Using Speech Technology to Enhance Isotope ID and Classification

Lane M. D. Owsley, *Member, IEEE,* Jack J. McLaughlin, Luca G. Cazzanti, *Member, IEEE,* and Saleem R. Salaymeh

***Abstract*—Scientific advances are often made when researchers identify mathematical or physical commonalities between different fields and are able to apply mature techniques or algorithms developed in one field to another field which shares some of the same challenges. The authors of this paper have identified similarities between the unsolved problems faced in gamma-spectroscopy for automated radioisotope identification and the challenges of the much larger body of research in speech processing. In this paper we describe such commonalities and use them as a motivation for a preliminary investigation of the applicability of speech processing methods to gamma-ray spectra. This approach enables the development of proof-of-concept isotope classifiers, whose performance is presented for both simulated and field-collected gamma-ray spectra.**

I. Introduction

HANDHELD NaI(Tl) detectors are the most prevalent radioisotope identifying instrument on the frontlines of nuclear homeland security. However, the embedded algorithms in commercial detectors perform poorly at isotope identification in laboratory environments and even worse in the field. Correct isotope identification is less than 35% in commercial detectors under the best conditions [1],[2]. The ANSI standard for NaI(Tl) detectors [3] requires an 80% correct identification rate for 10k count gamma-ray spectra and an 80% success of identifying 50-50 mixtures of various isotopes with special nuclear materials (SNM).

Much of the algorithmic design for radioisotope identification is not in the open literature, due both to government security limitations and to proprietary concerns of commercial developers. It can be inferred, however, that many of the latter's algorithms are based on peak-picking methods [4],[5] which are automated forms of the heuristic approaches taken by human spectroscopists. In an evaluation of research-grade algorithms Nelson and Sokkappa [6] find that common signal processing methods (template matching, maximum likelihood, principal component analysis) perform vastly better than commercial off-the-shelf hardware embedded algorithms. In controlled environments, these research grade algorithms would likely meet the 80% identification rates specified by the ANSI 42.34 and 42.12 standards [7] , however, the authors express concern about the "lack of adaptability to real world conditions," where the algorithms perform much worse (at least 25% lower correct isotope identifications) when subjected to data taken in the field.

Achieving the ANSI standard would be a great improvement for low-resolution gamma-ray spectroscopic identifier algorithms. However, these performance specifications are based upon controlled, laboratory data and unfortunately, homeland security field applications present much more complicated spectra with large and varying backgrounds and multi-component isotope mixtures among other real-world issues. The 2006 Government Accountability Office (GAO) report on Combating Nuclear Smuggling makes specific mention of the fact that Customs and Border Protection officers reported that the radiation isotope identifying devices (RIIDs) had "difficulty recognizing radioactive material and correctly identifying the material they did recognize" [8]. Standard low-resolution spectroscopy signal processing has been developed by academia, industry and DOE's National Laboratories; after the decades of work that have gone into programs such as Gamma Designer, Peak Finder, GADRAS and GAMANAL the approaches underlying these systems are quite mature. We feel that at this point, if significant additional gains are to be made with sodium-iodide detectors, it is going to come from non-traditional approaches.

A. Radioisotope Identification Research

The open literature contains many approaches to the automated isotope identification problem which are based on the use of more of the spectrum than simply the peaks. R. Estep *et al.* have been developing the multiple isotope material basis set, MIMBS, for a decade [9],[10]. The MIMBS method simultaneously solves for the isotope producing the radiation and the intervening shielding, addressing the attenuation effect in gamma-ray spectrum analysis. In a similar manner, the MIMBS method has also been applied to the problem of gain drift [11]. R. Runkle and D. Pfund have recently demonstrated the benefits of their algorithm, the nuisance-rejection spectral comparison ratio anomaly detection, N-SCRAD [12],[13]. The N-SCRAD algorithm was developed to work on a time series of data and employs an intelligent data binning approach coupled with an optimized region-of-interest (ROI) selection. The ROI selection is designed to maximize SNM detection while minimizing nuisance alarms. D. Stromswold, J. Ely and R. Kouzes and others have looked at signal processing in radiation portal monitors [14],[15],[16]. They show significant

Manuscript received November 13, 2009. This work was supported in part by the U.S Department of Defense, Defense Threat Reduction Agency. This work was also sponsored in part by the National Nuclear Security Agency Office of Nonproliferation Research and Development NA-22.

L. M. D. Owsley, J. J. McLaughlin, and L. G. Cazzanti are with the Applied Physics Laboratory, University of Washington, Seattle, WA 98105 USA (telephone 206-685-3592, e-mail lane@apl.washington.edu).

S. R. Salaymeh is with Savannah River National Laboratory, Aiken, SC 29808 USA (telephone 803-725-1628, email saleem.salaymeh@srnl.doe.gov).

gains in SNM versus naturally-occurring source discrimination via the use of energy windowing, *i.e.* subdividing the spectral content into pieces and making comparisons between the ratios of these energy "windows". Gosnell [17] and separately K. Nelson and P. Sokkappa [6] have investigated and shown the promise of principal component analysis (PCA). Principal component analysis is an analysis method that constructs a set of orthogonal vectors that linearly transform a dataset into a lower-dimensional basis of maximum variance. Gosnell *et al.* set up the framework for PCA radioisotope identification and Nelson and Sokkappa implemented a PCA "toy model" that worked well in identifying mixtures from a library of synthetic spectra.

B. Speech Processing Overview

The physics of speech and hearing have been areas of active research for over a century, and automated processing of the speech signal and its resulting spectrum predate the development of the integrated circuit in the 1960s. Because the application (and market) potential for computer systems that can model and automatically identify speech are so vast, the research activity level has also been vast and has resulted in highly robust fielded systems for many applications [18]. Speech recognition systems are commonly used in automated phone services, hands-free phone and vehicle control systems [19], and a wide variety of other specialty applications including handicap accessibility software [20] and auto-dictation systems [21]. In addition, speaker identification systems–in which the end goal is not to determine *what* is being said but to determine (or verify) *who* is saying it–have found use in border security [22], biometric control of access to secure facilities, and to automated analysis of recorded conversations and media broadcasts. In the last twenty years, such systems have moved out of the laboratory and into general use. One of the keys to this transition has been the research focus on maximizing the robustness of the systems to the wide variety of acoustic environments, recording conditions, and microphone responses which are inevitably faced in the field [23]. At the same time, of course, the algorithm must maintain a high level of specificity to enable distinctions to be made among a large set of speakers (for speaker identification systems) or a vocabulary which may run to tens of thousands of words. It is the goal of the remainder of this paper to introduce two methods which have found success in speech processing, justify their use in the radioisotope spectrum scenario based on commonalities in the spectra and the variations observed in those spectra in the field, and test their performance in a proof-of-concept classification scenario.

II. Cepstral Processing for Isotope Identification

A cepstrum is a nonlinear transform of a spectrum resulting in a representation which has several qualities of great value in a variety of applications where a signal passes through an environment which filters it and where the signal (and often the environmental effects) are of individual interest [18]. Cepstra have been widely used in a variety of disciplines, including medical imaging [24], microwave detection of objects behind walls[25], and speech processing [18] as features for signal enhancement and classification.

Cepstral representations of speech signals are the dominant feature set for almost all computer analysis of speech in modern systems. Among the many reasons for this are three that are particularly relevant to the radioisotope identification problem:

1) Cepstral features are very useful at separating the effects of the channels (the vocal tract, the room acoustics, the recording device, and the transmission lines) from the representation of the underlying source (which may be visualized as a pulse train or a more wideband noise source, depending on the type of speech). Our interest is in the possibility that a similar approach can be used to separate the effects of shielding, backscatter sources and the detector itself from the characteristics of the source spectrum.
2) Cepstral processing for speech contains a number of stages which are specifically designed to mimic the original (and still the best) "detector" of speech: the human ear. One characteristic of this detector is that subjective perceptual experiments show that humans perceive magnitudes of pitch differences in a manner which is a function of acoustic frequency. As a result, cepstral processing is performed using a filter bank with a *mel scale* [26] to maximize the efficiency in which the information available to the human ear is processed by the computer. NaI (Tl) gamma-ray detectors have a similarly varying resolution as a function of their energy axis, and thus a similar processing approach may be of value.
3) Another feature of the human ear is its sensitivity to lower signal magnitudes at high acoustic frequency. This results in a high dynamic range and a loss of processing power if all frequencies are treated equally. As a result, a pre-emphasis stage that flattens out the spectrum is employed [27]. This is reminiscent of the challenge faced in gamma spectroscopy in which peaks and other features at higher energy are of lower absolute count, and a similar preemphasis (tailored to the actual gamma spectral shape) could be included in the processing.

A. Mathematical Realization

In speech processing, cepstra are calculated from a time series by first computing the Fourier transform, taking the logarithm of the result, and then computing the inverse Fourier transform to arrive at the cepstrum. If the phase of the Fourier representation is dropped prior to taking the logarithm, then the quantity calculated is properly known as the "real cepstrum." To compute a cepstrum for a radiation detector output, we substitute the counts in the energy channels as the magnitude of the positive half of an spectrum with no phase. Taking the log and the inverse Fourier transform then yields the real cepstrum as summarized in Figure 1.

In speech, the next step is truncation by zeroing out of higher order cepstra, and results in a smoothing of the peaks in the energy spectrum (if the cepstrum were to be inverse

transformed back into the spectral domain). This reduction in dimensionality is desirable from the classification standpoint so long as peaks remain sufficiently resolved to be distinguished.

From the preceding discussion, we can now see how cepstral features can facilitate the removal of channel effects and other distortions which are convolutional in the time domain. Equation 1, where $z[n]$ is the undistorted signal and $c[n]$ is the distorting function, shows this relationship. Equations 2 and 3 show this same relationship in the spectral and cepstral domains, respectively, where the multiplicative distortion becomes additive due to the log operation used in computing the cepstra. Removal of the distortion is simplest in the cepstral domain since simple subtraction can be applied providing an estimate of the distortion is available.

$$m[n] = z[n] * c[n] \tag{1}$$

$$M(\omega) = Z(\omega)C(\omega) \tag{2}$$

$$\hat{m}[n] = \hat{z}[n] + \hat{c}[n] \tag{3}$$

There is some evidence that the effect of shielding can be addressed by the process of cepstral subtraction. The MIMBS approach to modeling shielding assumes that shielding has a multiplicative effect on the underlying spectrum [9]. This would become additive in the cepstral domain allowing us to subtract it out, and leaving us with the unshielded isototope spectrum.

A speech processing method can also be used to compensate for varying sensitivity (as a function of energy) of radiation detectors. Figure 2 shows a series of filters typically applied to a speech spectrum prior to calculation of cepstra. Similarly, some filtering of the output of a radiation detector could serve to counteract the sensitivity problem.

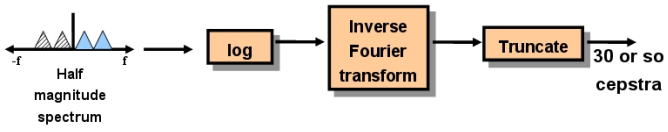

Fig. 1. Steps taken to compute cepstral coefficients from radiation detector data. After creating a full spectrum from the half spectrum, the left-hand side of this diagram is $M(\omega)$ as it appears in Equation 2. The cepstra that are output are the $\hat{m}[n]$ of Equation 3. These are also the inputs to the modeling process shown in Figure 4.

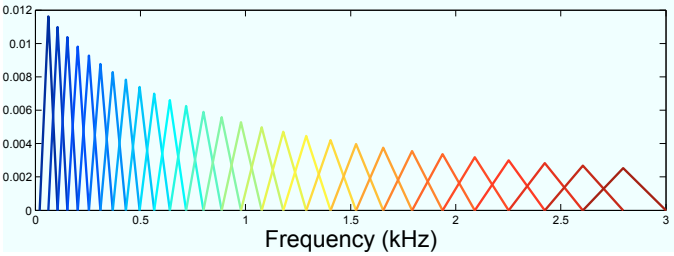

Fig. 2. Mel filters which are typically applied to a speech spectrum prior to calculation of cepstral coefficients. This filtering simulates that performed by the human ear.

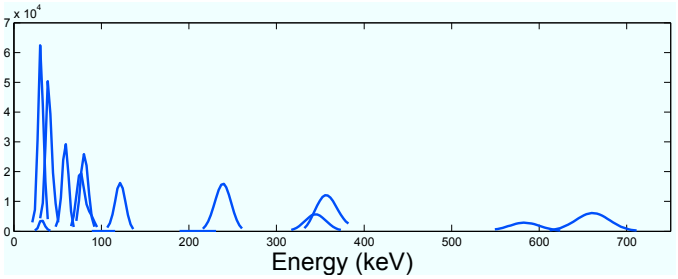

Fig. 3. Spectral peaks taken from the gamma spectra of isotopes considered in this study. Comparison of these peaks to the filters in Figure 2 suggests that the axis-dependent filter width approach used in cepstral processing could be useful in creating a compact representation of the information present in gamma spectra.

It is clear from this discussion that there are many commonalities between speech spectra (and their variations) and gamma-ray spectra. It is also clear that many of the ways of processing these spectra (mel filtering, pre-emphasis, cepstral-coefficient selection, and channel normalization) are application-specific. The scope of the current research did not allow for adaptation of these features to the gamma-ray spectrum, and as a result we must evaluate the resulting feature set in the context of that shortcoming.

B. Model Creation and Testing

Automatic classifiers comes in two basic types: generative and discriminative. The discriminative classifier takes in all the features vectors for all classes, and produces a single "model" which delineates the boundaries between all classes. Generative classifiers take all the feature vectors for each class in turn, and produces one model per class each of which characterizes the arrangement of that classes' feature vectors in the feature space. The primary advantage of discriminative classifiers, in theory, is that they can better separate classes because the classifier works with all the training data at once. Generative classifiers, on the other hand, are more easily extensible because if a new class is added, a model can be built with just the training data from that new class whereas a discriminative classifier must be retrained with the data from all classes.

A conceptual diagram of the training of a generative classifier is shown in Figure 4. Selection of training data is critical in assuring that models are robust across all conditions that may be encountered during testing. For example, if all training data come from detectors with energy axes which are perfectly calibrated then test examples from detectors which are not similarly calibrated are apt to be misclassified – absent any post-processing to correct the calibration.

To test the cepstral features, we employed a generative classifier, the Gaussian Mixture Model. The GMM is simply a sum of weighted Gaussians which can be described by

$$p(\vec{\hat{m}}|\lambda) = \sum_{i=1}^{M} p_i b_i(\vec{\hat{m}}) \tag{4}$$

where $\vec{m}$ is an N-dimensional random vector (the cepstral feature vector), b_i are the M component densities, and p_i are

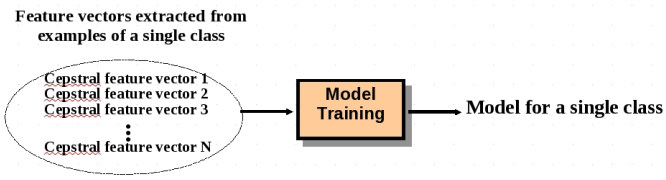

Fig. 4. Training of a generative classifier. Feature vectors from many examples of a single isotope are used to create a probabilistic model for that isotope.

the component weights. Each component is an N-dimensional Gaussian of the form

$$b_i(\vec{m}) = \frac{1}{(2\pi)^{N/2}|\Sigma_i|^{1/2}} exp\{\frac{1}{2}(\vec{m}-\vec{\mu}_i)^T\Sigma_i^{-1}(\vec{m}-\vec{\mu}_i)\} \quad (5)$$

where $\vec{\mu}_i$ and Σ_i are the mean and covariance of the component.

Weights are scaled such that they sum to one making the GMM a proper probability density function. Calculation of the probability of a test vector given a GMM is straightforward. The isotope type of the test vector is determined by calculating the probability of the test vector given models for many putative isotopes, and selecting the isotope whose model scores the highest probability.

The popularity of GMMs is largely due to the small number of parameters that must be estimated from training data to determine the components; the ability of GMMs to accurately describe training data in a high dimensional feature space through the inclusion of additional components; and the existence of the Expectation Maximization (EM) algorithm [28], an iterative process enabling rapid calculation of GMM parameters from training data.

III. Spectral Nearest-Neighbors for Isotope Identification

Instead of estimating probabilistic models of features $\vec{m}$ derived from the spectrum $M(\omega)$, a different method for identifying isotopes is based on a direct comparison of the spectra. A properly-chosen function computes the distance between pairs of spectra, and spectra close to each other are assumed to have been generated by the same physical process. In speech processing, the Itakura-Saito spectral distance [29], [30], dating back to the 1970s has been successfully used to compare speech spectra. More recently, it was adapted to the problem of automatically identifying musical recordings [31], [32].

Another spectral distance successfully employed in speech processing is the Kullback-Leibler (KL) divergence. The KL divergence is the standard information-theoretic method for comparing probability distributions [33]. It quantifies the "surprise" of observing the distribution M given the hypothesized distribution R: more similar distributions will be less surprising, and thus more likely to have been generated by the same physical process. Because of its solid underpinnings in information theory, the KL divergence pervasively appears in applications where probability distributions are compared.

In particular, for speech processing, the KL divergence has shown robustness to the types of variations common within a speaker uttering different words, but still sensitive to the key spectral differences between speakers [34], [35], [36]. The analogy to the radioisotope identification problem suggests that gamma-ray spectra characterized by decay peaks and broad Compton shelves may also be successfully compared, even in the presence of shielding and varying backgrounds.

Because of its demonstrated robustness in speech processing and its broad applicability, in this work we adopt the (symmetric) KL divergence as the spectral distance for radioisotope identification, The KL divergence between two spectra M and R is defined as

$$d_{KL}(M,R) = \sum_i M(i)\log\frac{M(i)}{R(i)}, \quad (6)$$

where $M(i)$ is the radiation energy in the i-th bin of the spectrum, and where we take all spectra as normalized so that $\sum_i M(i) = 1$. The symmetric KL divergence is

$$d(M,R) = \frac{d_{KL}(M,R) + d_{KL}(R,M)}{2}. \quad (7)$$

We adopt the nearest-neighbor classifier [37] to identify unknown spectra based on their KL divergences. Near-neighbor classifiers classify an unknown spectrum M as the class $\hat{\lambda}$ of its nearest neighbor, chosen among a set $\mathcal{R}$ of reference spectra whose class is known, or more generally as the most frequently occurring class in the size-k neighborhood $\mathcal{N}_k(M) \subset \mathcal{R}$. In its simplest form, which is the form adopted in our work, the k-nearest neighbor (k-NN) classifier is written

$$\hat{\lambda} = \arg\max_{\lambda} \frac{1}{k} \sum_{R\in\mathcal{N}_k(M)} I_{(\Lambda_R=\lambda)}, \quad (8)$$

where $I_{()}$ is the indicator function, which is one when its argument is true and zero otherwise, and Λ_R is the class label associated with the spectrum R.

Near-neighbor classifiers are special discriminative classifiers that are local and model-free. They are local because the classification is based on local neighborhoods of the unknown spectrum, and they are model-free because no model training is required. Instead, unknown spectra are compared to a reference database of known spectra by means of an appropriate distance function – in our case the KL divergence. Locality makes near-neighbor classifiers generally more robust to model bias than standard generative classifiers – at the cost of an increase in variance – and model-free classification allows the flexible extension of the reference database without costly model retraining. Low bias and high flexibility, coupled with the intuitive notion of nearest neighbors, have made the k-NN classifier a popular choice for many practical tasks [37].

IV. Data

To test the algorithms described in this paper, we used a combination of field-collected and synthesized data from six available isotopes. The collected data were obtained using a handheld NaI detector, the Identifinder NG. SRNL employed approximately 60-second live times for each collection, with samples placed at one foot and one meter distances from the

TABLE I
SOURCE RADIOACTIVITY FOR SRNL-COLLECTED FIELD DATA

Target	Radioactivity
^{60}Co	3.65 μCi
^{228}Th	0.51 μCi
^{133}Ba	5.99 μCi
^{241}Am	50.30 μCi
^{152}Eu	7.26 μCi

detector. The spectra were collected in a variety of background environments, including grass, concrete, and gravel, located on the SRNL grounds. Reference background spectra were also collected. The source radioactivities of the available samples are compiled in Table I.

SRNL created simulated data using GADRAS [38], with a detection parameter file chosen to match the detector. These spectra were scaled so that their expected values corresponded with sources of different radioactivities at one meter distance. The source spectra were combined additively with field-collected background spectra.

The scope of this work does not include improvements to the state of the art in detector calibration; instead we assume that an available calibration method has been employed. However, no calibration method is perfect and thus some robustness to variation is required. To include such variation in these preliminary experiments we added small perturbations to the offset, slope, and curvature of the spectra; this is referred to as a "warping" of the spectral axis.

V. CLASSIFICATION EXPERIMENTS

A. Cepstral Processing

Using the calibration warping algorithm and Poisson resampling, approximately 1000 simulated examples of each isotope type (^{241}Am , ^{133}Ba, ^{60}Co, ^{152}Eu and ^{228}Th) were produced. Field-collected background radiation spectra were added to these simulated data to yield an effective source-detector distance of 1 meter and isotope strength of 20 μCi. GMMs were trained for each class (each isotope) using roughly 70% of these with the remaining 30% being using for testing. These choices were driven primarily by the dimensionality of the feature set used (30 cepstral coefficients per isotope spectrum), and the number of parameters in the model (means, covariances and weights for a 15-component GMM). Test examples were classified with an accuracy of 98.4% which is not unexpected given the homogeneity of the training and test sets.

It would be ideal if a classifier could classify field data well after being trained only on simulated data as this would obviate the need to do data collection. Accordingly, a classifier was trained using simulated data (at a 1-meter detector distance with a 20 μCi source strength) for the five classes with no background data added. To classify spectra from field data, we then applied the trained system to collected spectra with a time-normalized reference background spectrum removed. For distance from the source yielded an accuracy of 75.4% correct, with the full confusion matrix shown in Table II. As discussed in Section II, this baseline system is has its pre-emphasis, filtering, and coefficient selection tuned to the speech processing application and a second generation of this system could be expected to perform significantly better when it was accurately matched to the emissions physics and to the NaI (Tl) detector's specific response.

	^{241}Am	^{133}Ba	^{60}Co	^{152}Eu	^{228}Th
^{241}Am	0	0	0	0	100
^{133}Ba	0	65.28	0	0.28	34.44
^{60}Co	0	0	100.00	0	0
^{152}Eu	0	0	1.67	98.33	0
^{228}Th	0	0	0	0	100

TABLE II
CONFUSION MATRIX FOR CEPSTRAL FEATURE EXPERIMENT; EMBEDDED SYNTHETIC DATA CLASSIFIED USING SYNTHETIC TEMPLATES.

	^{241}Am	^{133}Ba	^{60}Co	^{152}Eu	^{228}Th
^{241}Am	67.31	27.26	0	5.40	0.02
^{133}Ba	0.30	98.26	0	0.77	0.66
^{60}Co	0	0	75.89	24.06	0.04
^{152}Eu	0	0	15.06	84.94	0
^{228}Th	0	0	0	0	100

TABLE III
CONFUSION MATRIX FOR KL EXPERIMENT; EMBEDDED SYNTHETIC SPECTRA CLASSIFIED USING 5 TEMPLATE SPECTRA.

B. KL Divergence

In this experiment we tested the robustness of a nearest-neighbor classifier based on the KL divergence to typical spectral warpings. Five simulated spectra for the Identifinder detector are considered the "prototype" spectra against which warped spectra are compared. For each isotope, the prototype spectrum was warped to represent realistic spectral distortions that arise when the spectra are measured in the field. Each isotope's spectrum was warped in 40,000 different ways. The warped spectra were then compared to the prototype spectra using the KL divergence entropy and classified as the isotope of its nearest prototype. Thus, in the confusion matrix in Table III, each row corresponds to 40,000 comparisons. The significance of this result is that the KL divergence offers great promise for identifying distorted spectra even when using a very small database of pristine simulated reference spectra. This can enable practical deployment of the algorithms in the field because of the low required complexity of the algorithm.

The overall correct isotope classification percentage for the above experiment is 85.25%. The performance can be improved by expanding the database of reference spectra to include a few of the warped spectra. To this end, we randomly picked 10 simulated warped spectra for each isotope and added them to the pristine reference database, so that the expanded database consists of a 55 spectra, 11 per isotope. For each isotope, the remaining 39,990 warped spectra were used to test the approach. The results are in Table IV. The overall correct classification for this case is 99.45%. This extremely high classification rate shows that with a very small increase in the database size, the KL divergence can correctly account

	^{241}Am	^{133}Ba	^{60}Co	^{152}Eu	^{228}Th
^{241}Am	97.59	2.41	0	0	0
^{133}Ba	0	100	0	0	0
^{60}Co	0	0	99.65	0.35	0
^{152}Eu	0	0	0	100	0
^{228}Th	0	0	0	0	100

TABLE IV
CONFUSION MATRIX FOR KL EXPERIMENT; EMBEDDED SYNTHETIC SPECTRA CLASSIFIED USING 55 TEMPLATE SPECTRA (INCLUDING ADDED CALIBRATION NOISE).

	^{241}Am	^{133}Ba	^{60}Co	^{152}Eu	^{228}Th
^{241}Am	100	0	0	0	0
^{133}Ba	0	86.67	0	13.33	0
^{60}Co	0	0	89.29	10.71	0
^{152}Eu	0	0	0	100	0
^{228}Th	0	0	0	30	70

TABLE V
CONFUSION MATRIX FOR KL MEASURE; 148 FIELD-MEASURED SPECTRA CLASSIFIED USING 55 TEMPLATE SPECTRA (INCLUDING ADDED CALIBRATION NOISE).

for all warping effects in single-isotope spectra. Also note that including the additional spectra in the reference spectra required no classifier re-training, which would instead be a requisite for generative classifiers. The practical implication is that portable automatic detectors based on nearest-neighbor classification can be enhanced simply and inexpensively by expanding their available memory.

We then tested the KL divergence approach on field data. The reference database of spectra consists of the same 55 simulated pristine and warped spectra as in the previous experiment. The test spectra, however, are 148 single-isotope spectra measured with the Identifinder detector in a variety of background environments (asphalt, grass, etc). From each test isotope-and-background spectrum, the corresponding background-only spectrum is subtracted channel-by-channel, resulting in an estimated isotope-only spectrum which is compared to the reference database. The results are in Table V.

The overall correct classification rate is 89.19%. This result is very significant because it shows that the simple KL divergence approach can achieve very high classification rates even when field-measured data are compared to a small database of prototype simulated data. Thus, this approach is a promising path for possible implementation in portable detectors which have limited computational power.

Although the generalization to shielded sources was outside the scope of this work, we performed one preliminary experiment to test the robustness to shielding effects of the KL divergence approach. We used the same 55 simulated prototypes as in the previous experiments, but used 114 simulated shielded spectra for ^{60}Co. This isotope was correctly identified 85% of the times when using the nearest-neighbor approach used in all other experiments. When taking a majority vote of the 31 nearest prototype spectra, 100% accuracy was achieved. This result is important because it shows that single-isotope shielded spectra can be successfully identified using a small database of simulated warped and unwarped spectra.

VI. NEXT STEPS

The results presented in this paper demonstrate the ability of algorithms developed to maturity in fields outside of the field of gamma ray spectroscopy to provide new and important directions of research and performance improvement. However, it is also clear that an understanding of and incorporation of the specific physics and statistics of radioisotope spectra is crucial if such novel algorithms are to prove robust outside the laboratory. The authors are currently working closely with nuclear scientists to achieve this objective by structuring these algorithms for the target application. We are also operating on more relevant datasets including a much wider variety of isotopes and isotopes in shielded and masked environments as well as scenarios with multiple target isotopes. Algorithmically, we are developing a probabilistic approach to frame the classification problem and explicitly modelling and parameterizing the effects of shielding to reduce the search space for our algorithms.

ACKNOWLEDGMENT

The authors thank Alex Athey at the Applied Research Laboratory, University of Texas, for his contribution to the literature survey.

REFERENCES

[1] J.M. Blackadar, C.J. Sullivan, B. G. Rees, S.E. Garner, and D. J. Mercer, "Continuing evaluation of isotope identifiers," in *Proceedings of INMM 45th Annual Meeting*, 2004.

[2] L. Pibida, M. Unterweger, and L. R. Karam, "Evaluation of handheld radionuclide identifiers," *Journal of Research of the National Institute of Standards and Technology*, vol. 4, no. 4, pp. 451, 2004.

[3] American National Standards Institute, *ANSI Standard N42.34: Performance Criteria for Hand-held Instruments for the Detection and Identification of Radionuclides*, 2003.

[4] Canberra, Inc, *Genie 2000 Spectroscopy Software Reference Manual version 3.0.*

[5] Gordon Gilmore and John Hemingway, *Practical Gamma-Ray Spectrometry*, Wiley, 1995.

[6] K. Nelson and P. Sokkappa, "Foundations for improvements to passive detection systems - final report," Tech. Rep. Publication UCRL- TR-207129, Lawrence Livermore National Laboratory, 2004.

[7] American National Standards Institute, *ANSI Standard N42.12: Calibration and Usage of Thallium-Activated Sodium Iodide Detector Systems for Assay of Radionuclides*, 1994.

[8] "Combating nuclear smuggling: Dhs has made progress deploying radiation detection equipment at u.s. ports-of-entry, but concerns remain," Tech. Rep. GAO-06-389, Government Accounting Office, March 2006.

[9] R.J. Estep, C.W. McCluskey, and B.A. Sapp, "The multiple isotope material basis set (MIMBS) method for isotope identification with low- and medium-resolution gamma-ray detectors," *Journal of Radioanalytical and Nuclear Chemistry*, vol. 276, no. 3, pp. 737–741, 2008.

[10] R.J. Estep, M. Rawool-Sullivan, and D. Miko, "A method for correcting nai spectra for attenuation losses in hand-held instrument applications," *IEEE Transactions on Nuclear Science*, vol. 45, no. 3, pp. 1022n1028, 1998.

[11] R.J. Estep and B.A. Sapp, "Energy calibration algorithms for the multiple isotope material basis set (mimbs) isotope identification method," in *Proceedings of the IEEE Conference on Technologies for Homeland Security*, 2008.

[12] D.M. Pfund, R.C. Runkle, K.K. Anderson, and K.D. Jarman, "Examination of count-starved gamma spectra using the method of spectral comparison ratios," *IEEE Transactions on Nuclear Science*, vol. 54, no. 4, pp. 1232–1238, Aug 2007.

[13] R.C. Runkle, D.M. Pfund, K.D. Jarman, K. Hanlen, and K.K. Anderson, "Anomaly detection algorithms for nuclear threats: A final report for the advanced decision making for low-count spectra project," Tech. Rep. PNNL-16978, Pacific Northwest National Laboratories, 2007.

[14] D. C. Stromswold, J. W. Darkoch, J. H. Ely, R. R. Hansen, R. T Kouzes, B. D Milbrath, R. C. Runkle, W. A. Sliger, J. E. Smart, D. L. Stephens, L. C. Todd, and M. L. Woodring, "Field tests of a nai(tl)-based vehicle portal monitor at border crossings," in *Proceedings of 2004 IEEE Nuclear Science Symposium Conference*, 2004, vol. 1, pp. 196–200.

[15] J. H. Ely, R.T. Kouzes, B. D. Geelhood, J.E. Schweppe, and R. A. Warner, "Discrimination of naturally occurring radioactive material in plastic scintillator material," *IEEE Transactions on Nuclear Science*, vol. 51, no. 4, pp. 1672–1676, Aug 2004.

[16] J. H. Ely, R.T. Kouzes, J.E. Schweppe, E. Siciliano, D. Strachan, and D. Weier, "The use of energy windowing to discriminate snm from norm in radiation portal monitors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 560, no. 2, pp. 373–387, May 2006.

[17] T. B. Gosnell, J. M. Hall, C. L. Jam, D. A. Knapp, Z. M. Koenig, S. J. Luke, B. A. Pohl, A. Schach von Wittenau, and J. K. Wolford, "Gamma-ray identification of nuclear weapon materials," Tech. Rep. Publication UCRL-ID-127436, Lawrence Livermore National Laboratory, February 1997.

[18] T.F. Quatieri, *Discrete-Time Speech Signal Processing: Principles and Practice*, Prentice Hall, 2001.

[19] Adam Boretz, "Sensory launches truly hands-free voice control for bluetooth devices," *Speech Technology (online)*, June 2009.

[20] Paul C. Judge, "Q &a with kurzweil's ray kurzweil," *Business Week*, February 1998.

[21] Paul Mah, "Smbs are using speech recognition for productivity gains," *ITBusinessEdge*, September 2009.

[22] "Department of homeland security expands collection of biometrics for visitors," *DHS Press Release*, December 2008.

[23] Ji Ming, T. J. Hazen, J. R. Glass, and D. A. Reynolds, "Robust speaker recognition in noisy conditions," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 5, pp. 1711–1723, 2007.

[24] S. C. Sekhar and M. Unser, "Performance analysis of the cepstral technique for frequency-domain optical-coherence tomography," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 557–560, April 2008.

[25] H. Yacoub andT. K. Sarkar, "A homomorphic approach for through-wall sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1318–1327, May 2009.

[26] Stanley Smith Stevens, John Volkman, and Edwin Newman, "A scale for the measurement of the psychological magnitude of pitch," *Journal of the Acoustical Society of America*, vol. 8, no. 3, pp. 185–190, 1937.

[27] P. Zolfaghari and T. Robinson, "Formant analysis using mixtures of gaussians," in *Proceedings of theFourth International ICSLP Conference on Spoken Language*, October 1996, vol. 2, pp. 1229–1232.

[28] Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*, Wiley-Interscience, 2001.

[29] F. Itakura, "Minimum prediction residual principle applied to speech recognition," *IEEE Trans. Acoustics, Speech, Sig. Proc.*, vol. ASSP-23, no. 1, pp. 67–72, February 1974.

[30] A. H. Gray and J. D. Markel, "Distance measures for speech processing," *IEEE Trans. Acoustics, Speech, Sig. Proc.*, vol. ASSP-24, no. 5, pp. 380–390, October 1976.

[31] V. Venkatachalam, L. Cazzanti, N. Dhillon, and M. Wells, "Automatic identification of sound recordings," *Signal Processing Magazine, IEEE*, vol. 21, no. 2, pp. 92–99, Mar 2004.

[32] L. Cazzanti, V. Sielert, and M. R. Gupta, "Reverse engineering the sound of the jazz saxophone," Tech. Rep. UWEETR-2006-0011, University of Washington, Seattle, WA, June 2006.

[33] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, 1991.

[34] J. Ramirez, J. C. Segura, C. Benitez, A. de la Torre, and A. J. Rubio, "A new Kullback-Leibler VAD for speech recognition in noise," *IEEE Signal Proc. Letters*, vol. 11, no. 2, pp. 266–269, February 2004.

[35] Pedro J. Moreno, Purdy P. Ho, and Nuno Vasconcelos, "A kullback-leibler divergence based kernel for svm classification in multimedia applications," in *In Advances in Neural Information Processing Systems 16*, 2003.

[36] W. M. Campbell, W. M. Campbell, D. E. Sturim, D. E. Sturim, D. A. Reynolds, and D. A. Reynolds, "Support vector machines using gmm supervectors for speaker verification support vector machines using gmm supervectors for speaker verification," *Signal Processing Letters, IEEE*, vol. 13, no. 5, pp. 308–311, 2006.

[37] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer-Verlag, New York, 2001.

[38] J. Mattingly, D. J. Mitchell, L.T. Harding, E.S. Varley, and N. R. Hilton, "Fy2008 report on gadras radiation transport methods," Tech. Rep. SAND2008-7084, Sandia National Laboratories, 2008.