A Review of Automated Image Understanding within 3D Baggage Computed Tomography Security Screening

Andre Mouton[a,*], Toby P. Breckon[b]

[a] *School of Engineering, Cranfield University Bedfordshire U.K.*
[b] *School of Engineering and Computing Sciences, Durham University, U.K.*

Abstract

Baggage inspection is the principal safeguard against the transportation of prohibited and potentially dangerous materials at airport security checkpoints. Although traditionally performed by 2D X-ray based scanning, increasingly stringent security regulations have led to a growing demand for more advanced imaging technologies. The role of X-ray Computed Tomography is thus rapidly expanding beyond the traditional materials-based detection of explosives. The development of computer vision and image processing techniques for the automated understanding of 3D baggage-CT imagery is however, complicated by poor image resolutions, image clutter and high levels of noise and artefacts. We discuss the recent and most pertinent advancements and identify topics for future research within the challenging automated image understanding for baggage security screening CT.

Keywords: Baggage screening, Automated image understanding, Dual-energy computed tomography, Computer vision

*Principal Corresponding Author
Email address: amouton@bath.ac.uk;andremouton.email@gmail.com (Andre Mouton)

July 5, 2015

1 Introduction

Baggage screening plays a central role within the aviation security domain. From September 2014, all new installations of baggage screening systems in Europe will be required to meet the European Civil Aviation Conference (ECAC) Standard 3 screening regulations (the highest standard for European airports). Owing to these stringent regulations, 3D X-ray Computed Tomography (CT), which has enjoyed much success in medical imaging, has been introduced to the security-screening domain in an attempt to mitigate the limitations of conventional 2D X-ray based scanners (e.g. occlusion, clutter and density confusion) [116, 3]. This has, in turn, led to an increased interest in the development of image-processing and computer-vision techniques to advance the automated analysis of baggage imagery and ultimately change the manner in which baggage scanning is performed by improving efficiency and reducing the frequency with which manual inspection (by humans and/or sniffer dogs) is required.

The foremost application of CT within the security-screening domain has been the materials-based detection of explosives [90]. Dual-Energy Computed Tomography (DECT) [48], whereby objects are scanned at two distinct energies, has been shown to provide a more effective means for performing materials-based discrimination based on effective atomic numbers than do the conventional single-energy CT scanners [90, 93, 111]. Owing to this primary explosives detection-based objective of imaging within the aviation-security domain, DECT machines are increasingly becoming the baggage-CT scanners of choice. The primary, non-object recognition-based objective of typical dual-energy baggage-CT scanners, however, coupled with the demand for high throughput, means that 3D baggage-CT imagery typically presents with substantial noise, metal-streaking artefacts and poor voxel resolution and is thus generally of a poorer quality than medical-CT imagery [90] (Fig. 3), which complicates the implementation of image processing and computer vision-based solutions for automated baggage-CT image analysis.

The proprietary nature of deployed baggage-CT scanners makes it impossible to determine exactly what is implemented in practice. Nonetheless, the current certification standards (which do not provide clear regulations for automated image analysis systems beyond explosives detection)[1], the information available in published patents and literature and the existence of ongoing government-funded initiatives such as the ALERT project [23, 22],

[1] As defined by governmental organisations such as the Transport Security Administration (TSA), the European Union (EU) and the Civil Aviation Administration of China (CAAC).

suggest that automated baggage-CT image analysis techniques have yet to be widely adopted within current baggage-scanning protocols and the topic remains an open and active research area.

We present a review of the recent research addressing the challenges associated with image processing and computer vision-based solutions for the automated analysis of baggage security-CT imagery. Much of this work has been conducted based on the assumption that the final pipeline will adhere to the traditional automated object recognition or image classification pipeline adopted in the broader computer vision literature (Fig. 1) [94].

Image pre-processing typically involves some form of noise filtering or image enhancement algorithms, designed to ease the implementation of subsequent operators. Segmentation is performed to subdivide the input into meaningful parts or regions. From these regions features may be extracted which essentially encode information characterising the contents of the region which are subsequently passed to a classifier which assigns semantic labels to the regions or the input as a whole. While this is a somewhat crude description of a typical automated image understanding pipeline, it is sufficient for the purposes of contextualising the literature discussed in this review. For a more comprehensive overview of general and specific machine learning concepts in computer vision, the reader is referred to the broad resource of related literature (e.g. [94]).

The individual studies presented here typically address only a single block in this pipeline and the review thus focusses on the following key areas: 1) image quality improvement, whereby denoising and artefact-reduction techniques are used to reduce the characteristically high levels of noise and artefacts in baggage-CT imagery (Section 3.2); 2) dual-energy techniques, which are used to extract information related to the chemical characteristics of the materials present in a scan (Section 4); 3) unsupervised segmentation of objects in cluttered volumetric-CT imagery (Section 5); 4) 3D object classification for threat and contraband detection (Section 6) and 5) Threat Image Projection (TIP), whereby artificial threat items are automatically inserted into routine scans for online operator training and performance monitoring (Section 7). Finally, given the early developmental stages of the majority of these techniques, their integration into existing scanning protocol(s) is not necessarily obvious and may be viewed as a separate problem, which is beyond the scope of this review.

2 3D Baggage CT Imagery

There exist several significant differences between typical medical-CT data and that encountered in the aviation-security domain. These differences mean that automated techniques such as segmentation and classification, which have been successfully applied to medical imagery, are not guaranteed to be met with the same degree of success when applied to baggage-CT data. The most pertinent of these differences are discussed below.

Image quality: The nature of dual-energy baggage-CT scanners and the demands for higher scan speeds in the aviation-security domain (compared to the medical domain), lead to compromises in image quality in terms of noise and resolution [90]. While sub-millimetre isotropic resolutions in all three dimensions are common in medical-CT scanners [2,1] (Fig. 2), typical volumetric baggage-CT data is characterised by comparatively low anisotropic voxel resolutions (Fig. 3). Anisotropic voxel resolution and poor resolution in the axial plane in particular are additionally known to compound the effects of image noise and artefacts [53].

***A priori* information:** In the medical domain, *a priori* knowledge related to the geometric properties, the X-ray attenuation characteristics and the spatial relations of the anatomical structures being scanned exists. The exploitation of such *a priori* knowledge allows for the development of algorithms designed or fine-tuned for particular tasks or anatomical structures [53]. In contrast, the contents of baggage scans are unknown prior to scanning and exhibit considerable variability in shape, size, material and spatial context (Fig. 3), making the fine-tuning of algorithms significantly more challenging.

Image complexity: In addition to the availability of *a priori* knowledge, most medical CT scans exhibit relatively low degrees of complexity and clutter (i.e. they are fairly homogeneous). Checked baggage, on the other hand, is generally tightly packed and thus extremely cluttered/complex (Fig. 3 (b)) with no *a priori* knowledge of the number of objects in the bag. It is well documented that complexity and clutter negative impact human and computer detection rates [90].

Poor-resolution, image complexity and a lack of prior knowledge thus sets baggage-CT imagery apart from medical-CT data. Computer-vision and image-processing techniques designed for baggage-CT applications are required to be independent of the number of objects in an image as well as the composition of these objects, making their development particularly challenging.

3 Image Quality Improvement

Based on the suboptimal performance of early works in the presence of artefacts and noise [34, 67], image denoising and artefact removal have conventionally been considered mandatory precursors to automated image understanding tasks within baggage-CT screening and have demonstrated quantitative improvements in tasks such as 3D object segmentation [42, 72] and classification [71].

The origins, impact and removal of image noise and artefacts in baggage-CT imagery have been discussed extensively in prior works [75–77, 74]. An overview of the most important observations made in these prior studies is presented here and summarised in Table 1.

3.1 Denoising

Although the topic of digital-image denoising has been studied extensively, the denoising of complex volumetric baggage-CT imagery, has received comparatively little attention.

Zhou *et al.* [117] use image enhancement to remove background noise and improve the resolution of 2D (cross-sectional) baggage-CT imagery. Initial noise removal is performed by thresholding. A second threshold is used to subdivide the denoised image into an upper and lower image, containing brighter and darker regions respectively. The two intensity thresholds are computed as scalar multiples of the mean intensity of the input images (termed Alpha-Weighted Mean (AWM) thresholds) and are chosen empirically. The upper and lower sub-images are enhanced by intensity clipping and Histogram Equalisation (HE) [94] respectively. The final, enhanced CT image is computed as the summation of the enhanced sub-images. Performance analysis in the study focusses on the improvement in image contrast and little mention is made regarding the effectiveness of the denoising stage of the technique (particularly from a quantitative perspective). Furthermore, the images used in the study are largely free of streaking artefacts and the efficacy of the method in the presence of metal artefacts is thus unclear.

There is evidence in the medical literature that simple denoising operations can significantly improve the quality of CT images [45, 87]. Mouton *et al.* [76] have thus evaluated the efficacy of several popular denoising techniques when applied to low-quality, baggage-CT imagery. In particular, the denoising performance of the aforementioned baggage-CT approach of Zhou *et al.* [117,116] is compared to anisotropic diffusion [81]; Total Variation (TV) denoising [84]; bilateral filtering [114]; translation invariant wavelet shrinkage [21] and Non-Local Means (NLM) filtering [17]. Extensive qualitative and quantitative anal-

yses indicate that wavelet thresholding and NLM filtering produce the best results. The benefits of simple denoising as a pre-processing step for 3D object classification in baggage-CT imagery are demonstrated further in a recently developed end-to-end classification pipeline (see Table 4), where NLM filtering produces comparable classification results to more complex artefact reduction techniques [74] at significantly lower processing times.

3.2 Metal Artefact Reduction (MAR)

The most widely implemented CT-reconstruction technique is the analytical Filtered Back Projection (FBP) algorithm, where transformations between the sinogram (projection) and image domains are performed using the forward and inverse Radon transforms respectively [28]. FBP yields fast and accurate reconstructions of the attenuation function for ideal (or near ideal) projections which contain a sufficient number of projection samples and low degrees of noise, beam hardening and any other imperfections [45]. When the projections are incomplete for a certain range of arguments, however, the FBP reconstructions may become corrupted by artefacts. In reality, projections are only approximations of the ideal case. While FBP generally produces satisfactory reconstructions when these approximations are small (as is often the case), when the errors become large the reconstructions become corrupted by artefacts [9].

Metal objects in particular cause significant artefacts in CT images. In an extensive simulation study, De Man *et al.* [60] cite beam hardening (the preferential attenuation of low-energy photons in a polychromatic X-ray beam), scattered radiation, Poisson noise (resulting from the quantum nature of X-ray photons) and the exponential edge-gradient effect (trans-axial non-linear partial volume effect) as the predominant causes of metal-streaking artefacts in high resolution 2D fan-beam CT images.

The problem of Metal Artefact Reduction (MAR) in CT has been widely studied with over 100 publications in the last 10 years. The vast majority of MAR-based CT literature is found in the medical domain and techniques typically fall into one of three categories: 1) sinogram (or projection) completion methods, whereby artefact correction is performed by replacing corrupted projection data prior to CT reconstruction [50, 69, 115, 74]; 2) iterative methods, whereby iterative reconstruction techniques are used to generate superior quality reconstructions [107, 39, 92] and 3) hybrid methods, using combinations of (1) and (2) [58, 13, 61]. While many of these published techniques claim substantial improvements to previous methods, these claims are often based on limited comparisons. The develop-

ment of novel MAR techniques or the evaluation of existing, medical MAR techniques on baggage-CT imagery are limited [75, 74, 77, 52].

Mouton *et al.* [77] present a review of state-of-the-art MAR techniques, drawn predominantly from the medical-CT literature. The review is supported by an experimental comparison of 12 techniques using a software generated medical phantom, clinical CT data and for the first time baggage-CT imagery. Two important observations regarding the efficacy of the state-of-the-art medical MAR techniques when applied to baggage-CT imagery are highlighted: 1) medical MAR techniques relying on prior information suffer a decline in performance due to the increased difficulty in generating accurate priors and 2) the widely implemented sinogram-completion-based approaches are sensitive to input parameters, require manual tuning and result in a characteristic blurring of the regions surrounding high-density objects.

Mouton *et al.* [74] present an extension to their earlier intensity-limiting sinogram completion-based MAR algorithm [75] developed specifically for baggage-CT images containing multiple metal objects. A novel weighting scheme is proposed, whereby the intensities of the MAR-corrected pixels are modified based on their spatial locations relative to the metal objects. Pixels falling within straight-line regions connecting multiple metal objects are subjected to less intensive intensity-limiting, thereby compensating for the characteristic dark bands occurring in these regions [77]. The study demonstrates an improvement (relative to [75]) in the restoration of the underestimated intensities occurring in the regions connecting multiple metal objects in both medical and baggage-CT data but still suffers from the aforementioned characteristic blurring.

Karimi *et al.* [52] present a dedicated baggage-CT MAR technique which is the first to successfully incorporate prior images in the security-screening domain. Prior-images are constructed as solutions to constrained numerical optimisation problems. Particularly, a regularised Weighted Least Squares (WLS) error is minimised, where the regularisation is performed via the TV norm. Artefact reduction is predominantly achieved via the weighting scheme (which is chosen as to de-emphasises metal) and the chosen constraint (which exploits the fact low-frequency metal artefacts are caused by beam hardening and scattered radiation). To reduce computational overhead, the size of the convex problem is decreased by solving for a smaller image (reduced in size by a factor of four in each dimension). The sinograms are thus filtered and downsampled by a factor of four in views and samples. The employment of this so-called 'miniature image' is shown to lead to a

reduction in reconstruction time by a factor of 16^3. Once the miniature image has been constructed, a second miniature image is constructed in a similar manner but ignoring the aforementioned weights and constraints. This second image represents the original, artefact corrupted image in the same coordinate space as the first miniature image. An artefact-only miniature image is then computed as the difference between the two miniature images and upsampled to the original dimensions, yielding the so-called prior image. The sinogram of this prior image is computed and used to guide the replacement of the metal trace in the original corrupted sinogram (using a standard interpolation-based approach). The technique is shown to outperform the linear-interpolation-based approach of Kalender *et al.* [49] as well as a more recent iterative projection replacement method [102] - particularly in terms of preservation of image details. As with all interpolation-based approaches, some degree of blurring is observed. It is also worth noting that experiments were performed using medical-grade imagery which is not representative of that encountered in the aviation security domain. Nonetheless, this study represents the current state-of-the-art in sinogram-completion based MAR in the baggage-CT domain.

Iterative reconstruction techniques [107] are known to produce CT images with fewer artefacts [77]. Despite the development of optimised approaches such as Ordered Subset Expectation Maximisation (OSEM) [46], the Row-Action Maximum Likelihood Algorithm (RAMLA) [16], Model-Based Iterative Reconstruction (MBIR) approaches [113], Iterative Coordinate Descent (ICD) optimisation [14,95], Block-Iterative (BI) modifications [18] and numerous hybrid methods [58, 13, 61], high computational cost remains the major factor preventing the universal implementation of such techniques in commercial CT machines (even in the medical domain where the demand for high throughput are not paramount). Nonetheless, the Awareness and Localization of Explosives-Related Threats (ALERT) initiative (tasked with promoting academic and third party research in security-screening) [22], has investigated the role of non-iterative and iterative reconstruction techniques in explosives detection in baggage-CT imagery.

In particular, nine independent medical research groups were tasked with developing tomographic reconstruction algorithms to improve image quality and explosives detection in baggage-CT imagery. Of these nine groups, eight used raw projection data to directly develop reconstruction techniques, while the ninth group was tasked with developing simulation tools to mitigate the computational expense of complex reconstruction techniques. The initiative has led to the development of several novel/modified reconstruction algo-

rithms based on iterative-reconstruction techniques, sinogram pre-processing, modifications to the FBP process and dual-energy decomposition techniques. These techniques (which employ both single-energy and dual-energy techniques) are shown to offer varying degrees of improvements in image quality (in terms of metal artefact reduction and contrast enhancement). The degree of quality improvement is, however, shown to correlate with the complexity of the reconstruction technique (i.e. better image quality comes at an increased computational expense). The reconstruction techniques are further shown to improve subsequent explosive detection rates. This is largely attributed to the associated reduction in streaking artefacts and the improvements in image contrast. Similarly to the MAR study of Karimi *et al.* [52], the initiative employed medical-grade CT imagery.

3.3 Summary

The overwhelming consensus in the literature is that iterative-reconstruction techniques provide superior image quality to conventional FBP reconstruction, particularly in terms of artefact reduction. Furthermore, despite the broad range of existing medical MAR techniques, these perform comparatively poorly in the security-screening domain. This indicates that future baggage-CT systems will benefit from improved CT reconstruction as opposed to the development of further sinogram completion-based MAR techniques. In terms of baggage-CT image quality, the following areas are outlined as directions for future work:

- The development of iterative-reconstruction techniques, with a particular focus on minimising processing times.
- An exhaustive optimisation of the MAR parameter space and/or the development of techniques to automatically determine optimal MAR algorithmic parameters to aid the generation of accurate prior images.
- The evaluation of the state-of-the-art prior-based techniques of Karimi *et al.* [52] on CT data that is representative of that encountered in the security-screening domain (as opposed to medical-grade CT data).

4 Dual-Energy Computed Tomography (DECT)

Explosive Detection Systems (EDS) [90] are currently the only commercially approved application of CT within the aviation security infrastructure. Here, we review various appli-

cations of DECT (including, but not limited to EDS) within the aviation security-screening domain. The most prominent recent contributions in this domain are summarised in Table 2. The reader is referred to the work of Singh [90] and Smith *et al.* [93] for more extensive reviews of the particular role of EDS within the baggage security screening context.

Single-energy CT systems produce reconstructions representative of the Linear Attenuation Coefficients (LAC) of the object under investigation. That is to say, the greyscale intensities (i.e. CT numbers, in Hounsfield Units (HU)) in the CT image are dependent on the LAC of the scanned object. Consequently, it is challenging to distinguish between materials with similar LACs. In contrast, Dual-Energy CT (DECT) techniques, whereby attenuation data is captured using two distinct X-ray spectra, offer a means for characterising the chemical composition (e.g. atomic number and/or density) of the material under investigation based on its response under these different spectral conditions. DECT techniques typically fall into one of three categories [47]: 1) post-reconstruction techniques; 2) pre-reconstruction techniques and 3) iterative-reconstruction techniques.

4.1 Post-Reconstruction Techniques

Post-reconstruction (or image-based) DECT techniques, although the most straightforward and widely used techniques in medical imaging [43, 48], have traditionally not been considered in the baggage-CT domain. This is most likely due to their susceptibility to artefacts in the reconstructed images [88]. Only very recently has the Dual-Energy Index (DEI) [48] (a post-reconstruction DECT metric that offers a crude estimate of the effective atomic number of a material) been used for the segmentation of 3D baggage-CT imagery (see Tables 3 and 4).

4.2 Pre-Reconstruction Techniques

Alvarez and Macovski [5] modelled the total attenuation of X-rays is modelled as a linear combination of the photoelectric absorption and Compton scattering coefficients [26] using a non-linear polynomial approximation of the polychromatic measurement models. This early work has formed the basis for a broad range of techniques known as *basis material decomposition* methods [20, 51, 78, 111].

DECT imaging relies on the energy dependence of the interaction of X-ray photons with matter. Within a photon energy range of approximately 30 keV to 200 keV, these interactions are known to be dominated by the photoelectric effect and Compton scattering [5].

The dual-energy decomposition problem is to determine the Compton scatter and the photoelectric absorption coefficients of the material from the measured high and low-energy projections. Alternatively, the attenuation coefficients for any material may be expressed as a linear combination of the coefficients of two basis materials, provided that the two chosen materials are sufficiently different in their atomic numbers (and hence in their Compton and photoelectric coefficients) [51]. The decomposed dual-energy data may then be used to compute the effective atomic numbers and electron densities of the materials present in a scan [111] (this is the basis for traditional explosives detection systems in aviation security-screening [90]).

While basis material decomposition may be solved by direct approximation [33, 15], the more popular approach is to approximate the relationship between the high and low-energy projections and a set of decomposed projections as polynomial functions. While various techniques have been presented to achieve these approximations, those based on Look-Up Table (LUT) procedures [20, 51] have been shown to be faster, less sensitive to the numerical procedure and less susceptible to noise than the original Newton-Raphson approximation method of Alvarez and Macovski [5]. The main drawbacks of the majority of pre-reconstruction DECT techniques are their dependence on an intensive calibration procedure and a tendency to produce large approximation errors [78].

Although the majority of dual-energy decomposition algorithms have been motivated by medical applications, there has been a growing interest in the application of similar techniques for explosives detection in baggage-screening systems [111,78,88]. The fundamental objectives of dual-energy decomposition in the medical and security-screening domains differ. In the medical domain, the primary goal is to generate high-quality images to facilitate the diagnostic procedure, while in the security-screening domain, the main objective is the determination of the atomic properties of the objects in a scan to allow for materials-based explosives detection.

Traditional dual-energy decomposition using basis materials requires relatively accurate estimates of the combinations of basis material thicknesses to use in the calibration procedures. This complicates its implementation in the security-screening domain, where the range of possible materials encountered is much broader and more unpredictable than those encountered in medical scans (Section 2). Ying *et al.* [111] have further highlighted several limitations of traditional medical dual-energy decomposition methods (e.g. [5,51,20]) when used for explosives detection. These limitations include: high polynomial approximation

errors ($> 200\%$ for [5]), caused by the large dynamic range of the photoelectric coefficients (resulting from the broad spectrum of materials encountered in baggage scans); a lack of boundary constraints in dual energy decomposition; image artefacts and X-ray spectral drifts.

Despite these fundamentally differing objectives and the increased complexity of baggage data, the dual-energy techniques designed specifically for baggage screening have been fairly similar to their medical counterparts [111, 78]. Naidu *et al.* [78], present a dual-energy decomposition approach using a multi-step fitting procedure and the iso-transmission method of [20]. In contrast to prior medical studies [20], the required calibration data is generated using simulated low and high-energy spectra (as opposed to the measured spectra). The literature indicates that this is a more commonly adopted approach in the security-screening domain [78, 111, 112]. Note, however, that although simulated spectra are used, CT images of known materials, obtained on the CT scanner under investigation, are generally used to calibrate the simulated spectra [111].

Ying *et al.* [111] propose a pre-reconstruction basis material decomposition method, whereby the photoelectric and Compton-equivalent sinograms are obtained by solving a constrained least squares minimisation problem. Additionally, techniques for adaptive scatter correction [38], destreaking by nonlinear filtering and real-time image-based correction for X-ray spectral drifts are incorporated into the proposed framework. The so-called Constrained Decomposition Method (CDM) is evaluated on simulated and real baggage-CT data (containing known explosive and non-explosive materials). CDM is shown to yield numerically stable and physically meaningful solutions for the photoelectric and Compton-equivalent line integrals and a significant reduction in the approximation and boundary constraint errors common to earlier methods [5, 20].

4.3 Iterative-Reconstruction Techniques

Semerci and Miller [88] present a polychromatic DECT algorithm, tailored particularly for the detection of objects in cluttered baggage-CT images. The availability of *a priori* information regarding the Compton scatter and photoelectric absorption coefficients of the objects of interest is assumed. This prior information is incorporated (as a series of constraints) into a variational framework, using the Levenberg-Marquardt algorithm [62] for minimisation. The photoelectric and Compton scattering parameters are modelled as the superposition of a parametrically defined object of interest and a non-parametric

background. The object model contains a geometric component (equal in the photoelectric and Compton images) and a contrast component (specific to the photoelectric and Compton images) and is based on a parametric level-set representation of the characteristic function of the object. The proposed approach provides simultaneous solutions to the problems of object detection and background reconstruction. Tested on simulated data only, the algorithm is shown to successfully detect, locate and determine the geometric characteristics of objects of interest, while simultaneously producing reasonable background reconstructions.

4.4 Computation of the Effective Atomic Number

In the baggage-screening context, the predominant application of DECT has been the determination of effective atomic numbers Z_{eff} and electron densities to be used for materials-based detection of explosives [90]. The detection of explosives in baggage scans is essentially based on two assumptions [83]: 1) the majority of explosives may be characterised as organic substances with effective atomic numbers of approximately $Z_{\text{eff}} = 8$ and densities of $1.15 \leq \rho \leq 1.85$ g/cm^3 and 2) the majority of (non-metallic) innocuous items (e.g. clothing, toiletries, books etc.) have densities of $\rho < 1.0$ g/cm^3. Fig. 4 illustrates Z_{eff} as a function of density for common substances (and several illicit materials) found in packed luggage [32]. Innocuous materials include organic substances (e.g. books, sausages, alcohol, leather, cotton etc.); inorganic substances (e.g. salt, PVC, plastic) and metals (e.g. iron, copper). The illicit narcotics plotted are heroine and cocaine, while the explosives include C4, TNT, Semtex and Detasheet [32]. Such plots have traditionally been used by the US Federal Aviation Association (FAA) to evaluate the detection capabilities of a given scanner. Most importantly, the plot demonstrates that typical explosive materials (as well as illegal narcotics) are easily clustered and hence distinguished from other innocuous organic and/or inorganic materials. Based on these observations, it is theoretically possible to distinguish between illicit and innocuous items by computing the effective atomic numbers and densities of the materials in a scan [83]. Traditionally, this has been achieved via a simple calibration and interpolation procedure [93] (Fig. 5).

A set of reference materials, whose effective atomic numbers and densities span the expected range of the materials of interest, are chosen. The low and high-energy Linear Attenuation Coefficients (LACs) for each of the reference materials are then measured on the scanner under investigation and the relationship between the known Z_{eff} and the mea-

sured LAC ratio (μ_H/μ_L) is approximated by an interpolating polynomial. Z_{eff} for any unknown material may be interpolated from the measured LAC ratio and the approximation polynomial (Fig. 5). This approach assumes a perfect measurement and observation of the constituents of the illicit and innocuous substances [59]. Such ideal conditions are rarely encountered in reality. Furthermore, various innocuous substances that have similar chemical properties to common explosives (e.g. honey and chocolate), are typically not included in the evaluation of the detection capabilities of a system (as they would fall within very similar regions as the explosives in Z_{eff} vs. density plots) [59, 32]. Therefore, even under the assumption of ideal conditions, the discrimination of explosives by the interpolation of the Z_{eff} vs. density curve is at best a crude approximation.

Several more robust DECT-based techniques exist for computing the effective atomic number [111, 93, 88]. The majority of these rely on the decomposition of the low and high-energy data into equivalent Compton scatter and photoelectric absorption coefficients using the methods described in Section 4.2. Given the Compton scatter coefficient a_c and the photoelectric absorption coefficient a_p, the most common approach for computing Z_{eff} is given by:

$$Z_{\text{eff}} = K' \left(\frac{a_p}{a_c} \right)^{\frac{1}{n}} \tag{1}$$

where K' and n are constants [5]. This approach requires two separate reconstructions to obtain the photoelectric and Compton images. Ying *et al.* [111] present a faster approximation of Z_{eff} by eliminating the computation of the Compton reconstruction image:

$$Z_{\text{eff}} = K \left(\frac{a_p}{a_{hct}} \right)^{\frac{1}{n}} \tag{2}$$

where K and n are constants and a_{hct} is the CT number of the scanned materials (obtained from the high-energy CT image). The division of the photoelectric image by the high-energy CT image is also shown to reduce the partial volume effect in the Z_{eff} image.

Eger *et al.* [31, 30] demonstrate superior explosives detection using machine learning and Multi-Energy Computed Tomography (MECT) techniques. MECT is claimed to provide superior characterisation of the chemical composition of the materials in a scan relative

to conventional DECT. As discussed, the traditionally the photoelectric and Compton scatter coefficients of a material are estimated via linear projections of the material attenuation curve onto two fixed energy-dependent curves associated with the photoelectric and Compton scattering processes. Eger *et al.* [31, 30] propose a linear decomposition of the material attenuation using a different set of basis functions (i.e. not the photoelectric and Compton models). The resulting projection coefficients are used as features within a Support Vector Machine (SVM) [101] classification framework used for explosives detection. and shown to outperform the traditional photoelectric and Compton coefficients (particularly when more than two features are used). The studies suggest improved detection performance relative to conventional dual-energy X-ray systems.

4.5 Summary

For the most part, DECT techniques may be grouped into one of three categories: post-reconstruction techniques; pre-reconstruction techniques and iterative-reconstruction techniques. Post-reconstruction techniques operate directly on the low and high-energy scans and are the most straightforward and computationally efficient approaches. The literature does however, indicate that the effectiveness of post-reconstruction techniques are limited by artefacts and noise and provide comparatively little discriminative power (compared to more advanced techniques). Pre-reconstruction DECT techniques are the most widely implemented in the security-screening domain. In particular, DECT decomposition allows for the computation of effective atomic numbers which are then used for materials-based explosives detection [83, 90, 93, 111, 78]. Pre-reconstruction techniques are limited by their susceptibility to artefacts and the high computational cost associated with DECT decomposition (two FBP reconstructions required per image).

Similarly to single-energy CT, DECT using iterative-reconstruction techniques results in superior quality images in terms of artefacts. Improved performance, however, comes at an increase in computational demand. Nonetheless, iterative reconstruction is gaining popularity with the ever increasing computational power of modern hardware. Recent studies [31, 30] indicate superior materials-based discrimination using Multi-Energy CT (MECT), although further studies are limited. Iterative reconstruction and MECT techniques are most likely to be fruitful avenues for future work.

5 Segmentation

Image segmentation is a core problem in computer vision and has been the source of an extensive resource of literature. In the context of baggage screening, the objective of segmentation is typically pragmatic - that is to say, its implementation is intended to aid the subsequent detection of threats. Despite its importance, prior work addressing the automated segmentation of unknown 3D objects from low-resolution, cluttered baggage-CT imagery is limited. An overview of existing studies is presented in Table 3.

The differences in the quality and nature of security-screening imagery compared to medical imagery have limited the success of medical-segmentation techniques in this domain [64]. Furthermore, the dependence of the majority of the state-of-the-art medical-segmentation techniques on *a priori* information in particular, detracts from their suitability to the security-screening imagery, where the segmentation of multiple, unknown objects is required. This hypothesis has been verified by Megherbi *et al.* [64], who investigated the effectiveness of classical medical-segmentation techniques when applied directly to low-quality baggage-CT scans. Four methods were evaluated on the task of segmenting bottles and handguns from baggage-CT imagery: 1) confidence connected region growing [82]; 2) fuzzy connectedness [98]; 3) the watershed transform [103] and 4) fast marching [89]. Successful segmentation is shown to be dependent on careful per-case parameter tuning. Furthermore, even after parameter optimisation, the effects of image noise (despite pre-filtering), clutter and the lack of prior knowledge result in significantly poorer segmentations than those observed in the medical domain.

The bulk of the prior literature addressing baggage-CT segmentation in particular has its origins in a recent collaborative initiative between the US Department for Homeland Security (DHS) and the Awareness and Localization of Explosives-Related Threats (ALERT) Center of Excellence [23]. The initiative (tasked with promoting academic and third party research in security-screening) led to the development of the following five baggage-CT-segmentation algorithms (two of which have appeared in peer-reviewed publications [109, 42]).

Wiley *et al.* [109] present a 3D region-growing method based on the Stratovan Tumbler medical-segmentation technology [108]. The technique is composed of five stages: 1) definition of a 3D kernel; 2) determination of the kernel movement criteria; 3) seed initialisation; 4) flood-fill and 5) splitting and merging. Optimal results are obtained using a spherical kernel, provided the size of the kernel (determined automatically, based on the amount

of local clutter) is smaller than the object being segmented and larger than any expected holes in its boundary. The movement criteria for a given kernel are determined automatically using a training procedure, whereby an initial criteria are matured by manually improving errant segmentations and adding each improvement to a central training file. A polynomial is fitted to the training points and used to determine the future movement criteria at any voxel. Seed-points are determined according to a voxel ordering method which ensures that large kernel sizes, high intensity voxels and voxels in the centres of objects are considered first. The 3D kernel traverses a volume in a flood-fill manner provided the traversed voxels satisfy the movement criteria. Composite objects are represented by hierarchical tree-like models. In particular, objects are initially segmented into multiple parts and pairs of segmented parts are merged if their degree of overlap exceeds a threshold. The study demonstrates high-quality segmentations for homogeneous objects and results in good separation of touching objects. Performance deteriorates for low-contrast objects, thin objects and in the presence of artefacts. It is also indicated that high-quality segmentations rely on near isotropic voxel resolutions in all three dimensions.

Song *et al.* (TeleSecurity Sciences, Inc.) [23] present a sequential approach composed of three stages: 1) pre-processing (by 2D bilateral filtering [96]; 2) object segmentation and 3) post-processing. Object segmentation is achieved using a sequential 'Segment-and-Carve' (SC) approach, operating on the principal that *easy* objects should be segmented first. The objects segmented in each stage are *carved* out of the image before proceeding to the next stage. Segmentation is performed using the Symmetric Region-Growing (SymRG) [106] technique - a seedless (i.e. unsupervised) region-growing technique based on a symmetric function and which is invariant to starting conditions. In total, five SC stages are proposed, each targeting objects with different characteristics: 1) homogeneous, bulk objects; 2) homogeneous, medium thickness objects; 3) homogeneous, sheet-like objects; 4) homogeneous, metal objects and 5) heterogeneous objects. Each stage is composed of five steps: 1) binary mask generation by thresholding; 2) mask pre-processing; 3) segmentation by SymRG; 4) boundary correction and 5) object carving. The five steps each require parameter tuning, with parameters differing for each SC stage. On completion of the five-stage sequential SC procedure, the segmented objects from each stage are subjected to extensive post-processing operations to correct for over and under-segmentations. Particularly, object-splitting is performed in four stages: 1) splitting by histogram analysis; 2) splitting by RANSAC; 3) splitting by recursive k-means clustering and 4) splitting by morphologi-

cal opening. Object-merging is performed based on three thresholds: 1) spatial proximity; 2) mean intensity and 3) object type. While the study demonstrates high-quality segmentations for selected objects, the results for complete scans are not presented. The approach is extremely convoluted (with a large parameter set) and optimal performance requires careful parameter tuning.

Grady *et al.* [42] present a graph-based segmentation technique composed of three stages: 1) foreground identification; 2) candidate splitting and 3) candidate refinement. Foreground identification is performed by applying a Mumford-Shah functional [41] to artefact-reduced volumes (obtained by linear interpolation-based MAR [97]), producing labelled volumes (voxels labelled as foreground or background). Connected component analysis is applied to the labelled volumes. Each of the connected components in the foreground are recursively partitioned into candidate segments using the Isoperimetric Distance-Tree (IDT) algorithm [40]. Recursions are driven by a novel Automated QUality Assessment (AQUA) metric, which automatically computes the quality of a segmentation without *a priori* knowledge of the object being segmented. Computational expense is optimised by performing coarse-to-fine segmentation (i.e. the segmentation from the previous level is used as the initial mask for further splitting). High-quality segmentations are demonstrated for challenging cases. Manageable run-times of approximately four minutes per volume (on an Intel Core 2 Duo 2.8 GHz CPU) are presented. Low-density and sheet-like objects present the greatest challenges and it is suggested that superior MAR would be beneficial.

Harvey *et al.* (University of East Anglia) [23], present a technique based on the multi-scale sieves class of algorithms [7, 8]. Sieves function by filtering input signals to remove the intensity extrema at specific scales. In the context of image segmentation, *semantically meaningful* objects are removed at specific (typically higher) scales. The proposed approach is composed of four steps: 1) sieve the input volume to four logarithmically-spaced scales; 2) compute four *channel* volumes; 3) label the channel volumes using Sedgewick's connected component analysis [86] and 4) merge the labelled channel volumes into a single labelled volume. Merging is performed by determining the similarities between the density histograms for each labelled object in each channel volume using the Kolmogorov-Smirnov (K-S) test [55], which computes the probability that the histograms have been drawn from the same distribution. The specific strengths and weaknesses of the approach are not addressed in any great detail [23]. An interesting observation is that due to sieves

segmenting all objects at all scales, at least one channel will always contain a segmentation of an object. Harvey *et al.* [23] thus propose that a more suitable approach (compared to the merging of channels), would be to pass the channel volumes directly into some artificial intelligence system (e.g. a classifier, object detector or salient region detector). The decision to merge the channels was dictated by the specifications of the ALERT initiative [23]. The computational complexity of sieves is approximately $N \log p$, where p is image dependent and is proportional to the number of *flat-zones* (the largest connected components where the signal is constant) in the image.

Feng *et al.* (Marquette University) [23] present a true 3D (as opposed to per-slice) technique which, although not explicitly specified, draws significantly from the automatic segmentation and merging technique of Ugarriza *et al.* [99]. The approach is composed of three stages: 1) seed map generation; 2) adaptive region-growing and 3) merging. Seed maps are generated by locating sufficiently large homogeneous regions in the input volume. Homogeneous regions are determined by thresholding of the Sobel gradient map of the volume [94], while region size is determined by connected component analysis. Seed regions are grown by dynamic region-growing [99], where the region-growing threshold is not constant. To compensate for the variation of intensities within objects (due to CT artefacts), the region-growing threshold is modelled as a non-linear function of the mean intensity of the region. On completion of the region growing, pairs of touching objects (i.e. those sharing a common edge) are merged based on their similarity in a 2D feature space (characterising mean texture and intensity). This merging heuristic is applied recursively. The technique is shown to be sensitive to parameter tuning and susceptible to under-segmentations (occurring in approximately 15% of cases).

Each of the five aforementioned baggage-CT segmentation techniques were developed and evaluated using a fully labelled volumetric baggage-CT data set captured on a single-energy medical-CT scanner with a resolution of $0.98 \times 0.98 \times 1.29$ mm. This data is not representative of the current benchmark in baggage screening, where data is typically captured on dual-energy scanners and are characterised by considerably poorer voxel resolutions. The development of segmentation algorithms for such data has not been considered previously.

Mouton and Breckon [72] have recently presented a materials-based technique for the segmentation of unknown 3D objects from low-resolution, cluttered baggage-CT imagery. Initial coarse segmentations are generated from metal artefact reduced [74] images using

the Dual-Energy Index (DEI) [48] (Section 4.1), thresholding and connected component analysis. The quality of the individual components of the coarse segmentations are determined using a novel random-forest-based evaluation metric (the Random Forest Score (RFS)), trained to recognise high-quality, single-object segments. Low-quality individual object segments are subjected to an object-partitioning procedure which splits fused objects at automatically-detected regions using morphological operations and connected component analysis. In a comparative evaluation, the technique is shown to produce segmentations comparable to the state-of-the-art [40] at a significant reduction in processing time. A lack of ground-truth information is highlighted as a limitation in the performance evaluation of the study.

5.1 Segmentation Summary

The dependence of medical-segmentation techniques on *a priori* knowledge detracts from their suitability in security screening where unsupervised segmentation of multiple, unknown objects is required [64].

The ALERT segmentation initiative [23] has resulted in the bulk of the baggage-CT segmentation literature. Although leading to five dedicated baggage-CT segmentation techniques, evaluations are performed on medical-grade CT data which is not representative of that encountered in the security-screening infrastructure. Thin sheet-like objects, low-density objects and metal streaking artefacts are shown to negatively impact the performance of each of the techniques. The development of dedicated techniques for handling these objects is thus suggested as an area for future work. Furthermore, due to the significant differences in the quality of medical and baggage-CT imagery (Section 2), it is important that future work considers security-grade CT imagery.

Mouton and Breckon present a dual-energy materials-based 3D baggage-CT segmentation technique, which is the first to consider low-resolution, cluttered data obtained on a true baggage-CT scanner (see Table 3). Although the computational efficiency of the technique makes it well-suited to the security-screening domain, it is based on a crude material representation offered by the DEI. Future work is likely to benefit from a more accurate materials characterisation using pre-reconstruction and iterative reconstruction DECT techniques (Section 4).

Finally, we emphasise the importance of ground-truth data for reliable performance evaluation and suggest a meticulous approach to future data-gathering procedures, ensuring

accurate and detailed documentation of the contents and properties of scans.

6 Classification

In this section we review the literature specifically addressing the automated classification of 3D objects in baggage-CT data using machine learning techniques. We do not consider the detection of explosives, as this is traditionally performed using dual-energy techniques and has been addressed in Section 4. Table 4 presents an overview of the proposed classification methodologies.

While the algorithmic details of deployed baggage-CT scanners fall under vendor proprietary and are thus not publicly available, the literature addressing the automatic classification of objects within this domain is surprisingly limited. Chen *et al.* [11] address the classification of pistols in Dual-Energy CT (DECT) imagery. DECT decomposition is performed using High-Low (HL) energy curves and look-up tables constructed from 28 calibration elements. In this way the effective atomic numbers and electron densities of the scans are determined. The problem is reduced to 2D by considering only the central, cross-sectional slice of the volume in classification. Classification is performed by boosting 2D Haar-like features [104]. The study considers simplified (unrealistic) data containing only handguns with no background clutter, noise or artefacts and does not present any experimental results. Further work by the same author [12] proposes a methodology for the detection of planar materials within CT-baggage imagery using a 3D extension to the Hough transform [6].

Megherbi *et al.* [66, 67] present a comparison of classifier-based approaches using volumetric shape characteristics for the classification of pre-segmented 3D objects in cluttered baggage-CT imagery. Various combinations of three shaped-based descriptors (3D Zernike descriptors [79]; Histogram-of-Shape Index (HSI) [29] and a combination of the two) and five classifiers (Support Vector Machines (SVM) [10]; neural networks [105]; decision trees [85]; boosted decision trees [25] and random forests [25]) are compared. Correct classification rates in excess of 98.0% are achieved on a limited dataset using the HSI descriptor with an SVM or random-forest classifier. The reduction of image noise and/or artefacts are not considered.

Flitton *et al.* [35] compare the classification performance of four 3D interest-point descriptors of varying complexities (Density Histograms (DH) [35]; Density Gradient Histograms (DGH) [35]; 3D SIFT [34] and 3D RIFT [56, 57]) sampled at 3D SIFT interest points [34].

It is shown that the simpler density statistics-based descriptors (DH and DGH) outperform the more complex 3D descriptors (SIFT and RIFT) in the classification of known rigid objects within realistic baggage-CT imagery. The comparatively poor performance of the SIFT and RIFT descriptors are attributed to the low, anisotropic voxel resolution and high level of noise and artefacts characteristic to this type of imagery.

Extending upon their earlier work [35], Flitton *et al.* [37] investigate the suitability of the visual codebooks [91] for the detection of 3D threat objects in volumetric baggage-CT imagery. Combinations of four 3D descriptors (DH, DGH, SIFT and RIFT) and three codebook assignment methodologies (hard, kernel and uncertainty) are compared within a 3D sliding-window classification framework. Optimal correct classification rates ($\sim$ 89% for bottles; $\sim$ 97% for handguns) are obtained using an uncertainty assignment protocol [100] in conjunction DH descriptors [35] sampled at 3D SIFT interest points [34]. The impact of the classifier type, the clustering method and the interest point detection protocol are not considered. Poor resolution, image noise and metal-streaking artefacts are shown to negatively impact the efficacy of the 3D descriptors, leading to a decline in classification performance. Explicit image denoising and artefact reduction are, however, not considered.

A 3D extension to the hierarchical visual cortex model for object classification has demonstrated success in 3D object classification in baggage-CT imagery [36]. The approach outperforms a traditional codebook approach with correct classification rates in excess of 95% and low false-positive rates. Performance is hindered by image noise, streaking artefacts and background clutter. Denoising and artefact reduction are again not considered. Furthermore, an extremely high computational cost is associated with the construction of the cortex model and the adoption of a sliding-window-like approach to classification (as pre-segmentation is not performed).

Mouton [73] demonstrates an improvement over the 3D visual cortex model in terms of classification accuracy and processing time using a codebook approach based on Extremely Randomised Clustering (ERC) forests [70], a dense feature sampling strategy [80] and an SVM classifier [101]. In particular, correct classification rates in excess of 98% and false positive rates of less than 1%, in conjunction with a reduction of several orders of magnitude in processing time are demonstrated for the 3D object classification in manually segmented data. Similarly to related studies [34–36, 67, 37], evaluation is limited to two object classes. It is suggested that the true benefits of random-forest-based classification

[24] are thus not exploited and it is likely that such techniques will be of increased value in multiclass problems.

Each of the aforementioned techniques perform classification in a 3D sliding-window fashion or rely on manual pre-segmentations and do not consider techniques for the reduction of image noise and/or artefacts. Mouton [71] addresses these limitations by combining various methodologies for noise and artefact reduction, unsupervised segmentation and classification. More particularly, noise and metal streaking artefacts are reduced by NLM filtering and distance-driven MAR [17, 74]; 3D segmentation of the data is performed via materials-based dual-energy techniques [72] and the segmented data is classified using ERC-forest codebooks [73]. Correct classification rates in excess of 97% with false-positive rates of less than 2% are obtained at low computational costs for the classification of two object classes (handguns and bottles) in realistic baggage-CT imagery (Section 2). Perhaps the most significant observation is the negligible performance gains (in terms of classification accuracy) offered by the computationally expensive MAR over simple NLM filtering.

Finally, Chermark *et al.* [19] present a method for liquid detection within 3D baggage-CT imagery as a precursor for image-based explosives detection. The proposed approach is based on two stages of geometric fitting which exploit the geometric properties of liquids within 3D space. Firstly, a 3D plane is fitted to the horizontally-orientated surface of the liquid based on the self-levelling property of contained liquids. A least squares elliptical fitting procedure is then performed to emphasise the shape of the liquid surface at a given level within the container. The proposed approach exploits the fact that, given the cluttered nature of baggage-CT imagery, the occurrence of perfectly aligned horizontal planes are statistically unlikely and are thus indicative of liquid presence. The proposed methodology, which is based purely on the geometric properties of liquids and robust geometrical shape detection and thus does not require supervised training, demonstrates a correct liquid detection rate of 8598%.

6.1 Classification Summary

Despite the limited work addressing the problem of 3D object classification in low-resolution baggage-CT imagery, numerous studies have demonstrated encouraging results using traditional feature-based discriminative classification models.

A common limitation of each of the studies reviewed here is data related. Particularly,

evaluations are typically limited to one or two object classes and the particular target objects considered (usually handguns and/or bottles) are relatively easy to classify. In order for a baggage-CT classification tool to be of value in industry, a broad (and ultimately universal) range of materials must be detectable. Despite the likely challenges in data gathering arising due to the sensitivity of security-screening data, the most important direction for future work is thus an extension to multiclass problems.

7 Threat-Image Projection (TIP)

Threat Image Projection or TIP is a software tool which allows for the automated insertion of realistic threat items into routine baggage imagery with the objective of assessing and monitoring the performance of screening personnel [90,44,27]. TIP techniques typically fall into one of two categories: 1) Fictional Threat Image (FTI) projection and 2) Combined Threat Image (CTI) projection [27]. The former inserts an image of a threat (obtained from a 'threat image database' for example) into a real passenger bag, while the latter inserts an entire threat-containing bag into the screening pipeline. CTI is only feasible in scenarios where the screeners are unaware of the actual bags being scanned and is thus of limited practical use. In this review, we focus on recent developments in FTI-based projection.

The use of TIP technology is currently limited to 2D X-ray baggage-screening systems [90, 44, 27]. 2D TIP is relatively straightforward, requiring only a simple superimposition of the threat item onto the existing image. 3D TIP is considerably more challenging, as image clutter complicates the task of determining a viable insertion location that does not violate the existing contents of the bag (see Table 5 for an overview of techniques). Furthermore, in order for the insertion to appear realistic, metal streaking artefacts need to be incorporated [63].

Megherbi *et al.* [63, 65] present a fully-automated three-stage approach for realistic threat image projection in cluttered 3D baggage-CT imagery: 1) void determination; 2) threat insertion location determination and 3) Metal Artefact Generation (MAG).

Void determination is performed via intensity thresholding as low-density objects are usually not visible to human operators due to the high dynamic range of baggage-CT imagery [117]. The optimal insertion location is found by examining the cubic neighbourhoods of every voxel in the volume and selecting that with the highest ratio of void to non-void voxels (provided this ratio is larger than a predefined threshold). A novel Metal

Artefact Generation (MAG) approach [63,65], based on the inverse task of Metal Artefact Reduction (MAR) [77], is presented to incorporate streaking artefacts into the insertion. Drawing from sinogram completion-based MAR [77], artefacts are generate by corrupting the metal trace in sinogram space using an empirically designed function. The corrupted sinograms are back-projected (via Filtered Back-Projection (FBP) [28]) yielding the final TIP image. The study demonstrates more realistic 3D TIP than prior works, where metal artefact generation was not considered [110, 27]. Two main limitations are highlighted: 1) high computational cost and 2) frequent violation of low-density objects by the inserted threats due to the thresholding-based void determination. It is thus suggested that future work focusses on computational optimisation and improved void determination (e.g. using geometric constraints and/or 3D spatial reasoning) [63].

8 Conclusions

The introduction of 3D X-ray CT to the aviation security-screening domain has led to an increased interest in the development of automated techniques for the analysis of low-resolution baggage-CT imagery. We have presented a review of this literature focussing on the five most prominent areas: image quality improvement; dual-energy techniques; segmentation; classification and Threat-Image Projection (TIP). Despite the relative novelty of 3D X-ray CT in security-screening, a broad range of techniques have demonstrated clear potential. The development of techniques for artefact reduction [74, 52], dual-energy decomposition [111] and segmentation [109,42] have allowed for successful detection of explosives and threats [90,36] and online operator performance monitoring via fully-automated 3D threat-image projection [63].

Despite these promising results, the majority of existing studies utilise limited and/or unrealistic (e.g. medical-grade or uncluttered) data. Due to the sensitivity of security-screening data, datasets containing images which are truly representative of those encountered in the security-screening domain are limited - both in terms of size as well as variability. These limitations, in addition to low scan speeds and high costs, have meant that baggage-CT scanners and related technologies still fall short of the demands laid out by the aviation security domain. There thus remains a substantial gap between the published literature and commercialisation.

We have highlighted several important areas for future work. The poor quality of baggage-CT data remains one of the foremost factors limiting the successful implementation of

computer-vision and image-processing techniques in this domain. Future work is likely to benefit from the development of computationally-efficient iterative reconstruction techniques. Since explosives detection remains a crucial component of baggage inspection, the recent multi-energy CT and feature-based classification techniques (which have demonstrated improved detections) [31,30] are worth pursuing. Existing 3D object segmentation techniques have used single-energy medical-grade data [109,42,23]. Future work in this regard is likely to benefit from exploiting dual-energy information. Although several studies have demonstrated high threat classification rates [35,36], these are limited to two object classes - an extension to multi-class threat detection is thus necessary. Perhaps the most important direction for future work is the construction of a representative and variable dataset, which would likely benefit all facets of future work.

Acknowledgements

This study was funded under the Innovative Research Call in Explosives and Weapons Detection (2010), sponsored by the Home Office Scientific Development Branch, the Department for Transport, the Centre for the Protection of National Infrastructure and the Metropolitan Police Service. The authors thank Reveal Imaging Technologies Inc. (USA) for additional support.

References

[1] GE Healthcare discovery CT750 HD `http://www.gehealthcare.com/euen/ct/products/`, [Jan. 01, 2012].

[2] Toshiba Medical Systems Corporation Aquilion 32 `http://www.toshiba-medical.co.uk/ct-systems.asp`, [Jan. 01, 2012].

[3] B. R. Abidi, Y. Zheng, A. V. Gribok, and M. A. Abidi. Improving weapon detection in single energy X-ray images through pseudocoloring. *IEEE Transactions on Systems, Man, and Cybernetics*, 36(6):784–796, 2006.

[4] D.G. Altman and J.M. Bland. Measurement in medicine: the analysis of method comparison studies. *The Statistician*, 32:307–317, 1983.

[5] R E Alvarez and A Macovski. Energy-selective reconstructions in X-ray computerised tomography. *Physics in Medicine and Biology*, 21(5):733–744, 1976.

[6] DH Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.

[7] J. A. Bangham, P. Chardaire, C. J. Pye, and P. D. Ling. Multiscale nonlinear decomposition: The sieve decomposition theorem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):529–539, 1996.

[8] J. A. Bangham, J. R. Hidalgo, R. Harvey, and G. Cawley. The segmentation of images via scale-space trees. In *Proceedings of British Machine Vision Conference*, volume 1, pages 33–43, 1998.

[9] J. F. Barrett and N. Keat. Artifacts in CT: Recognition and avoidance. *Radiographics*, 24(6):1679–1691, 2004.

[10] A. Ben-Hur and J. Weston. A user's guide to support vector machines. *Methods in Molecular Biology*, 609:223–239, 2010.

[11] W. Bi, Z. Chen, L. Zhang, and Y. Xing. A volumetric object detection framework with dual-energy CT. In *Proceedings of the IEEE Nuclear Science Symposium Conference Record*, pages 1289–1291, 2008.

[12] Wenyuan Bi, Zhiqiang Chen, Li Zhang, and Y. Xing. Fast detection of 3D planes by a single slice detector helical CT. In *Proceedings of the IEEE Nuclear Science Symposium Conference Record*, pages 954–955, 2009.

[13] F. E. Boas and D. Fleischmann. Evaluation of two iterative techniques for reducing metal artifacts in computed tomography. *Radiology*, 259(3):894–902, 2011.

[14] C. A. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Transactions on Image Processing*, 5(3):480–492, 1996.

[15] W. R. Brody, G. Butt, A. Hall, and A. Macovski. A method for selective tissue and bone visualization using dual energy scanned projection radiography. *Medical Physics*, 8:353–357, 1981.

[16] J. Browne and A. B. de Pierro. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Transactions on Medical Imaging*, 15(5):687–699, 1996.

[17] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *SIAM Multiscale Modeling and Simulation*, 4(2):490–530, 2005.

[18] C. L. Byrne. Convergent block-iterative algorithms for image reconstruction from inconsistent data. *IEEE Transactions on Image Processing*, 6(9):1296–1304, 1997.

[19] Lounis Chermak, Toby P Breckon, Greg T Flitton, and Najla Megherbi. Geometrical approach for automatic detection of liquid surfaces in 3D computed tomography baggage imagery. *The Imaging Science Journal*, 2015 - to appear.

[20] K.S. Chuang and H.K. Huang. A fast dual-energy computational method using isotransmission lines and table lookup. *Medical Physics*, 14:186–192, 1987.

[21] R. R. Coifman, D. L. Donoho, A. Antoniadis, and G. Oppenheim. Translation-invariant de-noising. *Wavelets and Statistics*, pages 125–150, 1995.

[22] C.R. Crawford, C. Karl, and H.E. Martz. Research and development of reconstruction advances in CT-based object detection systems - final report. Northeastern University, Boston. This report can be found at: `https://myfiles.neu.edu/groups/ALERT/strategic_studies/TaskOrder03/TO3_FinalReport.pdf`.

[23] C.R. Crawford, H.E. Martz, and H. Piena. Segmentation of objects from volumetric CT data, 2013. Northeastern University, Boston. This report can be found at: `https://myfiles.neu.edu/groups/ALERT/strategic_studies/SegmentationInitiativeFinalReport.pdf`.

[24] Antonio Criminisi. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends in Computer Graphics and Vision*, 7(2-3):81–227, 2011.

[25] Antonio Criminisi and Jamie Shotton. *Decision Forests for Computer Vision and Medical Image Analysis*. Springer Publishing Company, Incorporated, 2013.

[26] T. S. Curry, J. E. Dowdey, R. C. Murry, and E. E. Christensen. *Christensen's physics of diagnostic radiology*. Lippincott Williams and Wilkins, 4 edition, 1990.

[27] Victoria Cutler and Susan Paddock. Use of threat image projection (tip) to enhance security performance. In *IEEE International Carnahan Conference on Security Technology.*, pages 46–51, 2009.

[28] Stanley R. Deans. *The Radon Transform and Some of Its Applications.* Krieger Publishing Company, 1993.

[29] Chitra Dorai and Anil K Jain. COSMOS-a representation scheme for free-form surfaces. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1024–1029, 1995.

[30] Limor Eger, Synho Do, Prakash Ishwar, W Clem Karl, and Homer Pien. A learning-based approach to explosives detection using multi-energy X-ray computed tomography. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2004–2007, 2011.

[31] Limor Eger, Prakash Ishwar, WC Karl, and Homer Pien. Classification-aware dimensionality reduction methods for explosives detection using multi-energy X-ray computed tomography. In *SPIE Electronic Imaging*, pages 78730Q–78730Q. International Society for Optics and Photonics, 2011.

[32] Richard F Eilbert and Kristoph D Krug. Aspects of image recognition in vivid technologies' dual-energy X-ray system for explosives detection. In *Applications in Optical Science and Engineering*, pages 127–143. International Society for Optics and Photonics, 1993.

[33] A. Fenster. Split xenon detector for tomochemistry in computed tomography. *Journal of Computer Assisted Tomography*, 2(3):243–252, 1978.

[34] G. Flitton, T.P. Breckon, and N. Megherbi. Object recognition using 3D SIFT in complex CT volumes. In *Proceedings British Machine Vision Conference*, pages 11.1–11.12, 2010.

[35] Greg Flitton, Toby P Breckon, and Najla Megherbi. A comparison of 3D interest point descriptors with application to airport baggage object detection in complex CT imagery. *Pattern Recognition*, 2013.

[36] G.T. Flitton, T.P. Breckon, and N. Megherbi. A 3D Extension to Cortex Like Mechanisms for 3D Object Class Recognition. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pages 3634–3641, June 2012.

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

[38] GH Glover. Compton scatter effects in CT reconstructions. *Medical Physics*, 9:860, 1982.

[39] G. Van Gompel, K. Van Slambrouck, M. Defrise, K. J. Batenburg, J. de Mey, J. Sijbers, and J. Nuyts. Iterative correction of beam hardening artifacts in CT. *Medical Physics*, 38:S36, 2011.

[40] L. Grady. Fast, quality, segmentation of large volumes - isoperimetric distance trees. In *Proceedings European Conference on Computer Vision*, volume 3, pages 449–462, 2006.

[41] L. Grady and C. V. Alvino. The piecewise smooth Mumford-Shah functional on an arbitrary graph. *IEEE Transactions on Image Processing*, 18(11):2547–2561, 2009.

[42] L. Grady, V. Singh, T. Kohlberger, C. Alvino, and C. Bahlmann. *Automatic segmentation of unknown objects, with application to baggage security*, pages 430–444. European Conference on Computer Vision. Springer, 2012.

[43] Anno Graser, Thorsten RC Johnson, Hersh Chandarana, and Michael Macari. Dual energy CT: preliminary observations and potential clinical applications in the abdomen. *European Radiology*, 19(1):13–23, 2009.

[44] Franziska Hofer and Adrian Schwaninger. Using threat image projection data for assessing individual screener performance. *WIT Transactions on the Built Environment*, 82:417–426, 2005.

[45] J. Hsieh. *Computed tomography: principles, design, artifacts, and recent advances.* SPIE and John Wiley and Sons, 2003.

[46] H. M. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609, 1994.

[47] Yannan Jin. Implementation and optimization of dual energy computed tomography [PhD Thesis], University of Erlangen-Nuremberg, 2011.

[48] Thorsten RC Johnson. *Medical radiology/diagnostic imaging: dual energy CT in clinical practice*. Springer, 2011.

[49] W. A. Kalender, R. Hebel, and J. Ebersberger. Reduction of CT artifacts caused by metallic implants. *Radiology*, 164(2):576, 1987.

[50] W.A. Kalender. *Computed Tomography*. Wiley, 2011.

[51] W.A. Kalender, W.H. Perman, J.R. Vetter, and E. Klotz. Evaluation of a prototype dual-energy computed tomographic apparatus. i. phantom studies. *Medical Physics*, 13:334–339, 1986.

[52] SS Karimi, PC Cosman, and HE Martz. Metal artifact reduction for CT-based luggage screening, October 2013. Lawrence Livermore National Laboratory Technical Report, LLNL-TR-645205.

[53] A. F. Kopp, K. Klingenbeck-Regn, M. Heuschmid, A. Kuttner, B. Ohnesorge, T. Flohr, S. Schaller, and C. D. Claussen. Multislice computed tomography: basic principles and clinical applications. *Electromedica-Erlangen*, 68(2):94–105, 2000.

[54] B. Kratz, S. Ens, J. Mller, and T. M. Buzug. Reference-free ground truth metric for metal artifact evaluation in CT images. *Medical Physics*, 38:4321–4328, 2011.

[55] Y. Lan, R. Harvey, and J. R. Perez Torres. Finding stable salient contours. *Image and Vision Computing*, 28(8):1244–1254, 2010.

[56] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using affine-invariant regions. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, volume 2, pages II–319 – II–324, 2003.

[57] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1265–1278, 2005.

[58] Q. Li, S. Ahn, and R. Leahy. Fast hybrid algorithms for PET image reconstruction. In *Nuclear Science Symposium Conference Record, IEEE*, volume 4, pages 1851–1855, 2005.

[59] Q. Lu. The utility of X-ray dual-energy transmission and scatter technologies for illicit material detection [PhD Thesis] Virginia Polytechnic Institute and State University, 1999.

[60] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens. Metal streak artifacts in X-ray computed tomography: a simulation study. In *IEEE Transactions on Nuclear Science*, volume 46, pages 691–696, 1999.

[61] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens. Reduction of metal streak artifacts in X-ray computed tomography using a transmission maximum a posteriori algorithm. *IEEE Transactions on Nuclear Science*, 47(3):977–981, 2000.

[62] Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[63] N. Megherbi, T.P. Breckon, G.T. Flitton, and A. Mouton. Fully automatic 3D threat image projection: application to densely cluttered 3D computed tomography baggage images. In *Proceedings of the IEEE International Conference on Image Processing Theory, Tools and Applications*, pages 153–159, 2012.

[64] N. Megherbi, T.P. Breckon, G.T Flitton, and A. Mouton. Investigating existing medical CT segmentation techniques within automated baggage and package inspection. In *Proceedings SPIE Security and Defence: Optics and Photonics for Counterterrorism, Crime Fighting and Defence*, volume 8901, pages 1–8, 2013.

[65] N. Megherbi, T.P. Breckon, G.T. Flitton, and A. Mouton. Radon transform based metal artefacts generation in 3D threat image projection. In *Proceedings of SPIE Optics and Photonics for Counterterrorism, Crime Fighting and Defence*, volume 8901, pages 1–7, 2013.

[66] 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 *Proceedings of the IEEE International Conference on Image Processing*, pages 1833–1836, 2010.

[67] Najla Megherbi, Jiwan Han, Toby P Breckon, and Greg T Flitton. A comparison of classification approaches for threat detection in CT based baggage screening. In *Proceedings of the IEEE International Conference on Image Processing*, pages 3109–3112, 2012.

[68] Rajiv Mehrotra, Kameswara Rao Namuduri, and Nagarajan Ranganathan. Gabor filter-based edge detection. *Pattern Recognition*, 25(12):1479–1494, 1992.

[69] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelriess. Normalized metal artifact reduction (NMAR) in computed tomography. *Medical Physics*, 37:5482, 2010.

[70] Frank Moosmann, Bill Triggs, and Frederic Jurie. Fast discriminative visual codebooks using randomized clustering forests. *Advances in Neural Information Processing Systems 19*, pages 985–992, 2007.

[71] A. Mouton. On artefact reduction, segmentation and classification of 3D computed tomography imagery in baggage security screening [PhD Thesis], Cranfield University UK, 2014.

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

[73] A. Mouton, T.P. Breckon, and G.T. Flitton. 3D object classification in complex volumes using randomised clustering forests. In *IEEE International Conference on Image Processing*, pages 5202–5206, 2014.

[74] A. Mouton, N. Megherbi, T.P. Breckon, K. Van Slambrouck, and J. Nuyts. A distance driven method for metal artefact reduction in computed tomography. In *Proceedings IEEE International Conference on Image Processing*, pages 2334–2338, 2013.

[75] A. Mouton, N. Megherbi, G.T. Flitton, and T.P. Breckon. A novel intensity limiting approach to metal artefact reduction in 3D CT baggage imagery. In *Proceedings of the IEEE International Conference on Image Processing*, pages 2057–2060, 2012.

[76] A. Mouton, N. Megherbi, G.T. Flitton, and T.P. Breckon. An evaluation of CT image denoising techniques applied to baggage imagery screening. In *Proceedings IEEE International Conference on Industrial Technology*, pages 1063–1068, 2013.

[77] A. Mouton, N. Megherbi, K. van Slambrouk, J. Nuyts, and T.P. Breckon. An experimental survey of metal artefact reduction in computed tomography. *Journal of X-Ray Science and Technology*, 21(2):193–226, 2013.

[78] R. Naidu, I. Bechwati, and C.R. Crawford. Decomposition of multi-energy scan projections using multi-step fitting. *US Patent 7,197,172 B1*, Filed : July 1, 2003.

[79] Marcin Novotni and Reinhard Klein. Shape retrieval using 3D zernike descriptors. *Computer-Aided Design*, 36(11):1047–1062, 2004.

[80] E. Nowak, F. Jurie, and B. Triggs. Sampling strategies for bag-of-features image classification. In *European Conference on Computer Vision*, pages 490–503, 2006.

[81] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[82] M. Petrou and P. Bosdogianni. *Image Processing: The Fundamentals*. John Wiley and Sons, 1999.

[83] Fredrick L Roder. Explosives detection by dual-energy computed tomography (CT). In *Techincal Symposium East*, pages 171–178. International Society for Optics and Photonics, 1979.

[84] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992.

[85] S Rasoul Safavian and David Landgrebe. A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man and Cybernetics*, 21(3):660–674, 1991.

[86] R. Sedgewick. *Algorithms in C*. Addison-Wesley Professional, 3 edition, 1998.

[87] E. Seeram. *Computed tomography: physical principles, clinical applications, and quality control*. WB Saunders Philadelphia, 2001.

[88] O. Semerci and E.L. Miller. A parametric level-set approach to simultaneous object identification and background reconstruction for dual-energy computed tomography. *IEEE Transactions on Image Processing*, 21(5):2719–2734, 2012.

[89] James Albert Sethian. *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, volume 3. Cambridge university press, 1999.

[90] S. Singh. Explosives detection systems (EDS) for aviation security. *Signal Processing*, 83(1):31–55, 2003.

[91] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1470–1477, 2003.

[92] K. Van Slambrouck and J. Nuyts. Metal artifact reduction in computed tomography using local models in an image block-iterative scheme. *Medical Physics*, 39(11):7080, 2012.

[93] J.A. Smith, H.E. Martz, and J.S. Kallman. Case for an improved effective atomic number for the electronic baggage screening program, December 2011. Lawrence Livermore National Laboratory Technical Report, LLNLTR-520312.

[94] C.J. Solomon and T.P. Breckon. *Fundamentals of digital image processing: a practical approach with examples in Matlab*. Wiley-Blackwell, 2010.

[95] J. B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh. A three-dimensional statistical approach to improved image quality for multislice helical CT. *Medical Physics*, 34:4526, 2007.

[96] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 839–846, 1998.

[97] H. K. Tuy. A post-processing algorithm to reduce metallic clip artifacts in CT images. *European Radiology*, 3(2):129–134, 1993.

[98] Jayaram K Udupa and Supun Samarasekera. Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. *Graphical models and image processing*, 58(3):246–261, 1996.

[99] L. Garcia Ugarriza, E. Saber, S. R. Vantaram, V. Amuso, M. Shaw, and R. Bhaskar. Automatic image segmentation by dynamic region growth and multiresolution merging. *IEEE Transactions on Image Processing*, 18(10):2275–2288, 2009.

[100] J. C. van Gemert, C. J. Veenman, A. W. M. Smeulders, and J. M. Geusebroek. Visual word ambiguity. *IEEE transactions on pattern analysis and machine intelligence*, 32(7):1271–1283, 2010.

[101] V. N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York Inc, 2000.

[102] Joost M Verburg and Joao Seco. CT metal artifact reduction method correcting for beam hardening and missing projections. *Physics in Medicine and Biology*, 57(9):2803, 2012.

[103] Luc Vincent and Pierre Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE transactions on pattern analysis and machine intelligence*, 13(6):583–598, 1991.

[104] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, volume 1, 2001.

[105] Eric A Wan. Neural network classification: A Bayesian interpretation. *IEEE Transactions on Neural Networks*, 1(4):303–305, 1990.

[106] S. Y. Wan and W. E. Higgins. Symmetric region growing. *IEEE Transactions on Image Processing*, 12(9):1007–1015, 2003.

[107] G. Wang, D. L. Snyder, J. A. O'Sullivan, and M. W. Vannier. Iterative deblurring for CT metal artifact reduction. *IEEE Transactions on Medical Imaging*, 15(5):657–664, 1996.

[108] D. F. Wiley. Analysis of anatomic regions delineated from image data. *US Patent 8,194,964 (US App. No. 12/430,545)*, 2009.

[109] D. F. Wiley, D. Ghosh, and C. Woodhouse. Automatic segmentation of CT scans of checked baggage. In *Proceedings of the 2nd International Meeting on Image Formation in X-ray CT*, pages 310–313, 2012.

[110] Yesna O Yildiz, Douglas Q Abraham, Sos Agaian, and Karen Panetta. 3d threat image projection. In *Electronic Imaging 2008*, pages 680508–680508. International Society for Optics and Photonics, 2008.

[111] Z. Ying, R. Naidu, and C. R. Crawford. Dual energy computed tomography for explosive detection. *Journal of X-ray Science and Technology*, 14(4):235–256, 2006.

[112] Zhengrong Ying, Ram Naidu, Sergey Simanovsky, and Carl R Crawford. Method of and system for computing effective atomic number images in multi-energy computed tomography. *US Patent 7,190,757*, Filed: March 13, 2007.

[113] Z. Yu, J. B. Thibault, C. A. Bouman, K. D. Sauer, and J. Hsieh. Fast model-based X-Ray CT reconstruction using spatially nonhomogeneous ICD optimization. *IEEE Transactions on Image Processing*, 20(1):161–175, 2011.

[114] M. Zhang and B. K. Gunturk. A new image denoising framework based on bilateral filter. In *Proceedings SPIE*, volume 6822, pages 68221B–68221B–8, 2008.

[115] S. Zhao, K. T. Bae, B. Whiting, and G. Wang. A wavelet method for metal artifact reduction with multiple metallic objects in the field of view. *Journal of X-Ray Science and Technology*, 10(2):67–76, 2002.

[116] Y. Zhou, K. Panetta, and S. Agaian. 3D CT baggage image enhancement based on order statistic decomposition. In *Proceedings of the IEEE International Conference on Technologies for Homeland Security*, pages 287–291, 2010.

[117] Y. Zhou, K. Panetta, and S. Agaian. CT baggage image enhancement using a combination of alpha-weighted mean separation and histogram equalization. In *Proceedings SPIE*, volume 7708, page 77080G, 2010.

Tables

Table 1: Overview of denoising and MAR within baggage-CT. For each study, the study objective, the methodology, the type of data, the evaluation domain, the Quantitative Evaluation (QE) measures used and the main conclusions are provided.

Study	Aim	Method(s)	Data	QE	Notes
Zhou *et al.* [116]	Contrast enhancement	Alpha-weighted separation; histogram equalisation	2D & 3D (source unknown)	Second-derivative-like measure of enhancement	Contrast improved but images have few artefacts and noise
Mouton *et al.* [76]	Denoising	Comparison of 6 techniques	2D & 3D; DECT baggage scanner	3D SIFT measure	NLM filtering [17] and wavelet shrinkage [21] quantitative and qualitative improvements
Mouton *et al.* [77]	MAR	Comparison of 12 techniques	2D; DECT baggage scanner; medical phantom [60]; medical-CT scanner	NRMSE; Reference-free measure [54]; combined error-time measure	techniques require priors; sensitive to parameters; perform poorly on baggage data
Mouton *et al.* [74]	MAR	NLM filtering; sinogram completion; distance-weighted correction	2D; DECT baggage scanner; medical phantom [60]; medical-CT scanner	NRMSE; Reference-free measure [54]; Bland-Altman plots [4]	Improved removal of shading artefacts; blurring
Karimi *et al.* [52]	MAR	Constrained optimisation to construct prior; miniaturisation; sinogram completion	2D; single-energy medical CT scanner	None	First to employ prior; computationally efficient; non-representative data
ALERT [22]	Reconstruction	FBP; iterative reconstruction; DECT	2D & 3D; DECT medical scanner	Various	Iterative reconstruction superior but computationally expensive; non-representative data

Table 2: Overview of DECT decomposition within baggage-CT. For each study, the study objective, the basic methodology, the type and source of data and the main conclusions are provided. Only studies since 2003 are included.

Study	**Aim**	**Method(s)**	**Data**	**Notes**
Naidu *et al.* [78]	DECT decomposition	Multi-step fitting; iso-transmission lines [20]	Unknown	Patented method; requires calibration
Ying *et al.* [111]	Explosives detection	DECT decomposition by constrained least squares minimisation; adaptive scatter correction; destreaking by nonlinear filtering; spectral drift correction	Simulated; DECT baggage scanner	Improved Z_{eff} images (w.r.t. mean atomic number, cupping effect, SNR); faster computation of Z_{eff} using high-energy data [112]; requires calibration
Smith *et al.* [93]	Z_{eff} computation	Physical model of compound reference materials	DECT baggage scanner	Qualitative and quantitative improvements in Z_{eff} for reference materials scanned on commercial scanner; requires calibration
Semerci and Miller [88]	Reconstruction for object detection	Polychromatic DECT iterative reconstruction using parametric level-sets	2D simulations	Successful object detection, object shape and location determination, background reconstruction; not tested on real data; requires calibration
Eger *et al.* [31, 30]	Explosives detection	MECT techniques; projection coefficient features; SVM classifier	Database of explosive and non-explosive LACs (source unknown)	Improved classification performance over Compton and photoelectric coefficients

Table 3: Overview of 3D object segmentation within baggage-CT. For each study, the core methodology, the type and source of data, the Quantitative Evaluation (QE) measures and the main conclusions are provided.

Study	Method(s)	Data	QE	Notes
Wiley *et al.* [109]	3D flood-fill region growing; merging by overlap	Single-energy medical-CT [23]	None	Kernel movement inferred from training data; performance deteriorates for low-density and sheet-like objects; non-representative data
Song *et al.* [23]	Bilateral filtering; sequential symmetric region growing [106]; 4-stage object splitting; 3-stage object merging	Single-energy medical-CT [23]	Wholeness and exclusiveness (comparison to ground truth) [23]	Large parameter space; sensitive to parameter tuning; performs poorly on low-density and sheet-like objects; non-representative data
Grady *et al.* [42]	MAR [97]; Mumford-Shah functional [41]; recursive IDT partitioning [40] - driven by AQUA measure	Single-energy medical-CT [23]	% Overlap with ground truth	Segmentation of cluttered bags in < 4 minutes; performance deteriorates on low-density and sheet-like objects; non-representative data
Harvey *et al.* [23]	Multi-scale sieves [7]; object merging by Kolmogorov-Smirnov histogram analysis [55]	Single-energy medical-CT [23]	None	Computationally efficient; can use intermediate results directly in classification system; fails on iso-density touching objects; non-representative data
Feng *et al.* [23]	Gradient-based seed map generation; adaptive region growing [99]; feature-based merging	Single-energy medical-CT [23]	Comparison to ground truth (size; mean intensity etc.)	Sensitive to parameter tuning; prone to under-segmentations; performs poorly in presence of streaking artefacts; non-representative data
Megherbi *et al.* [64]	Comparison of existing techniques [82, 98, 103, 89]	Dual-energy baggage-CT	None	Sensitive to parameter tuning; performance limited by lack of *a priori* information, image noise, artefacts and clutter
Mouton [72]	MAR [74]; DEI materials-based coarse segmentation; refinement guided by random forest score	Dual-energy baggage-CT	Random forest-based performance measure	Computationally efficient; realistic data; performs poorly on thin sheet-like objects

Table 4: Overview of 3D object classification within baggage-CT. For each study, the target objects, the core methodology, the type and source of data, the Quantitative Evaluation (QE) measures and the main conclusions and approximate performance are provided.

Study	Target	Method(s)	Data	QE	Notes
Chen *et al.* [11]	Pistols	DECT decomposition by HL curves and LUTs; boosted 2D Haar features	Isolated pistols (source unknown)	None	Classification reduced to 2D; oversimplified data; no experimental results
Megherbi *et al.* [66, 67]	Bottles	Histogram-of-shape index; SVM classifier	Manually segmented baggage-CT	Accuracy; precision; TNR; recall	Requires manual segmentation; limited dataset; classification accuracy $> 98\%$
Flitton *et al.* [35]	Handguns & bottles	Density Histograms (DH) descriptors; k-means encoding; SVM classifier	Manually segmented baggage-CT	ROC curves; mean correspondence	DH descriptors outperform 3D SIFT [34] and RIFT [56]; requires manual segmentation; performance limited in presence of streaking artefacts
Flitton *et al.* [36, 37]	Handguns & bottles	Dense Gabor features [68]; 3D visual cortex model; SVM classifier	Manually segmented baggage-CT	TPR; FPR; precision; recall	Outperforms BoW model using 3D SIFT and SVM; TPR$\sim$ 96%; FPR$\sim$ 1%; computationally expensive; requires manual segmentation; performance limited in presence of streaking artefacts
Mouton [73]	Handguns & bottles	Dense DH descriptors [35]; random forest feature encoding [70]; SVM classifier	Manually segmented baggage-CT	TPR; FPR; precision; time	Outperforms visual cortex [36]; TPR$\sim$ 99%; FPR $< 1\%$; computationally efficient; robust to noise; requires manual segmentation
Mouton [71]	Handguns & bottles	MAR/denoising [74, 17]; DEI materials-based segmentation; dense DH descriptors [35]; random forest encoding [70]; SVM classifier	Dual-energy baggage-CT	TPR; FPR; accuracy; precision; time	First end-to-end classification pipeline; TPR$>$ 97%; FPR$<$ 2%; NLM gives similar results to MAR at massive reduction in computational cost
Chermark *et al.* [19]	Liquids	Thresholding and slice differencing; 3D planar fitting; 2D elliptical fitting	Dual-energy baggage-CT	Accuracy, precision, recall	TPR$>$ 85%; no training required.

Table 5: Overview of 3D Threat Image Projection (TIP). For each study, the core methodology, the type and source of data, the evaluation techniques and the main conclusions are provided.

Study	**Method(s)**	**Data**	**Evaluation**	**Notes**
Yildiz *et al.* [110]	Manual void determination; direct threat insertion (with original artefacts)	Source unknown	None	Algorithmic details limited; no results presented
Megherbi *et al.* [63]	Void determination; threat insertion location determination; sinogram-space metal artefact generation [65]	Cluttered baggage-CT	TIP and non-TIP images examined and classified by 3 expert screeners	Realistic threat insertion; experts unable to distinguish TIP from non-TIP; limited void determination; computationally expensive

Figure Captions

Fig. 1: Conventional classification pipeline.

Fig. 2: Medical-grade CT scans with sub-millimetre isotropic resolution [2, 1].

Fig. 3: Baggage-CT scans illustrating poor image quality, low resolution ($1.56 \times 1.61 \times$ 5mm), artefacts and clutter (obtained on Reveal CT80-DR dual-energy baggage scanner).

Fig. 4: Z_{eff} as a function of density for common materials found in packed luggage [32].

Fig. 5: Determination of Z_{eff} by interpolation of approximating polynomial [93].

Figures

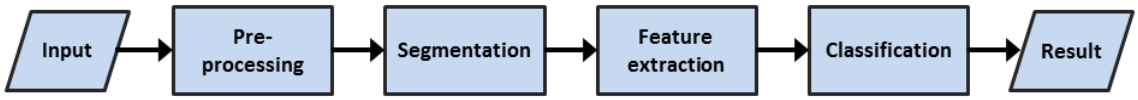

Fig. 1: Conventional classification pipeline.

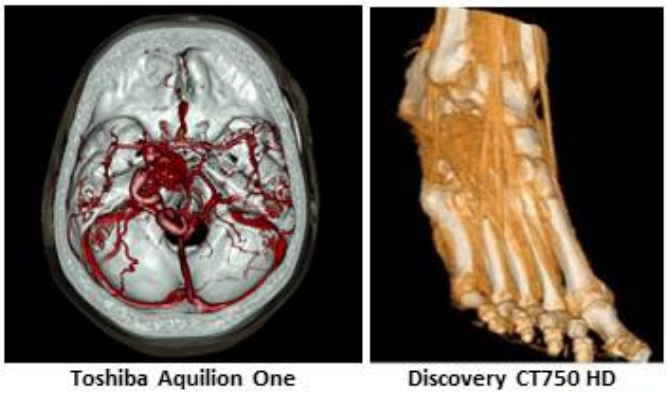

Fig. 2: Medical-grade CT scans with sub-millimetre isotropic resolution [2, 1].

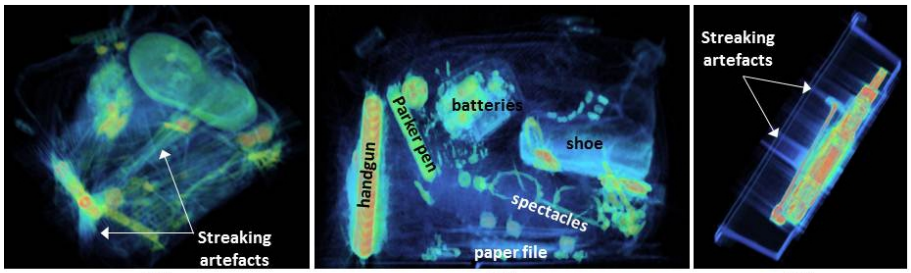

Fig. 3: Baggage-CT scans illustrating poor image quality, low resolution ($1.56 \times 1.61 \times$ 5mm), artefacts and clutter (obtained on Reveal CT80-DR dual-energy baggage scanner).

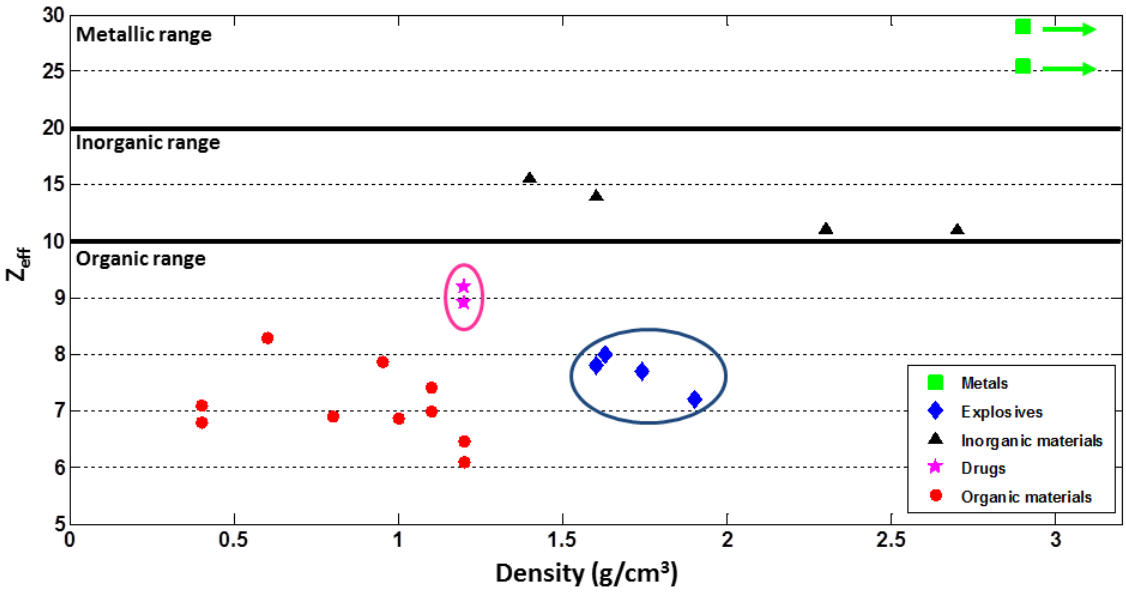

Fig. 4: Z_{eff} as a function of density for common materials found in packed luggage [32].

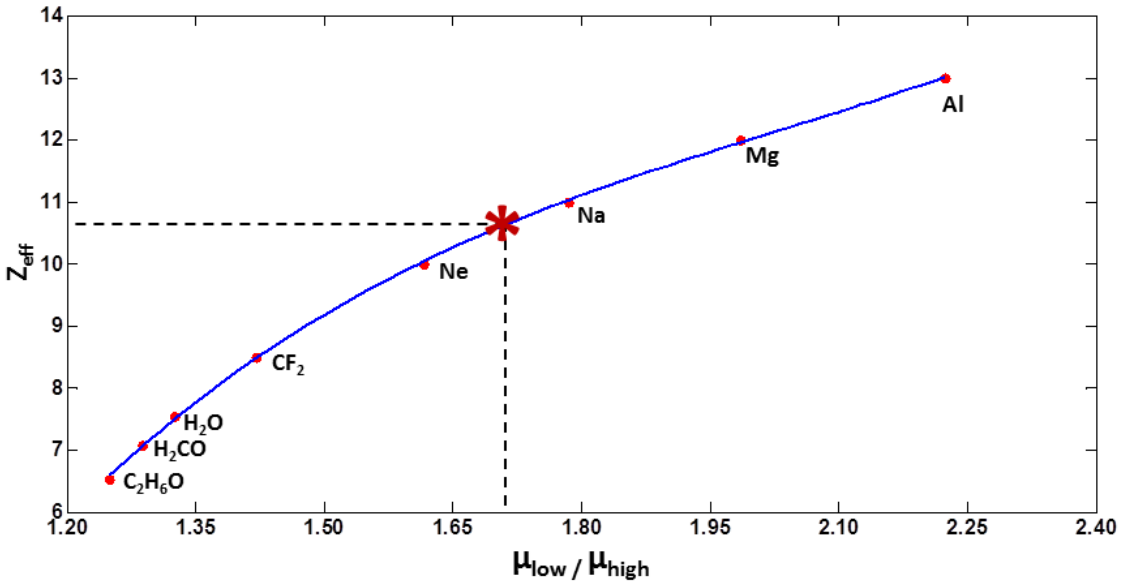

Fig. 5: Determination of Z_{eff} by interpolation of approximating polynomial [93].