Target and Background Signatures III, 11-14 September 2017, Warsaw, Poland.
DOI: 10.1117/12.2277190

Automatic X-ray Image Segmentation and Clustering for Threat Detection

Odysseas Kechagias-Stamatis*, Nabil Aouf, David Nam, Carole Belloni
Signals and Autonomy group, Centre for Electronic Warfare Information and Cyber,
Cranfield University, Defence Academy of the United Kingdom, Shrivenham SN6 8LA,UK

ABSTRACT

Firearms currently pose a known risk at the borders. The enormous number of X-ray images from parcels, luggage and freight coming into each country via rail, aviation and maritime presents a continual challenge to screening officers. To further improve UK capability and aid officers in their search for firearms we suggest an automated object segmentation and clustering architecture to focus officers' attentions to high-risk threat objects. Our proposal utilizes dual-view single/ dual-energy 2D X-ray imagery and is a blend of radiology, image processing and computer vision concepts. It consists of a triple-layered processing scheme that supports segmenting the luggage contents based on the effective atomic number of each object, which is then followed by a dual-layered clustering procedure. The latter comprises of mild and a hard clustering phase. The former is based on a number of morphological operations obtained from the image-processing domain and aims at disjoining mild-connected objects and to filter noise. The hard clustering phase exploits local feature matching techniques obtained from the computer vision domain, aiming at sub-clustering the clusters obtained from the mild clustering stage. Evaluation on highly challenging single and dual-energy X-ray imagery reveals the architecture's promising performance.

Keywords: Object Clustering, Object Segmentation, Threat Detection, X-ray Images

1. INTRODUCTION

One of the major worldwide border considerations is firearm smuggling. A common method to detect such activities is via X-ray screening the luggage coming into each country via rail, aviation and maritime. This screening operation is performed by specially trained offices. Considering the enormous amount of luggage entering each country every day, the high number of X-ray images produced presents a continual challenge to screening officers. Therefore, current literature aims at aiding officers in their search for firearms within congested luggage either by clustering the objects [1,2] or by introducing fully automated threat detection systems in the 2-dimensional (2D) [3,4] or 3-dimensional space (3D) [5–9].

Current X-ray clustering techniques are either applied on 3D Computed Tomography (CT) data [1] or on 2D imagery [2]. Even though 3D data can afford higher levels of information completeness, as it reveals the underlying structure of the contained objects [10], it has a number of disadvantages. First, it requires quite a few X-rays in order to construct a single 3D CT representation. In addition, 3D data processing is computationally deficient compared to 2D data processing. Regarding the current 2D X-ray clustering techniques for threat detection, literature relies on single image concepts utilizing standard image processing techniques [2].

Independently of the application, common object clustering techniques include k-means [11], hierarchical clustering [12], DBSCAN [13], multivariate Gaussian distributions using the Expectation-maximization (EM) algorithm [14] and sparse models [15]. Although these standard clustering techniques are quite widespread, they fail clustering overlapping objects, which is a very common situation for luggage contents. Although recent techniques from the biomedical community [16,17] can afford clustering of overlapping objects, these techniques are constrained to simplistic elliptical shapes rather than complex unknown objects which is the case investigated in this paper.

Urged by the deficiencies of current clustering techniques, in relation to complex object clustering, we propose an automated X-ray image segmentation and clustering method that suits the needs of border security. In specific, we suggest a multi-discipline multi-level X-ray image segmentation and clustering architecture driven by radiology, image

*o.kechagiasstamatis@cranfiled.ac.uk

processing and computer vision concepts. Our architecture is applicable on single/ dual-energy dual-view X-ray images affording enhanced segmentation and clustering performance, even in case of overlapping or highly connected objects. This is achieved by forcing disjoining objects based on a soft and a hard clustering procedure that relies on state-of-the-art local feature matching techniques.

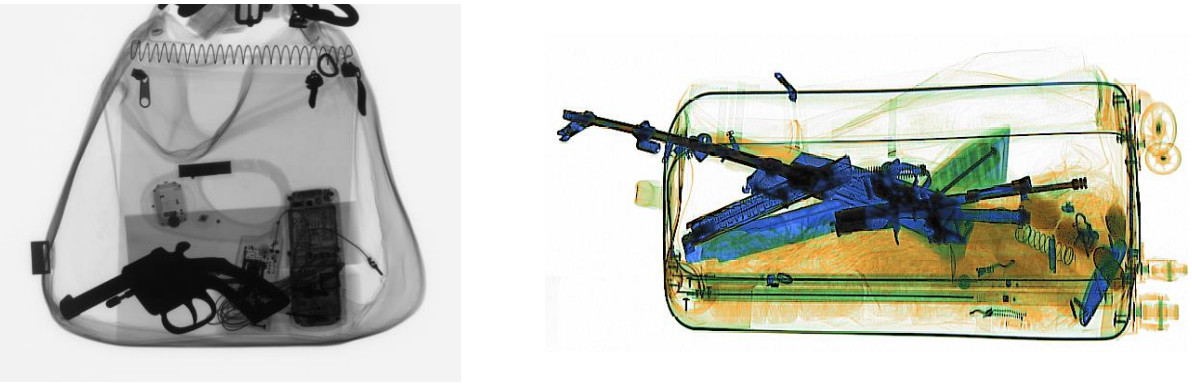
The rest of this paper proceeds as follows. Section 2 describes the single/ dual energy x-ray imagery and the dataset used. Section 3 presents the proposed automated X-ray image segmentation and clustering architecture and Section 4 concludes this paper.

2. SINGLE/ DUAL ENERGY X-RAY IMAGERY

X-ray imagery has the ability to differentiate objects of various materials based on the interaction between X-ray photons and the atomic structure of the target material. Depending on the target being scanned, different energy levels of X-rays are used starting from 140 – 160 keV for baggage and increase up to MeV for other types of cargo. X-ray machines typically have two modes, single energy and dual energy. Single energy systems emit photons at a single energy level and produce a grayscale image with higher density materials being darker. Dual energy systems emit photons at two distinct energy levels. These systems are able to approximate the effective atomic number Z_{eff} of different materials, by comparing the measurements relative to a baseline. This capability allows the system to generate a false color image of the object, giving the operator the ability to discriminate between different materials. Table 1 outlines the typical color scheme used while an example is illustrated in Figure 1.

Table 1: Color scheme for different material densities in a dual-energy X-ray

Z_{eff}	Color	Typical materials
Low	Orange	Organic material, natural substances
Mid	Green	Plastics, alloys, ceramics, light metallic elements
High	Blue/Black	Hard, dense materials e.g. steel, lead



(a) (b)

Figure 1. Example of a (a) single energy X-ray image of a handbag and (b) dual energy X-ray of a luggage containing a firearm. (single-energy X-ray courtesy of [18])

Typical X-ray security scanning devices have an object passing through the X-ray target zone where multiple images are captured and collated to produce a single 2D image. Although multiple images are used, this procedure is different to a CT (Computed Tomography) X-ray, which provides a full 3D image of the target object. Technological advancements have increased the number of dual energy images from single one up to two or even four that are created by a single pass of the scanned object through the X-ray target zone. This means that in a single pass, two or four perspectives can be viewed.

3. PROPOSED ARCHITECTURE

The proposed object segmentation and clustering pipeline is a blend of radiology, image processing and computer vision concepts applied on a dual-view imaging scheme. Our architecture is presented in Figure 1 and will be analyzed in the following paragraphs. During that analysis, dual-energy X-rays are considered as baseline images while this section concludes by tackling the challenging single-energy imagery.

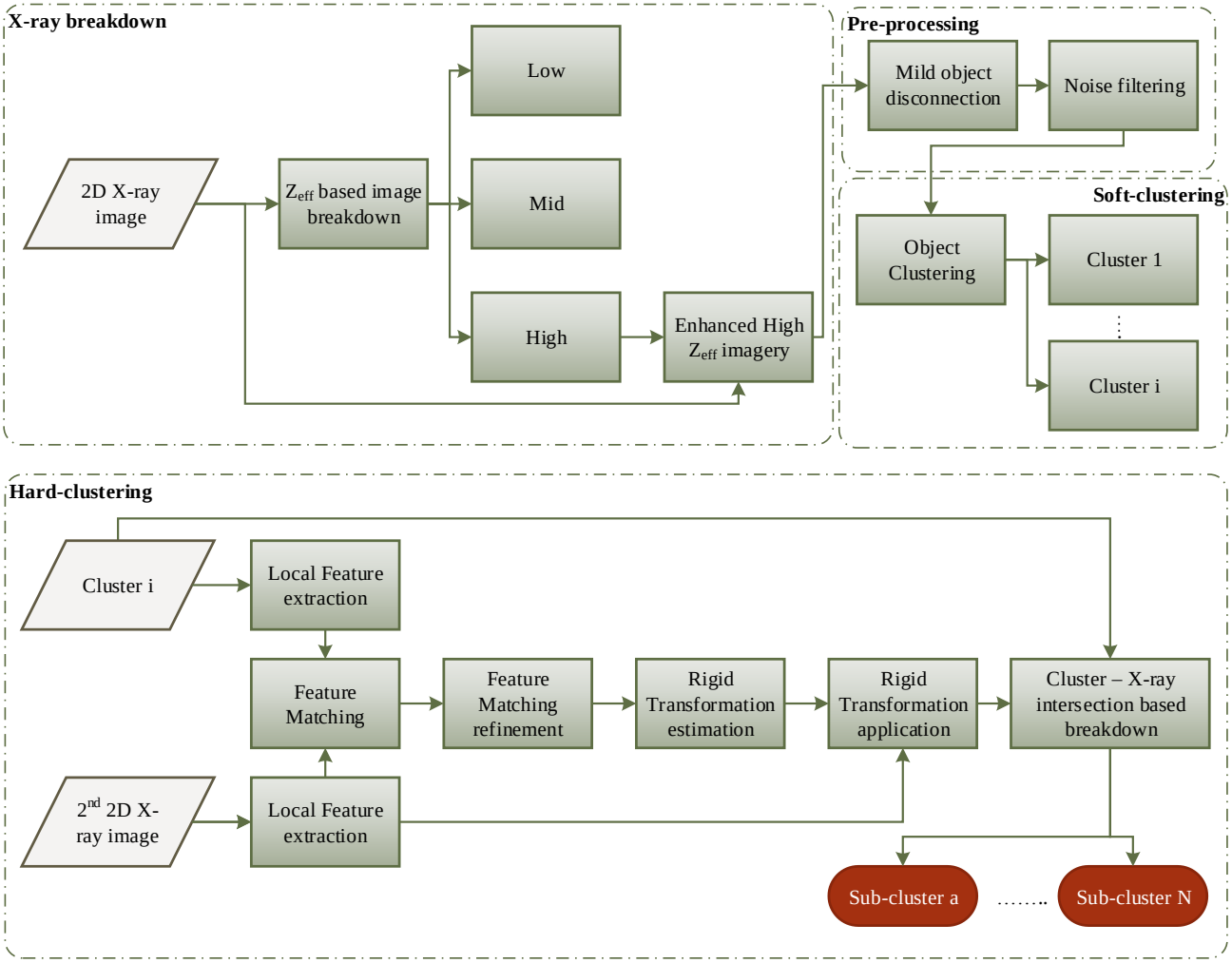


Figure 2. Proposed object segmentation and clustering architecture.

3.1 Dual-Energy X-ray image breakdown

The main information obtained from a dual-energy 2D X-ray image f is the type of materials it contains. Objects are distinguished based on their effective atomic number Z_{eff} (Table 2), which we exploit to establish three image subcomponents f_{low}, f_{mid}, f_{high} that correspond to the low, medium and high Z_{eff} values.

Table 2: RGB color space response based on Z_{eff}

Z_{eff}	Color	RGB color space response
Low	Orange	Red
Mid	Green	Green
High	Blue/Black	Blue

The usual case is that every pixel is a mixture of all RGB color space responses with one of the colors gaining the highest response. Thus, we correlate the maximum RGB response of a pixel with $f(i,j,c), c \in \{R,G,B\}$ and i, j are the X-ray image coordinates, with the dominant material at position i, j according to Table 2 and given by:

$$f_c(i,j) = \begin{cases} 1 & if \ \left(\max\left(f(i,j) \right) \geq 0.33 \rightarrow \underset{c}{\operatorname{argmax}}\left(f(i,j,c) \right) \right) \vee \left(\max\left(f(i,j,c) \right) < 0.33 \rightarrow c = 1 \right) \\ 0 & elsewhere \end{cases} \quad (1)$$

In the event a pixel has the same maximum response in two color-bands, then that pixel will belong in both subcomponent images. From the three-subcomponent images created, the f_B contains the hard and dense materials which will likely contain the firearm and therefore this sub-component is further propagated through the suggested pipeline. For better performance, objects in f_B are enhanced by incorporating the low/ mid Z_{eff} materials enclosed as extracted from f (Figure 3):

$$f_{high} = f \circ f_B \quad (2)$$

where $\circ$ is the *Hadamard* product [19].

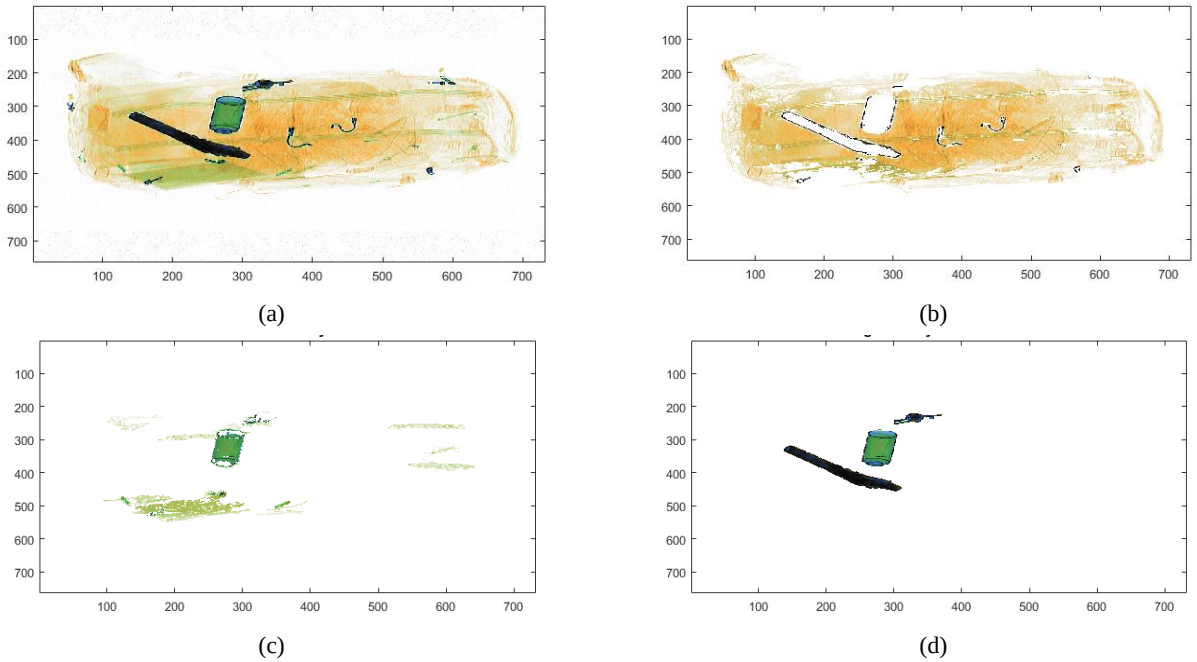


Figure 3. X-ray of baggage, split into its three subcomponents based on the effective atomic number Z_{eff} of each material contained (a) dual-energy X-ray image (b) f_R (c) f_G (d) f_{high}.

3.2 Soft object clustering

Objects contained in the high-density image are then disconnected by applying a fine morphological operation of image opening and closing. For that process, we use a disk structuring element s with one pixel radius in order to preserve, as possible, the initial shape of each object:

$$f_{high}^{'} = \left(f_{high} * s \right) \bullet s \tag{3}$$

Where $*$ denotes image opening and $\bullet$ image closing. This methodology can afford disjoining two objects that are not overlapping or that are not heavily connected. During our experiments, we evaluated several structuring element shapes e.g. square and sizes, which provided though a poorer soft object clustering performance. This is because, for a given radius, the disk has the smallest area and thus decimates less the objects within f_{high}. Similarly, increasing the size of the structuring element corrupts the information contained in f_{high}.

On the manipulated high-density image $f_{high}^{'}$, artefacts smaller than 400pixels sized are discarded as noise. This size is properly selected such as to balance preserving weapon sub-components while rejecting small clutter objects. An example is shown in Figure 4, where the effectiveness of the soft object clustering is indicated with green arrows. The same image highlights with a red arrow the poor performance on highly connected objects, highlighting the requirement for additional more sophisticated approaches that are presented in Section 3.4.

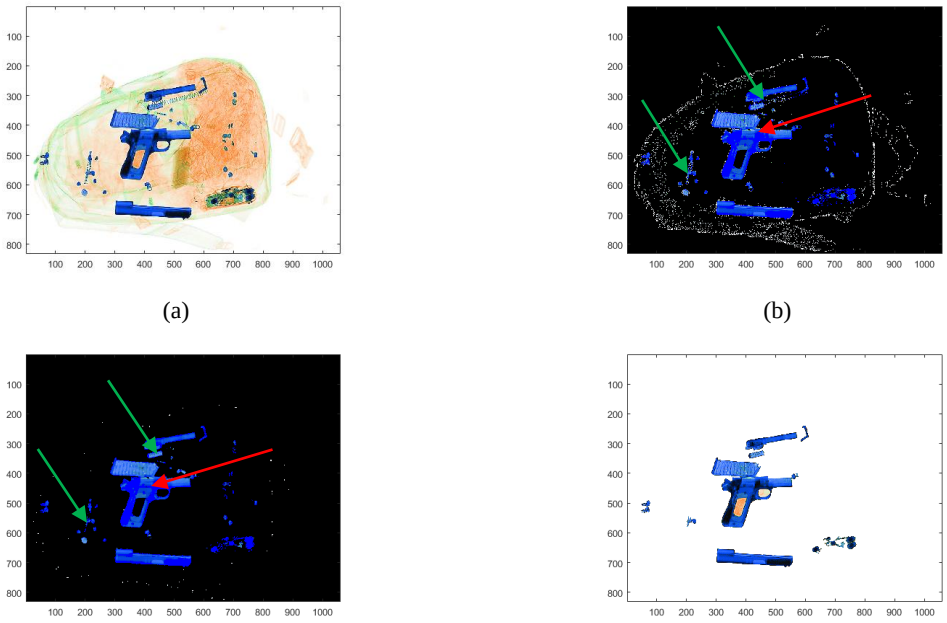


Figure 4. (a) Original dual energy 2D X-ray imagery (b) High-density subcomponents, red arrows highlight connected regions of different objects (c) Image after morphological opening and closing (d) final high-density imagery after discarding small artefacts

3.3 Object clustering

Given the processed high-density image $f_{high}^{'}$, we cluster the contents based on a pixel-pair connectivity metric. Pixel-pairs that meet the 8-connectivity requirement are assigned under the same cluster Cl_k, otherwise to a different cluster. Figure 5 depicts a clustering example where nine clusters are identified.

At this stage, the suggested object segmentation and clustering pipeline exploits one dual-energy 2D X-ray image. It is possible though that high-density objects within the baggage are overlapping or are highly connected. To compensate that we force disjoining the objects belonging to the same cluster by exploiting concepts from the computer vision domain.

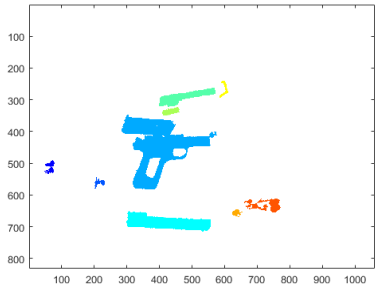


Figure 5. Clustering example on the processed high-density image

3.4 Hard object clustering

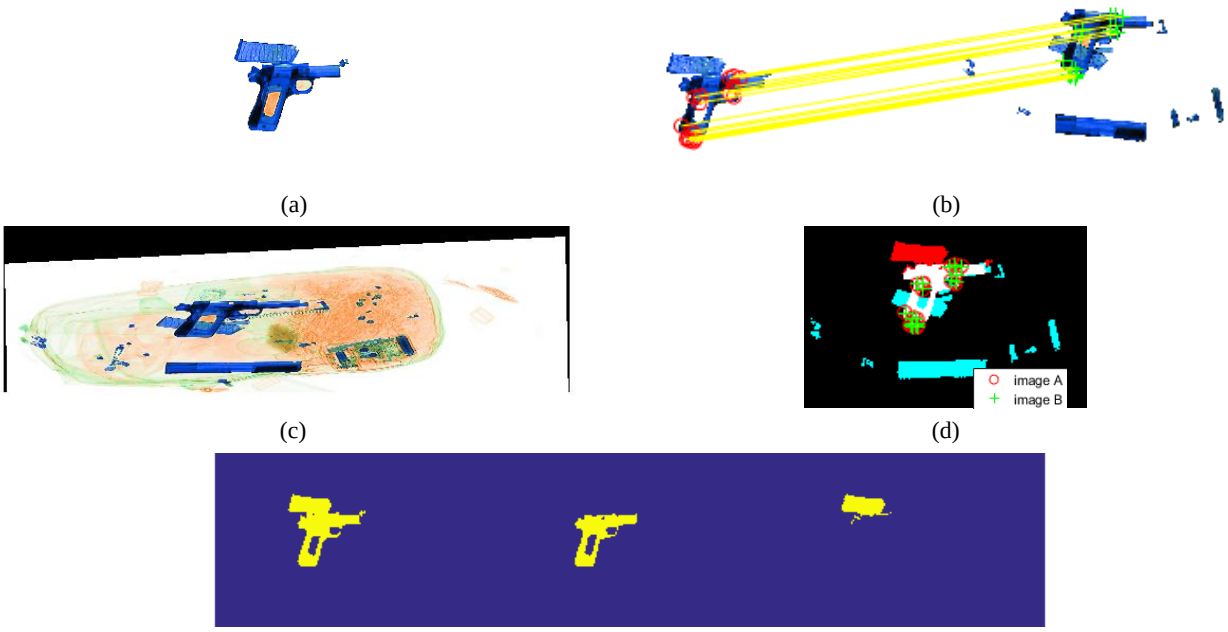
We force clustering overlapping/ heavily connected objects by introducing a hard object clustering process that takes advantage of current computer vision concepts. These concepts are applied on a second complimentary view $f^{'}$ of the same luggage image f .

Specifically, we propagate $f^{'}$ via the pipeline presented in Sections 3.1 and 3.2 and create image $I_{high}^{'}$. Then $I_{high}^{'}$ along with the clusters Cl obtained from image f after the soft disjoining and clustering modules (Sections 3.1 – 3.3), are input to the hard object clustering module (Figure 2). The first stage of that module considers applying a local feature based description technique on $I_{high}^{'}$ and on each cluster Cl_i . Although a number of such description methods exist, we choose SURF [20] as it balances performance with processing burden. Due to the modality difference between a true RGB imagery, where SURF was developed in, and the X-ray imagery, we set a low keypoint detection threshold of 10^{-3}, within three Octaves and six scale levels. After SURF keypoint detection and description, the SURF features of $I_{high}^{'}$ are matched against the corresponding ones of each cluster Cl_i using a Nearest Neighbor Distance Ratio [21] criterion. Finally, matches are refined based on RANSAC [22]. Figure 6 (a) shows cluster Cl_3 and Figure 6 (b) the refined matches between Cl_3 and $I_{high}^{'}$.

The matched keypoints leverage a rigid transformation T between Cl_i and $I_{high}^{'}$ which is estimated by RANSAC during the keypoint matching refinement stage. Estimating T requires at least three matches and in the event RANSAC does not qualify these, then Cl_i cannot be sub-clustered and the process terminates for cluster Cl_i . If $T=\{r,t\}$ is estimated, it is applied to $I_{high}^{'}$ creating the transformed image I_{high}^{T} (Figure 6 (c)). Finally, I_{high}^{T} and Cl_i are fused via an intersection scheme (Figure 6 (d)) and the two sub-clusters Cl_i^A, Cl_i^B of Cl_i presented in Figure 6 (e) are given by:

$$Cl_i^A = Cl_i \bigcap I_{high}^T \ \ and \ \ Cl_i^B = Cl_i - Cl_i^A \tag{4}$$

It is evident that the quality of the hard object clustering process is related to the complimentary nature of the two X-ray images used. Therefore, it is quite important to obtain a few X-ray images aiming at capturing firearms from complimentary poses. Care has to be taken that firearms contained have a pairwise pose change that does not exceed the 30° affine transformation limit that SURF can handle [20]. A solution to this consideration might be using a quad-view X-ray scanning systems and exploit the optimum combination to achieve a valuable firearm clustering.



(a) (b) (c) (d) (e)

Figure 6. Hard object clustering pipeline (a) cluster Cl_i of image f'_{high} (b) SURF feature matching of Cl_i and I'_{high} (c) transformed image I^T_{high} based on the matched SURF features (d) image fusion of Cl_i and I^T_{high} highlighting the matched keypoints (e) sub-clusters Cl_i^A, Cl_i^B as given by Equation 4

Our architecture is also applicable to single-energy dual-view X-ray imagery. A major disadvantage of such imagery is that it relies on a single color band i.e. grayscale, rather than three-color variants as the dual-energy does i.e. RGB domain. This leads to poor or even absent texture information of the objects contained and therefore, single-energy object clustering is more challenging compared to the dual-energy.

To enhance performance we pseudo-color the single-energy X-rays by remapping them from the grayscale to a RGB color scheme. Thereafter, the manipulated X-ray images are input to the architecture proposed in Section 3. An example of the suggested technique on single-energy X-ray images from [18] is presented in Figure 7. Despite the weapon being texture-less our enhancement strategy still afforded disjoining the weapon from the clutter object. It should be noted though that compared to the dual-energy dual-view X-ray example, the sub-clustering accuracy is slightly reduced as a part of the clutter object is still connected. This is because SURF does not provide good quality keypoints and so the transformation established based on the matched SURF features of the two X-ray images does not optimize the sub-clustering procedure. Despite that, our architecture still achieves clustering the weapon out of its initial cluster offering enhanced weapon detection capability either for manual by the border officer or for automated threat detection purposes.

4. CONCLUSION

We suggest an automated object segmentation and clustering architecture to aid security officers while searching luggage or bags for high-risk threat objects. Our proposal utilizes dual-view single or dual-energy 2D X-ray imagery and exploits concepts from the radiology, image processing and computer vision domains. It consists of a triple-layered segmentation and clustering scheme capable of disjoining overlaying objects. Evaluation on highly challenging X-ray imagery reveals the architecture's promising performance.

Future work shall focus on identifying the optimum number of X-rays such as to optimize the complimentary nature of the two images used in our suggested architecture.

(a) (b)

Figure 7. (a) Single-energy X-ray imagery of a bag (b) weapon cluster disjoint (X-ray image from [18])

ACKNOWLEDGEMENTS

The authors would like to thank the Home Office for funding this work and providing dual-energy X-ray images. Views contained within this paper are not necessarily those of the Home Office.

REFERENCES

[1] Mouton, A.., Breckon, T. P., "Materials-based 3D segmentation of unknown objects from dual-energy computed tomography imagery in baggage security screening," Pattern Recognit. **48**(6), 1961–1978 (2015).

[2] Jimin Liang., Abidi, B. R.., Abidi, M. A., "Automatic X-ray image segmentation for threat detection," Proc. Fifth Int. Conf. Comput. Intell. Multimed. Appl. ICCIMA 2003, 396–401, IEEE Comput. Soc.

[3] Turcsany, D., Mouton, A.., Breckon, T. P., "Improving feature-based object recognition for X-ray baggage security screening using primed visualwords," Proc. IEEE Int. Conf. Ind. Technol., 1140–1145 (2013).

[4] Kundegorski, M. E., Akcay, S., Devereux, M., Mouton, A.., Breckon, T. P., "On using Feature Descriptors as Visual Words for Object Detection within X-ray Baggage Security Screening," Proc. Int. Conf. Imaging Crime Detect. Prev., 1–6 (2016).

[5] Mouton, A., Breckon, T. P., Flitton, G. T.., Megherbi, N., "3D Object Classification in Baggage CT Imagering using Randomised Clustering Forests," Int. Conf. Image Process., 5202–5206 (2014).

[6] Chermak, L.., Breckon, T., "Geometrical approach for automatic detection of liquid surfaces in 3D computed tomography baggage imagery," Imaging Sci. J. **63**(8), 1–9 (2015).

[7] Megherbi, N., Breckon, T. P.., Flitton, G. T., "Investigating existing medical CT segmentation techniques within automated baggage and package inspection," 89010L (2013).

[8] Flitton, G., Mouton, A.., Breckon, T. P., "Object classification in 3D baggage security computed tomography imagery using visual codebooks," Pattern Recognit. **48**(8), 2489–2499 (2015).

[9] Bastan, M., Byeon, W.., Breuel, T. M., "Object Recognition in Multi-View Dual Energy X-ray Images," Br. Mach. Vis. Conf., 1–11 (2013).

[10] Kechagias-Stamatis, O., Aouf, N.., Richardson, M. A., "3D automatic target recognition for future LIDAR missiles," IEEE Trans. Aerosp. Electron. Syst. **52**(6), 2662–2675 (2016).

[11] Hartigan, J. A.., Wong, M. A., "A K-Means Clustering Algorithm," Appl. Stat. **28**(1), 100 (1979).

[12] Leonard, K.., Peter, R., Finding groups in data: an introduction to cluster analysis., John Wiley Sons (2009).

[13] Ester, M., Kriegel, H.-P., Sander, J.., Xu, X., "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," Proc. KDD, 226–231, AAAI Press (1996).

[14] Dempster, A. P., Laird, N. M.., Rubin, D. B., Maximum Likelihood from Incomplete Data via the EM Algorithm, J. R. Stat. Soc. Ser. B **39**(1) (1977).

[15] Yuhan Fang., Ruojing Jiang., Chenguang Li., "Learning a mixture of sparse models by EM algorithm for object clustering," 2015 4th Int. Conf. Comput. Sci. Netw. Technol., 594–597, IEEE (2015).

[16] Song, J., Xiao, L.., Lian, Z., "Boundary-to-Marker Evidence-Controlled Segmentation and MDL-Based Contour Inference for Overlapping Nuclei," IEEE J. Biomed. Heal. Informatics **21**(2), 451–464 (2017).

[17] Zafari, S., Eerola, T., Sampo, J., Kalviainen, H.., Haario, H., "Segmentation of Overlapping Elliptical Objects in Silhouette Images," IEEE Trans. Image Process. **24**(12), 5942–5952 (2015).
[18] Mery, D., Riffo, V., Zscherpel, U., Mondrag??n, G., Lillo, I., Zuccar, I., Lobel, H.., Carrasco, M., "GDXray: The Database of X-ray Images for Nondestructive Testing," J. Nondestruct. Eval. **34**(4), 1–12 (2015).
[19] Horn, R. A., "The Hadamard product," Proceeding Symp. Appl. Math., 87–169 (1990).
[20] Bay, H., Ess, A., Tuytelaars, T.., Gool, L. Van., "Speeded-Up Robust Features (SURF)" (2008).
[21] Mikolajczyk, K.., Schmid, C., "A performance evaluation of local descriptors," 2003 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition, 2003. Proceedings. **2**, II-257-II-263, IEEE Comput. Soc.
[22] Fischler, M. a., Bolles, R. C., "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Commun. ACM **24**(6), 381–395 (1981).