Nuclear Instruments and Methods in Physics Research A 784 (2015) 298–305

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Validation of a Bayesian-based isotope identification algorithm

C.J. Sullivan, J. Stinnett *

University of Illinois at Urbana-Champaign, Department of Nuclear, Plasma, and Radiological Engineering, Urbana, Illinois, USA

ARTICLE INFO

Available online 22 December 2014

Keywords:
Algorithms
Isotope identification
Wavelets
Low resolution gamma-ray spectra

ABSTRACT

Handheld radio-isotope identifiers (RIIDs) are widely used in Homeland Security and other nuclear safety applications. However, most commercially available devices have serious problems in their ability to correctly identify isotopes. It has been reported that this flaw is largely due to the overly simplistic identification algorithms on-board the RIIDs. This paper reports on the experimental validation of a new isotope identification algorithm using a Bayesian statistics approach to identify the source while allowing for calibration drift and unknown shielding. We present here results on further testing of this algorithm and a study on the observed variation in the gamma peak energies and areas from a wavelet-based peak identification algorithm.

1. Introduction

The Department of Homeland Security (DHS), the International Atomic Energy Agency (IAEA), and others rely on handheld radio-isotope identifiers (RIIDs) to detect nuclear threats, track and identify radioactive materials, and more. Most of these RIIDs use sodium iodide (NaI) detectors because they are relatively cheap, decently stable over a wide range of operating conditions, and reasonably efficient. However, published evaluations of these commercial detectors report poor isotope identification results [1–3]. Because a trained spectroscopist can use the spectra taken by these detectors to make accurate identifications, several authors have suggested that the performance issues are largely due to poor identification algorithms [4,5].

To combat many of these issues, we have previously proposed a new algorithm that uses peak energies and areas to compute the posterior probabilities for each isotope in our library given the data set [6]. An adaptive library [7] was used that includes all of the isotopes in the ANSI N42.34 standard for handheld detectors [8], plus ^{152}Eu and ^{177m}Lu for interesting test cases. A wavelet-based algorithm identifies peak energies and areas from a gamma-ray spectrum, and a Bayesian classifier algorithm identifies the radioisotope. This approach is capable of accurate identifications with unknown shielding scenarios and significant calibration drift and yet is computationally cheap enough to be feasible for implementation on handheld detectors. Running within MATLAB on a laptop with an i7-4700HQ processor, early versions of the wavelet peak identification algorithm run in under 25 s, and the Bayesian isotope identification algorithm runs in 0.004 s. In this paper, we report on a series of experiments conducted to validate and optimize the models presented in [6].

2. Peak detection and localization with wavelets

Many algorithms for isotope identification must first employ a peak finding method to quantify the centroid energy and possibly the area of all photopeaks in a spectrum [9]. Various methods for peak detection have been explored, all looking to improve peak centroid and area determinations [10–13].

A wavelet transform algorithm is implemented for automated peak extraction and is combined with a non-negative least squares (NNLS) method to more accurately quantify peak centroids and areas [14,15]. This approach minimizes the contribution of the Compton continuum and the high-frequency Poissonian noise typical of low-count spectra. Further, this method is capable of resolving overlapping peaks that would be missed by some traditional peak identification methods [14]. In Section 5, we present the results of an experiment that demonstrates the accuracy and precision of this peak identification method.

3. Spectral library algorithm

Our isotope identification algorithm uses a spectral library, containing a list of peak energies and branching ratios convolved with the detector efficiency. To form this library, for each isotope we begin with a complete list of all emitted gamma-rays [16] for all isotopes in the ANSI N42.34 standard for handheld detectors

* Corresponding author.
E-mail address: stinnettjacob@gmail.com (J. Stinnett).

http://dx.doi.org/10.1016/j.nima.2014.11.113
0168-9002/

[8]. However, many of these peaks would never be resolvable with a NaI detector, as demonstrated in Fig. 1.

When trying to fit a region with a Gaussian peak, the presence of overlapping peaks can distort the measured centroid to a different location than the true photopeak energies, which can be approximated as a weighted average of the overlapping peaks. The exact measured centroid will generally be a function of the peak-detecting algorithm used and detector resolution. Some algorithms may be capable of resolving some of the overlapping peaks while others may only be able to locate a single peak within an energy range more closely representing that area-weighted average. Further, gamma rays with small branching ratios will be difficult to detect, especially in the presence of noise. Therefore, it is necessary to create a custom library optimized to the response of both of the detector and the peak detection algorithm.

To form a more appropriate library for a NaI detector, we use the library algorithm of [7]. Because significantly overlapping peaks in a spectrum effectively result in a single peak, this library algorithm combines overlapping peaks to form effective peaks. First, a study on the wavelet/NNLS algorithm was conducted to determine the minimum energy gap that the algorithm can resolve (the minimum resolvable energy, or MRE). The MRE is a function of the relative amplitudes of the neighboring peaks, as similarly sized overlapping peaks will be easier to resolve than overlapping peaks of significantly different amplitudes. By finding the MRE at different relative amplitudes and different signal-to-noise ratios, overlapping peaks can be combined to form "effective" peaks that are more representative of the measured centroid positions in a spectrum taken with the specific 2 in. × 2 in. NaI detector used to collect the spectra in this work.

The resulting libraries are smaller, which increases computational efficiency, but also more realistic, as they do not contain peaks that would not be observed. Further, the resulting libraries are specifically tuned for the wavelet algorithm and the detector used for this work. Although many of the centroids in this adapted library are not the true photopeak energies of each isotope, they are instead a reflection of the resolving power of the wavelet peak detection method and the response function of NaI detectors, given the wavelet algorithm and the specific NaI detector used in this research. The adapted library for ^{152}Eu is shown in Fig. 2, which is significantly more practical than the complete library shown in Fig. 1. The reduced libraries improve the identification accuracy of the isotope identification algorithm discussed in Section 4, as the identification algorithm would not penalize isotopes for peaks that are not resolvable.

4. Bayesian isotope identification algorithm

Given a data set containing the centroids and areas found by the peak identification algorithm, a naive Bayes classifier is used to quantify the posterior probability that an isotope is present given the data. Naive Bayes classifiers compute the posterior probabilities for each class (in this case, for each isotope) by assuming that the components of the data set are independent. While this is generally an unrealistic assumption, these classifiers work remarkably well and often compete with more sophisticated techniques [17,18].

Our Bayesian identification algorithm compares the data set to each isotope in our library to compute the posterior likelihoods [6]. For this problem, Bayes' theorem says that for a given library isotope M and data set D, the posterior probability of M given D is

$$P(M|D) \propto P(D|M)P(M). \tag{1}$$

We previously proposed approximate models to compute the likelihood $P(D|M)$ that accounted for the peak energies, ratios of areas of neighboring peaks, how many of the data peaks correspond to the library isotope, and how many of the library peaks correspond to the data set [6]. In this paper, we present results validating and optimizing these models.

Before calculating the likelihood and the posterior distributions, the algorithm first matches each peak in the data set with the nearest peak (by energy) in the library isotope M. To reasonably restrict the size of the neighborhood considered for matching, only matches for the library peak of energy E_L are considered if the data peak energy is within the interval $(E_L-\theta_w(E_L), E_L+\theta_w(E_L))$, where $\theta_w(E_L)$ is the limit on the maximum distance allowed for peak matching:

$$\theta_w(E_L) = \theta_w^{min} + \frac{E_L}{3\ \text{MeV}}(\theta_w^{max} - \theta_w^{min}). \tag{2}$$

Using a neighborhood that grows with energy reduces the effect of calibration drift on the identifications.

After the peak matching stage, the likelihood $P(D|M)$ is approximated as the product of four terms:

$$P(D|M) = f_{LPI} f_{DPI} f_{PP} f_{AR} \tag{3}$$

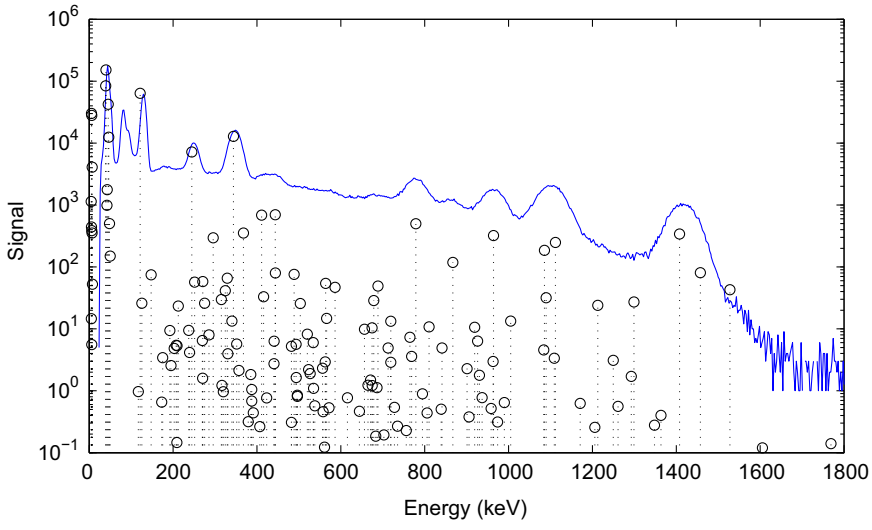

Fig. 1. A NaI spectrum of ^{152}Eu, with an overlay showing all of the gamma-rays emitted. The original library peaks shown are scaled by branching ratios times the detector efficiency. Note that the low energy peaks unaccounted for by ^{152}Eu are Pb K-shell x-rays.