International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:6, No:4, 2012

Color View Synthesis for Animated Depth Security X-ray Imaging

O. Abusaeeda, J. P. O Evans and D. Downes

***Abstract*—**We demonstrate the synthesis of intermediary views within a sequence of color encoded, materials discriminating, X-ray images that exhibit animated depth in a visual display. During the image acquisition process, the requirement for a linear X-ray detector array is replaced by synthetic image. Scale Invariant Feature Transform, SIFT, in combination with material segmented morphing is employed to produce synthetic imagery. A quantitative analysis of the feature matching performance of the SIFT is presented along with a comparative study of the synthetic imagery. We show that the total number of matches produced by SIFT reduces as the angular separation between the generating views increases. This effect is accompanied by an increase in the total number of synthetic pixel errors. The trends observed are obtained from 15 different luggage items. This programme of research is in collaboration with the UK Home Office and the US Dept. of Homeland Security.

***Keywords*—**X-ray, kinetic depth, view synthesis, KDE

I. Introduction

AIRPORT luggage checks are a recognized concern for various reasons. The increasing sophistication of concealed devices is compounded by the difficult task of distinguishing between everyday objects within a 'typical' suitcase. The attenuation of the screeners' attention when watching the display and examining repetitively the same category of benign objects. This complex visual task is exacerbated by the lack of cues to depth in imagery, which has been produced by transmitted radiation. Achieving higher detection rates coupled with low false alarm rates during airport luggage inspection is a prime challenge for aviation security personnel. Past work by the University team in collaboration with the Centre for Applied Science and Technology (CAST), formally the UK Home Office Scientific Development Branch (HOSDB) has produced a novel binocular stereoscopic X-ray technique [1]–[3] which aids the detection and identification of objects under inspection. Imaging technology based on this early work is now commercially available. More recently through ongoing collaboration with the CAST and the US Department of Homeland Security (DHS), the University team has developed multiple view techniques that produce three-dimensional imagery termed kinetic depth X-ray imaging or KDEX.

O. Abusaeeda is a PhD student at The Imaging Science Group, College of Arts and Science Nottingham Trent University (phone: +44(0)1158488409; e-mail: omar.abusaeeda@ntu.ac.uk).

J.P.O. Evans is a professor of Applied Imaging Science and the head of The Imaging Science Group at College of Arts and Science, Nottingham Trent University ((phone: +44(0)1158488342; e-mail: paul.evans@ntu.ac.uk).

D. Downes is a Lecturer/Senior Lecturer at Interactive Arts, he is also researcher at The Imaging Science Group, College of Arts and Science Nottingham Trent University (phone: +44(0)1158482050; e-mail: david.downes02@ntu.ac.uk).

The KDEX approach involves capturing a series of perspective views from different station points. The resultant image sequence exhibits sequential parallax in a visual display to enable the perception of depth through motion. Threat objects which are imperceptible in some views may become visible in other views. This imagery may also be viewed as a binocular stereoscopic sequence [4] but this approach is beyond the scope of this paper. The practical implementation of KDEX requires multiple views of the luggage to be acquired from an arrangement of linear (or folded linear) X-ray detector arrays. Each array is illuminated by a thin curtain of X-rays originating from a single X-ray source. The views are captured during a linear translation of the object under inspection through the interrogating X-ray beams.

The research presented in this paper explores the replacing intermediary linear X-ray detector arrays by employing image synthesis. This approach will enable 'in-between' detectors and their associated hardware to be discarded. Earlier work undertaken by the University team [5], [6] has developed an image synthesis algorithm based on correlation matching operating on grey scale imagery. The research reported here explores the potential of SIFT as a powerful feature extraction algorithm applied to color encoded (materials discriminating) X-ray images. SIFT was firstly proposed by David Lowe in 2004 [7]. In this work SIFT is additionally bounded by epipolar and disparity window criteria. The output from the SIFT is employed by material based morphing algorithm to produce color synthetic views, which conserve the encoded material classes. Ultimately, high quality synthetic images are combined together with the detector images to produce the resultant KDEX sequence. Limitations of the approach are investigated by employing ground truth imagery for comparative error analysis.

II. Background

An early example of the automated matching of corresponding points of interest in images is the corner detector proposed by Moravec [8]. This effort was later improved by Harris and Stephens [9]. As a result the Harris corner detector has since been widely used for numerous other image matching tasks. A natural extension of this work is to consider more other local features and their identification under conditions of scale invariance [10]. In recent times, there has been significant interest in extending local feature identification to be invariant to full affine changes [11], [12]; SIFT is one such approach [7], which extracts distinctive invariant features from images and matches them across a substantial range of features. The aim is to establish robust performance against changes in viewpoint. SIFT may also be combined with principal components analysis (PCA-SIFT)

[13] which is able to produce more compact descriptors in comparison with the standard SIFT. But, it can increase blurring along features [14]. Another local feature descriptor namely, speeded-up robust features (SURF) [15] has a similar performance to SIFT although it is not invariant to rotation and strong illumination changes [14]. It has been demonstrated recently that features identified by SIFT are highly distinctive and invariant to image scales and rotations, and partially invariant to a change in illumination [16]. The latter aspect is important consideration for the X-ray transmission imagery dealt with in this paper. Prior work by the University team has highlighted the potential for SIFT in searching for correspondences in X-ray images [17]. In this paper, SIFT provides key points for a morphing algorithm, which generates the resultant synthetic view.

Typically, morphing methods combine a geometric warp with an intensity blend or cross-fade [18], [19]. Numerous variations adopting these methods are documented [20]–[23]. The majority of warping methods are similar in that they require a set of corresponding features in a pair of perspective views while the remaining correspondences are determined automatically by interpolation. Features in the original image are shifted incrementally with respect to those in the final image by applying the mapping and vice versa. However, they often differ in how they locate features and the method of interpolating them. In practice, it emerges that there is no extensive criterion in which the quality or practicality of the morph can be assessed [18], [24]. For example, unnatural image transitions can arise due to the three-dimensional shape of objects not being conserved. This situation is not surprising given the three–dimensional coordinate positions of the features are not taken into account or known in the majority of warping procedures.

III. Methods And Materials

The main methods involve the production of the raw color encoded X-ray imagery; the algorithmic development and, the comparative analysis of the synthetic imagery with ground truth. Each of these aspects is discussed in the following sections.

A. Production of X-ray images

The color encoded X-ray images employed in this investigation were produced by a multiple view X-ray scanner housed in the University's Imaging Science Laboratory. The luggage imaged contained a mixture of 'typical' objects composed of different materials, thicknesses and shape [6]. Care was taken to ensure that partially overlapping and fully overlapping objects provided the multi-layered translucency typical encountered in security scans of luggage.

When the X-ray photons pass through matter, measurements are taken at two X-ray energy levels '75keV and 150keV' to generate low (Lo) and high energy (Hi) X-ray signals. The relative difference in magnitude between (Hi) and (Lo) is exploited to broadly discriminate an inspected object into three material classes as function of atomic number, Z. The organic class has an atomic number of $Z \leq 10$, a mixture class $10 < Z < 20$ and the metallic class $Z \geq 20$ [25], [26]. The resultant discrimination information is presented to the human operators by color encoding the X-ray images. Fig. 1 presents an X-ray images of printed circuit where (a) and (b) are the low (L)and high (H) energy grayscale X-ray images while (c) is the color encoded X-ray image.

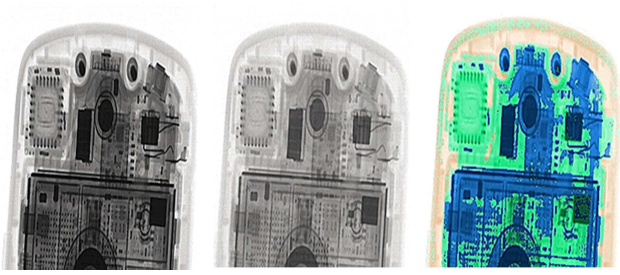


Fig. 1 Production of color encoded image from (a) Low energy image, (b) High energy image and (c) Color encoded X-ray image

The resultant color coded information provides a broad discrimination of the objects under inspection into three material classes. This materials discrimination technique has afforded an elegant segmentation solution to simplify highly cluttered images

Typically organic, mixture, and metallic material is displayed as orange, green and blue color respectively as shown in figure 1 (c). Security X-ray screening systems utilizing dual-energy materials discrimination are widely deployed in the major airports throughout the world [26].

In this study 7 sequential perspective images are acquired and stored for each luggage item or bag under inspection. The angular separation between each successive view is 1º taken over an angular range of ±3º about the 0º or normal view position as illustrated in figure 2. A total of 105 perspective images from 15 different bags termed; b1, b2…b15 were employed in this study.

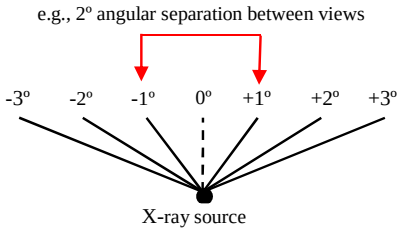


Fig. 2 Illustration of angular separation between perspective views

B. Image Matching

The SIFT algorithm adopts the fast nearest-neighbour method to identify the best match for a particular feature from a database of features. Lowe suggested that all matches in which the distance ratio between the closest neighbour to that of the second closest neighbour > 0.8 be discarded. He also noted that the application of this criterion can remove around 90% of the incorrect matches while discarding less than 5% of the true matches [7]. To maximise the potential applicability of SIFT Lowe originally did not impose a disparity limit. In

our work, we can employ additional bounding criteria of a disparity window and an epipolar line constraint as illustrated in the Flowchart in figure 3. The former, arises from consideration of the inspection tunnel of the X-ray machine, defining a near and far boundary to the inspection volume, while the latter concerns the nominally zero vertical disparity exhibited by the KDEX imagery i.e. the epipolar line is along the image y-axis (horizontal in the display). The adoption of these new criteria tighten the support of SIFT. The

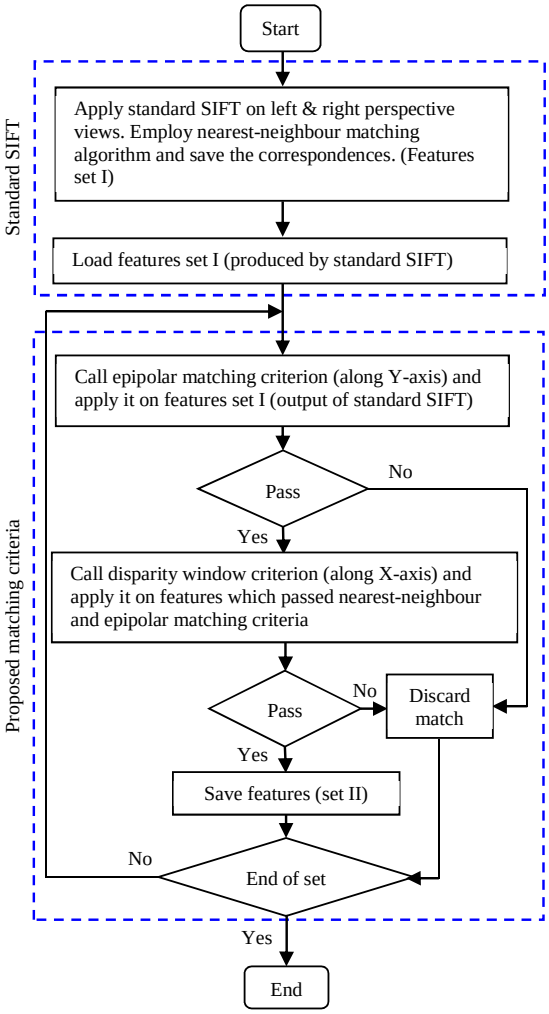
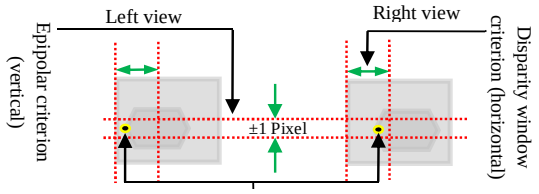


Fig. 3 Flowchart for the proposed matching route

Fig. 4 is a hypothetical example of X-ray image where the feature search space bounded by applying a maximum disparity and an epipolar line constraint. The adoption of these criteria, determined by the spatiotemporal X-ray machine design, tighten the support of SIFT.



Point of interest (left view) and its corresponding (right view)

Fig. 4 Pixel search space determined by the epipolar line and maximum disparity criteria

Matched features are categorised into two groups; negative matches, which satisfy the standard SIFT criterion but violate either one of the new criteria, while positive matches are grouped as features in which standard SIFT and both new criteria are satisfied. Taking into account various practical fluctuations a tolerance of ±1 pixel deviation in the y-axis coordinate position (vertical in the display) is employed to accommodate a practical epipolar line criterion[17]. To further limit the search space a disparity window criterion is introduced. The window size in pixels is determined by the spatiotemporal X-ray machine design. It should be noted that for comparative purposes the new criteria are applied to corresponding pairs that have already satisfied the standard SIFT criteria.

C. Color encoded image synthesis

Image morphing is the transformation of one image into another image. The morph of two X-ray imagesI_0 and I_1 requires two correspondence maps M_0 and M_1 where $(M_0: I_0 \rightarrow I_1)$ and $(M_1: I_1 \rightarrow I_0)$. In practice, M_0 and M_1 may be derived manually by the user where critical points on two images are selected on two images side by side and then finally provides a set of common features on two images[18], [24]. The remaining correspondences are established automatically by interpolation [27], [28]. Shapes of objects are destroyed if M_0 and M_1 are inaccurately identified. In our work automated morphing is an essential prerequisite and M_0 and M_1 are established by applyingthe SIFT bounded by epipolar and disparity constraints as previously described. As soon as the pixels in I_0 and I_1 are interpolated, the process of generating the morph is computed by applying the mapping incrementally on $x\text{-}axis$ $(motion\ axis)$direction only to shift the features in the original image to those in the final image and vice versa.

As it is well known, morphing is a combination between geometric warp and an intensity blend (cross-fade). However, blending intensities might cause generation of new colors. This procedure is applicable in visible light images as there is no color boundary conditions but this scenario cannot directly applied onto color encoded X-ray images as it might lead to generate new material class out of the standard material classes (organic, mixture and metallic). As a result, the material segmented morphing is proposed. It is new approach of color encode the synthetic X-ray image according to the perspective views used to generate it. For instance, if a left pixel is classified as organic (orange) and the corresponding

right pixel is classified as metallic (blue), then the resultant synthetic pixel is classified as a mixture (green). This is an important consideration as it ensures that each synthetic pixel belongs to one of the calibrated material classes. This way of color encode avoids interpolated colors, representing false material information, which could impinge upon the detection and identification of a threat object.

D. Comparison between synthetic and ground truth views

The quality of the synthetic images was determined by comparison with the ground truth according to the following formula:

$$Error = \left\| \left\{ |P \in M| abs\big(G(P) - S(P)\big) > \sqrt{G(P)} \right\} \right\| \qquad (1)$$

Where M depends on the image size, $G(P)$ is the intensity of the ground truth image at P and $S(P)$ is the intensity of the synthetic image at position P. This approach [6] is designed to accommodate ground truth imagery, which is subject itself to the concatenation of noise sources within a complex X-ray imaging chain.

The errors recorded for each bag are normalised with respect to the minimum error recorded for that bag. The minimum error occurs when the angular separation between the successive views is also at a minimum, which in this study is 2º. This approach enables the relative error behaviour associated with each individual bag to be presented and studied more easily.

IV. Results and Discussion

A. Matching results

Fig. 5 graphically illustrates the number of normalised correspondence matches for 15 bags as a function of the angular separation between views i.e. the angle between the interrogating X-ray beams employed to generate each view.

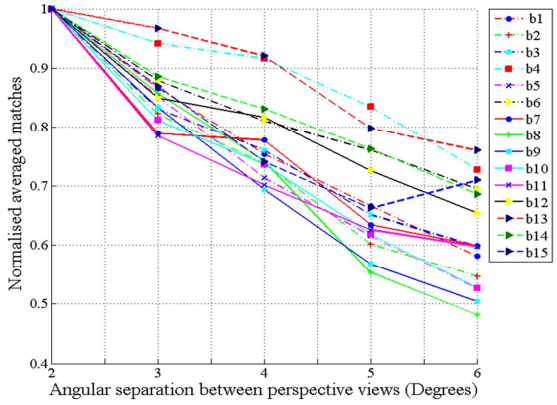


Fig. 5 Normalized matches for 15 bags at 2º, 3º, 4º, 5º and 6º of separation

It can be observed that there is a strong and consistent trend for the total number of correspondence matches to decrease as the beam angle is increased. This observation was expected since in general, the dissimilarity in views increases with increasing angular separation between the views.

The average trend in matching performance is represented by the central red color curve in figure 6. The upper and lower bounds of the maximum and minimum number of matches recorded for any of the 15 bags is also illustrated graphically for completeness.

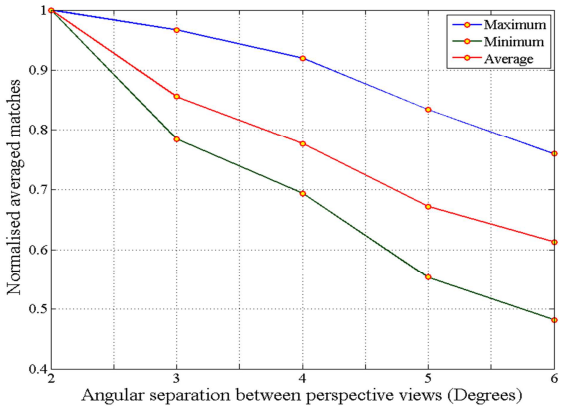


Fig. 6 Maximum, minimum and average number of matches for 15 bags as a function of angular separation

B. Synthetic results

Fig. 7 shows ground truth image and its synthetic view by the proposed algorithm at 2º where the ROI, highlighted by red rectangle, is magnified in figure 8 for further discussion.

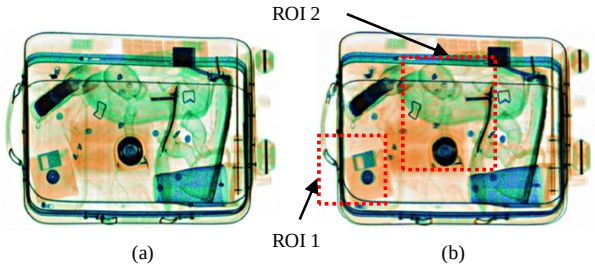


Fig. 7 (a) Ground truth view, (b) synthetic view generated at 2^0 separation between views

Images in figure 8 and 9 are arranged to conveniently enable the quality comparison of the synthetic results with regard to the ground truth.

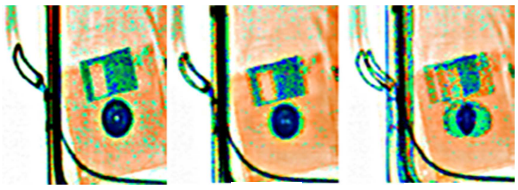
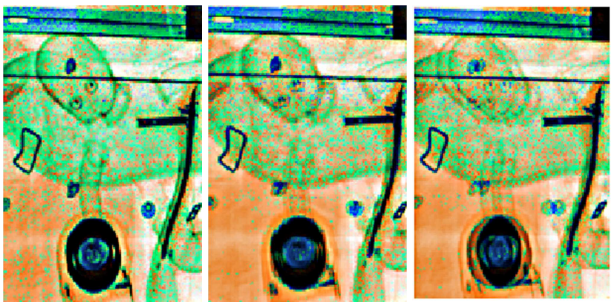


Fig. 8 (a) Ground truth ROI 1, (b) and (c) are the synthetic views at 2^0 and 6^0 respectively

It is notable that the synthetic algorithm is capable of preserving good edges and is also capable of preserving textured structure. It can also be observed that the hub ring of a floppy disk is blurred when the angle between views is increased. It is interesting to note that this blurred artifact has been highlighted in green color due to material segmented morphing. It can be appreciated from figure 8 (b) and (c) that the synthetic view deteriorates as the angle separation between views is increased. Fig. 9 is another region of interest extracted from figure 7 (b).



(a) (b) (c)

Fig. 9 (a) Ground truth ROI 2, (b) and (c) are the synthetic views at 2^0 and 6^0 respectively

Image of figure 9 (a) represents good example of overlap structure where the baby toy is overlapped with heels shoe, metal horizontal line and other mixture material shown in green. On one hand, the synthetic results of this portion have shown appreciated results in terms of shape preservation 'metal material embedded in the heels shoe', vertical and horizontal lines. Synthetic views of ROI 2 have also indicated that synthetic algorithm is capable of preserving object shapes in terms of curve edges. On the other hand, some objects edges tend to blur when the angle separation between views is increased. This observation is highlighted in the edges of the ring of the speaker 'shown in black at the bottom of the image' as well as the face of the toy in figure 9 (c).

The absolute pixel error values recorded for 15 bags are presented as a function of the angular separation between views in the graph of figure 10.

It is noticeable from figure 10 that the number of pixel errors increase as the angular separation between the views is increased. This effect is expected as the 'generating' perspective images will exhibit increasing variations in shape, overlap and pixel intensity. For this reason the absolute amount of errors may be highly sensitive to image content and the normalization (as it is discussed in Section III *D*) of results enables trends in the data to be readily identified. This observation is further illustrated in figure 11 where the pixel error values are normalized with respect to the values obtained at 2^0 of angle separation between views.

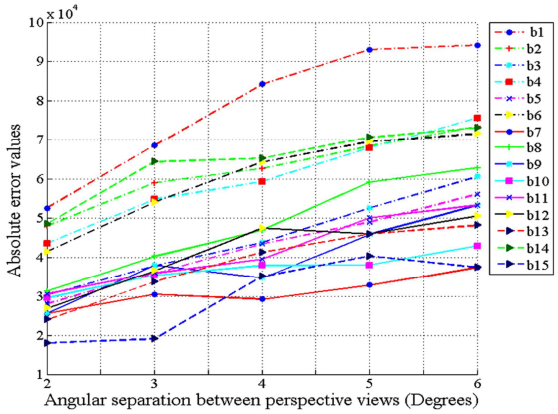


Fig. 10 Absolute pixel error for 15 bags at 2º, 3º, 4º, 5º and 6º of separation

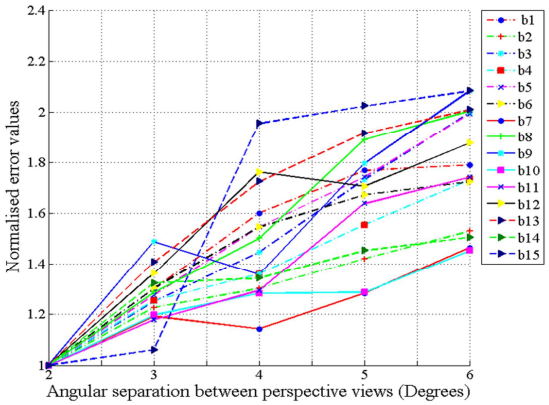


Fig. 11 Normalized pixel error for 15 bags at 2º, 3º, 4º, 5º and 6º of separation

The upper and lower bounds of the maximum and minimum as well as the average number of errors recorded for any of the 15 bags are also illustrated graphically in figure 12 for completeness.

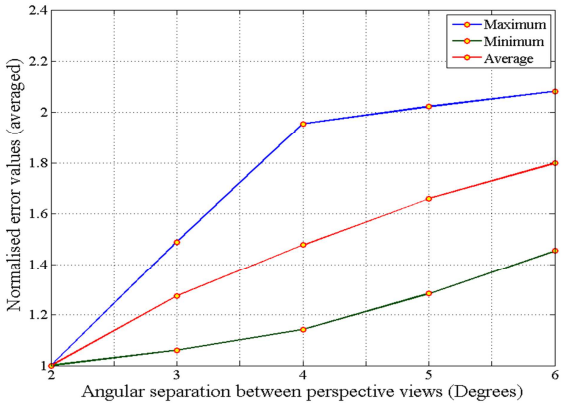


Fig. 12 Maximum, minimum and average number of normalized pixel errors for 15 bags as a function of angular separation

V. Conclusion and Future Work

A novel multiple view technique incorporating image synthesis is presented. The potential of SIFT to locate correspondences in X-ray images is established and quantified for 15 different bags. The performance of SIFT is significantly enhanced by applying two additional criteria namely; a disparity window and an epipolar line constraint. The appropriateness of these criteria is supported by the matching results organised in Section (IV *A*). Material segmentation morphing is proposed to generate intermediary views. This approach is supported by the results reported in Section (IV *B*). Overall, the error analysis indicates that the fidelity of the synthetic imagery is adversely affected by the variations in shape, overlap and material classes occurring in increasingly disparate 'generating views'. The implications for this work are significant as for example, it renders the middle X-ray sensor array in a group of 3 redundant. Therefore, in the case of a 29-view system the overall requirement reduces to 15 X-ray detector arrays in combination with the ability to synthesize the remaining 14 views. This approach will increase the cost effectiveness of systems capable of producing high quality KDEX imagery. Future work will explore an increased sample luggage set to assist in optimizing the algorithmic approach. In due course, high speed implementation of the algorithm will also be required to support operation in the field.

Acknowledgment

The EPSRC Grant (EP/F017596/1) funds this research in collaboration with the Centre for Applied Science and Technology, CAST (formally the UK HOSDB) Sandridge, and the US DHS: Transportation Security Laboratory Human Factors Program. The authors would like to thank Professor Dick Lacey Chief Scientist CBRN of the UK HOSDB and Dr Josh Rubinstein of the US DHS for their ongoing support of this programme of research.

References

[1] J.P.O. Evans, "Stereoscopic imaging using folded linear dual-energy X-ray detectors," *Institute of Physics (IOP) Journal Meas. Sc. & Tech*, vol. 13, no. 9, pp. 1388-1397, 2002.

[2] J.P.O. Evans, M. Robinson, and S.X. Godber, "A new stereoscopic X-ray imaging technique using a single X-ray source: theoretical analysis," *Non-Destructive Testing and Evaluation Int. (NDT&E)*, vol. 29, no. 1, pp. 27-35, 1996.

[3] J.P.O. Evans, and M. Robinson, "Design of a stereoscopic X-ray imaging system using a single X-ray source," *Non-Destructive Testing and Evaluation Int. (NDT&E)*, vol. 33, no. 5, pp. 325-332, 2000.

[4] J.P.O. Evans and H. W. Hon, "Dynamic stereoscopic X-ray imaging," *J. of NDT&E*, vol. 35, no. 5, pp. 337-345, 2002.

[5] J.P.O. Evans, Y. Liu and J. W. Chan, "Depth from motion 3D X-ray imaging for security screening," *in IEE ICDP*, pp. 5-8, 2005.

[6] J.P.O. Evans, Y. Liu, J. W. Chan, and D. Downes, "View synthesis for depth from motion 3D X-ray imaging," *Pattern Recognit. Lett*, vol. 27, no. 15, pp. 1863–1873, 2006.

[7] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[8] H. Moravec, "Rover visual obstacle avoidance," *International Joint Conference on Artificial Intelligence Vancouver, Canada*, pp. 785-790, 1981.

[9] C. Harris, "Geometry from visual motion," *Active vision, MIT Press*, pp. 263- 284, 1993.

[10] Z. Zhang, R. Deriche, O.Faugeras, and Q.T. Luong, "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry," *Artificial Intelligence*, vol. 78, no 1-2, pp. 87-119, 1995.

[11] K. Mikolajczyk, and C. Schmid, "An affine invariant interest point detector," *European Conference on Computer Vision*, pp. 128–142, 2002.

[12] M. Brown, and D.G. Lowe, "Invariant features from interest point groups," *British Machine Vision Conference, Cardiff, Wales, Citeseer*, pp. 656-665, 2002.

[13] Y. Ke, and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," *Conference on Computer Vision and Pattern Recognition, Washington, USA*, pp. 511-517, 2004.

[14] L. Juan, and O. Gwon, "A Comparison of SIFT, PCA-SIFT and SURF," *International Journal of Image Processing (IJIP)*, vol. 3, no. 4, pp. 143-152, 2009.

[15] H. Bay, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, 2008.

[16] Z. Feng, B.Yang, Y.Chen, Y. Zheng, and T. Xu, "Features extraction from hand images based on new detection operators," *Pattern Recognition*, vol. 44, no. 5, pp. 1089-1105, 2010.

[17] J. W. Chan O. Abusaeeda, J.P.O. Evans, D. Downes, X. Wang and L.Yong, "Feasibility of SIFT to Synthesise KDEX Imagery for Aviation Luggage Security Screening," *IET The Crime and Security Conference: Imaging for Crime Detection and Prevention*, ISBN 978-1-84919-207-1, 2009.

[18] S. Seitz, and C. Dyer, "View morphing," *the 23rd annual conference on Computer graphics and interactive techniques ACM SIGGRAPH*, pp. 21–30 1996.

[19] E. Shechtman, A. Rav-Acha, M. Irani, and S.M. Seitz, "Regenerative morphing," *IEEE computer society conferene on Computer Vision and Pattern Recognition, San Francisco, CA, USA*, pp. 1-8, 2010.

[20] B. Zitova, and J. Flusser, "Image registration methods: a survey," *Image Vision Comput*, vol. 21, no. 11, pp. 977-1000, 2003.

[21] C.A.Glasbey, and K.V. Mardia, "A review of image-warping methods," *Journal of applied statistics*, vol. 25, no. 2, pp. 155-172, 1998.

[22] G. Wolberg, "Image morphing: a survey," *The Visual Computer*, vol. 14, no. 8, pp. 360-372, 1998.

[23] Z. Zhang, L. Wang, B. Guo, and H.Y. Shum, "Feature-based light field morphing," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 457-464, 2002.

[24] R. Manning, and C. Dyer, "Interpolating view and scene motion by dynamic view morphing," *Proc. CVPR*, pp. 388-394, 1999.

[25] J.P.O. Evans, Y. Liu and J. W. Chan, "Depth from motion 3D X-ray imaging for security screening," *Conference, in IEE ICDP*, pp. 5-8, 2005.

[26] Wang T. W. and Evans J. P. O., "Stereoscopic dual-energy x-ray imaging for target materials identification", *Conference, in IEE Proceedings Vision Image Processing*, Vol. 150, no. 2, pp. 122-130, April 2003.

[27] J. Gomes, L. Darsa, B. Costa, and L. Velho, "Warping and Morphing of Graphical Objects," *Morgan Kaufmann, San Francisco, Calif.*, (1998).

[28] Z. Zhang, L. Wang, B. Guo, and H.Y. Shum, "Feature-based light field morphing," *ACM Transactions on Graphics*,Vol. 21, no. 3, pp. 457-464, 2002.