# A Modular NMF Matching Algorithm for Radiation Spectra

Melissa L. Koudelka
Sensor Exploitation Applications
Sandia National Laboratories
mlkoude@sandia.gov

Daniel J. Dorsey
Systems Technologies
Sandia National Laboratories
djdorse@sandia.gov

## Abstract

*In real-world object identification systems, the operational mission may change from day to day. For example, a target recognition system may be searching for heavy armor one day, and surface-to-air assets the next, or a radiation detection system may be interested in detecting medical isotopes in one instance, and special nuclear material in another. To accommodate this "mission of the day" type scenario, the underlying object database must be flexible and able to adjust to changing target sets. Traditional dimensionality reduction algorithms rely on a single basis set that is derived from the complete set of objects of interest, making mission-specific adjustment a significant task. In this work, we describe a method that uses many limited-size individual basis sets to represent objects of interest instead of a single unifying basis set. Thus, only the objects of interest for the mission at hand are used at any given time, and additional objects can be added to the system simply by training a basis for the new object. We demonstrate the modular identification system on the problem of identifying radioisotopes from their gamma ray spectra using nonnegative matrix factorization.*

## 1. Introduction

Modern identification systems are often required to be agile, and able to include or exclude objects of interest on the fly as the mission of the day dictates. Traditional subspace methods used for object identification have monolithic databases that require recalculation to include additional objects, or remove unneeded objects. The technique presented here avoids that shortcoming by producing a set of compact, modular object representations that can be added or removed as needed. We use nonnegative matrix factorization to produce physically realizable basis sets for each object of interest. Incoming test data is scored against desired templates to produce a match score that can be thresholded to produce a "match or no match" decision. Performance is demonstrated on the problem of identifying radioisotopes

from their gamma ray spectra.

Section 2 outlines previous work with subspace techniques as well as radioisotope identification. Section 3 describes our modular nonnegative matrix factorization technique. Section 4 discusses the application of the technique to radioisotope identification, and Section 5 summarizes the work and discusses areas for future research.

## 2. Previous work

Matrix factorization is commonly used for data reduction in applications such as face recognition [1], [2], [3], [4]. In typical applications, a large database of training images represented in matrix form is decomposed and represented with limited loss in fidelity by a small set of representative vectors called basis vectors. The form of those basis vectors is dictated by constraints applied to the decomposition, e.g., orthogonality of basis vectors in principal components analysis (PCA) [5], statistical independence in independent component analysis (ICA) [2], or nonnegativity of matrix elements in nonnegative matrix factorization (NMF) [6]. NMF is of particular interest in applications such as radiation spectra identification (where the signatures represent counts of gamma ray emissions, and therefore cannot be negative) because the basis vectors are physically realizable instances of the real-world data. In contrast, PCA would allow for negative elements, which are not possible in the case of physical measurements of radiation counts. In addition, NMF constraints tend to yield a representation that is "parts-based" instead of holistic, corresponding to the intuitive notion that each basis vector should capture discrete parts that can be selectively recombined to yield a reconstruction.

Another key element of previous work with matrix factorization techniques for target identification is the fact that these algorithms seek global basis sets that apply to a large group of objects, e.g., human faces. This is non-ideal if the set of objects the system is seeking changes, as in a "mission of the day" type scenario. Adding a new object with distinct features to the database means recomputing the basis set. This is impractical in real-

world object identification systems, which may be rather fluid in the objects they seek to identify, meaning they must be agile in terms of updateability.

To date, several researchers have addressed the problem of radiation spectra identification [7], [8]. In contrast to the work discussed here, previous research has generally focused on high fidelity data and a limited number of shielding configurations and source-to-detector geometries. One of the most widely used and effective techniques uses calculated spectra as exemplars for materials of interest [9]. While the technique is effective, changes in shielding configuration and detector-to-source geometry alter the spectrum and can reduce the quality of a match against an exemplar if a spectrum from a similar configuration is not in the library. Other techniques have used principal components analysis (PCA) to represent a library of isotopes [10], or maximum likelihood estimators to determine the probability that a particular source produced the given spectrum [7]. Peak search algorithms [11] are another common technique for identifying high fidelity radiation spectra, however, these techniques are less effective when applied to systems with lower resolution spectra acquired by sensors such as those considered here.

## 3. Modular nonnegative matrix factorization

Most subspace-based matching algorithms put all of the training data for all of the objects of interest into one matrix and generate a single basis vector set that spans all objects. The coefficients on those vectors that best fit a particular piece of test data are then used for discrimination. In contrast, we treat each object individually, generating a separate basis vector set for each one. The ability of a basis set to accurately reconstruct a piece of test data is then used for discrimination. We refer to this technique as nonnegative matrix factorization matching, or NMFM. Training data that is representative of the variation expected in real-world data is used. For example, with high range resolution radar signals, one template spans a narrow range of viewing angles; with gamma ray spectra one template might span a range of material shielding thicknesses.

### 3.1 Template training and basis determination

The training data examples make up the columns of a matrix, $\mathbf{V}$. Nonnegative matrix factorization attempts to approximate $\mathbf{V}$ with the product of two matrices, $\mathbf{W}$ and $\mathbf{H}$, with $\mathbf{W}$ the basis vector set and $\mathbf{H}$ the coefficients or encoding vectors. Typically, the dimension of $\mathbf{W}$ is significantly lower than that of $\mathbf{V}$.

$$\mathbf{V} \approx \mathbf{WH} \qquad (1)$$

The entries in $\mathbf{V}$, $\mathbf{W}$, and $\mathbf{H}$ are all nonnegative, making this method well-suited to datasets representing physical phenomena that cannot contain negative values, such as gamma ray spectra. We use the divergence update rule from Lee and Seung to iteratively determine the basis and coefficients that best approximate $\mathbf{V}$. [12]

$$\left[\mathbf{H}_{(n)}\right]_{i,j} = \left[\mathbf{H}_{(n-1)}\right]_{i,j} \frac{\left[\mathbf{W}'_{(n-1)}\mathbf{V}\right]_{i,j}}{\left[\mathbf{W}'_{(n-1)}\mathbf{W}_{(n-1)}\mathbf{H}\right]_{i,j}} \qquad (2)$$

$$\left[\mathbf{W}_{(n)}\right]_{i,j} = \left[\mathbf{W}_{(n-1)}\right]_{i,j} \frac{\left[\mathbf{VH}'_{(n)}\right]_{i,j}}{\left[\mathbf{W}_{(n-1)}\mathbf{H}_{(n)}\mathbf{H}'_{(n)}\right]_{i,j}} \qquad (3)$$

where subscript ($n$) indicates the solution obtained at the $n$th iteration.

For analysis purposes, it is necessary that the basis generation process be repeatable. Hence, we use a deterministic algorithm to initialize the basis set $\mathbf{W}$ and the coefficients or encoding vectors $\mathbf{H}$. Namely, we use the nonnegative double singular value decomposition method in [13] as a starting point.

### 3.2 Scoring test signatures

The "template" for a particular object is the set of basis vectors $\mathbf{W}$ that represent it. To score a piece of test data, $\mathbf{x}$, against a template, the test signature is decomposed in the basis $\mathbf{W}$ to yield an approximate reconstruction. Decomposing $\mathbf{x}$ thus amounts to selection of a single vector of coefficients $\mathbf{h}$ with a fixed basis $\mathbf{W}$, a fixed test signature $\mathbf{x}$ and a residual term $\mathbf{r}$.

$$\mathbf{x} = \mathbf{Wh} + \mathbf{r} \qquad (4)$$

The decomposition can then be posed as an optimization problem with the following objective function:

$$g(\mathbf{x}, \mathbf{Wh}) = \sum_{i=1}^{K} (x_i - (Wh)_i)^2 \qquad (5)$$

This objective function allows $\mathbf{h}$ to be determined using computationally efficient least-squares optimization techniques, an important distinction given the large number of signatures that must generally be scored against templates in the types of systems discussed here.

Once the coefficients are determined, they are used to reconstruct the signature. The mean-squared error between that reconstruction and the original test signature then provides a measure of the match quality in the form of match score $S$:

$$S = \left[ \frac{\sum_{i=1}^{K}(x_i - (Wh)_i)^2}{\sum_{i=1}^{K} x_i^{\,2}} \right]^{\frac{1}{2}} \qquad (6)$$

Since minimum error is the goal, smaller scores indicate stronger matches. Note that signature amplitude scaling is performed to mitigate the impact of variability across signatures due to physical phenomena such as signal amplitude and object length.

## 3.3 Modularity considerations

It is key that the basis size remain small for this type of modular system to work effectively. The algorithm relies on the fact that only close matches can be reconstructed with the given basis set. Hence, we use a very small basis size, typically 2 or 3 vectors, which tends to force multiple features to be represented in a single vector. This allows the basis to capture some intra-target variation, while making it very difficult to reconstruct out-of-class targets due to the limited amount of flexibility. If the basis size was to grow and features to become separated between the different vectors, it becomes easier to reconstruct signatures that are not from the intended object.

Figure 1 shows an example system block diagram with N objects. To identify the presence or absence of a particular object in a piece of test data, one need only run that piece of test data against that object's template. Thus, only the objects of interest for the mission at hand are used at any given time, and additional objects can be added to the system simply by training a basis for the new object.
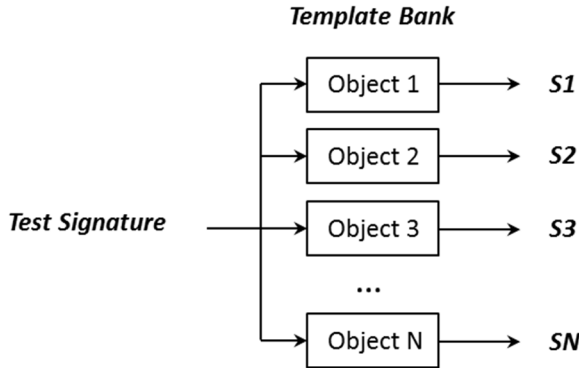


Figure 1: Modular NMFM basic system diagram

## 4. Radioisotope identification application

We have applied the modular NMFM technique to the problem of identifying radioisotopes from their gamma ray spectra. Spectroscopic radiation detection systems capture data in the form of gamma ray spectra and may operate at points of interest to detect specific materials.

Much of the data that these systems acquire contains ambient background or naturally occurring materials that are of little interest. When objects of interest are observed, however, immediate identification is often required. Given the large amount of data and the limited availability of trained spectroscopists, automated methods are needed to analyze the data and identify radioisotopes that may be present. In the most challenging scenario, spectra are collected from highly-shielded radioactive sources with very limited total counts due to short measurement times.

Radiation detection systems detect gamma rays emitted when an isotope undergoes decay. The radiation spectrum is then a 1D signature that indicates the number of detected gamma rays as a function of energy. The longer the system is able to count (and thus, the greater the total number of counts in the measurement), the higher the quality of the signature. Radiation spectra that can be used to train and test the algorithms under development will be required in this effort. For this purpose, we use the Sandia National Laboratories' Gamma Detector Response and Analysis Software (GADRAS) [14] to compute radiation spectra.

## 4.1 Background handling

Radiation spectra contain background radiation from the ambient environment. Traditionally, an estimate of this background has been subtracted from the signature prior to matching. In the NMFM algorithm, we treat the current background radiation estimate as an additional basis vector, $\mathbf{b}$. We can then find the best fit coefficients $\mathbf{h}$ to represent a 1D test signature $\mathbf{t}$ in isotope basis $\mathbf{W}$ (with basis vectors in columns, $\mathbf{w_i}$), and background estimate $\mathbf{b}$.

$$\mathbf{t} \approx h_1\mathbf{w_1} + h_2\mathbf{w_2} + \beta\mathbf{b} \qquad (7)$$

Once the coefficients have been determined, the NMFM-fitted background level can be subtracted from the test signature, and the remaining energy is decomposed to give a foreground only match score. By doing this subtraction, we avoid the problem of matching background only and producing a strong match score.

$$\mathbf{t_w} = \mathbf{t} - \beta\mathbf{b} \qquad (8)$$

$$\widehat{\mathbf{t_w}} = h'_1\mathbf{w_1} + h'_2\mathbf{w_2} \qquad (9)$$

$$S = |\mathbf{t_w} - \widehat{\mathbf{t_w}}| \qquad (10)$$

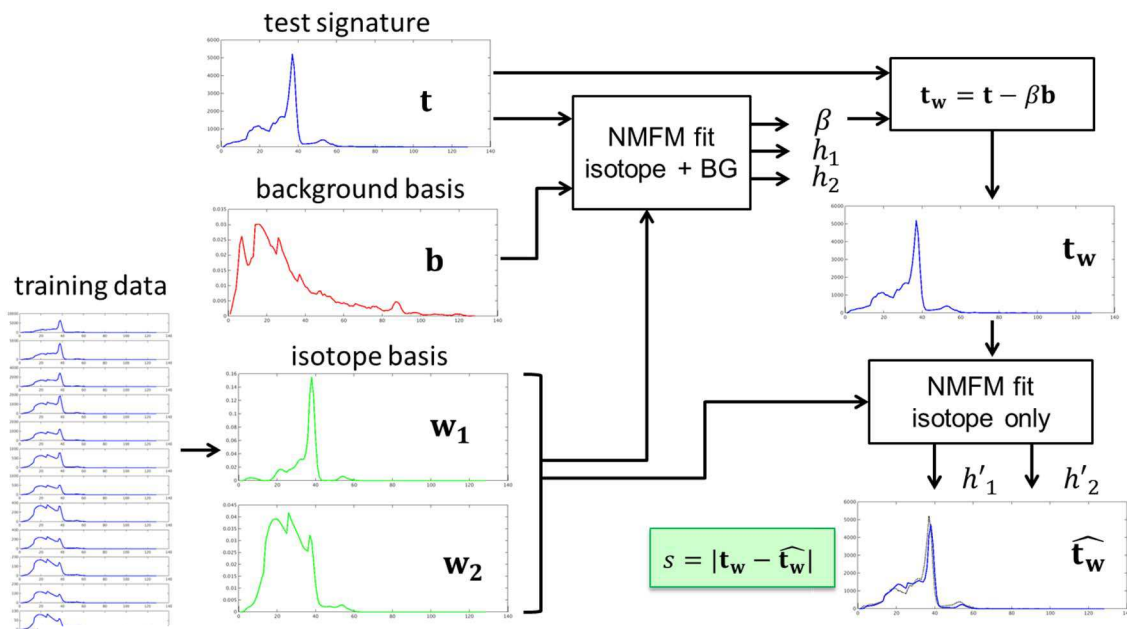Figure 2 shows this background handling and scoring procedure in the form of a block diagram.

**Figure 2: Data flow for background handling and scoring**

## 4.2 Experimental set-up

We have generated a set of training data that can be used to generate templates for a number of radioisotopes under different shielding conditions. In addition, we have generated a large set of independent test data to exercise those templates. The available tools for generating gamma ray spectra are highly accurate at predicting the appearance of these signatures, and so this is considered to be a very realistic test [14]. Table 1 shows the list of isotopes and shieldings used. Shieldings are specified in terms of an atomic number (e.g., Z10).

**Table 1: Isotopes and shieldings included in experiments**

| Isotope | Shielding |
|---------|-----------|
| Ba133 | Z10, Z26, Z50, Z74 |
| I131 | Z10, Z26, Z82 |
| Tc99m | Z7, Z10, Z13, Z26 |
| Cs137 | Z6, Z26, Z82 |

For each material and shielding in the list, a separate modular NMFM template was constructed from the available training data using the process outlined above. The training data was generated assuming a 100 second collection time, meaning the resulting profiles are extremely low noise. The amount of shielding was varied across the test set, with typically 60 different thicknesses included in a single training set. Each signature contains 1024 bins of data with the majority of the energy in the first 500 bins.

The test data was then run against all of the available templates to produce match scores. The test data is comprised of 2000 novel instances of the in-class materials in 2-4 different shielding configurations. The collection times for these signatures were varied from 0.1 second to 10 seconds, meaning the total gamma ray emission counts in each signature varied from the low hundreds to more than 10,000 counts. The noise in these signatures is Poisson distributed [8]; hence a lower count value translates to greater noise. Figure 3 shows an example training signature with 100 second collection time, and example test signatures with collection times of 0.1, 1, and 4 seconds to demonstrate the amount of noise present in very short collection time signatures. Note the total counts for each signature are indicated in the figure.
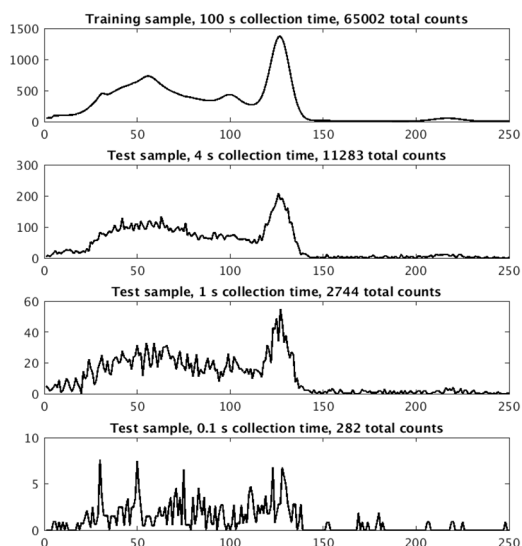


**Figure 3: Sample training data (top) and test data for I131 at 4 seconds, 1 second, and 0.1 second collection times.**
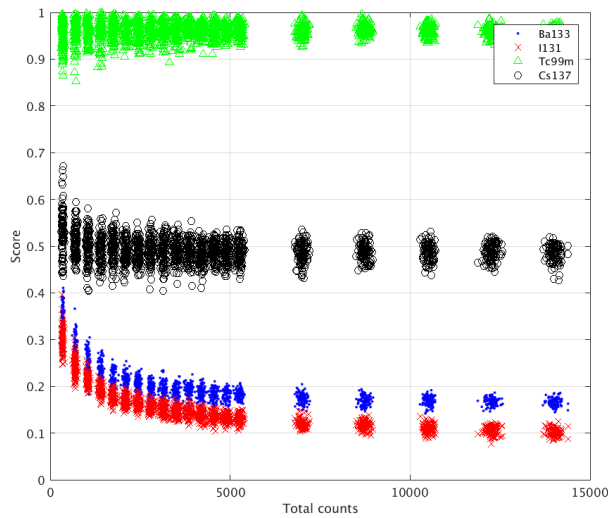
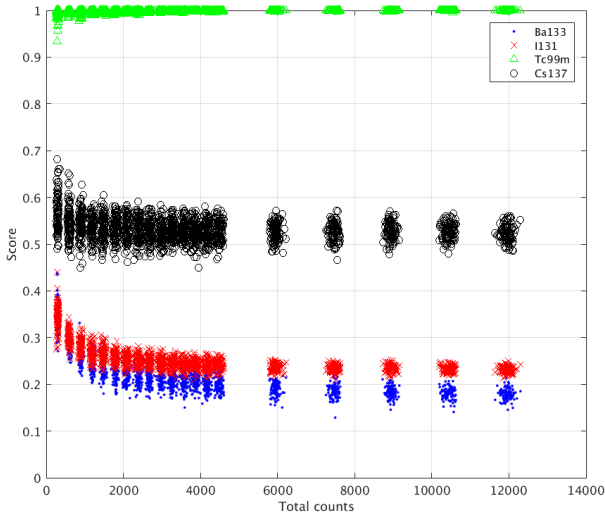**Figure 5: NMFM score for each template vs. total counts in the test signature for an I131 test source.**



**Figure 4: NMFM score for each template vs. total counts in the test signature for a Ba133 test source.**

## 4.3 Results

The ability to discriminate between different materials is heavily dependent on the total number of counts in the signature, and thus the measurement time and shielding.

Figure 4 and Figure 5 demonstrate this, showing the NMFM score achieved by each template vs. the total counts in the test signature for an I131 test source and a Ba133 test source respectively. In these examples, each template has 2 basis vectors in addition to the background estimate. Note that in both cases, the ability to discriminate the correctly matched isotope from the remaining isotopes, indicated by the separation in scores, improves as the signature quality improves. Even with a
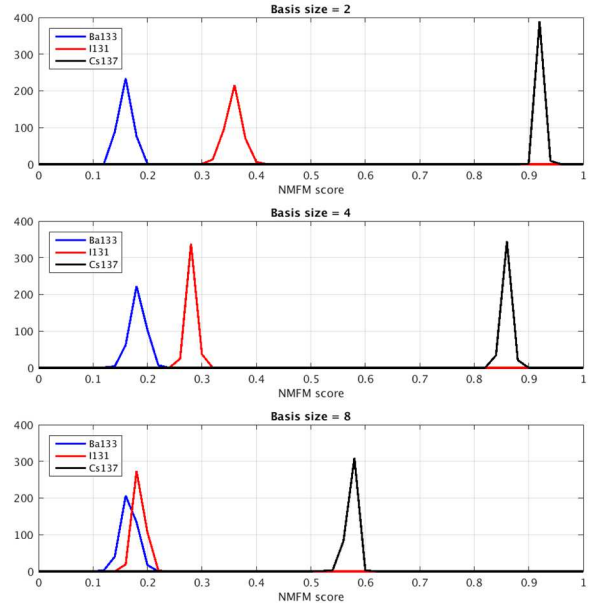


**Figure 6: Score histograms for Ba133 test source vs. Ba133, I131, and Cs137 templates with increasing basis size.**

very small number of counts, the field of possible matches is significantly reduced by the resulting scores. For example, with only hundreds of counts, it is clear that the I131 test source is not Tc99m or Cs137, although the scores for the correct match of I131 and the close confuser Ba133 are both reasonable.

We stated in Section 3.3 that the size of the basis is very important to the ability to discriminate objects. While a more accurate reconstruction of the correctly matched object may be achievable with a larger basis, it also becomes easier to reconstruct confusers. Figure 6 demonstrates this point, showing the NMFM score histograms for a test source of Ba133 against the Ba133, I131, and Cs137 templates, using only signatures with at least 7000 counts. Note that with a basis size equal to 2, the separation between Ba133 and I131 is very good, meaning that the two radioisotopes are easily distinguished. As the basis size increases to 4 and then 8 vectors, the separation between the scores for the Ba133 and I131 templates gets smaller and smaller. The Cs137 template's scores move closer to the in-class distribution as well.

## 5  Summary and Future Work

We have presented a modular system for object identification that uses nonnegative matrix factorization to construct individualized templates for each object of interest. We have demonstrated the system on radiation spectra identification, indicating that we are able to separate different materials with strong target separation in some cases.

To date, the system has been tested on synthetically-

generated data. While synthetically-generated data for this application is generally accepted as realistic, a real-world test would still be highly informative. We intend to apply the system to real-world data in the near future. In addition, we will consider the feasibility of quantifying variables including the amount and type of shielding applied to the object under test. We will also take the existing Matlab-based research code and develop it into a more user-friendly system to aid spectroscopists in analyzing field data.

## References

[1] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions.," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994.

[2] A. Hyvarinen and E. Oja, "Independent component analysis: Algorithms and applications.," *Neural Networks,* vol. 13, no. 4, pp. 411-430, 2000.

[3] D. Lee and H. Seung, "Algorithms for non-negative matrix factorization.," *Advances in Neural Information Processing Systems,* vol. 13, pp. 556-562, 2001.

[4] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1991.

[5] G. Strang, Linear algebra and its applications, Academic Press, Inc., 1980.

[6] D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," *Advances in Neural Information Processing Systems,* vol. 13, pp. 556-562, 2001.

[7] T. Burr and M. Hamada, "Radio-isotope identification algorithms for NaI gamma spectra," in *Algorithms 2009*, 2009.

[8] J. Powell and C. Nelson, "An analysis of radiological threat identification algorithms," Sandia National Laboratories, SAND2009-5986, Albuquerque, NM, 2009.

[9] D. Mitchell, "DHSIsotopeID and full-spectrum analysis methods in GADRAS," Sandia National Laboratories, SAND2009-2544P, Albuquerque, NM, 2009.

[10] S. Labov and et. al., "Foundations for improvements to passive detection systems -- final report," Lawrence Livermore National Laboratory, UCRL-TR-207129, 2004.

[11] G. Phillips and K. Marlow, "Automatic analysis of gamma-ray spectra from germanium detectors," *Nuclear Instruments and Methods,* vol. 137, pp. 525-536, 1976.

[12] D. Lee and H. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature,* vol. 401, pp. 788-791, 1999.

[13] C. Boutsidis and E. Gallopoulos, "SVD based initialization: A head start for nonnegative matrix factorization," *Pattern Recognition,* vol. 41, no. 4, pp. 1350-1362, 2008.

[14] D. Mitchell and J. Mattingly, "Overview of GADRAS Gamma Detector Response and Analysis Software," Sandia National Laboratories, Albuquerque, March 2009.

[15] J. A. Richards and M. L. Koudelka, "High Range Resolution Radar Automatic Target Recognition," Sandia National Laboratories, SAND2008-5821, Albuquerque, NM, 2008.