# Object Recognition in Multi-View Dual Energy X-ray Images

Muhammet Başstan[1]
mubastan@gmail.com, mbastan@turgutozal.edu.tr
Wonmin Byeon[2]
byeon@rhrk.uni-kl.de
Thomas M. Breuel[2]
tmb@cs.uni-kl.de

[1] Department of Computer Engineering, Turgut Özal University, Ankara, Türkiye (work done while at IUPR)
[2] Image Understanding and Pattern Recognition Group (IUPR), Technical University of Kaiserslautern, Germany

Object recognition in X-ray images is an interesting application of machine vision that can help reduce the workload of human operators of X-ray scanners at security checkpoints. However, automatic inspection systems using machine vision techniques are not yet commonplace for generic threat detection in X-ray images. Moreover, this problem has not been well explored by machine vision community due to the lack of publicly available X-ray image datasets. This paper aims to fill in this gap.

We first present a comprehensive evaluation of image classification and object detection in X-ray images using standard local features in a BoW framework with (structural) SVMs. Then, we extend the features to utilize the extra information available in dual energy X-ray images. Finally, we propose a multi-view branch-and-bound algorithm for multi-view object detection. Through extensive experiments on three object categories (laptops, guns, bottles), we show that the classification and detection performance substantially improves with the extended features and multiple views.

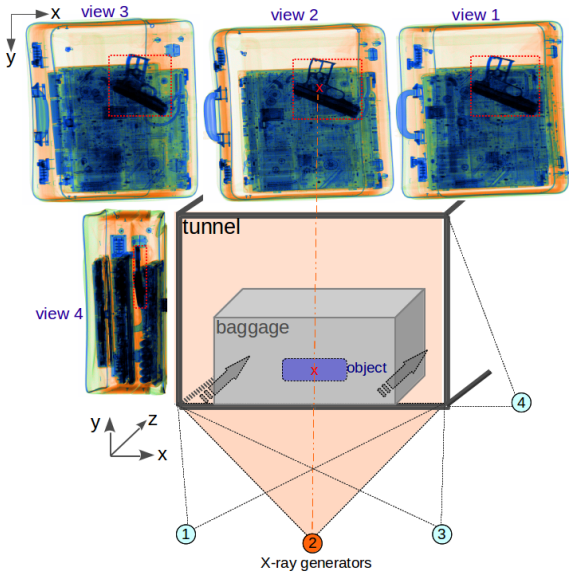


Figure 1: Simplified model of a 4-view X-ray scanner. As the baggage goes through the tunnel (in $z$ direction), 4 X-ray generators scan one slice of the baggage ($x-y$ plane) to generate two energy images for each of the 4 views. The energy images are combined through a look-up table to obtain the pseudo color RGB images for better viewing for the machine operators. The pseudo colors encode the material information (e.g., blue: metals, orange: organic). We make use of two main properties of these X-ray images to boost the object recognition performance: (1) material information, (2) multiple views.

X-ray images contain much ***less texture*** compared to regular photographic images, with the implication of harder object recognition. X-ray images may get extremely ***cluttered***, with many overlapping, possibly high density objects, making the inspection/recognition impossible even for the human screeners (in which case manual inspection is required). Objects usually undergo in– and out-of-plane ***rotations***, which is another major difficulty for recognition. On the other hand, the problem of change of object ***scale*** and ***illumination*** in regular images is not a big issue in X-ray images, since the machine geometry and energy levels are fixed. More importantly, the material information provided by ***dual energy imaging*** and multiple view images provided by ***multi-view imaging*** are both very helpful for improving the performance of object recognition.

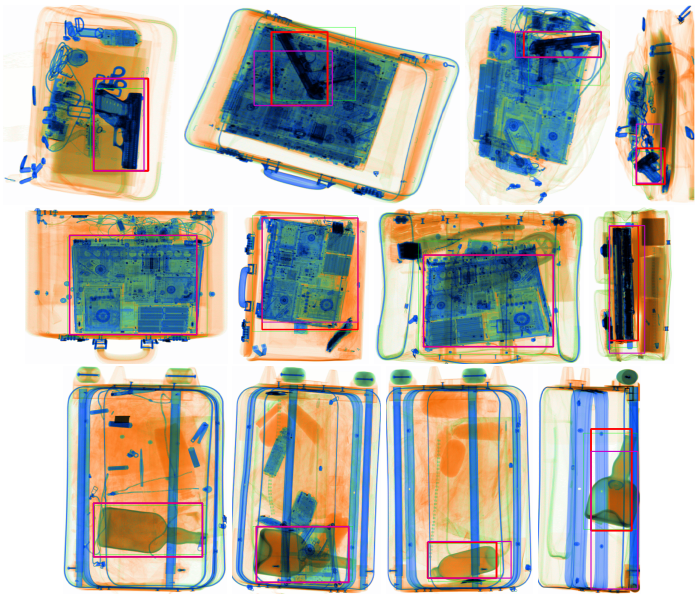

Figure 2: Sample single/multi-view object (handgun, laptop, glass bottle) detections. Green box (thin): ground truth, purple box (medium thickness): single view detection, red box (thick): multi-view detection.

We show that using multiple local color and texture features (different point detectors + descriptors) improves classification and detection performance. Furthermore, utilizing the material information, encoded in the dual energy images or pseudo color images, with extended SPIN descriptors (ESPIN, CSPIN) significantly improves the performance. For textureless object classes (e.g., bottles), contour and super pixel based features are better than sparse texture features. We present the results of extensive experiments with combinations of various features.

Multi-view imaging helps the screeners examine the baggage contents better from multiple viewing angles. We show that multi-view object detection improves over single view object detection [1, 3], especially when the single view detection does not work well. More specifically, we propose a multi-view branch-and-bound search algorithm to estimate the 3D location of an object passing through the scanner tunnel. To do so, we approximate the 3D bounding box of an object with 5 parameters (instead of 6, so that the B&B search converges in reasonable time) by making use of the machine geometry. We project the 3D box into 2D views and use the local features from 2D view images directly.

We show that our single wiew detector with multiple local features (BoWs + structural SVM + single view B&B search) works better than the HOG-based single view detector (HOG + linear SVM + sliding-rotating windows) proposed in [2]. Our multi-view object detector also works better than the multi-view detector of [2], which fuses the HOG-based single view detections in 3D to reinforce the geometrically consistent detections while suppressing the false detections.

[1] M.B. Blaschko and C.H. Lampert. Learning to Localize Objects with Structured Output Regression. In *ECCV*, 2008.

[2] T. Franzel, U. Schmidt, and S. Roth. Object Detection in Multi-view X-Ray Images. In *DAGM*, 2012.

[3] C.H. Lampert, M.B. Blaschko, and T. Hofmann. Efficient Subwindow Search: A Branch and Bound Framework for Object Localization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009.