions that confounded collecting high-granularity data with high statistics, data dropouts that were not correctly recognized as part of the deadtime calculations, correction for intermittent noisy pixels by raising their threshold which confounded attempts to map the detector with localized pinhole data, and apparent changes to the event distributions as a function of localized event rate (pinhole beams), even though the localized rate was easily sustainable across the full detector.

At the start of the current reporting period, most of these issues had been resolved or their impact mitigated. For instance, random restarts of the H420 were greatly reduced when it was uncovered that some commands needed to be sent with long (many second) wait times. Software updates were provided by H3D, Inc. to allow disabling automated threshold changes so that localized data could be collected successfully. Deadtime corrections were also improved by including simple checks in the analysis routines that look for anomalously long times between events. H3D, Inc. has also recognized that for coded-aperture imaging, detector uniformity is important and has started including a quality check on the crystals that are used in the H420. (In fact they voluntarily replaced the crystals in one of the ORNL imagers after recognizing they did not meet their new specifications.)

1.3.6 Coded Aperture Imaging

At gamma-ray energies, focusing optics are generally ineffective, and while images can be made with pinhole cameras, such devices are very inefficient. Coded-apertures are a means of overcoming the inefficiency of pinhole cameras by replacing the blocking sheet containing the single pinhole with one that contains many pinholes [14]. In fact, for weak sources, the best sheet design is 50% open area, greatly increasing the efficiency of the imager [15]. While each opening in the sheet projects an image onto the position-sensitive detector, the images overlap to the point that the shadowgram seen by the detector does not represent an identifiable image of the source. Instead one can think of the mask encoding the scene onto the detector, with different source configurations generating different shadow patterns. Modern coded-aperture designs allow simple mathematical deconvolution of the recorded shadowgram to create artifact-free images of the source field. With the scene encoded as variations in the count rate seen across the face of the detector, one must be careful to remove any count-rate variations that are not due to the shadow cast by the mask. Unfortunately, such variations are always present, either due to radiation that scatters around the mask, background sources, or non-linearities in the detector piling counts up in some regions of the detector. Fortunately, one can measure such structures *in situ*, without sacrificing measurement time by using an anti-mask that exchanges the open and closed holes in the mask and taking equal-time integrations with the mask and anti-mask coded-apertures [16]. The difference between these two data sets have non-mask-induced count rate variations removed and can be deconvolved directly to generate the images. In essence, one measures the background with half of the detector and source field plus background with the other half. Halfway through a measurement, one exchanges which is collected by each part of the detector. Operationally this is achieved automatically by using mask patterns that are anti-symmetric when rotated about their centers [13].

In prior years, extant software was adapted to allow generating coded-aperture images with data from the H420. Software was also modified to allow automated data collection with the system for the various calibration tasks discussed in Section 1.3.4 above. To generate high-quality images requires that the detector uniformly samples the full coded-aperture mask pattern. However, the H420 has gaps between the four CZT detectors used as its focal plane. The software had to be adapted to compensate for partial coverage of some of the pixels in the shadowgram. In addition, as already discussed, we modified the coded-aperture mask to move the open and closed rows away from the center of the mask where they interacted unfavorably with the gaps between the crystals.

A goal of this project is to project coded-aperture images onto 3D contextual data to generate 3D gamma-ray images. The distribution of sources in depth is important to allow correcting for the change in imager solid angle as a function of source distance. Prior work along these lines included simply projecting the 2D images onto a 3D volume from multiple directions. If one considers a single point source, then the image from each direction represents a beam through the volume and the beams will overlap at the source location. However, while adequate to determine the distance to the source, it has an inherently distributed response function (e.g., each beam leaves source intensity along its entire track through the volume) and requires many views to blur out the response function to the point that it is not a major component of the image.

As a first attempt to improve on this approach a Model Based Iterative Reconstruction (MBIR) technique [17, 18] was applied to simulated data. The data represented a distributed source that was imaged from a number of directions around the source. To generate a 3D reconstruction, the 2D reconstructed coded-aperture images from each direction were projected into the reconstruction volume. This approach takes advantage of the fact that iterative techniques can collapse the source strength to specific locations, removing the impact of the beams seen with simple projection. Results from that work demonstrated the efficacy of the approach and that a limited number of views could be used to obtain reasonable results. However, the projection of processed images onto the scene is not ideal for more than simple source

configurations in that the system response changes with source distance and it is difficult to make use of that information once the 2D reconstruction has been applied.

2. FY2023 PROGRESS

With the disruption to the project by the LBNL staffing issues, the original tasking in the PWP does not provide a logical framework for a narrative on progress in the project. Instead the narrative will be framed around the major intellectual goals of continued development of coded-aperture imaging. This includes sub headings of 3D image reconstruction and improved calibration of the detector response, which is where the bulk of the effort occurred. Tasks 2.1, 2.3, and 2.4 were all concerned with transferring the project code-base to H3D and because the project will now rely heavily on internal H3D software, the emphasis has shifted to ensuring that those software tools will provide support to the quantitative 3D coded-aperture imaging goals. This includes making sure that calibration of the H420 imagers is sufficient to support quantitation, and that the contextual data provides sufficient information to accurately determine the imager pose during a measurement and a suitably voxelized representation of the measurement locale. Those items are covered under the Software Integration heading that leads the presentation.

2.1 SOFTWARE INTEGRATION

The project advances will be implemented in the H3D, Inc. Visualizer software suite for transfer to the IAEA and other end users. However, because it was unclear during the year whether or not the LBNL or H3D, Inc. contextual data tools would be used, detailed work beyond identifying software requirements was not performed.

The Visualizer software includes tools to replay the list-mode data created by the instrument, with options for the user to select individual and multiple data collection sessions, as well as specific temporal portions of those data. Specific spectral regions of interest can then be used to create coded-aperture and Compton images that can be viewed as false color heat maps, either independently, or projected onto visible-light images of the scene. In addition, an advanced version of the code (Visualizer 3D) includes tools to handle and display contextual data, including, the imager track through the scene and 3D projection of Compton image data onto the scene.

The quantitative, coded-aperture imaging algorithms will have to be added to these tools and the user interface expanded to allow simple selection of spatial regions of interest from which material masses will be determined. For simple usage, such as measuring known containers at a fixed distance, the quantitative data can be projected and harvested from the simple 2D overlay maps. For more complex scenes, projection and integration of 3D volumes will be used. Tools already exist to perform both of these functions, although the underlying image generation algorithms will have to be updated with the quantitative versions under development. A key component of this work has been identified as the addition of hyperspectral calibration data to the code base. These data represent a pixel-by-pixel (2D) or voxel-by-voxel (3D) response of the imager to a point source at each location in the field of view and at regular energy intervals. In addition, tools to interpolate to different distances and individual spectral lines will also be required as part of the calibration structure. Interpolation routines to allow registration of the precomputed calibration data to the fixed 3D scenes voxels will also be required (see 2.2.2). Finally, to include and use the hyperspectral calibration data, the point-cloud-based 3D representations of the measurement environment currently used by Visualizer 3D will have to be updated to generate a voxelized version of the map. This map must be labeled to identify and keep only those voxels that can contain source, e.g. source material can only reside within enclosed volumes or on surfaces.

2.2 CODED-APERTURE IMAGING

2.2.1 Detector Calibration

At the heart of obtaining quantitative results is knowing the detector response function as both a function of energy and location. For simple 2D reconstruction at a fixed distance, this represents a hyperspectral response cube comprising a pixel-by-pixel map of the number of counts reconstructed from an image of a source in each pixel of the image. Such maps can be generated at specific energies using a known radioactive source progressively scanned across the field of view [19]. However, because the response varies as a function of energy, generating such maps with enough granularity to allow successful interpolation for any arbitrary energy is problematic. Instead, the data are used to validate Monte Carlo simulation models that can then be run at arbitrary energies.

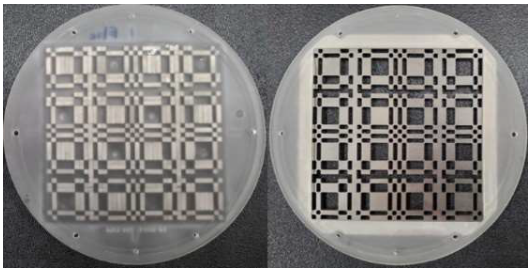


Fig. 1. Left, potted mask and right, machined mask with clear openings.

As mentioned above, there have been persistent discrepancies between the simulated and measured response maps of the imager. Throughout the year understanding and correcting these discrepancies has been a significant part of the work performed. One concern, addressed by this effort, was the construction of the coded-aperture mask in the H420 imager. The manufacturer creates the mask by placing small tungsten elements into a holder to create the overall mask pattern and then potting that assembly into a monolithic structure. While these design elements were included in the simulations, there was concern that the "open" mask elements were filled with epoxy and there could be some error in the simulations of that material. To explore this, machined masks created from a sheet of tungsten were purchased and used in place of the original mask (Fig. 1). Unfortunately, experimental data collected with these monolithic masks did not change the overall shape of the experimental response maps (Fig. 2). However, they do have the benefit that the open holes are now truly open and calibration data taken at 122 keV with a ^{57}Co source showed an 8% increase in sensitivity.

As with all detectors, the CZT crystals in the H420 introduce some measurement error into the data collected. In general, there is a spatial blurring that softens the edges of a sharp shadow pattern. This blurring is worse at the edges of the CZT crystals then at the center and given that the sensitivity plots for different imagers were similar, our initial line of research was to implement a blurring function in the simulated data that matched this response. To measure amount of blurring we have tried using a collimated source to create a pencil beam and then a narrow slit to create a linear beam. In both cases, we found that the whole face intensity maps created from the aggregate collimated data did not match those obtained with a flood-field illumination. This was ascribed to local charging in the CZT distorting the response function and so we sought a technique to collect the data with a lower intensity source, without commensurately increasing the collection time. To achieve this, a multi-slit mask was used. This allowed the use of a weaker source for a longer interval without increasing the overall measurement time because we obtained data from multiple regions of the detector simultaneously (Fig. 3).

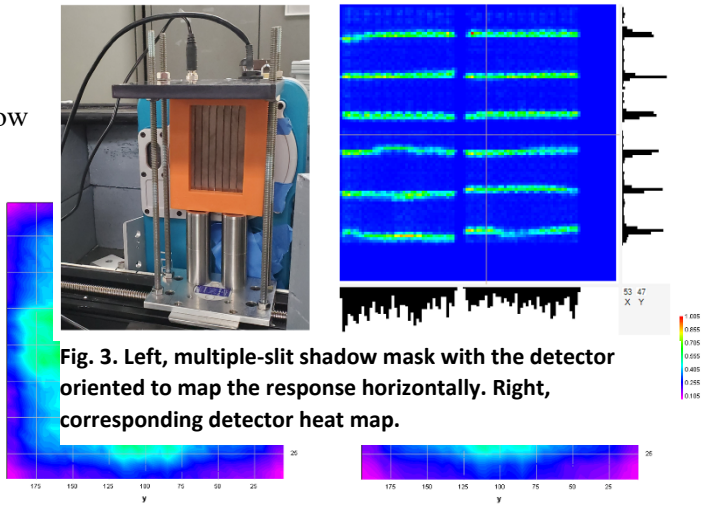


Fig. 3. Left, multiple-slit shadow mask with the detector oriented to map the response horizontally. Right, corresponding detector heat map.

Fig. 2. Normalized sensitivity maps obtained with: (left) the original H3D mask with potting in the holes, and (right) metal mask with clear holes. The responses maps are basically identical, although the overall sensitivity of the unpotted mask is ~ 8% higher.

Initially, the multi-slit-scan data were analyzed to generate a global blurring function that increased the blurring at the crystal edges. However, when that function was applied in the simulations, the overall sensitivity map was no closer to that obtained from the experimental data.

Finally, it was decided to explore the impact of the spatial distortions in a single H420 detector plane to see if that was the source of the measurement/simulation discrepancies. Such distortions are seen as curvature of the slit-scan data in Fig. 3, and arise from distortion of the electric fields within the CZT that originate from crystal imperfections. With the multi-slit data, it was clear that in some portions of the crystals, the deviations were comparable to the size of the mask features and thus large enough to impact the sensitivity of the coded-aperture response maps.

When the detector-specific detailed calibration was applied to the simulation data we finally obtained a simulated response map comparable to that obtained from laboratory data. Based on this, calibrations were performed for several additional detectors (Fig. 4) and the results confirmed, that while the overall experimental response of the different imagers were similar, they were due to the individual distortions of each detector.

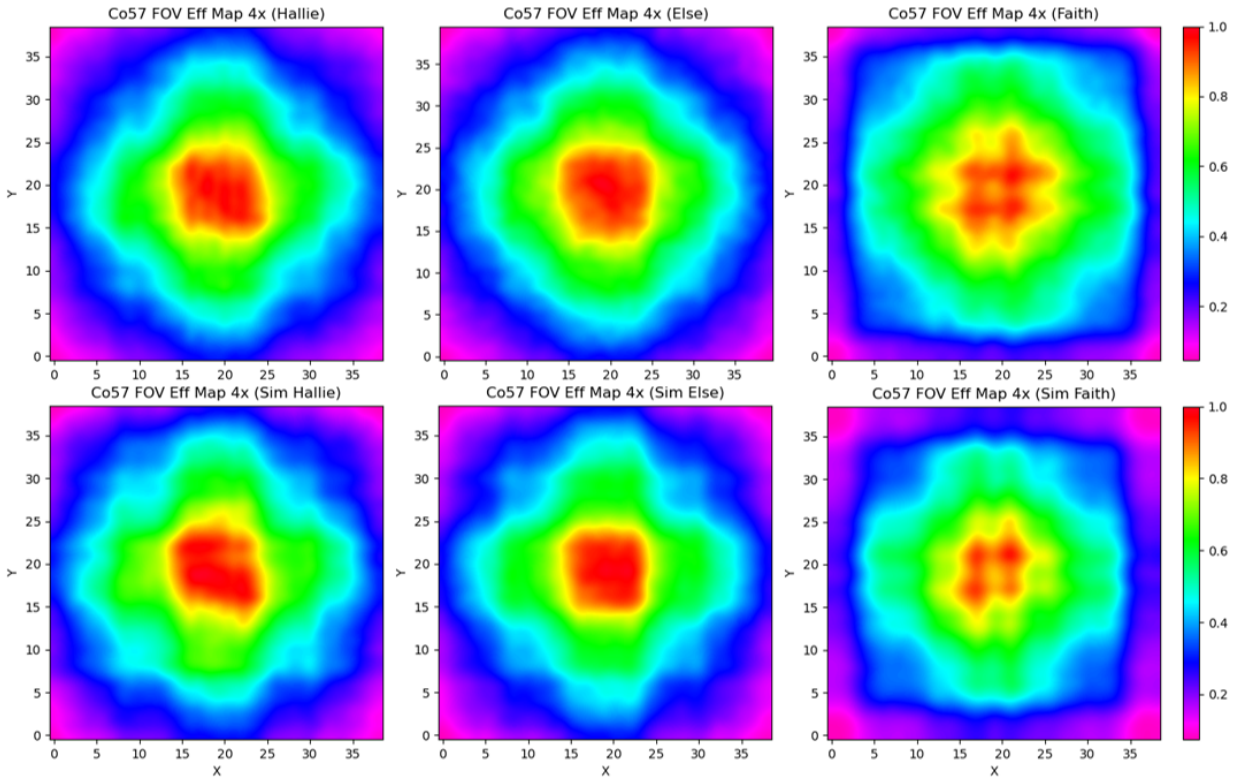


Fig. 4. Experimental (top) and simulated (bottom) sensitivity maps for three different imagers. The two on the left use the new phase mask developed by this project, while the one on the right uses the original manufacturers mask design.

Based on these results, it is clear that each detector will have to be individually calibrated to map the spatial distortions inherent in its detector plane. To facilitate and improve on that work, a multi-pinhole mask was ordered to collect further data. The use of a multiple pinhole mask has the same multiplexing advantages of the multi-slit mask, but will allow simultaneous remapping of the detector distortions in two dimensions. With the slits, the 2D mapping was inferred by integrating small lengths of each slit and matching the independent x- and y-slit results, resulting in some uncertainty in the combined 2D mapping.

With an improved x-y mapping available from the multiple pinhole data there is an additional approach to reconciling the experimental and simulated response maps that will be explored. Specifically, the improved calibration data may allow correcting the experimental data to remove the linear distortions to the shadowgrams collected by an imager before the images are reconstructed. This has the potential benefit of allowing a single set of hyperspectral simulation data to be used for all detectors, and should improve the results from simple cross-correlation analysis used in on-line image generation and for simple 2D analysis. Its benefit to 3D MLEM reconstruction is not as obvious because that approach directly uses the distribution of the collected data and compares it to the redistributed simulated data. When the experimental data is linearized, invariably some information will be lost.

With the successful results, the overall system calibration procedure is now validated and will include the following steps,:

1) Electronic calibration at multiple energies. This calibration is currently performed by the manufacturer at a single energy and balances the interpolation routine used to locate an event to sub-pixel accuracy. Software exists to allow incorporating terms taken at multiple energies.

2) Multi-Pinhole calibration data collected at multiple energies. Those data are analyzed to create energy-specific response maps with interpolation procedures developed for use at other energies. Because the spatial remapping is due to electric-field distortions originating from crystal imperfections, it is not expected to vary significantly with energy.
3) Monte Carlo simulation data created at regular energy intervals. Together with detector-specific blur maps, these data will form the core of the hyperspectral calibration data.

For 2D analysis the hyperspectral calibrations will include a pixel-by-pixel quantitative response of the imager that can be corrected for distance to the source plane. For 3D analysis the hyperspectral calibration will include the quantitative shadowgram of a source in a regular voxel map covering the volume in front of the imager. The final form of the 3D calibration data is still under consideration and will comprise either spatially blurred or spatially blurred and distorted maps, depending on the advantages and system performance if the detector data are linearized before analysis.

2.2.2 3D Image Reconstruction

Significant progress has been made in developing a 3D coded-aperture image reconstruction algorithm that works with the detector data rather than the processed images. The Maximum Likelihood Expectation Maximization technique (MLEM) [20] leverages Bayes' Theorem, which describes the probability of an event occurring given prior knowledge. Specifically, it uses the system response matrix, $\boldsymbol{M}$, comprised of the set of shadow patterns expected for a single source sequentially placed in every pixel of the imager's field of view. Then given the image $\boldsymbol{I^S}$, of the current iteration, the algorithm calculates the detector that should result from that image, $\boldsymbol{D^S}$, by adding the intensity-weighted sum of response terms for all source locations:

$$D_i^s = \sum_{k=1}^{m} M_{ik} I_k^s.$$

These values are compared to the actual observed counts, d_i, as a ratio R_i:

$$R_i = \frac{d_i}{D_i^S} = \frac{d_i}{\sum_{k=1}^{m} M_{ik} I_k^s},$$

that generates a correction detector. That correction factor is projected back through the response matrix and multiplied by the current image to generate a new image

$$I_j^{s+1} = \frac{I_j^s}{\sum_{i=1}^{n} M_{ij}} \sum_{i=i}^{n} M_{ij} R_i = \frac{I_j^s}{\sum_{i=1}^{n} M_{ij}} \sum_{i=i}^{n} \frac{M_{ij} d_i}{\sum_{k=1}^{m} M_{ik} I_k^s}$$

Where the denominator under the current image is a normalization factor for the overall summed strength of the system response matrix.

Use of this standard image reconstruction algorithm is somewhat hampered in coded-aperture imaging by the fact that as a ratio-based approach, it does not handle negative numbers, so background subtraction is more complex. Specifically, the mask minus anti-mask technique described in 1.3.6 above, will result in negative values for some detector locations. To allow the use of mask and anti-mask data in MLEM reconstruction, we use the formalism proposed by Brubaker [21], where the data and system response matrices are expanded to handle the mask and the anti-mask data separately—a capability allowed by the time-tagged nature of the data coming from the instrument. To include the presence of background events that are not encoded by the mask, an additional response term is used that links the same mask and anti-mask detector pixels to each other. Simulated sample mask and anti-mask detector data from a point

source are shown in Fig. 5. The data also include a complex (spiral) background term that is the same in both data sets.

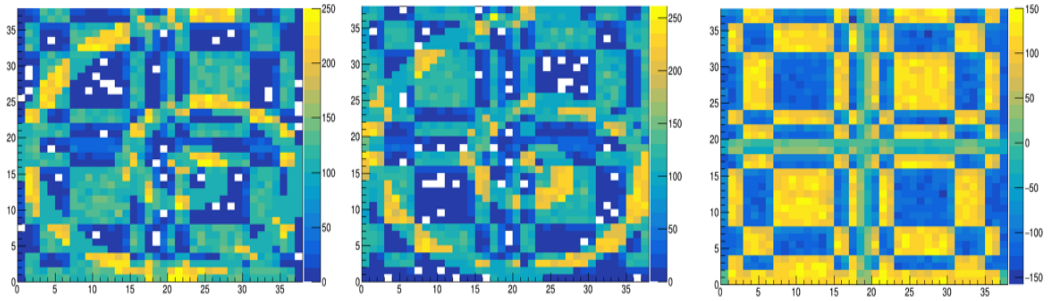


Fig. 5. Simulated data of a point source and a spiral background. In the standard approach the anti-mask data (center) is subtracted from the mask data (left) to remove the background (right), but yields negative detector values.

In practice, the mask and anti-mask terms are stacked to create a single rectangular detector that includes both the mask and anti-mask data. The data in that configuration, and a sample iterative detector together with the source-only and background-only detectors determined from an MLEM reconstruction are shown in Fig. 6.

The software developed using this approach has been successfully tested with data collected late in FY 22 where a source holder with several different source configurations was observed from multiple angles (Fig. 7). A single-source reconstruction with different numbers of iterative steps is shown in Fig. 8.

Because coded-aperture imaging is an indirect imaging technique, one must include the response from all of the detector elements for all of the voxels in the scene. This means that the MLEM reconstruction can be computationally intensive since each iteration includes a forward (image to detector) and backward (detector to image) calculation through the response matrix. The computational burden can be reduced by using contextual data that restricts image locations to only those where source material can reside, e.g. on surfaces or in enclosed volumes. With this information many of the voxels in the field of view can be removed from the iteration.

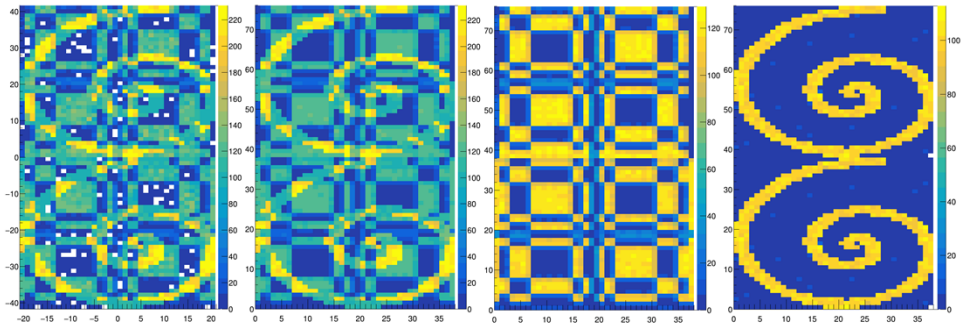


Fig. 6. MLEM approach. The left image is the same simulated data from Fig. 1, but now displayed as a single larger detector with the mask and anti-mask detectors "stacked" into a single structure. The MLEM detector after fitting (2nd from left), matches the composite detector well, and can be split into terms that include only the coded-aperture-modulated signal (3rd from left) and the background (right).

Limiting the number of available source voxels has the additional benefit that it reduces the number of possible ghost-source locations that arise from collecting data from only a few angles. Consider that if one views two sources from two different directions, then because little distance information is available from a single imager, sources could be located wherever lines connecting the imager to each of the sources intersect. Referring to Fig. 9, this includes the true locations and erroneous locations A and B. One solution is to add a third view angle that would not see a source in either the A or B direction. Alternately, if A and B can be excluded from the reconstruction volume because they do not correspond to possible source locations, then one achieves the same end result without increasing the number of basis points for the third view angle. This is shown in Fig. 10 where two sources imaged from only two directions incorrectly reconstruct to three locations. However, when the solution is constrained to lie on the surface of the L-shaped source holder and a back wall only 5 cm behind the ghost location, most of the source strength correctly returns to the true source locations.

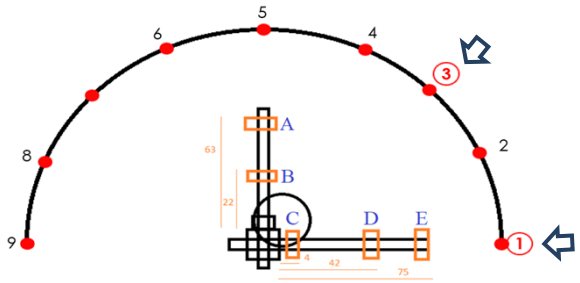


Fig. 7. Measurement setup for comprehensive 3D data. Sources can be placed at any of the labeled orange rectangles (A-E) with data collection locations at a radial distance 1.5 m from the center of the source holder (approximately where the arms cross). The nine measurement locations cover 180° and are uniformly spaced at 22.5°. The sample MLEM results shown in Figs. 8 and 10 are reconstructed using data from locations 1 and 3 as indicated by the blue arrows.

While the MLEM results obtained so far are promising, yielding source locations and quantities that are close to those used in the measurements, they are still outside the accuracy we hope to achieve during the project. For the moment, the discrepancies are attributed to the fact that the system response matrix used in the calculations was based on the results of the original Monte Carlo simulations that did not include the individual detector distortions as reported above.

In an effort to understand the limits of the MLEM reconstruction technique, a study is being conducted that removes the detector performance issues by using both simulated data and basis sets. That work is continuing into FY 2024 and is studying such things as the number of iterations, the source size and shape, the resolution of the imager, the data integration time, and the reconstruction voxels, as well as the impact of background. Preliminary results show that developing a convergence criterion to limit the number of iterations is important because with a large number of iterations, although the overall source

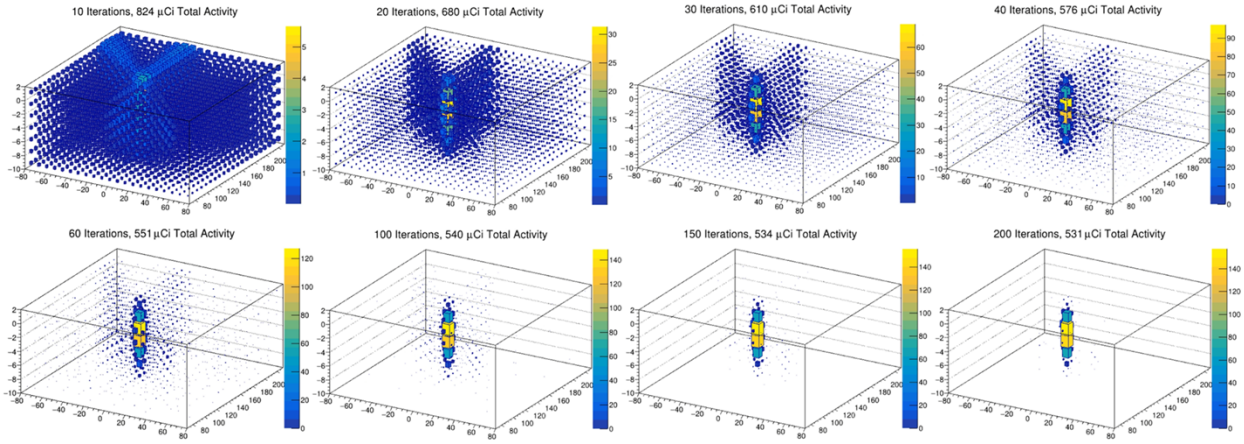


Fig. 8. MLEM reconstruction from two angles of a single point source. From left to right, the upper row shows the images obtained after 10, 20, 30, and 40 iterations. Similarly, the bottom row shows the results after 60, 100, 150, and 200 iterations. The cylindrical nature of the source is an artifact of the exaggerated scale of the z-axis

intensity is preserved, the intensity across a uniform source is not maintained. The source takes on a mottled form with hot and cold pixels developing at seemingly arbitrary locations. The extent of the non-uniformity also seems to be impacted by the resolution of the reconstruction volume compared to that of the native resolution of the imager.

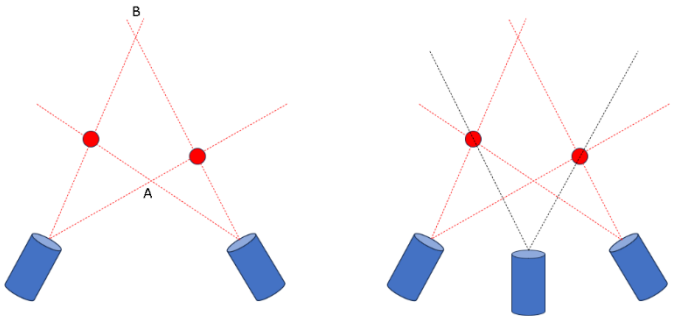


Fig. 9. Left, two sources (red disks) imaged from two locations could have sources located at the the true source locations or where lines connecting each source to each imager intersect (locations A and B). Right, adding a third imager removes the degeneracy. Alternately, one could restrict the voxels included in the image reconstruction to disallow location A and B if they are not associated with a surface or an enclosed volume.

Overall, the only concerning issue uncovered at the time of this reporting is a modest loss of source counts to background counts when extended sources and hot pixel backgrounds are included. This loss is proportional to the size of the source, with the strength of larger sources being consistently under reported. The origin of this source "leak" presumably originates in the MLEM characteristic of sharing radiation strength to any location commensurate with the data. For a point source, the response of the mask and anti-mask data are opposite at all locations across the detector (Fig. 11). However, for extended sources, regions of both the mask and anti-mask detectors have equal numbers of counts. In those regions, counting statistic variations allow leaking source counts to background counts, which also manifest as equal counts in the mask and anti-mask detector pixels. Originally detected when only hot pixel backgrounds were included, it appears that this effect is less important if a more realistic continuous background is also included. This effect will be further studied early in FY 2024.

As an alternate to MLEM, a study of 3D reconstruction has also begun using linear-regression algorithms. Specifically, the source distribution across the field of view will contribute to the detector based on an intensity scaled version of the shadow patterns in the system response matrix. The best image will be the linear combination of intensity-scaled shadow patterns that most closely fit the observed detector pattern. An immediate advantage of this approach is that it does not include a positivity constraint so that the basis functions can be made of the mask minus anti-mask patterns generated from each possible source location. This removes concerns about background handling and significantly reduces the number of basis sets and data locations since a single mask minus anti-mask response replaces the separate mask and anti-mask responses used by MLEM, while the linked mask and anti-mask background terms are also not required. In principle, the technique should also be computationally less intensive since one does not have to repeatedly project between image and detector space. The system is sufficiently complex that a direct solution is not feasible and different iterative minimizers are being explored to generate the best fit to the data.

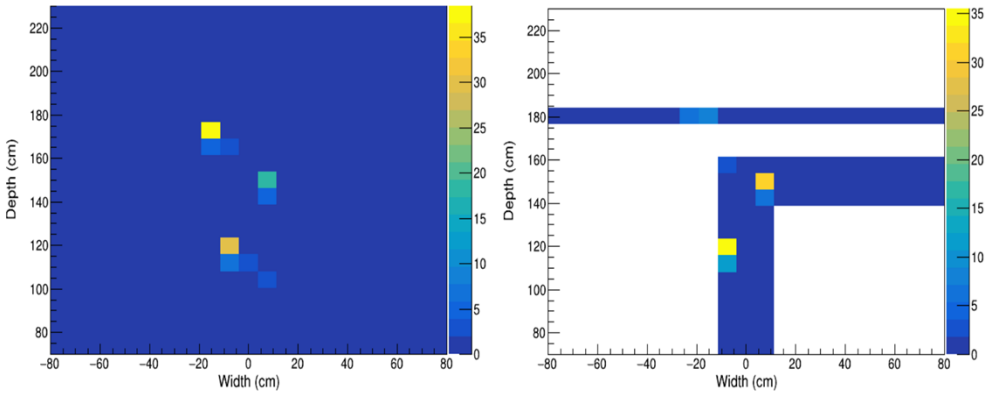


Fig. 10. Top-down view of an MLEM reconstruction of a two-source configuration. Left, using the full reconstruction volume and right, using a reconstruction volume including only the L-shaped source holder and an artificial back wall at 180 cm. Despite the proximity of the ghost source (~ 20, 175) to the back wall at 180 cm, it is almost completely absent from the source-limited reconstruction.

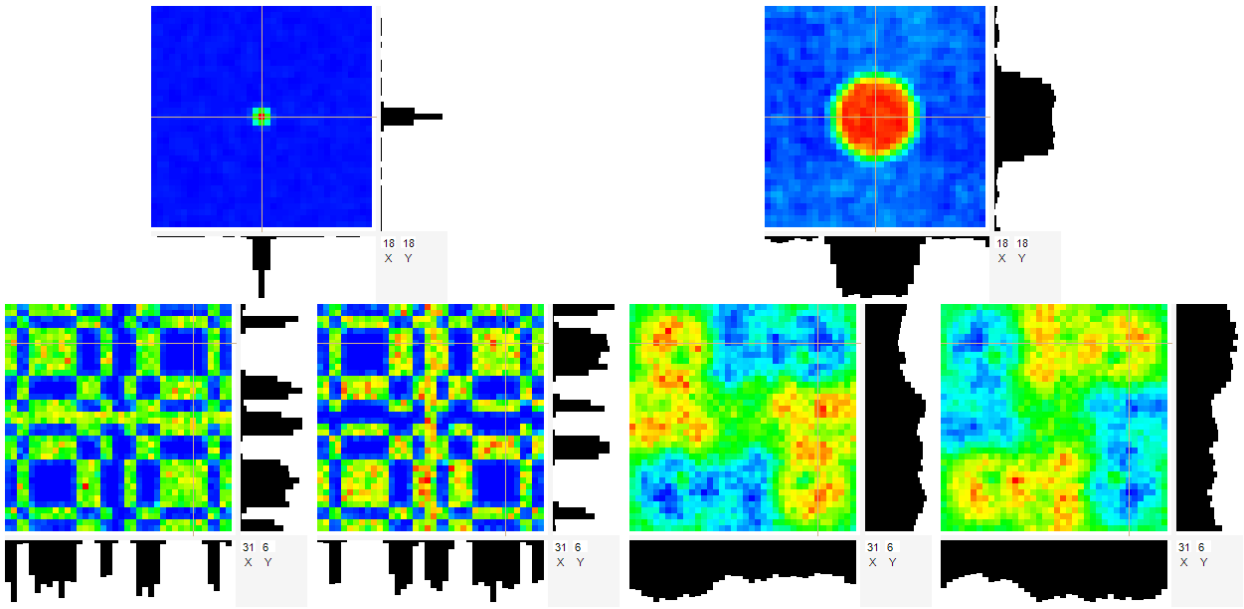


Fig. 11. Idealized simulations of a point source (left) and an extended source (right) showing how the detector (bottom row) changes for mask (left and third panel) and anti-mask (second and right panel) configurations. In the point source there are only counts in one region or the other, while for the extended case there is a global offset, similar to a uniform background illumination. In the regions with green counts, the two have similar count rates where small non-uniform backgrounds cannot be distinguished from structure induced by the source.

Regardless which reconstruction technique is ultimately selected for use, an accurate imager response matrix function is required. While the overall progress on developing such a response is discussed in 2.2.1 above, there is an additional complication that must be considered in applying the response. Specifically, the precomputed matrix comprises the response for a source located in each of the set of voxels in front of and aligned to the imager axes. However, in a general measurement it is unlikely that those precomputed voxels will exactly match those derived from voxelizing the scene, particularly if more than one imager pose is used, because that would require careful coalignment of the multiple poses. Since obtaining accurate response functions will be computationally intensive, it will not be viable to recompute the response matrix for every measurement. If a direct imager (such as a camera) were used, the simple solution would be to derive the needed response matrix by interpolating the response at the actual voxels from the nearby precomputed locations so that they match a given pose. The situation appears more complicated for coded-aperture response matrices due to the indirect nature of coded-aperture imaging. However, a simple solution has been developed and tested with the MLEM algorithm. It is based on the precepts that small changes in a voxel's location at a given distance correspond to a shift in the phase of the shadowgram observed by the detector, while small changes in distance correspond to small shifts in the magnification of the pattern.

The computation is performed for every scene location $Q(x_2, y_2, z_2)$ using the nearest precomputed response, $P(x_1, y_1, z_1)$. The objective of the interpolation algorithm is to use the response from P to obtain the appropriately shifted and scaled response at Q. From the detection location and the original source location P, the equations of a line in 3D allow us to calculate the location in the mask that the gamma ray passed through. The new gamma-ray detection location can then be calculated using the source location Q

and the same mask location calculated in the previous step. Shifting these gamma detections within the detector crystal produces a detector response function for the new source location.

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}$$

This procedure is performed for both imager poses and repeated for each point provided by the contextual information. The final result is a basis set consisting of the multi-pose detector response to individual point sources regularized to the measurement environment based on contextual information.

3. SUMMARY

The project has made significant strides toward providing contextually enabled, quantitative, 3D gamma-ray images. In particular, the long-standing discrepancies between simulated and laboratory-based imager efficiency maps have been resolved by mapping the spatial distortions of individual detectors. Going forward, such maps will need to be generated by the manufacturer as part of their quality and assurance program. 3D coded-aperture quantitative images have also been successfully generated using an iterative, MLEM-based reconstruction algorithm. The overall performance of that approach will be explored early in the next Fiscal Year as we can now generate and apply improved simulated detector response maps to each imager. Finally, uncertainties associated with the contextual information approach due to staffing issues at one of the partner laboratories have been resolved by using the manufacturer's newly developed contextual capabilities, simplifying the algorithmic transfer requirements of the project.

4. REFERENCES

1. PHDS, Inc. GeGI imager, 3011 Amherst Road, Knoxville, TN 37921, phdsco.com/products/gegi
2. H3D, Inc., H420 imager, 812 Avis Dr. Ann Arbor, MI 48108, h3dgamma.com/h100.php
3. K.P. Ziock, M. Blackston, I. Garishvili, K. Schmitt, R. Venckataraman, "Quantitative Holdup Determination Using Coded-Aperture Gamma-Ray Imaging, Proc. INMM & ESARDA Joint Annual Meeting, Vienna, Austria (Virtual), 2021.
4. Google Cartographer information available from: google-cartographer.readthedocs.io/en/latest/
5. K. Knecht, et al., "Scene-informed optimal measurement positions for quantitative safeguards measurements," Proc. INMM & ESARDA Joint Annual Meeting, Vienna, Austria, 2023.
6. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," IEEE Trans on Evolutionary Computation, vol. 6, no. 2, pp. 182–197, 2002.
7. V. Schönfelder, A. Hirner, K. Schneider, "A Telescope for Soft Gamma-Ray Astronomy," Nucl. Inst. Meth., **107**, 385-394, 1973.
8. L. Parra and H. H. Barrett, "List-Mode Likelihood: EM Algorithm and Image Quality Estimation Demonstrated on 2-D PET," IEEE Trans. on Med. Imaging, vol. 17, no. 2, pp. 228–235, 1998.
9. D. Hellfeld, et al., "Quantitative Compton Imaging in 3D," Proc. INMM 63rd Ann Meeting, July 24-28, 2022, virtual.
10. D. Xu and Z. He, "Gamma-ray Energy-Imaging Integrated Spectral Deconvolution," Nucl. Inst. Meth., **A574**, 98-109, 2007.
11. D. Mitchell, J. Mattingly, "Rapid computation of gamma-ray spectra for one-dimensional source models," Trans. Am. Nucl. Soc. **98**, 565-566, 2008.
12. S. Agostinelli et al., "GEANT4—a simulation toolkit," Nucl. Instr. Meth., **A506**, 250-303, 2003.
13. R. Gottesman, E.E. Fenimore, "New family of binary arrays for coded aperture imaging," Applied Optics, **28**, 4344-4352, 1989.

14. E.E. Fenimore, T.M. Cannon, “Coded aperture imaging with uniformly redundant arrays,” Applied Optics, **17**, 337-347, 1978.
15. E.E. Fenimore, “Coded aperture imaging: predicted performance of uniformly redundant arrays,” Applied Optics, **17**, 3562-3570, 1978.
16. J. Braga, et al., “A New Mask-Antimask Coded-Aperture Telescope for Hard X-Ray Astronomy,” Exp. Astro., **2**, 101-113, 1991.
17. S.V. Venkatakrishnan, et al, “A Model Based Iterative Reconstruction Algorithm for High Angle Annular Dark Field-Scanning Transmission Electron Microscope (HAADF-STEM) Tomography,” IEEE Trans. Image Proc., **22**, 4532-4544, 2013
18. S.V. Venkatakrishnan, K.A. Mohan, A.K. Ziabari, C.A. Bouman, “Algorithm-Driven Advances for Scientific CT Instruments,” Signal Processing for Advanced Materials, 32-43, January 2022, DOI: 10.1109/MSP.2021.3123594.
19. K.P. Ziock, et al., “Efficiency Calibration of the H3D H420 Gamma-Ray Imager,” Joint Annual INMM & ESARDA Meeting, Vienna, Austria, May 2023.
20. J.L. Tain, C. Cano-Ott, “Alogrithms for the analysis of B-decay total absorption spectra,” Nucl. Inst. Meth. in Phys. Res., **A571**, pp. 728-738, 2007, doi:10.1016/j.nima.2006.10.098.
21. E.M. Brubaker, “A Maximum Likelihood Expectation Maximization Iterative Image Reconstruction Technique for Mask/Anti-mask Coded Aperture Data,” Proc. IEEE Nucl. Sci. Symp. and Med. Imaging Conf., 2013.