https://doi.org/10.2352/ISSN.2470-1173.2017.17.COIMG-420


# Joint Segmentation and Material Recognition in Dual-Energy CT Images

***Parisa Babaheidarian and David Castañón; Boston University; Boston, MA/U.S.A.***

## Abstract

*The problem of identifying materials in dual-energy CT images arises in many applications in medicine and security. In this paper, we introduce a new algorithm for joint segmentation and classification of material regions. In our algorithm, we learn an appearance model for patches of pixels that captures the correlation in observed values among neighboring pixels/voxels. We pose the joint segmentation/classification problem as a discrete optimization problem using a Markov random field model for correlation of class labels among neighboring patches, and solve the problem using graph cut techniques. We evaluate the performance of the proposed method using both simulated phantoms and data collected from a medical scanner. We show that our algorithm outperforms the alternative approaches in which the appearance model is based on pixel values instead of patches.*

## Introduction

Automatic material recognition in 2D and 3D CT images has a wide range of applications both for medical and security purposes. In the medical domain, an automatic recognition algorithm helps the radiologist in diagnosis of tumors and pathology in tissues, particularly when the image contrast is poor. In security, automated recognition algorithms assist operators in detection of hazardous materials and reduces the need for human inspections.

Existing CT image recognition methods can be categorized into two general approaches. In the first approach, reconstructed images are segmented into objects. Features extracted from the segmented objects are then used to classify the material in each object. The advantage of this approach is that the features extracted from the segment information includes information from the global region, such as value histograms, spatial extent and other similar features. This approach is commonly used in detection of regions of interest in security, as detailed in the results [1–4].

The second approach performs the recognition task in two steps: image reconstruction and joint segmentation/classification as in Figure 1. In this approach, the goal is to segment regions with homogeneous classification. The idea behind this approach is that comparing the reconstructed values with the values expected for the different classes of materials will lead to improved segmentation and recognition. However, this approach has difficulty capturing regional information such as texture, mass or shape. Examples of this approach are found in [5, 6].

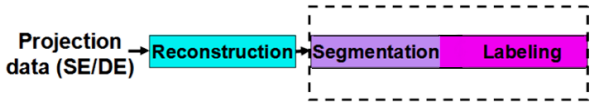


Figure 1: General steps in a joint segmentation/classification algorithm for CT images; SE refers to single energy and DE refers to dual-energy measurements.

In [5], Martin et al., proposed a joint segmentation/classification method for dual-energy CT images. In their model, they learned the appearance model using supervised training images. The training images are collection of filtered back projection (FBP) images reconstructed using two different source spectra. For simplicity, we consider 2D images in our discussion. Each pixel in the training 2D images is labeled for material classes of interest and background. To learn the appearance model they assume conditional independence of the intensity vectors at different pixels, given the label values. Hence, the training data is used to learn the pixel appearance model for each label, using the labeled training images.

In this paper, we introduce a new appearance model based on patches. Instead of learning an appearance model for each pixel, for each labeled class, we learn a joint appearance model for the collection of pixels in a patch. The goal is to capture correlations in nearby pixel values in order to represent better the texture and regional information. We model the prior information on patch labels using a Markov random field (MRF). We then pose the joint segmentation/classification problem as a maximum a posteriori estimation problem for the discrete patch labels. We solve the resulting optimization problem using discrete optimization techniques based on graph cuts [7, 8]. We evaluate our algorithm on both simulated dual energy data phantoms and data collected using a medical scanner. Our results show that our patch-based segmentation/classification approach outperforms alternative approaches based on pixel-based likelihood models.

The rest of this paper is organized as follows: In the next section, we provide an overview of the previous pixel-based joint segmentation/classification algorithm in [5]. Subsequently, we describe the maximum a posteriori formulation for our joint segmentation/classification problem, and describe our proposed algorithm in detail. The next section discusses our experiment results where we compare the performance of our algorithm to alternative algorithms. We conclude the paper with a brief discussion of directions for future investigation.

This material is based upon work supported by the U.S. Department of Homeland Security, Science and Technology Directorate, Office of University Programs, under Grant Award 2013-ST-061-ED0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

## Background work

In this section, we review the joint segmentation and classification method of [5], which forms the background for our work. In [5], dual-energy labeled CT images are used as training data for learning an appearance model for the different classes. For each label class of interest and background, roughly, the same number of pixels were collected to be included in the training set. Using this training data, the appearance model that is learned is a probability density for the random values of the pair of intensity values at high and low energies, $\mathbf{x}_i$, associated with voxel $i$ given that its label is $\ell_i$. We denote this density as $p(\mathbf{x}_i|\ell_i)$. To model the joint distribution of pixel value pairs over the full image, one assumes a naive Bayes model where the values at different pixels are assumed conditionally independent given the labels of the pixels. That is,

$$p(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_T|\ell_1,\ell_2,\ldots,\ell_T)=\prod_{i=1}^{T}p(\mathbf{x}_i|\ell_i) \tag{1}$$

where $T$ is the total number of pixels in the 2D image. In (1), $\mathbf{x}_i$ denotes the attenuation vector associated with pixel $i$, $T$ is the total number of pixels in the 2D image, and $\ell_i$ is the label assigned to pixel $i$.

Instead of learning a parametric model $p(\mathbf{x}_i|\ell_i)$, the approach in [5] used a K-nearest neighbor likelihood estimation for estimating the conditional likelihood $p(\mathbf{x}_i|\ell_i)$ for each $\mathbf{x}_i$, with $K=19$:

$$p(\mathbf{x}_i|\ell)=C\frac{1}{K}\sum_{j\in\mathscr{N}(\mathbf{x}_i)}1_{\{\ell_j=\ell\}} \tag{2}$$

In (2), $\mathscr{N}(\mathbf{x}_i)$ defines the K-nearest neighbors in the training data to the observed data $\mathbf{x}_i$, based on Euclidean distance among attenuation vectors. $1_{\{\ell_j=\ell\}}$ is the indicator function. The constant $C$ is a normalizing factor for the density.

In addition to the appearance model, the prior information on the label field $\ell_i, i=1,\ldots,T$ is given by a Markov random field, in the form

$$P(\ell_1,\ldots,\ell_T)=C_1e^{-\lambda\sum_{i=1}^{T}\sum_{j\in\mathscr{N}(i)}1_{\{\ell_i\neq\ell_j\}}\phi(\mathbf{x}_i,\mathbf{s},i,j)}$$

where the function $\phi(\mathbf{x}_i,\mathbf{s},i,j)$ will be discussed later.

With this background, the problem of joint segmentation and classification was formulated as the minimization of the negative logarithm of the probability of the field of object labels given the observed attenuation vectors per pixel. After discarding constants with respect to pixel labels, the discrete optimization becomes

$$\min_{\substack{\{\ell_i\}_{i=1}^{T}\\ \ell_i\in\{1,\ldots,M\}}}\sum_{i=1}^{T}v_i\ln P(\mathbf{x}_i|\ell_i)+\lambda\sum_{i=1}^{T}\sum_{j\in\mathscr{N}(i)}1_{\{\ell_i\neq\ell_j\}}\phi(\mathbf{x}_i,\mathbf{s},i,j) \tag{3}$$

where the variables $v_i$ and the function is defined to alleviate the effects of metal artifacts. In a preprocessing step on the higher energy image, pixels corresponding to metal are identified. The weights $v_i$ increase to 1 as the distance to the nearest metal pixel increases, thereby placing less emphasis on matching the appearance distribution for pixels near metal. Also, function $\phi(\mathbf{x}_i,\mathbf{s},i,j)$ depends on an estimate of the boundary field among objects, $\mathbf{s}$, which is obtained based on the gradients of the high-energy attenuation image. This estimate assigns high penalty values when neighboring pixels $i,j$ have different labels when there is little likelihood of boundaries between them. It assigns low penalty values when the likelihood of boundaries between them is high. This requires a complicated estimate of the boundary field $\mathbf{s}$, obtained by solving an inverse problem.

The solution to (3) is obtained using a multi-class graph cut optimization algorithm, as in [8]. Graph cut techniques [7] are commonly used in computer vision to solve discrete optimization problems associated with segmentation and classification. Formulating these minimization problems as graph cut optimization problems requires special structure in the objective functions, including sub-modularity [7]. The original graph cut approach was developed for optimization problems with binary-valued decision variables, and was subsequently extended in [8], to problems with expanded discrete values, corresponding to multiple labels.

The graph cut algorithm maps the objective function of interest into a graph where nodes are categorized as either sites or terminal nodes. Site nodes often represent pixels in the image, and terminal nodes are possible labels. Edges are divided into two categories known as $n$-links (edges between two sites) and $t$-links (edges from a site node to a terminal node). The capacity assigned to an $n$-link captures the penalty for the Markov random field term for the log-likelihood of the two label classes they were assigned. The capacity of a $t$-link represents the penalty for assigning the associated label to the associated site. The general form of the objective function used for graph cut algorithms is represented as

$$E(L)=\sum_{p}f_p(L_p)+\sum_{(p,q)\in\mathscr{N}}V_{p,q}(L_p,L_q) \tag{4}$$

where $L=\{L_p|p\in\mathscr{P}\}$ is a labeling (a partitioning) on image $\mathscr{P}$. $f_p(L_p)$ is a data penalty function which determines the cost of $t$-links, $V_{p,q}$ is an interaction potential which determines the cost of $n$-links, and $\mathscr{N}$ is a set of all pairs of neighboring pixels.

In [8], The $\alpha$-expansion algorithm was developed to solve the multi-label minimization when the interaction potential is a metric on the space of labels. This algorithm finds provably good approximate solutions within a known factor of the global minima by iteratively running a binary min-cut/max-flow algorithm [7] on appropriate graphs. In case $V_{p,q}$ is a semi-metric, one can use the $\alpha$-$\beta$ swap algorithm [8]. This latter algorithm does not have any similar guaranteed optimality as opposed to the $\alpha$-expansion algorithm.

## Proposed algorithm

Our algorithm is based on learning an appearance model for patches instead of pixels. Patches are extracted from the training data, where each patch is characterized by a vector of labels associated with the pixels in the patch. We work with dual-energy CT images, generated from reconstructions using filtered back projection algorithms at high and low source spectra. Each pixel in a patch contains an estimate of effective linear attenuation coefficients at both high and low energies. Thus, for a patch of size $\sqrt{K}\times\sqrt{K}$, the feature vector associated with the patch is of length $2K$.

We restrict our model to patches where all of the pixels have the same label. This represents a loss of resolution in the possible

labels, but results in a major reduction in computation requirements and complexity. Using training data from patches with this property, we estimate the marginal conditional probability density for the patch of linear attenuation coefficients given the patch label either using a Gaussian model, as

$$p(\mu_1^H,\mu_1^L,\ldots,\mu_K^H,\mu_K^L|\ell) \sim \mathcal{N}(\vec{\mu}_\ell,\Sigma_\ell), \tag{5}$$

or using a kernel density estimator with Gaussian kernels of the form

$$P(\mathbf{x}|\ell) \sim \frac{1}{N_\ell\, h}\sum_{i=1}^{N_\ell}\mathcal{N}(\mathbf{x}-\vec{\mu}_i,h\mathbf{I}) \tag{6}$$

where vector $\mathbf{x} \triangleq (\mu_1^H,\mu_1^L,\ldots,\mu_K^H,\mu_K^L)^T$ and $N_\ell$ is the number of training patches with label $\ell$, and $h$ is a kernel's bandwidth selected using cross-validation techniques.

We subsequently divide the image into $N^2/K$ non-overlapping patches. We assume conditional independence of the patch measured features given the patch labels, so that

$$p(\mathbf{x}_1,\ldots,\mathbf{x}_{N^2/K}|\ell_1,\ldots,\ell_{N^2/K}) = \prod_{n=1}^{N^2/K} p(\mathbf{x}_n|\ell_n) \tag{7}$$

In the second step, we cast the joint segmentation and classification problem as a discrete minimization whose solution is a set of labels for each non-overlapping patch in the test image of size $N\times N$, as

$$\min_{\{\ell_n,n=1,\ldots N^2/K\}} \sum_{n=1}^{N^2/K} -\log\Big(p(\mu_{1,n}^H,\mu_{1,n}^L,\ldots,\mu_{K,n}^H,\mu_{K,n}^L|\ell_n)\Big)$$

$$+\lambda\sum_{n=1}^{N^2/K}\sum_{j\in\mathcal{N}(n)}\phi_{n,j}(\vec{\ell}_n,\vec{\ell}_j) \tag{8}$$

where $\lambda$ is a smoothing parameter in (8), and the second term will be defined below to represent a Markov random field as the prior information for the field of patch labels.

In our model, we consider 4-way neighborhood for each patch including left, right, top, and bottom adjacent patches. We consider a Markov random field (MRF) over patches using Potts model as

$$\sum_{j\in\mathcal{N}(n)}\phi_{n,j}(\ell_n,\ell_j) \triangleq \sum_{j\in\mathcal{N}(n)} 1_{\{\ell_j\neq\ell_n\}} \tag{9}$$

Expression (9) penalizes mismatch between two neighboring patches when the associated labels are different. Figure 2 demonstrates the patch structure used for patches of size $2\times 2$.

Note that there is a trade-off between patch size and resolution used in segmentation/classification. Bigger patch sizes lead to better representation of regional information and faster computation, but also lead to loss of resolution. In our experimental results, we observed a good performance for patches of size $2\times 2$, i.e., $K=4$, which is the size we will use in the rest of this paper.

Note that the Potts model (9) for the Markov random field results in an energy function that satisfies the metric property. In addition, for labels $\beta,\gamma,\alpha$ and neighbor pixels $n_1,n_2$, we have

$$\phi_{n_1,n_2}(\beta,\gamma)+\phi_{n_1,n_2}(\alpha,\alpha) \le \phi_{n_1,n_2}(\beta,\alpha)+\phi_{n_1,n_2}(\alpha,\gamma), \tag{10}$$

which implies that the cost function in (8) is sub-modular, and can be minimized effectively using the $\alpha$-expansion algorithm [8].

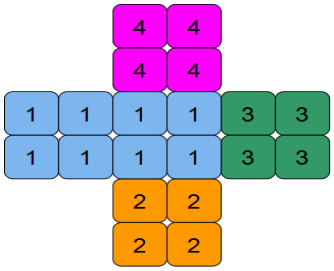


Figure 2: An example of a neighborhood for a given patch with label tuple $(1,1,1,1)$.

In the resulting graph cut algorithm, each site node in the graph represents a non-overlapping patch in the test image, the total number of site nodes in the graph is $\frac{N^2}{K}$ for a test image of size $N\times N$ pixels with patches of size $\sqrt{K}\times\sqrt{K}$. In addition, the number of terminal nodes corresponds to the number of labels $M$, similar to the pixel-based model described in the previous section.

## Experiments

To illustrate the performance of our algorithm, we conducted experiments using both simulated data and data collected from a medical scanner. As a measure of performance, we computed the percentage of pixels that have a different estimated label from the ground truth image label, defined as

$$N_e = \sum_{k=1}^{N^2} 1_{\ell_k\neq\hat{\ell}_k}$$

where $\ell_k$ is the true label of pixel $k$ and $\hat{\ell}_k$ is the label assigned by our algorithm.

In addition, for each material class of interest $m$, we compute measures of precision and recall, defined as follows: Let $\mathrm{TP}^m$ denote the number of true positive pixels of class $m$, defined as

$$\mathrm{TP}^m = \sum_{k=1}^{N^2} 1_{\ell_k=m}1_{\hat{\ell}_k=m}$$

Let $\mathrm{FP}^m$ denote the number of false positive pixels of class $m$, defined as

$$\mathrm{FP}^m = \sum_{k=1}^{N^2} 1_{\ell_k\neq m}1_{\hat{\ell}_k=m}$$

Let $\mathrm{FN}^m$ denote the number of false negative pixels of class $m$, defined as

$$\mathrm{FN}^m = \sum_{k=1}^{N^2} 1_{\ell_k=m}1_{\hat{\ell}_k\neq m}$$

Then, the precision $\mathrm{Pr}^m$ and recall $\mathrm{Re}^m$ for class $m$ are defined as

$$\mathrm{Pr}^m \triangleq \frac{\mathrm{TP}^m}{\mathrm{TP}^m+\mathrm{FP}^m};\quad \mathrm{Re}^m \triangleq \frac{\mathrm{TP}^m}{\mathrm{TP}^m+\mathrm{FN}^m} \tag{11}$$

In the first experiment, we created four variations of the FORBILD head phantom with metal implants using the tools provided

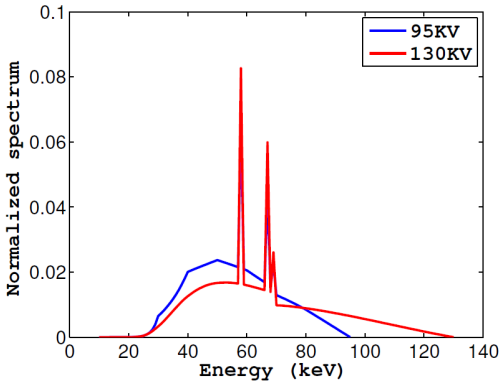


Figure 3: System spectra used when simulating dual-energy measurements [9].

in [10]. To generate simulated FBP images at lower and higher energies, we approximated the Imatron C300 source spectra operating at 95 Kvp and 130 Kvp voltages. The normalized spectral weighting functions for the Imatron source is displayed in Figure 3. We sampled the spectra at 40 points between 20 Kev to 100 Kev energies to obtain the energy weights associated with the two source spectra, $W_m(.), m = 1,2$. The materials in the created phantoms include water, air, bone, and iron. The linear attenuation coefficients for each material at different energy levels were derived from the NIST reference in [11]. We considered $I_0 = 3 \times 10^4$ photons per path as the initial flux. The received photon counts were simulated as a Poisson process as in the following:

$$c_i^m = \text{Poi}\left(\sum_{j=1}^{40} I_0 W_m(j) e^{-\int_{L_i} \mu(\vec{r},j) d\ell}\right) \quad m = 1,2 \tag{12}$$

In this equation, $c_i^m$ denotes the photon counts associated to path $L_i$ and source spectrum $m$, which is a Poisson random variable. Also, $\mu(\vec{r}, j)$ is the linear attenuation coefficient at energy level $j$ and is a function of spatial location for a fixed $j$. Lastly, $\int_{L_i} \mu(\vec{r}, j) d\ell$ denotes the line integral along path $L_i$ and at energy level $j$.

The sinogram value for each path $L_i$ is given by

$$y_i^m = -\log\left(\frac{c_i^m}{I_0}\right) \tag{13}$$

To form the image, we generated 200 views evenly distributed from 0 to 180 degrees, with 100 detectors per view. After computing the simulated sinogram vector for both source spectra, we reconstructed the images using the filtered backprojection (FBP) algorithm. Figure 4 displays the FBP images of the four phantoms reconstructed for the 130 Kvp source spectrum. Due to the presence of iron, there are visible metal artifacts in the reconstructed images. With these phantoms, we ran three experiments. The fourth phantom in Figure 4 was used as a cross-validation image to tune the smoothing parameter $\lambda$ in (8). From the first three phantoms, at each experiment, two phantoms randomly were selected as training images and the remaining phantom was used as the test image. In the three experiments, we used multi-variate Gaussian distributions to estimate the patch-based conditional likelihoods. Figure 5 displays the visual result of applying our algorithm to the tested phantom. We included a

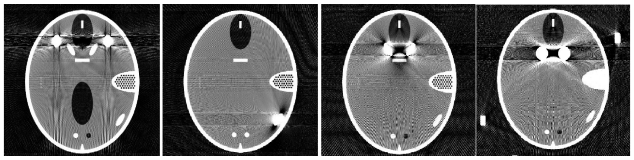

Figure 4: Sampled FBP reconstructed images at 130 Kvp source spectrum for the created phantoms.

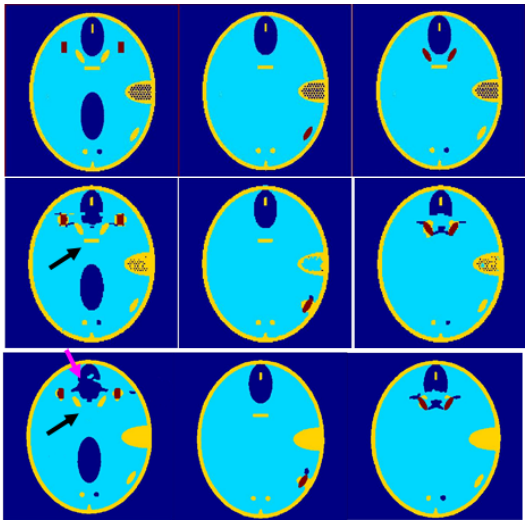

Figure 5: Labeled result for the tested phantoms. The top row is the ground truth with materials including iron=red, water=light blue, bone=yellow, and background (air)=dark blue. The second row displays our algorithm's result, and the bottom row displays the result of pixel-based KNN with MRF spatial smoothing.

pixel-based alternative's result in Figure 5, too. Specifically, we compared our algorithm against a pixel-based K-nearest neighbor (KNN) with spatial smoothing. The spatial smoothing in the latter approach was applied through a MRF over pixels with a similar penalty model as in (9). The pixel-based algorithm required heavier regularization to smooth out the result in the uniform regions. Therefore, it loses some information due to over-smoothing. Our algorithm's result also shows some error in the labeled images especially in the left ear of the phantoms. This is due in part to the loss of resolution in the patch-based model. Table 1 summarizes the numerical result for the tested phantoms in terms of the percentage of pixels with wrong labels. We compared the performance of our algorithm against three pixel-based alternatives. Alternative 1 is the pixel-based KNN with MRF. The second alternative uses pixel-based KNN with MRF and additional boundary information as in [5]. Finally, alternative 3 uses pixel-base Gaussian distribution to estimate the appearance model along with MRF over pixels. The table shows that the advantage of our algorithm over the alternatives becomes more significant when it is tested on harder cases, e.g., the third phantom in which iron pieces are placed closer. In other cases, our algorithm's performance is comparable with alternative 2 which utilizes the additional boundary field information. We further computed the average recall and precision of the materials in the ground truth for our algorithm, the pixel-based KNN+MRF, and the algorithm in [5]. The result is displayed in Figure 6. Our algorithm has higher precision and

| Algorithm | Our algorithm | alternative 1 | alternative 2 | alternative 3 |
|---|---|---|---|---|
| Phantom 1 | 2.42 | 2.62 | 2.38 | 3.19 |
| Phantom 2 | 2.32 | 2.35 | 2.23 | 2.4 |
| Phantom 3 | 3.38 | 4.15 | 4.05 | 5.21 |

Table 1: Performance of our algorithm versus the pixel-based alternatives on the tested phantoms with respect to the percentage of pixels labeled incorrectly compared to their ground truth.

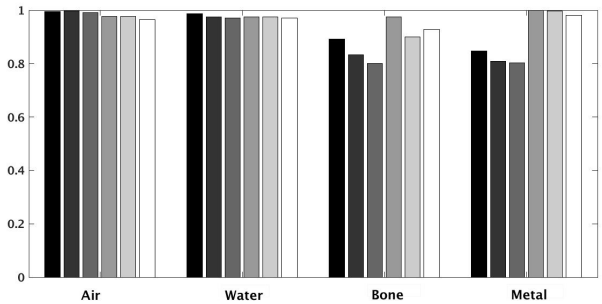


Figure 6: The bar plot demonstrates the performance of our algorithm against two pixel-based alternatives in terms of average precision and recall values. For each material, from left to right, the first three bars show the average precision over our algorithm's results, the results of the algorithm in [5], and the pixel-based KNN+MRF, respectively. The remaining three bars shows the average recall values for each material class using the same algorithms, respectively.

recall for each of the four materials than the competing alternatives.

In the next experiment, we evaluated our algorithm and the pixel-based alternatives on the data collected from bag scans in ALERT Task Order 3 data-set [12]. The bags were scanned using the GE Imatron 300 scanner at two voltage levels, 95 Kvp and 130 Kvp, which we refer to as low and high energy sources, respectively. We used 30 slices from bag 1 to serve as the training images. The bags contained a diverse set of background objects, as well as specific objects of interest such as water, saline water, metal, and rubber sheets. Including the background class, the experiment had five classes of interest. For each of these classes, we selected roughly the same number of pixels and patches from the 30 slices to use as training data.

In this experiment, the background class can have many types of objects with a wide range of linear attenuation coefficients. Since this would be a poor fit to a Gaussian distribution, we used KDE with normal kernels in our algorithm to estimate the patch-based conditional likelihood for each class. The bandwidth parameter $h$ in (6) was tuned using cross-validation slices from bag 2 which were different from the tested slices.

Figures 7 and 8 display the result of our patch-based algorithm and the pixel-based alternatives on two slices of bag 2. Fig. 7 shows that using KDE models for patch likelihoods is superior to using KNN models, Comparing the patch-based KDE reconstruction in the bottom right to the ground truth labeling in the top left illustrates that the patch based algorithm loses some resolution, as seen in the merging of metal pieces that are close by, and in the partial filling of the water in the bottle. However, the

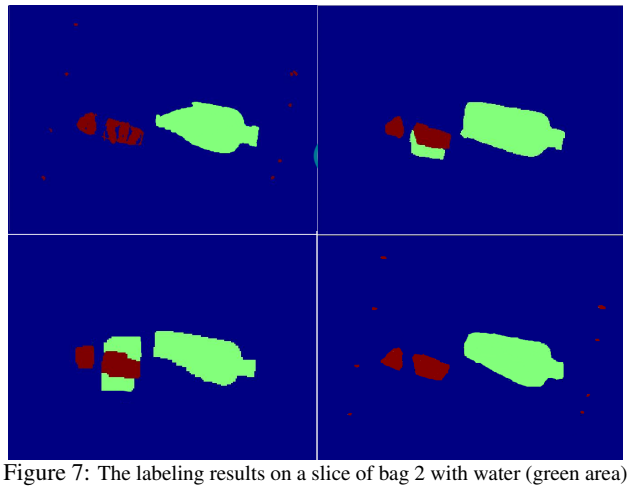

Figure 7: The labeling results on a slice of bag 2 with water (green area) and metal regions (red areas). The top row from left to right shows the ground truth labeling and the result using the algorithm in [5], respectively. The bottom row from left to right shows the patch based result when using KNN and KDE, respectively.

pixel-based algorithm of [5] also loses that resolution, and creates artifacts near those metal regions, mislabeling background pixels as water. The KNN patch-based algorithm also mislabels the background as water, indicating that the KNN model is not appropriate in these cases.

Fig. 8 displays the result of three labeling algorithms on a different slice, comprised primarily of metal regions, background and rubber. The image on the left corresponds to a pixel-based (patch size $K = 1$) using KNN conditional likelihood models; the center image was generated by the algorithm in [5], and the left image is our algorithm using KDE models with $2 \times 2$ patches of size. The pixel-based algorithm is very susceptible to pixel label errors in the neighborhood of metal. The algorithm in [5] avoids this using boundary information to avoid this problem, while still using KNN models. Our patch-based algorithm provides comparable results without using any extracted boundary information; however, the labeling shows some loss in resolution due to the use of patches vs pixels in the assignment of labels, as was expected.

## Conclusion

In this paper, we introduced a new algorithm for joint segmentation/classification of dual energy CT images using appearance models based on patches. These models are learned from training data, and capture the spatial correlation among neighboring pixel values to represent regional texture information. We modeled the prior information on patch labels using a Markov Random Field approach, and to pose the joint segmentation/classification problem as a maximum a posteriori estimation problem for the discrete patch labels. We developed an algorithm to solve the estimation problem using graph cut techniques, and evaluated the performance of this algorithm versus that of competing alternatives on both simulated data and on data collected using a medical scanner. Our results indicate that our patch-based segmentation/classification approach outperforms alternative approaches based on pixel-based appearance models.

There are several directions of interest for future investigation. First, we would like to explore whether patch-based appear-

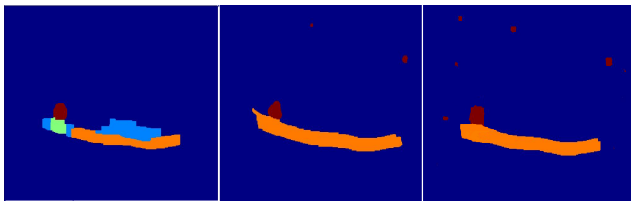

Figure 8: The labeling results on a different slice of bag 2 with rubber-sheets (orange area) and metal regions (red areas). From left to right shows the results of the pixel-based KNN+MRF algorithm, the algorithm in [5], and our algorithm, respectively.

ance models provide improvement when using different feature representations. In particular, we are interested in systems that use a single excitation spectra rather than dual energy systems, making the separation between object classes harder. We are also interested in exploring the effect of different representations for dual-energy reconstruction, such as the photoelectric and Compton reconstructions or the material basis reconstructions that have been proposed to alleviate metal artifacts in reconstruction. In addition, the emergence of less expensive photon-counting detectors can provide alternative features for reconstruction by collecting measurements over multiple energy bands. We would like to extend our patch based techniques to systems that use such features.

Along different directions, a weak point of our approach is the assumption that training data will be available on the materials of interest and on the scanner of interest. We would like to investigate whether such training data can be replaced by appropriately specifying appearance models based on material characteristics that may be generalized across scanners. These directions are currently under investigation in our ongoing research efforts.

## References


[1] S. Singh and M; Singh, Explosives detection systems for aviation security: A review, Signal Processing 83, 1 (2003).

[2] N. Megherbi, G. Flitton, T. Breckon, A classifier based approach for the detection of potential threats in CT based baggage screening, Proceedings of the IEEE International Conference on Image Processing (2010).

[3] Engineering Aviation Security Environments–Reduction of False Alarms in Computed Tomography-Based Screening of Checked Baggage, National Academies Press (2013)

[4] C. Crawford, Advances in automatic target recognition (ATR) for CT based object detection system–Final report, Dept. Homeland Security Center Excellence, Task order number HSHQDC–12–J–00429, Boston, MA, 2014.

[5] L. Martin, A. Tuysuzoglu, W.C. Karl, and P. Ishwar, Learning–Based Object Identification and Segmentation Using Dual–Energy CT Images for Security, IEEE Transactions on Image Processing 24, 11 (2015).

[6] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, A survey of MRI–based medical image analysis for brain tumor studies, Phys. Med. Biol., 58, 13 (2013).

[7] Y. Boykov, and Olga Veksler, Graph cuts in vision and graphics: Theories and applications, In Handbook of mathematical models in computer vision, Springer US, 2006, pp. 79-96.

[8] Y. Boykov, O. Veksler, and R. Zabih, Fast approximate energy minimization via graph cuts, IEEE Trans. Pattern Anal. Mach. Intell., 23, 11, (2001).

[9] K. Cranley, B. J. Gilmore, G. W. A. Fogarty, L. Desponds, and D. G. Sutton, Catalogue of diagnostic X-ray spectra and other data, Inst. Phys. Eng. Med. (IPEM), York, U.K., Tech. Rep. 78, (1997).

[10] Z. Yu, F. Noo, F. Dennerlein, A. Wunderlich, G. Lauritsch and J. Hornegger, Simulation tools for two-dimensional experiments in x-ray computed tomography using the FORBILD head phantom, Phys Med Biol., 57, 13 (2012).

[11] M. Berger et al, XCOM: Photon Cross Sections Database, Available online: http://physics.nist.gov/xcom (2005).

[12] C. Crawford, H. Martz, and W. C. Karl, Research and development of reconstruction advances in CT-based object detection systems–Final report, Dept. Homeland Security Center Excellence, ALERT, Boston, MA, Tech. Rep. HSHQDC-12-J-00056, (2013).


## Author Biography

*Parisa Babaheidarian received the B.Sc. degree in electrical engineering from University of Tehran in 2008. She graduated with the M.Sc. degree in electrical engineering from Sharif University of Technology in 2011. Currently, she is a Ph.D. candidate in electrical engineering program at Boston University. Her research interests include image recognition, analyzing CT images, signal processing, machine learning, and information theory.*

*David Castañón is Professor of Electrical and Computer Engineering at Boston University. He received his Ph.D. in Applied Mathematics from MIT (1976). He was Chief Scientist at ALPHATECH before joining Boston University (1990). He has been ECE Department Chair, co-director of the Center for Information and Systems Engineering, President of the IEEE Control Systems Society, and member of Air Force Scientific Advisory Board. His interests include control, estimation, optimization, inverse problems and image understanding.*