R4-B.4: Integrated Reconstruction/Recognition

Abstract — In CT-based security screening, a challenging problem is to correctly identify and label objects in a scene from X-ray projection data. The presence of artifacts in the reconstructed image confounds traditional segmentation and labeling methods. In this work, we focus on integrating machine learning, physical modeling and Bayesian inference in a framework for direct material identification and labeling. Such an approach can mitigate image artifacts, reduce the number of corner cases and minimize false alarms.

<table>
<thead>
<tr>
<th>Name</th>
<th>Title</th>
<th>Institution</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clem Karl</td>
<td>Co-PI</td>
<td>BU</td>
<td><a href="mailto:wckarl@bu.edu">wckarl@bu.edu</a></td>
</tr>
<tr>
<td>David Castañón</td>
<td>Co-PI</td>
<td>BU</td>
<td><a href="mailto:dac@bu.edu">dac@bu.edu</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Degree Pursued</th>
<th>Institution</th>
<th>Month/Year of Graduation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmet Tuysuzoglu</td>
<td>PhD</td>
<td>BU</td>
<td>6/2014</td>
</tr>
<tr>
<td>Limor Martin</td>
<td>PhD</td>
<td>BU</td>
<td>6/2014</td>
</tr>
</tbody>
</table>

II. PROJECT OVERVIEW AND SIGNIFICANCE

In CT-based security screening, a challenging problem is to correctly identify and label objects in a scene from X-ray projection data. Conventionally, reconstruction and labelling are performed as two decoupled steps. Image artifacts induced from metal and other clutter cause variation in apparent material density as well as streaking that can break up homogeneous objects, making their correct identification and assessment challenging. In this project methods are developed that incorporate the tools of machine learning, physical modeling, and Bayesian inference into a unified framework for direct material identification and labeling. This approach can mitigate image artifacts and reduce the number of corner cases, which can in turn reduce false alarms and the need for On-Screen Alarm Resolution Protocol (OSARP) and manual inspection.

III. RESEARCH ACTIVITY

A. State-of-the-art and technical approach

Traditional approaches perform decoupled steps in obtaining material labels. The first step performs image processing, such as filtering and denoising, in an attempt to reduce noise and image artifacts. This processed image is then labeled or segmented in an attempt to find objects and determine their materials. In the current period, we have developed a new direct, dictionary-based method for material labeling from Multi-Energy CT (MECT) measurements. This new method takes into account the system model information together with a dictionary-based model of the measured sinograms. The linear attenuation coefficients (LACs) of the materials of interest are used as dictionary elements. The typically decoupled approach and the new proposed joint approach are illustrated in Figure 1 on the next page.
The observed normalized log-sinogram data in MECT sensing follows the non-linear Beer-Lambert law of X-ray imaging:

\[ I_s(x) = -\ln \left( \frac{\int w_s(E)e^{-\int \mu(x,E)dx} dE}{\int w_s(E) dE} \right) \]

where \( I_s(E) \) is the measurement along ray-path for spectral weighting \( s \), \( w_s(E) \) is the spectral weighting function \( S \) used in the measurement, and \( \mu(x,E) \) is the LAC of the material at spatial location \( X \) and energy \( E \), which identifies the material present at that location. Examples of LAC curves and spectral weighting functions are shown in Figure 2.

The characteristics of the material at spatial location \( X \) are captured through the energy dependent function \( \mu(x,E) \). Typically, this function is approximated as a linear combination of only two basis functions, such as the photoelectric and Compton basis functions. In the security application, however, we are less concerned with accurate representation of the function than with accurate identification of the material at that spatial location. In this work we focus on this identification aspect.

To that end, one aspect of this project has developed a unified approach for the estimation of the material label image from MECT measurements. The approach takes into account the system model information and the way the materials are observed via the Beer-Lambert law. A dictionary-based model of the materials present in the scene is used, with the LACs of materials of interest used as the dictionary elements. Estimated dictionary coefficients are used to identify material labels that match the observed multi-energy observations. A Markov random field (MRF) type model, which captures our belief that the label field should display spatial
coherence, is used to suppress artifacts. This spatial coherence reflects the behavior of objects in the scene and serves to further our goal of preventing splitting of objects.

Specifically, we create a dictionary of LACs of the few materials of interest: \( D = [\mu_1, \mu_2, \ldots, \mu_m] \) where \( \mu_k \) is a vector representing the LAC of material label \( k \) as a function of (discretized) energy level. This dictionary is an overcomplete basis for representing the material LAC in the scene. Let \( c_x \) be the LAC composition vector for the pixel at location \( x \) in the scene. The LAC of the material at pixel location \( x \) is then given by the linear combination of dictionary elements specified by the product \( Dc_x \). Since we believe there is only a single material at pixel \( x \) we assume sparsity of the vectors \( c_x \) such that only a single entry of \( c_x \) will be 1 and the rest 0. We find the material labels by solving a unified optimization problem with this dictionary-based model in the previous Beer-Lambert X-ray model. The overall formulation is:

\[
\min_{\{c_1, c_2, \ldots, c_N\}} \left\{ \sum_{x, y} \left[ - \ln \left( \frac{\int w_x(E)e^{-\int Dc_x(E)dx} \, dE}{\int w_x(E) \, dE} \right) - I_x(\ell_x) \right]^2 + \lambda \psi_{\text{MRF}}(c_1, c_2, \ldots, c_N) \right\}
\]

The first summation in the formula above is the data fidelity term. It is defined as the squared error between the measured sinogram and the dictionary-based model of it based on the physical Beer-Lambert law. The second term is a Potts-type MRF prior term. This term penalizes any differences in dictionary coefficient vectors for pixels in a small neighborhood. Different coefficient vectors correspond to different labels and the desire is to increase the coherency of the label field. The minimization is performed subject to the constraints that only a single entry of any coefficient vector \( c_x \) at spatial location \( x \) is equal to 1. These constraints reflect the belief that there is only one material at each spatial location, and this material is associated to a particular column of the material dictionary \( D \). A solution is obtained by an iterative process that solves for each \( c_x \) in sequence while keeping the others fixed. All the material possibilities are tested and the one that provides the lowest value of the cost function is selected. Note that this new approach is performing direct, integrated segmentation and labeling in contrast to conventional ad-hoc multi-step processes. We call this method the “dictionary labeling” method.

While the above dictionary-based direct labeling method exploits the complete physical sensing model, as well as knowledge of material X-ray behavior; it requires detailed knowledge of system parameters, such as sensing geometry, spectral shapes, material LACs, etc., and can be computationally demanding. As a result, another thrust of our research in this project has focused on the development of efficient data driven approaches using machine learning and graph-cut methods. In this approach, we start from conventionally formed effective attenuation images for each multi-energy experiment and then directly learn the conditional appearance model for each material of interest from a set of training data. Given this learned appearance model, a discrete label-based optimization is performed. As a result, explicit physical models of the tomographic system and detailed information on material LACs are not needed.

To develop the method, let \( \mu^H \) and \( \mu^L \) denote the conventionally created effective attenuation images for a high energy and low energy X-ray CT scan, respectively. At any spatial location \( x \), the appearance of a material in these two scans is modeled as a probability density

\[
P \left( \begin{bmatrix} u_{H,x} \\ u_{L,x} \end{bmatrix} \left| l_x \right. \right)
\]

conditioned on the label \( l_x \) the material at that location. This density captures the potential variability in ma-
terial appearance and is directly learned using machine learning kernel density estimation techniques from training data. In this way, the method directly captures appearance uncertainty and does not require detailed physical material knowledge.

Unfortunately, conventionally created effective attenuation images may contain streaks which cause splitting of objects and corruption of attenuation values. The probabilistic appearance model can capture and model a portion of this variability, providing some mitigation, but direct incorporation of additional prior information can improve the final result. For example, a large source of artifacts in luggage scans is the presence of metal objects. Further, such metal artifacts are often stronger the closer the pixels are to the metal object. To reflect this insight, observed data near metal objects is assumed less reliable and, therefore, down weighted. In addition, to prevent object splitting and reduce attenuation variability in homogeneous regions, an object boundary field is used.

Material labels at each pixel location and object segmentation is obtained jointly as the solution of the following optimization problem:

\[
\min_{(l_1, l_2, \ldots, l_N)} \left\{ \sum_x v_x \left( -\ln p \left( \begin{bmatrix} \bar{\mu}_j^L \\ \bar{\mu}_j^H \end{bmatrix} \mid l_x \right) \right) + \lambda g_{\text{MRF}} (l_1, l_2, \ldots, l_N, s) \right\}
\]

subject to \( l_x \in \{1, 2, \ldots, M\} \ \forall x \)

In this framework \( \bar{\pi}^L \) and \( \bar{\pi}^H \) are the conventionally formed effective attenuation images obtained from measurements with two different (high and low) spectral weightings, \( l_x \) is the material label in pixel \( x \),

\[
P \left( \begin{bmatrix} u_{1,x} \\ u_{2,x} \end{bmatrix} \mid l_x \right)
\]

is the appearance model for material label \( l_x \) at pixel \( x \), \( v \) are data weights which down-weight data points in the vicinity metal, \( \lambda \) is a non-negative regularization parameter and \( g_{\text{MRF}} \) is an MRF smoothing term, which is based on an estimate of the image boundary field \( s \). This MRF model captures local coherence to material labels and takes into account an estimate of object boundaries to further homogeneity within an object. Figure 3, on the next page, shows the main components of the method for an example slice from the Imatron scans database. The resulting optimization problem is a non-convex, discrete label problem, which are in general challenging to solve. To accomplish this optimization, an efficient graph-cut based method has been developed. Such graph-cut methods have been popular in the computer vision land discrete optimization literature, but have not been used in this domain. These methods map the original optimization problem to an equivalent graph flow problem and a minimal cut of this graph provides the optimal solution. These methods have shown great success in producing efficient near optimal solutions of very challenging discrete problems, and are well suited to our application. We call this overall method “learning-based graph-cut labeling.”
B. Major contributions

In this period, we have demonstrated preliminary results for creating direct labeling of material from dual-energy CT data on a slice-by-slice basis. We implemented the methods described above and tested them on real dual-energy data from the Imatron scans database. These are dual-energy scans of different objects in bags obtained with 95 kVp and 130 kVp source spectra.

For the dictionary labeling method, we constructed a dictionary with the LACs of graphite, magnesium and silicon, providing four possible material labels: graphite, magnesium, silicon and background (air). We applied the method to a slice that had graphite, magnesium and silicon rods, and which were placed in foam inside a plastic case. The results are shown in Figure 4. The labeling results seem to be accurate and we see that the MRF prior helps maintain object homogeneity.

Figure 3: Illustration of the main components of the learning-based graph-cut labeling method. The top left figure shows the learned material appearance models. The top right figure shows an example of the data weighting scheme. Data points close to metal are given lower weights. The bottom figure shows the boundary-field used in the smoothing term.

B. Major contributions

In this period, we have demonstrated preliminary results for creating direct labeling of material from dual-energy CT data on a slice-by-slice basis. We implemented the methods described above and tested them on real dual-energy data from the Imatron scans database. These are dual-energy scans of different objects in bags obtained with 95 kVp and 130 kVp source spectra.

For the dictionary labeling method, we constructed a dictionary with the LACs of graphite, magnesium and silicon, providing four possible material labels: graphite, magnesium, silicon and background (air). We applied the method to a slice that had graphite, magnesium and silicon rods, and which were placed in foam inside a plastic case. The results are shown in Figure 4. The labeling results seem to be accurate and we see that the MRF prior helps maintain object homogeneity.

Figure 4: Material label images using the dictionary labeling method for a slice with graphite, magnesium and silicon rods placed in foam inside a plastic container. The magnesium rod is on the left, the graphite rod is in the middle and silicon is on the right. The dictionary was composed of graphite, magnesium and silicon LACs. Graphite is labeled in light blue, magnesium is labeled in yellow, silicon is labeled in red and background is labeled in dark blue. Top: initialization. Bottom left: result with no MRF prior. Bottom right: result with an MRF prior. The method achieves accurate material labeling and using the MRF prior helps to get more homogeneous results.
For the learning-based graph-cut labeling method, training data was obtained for the following material labels: water, doped water, rubber and metal. Figure 5 shows the original high-energy input image and the corresponding directly labeled output on the right. Despite the presence of metal (orange) and its associated streak artifacts, the objects of interest, in this case rubber sheets objects (green) and water (blue), are correctly labeled without splitting or breaks. Overall, image quality is improved.

Overall, these methods provide physics-based principled approaches for direct material labeling that contrasts current ad hoc, two-step decoupled approaches. The potential of such optimal methods are reduced artifacts, lower false alarms and improved material labeling.

C. Future plans

Future plans include further development of the dictionary-based approach to include more materials and application to more complex scenes. For the learning-based graph-cut labeling method, further development of metal artifact reduction techniques is planned. In addition, development of more efficient solution methods of both formulations will be done. We anticipate that an ADMM or augmented Lagrangian approach can be used, incorporating both a label field, as well as an attenuation field. These will lead to flexibility in using both projection and image fields in the same formulation and provide faster convergence rates. We also plan to include more accurate physical modeling of the materials through the inclusion of more material classes and possibly more complicated feature spaces. In addition, we plan to extend the methods from 2D to 3D, and to examine the benefit of using a fully 3D formulation to allow more integrated spatial information to be included.

IV. EDUCATION & WORKFORCE DEVELOPMENT ACTIVITY

- W. C. Karl gave an invited seminar at the CIMI workshop “Optimization and Statistics in Image Processing” in Toulouse.
- L. Martin gave a presentation at the ALERT Task Order 3 Symposium.

V. RELEVANCE AND TRANSITION

A. Relevance of your research to the DHS enterprise

This project is of relevance to the DHS enterprise because it is developing a method to mitigate image artifacts in CT scanning of baggage. This approach can reduce the number of corner cases and false alarms, which in turn can reduce the need for OSARP and manual inspection.

B. Anticipated end-user technology transfer

Vendors could incorporate the methods being developed in this project into their ATR chains. Several vendors at the symposium for the Task Order 3 “Research and Development of Reconstruction Advances in CT-based Object Detection Systems” effort supported by DHS Task Order Number HSHQDC-10-J00396, (e.g. L3) commented that they had not thought that results this good could be obtained directly from dual-energy data. In addition, TSL/S&T personnel (C. Love, R. Krauss, R. Klueg) expressed interest in collaborating to see how these methods would perform on laboratory material samples. This ongoing collaboration has led to
TSL sharing some dual-energy data with our laboratory for experimental use.

VI. LEVERAGING OF RESOURCES
DHS BAA 13-05 funding was pursued and we are currently in negotiation for a contract.

VII. PROJECT DOCUMENTATION AND DELIVERABLES

A. Peer reviewed journal articles
Pending-


B. Peer reviewed conference proceedings


C. Student theses or dissertations produced from this project


D. Software developed

1. Datasets
   a. TO3 and TO4 Data resources.

E. Requests for assistance/advice

1. From DHS

VIII. REFERENCES


