This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNS.2017.2676045, IEEE Transactions on Nuclear Science

Uncertainty Analysis of Wavelet-Based Feature Extraction for Isotope Identification on NaI Gamma-Ray Spectra

J. Stinnett, M. Watson, C.J. Sullivan, and H. Xiong

***Abstract*—Low-resolution isotope identifiers are widely deployed for nuclear security purposes, but these detectors currently demonstrate problems in making correct identifications in many typical usage scenarios. While there are many hardware alternatives and improvements that can be made, performance on existing low resolution isotope identifiers should be able to be improved by developing new identification algorithms. We have developed a wavelet-based peak extraction algorithm and an implementation of a Bayesian classifier for automated peak-based identification. The peak extraction algorithm has been extended to compute uncertainties in the peak area calculations. To build empirical joint probability distributions of the peak areas and uncertainties, a large set of spectra were simulated in MCNP6 and processed with the wavelet-based feature extraction algorithm. Kernel density estimation was then used to create a new component of the likelihood function in the Bayesian classifier. Identification performance is demonstrated on a variety of real low-resolution spectra, including Category I quantities of special nuclear material.**

S : = CWT coefficient matrix
W : = Wavelet transform tensor
X : = Spectrum vector
E : = Wavelet centroid parameter
a : = Wavelet scale parameter
B : = Wavelet basis matrix
B_1 : = Optimal submatrix of B
k : = Fit vector, wavelet representation of X
C_S : = Signal covariance
C_k : = Fit covariance

I. Introduction

Since the September 11th terrorist attacks, the United States has spent on the order of $1 trillion on homeland security and anti-terror efforts; this includes approximately $150 billion on either failed projects or equipment that does not work as well as intended [1]. These failed projects range from $3.5 billion on motion sensor and camera networks along the US-Mexico border [2] to $230 million on prototype tests for new radiation portal monitors [3]. This also includes many millions spent on handheld radioisotope identifiers.

Handheld radioisotope identifiers (RIIDs) are used to detect and automatically identify radioactive material in a wide variety of settings. From shipping ports to explosive ordnance disposal to border security checkpoints, these identifiers are relied upon for security. Generally, these detectors use a sodium iodide (NaI) spectrometer because of their relatively low cost and reasonable efficiency over a broad range of field conditions.

There is a need to improve the identification performance of these detectors, ideally without substantially increasing their costs. Published evaluations of these commercial detectors have demonstrated that their identifications are not generally accurate [4], [5], [6]. Unfortunately, more recent evaluations of these commercial detectors have not been published, likely due to security or commercial concerns.

It has been suggested [4] that the focus for performance improvements should be the isotope identification algorithms used by these detectors. While the spectra they take are low-resolution, a trained spectroscopist can generally make highly accurate identifications using these spectra, implying that the ID algorithms are potentially problematic [7], [8]. The U.S. Department of Energy (DOE) requires trained spectroscopists to be on call at all times to analyze unknown spectra and resolve alarms [7], further suggesting that these identification algorithms are out-performed by knowledgeable humans.

The lack of recent and consistent performance benchmarks along with the relatively small number of papers published in isotope identification make it difficult to identify at particular state-of-the-art approach. Existing algorithms for automated isotope identification can roughly be categorized [7] into library comparison methods [9], [10], region of interest techniques, template matching [11], eigenvector methods (such as principal components analysis [12]), and more general machine learning techniques (such as neural networks [13], [14], [15]).

There are several challenges for any automated identification algorithm. First, these handheld NaI spectrometers have relatively low energy resolution, making it impossible to resolve many of the photopeaks in a spectrum [5] and limiting both the performance of the identification algorithms and the methods for identification that are even applicable [16]. While other scintillator materials have much better energy resolutions, their increased costs and, in some cases, their decreased absolute detection efficiency make them infeasible for the same wide-scale deployment as NaI detectors.

Second, NaI spectrometers are susceptible to temperature changes [17]. As the temperature fluctuates, the energy calibration will change and the centroid of the peaks will shift significantly. Some identification algorithms are unable to handle the errors in peak centroid detection, leading to incorrect

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNS.2017.2676045, IEEE Transactions on Nuclear Science

identifications.

Third, in most use cases for these detectors there is an unknown amount of background radiation from terrestrial naturally occurring radioactive material (NORM), from cosmic sources, and from some manufactured materials. This can obfuscate the spectrum and generally makes feature identification in a spectrum more difficult. Some methods for isotope identification rely on background subtraction, in which a background measurement is taken separately from the source measurement. This introduces further uncertainties in the process and is not always feasible for the implementations of these detectors.

Fourth, an issue for all isotope identification methods is that the source activity and attenuating materials between the source and detector are generally unknown. Attenuating materials have multiple effects on spectra. They reduce the total number of gamma-rays that reach the detector, making counting statistics worse and photopeaks harder to resolve. Shielding materials also preferentially attenuate lower energy gamma-rays, often to the point of removing lower-energy photopeaks from a spectrum. Finally, since the shielding between the detector and the source is generally not the same as the shielding between the detector and sources of background radiation, intermediate materials will generally decrease the signal to background ratio, further compounding correct identification.

Finally, handheld RIIDs have limited computational capabilities. These detectors are expected to be operable for several hours on battery power, and as a result are limited on the computational complexity of the identification algorithm. The limitation can be significantly worse for applications where a large isotope library is needed. Methods for identification that might be feasible with a desktop computer may be too time-consuming for portable applications.

The poor performance of modern isotope identification algorithms prompts the development of new approaches, particularly ones more similar to the process a spectroscopist would use to make manual identifications. To this end, we have previously proposed a feature extraction algorithm using a wavelet/non-negative least squares (NNLS) approach [18], [19], [20], [21] and a Bayesian classifier identification algorithm [22], [23] for isotope identification.

To improve this isotope identification methodology, a novel extension to the feature extraction algorithm to compute uncertainty in the area fit has been developed. While various methods for including this new information to the Bayesian classifier are being explored, an empirical method using kernel density estimation (KDE) is presented here. Finally, example identifications on a set of real sources are presented.

II. Feature Extraction with Wavelets/NNLS

Excellent introductions to wavelet analysis exist elsewhere [24], [25]; only a brief overview is presented here. The wavelet transform can be formulated as a convolution integral or as a tensor product. Both methods are presented for clarity. A wavelet $\psi(t)$ is a zero-area, square-integrable signal that is non-zero over a finite domain. An example wavelet (the 'bior2.6' wavelet, a member of the biorthogonal family [26]) is shown in Figure 1. The wavelet is shifted by parameter E and scaled by parameter a, as:

$$\psi_{E,a}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-E}{a}\right). \tag{1}$$

The continuous wavelet transform (CWT) of a signal $f(t)$ is the convolution integral of the signal and the wavelet:

$$T(E,a) = \int_{-\infty}^{\infty} f(t)\cdot\frac{1}{\sqrt{a}}\psi^*(\frac{t-E}{a})dt. \tag{2}$$

By computing the CWT of a spectrum over a range of scales and shifts, peak information can be extracted. Consider the perfectly Gaussian signal and its CWT in Figure 2. Maxima in the CWT are colocated with the peak. By finding the local maxima over the shift parameter, wavelet transform modulus maxima (WTMM) lines are formed. Finally, finding maxima along the WTMM lines gives peak width information.

In reality, some of the WTMM lines in a CWT are not useful for peak detection. There are many approaches for filtering out these unneeded maxima [20]. Most importantly, a photopeak in a spectrum that was collected with an NaI detector has a width that is determined by the photopeak's energy. Because this relationship is known, an optimal scale curve can be computed. The optimal scale $\alpha(E)$ is the scale that produces a maximum in the CWT of a perfectly Gaussian photopeak with mean E and full-width at half-max (FWHM) that matches the detector's FWHM vs E curve.

To get better estimates of peak centroid and to obtain area information, non-negative least squares fitting is performed with a predetermined basis matrix. However, the NNLS is more easily defined using the algebraic formulation of the wavelet transform approach. Let S be the (1024x256) matrix of values from the CWT of X. Then the wavelet transform can be expressed as

$$S_{i,j} = W_{i,j}X \tag{3}$$

where X is the spectrum (a 1024x1 vector), and $W_{i,j}$ is a 1x1024 row vector of the wavelet transform with scale index i and centroid index j. Compared to the integral formulation before:

$$S_{i,j} = T(E_i, a_j). \tag{4}$$

Next, define:

- $G_j :=$ a normalized (area equals one) Gaussian function with centroid (mean) in channel j and a standard deviation that matches the full width at half-max (FWHM) response of the detector. A 1024x1 vector.
- $[CWT(f)] :=$ the CWT of signal f along the optimal scale only. A 1024x1 vector.
- $B_{ij} := [CWT(G_j)]_i$, the basis matrix. For this work, B is a 1024x256 matrix [18].

Given the basis matrix B, NNLS is performed to find the fit vector k, whose nonzero elements ideally represent the identified peaks. In NNLS, a submatrix of B is defined by setting some columns of the basis matrix to zero to minimize the fit vector error subject to the nonnegative constraint. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNS.2017.2676045, IEEE Transactions on Nuclear Science

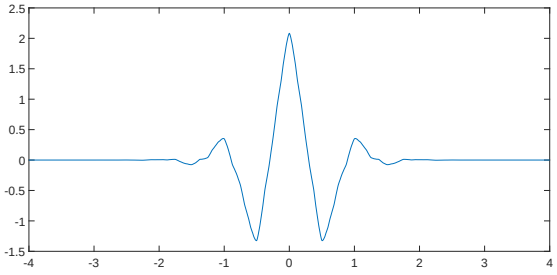

Fig. 1. The bior2.6 wavelet, or $\psi(\frac{t-E}{a})$.

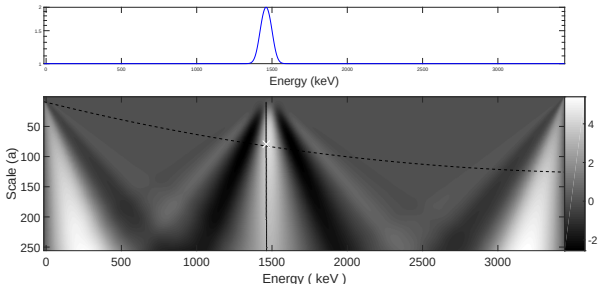

Fig. 2. Top: a Gaussian signal. Bottom: The wavelet transform of the signal. The WTMM line (solid) and the maximum along the WTMM line (white 'x') give information about the peak centroid and peak width. Comparing the scale at maximum transform value against the optimal scale curve (dashed) gives a sanity check for peak widths.

window of nonzero columns is chosen as a function of peak width. This window must be suitably wide to have enough peak information for accurate fits, but small enough that other spectral features (Compton edges, other significant peaks, etc.) are not included if feasible. Let B_1 be this submatrix of B.

$$Bk = S \quad (5)$$
$$B_1^T B_1 k = B_1^T S. \quad (6)$$

If the inversion exists, the solution can be written explicitly as:

$$k = (B_1^T B_1)^{-1} B_1^T S = OS \quad (7)$$
$$O = (B_1^T B_1)^{-1} B_1^T. \quad (8)$$

However, the matrix $B_1^T B_1$ is singular. To avoid this issue, truncated singular value decomposition (TSVD) is used here. In TSVD, B_1 is decomposed as UEV^T, where U and V are unitary matrices and E is a diagonal matrix of non-negative real numbers. The pseudoinverse of B_1 is then VE^+U^T, where E^+ is a pseudoinverse of E. The diagonal elements of E^+ are simply the reciprocal of the corresponding elements in E. In TSVD (vs the usual singular value decomposition), these diagonal elements are set to zero if they are below a user-defined threshold.

$$E^+_{i,i} = \begin{cases} E^{-1}_{i,i} & \text{if } E^{-1}_{i,i} \geq 0.01 \\ 0 & \text{if } E^{-1}_{i,i} < 0.01 \end{cases} \quad (9)$$

Here, 0.01 is a manually defined threshold. Finally, the fit vector is computed by:

$$B_1 = UEV^T \quad (10)$$
$$k = VE^+U^T S. \quad (11)$$

A. Area Uncertainty

To augment the isotope identification algorithm, the variance of the fit vector k is computed. This information can be used as a peak filter (e.g. reject peaks with peak uncertainties over a threshold) or as additional information in an isotope identification algorithm (e.g. weight peaks with a lower uncertainty more than high uncertainty peaks in a peak scoring algorithm). Let C_S be the covariance of the coefficient matrix S and C_X be the covariance of the signal X. If S is uncorrelated, C_S can be calculated simply as:

$$C_s = \sigma_s^2 I \quad (12)$$
$$\sigma_s = \frac{1}{m-n} S^T (I - B_1 O) S. \quad (13)$$

where $m-n$ is the degrees of freedom of the submatrix B_1 and σ_s^2 is the variance of the signal. This variance is primarily due to Poissonian noise, and can be estimated using the coefficients of S at small scales.

However, C_S is generally not uncorrelated and will be a function of the covariance of the signal C_X. Suppose $C_X = I\sigma_n^2$. Then:

$$C_S = W C_x W^T \quad (14)$$
$$C_k = O C_s O^T \quad (15)$$

and the variance of vector k is $diag(C_s)$.

In this approach, it is possible that the fits obtained through the NNLS process are not accurate representations of true peaks. In this case, the area uncertainties calculated may not have the same interpretation due to incorrect assumptions in their derivation. However, in practice when this happens the computed uncertainties are generally significantly larger than those for true peaks. These uncertainties can then be used as an indicator that a fit is not good, and the poor fits can be rejected with a threshold filter.

III. Feature Extraction Demonstration

To demonstrate the feature extraction code, a ^{154}Eu spectrum with excellent counting statistics and a ^{152}Eu spectrum with poor counting statistics were processed. These two spectra are shown in Figure 3.

First, the smooth spectrum of ^{154}Eu was processed for peak information. ^{154}Eu has a over a hundred gamma-ray emissions, only a few of which are resolvable with an NaI detector. A comparison of the larger gamma-ray emissions [27] and the peaks produced by the feature extraction code are shown in Table I. With the exception of the 38.6 keV peak, the uncertainties of all of the detected peaks are small. For the 38.6 keV peak, a larger uncertainty was found due primarily to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNS.2017.2676045, IEEE Transactions on Nuclear Science

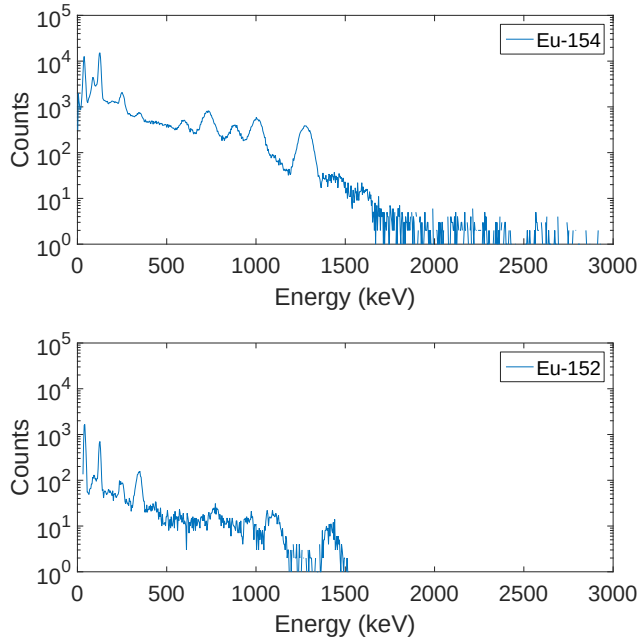

Fig. 3. A 60 s spectrum of ^{154}Eu (top) and a 3 s spectrum of ^{152}Eu (bottom) used to demonstrate the feature extraction performance.

the small fitting window and the low peak area to background ratio at these low energies.

Of the peaks that were not detected, the 756.8 keV peak is overshadowed by the much larger 723.4 keV. The 996.3 keV peak is too close to the 1004.8 keV peak to be resolvable with an NaI-measured spectrum. Finally, the highest energy peaks were not detected, though a peak around 1595 keV is visible to a human observer in the spectrum.

Next, a 3 second spectrum of ^{152}Eu is analyzed for peak information. Similarly to the previous example, ^{152}Eu has 138 gamma-ray emissions above 50 keV, but most of these are not resolvable with an NaI detector. In this example, much larger uncertainties are observed. First, while a peak at 569.3 keV was reported, the uncertainty of 50.2% is large enough to suggest it may not be a real peak. However, a peak was also detected at 964.5 keV with an uncertainty of 47.3%, and this peak clearly does correspond to a true peak. In this case, the uncertainty alone is not a suitably informative indicator to determine the validity of peaks without other information.

TABLE I
EU-154 FEATURE EXTRACTION RESULTS

Expected Energy	Branching Ratio	Fit Energy	Fit Area	Fit Uncertainty
43.0	0.132	38.6	61079	0.306
123.1	0.405	123.7	1.24E5	0.153
248.0	0.066	249.5	54500	0.012
-	-	346.8	26016	0.010
591.8	0.0484	594.6	22059	0.009
723.4	0.197	732.4	26645	0.012
756.8	0.0434	-	-	-
996.3	0.1029	-	-	-
1004.8	0.179	1004.0	17189	0.027
1274.5	0.355	1273.5	11350	0.005
1593.0	0.0103	-	-	-
1596.6	0.0183	-	-	-

IV. MODIFYING THE IDENTIFICATION ALGORITHM

Previously, we developed a Bayesian classifier approach to isotope identification [23], [28], [29]. Using the peak energies and areas found with the feature extraction code, this algorithm scored each of these features against an isotope library containing peak energies and branching ratios only. These scores are combined to compute the likelihood $P(D|M_i)$ of data D for each model M_i in the isotope library.

After computing the likelihoods for each model, the posterior probability $P(M_i|D)$ of model M_i is computed with Bayes' Theorem:

$$P(M_i|D) = \frac{P(D|M_i)\pi(M_i)}{\sum_{i=1}^{m} P(D|M_i)\pi(M_i)}. \tag{16}$$

where $\pi(M_i)$ is the prior probability of model M_i, or how probable that model M_i is correct without accounting for the data D. There are many options for constructing the prior. Currently, a simple prior that depends only on the number of isotopes in a model is used, with the prior decreasing as more isotopes are added to a model. This prior does not weight any combination of isotopes more strongly than another combination of the same number of isotopes. While this prior is almost non-informative, it is used to demonstrate that identification results are not relying on a prior distribution that unfairly biases it towards the desired answer.

This yields a probability measure for each model in the library. The scoring functions are built such that the models are orthogonal, e.g. $P(D|M_1$ and $M_2) = 0$. Practically, this means that if M_1 is the model of a single source of ^{241}Am, then $P(M_1|D)$ is the probability that ^{241}Am is the only source present for the spectrum measurement.

To account for the area uncertainty information, the previously-developed likelihood calculation must be modified [29]. The likelihood is approximated as the product of four scoring functions:

$$P(D|M_i) \approx f_{LR} f_{DR} f_{PP} f_{AR}. \tag{17}$$

TABLE II
EU-152 FEATURE EXTRACTION RESULTS

Expected Energy	Branching Ratio	Fit Energy	Fit Area	Fit Uncertainty
121.8	0.284	121.2	4777.1	0.104
244.7	0.075	250.6	1760.1	0.047
344.3	0.266	342.5	2392.6	0.032
411.1	0.022	-	-	-
443.9	0.028	-	-	-
-	-	569.3	562.7	0.502
778.9	0.130	772.3	667.7	0.156
867.4	0.042	-	-	-
964.1	0.145	964.5	477.6	0.473
1085.9	0.099	-	-	-
1089.7	0.017	1096.6	595.2	0.048
1112.1	0.136	-	-	-
1408.0	0.208	1416.1	316.9	0.073
1457.6	0.050	-	-	-
1528.1	0.027	-	-	-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNS.2017.2676045, IEEE Transactions on Nuclear Science

Each of these scoring functions considers a different subset of the data and is described in depth elsewhere [23], [29].

- f_{LR}: scores which peaks in the library model M_i are possibly represented in the dataset D
- f_{DR}: scores which peaks in the data D are possibly represented in the library model M_i
- f_{PP}: scores how well the data peak energies match the library peak positions in energy
- f_{AR}: scores the peak areas by examining the ratios of areas of neighboring peaks

The incorporation of the area uncertainty information involves reconstructing only one term, f_{AR}. The modularity of this approach to isotope identification is a major advantage; it allows for simple modifications to the algorithm to account for new information.

A. Constructing f_{AR}

In the identification algorithm, peak area information is incorporated by scoring the ratios of peak areas of neighboring peaks. Using the area ratios largely eliminates the need for information about the source activity and geometry and reduces the effects of shielding somewhat when the peaks are close in energy.

Let A_1 and A_2 be the peak areas found with the wavelet/NNLS algorithm with peak energies E_1 and E_2 respectively ($E_1 < E_2$), and let u_1 and u_2 be the corresponding area uncertainty percentages. Then the area ratio $r = A_1/A_2$ is scored against the expected area ratio R. In an ideal spectrum with perfect counting statistics and no attenuation, the observed ratio r should be the same as the expected R. In reality, intermediate materials will reduce this ratio, the significance of which depends on the materials and the peak energies. Poor signal-to-noise ratio (SNR) can dramatically increase or decrease the observed R, depending on the areas of each of the peaks.

Ideally, the effect of the SNR will be captured in the peak uncertainties u_1 and u_2. For small uncertainties, this effect should be minimal, and the observed ratio should be the same (little effect from shielding) as the expected ratio or be somewhat less (more effect from shielding materials). When u_1 is large but u_2 is small, only some additional deviation in the observed r is reasonable, as varying A_1 by 20% changes r by 20%. However, when u_2 is large, the observed ratio may be wildly different than the expected ratio, and the peak area information should only contribute weakly to the isotope identification.

To build this scoring function, 200,000 spectra were simulated in MCNP6. The simulation geometry assumed a model of an Ortec 905-3 NaI detector [30]. For each simulated spectrum, random energies and branching ratios were sampled from uniform distributions, and a point source with these characteristics was created. Cosmic and terrestrial backgrounds were not included in the simulations. Including these backgrounds would either increase the area uncertainties slightly or, in some cases of source photopeaks overlapping with background peaks, cause peaks to be missed by the wavelet/NNLS.

The wavelet/NNLS algorithm then extracted peak information from each of these simulated spectra. For each spectrum, this consists of a pair of peak energies, areas, and area uncertainties.

Finally, the probability density was estimated using kernel density estimation (KDE). During isotope identification, this expected area ratio R is the ratio of the branching ratios for the two photopeaks, convolved with detector efficiency. However, at this stage, the expected area ratio R is known from the simulation stage.

KDE is used to directly estimate the probability $P(r, u_1, u_2, R)$. However, an isotope should not be penalized for the area uncertainties, but instead should be evaluated on the observed data ratio r as a function of the area uncertainties and the library ratio. Further, it is actually desired $\max f_{AR}(r, u_1, u_2, R) = 1$ instead of being a normalized probability density (integrating to one). An arbitrary scaling factor can be introduced without breaking the identification algorithm; with the normalization term in Bayes' Theorem (Equation 16), the normalization constant will drop out. This scaling factor makes tuning the individual terms in the likelihood estimation easier, but could be excluded if desired.

The area ratio score is set to be proportional to the conditional density $P(r|u_1, u_2, R)$:

$$f_{AR}(r, u_1, u_2, R) := \frac{P(r|u_1, u_2, R)}{\max_\rho P(\rho|u_1, u_2, R)}. \qquad (18)$$

An example of this scoring function are shown in Figure 4. Ultimately, the shape of this function is similar to previous models for the area ratio [23], [29], but the area uncertainties can scale and distort the score. In general, if the observed area ratio r is close to the expected (library) area ratio R, a high score (no penalty) is assigned. As the difference between r and R increases, the score decreases (penalty increases). The penalty is much more severe for $r \gg R$ than for $r \ll R$; the effect of attenuating materials is to reduce the observed ratio r because lower energy peaks are more strongly attenuated than high energy peaks.

It is important to note that while this scoring function was constructed in the absence of background radiation, it can be used for identifications on spectra with the backgrounds included. The true distribution for area ratios and uncertainties depends on many unknown variables, such as the geometry, background concentrations of each NORM isotope, and more. The empirical distribution found here is essentially a first order approximation, where all of these unknown variables are fixed. The significance of these fixing these variables and the effects on the empirical distribution are to be explored in the future.

V. Example Identifications

Example spectra have been collected with an Ortec 905-3 2"x2" NaI detector; spectra are shown in Figure 5, and identification results are shown in Table III. Background subtraction was not performed on any of these test cases; this often leads to the identification of ^{232}Th decay chain and/or ^{40}K in addition to the radiation source of interest. The backgrounds are also not the same across all test cases; the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNS.2017.2676045, IEEE Transactions on Nuclear Science

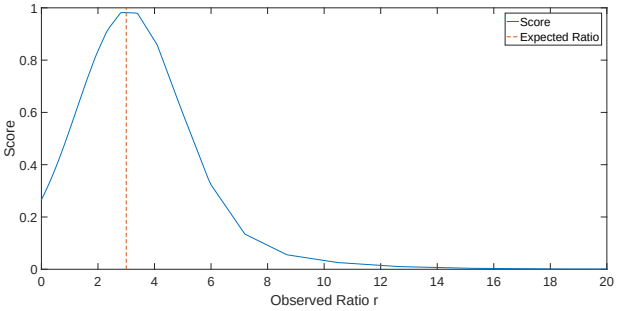

Fig. 4. An example of the area scoring function for library ratio $R = 3$ and $u_1 = u_2 = 10\%$. If the observed ratio r was the same as the library ratio R, a high score (no penalty) would be assigned. However, if the observed ratio is significantly different, a lower score (larger penalty) is given, with the severity dependent on the difference between the library and observed ratios.

first two spectra were measured in Los Alamos, NM, while the latter three were measured at the Nevada National Security Site. Identification results on test spectra are promising, and suggest several routes for algorithm improvement.

In the first test case, the large 136.0, 264.6, and 400.7 keV peaks of ^{75}Se are easily detected. The smaller peak at 303.9 keV was also detected; however, the area of this peak was found to be larger than expected. Because ^{67}Ga also has a peak at 300.2 keV, the combination of ^{67}Ga and ^{75}Se would result in a larger peak at 300 keV, which lead to a partially incorrect identification overall. A more accurate fit of the 303.9 keV peak would improve this result, as would a more informed prior distribution. Alternatively, a cost function that more strongly penalizes adding extra isotopes would likely fix this test case.

In the second example, in addition to the correct source identification, ^{232}Th is predicted. While ^{232}Th is often correctly identified as a component of terrestrial background radiation, the production of ^{233}U includes a small ^{232}U impurity due to various (n, 2n) reactions, which eventually decay to ^{232}Th. This leads to a higher concentration of ^{232}Th and its decay chain in the spectrum.

In the last three examples, the sources have significant self-attenuation due to their size, as well as additional shielding materials in the third and fifth examples. While the shielded cases were identified correctly, the unshielded weapons-grade plutonium (WGPu) also included a false identification of ^{125}I in addition to ^{241}Am and ^{239}Pu due to the peak at approximately 26 keV. ^{125}I has significant emissions at 27.2, 27.4, 31.0, and 31.9 keV, which the identification code considers a reasonable match for 26 keV. This would be corrected in hardware by increasing the lower-level discriminator to approximately $40 - 50$ keV.

There are other areas for improving identification performance. First, further tuning of the parameters of the wavelet code should be explored in order to improve sensitivity to small or overlapping peaks and to reduce false alarms.

Second, further model refinement for the likelihood calculation in the isotope identification code is needed. This includes increasing the number of spectra used in the empirical modeling of the area scores, incorporating other data features into the identification process, and a parameter optimization across all of the models. Besides generally improving performance, the identification code could be tuned to allow for a higher false alarm rate in the wavelet/NNLS code, which will allow the sensitivity of the wavelet/NNLS algorithm to be increased further.

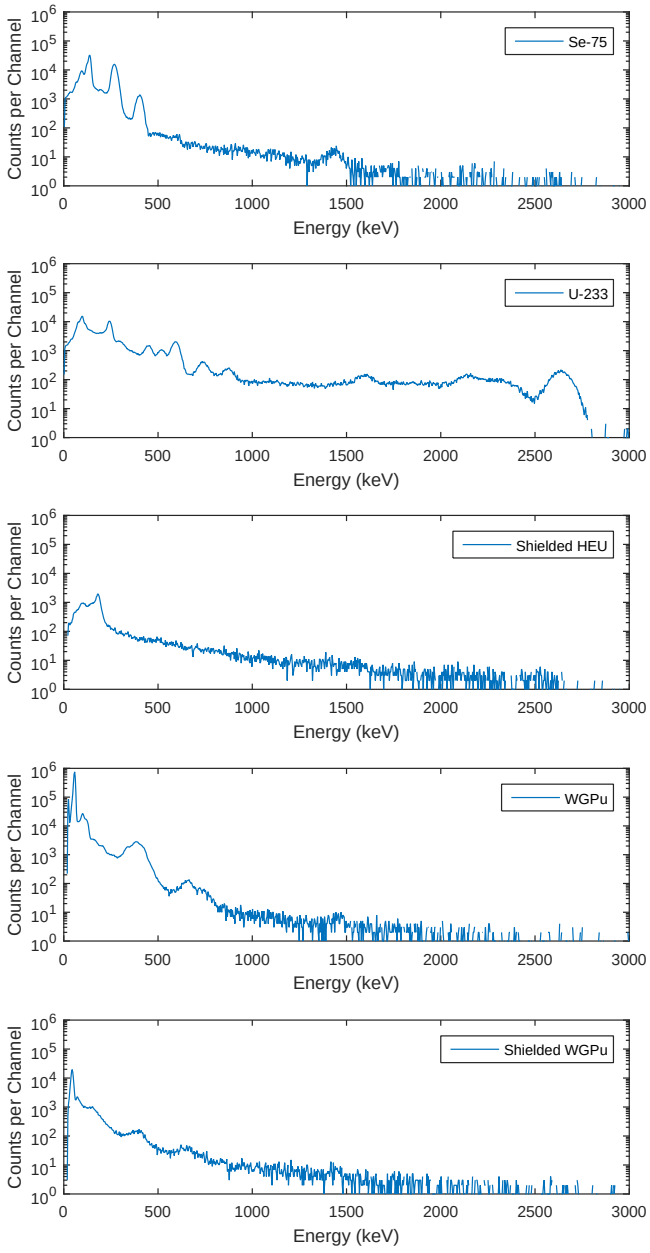

Fig. 5. Example spectra used for test identifications. With the exception of the ^{233}U spectrum (300 s), all spectra were measured for 60 s. The spectra were also not all collected in the same location, leading to different levels of background radiation from ^{40}K and the ^{232}Th decay chain.

TABLE III
EXAMPLE IDENTIFICATION RESULTS. ALL SPECTRA WERE MEASURED WITH AN ORTEC 905-3 2"x2" NAI DETECTOR AND ARE NOT BACKGROUND-SUBTRACTED.

#	Source	Amount	Shielding	Dist	Time	ID
1.	Se-75	0.54 mCi	-	50 cm	60 s	Se-75, Th-232, Ga-67
2.	U-233	1 g	-	100 cm	300 s	U-233, Th-232
3.	HEU: 93.2% U-235	13 kg	1.2 cm Fe	120 cm	60 s	U-235
4.	WGPu: 93.7% Pu-239 557 ppm Am-241	4.5 kg	-	120 cm	60 s	Pu-239, Am-241, I-125
5.	WGPu: 93.7% Pu-239 557 ppm Am-241	4.5 kg	1.2 cm Pb	120 cm	60 s	Pu-239, Am-241

Finally, the Bayesian classifier utilizes a simple prior probability distribution and uniform cost function. These could be changed to improve performance towards specific benchmarks or applications, but their simple forms have been kept for developing the rest of the identification code and to demonstrate that our performance is not being specifically driven by a specific choice of prior distribution or cost function.

VI. CONCLUSIONS

A wavelet/NNLS approach to peak detection and quantification has been developed. This method can extract peak energy, area, and area uncertainty measurements from a spectrum. Modifications of the isotope identification code have been developed to incorporate the new area uncertainty information into the likelihood modeling using an empirical approach. Future work on the feature extraction algorithm will focus on improving sensitivity to overlapping and small peaks, while development of the isotope identification algorithm will center on improving the new area uncertainty modeling and on parameter optimization.

REFERENCES

[1] S. Brill, "Is America Any Safer?" September 2016. [Online]. Available: http://www.theatlantic.com/magazine/archive/2016/09/are-we-any-safer/492761/

[2] United States Government Accountability Office, "Secure Border Initiative: Technology Deployment Delays Persist and the Impact of Border Fencing Has Not Been Assessed," September 2009. [Online]. Available: http://www.gao.gov/products/GAO-09-896

[3] ——, "Combating Nuclear Smuggling: Recent Testing Raises Issues About the Potential Effectiveness of Advanced Radiation Detection Portal Monitors," November 2009. [Online]. Available: http://www.gao.gov/products/GAO-10-252T

[4] J. Blackadar, J. Bounds, P. Hypes, D. Mercer, and C. Sullivan, "Evaluation of Handheld Isotope Identifiers," *LA-UR-03-2742*, April 2003.

[5] M. Swoboda, R. Arlt, V. Gostilo, A. Lupilov, M. Majorov, M. Moszynski, and A. Syntfeld, "Spectral Gamma Detectors for Hand-held Radioisotope Identification Devices (RIDs) for Nuclear Security Applications," *Nuclear Science Symposium Conference Record, 2004 IEEE*, vol. 7, pp. 4296–4302, 2004.

[6] J. Blackadar, "Automatic Isotope Identifiers and Their Features," *IEEE Sensors Journal*, vol. 5, no. 4, pp. 589–592, August 2005.

[7] T. Burr and M. Hamada, "Radio-Isotope Algorithms for NaI Gamma Spectra," *Algorithms*, vol. 2, pp. 339–360, March 2009.

[8] C. J. Sullivan, S. E. Garner, M. Lombardi, K. B. Butterfield, and M. A. Smith-Nelson, "Evaluation of Key Detector Parameters for Isotope Identification," *IEEE Nuclear Science Symposium Conference Record*, pp. 1181–1184, 2007.

[9] M. Alamaniotis, A. Heifetz, A. C. Raptis, and L. H. Tsoukalas, "Fuzzy-logic radioisotope identifier for gamma spectroscopy in source search," *IEEE Transactions on Nuclear Science*, vol. 60, no. 4, pp. 3014–3024, August 2013.

[10] M. T. Batdorf, W. K. Hensley, C. E. Seifert, L. J. Kirihara, L. E. Erikson, and D. V. Jordan, "Isotope Identification in the Gamma-Tracker Handheld Radioisotope Identifier," *IEEE Nuclear Science Symposium Record*, 2009.

[11] D. J. Mitchell and L. T. Harding, "GADRAS Isotope ID User's Manual for Analysis of Gamma-Ray Measurements and API for Linux and Android," *SAND2014-3933*, May 2014.

[12] R. Runkle, "Analysis of Spectroscopic Radiation Portal Monitor Data using Principal Components Analysis," *IEEE Transactions on Nuclear Science*, vol. 53, no. 3, pp. 1418–1423, 2006.

[13] P. Olmos, J. Diaz, J. Perez, and G. Garcia-Belmonte, "Application of Neural Network Techniques in Gamma Spectroscopy," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometors, and Associated Equipment*, vol. 312, pp. 167–173, 1992.

[14] S. Iwasaki, H. Fukuda, and M. Kitamura, "High-Speed Analysis Technique for Gamma-Ray and X-Ray Spectra Using an Associative Neural Network," *International Journal of PIXE*, no. 3, pp. 267–273, 1993.

[15] P. E. Keller, L. J. Kangas, G. L. Troyer, S. Hashem, and R. T. Kouzes, "Nuclear Spectral Analysis via Artificial Neural Networks for Waste Handling," *IEEE Transactions on Nuclear Science*, vol. 42, no. 4, pp. 709–715, August 1995.

[16] D. K. Fagan, S. M. Robinson, and R. C. Runkle, "Statistical Methods Applied to Gamma-Ray Spectroscopy Algorithms in Nuclear Security Missions," *Applied Radiation and Isotopes*, vol. 70, pp. 2428–2439, 2012.

[17] G. Knoll, *Radiation Detection and Measurement*, 4th ed. John Wiley and Sons, 2010.

[18] C. J. Sullivan and J. Lu, "Automated Photopeak Detection and Analysis in Low Resolution Gamma-Ray Spectra for Isotope Identification," *IEEE Nuclear Science Symposium Conference Record*, 2013.

[19] X. Kong, "Advanced Adaptive Library for Gamma-Ray Spectrometers," Master's thesis, University of Illinois at Urbana-Champaign, 2014.

[20] H. Xiong, "Automated Wavelet Analysis of Low Resolution Gamma-Ray Spectra and Peak Area Uncertainty," Master's thesis, University of Illinois at Urbana-Champaign, 2015.

[21] K. Vaughan, R. Turner, and P. Kendall, "Wavelet analysis of shielded radiological material: The applicability of wavelet analysis for national security," *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference*, 2014.

[22] J. Stinnett and C. J. Sullivan, "An Automated Isotope Identification Algorithm Using Bayesian Statistics," *IEEE Nuclear Science Symposium Conference Record*, 2013.

[23] J. Stinnett, "Bayesian Algorithms for Automated Isotope Identification," Master's thesis, University of Illinois at Urbana-Champaign, 2014.

[24] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological Society*, vol. 79, no. 1, pp. 61–78, 1998.

[25] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[26] The MathWorks Inc., *MATLAB and Wavelet Toolbox Release 2016a*, 2016.

[27] E. Browne and R. B. Firestone, *Table of Radioactive Isotopes*, V. S. Shirley, Ed. Wiley-Interscience Publications, 1986.

[28] C. J. Sullivan and J. Stinnett, "Validation of a Bayesian-Based Isotope Identification Algorithm," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometors, and Associated Equipment*, 2014.

[29] J. Stinnett and C. J. Sullivan, "Performance of an automated isotope identification algorithm for handheld NaI detectors," *IEEE Nuclear Science Symposium Conference Record*, 2015.

[30] *905 Series NaI(TI) Scintillation Detectors*, Ortec, 2016.