Geo-EM-ML: Geometrically Constrained EM-ML Algorithm for Multi-Source Localization in Pipeline Systems

Authors: Zijia, Dr. Kuang, wang, Ms. wei, Bi, Dr. Yuanjie, Dr. Rui Qiu, Wang, Chuan-Gao, Dr. Senlin Liu, Bi, Dr. Yuanjie

Date: 2025-03-11T00:00:00+00:00

Abstract

In nuclear facility maintenance, accurately locating multiple radiation sources within complex pipeline systems is paramount for ensuring the safety of maintenance personnel and optimizing work routes to minimize radiation exposure. However, existing localization methods often rely on two-dimensional grid spaces or are limited to a few sources, and they frequently neglect the attenuation effects caused by pipeline materials, leading to reduced accuracy and increased computational complexity. This study introduces the Geo-EM-ML algorithm, a novel hybrid approach that integrates the Expectation-Maximization (EM) algorithm and Maximum Likelihood Estimation (MLE), enhanced by geometric constraints tailored to the three-dimensional continuous space of pipeline systems. Additionally, a ray tracing program is employed to model the attenuation effects accurately, ensuring reliable detector response calculations. Experimental results demonstrate that Geo-EM-ML achieves a high success rate exceeding 94% in scenarios with up to six radiation hotspots, maintaining average position error and relative activity error below 8.58 mm and 3.48%, respectively. The algorithm exhibits robustness across varying pipeline shielding materials, wall thicknesses, source intensities, and detector configurations. The Geo-EM-ML algorithm represents a significant advancement in multi-source localization, offering a scalable and precise solution for complex pipeline environments in nuclear facilities, thereby mitigating safety risks and optimizing maintenance workflows.

Full Text

Preamble

In nuclear facility maintenance, accurately locating multiple radiation sources within complex pipeline systems is paramount for ensuring the safety of main-

tenance personnel and optimizing work routes to minimize radiation exposure. However, existing localization methods often rely on two-dimensional grid spaces or are limited to a few sources, and they frequently neglect the attenuation effects caused by pipeline materials, leading to reduced accuracy and increased computational complexity. This study introduces the Geo-EM-ML algorithm, a novel hybrid approach that integrates the Expectation-Maximization (EM) algorithm and Maximum Likelihood Estimation (MLE), enhanced by geometric constraints tailored to the three-dimensional continuous space of pipeline systems. Additionally, a ray tracing program is employed to model the attenuation effects accurately, ensuring reliable detector response calculations. Experimental results demonstrate that Geo-EM-ML achieves a high success rate exceeding 94% in scenarios with up to six radiation hotspots, maintaining average position error and relative activity error below 8.58 mm and 3.48%, respectively. The algorithm exhibits robustness across varying pipeline shielding materials, wall thicknesses, source intensities, and detector configurations.

The Geo-EM-ML algorithm represents a significant advancement in multi-source localization, offering a scalable and precise solution for complex pipeline environments in nuclear facilities, thereby mitigating safety risks and optimizing maintenance workflows.

Keywords: Multiple sources localization, Maximum Likelihood Estimation, Expectation-Maximization Algorithm, Shielding scenario

Introduction

Complex pipeline systems containing radiation hotspots present a significant threat to maintenance and experimental personnel in nuclear facilities. Additionally, work routes cannot be planned without knowledge of the hotspots' locations and activity, resulting in increased radiation exposure to personnel.

Currently, there are two methods for obtaining radiation hotspot information: (1) estimation based on staff experience and (2) direct measurement near the pipeline or sampling from the pipeline. The first method can lead to significant deviations due to the subjective nature of estimates. The second method, while more direct, exposes personnel to potentially high radiation doses. Additionally, due to spatial constraints, many locations may be inaccessible to close-range detectors. Therefore, developing a multi-source localization algorithm for complex pipeline systems helps workers locate radiation sources quickly and accurately.

Multi-source localization is an inverse problem that aims to determine the location and activity of sources based on the count data collected by a group of detectors positioned around the sources. In contrast to the search for a lost single source in nuclear security scenarios, multi-source localization algorithms are inherently more complex and challenging to develop. Currently, several algorithms are used to solve the multi-source localization problem, mainly including least squares estimation, maximum likelihood estimation/expectation maximization, Bayesian estimation, and neural network.

a. Least Squares Estimation: Chen et al. proposed a multi-source localization algorithm based on the least squares method, which can locate up to two radiation sources in a three-dimensional environment. They designed a screening mechanism to determine whether the detector participates in the calculation through geometric relationships, which improves the localization accuracy and algorithm robustness in the presence of shielding[1].

b. Maximum Likelihood Estimation/Expectation Maximization: Deb modeled the multi-source localization problem as an optimization problem of a high-dimensional function and solved the maximum likelihood estimation (MLE) through Fisher score iteration. Since the multi-source localization problem is a non-convex optimization problem, finding a reasonable initial estimate is an inherent need. Deb proposed an expectation-maximization-based algorithm to find the approximate distribution of the source intensity in space. Then, the local maxima of the spatial distribution are identified as the initial estimates for the maximum likelihood estimation[2]. In addition, Deb et al. implemented the mixed use of directional and non-directional detectors under the expectation-maximization framework so that the radiation source localization work can consider both cost and efficiency[3]. Hellfeld et al. improved the expectation-maximization algorithm and proposed the Point Source Localization (PSL) algorithm. The expectation-maximization algorithm usually divides the space into grids and estimates the radiation source activity at each grid point. The PSL algorithm treats the position and radiation source activity as continuous variables.

The PSL algorithm is suitable for the localization of a single radiation source. To solve the multi-source localization problem, Hellfeld et al. proposed the Additive Point Source Localization (APSL) algorithm. The APSL algorithm estimates the number of radiation sources by gradually adding new radiation sources and uses the Bayesian information criterion as the stopping criterion, thereby improving the robustness of the algorithm[4, 5]. Subsequently, Bandstra et al. improved the PSL algorithm to enable it to estimate the position of radiation sources accurately in the presence of shielding. They used LiDAR to build a voxel model of the scene and calculated the shielding effect through a ray tracing algorithm. They regarded the entire scene as composed of the same material and added the linear attenuation coefficient of this material as an estimated parameter to the PSL algorithm, thereby achieving better localization accuracy in practical tests than the original PSL algorithm[6]. Abdelhakim proposed heuristic techniques to enhance the efficiency of maximum likelihood estimation for localizing radioactive sources[7].

Anderson et al. proposed a new method for radiation measurement using mobile robots, which can accurately locate and characterize radioactive sources in real scenarios containing multiple radiation sources and complex environments through recursive Bayesian estimation combined with attenuation modeling. In addition, their system implemented autonomous isotope identification and measurement point optimization based on Fisher information, significantly improving measurement efficiency and accuracy[8].

c. Bayesian Estimation: Morelande et al. proposed a Bayesian estimation framework that transformed the task of estimating radiation source parameters into a model selection problem and utilized partial Bayes factors (PBF) to compare the evidence of different models for determining the number of radiation sources. Their work utilized the sequential Monte Carlo method, especially by introducing the progressively corrected importance sampling (PCIS) technique, effectively approximating the posterior distribution of radiation count data collected by detectors when multiple radiation sources exist[9–12]. Chin et al. proposed an efficient Bayesian estimation framework combining the particle filter and mean shift technique. Their proposed Bayesian framework ensures constant computational complexity, accounts for the influence of shielding, and enables parallel processing, making it applicable to large-scale sensor networks[13].

d. Neural Network: The multi-source localization methods based on neural networks are mainly divided into recurrent and feedforward neural networks. Wacholder et al. modeled the multi-source localization problem as a combinatorial optimization problem and solved it using the Hopfield recurrent neural network. The Hopfield recurrent neural network does not require pre-training of network parameters but updates the state of neurons iteratively to minimize the energy function and obtain the optimal solution[14]. Mendes et al. transformed the multi-source localization problem in two-dimensional space into an object detection problem and detected the possible positions of radiation sources from the radiation distribution map using a pre-trained convolutional neural network[15]. Abdelhakim proposed a feature extraction technique to convert the position information and count data of the detectors into feature vectors and then used the decision tree regression algorithm to learn how to predict the activity and position of the radiation sources from the feature vectors[16]. Hao et al. designed a Source Distribution Inversion Convolutional Neural Network (SDICNN) to obtain the distribution information of complex source terms from radiation parameters in space. The SDICNN comprises a fully connected network (FCN) and a convolutional neural network (CNN). The FCN obtains low-resolution source distribution parameters from a single sampling point in the radiation field, and the CNN uses a structure similar to the super-resolution CNN (SRCNN) to complete the delicate reconstruction of the source distribution[17]. Okabe et al. designed a detector pixel layout inspired by Tetris and increased the response contrast between pixels by adding inter-pixel padding materials between pixels. They used a neural network composed of filtering layers and a deep U-net to predict the direction of radiation sources. Based on the predicted direction and the incident radiation intensity, they used posterior probability maximization to estimate the position of the radiation source[18].

However, the existing research encountered challenges when applied to the localization of multiple radiation sources in complex pipeline systems. Firstly, most previous methods focused on multi-source localization in two-dimensional planes or discrete grid spaces, whereas radiation sources in complex pipeline systems are continuously distributed in three-dimensional space[2, 3, 7, 9–16]. Besides, most previous studies primarily considered attenuation from the air, whereas radiation sources in complex pipeline systems are shielded by the pipeline wall, leading to errors in estimating the radiation source position[1–5, 7–12, 14–16, 18]. Moreover, previous studies primarily focused on localizing only a few radiation sources (fewer than four), whereas complex pipeline systems typically involve a much larger number of radiation sources, exacerbating the ill-posedness of the inverse problem[1, 3, 6–12, 14–18].

This paper proposes a geometrically constrained EM-ML algorithm (Geo-EM-ML) to solve the multi-source localization problem in complex pipeline systems. At the initial stage, the expectation-maximization (EM) algorithm is used to estimate the position and activity of the radiation sources quickly. Maximum likelihood estimation (MLE) is used to optimize the initial estimation provided by the EM algorithm. During the optimization process, geometric constraints are introduced to narrow the algorithm's search space and improve the convergence efficiency. Meanwhile, a parameter filter is designed to estimate the number of radiation sources progressively. Finally, the algorithm continuously optimizes and adjusts the position, number, and activity of the radiation sources until the estimated number of radiation sources no longer changes.

The primary contributions of the paper are as follows: - The proposed hybrid algorithm can rapidly provide an approximate orientation of the radiation sources within the pipeline system and efficiently compute the number, positions, and activity of the radiation sources in 3min. Tests with up to twelve hotspots were conducted, yielding excellent results. - Our algorithm introduces geometric constraints, which narrows the search space of radiation sources and improves the robustness and convergence speed of the algorithm. - We designed a filter strategy to progressively optimize the estimation of the number of radiation sources, improving the stability and accuracy of the algorithm while reducing the computational burden in the later stages. - We used ray tracing to accurately model the radiation shielding effect, ensuring the accuracy and stability of the algorithm in complex pipeline systems. The stainless steel pipelines with a wall thickness of up to 15 cm were tested in this study, achieving a success rate of around 95%.

The rest of this paper is organized as follows: Sec. II introduces the details of the Geo-EM-ML algorithm; Sec. III tests the algorithm under different experimental settings to demonstrate the feasibility of the algorithm, and conducts ablation experiments on various sub-modules of the algorithm to illustrate their usefulness; Sec. IV summarizes the work of this paper.

II. Method

A. Preliminaries

[Figure 1: see original paper] Pipeline system with detectors and sources. The red balls represent the sources. The blue balls represent the detectors.

The problem to be solved by our algorithm is that there are multiple point sources in the pipeline system, which are composed of the same radioactive nuclide and emit gamma rays, as shown in Figure 1. The total number of these point sources M, the position r_m and the activity λ_m of each point source are unknown. Therefore, the parameter distribution to be estimated is represented as shown in Eq. (1).

 = {r_m, λm}{m=1}^M

N detectors are deployed in the free space outside the pipeline system, and their positions r_n are known. During the detection time t, the photon peak counts received by detector n are also known and denoted as C_t^n.

A well-studied forward problem is formulated as determining the count rate c_n( ) of the photon peak at the position r_n of detector n based on the distribution of radiation sources and the geometry and material of the shielding in the environment. In the case where the radiation source distribution is known, the count rate of the photon peak on the n-th detector contributed by the m-th radiation source is as shown in Eq. (2).

c_{n,m}( ) = α{n,m}λm = e^{- $\bar{\rho}${n,m} *d*{n,m}} λ_m

Where α{n,m} *is the attenuation coefficient from the m-th radiation source to the n-th detector, A is the detector area, is the detector detection efficiency,* $\bar{\rho}${n,m} is the average linear attenuation coefficient from the detector to the radiation source, and d_{n,m} = ||r_n - r_m|| is the distance between the detector and the radiation source. The detectors used in this study have consistent detector area and detection efficiency, and are isotropic detectors. Since the influence of each radiation source on the detector is independent, the count rate of the photon peak on the detector is as shown in Eq. (3).

c_n( ) = *{m=1}^M c*{n,m}( ) = *{m=1}^M* α{n,m}λ_m

The problem studied in this paper is the inverse problem of the above forward problem, i.e., determining the distribution of radiation sources based on the positions of the detectors and the count of the photon peaks {r_n, C_t^n}_{n=1}^N and the geometry and material of the shielding in the environment. The inverse problem is generally solved using optimization methods, i.e., finding the optimal parameters * in the solution space such that the difference between the estimated count rate {c_n(*)}*{n=1}^N and the actual count rate {c_n}*{n=1}^N of the photon peaks detected is minimized. The entire solution process essentially involves continuously sampling the parameters in the solution space, calculating the corresponding count rate {c_n( )}_{n=1}^N, determining the direction of the next sampling based on the estimated count rate and the actual count rate of the photon peaks detected, until the optimal parameters * are found.

When solving the inverse problem using the least squares estimation (LSE), the algorithm calculates the square difference between the estimated count rate and the actual count rate {c_n}_{n=1}^N as the loss function L( ), and selects the parameter * that minimizes L( ) as the distribution of the radiation sources, as shown in Eq. (4).

* = arg min L_{LSE}( ) = arg min _{n=1}^N (c_n( ) - c_n)^2

When solving the inverse problem using the maximum likelihood estimation (MLE), the algorithm calculates the probability of the photon peak counts received by each detector as the likelihood function, and selects the parameter * that maximizes the likelihood function as the distribution of the radiation sources, as shown in Eq. (5).

* = arg max L_{MLE}( ) = arg max _{n=1}^N P(C_t^n |)

Where P(C_t^n |) is the probability of the photon peak counts received by detector n when the source distribution is , which follows the Poisson distribution, as shown in Eq. (6).

P(C_t^n |) = (c_n( )t)$^{\{C_t}$n} e^{-c_n( )t} / (C_t^n)!

B. Hybrid Algorithm Structure and Key Techniques

1. Algorithm Structure

We propose a hybrid algorithm, Geo-EM-ML, which combines the expectation-maximization algorithm and the maximum likelihood estimation method, incorporating the geometric information of complex pipeline systems to address the multi-source localization problem. The algorithm pipeline is shown in [Figure 2: see original paper].

An introduction to each part of the algorithm framework is as follows:

Initial Estimation: At the beginning of the algorithm, the pipeline system is uniformly sampled, and the Expectation-Maximization algorithm is used to estimate the activity of the radiation sources at these sampled positions, denoted as EM. Additionally, running the full Geo-EM-ML algorithm takes a relatively long time (around 3 minutes), while the EM algorithm can quickly provide a good initial estimate (in about 2 seconds), helping to identify potential radiation source areas.

Pre-optimization: Then, the maximum likelihood estimation is used to optimize the results of the EM to obtain more accurate estimates of the activity of the radiation sources, denoted as MLE-1. Subsequently, a local maximum search is performed on the estimated results to exclude radiation sources with smaller activity values. After that, the maximum likelihood estimation is used again to optimize the activity at the local maximum, denoted as MLE-2.

Joint Optimization: Finally, joint optimization is performed to iteratively estimate the position, activity, and quantity of radiation sources until the termination criteria are satisfied. In each iteration, the maximum likelihood estimation method is alternately applied without and with geometric constraints to estimate the position and activity of the radiation sources. The maximum likelihood estimation method without geometric constraints is denoted as MLE-3, and the one with geometric constraints is denoted as MLE-4. After each iteration, a local maximum search is performed, and the radiation sources around the local maximum are merged to reduce the estimated number of radiation sources.

2. Geometric Constraints

In the Geo-EM-ML algorithm, geometric constraints are used to provide prior information on the position of the radiation sources, thereby improving the convergence speed and accuracy of the algorithm. Geometric constraints are used in two modules of the Geo-EM-ML algorithm. The initial estimation module uses the Expectation-Maximization algorithm to estimate the activity of the radiation source, requiring geometric constraints to restrict the sampling positions of the radiation source. In the joint optimization module, geometric constraints are added to the likelihood function in the form of probabilities to constrain the position of the radiation sources.

In the initial estimation module, we propose two sampling methods for the position of the radiation sources, denoted as PIPE and SPACE. The PIPE sampling method uses the geometric model of the pipeline system as prior information, while the SPACE sampling method does not use any prior information. Figure 3: see original paper shows that when using the PIPE sampling method, the radiation sources are uniformly distributed within the pipeline system. Figure 3: see original paper shows that when using the SPACE sampling method, the radiation sources are uniformly distributed within the free space around the pipeline system.

In the joint optimization module, geometric constraints are incorporated into the likelihood function through probabilistic modeling. For each radiation source position r_m, we calculate its minimum distance to the pipeline midline $D(r_m)$. This distance's probability density distribution is approximated using a Gaussian function:

$$f(D(r_m)) = 1/(\sqrt{(2\pi)}\ d_{min}(S))\ e^{-D(r_m)2\ /\ (2\ d_{min}(S)^2)}$$

Where $d_{min}(S)$ represents the minimum possible distance between radiation sources. This constraint term combines with the detector response probability to form the constrained likelihood function:

$$L_{MLE}(\ ) = {}_{n=1}^{N} P(C_t^n|\ ) \cdot {}_{m=1}^{M} f(D(r_m))$$

This probabilistic formulation allows geometric constraints and detector responses to be optimized on the same scale, ensuring both geometric compliance and numerical stability. The Gaussian variance setting ($d_{min}(S)$) balances constraint strength with optimization smoothness.

3. Source Fusion and Filtering

In the Geo-EM-ML algorithm, source fusion and filtering are employed to reduce the number of estimated radiation sources, thereby improving the algorithm's convergence speed and accuracy while correctly determining the number of radiation sources.

These techniques are applied in two main stages of the algorithm: the pre-optimization stage and the joint optimization stage.

In the pre-optimization stage, specifically in the MLE-1 and MLE-2 modules, local maxima search and activity threshold filtering are primarily used to reduce the number of candidate radiation sources. The local maxima search identifies peaks in the activity distribution to eliminate pseudo-sources with lower activity values, while activity threshold filtering directly removes candidate sources below a preset threshold.

In the joint optimization stage, involving the MLE-3 and MLE-4 modules, a dynamic fusion strategy is implemented. After each iteration, the algorithm performs local maxima clustering, merging radiation sources with spatial distances less than the minimum possible distance $d_{min}(S)$ into a single source. The position of the merged source is determined through activity-weighted averaging, while its activity is the sum of the activities of the merged sources.

Additionally, the algorithm employs an adaptive activity filtering strategy. In the early stages of iteration, activity threshold filtering is not applied to retain all potential radiation sources. As iterations progress and the fusion strategy becomes less effective, activity threshold filtering is gradually introduced to eliminate potential pseudo-sources. The algorithm is considered to have converged when the number of radiation sources remains unchanged between consecutive iterations.

This multi-stage source fusion and filtering strategy significantly enhances the algorithm's performance. It not only reduces the number of candidate sources from $O(10^2)$ in the initial EM stage to $O(10^0)$ in the final stage but also improves localization accuracy by eliminating duplicate estimates and reducing the impact of measurement noise. Furthermore, the dynamic adjustment strategy effectively prevents the algorithm from converging to local optima, ensuring stable convergence.

4. Shielding Effects

The key to accurately calculating the detector response for environments with various materials is determining the distance radiation travels through each material from the source to the detector. Specifically, for complex pipeline systems, the task involves calculating the distances traveled by the line segment between the n-th detector and the m-th radiation source through the pipeline material and air, $d_{n,m,pipe}$ and $d_{n,m,air}$. [Figure 4: see original paper] shows the key steps of the calculation, including listing the intersection points and classifying the line segments.

The pipeline system is first modeled as a water-tight triangle mesh to ensure no holes or cracks on its surface and that the internal and external spaces are clearly defined. Then, the set of line segments to be solved, denoted as $(r_n,$

r_m), n = 1, …, N, m = 1, …, M, is determined based on the positions of the detector and the radiation source.

Next, the position of all intersection points of each line segment with the pipeline model is obtained using the list_{intersections} function in the Open3D library, as denoted in Eq. (9) and shown in Figure 4: see original paper[19].

R_{n,m} = {r_{n,m,1}, …, r_{n,m,I_{n,m}}}

Where I_{n,m} represents the number of intersection points of the line segment with the pipeline model, and r_{n,m,i} represents the position of the i-th intersection point. The intersection points in the set are sorted in ascending order of distance, that is, ||r_{n,m,i+1} - r_m|| > ||r_{n,m,i} - r_m||, i {1, …, I_{n,m} - 1}.

[Figure 4: see original paper] Schematic diagram of calculating d_{n,m,pipe} and d_{n,m,air}, including steps such as listing the intersection points and classifying the line segments.

Since the detector is positioned outside the pipeline model, the segment of the path from the final intersection point, r_{n,m,I_{n,m}}, to the detector position, r_n, must traverse through the air. Additionally, the segment between the last two intersection points, r_{n,m,I_{n,m}-1} and r_{n,m,I_{n,m}}, lies within the pipeline model. This alternation between air and pipeline can be determined by the parity of the total number of intersection points, I_{n,m}, allowing us to classify each segment as occurring in air or within the pipeline model, as shown in Eq. (10) and illustrated in Figure 4: see original paper.

(r_{n,m,i}, r_{n,m,i+1}) = { _{pipe}, if i mod 2 ≠ I_{n,m} mod 2; _{air}, if i mod 2 = I_{n,m} mod 2}

Where i {0, …, I_{n,m}}, _{pipe} and _{air} represent the attenuation coefficients of the pipeline model and air, respectively.

The starting and ending points of the line segment are defined as r_{n,m,0} = r_m and r_{n,m,I_{n,m}+1} = r_n. Finally, the average attenuation coefficient of the line segment, $\bar{\rho}$_{n,m}, can be calculated as shown in Eq. (11).

$\bar{\rho}$*{n,m}* = *(* {i=0}^{I_{n,m}} (r_{n,m,i}, r_{n,m,i+1}) ||r_{n,m,i+1} - r_{n,m,i}||) / ||r_n - r_m||

III. Experiments

A. Experimental Settings

[Figure 5: see original paper] Detectors with different distribution methods.

The overall size of the pipeline system is 3.5655m × 0.495m × 1.3628m, with different shielding materials or pipe wall thicknesses designed according to the specific model. Detectors are distributed throughout the entire pipeline model space using the GRID or RANDOM method, as shown in [Figure 5: see original paper]. The GRID distribution method divides the overall space of the pipeline model into multiple equally sized small cubes, with the center point of each small cube as the position of a detector. The RANDOM distribution method randomly generates the position of a detector within each small cube to ensure that the detectors are evenly distributed throughout the space.

One hundred sets of radiation source positions and activity parameters were randomly generated for each experimental setting. The detector responses for each set of source parameters were generated using two different methods. The first method involved calculating the detector response using Point Kernel Integration, that is, calculating the detector response according to Equations (2) and (3). The second one used the Monte Carlo Simulation to generate the detector response. The simulation process involved building the pipeline model in Geant4 using boolean entities or imported triangular mesh files, positioning isotropic particle guns at the radiation source locations, determining emission probabilities based on the activity ratios of the radiation sources, randomly sampling and emitting particles, simulating particle transport within the pipeline model, and calculating the particle counts at the photon peak to obtain the detector response.

The point kernel integration method is more suitable for theoretically illustrating the best possible algorithm's performance. In contrast, the detector response data generated by the Monte Carlo method, which introduces random noise, is closer to the actual detector response and is more suitable for illustrating the performance of the Geo-EM-ML algorithm in practical applications.

[Figure 6: see original paper] Point Kernel Integration and Geant4 generated the detector response under the same experimental settings, and the relative deviation between the two data sets was also recorded.

The detector responses are obtained under the same source distribution parameters and pipeline model, using different generating methods, as shown in [Figure 6: see original paper]. Despite the statistical errors the Monte Carlo results introduced, the overall trends are consistent. In addition, all high-count detectors have a relative deviation of less than 5%. Due to the more significant impact of high-count detectors on the algorithm, the fact that all high-count detectors have low deviation helps narrow the gap between point kernel integration and Geant4. The result also demonstrates the effectiveness of the detector response calculation for pipeline models with shielding materials proposed in Sec. II B 4.

The final results produced by the Geo-EM-ML algorithm are denoted as $\_\{opt\} = \{(r\_\{opt\}^m, \lambda\{opt\}^m), m = 1, 2, \ldots, M\{opt\}\}$, and the true source parameters are denoted as $\_\{gt\} = \{(r\_\{gt\}^m, \lambda\{gt\}^m), m = 1, 2, \ldots, M\{gt\}\}$. The absolute position error is defined as $\Delta r\_m = ||r\_\{gt\}^m - r\_\{opt\}^m||$. The standard distance is defined as the distance from each radiation source to the nearest detector, that is, $d\_\{standard\} = \min\_n d\_\{n,m\}$. The relative position error is defined as the ratio of the absolute position error to the standard distance, that is, $\delta r\_m = \Delta r\_m / d\_\{standard\}$. The absolute activity error is defined as $\Delta\lambda m$ = /λ{gt}^m - λ_{opt}^m|. The relative activity error is defined as the ratio of the absolute activity error to the true activity, that is, $\delta\lambda$_m = $\Delta\lambda m$ / λ{gt}^m.

In all experiments, the criteria for evaluating the algorithm's success are defined by the following three conditions. First, the estimated number of radiation sources should be correct, i.e., M_{opt} = M_{gt}. Second, the maximum relative position error of the estimated radiation sources should be less than 50%, i.e., max_m δr_m < 50%. Finally, the maximum relative activity error of the estimated radiation sources should be less than 50%, i.e., max_m $\delta\lambda$_m < 50%.

The success rate p is defined as the ratio of successful experiments to the total number of experiments, where an experiment is considered successful if it meets the three conditions mentioned above.

B. Experimental Results and Analysis

1. Implementation and Basic Performance

The parameters of the baseline experiment were set as follows: the shielding material of the pipeline model was stainless steel, the wall thickness of the pipe was 5 mm, the number of radiation sources was 6, the ratio of the minimum to the maximum activity of the radiation sources was 1:2, the number of detectors was 164, and the distribution of detectors was GRID.

[Figure 7: see original paper] provides a comparison of the initial estimation, final estimation, and ground truth obtained by the Geo-EM-ML algorithm. As shown in Figure 7: see original paper, the EM sub-module roughly identifies the area where the hotspot is located, helping users identify the possible location of the radiation source. Comparison of Figure 7: see original paper and Figure 7: see original paper shows that the final result accurately reflects the radiation sources' number, position, and activity.

The baseline experiment was run on an Intel(R) Core(TM) i5-12600 CPU @ 4.80 GHz with 128 GB of memory. The time distribution of each sub-module and the total time distribution in the baseline experiment are shown in [Figure 8: see original paper]. The average time to obtain the final result using the point kernel integration method and Geant4 simulation was 172.22 s and 174.28 s, respectively. The time consumption of each sub-module satisfies the order MLE-1 > MLE-3 > MLE-4 > EM > MLE-2. The MLE-1 sub-module contributed the most time, accounting for approximately 65% of the total time. The EM sub-module can provide an initial estimate in 2s, enabling users to identify the area where radiation sources may exist quickly. The computation time of MLE-4 is between one-quarter and one-third of that of MLE-3. Since the number of calls for MLE-3 and MLE-4 is the same, using geometric constraints has enhanced convergence efficiency.

2. Parameter Influence Analysis

To comprehensively evaluate the performance and applicability of the algorithm, experiments were conducted to assess the algorithm's performance under different pipeline shielding parameters (material and thickness of pipeline wall), source parameters (number and intensity distribution range), and detector parameters (number and distribution method). [Figure 9: see original paper] provides the experimental results and analysis.

While six radiation sources are randomly placed in the pipeline model, with a minimum-to-maximum ratio of 1:2 for the source activity parameters and 164 detectors distributed in a GRID pattern, the success rate p, average relative position error Δr, and average relative activity error $\delta\lambda$ of the algorithm are tested under different pipeline wall shielding materials and thicknesses. The test results are shown in Figure 9: see original paper. As the linear attenuation coefficient of the shielding material or the wall thickness increases, the algorithm's performance slightly decreases. However, the success rate remains above 92%, and the average position error and the average activity relative error are also maintained below 8.58 mm and 4.96%, respectively. The result indicates that, with accurate modeling and calculation of the attenuation effects, the Geo-EM-ML algorithm has a certain degree of robustness to changes in the shielding material and wall thickness.

With the shielding material as stainless steel, a wall thickness of 5 mm, and 164 detectors distributed in a GRID pattern, the success rate p, average relative position error Δr, and average relative activity error $\delta\lambda$ of the algorithm are tested under different numbers of radiation sources and intensity distribution ranges. The test results are shown in Figure 9: see original paper. As the number of radiation sources and the range of relative intensity increase, the algorithm's performance is affected, with a decrease in the success rate and an increase in the average position error and the average activity relative error. Due to the increase in the number of radiation sources and the range of relative activity, the parameter space for searching for the optimal solution to the inverse problem expands, increasing the ill-posedness of the inverse problem. Apart from that, as the range of relative activity increases, the likelihood that the detector response from the sources with higher activity masks the one from weaker sources increases. These factors lead to a decrease in the algorithm's success rate and an increase in the position error and activity relative error.

With stainless steel as the shielding material of the pipeline model, a wall thickness of 5 mm, six radiation sources randomly placed, and a minimum-to-maximum ratio of 1:2 for the source activity parameters, the success rate p, average relative position error Δr, and average relative activity error $\delta\lambda$ of the algorithm are tested under different numbers of detectors and distribution methods. The test results are shown in Figure 9: see original paper. As the number of detectors increases, the constraint equations for the localization task increase, reducing the ill-posedness of the inverse problem and improving the algorithm's performance. Whether the detectors are arranged in a GRID or RANDOM pattern, the success rate of localizing six radiation sources exceeds

92% when the number of detectors exceeds 135. The result indicates that the algorithm does not rely on a structured detector arrangement.

In most cases, when using the same pipeline model, radiation source parameters, number of detectors, and distribution method, the algorithm performs better with detector response generated by the point kernel integration compared to that generated by Geant4, as shown in [Figure 9: see original paper]. The result is consistent with the analysis in Sec. III A.

C. Ablation Study

In this section, ablation experiments are conducted on each module of the Geo-EM-ML algorithm to demonstrate the necessity of using these modules. All ablation experiments use the same experimental settings, which are the same as the settings of the baseline experiment in Sec. III B 1. The algorithm settings and results of the ablation study are shown in , with the first row showing the results of the baseline experiment.

Compared to the baseline experiment, the results of the second group show a significant decrease in both success rate and positioning accuracy. The results of the third group show a slight decrease in both success rate and positioning accuracy compared to the baseline experiment. The algorithms used in the second and third groups replace the maximum likelihood estimation algorithm with the least squares method in the pre-optimization and joint optimization stages, respectively. The results indicate that using the maximum likelihood estimation algorithm can achieve better positioning results.

The results of the fourth and fifth groups of experiments show that using the maximum likelihood estimation algorithm, either with or without geometric constraints alone, will significantly increase positioning error when using the detector response data generated by point kernel integration. However, for the detector response data generated by Geant4, the change in the results of these two groups of experiments compared to the baseline experiment is insignificant. In the sixth and seventh groups of experiments, the radiation source positions sampled by the SPACE method in the initial estimation and pre-optimization stages, and in the seventh group of experiments, the geometric constraints are not used in the joint optimization stage. The results show that the radiation source positions sampled by the SPACE method significantly reduce the success rate and positioning accuracy of the algorithm. Furthermore, not using geometric constraints in the joint optimization stage further significantly reduces the performance of the algorithm. The PIPE sampling method utilizes geometric constraints and samples fewer positions (SPACE samples 1540 positions, while PIPE samples 916 positions). By comparing experiments 4, 5, 6, 7, and the baseline experiment, it can be seen that using geometric constraints significantly improves the performance of the algorithm.

Ablation Study of Geo-EM-ML Algorithm

Index Algorithm Settings Ray Casting Initial Positions Sample Method Pre-optimization Joint Optimization MLE-1 and MLE-2 MLE-3 and MLE-4 MLE-3 and MLE-4 MLE-1 and MLE-2 LSE and MLE-4 MLE-1 and MLE-2 MLE-4 MLE-1 and MLE-2 MLE-3 SPACE MLE-1 and MLE-2 MLE-3 and MLE-4 SPACE MLE-1 and MLE-2 MLE-3 MLE-1 and MLE-2 MLE-3 and MLE-4 Experiments Settings Detectors Generation Method Geant4 Geant4 Geant4 Geant4 Geant4 Geant4 Geant4 Geant4

Results

p (%) ↑ Δr (mm) ↓ $\delta\lambda$ (%) ↓ a In the data of the fourth group of experiments, the average position error and the average activity relative error of the point kernel integration results are more significant than those of Geant4. The issue is caused by the failed cases having more significant position errors and activity relative errors, increasing the average value. For successful cases, the average position error of the point kernel integration result is 6.12 mm, and the average activity relative error is 2.57 %, while the results of Geant4 are 8.69 mm and 3.91 %, respectively. b For successful cases, the average position error of the point kernel integration result is 4.74 mm, and the average activity relative error is 2.26 %, while the results of Geant4 are 7.27 mm and 3.76 %, respectively.

The results of the eighth group of experiments show that neglecting the shielding effects in a shielded scenario significantly decreases both positioning accuracy and success rate. Compared to previous studies, the Geo-EM-ML algorithm accurately models the attenuation based on ray tracing, effectively improving positioning accuracy and success rate.

IV. Conclusions

This study addresses the problem of multi-source localization in complex pipeline systems and proposes a novel Geometrically Constrained Expectation-Maximization Maximum Likelihood (Geo-EM-ML) algorithm. The algorithm integrates the Expectation-Maximization algorithm and Maximum Likelihood Estimation, incorporating constraints based on the geometric structure of the pipeline system. The algorithm effectively addresses the challenges of limited solution spaces, shielding effects, and the ill-posed nature of inverse problems encountered by traditional methods in multi-source localization.

This study proposed a multi-source localization algorithm effectively integrating geometric constraints, breaking through the limitations of traditional two-dimensional or discrete grid spaces, and achieving efficient localization in three-dimensional continuous space. The algorithm can handle more radiation hotspots, expanding its applicability in practical applications.

It uses ray tracing to model radiation shielding effects accurately, ensuring the accuracy and stability of the algorithm in complex pipeline systems. Additionally, the algorithm designs an effective source fusion and filtering strategy, which reduces the computational complexity of the algorithm while accurately estimating the number of radiation sources, providing faster and more accurate localization results for practical applications.

The experimental results of this study show that the Geo-EM-ML algorithm exhibits high success rates and precise localization capabilities under different pipe thicknesses, materials, source activities, and detector configurations. In addition, ablation studies further validate the necessity of each module, demonstrating the critical role of geometric constraints and accurate attenuation calculations in improving algorithm performance.

The Geo-EM-ML algorithm has crucial practical value in nuclear facility maintenance, enabling rapid and accurate localization of radiation sources, significantly reducing the risk of radiation exposure to workers, and optimizing maintenance path planning. Its efficiency and robustness make it highly scalable and applicable in large-scale sensor network environments.

Future work will further optimize the Geo-EM-ML algorithm, develop geometric constraint methods for more general scenarios, and apply it to multi-source localization problems in arbitrary three-dimensional environments. In addition, other optimization algorithms, such as genetic algorithms and simulated annealing algorithms, will be explored to improve the algorithm's global search capabilities and robustness.

References

[1] L. Chen, N. Van Thai, H. Shyu, et al., In situ clouds-powered 3-D radiation detection and localization using novel color-depth-radiation (CDR) mapping. Advanced Robotics 28, 841–857 (2014). doi: 10.1080/01691864.2014.894942

[2] B. Deb, Iterative Estimation of Location and Trajectory of Radioactive Sources With a Networked System of Detectors. IEEE Transactions on Nuclear Science 60, 1315–1326 (2013). doi: 10.1109/TNS.2013.2247060

[3] B. Deb, J.A.F. Ross, A. Ivan, et al., Radioactive Source Estimation Using a System of Directional and Non-Directional Detectors. IEEE Transactions on Nuclear Science 58, 3281–3290 (2011). doi: 10.1109/TNS.2011.2165558

[4] D. Hellfeld, T.H.Y. Joshi, M.S. Bandstra, et al., Gamma-Ray Point-Source Localization and Sparse Image Reconstruction Using Poisson Likelihood. IEEE Transactions on Nuclear Science 66, 2088–2099 (2019). doi: 10.1109/TNS.2019.2930294

[5] J.R. Vavrek, D. Hellfeld, M.S. Bandstra, et al., Reconstructing the Position and Intensity of Multiple Gamma-Ray Point Sources With a Sparse Parametric Algorithm. IEEE Transactions on Nuclear Science 67, 2421–2430 (2020). doi: 10.1109/TNS.2020.3024735

[6] M.S. Bandstra, D. Hellfeld, J.R. Vavrek, et al., Improved Gamma-Ray Point Source Quantification in Three Dimensions by Modeling Attenuation in the Scene. IEEE Transactions on Nuclear Science 68, 2637–2646 (2021). doi: 10.1109/TNS.2021.3113588

[7] A. Abdelhakim, Heuristic techniques for maximum likelihood localization of radioactive sources via a sensor network. Nuclear Science and Techniques 34,

127 (2023). doi: 10.1007/s41365-023-01267-3

[8] R.B. Anderson, M. Pryor, A. Abeyta, et al., Mobile Robotic Radiation Surveying With Recursive Bayesian Estimation and Attenuation Modeling. IEEE Transactions on Automation Science and Engineering 19, 410–424 (2022). doi: 10.1109/TASE.2020.3036808

[9] M. Morelande, et al., Detection and parameter estimation of multiple radioactive sources, 2007 10th International Conference on Information Fusion (2007)

[10] B. Ristic, et al., A controlled search for radioactive point sources, 2008 11th International Conference on Information Fusion (2008)

[11] M.R. Morelande, B. Ristic, Radiological Source Detection and Localisation Using Bayesian Techniques. IEEE Transactions on Signal Processing 57, 4220–4231 (2009). doi: 10.1109/TSP.2009.2026618

[12] B. Ristic, M. Morelande, A. Gunatilaka, Information driven search for point sources of gamma radiation. Signal Processing 90, 1225–1239 (2010). doi: 10.1016/j.sigpro.2009.10.006

[13] J. Chin, et al., Efficient and Robust Localization of Multiple Radiation Sources in Complex Environments, 2011 31st International Conference on Distributed Computing Systems (2011)

[14] E. Wacholder, E. Elias, Y. Merlis, Artificial neural networks optimization method for radioactive source localization. Nuclear Technology 110, (1995). doi:

[15] F. Mendes, M. Barros, A. Vale, et al., Radioactive hot-spot localisation and identification using deep learning. Journal of Radiological Protection 42, 011516 (2022). doi: 10.1088/1361-6498/ac1a5c

[16] A. Abdelhakim, Machine learning for localization of radioactive sources via a distributed sensor network. Soft Computing 27, 10493–10508 (2023). doi: 10.1007/s00500-023-08447-8

[17] Y. Hao, Z. Wu, Y. Pu, et al., Research on inversion method for complex source-term distributions based on deep neural networks. Nuclear Science and Techniques 34, 195 (2023). doi: 10.1007/s41365-023-01327-8

[18] R. Okabe, S. Xue, J.R. Vavrek, et al., Tetris-inspired detector with neural network for radiation mapping. Nature Communications 15, 3061 (2024). doi: 10.1038/s41467-024-47338-w

[19] Qian-Yi Zhou, Jaesik Park, Vladlen Koltun, Open3D: A Modern Library for 3D Data Processing. Dig. arXiv (2018). doi: 10.48550/arXiv.1801.09847

Note: Figure translations are in progress. See original paper for figures.