A COMPARISON OF CLASSIFICATION APPROACHES FOR THREAT DETECTION IN CT BASED BAGGAGE SCREENING

Najla Megherbi, Jiwan Han, Toby P. Breckon, Greg T. Flitton
School of Engineering, Cranfield University, Bedfordshire, UK.

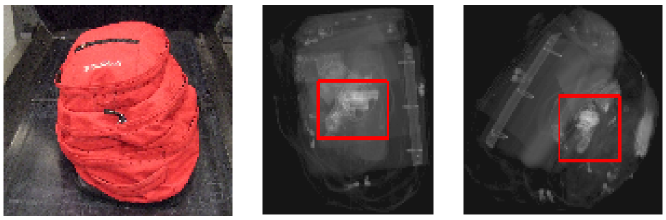

Fig. 1. Exemplar threat item in 2D X-ray imagery

ABSTRACT

Computed Tomography (CT) based baggage security screening systems are of increasing use in transportation security. The ability to automatically identify potential threat item is a key aspect of current research in this area. Here we present a comparison of varying classification approaches for the automated detection of threat objects in cluttered 3D CT imagery from such security screening systems. By combining 3D medical image segmentation techniques with 3D shape classification and retrieval methods we compare five varying final classification stage approaches and present significant performance achievements in the automated detection of specified exemplar items.

***Index Terms*—** aviation security, 3D medical image segmentation, 3D Zernike descriptors, histogram of shape index, automated classification, 3D object recognition.

I. INTRODUCTION

Advanced security screening systems are increasingly employed to aid human screening operators in the detection of potential threat items in a transport, notably aviation, security screening [1].

Although several X-ray technology based automatic systems exist for threats detection [2], only a few of these systems make use of the well established pattern recognition and classification techniques [3], [4], [5], [6], [7]. This prior work is predominantly in focused on 2D X-ray images and does not take advantage of information now offered by more recent systems using 3D CT imagery [8], [9]. In practice, 2D X-ray images obtained from the current luggage screening systems are not able to unambiguously reveal exact luggage content and hence make the process of screening operator interpretation of 2D X-ray images a difficult task - particularly for cluttered luggage (see Figure 1). As such, the need to use 3D CT imagery is of great importance to readily provide improved viewing and interpretation where we have an explicit 3D voxel-based representation of each and every object present (see Figure 2).

In prior work we have introduced a method for threat detection which deals with 3D volumes originating from CT based security screening systems [9]. This work follows a global-based 3D object recognition approach combining medical CT image segmentation approaches for object extraction with 3D shape descriptors [10], [11], developed within the established domain of 3D content retrieval systems. This global approach to object recognition with the threat detection context contrasts with contemporary work following a 2D or 3D Scale Invariant Feature Transform (SIFT) driven *"bag of visual words"* paradigm [7], [8]. In this paper we explicitly explore the use of an extended set of classification approaches for this task on the consideration of Support Vector Machines [12], Artificial Neural Network [13], Decision Trees, Boosted Decision Trees [14] and Random Forests [15].

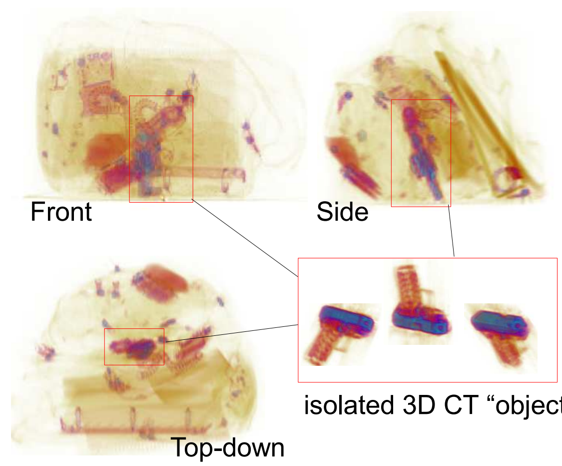

Fig. 2. Exemplar threat item in 3D CT imagery

II. FEATURE EXTRACTION AND PREPROCESSING

Our object detection approach takes 3D objects pre-segmented from the CT imagery using the approach of [16] from the domain of medical imagery. Prior studies

This project is funded under the Innovative Research Call in Explosives and Weapons Detection (2007), a cross-government programme sponsored by Home Office Scientific Development Branch (HOSDB), Department for Transport (DfT), Centre for the Protection of National Infrastructure (CPNI) and Metropolitan Police Service (MPS).

[9] identify [16] as outperforming several other CT segmentation methods [17] in our cluttered baggage data - especially for thin objects such as weaponry. Objects are preprocessed, normalized and for each 3D object instance a feature vector representation is constructed. For initial (off-line) classifier training this methodology is used to construct a set of positive and negative feature vectors examples for class differentiation. In the on-line classification/recognition phase a new unseen cluttered CT baggage volume is first segmented into multiple connected objects from which a feature vector representation is similarly constructed for each to enable classification via a given classifier methodology.

II-A. Feature Vector Construction

Our 3D segmented objects are described by a 3D shape descriptor made up of rotational-invariant 3D Zernike descriptors [11] and Histogram of Shape Index (HSI) [10] following the comprehensive review of [18], [19] within the domain of available 3D global shape descriptors.

3D Zernike descriptors are obtained from a voxelized model of 3D objects and are rotation invariant by design. In comparison to many other 3D shape descriptors, 3D Zernike descriptors are robust to noise [11]. Histogram of shape index is a local surface shape measure based on the curvature of the object. It is constructed by segmenting the range of the shape index curvature measure into equal sized histogram bins. The shape index itself, $SI(p)$, represents the shape of a local surface, at point p, by a single-value angular measure. It is defined as follows:

$$SI(p) = \frac{1}{2} - \frac{1}{\pi}.tan^{-1}\frac{k_1(p)+k_2(p)}{k_1(p)-k_2(p)} \tag{1}$$

Where $k_1(p)$ and $k_2(p)$ are the maximum and minimum curvature at point p of the surface. Each value of the shape index, which is in the range $[0 \rightarrow 1]$, corresponds to a distinct surface shape except for an exact plane where it is undefined. In contrast to 3D Zernike descriptors, the shape index is derived from a surface mesh model [10] and derived from curvature (known to be susceptible to noise).

As with most of shape representation methods, 3D Zernike descriptors and histogram of shape index require prior pose normalization in which the 3D object is normalized to achieve invariance with respect to geometrical transformations (i.e. rotation, scale, and translation). 3D Zernike descriptors [11] are rotation invariant but do not have scale and translation invariance properties. By contrast, the histogram of shape index descriptor is rotation and translation invariant [10]. Thus our 3D objects are pose-normalized with respect to scale and translation by translating and rescaling (i.e. voxel re-sampling) each object based on its identifiable 3D bounding box, from prior segmentation, within the original CT baggage image.

To generate Zernike descriptors, we use a maximal order of 20 to yield a Zernike descriptors feature vector in $\mathbb{R}^{121}$. The step bin of the shape index used to compute the histogram is 0.005 which results in a histogram vector in $\mathbb{R}^{200}$ over a shape index value range of $[0 \rightarrow 1]$. The resulting combined feature vector of concatenated Zernike descriptors and histogram of the shape index is thus in $\mathbb{R}^{321}$.

III. CLASSIFICATION

Five classification approaches are compared for the classification of such items within this study:- *Support Vector Machines [12], Artificial Neural Network [13], Decision Trees [14], Boosted Decision Trees [20] and Random Forests [15].* Our support vector machine approach utilizes grid search [12] to identify a maximally performing kernel and parameter set (RBF kernel) whilst empirical experimentation identifies a maximal optimal 200-hidden node configuration for our Neural Network approach trained via backpropogation [13]. For comparison we additionally present a tryptic of decision tree classification approaches from basic (post-pruned) tree construction using CART *[14]*, a boosting based ensemble approach *[20] (discrete Adaboost, 100 weak classifiers)* and the seminal contemporary image classification approach of Random Forests *(100 tree classifiers).* Each approach is trained, using a cross-validation approach as appropriate, to give a binary detection (present/not present) output for a given object class.

IV. EXPERIMENTAL RESULTS

Our evaluation is based on a CT baggage imagery data-set containing two threat object classes: firearms and bottles [8], [9] (236 CT images containing a set of 35 bottles and 7 firearms variants). Within current air travel security regulations, bottles represent interesting target objects to examine in any luggage screening process. Due to the limited availability of firearm examples we use bottles as the class object with which to present our results. Our training set consists of 24 bottle example volumes (positive samples) and 55 non-bottle volumes (negative samples) segmented and normalized to a scale of 60^3 voxels. In Figure 3 we show a representative set of the positive training examples used. Negative samples are randomly generated from negative CT volumes (without bottles) and subsequently normalized (see pre-normalisation examples in Figure 4). We compare the performance of varying classification approaches in identifying this object class using both isolated Zernike ($\mathbb{R}^{121}$) or histogram of shape index (HSI, $\mathbb{R}^{200}$) descriptors and additionally the (full) combined $\mathbb{R}^{321}$ Zernike/histogram of shape index feature vector as outlined in Section II-A.

Our test data set consists of 61 positive and 65 negative variant bottle object examples within which 19 of the physical bottle types present have not been used in training set of volumes. From Table I to Table V we summarize the results considering Accuracy, Precision, True Negative Rate (TNR) and Recall as defined in Equations 2 to Equation 5 and subsequently scaled to a percentile range:-

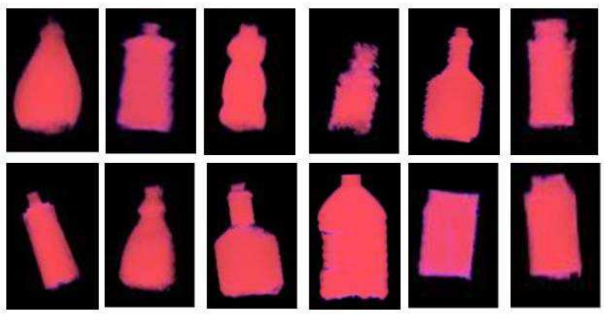

Fig. 3. A representative set of 3D positive object examples

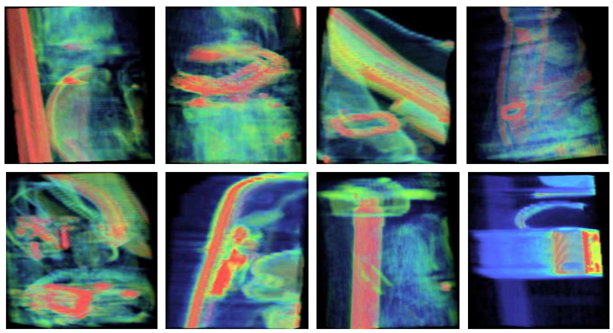

Fig. 4. A representative set of 3D negative object/clutter examples

$$Accuracy = \frac{tp + tn}{tp + fp + tn + fn} \quad (2)$$

$$Precision = \frac{tp}{tp + fp} \quad (3)$$

$$TrueNegativeRate(TNR) = \frac{tn}{tn + fp} \quad (4)$$

$$Recall = \frac{tp}{tp + fn} \quad (5)$$

where tp, tn, fp and fn are defined as the recorded number of true positive, true negative, false positive and false negative occurrences respectively.

From Tables I - V it can been seen that the varying classifiers offer good performance on both the Histogram of Shape Index (HSI) based and combined Zernike/HSI feature vector input (outperforming isolated Zernike descriptors in all cases). Notably the performance of the simple decision tree classifier and the boosted tree ensemble classifier is identical on these two feature vector input cases (Table III/IV) with marginal out-performance by the Random Forest

Feature Vector	Accuracy	Precision	TNR	Recall
Isolated Zernike	93.65	88.52	90.14	98.18
Isolated HSI	98.41	96.72	97.01	100
Combined	98.41	96.72	97.01	100

Table I. Performance of Support Vector Machine classifier (%)

Feature Vector	Accuracy	Precision	TNR	Recall
Isolated Zernike	80.95	91.11	93.85	67.21
Isolated HSI	89.68	96.15	96.9	81.97
Combined	97.61	100	100	95.08

Table II. Performance of Neural Network classifier (%)

Feature Vector	Accuracy	Precision	TNR	Recall
Isolated Zernike	70.63	81.58	89.23	50.82
Isolated HSI	98.41	100	100	96.72
Combined	98.41	100	100	96.72

Table III. Performance of Decision Tree classifier (%)

Feature Vector	Accuracy	Precision	TNR	Recall
Isolated Zernike	89.68	92.86	93.85	85.25
Isolated HSI	98.41	100	100	96.72
Combined	98.41	100	100	96.72

Table IV. Performance of Boosted Decision Tree classifier (%)

approach (Table V). The support vector machine classifier outperforms all other approaches on the combined descriptor (Table I) offering comparable performance on the isolated histogram of shape index feature vectors to the Random Forest approach. We can conclude the histogram of shape index is a maximally discriminative feature, over and above 3D Zernike descriptors, for this classification task over a range of independent classification approaches.

Some representative detection results obtained by our approach using both feature vectors are depicted in Figure 5. Bottles are detected independently of their segmentation quality which can depend on the bottle background in the bag which can contain a significant amount of clutter and exhibit a huge variability over the data-set. In addition, the segmentation results can depend on the bottle liquid fill level with the liquid inside the bottle taking several forms in the CT image depending on the position of the bottle in the baggage item and also the position of the item within the CT security scanner [9]. This is a synonym to the 3D occlusion case in non-CT based 3D object retrieval but using our method we illustrate bottles with varying liquid fill levels are still detected (Figure 5).

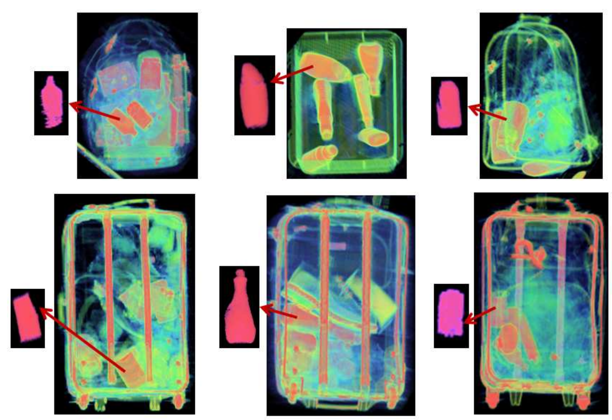

Fig. 5. Correct detection with all classifiers using combined histogram of shape index and Zernike descriptors

In Figure 6 we show some negative samples which have

Feature Vector	Accuracy	Precision	TNR	Recall
Isolated Zernike	89.95	93.02	95.38	65.57
Isolated HSI	100	100	100	100
Combined	98.41	100	100	96.72

Table V. Performance of Random Forest classifier (%)

been classified by the classifiers in use as bottles with isolated 3D Zernike descriptors. All these false detections have been subsequently correctly classified by all approaches, other than the neural network, when the 3D Zernike descriptors are combined with the histogram of shape index.

Figure 7 shows some typical missed detections made by the classifier approaches in use when isolated Zernike descriptors are used as a feature vector. Although all the bottles which are not detected are the same as types used in training set, they have not been recognized in the testing phase. However, for the support vector machine approach all of the 19 bottle types not used in training have are correctly detected using both combined and isolated feature vector combinations.

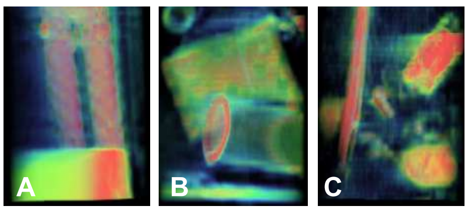

Fig. 6. False detections with isolated Zernike descriptors: (A) and (B) neural network, decision tree and random forest, (C) boosted decision tree and support vector machine

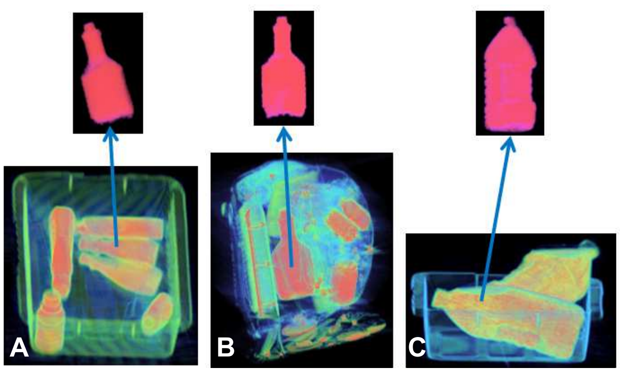

Fig. 7. Missed detections with isolated Zernike descriptors: (A) neural network, decision tree and boosted decision tree, (B) decision tree and boosted decision tree, (C) random forest and support vector machine

V. CONCLUSION

In this paper we compare a range of classification approaches for the detection of potential threats in 3D CT baggage imagery based on the use of two popular 3D shape descriptors (3D Zernike descriptors and histogram of shape index). The results suggest that isolated histogram of shape index descriptors consistently outperforms 3D Zernike descriptors alone. This is despite the reliance of the former on surface curvature which is known to be susceptible to noise. Maximal classification success was achieved using feature descriptor in combination with a Support Vector Machine or Random Forest classifier approach. Future work will investigate the extension of this technique to larger datasets and more sophisticated threat objects such as weapons.

VI. REFERENCES

[1] N.E.L. Shanks and A.L.W. Bradley, *Handbook of Checked Baggage Screening: Advanced Airport Security Operation*, Wiley Blackwell, 2004.

[2] S. Singh and M. Singh, "Explosives detection systems (EDS) for aviation security," *Signal Processing*, vol. 83, no. 1, pp. 31–55, 2003.

[3] R. Gesick, C. Saritac, and C. Hung, "Automatic image analysis process for the detection of concealed weapons," in *Proc. of the 5th Annual Workshop on Cyber Security and Information Intelligence Research*, 2009.

[4] V. Muthukkumarasamy, M. Blumenstein, J. Jo, and S. Green, "Intelligent illicit object detection system for enhanced aviation security," in *Int. Conf. on Simulated Evolution and Learning*, 2004.

[5] X. Shi, *Improving object classification in X-ray luggage inspection*, Ph.D. Thesis, Dept. Computer and Electrical Engineering, Virginia Tech. and State University, 2000.

[6] T. Feather, I. Guan, A. Lee-Kwen, and R.B. Paranjape, "Caxss: an intelligent threat detection system," *SPIE - Applications of Signal and Image Processing in Explosives Detection Systems*, vol. 1824, pp. 152–161, 1993.

[7] Muhammet Baştan, Mohammad Reza Yousefi, and Thomas M. Breuel, "Visual words on baggage x-ray images," in *Proc. Int. Conf. on Computer Analysis of Images and Patterns*, 2011, pp. 360–368.

[8] G. Flitton, T.P. Breckon, and N. Megherbi, "Object recognition using 3D SIFT in complex CT volumes," in *Proc. British Machine Vision Conf.*, 2010, pp. 11.1–12.

[9] N. Megherbi, G.T. Flitton, and T.P. Breckon, "A classifier based approach for the detection of potential threats in CT based baggage screening," in *Proc. Int. Conf. on Image Processing*, 2010, pp. 1833–1836.

[10] C. Dorai and A.K. Jain, "COSMOS-a representation scheme for free-form surfaces," in *Fifth Int. Conf. on Computer Vision*. IEEE, 1995, pp. 1024–1029.

[11] M. Novotni and R. Klein, "Shape retrieval using 3D Zernike descriptors," *Computer Aided Design*, vol. 36, no. 11, pp. 1047–1062, 2004.

[12] C. Hsu, C. Chang, and C. Lin, "A practical guide to support vector classification," *Bioinformatics*, vol. 1, no. 1, pp. 1–16, 2010.

[13] C.M. Bishop, *Neural Networks for Pattern Recogntion*, Oxford University Press, 1995.

[14] L. Breiman, *Classification and Regression Trees*, Chapman & Hall/CRC, 1984.

[15] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[16] J.K. Udupa and S. Samaresekera, "Fuzzy connectedness and object definition: Theory, algorithms and applications in image segmentation," *Graphical Models and Image Processing*, vol. 58, no. 3, pp. 246–261, 1996.

[17] O. Wirjadi, "Survey of 3D image segmentation methods," 2007, Technical report No. 123, Fraunhofer ITWM, Kaiserslautern, Germany.

[18] N. Iyer, S. Jayanti, Kuiyang Lou, Y. Kalyanaraman, and K. Ramani, "Three dimensional shape searching: State-of-the-art review and future trends," *Computer-Aided Design*, vol. 37, no. 5, pp. 509–530, 2005.

[19] Z. Qin, J. Jia, and J. Qin, "Content based 3D model retrieval: A survey," in *Int. Workshop on Content-Based Multimedia Indexing*. IEEE, 2008, vol. II, pp. 249–256.

[20] H. Drucker and C. Cortes, "Boosting decision trees," *Advances in Neural Inf. Proc. Sys.*, pp. 479–485, 1996.