Proximity-based Sensor Fusion of Depth Cameras and Isotropic Rad-detectors

Kristofer Henderson, Xiaomeng Liu, Kelsey Stadnikia, Allan Martin, Andreas Enqvist, and Sanjeev J. Koppal
University of Florida
khenderson5@ufl.edu

***Abstract*—Finding and tracking radioactive sources has numerous security applications in civilian energy installations, military facilities and ports of entry. The price of radiological sensors varies proportionally to size and imaging characteristics such as angular resolution, and the cheapest devices are nearly isotropic - i.e. they integrate radiation from a sphere of directions centered at the sensor. While many radiation sensors have high aspect rations or odd shapes, the sensors used here are right cylinders, with near identical directional efficiency such that for analysis purposes, other aspects such as counting statistics would make non-isotropy of the sensor negligible. In this paper, we propose a simple and robust way to integrate measurements from both isotropic radiological sensors and depth sensors, whose reliability and resolution benefit from recent advances in computer vision and imaging. Our key idea is to convert all sensor measurements into proximity signals based on radial distance variations over time. Based on this sensor fusion model, we show that for *moving* radiological sources even a simple Kalman filter can trade-off the complementary strength of high-resolution depth sensors and isotropic radiological sensors. We show novel results with a LIDAR sensor and a thermal stereo pair, and demonstrate applications such as tracking and rendering non-line-of-sight imagery behind obstacles and detecting multiple radiological sources in the same scene.**

I. Introduction

Nuclear material trafficking at national ports of entry is a major security risk. Trafficked nuclear material can be used to inflict harm on people as a radiation dispersal device (RDD) or nuclear weapon. Even if no such attack occurs, the contraband material exposes those nearby to high radiation dosage [1]. Less than 1 liter of special nuclear material (SNM) is required to create a nuclear weapon, and is easily stored on a person or in their luggage. Nuclear material trafficking is not an issue of the past; 2500 incidents of reported nuclear material trafficking are recorded in the International Atomic Energy Agency (IAEA) incident and trafficking database [2], since the creation of the organization in 1957.

Radiation detectors are used to measure the neutrons, gamma-ray photons and other nuclear radiation emitted by trafficked material. High-end detectors can differentiate between multiple sources, and detect the direction from the detector to the source. However, these high-end detectors are expensive, a single detector can cost over $100,000. There are detectors in the $1,000-$5,000 range with spectroscopic capabilities but without directional imaging capabilities, these are problematic because of they lack the acuity to spatially differentiate between multiple sources and cannot determine source direction, unless leveraging shielding or multiple detectors.

Sensor researchers have tried to use inexpensive detectors with depth-based vision sensors, in order to reduce costs and exploit relatively mature vision capabilities such as detection and tracking in simple environments. Such systems fuse data across multiple sensors and include techniques such as satellite imagery and laser mapping [3]–[7]. The main technique has been to increase angular resolution by using an array of isotropic radiation detectors with a coded aperture (made of gamma-ray or neutron attenuating slabs). However, this increases the cost of the system dramatically $\approx$ $\$100,000$ and requires additional calibration. Both of these reasons prohibit such systems from being deployed to monitor large areas. The coded aperture needs to be applied to each of the sensors to encode directionality. To monitor a large area, a larger number of detector systems must be deployed, which would be very expensive. In addition, each detector must be extensively calibrated and this will take a large amount of time. On the other hand, reducing the number of sensor nodes in these systems increases the chances of occlusion, especially in crowded and visually cluttered environments such as airports and ports. This paper is about robust, reliable and inexpensive strategies for fusion between depth sensors and a small number of cheap individual radiological sensors. The data fusion described can make use of separate gamma-ray and neutron data streams (if the detector utilized has pulse shape discriminating (PSD) capabilities), or as mostly shown in the paper here [8], combined neutron and gamma-ray detection data, which would be the only data stream available for non-PSD capable fast neutron detectors.

Stationary Sources: There are also ways to locate stationary sources when having many suitably located stationary or mobile radiation sensors, however that scenario is not one we are exploring here, focusing on the data fusion with vision sensors and dynamic environments which is a very different scenario [9].

A. Fusion with Radial Trajectories

We have recently attacked this problem by abstracting both vision and radiological sensor measurements into simple, 1D radial trajectories based on the proximity of the object of interest to the sensor [10], [11]. For depth-based vision sensors, if there are I moving objects in a scene over time, the i^{th} moving object is represented as $R^i_{vision}(t) = \sqrt{X_i^2(t) + Y_i^2(t) + Z_i^2(t)}$, where $(X(t), Y(t), Z(t))$ is some measure of overall location of this object, such as the centroid.

For a radiological sensor detecting a single source in the scene, $R_{sensor}(t) \propto \frac{1}{\sqrt{count\_rate(t)}}$ where $count\_rate(t)$ is the radiation counts measured over time. The sensor fusion problem is therefore simplified into a 1D signal matching/decomposition problem for $R^i_{vision}(t)$ and $R_{sensor}(t)$ across objects i and time t.

In contrast to [10] where calibration of the system was the goal, here we focus on sensor fusion complications that occur when there are multiple sources, occlusions and noise in the data, and address both passive and active depth sensors. This paper extends our previous work [11], and we show new results on non-laboratory weapon-grade radiological sources, introduce new computational imaging algorithms to render non-line-of-sight (NLOS) imagery, add noise analysis to illustrate our dependence on radiological counts and demonstrate the applicability of our methods on both a passive, thermal stereo pair and an active LIDAR system.

B. Context and Contributions

Sensor fusion between radiological and depth sensors is a worthwhile goal since they have complementary strengths. The radiological sensors, despite being isotropic to minimize cost, have the advantage of potentially tracking sources through visual obstacles such as walls. The depth sensors, while requiring line-of-sight (LOS), have higher angular resolution. For example, many commercial LIDAR systems have $360°$ field of view (FOV) that enables easy calibration. Further, the novel passive thermal stereo pair that we introduce in this paper has the advantages of people tracking in visually degraded environments that might incapacitate active sensors such as a LIDAR. We compromise slightly on the cost of the vision-based depth sensors (on the order of $\$10,000$s for LIDAR and thermal sensors, and whose costs are dropping rapidly), while noting that this is far lower than competing radiological techniques. We achieve the following functionality with our sensor fusion suit (please see the accompanying video material website [12] for a full summary),

- **Single source tracking:** We show that a depth sensor paired with a single radiological detector can easily identify a single moving source in a cluttered and fast moving environment, without knowing the source strength. We show this for a lab environment, as well as for tracking in a real-world military installation.
- **Multiple source tracking:** We also demonstrate, for the first time, a simple and robust non-blind signal separation can detect and track multiple radiological sources, estimating the relative unknown strength of each.
- **Blind tracking:** We investigate a variety of occlusion-resistant applications by combining the vision and radiological measurements in the Kalman filter framework. By adding additional radiation detectors to our system, we demonstrate a new capability of tracking a radiation source through occlusions.
- **Blind tracking with novel depth sensor:** We replicate single radiological source detection, localization and tracking with a novel catadioptric thermal stereo pair. The novel optics reduce cost since only one thermal camera is needed, and the passive sensor addresses some of the limitations of the LIDAR setup.
- **Rendering NLOS views using blind tracking:** NLOS research has given us many ways of sensing through and beyond obstacles [13]–[15]. Using our radiological sensing as an example of NLOS sensing, we propose a method to render unseen human motion by building a motion graph of the tracked individual prior to occlusion. We use the estimated blind tracking to drive rendering of images of the scene as if it was viewed through the obstacle.

C. Related Work

3D Vision and Radiological Sensing: Most efforts [3], [4] at LIDAR and rad-detector fusion focus on rigidly constructed gantries in static scenes. Other efforts use coded apertures (where the random pattern is made of lead squares) to encode directionality [5] in the isotropic sensor, to enable stereo reconstruction [7] or to reduce noise [16]. Inferring material properties, and not just geometry, from the visual measurements allows for estimating background radiation [6]. We show how to replicate these types of capabilities, at a small fraction of the cost of the coded aperture systems, by inferring radioactive source strength and radial distance together, as byproducts of our fusion approach.

Sensor Fusion: Multi-modal sensor platforms allow the combination of thermal, acoustic, sonar, LIDAR, etc. This has a long history in 3D vision and recent efforts have had significant impact [17]. In the radiological sensor domain, using many rad-detectors have resulted in intelligent radiation sensor systems (IRSS) [18] which are based on distributed radiation sensors coupled with networked position data to detect and locate radiation sources, either using geometry [18], [19], statistical models [20]–[22], particle filtering [23], or combining rad-detection with electromagnetic induction data [24]. All of these efforts rely on a large number of radiation detectors and are prohibitively expensive, or suffer from low efficiency; our goal is to achieve similar results with just one or two rad-detectors. Mobile systems that fuse radiation detection with a high-resolution 3D vision sensor (eg Microsoft Kinect) are being used to localize stationary radiation sources in a 3D environment [25]–[27]. Closest to our work is [10] where sensor calibration is presented and its noise characteristics analyzed. In contrast, we show that a depth enables simple user-driven calibration and allows for results such as tracking multiple sources and tracking through walls.

Depth from Thermal Imaging: Thermal sensors are becoming more widespread, and commercial available smartphone sleeves exist $\approx \$200$ (for e.g. from FLIR). We propose a novel catadoptric system similar to that proposed for visible light cameras [28] using a single camera and a planar hot mirror. This enables, for half the cost, a thermal stereo pair, allowing the use of well-known thermal vision algorithms for successful detection, segmentation and tracking, such as [29].

Motion Graph for NLOS Rendering: Many NLOS [13], [14] sensors have been proposed to recover imagery beyond visual obstacles. In particular we note that [14] passively recover a 1D signal of movement beyond the visual field-of-view. Rendering views from such data is a fascinating direction of research. We propose to do this by exploiting motion models that can be generated before occlusion occurs. Such models exist to understand human motion in videos [30], [31]. Another approach is to use recorded human motion as an example of motion to synthesize unrecorded motion. We utilize the motion graph [32] which uses graph topology to link sections of motion capture data at similar frames to generate human motion not present in the data set. While most motion graphs use motion capture data [33], [34], there have been attempts to utilize it with 3D video [35], as we do in this paper, in the context of radiological sensing. The resulting technology enables security operators to see a visualization of identified threats even during times of NLOS for both the operator and his sensors, greatly aiding response-time and accurate decision making

II. Overview, Setup and Kalman Filtering

Our hybrid sensor suite produces radiological measurements and 3D vision sensor-based tracking of moving objects in the scene. Both of these streams of information can be converted into *trajectories*, i.e. 1D signals that denote change in radial distance over time. As described previously, the radial trajectory for each object i is $R^i_{vision}(t)$ and for the fall-off from a single radiological source is $R_{sensor}(t)$.

The core problem of sensor fusion in our application becomes analysis of these signals to find those tracked objects that carry or contain radioactive material or sources. For N multiple sources, indexed by n, the response at the sensor is a weighted combination of the count rates, which are related to radial trajectories through the inverse square law $\sum_{n=1}^{N} w_n(t) * \frac{1}{(R^n_{sensor}(t))^2}$.

The measured trajectories and the unknown, estimated trajectories depend on the application. For example, in an occluded environment, the vision trajectories are unknown but the radiological trajectories are known. In contrast, in a multiple source situation in open space, the vision trajectories are known but the radiological weights w_n are unknown.

Teasing apart the unknown components, given measured trajectories as priors, can be cast as non-blind signal separation [36]. With multiple radiological sources, the problem specifically becomes a non-negative least squares analysis of the radiological trajectory, and we discuss heuristics for robust performance in the face of radiological interreflections, occlusions, noise and radiological background noise.

Our focus in this paper is the novel radiological-depth sensor combination, along with new depth sensor designs and algorithms for rendering NLOS data. We retained only the relatively simple non-blind signal separation algorithm on the raw data since it robustly produced useful results in our context. More sophisticated algorithms, such as say convex optimization [37] or compressive reconstruction [38], may provide additional performance at the cost of extra computation. As an example, we used the transformation metric in [39] to make radiological trajectories from similar objects appear closer to one another, and we found only a small improvement in single source experiments. Further analysis of the recovery algorithm could be a topic of future work.

While sparsity analysis [38] might seem useful when we wish to track a few sources among many people and objects in a scene, in our experiments we found that, practically, non-negative least squares with a sparsifying threshold worked well and did better than sparse signal reconstruction by using L1 optimization. In most of our experiments, we processed either the trajectories in L2 space or by using a similarity transform applied to appearance profiles described in [39]. Finally, a small number of additional false positives are acceptable when determining the identity of the trafficker as long as the correct trajectory is found, i.e. the right person is apprehended.

A. Experimental Setup

Radiological preliminaries: The algorithms we developed are demonstrated on a physical experimental setup that uses EJ-309 scintillator detectors like the ones shown in Fig. 1. The main radiation source used in this paper is a Californium-252 fission source (131-64 μCi strength, as depending the date of each specific measurement data set used, which is specified on this project's website). A Plutonium Beryllium (PuBe) (α,n) source (1 Ci) was used alongside the Californium source for the two-source experiments. Both sources are isotropic, with the fission source having a slightly softer spectrum, thus having a slightly reduced detection efficiency per emitted neutron. Both sources emit a broad spectrum of mostly detectable gamma-ray energies. In the device assembly facility (DAF) experiments, a BeRP 4.484kg weapons-grade plutonium sphere was used, also emitting a fission spectrum of neutrons and gamma-ray energies.

Fig. 1: Two EJ-309 scintillator detectors (3-by-3 and 8-by-5 inch, respectively) as examples of nearly isotropic radiological sensors.

Our rad-detector uses organic liquid compound EJ-309 [40] that has both a high flashpoint and low chemical toxicity when compared to other detector liquids. A cylindrical alumina cell holds the liquid, of dimensions 3-by-3 inches (diameter by height). The diameter and height of the scintilation chamber are the same, causing this type of detector to be nearly isotropic. Another size of EJ-309 scintillation detector was used, having dimensions of 8-by-5 inches (diameter by height). This detector, because of its size, is directional, but was placed at the boundary of the experiment area so the source would hit the detector in a way that the directionality of the detector was minimized. Scintillation in the liquid due to radiation depositing energy is absorbed by a photomultiplier tube (ET-Enterprises 9821B), and then converted into an electronic signal with high gain. A 14 bit, 250 MHz, 16-channel digitizer data acquisition system (Struck SIS3316) was used, which has no input buffer and can continuously collect data.

Radiological safety: Safety is of the utmost importance when performing these experiments. During the experiments performed in our lab at the University of Florida, all participants carrying sources were trained in radiation safety and given dosimetry badges to monitor their dosage and keep it at safe levels. All badges used came back with below detectable limit results, emphasizing minimized dose to experimenters.

Proper safety was observed while measuring at the device assembly facility as well. Prior to arriving at the DAF, every participant completed the DAF training, which includes safety training. Only properly trained DAF personnel handled the sources while performing experiments at the DAF, and the strong sources were not stored or carried on the body, in accordance with DAF regulations. Instead, the sources were placed on rolling carts, and these carts were pushed by the trained DAF personnel for the experiments.

LIDAR preliminaries: We use the Velodyne HDL-32E LIDAR which generates a point cloud with 70k points per frame at 10Hz with a 360° horizontal field of view (FOV) and a vertical FOV between $+10.67^\circ$ and -30.67° from the azimuth. The vertical angular resolution is $\approx 1.33^\circ$, and the horizontal angular resolution is $\approx 0.16^\circ$. This LIDAR has an accuracy of 2cm at a distance of 25m. Tracking moving humans is a core problem of computer vision, and the number of efforts in this space are too numerous to list here; see Ref. [41] for a good survey. We are focusing on situations where people are carrying radiological sources. To reduce the computational burden and decrease response times, we project the 3D data from the 360° FOV LIDAR onto a 2D plane where 1 pixel is equal to 2.5cm. This effectively groups the 2D point cloud into clusters of closely spaced points, and since the pixel size is a little larger than the accuracy of the LIDAR, it eliminates uncertainty in the measurements of objects within 25m of the LIDAR. It also sets the vertical center of all objects in the scene to the height of the LIDAR, a good assumption for source location if there is no prior knowledge of its location. We use background subtraction to separate the static point clusters from the moving clusters. The moving clusters are considered moving objects of interest and these are fed into a Kalman filter for object tracking [42]. Transforming the point cloud into a 2D representation of the data reduces the computational burden of grouping the data points into clusters. It also makes applying the Kalman filter less computationally intensive because the Kalman filter only needs to track the object in 2 dimensions. The loss of height information could be a problem in areas where there is significant change in elevation, but we performed our experiments in a mostly level environment because we thought it would be a good representation of the most traveled areas in airports and other areas this system could be deployed.

The Kalman filter uses prior information to estimate the future state of the objects being tracked, based on a model used to describe objects' motion. It is used in many applications to track moving objects in video [43]. The linear motion model allows the Kalman filter to track randomly moving objects in a scene [43], and this enables the system to track the moving people used to carry sources in our experiments.

Before applying fusion, it is important to locate the radiological sensor in the LIDAR frame of reference. We envision a scenario where a space to be monitored (airport, docks, etc.) is instrumented with a small number of reasonably priced omnidirectional radiological sensors and a larger number of cheaper vision sensors. These differences in number are because visual sensors are affected by opaque obstacles (e.g. walls) and require proper placement for full coverage.

Most previous efforts have focused on highly engineered setups [3], [4], [44] to estimate the relative pose of the LIDAR and radiological sensor. In the unstructured domain, [10] have proposed a calibration technique with a continuous wave time-of-flight (TOF) sensor. Since the method requires blind estimation of the radiological sensor location (i.e. self-calibration) errors in the optimization can result in location errors of the order of 1 meter. In contrast, a 360° FOV vision system such as the LIDAR pictured in Fig. 1, can, in fact, directly view the radiological sensor. The only problem is disambiguating the radiological sensor in the field-of-view. To make the problem easier, we currently require that the user clicks on the location of the sensor, in the top-down 2D projected view. From this point on, we will assume that the location of the LIDAR is at the origin $(0, 0, 0)$ and the location of the radiological sensor is at coordinate point (S_x, S_y, S_z).

Catadioptric thermal stereo preliminaries: Catadioptric camera systems use a single camera and a mirror to create a rectified stereo pair. While the Kalman filter framework is also used for the thermal stereo experiments, this device is a passive sensor when compared to the LIDAR and therefore has well-known advantages. These sensors can also be used at night where traditional cameras cannot. Thermal cameras can be used to spot aircraft at night, seacraft by night, and segment people from the background trivially due to the difference in temperature. Long Wavelength Infrared (LWIR) cameras ($\approx 9 - 14\mu m$) can also be used in rain, fog, dust conditions where the Short Wave Infrared (SWIR) LIDAR would show degraded ability.

Uncertainty: The LIDAR has a 2cm accuracy at a range of 25m, which introduces some uncertainty into the system. This uncertainty will translate into uncertainty in the location of the moving objects' locations and create noise in the radial trajectories measured by the LIDAR, creating potential problems when localizing the source, and when using signal reconstruction for tracking multiple sources.

Uncertainty also lies in the measurements from the radiological detectors. The decay process of radiological source is non-deterministic, causing uncertainty in the radial trajectory calculated from the radiological counts (detection statistics uncertainty). This translates to uncertainty in the radial distance measurement calculated from the detector, which will negatively impact the localization of the source and the signal reconstruction used to locate multiple sources simultaneously. Not having prior knowledge of the source being trafficked is a major source of uncertainty, requiring us to develop a model of the intrinsic efficiency on the fly. Lacking prior knowledge of the source's location on a person adds uncertainty to the radiological data from the detector.

The uncertainty in the LIDAR and the radiological detector combine results in uncertainty in the radiological trajectory signal, and in the object trajectory signals. As the trajectories from the LIDAR and the radiological detector become more dissimilar due to uncertainty, the person carrying the source is more difficult to be correctly identified for both the single source and two source localization. For blind tracking, the uncertainty in the sensors result in uncertainty for the location of the occluded source, and uncertainty in the mapping of intrinsic efficiency $\lambda(R)$ for each detector. The motion graph is affected by this uncertainty and it results in uncertainty in the path reconstruction the motion graph uses to visualize the source behind occlusions.

Sensor suite setup: To set up the sensor suite, we place one or more EJ-309 scintilator detectors in the room, and the LIDAR is placed on a table at a height about chest height, directly over one of the scintilator detectors. We measure the distance between the LIDAR and each scintilator detector either with the LIDAR, or by hand in the case of detectors that are occluded from the LIDAR's view [45]. When we convert the radiological measurements from counts to R_{norm}, we use simple vector addition to transform the trajectories measured by the LIDAR into the point of view of each detector. We take background measurements to calculate the constant background radiation.

III. Single Source Tracking

In this section, we tackle the problem of finding and tracking a single *moving* radiological source. One of the main issues when tracking such a radiological source is that the measurements at the radiological sensor (usually called counts) depend on both intrinsic efficiency (λ) and the source-sensor distance (R), i.e. the count rate is

$$C = \lambda(x, y, z) * e^{-\sigma R} * \frac{1}{R^2}. \quad (1)$$

Exp	GT	Unaltered Signal			Transformed Signal		
		Det	Cor	y/n	Det	Cor	y/n
1	1	1	0.984	Yes	1	0.840	Yes
2	1	1	0.986	Yes	1	0.842	Yes
3	1	1	0.992	Yes	1	0.595	Yes
4	3	3	0.991	Yes	3	0.686	Yes
5	1	2	0.991	No	2	0.915	No
6	3	1	0.993	No	3	0.846	Yes
7	2	2	0.988	Yes	2	0.933	Yes
8	1	1	0.982	Yes	1	0.726	Yes
9	3	3	0.990	Yes	3	0.815	Yes
10	3	3	0.994	Yes	3	0.956	Yes
11	2	2	0.994	Yes	2	0.703	Yes
12	3	3	0.989	Yes	3	0.838	Yes
13	3	3	0.994	Yes	3	0.779	Yes
14	2	2	0.993	Yes	2	0.834	Yes
15	2	2	0.999	Yes	2	0.985	Yes

TABLE I: Single source detection results (131μuCi Cf-252) both with and without the transformation from [39]. The trajectory indicated by GT (Ground Truth) is carrying the radiation source. The trajectory indicated by Det (Detected Signals) is *detected* to be carrying the radiation source. Cor indicates correlation between the original radiological signal and the reconstructed signal. A correct detection is given by "Yes" in the y/n column.

Intrinsic efficiency (λ) is a function of source strength, along with the size of the radiological sensor and its internal efficiency, and the room geometry. (x, y, z) is the 3D location of the source and $R = \sqrt{(x - S_x)^2 + (y - S_y)^2 + (z - S_z)^2}$ is the radial distance from source to radiological sensor. σ is the scattering and absorption of radiation in the medium. This will be dependant on the material between the source and the detector. Most of the time during our experiments, it is air, but occasionally changes when a person or wall comes between the source and the detector. For fast neutron radiation data the attenuation is only about 1% per meter of air distance travelled. Since we are measuring the radiation sources at close distances (<10 meters) we can assume that the attenuation from the $e^{-\sigma R}$ term will be negligible, which then results in a simplified equation $C = \lambda(x, y, z) * \frac{1}{R^2}$. Riley et al [10] also assumes that the exponential term is negligible for air at close distances for localization and tracking.

Note that this equation is similar to the near-lighting model [46] in the case where the *camera directly images the source*, where the numerator would be the light-source intensity. The dependence of the "source strength" $\lambda(x, y, z) * e^{-\sigma R}$ on the scene is due to radiation's propensity to penetrate material and scatter about the scene. This produces different counts, especially if the source or sensor is near the floor or walls. Even in the absence of large objects or boundaries, there is a weak dependence on scattering through air, governed by the normally small parameter σ.

Of course, the other issue is that the low cost isotropic radiological sensors used in the experiments in this paper have no angular resolution and cannot tell where in the scene the detected material is located. Before we explain how to exploit the 360° FOV LIDAR to solve this problem, we make one more assumption about how the model for the intrinsic efficiency $\lambda(x, y, z)$ can be made simpler, similar to previous models such as from [10]. We approximate the numerator in Eq. 1 as a function that slowly varies with R,

$$C = \lambda(R) * \frac{1}{R^2}. \quad (2)$$

Consider a visually cluttered scenario where there is only a single source in the scene. From the Kalman filter-based tracking, let us assume we have tracked all objects in the scene $O_1, O_2, ..O_I$ throughout the time of the experiment T. Each object would have a (x, y, z) trajectory in our 3D LIDAR space, for example O_i : $((x_{i1}, y_{i1}, z_{i1}), (x_{i2}, y_{i2}, z_{i2}), ...(x_{iT}, y_{iT}, z_{iT}))$, and this can easily be converted into radial distance measured *from the radiological sensor's location* at (S_x, S_y, S_z), for example $O_i : (R_{i1}, R_{i2}, ...R_{iT})$.

Now consider the count measurements C_t at each time instance t over the time interval T, denoted as $C_1, C_2, ..., C_T$ from the radiological sensor itself. These must, of course, be only dependent on the movement of the person carrying the single source and invariant to the other people in the scene. Inverting Eq. 2, we can convert these count measurements into radial measurements, modulated by the unknown source strength λ. Therefore the counts $C_1, C_2, ..., C_T$ from the rad-detector can be converted into radial distances $(R_{rad1}, R_{rad2}, ...R_{radT})$, given by $R_{radt} = \sqrt{\frac{\lambda}{C_t}}$.

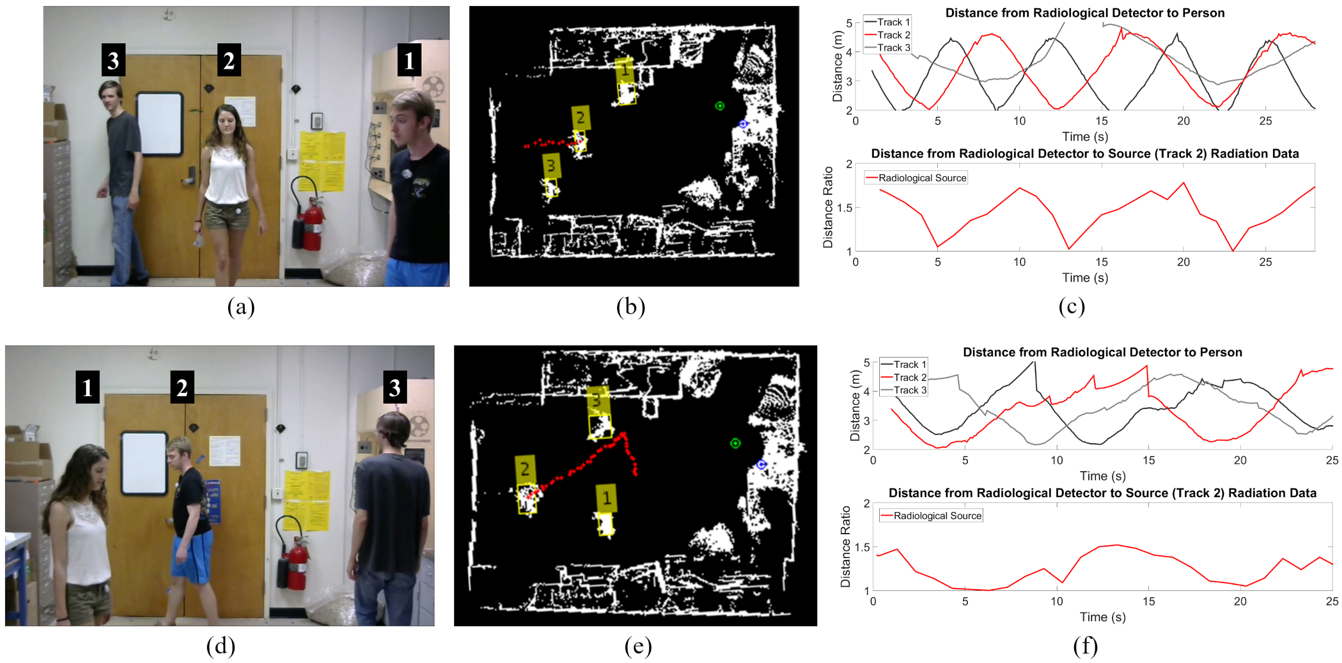


Fig. 2: Results from two single-source tracking experiments. The person carrying the radiological source is Person 2. (top) The detection and tracking of a radiological source during periodic motion. (lower) The detection and tracking of a radiological source during random motion. (a,d) RGB image of the scene. (b,e) An Image of the LIDAR data's 2-D representation. (c,f) A plot of the distance to each person above a plot of the radiation data.

We can remove the dependence on intrinsic efficiency by normalizing the R_{radt} values by taking the ratio of the radial measurements with, say, the minimal radial measurement corresponding to the shortest distance from the source to the radiation detector. We chose this because it corresponds to the data point with the highest count rate and therefore the best statistics. $R_{normt} = \frac{R_t}{R_{min}} = \sqrt{\frac{C_{max}}{C_t}}$. To get this result, we must assume that $\lambda(R)$ is a slowly varying function of R, and that since R remains under 10m, any change in $\lambda(R)$ caused by a change in R will be negligible. Similar normalization can also be applied in the objects tracked by the Kalman filter, for example $O_{normi} : (\frac{R_{i1}}{R_{imin}}, \frac{R_{i2}}{R_{imin}}, \ldots \frac{R_{iT}}{R_{imin}})$. These trajectories are now independent of intrinsic efficiency and depend only on geometry. We can then find the Kalman filter trajectory that minimizes the cosine distance [47] between the radiological measurements to determine which moving Kalman filter object is carrying the radiological source.

The radiological data is dealt with in 3D space. The point cloud we collect from the LIDAR is converted to a 2D image. By doing so, not only does this make point clustering and grouping computationally less intensive, it also sets the z value of the objects all on a single plane, and that plane is equal to the height of the LIDAR. Since we set the LIDAR at a height roughly approximating the mean height of a person, we are effectively tracking the center of the person in 3D even when using the 2D LIDAR representation to track the object. The center of a person is a good assumption of the location of the source given no prior knowledge of the sources location. This causes an error because the center of the person and the center of the source are not guaranteed to coincide, but our results demonstrate that the error does not impact our tracking ability. If we do know where on the person the source is, then we can change the z plane to match the expected location of the source. Meaning that $R_{rad} \approx R$ for the people carrying the source.

From this above discussion, it is clear that we are tracking with the vision sensor in 2D, not 3D, to save computation. With additional computation, this can easily be changed. However, we note that the radiological sensing and distance estimates happen in 3D. In other words, the 2D tracking does not affect the *attribution* of the vision trajectory to the radiological trajectory.

Count rates: For the detectors and sources used here the gamma-ray background was approximately 45 cps and the neutron background was 1 cps. When measuring with a source the average count rate generating a radiological trajectory was 19 cps and 4 cps for gamma-rays and neutrons, respectively.

Since the trajectories only depend on geometry, finding the trajectory that best matches the radiological measurements does not necessarily require both signals be unfiltered. If both the radiological measurement data and the trajectories are passed through the same filter, the cosine distance should still find the best match between the radiological measurements and the n Kalman filter trajectories of the moving radiological sources. This lets us apply filtering to the data in an attempt to improve the detection results. We apply the filter specified by Koppal et al [39], to our radiological and trajectory data in an attempt to improve the detection results. This filter takes the distance signal as an input and outputs a signal that is increasing at a constant rate when the slope of the signal is positive, and decreases at a constant rate when the slope of the input signal is negative. When the signals slope is zero, the value of the output signal is set to zero as well. Note that while the maximum cosine distance may decrease with this filtering technique, the other cosine distances that do not correspond to a positive detection will decrease by a larger amount, potentially resulting in more accurate detection.

Success and failure metrics: We consider a positive or correct detection when the person carrying the source is identified as carrying the source by the sensor suit and algorithm. For our single source experiments, we assume there was only one source in the scene. The failure case is when the system and algorithm incorrectly identifies a person who is not carrying the source to be carrying the source. We recorded the person carrying the source in every experiment, and used this information to verify the accuracy of the system, but never used it as prior knowledge input for the system. The reason for the failure of the two failed experiments in table I was because two or more trajectories R were similar to each other, and one of these happened to be carrying the source. If the source carrier is occluded from the LIDAR of a long period of time, the Kalman filter cannot track the person, so the algorithm cannot locate the source. It is true that if more than one trajectory is similar to R_{norm}, then it is not guaranteed that the correct trajectory will be selected as having the highest value. However, this does not mean that the value associated with the correct object was much lower, it simply means that there was another track that was higher by chance. Given this fact that there are two high values, there are easy ways to combat this failure in practice. Since both trajectories will have high correlation values, it can be on the security personnel to apprehend multiple people with a high correlation score, and inspect each of them. This will still lead to the apprehension of the person carrying the source, it will also result in some false positives.

Evaluation: Table I shows the results from 15 experiments. Each experiment has a group of two or three people moving in a random motion in a laboratory environment. The Cf-252 source strength in these experiments was 107-113 $/mu$Ci. In 13 of those experiments, we correctly detected the radiological source holder, giving an accuracy of 86.67%. In Fig. 2 we show two of the experiments summarized in the table. In each of these two experiments Person 2 has the source. The final column shows plots of the trajectories (distance vs. time) for all three people in the scene, and under that a plot of the radiation data. There is only one trajectory (Person 2) that matches the radiation data's waveform.

Different moving objects have the potential to generate the same radial trajectory, and this could cause false positives in localizing and tracking the source. to explore this in simulation, we simulated 10 simultaneous people randomly walking in a large featureless room for 40 seconds. We then systematically chose each person to be the source carrier by selecting their radial trajectory as the radiological data and performing the dot product with each of the trajectories in the scene. After we used the unaltered trajectory as the radiological data, we then applied additive white Gaussian noise and down sampling to the trajectories to simulate the noise in the real radiological data. We performed 16 experiments this way, and include the most important results in this paper. For experiment 1, we used the unaltered trajectory as the radiological data. Experiment 2 we down sampled the trajectory by a factor of 10. For Experiments 3 and 4, we added white Gaussian noise with mean 0 variance 1, and mean 0 variance 9, respectively. For experiments 5,6,7 we used both additive white Gaussian noise and down sampling simultaneously; using AWGN with mean 0 and variance 1,4,9 respectively. For all 3 experiments (experiment 5, 6, 7) we down sampled by a factor of 10. The results of this can be found in figure 3; the dot product of each trajectory with the radiological data is show in the tables in Fig. 3.

For experiments 1-4 there was no errors in localizing the source. For experiments 5, 6, 7 many of the sources were correctly detected, but some were not. When this occurs, the dot product value tables in figure 3 show that the correct trajectory had a similarly high dot product value. Trajectories caused by the noise and uncertainty of the detector can cause the detector to select the wrong trajectory as the source, and there are ways to overcome this we would like to explore, such as increasing the length of the trajectories in time, and selecting all trajectories with high dot product values to apprehend and inspect for radiological material.

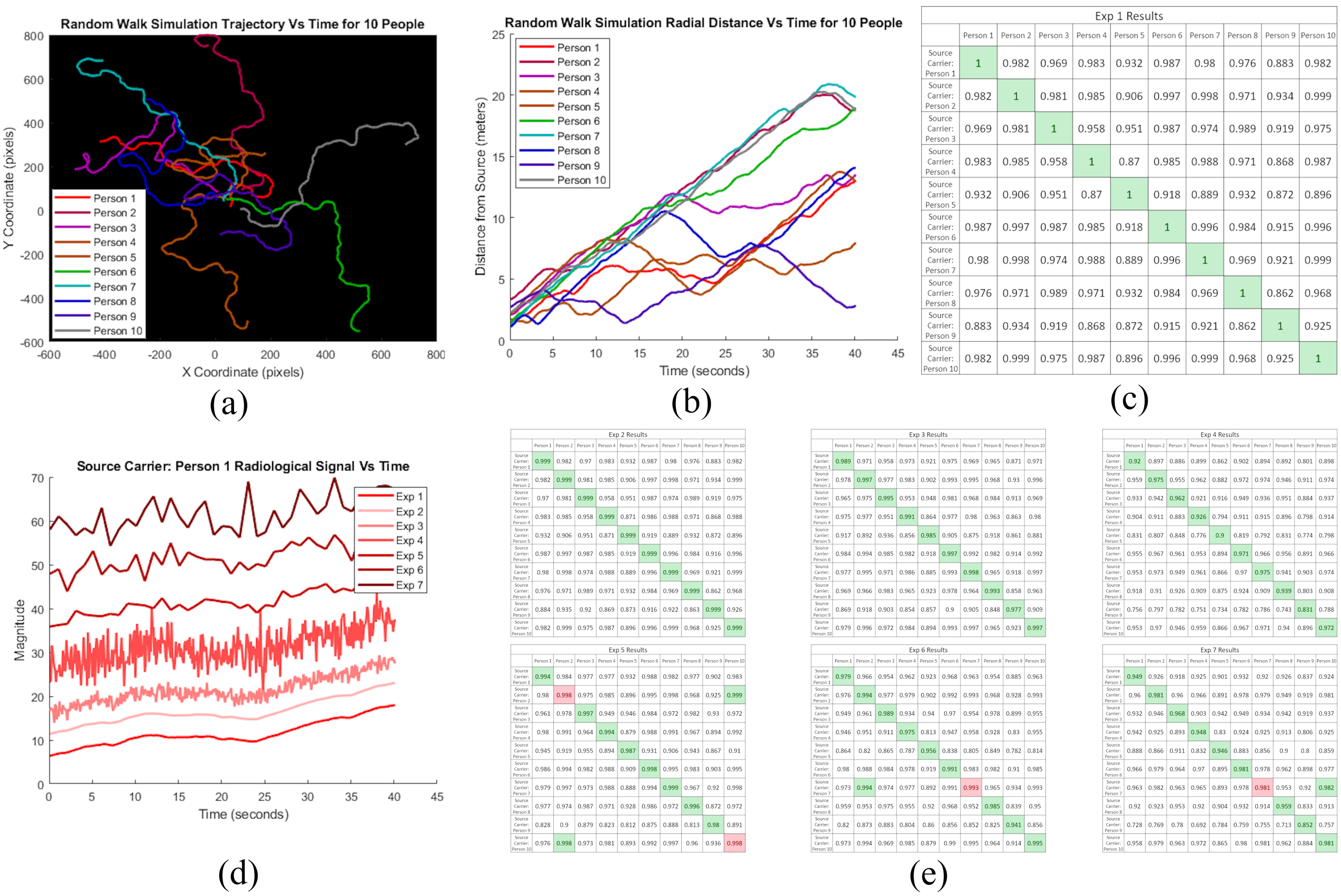
IV. Multiple Source Tracking

In many environments, such as nuclear power plants or nuclear medicine clinics, there may be more than one radiological source being transported. Imagine a coordinated attack in which the nuclear material has been distributed among many individuals in an attempt to make each piece of the material smaller, lighter, and easier to transport. Each piece of material could also require less shielding than the large source, reducing the amount of shielding around each of the smaller sources. While spectral information is available from these sensors, using it to differentiate between multiple sources is only viable if we have prior knowledge of the material makeup of each source. It cannot be assumed that in a real-life trafficking situation the material makeup of each source is known, so we propose to use a vision sensor to provide additional information. To allow for better monitoring and awareness, it is useful to consider how to track multiple such sources. The measurements in the radiological sensor are linear in the number of counts C in Eq. 2, so, at some time t, given N sources $(1, 2, ...N)$, the counts measured would be,

$$C(t) = \sum_{n=1}^{N} C_n(t) = \sum_{n=1}^{N} \lambda_n * \frac{1}{R_n(t)^2}, \tag{3}$$

where $R_n(t)$ is the radial distance from the n^{th} source to the radiological sensor. Now consider again, from the Kalman filter-based tracking, we have tracked all objects in the scene $O_1, O_2, ..O_I$ throughout the time of the experiment T. For each object, we can convert the radial distance from the radiological sensor's location (S_x, S_y, S_z), into a series of inverse-square fall-off terms, $O_i : (\frac{1}{R_{i1}^2}, \frac{1}{R_{i2}^2}, ...\frac{1}{R_{iT}^2})$.

If we collect the radiological counts as a $T \times 1$ vector $\mathbf{C}$, and the radial inverse-square fall-off terms for each object as a $T \times I$ matrix $\mathbf{O}$, then



Exp 1 Results

	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Person 7	Person 8	Person 9	Person 10
Source Carrier: Person 1	1	0.982	0.969	0.983	0.932	0.987	0.98	0.976	0.883	0.982
Source Carrier: Person 2	0.982	1	0.981	0.985	0.906	0.997	0.998	0.971	0.934	0.999
Source Carrier: Person 3	0.969	0.981	1	0.958	0.951	0.987	0.974	0.989	0.919	0.975
Source Carrier: Person 4	0.983	0.985	0.958	1	0.87	0.985	0.988	0.971	0.868	0.987
Source Carrier: Person 5	0.932	0.906	0.951	0.87	1	0.918	0.889	0.932	0.872	0.896
Source Carrier: Person 6	0.987	0.997	0.987	0.985	0.918	1	0.996	0.984	0.915	0.996
Source Carrier: Person 7	0.98	0.998	0.974	0.988	0.889	0.996	1	0.969	0.921	0.999
Source Carrier: Person 8	0.976	0.971	0.989	0.971	0.932	0.984	0.969	1	0.862	0.968
Source Carrier: Person 9	0.883	0.934	0.919	0.868	0.872	0.915	0.921	0.862	1	0.925
Source Carrier: Person 10	0.982	0.999	0.975	0.987	0.896	0.996	0.999	0.968	0.925	1

Fig. 3: We performed simulations of random walks. (a) is the random walk of each simulated person in 2D. (b) is the corresponding Radial trajectory of each random walk. (c) is the result using the unaltered radial distance signal as the radiological signal when comparing with the radial trajectory (Exp 1). Notice how each source carrier is detected correctly as carrying the source. (d) visualizes the simulated radiological data we created by altering the original radial distance signal. (e) shows the results for the each of the 6 altered signals. A green entry means the system detected that trajectory corresponds to teh source carrier. If there is a red entry the row, the detected source carrier is not the true source carrier.

$$\mathbf{C} = \mathbf{O}\ \mathbf{\Lambda}, \tag{4}$$

where $\mathbf{\Lambda}$ is an $I \times 1$ vector such that

$$\Lambda(i) = \begin{cases} \lambda_i & \text{if } O_i \text{ is carrying a radiation source} \\ 0 & \text{otherwise} \end{cases}$$

Use of L2 norm: While radiological counts are usually modeled by Poisson distributions, we note that the radiological trajectories combine the smooth motion of the source with the Poisson distribution of the counts, i.e. any measurement is a combination of the source-sensor distance and the source strength. Additionally, the frame rates of the radiological counts and the vision sensor differ by many orders of magnitude. Further, the radiological counts show significantly more noise than the visual sensors. The end result is that linear smoothing techniques are applied to the radiological sensor measurements so the counts can be modelled by a Gaussian distribution. Modeling radiological counts (and for our purposes, the corresponding calculated radial distances) using a Gaussian distribution instead of a Poisson distribution is acceptable when the expected value is large. This technique has been used before to localize radiological sources [45], [48]. In particular, Poisson works best when the number of samples is low. This is the main reason why L2 norm (which optimizes Gaussian noise in a least squares sense) is a reasonable approximation for our scenario.

We used non-negative least-squares matrix factoring to solve Eq. 4, and sparsified $\mathbf{\Lambda}$ by setting all the values lower than a user defined threshold to zero, and declaring the trajectories corresponding to those remaining non-negative values as the ones with the radiological sources. The user defined threshold was set as some fraction of the largest coefficient in an attempt to excluding all coefficients much smaller than the largest coefficient. In most cases, the threshold was set at $\frac{1}{2}$. In practice, we found that threshold setting was easy, and could perhaps be automated. We did not need to add additional constraints to enforce the sparsity of $\mathbf{\Lambda}$ during its estimation, although these can be used to increase robustness in the future.

Evaluation: In Table II we show the first set of simulated

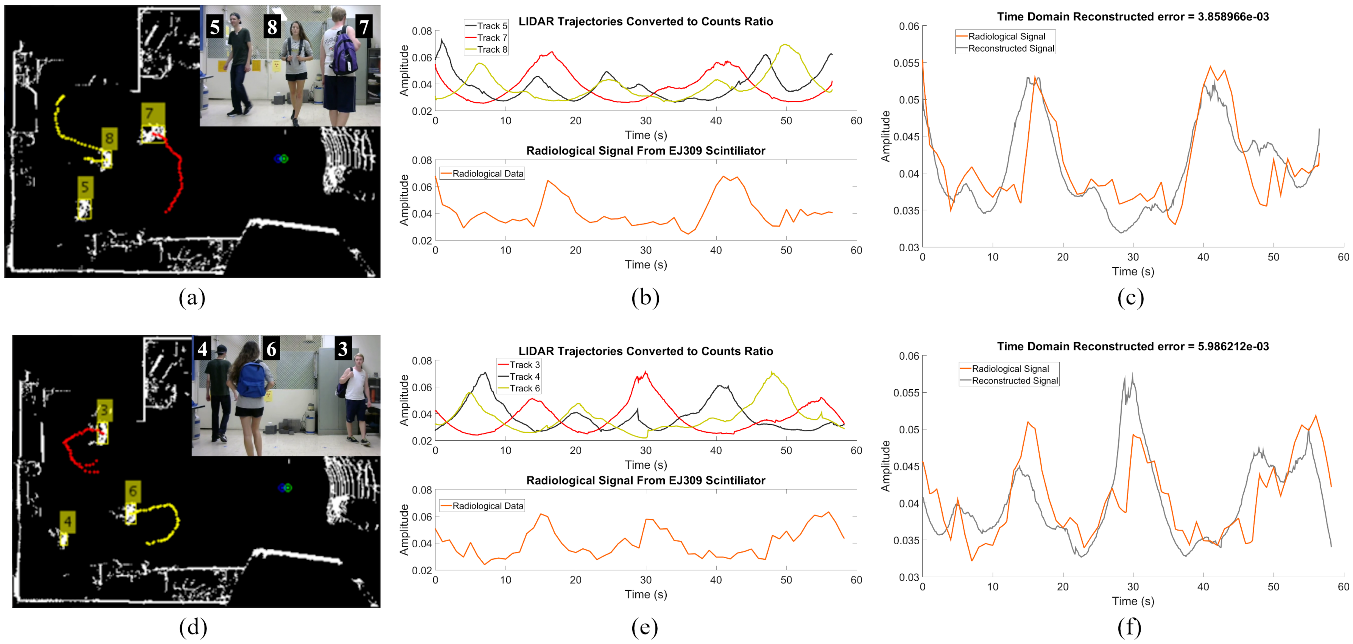


Fig. 4: Results from two-source tracking experiments. Red and yellow indicate the people carrying radiation sources. (inset) RGB image of the scene. (a,d) An image of the LIDAR data's 2-D representation. (b,e) A plot of the distance to each person and a plot of the radiation data. (c,f) A plot of the radiation signal (orange) and the reconstructed signal (grey).

Exp	GT	Unaltered Signal		Transformed Signal	
		Det.	y/n	Det.	y/n
1	2, 6	2, 6	Yes	2, 6	Yes
2	2, 5	2, 5	Yes	2, 5	Yes
3	2, 6	2, 6	Yes	2, 6	Yes
4	2, 6	2, 6	Yes	2, 6	Yes
5	2, 5	2, 5	Yes	2, 5	Yes
6	3, 5	3, 5	Yes	3, 5	Yes
7	3, 6	3, 6	Yes	3, 6	Yes
8	3, 6	3, 6	Yes	3, 6	Yes
9	3, 5	3, 6	No	3, 5	Yes
10	2, 5	2, 5	Yes	2, 5	Yes
11	2, 5	2, 5	Yes	2, 3	No
12	2, 4	2, 4	Yes	2, 4	Yes
13	3, 6	3, 6	Yes	3, 6	Yes
14	3, 5	3, 5	Yes	3, 5	Yes
15	3, 5	3, 5	Yes	3, 5	Yes

TABLE II: Two source detection results using simulated data for both unaltered and transformed data. Both sources have the same strength in these simulated data sets. The trajectories indicated by GT (Ground Truth) are carrying a radiation source. The trajectories indicated by Det (Detected Signals) are *detected* to be carrying a radiation source. A correct detection is indicated by a "Yes" in the y/n column.

Exp	GT	Unaltered Signal			Transformed Signal		
		Det.	Cor.	y/n	Det.	Cor.	y/n
1	1, 2	1, 2	0.996	Yes	1, 2	0.598	Yes
2	2, 1	2, 1	0.995	Yes	2, 1	0.579	Yes
3	1, 2	2, 1	0.993	Yes	1, 2	0.487	Yes
4	2, 3	3, 2	0.996	Yes	2, 3	0.607	Yes
5	1, 3	3, 1	0.991	Yes	1, 3	0.334	Yes
6	2, 3	3, 2	0.994	Yes	3, 2	0.102	Yes
7	1, 3	1, 3	0.996	Yes	1, 3	0.479	Yes
8	1, 3	1, 3	0.992	Yes	1, 2	0.613	No
9	1, 3	1, 3	0.996	Yes	1, 3	0.461	Yes
10	1, 3	1, 3	0.996	Yes	1, 3	0.462	Yes
11	2, 3	2, 3	0.994	Yes	2, 3	0.603	Yes
12	2, 3	2, 3	0.997	Yes	2, 3	0.617	Yes

TABLE III: Two source detection results using real experimental data both with and without the transformation from [39]. The trajectories indicated by GT (Ground Truth) are carrying a radiation source. The trajectories indicated by Det (Detected Signals) are *detected* to be carrying a radiation source. A correct detection is indicated by a "Yes" in the y/n column. The first number in both the GT and Det columns corresponds to the stronger sources and the second number corresponds to the weaker source.

experiments that we ran to test our approach. These were "real" simulations in the sense that we took real radiological counts from two actual experiments and added these using Eq. 3, and attempted to recover the ground-truth. The table has the ground-truth radiological pair, and the detected pair. Since this was a simulated setting, we could create groups of up to 7 "people" using vision trajectories from other experiments, to see how well our method can detect the two sources in a cluttered environment. A few additional false positives are acceptable if every trafficker is correctly identified. This means that a correct detection can be achieved despite a detected false positive as long as all ground truth trajectories are detected. In 15 experiments, we achieved only one incorrect detection, resulting in 93% separation and detection rate.

We then performed 12 actual experiments using two radiological sources, as shown in Table III. The table shows, again the ground-truth radiological pair, and the detected pair, as well as the reconstructed distance measure from the cosine distance metric. Running our algorithm with the unaltered data yielded a detection rate of 100%. Fig. 4 shows two experiments from the set of 12. Each person wearing a backpack is carrying a radiation source. The center image shows three vision trajectories, one for each person in the scene, and the single radiological response (C in Eq. 3). The final column compares the original radiation signal shown in orange to the reconstructed signal using the weighted trajectories (C_i and C_j in Eq. 3) that have been inferred from the single radiological response C.

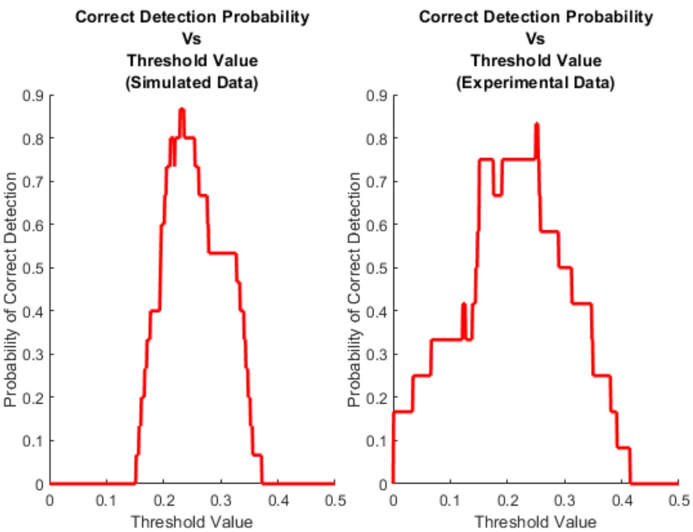


Fig. 5: Graphs of the probability of getting an absolutely correct detection (i.e. no false positives nor false negatives) at any threshold value. Simulated data is represented on the left, real experimental data on the right.

False alarm rates: The success rates of tracking multiple sources needs some qualification. If the correct threshold value is known prior to separation, we are able to achieve correct detection in 93% of the 15 simulated multi source experiments and 100% of the 12 real experiments. However, we must consider the scenario where the threshold for separation is not known prior to separation. The weights from the non-negative least-squares matrix factoring were taken from the 15 simulated experiments and the 12 real experiments. To determine the threshold values that result in an absolute correct detection (i.e. no false positives and no false negatives) for each data set, the threshold was increased from 0 to 1 in steps of 0.001 and a binary metric was returned indicating if the detection was absolutely correct. Separating the data into real and simulated data, we averaged the resulting binary signal across all experiments, resulting in the probability of any given threshold resulting in a detection with no false positives nor false negatives. This result can be seen in figure 5. The optimal threshold for the simulated data is between 0.229 and 0.235, resulting in 87% of the experiments having absolutely correct detection. The optimal threshold for the real experimental data is between 0.250 and 0.253, resulting in 83% of the experiments having absolutely correct detection. We also find that setting the threshold higher than the optimal setting results in a higher chance of false negatives, while setting the threshold lower results in a higher chance of false positives.

V. Blind Tracking

Radiological sources emit radiation that can pass through many materials that are opaque to light. There are light opaque materials that radiation cannot pass through: gamma rays cannot pass through high Z materials such as lead, and neutrons cannot pass through low Z materials such as polyethylene and other organic materials. Even if the radiation can pass through a material, passing through any material will result in the radiation being attenuated to some degree, but a detector can still measure radiation from a source through many light-opaque objects. Therefore, there is potential for tracking radiological sources through opaque obstacles, perhaps even far beyond visual range. However, the reason this is not done in the radiological community is because the constant that we made in the earlier section does not hold. In other words, the source strength varies with radial distance, over time t, as

$$C(t) = \lambda(R(t)) * \frac{1}{R(t)^2} = F(R(t)). \tag{5}$$

Since counts are now a non-linear combination of source strength and distance, tracking single sources over long distances with a single radiological sensor is almost never done, since it would imply finding $F^{-1}(C(t))$ or inverting the above equation. In our case, however, we are using a combination of a radiological sensor and an $360°$ FOV LIDAR, which can allow us to break the dependency between source strength λ and radial distance R. The radiation signal would be attenuated through occlusion materials, but unless the occlusion is of variable geometry the change can be approximated by a fixed step-function change during the occlusion. This allows some ability to estimate that change when tracking a person visually and knowing when they become occluded. Additionally, in many security settings, thinner privacy walls are common, which have only minor impact on the radiation transport.

Consider a scene with a single moving radiological source, which we have identified using the previous methods. For a given interval of time, we now know pairs of $(C(t), R(t))$, or $(F(R(t)), R(t))$. This is because the radial distances $R(t)$s are estimated directly by the LIDAR, and the counts come from the radiological sensor. We can therefore fit a parametric model (always linear in the results shown) to the $(F(R(t)), R(t))$ pairs. This data-driven approach makes it trivial to invert F, since we can extrapolate the model to predict the $R(t)$ associated with any $C(t)$ in any future time instance t.

The only problem remaining is tracking a person with a radial distance $R(t)$, since the person could be anywhere in a circle around the radiological sensor. To break this ambiguity, we have two strategies. In the case of short periods of occlusion, we use the previous velocity vector estimated by the Kalman filter, and intersect the circle with the ray defined by this velocity vector and the last detected LIDAR position. In the top row of Fig. 6 we show an example of a person disappearing behind a large cubicle wall. The first image shows the color frame at the moment of occlusion. The second image is a 2-D representation of the LIDAR data before

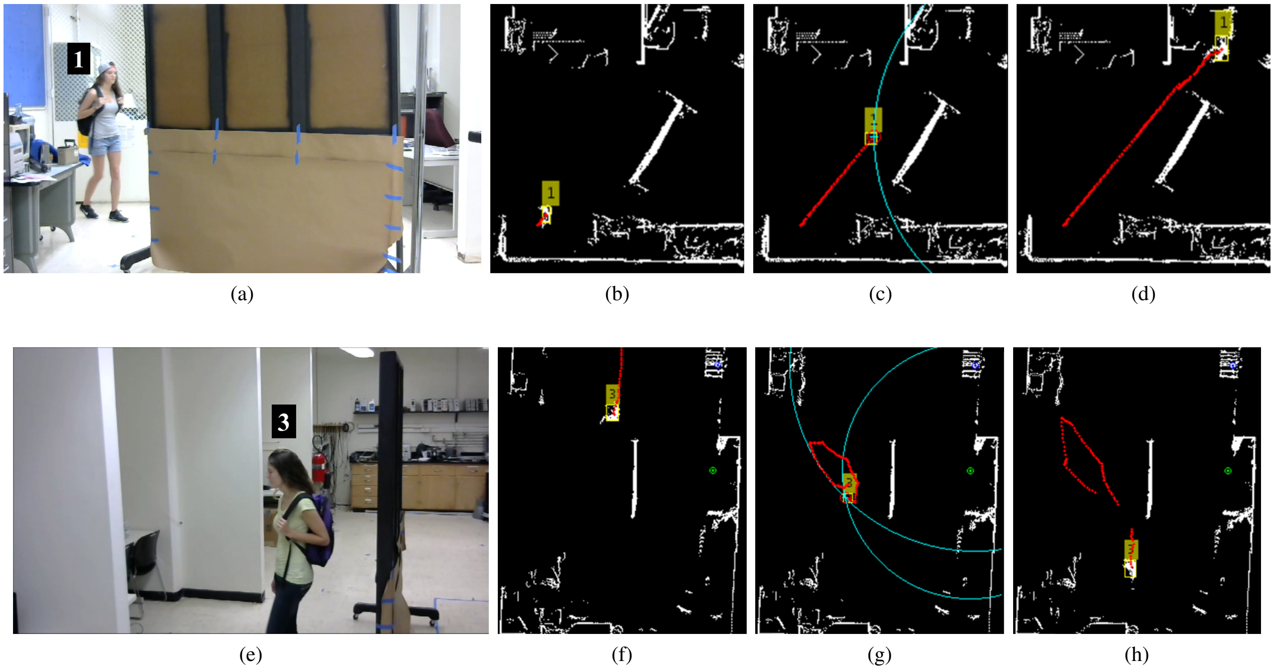


Fig. 6: Blind tracking results using a single detector (top) and two detectors (lower). (a,e) RGB image of the scene. (b-d,f-h) A sequence of images from the 2-D LIDAR data's representation showing the person being tracked through the occlusion. The blue circles indicate the unidirectional distance from each detector to the person (Cf-252 source 102-113 μCi).

the person is occluded. In the third image, we intersect the circle (depicted in blue) with the motion ray. This position is fed into the Kalman filter. The final image shows that the person reappears at the other end of the wall, with their Kalman filter label intact, i.e. we have tracked the same person through visual occlusion.

Obviously, the previous method fails if the person's motion deviates from the most recent Kalman filter estimated velocity vector. To go beyond this, we simply add a second radiological sensor and apply the same modeling algorithm as discussed earlier. Given the two radiological counts from each sensor, we can convert them into two estimated radial distances. These (in 2D) will intersect at two points, and we use the point closest to the last seen LIDAR position, followed by Kalman filtering to choose between the two intersection points. In the second row of Fig. 6 we show an example of a person going behind a wall, and then walking in a circle. The circular motion happens completely behind the wall and is not visible by the LIDAR. Yet, using the detector and LIDAR to estimate the $(C(t), R(t))$ map for each detector allows us to predict the location of the person. Notice the two green circles intersect behind the wall, and the red dots (depicting recent motion) show a circular path.

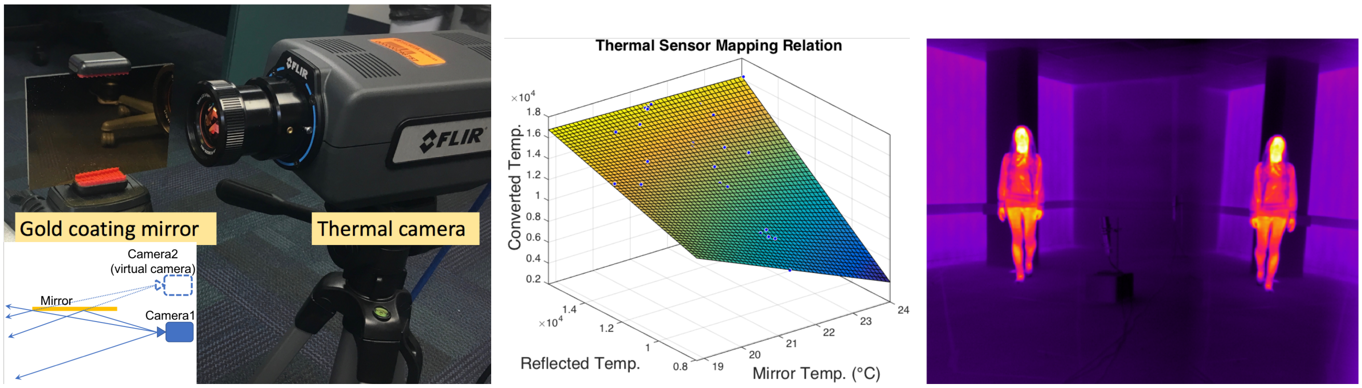
A. Blind Tracking with Thermal Stereo

Thermal sensor technology is becoming widely accessible and impacts many applications, such as finding the heat from overloaded circuitry and other hazards in the home and industrial areas, finding a hidden person or animal in hard to see or dark places, etc. Compared with traditional cameras, thermal camera have more advantages when performing people detection because the infrared radiation from the human body is within a certain range, allowing the person to be easily segmented from the background.

To reduce the cost even further, we propose a novel stereo method for thermal imagery. We use a catadioptric system to optically create a rectified stereo pair using a single camera and planar mirrors as shown in Fig. 7. Several researchers have demonstrated different visible-light catadioptric designs [49]. While the idea has been used with single thermal detectors [50] we believe that we are the first to apply this to full resolution thermal cameras.

Our main algorithmic contribution to accompany our novel thermal catadioptric design is a calibration function that incorporates the mirror temperature. In particular, we noticed that, between the stereo pair images, there existed a temperature differential that depends on the mirror temperature (i.e. the ambient temperature of the scene). Further, the extent of this differential is scene dependent; the reflected temperature for the high-temperature (about higher than 40°C) object would be decreased, while it would be increased for low-temperature object (lower than 18°C). We propose a calibration step where we acquire the relationship between the ambient (mirror) temperature, reflected temperature and ground truth object temperature (Fig. 7). Given this curve for certain camera settings, if an object of known temperature is observed (for e.g. healthy human) then we can infer the scene temperature, much like a thermometer. Conversely, if the ambient temperature is known from an on-board thermometer, the reflected image can be corrected to real-world thermal radiation.

Experiment setup: Several designs for the catadioptric systems have been proposed from previous research [49].



(a) (b) (c)

Fig. 7: (a) The thermal distance-detection sensor setup and optical path. (b) Temperature mapping of the thermal distance-detection sensor. (c) Scene and reflected scene are captured in the same frame. The left half is from the scene and right half is from reflected scene. Note that the reflected temp. and converted temp. are in raw linear thermal camera measurements, and can be converted to conventional temperature with a calibration step.

In this experiment, a single planar mirror was used. The 75mm$\times$100mm protected gold flat mirror gave over 96% reflection for the wavelength of 700nm-$10,000$nm. The mirror was placed parallel to the optical axis of the lens with the gold coated surface facing to the incident light, which is the only design for single mirror catadioptric system. The size and position of the mirror also influence the field of view. The thermal camera involved in the experiment is FLIR A6751sc with a 13mm f/2.5 thermal lens sensitive to 3-5μm wavelength. The images of scene and mirror reflected scene could be shown in one picture that the camera captured (shown in figure 8). For the measurements, we use a non-contact infrared thermometer (LaserGrip 630 from ETEKCITY) to read the actual surface temperature of object and mirror in Celsius degree. And use the raw numbers(under default setting) that were given by Software ResearchIR from FLIR company to note the object temperature and reflected temperature.

Temperature Mapping: To generate the mapping relation, we give the definitions of the temperatures. We defined object temperature T_o as the temperature given by the camera, which is a unitless quantity. Then reflected temperature T_r was also a unitless number from the thermal camera that represents the temperature of object reflected by mirror. And mirror temperature T_m was the body temperature (in°C) of mirror that was measured by infrared thermometer. The temperature mapping could be shown as

$$T_o = f(T_m, T_r). \tag{6}$$

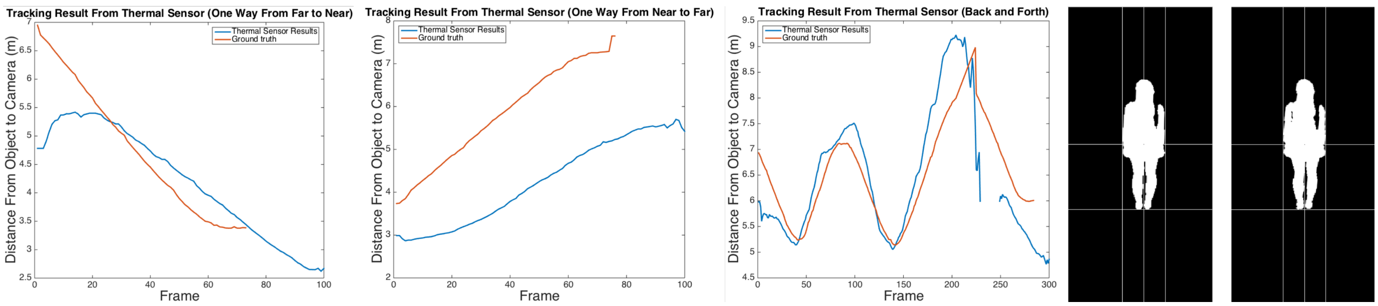
Then 70 measurements were taken, and 50 of them were randomly selected as examples while the others were used as validations. For each measurement, we recorded the values of T_o, T_r and T_m. The fitting surface with lowest mean square error was given in Fig.7 and we achieved this with a quadratic in two variables. With this calibration surface, we could derive any one of the three factors from the other two. For example, we use T_m and T_r to compute 'converted temperature' $T_o^{'}$. It should be represented as

$$T_o^{'} = p_{00} + p_{10} * T_m + p_{01} * T_r + p_{11} * T_m * T_r + p_{02} * T_r^2. \tag{7}$$

where p_{00}=6.207e_4, p_{10}=-2925, p_{01}=-2.737, p_{11}=0.1767, p_{02}=1.339e_{-6}.

Camera calibration and Person Segmentation: Before calibration, we note, as shown in Fig.7, scene and reflected scene were captured in the same frame. The ratio of scene to reflected scene did not equal 1. So we manually set the boundary to split the image in two. The left image was cropped to the same size as the right image. For calibration, a 5×4 hollowed-out checkerboard was 3D printed for the camera calibration, with 40mm edge length for each square. The checkerboard was chilled in a freezer so it could be imaged by the thermal sensor. Then Matlab Stereo Camera Calibration app was introduced to generate the intrinsic and extrinsic parameters for the stereo thermal system. Person segmentation was performed by assuming that body temperature fluctuates within a certain range. By looking for pixels within this range of temperatures, we can easily and accurately segment the human from the background, as the background will be unlikely to contain pixels within the human body temperature range. We used a mean filter to remove noise. We conducted background subtraction for the image and flipped reflected-image to reduce unnecessary noise. Then input the flipped reflected image as T_r and mirror temperature T_m to Eq.7 to obtain a converted image. The significance of this step was to reduce the influence of the mirror, so the reflected temperature could be converted to a similar range with object temperature. Then according to T_o for human in the original image (left half of Fig7 (c)), we set the thresholds of human body temperature.

Evaluation: After the segmentation, the rectified images (Fig.8(d)) were generated. We used the centroid of the temperature-segmented human for tracking. The distance from human to camera could be obtained by the well-known rectified stereo equation $D = \frac{B*C}{d}$, where B and C represented baseline and focal length that were included in the parameters from camera calibration, and d is the centroid disparity. We used a moving average filter to smooth the curve and obtained the results shown in Fig.8(a)-(c). We compared the thermal distance-detection sensor with LIDAR, which we took as the



(a) (b) (c) (d)

Fig. 8: (a) Comparison between the results from thermal distance-detection sensor and LIDAR. Person walked from far place to near. (b) Person walked from near place to far. (c) Person walked back and forth, and disappear when behind a pillar which cause the absence of thermal data. (d) Image and reflected image are captured in the same frame.

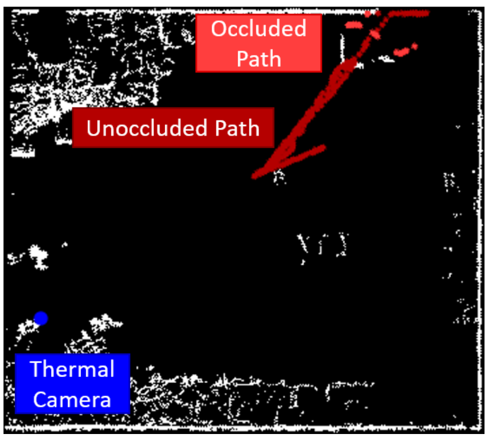


Fig. 9: This figure shows the results of blind tracking using the thermal stereo instead of the LIDAR. The results are represented by a top-down view of the room. Blue area shows where the thermal camera is located. The dark red shows the path of the source carrier while within the field of view of the thermal camera. The light red shows the path of the source carrier while occluded from the thermal camera.

ground truth. The errors for each measurement were 0.9m, 1.2m and 0.6m respectively, averaged over the length of the experiment. In the next section we evaluate the thermal stereo pair for detecting radiological sources. Despite the relatively large error in the measurements when compared to LIDAR, the sensor is able to allow for robust blind tracking since the depth ordering and trends in the measurements are correct.

B. Thermal Blind Tracking

We applied the blind tracking algorithm described previously with the thermal stereo camera system, as in Fig. 9 (Cf-252 source, 64 μCi). The calibration of the entire setup was done with a one-time single static LIDAR sensor scan that was not used in subsequent measurements of the dynamic scene. As before the radiation model λ for the room is estimated from the depth information provided by the thermal stereo camera, when the source carrier is visible. When the source carrier exits the thermal camera's field of view, the source's location can still be acquired by using the radiation model of each detector to calculate a distance from each detector to the source based on the number of counts each detector receives. These radial distances allow the inference of source location, using the Hampel filter [51] to replace outliers. We leverage prior knowledge of the room to remove impossible motion estimate that go beyond the room walls. Figure 9 shows a person walking in a straight line behind a small vertical obstacle in the top right of the scene. We are able to reliably track the person through the obstacle and back along the same path.

VI. Motion Graph Visualization of Blind Tracking

In security and military applications, it may be useful to provide visualizations of the blind tracking described previously to assist with suspect apprehension. This requires rendering realistic looking and believable motion of the person carrying the source that also is faithful to the blind tracking behind the obstacle. In this section, we achieve this by building a motion graph from a library of poses and actions. The key assumption is that the person has been imaged from a significant number of views and poses prior to occlusion.

Traditional motion graphs use high-fidelity 3D motion capture data. In practice such data might be obtained with high-resolution arrays of time-of-flight cameras. Even the LIDARs in our experiments are too sparse to use for conventional motion graph algorithms. Instead, we demonstrate how to use sparse 3D tracking information to reconstruct a motion graph of 2D RGB images. We fuse each image captured prior to occlusion with a 3D feature vector containing the object centroid and current velocity (obtained from the Kalman filter).

The idea of a motion graph was proposed by Kovar [32] and aims to automatically generate believable human motion that follows a given path through 3D space by using 3D motion capture data. We have modified the original design so it can be used with 2D images instead of 3D motion capture data. By using a traditional RGB camera with the LIDAR, we are able to capture pose information with the camera, and position/velocity information with the LIDAR. We collect 2D RGB image data containing pose information as well as the person's direction and velocity. With all of this information, we can generate a motion graph to reconstruct

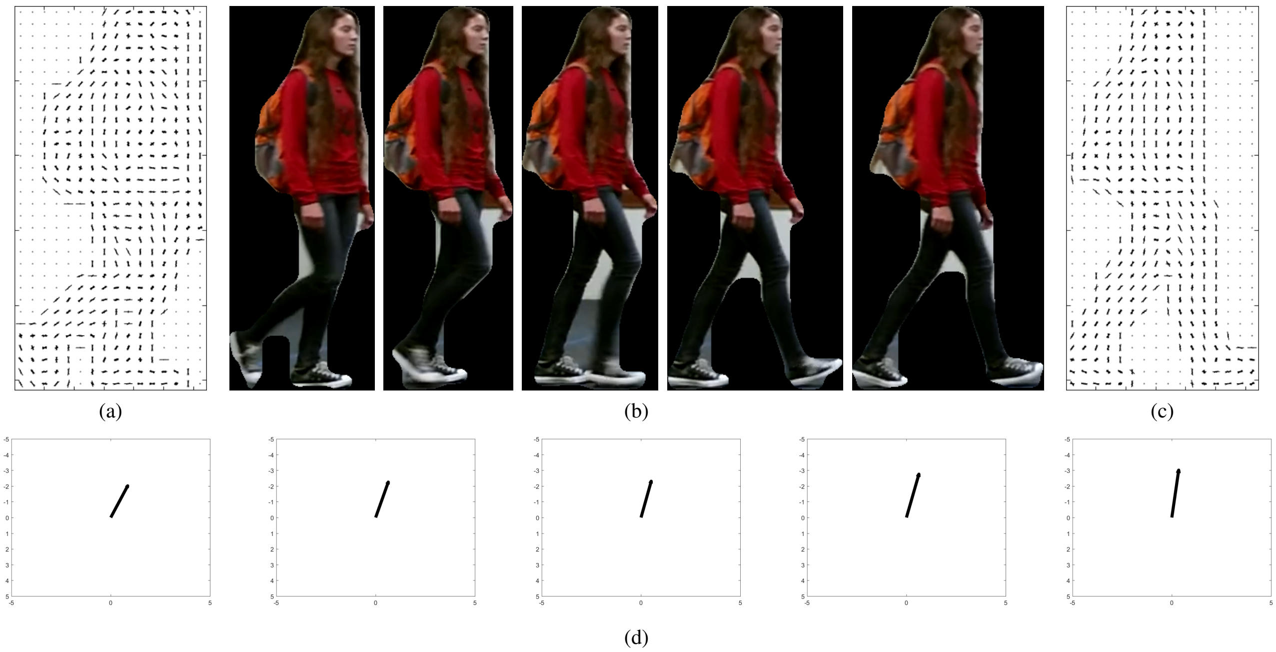
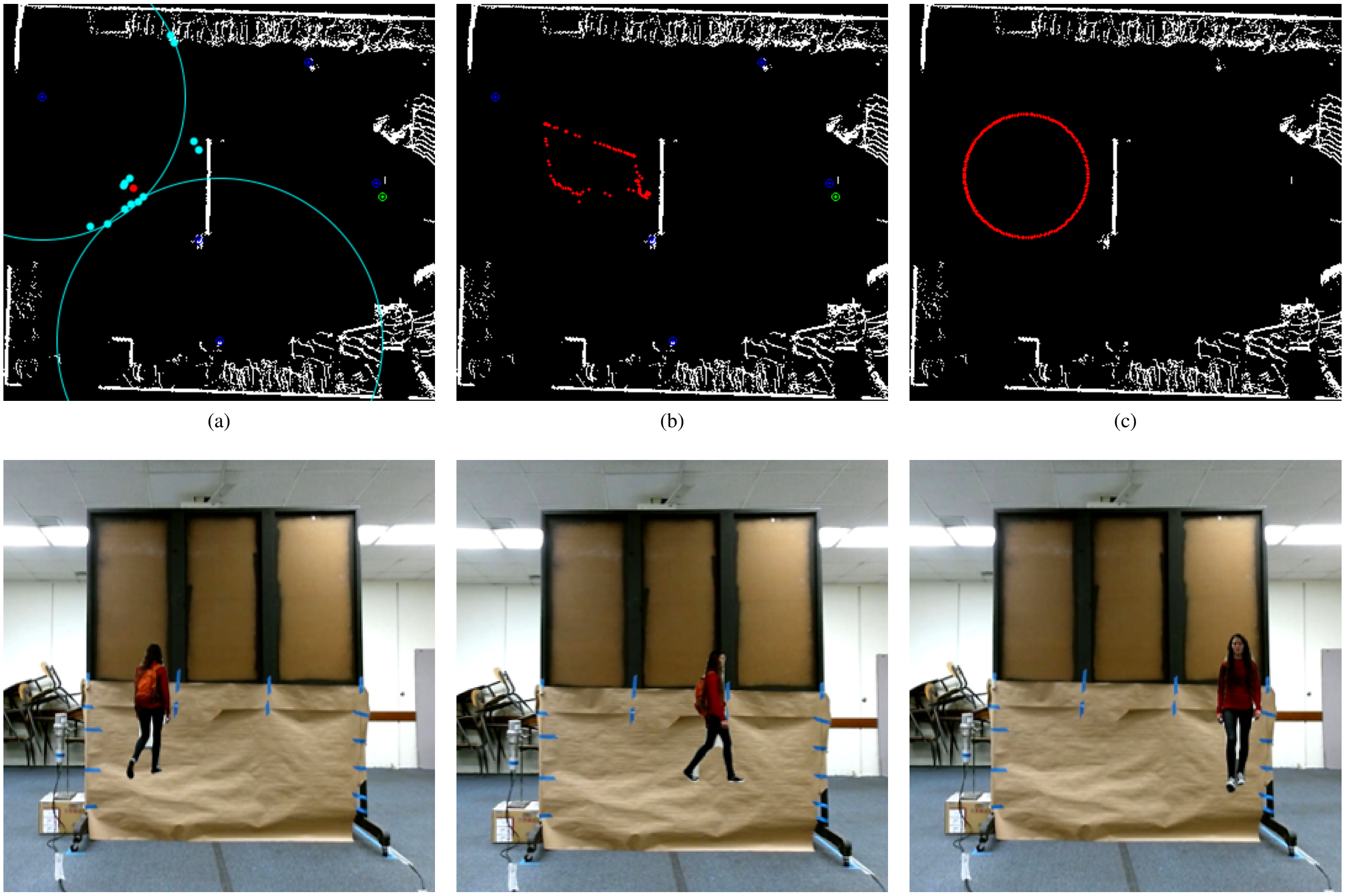


Fig. 10: A visual representation of part of the motion graph. (a-d) depict a single edge connecting two nodes. (a) is a visual representation of the HOG features that describe the beginning node. (b) contains the five frames of animation that connect the two nodes. (c) is the visualization of the HOG features that describe the termination node. (d) represents the velocity and direction of motion in each frame of the edge animation that connects the two nodes.

believable human motion. We use background subtraction to remove the person of interest from the background in the RGB video, and then break up the data into the two necessary sets for creating the motion graph; nodes, and edges. Each node contains the direction and velocity information of the person at that time instance, but also contain the person's pose information obtained by These key frames are run through a histogram of oriented gradients (HOG) feature extractor to quantify the pose information of the node as others have done for human pose detection, examples are [52], [53]. Each node is a HOG representation of a single key frame and each node contains direction and velocity information of the moving person captured by the depth sensor in that frame, as shown in Fig. 10. The orientation and motion information related to that single frame. The nodes are linked by edges containing animation information (i.e. a video clip) along with the person's velocity and direction in each frame. These sets of velocity and direction information form short paths that represent the path taken by the person, and these paths are linked to short animation information (i.e. videos). To generate motion, we search for paths that best describe the input path while simultaneously optimizing pose similarity between the nodes. Between nodes (video clips) the start and ending frames provide a distance metric based on the similarity between their HOG-based feature vectors. The motion graph is constructed from these nodes and edges and, to generate motion, the search algorithm attempts to reconstruct a given path through the scene in a least squares sense.

We consider the blind tracked path obtained from the radiological detectors as the desired goal, and the motion graph search algorithm reconstructs a path using the velocity and location information contained in each node. For each segment in the blind track, the match in the motion graph was made using cross-correlations of person velocity across the node or video clip. Simultaneously the pose data stored in the RGB image at the beginning of each blind track segment was converted in HOG and compared to the HOG of the person in the previous segment. The path information from the motion graph was then used to calculate the new position of the character. Frame skipping and frame repeating was used in cases where the motion graph data could not match the track closely, to avoid visual artifacts.

For the motion graph to perform its visualization, it must have an input path to attempt to reconstruct. We recover the person's location from behind the vertical white occluder using the radiological data from the detectors. The location of the person carrying the source is found by intersecting the R_{rad} from each radiological sensor in the room. Recovering the intrinsic efficiency $\lambda(R)$ at each sensor is accomplished on the fly when the source is visible to the depth sensor. Since the mapping of intrinsic efficiency $\lambda(R)$ is not calculated beforehand, we are unlikely to obtain a complete mapping of how intrinsic efficiency varies with R for the entire room. This mimics a real-world tracking scenario, where we have no prior knowledge of the source being trafficked and the only way to find intrinsic efficiency is measuring it on the fly. The incomplete mapping of intrinsic efficiency means that there will be uncertainty in the location data, which causes there to be noise in the location data. To minimize the effect of this noise, we use the Hampel filter to identify and remove outliers, and then average the remaining location points to recover an approximation of the source location. For this tracking, the



(a) (b) (c)

(d) (e) (f)

Fig. 11: We recover the person location behind the vertical white occluder in (a)-(c). In (a), the location of the person carrying the source is found by intersecting 6 radiological sensors. Since the mapping of $\lambda(R)$ for each sensor is performed on the fly when the source is visible to the depth sensor, there is noise in the recovered candidate intersection points, shown in cyan. The red dot is the mean location of the source once outliers have been replaced with the Hampel filter. In (b) the tracked locations of the radiological source as it moves behind the visual occlusion in each frame are shown in red and can be seen to form a closed shape. In (c) a circle is fitted to the source locations to generate a smoother closed shape for use in the motion graph. (d,e,f) are 3 frames of the video created by the motion graph showing the source carrier walking behind the visual occlusion. Please see the full video on our website [12].

resulting source locations form a closed shape, but using a motion graph to reconstruct this noisy path without smoothing will result in bizarre motion that is difficult for a human to interpret visually. A circle is fitted to the source locations to generate a smoother closed shape path to use as input to the motion graph. There is a visual representation of this process shown in figure 11. For video results, please see the full video on our website [12].

The motion graph must provide an accurate visual representation of the source's motion if it is going to be a useful tool for security personnel. To explore the accuracy of the motion graph, we seek to measure how accurately the motion graph reconstructs an input path. There are two metrics we would like to measure for the reconstructed path. The first is how well the motion graph reconstructs a noiseless path; this will be measured by cross correlation between the ground truth input path and the output path, and will show how accurately the reconstructed path resembles the ground truth path taken by the source. The second metric is how well the motion graph reconstructs an input path with added Gaussian noise; this will show how accurately the motion graph reconstructs the input path given to the algorithm. The goal here is to see the effect of noise in our radiological measurements, and the robustness of the motion graph as a smoothing operator. After fitting a circle to the closed-shape radiation data, we pass it as the input path to the motion graph, varying the Gaussian noise added to the input path, and calculating the cross correlation. Refer to figure 12 for the results of this experiment. The result is that the cross correlation between the reconstructed path and the noisy input path is near 1 at all SNR, meaning the motion graph is successfully reconstructing the input path, even at low SNR. The increase in cross correlation between the ground truth path and the output path means that the motion graph has trouble reconstructing the ground truth path if the input path is very noise (i.e. SNR $< 20dB$). The conclusion is that, while some noise can be compensated for, large errors in radiological sensor measurements result in large errors in blind tracking.

Experiments: For the motion graph experiments, we

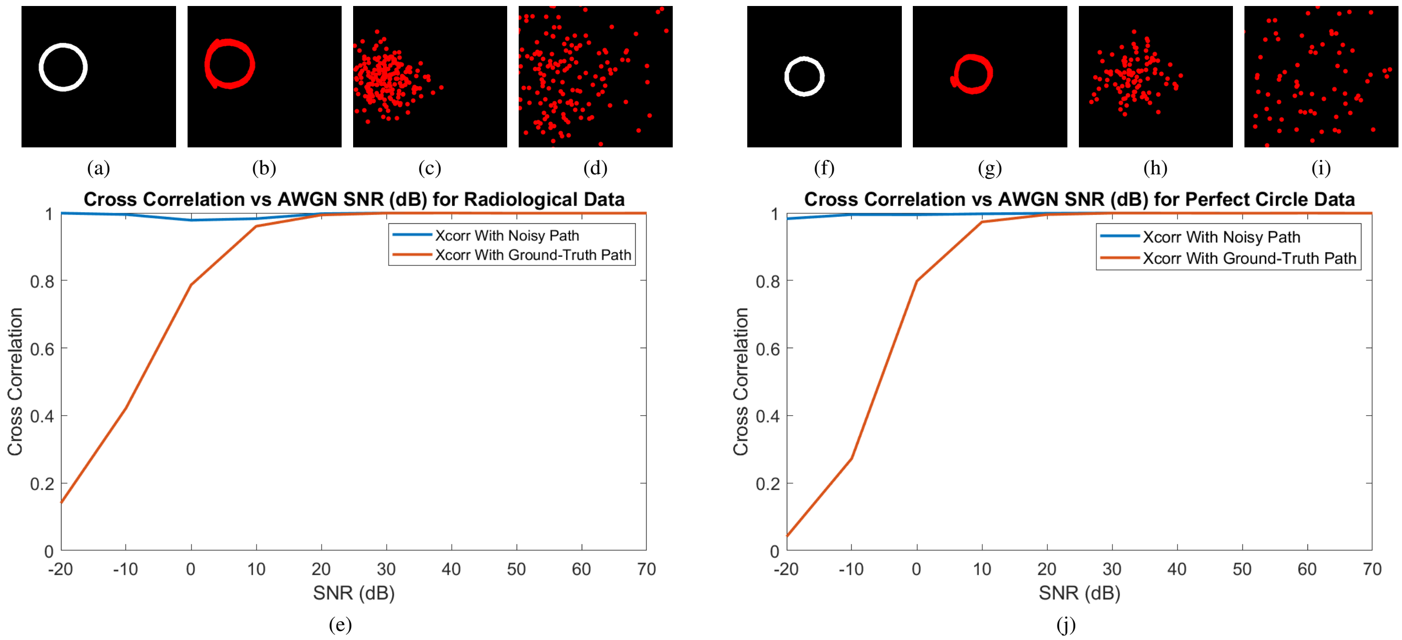


Fig. 12: This figure illustrates how well the motion graph reconstructs a given path. On the left are experiments where a circle has been fitted to real radiological data, and on the right is the pure simulation result. The goal here is to see the effect of noise in our radiological measurements, and the robustness of the motion graph as a smoothing operator. (a,f) shows the circle fitted to the radiation data, and (b,g) shows the reconstructed path with no noise added to the fitted circle. (c,h) shows the reconstructed path when noise has been added to the input path so that the SNR is 10 dB, and (d,i) shows the reconstructed path when the SNR is 0dB. The cross correlation where there is maximal pathway overlap is used to measure how well the motion graph reconstructs the given pathway. (e,j) shows a plot of the cross correlation between the reconstructed path, and either the fitted circle with added noise (Blue) or the ground truth fitted circle (Red) as SNR changes.

wanted to improve the SNR of the location data found using radiological measurements. We use the same method of blind tracking as in section IV, but we used the maximum number of EJ-309 detectors available to our lab (six) to gather much more location information. We utilized a Hampel filter ($\sigma = 1$) to identify and replace outliers in the positioning data before taking the average of the locations. Adding detectors will results in more location data contributing to the average, improving the SNR of the calculated source location. Conversely, using fewer detectors will result in fewer location data points and decrease the SNR of the calculated source location. A single detector can be used for blind tracking if we assume the source carrier does not deviate from their pre-occlusion path, but using 2 detectors is required to track the source if this assumption does not hold. Despite the resulting SNR improvement, we need further noise reduction in order for the blind track to drive a motion graph visualization. This is because the conversion from counts to distance is affected by dynamic shielding, as the relative room geometry around the moving source changes in the local frame of the source. While overcoming this can be a good direction for future work, here we constrain the blind track to remain inside the room boundary and further smooth the data by applying low parameter shapes (circle, triangle, square etc.) to the fitted blind tracks behind the obstacle, to improve the quality of the visualization.

Evaluation: In Fig. 12 we show evaluations of our method on simulated data. On the left are results for fitting a low parameter shape to radiological data, and on the right are results for a perfect circle with no radiological influence, to evaluate the motion graph approach. In these experiments we showed that the motion graph reconstruction is sensitive to noise and that at least 10dB SNR in the radiological counts are needed to recover the shape of the blind track effectively. In Fig. 11 and on our website [12] we show qualitative evaluations of our method on real radiological data from six detectors, where the person carrying the source walked in a closed path. The figure shows screenshots and the full video is available at [12].

VII. Conclusion and Limitations

Our work, including the previous conference publication [11], represents a novel approach to fuse depth sensors with isotropic radiological detectors. We are able to successfully localize and track a single moving source in 86.67% of the 15 experiments performed in our lab. We were able to successfully localize and track two moving sources in 100% of our 12 simulated experiments, and in 93% of our 12 real world experiments.

We now sketch proof-of-concept experiments performed at the Device Assembly facility, where we show the potential of our method for applications beyond homeland security, such as tracking in military and energy installations.

Exp	AIS	GT	Det	Cor	y/n
1	3	2	2	0.982	Yes
2	4	3	3	0.997	Yes
3	5	4	4	0.951	Yes
4	5	4	4	0.958	Yes
5	2	1	1	0.979	Yes
6	2	1	1	0.997	Yes
7	2	2	2	0.996	Yes
8	3	3	3	0.996	Yes
9	3	3	3	0.990	Yes
10	4	4	4	0.996	Yes
11	4	4	4	0.995	Yes
12	5	5	5	0.998	Yes
13	5	5	5	0.998	Yes

TABLE IV: Single source results from the DAF. AIS stands for "Actors in Scene" and indicates the number of people present in the scene during data collection. The trajectory indicated by GT (Ground Truth) is carrying the radiation source. The trajectory indicated by Det (Detected Signals) is *detected* to be carrying the radiation source. Cor indicated the correlation between the original radiological signal and the reconstructed signal. A correct detection is given by "Yes" in the y/n column.

A. *Towards military-grade tracking*

We took our experimental setup to the Device Assembly Facility in Nevada, USA to test with sources beyond our lab-grade Californium and Pu-Be sources. These lab sources do not represent the properties of the weapons-grade nuclear material we would expect a real-life trafficker to be carrying.

We performed experiments with our sensor fusion setup and captured radiological data from the Device Assembly Facility (DAF) in Nevada, using a weapons-grade radiological source and a larger space to perform experiments. Single source data was collected using the BeRP ball as the radiological source. **Evaluation:** The tracking results using the new sources and more people are in Table IV. Of the 13 single source experiments, the correct source was located, detected and tracked correctly in 100 percent of the experiments. This is an improvement over the results we obtained in the lab. The improvement is most likely caused by the increased source strength that comes from using weapons grade nuclear material and that results in higher SNR.

B. *Limitations*

The impact of this work is that such a combination can be used in airports, ports, commercial areas and battlefields with net security benefits. However, we point out a few limitations that we hope to address in future work:

Source strength model: In single source tracking we assumed the source strength λ is constant, and in the blind tracking experiments we assumed it was modeled with a radial dependency $\lambda(R)$. Both of these approximations can break with scenes that have more interesting geometries than our laboratory settings, with complex material properties. A future goal is to learn the scene radiation background prior, from the scene geometry itself.

NLoS imaging: We are confident that our motion graph-based NLoS imaging will provide compelling single image insights into what is happening behind the occluders. As we show in this paper, for simple shapes and trajectories, the video results are also insightful. However, for more complex trajectories, we would require depth information prior to occlusion, as in the original motion capture methods. In particular in this application it could allow tracking of covert material hand-overs between traffickers, without visual information, which is highly attractive. Finally, all our results assume the LIDAR does view the target at some point. If there is no viewing of the target at all, new algorithms incorporating NLoS imaging must be used.

Complex sources: Our algorithm assumes that the sources are point sources. With different radiological materials, or a partitioned source in several locations on a person or in luggage, these assumptions must change with new models.

Static source and switching sources: Our algorithms require moving radiation sources. A static source becomes part of the background and is removed by background subtraction. This may cause problems if people carrying the sources "drop off" the package. Of course, we can detect when the radiological counts become constant, and can pick up the radiological trajectory again as the source starts moving. This is a fascinating direction for future work.

Failure Cases: Single-source and multi-source tracking failure cases are caused when two object trajectories are similar. This happens due to the different selected sampling rate of radiological (1Hz) and vision (10Hz) sensors. Combined with Poisson noise in the radiation counts, this explains visual discrepancies in Fig.2 and Fig.4. Notice that the correct match still shares trajectory extrema (maxima and minima), which allows signal correlation to find the answer.

We would like to show in future work that our simplified forms of the counts equations allow us to know the linear relationship between radial distance $R(t)$ and detector counts $C(t)$, and radiological source position. Therefore, we can apply the principals from Arulampalam, et al [54] and find that the Kalman filter is the optimal estimation for the tracking of the moving source

REFERENCES

[1] T. B. Cochran and M. G. McKinzie, "Detecting nuclear smuggling," *Scientific American*, vol. 298, no. april, pp. 98 – 104, 2008. [Online]. Available: http://www.nature.com/scientificamerican/journal/v298/n4/full/scientificamerican0408-98.html

[2] "Application of nuclear forensics in combating illicit trafficking of nuclear and other radioactive material," in *IAEA TECDOC Series 1730*. IAEA, 2014.

[3] L. Mihailescu, K. Vetter, and D. Chivers, "Standoff 3d gamma-ray imaging," *Nuclear Science, IEEE Transactions on*, vol. 56, no. 2, pp. 479–486, 2009.

[4] A. C. Raffo-Caiado, K.-P. Ziock, J. Hayward, S. Smith, J. Bogard, and C. B. Boehnen, "Combining measurements with three-dimensional laser scanning system and coded-aperture gamma-ray imaging systems for international safeguards applications," in *2010 IAEA Symposium on International Safeguards*, 2010, pp. 1–5.

[5] K. P. Ziock, W. W. Craig, L. Fabris, R. C. Lanza, S. Gallagher, B. K. Horn, and N. W. Madden, "Large area imaging detector for long-range, passive detection of fissile material," *Nuclear Science, IEEE Transactions on*, vol. 51, no. 5, pp. 2238–2244, 2004.

[6] B. Bernacki, J. Schweppe, S. Stave, D. Jordan, J. Kulisek, T. Stewart, and C. Seifert, "Estimating radiological background using imaging spectroscopy," in *SPIE Defense+ Security*. International Society for Optics and Photonics, 2014, pp. 90 880L–90 880L.

[7] K.-P. Ziock, E. C. Bradley, A. Cheriyadat, M. Cunningham, L. Fabris, C. Fitzgerald, J. Goddard, D. E. Hornback, R. A. Kerekes, T. P. Karnowski *et al.*, "Performance of the roadside tracker portal-less portal monitor," *Nuclear Science, IEEE Transactions on*, vol. 60, no. 3, pp. 2237–2246, 2013.

[8] K. Stadnikia, A. Martin, K. Henderson, S. Koppal, and A. Enqvist, "Data-fusion for a vision-aided radiological detection system: Sensor dependence and source tracking," *EPJ Web of Conferences*, vol. 170, p. 07013, 01 2018.

[9] J. M. Hite, J. K. Mattingly, K. L. Schmidt, R. Ştefănescu, and R. Smith, "Bayesian metropolis methods applied to sensor networks for radiation source localization," in *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, Sep. 2016, pp. 389–393.

[10] P. Riley, A. Enqvist, and S. J. Koppal, "Low-cost depth and radiological sensor fusion to detect moving sources," in *3D Vision (3DV), 2015 International Conference on*. IEEE, 2015, pp. 198–205.

[11] K. Henderson, K. Stadnikia, A. Martin, A. Enqvist, and S. Koppal, "Tracking radioactive sources through sensor fusion of omnidirectional lidar and isotropic rad-detectors," in *3D Vision (3DV)*. IEEE.

[12] "3d vision and radiological sensor fusion." [Online]. Available: http://focus.ece.ufl.edu/3d-vision-and-radiological-sensor-fusion-2/

[13] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nature communications*, vol. 3, p. 745, 2012.

[14] K. L. Bouman, V. Ye, A. B. Yedidia, F. Durand, G. W. Wornell, A. Torralba, and W. T. Freeman, "Turning corners into cameras: Principles and methods," in *International Conference on Computer Vision*, vol. 1, no. 2, 2017, p. 8.

[15] D. B. Lindell, M. O'Toole, and G. Wetzstein, "Single-photon 3d imaging with deep sensor fusion," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, p. 113, 2018.

[16] H. Barrow and J. Tenenbaum, "Computer vision systems," *Computer vision systems*, p. 2, 1978.

[17] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison *et al.*, "Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 2011, pp. 559–568.

[18] D. Cooper, R. Ledoux, K. Kamieniecki, S. Korbly, J. Thompson, J. Batcheler, S. Chowdhury, N. Roza, J. Costales, and V. Aiyawar, "Intelligent radiation sensor system (irss) advanced technology demonstrator (atd)," in *Homeland Security (HST), 2012 IEEE Conference on Technologies for*, Nov 2012, pp. 511–516.

[19] J.-C. Chin, D. K. Yau, N. S. Rao, Y. Yang, C. Y. Ma, and M. Shankar, "Accurate localization of low-level radioactive source under noise and measurement errors," in *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '08. New York, NY, USA: ACM, 2008, pp. 183–196. [Online]. Available: http://doi.acm.org/10.1145/1460412.1460431

[20] A. Sundaresan, P. Varshney, and N. Rao, "Distributed detection of a nuclear radioactive source using fusion of correlated decisions," in *Information Fusion, 2007 10th International Conference on*, July 2007, pp. 1–7.

[21] A. Liu, J. Bunn, and K. Chandy, "An analysis of data fusion for radiation detection and localization," in *Information Fusion (FUSION), 2010 13th Conference on*, July 2010, pp. 1–8.

[22] M. Morelande and A. Skvortsov, "Radiation field estimation using a gaussian mixture," in *Information Fusion, 2009. FUSION '09. 12th International Conference on*, July 2009, pp. 2247–2254.

[23] N. S. Rao, S. Sen, N. J. Prins, D. A. Cooper, R. J. Ledoux, J. B. Costales, K. Kamieniecki, S. E. Korbly, J. K. Thompson, J. Batcheler, R. R. Brooks, and C. Q. Wu, "Network algorithms for detection of radiation sources," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 784, pp. 326 – 331, 2015, symposium on Radiation Measurements and Applications 2014 (SORMA XV). [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900215000686

[24] Z. Long, W. Wei, A. Turlapaty, Q. Du, and N. Younan, "Fusion of radiation and electromagnetic induction data for buried radioactive target detection and characterization," *Nuclear Science, IEEE Transactions on*, vol. 60, no. 2, pp. 1126–1133, April 2013.

[25] R. Barnowski, A. Haefner, L. Mihailescu, and K. Vetter, "Scene data fusion: Real-time standoff volumetric gamma-ray imaging," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 800, pp. 65 – 69, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016890021500950X

[26] D. Kim, H. Woo, Y. Ji, Y. Tamura, A. Yamashita, and H. Asama, "3d radiation imaging using mobile robot equipped with radiation detector," in *2017 IEEE/SICE International Symposium on System Integration (SII)*, Dec 2017, pp. 444–449.

[27] K. Vetter, R. Barnowski, J. W. Cates, A. Haefner, T. H. Joshi, R. Pavlovsky, and B. J. Quiter, "Advances in nuclear radiation sensing: Enabling 3-d gamma-ray vision," *Sensors*, vol. 19, no. 11, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/11/2541

[28] E. Hecht and A. Zajac, in *Optics*, 1974.

[29] A. Treptow, G. Cielniak, and T. Duckett, "Real-time people tracking for mobile robots using thermal vision," *Robotics and Autonomous Systems*, vol. 54, no. 9, pp. 729 – 739, 2006, selected papers from the 2nd European Conference on Mobile Robots (ECMR '05). [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0921889006000832

[30] Y. Li, T. Wang, and H.-Y. Shum, "Motion texture: A two-level statistical model for character motion synthesis," in *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '02. New York, NY, USA: ACM, 2002, pp. 465–472. [Online]. Available: http://doi.acm.org/10.1145/566570.566604

[31] J. Chai and J. K. Hodgins, "Performance animation from low-dimensional control signals," in *ACM SIGGRAPH 2005 Papers*, ser. SIGGRAPH '05. New York, NY, USA: ACM, 2005, pp. 686–696. [Online]. Available: http://doi.acm.org/10.1145/1186822.1073248

[32] L. Kovar, M. Gleicher, and F. Pighin, "Motion graphs," in *ACM SIGGRAPH 2008 Classes*, ser. SIGGRAPH '08. New York, NY, USA: ACM, 2008, pp. 51:1–51:10. [Online]. Available: http://doi.acm.org/10.1145/1401132.1401202

[33] O. Arikan and D. A. Forsyth, "Interactive motion generation from examples," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 483–490, Jul. 2002. [Online]. Available: http://doi.acm.org/10.1145/566654.566606

[34] L. Zhao and A. Safonova, "Achieving good connectivity in motion graphs," *Graphical Models*, vol. 71, no. 4, pp. 139 – 152, 2009, special Issue of ACM SIGGRAPH / Eurographics Symposium on Computer Animation 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1524070309000174

[35] P. Huang, A. Hilton, and J. Starck, "Human motion synthesis from 3d video," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, June 2009, pp. 1478–1485.

[36] H. H. Yang, S.-I. Amari, and A. Cichocki, "Information–theoretic approach to blind separation of sources in non-linear mixture," *Signal Processing*, vol. 64, no. 3, pp. 291–300, 1998.

[37] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[38] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE signal processing magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[39] S. J. Koppal and S. G. Narasimhan, "Clustering appearance for scene analysis," in *Conference on Computer Vision and Pattern Recognition*. IEEE, 2006, pp. 1–8.

[40] S. T. . Eljen Technology, P. O. Box 870, "http://www.eljentechnology.com/images/stories/Data_Sheets/Liquid_Scintillators/EJ309datasheet.pdf," 2012.

[41] C. Raghavachari, V. Aparna, S. Chithira, and V. Balasubramanian, "A comparative study of vision based human detection techniques in people counting applications," *Procedia Computer Science*, vol. 58, pp. 461–469, 2015.

[42] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[43] J. M. Pak, C. K. Ahn, P. Shi, and M. T. Lim, "Self-recovering extended kalman filtering algorithm based on model-based diagnosis and resetting using an assisting fir filter," *Neurocomputing*, vol. 173, pp. 645 – 658, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231215011558

[44] A. C. Raffo-Caiado, K.-P. Ziock, J. Hayward, S. Smith, A. Solodov, L. Mihailescu, K. Vetter, A. Dougan, M. Burks, J. G. Gonçalves *et al.*, "Investigation of combined measurements with three-dimensional design information verification system and gamma-ray imaging systems for international safeguards applications," in *Proceedings 50th INMM Annual Meeting*, 2009, pp. 12–16.

[45] K. Stadnikia, K. Henderson, A. Martin, P. Riley, S. Koppal, and A. Enqvist, "Data fusion for a vision-aided radiological detection system: Calibration algorithm performance," *Nuclear Instruments and Methods*

in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 890, pp. 8–17, 2018.

[46] J. J. Clark, "Active photometric stereo," in *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on*. IEEE, 1992, pp. 29–34.

[47] P.-N. Tan *et al.*, *Introduction to data mining*. Pearson Education India, 2006.

[48] A. Gunatilaka, B. Ristic, and R. Gailis, "On localisation of a radiological point source," in *2007 Information, Decision and Control*, Feb 2007, pp. 236–241.

[49] J. Gluckman and S. K. Nayar, "Catadioptric stereo using planar mirrors," *International Journal of Computer Vision*, vol. 44, no. 1, pp. 65–79, 2001.

[50] V. Da Deppo, G. Naletto, G. Cremonese, and L. Calamai, "Optical design of the single-detector planetary stereo camera for the bepicolombo european space agency mission to mercury," *Applied optics*, vol. 49, no. 15, pp. 2910–2919, 2010.

[51] R. K. Pearson, Y. Neuvo, J. Astola, and M. Gabbouj, "Generalized hampel filters," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, p. 87, 2016.

[52] G. Rogez, J. Rihan, S. Ramalingam, C. Orrite, and P. H. S. Torr, "Randomized trees for human pose detection," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, June 2008, pp. 1–8.

[53] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, June 2005, pp. 886–893 vol. 1.

[54] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb 2002.