Journal of Academic Research
Vol 28, Issu 2, 2024

Framework for Feature Matching in Synthetic Multi-View X-ray Imaging Using SIFT

Eltohami Muftah Elghoul, Bashir Milad Elkharraz, and Afif Saleh Abugharsa

e.elghoul@it.misuratau.edu.ly, bashir.elkharraz@it.misuratau.edu.ly, afif.abugharsa@gmail.com
Lecturers at Faculty of Information Technology, Misurata University

Article information **Abstract**

Key words

- SIFT (Scale Invariant Feature Transform)
- X-ray imaging
- Feature matching
- KDEX (Kinetic Depth X-ray imaging)
- Stereoscopic imaging

Received **02**/07/ 2024, Accepted **15** / **07** / 2024, Available online **20** / **07** /**2024**

Abstract:
This paper presents a framework for evaluating the performance of the Scale Invariant Feature Transform (SIFT) algorithm in matching features within synthetic multi-view X-ray images. The research supports the development of the kinetic depth X-ray imaging (KDEX) technique, which enhances luggage scan interpretation at airport security checkpoints.
Two experiments were conducted with identical input imagery. The first experiment applied SIFT directly to color-coded X-ray views, while the second experiment involved segmenting the imagery into material classes before applying SIFT to each class. The matching results from both experiments were analyzed to assess the algorithm's effectiveness as the angular separation between X-ray views increased. This study lays the groundwork for further research into automated feature matching in security screening applications, offering a promising approach to improving X-ray image interpretation.

1. Introduction

From the standpoint of signal and image processing, a rather well-defined problem is to identify the appropriate features in the actual X-ray image and find the methods and tools to quantify these features. Feature matching has been a key area of research in computer vision and image processing, and comparisons of synthetic and actual images have provided a controlled setting for verification of various feature extraction and matching algorithms. While advances in these algorithms have potential widespread impact, the medical imaging community is a natural beneficiary, and the development of feature-based tools for image assessment and diagnosis is another active research area. [1]
Therefore, to address the problems faced by the 2D X-ray system technology, researchers and security experts are proposing 3D X-ray (multi-view X-ray) as a new technology for bag screening. The 3D X-ray system provides more detailed information not seen in a 2D system, such as hidden within bags. This technology also offers the advantage of reducing the human interface interpretation during the check. Moreover, this innovative technology also reduces bag handling, improving security via checkpoint throughput and reducing the need for manual bag checks[2] .
Maintaining security has become an essential part of air travel safety due to the increasing number of air passengers, air cargo, and terrorism activities. To date, the technology of

e.elghoul@it.misuratau.edu.ly, bashir.elkharraz@it.misuratau.edu.ly, afif.abugharsa@gmail.com

screening bags is performed by a two-dimensional (2D) X-ray system. However, this technology is unable to provide information about the overlapping layers of items in a bag. Due to the inability of the 2D X-ray system to provide information on overlapping layers of items, it has led to errors and security breaches in identifying prohibited and harmful items in the bag. Besides, the current 2D X-ray system requires a longer waiting time for detailed check. This will reduce the effectiveness of the security check of bag screening as the human eyes cannot distinguish items within a short period[3] .

Interpretation of X-ray images by humans primarily relies on 3D information, which is likely difficult to perceive under monoscopic settings. According to one preeminent viewpoint, the object recognition limitations of X-ray security agents result from the lack of clear depth perception in 2D X-ray images. Should 3D (stereo or stereoscopic) depth perception be prominently more enabled, threat detection results would ostensibly improve. When considering both the perception and anticipation aspects of threat detection, the law enforcement and military system effectiveness are fundamentally augmented in their operating experiences through highly-developed tangible performance improvements. [4] Consequently, airport checkpoint security outcomes could be significantly enhanced. The addition of 3D data would permit analysts to make more informed decisions by enabling depth perception, reducing the number of hidden details, and mitigating the accumulation of visual search displays, ultimately culminating in efficacious security screening implementations[5] .

X-ray technologies that are widely used in the aviation security setting include backscatter and transmission methods. As of 2019, the security screening process involving X-ray scanners primarily takes monocular 2D images. The automated analysis of these 2D images is a critical component for ensuring that the machines' capabilities do not result in significantly lower security effectiveness, leading to hazardous items not being found. In baggage screening, by default, the checkpoint operator will examine the 2D images with monocular views of the X-ray scans. These security agents have to consistently maintain their high levels of attention and make the correct and instant decisions without making errors while examining a huge amount of luggage in a very short time to secure the aviation transportation network[6] .

For luggage inspection, the challenge lies in the high transmission (penetration) of X-ray beams. Available gamma-ray densitometers are used to inspect big items instead. The use of X-ray beams for small-item inspection is aimed at detecting dangerous objects, hazardous items, bombs, weapons, and narcotics concealed in hand luggage, though not visible through penetration cameras. At present, operators would provide a few X-ray images with different perspective views for the observation and analysis of suspicious items in hand luggage. However, X-ray images are a two-dimensional projection of the luggage, which does not provide three-dimensional (3D) information about the location and orientation of the suspicious items in the luggage. This is a significant disadvantage in luggage inspection for the present security control system. Although one may obtain a 3D profile of the pocket (contour) on X-ray imaging, it requires a search of every single pixel, a process that is slow and monotonous in the eyes of the operator/reviewer. This is especially true when the luggage volume increases rapidly due to the growth in the number of passengers traveling through airports. The human factor underlined here can be present operational delays. [7]

2. Stereoscopic X-ray imaging

Binocular stereoscopic X-ray imaging provides airport security personnel with an effective detection tool. The principles of stereoscopy are rooted in the fundamentals of the human visual system. Implementing these techniques allows security staff to differentiate the spatial characteristics of each item within scanned baggage. Additionally, the university team has developed methods for generating color-coded imagery, facilitating the identification of various materials penetrated by the X-ray beam. Consequently, binocular stereoscopic systems can substantially enhance the screener's comprehension of the true nature of the inspected items[8] .

As depicted in Figure 2.1, the developed stereoscopic system utilizes a single X-ray source, two linear dual-energy X-ray arrays, and two CCD cameras. The pair of collimated X-ray beams which configured to radiate towards the linear detector arrays. The intersecting fields of view of the stereoscopic system create the inspection volume. The overlapping fields of view (FOV) of the left and right image-detecting elements form a volumetric element, or voxel. Figure 2.2 illustrates the angle σ between the two crossing lines of sight of the X-ray detecting elements, where δz represents the separation between depth planes, and δx denotes the motion axis resolution. [9]

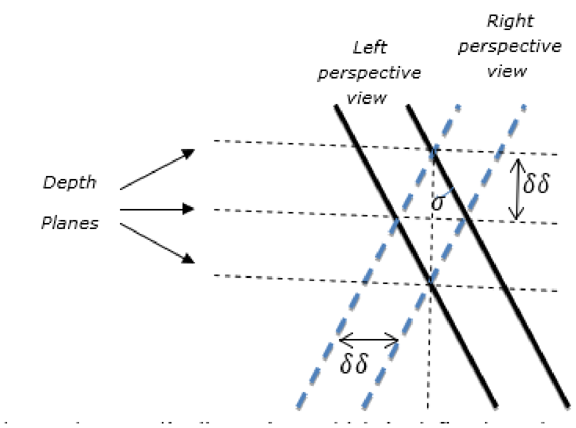

Figure 2.1 Stereoscopic X-ray system with single source, dual arrays, and CCD cameras.

Recently, a combination technique has been developed to integrate binocular stereoscopic imaging with the motion or kinetic depth effect (KDE). A vivid appearance of depth can be produced by moving or rotating the inspected object, achieved by merging a set of different views of the same object. This integration allows viewing the object from multiple viewpoints, revealing features that are not clear in a two-dimensional view[10] .

The KDEX technique has been implemented using two X-ray machines to produce image sequences: the flatbed multiple views X-ray machine and the folded array multiple view X-ray machine. This project focuses on finding feature matches for digital X-ray images captured by the folded array linear X-ray scanner (see Figure 2.2) [11] .

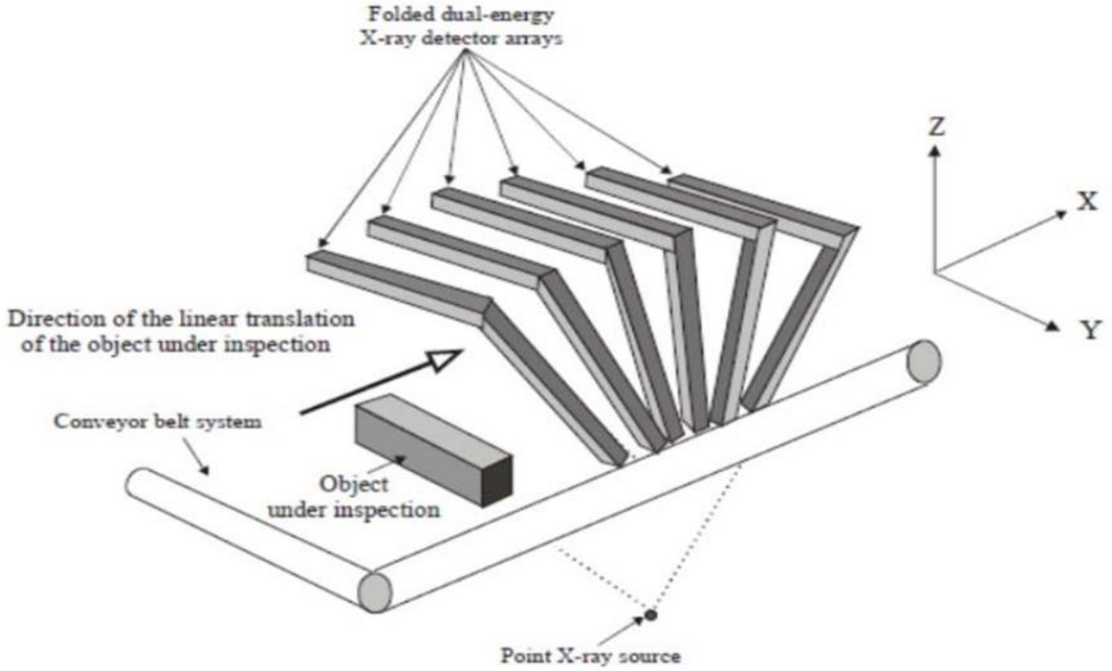

Figure (2.2) an illustrative for a multi-sensor array X-ray scanner.

The scanner, depicted in Figure 2.2, features an inspection tunnel with dimensions of 40 cm x 60 cm, allowing for the scanning of full-sized baggage. The X-ray sensor array is folded around the conveyor belt to facilitate the collection of multiple views from different perspectives. This experimental machine was constructed at the Home Office Scientific Development UK laboratories according to a design proposed to investigate depth from motion binocular stereo imaging.

The process of creating an image from some form of image description is known as image synthesis, which can generally be classified into model-based or image-based techniques. Model-based techniques involve computing and manipulating a three-dimensional mathematical representation of the scene, while image-based techniques use the matching process to address correspondences between two or more images of the same scene point . [12]

The correspondence problem involves finding a set of points in one image that can be recognized as the same points in another image, typically occurring in stereo correspondence. Two main approaches to finding correspondence between images are correlation-based, which checks if one position in the first image is similar to another position in the second image, and feature-based, which focuses on identifying features in the image and determining if the layout of a subsection of features is the same in both images. [13][14]

Correlation-based approaches rely on statistical correlations between local subsections of images to measure similarity. Various types of statistical correlation are implemented between intensity or color patterns in the local support window. These methods have been effectively applied to stereo images with good features and smooth surface variation. However, correlation-based methods assume that all elements in the local window have the same depth, an assumption that is violated at depth discontinuities.

In feature matching, algorithms start by extracting noticeable primitives in the images, such as edges or contours, and then match them in two or more views. Two different maps describe the image: primitives that define the nodes and geometric relations that define the links. The registration of the two maps becomes the mapping of the two graphs. To reduce complexity, heuristic techniques are often used, such as assuming affine transformation. Because only a small portion of the image pixels is considered, these techniques are relatively fast. However, they may fail if the selected primitives are not adequately detected in the image[15] .

One of the most widely used techniques in feature matching is the point matching method, which can be categorized into two approaches: intensity distribution-based methods and gradient distribution-based methods. Descriptors are employed to represent the intensity distribution or gradient distribution in local image regions. Previous works in 1995 and 1997 utilized traditional methods, such as cross-correlation and relaxation, to identify an initial set of matches. Subsequently, a robust technique, the Least Median of Squares (LMedS), was used to eliminate incorrect matches from this set. [16] While these descriptors are based on intensity distribution, the spin image represents a more distinctive descriptor of this type. It is a data-level shape descriptor used for matching surfaces represented as surface meshes by aligning points.

D.G. Lowe proposed the Scale Invariant Feature Transform (SIFT), which is considered one of the most powerful gradient distribution-based descriptors. In this approach, local image gradients are measured at a specific scale in the region around each keypoint. A similar variant, based on shape context, involves calculating similarity after solving for correspondences between points in the two images. These correspondences are then used to approximate an aligning transform[17] .

The SIFT approach converts image data into scale-invariant coordinates related to local features and minimizes the cost of feature extraction through cascade filtering. The main computational steps for creating a set of features in the SIFT approach are as follows:

- Scale-space extrema detection.
- Keypoint localization.
- Orientation assignment.
- Keypoint descriptor.

3. Proposed framework

All matching techniques discussed in section two are applied to address the correspondence problem in visible light images. However, the correspondence problem in multiple-view X-ray images is similar to that in visible light images, though the transparency of X-ray images exacerbates the issue.

This project is part of ongoing efforts to develop the KDEX technique by generating synthetic views to reduce hardware requirements. The most significant challenge in image synthesis is the correspondence problem, a crucial aspect of many computer vision problems. It involves identifying correspondences between two different images obtained from different angular views.

As highlighted in the literature review, the Scale Invariant Feature Transform (SIFT) has recently played a crucial role in addressing the correspondence problem. Consequently, The objective of this research is to utilize SIFT as a method for matching images in order to identify similar features across different perspectives. Specifically, this technique will be used to analyze color-encoded X-ray images. The matching process will be structured into two primary stages.

In the first phase, the SIFT algorithm will be applied to color-encoded X-ray images, where the three material classes are visualized within the same image (see Figure 3.1).

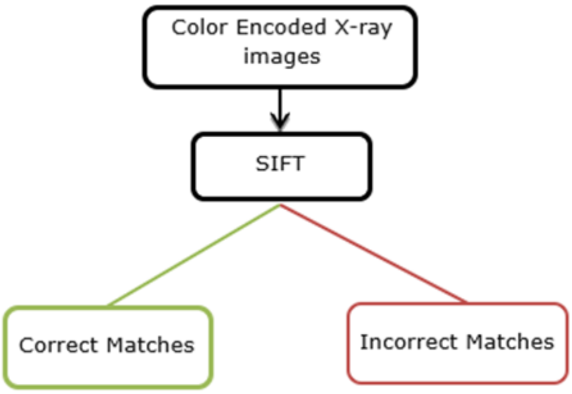

Figure (3.1) SIFT algorithm will be utilized on color-encoded X-ray images

In the second phase, as illustrated in Figure 3.2, the color-encoded X-ray images will be segmented into four distinct images: Organic Material Image, Mixture Material Image, Metal Material Image, and Greyscale Image. SIFT will then be applied individually to each material type. Subsequently, all correct and incorrect matches will be collected and compared.

These two procedures will be applied to 30 sets of color X-ray images, with each set representing the structure of a unique bag. To evaluate the algorithms, the angular separation between views will be increased to cover up to 6 degrees. Additional procedures will include combining the correct and incorrect matches and visualizing the behavior of the matching algorithms.

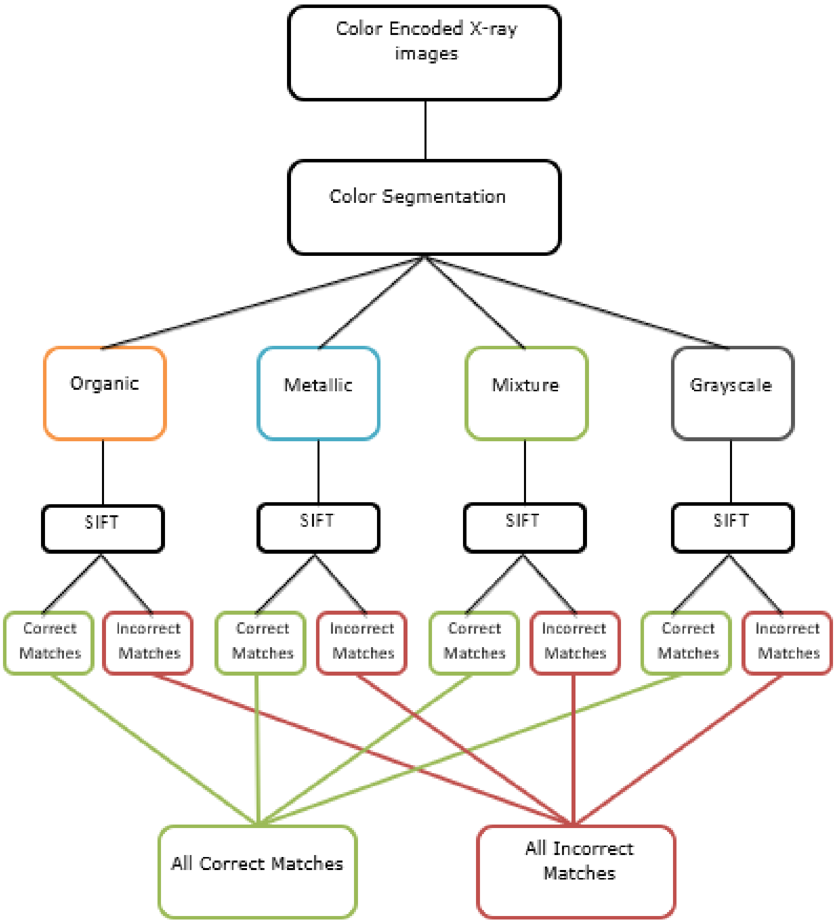

Figure (3.2) Color-encoded X-ray images are divided into four separate images

4. Conclusion

We anticipate the algorithm will be proficient in matching features. This research examined the invariant feature transform (SIFT) as a potent method for feature matching in synthetic X-ray images from multiple viewpoints. By categorizing color-coded X-ray images into distinct material classes and then employing SIFT, its effectiveness in identifying corresponding features across different views is evident.

The experiments can also be applied to a series of X-ray images, each depicting unique luggage structures, show casing the algorithm's capability to handle angular separations of up to 6 degrees. The proposed approach underscores the importance of resilient feature matching methods in enhancing the interpretation and analysis of X-ray images in security screening applications. Future research efforts can further refine these methods and investigate enhancements necessary for effectively handling transparent and complex materials.

Overall, integrating SIFT into multi-view X-ray imaging would demonstrates potential in developing automated detection and inspection capabilities in security settings, thus contributing to safer and more efficient inspections.

References:

[1] M. Elwan, A.S. Amein, A. Mousa, et al., "SAR image matching based on local feature detection and description using convolutional neural network," Security And ..., vol. 2022, Wiley Online Library, 2022. wiley.com

[2] D. Mery, D. Saavedra, and M. Prasad, "X-ray baggage inspection with computer vision: A survey," Ieee Access, 2020. ieee.org

[3] A. Chang, Y. Zhang, S. Zhang, L. Zhong, et al., "Detecting prohibited objects with physical size constraint from cluttered X-ray baggage images," Knowledge-Based Systems, vol. 2022, Elsevier, 2022. [HTML]

[4] M. Rafiei, J. Raitoharju, and A. Iosifidis, "Computer vision on x-ray data in industrial production and security applications: A comprehensive survey," Ieee Access, 2023. ieee.org

[5] M. Mobini, M. Sabzehparvar, "Visual landing based on the human depth perception in limited visibility and failure of avionic systems," Computational Intelligence and ..., Wiley Online Library, 2022. wiley.com

[6] BKS Isaac-Medina, CG Willcocks, et al., "Multi-view object detection using epipolar constraints within cluttered x-ray security imagery," in 2020 25th International Conference, 2021. worktribe.com

[7] V. Ponnusamy and D. Ekambaram, "Image analysis approaches for fault detection in quality assurance in manufacturing industries," Computational Intelligence based ..., vol. 2024, Taylor & Francis, 2024. [HTML]

[8] A. Cordova, "Technologies for primary screening in aviation security," Journal of Transportation Security, 2022. [HTML]

[9] X. Ou, X. Chen, X. Xu, L. Xie, X. Chen, Z. Hong, H. Bai, et al., "Recent development in x-ray imaging technology: Future and challenges," Research, 2021. science.org

[10] D. Gradolewski, D. Dziak, D. Kaniecki, A. Jaworski, et al., "A runway safety system based on vertically oriented stereovision," Sensors, vol. 2021, mdpi.com, 2021. mdpi.com

[11] A. Ali and I. Mohamed, "A Survey Of X-Ray Multiview Object Detection," Int. J. Innov. Res. Adv. Stud, 2020. ijiras.com

[12] Z. Qin, Z. Liu, P. Zhu, and Y. Xue, "A GAN-based image synthesis method for skin lesion classification," Computer Methods and Programs in Biomedicine, vol. 185, Elsevier, 2020. [HTML]

[13] P. Zhang, B. Zhang, D. Chen, et al., "Cross-domain correspondence learning for exemplar-based image translation," in Proceedings of the ..., 2020, openaccess.thecvf.com. thecvf.com

[14] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," International Journal of Computer …, vol. 2021, Springer, 2021. springer.com

[15] Y. Zhao and P. A. Vela, "Good feature matching: Toward accurate, robust vo/vslam with low latency," IEEE Transactions on Robotics, 2020. ieee.org

[16] C. Li, W. Jia, and D. Wei, "Study on Elimination Algorithms for Line Segment Mismatches," Canadian Journal of Remote Sensing, 2022. [HTML]

[17] EG Andrianova, LYA Demidova, "An approach to image matching based on sift and orb algorithms," 2021 3rd International …, 2021. [HTML]

إطار عمل لمطابقة الميزات في التصوير بالأشعة السينية الاصطناعية متعددة المشاهدة باستخدام SIFT

التهامي مفتاح الغول بشير ميلاد الخراز عفيف صالح أبوغرسة
كلية تقنية المعلومات، جامعة مصراتة، ليبيا

الملخص

يقدم هذا البحث إطاراً لتقييم أداء خوارزمية تحويل الميزات المتناسبة مع الحجم (**SIFT**) في مطابقة الميزات داخل صور الأشعة السينية الاصطناعية ذات المناظر المتعددة. يدعم البحث تطوير تقنية التصوير الإشعاعي ثلاثي الأبعاد (**KDEX**) الذي يعزز تفسير فحص الأمتعة في نقاط تفتيش أمنية بالمطارات.

يسعى الباحثون لاجراء تجربتان باستخدام صور متطابقة كإدخال. ففي التجربة الأولى، يتم تطبيق تحويل **SIFT** مباشرةً على صور الأشعة السينية الملونة، بينما في التجربة الثانية، يتم تقسيم الصور إلى فئات مواد قبل تطبيق **SIFT** على كل فئة. واستطلعت نتائج المطابقة من كلا التجربتين لتقييم فعالية الخوارزمية مع زيادة الفجوة الزاوية بين مناظر الأشعة السينية والتي يتوقع انها نتائج معتبرة.

يضع هذا الدراسة الأساس لإجراء مزيد من البحوث في مجال مطابقة الميزات التلقائية في تطبيقات فحص الأمان، مما يقدم نهجاً واعداً لتحسين تفسير صور الأشعة السينية.

استلمت الورقة بتاريخ 2024/7/002، وقبلت بتاريخ **15** / **07**/2024 ، ونشرت بتاريخ **20** / **07** /**2024**.

الكلمات المفتاحية:
• تحويل الميزات المتناسب بالحجم (SIFT)
• تصوير الأشعة السينية
• تقنية التصوير الإشعاعي ثلاثي الأبعاد الحركي (KDEX)