U.S. DEPARTMENT OF ENERGY

PNNL-21972

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Low-Intrusion Techniques and Sensitive Information Management for Warhead Counting and Verification: FY2012 Annual Report

KD Jarman
BS McDonald
SM Robinson
AJ Gilbert
TA White
WK Pitts
AC Misner
A Seifert

October 2012

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

**Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov**

**Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm**

This document was printed on recycled paper.
(9/2003)

PNNL-21972

Low-Intrusion Techniques and Sensitive Information Management for Warhead Counting and Verification: FY2012 Annual Report

KD Jarman*
BS McDonald
SM Robinson
AJ Gilbert
TA White
WK Pitts
AC Misner
A Seifert

October 2012

Prepared for the Department of Energy Office of Non-Proliferation and Verification Research and Development (NA-22) under U.S. Department of Energy Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

* Corresponding author: kj@pnnl.gov

Executive Summary

Progress in the second year of this project is described by the series of technical reports and manuscripts that make up the content of this report. These documents summarize successes in our goals to develop our robust image-hash templating and material-discrimination techniques and apply them to test image data.

Building on our efforts in FY2011 to survey radiological imaging technology, we convened a series of panel discussions with imaging and arms control verification experts at PNNL to define the set of attributes, and imaging technologies to confirm those attributes, on which this project will focus. Two manuscripts have been drafted on the outcomes, which emphasize the fact that imaging techniques can increase confidence in verification inspections by confirming additional attributes outside the reach of historically considered, aggregate methods such as gamma-ray spectroscopy. These manuscripts will be completed in FY2013 and delivered as technical reports. In the first manuscript we consider attributes specifically accessible by imaging for the declared AT400-R materials-storage configuration and list imaging methods capable of measuring these attributes. We estimate the ability of passive or active, neutron or photon imagers to measure different attributes and note complementary system combinations. We also identify several nominal and existing multi-modal systems for confirming an expanded set of attributes and suggest avenues of further hardware and algorithm development. The second, classified manuscript addresses the applicability of attributes specific to warheads rather than the AT400-R, and techniques for a nuclear warhead counting regime. Our informal ranking of technologies for confirming these attributes in both cases strengthened an emphasis on ORNL neutron imaging systems as a target for testing and further development of our algorithms.

For templating, we developed and tested components of a general framework for robust image hash templates and demonstrated a high degree of robustness to variation in images of like objects and discriminative power between images of different objects. The approach was applied to x-ray image data collected at PNNL on a scale model of the AT400-R material storage container, and to a generic set of x-ray data with spherical objects of varying material in a plastic container with a varied internal structure. Results on the scale model AT400-R container are summarized in an INMM paper, which is being extended to a manuscript for submission to an arms-control-related journal. Results on the generic images are summarized in a manuscript under revision for submission to a journal related to image processing and/or secure information transmission.

For our attribute-based material-discrimination, an iterative regression technique was developed to address the challenge of similarity in attenuation coefficients for different materials. The technique was demonstrated on simulated x-ray images of the AT-400R container. Results can be used to verify declared material configurations as well as provide a mass estimate for SNM contained within the inspected object. Comparison with other mass estimates (e.g., Pu mass estimates provided by gamma spectroscopy or neutron multiplicity counting) behind an information barrier enable improved verification of material presence without disclosure of sensitive information. Results are summarized in an INMM paper.

Our university collaborators include experts in inverse problems, and they are directing a PhD student, Andy Gilbert (U. Texas at Austin), on methods to improve the attribute-based, material-discrimination techniques. Andy successfully presented his research proposal and relocated to PNNL where he is completing his research under this project. Accomplishments in FY12 include testing of

various minimization techniques to achieve the best performance and regularization of the objective function (a measure of goodness of fit) to improve results when attempting to find few materials using noisy image data. A manuscript is being completed on the application of inverse problems methods to baggage inspection using commercially available spectral detectors. (The baggage-inspection task is a surrogate for the treaty-verification task using the same mathematical and physical models, potentially easing classification concerns. All publications from this project will be properly reviewed by our ADC office.) The utility of these methods has also been tested for application to nuclear-material-container inspections and a complementary paper is to be written for this application. The inverse-problems-based approach will be applied to complementary forms of imaging data in the next year (e.g., photon and neutron radiography). Andy Gilbert's PhD proposal contains further details, which will be provided to NA-22 as a separate report.

A review of the project by an external panel took place on January 13, 2012. The panel consisted of Keith Tolk (Milagro Consulting), Helen White (AWE), Todd Peterson (Vanderbilt Univ.), and Dick Kroeger (SPAWAR). A report on the review has been provided through PMIS.

In this final year of the project, our attribute-based material discrimination and template-based robust-hashing techniques will be refined with a focus on application to additional image data that are representative of verification challenges. In particular, we will pursue application to relevant images from colleagues identified at ORNL (e.g. the Nuclear Materials Identification System) and LLNL (Compton camera) in addition to testing on x-ray images of mock storage and other scenarios. Ideally, our techniques will be applied to images collected under the recent NA-24 Weapons Measurement Campaign, if access to those images can be granted to PNNL. Otherwise, an alternative and benign test object should be identified for imaging with the systems to be used in the warhead counting campaign. One candidate, for example, would be the criticality safe cylinders used for system development at ORNL and Y-12.

With guidance from University of Texas (Austin) advisor Dr. Mark Deinert and other University of Texas professors and in collaboration with PNNL project staff, Mr. Gilbert will continue work on algorithms for using energy-dependent x-ray data to noninvasively determine the presence of declared materials and extend the algorithms to multiple, complementary image data sources.

Acknowledgments

This research was sponsored by the U.S. Department of Energy National Nuclear Security Administration Office of Nonproliferation and Verification Research and Development (NA-22). Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle under contract DE-AC05-76RL01830.

Acronyms and Abbreviations

FY	fiscal year
INMM	Institute of Nuclear Materials Management
LLNL	Lawrence Livermore National Laboratory
ORNL	Oak Ridge National Laboratory
PMIS	Project Management Information System
PNNL	Pacific Northwest National Laboratory
SNM	special nuclear material
AWE	Atomic Weapons Establishment (UK)
SPAWAR	Space and Naval Warfare Systems Command

Contents

Publications

T.J. Janik , K.D. Jarman, S.M. Robinson, A. Seifert, B.S. McDonald, and T.A. White, "Image Hashes as Templates for Verification", Proc. INMM Annual Meeting, July, 2012. (attached)

T.J. Janik, K.D. Jarman, and T.A. White, "Robust and Discriminative Image Hashing for Secure Object Verification," submitted to IEEE T. Inf. Foren. Sec., 2012, in revision for resubmission. (draft attached)

T.J. Janik and K.D. Jarman, "Secure and Robust Template Methodology Using Image Hashing for Nuclear Arms Control Verification," in preparation.

S.M. Robinson, K.D. Jarman, A. Seifert, B.S. McDonald, A.C. Misner, T.A. White, E.A. Miller, and W.K. Pitts, "Image-Based Material Discrimination Algorithms for Arms Control", Proc. INMM Annual Meeting, July, 2012. (attached)

S.M. Robinson SM, K.D. Jarman, W.K. Pitts, A. Seifert, A.C. Misner, M.L. Woodring, and M.J. Myjak. "Imaging for dismantlement verification: Information management and analysis algorithms," NIM-A, 662(1):81-89.

Presentations

B.S. McDonald, A. Seifert, T.A. White, S.M. Robinson, E.A. Miller, K.D. Jarman, A.C. Misner, and W.K. Pitts, "Image-Based Verification: Some Advantages, Challenges, and Algorithm-Driven Requirements," IEEE Richland Nuclear & Plasma Science Society Meeting, Richland, WA, December 13, 2011.

B.S. McDonald, "Enabling Radiation Imaging Techniques for Arms Control Verification", Project on Nuclear Issues (PONI) Conference, Richland, WA, April 18-19, 2012.

K.D. Jarman, "Low-Intrusion Techniques and Sensitive Information Management for Warhead Counting and Verification", presented by W.K. Pitts at RadSensing 2012, June 2012, PNNL-SA-88354.

A.J. Gilbert, "Noninvasive Material Discrimination Using Spectral Radiography and an Inverse Problem Approach," UT-Austin Nuclear Engineering PhD Proposal, August 2012.

A summary of our work was included in a presentation by Arden Dougan at the ESARDA meeting in October 2011.

Dr. James Fuller (Dept. of State, Bureau of Arms Control, Verification and Compliance, Special Govt. Employee) requested material from this project to incorporate into a briefing for Dr. Rose Gottemoeller at the Department of State in April as well as to colleagues at Princeton and Microsoft. We provided Dr. Fuller with our INMM papers from 2011 and slides from an internal presentation on hash image templates to inform his presentations.

Appendix: Publications Summarizing FY12 Results

1. T.J. Janik , K.D. Jarman, S.M. Robinson, A. Seifert, B.S. McDonald, and T.A. White, “Image Hashes as Templates for Verification”, Proc. INMM Annual Meeting, July, 2012. PNNL-88946.

2. T.J. Janik, K.D. Jarman, and T.A. White, “Robust and Discriminative Image Hashing for Secure Object Verification,” submitted to IEEE T. Inf. Foren. Sec., 2012, in revision for resubmission. PNNL-89103.

3. S.M. Robinson, K.D. Jarman, A. Seifert, B.S. McDonald, A.C. Misner, T.A. White, E.A. Miller, and W.K. Pitts, “Image-Based Material Discrimination Algorithms for Arms Control”, Proc. INMM Annual Meeting, July, 2012. PNNL-88945.

PNNL-SA-88946

Image Hashes as Templates for Verification

Tad Janik, Ken Jarman*, Sean Robinson, Allen Seifert, Ben McDonald, and Tim White
Pacific Northwest National Laboratory
July 17, 2012

Abstract

Imaging systems can provide measurements that confidently assess characteristics of nuclear weapons and dismantled weapon components, and such assessment may be needed in future verification for arms control. Yet imaging is often viewed as too intrusive, raising concern about the ability to protect sensitive information. In particular, the prospect of using image-based templates for verifying the presence or absence of a warhead, or of the declared configuration of fissile material in storage, may be rejected out-of-hand as being too vulnerable to violation of information barrier (IB) principles. Development of a rigorous approach for generating and comparing reduced-information templates from images, and assessing the security, sensitivity, and robustness of verification using such templates, are needed to address these concerns. We discuss our efforts to develop such a rigorous approach based on a combination of image-feature extraction and encryption-utilizing hash functions to confirm proffered declarations, providing strong sensitive data security while maintaining high confidence for verification. The proposed work is focused on developing automated techniques that may enable the comparison of non-sensitive hashed image data outside an IB. We present an assessment of the performance of our techniques on the basis of a methodical and mathematically precise framework.

Introduction

Modern imaging technology provides an exceptional capability for providing and quantifying detailed properties of imaged objects. The challenge in the area of Nuclear Arms Control Verification for imaging, as with any measurement technology, is to collect necessary and sufficient evidence in such a way that it can be used to verify a proffered declaration about a weapon or a weapon component without compromising sensitive information [9, 10, 16]. In such settings, there is a host undergoing inspection and a monitor who carries out the inspection. The monitor must be able to trust that verification is accurate, and the host must be able to trust that no sensitive information is disclosed. Information barriers consisting of a combination of software and hardware mechanisms designed to protect sensitive information, are a crucial part of verification. See [3, 5, 11, 13, 14, 17, 18], for example, for more detailed definition and discussion on the roles of host and monitor, the concept of information barriers, and the challenges of certification, authentication, and trust.

One approach to address this challenge is to use data templates. In template matching methods, a measurement of a trusted item serves as a reference to which measurements of inspected items are compared, and a match provides verification. Here we refer to a trusted item as one believed by the monitor to be consistent with the declaration to be verified. Storing detailed and thus highly sensitive reference data in non-volatile memory would likely present an unacceptable risk of disclosure, and data hashing methods have been used in the past to address this concern. For instance, the TRIS measurement system developed by Sandia National Laboratory relies on the generation of (non-imaging) template measurements that are hashed and test measurements are compared to the decrypted template behind a formal information barrier [15]. Similar approaches

*Corresponding Author: tel: (1) 509.375.6360, email: kj@pnnl.gov

could potentially be used for imagery, but images may heighten concern over sensitive information management.

An alternative is to incorporate data reduction, feature extraction, and hashing techniques to isolate the content in images that is sufficient for verification as a template, while prohibiting the extraction of sensitive information from the template. Ideally, such a template might even be compared and stored outside an IB. In this paper we present procedures to support Nuclear Arms Control Verification based on coupling image data reduction with encryption-utilizing robust hash functions intended for use in template matching along with IBs. The work is focused on developing secure, robust, tamper-sensitive and automatic techniques that process all the sensitive measurements behind the information barrier and produce a non-invertible template to be used for comparisons outside of the IB (see also [4]).

Robust (a.k.a. perceptual) hashing is a transformation that maps high-dimensional content of an object (e.g., image, document, biometric template) into a low-dimensional vector space of short bit strings to enable fast comparison and searches. In contrast to conventional purely cryptographic hash functions (e.g., MD5, SHA-2) which are highly sensitive to every bit of input data, robust hashing is sensitive to an object's content rather than the integrity of all of the object's data bits (see [2] for the review of major perceptual hashing algorithms). Many of these algorithms rely on the correspondence between perceptual similarity of images and coarse image representation based on several standard image processing techniques such as Discrete Cosine Transform, Fourier-Mellin Transform, and Singular Value Decomposition. However, the images of items subject to arms-control verification may preserve perceptual similarity even when items are altered, for instance, to cover up diversion of fissile material. The hashing techniques based on the image coarse representations applied to such tampering scenarios don't have enough discriminative power to successfully support verification decision processes.

Thus, while perceptual or robust image hashing provides a starting point, we must extend the concept to find a proper balance between robustness, discriminability and security, as well as simplicity. To be precise we define robustness, as above, to be an insensitivity to non-content variation in the data; discriminability to be accurate differentiation between objects that are as declared and objects that are not; and security to be the inability to obtain sensitive information from the output of image hashing. Discriminability here is considered in terms of two possible cases of host tampering with the imaged object. The first case is represented by simple removal or replacement of a portion of the object. The second tampering case is represented by a more elaborate attempt to replicate the "correct" image hash by learning from observation and knowledge of the hash algorithm one or more variations on the objects that produce essentially the same image hash. To achieve the robustness, discriminability, and security objectives we have proposed histogram-based techniques which exploit the invariance of the relative frequencies of pixel intensities in histograms [4]. To reduce further the risk of disclosure of sensitive information to the monitor, we introduce relation-based data reduction and a joint key strategy (a codebook construction) that relies on random permutations of histogram bins and reduction to short bit strings. This process increases irreversibility and unpredictability of the histogram-based hash values. The goal is to develop a procedure which minimizes the ability of a potential attacker to learn details of the full image and thus the imaged item from the observed hash values.

Hash-Template-Based Verification Process

A schematic of a hash-template-based verification concept is presented in Fig. 1.

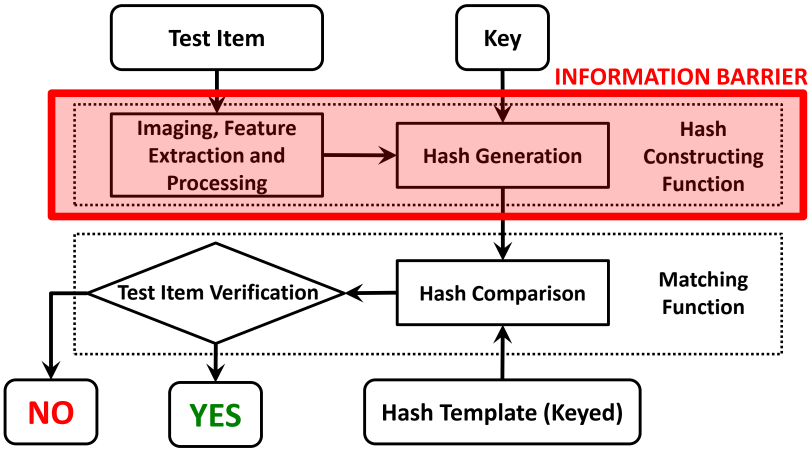


Figure 1. Basic elements of the hash-based verification concept.

In this proposed system, all sensitive processes such as imaging an inspected item, discriminative feature extraction and processing, and generation of a hash value (a relatively short binary string) are performed behind an Information Barrier. For verification, the hash of the image data from a trusted item is pre-computed using a joint key, stored as a template, and compared with the hash of an inspected item's image. Details of joint key construction are beyond the scope of this paper; an example can be found in [15]. In practice, the hash value produced for two images of exactly the same object would not be identical, due to typical distortions in image acquisition and processing. A distance metric between the hash of the inspected item's image and the original template is computed and compared with a threshold as a test for compliance with declaration. Since the hash values are binary strings we use the standard Hamming distance (bit error rate) as the distance metric.

Secure, Robust, and Discriminative Hashing: Definition and Desired Properties

By hash function here we mean the combination of transformation of measurements (e.g. feature extraction, here represented by histograms) followed by reduction to a short bit string. In general, the hash function that we propose takes two inputs, an image and a key to produce (preferably at low complexity) a binary string of length q (preferably relatively small). That is, denoting a set of images by I and a set of keys by K, the hash function H is defined as

$$(I, K) \ni (i, k) \rightarrow h = H(i, k) \in \{0, 1\}^q. \qquad (1)$$

To support arms control verification the hash functions need to satisfy several desired properties. To formally define these properties let $i_{ident} \in I$ denote the image of an item essentially identical to the trusted item, whose image is $i \in I$. In other words, i_{ident} is a slightly distorted (rotated, compressed, noisy, etc.) version of i. We will also use the notation $i_{ident} \approx i$. Moreover, let $i_{diff} \in I$ denote an image of an item that is distinct from the trusted item. That is, $i_{diff} \neq i$ may be the image of an altered item. An example would be a material storage container with the correct material but in different chemical form than declared.

The *robustness* property requires that the hash values of images (subjected to insignificant or legitimate global distortions) that represent the same item are close to each other or, in other words, identical with high probability:

$$Prob\{\ H(i, k) = H(i_{ident}\,, k\ )\ \} \geq 1 - \theta_1, \text{ for all } i,\ i_{ident} \in I,\ i_{ident} \approx i,\ k \in K,\ 0 < \theta_1 < 1. \quad (2)$$

The *discriminability* (collision-resistant) property requires that the hash values of any pair of images in *I* for distinctive (e.g., tampered) items must be different with high probability:

$$Prob\{\ H(i, k) \neq H(i_{diff}\,, k)\ \} \geq 1 - \theta_2, \text{ for all } i,\ i_{diff} \in I,\ i_{diff} \neq i,\ k \in K,\ 0 < \theta_2 < 1. \quad (3)$$

The property (3) is very important since it must extremely difficult for the host to tamper with the inspected item and yet obtain a hash value very close to that of the trusted item. Another property supporting tamper resistance as well as information *security* of the hash function is its *unpredictability*, requiring that the output hash value must be approximately uniformly distributed among all possible *q*-bit outputs when the key varies over *K* for a fixed input image *i*:

$$Prob\{\ H(i, k) = h\ \} \approx 1/2^q, \text{ for all } h \in \{0, 1\}^q. \quad (4)$$

Also for hash *security*, i.e., inability of the monitor to deduce detailed knowledge about the items being imaged based on the observed hash values needs to include the *one-way* hashing or *non-invertibility* property: a high degree of computational difficulty in identifying image data *i* that produce a given hash value *h* of an imaged item. This property can be expressed as follows, borrowing from the literature on cryptographic hashing. First, it must be difficult to find a pre-image i^* that produces a given hash value:

$$Prob\{ \text{ find } i^* \in I \text{ such that } H(i^*, k) \approx h\ \} \leq \theta_3, \text{ for a given } h \in \{0, 1\}^q,\ 0 < \theta_3 < 1. \quad (5)$$

Second, it must be difficult to find a pre-image $i^{\#}$, strictly different from the image of the trusted item, that produces a hash value matching that produced by the trusted item:

$$Prob\{ \text{ find } i^{\#} \in I,\ i^{\#} \neq i \text{ such that } H(i^{\#}, k) \approx H(i, k)\ \} \leq \theta_4, \text{ for a given } i \in I,\ 0 < \theta_4 < 1. \quad (6)$$

The required hash properties (2) – (6) clearly conflict with each other. For example, property (2) calls for robustness under insignificant image data perturbations while (3) requires minimization of collision (matching hashes) probabilities for distinctive images. As an illustration, using very crude features can yield high robustness but also high probability of encountering matches (collisions) between images of distinct items. Conversely, perfect randomization of the hash values would virtually eliminate collisions, but also makes the hash much less robust. Depending on particular applications secure hash functions need to satisfy these conflicting properties to some extent and/or facilitate the trade-offs [8, 21]. The accuracy parameters θ_1, θ_2, θ_3, and θ_4 defined in (1-6) provide a quantitative measure for overall performance and must be made as small as possible, optimized with respect to requirements of a given verification scenario.

Histogram-Based Image Feature Hashing Procedures

In order to resolve the trade-off between discriminability and noninvertibility of the hash function, we implement and analyze hash procedures generated using the relative quantities extracted from gray-level (pixel intensity) image histograms [19, 20]. The image histograms are invariant to scaling (up to a multiplicative factor), rotation, and translation, and they should preserve their basic shape under other moderate image distortions (noise, compression, etc.) Defining an image histogram as:

$$G_q = \{\ g_q(j) \text{ such that } j = 1,\dots,\ q\}, \text{ where } g_q(j) \geq 0 \text{ is the number of pixels in the } j^{th} \text{ bin} \qquad (7)$$

(with bins formed in the usual manner of ranges of pixel intensity) we compute a hash value by comparing numbers of pixels in consecutive histogram bins:

$$h(j) = 1 \text{ if } g_q(j) > g_q(j+1) \text{ for } j = 1,\dots,q\text{-}1,\ h(q) = 1 \text{ if } g_q(q) < g_q(1),\ h(j) = 0 \text{ otherwise.} \qquad (8)$$

A binary string with length equal to the total number of bins results. This procedure is generalized to include concatenation of hash values evaluated for subimages of a partition of the original image. An implementation of a keyed encryption scheme may involve a random permutation of the histogram bins as well as random shuffling of the subimages. Many other keying techniques (e.g., random projections or dithering and distributed source coding) can be also utilized [6]. A schematic of the process of generating values of a keyed hash function is presented in Fig. 2. First an image is partitioned into a number of subimages, on each of which a histogram is formed; then the bins are rearranged and combined and a hash is generated on the results according to the number of pixels counted in each adjacent bin.

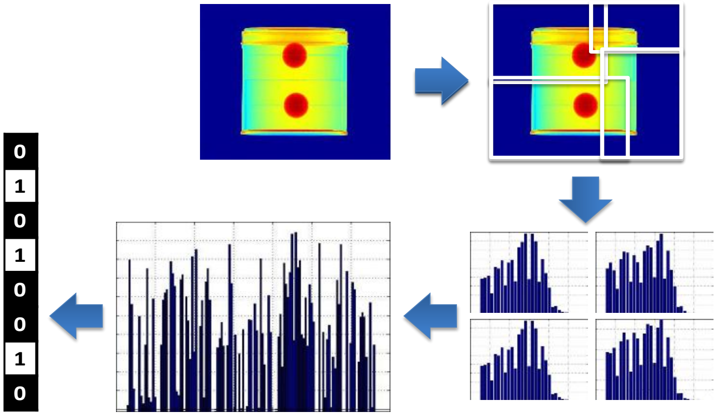


Figure 2. Process of generating values of a keyed hash function (random tiling of image data, histogram calculation, random bin permutation, binary hash extraction).

Experimental Setting

To illustrate application of the proposed image hash template techniques and analysis of the desired properties, we use a set of 300 x-ray images of a container designed to mimic a scale model of a special nuclear material storage container. The AT400R container was designed for use at the Mayak Fissile Material Storage Facility and in that purpose is declared to contain 2 2-kg spheres of plutonium. In our example, spheres of glass and ceramic are placed in the mock AT400R to represent nominal as-declared and simple diversion scenarios. This model is placed between a strong x-ray source and a large-area imaging detector. An image is then formed from the transmitted photon flux. The image is representative of the attenuation of the x-ray source through the intervening materials (including the container and contents). The digital images are formed with pixel values ranging from 0 to 255 (8 bits) and belong to 3 groups (each with 100 slightly different images) divided according to their content: group A, representing images of "as declared" items, consists of images representing a container with 2 ceramic balls; groups B and C, representing diversion, have images of a container with two glass balls and one ceramic and one glass ball, respectively. In order to obtain slight variations of single images of the same object for comparison, the 100 images within each group were created while rotating a roundtable with the container. Figure 3 shows the mock AT400R with a few objects that can be placed inside, and an example image from each group using a color scale from blue, corresponding to low attenuation, to red, corresponding to high attenuation. Greater translation of each item was evident among the full set of original images; a very simple automated cropping step was used to produce the images in the figure. Note that with the bare eye the difference between the images in Fig. 3 is hardly noticeable; perceptually we might say they are the same.

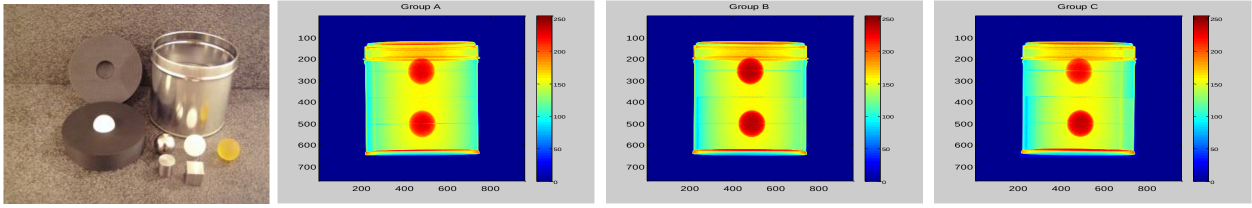


Figure 3. Mock AT400R container and three categories of images used to assess the quality of proposed hashing techniques. Two ceramic balls, two glass balls, and one ceramic and one glass ball (second, third, and fourth panels, respectively).

Robustness and Discriminative Power

We computed hash values for all the images using the procedure (7-8) and computed intra- and intergroup Hamming distances between each pair of hash values. For this test we use $q_1 = 32$ histogram bins and 4 subimages (tiles) which yield hash binary strings of length $q = 4q_1 = 128$. This is a significant reduction of information that can greatly reduce the ability to extract sensitive details of the original image. The results of the comparison are presented in Fig. 4.

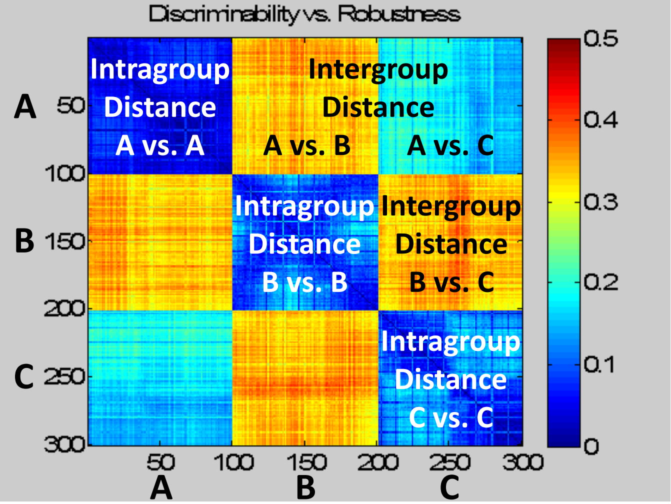


Figure 4. Graphical representation of high discriminative and reconstructive (robustness) power of the proposed hashing technique applied on sample data (300 images).

The displayed *300*-by-*300* array represents intra- and intergroup Hamming distances (bit error rates) between the computed hashes for 300 tested images, i.e., each *(i, j)* – element of this array is a color representation of the distance between hashes of the i^{th} and j^{th} images. In this figure, images numbered from 1 to 100 belong to group A, images numbered from 101 to 200 to group B, and images numbered from 201 to 300 to group C. There are apparently strong dissimilarities between the "as declared" group A and "diversion" group B (blue color in the figure corresponding to strong similarity and red to strong dissimilarity). Greater similarity is evident between groups A and C, indicating higher likelihood of error in distinguishing between items in these two groups.

The test indicates both robustness and high discriminative power of the proposed hashing scheme applied to the mock AT400R container and its content. A decision on the consistency of an inspected item with declaration is made by comparing the distance between hash values with a threshold. Therefore, the quality of the decision depends on the separation between the hash distances of the same object and hash distances of different objects. The overall robustness of the procedure can be measured by the maximum intragroup distance between hash values derived from different images of the same object. The maximal intragroup distances are $d_{maxA} = 0.1719$, $d_{maxB} = 0.2266$, and $d_{maxC} = 0.2031$. The overall discriminative power of the hashing technique can be derived from the minimal intergroup distances which are $d_{minAB} = 0.2891$, $d_{minAC} = 0.1016$, and $d_{minBC} = 0.2500$. The only overlap is between hash distances for images in group A and C. To provide a quantitative assessment in the verification decision problem, we may use the False Accept Rate (FAR) and False Reject Rate (FRR). Here, FAR represents the fraction of inspected items that are falsely determined to be "as declared," and FRR represents the fraction of inspected items that are falsely determined to be "not as declared". Selecting a threshold between 0.1719 and 0.2891 leads to zero error in discriminating between group A and B items.

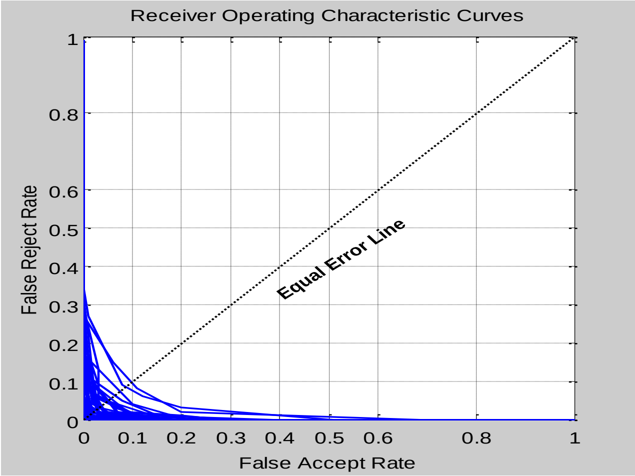


Figure 5. Error curves for image hashes of items in Groups A and C.

The FRR approximates the probability that two images of the nominal item are determined to be different, providing an estimate of θ_1 in formula (2). Similarly, the FAR provides an estimate of θ_2 in formula (3). The probability of falsely categorizing items in the decision process involving groups A and C can be estimated as a function of a threshold c by plotting the ordered pair *(FAR(c), FRR(c))* when each item in group A is verified against the remaining items A and all the items of group C. The resulting 100 curves (a version of Receiver Operating Characteristic curves) provide a basis for choosing a threshold to balance robustness and discriminability through a corresponding choice of θ_1 and θ_2. A standard trade-off is to select a threshold that leads to similar FAR and FRR. The curves displayed in Fig. 5 indicate that this trade-off yields $\theta_1 \approx \theta_2 < 0.1$ and the average error rates over 100 tests are approximately equal to *0.02*.

Unpredictability of the Image Hash Template

In this section we assess unpredictability. This relates to both the risk described above of a host altering an inspected item in such a manner as to closely match the template hash value, and to the risk of a monitor "inverting" the hash value to obtain sensitive information. In our proposed scheme, both parties could have knowledge of the hashing algorithm and the image hash value, and the host has knowledge of the items being inspected. A joint key, unknown to both parties, may be generated by combining a host-selected key and a monitor-selected key, for example [15]. An altered item could be rearranged in such a way that the gray-level histogram of a new image is close to one produced from the declared item and simultaneously attempting to deceive the complementary verification procedure (e.g., gamma spectrum). Thus, under this scenario, tamper resistance depends on a level of changeability of the scrambled histogram bin values (and, more importantly, their ratios) when the joint key varies, i.e., when the histogram bins of a subimage's random shuffling are randomly permuted.

The factorial growth characteristic of permutations suggests their potential for increased unpredictability of the produced hash values, thereby providing greater defense against this type of attack [1]. However, the unpredictability property of the hash requires that the output hash value must be approximately uniformly distributed among all 2^q possible q-bit outputs when the histogram bins are permuted and the bins ratios yield the bits distribution.

Figure 6 shows an example analysis of unpredictability. In each panel, all 8! (40,320) permutations of the bins of an example 8-bin histogram are generated, and from each result the hash is computed (based on the relative magnitude in permuted bins as described previously); the number of each possible 8-bit hash value is then plotted. The ideal result would be that each hash value is equally likely, which would be indicated by a uniform distribution (the red line). The first panel (left) represents the most typical case of different numbers of pixels in each pixel intensity histogram bin, and the other panels represent less likely cases in which either two or four of the bins have exactly the same number of pixels. Note that it is only the relative magnitude (number of pixels) in each adjacent bin that matters in calculating the hash value, so the result will be the same as in the first panel for any image histogram with the same relative histogram pattern. From the figure it is clear that the hash outputs are not uniformly distributed. Further analysis is needed to quantify how far from uniform the distributions are, and more importantly, how that impacts the hash unpredictability.

This analysis does not account for partitioning the image randomly into subimages, which increases unpredictability. To further improve the theoretical security of the hashing technique, partitioning and/or bin distribution for each subimage may be optimized to obtain the highest variability of the hash values measured by an information measure such as differential (Shannon) entropy. Differential entropy is a well-established measure of the complexity of the relationship between the image, its features (in our case, histograms), and the key [6, 7, 12]. Although we have not performed formal optimization of the differential entropy associated with the proposed hashing technique, ad-hoc computations suggest that the *128*-bit histogram/random tiling/bin permutation-based hash function used in our computations yields relatively high entropy values.

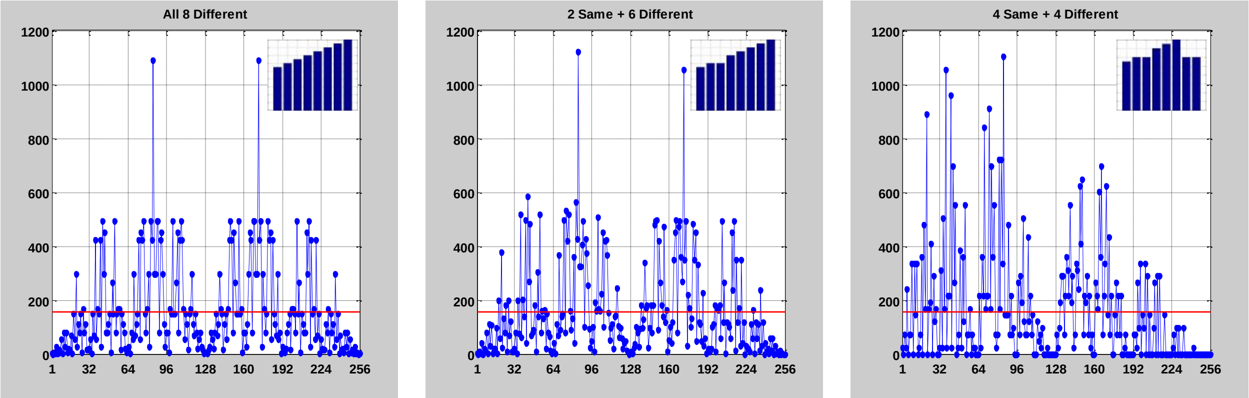


Figure 6. Distribution of hash values under permutations of the 8 bin values, for three cases. Left: different number of pixels in each bin; middle: same number of pixels in exactly two bins; right: same number of pixels in exactly four bins. The red line indicates the uniform distribution ($40320 / 256 = 157.5$ possible outcomes per permutation).

Conclusions

The main contribution of this work is a design and implementation of a simple histogram-based hash scheme for templating based on image data for arms control verification tasks. The scheme is designed to address the key criteria of robustness, discriminability, and security, for which we provided formal mathematical definitions. Analysis methods for evaluating performance in terms of these criteria were demonstrated and applied. Further analysis is needed to complete a framework for evaluating the security of image templates. The success of the approach on the example studied here is by virtue of the insensitivity of the histogram shape and its bin ratios to moderate image distortion as well as high sensitivity to even minor gray-level image variation indicative of item alteration. As the approach is developed for practical application, formal vulnerability assessment beyond purely mathematical arguments will be needed to test the ability to protect sensitive information against a variety of attacks.

The general concept of non-cryptographic image hashing for generating secure templates differs from previous verification techniques in several ways. Primarily, the goal is a template that can be stored and used outside of an IB. This is a significant advance over previously developed template techniques if it can be achieved. Second, the general concept applies, with modification of the hashing algorithm, to any type of imaging, not just ionizing-radiation-based imaging as in the example used here. Third, the level of feature extraction—here represented by histograms—that is applied prior to hashing could be modified to allow more or less detail in extracted features to suit specific agreements between host and monitor.

Acknowledgements

This research is supported by the U.S. Department of Energy National Nuclear Security Administration Office of Nonproliferation and Verification Research and Development (NA-22). The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy.

References

[1] Betz V., Ueltschi D., and Velenik Y., Random permutation with cycle weights, Ann. of Appl. Prob., 21 (1), 312-331, 2011.

[2] Han S. and Chu C, Content-based image authentication: current status, issues, and challenges, Int. J. Inform. Security, 9, 19-32, 2010.

[3] Harahan J.P., On-Site Inspections Under the INF Treaty, A history of the On-site Inspection Agency and Treaty Implementation, 1988-1991: Treaty History Series, Government Printing Office, 1993.

[4] Jarman K., Robinson S., Seifert A., McDonald B., Misner A., White T., Miller E., and W.K. Pitts, Non-invertible transforms for image-based verification, Proc. INMM Annual Meeting,, Palm Desert CA, 2011.

[5] Kouzes R.T. and Fuller J.L., "Authentication of monitoring systems for nonproliferation and arms control," Proc. Symposium on Int. Safeguards: Verification and Nuclear Material Security, IAEA Vienna, Austria, 2001.

[6] Koval O., Voloshynovskiy S., Beekhof F., and Pun T., Security analysis of robust perceptual hashing, Proc. SOIE-IS&T Electronic Imaging, SPIE 6819, 2008.

[7] Mao Y. and Wu M., Unicity Distance of Robust Image Hashing, IEEE Trans. Inf. Forensics and Security, 2 (3), 215-230, 2007.

[8] Monga V. and Evans B.L., Perceptual image hashing via feature points: Performance evaluation and tradeoffs, IEEE Trans. Image Proc., 15 (11), 2006.

[9] Pitts W.K., Jarman K.D., Miller E.A., McDonald B.S., Misner A.C., Myjak M.J., Robinson S.M., Seifert A., Seifert C.E., and Woodring M.L., "Advantages of dual mode imaging for managing sensitive imaging information," Proc. INMM Annual Meeting, Baltimore, MD, 2010.

[10] Robinson S.M., Jarman K.D., Pitts W.K., Seifert A., Misner A.C., Woodring M.L., and Myjak M.J., Imaging for dismantlement verification: Information management and analysis algorithms, Nucl. Instr, Meth. Phys. Res. Sec. A, 662 (1), 81-89, 2011.

[11] Smith M., MacArthur D.W., Karpius P., Vo D.T., Thron J.L., Frame K., Johansen N., Valdez J., and Williams R., Next Generation Attribute Measurement System, Proc. INMM Annual Meeting, Nashville, TN, 2008.

[12] Swaminathan A., Mao Y., and Wu M., Robust and Secure Image Hashing, IEEE Trans. Inf. Forensics Security, 1 (2), 215-230, 2006.

[13] Technology R&D for Arms Control, Tech. Rep. NNSA/NN/ACNT-SP01, Office of Nonproliferation Research and Engineering (ONRE), U.S. Department of Energy (Spring 2001).

[14] Tolk K.M., Trusting the Data in Arms Control and International Safeguards, Proc. INMM Annual Meeting, Palm Desert CA, July 2011.

[15] Tolk K.M., Lucero R.L., Seager K.D., Mitchell D.J., Laub T.W., and Insch K.W., Trusted radiation identification system, Proc. INMM Annual Meeting, Indian Wells, CA, 2001.

[16] White H., Chambers D.M., Keir D., Allen K., Burjan A., and Owen M., Research into nuclear Arms Control Verification at the UK Atomic Weapons Establishment, Proc. INMM Annual Meeting, Tucson, AZ, 2009.

[17] Whiteson R. and MacArthur D.W., "Information Barriers in the Trilateral Initiative: Conceptual description," Los Alamos National Laboratory, LAUR-98-2137 (Unlimited distribution, available from the United States Department of Energy, Office of Scientific and Technical Information website, www.osti.gov), 1998.

[18] Wolford, Jr., J.K. and White G.K., "Progress in gamma ray measurement information barriers for nuclear material transparency monitoring," Proc. INMM Annual Meeting, New Orleans, LA, 2000.

[19] Xiang S., Kim H, Huang J., Histogram-based image hashing scheme robust against geometric deformations, Proc. of the 9th Workshop on Multimedia and Security, Dallas, TX, 2007.

[20] Xie N., Ling H., Hu W., and Zhang X., Use bin-ratio information for category and scene classification, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR), San Francisco, CA, 2010.

[21] Zauner C., Implementation and benchmarking of perceptual image hash functions, PhD Thesis, University of Applied Sciences Hagenberg, Austria, 2010.

Robust and Discriminative Image Hashing for Secure Object Verification

Tadeusz J. Janik, *Member, IEEE*, Kenneth Jarman, and Timothy White

Abstract—**Imaging systems can provide measurements that confidently assess characteristics of objects under a process of secure verification. Secure verification is defined as a confirmation that the object is in agreement with the provided declaration without compromising sensitive information. Yet imaging is often viewed as too intrusive, raising concern about the ability to protect sensitive information. For example, the prospect of using image-based templates for arms control treaty verification, personal identity management systems or airport security may be rejected out-of-hand as being too vulnerable to violation of information barrier (IB) principles. Development of a rigorous approach for generating and comparing reduced-information templates from images, and assessing the security, sensitivity, and robustness of verification using such templates, are needed to address these concerns. We discuss our efforts to develop such a rigorous approach based on a combination of image-feature extraction and encryption-utilizing hash functions to confirm proffered declarations, providing strong sensitive data security while maintaining high confidence for verification. The proposed work is focused on developing automated techniques that may enable the comparison of non-sensitive hashed image data outside an IB. We present an assessment of the performance of our techniques on the basis of a methodical and mathematically precise framework.**

Index Terms— **robust object verification and authentication, image hashing, data security**

I. Introduction

WHEN protection of sensitive information is critical, the detection, characterization, or prediction of phenomena of interest requires methods for hiding the very information that produces a signature of those phenomena. Concurrently, the potential for illicit manipulation of information demands methods for detection, characterization, or prediction that can withstand and detect various attack strategies. Often these two problems occur in tandem. In our work we focus on techniques for the secure object verification defined as measurements and procedures providing confidence that the declarations concerning the object under consideration are true. The security of this verification refers to the fact that no sensitive information can be revealed to the inspector or a third party. Settings include arms control treaty verification (e.g. using radiation detection equipment to verify that dismantled nuclear weapons components in storage without revealing sensitive design information) [3], airport security (e.g. masking portions of millimeter-wave images used to detect concealed weapons, and destroying the images afterward), identity management systems (e.g. biometric templates for access to personal information), satellite data downloads, and illicit radiological material detection at U.S. ports of entry (e.g. stripping personally identifiable information [PII] from archived data for analysis). Other applications of signatures such as the use of financial records for loan or credit card applications could benefit from the introduction of secure, tamper-detecting transformations of sensitive information. Simple security approaches like checksums, data sensoring (e.g. providing binary rather than feature value output based on thresholding) or data masking (mainly replacement of real data with fake data for testing and development) are insufficient. Traditional secure transmission through encryption is not enough, as decryption is precluded in many cases. Requiring all analysis to take behind formal software and hardware tamper-sensitive information barriers (IBs), and thus never viewed by a human, can address the need in some applications, but is severely limiting. The ideal use of newly discovered and/or dynamic signatures would be outside IBs. What is needed is a methodology for constructing signatures from raw data or derived features that (1) hide sensitive information, (2) communicate crucial information, (3) can be analyzed outside formal IBs, and (4) are resistant to and detect tampering with the features, data, or underlying phenomena. In the paper we present this methodology using image-based hash templates, a process in which a choice of feature extraction methods and cryptographic hashing techniques are combined to produce a signature that is secure, tamper-resistant, and robust according to and guided by the specific scenario and phenomena of interest, within a rigorous mathematical formulation and evaluation framework.

Modern imaging technology offers an exceptional capability for providing and quantifying detailed properties of imaged objects which can be used for their verification [8]. The challenge in many application areas is to collect necessary and

Manuscript received July 2, 2012. This research is supported by the U.S. Department of Energy National Nuclear Security Administration Office of Nonproliferation and Verification Research and Development. The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy.

The authors are with the Pacific Northwest National Laboratory, Richland WA, 99352 USA (Corresponding author: T.J. Janik, phone: 509-375-6513; fax: 509-555-5555; e-mail: tadeusz.janik@pnnl.gov; other emails: kj@pnnl.gov, timothy.white@pnnl.gov).

sufficient evidence in such a way that it can be used to verify a proffered declaration about an object or its component without compromising sensitive information. One approach to address this challenge is to use data templates. In template matching methods, a measurement of a trusted item serves as a reference to which measurements of inspected items are compared, and a match provides positive and objective evidence in the verification process. To protect sensitive information we propose to incorporate data reduction, feature extraction, and hashing techniques to isolate the content in images that is sufficient for verification as a template, while prohibiting the extraction of sensitive information from the template. The work is focused on developing secure, robust, tamper-sensitive and automatic techniques that process all the sensitive measurements behind the information barrier and produce a non-invertible template to be used for comparisons outside of the IB.

Robust (a.k.a. perceptual) hashing is a transformation that maps high-dimensional content of an object (e.g., image, document, biometric template) into a low-dimensional vector space of short bit strings to enable fast comparison and searches. In contrast to conventional purely cryptographic hash functions (e.g., MD5, SHA-2) which are highly sensitive to every bit of input data, robust hashing is sensitive to an object's content rather than the integrity of all of the object's data bits (see [2] for the review of major perceptual hashing algorithms). Many of these algorithms rely on the correspondence between perceptual similarity of images and coarse image representation based on several standard image processing techniques such as Discrete Cosine Transform, Fourier-Mellin Transform, and Singular Value Decomposition. However, the images of items subject to real-life object verification may preserve perceptual similarity even when items are altered, for instance, to cover up diversion of used material. The hashing techniques based on the image coarse representations applied to such tampering scenarios don't have enough discriminative power to successfully support decision processes of verification.

Thus, while perceptual or robust image hashing provides a starting point, we must extend the concept to find a proper balance between robustness, discriminability and security, as well as simplicity. To be precise we define robustness, as above, to be an insensitivity to non-content variation in the data; discriminability to be accurate differentiation between objects that are as declared and objects that are not; and security to be the inability to obtain sensitive information from the output of image hashing. Discriminability here is considered in terms of two possible cases of host tampering with the imaged object. The first case is represented by simple removal or replacement of a portion of the object. The second tampering case is represented by a more elaborate attempt to replicate the "correct" image hash by learning from observation and knowledge of the hash algorithm one or more variations on the objects that produce essentially the same image hash. To achieve the robustness, discriminability, and security objectives we have proposed histogram-based techniques which exploit the invariance of the relative frequencies of pixel intensities in histograms [4]. To reduce further the risk of disclosure of sensitive information to the monitor or the third party, we introduce relation-based data reduction and a joint key strategy (a codebook construction) that relies on random permutations of histogram bins and reduction to short bit strings. This process increases irreversibility and unpredictability of the histogram-based hash values. The goal is to develop a procedure which minimizes the ability of a potential attacker to learn details of the full image and thus the imaged item from the observed hash values.

II. Hash-Template-Based Verification Process

A. Overview of the Concept

A schematic of a hash-template-based verification concept is presented in Fig. 1.

In the proposed system, all sensitive processes such as imaging an inspected item, discriminative feature extraction and processing, and generation of a hash value (a relatively short binary string) are performed behind an Information Barrier.

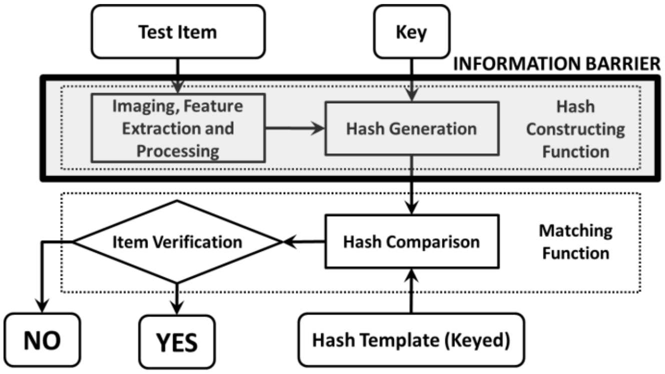


Figure 1. Basic elements of the hash-template-based verification concept.

For verification, the hash of the image data from a trusted item is pre-computed using a joint key, stored as a template, and compared with the hash of an inspected item's image. Details of joint key construction are beyond the scope of this paper; an example can be found in [10]. In practice, the hash value produced for two images of exactly the same object would not be identical, due to typical distortions in image acquisition and processing. A distance metric between the hash of the inspected item's image and the original template is computed and compared with a threshold as a test for compliance with declaration. Since the hash values are binary strings we use the standard Hamming distance (bit error rate) as the distance metric.

B. Secure, Robust, and Discriminative Hashing: Definition and Desired Properties

By hash function here we mean the combination of transformation of measurements (e.g. feature extraction, here represented by histograms) followed by reduction to a short bit string. In general, the hash function that we propose takes two inputs, an image and a key to produce (preferably at low complexity) a binary string of length q (preferably relatively small). That is, denoting a set of images by I and a set of keys by K, the hash function H is defined as

$$(I, K) \ni (i, k) \rightarrow h = H(i, k) \in \{0, 1\}^q. \qquad (1)$$

To support verification process the hash functions need to satisfy several desired properties. To formally define these properties let $i_{ident} \in I$ denote the image of an item essentially identical to the trusted item, whose image is $i \in I$. In other words, i_{ident} is a slightly distorted (rotated, compressed, noisy, etc.) version of i. We will also use the notation $i_{ident} \approx i$. Moreover, let $i_{diff} \in I$ denote an image of an item that is distinct from the trusted item. That is, $i_{diff} \neq i$ may be the image of an altered item. An example would be a material storage container with the correct material but in different chemical form than declared. The *robustness* property requires that the hash values of images (subjected to insignificant or legitimate global distortions) that represent the same item are close to each other or, in other words, identical with high probability:

$$Prob\{\ H(i, k) = H(i_{ident}, k)\ \} \geq 1 - \theta_1,$$

$$\text{for all } i, i_{ident} \in I, i_{ident} \approx i, k \in K, 0 < \theta_1 < 1. \qquad (2)$$

The *discriminability* (collision-resistant) property requires that the hash values of any pair of images in I for distinctive (e.g., tampered) items must be different with high probability:

$$Prob\{\ H(i, k) \neq H(i_{diff}, k)\ \} \geq 1 - \theta_2,$$

$$\text{for all } i, i_{diff} \in I, i_{diff} \neq i, k \in K, 0 < \theta_2 < 1. \qquad (3)$$

The property (3) is very important since it must extremely difficult for the host to tamper with the inspected item and yet obtain a hash value very close to that of the trusted item. Another property supporting tamper resistance as well as information *security* of the hash function is its *unpredictability*, requiring that the output hash value must be approximately uniformly distributed among all possible q-bit outputs when the key varies over K for a fixed input image i:

$$Prob\{\ H(i, k) = h\ \} \approx 1/2^q, \text{ for all } h \in \{0, 1\}^q. \qquad (4)$$

When the keyed hashing algorithm is used as an object verification code, the hash value needs to be highly dependent on the key [13]. If two different keys are used for the same image, the corresponding hash values should be completely different as if they correspond to different content.

Also for hash *security*, i.e., inability of the monitor or a third party to deduce detailed knowledge about the items being imaged based on the observed hash values needs to include the *one-way* hashing or *non-invertibility* property: a high degree of computational difficulty in identifying image data i that produce a given hash value h of an imaged item. This property can be expressed as follows, borrowing from the literature on cryptographic hashing. First, it must be difficult to find a pre-image i^* that produces a given hash value:

$$Prob\{\ \text{find } i^* \in I \text{ such that } H(i^*, k) \approx h\ \} \leq \theta_3,$$

$$\text{for a given } h \in \{0, 1\}^q, 0 < \theta_3 < 1. \qquad (5)$$

Second, it must be difficult to find a pre-image $i^{\#}$, strictly different from the image of the trusted item that produces a hash value matching that produced by the trusted item:

$$Prob\{\ \text{find } i^{\#} \in I, i^{\#} \neq i \text{ such that } H(i^{\#}, k) \approx H(i, k)\ \} \leq \theta_4,$$

$$\text{for a given } i \in I, 0 < \theta_4 < 1. \qquad (6)$$

The required hash properties (2) – (6) clearly conflict with each other. For example, property (2) calls for robustness under insignificant image data perturbations while (3) requires minimization of collision (matching hashes) probabilities for distinctive images. As an illustration, using very crude features can yield high robustness but also high probability of encountering matches (collisions) between images of distinct items. Conversely, perfect randomization of the hash values would virtually eliminate collisions, but also makes the hash much less robust. Depending on particular applications secure hash functions need to satisfy these conflicting properties to some extent and/or facilitate the trade-offs [7, 14]. The accuracy parameters θ_1, θ_2, θ_3, and θ_4 defined in (1-6) provide a quantitative measure for overall performance and must be made as small as possible, optimized with respect to requirements of a given verification scenario.

C. Histogram-Based Image Feature Hashing Procedures

In order to resolve the trade-off between discriminability and noninvertibility of the hash function, we implement and analyze hash procedures generated using the relative quantities extracted from gray-level (pixel intensity) image histograms [11, 12]. The image histograms are invariant to scaling (up to a multiplicative factor), rotation, and translation, and they should preserve their basic shape under other moderate image distortions (noise, compression, etc.) Defining an image histogram as:

$$G_q = \{\ g_q(j) \text{: } g_q(j) \text{ is the pixel number in } j^{th} \text{ bin, } j = 1,\ldots,q\ \} \quad (7)$$

(with bins formed in the usual manner of ranges of pixel intensity) we compute a hash value by comparing numbers of pixels in consecutive histogram bins:

$$h(j) = 1 \text{ if } g_q(j) > g_q(j+1) \text{ for } j = 1,\ldots,q\text{-}1,$$

$$h(q) = 1 \text{ if } g_q(q) < g_q(1), \qquad (8)$$

$$h(j) = 0 \text{ otherwise.}$$

A binary string with length equal to the total number of bins results. This procedure is generalized to include concatenation of hash values evaluated for subimages of a partition of the original image. An implementation of a keyed encryption scheme may involve a random permutation of the optimally distributed histogram bins (to maximize Shannon entropy) as well as random shuffling of the subimages. Many other keying techniques (e.g., random projections or dithering and distributed source coding) can be also utilized [5]. A schematic of the process of generating values of a keyed hash function is presented in Fig. 2.

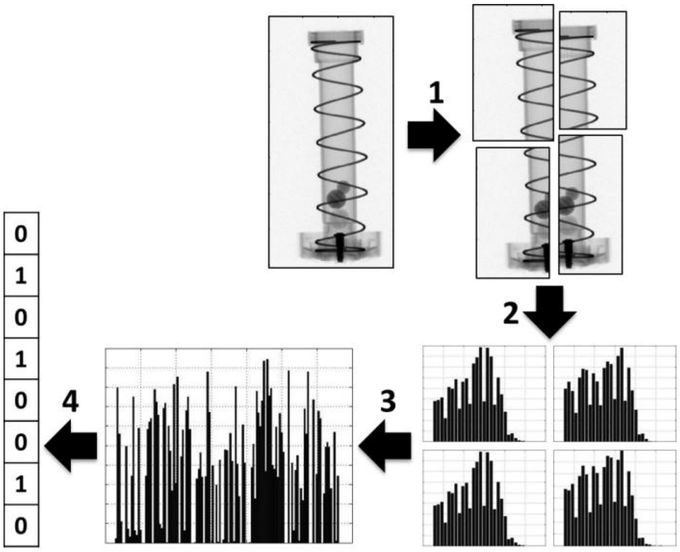


Figure 2. Process of generating values of a keyed hash function (1. random tiling of image data, 2. histogram calculation, 3. random bin permutation, 4. binary hash extraction).

First an image is partitioned into a number of subimages, on each of which a histogram is formed; then the bins are rearranged and combined and a hash is generated on the results according to the number of pixels counted in each adjacent bin.

III. Hash-Template-Based Verification Process

In this section we present experiments to illustrate application of the proposed image hash template techniques and analysis of the desired properties.

A. Experimental Setting

We use a set of 150 x-ray images of a container designed to mimic a scaled-down special material storage container. In our example, 4 spheres of plastic, ruby, aluminum, and jasper are placed in the mock container to represent nominal as-declared and simple diversion scenarios (Fig. 3).

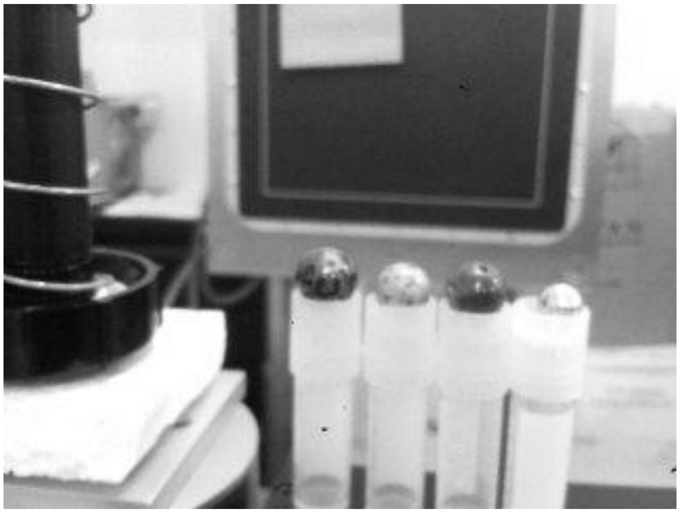


Figure 3. Components to build a mock special container. From left to right: sprinkler head and the plastic, jasper, ruby, and aluminum spheres. The first three are beads and have holes through the center. Detector can be seen in the background.

This model is placed between a strong x-ray source (160kVp) and a large-area imaging detector. An image is then formed from the transmitted photon flux. The image is representative of the attenuation of the x-ray source through the intervening materials (including the container and contents). The digital images are formed with pixel values ranging from 0 to 255 (8 bits) and belong to 3 groups (each with 50 slightly different images) divided according to their content: group A, representing images of “as declared” items, consists of images representing a container with 3 balls (plastic, ruby, aluminum); groups B (4 balls: plastic, jasper, ruby, aluminum) and C (3 balls as in A with swapped order of ruby and aluminum spheres) represent diversion. In order to obtain slight variations of single images of the same object for comparison, the 50 images within each group were created while rotating a roundtable with the container. Table 1 summarizes the distortion level occurring within the images of group A.

TABLE I. Image Distortion Levels in Group A

Distortion Type	Distortion Level		
	Range	*Average*	*St. Dev.*
Container Translation [pixels]	0 – 6	2.1	1.4
Container Rotation [degrees]	0 – 2.4	1.4	0.7
Relative Distance 2 Top Balls [pixels]	0 – 15	7.9	4.5
Relative Rotation 2 Top Balls [degrees]	0 – 40.7	12.5	10.9
Pixel Variability [gray levels]	0 – 5	2.6	1.7

Figure 4 shows an example image from each group using a gray scale from white, corresponding to low attenuation, to black, corresponding to high attenuation. Greater translation of each item was evident among the full set of original images; a very simple automated cropping step was used to produce the images in the figure. Note that with the bare eye the difference between the images in Fig. 4 (especially between group A and C) is hardly noticeable; perceptually we might say they are the same.

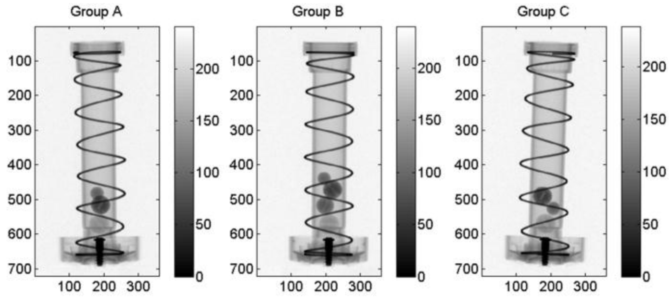


Figure 4. Three categories of images used to assess the quality of proposed hashing techniques. Left: a container with 3 balls (from top to bottom: aluminum, ruby, plastic), middle: the same container with 4 balls (t-b: aluminum, ruby, jasper, plastic) and right: the same container with 3 swaped balls (t-b: ruby, aluminum, plastic).

B. Robustness and Discriminative Power

We computed hash values for all the images using the procedure (7-8) and computed intra- and intergroup Hamming distances between each pair of hash values. For this test we use q_1 = 32 histogram bins and 4 subimages (tiles) which yield hash binary strings of length $q = 4q_1 = 128$. This is a significant reduction of information that can greatly decrease the ability to extract sensitive details of the original image. The image partitioning and bin distribution for each subimage have been optimized to obtain the highest variability of the hash values measured by the Shannon entropy (see also Section III.D.). The results of the comparison are presented in Fig. 5.

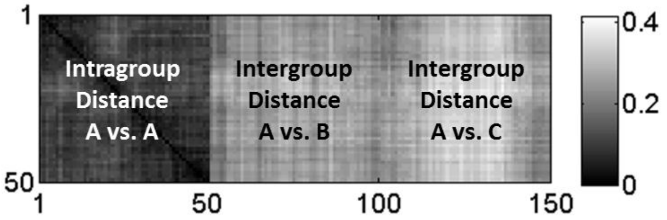


Figure 5. Graphical representation (confusion matrix) of high discriminative and reconstructive (robustness) of the proposed hashing technique applied on sample data (150 images).

The displayed 50-by-150 array represents intra- and intergroup Hamming distances (bit error rates) between the computed hashes for 150 tested images, i.e., each (i, j) – element of this array is a color representation of the distance between hashes of the i^{th} and j^{th} images. In this figure, images numbered from 1 to 50 belong to group A, images numbered from 51 to 100 to group B, and images numbered from 101 to 150 to group C. There are apparently stronger dissimilarities between the "as declared" group A and "diversion" group C (black color in the figure corresponding to strong similarity and white to strong dissimilarity). Greater similarity is evident between groups A and B, indicating higher likelihood of error in distinguishing between items in these two groups.

The test indicates both robustness and high discriminative power of the proposed hashing scheme applied to the mock container and its content. A decision on the consistency of an authenticated item with declaration is made by comparing the distance between hash values with a threshold. Therefore, the quality of the decision depends on the separation between the hash distances of the same object and hash distances of different objects. The overall robustness of the procedure can be measured by the maximum intragroup distance between hash values derived from different images of the same object. The maximal intragroup distances are d_{maxA} = 0.2031, d_{maxB} = 0.1875, and d_{maxC} = 0.2891. The overall discriminative power of the hashing technique can be derived from the minimal intergroup distances which are d_{minAB} = 0.1281 and d_{minAC} = 0.1484. The overlaps in hash intra- and intergroup distances indicate possibility of errors in the verification process. To provide a quantitative assessment in the verification decision problem, we may use the False Accept Rate (FAR) and False Reject Rate (FRR). Here, FAR represents the fraction of inspected items that are falsely determined to be "as declared," and FRR represents the fraction of inspected items that are falsely determined to be "not as declared".

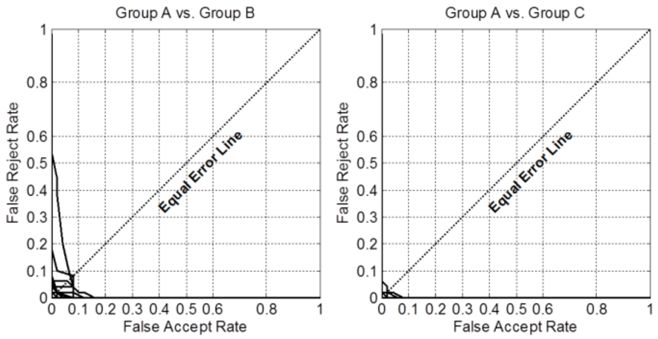


Figure 6. Error curves for image hashes of items in Groups A and B (left) and A and C (right).

The FRR approximates the probability that two images of the nominal item are determined to be different, providing an estimate of θ_1 in formula (2). Similarly, the FAR provides an estimate of θ_2 in formula (3). The probability of falsely categorizing items in the decision process involving groups A and B (A and C) can be estimated as a function of a threshold c by plotting the ordered pair ($FAR(c)$, $FRR(c)$) when each item in group A is verified against the remaining items A and all the items of group B (C). The resulting 50 curves (a version of Receiver Operating Characteristic curves) provide a basis for choosing a threshold to balance robustness and discriminability through a corresponding choice of θ_1 and θ_2. A standard trade-off is to select a threshold that leads to similar FAR and FRR. The curves displayed in Fig. 6 indicate that this trade-off yields $\theta_1 \approx \theta_2 < 0.08$ for groups A and B, $\theta_1 \approx \theta_2 < 0.03$ for A and C, and the average error rates over 50 tests are approximately equal to 0.0085 for A and B and 0.0015 for A and C.

C. Unpredictability of the Image Hash Template

In this section we assess unpredictability. This relates to both the risk described above of a host altering an inspected item in such a manner as to closely match the template hash value, and to the risk of a monitor or third party “inverting” the hash value to obtain sensitive information. In our proposed scheme, all parties may have knowledge of the hashing algorithm and the image hash value, and the host has knowledge of the items being authenticated. A joint key, unknown to both parties, may be generated by combining a host-selected key and a monitor-selected key, for example [10]. An altered item could be rearranged in such a way that the gray-level histogram of a new image is close to one produced from the declared item and simultaneously attempting to deceive the complementary verification procedure (e.g., gamma spectrum). Thus, under this scenario, tamper resistance depends on a level of changeability of the scrambled histogram bin values (and, more importantly, their ratios) when the joint key varies, i.e., when the histogram bins of a subimage’s random shuffling are randomly permuted.

The factorial growth characteristic of permutations suggests their potential for increased unpredictability of the produced hash values, thereby providing greater defense against this type of attack [1]. However, the unpredictability property of the hash requires that the output hash value must be approximately uniformly distributed among all 2^q possible q-bit outputs when the histogram bins are permuted and the bins ratios yield the bits distribution.

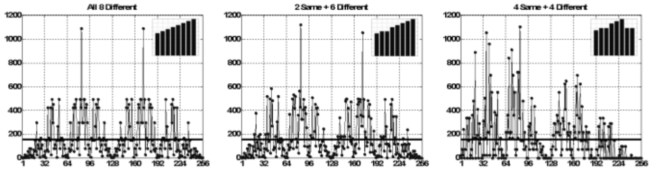


Figure 7. Distribution of hash values under permutations of the 8 bin values, for three cases. Left: different number of pixels in each bin; middle: same number of pixels in exactly two bins; right: same number of pixels in exactly four bins. The horizontal line indicates the uniform distribution (40320 / 256 = 157.5 possible outcomes per permutation curves).

Figure 7 shows an example analysis of unpredictability. In each panel, all 8! (40,320) permutations of the bins of an example 8-bin histogram are generated, and from each result the hash is computed (based on the relative magnitude in permuted bins as described previously); the number of each possible 8-bit hash value is then plotted. The ideal result would be that each hash value is equally likely, which would be indicated by a uniform distribution (the red line). The first panel (left) represents the most typical case of different numbers of pixels in each pixel intensity histogram bin, and the other panels represent less likely cases in which either two or four of the bins have exactly the same number of pixels. Note that it is only the relative magnitude (number of pixels) in each adjacent bin that matters in calculating the hash value, so the result will be the same as in the first panel for any image histogram with the same relative histogram pattern. From the figure it is clear that the hash outputs are not uniformly distributed. Further analysis is needed to quantify how far from uniform the distributions are, and more importantly, how that impacts the hash unpredictability.

This above analysis does not account for partitioning the image randomly into subimages, which increases unpredictability. To further improve the theoretical security of the hashing technique, we have optimized the partitioning and bin distribution for each subimage to obtain the highest variability of the hash values measured by an information measure such as differential (Shannon) entropy. Differential entropy is a well-established measure of the complexity of the relationship between the image, its features (in our case, histograms), and the key [5, 6, 9]. The optimal 128-bit histogram/random tiling/bin permutation-based hash function used in our computations yields Shannon entropy equal to 6.9143.

D. Key Dependence test

The discussed tests use a fixed key (a fixed seed for random tiling and permutations). When the keyed hashing algorithm is used to generate an object verification code, the hash value needs to be highly dependent on the key [13]. In this test, we use all 150 images to validate the key dependence property. For each image, we generate 100 hash values using different seeds for random histogram bin permutations. They are pairwise compared hence there are 4950 hash comparisons for each image and 742500 comparisons in total. For different keys, the corresponding hash values should be as different as possible (Hamming distance closed to 0.5) as if they correspond to different content [13]. The average hash distances for all 150 images are plotted in Fig. 8. All the average hash distances are localized around 0.5 with a very small dynamic range (0.4969, 0.5017). This demonstrates a good randomization mechanism of the proposed keyed hashing procedure.

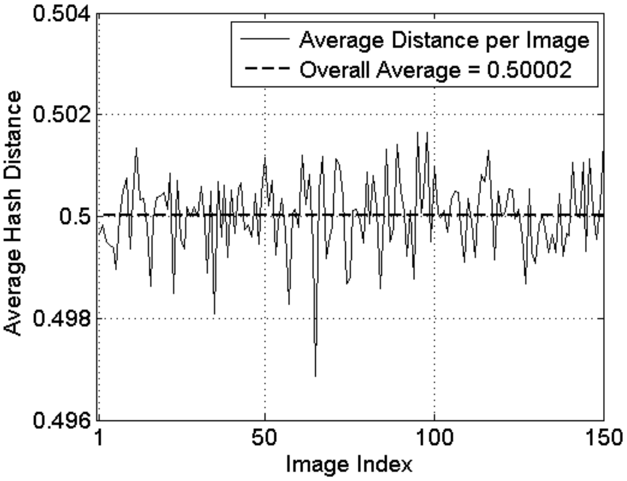


Figure 8. Key dependence test. Average hash distance between hash values by different keys.

IV. Conclusions

The main contribution of this work is a design and implementation of a simple histogram-based hash scheme for templating based on image data for various verification tasks. The scheme is designed to address the key criteria of robustness, discriminability, and security, for which we provided formal mathematical definitions. Analysis methods for evaluating performance in terms of these criteria were demonstrated and applied. Further analysis is needed to complete a framework for evaluating the security of image templates. The success of the approach on the example studied here is by virtue of the insensitivity of the histogram shape and its bin ratios to moderate image distortion as well as high sensitivity to even minor gray-level image variation indicative of item alteration. As the approach is developed for practical application, formal vulnerability assessment beyond purely mathematical arguments will be needed to test the ability to protect sensitive information against a variety of the of attacks.

The general concept of non-cryptographic image hashing for generating secure templates differs from previous verification techniques in several ways. Primarily, the goal is a template that can be stored and used outside of an IB. This is a significant advance over previously developed template techniques if it can be achieved. Second, the general concept applies, with modification of the hashing algorithm, to any type of imaging, not just ionizing-radiation-based imaging as in the example used here. Third, the level of feature extraction—here represented by histograms—that is applied prior to hashing could be modified to allow more or less detail in extracted features to suit specific agreements between the interested parties.

References

[1] Betz V., Ueltschi D., and Velenik Y., Random permutation with cycle weights, Ann. of Appl. Prob., 21 (1), 312-331, 2011.

[2] Han S. and Chu C, Content-based image authentication: current status, issues, and challenges, Int. J. Inform. Security, 9, 19-32, 2010.

[3] Harahan J.P., On-Site Inspections Under the INF Treaty, A history of the On-site Inspection Agency and Treaty Implementation, 1988-1991: Treaty History Series, Government Printing Office, 1993.

[4] Jarman K., Robinson S., Seifert A., McDonald B., Misner A., White T., Miller E., and W.K. Pitts, Non-invertible transforms for image-based verification, Proc. INMM Annual Meeting,, Palm Desert CA, 2011.

[5] Koval O., Voloshynovskiy S., Beekhof F., and Pun T., Security analysis of robust perceptual hashing, Proc. SOIE-IS&T Electronic Imaging, SPIE 6819, 2008.

[6] Mao Y. and Wu M., Unicity Distance of Robust Image Hashing, IEEE Trans. Inf. Forensics and Security, 2 (3), 215-230, 2007.

[7] Monga V. and Evans B.L., Perceptual image hashing via feature points: Performance evaluation and tradeoffs, IEEE Trans. Image Proc., 15 (11), 2006.

[8] Robinson S.M., Jarman K.D., Pitts W.K., Seifert A., Misner A.C., Woodring M.L., and Myjak M.J., Imaging for dismantlement verification: Information management and analysis algorithms, Nucl. Instr, Meth. Phys. Res. Sec. A, 662 (1), 81-89, 2011.

[9] Swaminathan A., Mao Y., and Wu M., Robust and Secure Image Hashing, IEEE Trans. Inf. Forensics Security, 1 (2), 215-230, 2006.

[10] Tolk K.M., Lucero R.L., Seager K.D., Mitchell D.J., Laub T.W., and Insch K.W., Trusted radiation identification system, Proc. INMM Annual Meeting, Indian Wells, CA, 2001.

[11] Xiang S., Kim H, Huang J., Histogram-based image hashing scheme robust against geometric deformations, Proc. of the 9th Workshop on Multimedia and Security, Dallas, TX, 2007.

[12] Xie N., Ling H., Hu W., and Zhang X., Use bin-ratio information for category and scene classification, Proc. IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR), San Francisco, CA, 2010.

[13] L. Weng and B. Preneel, A secure perceptual hash algorithm for image content aythentication, Proc. 12th IFIP International Conference on Communications and Multimedia Security, Ghent, Belgium, 108-121, 2011.

[14] Zauner C., Implementation and benchmarking of perceptual image hash functions, PhD Thesis, University of Applied Sciences Hagenberg, Austria, 2010.

IMAGE-BASED MATERIAL DISCRIMINATION ALGORITHMS FOR ARMS CONTROL

Sean Robinson, Andrew Gilbert, Ben McDonald, Tim White, Ken Jarman, Alex Misner
Pacific Northwest National Laboratory
Richland, WA 99354

ABSTRACT

The Pacific Northwest National Laboratory is developing and evaluating active radiographic image analysis techniques for verifying sensitive objects in an arms control, material control, or warhead counting regime in which sensitive information may be processed. Material discrimination algorithms which attempt to estimate the amount of specific materials present in each pixel of an image can be used to verify pertinent non-sensitive or declared attributes (e.g., the presence of special nuclear material (SNM) within an object of interest) with all image analysis performed behind an information barrier stage, allowing for reporting and storage of non-sensitive attributes only. Techniques proposed here employ spectroscopic detectors to determine the materials present between the source and a single detector pixel, and operate by fitting the attenuated spectrum to a set of expected attenuation spectra for an ensemble of materials. Practical limits on single-pixel material discrimination are defined and suggestions for the optimal use of this technique within the context of arms control or cargo scanning is presented. A limited number of free parameters are expected for analysis on a single pixel, limiting the context of material discrimination to a few representative materials. SNM may be estimated by assuming that two materials lie along the ray from source to detector: one material is parameterized by the attenuation coefficient of plutonium, and the other material is chosen such that a best-fit metric to measured data is optimized. This discrimination method may be used to verify declared material configurations, as well as for mass estimates of SNM contained within an inspected object. Results may be compared with other mass estimates (e.g., Pu mass estimates provided by gamma spectroscopy or neutron multiplicity) behind an information barrier. As only the agreement between mass estimates would be returned, this technique may allow for verification of material presence without disclosure of sensitive information.

INTRODUCTION

Verification of sensitive objects is important to a variety of nonproliferation tasks. Inspection regimes may require the verification that a given object is as expected, without revealing any specific information regarding the object in excess of previously agreed-upon non-sensitive attributes. In this approach, object analysis is performed behind an information barrier, and only simplified information is retained, such as the presence or absence of an isotope or whether the estimated amount of a material passes a threshold [1, 2, 3]. Earlier work has included material discrimination approaches utilizing active radiographic imaging, based on fitting effective geometry and attenuation parameters to images generated for a known geometric shape behind an information barrier [4]. Material discrimination approaches have been employed in which a physical model of photon attenuation provides an approximate means for estimating the physical properties (e.g., thickness or areal density) of materials along that path. These approaches are expected to allow for several metrics, such as Pu mass estimates, that could be very specific and useful in an arms control context.

Further work has shown that an imaging system capable of producing images as a function of energy as well as position would have the additional capability of discrimination on a single-pixel basis by using those spectral differences [5]. In this technique, a nonnegative least squares (NNLS) regression method (in which solutions containing negative material quantities are excluded) is used to estimate the presence of materials within a physical object containing a variety of materials. Here, we assume a detector that can discriminate by energy *E*, given a count spectrum *C*(*E*) in each pixel that corresponds to a line from the source through the object. We consider a set of materials to be estimated in each pixel, denoted by the subscript *i*. Our forward estimate of the count spectrum in a given pixel is given by:

$$C(E) = S(E)D(E)\exp\left[-\sum_i \mu_i(E)\rho_i\right], \tag{1}$$

where *S* is the source emission term, *D* is a detector response function, the *μ* terms are the mass attenuation coefficients for each material (cm^2/g), and the *ρ* terms represent the areal density (g/cm^2) of each material between the source and pixel. Counts are accumulated in N energy bins with bin center $\overline{E}_j$. The baseline spectra $C_0\left(\overline{E}_j\right)$ are estimated from the detector response in a pixel with no intervening materials. The image is then evaluated pixel by pixel, estimating the areal densities for each material:

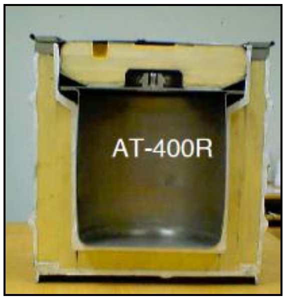
$$\frac{C\left(\overline{E}_j\right)}{C_0\left(\overline{E}_j\right)} = \frac{\int_{E_{j1}}^{E_{j2}} S(E)D(E)\exp\left[-\sum_i \mu_i(E)\rho_i\right]dE}{\int_{E_{j1}}^{E_{j2}} S(E)D(E)\,dE} \approx \exp\left[-\sum_i \mu_i\left(\overline{E}_j\right)\rho_i\right]. \tag{2}$$

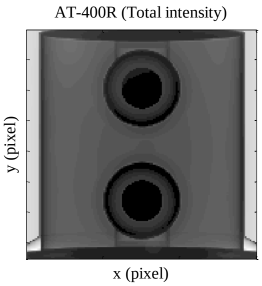
With some simplifying approximations, a linear regression model may suffice for roughly estimating densities. The results may be useful in themselves or as an initial guess for a fully nonlinear least squares approach. Given the material attenuation functions chosen (the “material set”), the vector of areal densities $\left(\rho_1, \rho_2, \ldots, \rho_n\right)$ for *n* materials are estimated using NNLS. This approximation enables relatively simple calculations for the estimation problem, and includes all contributions to the attenuation of a pencil beam of radiation, including attenuation due to photoelectric absorption and photons scattered out of the beam.

The multi-energy, single-pixel approach allows for a determination of the presence of materials of interest regardless of location or shape. In medical imaging [6] and some explosives detection applications, K-edges in the mass attenuation coefficients, which show up in the energy range from around 0.1 to around 125 keV [7], are highly dependent upon the material and so provide strong discriminatory power between imaged materials when they can be used. Our test items of interest include high-Z objects, so although we do not include downscattering, it is expected that the limited penetration of counts in lower energy ranges will render information in that part of the spectrum not useful for these analyses. The higher energy regions of the attenuation functions are less distinct, producing material confusion when many materials are included together in a fit. The lack of K-edge and low-energy information, as well as attenuation and down-scattering in thick regions of an object [5], present a fundamental limitation to the material discrimination technique, as only limited variations due to Compton scattering and pair production at higher energies are observable [8].

To illustrate this material confusion, the material discrimination methodology was previously applied to a real object under idealized conditions [5]. As a specific example relevant to material verification, the AT-400R Pu storage container furnished for use at the Mayak Fissile Material Storage Facility in a declared configuration is considered as an object for inspection. The AT-400R container holds 2-kg Pu spheres within a storage container primarily composed of Fe and Polyethylene (Figure 1). Throughout this work, this Pu storage object and elements of the AT-400R storage container are used to provide examples relevant to detection of potential material diversion. For the purpose of spanning likely materials and effective Z values, polyethylene (poly), iron (Fe), aluminum (Al), and plutonium (Pu) are used as the material set for estimation. A ray-tracing approach in the MatLab® code [9] used Beer's law for attenuation in each pixel to estimate the transmitted detected flux of a parallel-beam bremsstrahlung source incident on the configuration, with maximum energy of 450 keV, as measured by a detector with ideal energy resolution (i.e., the energy of incoming photons are always measured perfectly). MatLab results were validated against the Monte Carlo N-Particle radiation transport code (MCNP), and no significant differences were found between these two approaches [10].

Material density estimates were translated into estimated thicknesses in each pixel by the use of nominal densities for each material (Figure 1). In these figures, the grayscale represents the estimated density of each material. Although very little Al is present in this configuration, the use of this material discrimination analysis shows significant material confusion when a complex configuration of materials is present. Furthermore, the strong attenuation in the center of the Pu objects (coupled with the relatively low energy bremsstrahlung source) causes a number of energy bins with zero total counts in those pixels. As a consequence, no results are available for those regions with sufficiently high density. It is expected that a real system for interrogation of this sort of material configuration will require higher-energy gamma rays for interrogation, as considered later in this work.





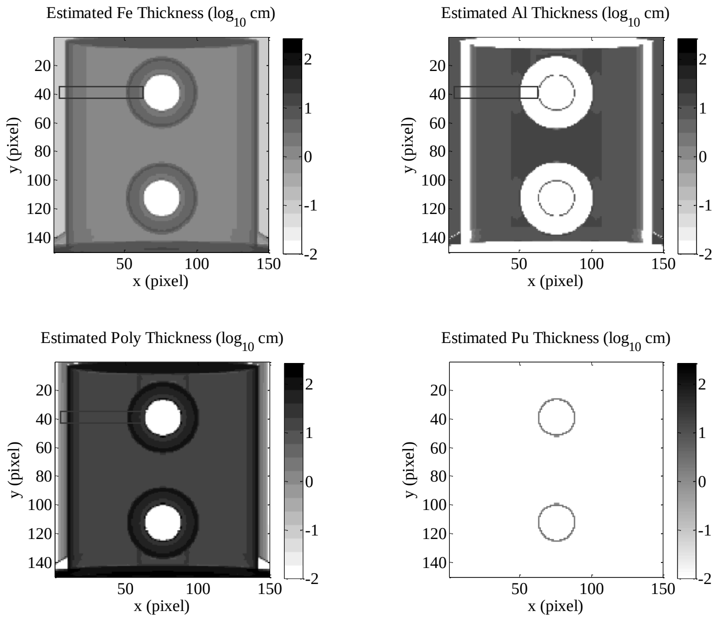


Figure 1. Top two plots: The AT-400R Pu storage container, containing two 2-kg Pu spheres, with total radiographic transmission image. Middle and bottom plots: Estimated material components for four materials (units are cm of estimated material, log scale).

Even for this configuration, (where the materials exactly match the material functions chosen for analysis) the use of four materials produces results with significant material confusion in cases without noise. In particular, a substantial amount of Al is estimated in regions where it is not actually contained in the AT-400R. In order to quantify the loss of discrimination capability as the number of materials increases, an investigation of the similarity of material attenuation functions is made.

OPTIMAL DISCRIMINATION - PRINCIPAL COMPONENT ANALYSIS

We attempt to quantify the information obtainable from a single pixel by the NNLS regression method by estimating the total number of different materials that could be differentiated by gamma radiography in an ideal case. Ideally, we could rigorously obtain this upper bound by simulating gamma radiography with a broad beam source, and considering only a single pixel at a time. To provide a best case estimate for this sort of analysis, we assume an idealized bremsstrahlung source with a maximum energy of 9 MeV, and a spectroscopic gamma imaging detector with energy spacing of 5 keV. Loss of photons due to attenuation and scattering are assumed, but no downscattering is accounted for. We consider mass attenuation coefficients for Z=3 to 100 across the measured energy range of 1 MeV to 9 MeV [7].

We address the collinearity of the mass attenuation coefficients by using a standard principal component analysis (PCA) approach. PCA indicates that 85.70%, 99.84%, and 99.99% of the variability in the mass attenuation coefficients is described by the first one, two, and three principle components, respectively. This provides a strong suggestion (though not a proof) that at most two materials can be discriminated using this estimation process. A primary conclusion from the PCA method is that no more than two components describe the variation in material attenuation where the noise from the data is above a few percent, in an otherwise ideal case. This is a compelling argument for an extremely limited material set possible without material confusion, as additional limitations (higher image noise, reduced energy range, or the effect of downscattering) would tend to reduce rather than enhance overall discrimination power.

Using the Principal Component functions instead of the individual material attenuation functions as bases could also produce a useful material discrimination method. However, PCA often produces components that do not have any particular physical relevance, while the NNLS fit utilizing two components has the familiar "material depth" interpretation in each pixel. We will therefore concentrate on an NNLS approach employing a two-material set for this investigation.

The best material set to use is expected to be driven by the context of the investigation – for example, if the declared configuration is just a single material (e.g., only Fe is declared to be present), then the illicit presence of Pu might be tested for using $\rho=\{\rho_{Pu}, \rho_{declared}\}$ (where $\rho_{declared}$ represents the declared material) as a set of unknown parameters to be fit by the NNLS approach. This would form an effective technique for confirming the absence of Pu in the object. However, a more complex or unknown object would be much more difficult to evaluate in this way. To address material discrimination in an arms control context, we frame the issue of material discrimination as the need to estimate the Pu present in an arbitrary object. This allows for a comparison of approaches pertinent to an investigation of material diversion.

OPTIMAL DISCRIMINATION – "BOX OF TOOLS" EXAMPLE

As a first example of this discrimination technique, we consider a diversion scenario in which some material is diverted by being placed into an otherwise declared and benign object. We consider a simple "box of tools" model, in which steel (modeled here as iron) tools are contained in an arrangement not ordinarily able to be investigated due to sensitivity issues. The potential for material diversion suggests inspection of this configuration be performed behind an information barrier in order to verify the declaration.

This scenario is modeled in the Matlab code as a simple "wedge" of iron (representing an unknown configuration) potentially containing one of the Pu objects from the AT-400R storage container shown in Figure 2. The resulting radiographic images were simulated using the MatLab code [4] and validated with MCNP [10]. The MatLab approach used Beer's law for attenuation in each pixel to estimate the transmitted detected flux of a parallel-beam bremsstrahlung source incident on the configuration, as measured by a detector with perfect energy resolution. The resulting configuration is analyzed with the NNLS approach using a material set of $\rho=\{\rho_{Pu}, \rho_i\}$ where i is either Fe, polyethylene (Poly), or Pb (lead) (See Figures 2 and 3). To simulate a detector with realistic but limited spatial resolution, we consider 50x50 pixel images (each pixel 0.2 cm in size), and an energy range between 1 MeV and 9 MeV, covered by linear energy bins 5 keV in width. For the purpose of this exercise, the detector efficiency is simulated to be perfect, and the flux of a parallel-beam bremsstrahlung source incident on the configuration is simulated with a maximum energy of 9 MeV.

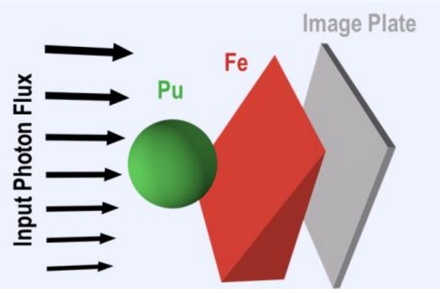


Figure 2. The "Box of Tools" physical model.

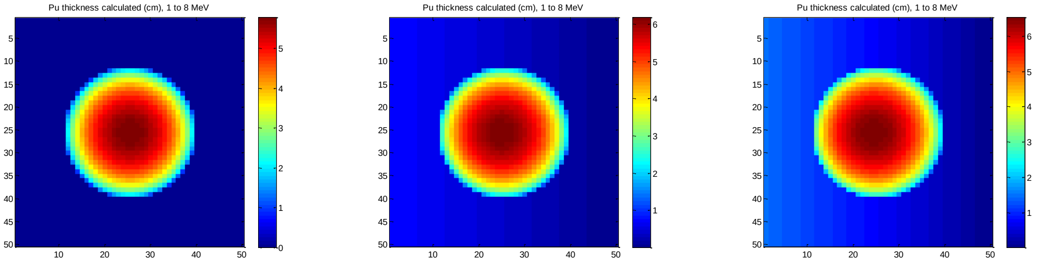


Figure 3. The "box of tools" model, Pu estimation for three different material sets, noiseless case (color scale represents estimated material thickness in cm). Left: {Pu, Fe}, Middle: {Pu, Poly}, Right: {Pu, Pb}.

Pu estimates for these three choices of second material were made by summing the estimated Pu in each pixel, and are as follows: For {Pu,Fe}, 2.02 kg of Pu are estimated, for {Pu, Poly} 2.79 kg of Pu are estimated, and for {Pu,Pb}, 3.47 kg of Pu are estimated. In the {Pu, Fe} case, Pu is estimated to be present only where it actually is, and in the correct quantity (a 2 kg sphere, accurate to around 1%), while no Pu is estimated in regions not actually containing any. In an otherwise ideal case, the use of an "incorrect" second material leads to material confusion, and an error in the estimate of Pu quantity.

To generalize this result, we performed the same analysis with a second range of materials, and in each case noted the total estimated Pu present in the configuration (Figure 4). This is compared with the real amount to determine the effectiveness of several choices of material for estimation. As expected, the correct amount of Pu is estimated when the Z of the second material matches the actual second material (Fe) contained in the image, while the amount deviated strongly with an incorrect second material. Use of a second material with a Z lower than iron causes some overestimate of Pu, as the method estimates the real Fe as being composed partly of Pu and partly of the light material. Likewise, using the higher Z materials resulted in the pixels containing Pu and Fe being evaluated as more like the second material, ultimately underestimating Pu. However, when the second material is sufficiently close to Pu, the collinearity between material attenuation functions is sufficient to produce significant material confusion, leading to an overestimate of Pu, as slight variations in lower-energy attenuation dominate the material estimate. With no foreknowledge of an appropriate second material to use (e.g., with an undeclared material configuration), estimates for Pu may be unreliable.

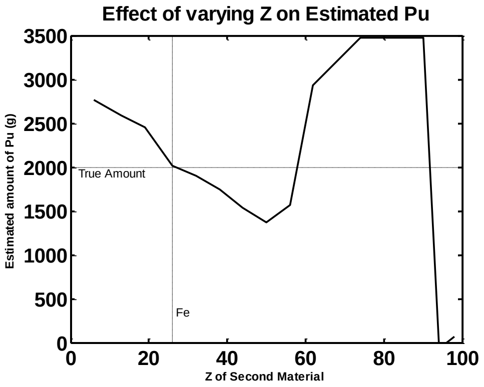


Figure 4. Estimated Pu vs. Z (second material set), actual mass of AT-400R object noted.

ADAPTIVE SINGLE-PIXEL TECHNIQUE

While a single material declaration would allow for a useful evaluation of Pu presence (and a confirmation that an inspected object is as declared), generally more complex situations may be expected from arms control scenarios. In these cases significant material confusion may be expected with the NNLS method as described thus far. An optimal second material (i.e., the choice of a second material in the fit) to use for analysis is difficult to identify when no single material represents the entire configuration. However, the choice of Pu and a single second material in each pixel may allow for accurate Pu estimation if the second material is representative of the remaining attenuation. A desirable result is therefore an estimate for the “effective” Z present in all the non-Pu materials present in a single pixel.

The fitting error (the sum of square deviations between the best NNLS fit and observed attenuated spectra) offers a potential figure of merit for making a choice of effective Z in a single pixel material estimation. As an example of this approach, the central pixel in the “box of tools” configuration (containing at most Pu and Fe) is analyzed using the material discrimination method. The fitting error (that is, the root mean squared error between the fit and observed spectra) is displayed as a function of the Z of the second material (Figure 5).

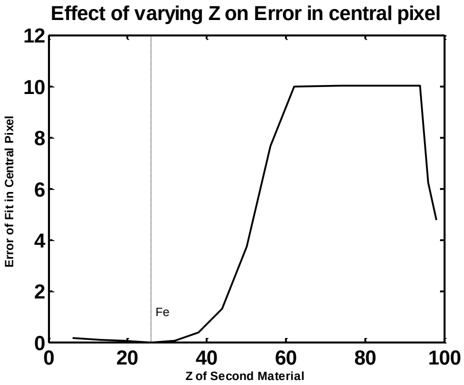


Figure 5. Fitting error (root mean squared error) vs. Z of second material, for the central pixel in the "box of tools" model.

A minimum error is found at the appropriate second material (Fe), suggesting the fitting error as a metric to determine a best second material to use for the fit in each pixel, even without knowing the actual second material *a priori*. Extending this result, we form an adaptive technique by fixing one material for estimation to be Pu, iterating over the second material and optimizing the goodness of fit metric to choose a second material in each pixel. This approach is expected to produce good results for Pu presence even in the presence of other more complex material configurations or scenarios containing noise.

The issues of noise and complex attenuators will be approached by considering the effect on the estimated Pu present in each pixel of a noisy radiograph. A 3-slab model is implemented to represent a more complex structure that may not be declared before inspection. In this model the goal of estimating Pu presence to detect material diversion is the same (Figure 6).

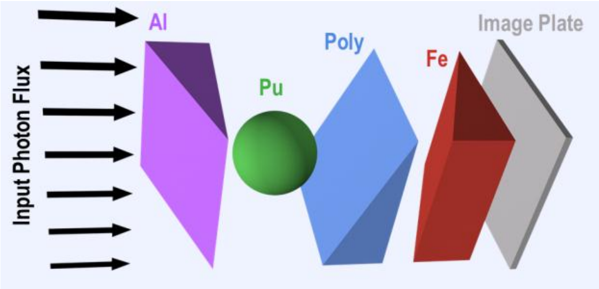


Figure 6. The 3-slab model with Pu object from AT-400R storage container (2 kg sphere).

To test the adaptive fitting method, the "box of tools" model and the 3-slab model are both investigated with added noise. The 3-slab Model results are shown in Figure 7. To characterize the noise, we scale all the pixels such that the highest value in the highest background pixel (for a single 5 keV energy bin) is 100,000 total counts, representing the use of a realistic emission/detection system. We then add Poisson distributed noise to each energy/spatial bin independently. An approximately correct estimate for Pu presence is obtained, and errors in the estimated mass of Pu are relatively small (~10%).

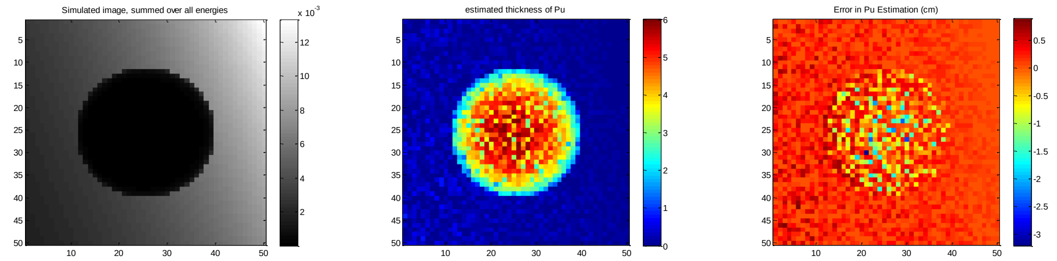


Figure 7. Results of the method for 3-slab model, noisy case. Left: total attenuation image (grayscale), units are counts per pixel. Middle: Estimated Pu, units are cm. Right: Pu error, units are cm. The estimated total Pu amount for this configuration is 2.16 kg, close to the correct answer.

CONCLUSIONS

The use of the adaptive single-pixel technique (in which two materials are used as a material set) exploits the information present in a single image pixel to estimate the presence of Pu. This technique is capable of several useful tasks pertinent to an arms control regime. First, it is capable of verifying a declared material configuration, even for overlapping or complex attenuating bodies. Also, this technique can confirm the absence of SNM with a declared or undeclared configuration. Finally, merging or comparing this Pu mass estimate with a passive radiation-measurement-based mass estimate behind an information barrier would allow for “yes/no” verification without the disclosure of sensitive information. This approach does not strictly depend on spatial imaging resolution, and is expected to work with dual-energy or multi-endpoint radiography systems, provided that at least several energy bins/endpoints are available.

The PCA result suggests that the adaptive single-pixel approach incorporates all useful information, as most or all of the available discriminating power is contained within the first two free parameters of the fit. The optimality of this method is not rigorously shown, but it is still expected that further work in regard to an optimal material set will not radically improve the results, given this outcome. Other approaches are possible, such as letting both material bases vary, but these are not expected to further enhance results if SNM is the desired target of measurement. Therefore, the methods by which this approach may be improved are expected to include some aspect of constraint on the fitting method used. This sort of “constrained inversion” can be approached in a few ways. Spatial constraints on object size or shape, as well as examination of the edge transition characteristics of objects may be used to enhance results. The performance of material discrimination under limitations of the field of view near edges in an image may also be improved by utilizing an iterative technique for estimating similar regions within an image. These will be the topic of further research.

The relatively high energies assumed by this work (1 MeV to 9 MeV) are intended to represent requirements for the gamma ray generation and detection system necessary to achieve positive material discrimination with the dense and high-Z materials expected by an arms control inspection regime. Beyond this requirement, materials sufficiently close in Z to Pu or heavy attenuation may still make Pu discrimination more difficult. The results from material discrimination approaches can be directly interrogated for the presence of nuclear materials (provided sufficient counting statistics in individual pixels). This allows for a simple “yes/no” metric for SNM detection to be developed. However, only Pu is considered in this work, and material confusion between U and Pu is a potential consideration. With further study to address the challenges noted above, these methods may prove useful for the verification of objects in both warhead counting and dismantlement verification settings.

ACKNOWLEDGEMENTS

This research is supported by the U.S. Department of Energy National Nuclear Security Administration Office of Nonproliferation and Verification Research and Development. The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy.

REFERENCES

[1] Kouzes RT and JL Fuller, Authentication of Monitoring Systems for Non-Proliferation and Arms Control, PNNL-SA-35296. Prepared for the IAEA Vienna Symposium (October 2001).
[2] Whiteson R and DW MacArthur, “Information Barriers in the Trilateral Initiative: Conceptual description,” Los Alamos National Laboratory, LAUR-98-2137 (Unlimited distribution, available from the United States Department of Energy, Office of Scientific and Technical Information website, www.osti.gov), 1998.
[3] Wolford, Jr., JK and GK White, “Progress in gamma ray measurement information barriers for nuclear material transparency monitoring,” Proc. INMM Annual Meeting, New Orleans, LA, 2000.
[4] Robinson, SM, KD Jarman, WK Pitts, A Seifert, AC Misner, ML Woodring, MJ Myjak, Imaging for dismantlement verification: Information management and analysis algorithms, NIM-A, Vol. 662, 1, pp. 81-89 (2011).
[5] Robinson S, KD Jarman, A Seifert, B McDonald, AC Misner, T White, EA Miller, WK Pitts, Image-Based Verification Algorithms for Arms Control, Proceedings of the 2011 INMM Conference, Baltimore, MD.
[6] Alvarez, RE and A Macovski, Energy selective reconstructions in x-ray computed tomography, Phys. Med. Biol. 21(5):733-744, 1976.
[7] Chadwick MB, et al., ENDF/B-VII.0: Next Generation Evaluated Nuclear Data Library for Nuclear Science and Technology, Nuclear Data Sheets, Volume 107, Issue 12, December 2006, Pages 2931-3060, ISSN 0090-3752, 10.1016/j.nds.2006.11.001.
[8] Wang Xue-Wu et al, Material Discrimination by High-Energy X-Ray Dual-Energy Imaging, High Energy Physics and Nuclear Physics 31(11) (Nov. 2007).
[9] MATLAB, version 7.8.0, Natick, MA, The Mathworks, Inc. (2009).
[10] MCNP X-5 Monte Carlo Team, MCNP—A General Purpose Monte Carlo N-Particle Transport Code, Version 5, LA UR 03 1987, Los Alamos National Laboratory, Apr. 2003. The MCNP5 code can be obtained from the Radiation Safety Information Computational Center (RSICC), P. O. Box 2008. Oak Ridge, TN, 37831-6362.

Pacific Northwest
NATIONAL LABORATORY

U.S. DEPARTMENT OF
ENERGY

Proudly Operated by Battelle *Since 1965*

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99352
1-888-375-PNNL (7665)
www.pnl.gov