Simulation Study of Muon Scattering For Tomography Reconstruction

Debasis Mitra, *Senior Member, IEEE*, Avishek Banerjee, Sammy Waweru, Scott White, Kondo Gnanvo, *Member, IEEE*, Marcus Hohlmann, *Member, IEEE*

Abstract: **Nuclear materials that pose a homeland security threat typically have high atomic numbers (Z > 82). It is of vital importance to develop smart, efficient, and inexpensive systems to detect such high-Z materials without opening a container. Muon Tomography (MT) provides a non-invasive channel for such investigation. We have been investigating such muon scattering with numerical simulation with GEANT4 for some time [6]. In this article we report the development of an efficient clustering algorithm for detecting threat objects in a probed volume.**

INTRODUCTION

Muons produced by primary cosmic rays at the upper atmosphere provide an excellent source as passive probes for discriminating materials with different atomic numbers (Z), without any extra radiation or incurring any extra cost for the probe generation [5]. Highly penetrating muon tracks may suffer from multiple scatterings by Coulomb interaction with nuclei of atoms on its path. The angle of scattering depends on the charge Z of the corresponding nucleus [3] (Fig. 1). The incoming and outgoing tracks for each muon may be detected by appropriate sensors.

Muon Tomography (MT), based on the measurement of multiple scattering of atmospheric cosmic ray muons traversing cargo or vehicles, is a promising technique for Homeland security application of nuclear contraband detection. Many groups around the world are studying and developing MT systems. For the purpose of reconstruction the standard iterative algorithms, where the probed volume is voxelized, are adapted from medical imaging [7, 10]. These algorithms often does not run in real time and tend to be slow for the homeland security purpose.

However, since in this application both the incoming and outgoing track of a scattered muon is known, a special feature of the underlying scattering event in MT is the existence of a point of closest approach (*poca*) between the two tracks, which represents the scattering region quite well. A geometrical algorithm (POCA-algorithm) that detects *poca* points and their corresponding angle of scattering is surprisingly efficient in identifying most of the metallic targets in a probed volume of a MT station. Expectation Maximization algorithms based on maximum likelihood estimation [9] are very useful in estimating scattering parameter of each voxel, but have slow convergence and typically computes the parameters for all voxels even when many of them are not of interest.

In this work we studied the muon scattering properties for possible target materials and tried to address the question on how to effectively utilize the *poca* points generated by the tracks. We have used GEANT4 simulation software, enhanced with the CRY package as the source for cosmic ray generated muons, for this purpose. Empirical studies of muon scattering from a target, with a MT-station like geometry, is essential in developing tomography systems and reconstruction algorithms. One of our objectives is to develop efficient real-time algorithm for the reconstruction purpose for MT systems. We expect our work will make a broader impact on some transmission tomography applications including those used in medical areas.

EXPERIMENTAL SET UP

A typical geometry for our simulation using GEANT4 software [2 Allison] is shown in Fig. 1. In our simulations, *z*-axis is the vertical direction, *x*-axis is the axial direction (cargo movement), and *y*-axis is perpendicular to the *xz*-plane. Sensor arrays parallel to the *xy*-plane (typically three) are deployed above and below the probed volume. The CRY Monte Carlo generator [11] produces muons over a 5×5 m^2 square plane. Underneath this CRY plane a top and a bottom detector array - each comprising 3 detector planes - sandwich a "floating" Uranium target cube of volume $40\times40\times20$ cm^3. The detector planes have an area of 4×4 m^2 and encloses the probed volume of 3m height in the *z*-direction.

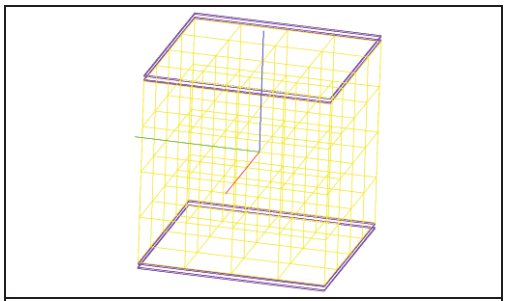

Figure 1: Geometry of the probed volume

The geometric *POCA-algorithm* [8] is used to locate the point-of-closest-approach (*poca*) between the ingoing and outgoing vectors of a scattered muon track (Fig. 2). All unscattered tracks are ignored by this algorithm. For each such

Manuscript received November 14, 2009.

D. Mitra, S. Waweru, and S. White are with the Department of Computer Science, Florida Institute of Technology, Melbourne, FL USA.

A. Banerjee is with Indian Institute of Technology, Kharagpur.

K. Gnanvo, and M. Hohlmann are with the Department of Physics and Space Sciences, Florida Institute of Technology, Melbourne, FL USA (telephone: 321-674-7737, e-mail: dmitra@fit.edu).

poca point we also note the corresponding scattering angle (between the two vectors), and the distance-of-closest-approach (*doca*).

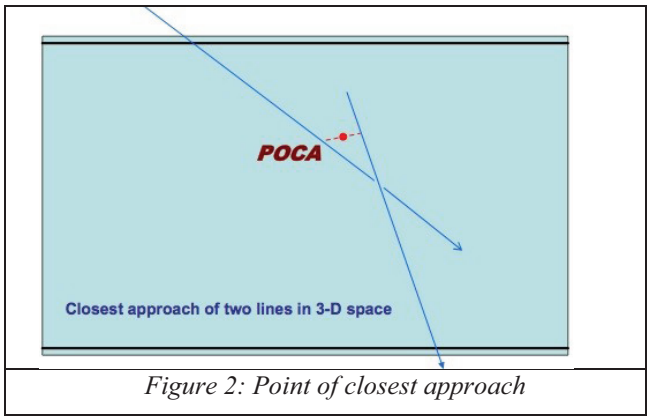

Figure 2: Point of closest approach

We use the *doca* values to filter more tracks. Higher *doca* values are considered to have scattered over a much larger region (for multiple scattering events). So, the relative confidence that the corresponding *poca* actually represents the scattering region is low for a large *doca* value. In order to test this hypothesis we collected unfiltered *poca* points and filtered *poca* points inside the Uranium target in this experiment. We calculated the *scattering density* λ (*definition:* square of the variance of the scattering angles over the collection of *poca* points per unit length of path in the target, approximated as per unit height of the target, for muons with nominal momentum, in miliradian square per mm) for each collection (Table 1). The unfiltered value is unreasonably higher - possibly because it involves more scattering events per *poca* than that for the filtered ones, as the large values of *doca* would suggest. In our subsequent works utilizing *poca* we have eliminated tracks with *doca* values greater than 1 mm. We have also not considered those tracks whose *poca*'s are outside the probed volume as they cannot represent the actual scattering region.

	Inside Target	Outside Target
DOCA filtered (1.0mm)	203.404900	3.859259
Unfiltered	808.751000	8.218831

Table 1: Scattering densities with and without doca cut

SCANNING STUDY WITH SLABS

Preliminary results from our study are shown below (Figures 3 and 4). We discretize the probed volume into slabs perpendicular to each axis. We studied by varying the width of each slab. In each slab we collect all the valid *poca's* (after the filtering of *poca's* as mentioned above) and compute the average scattering angle. The plots below show the values of average scattering angles (in degrees) against the mid-position of the slabs on the corresponding axis. Fig. 3 shows the plots over the *z*-axis or height, for three slab sizes 5cm, 3cm and 1 cm respectively. The target is actually located between (-10cm, +10cm) on the *z*-axis. Graphs only near the actual target position are shown here. All three plots can discriminate the target. Please note that the smaller the slab size more the computation time for a scanning algorithm is. An interesting phenomenon in these graphs is that a good number of high scattering-angle poca's converge just below the target, making it difficult to isolate the lower boundary of the target with high precision. This is an unexpected observation and could be an indirect cause of non-Gaussian (high-tailed) nature of the distribution of scattering angle known in the literature [3 Bethe], or alternatively an artifact from the simulation using GEANT4.

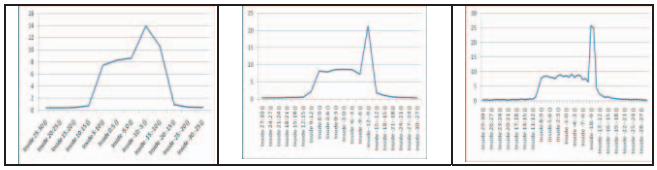

Figure 3: z-axis variation, 5, 3, 1 cm slab-sizes

Fig. 4 shows similar plots with slabs sliding along the *x*-axis. The target is located between (-20cm, +20cm) in this direction. The boundary discrimination with 3cm could be a good compromise for optimum slab-size. The precision of boundary appears good in these plots. For symmetry reason plots along *y*-axis is not shown here, but they are very similar in nature as that along the *x*-axis.

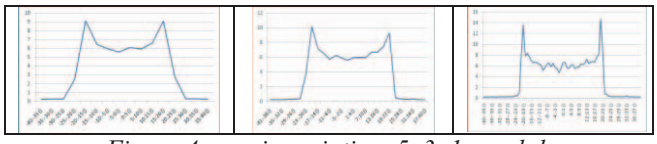

Figure 4: x-axis variation, 5, 3, 1 cm slab

For a very simple scenario as in the above experiment, a linear slab-scanning algorithm over the *poca* points may suffice for detecting a threat object. However, in any complicated real life scenario simple scan along three axes will not be enough. An example of a three-target (iron, tungsten, uranium) are shown in Figure 5. Discerning capability as well as localizing capability significantly reduces with the complicacy of the scenario.

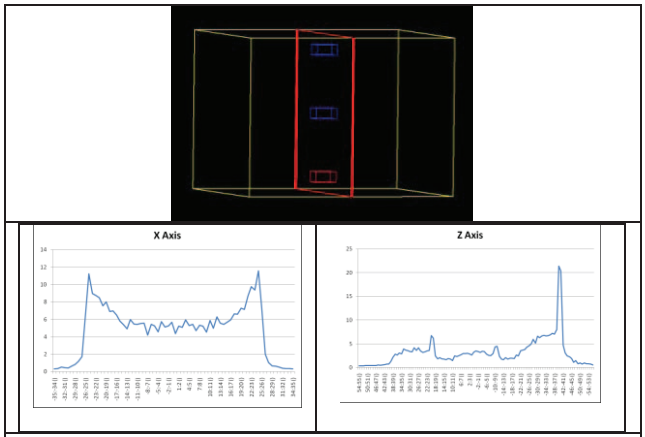

Figure 5: Scan over a three target scenario

The purpose of our above study was to mainly elucidate the power of local identification capability of *poca* points. An obvious next step is to go beyond axis-wise scan and cluster the *poca* points regionally. For this purpose, we developed a density based clustering algorithm and studied somewhat more complicated multi-target scenarios with it.

POCA CLUSTERING STUDY

Our previous studies with *poca* points [6] have shown that they tend to form dense clouds around high density or high atomic number materials. Also, higher atomic number materials tend to have more scattering angles. The scattering property of a material may be characterized by the *scattering density* λ as defined before that relates to the variance of the scattering angle distribution over the respective material (also inversely related to the *radiation length*). From these observations we developed the unsupervised density-based clustering algorithm that we call *POCLust* algorithm (Figure 6). POCLust returns dense region of *poca* points that are related to each other by respective scattering angles. As a cluster grows in size starting from a single *poca*, the variance of the scattering angles as well as the spatial density of *poca* points within a cluster is kept track of as the cluster quality parameter. POCLust is a real-time algorithm working with each track sequentially, thus having a linear time complexity (with respect to the number of tracks) and a low memory requirement. The asymptotic time-complexity is $O(m|C| + |C|^2)$, for m tracks and $|C|$ number of clusters.

Input: Geant4 output (list of all muon tracks and associated parameters)

1. *For each muon track {*
2. *Calculate the POCA pt P and scattering-angle*
3. *if (P lies outside container) continue;*
4. *Normalize the scattering angle (angle*p/3GeV).*
5. *C = Find-nearest-cluster-to-the (POCA pt P);*
6. *Update-cluster C for the new pt P;*
7. *After a pre-fixed number of tracks remove sporadic-clusters;*
8. *Merge close clusters with each-other }*
9. *Update λ (scattering density) of each cluster C using straight tracks passing through C*

Output: Clusters as regions of interest (ROI)

Figure 6: POCLust algorithm for clustering poca points

POCLust initially ignores non-scattered tracks. However, in stand-alone mode it needs to make a second scan over the tracks to include all non-scattered tracks passing through each cluster (identified in first pass) in order to update the *scattering density* λ computation, as the latter value would be very high if the zero-angle non-scattered tracks are ignored. POCLust may also be used in a pre-processing mode where only the clusters are output as regions of interest (ROI) for further analysis or for subsequent reconstruction purpose. In this mode only the first real-time pass is enough.

The main challenge in the algorithm is to avoid comparing each *poca* with all other ones. Also, cluster explosion needs to be tackled by filtering unpromising clusters from time to time (line 7, typically after processing each 5000 events).

In the following Figures 7-11 we show some of our results in using POCLust over multiple target scenarios. The scenario plot from GEANT4 is on the left hand side and corresponding clusters are shown on the right of each figure. Color of the cluster indicates the lambda value of each corresponding cluster after the second pass over the *poca* points (line 9, Algorithm in Figure 6).

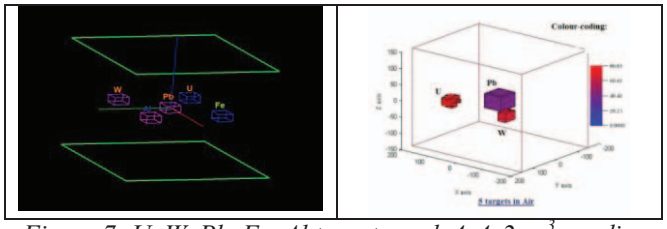

Figure 7: U, W, Pb, Fe, Al targets each 4x4x2cm^3, medium is air, 10 million total tracks equivalent to 10 min of exposure

POCLust parameters are adjusted to detect three of the targets U, W and Pb (Fig. 7), and these parameters are used in all subsequent scenarios. We could distinguish between all three of the targets, where distinction between U and W using iterative reconstruction algorithms remain a challenge in the literature. The algorithm takes 29.2 sec run time on a PC. Minimum number of points inside a cluster must be 100, and minimum distance between the centers of any pair of clusters must be 250mm. Scattering angles are normalized with corresponding momentum values associated with the tracks.

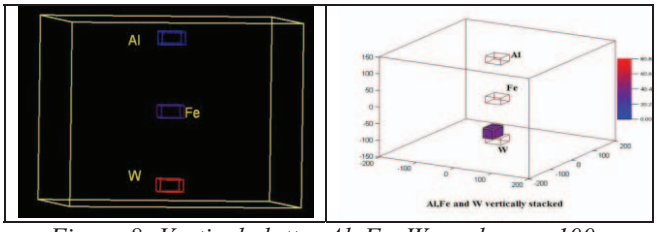

Figure 8: Vertical clutter Al, Fe, W, each gap=100cm

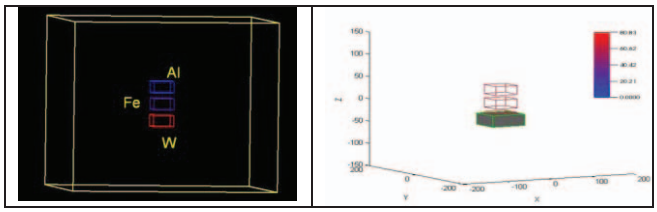

Figure 9: Vertical clutter Al, Fe, W, each gap=10cm

In the two above vertical clutter scenarios (Figs. 8 and 9) only W could be detected, and detection of Fe and Al are not possible with the POClust parameters that we are using. False positives (clusters) will increase significantly if we readjust the parameters to detect low-Z materials like Fe and Al. However, the proximity effect of the two low-Z targets above W box in Fig. 8 makes the cluster corresponding to the latter

much larger and with lower λ-value as opposed to that in Fig. 9.

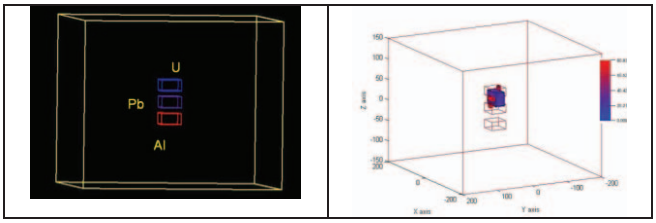

Figure 10: Reverse vertical clutter U, Pb, Al, each gap=10cm

This reverse vertical clutter scenario with small gap between each pair of targets (Fig. 10) is a classic example where reconstruction algorithms may be challenged because the higher-*Z* material above lower-Z ones may produce a "shadow" effect over the lower-*Z* materials. Even the *poca* points' cloud for the higher-Z target gets somewhat confusing and a slabbing-scan fails to detect the threat. Attempts to detect such scenarios may result in higher false positives.

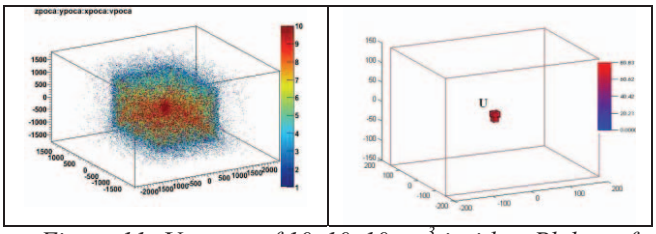

Figure 11: U target of 10x10x10 cm^3 inside a Pb box of 200x200x200 cm^3 with 10 cm thickness, surrounding medium is vacuum

Here the left side is *poca* visualization with color representing each *poca* point's scattering angle (Fig. 11). POCLust clears up the cloud and detects the threat object at the center. This is an example closer to reality (as a threat object is likely to be hidden in a box of some dense material like lead), where purely visualization of *poca* may not suffice, and some type of quantitative study (e.g., done with POCLust) is necessary.

POCLUST AS PREPROCESSOR

In this section we address the usefulness of POCLust as a preprocessor to a standard iterative reconstruction: *maximum likelihood expectation maximization* (MLEM) algorithm [7]. MLEM typically runs over the whole probed volume. In our case it is the volume surrounded by the detector arrays. It needs the volume to be discretized over voxels of pre-determined size (Fig. 1) and produces value of λ over each voxel over the iterations. As the targets are likely to be limited in space in some regions within this probed volume, the algorithm wastes significant amount of computing resources. Here, the POCLust preprocessing may hypothesize the *volume of interest (VOI)* as a minimum rectangular bounding box over the clusters it produced. The non-scattered tracks are not required for VOI detection, and so, a single pass algorithm (as mentioned before) is sufficient. MLEM becomes much more time and memory efficient as the volume over which it runs reduces. The following (Fig. 12) shows the result of such an experiment over the five-target scenario (as in Fig. 7). The VOI is 235x235x45 cm^3 in volume (19881 voxels, each voxel of size 5x5x5 cm^3) as opposed to the whole probed volume of 400x400x300 cm^3 (with 384000 voxels), almost 95% probed space reduction, resulting in 80% in time reduction. As seen in Fig. 12, the Pb is not well detected after MLEM run, but the two high-Z targets are clearly identified.

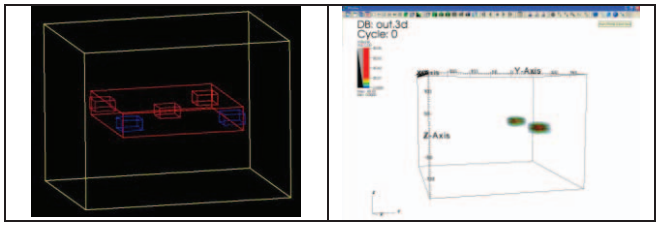

Figure 12: MLEM reconstruction over Volume of Interest extracted by POCLust

CONCLUSION

In this work we have developed an efficient density-based clustering algorithm for detecting high-Z targets with cosmic-ray generated muon tracks scattered from the targets within a probed volume. Results from numerical simulation using phantoms are presented here. The algorithm may also be used as a pre-processor to significantly reduce the computational resource requirements of a standard iterative statistical algorithm.

An open question remains if ever cosmic ray generated muons may be used for medical purpose, because of its low flux rate, even though its free source is very attractive. Also, the scattering of muons from low-Z biological materials may not be significant enough. To initiate some thought in this direction we simulated the scattering in low-Z materials like water and calcium. A simple *poca* plot from such a simulation is shown in Fig. 13.

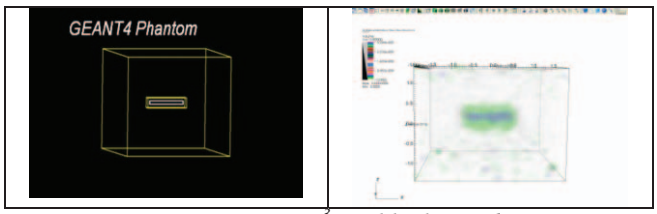

Figure 13: A 130x10x10 cm^3 Ca block, inside a 150x30x30 cm^3 H_2O block, in air background
20 million muon tracks, equivalent to 20 minutes of exposure

ACKNOWLEDGMENT AND DISCLAIMER

This material is based upon work supported in part by the U.S. Department of Homeland Security under Grant Award Number 2007-DN-077-ER0006-02. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or

implied, of the U.S. Department of Homeland Security. Patrick Ford has provided some computational support. Banerjee and Waweru were supported by NSF.

REFERENCES

[1] Agostinelli, S., and one hundred twenty nine others, (2003) "GEANT4 – a simulation toolkit," *Nucl. Instrum.Meth. A,* 506, 250-303.

[2] Allison, J., Amako, K., Apostolakis, J., Araujo, H., Dubois, P., Asai M., and others (2006) "GEANT4 developments and applications," *IEEE Trans. Nuclear Sc.,* 53(1), 270-278.

[3] Bethe, H. (1953) "Moliere's theory of multiple scattering," *Physical Review*, 89(6), 1256.

[4] Hagmann, C., Lange, D., and Wright D. (2007) "Cosmic-ray shower generator (CRY) for Monte Carlo transport codes," *Proc. IEEE Nucl. Sci. Symp.,* Honolulu, HI, 2, 1143-1146.

[5] Hogan, G.E., et al., (2004) "Detection of high-Z objects using multiple scattering of cosmic ray muons," AIP Conf. Proc., vol. 698, , pp. 755-758, presented at 8th Conference on the Intersections of Particle and Nuclear Physics (CIPANP 2003), New York, NY, 2003.

[6] Hohlmann, M., Ford, P., Gnanvo, K., Helsby, J., Pena, D., Hoch, R., and Mitra, D. (2009) "GEANT4 Simulation of a Cosmic Ray Muon Tomography System with Micro-Pattern GasDetectors for the Detection of High-Z Materials," IEEE Transactions on Nuclear Science, vol. 56, no. 3, pp. 1356-1363, June 2009.

[7] Schultz, L. J., Blanpeid, G. S., Borozdin, N., Fraser, A. M., Hengartner, N. W., Klimenko, A.V., Morris, C. L., Orum, J. C., and Sossong, M. J. (2007) "Statistical reconstruction for cosmic ray muon tomography," IEEE Trans. Image Processing, 16(8), pp. 1985-1993.

[8] Sunday. D. (2006) "Distance between Lines and Segments with Their Closest Point of Approach." *At*: http://geometryalgorithms.com/Archive/algorithm0106/algorithm0106.htm

[9] Verdi, Y., Shepp, L. A., and Kaufman, L. (1985) "A statistical model for positron emission tomography." *l. Of American Statistical Association*, 80(389), 8-20.

[10] Wang, G., and Qi, J. (2008) "Statistical image reconstruction for muon tomography using Gaussian scale mixture model." ICIP-2008, pp. 2948-2951, 15th IEEE International Conference on Image Processing, San Diego, CA.

[11] Wright D., and others from the Cosmic-ray Physics Team at the Lawrence Livermore National Laboratory (2006) "Monte Carlo Simulation of Proton-induced Cosmic-ray Cascades in the Atmosphere," Lawrence Livermore National Lab., CA, *Tech. Rep. LA-UR-06-8497*.