VIEW SYNTHESIS FOR DEPTH FROM MOTION 3D X-RAY IMAGING

YONG LIU

A thesis submitted in partial fulfilment of the requirements of
Nottingham Trent University for the degree of Doctor of Philosophy

School of Science and Technology
Nottingham Trent University,
Burton Street,
Nottingham, UK.

Collaborating Establishment:
Home Office Scientific Development Branch (H.O.S.D.B),
Sandridge, UK.

January 2009

Copyright statement

Abstract

The depth from motion or kinetic depth X-ray imaging (KDEX) technique is designed to enhance the luggage screening at airport checkpoints. The technique requires multiple views of the luggage to be obtained from an arrangement of linear X-ray detector arrays. This research investigated a solution to the unique problems defined when considering the possibility of replacing some of the X-ray sensor views with synthetic images. If sufficiently high quality synthetic images can be generated then intermediary X-ray sensors can be removed to minimise the hardware requirements and improve the commercial viability of the KDEX technique. Existing image synthesis algorithms are developed for visible light images. Due to fundamental differences between visible light and X-ray images, those algorithms are not directly applicable to the X-ray scenario. The conditions imposed by the X-ray images have instigated the original research and novel algorithm development and experimentation that form the body of this work.

A voting based dual criteria multiple X-ray images synthesis algorithm (V-DMX) is proposed to exploit the potential of two matching criteria and information contained in a sequence of images. The V-DMX algorithm is divided into four stages. The first stage is to aggregate matching cost among input images. Subsequently, a novel voting approach is developed for electing the "best" disparity prior to generation of synthetic pixels. A void filling routine is applied to complete the synthetic image generation.

A series of experiments, using real acquired images, investigated the fidelity of the synthesised images resulting from application of the V-DMX algorithm as a function of several parameters: number of input images, matching criterion, method of handling multiple images and X-ray beam separation. The performance measure is based on counting the number of pixel errors in the synthetic images relative to the ground truth images.

The V-DMX employs the widely adopted sum of squared differences (SSD) criterion and a novel criterion, which is derived from the laminography technique, termed laminography intensity (LamI). SSD is shown experimentally to have poor performance when the image contains repeating features, discontinuities and overlapping regions. While the overall performance of the LamI is found to be weaker than SSD, LamI consistently outperformed SSD in discontinuity and overlapping regions. This has spurred the use of LamI as a complement to SSD. Integration of the two criteria has demonstrably produced better results than using solely either of the criteria.

Limitations of the algorithm are assessed by increasing the angular separation between X-ray beams used to produce the perspective X-ray images. The resultant image fidelity degraded as the angular separation increases. This result was expected because the increase in angular separation meant a concomitant increase in images' dissimilarity and disparity window. Empirical evidence demonstrated that synthetic images may be satisfactorily produced by processing images produced by X-ray beams separated by angular increments up to 6º. This result is based on comparing the algorithm performance for four beam separations, which are 4°, 6°, 8° and 10°. This finding reveals that, for example, a 32-view X-ray scanner with 1° beam separation may be scaled down to a 7-view system with at least the same angular coverage.

The encouraging result has formed a basis for further research to extend the current algorithmic approach to the use of dual-energy X-ray data. The practical performance of the algorithm will be evaluated by conducting human factors investigation in collaboration with the US Department of Homeland Security.

Acknowledgement

I would like to extend my sincere gratitude to my director of studies, Professor Paul Evans for his guidance and invaluable support throughout the duration of the research programme. I would also like to convey my sincere appreciation to my second supervisor Dr. Richard Germon for his generous help.

Sincere thanks must go to Dr. Jer Wang Chan for his continuous generous help, guidance and encouragement.

Many thanks go to Dr. David Downes for helping to collect the experiment data and his positive impact on my thesis.

I would like to give my deepest appreciation to my mother Shu Ying Wang and my father Wen Zhen Liu for their unwavering support, without which, studying in UK would have been an impossible dream for me. Similarly my two brothers have provided me with continuous encouragement. I would like to dedicate this thesis to my grandmother who past away during the third year of my PhD degree in 2006; I know that she is blessing me, all of the time, from Heaven.

Last but not least, I would like to thank all of my friends that have supported me throughout my studies.

Table of contents

List of figures

List of tables

Chapter One Introduction

1.1 Background

The research programme seeks to improve the luggage screening at airport checkpoints. The unique arrangement of items in a "typical" suitcase or carry-on bag makes the interpretation of these images by X-ray machine security personnel extremely difficult. Under high volume and stressful working conditions, their screening task is exacerbated by the lack of visual cues to depth in an image that has been produced by transmitted radiation. The global government bodies have identified that the current way forward is the development of systems which produce a so-called "best image first" [10]. Academics at Nottingham Trent University have responded to this ethos by collaborating with the UK Home Office. Preliminary human factors work conducted by the Federal Aviation Administration (FAA) and the Defence Evaluation and Research Agency (now QinetiQ) in the UK has confirmed that the 3D images have important implications for improving the speed and the efficiency of security checks [15].

Past work by the University team in collaboration with the Home Office Scientific Development Branch (HOSDB) UK, over a twenty-year period, has produced a novel binocular stereoscopic X-ray technique, Evans [13,14], to aid the detection and identification of objects in X-ray scans of luggage. Imaging technology based on this early work is now commercially available [10]. More recently the University team in collaboration with the US Department of Homeland Security (DHS) and the Home Office Scientific Development Branch UK has developed X-ray imaging techniques that combine binocular stereoscopic imagery with motion or kinetic depth effects. The technique produces "moving" binocular stereoscopic image sequences from a static configuration of multiple line-scan X-ray sensors and a single X-ray source [1,2,3,4,5]. Movement or rotation of an object relative to the observer can produce a vivid depth in a two dimensional display. The ability effectively to "look around" an object under inspection using multiple views is particularly advantageous in comparison with standard stereoscopic techniques [7,8,9,10]. Linear "motion parallax" refers to the differential angular velocities of retinal images of points moving laterally with the same speed, but at different distances (from the eye in the case of the real world, and from the sensors in the case of the X-ray scanner). Therefore, this effect can be used to produce *motion perspective*. Motion perspective enables a viewer to extract depth information

from continuous movement occurring in a visual display. Interestingly, the depth effect obtained from motion can exceed that produced by the binocular stereoscopic effect.

Motion provides a powerful visual cue to depth which greatly enhances the interpretation of spatially complex structures: termed kinetic depth effect [6] in shadowgraph images and identified as kinetic depth X-ray (KDEX) imaging in this work. However, to produce a smooth image rotation over sufficiently wide angles suitable for security screening applications does require a relatively large number of views (up to 32) [5]. The implementation of such a large number of folded array [7] detectors presents a number of serious practical problems for the construction of the X-ray collimators and configuration of the dual-energy sensor modules. This problem is exacerbated by the small angular increments, of the order of 1º, required between each successive view and the physical bulk of the sensor arrays. These physical constraints currently preclude the development of a "single pass" KDEX imaging system. Therefore, this research programme seeks to establish whether image synthesis [11,12] can be used to compute intermediary views of sufficiently high visual quality to enable the angular separation of the sensors to be increased. If this can be achieved then the world's first dual-energy X-ray scanner producing kinetic binocular stereoscopic dual-energy X-ray images can be realised.

The research programme is aimed at synthesising and introducing high quality intermediary images between successive pairs of practical (or detector derived) images in a sequence of transmission X-ray images. The fundamental requirement of such an image synthesis technique is to develop matching algorithms to solve the correspondence problem. However, the correspondence problem is ill-posed inherently, and is practically unsolvable. This problem is exacerbated by the transparency property of X-ray images, which makes them fundamentally different from visible light images. Therefore, photometric compatibility, continuity and uniqueness constraints that are commonly applied to visible light images are unsuitable for X-ray images.

Among the existing algorithms that use multiple images (i.e. more than two images) to produce a disparity map, there is no clear indication about the optimal number of images needed. The number of images that require processing is a critical practical consideration for the implementation of the KDEX technology, as it will strongly influence the hardware complexity, system cost and operational speed. This research

programme is designed to carefully evaluate the algorithm performance as a function of the number of images required and the angular separation between these images. These aspects are considered further in terms of the total angular coverage achievable as an indicator of the potential cost effectiveness of KDEX technology.

To the author's knowledge, there is no literature concerning image synthesis in a sequence of transmission (dual-energy or polychromatic) X-ray images. Previous work by the university team concerning the automatic extraction of 3D information from stereoscopic X-ray images, Sobania [15], utilised neighbourhood correlation analysis with a Mexican hat filter. However, the depth information extracted was used to render 3D representations of the objects under inspection. A major consideration in this respect is to apply the techniques developed in this research to dual-energy X-ray imaging. It is anticipated that image synthesis incorporating the dual-energy data would significantly increase the computational loading for a real time application, possibly requiring custom hardware. These wider considerations have prompted the authors to look at simple techniques detailed in this thesis as a first step in algorithm development.

To achieve the research aim and objectives set out in this thesis requires that an image synthesis algorithm, which specifically addresses the properties of transmission X-ray images and the associated correspondence problem, be investigated and developed.

1.2 Aim and objectives

The aim of the work is to investigate image synthesis techniques to reduce the hardware required to implement a novel kinetic X-ray imaging technique. To achieve this, the following objectives were set.

- To develop an image synthesis algorithm that utilises multiple perspective images.
- To devise a method to determine the optimal number of images required for producing high quality synthetic images.
- To investigate a new correspondences matching criterion suitable for transmission X-ray images.
- To formulate a performance measure for assessing the resultant synthetic images.
- To evaluate the limitations of individual criterion and combined criteria.

- To evaluate the algorithm performance as a function of the X-ray beam angles used to produce the input images and the total angular coverage achievable.

1.3 Research contributions

A list of the major contributions contained in this thesis is outlined as follows.

A novel matching criterion termed laminogram intensity criterion. This is derived from a novel interpretation of X-ray laminographic data developed in this work.

A weighted based matching cost aggregation method termed "sum of reciprocals" was devised to preferentially accentuate low matching cost values.

A ranking based voting system is developed to resolve the conflict between multiple disparities and thus decide the optimum disparity for pixels under consideration.

An integration method to combine two different but balancing criteria, by virtue of repeating matching process, which underpins the development of the V-DMX.

1.4 Structure of the report

This report consists of eight chapters:

- Chapter One Introduction: presents the background, scope and objectives of the thesis.
- Chapter Two Enabling technologies and theoretical considerations: presents an introduction to various X-ray techniques, overview of the existing image synthesis approaches, and followed by a discussion on considerations that have been taken to address the correspondence problem and X-ray transparency.
- Chapter Three Image synthesis algorithm: presents a detailed description of the developed algorithm and its supporting methodology.
- Chapter Four Experiment considerations and overview: presents the experiment methodology, experiment plan and image input considerations.

- Chapter Five Experiments and analysis: Laminogram intensity (LamI) criterion: reports the experiment results and analysis of the laminogram intensity criterion for different experiment conditions.
- Chapter Six Experiments and analysis; Sum of squared differences (SSD) criterion: reports the experiment results and analysis of the Sum of squared differences (SSD) criterion for different experiment conditions.
- Chapter Seven Development of the V-DMX algorithm: reports the experiment results and analysis to support the development of a V-DMX algorithm.
- Chapter Eight Summary, Conclusions and future Work: summarises the work completed, presents the conclusions and proposes the direction of future work.

Chapter Two Enabling technologies and theoretical considerations

2.1 Introduction

This chapter presents an overview of the enabling technologies and the theoretical considerations necessary for the design and investigation of the synthetic algorithm presented in Chapter 3.

The multiple view (stereoscopic) X-ray imaging techniques previously developed by the university team, and utilised in this work, are discussed briefly. Two different experimental X-ray scanners are presented as the practical source of the perspective image sequences used for the algorithm development and investigation. The basic principles of X-ray laminography are presented as laminograms are implemented to enhance the identification of corresponding or conjugate image points.

A review of computational methods for the identification of corresponding points and features in a sequence of perspective images is presented. Establishing the coordinate position of each synthetic pixel involves searching for corresponding pixels throughout the image sequence. The correspondence problem encountered in X-ray image sequences is ill posed and similar to its visible light counterpart, although the inherent transparency property of X-ray images exacerbates matching problems. Searching for the correspondences is only part of the synthesis technique investigated in this work. To ensure a realistic blend of practical and synthetic images requires that the intensity of each synthetic pixel be carefully computed.

In theory, image synthesis requires solving the correspondence problem as a precursor to establishing the relative location (in the x, y and z-axes) and the orientation (pitch, yaw and roll) of each object in the scene to accurately predict the content of the synthetic image. Additional factors to be considered include the amount of attenuation experienced by the X-ray beam and the properties of the X-ray detectors. The polychromatic nature of the X-ray source, installed in typical airport luggage security scanners, is also an important aspect of the imaging chain under consideration in this work.

Knowledge from enabling technologies has inspired some theoretical considerations prior to the development of synthetic algorithm. Considerations cover several issues related to transparency of X-ray images and some issues related to utilising multiple images for image synthesis.

2.2 Stereoscopic X-ray imaging

In 1895 Roentgen, the discoverer of X-rays, observed shadowgraph images by placing objects between an X-ray source and a fluorescent screen [16]. From this discovery to the present day, X-ray imaging has undergone constant development, particularly in the fields of medicine [17,18,19], non-destructive inspection [20] and security screening [21,22,23,24,25,26,27,28,29]. Modern systems incorporate many different techniques to produce X-ray images.

In aviation security screening, X-rays are routinely used to examine baggage contents to help detect dangerous or illegal items. Stereo techniques have been available to radiologists for decades, J. MacKenzie Davidson introduced stereoscopy to radiology in 1898 [30].

The Manual of photogrammetry [31] defines stereoscopy as:

"…… the science and art that deals with the use of images to produce a three-dimensional visual model with characteristics analogous to those of actual features viewed using true binocular vision."

Binocular stereoscopic X-ray imaging offers an effective screening method, used in airport security. The application of stereoscopy in X-ray imaging originates from the operating principles of the human visual system [32,33]. Binocular parallax is one of the most robust depth cues utilised by a human observer [34,35]. The utilisation of binocular stereoscopic imagery provides security personnel with a mechanism to understand the relative spatial position of cluttered items in X-ray scans of luggage. Previous work by the university team has developed techniques, now commercially realised, which also produce colour coded imagery to enable the observer to discriminate between different types materials. Thus binocular stereoscopic X-ray screening techniques [36,37] can greatly enhance the human observer's understanding of the true nature of the 3D scene under observation.

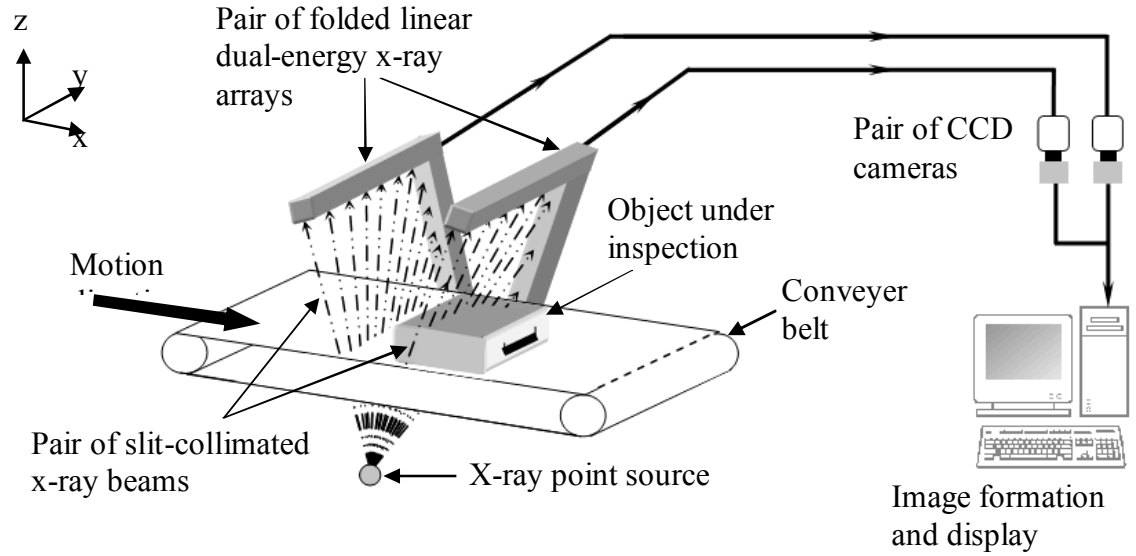


Figure 2.1 The schematic of the binocular stereoscopic folded dual-energy X-ray screening system invented and developed by the university team.

The stereoscopic imaging technique developed by the university team, utilises a single X-ray source, a pair of folded linear dual-energy X-ray arrays and a pair of CCD cameras as illustrated in Figure 2.1. The two slit-collimated X-ray beams are arranged to irradiate a left and a right folded configuration of linear detector arrays. This technique has formed the basis of several commercially available products, manufactured by 3D X-ray Ltd.

The inspection volume formed by the two overlapping fields of view of a stereoscopic system may be considered in terms of the spatial sampling pattern formed by individual sensing elements that comprise the image sensors. The intersecting field of view of a left image sensing element and a right image sensing element form a volume element or a voxel. The whole of the stereoscopic region is comprised of a 3D tessellation of voxels. Figure 2.2 illustrates a voxel as a function of the angle σ between the intersecting lines of sight of two X-ray sensing elements. The size of the voxel in the depth (or z-axis) is δZ and δX is the motion axis resolution. It is readily appreciated from the simple geometry that when the angle σ increases, the minimum detectable increment in object space δZ decreases. However, in binocular stereoscopic systems the maximum allowable σ is limited by the maximum permissible parallax in the resultant display, which in turn is determined by the maximum disparity that can be fused by an observer comfortably [38,39].

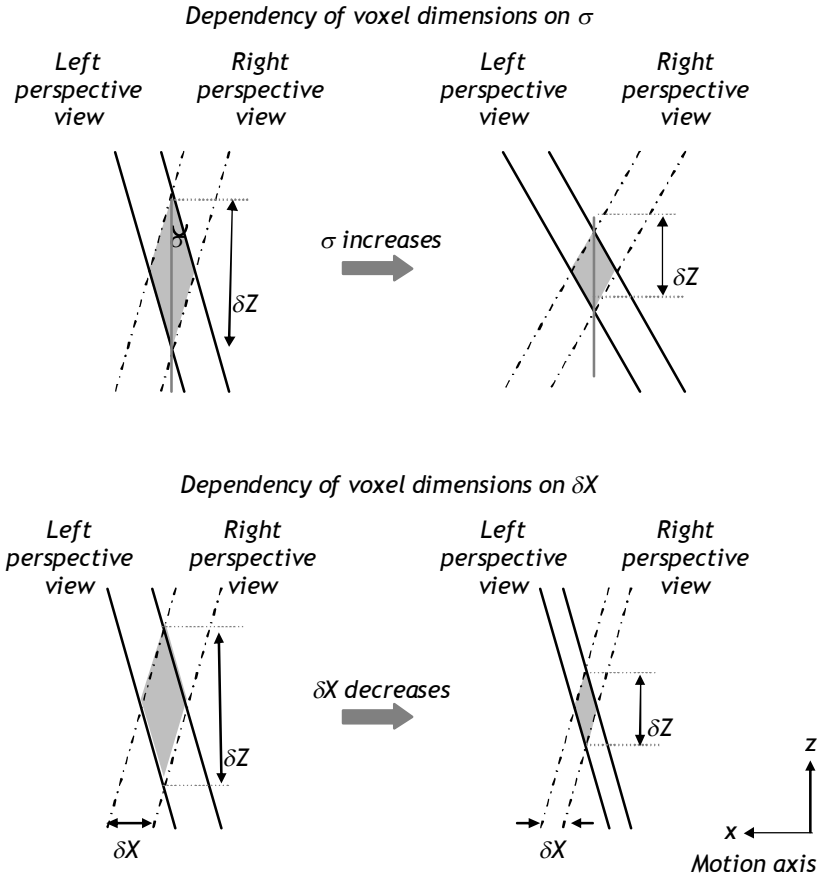


Figure 2.2 Dependency of voxel dimensions on δ and δX.

In stereo matching, the smaller δZ produced the larger the potential disparity window. This effect will increase the depth resolution potentially available but decrease the likelihood of identifying corresponding points, which enable the increase in depth resolution to be realised.

2.3 Multiple view X-ray imaging

More recent research by the university team [2,3] has developed novel techniques, which combine binocular stereoscopic imagery with motion or kinetic depth effects (KDE). Movement or rotation of an object relative to the observer can produce a vivid appearance of depth in a two dimensional display. This effect is achieved by collecting a number of different views (typically from 6 to 32 in this work) of the object under inspection. The resultant perspective images are presented in a specific sequence on a standard video display monitor. The linear "motion parallax", apparent in the displayed image sequence, is due to the differential angular velocities of points moving laterally with the same speed, but at different distances from the sensors in the X-ray scanner. The combination of binocular disparity and motion parallax allows the object to be

viewed from different viewpoints, thus revealing attributes of the object that may not be evident in a simple two-dimensional display.

Two different experimental X-ray machines have been used to produce the image sequences utilised by the research group. This research only uses image sequences produced by X-ray machine described in Section 2.3.2. Each system has a high degree of utility for experiments but is very different in terms of their physical construction and imaging capabilities.

2.3.1 Flatbed multiple view X-ray scanner

Initial research [5] concentrated on simulating complex multiple line-scan X-ray source configurations with an X-ray image intensifier system (see Figure 2.3). In order to store electronically the shadowgraph information projected onto the input window of the image intensifier, the output window is optically coupled to an area array camera. Thus as the object under inspection is translated through the X-ray beam, image information is produced by collecting and storing data from the selected photosite columns on the charge coupled device (CCD) array. The selected column on the area array maps to a line on the 9cm diameter input window of the image intensifier. In this way a novel line-scan system can be produced. In order to produce motion parallax in a sequential display of images requires that each successive pair of perspective images exhibit parallax as a function of range from the perspective centre (i.e. X-ray point source) of the imaging system.

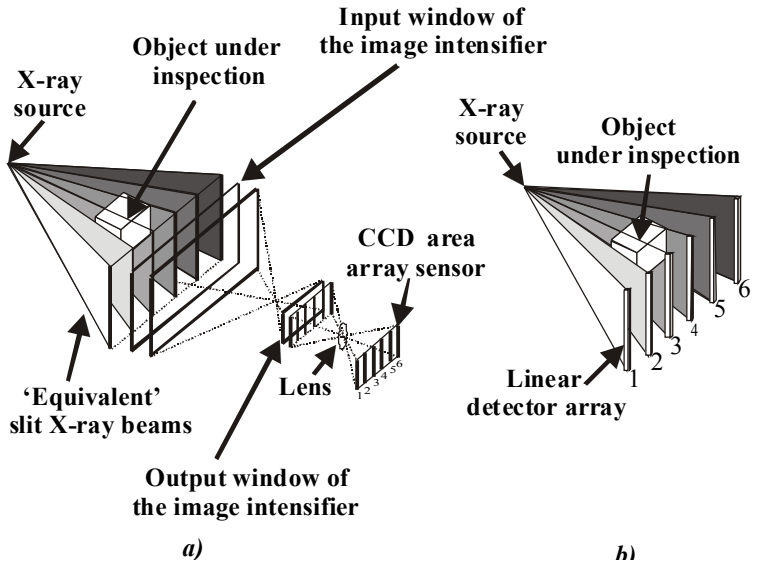


Figure 2.3 Line-scan principle utilising a) image intensifier X-ray system b) linear X-ray detector array system.

2.3.2 Folded array multiple view X-ray scanner

The machine illustrated in Figure 2.3 employs a 140kVP polychromatic X-ray source and a folded array of linear X-ray sensors depicted in Figure 2.4 [40]. It has an inspection tunnel, which is 40cm high by 60cm wide, and is designed to scan full size luggage.

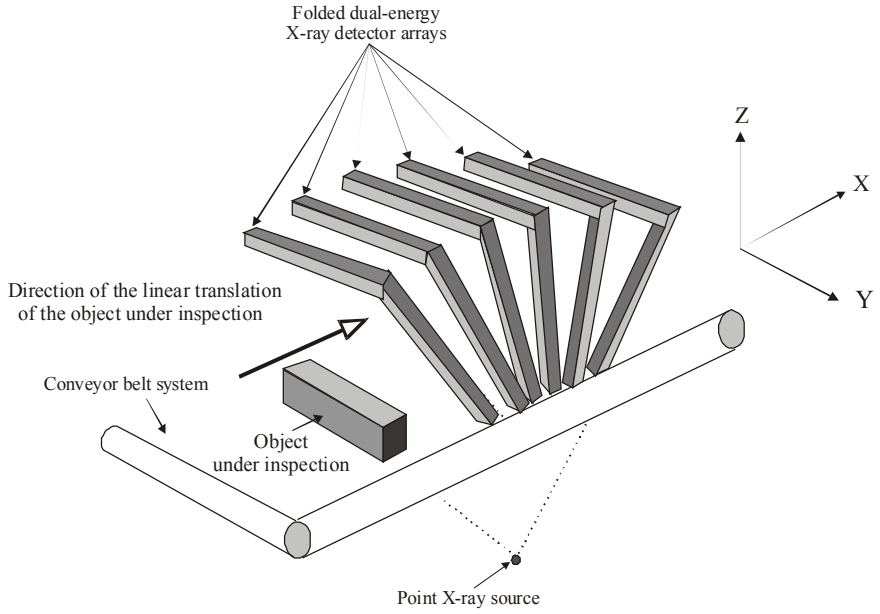


Figure 2.4 Depiction of the experimental system with a single folded array in multiple positions.

The X-ray source/sensor assembly may be rotated about the conveyor belt to enable the collection of the multiple views at different angular position Hon and Evans [41,42, 43,44,45]. In this way the image output of a single pass multiple view scanner can be simulated. Figure 2.5 illustrates the modular configuration of the folded linear X-ray detector

array.

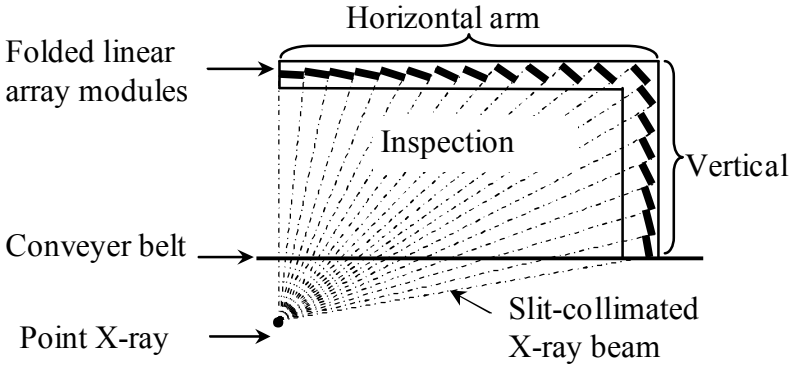


Figure 2.5 Modular configuration of the folded linear (dual-energy) X-ray detector array.

This experimental machine was designed by the university team for the investigation of depth from motion (or kinetic) binocular stereo imaging and constructed at the Ionising Laboratories at Sandridge of Home Office Scientific Development Branch UK, and is housed in the university's Imaging and Display Research Laboratories.

2.4 X-ray laminography

Plantes [46] introduced a slice imaging technique termed laminography [47] that was employed in medical diagnostics until the 1970s. Laminography utilises motion (or different relative positions) of the X-ray focal spot and X-ray detector to produce an image of a single layer or depth plane through the object under inspection. Spatial structures above and below the "in focus" depth plane produce blurring artefacts in the resultant laminogram (i.e. image of the layer of interest in the object under inspection). Laminography techniques may be categorised in accordance with the type of relative motion employed between X-ray source and X-ray detection plane, circular motion and linear motion being the most common (see Figure 2.6).

The Laminography principle is often based on the relative motion of the X-ray source and the detector, about a plane of interest in a stationary object.

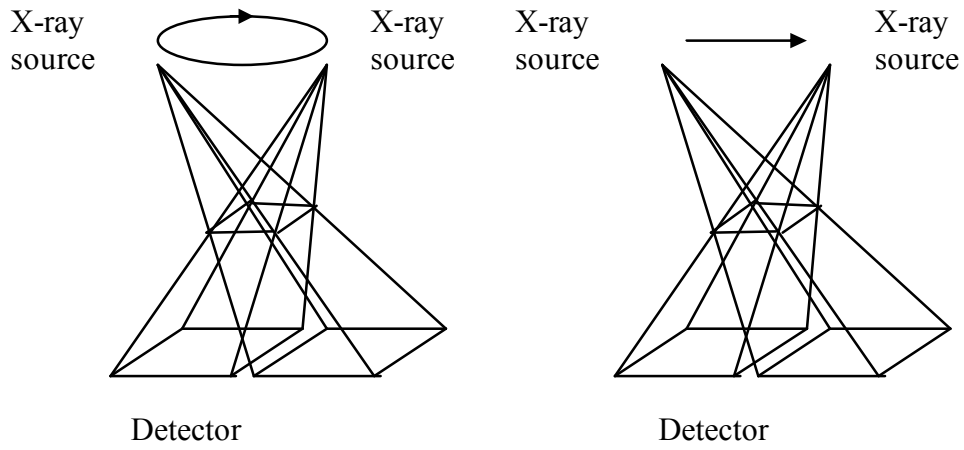


Figure 2.6 "Classical" laminography

The basic principles of X-ray laminography is described in [48]. The basic geometric principle of laminography is to maintain a one-to-one correspondence, throughout the

motion cycle, between the spatial structures in the focal slice and their absolute position on the detector plane. X-ray intensity summation occurs by conserving the ray paths from the X-ray source through specific locations in the focal slice to fixed positions on the detector. Circular motion, in comparison to say linear motion, reduces the likelihood of image artefacts resulting from preferred spatial structure orientation. Early laminography machines utilised radiographic film to produce an image of each focal slice. This cumbersome approach required a new scan together with a new radiographic film for each slice image.

Digital laminography [49] employs electronic X-ray detection elements to sample and store successive images in a digital format. The advantage of storing a sequence of images, acquired at different relative positions during the motion cycle, is that the data for successive slices is obtained in a single scan. Furthermore, the availability of digital imagery enables the reconstruction of the projection data using well-known CT reconstruction algorithms such as the Algebraic Reconstruction Technique (ART) [50]. This approach produces improved contrast resolution and helps to reduce the smearing artefacts prevalent in simple tomosynthesis [51].

Computed laminography (CL) [52] has been investigated and developed by the Fraunhofer Institute of Non-destructive Testing IZEP, to utilise linear translation of the object through a fan beam of X-rays as illustrated in Figure 2.7.

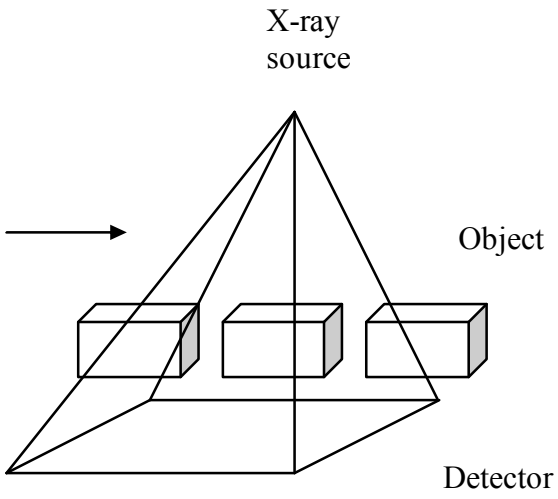


Figure 2.7 Limited angle CL.

The use of linear translation enables the X-ray source and detector to remain stationary. The image collection geometry is similar to the "flatbed" version of the KDEX technique developed by the university team. However, the Fraunhofer Institute's technique and the NTU's technique are based on very different design considerations and are quite different in practice.

2.5 Image synthesis

Image synthesis is the process of creating new images from some form of image description. The approach to synthesising images may be broadly categorized as model based or image based. Model based techniques are designed to compute and manipulate a 3D mathematical representation of the scene. Such techniques computationally reconstruct other viewpoints, as required, and are highly computationally intensive.

Image based techniques employ matching processes to identify correspondences between two or more input images that refer to the same scene point. Given two original images, a pixel in one image is the corresponding pixel in the other image if both pixels are projections along the lines of sight of the same physical scene element. Once these correspondences are known, the world coordinates of each image point can be reconstructed by triangulation.

If the two images are of the same scene at a different time, then computing the correspondence can help determine any motion in the scene. If the two images are taken simultaneously from different stationary points, then computing the correspondence may be used to determine the range or depth in the scene. In this research, new images are synthesized from a sequence of original perspective images obtained from an X-ray scanner.

2.5.1 Stereo matching

Stereo matching is the problem of identifying correspondences between two input images obtained with different angular views. It is a fundamental computer vision problem with a wide range of applications [53,54,55,56,57], and it has been extensively studied in the computer vision field for decades. Scharstein D and Szeliski R [58] present an extensive survey on recent stereo algorithms. Despite the advances in computing and electronics technology, the correspondence problem remains relevant and challenging. There are two basic techniques widely used, correlation-based, and

feature-based methods. Correlation-based methods attempt to establish a correspondence by matching image intensities while feature-based methods attempt to establish correspondence by matching a sparse set of image features. Each approach is discussed in the following text.

2.5.1.1 Correlation-based methods

Correlation-based methods usually rely on statistical correlations between local intensity regions to enable similarity measurements. Typically they implement various types of statistical correlation between colour or intensity patterns in the local support windows. By using local support windows, image ambiguity is reduced efficiently while the discriminative power of the similarity measure is increased. Correlation-based methods have been applied successfully to stereo images with good features and smooth surface variations [59, 60]. The method has the advantage of directly producing dense disparity maps but it tends to fail where there is lack of texture or where depth discontinuities occur. It is assumed that all pixels in a support window are from similar depth in a scene and, therefore, that they have similar disparities.

Correlation-based methods are easier to implement than feature-based methods and provide a dense disparity map, but they tend not to work well when the viewpoints are very different. A key issue in correlation-based methods is the appropriate selection of a window size for calculation. The kernel of the correlation-based methods for stereo matching lies in the underlying similarity criterion that determines optimal statistical correlation between windows around corresponding points. Correlation methods assume that all pixels in a correlation window have the same depth. This assumption is violated at depth discontinuities. When the comparison windows straddle a depth discontinuity, they represent projections of different surface regions. This effect can result in object borders being blurred and small details or objects being removed.

The key problem associated with this window-based approach is that the size of the correlation windows must be carefully chosen. Too small a window may not capture enough image information, and increases the noise problem and can lead to a decrease in correct matches. Too large a window tends to produce matching less susceptible to noise but increases the actual variations in image intensity. A larger window is not a cure-all, since it can result in a greater number of false positives in occlusion zones and increases smoothing of disparity across discontinuities. Adaptive searching windows

have been proposed to solve this problem. Adaptive-window methods [61,62,63, 64,65,66] seek to find an optimal support window for each pixel by adaptively changing the size and shape of a window. Kostkova and Sara [67] introduced a new method to find suitable matching windows based on disparity space found pre-matching. Geiger et al [68] and Fusiello et al [69] adopted a multiple window method where a limited number of distinct windows are implemented for each pixel and disparity, to identify the best correlation. Some multiple-window methods [70] select an optimal support window among the pre-defined windows, which are located at different positions with the same shape. This is also the idea behind spatially shiftable windows [71,72].

Rectangular-shaped and constrained-shaped windows, however, may be inappropriate for pixels near arbitrarily shaped depth discontinuities. To resolve this problem, segmentation-based methods [73] use segmented regions with arbitrary size and shapes as support windows. Methods [74] have been proposed that assign support-weights to the pixels in a support window while fixing its shape and size.

Most of the correlation-based methods use intensity to measure similarity. This approach has limitations when the image suffers from aperture problems, so non-parametric local transforms were developed as the base for correlation. Non-parametric local transforms rely on the relative ordering of local intensity values, and not the intensity values themselves. There are basically two non-parametric local transforms: the rank transform, which measures local intensity, and the census transform, which summarises local image structure. Bhat and Nayar [75,76] worked on ordinal measures, which also belong to this category. Ordinal measures are based on relative ordering of intensity values in windows, and have demonstrable robustness. Non-parametric transforms suffer from the limitation that the amount of information they associate with a pixel is not very large. Non-parametric transform based methods rely only on the rank, during the transform; a lot of information is lost as a result of this approach and the intensity variance is less in the transformed image.

2.5.1.2 Feature-based methods

Feature-based methods produce sparse disparity maps that work by matching characteristic special features of two images, such as corners or edges to produce a sparse disparity map [77]. This method matches more abstract features, rather than matching textured regions in the two images. Feature-based methods provide more

precise positioning for the matching results and are more reliable than correlation-based matching when good image features (e.g. corners, edges) can be extracted from the scene. They are also faster than correlation-based methods, and relative insensitive to illumination changes, but only provide sparse disparity maps. Because of the sparse and irregularly distributed nature of the features, the matching results must be augmented by an interpolation step if a dense map of the scene is desired. If a feature-based method is used, an extra stage is required for feature detection in the two images, which increases the computational cost. As feature-based methods can only produce a sparse disparity map, they are usually implemented together with other techniques to generate a dense disparity map.

Feature-based methods are widely used in wide-base stereo image matching. In methods [78,79,80], local features are extracted independently from the two images, which are then characterised by invariant descriptors and finally matched.

2.5.1.3 Other methods

Other types of stereo matching methods have also been developed such as pixel-based [81] and diffusion-based [82].

Segmentation techniques have been investigated to separate the image into several regions [83], since the regions contain much more information than individual pixels; the possibility of making a wrong decision concerning a particular region is greatly reduced. This approach is termed segment-based. Segmentation information is used in several recent stereo approaches [84,85,86,87] where the assumption that discontinuity only occurs at the boundaries of the segmented regions. These methods use features of the segmented areas in the matching process.

A stereo algorithm is termed a local method if it consists of matching cost computation, aggregation of cost, and disparity computation while a global method [88,89] consists of matching cost computation and disparity optimisation. The global matching is formulated as the minimisation of an energy term that takes into account the matching constraints induced by the local stereo algorithm. Fast, approximate minimisation of this energy is achieved through graph cuts [90,91,92], which is based on the max flow algorithm in graph theory. Dynamic programming [93,94] is another global method.

2.5.2 Multiple view stereo matching

Reconstruction of a 3D scene from sequence of multiple views is a fundamental challenge that has been extensively investigated in computer vision, and numerous applications [95,96,97,98] have been developed.

The multiple view stereo evaluation website [99] provides a significant amount of information concerning this topic. This website provides some standard sequences of images, to the input data for comparative studies. Dyer [100] gives a useful review of multiple stereo algorithms up to 2001. SM. Seitz et al [101] provided a recent and comprehensive review of multiple stereo methods where existing methods are categorized according to six fundamental properties: scene representation, photo consistency measure, visibility model, shape prior, reconstruction algorithm and initialization requirements.

The following section is a review of the various existing multiple view stereo matching methods according to three important aspects related to the algorithm developed in this work.

2.5.2.1 Correlation-based, feature-based and other methods

As previously mentioned, stereo matching methods can be divided into correlation-based and feature-based methods. The majority of the multiple view stereo matching methods share some knowledge with stereo matching and can be incorporated into this category.

Some methods compare two images at a time, and use window matching metrics such as Sum of squared differences or normalized cross correlation, B. Jia et al [102] introduced a method that computes the sum of squared-differences (SSD) values for each stereo pair in orthogonal-directions, then it simply adds the individual SSD to produce the sum of SSDs. The resulting function is called the OSSSD by the author.

When dealing with discontinuity and wide-base images, feature-based method is known to be better. Ferrari et al. [103] proposed an algorithm for identifying multi-view feature correspondences across an unordered set of widely separated views.

As in stereo matching, there are other ways to categorise matching methods, global methods such as graph cuts are also used frequently in multiple stereo matching [104,105].

Several methods can also be combined into a single application; Watanabe, M. Ohta, Y [106] developed a matching method that utilises correlation, feature-based matching using edges and the segment-based matching using connected edges. The method executes each matching process in parallel, with making use of information from other processes whenever necessary. Then, depending upon different image features, the best-suited algorithm is selected and applied automatically to those features. Each module evaluates the extent of confidence of its own processing, and the reliability score is attached when the result of processing is exchanged.

2.5.2.2 Scene representation

There are numerous ways to represent the geometry an object. Most multiple view algorithms use voxels, level-sets, polygon meshes, or depth maps. Some algorithms adopt a single representation; others employ different representations for various stages. Methods using multiple depth maps [107,108] are better suited to limited scene datasets, as the computation of a depth map has a minimum requirement of two images.

This research programme aims to produce a set of depth maps, and then generate virtual intermediary images. The angular coverage of the experimental X-ray scanner is not large enough for a reliable construction of 3D model of the luggage under inspection.

Some methods [108,109,110] compute a set of depth maps and then merge them together to obtain a final 3D representation of objects, As a set of depth maps are calculated, to ensure a single consistent 3D scene interpretation, these methods enforce constraints between depth maps [111] or merge the set of depth maps into a 3D scene as a post process. P. Merrell et al [110] advocates a two-stage process in which the first stage generates potentially noisy, overlapping depth maps from a set of calibrated images and the second stage fuses these depth maps to obtain an integrated surface with higher accuracy, suppressed noise, and reduced redundancy. Most algorithms, except [112,113], decide the potential depth of candidate points using two images, then refer to

other depth maps using geometrical consistency and other constraints and to obtain an optimum depth for the candidate points.

Correspondence matching for stereo images is inherently noisy, ambiguous and prone to multiple matches, hence identifying the correct disparity information from multiple disparity maps requires sophisticated optimisation methods. T. Kanade et al proposed an algorithm [113] simply accumulates the measures of the matching criterion SSD from all the stereo pairs into a single value, and then computes the corresponding point from these measures. The method accumulates matching measure values from all images for a final decision, rather than filtering and optimisation the set of depth maps.

M. Goesele et al [107] presents another way of using these depth maps. Each depth map is used to reconstruct a part of the structure that has high confidence correspondence points. Each map has different low confidence correspondence which relates to occlusions, low textured regions, discontinuity regions, e.g, most of these effects occur in different image regions in different views, so each map can fill in a part of the final structure, and improve the accuracy in regions that are reconstructed multiple times.

2.5.2.3 Image selection

As multiple images are available, when determining the disparity for one candidate point or one candidate feature, some algorithms try to only use those images that contain useful information during the matching process, for example, only those images in which the candidate point or the candidate feature is visible. Most of the multiple view stereo methods simply choose the 'nearest' images for each reference view, for example, C. Hernandez and F. Schmitt [114] proposed an algorithm that limits the matching process to clusters of nearby cameras.

M. Goesele et al [115] tried to intelligently match images on a pixel level and illustrates that such adaptive view selection enables robust performance even with dramatic appearance variability. The underlying rationale is that, given of a set of sequence images, there should be some subsets of images, which are captured under compatible lighting, weather, and exposure conditions, as well as sufficiently similar resolutions, so that the features in those images should have similar appearance. By automatically identifying such subsets, the problem can be simplified.

A specific point or feature may only be visible in some subsets of images. A visibility model can be used to specify which images to consider during the matching process. Because scene visibility can change dramatically with viewpoint, almost all modern multi-view stereo algorithms account for occlusions in some way or another.

G. Vogiatzis et al [116] use the current estimate of the geometry to predict visibility for every point on a surface. It aims to decide which scene structures are visible in which images. Others use the current estimate of the shape to compute the exact visibility of all points [117].

2.6 Common matching constraints

Constraints that are commonly used by stereo matching methods are listed as follows:

1 Epipolar constraint:

The corresponding point can only occur along the epipolar line in an image taken from a different station point of the same scene. This constraint reduces the potential 2D search space into 1D search. The epipolar constraint can be reliably applied only after the geometry of the system is known and a series of corresponding epipolar lines in both stereo images is estimated. Calibration is used to make two images satisfy this criterion. The epipolar constraint is one of the most fundamentally useful pieces of information, which can be exploited

2 Uniqueness constraint:

The uniqueness constraint [118] enforces a one-to-one mapping between pixels in two images. This states that in most cases, a pixel from the first image can correspond to only one pixel in the second image. An exception arises when two or more points lie on one ray originating from the first camera and can be seen as separate points from the second, which is called self-occlusion. This constraint holds for opaque surfaces, but fails if partially transparent surfaces are present in the scene.

3 Photometric compatibility constraint:

The compatibility constraint states that intensities of a point in the first and second images only change a little. The intensities may not be exactly the same due to many different effects such as the light source and surface normal, but the difference is expected to be relatively small. In correlation-based stereo matching algorithms, it implies that corresponding image widows have high cross-correlation intensity.

4 Geometric similarity constraint:

Geometric similarity assumes that the shape characteristic of the features found in the first and second images do not change too much.

5 Disparity smoothness constraint:

Smoothness constraint is developed from the assumption that a visible surface, and therefore the disparity of corresponding points, varies smoothly almost everywhere over the scene. In the presence of multiple visible surfaces with discontinuities, this constraint is invalid.

6 Ordering constraint:

This constraint says that for a surface of similar depth, corresponding feature points typically lie in the same order along the epipolar line.

2.7 Theoretical considerations

2.7.1 Transparency

The transmission of X-rays through an object to produce a shadowgraph which results in the transparency property commonly attributed to such imagery. As a result, an overlapping structure in an X-ray image may appear as an integral part of two or more spatially separate objects. This property makes X-ray images fundamentally different from visible light images and presents additional considerations and complexities for the development of image synthesis techniques.

2.7.1.1 Multiple correspondences

The intensity of each pixel in the resultant images is an aggregated intensity of all points along the ray path under consideration. Each pixel may have multiple correspondences

associated with a number of different overlapping structures encountered along the ray path. Figure 2.8 presents the scenario where objects under inspection are imaged by X-rays which are incident normal to the detector plane and alternatively at an angle to the detector plane. Two different interrogating rays image the section of the rectangular object in the diagram. The resultant pixels (PixAng1) and (PixVert1), produced by the inclined and the normal rays respectively, are a corresponding pair, although their intensities are different. However, PixVert1 is also the corresponding point for PixAng2. As a result, PixVert1 has two potential correspondences arising from the transparency property in the transmission image. In contrast, pixels in visible light images are not, in general, subject to such uncertainty for potential matches. Any error produced by the multiple correspondences has the potential to create "voids" in the resultant synthetic images, which requires further computational measures to be undertaken to produce acceptable synthetic imagery.

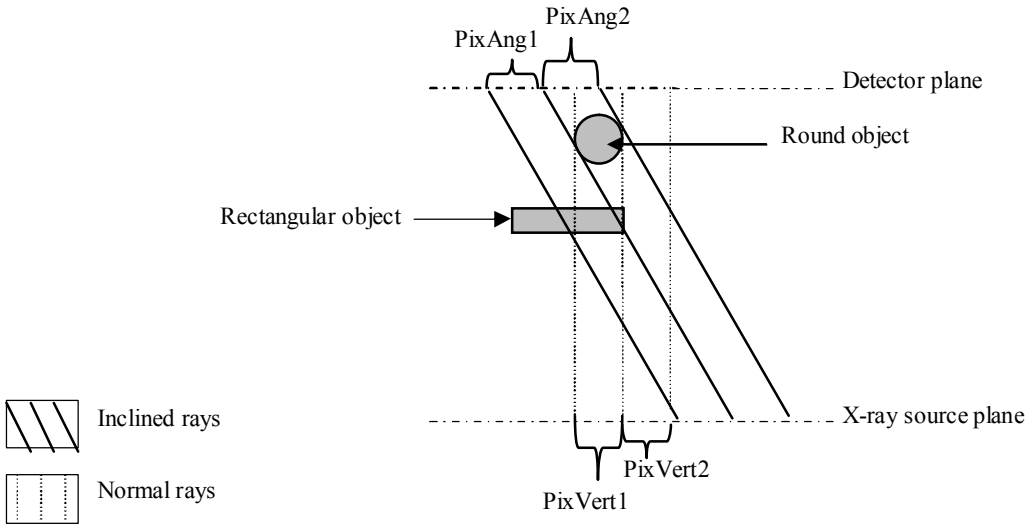


Figure 2.8 Illustration of multiple correspondences: PixAngl and PixAng2 record the X-ray attenuation from the inclined ray, while PixVertl and PixVert2 record the X-ray attenuation from the normal rays.

The three factors affecting the nature of multiple correspondences in X-ray images are described as follows.

a) Effective thickness

The "effective thickness" of an object is defined, in the context of this research, as the length of the ray path subtended by the object(s) under inspection. Thus the effective

thickness can change from view-to-view and also within a single view, the latter being known as structural unsharpness. Each effect is a fundamental property of transmission X-ray imaging and is discussed in the following text.

Variation in "effective thickness" exhibited by multiple views

The intensity of the pixels, which comprise the resultant image, is proportional to the amount of attenuation experienced by the X-ray beam. The distance travelled through a given material or object largely determines the amount of attenuation. When an object is acquired at different perspective angles, the commensurate change in the ray path through the object (or effective thickness) produces a relative change in the image intensity. In practice, the thicker the object, the greater the intensity fluctuation observed for different perspective views. Thus even a relatively straightforward correspondence-matching task can be problematic. On a practical note, very thin objects often produce negligible variations in intensity if they nominally occupy a plane parallel with respect to the detector plane. The complex overlapping patterns of objects routinely encountered in images of luggage further exacerbate the problems associated with integrated thicknesses.

Variation in the "effective thickness" exhibited within a single image (or structural unsharpness)

Soft or blurred edges are a fundamental property of transmission images. This effect is a natural consequence of transmission imaging, termed structural unsharpness, and is the result of the variation in ray paths through an imaged object. Typically, structural unsharpness may be observed near the boundaries of imaged objects. The effect tends to limit the performance of the view synthesis algorithm. A practical example of structural unsharpness is illustrated in Figure 2.9 (see included line profile). It is important to note that this effect should not be confused with other types of unsharpness arising from practical fluctuations or the limited resolution of the imaging system.

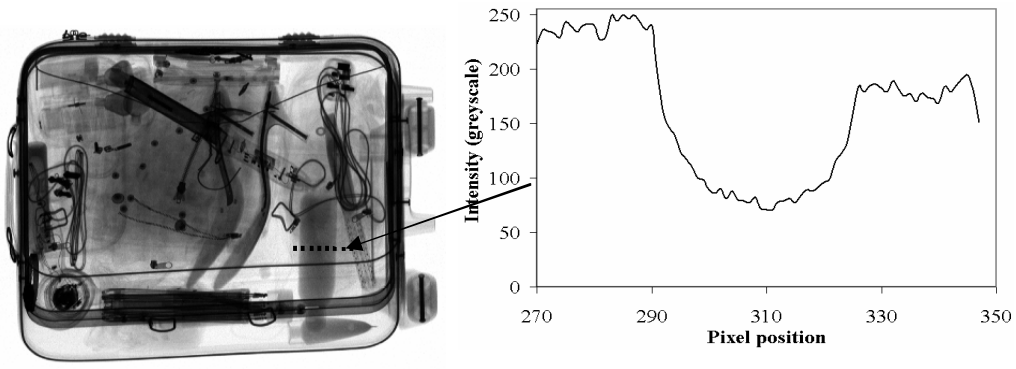


Figure 2.9 Illustration of structural unsharpness.

The shape of an object may dramatically change as a function of the viewing angle. This orientation factor (or shape transformation) is a known challenge in synthesising visible light images. Due to the transparency inherent in X-ray images, many overlapping structures may transform their shape simultaneously presenting highly complex and unpredictable scenarios.

b) Material composition

The chemical composition of an object plays an important role in its X-ray attenuation properties and the resultant image intensity. Assuming that the rectangular object depicted in Figure 3.1 is made of a highly attenuating dense material (e.g, metal) and the round object is made of a less dense material (e.g. acrylic), the difference in intensity between PixAng1 and PixVert1 could be marginal, leading to a potentially robust match. On the other hand, a switch of the material characteristics of the two objects would confound a matching process due to occlusion.

c) Angular separation of the perspective images

The greater the angular separation employed in the imaging geometry used to collect a sequence of perspective images, the greater the magnitude of the sequential parallax evident in the resultant imagery. As the angular separation is increased, the individual images tend to become increasingly dissimilar in terms of shape, overlapping features and intensity due to the factors discussed in a) and b) in the preceding text. In simple, uncluttered scenes it may be advantageous to acquire highly disparate imagery in which objects, or their salient features, have been conveniently separated, in the x-axis, by virtue of their relative location in range (or z-axis). It is equally relevant to consider

highly cluttered scenes whereby increased parallax results in highly dissimilar imagery in which it is difficult or impossible to automatically identify corresponding points or features.

Figure 2.10 and Figure 2.11 are organised to illustrate the effect of perspective angle on view synthesis. Figure 2.10 is a modification of Figure 2.8 by mainly adding a third set of lines of ray (named 2nd angular lines of ray) that has smaller angle separation from the vertical lines of ray, in relative to 1st angular lines of ray. In comparison, the intensity difference between PixVert1 and PixNew1 is smaller than the intensity difference between PixVert1 and PixAng1. The reduction in intensity difference is very favourable to matching correspondences and influencing the intensity of the synthetic pixels.

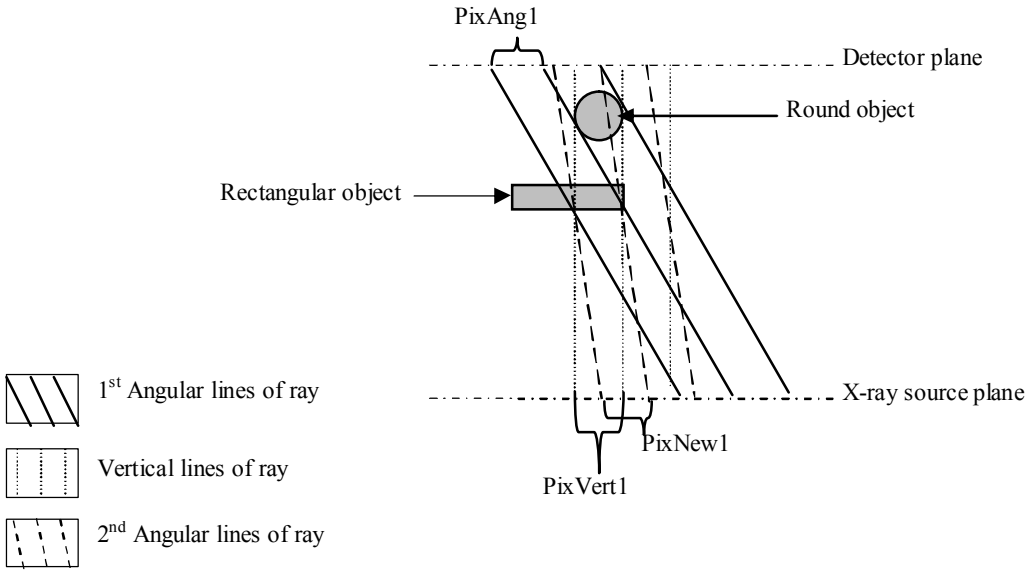


Figure 2.10 Relationship between perspective angle and intensity: Smaller Separation Smaller Intensity Difference. PixAngl, PixNew1 and PixVert1 are the first pixels that records the X-ray attenuation from the first angular lines of ray, the second angular lines of ray and the vertical lines of ray respectively.

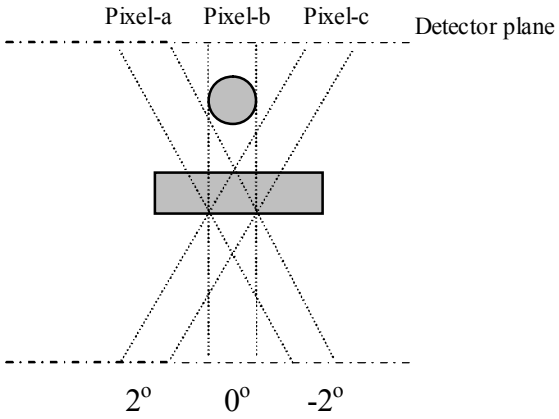


Figure 2.11 Relationship between perspective angle and intensity: Bigger Separation Smaller Intensity Difference.

However, a relatively increased angular separation between views may also contribute to positive outcome. Figure 2.11 depicts two objects are scanned by three perspective rays: 2°, 0° and -2°. It can be appreciated from Figure 2.11 that Pixel-a and Pixel-c may have very similar intensities although the angular separation is big.

The observation revealed that acquiring images of an object from multiple perspectives could offer valuable information to improve the performance of the search algorithm. The two scenarios illustrated by Figure 2.10 and Figure 2.11 might sound contradictory in terms of the superiority of employing different angular separations between adjacent views. The conflicting requirement is a central issue for stereo matching because there are only two images involved. In this research programme the potential benefits of a range of angular separations is explored using multiple perspective images.

2.7.1.2 Limitations of common matching constraints

Figure 2.12 illustrates a hypothetical example of an X-ray image and a visible light image consisting of a simple arrangement of three geometric objects. The search for corresponding points is limited within a maximum allowable disparity window, which in turn is determined by the design parameters of the experimental system. As can be deduced from Figure 2.12, the search for the point-p in the X-ray image, along its search direction, will encounter depth discontinuity, inconsistency and, variation in pixel values. Thus point-p may be associated with several correspondences thereby increasing the matching ambiguity. In contrast, the search in the visible light image does not encounter this effect (usually) as the square object occludes the other two objects. This situation can be further exacerbated in the X-ray case when the material composition of the different objects is taken into account. Common matching constraints have good

performance in visible light image matching. However, due to the relatively increased frequency of multiple correspondences in X-ray images and more discontinuity regions that arising from transparency, common matching constraints which are commonly used for visible light images such as uniqueness, smoothness, ordering constraints become at worst inappropriate and often produce unstable results when applied to X-ray images.

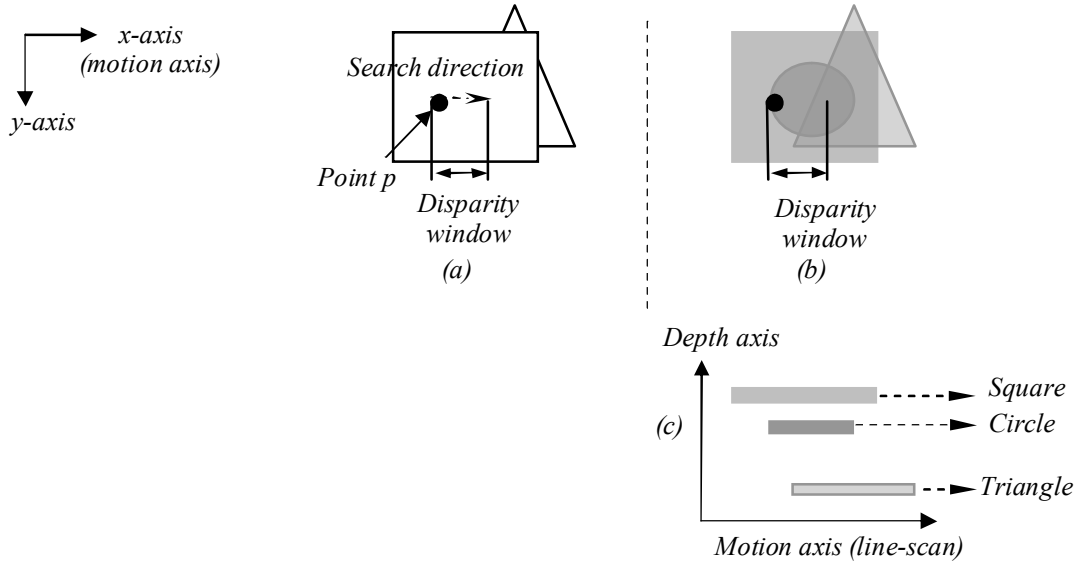


Figure 2.12 A hypothetical example of a square object, a circular object and a triangular object, where (a) represents the visible light image and (b) the X-ray image while (c) illustrates the object separation in the depth or range axis.

2.7.2 Multiple images

2.7.2.1 Image features under consideration

It is implied that using multiple images might help to solve some problems associated with stereo matching. However, more information may also carry the risk of increased ambiguities. The problems that are commonly suffered by stereo matching algorithms that apply correlation-based method for visible light images include repeating features, overlapping, discontinuity and homogenous areas. X-ray images exhibit modified versions of these problems. Amongst these problems, homogeneous areas were considered to be similar as repeating features. The following sub-sections are organised to illustrate the advantages and disadvantages of employing multiple images for addressing the remaining three common problems.

a) Repeating features

Figure 2.13 consists of a selection of X-ray images that were imaged at the angles of 6°, 2° and -2°. Object 1 and Object 2 are repeating features. By applying stereo matching criterion on any two of the three images, Object 1 would have an equal opportunity to match with either Object 1 or Object 2, and thus could produce an erroneous result. However, when all three perspective images are utilised in the matching process, it may decrease the likelihood of Object 1 matching with Object 2. For example, if Object 1 in the 2° image matches with Object 2 in the -2° image, then the corresponding point of Object 1 in the 6° image should appear at Position A, which is evidently a wrong coordinate. The example signifies that an appropriate use of the information extracted from multiple images may offer a significant advantage. On the other hand, as illustrated in Figure 2.14, adding more images to the matching process does not always convey useful data. Object 1 is overlapped by a dense object (black blob), which alters the intensity significantly. Under such circumstances, it is difficult to decide the coordinate of Object 1 in the 6° image.

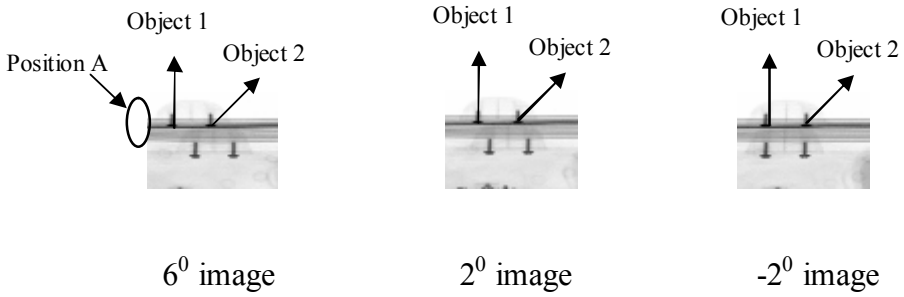


Figure 2.13 Effect of using multiple images for repeating features. Position A illustrates the possible location of Object 1 in 6° image as a result of stereo mismatching when only the 2° and -2° images are deployed.

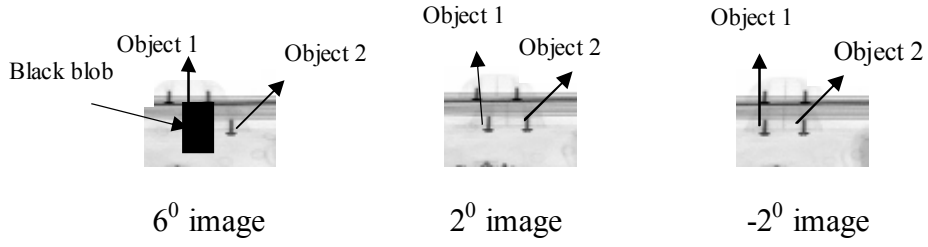


Figure 2.14 Effect of using multiple images on repeating features. A dense object tends to cause ambiguity in the search of correspondence.

b) Overlapping Structure

Over the past decades, extensive research has been devoted to solving the problems produced by overlapping structures. Occlusion destroys the parallax information

associated with overlapping foreground or background objects. The nature of occlusion in X-ray images is different from visible light images in that multiple object features may cause localised occlusions due to the X-ray beam not fully penetrating the overlapping objects. Acquiring images from multiple perspectives can improve the probability of obtaining image information produced from unoccluded ray paths. For example, as illustrated by Figure 2.15, if stereo matching is applied to the -2° and 2° images to synthesise a target intermediary image at 0°, the matching of the knife would be problematic because it is camouflaged by the umbrella in the 2° image. Alternatively, if the stereo pair -6° and -2°is employed then the matching of the knife would be successful. This observation indicates the utility afforded to the matching process in comparison to using only two views.

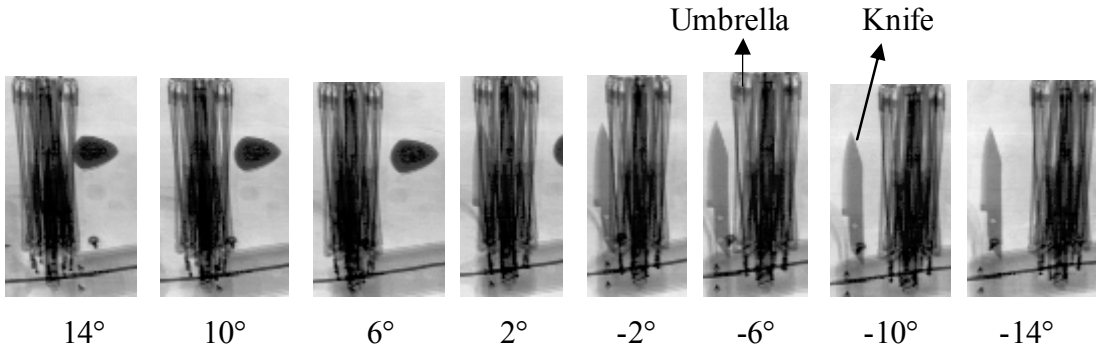


Figure 2.15 Effect of using multiple images on overlapping structures.

c) Discontinuity

Figure 2.16 exemplifies a discontinuity condition where the task is to create an intermediary image at the angle of 12°. If the 13° image and the 11° image are used for the correspondence search, the leftmost edge of Object 1 may suffer from unwanted distortion because it is too close to the umbrella's edge. Employing the 11° image and the 9° image will produce a more robust result. It is advantageous to identify the appropriate local conditions to facilitate a high integrity matching process.

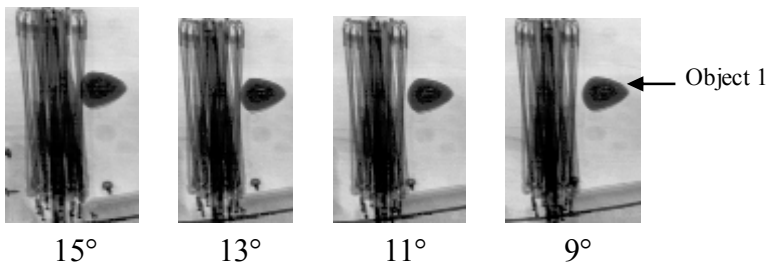


Figure 2.16 Effect of using multiple images on discontinuity.

2.7.2.2 Laminogram

Knowledge from Section 2.4 suggests that, with multiple perspective X-ray images available, a set of laminograms can be generated. The theory of laminography shows that objects located in different depth planes will focus in different laminograms. If a method can be designed to find out in which laminogram those objects are focused, then the information associated with the focused image features could offer a potential solution to the difficult correspondence problem. Inspired by this hypothesis, a laminogram intensity based matching criterion is introduced in this work.

Chapter Three Image synthesis algorithm

3.1 Introduction

The image synthesis algorithm developed in this research programme is presented fully in this chapter. The algorithm is termed the, "voting based dual criteria multiple X-ray images synthesis (V-DMX)". The algorithm is designed to accept X-ray detector images produced by the kinetic depth X-ray (KDEX) imaging techniques previously developed by the university team. The synthetic images produced by the algorithm may be used to replace X-ray detector images and therefore reduce the total number of detectors required to realise a KDEX sequence.

The algorithm employs three detector images to produce disparity information for generating the synthetic images. A novel approach has been proposed to first compute the matching cost [58] using data extracted from the input images and followed by the application of a rank based voting process to decide the "best" disparity for generating the synthetic pixels.

The algorithm employs two matching criteria namely SSD and laminogram intensity. SSD is a widely used matching criterion but can perform poorly for regions exhibiting overlapping structure and discontinuity each of which is particularly prevalent in X-ray images. The laminogram intensity criterion is based upon a novel interpretation of X-ray laminographic data developed in this work. Both criteria are described in Section 3.7.

Each criterion is applied in two separate, but potentially parallel, processes to produce two independent disparity information tables for a selected image. Both disparity information tables are used to produce the resultant synthetic image.

3.2 The algorithm

3.2.1 Algorithm overview

Among the existing algorithms that use multiple images (i.e. more than two images) to produce a disparity information table, there is no clear indication about the optimal number of images needed. The number of images that require processing is a critical practical consideration for the implementation of the KDEX technology, as it will

strongly influence the hardware complexity, system cost and operational speed. This research programme is designed to carefully evaluate the algorithm performance as a function of the number of images required and the angular separation between these images. These aspects are considered further in terms of the total angular coverage achievable as an indicator of the potential cost effectiveness of KDEX technology.

The experimental X-ray scanner produces a line-scan format sequence of images, which form the input data for the image synthesis algorithm. Ultimately, to evaluate the performance of the algorithm, the resultant synthetic images are compared with ground truth images produced at the appropriate X-ray beam angle by the scanner as depicted in the diagram below.

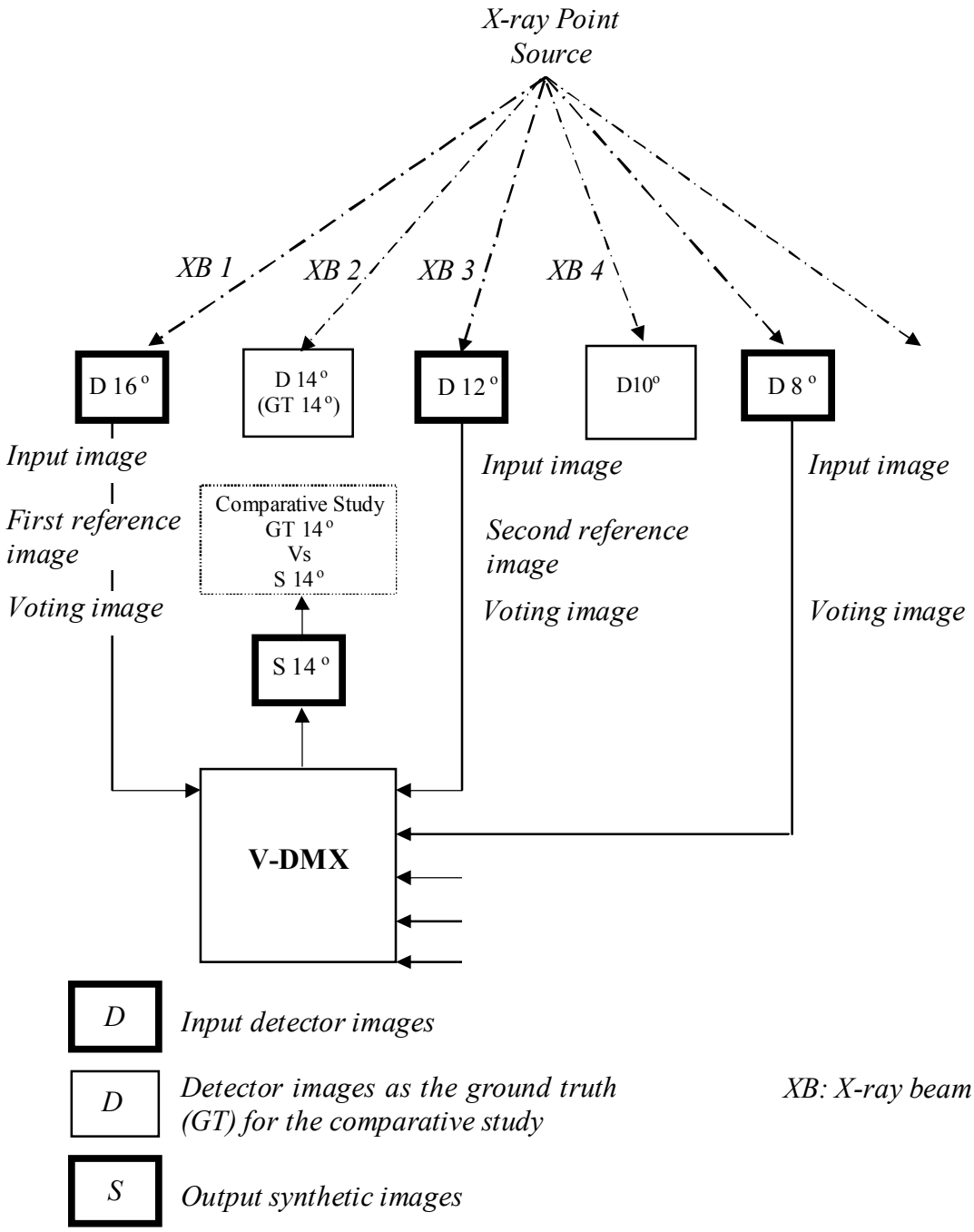


Figure 3.1 Conceptual overview of the image synthesis algorithm in relation to the input detector images and the output synthetic image(s).

Figure 3.1 illustrates the conceptual overview of the image synthesis algorithm in the context of the basic X-ray beam geometry that produces the synthetic image S14°. The algorithm is comprised of a number of different stages and processes. To comprehensively describe the approach developed requires that the images selected for, and processed by, the various algorithm stages be identified by appropriate names. This point is particularly important, as an individual image may form a part of several different sets of images, which are input to a number of different stages within the

algorithm. This section describes the naming convention adopted and presents it within the context of the overall structure of the algorithm.

a) Detector Images

The practical X-ray images that are produced by the linear X-ray detector arrays; nDetect is the acronym adopted to indicate the number of detector images.

b) Target Image

The image which is in the process of being synthesised, is termed the target image. Initially the target image contains no data. Its relative position in the sequence of detector X-ray images is commensurate with the position of the resultant synthetic image, which in turn is determined by the angular distribution of the detector images under consideration.

c) Reference Images

The images employed to produce a target image by virtue of its disparity information. For each target image, the two adjacent perspective images i.e. the images that are closest in angular separation with respect to the target image, in the image sequence, are selected as the reference images. The image on the left side (negative direction) of the target image will be selected as the first reference image, and the other is selected as the second reference image. The left image is an arbitrary choice.

d) Input Images

The detector images that have been predetermined to provide data for the computation of the matching cost; nInput is the acronym adopted to indicate the number of input images. nInput required to generate a single target image is three.

To generate different target images, three different input images are selected from the detector images. The criteria used to select the three images from detector images are:

- In the first case the images closest to the target image.

- In choosing the third input image i.e. there is more than one image that has equal angular separation to the target image, then the image nearest to the first reference image is chosen as the third input image.

e) Voting Images

Voting images are selected from the input images. Each voting image is utilised in the production of a single disparity information table. Disparity information is employed to determine the most appropriate or "best" disparity value obtainable from the reference images; nVote is the acronym adopted for indicating the number of voting images. nVote for the SSD criterion is two, and nVote for LamI criterion is three. The criterion used to select the voting images from the input images are the same as that used to select the input images from the detector images.

3.2.2 Algorithm diagrams

The image synthesis technique developed in this research programme is broadly divided into two distinct processes, as illustrated by the flowchart in Figure 3.2. The first process is to determine the correspondences for each pixel; the second process generates the synthetic image by interpolating the intermediary coordinate position and assigning the intensity to each synthetic pixel.

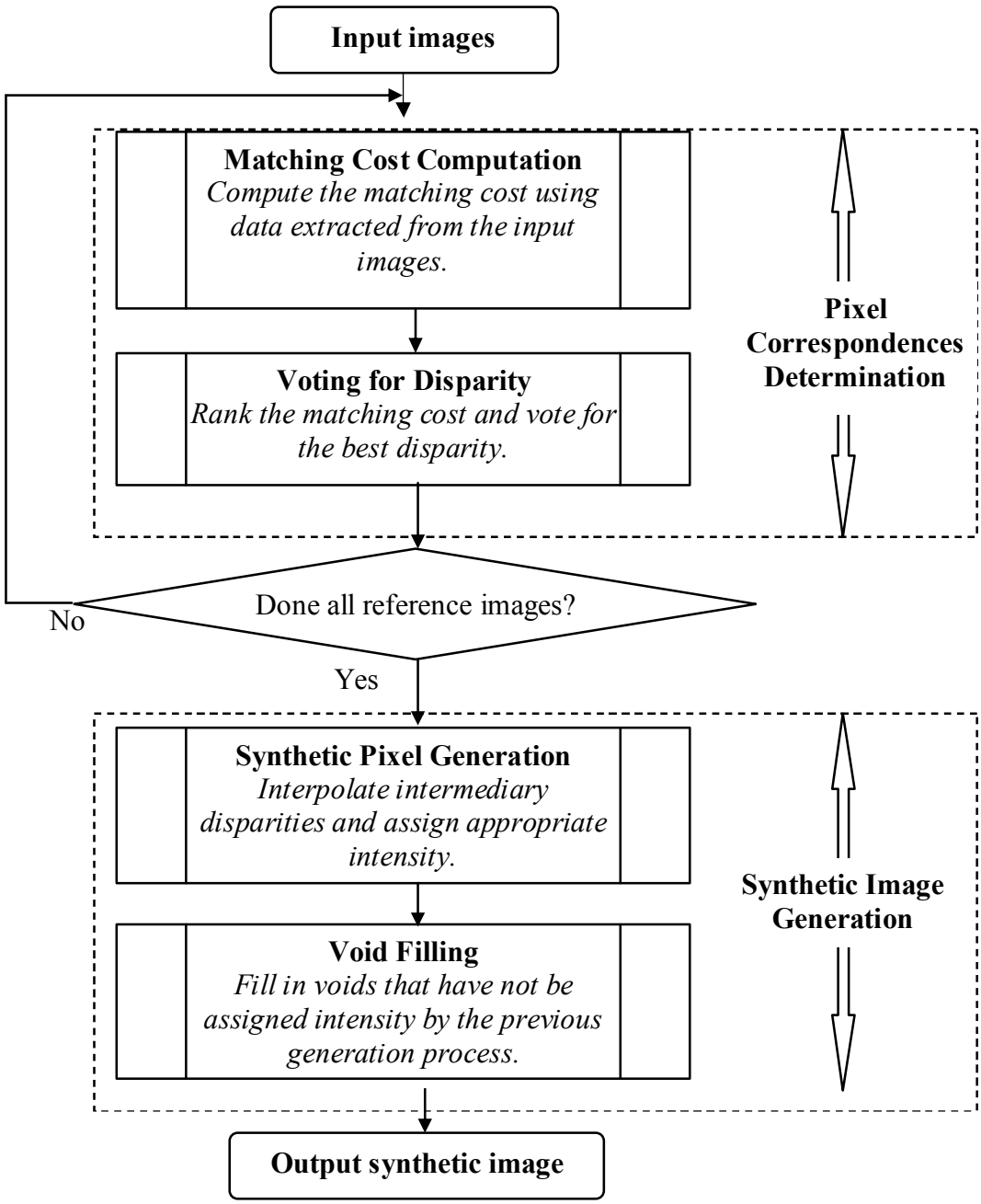


Figure 3.2 Flowchart of the image synthesis algorithm.

Due to differences in the operational nature of the two criteria employed, the prerequisite of the algorithm is a comparative study of the selection procedure that determines the optimum number of input images and voting images required by each criterion for image synthesis. The results of the empirical study are reported in Chapters 5, 6 & 7. To provide the reader a general idea of the operation of the developed V-DMX algorithm, Figure 3.3 is presented to illustrate the synthesis of a target image at 14^{o} (i.e, $S14^{o}$).

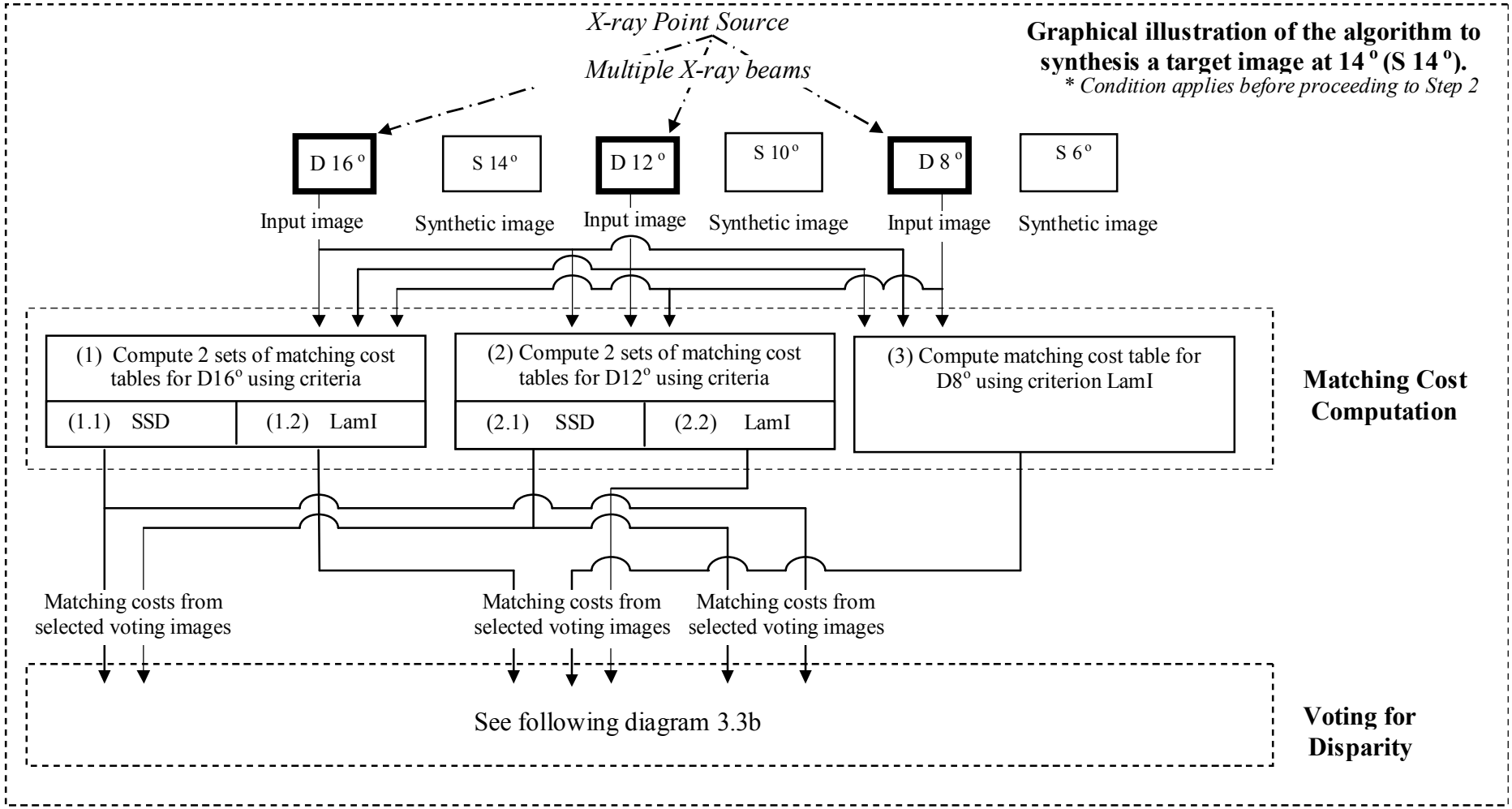


Figure 3.3a Graphical illustration of the algorithm to synthesis a target image at 14 o.

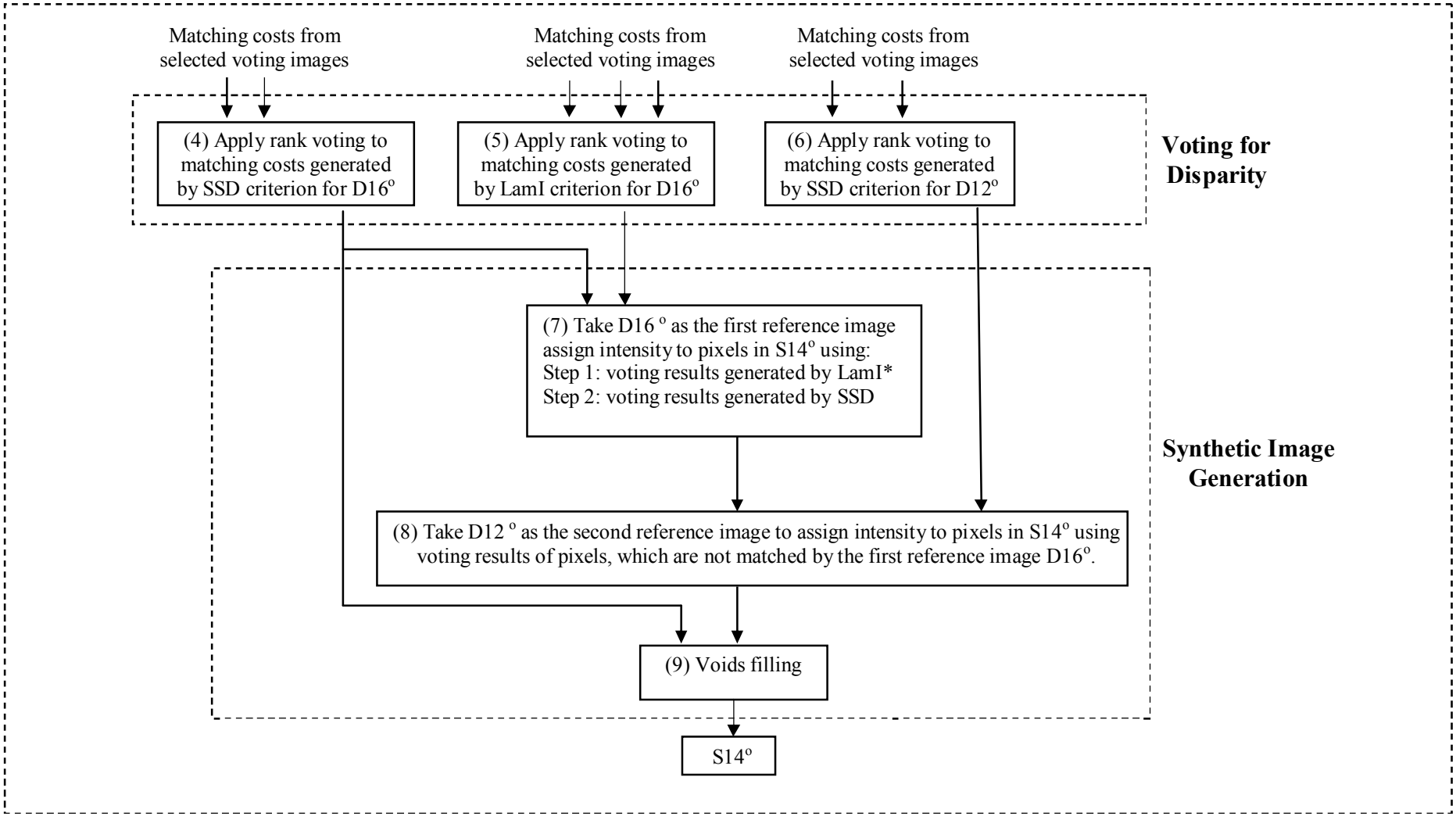


Figure 3.3b Graphical illustration of the algorithm to synthesis a target image at 14°.

In this chapter, each of the algorithmic processes is individually presented in the following sections to provide the reader with a coherent, localised and detailed description of the algorithm design.

3.3 Matching cost computation

Multiple views of the same scene can significantly enhance the identification of corresponding points in comparison to two views. Initially, a process that can track the potential correspondences throughout the image sequence, on a pixel-by-pixel basis, to improve the probability of identifying satisfactory matches was investigated.

The process requires satisfying two geometric constraints.

(a) Epipolar line constraint

It is assumed that any y-disparity (vertical in the display axis) in the KDEX images is negligible, thus the epipolar line is constrained to the x-axis (horizontal in the display axis) [5].

(b) Constant disparity constraint (i.e. from one pair of adjacent images to the next pair of adjacent images)

The rate of change of disparity, with respect to the z-axis (normal to the plane of linear translation) in object space, is constant from one pair of adjacent images to the next pair of adjacent images. In other words the number of depth or disparity planes and the separation between adjacent planes (i.e. depth resolution) in object space is nominally constant over the X-ray beam angles of interest in this research programme [5].

The first step in the algorithm is to compute several matching costs and aggregate those matching costs together. For example, in stereo matching, for one pixel, one disparity is associated with a single matching cost. When nInput images are involved, for one pixel, one disparity is associated with a nInput-1 individual matching cost. If all these matching costs are considered together then the matching process is akin to common stereo matching. The nInput images are defined by their selection for the matching cost aggregation process. The computation of the matching cost requires choosing initially a matching reference image. For example, in stereo matching, a stereo pair can be chosen to compute the matching cost for producing a disparity information table. In effect, two

disparity information tables can be generated for a single stereo pair. The subsequent process is to build a single or "best" disparity information table from the two individual disparity information tables. In this example, both perspective images, which comprise the stereo image, are used as voting images and implemented in a voting process to build the final disparity information tables. The following flowchart in Figure 3.4 illustrates the computation of the matching cost.

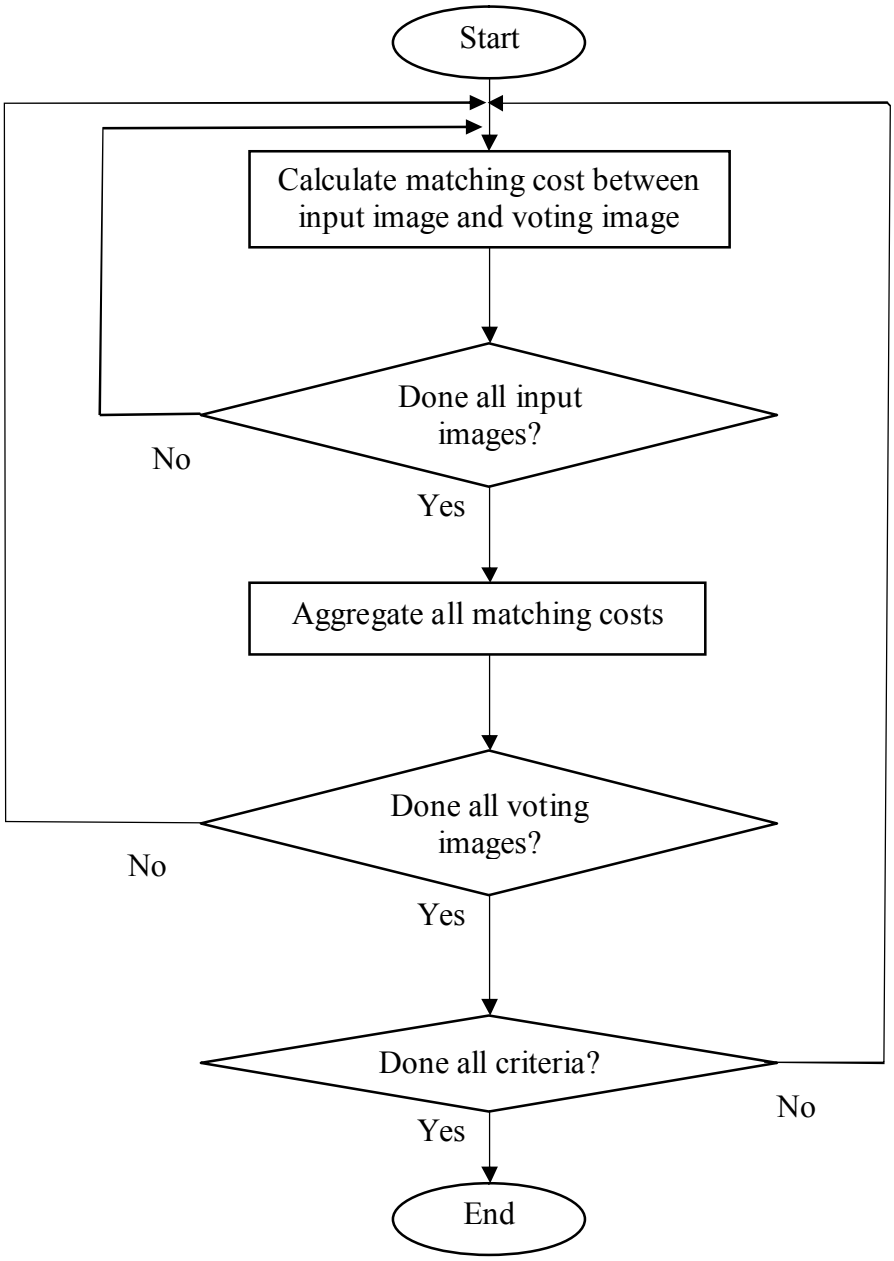


Figure 3.4 Flowchart of the matching cost computation.

The following paragraphs describe the two proposed computation methods, which are termed a) laminography and b) sum of reciprocals. Each complementary method plays an important role in the overall process. The LamI criterion forms the basis for a

laminography matching cost calculation method; the SSD criterion forms the basis for a sum of reciprocals matching cost calculation method.

3.3.1 Laminography

The laminography matching cost calculation method requires the production of a set of laminograms, and then matching cost is calculated between voting images and laminograms. Figure 3.5 illustrates the procedures of using laminography matching cost calculation method to calculate matching cost table for $D16^{o}$, during the process of generating target image $S14^{o}$ using the designed V-DMX algorithm. Matching cost tables for $D12^{o}$ and $D8^{o}$ are also calculated using laminography matching cost calculation method, the procedures are quite similar as in Figure 3.5.

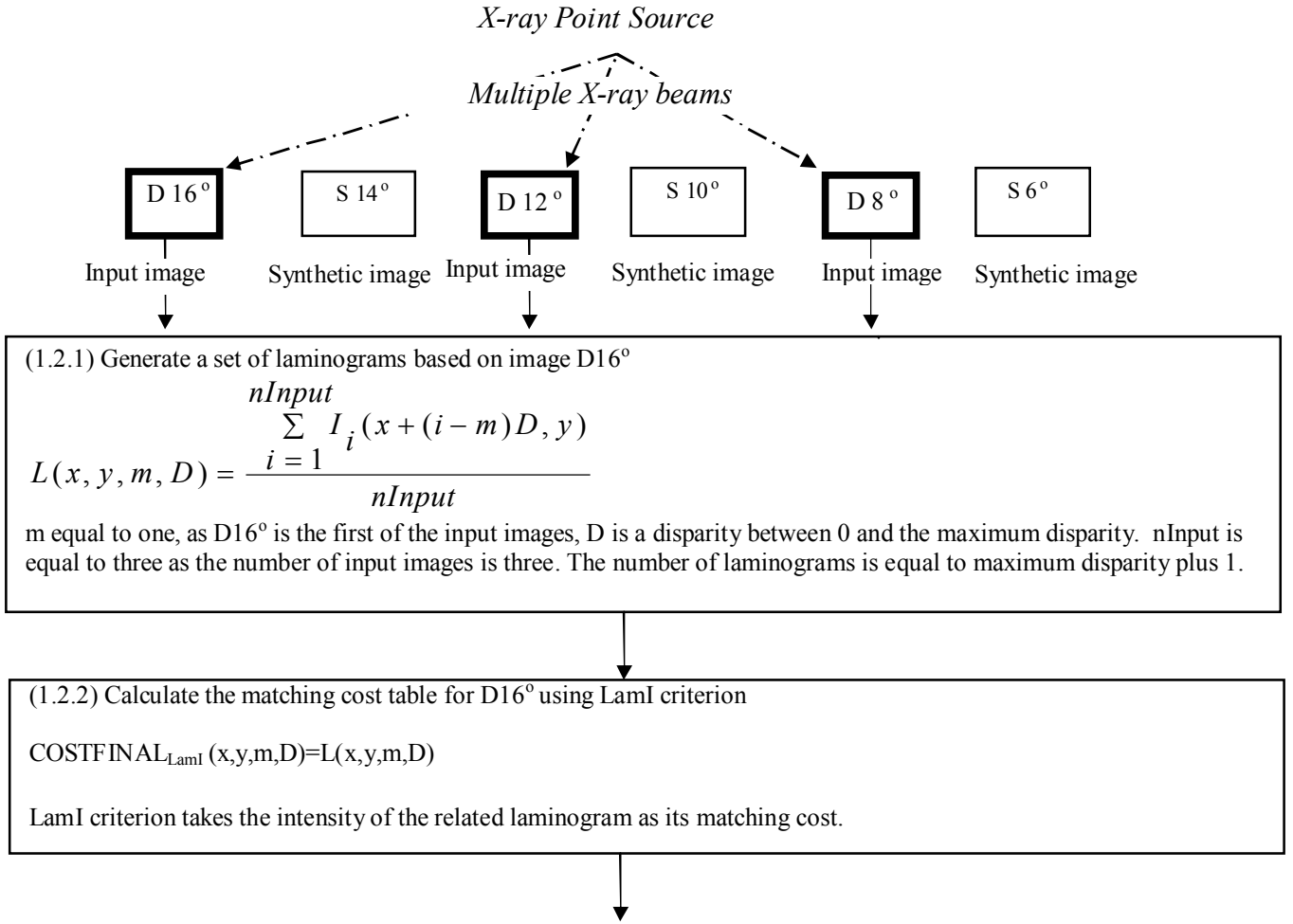


Figure 3.5 Graphical illustration of using laminography matching cost calculation method to calculate matching cost table for D16°, during the process of synthesising a target image at 14°.

By appropriate processing of the input images, a set of laminograms can be produced. The maximum number of laminograms that can be generated is determined by the disparity window size (D_{max}) of the reference image. If D_{max} is 10 pixels in size, then 11 laminograms can be generated. In this research programme, an averaging method is applied to generate a $D_{max}+1$ number of laminograms from nInput as governed by the following equation.

$$L(x,y,m,D) = \frac{\sum_{i=1}^{nInput} I_i(x+(i-m)D, y)}{nInput} \qquad \text{Equation 3.1}$$

$I_1 - I_8$ refer to detector images

$I_m(x,y)$ refers to the intensity of the pixel at position x,y in image m.

D is the possible disparity between two images.

L(x,y,m,D) refers to the intensity of the pixel at position x,y of laminogram calculated using disparity D based on voting image m.

Figure 3.6 illustrates a laminogram for a disparity value of 0, generated by 8 input images; in addition, Figure 3.7 presents a series of thumbnail laminograms to illustrate the effect of in focus and out of focus conditions, for a series of objects placed at different depths in the luggage. The objects in a black box are in focus. The in focus content of each laminogram represents a specific depth plane in the object space.

Figure 3.6 Laminogram of a luggage item for the disparity value 0.

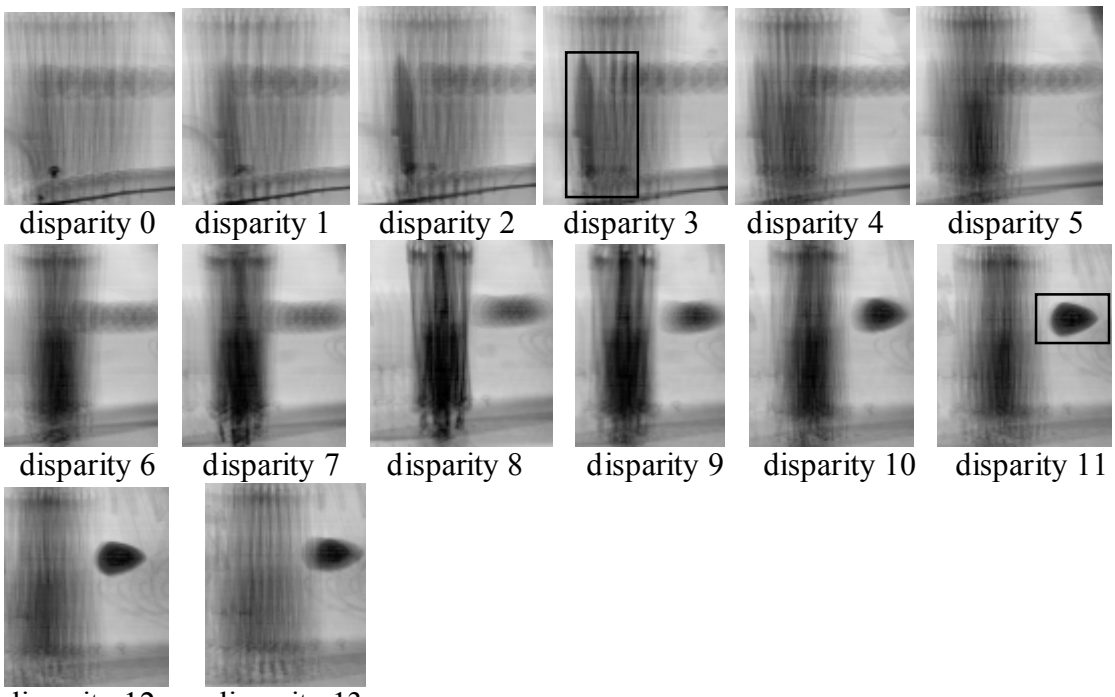


Figure 3.7 A sequence of 14 laminograms illustrating in focus and out of focus conditions for a series of objects at different depths.

Initially, for one pixel, one disparity is associated with a nInput-1 individual matching cost, however after a set of laminograms were created, for one pixel, one disparity is only associated with one matching cost, which is calculated between the voting image and the laminogram associated with the disparity under consideration.

For the laminogram intensity criterion, matching cost is given by:

$$COSTFINAL_{LamI}(x,y,m,D)=L(x,y,m,D) \qquad \text{Equation 3.2}$$

COSTFINAL(x,y,m,D) refers to the final matching cost for a pixel at position x,y for a voting image m with a disparity value of D.

3.3.2 Sum of reciprocals (SOR)

Equation 3.1 applies an equal weight to each input image to produce laminograms, matching cost is then calculated based on those laminograms as in Equation 3.2. Inevitably, the averaging nature of laminography tends to hide individual intensity of each input images. To compensate for this effect, a new matching cost calculation method based on the sum of reciprocals is proposed to preferentially emphasize each individual matching cost value calculated between voting image and input image. The method provides a weight to each matching cost by the simple expedient of converting the matching cost into its reciprocal. In this manner, low matching cost values will have

a higher score, by virtue of their higher reciprocal value. SSD criterion utilise SOR matching cost calculation method.

Figure 3.8 illustrates the procedures of using sum of reciprocals matching cost calculation method to calculate matching cost table for $D16^{o}$, during the process of generating target image $S14^{o}$ using the designed V-DMX algorithm. Matching cost table for $D12^{o}$ is also calculated using sum of reciprocals matching cost calculation method, the procedures are quite similar as in Figure 3.8.

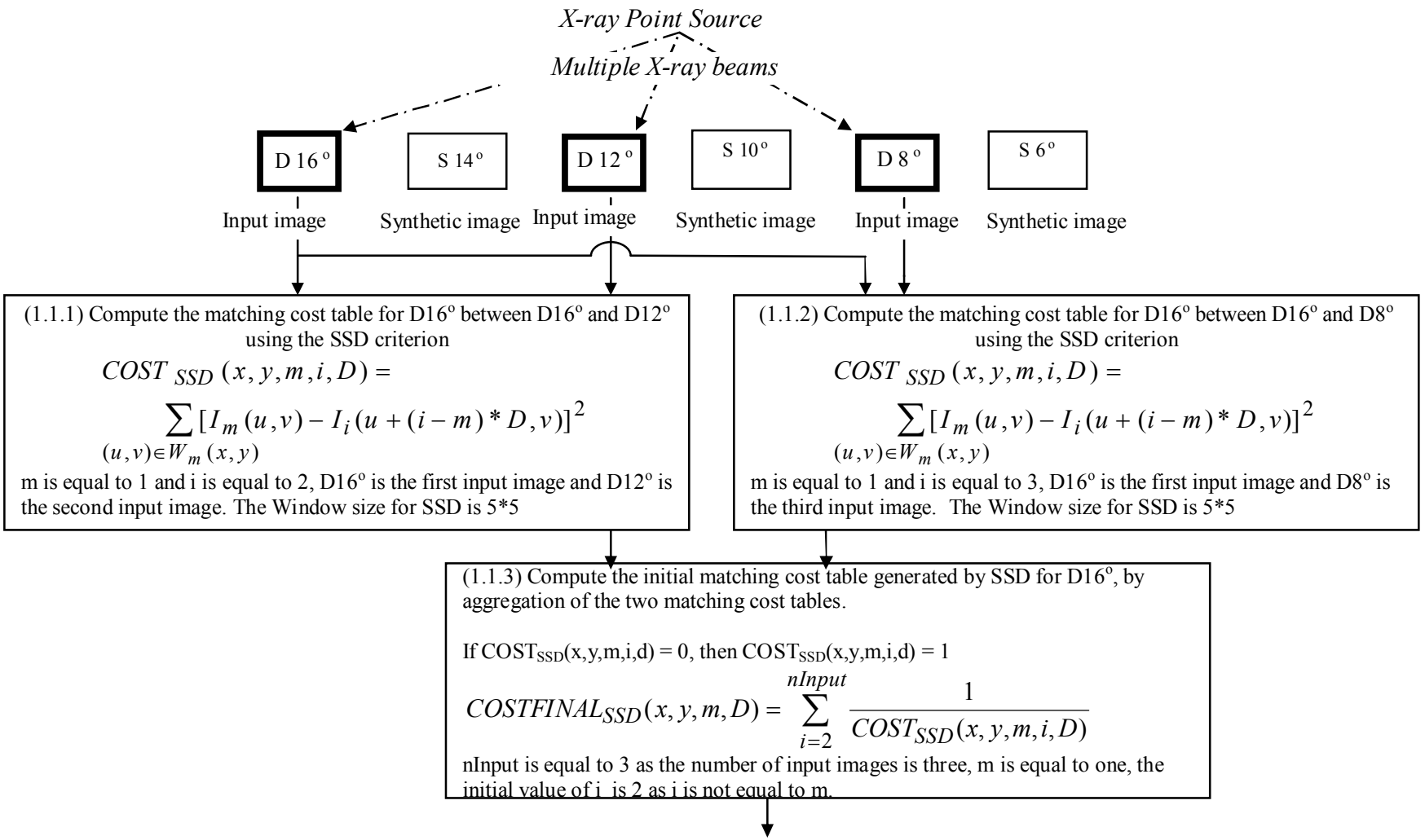


Figure 3.8 Graphical illustration of using the sum of reciprocals matching cost calculation method calculate matching cost table for D16° during the process of synthesising a target image at 14 °.

For one pixel, one disparity is associated with a nInput-1 individual matching cost, each individual cost is calculated as:

$$COST_{SSD}(x, y, m, i, D) = \sum_{(u,v)\in W_m(x,y)} [I_m(u,v) - I_i(u + (i-m) * D, v)]^2$$

Equation 3.3

COST(x,y,m,i,D) refers to the matching cost for a pixel at position x,y for voting image m with disparity D calculated between voting image m and input image i.

nInput-1 individual matching costs are aggregated according to:

If $COST_{SSD}(x,y,m,i,d)=0$, then $COST_{SSD}(x,y,m,i,d)=1$ Equation 3.4

$$COSTFINAL_{SSD}(x, y, m, D) = \sum_{i=1}^{nInput} \frac{1}{COST_{SSD}(x, y, m, i, D)} \quad (\text{i not equal m})$$

Equation 3.5

Thus for one pixel, one final matching cost is obtained for each disparity.

3.4 Voting for disparity

The result of the previous process is sets of matching cost value tables for all possible disparities for each voting image. These matching cost value tables can also be called disparity information tables, as from these matching cost values, a disparity can be decided for each pixel by applying a winner-take–all method [58]. As all voting images are of the same scene, the disparities of corresponding points in each voting image should, ideally, be identical but the ill-posed nature of the correspondence problem produces many potentially conflicting disparities. In this work a voting system is developed to resolve the conflict between multiple disparities. Rank voting is a correspondence disparity consistency voting process. Using other nVote-1 disparity information tables optimises each disparity information table; the voting image is defined as the image whose disparity information table is involved in the disparity consistency voting process. Applying the correspondence disparity consistency voting among voting images can optimise every disparity information table. In this work, only the disparity information tables for two reference images is required, so the output from this stage is two disparity information tables that have been optimised by applying disparity voting. Figure 3.9 describes the flowchart of the voting process.

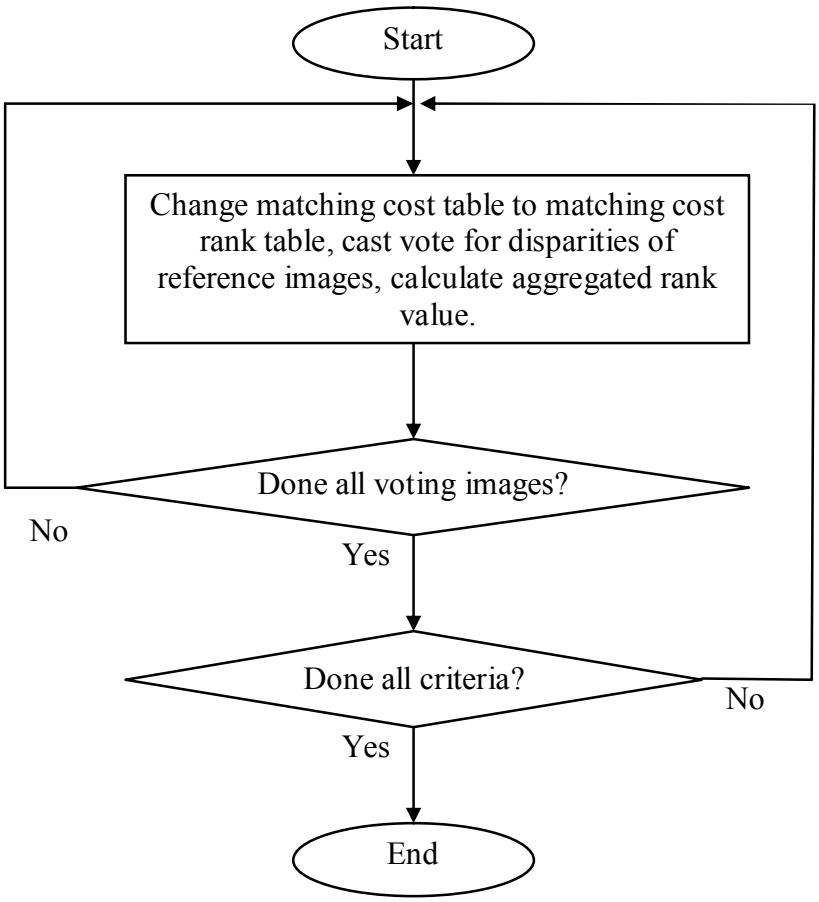


Figure 3.9 Flowchart of disparity voting.

Figure 3.10 shows the disparity voting procedure for matching cost table of D16^{o} generated by SSD criterion, and Figure 3.11 shows the disparity voting procedure for matching cost table of D16^{o} generated by LamI criterion, during the process of synthesising target image at 14^{o} using designed V-DMX algorithm. The voting procedures for matching cost table of D12^{o} generated by SSD criterion are quite similar as in Figure 3.10.

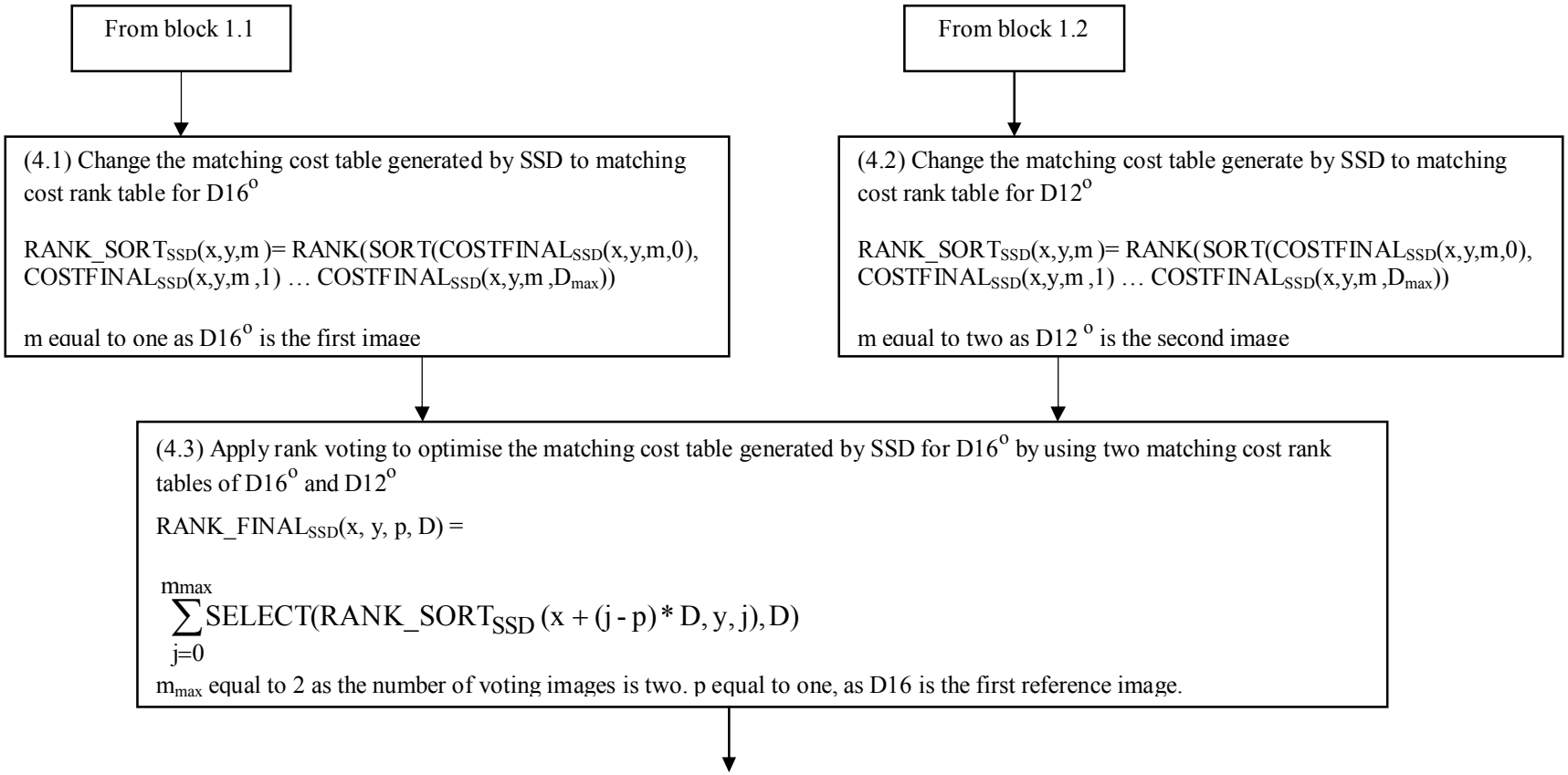


Figure 3.10 Graphical illustration of using disparity voting method for matching cost table of D16° generated by SSD during the process of synthesising target image at 14 °.

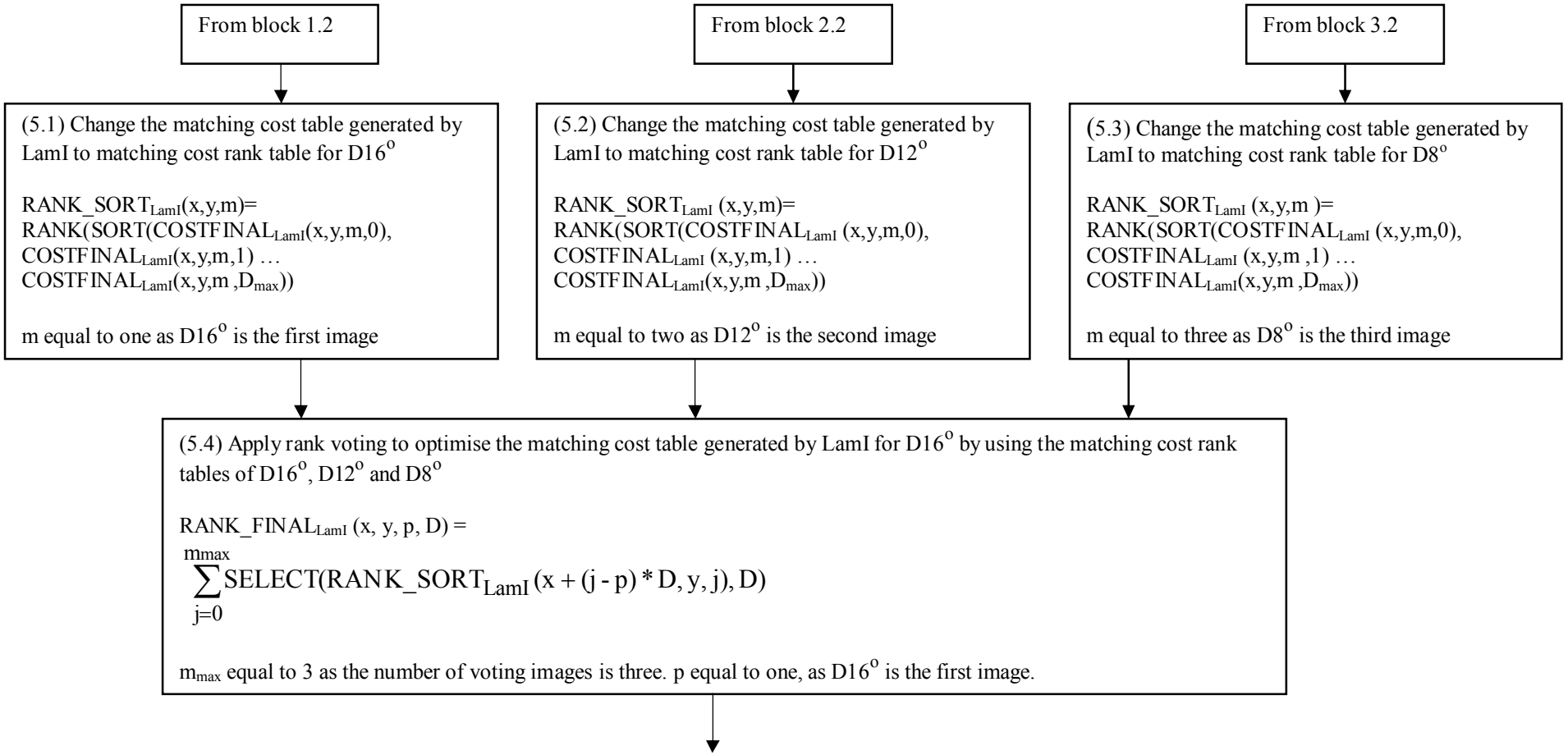


Figure 3.11 Graphical illustration of using disparity voting method for matching cost table of D16° generated by LamI during the process of synthesising a target image at 14 °.

Consider a pixel (x,y) in a perspective image $I_1(x,y)$ that has a correspondence at disparity d in a second perspective image $I_2(x+d,y)$. The search for correspondence $I_1(x,y)$ involves calculating the scores for a range of disparities in I_2. The disparity with the highest score will be identified as the "winner". The common approach to improve the integrity of the matching is to perform a reverse search, i.e, taking $I_2(x+d,y)$ as the reference and calculating the score of d from a disparity window in I_1. It should be noted that the search direction for $I_1(x,y)$ and $I_2(x+d,y)$ is different and involves a different set of pixels, thus the calculated scores will be different. Ideally, the likelihood of matching $I_2(x+d,y)$ from $I_1(x,y)$ should be the same as the likelihood of matching $I_1(x,y)$ from $I_2(x+d,y)$. In other words, both pixels must "vote" for each other as the corresponding point. The adoption of a "voting" concept in this work is associated with the election process that occurs in choosing the "winner". If each disparity represents a candidate, then the number of voting images is the number of voters eligible. Consider the example given here, if the scores produced by I_1 and I_2 are taken into account for deciding the "winner", and then both images are defined as the voting images. In the simplest form of stereo matching which uses only scores generated by I_1 to make decision, then only I_1 is defined as the voting image but both I_1 and I_2 are classified as input images. So, if the voting images indicate that a disparity has a very high matching likelihood (i.e, vote in favour for a disparity), and then the disparity will be very likely to be declared as the "winner". However, the application of the voting concept to multiple images offers a greater potential to alleviate the matching ambiguities.

A system is introduced to initially rank the aggregated matching cost of individual voting images before summing the ranks together. In the context of this work, for SSD criterion with SOR matching cost calculation, the disparity with the smallest score (i.e, the least influential disparity) is given the lowest rank i.e, 1, while the disparity with the largest score is assigned to the largest value, i.e, the disparity window size plus 1. In this manner, the ranking method gives an equal voting right to each voting image.

The ranking process is described by following relationship:

$$RANK\_SORT(x,y,m) = RANK(SORT(COSTFINAL(x,y,m,0), COSTFINAL(x,y,m,1) \ldots COSTFINAL(x,y,m,D_{max})) \quad \text{Equation 3.6}$$

RANK_SORT(x,y,m) operates on the matching cost values, sorts the values and then adjusts the results to give a rank value. The result is an array of tuples that consists of (position in sort, disparity value) for the pixel at x,y in voting image m.

For example, if the disparity window is 10, then each voting image can vote for a disparity by using a rank of 1 to 11. The integration of the ranking approach to enable fair treatment of the voters and the common voting concept to exploit multiple images, forms the methodological basis of the "rank voting" proposed in this research programme.

After matching cost is converted into rank, the voting process can be described by the following equation:

$$\text{RANK\_FINAL}(x, y, p, D) = \sum_{j=0}^{\text{nVote}} \text{SELECT}(\text{RAN K\_SORT}(x + (j - p) * D, y, j), D)$$

Equation 3.7

RANK_FINAL(x, y, p, D) refers to final adjusted rank value for pixel x,y with respect to reference image p and disparity value D.
SELECT(array, disparity) returns the rank value given the disparity D.

With the aid of Table 3.1 and Table 3.2, the procedure of determining pixel correspondences using the matching cost computation and the rank based voting processes is described. Consider eight detector images are available, three of them are selected as the input images and only two out of the three input images are used in the voting process. The task is to identify the correspondence of the pixel (3,0) in the reference image from a disparity window of two pixels by using the sum of reciprocals method and, for example, an absolute intensity difference matching criterion. The three input images are denoted as I_1, I_2 and I_3; the two voting images are represented by V_1 and V_2; the three disparities are symbolized as d_0, d_1 and d_2. Table 3.1 exemplifies the aggregation of the matching cost process for four pixels denoted by $V_1(3,0)$, $V_2(3,0)$, $V_2(4,0)$ and $V_2(5,0)$. $|I_1 - I_2|$ is the absolute intensity difference between I_1 and I_2. $V_1(3,0;d_0)$ is the sum of reciprocals matching cost of disparity 0 for the pixel (3,0) in the first voting image. The process of matching cost aggregation generates a set of costs for each pixel in the voting image as indicated by the "Cost aggregation" rows in Table 3.1. With an appropriate sorting of these costs, each disparity with respect to the reference

image pixel will have nVote of costs. Table 3.2(a) illustrates the sorting outcomes. To illustrate the direct summation of the costs, the cost aggregation functions in the table are replaced by arbitrary values in the "Example cost" rows, and correspondingly Table 3.2(b) is derived. The "Rank" rows in Table 3.1 and Table 3.2(c) are the results of the conversion. This example illustrates that the ranking approach grants the same number of votes to each cost produced by the matching cost computation stage. At the voting stage, the individual votes granted are pooled together to determine the best disparity candidate. The disparity with the most votes, i.e, largest total rank value is selected as the "best" correspondence. Comparatively, the rank voting system selected d_2 disparity as the best disparity candidate while the direct summation of costs indicates that d_1 is the better choice.

It can be appreciated that matching cost computation and rank voting processes work in tandem to form the essential part of the algorithm. The number of possible combinations of nInput and nVote depends on nDetector. For example, given the nDetector has a value of 8, then there are 35 possible combinations as tabulated in Table 3.3. The 3-2 pairing refers to using three input images and two voting images for the process. The performance of these pairs relies firmly to the nature of the criterion involved. Different criteria respond differently to the algorithm and thus produce a set of different results. So, one of the main tasks in this programme is to investigate which is the best pairing of the nInput and nVote for the selected criterion. Empirical results and associated discussion are organised in Chapter 5, 6 &7.

		$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$
$V_1(3,0)$	$\lvert I_1 - I_2 \rvert$	$\frac{1}{I_1(3,0) - I_2(3,0)}$	$\frac{1}{I_1(3,0) - I_2(4,0)}$	$\frac{1}{I_1(3,0) - I_2(5,0)}$
	$\lvert I_1 - I_3 \rvert$	$\frac{1}{I_1(3,0) - I_3(3,0)}$	$\frac{1}{I_1(3,0) - I_3(5,0)}$	$\frac{1}{I_1(3,0) - I_3(7,0)}$
Cost Final		$V_1(3,0;d_0)$	$V_1(3,0;d_1)$	$V_1(3,0;d_2)$
Example cost		10	2	15
Rank		**2**	**1**	**3**
$V_2(3,0)$	$\lvert I_2 - I_1 \rvert$	$\frac{1}{I_2(3,0) - I_1(3,0)}$	$\frac{1}{I_2(3,0) - I_1(2,0)}$	$\frac{1}{I_2(3,0) - I_1(1,0)}$
	$\lvert I_2 - I_3 \rvert$	$\frac{1}{I_2(3,0) - I_3(3,0)}$	$\frac{1}{I_2(3,0) - I_3(4,0)}$	$\frac{1}{I_2(3,0) - I_3(5,0)}$
Cost Final		$V_2(3,0;d_0)$	$V_2(3,0;d_1)$	$V_2(3,0;d_2)$
Example cost		2	12	5
Rank		**1**	**3**	**2**
$V_2(4,0)$	$\lvert I_2 - I_1 \rvert$	$\frac{1}{I_2(4,0) - I_1(4,0)}$	$\frac{1}{I_2(4,0) - I_1(3,0)}$	$\frac{1}{I_2(4,0) - I_1(2,0)}$
	$\lvert I_2 - I_3 \rvert$	$\frac{1}{I_2(4,0) - I_3(4,0)}$	$\frac{1}{I_2(4,0) - I_3(5,0)}$	$\frac{1}{I_2(4,0) - I_3(6,0)}$
Cost Final		$V_2(4,0;d_0)$	$V_2(4,0;d_1)$	$V_2(4,0;d_2)$
Example cost		100	120	1
Rank		**2**	**3**	**1**
$V_2(5,0)$	$\lvert I_2 - I_1 \rvert$	$\frac{1}{I_2(5,0) - I_1(5,0)}$	$\frac{1}{I_2(5,0) - I_1(4,0)}$	$\frac{1}{I_2(5,0) - I_1(3,0)}$
	$\lvert I_2 - I_3 \rvert$	$\frac{1}{I_2(5,0) - I_3(5,0)}$	$\frac{1}{I_2(5,0) - I_3(6,0)}$	$\frac{1}{I_2(5,0) - I_3(7,0)}$
Cost Final		$V_2(5,0;d_0)$	$V_2(5,0;d_1)$	$V_2(5,0;d_2)$
Example cost		20	2	22
Rank		**2**	**1**	**3**

Table 3.1 Example results of the sum of reciprocals method.

Note: Actual example cost values depend on the number of input images and other considerations; examples here are for illustrative purposes only.

3.2(a)

	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$
V_1	$V_1(3,0;d_0)$	$V_1(3,0;d_1)$	$V_1(3,0;d_2)$
V_2	$V_2(3,0;d_0)$	$V_2(4,0;d_1)$	$V_2(5,0;d_2)$

3.2(b)

	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$
V_1	10	2	15
V_2	2	120	22
Sum of costs	12	122	37

3.2(c)

	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$
V_1	2	1	3
V_2	1	3	3
Sum of ranks	3	4	6

Table 3.2 Tabulated data for voting process to decide the correspondence of the pixel (3,0) in the reference image.

	nVote							
nInput	2-1	2-2						
	3-1	3-2	3-3					
	4-1	4-2	4-3	4-4				
	5-1	5-2	5-3	5-4	5-5			
	6-1	6-2	6-3	6-4	6-5	6-6		
	7-1	7-2	7-3	7-4	7-5	7-6	7-7	
	8-1	8-2	8-3	8-4	8-5	8-6	8-7	8-8

Table 3.3 Possible combinations of input images (nInput) and voting images (nVote) for eight detector images.

3.5 Synthetic pixel generation

3.5.1 Interpolation of a virtual target image employing a single criterion

After the voting for disparity process is completed, the disparity value for each pixel can be decided by applying the winner-take-all method. Once the disparity values are determined, the next task is to synthesise the intermediary target images, for a single criterion, the process is described by the following Figure 3.12.

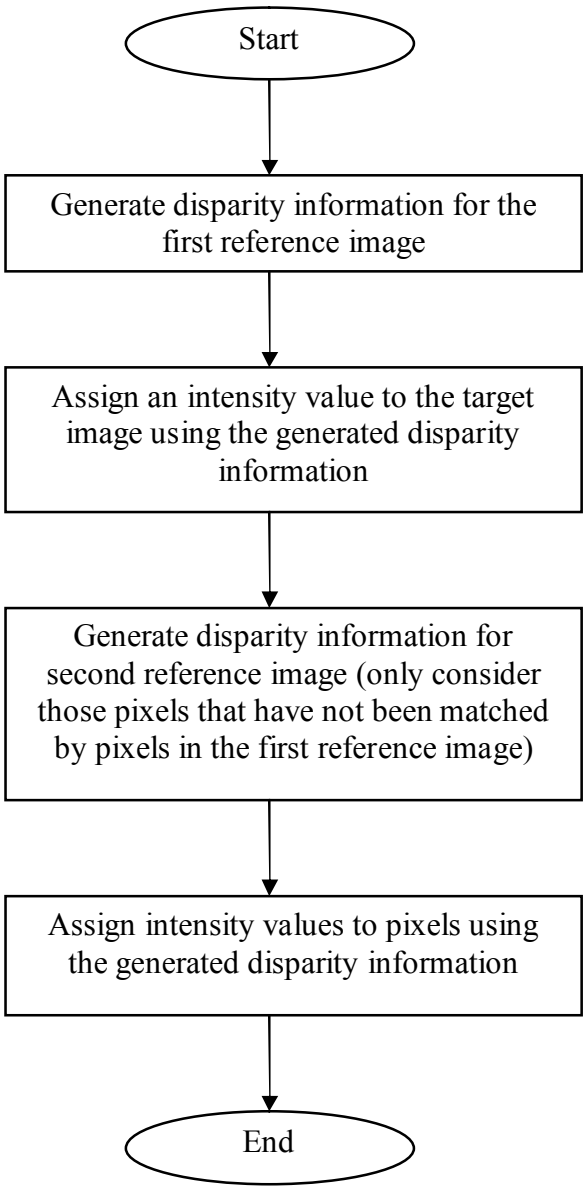


Figure 3.12 Flowchart of the synthetic pixels generation.

The process involves establishing disparities, interpolating the coordinate position of the intermediary target pixels and assigning intensity values to these pixels. The interpolation process requires ensuring that the pixel coordinates in the target images are identical to the detector images. The pixel position is a function of the angular separation between adjacent detector images. For a relatively small angular separation between the X-ray beams, the midpoint view is adequately represented by placing the target pixels at a coordinate position, which is half way between its corresponding pixels in the two adjacent detector images. The error introduced by this assumption is negligible in this work because of the relatively small angular increments between adjacent perspective views. Based on this approximation, the interpolation algorithm begins by using the disparity information table of the first reference image to calculate the coordinate values. The calculation of the absolute intensity of the intermediary

target pixel is reliant upon the solution of ill posed problems and in practical terms is impossible. Therefore, an approximation is obtained from the intensity of corresponding pixels in the reference image. For each pixel, computing half of the winning disparity and assigning the pixel intensity from the first reference image determine the intermediary coordinate position and intensity of the pixel. For example, if the winning disparity (i.e, disparity with most votes) for a reference pixel at the coordinate of (3,0) is 4, then the coordinate of the synthetic pixel will be (5,0).

Due to multiple corresponding features, each pixel in the reference image will exhibit a number of potential disparities, which in turn produce several intermediary pixels. It is difficult to establish a rationale for deciding how many disparities should be considered as potential correspondences other than the disparity with the most votes. The method proposed here, to address multiple correspondences, is to proceed to using the disparity information of the second reference image to assign intensity to pixels in the target image. Only pixels in the second reference image that are not matched by pixels in the first reference image are entitled to participate in the interpolation process. One reason why some pixels remain unmatched in the first reference image is due to "legitimate" multiple correspondences. In other words, some pixels in the first reference image should be expected to match to more than one pixel in the second reference image. However, the algorithm allows initially many to one pixel matches, so some pixels in the second reference image remain unmatched; by using those unmatched pixels in second reference image to synthesis the target image, the algorithm indirectly allows a pixel in first reference image to have two disparity values, which is a normal situation in X-ray images.

The transparency inherent in transmission images make it possible for a large number of pixels to be considered for the same coordinate position in the target image. In this work the lowest intensity pixel takes precedence over the higher intensity pixels when duplicate correspondences occur. The rationale is to consider a pixel that has duplicate correspondences as equivalent to a pixel "seeing" an overlapping structure. In theory, the pixel intensity associated with an overlapping structure is always lower than the intensity produced by individual object in the overlapping structure. The algorithm is developed to identify the lowest intensities as the best representatives of a potential match.

The first step of this process is to determine the final the maximum rank value for a pixel:

$$MAX\_RANK\_FINAL(x,y,p)=MAX(RANK\_FINAL(x,y,p,0), RANK\_FINAL(x,y,p,1) \ldots RANK\_FINAL(x,y,p,D_{max})) \quad \text{Equation 3.8}$$

Assign a disparity value to this pixel and employ this pixel to generate a pixel in the target image:

$$DISAPRITY(x,y,p)= GETDISPARITY(MAX\_RANK\_FINAL(x,y,p), RANK\_FINAL(x,y,p,0), RANK\_FINAL(x,y,p,1) \ldots RANK\_FINAL(x,y,p,D_{max})) \quad \text{Equation 3.9}$$

$$I_{target}(x+DISAPRITY(x,y,p)/2, y)= I(I_{ref}(x,y,p)) \quad \text{Equation 3.10}$$

MAX_ RANK_FINAL(x,y,p) refers to the maximum value of RANK_FINAL of pixel at position x,y for reference image p.

DISAPRITY(x,y,p) refers to the disparity of pixel at position x,y of reference image p.

GETDISPARITY(MAX_RANK, RANK_FINAL…) returns the disparity value associated with the MAX_ RANK in the array of RANK_FINAL.

$I_{target}(x,y)$ *refers to the intensity of pixel at position x,y in the target image.*

$I_{ref}(x,y,p)$ *refers to the intensity of the pixel at position x,y in the reference image p.*

For all pixels in the second reference image that have not been matched by the first reference image, employ the second reference image to assign intensity values to those pixels in the target image that have not already been assigned an intensity value:

$$I_{target}(x-DISAPRITY(x,y,p)/2, y)= I(I_{ref}(x,y,p)) \quad \text{Equation 3.11}$$

3.5.2 Interpolation in iterations employing a single criterion

This section describes how the interpolation can be divided into $D_{max}+1$ iterations as described in Figure 3.13.

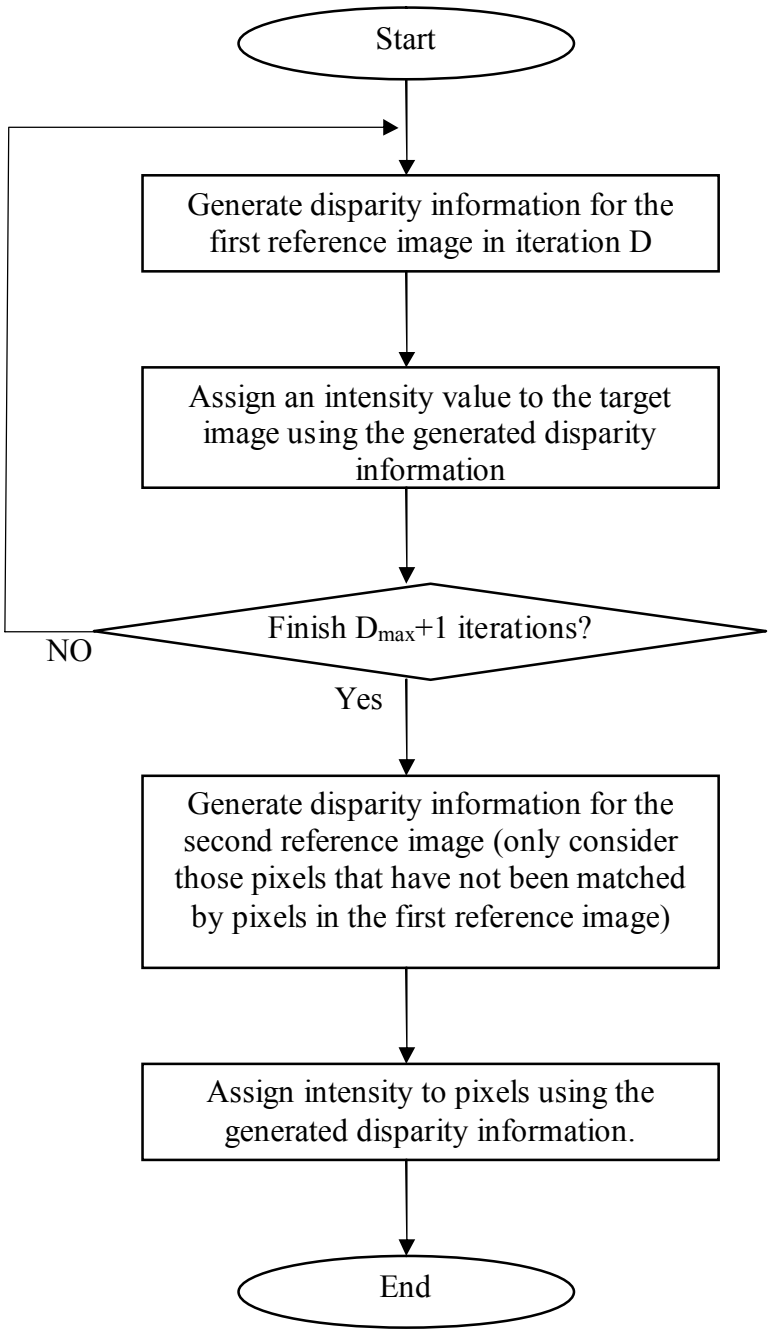


Figure 3.13 Flowchart of interoperation in iterations.

The outcome of the disparity voting is an aggregated rank voting value of the pixels in the reference images. If the maximum disparity is D_{max}, then for one voting image the maximum rank value it can vote for is D_{max}+1, the minimum value is 1. After all voting images cast a vote, the system chooses the disparity that has the maximum rank voting value to be the final disparity. The maximum value of MAX_ RANK_FINAL is

$$MAX(D_{max})=MAX(MAX\_\ RANK\_FINAL(x,y,p)) = (D_{max}+1)*nVote \qquad \text{Equation 3.12}$$

The minimum value of MAX_ RANK_FINAL is

$$MIN(D_{max})=MIN(MAX\_RANK\_FINAL(x,y,p))= 1* nVote \quad \text{Equation 3.13}$$

MAX(D_{max}) is the maximum possible rank value for a pixel with D_{max}.

MIN(D_{max}) is the minimum possible rank value for a pixel with D_{max}.

Theoretically, the MAX_ RANK_FINAL value can be any value between the minimum rank value $MIN(D_{max})$ to the maximum value $MAX(D_{max})$.

The interpolation process can be divided into $D_{max}+1$ iterations, assign each iteration D ($1=<D<= D_{max}+1$) an iteration value:

$$ITERATION(D)=MAX(D_{max})-(D-1)*MIN(D_{max}) \quad \text{Equation 3.14}$$

ITERATION(D) is the value for iteration D.

For each iteration D:

If (MAX_ RANK_FINAL(x,y,p)> =ITERATION(D) then

DISPARITY(x,y,p)=GETDISPARITY(MAX_RANK_FINAL(x,y,p),
RANK_FINAL(x,y,p ,0), RANK_FINAL(x,y,p ,1) … RANK_FINAL(x,y,p ,D_{max}))

Else go to next iteration.

For each iteration D, only pixels that satisfy the iteration value are assigned a disparity value. The target image contains many pixels with intensity unassigned in the early iterations. As the iteration progresses, the unassigned pixels will also be processed. The transparency property inherent in X-ray images makes it possible for a number of pixels to be considered for the same coordinate position in the synthetic image. While the pixels synthesised by the early iterations are assumed more reliable, the ideal solution should combine the intensity of all the associated pixels to provide a concluding intensity value. In practice, the generation of the intermediary pixel is a highly complicated process which is mainly due to ill-posed nature of the correspondence problem and the polychromatic nature of the X-ray source. Therefore, in this research, the lowest intensity pixel takes precedence when multiple correspondences occur. Interpolation in iterations is used in later Section 3.7.1.2 to improve the laminogram intensity criterion and in Chapter 7 for experiments.

3.5.3 Interpolation employing dual criterion

Section 3.5.1 describes the synthetic pixel generation process for one criterion; however, the V-DMX algorithm employs dual criteria. This section describes the synthetic pixel generation process employing dual criteria.

Independently, each criterion offers different but important solutions to address complimentary problems encountered in the algorithm development. The SSD criterion is a window based matching operator. Thus, it tends to create false matches when the criterion is applied in regions of a discontinuity and overlapping image structure. The LamI criterion provides an effective alternative to handling these types of image features, although the SSD criterion is superior to the LamI criterion overall; the performance comparison of these two criteria is presented in Section 7.1.

Both criteria are incorporated into the, "Dual Criteria Multiple X-ray Images Synthetic System". Each criterion is initially run independently of one another until the synthetic pixel generation process is reached. The LamI criterion utilises 3 input images and 3 voting images and stops at iteration one, consequently, only some pixels are assigned matching cost values in the matching cost table. The SSD criterion utilises 3 input images and 2 voting images and employs the SOR matching cost calculation.

Results from the two criteria are incorporated in synthetic pixel generation process, disparity information from LamI is used first to assign intensity to pixels in the target image, for those pixels in target image have not been assigned intensity, the disparity information from the SSD criterion is used to assign intensity values to those pixels. Finally, the disparity information for the second reference image generated by the SSD criterion is employed to assign intensity to pixels in target image that have not been assigned an intensity value by LamI criterion. The new interpolation process is described by Figure 3.14.

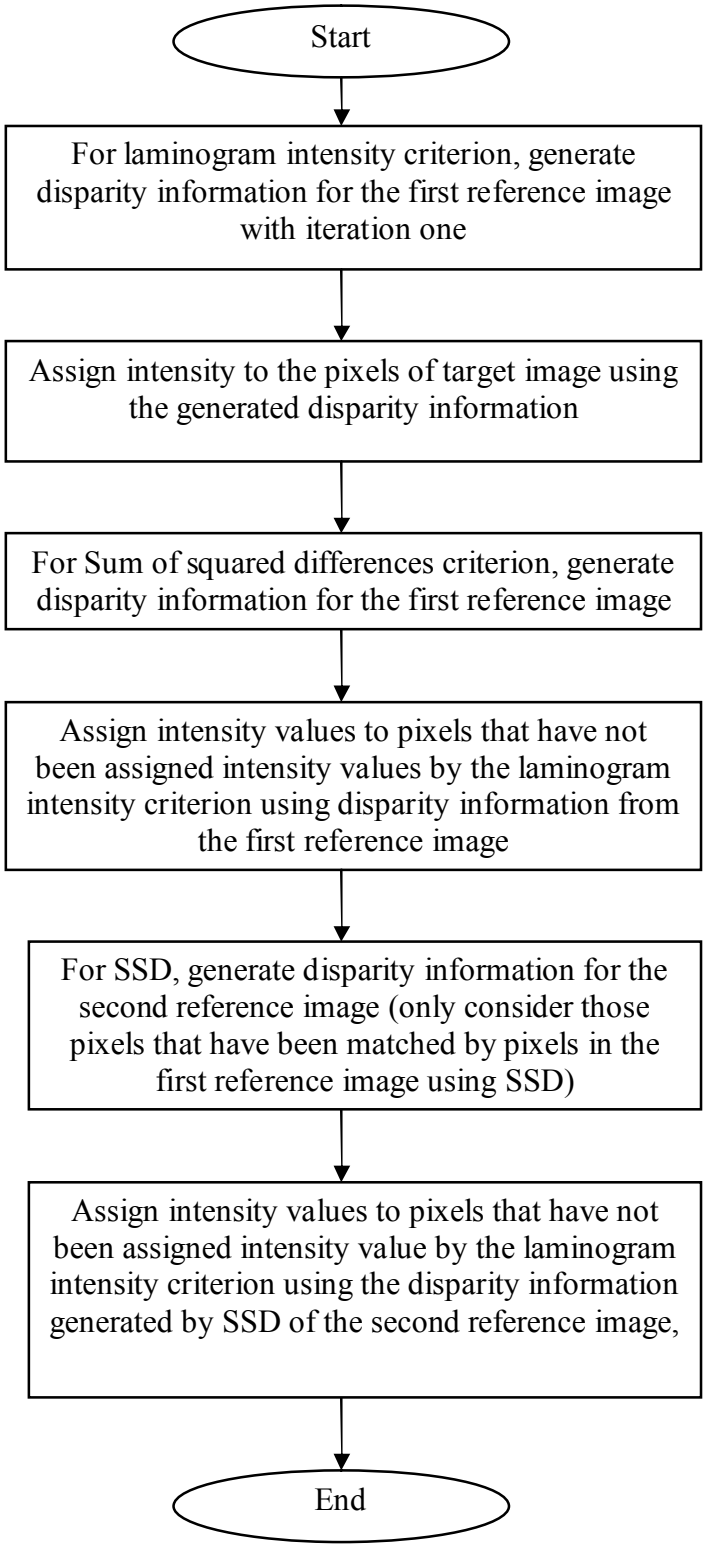


Figure 3.14 Flowchart of the synthetic pixels generation algorithm incorporating dual criterion.

Figure 3.15 and 3.16 shows the procedures of synthetic pixels generation to synthesise target image at 14^{o} using designed V-DMX algorithm.

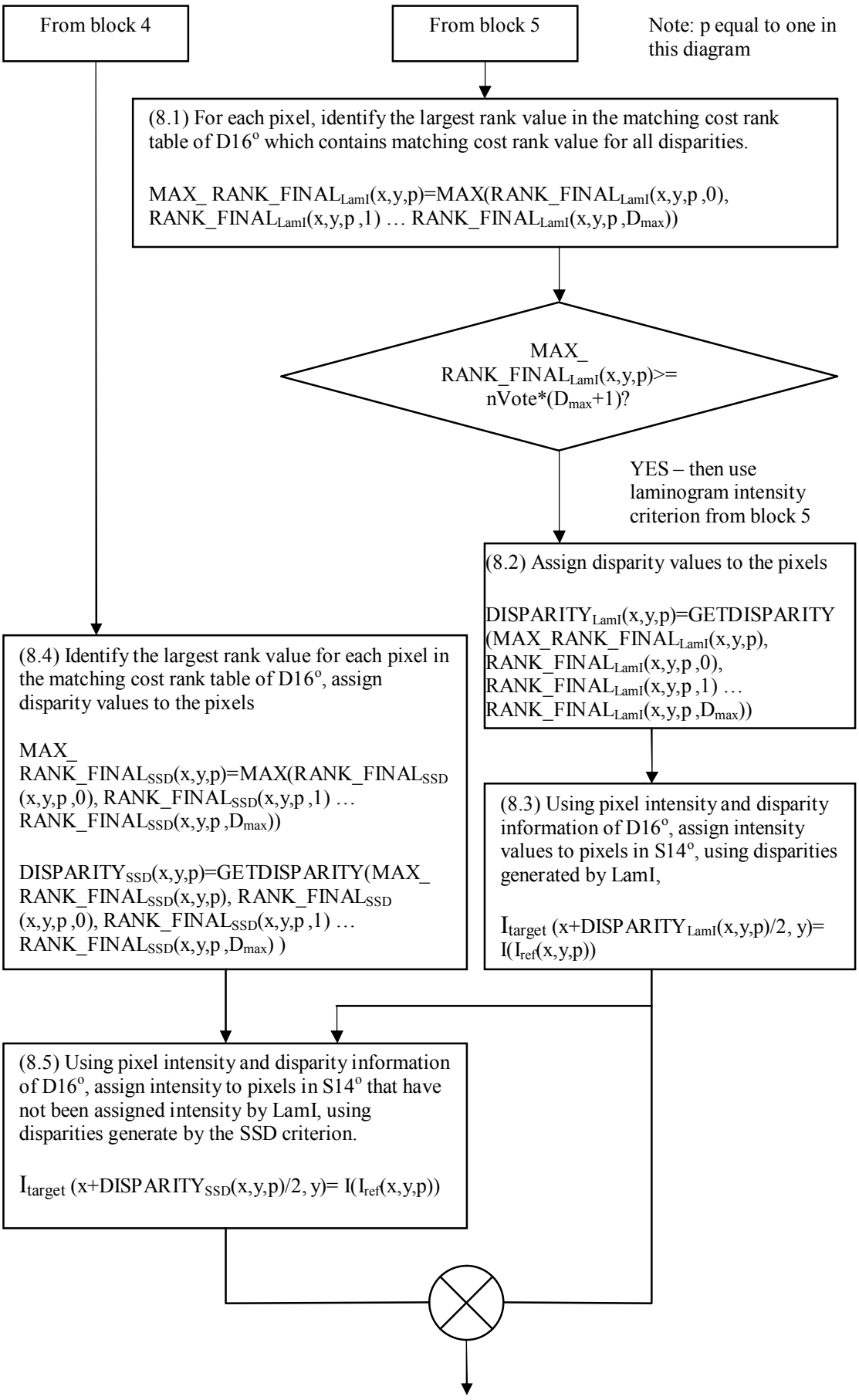


Figure 3.15 Graphical illustration of synthetic pixel generation using the first reference image to synthesis a target image at 14 °.

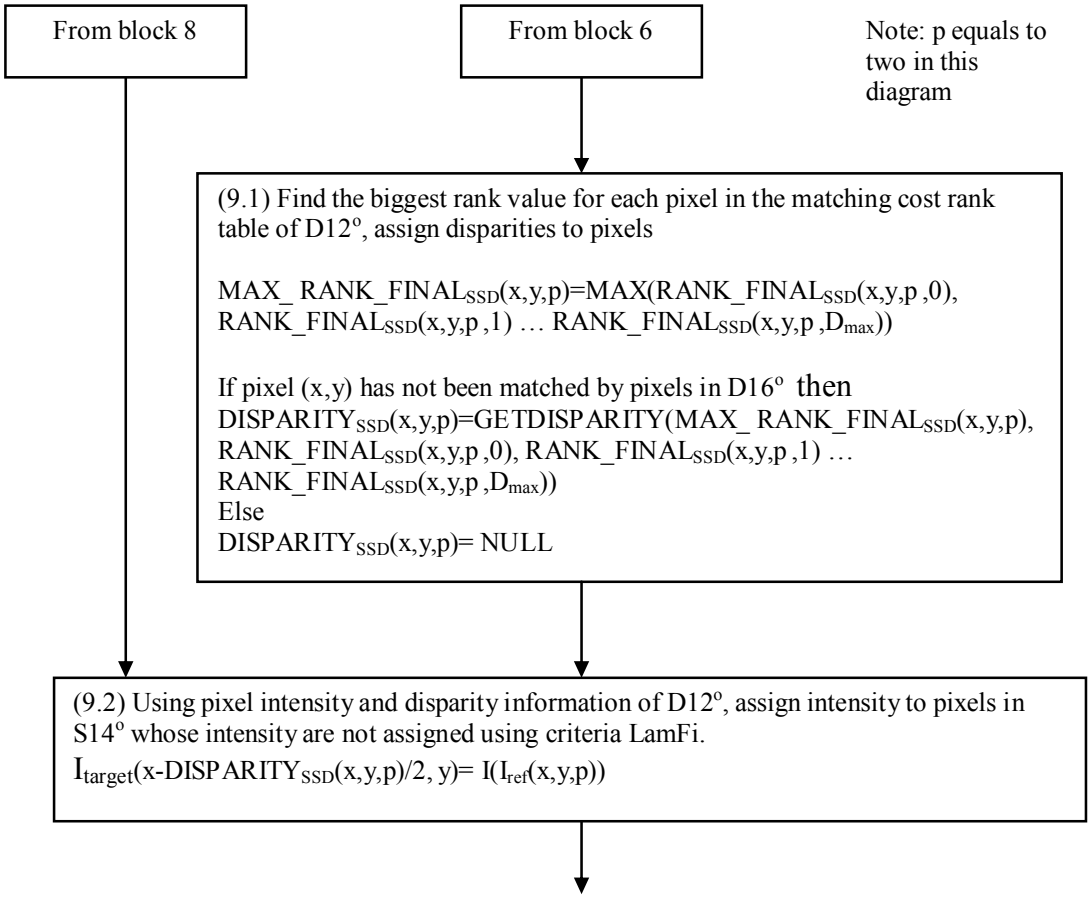


Figure 3.16 Graphical illustration of synthetic pixel generation using second reference image to synthesis target image at 14 °.

The process can be described by a set of equations: Equation 3.8 is used to find the maximum rank value, the following relationship checks whether the maximum rank value generated by LamI criterion is equal to the maximum possible rank value, if it is then a disparity value is assigned to the pixel, and it is used to generate a pixel in target image.

If (MAX_ RANK_FINAL$_{LamI}$(x,y,p)=nVote*(D_{max}+1):

DISPARITY$_{LamI}$(x,y,p)= GETDISPARITY(MAX_RANK_FINAL$_{LamI}$(x,y,p) , RANK_FINAL$_{LamI}$(x,y,p,0), RANK_FINAL$_{LamI}$(x,y,p,1) … RANK_FINAL$_{LamI}$(x,y,p,D_{max})) Equation 3.15

I_{target} (x+DISPARITY$_{LamI}$(x,y,p)/2, y)= I(I_{ref}(x,y,p)) Equation 3.16

For SSD criterion, Equation 3.8 is used to get maximum rank value, then disparity for every pixel can be assigned by following equation:

$$DISPARITY_{SSD}(x,y,p)= GETDISPARITY(MAX\_ RANK\_FINAL_{SSD}(x,y,p), RANK\_FINAL_{SSD}(x,y,p,0), RANK\_FINAL_{SSD}(x,y,p,1) \ldots RANK\_FINAL_{SSD}(x,y,p,D_{max}))$$ Equation 3.17

For all pixels in the target image that have not been assigned intensity by laminogram intensity criterion, using first reference image, assign intensity to pixels in target image:

$$I_{target}(x+DISPARITY_{SSD}(x,y,p)/2, y)= I(I_{ref}(x,y,p))$$ Equation 3.18

For all pixels in the target image, whose intensity is not assigned by the LamI criterion, using pixels in the second reference image that have not been matched by pixels in the first reference image, assign intensity to pixels in target image:

$$I_{target}(x-DISPARITY_{SSD}(x,y,p)/2, y)= I(I_{ref}(x,y,p))$$ Equation 3.19

3.6 Void filling

A void in the context of this work is defined as a pixel coordinate position that is not assigned an intensity value after the interpolation process using the disparity information of both reference images. The voids exist due to errors of the disparity information and any inadequate treatment of multiple correspondences. Consider T(x,y) is a void in the target image with a coordinate (x,y), the possible correspondences matching from the first reference image (I_{ref}) is $D_{max}/2$, where D_{max} is the disparity window size. So, the intensity set of T(x,y) is $\{I_{ref}(x- D_{max} /2,y) ,\ldots.I_{ref}(x-D/2,y),\ldots I_{ref}(x, y)\}$ where D = 0, 1…. D_{max}. In order for the pixels in the set to match with the void, each pixel has a different disparity and each disparity has a vote value. By comparing the vote values of these pixels, the disparity with the most votes is chosen and its intensity is used to fill the void. The process is repeated for every void until all voids are filled. The process is described with the aid of Figure 3.17.

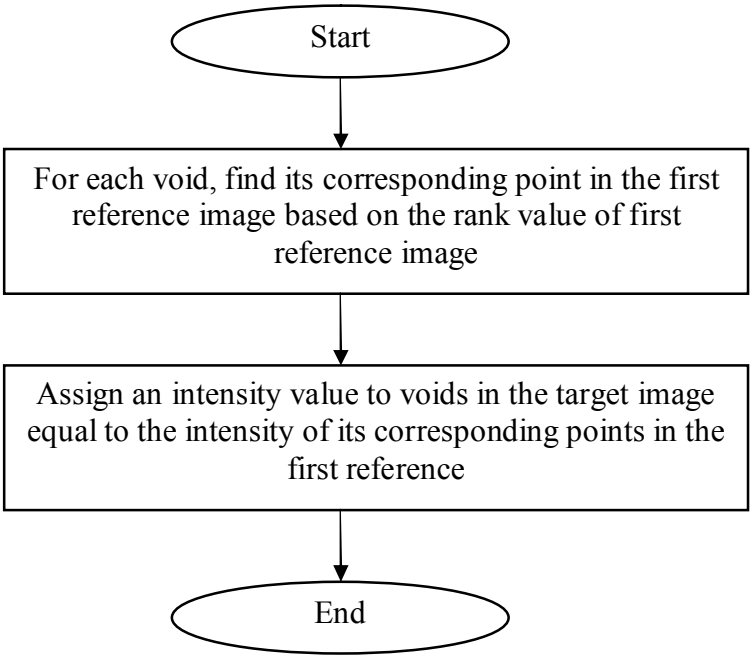


Figure 3.17 Flowchart of void filling.

Figure 3.18 shows the procedures of void filling to synthesise a target image at 14 ^{o} using designed V-DMX algorithm.

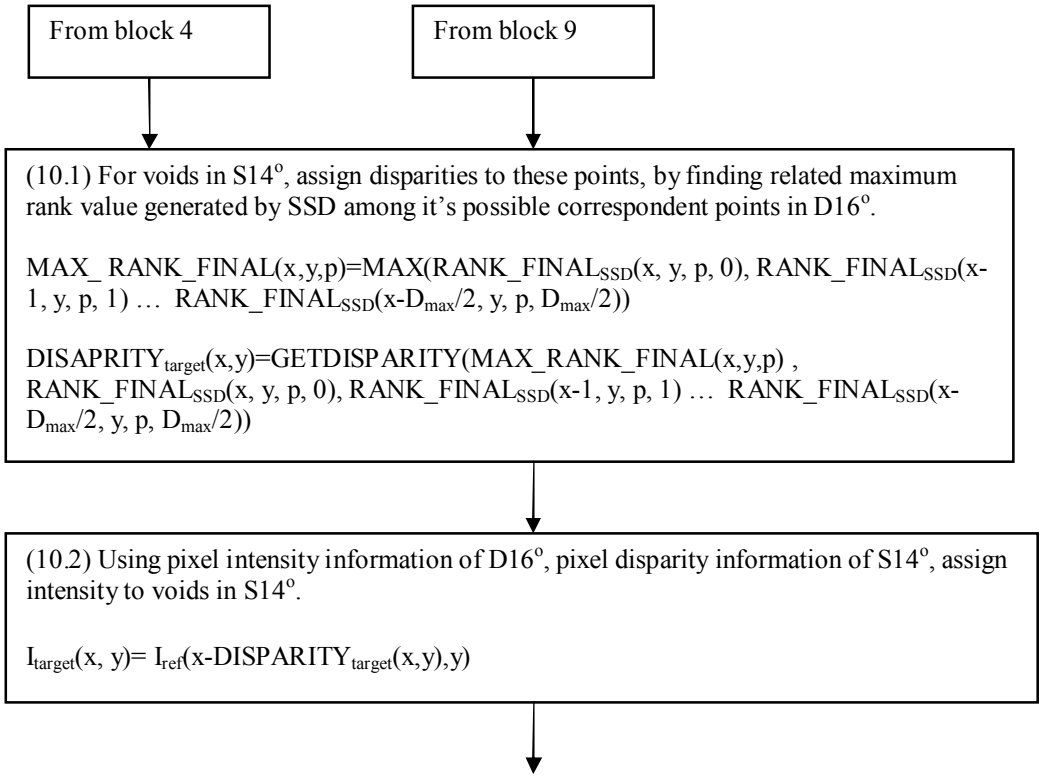


Figure 3.18 Graphical illustration of void filling to synthesis target image at 14^{o}.

The process can be described by following equations: the equation to find maximum rank value is:

$$MAX\_RANK\_FINAL(x,y,p)=MAX(RANK\_FINAL_{SSD}(x, y, p, 0), RANK\_FINAL_{SSD}(x-1, y, p, 1) \ldots RANK\_FINAL_{SSD}(x-D_{max}/2, y, p, D_{max}/2))$$

Equation 3.20

Assign a disparity value to pixels in target image:

$$DISAPRITY_{target}(x,y)= GETDISPARITY(MAX\_RANK\_FINAL(x,y,p) , RANK\_FINAL_{SSD}(x, y, p, 0), RANK\_FINAL_{SSD}(x-1, y, p, 1) \ldots RANK\_FINAL_{SSD}(x-D_{max}/2, y, p, D_{max}/2))$$

Equation 3.21

$DISPARITY_{target}(x,y)$ refers to the disparity of pixel at position x,y of the target image.

The equation to assign intensity to target image is:

$$I_{target}(x, y)= I_{ref}(x-DISPARITY_{target}(x,y),y,p)$$ Equation 3.22

$I_{ref}(x,y,p)$ refers to the intensity of the pixel with position x,y in reference image p.

3.7 Algorithm criteria

Besides developing the algorithm framework, a significant amount of effort was committed to identifying the criterion that can fully explore the potential of the framework. The two criteria utilised in the algorithm are described in the following paragraphs. Although the developed algorithm employs both criteria, each individual criterion may be used independently to synthesise a target image.

3.7.1 Laminogram intensity

3.7.1.1 Developing criterion

As can be deduced from Figure 3.19 that an object (e.g, object A in disparity 11) appears to exhibit its lowest intensity when in focus. This observation illustrates that the lowest intensity is a potentially useful matching criterion for laminograms. For a reference pixel of interest positioned at coordinate position (x,y), the program compares the intensity of all laminograms at the same coordinate position and chooses the disparity with the lowest intensity as an in focus point. This novel criterion is termed the "laminogram intensity" (LamI). A laminogram is a result of superimposing multiple

perspective images. So, the quality of laminograms depends on the number of perspective images used. In general, the more images involved the higher quality the laminograms. The criterion employs the intensity of the laminogram as the matching cost value. In comparison with the sum of reciprocals method, the laminogram intensity is equivalent to the "Example cost" rows in Table 3.1 in Section 3.4.

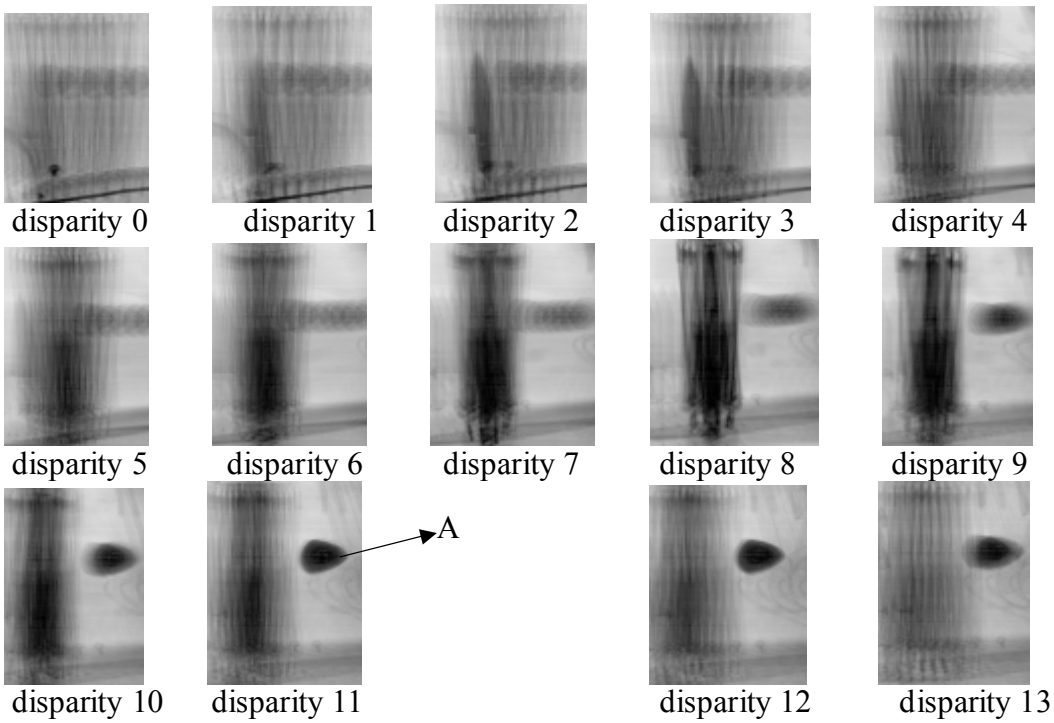


Figure 3.19 An example of laminograms.

In practice, when dealing with cluttered X-ray images, the laminogram intensity criterion will fail to perform correctly in certain circumstances. The principle of the criterion is based on choosing the disparity with the lowest intensity as the correspondence. When a bright area falls within the disparity window of a dark area, the relative lateral shifting of the perspective images to generate laminograms will cause the dark area, in visual terms to smear the relatively bright area. Figure 3.20 illustrates an example of this effect. If the magnitude of the smear (after averaging) is lower than the bright area when it is in focus, then the smearing will confound the algorithm and lead to an erroneous selection of disparity.

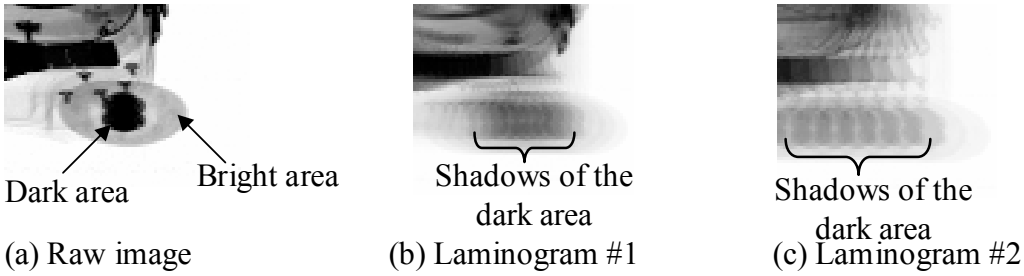


Figure 3.20 Illustration of the performance limitation of the basic laminogram intensity criterion.

The performance of this criterion with various combinations of input images and voting images is presented in Section 5.1.

3.7.1.2 Improved laminogram intensity criterion

It can be appreciated that the performance limitation of the criterion is a function of the degree of smearing. In concept, the removal of dark areas will enhance the bright areas and thus increases the matching likelihood of the bright area. To implement this concept, iterations are applied similar to the concept described in the Section 3.5.2, except that iterations are not only within the synthetic pixel generation process. The iteration includes matching cost computation, voting for disparity and synthetic pixel generation, because laminograms require to be regenerated. So for each iteration, matching cost computation, voting for disparity and synthetic pixel generation are repeated. In the interpolation process, for each iteration, pixels are assigned a disparity in the same manner as described in Section 3.5.2. Figure 3.21 shows the flowchart of Improved LamI criterion.

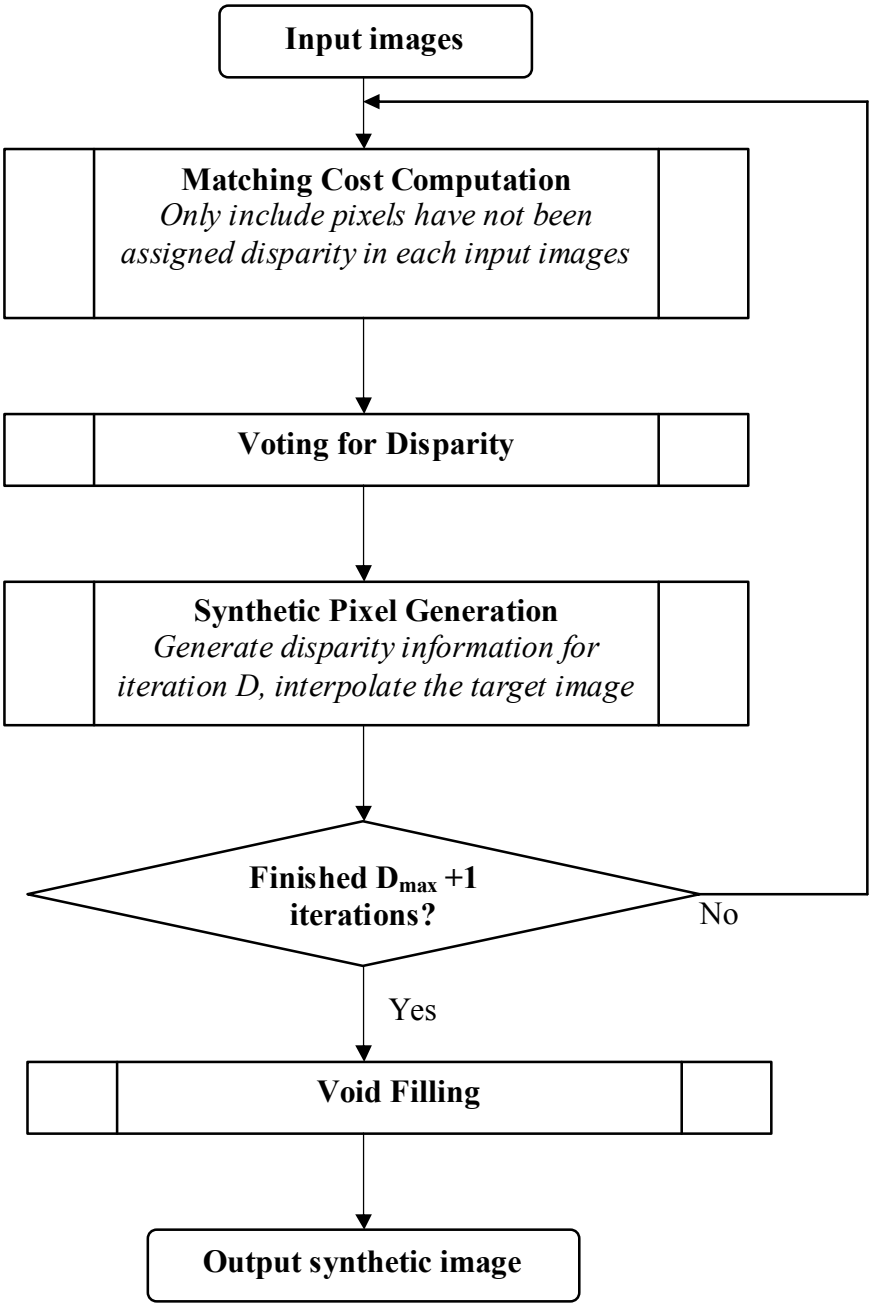


Figure 3.21 Flowchart of Improved LamI criterion.

The first iteration is considered to produce the most robust data. Using the Improved LamI criterion, in the first iteration, pixels with lower intensity are more likely to be matched. In the second iteration, pixels that have been assigned a disparity value will be excluded from the process of regenerating successive laminograms, although those pixels will still be included in the voting process. This approach implies that different pixels from the laminogram may be produced by a different number of input images. Effectively, the dark areas (i.e, pixels with lower intensity) are tending to be removed from the matching cost computation process. The pixel exclusion (i.e, dark area removal) procedure is iteratively implemented in which the iteration value is loosened for each

iteration to process unmatched pixels. The performance of the improved laminogram intensity criterion with various combinations of input images and voting images is presented in Section 5.2. A performance comparison of the improved laminogram criterion to the basic laminogram intensity criterion is presented in Section 5.3.

3.7.2 Sum of squared differences criterion

The SSD criterion is well known in stereo matching world for its utility. The performance of SSD using laminography and rank voting with various combinations of the number of input images and voting images is presented in Section 6.2. The performance of SSD using sum of reciprocals and rank voting with various combinations of input images and voting images is presented in Section 6.1. A performance comparison of the SSD using laminography and sum of reciprocals is presented in Section 6.3, results indicate that the SSD has a better performance when employing the sum of reciprocals matching cost calculation (also discussed in Section 6.3).

3.8 Summary

Each of the algorithmic processes employed in the developed V-DMX image synthesis algorithm is described in detail in this Chapter. The algorithm is developed to exploit the advantages afforded by a disparate sequence of X-ray detector images to synthesise intermediary images. The experiment plan presented in following Chapter 4 forms the basis for the empirical studies reported in Chapters 5, 6 & 7, which represents a critical part of the design and optimisation process for the V-DMX algorithm. The empirical work indicates that the optimum performance of V-DMX algorithm is achieved when three adjacent X-ray detector images are employed for image synthesis. This summary concludes with a mathematical overview of the algorithm presented in the following Figure 3.22.

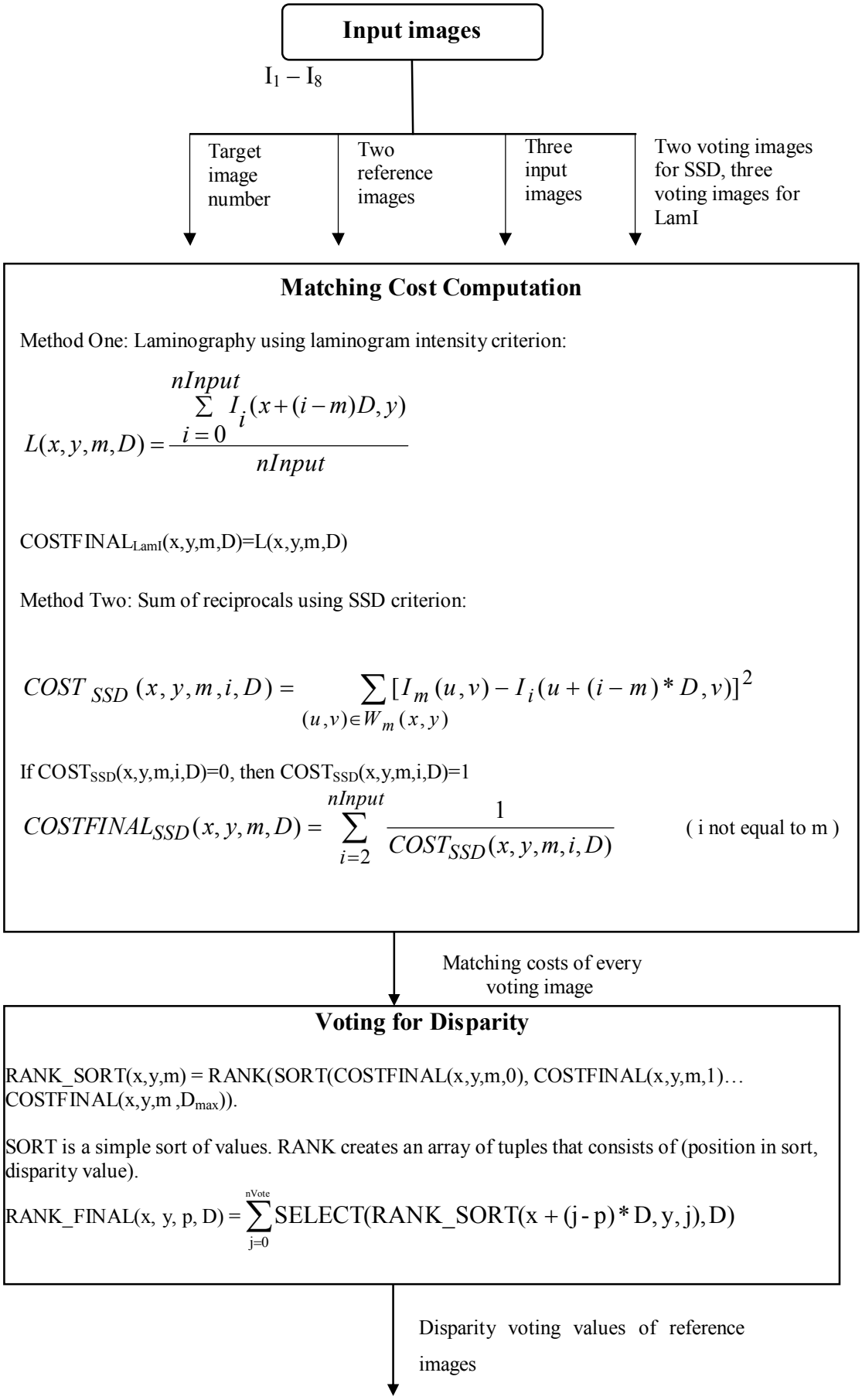


Figure 3.22a A mathematical overview of the V-DMX algorithm.

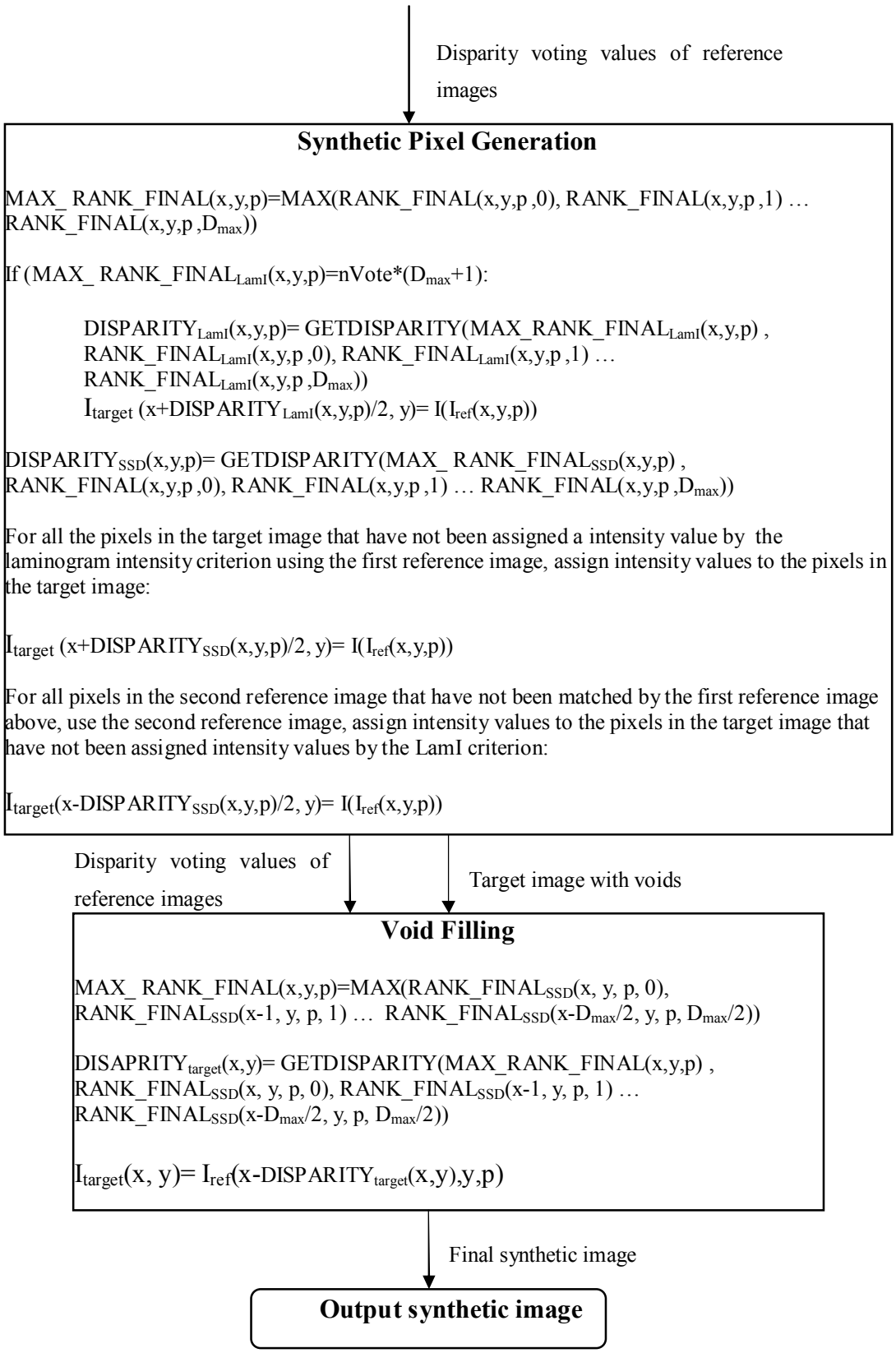


Figure 3.22b A mathematical overview of the V-DMX algorithm.

Chapter 4 Experiment considerations and overview

This chapter presents the methodology for a series of experiments and comparative studies presented in Chapters 5, 6 &7.

4.1 Methodology

To ensure realistic outcomes, real X-ray images were employed for the investigation. The images used for the indicative analysis presented in this thesis were chosen carefully to include objects composed of different material classes arranged in spatially complex scenes. The inclusion of dense amorphous structures together with overlapping high frequency details exhibits the multi-layered translucency typical of images routinely encountered at security checkpoints.

When real and synthetic images are viewed in succession, the resultant transition should, ideally, be natural to the observer. Therefore, the fidelity of the synthetic images is established by comparing them, pixel by pixel, with detector images produced at the appropriate X-ray beam angle by the scanner. This approach enables a direct measure on how well the synthetic images resemble the detector images. In the context of this research programme, these detector images that are used for comparison are defined as the ground truth (GT) images. It should be noted that GT images were solely used for comparative study but not as the input images of the synthesis algorithm. Both quantitative and qualitative measures were undertaken to aid making decision. An image subtraction method is used as the quantitative measure to compute the number of errors (NE) between the GT images and the synthetic images. The subtraction equation is presented in the next section. To account for various practical fluctuations (e.g, system noise), the NE computation includes a tolerance that has the value of the square root of the intensity. This rationale is based on the prior research work [119] on measuring the practical noise recorded by typical X-ray luggage scanner. The success of the synthetic image approach will ultimately be established by conducting a series of human factors investigations. While such work is ongoing in the university team, it is beyond the scope of this thesis. Nonetheless, besides comparing the NE produced, a visual inspection was undertaken to assist in concluding the results within the context of the KDEX system.

The proposed synthesis algorithm V-DMX is based on using two matching criteria namely SSD and LamI. While the algorithm presented in Chapter 3 is the ultimate delivery of this research programme, the growth of the V-DMX involved conducting an independent investigation of each individual criterion. This is because each criterion contributes to different but important aspects to address the problems defined in this research programme. The empirical results and analysis of the SSD and LamI investigations are organised in Chapter 5 and Chapter 6 respectively. These experiments are interrelated to form the basis for the investigation into the performance of the V-DMX, as presented in Chapter 7. Section 4.4 is prepared to explain the details of each experiment.

The two critical practical considerations for realising the KDEX implementation are the number of input images and voting images required for processing and the total achievable angular coverage. These considerations are important with regard to establishing the optimum X-ray sensor geometry, especially for systems using up to 32 images which are under consideration as future work. The ideal condition is to develop an algorithm that can cope with large angular separation between images to increase the total angular coverage with the use of minimal number of input images and voting images. However, the increase in angular separation causes a concomitant increase in the disparity window, which tends to introduce more matching ambiguity. Therefore, an investigation was instigated to assess the algorithm performance (or limitations) as a function of the X-ray beam angle between successive views.

4.2 Performance measure

Equation 4.1 is utilised as the quantitative factor to measure the performance of the synthetic algorithm under different experimental configuration.

$$\mathrm{NE} = \left\| \left\{ \mathrm{P}' \in \mathrm{M} \mid \mathrm{abs}\left(\mathrm{G}(\mathrm{P}') - \mathrm{S}(\mathrm{P}')\right) > \sqrt{\mathrm{G}(\mathrm{P}')} \right\} \right\|. \qquad \text{Equation 4.1}$$

Where M is dependant on the image size, $G(P')$ is the intensity of the GT image at P', and $S(P')$ is the intensity of the synthetic image at position P'. If the intensity difference between GT image and the synthesised target image is bigger than the square root of GT image's intensity, it is considered to be an error.

4.3 Experimental images and image selection

4.3.1 Experimental images

In each image sequence generated for the experiments, there are 32 real detector X-ray images with an angular separation between adjacent images of 1° covering an angular range of +16° to -15°. The experimental images were produced by the folded array multiple view X-ray scanner using line-scan imaging technique as described in Section 2.3.2. The resolution of the resultant line-scan images is 512 × 512. The selection of input images from the full sequence (32 detector images) depends on the angular separation involved, which is an important factor to establish the angular coverage of the practical system. Generally as the angular separation increases, the resultant images become more disparate necessitating an increase in the pixel disparity limit, which in turn increases the potential for spurious pixel matches. The result produced by the algorithm in the preliminary stage has indicated its superior capability by generating high quality synthetic images for angular separations smaller than 4°. This desirable result has therefore encouraged the author to focus on presenting investigation results of angular separations greater than 4° in this thesis. So, the four angular separations considered in this thesis are: 4°, 6°, 8° and 10°. In this thesis, results produced by the angular separation of 4° are utilised to illustrate the operational and performance of the developed algorithm. Figure 4.1 presents an example of eight sequential detector images with an angular separation between adjacent images of 4° from image set one.

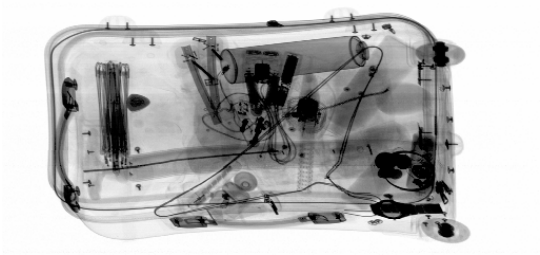


Image set one 14°

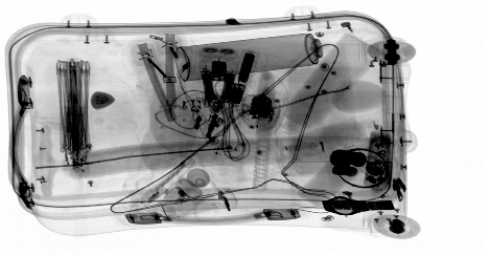


Image set one 10°

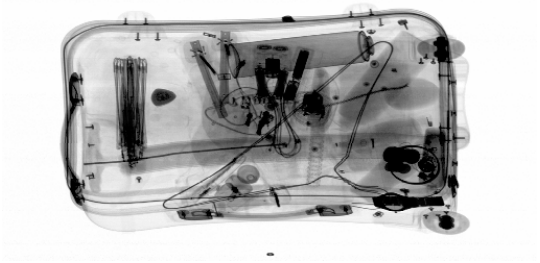


Image set one 6°

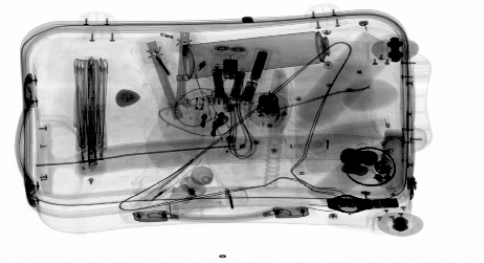


Image set one 2°

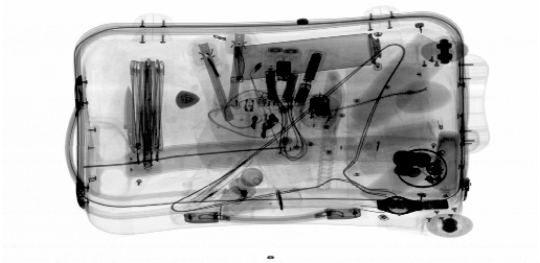


Image set one -2°

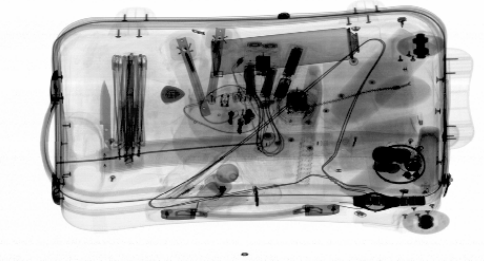


Image set one -6°

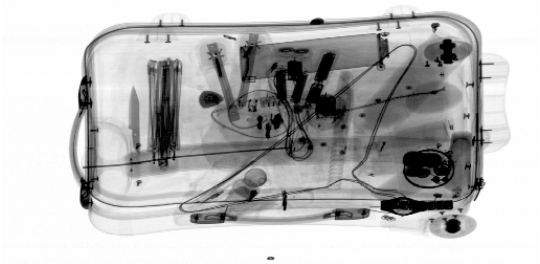


Image set one -10°

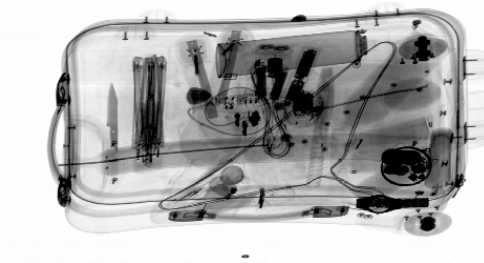


Image set one -14°

Figure 4.1 Examples of detector images for image set one where the angular separation between images is 4°.

All graphs and images in experiment chapter 5, 6 &7 are based on images from image set one unless specified.

4.3.2 Image selection

With an angular separation of 4°, eight out of the full sequence (32 detector images) are available to be selected as the input images. Thus, seven new intermediary images are

required generating, which gives a total of fifteen images in a complete sequence. As the number of detector images is eight, so the maximum number of input images is eight, the maximum number of voting images is eight.

$$nDetector \supseteq nInput \supseteq nVote$$

The total possible combinations of nInput and nVote for eight detector images are 35 as tabulated in Table 3.3.

When the angular separation increases, the nInput and thus nVote available for image synthesis are less in the sequence of 32 X-ray detector images. Therefore, the number of possible pairings is also reduced accordingly. Table 4.1 shows the nInput and the number of possible pairings for the angular separations considered in this work.

Angular Separation (°)	nInput	Number of Possible Pairings
4	8	35
6	6	20
8	4	9
10	4	9

Table 4.1 The nInput and the number of possible pairings for four angular separations under investigation.

The criterion to select reference images, input images and voting images from detector images is explained in Section 3.2.1.

4.4 Experimental plan

There are three experiment chapters in this thesis which are Chapter 5, 6 &7. The details of the experiments are illustrated by the diagram in Figure 4.2. To aid the description of the diagram, the (nInput, nVote) pairing for an angular separation of 4° is used as an example.

Chapter 5 reports the experiment results and analysis of the novel LamI criterion for different experiment conditions. It contains four individual experiment sections, which are designed to establish the performance of the LamI criterion. The first two experiments govern the independent investigations into two versions of the criterion, namely the Basic LamI and the Improved LamI. The major departure of the two versions is that the Improved LamI uses an iteration process to compute the matching cost while the Basic LamI involves no iteration in the cost computation. The iteration

process is to enable the removal of dark image areas and thus improving the matching likelihood of the bright image areas. The details of each criterion are described in Section 3.7. Since both criterion versions are fundamentally the same, the matching cost calculation method is also the same, as governed by Equation 3.1. The preliminary investigation into the two criterion versions was to determine the optimum pairing of each criterion. Then, a comparative study between the two criterion versions using best pairing was conducted to demonstrate the utility afforded by the Improved LamI. Finally, the practical significance of the Improved LamI is studied as a function of increasing angular separation.

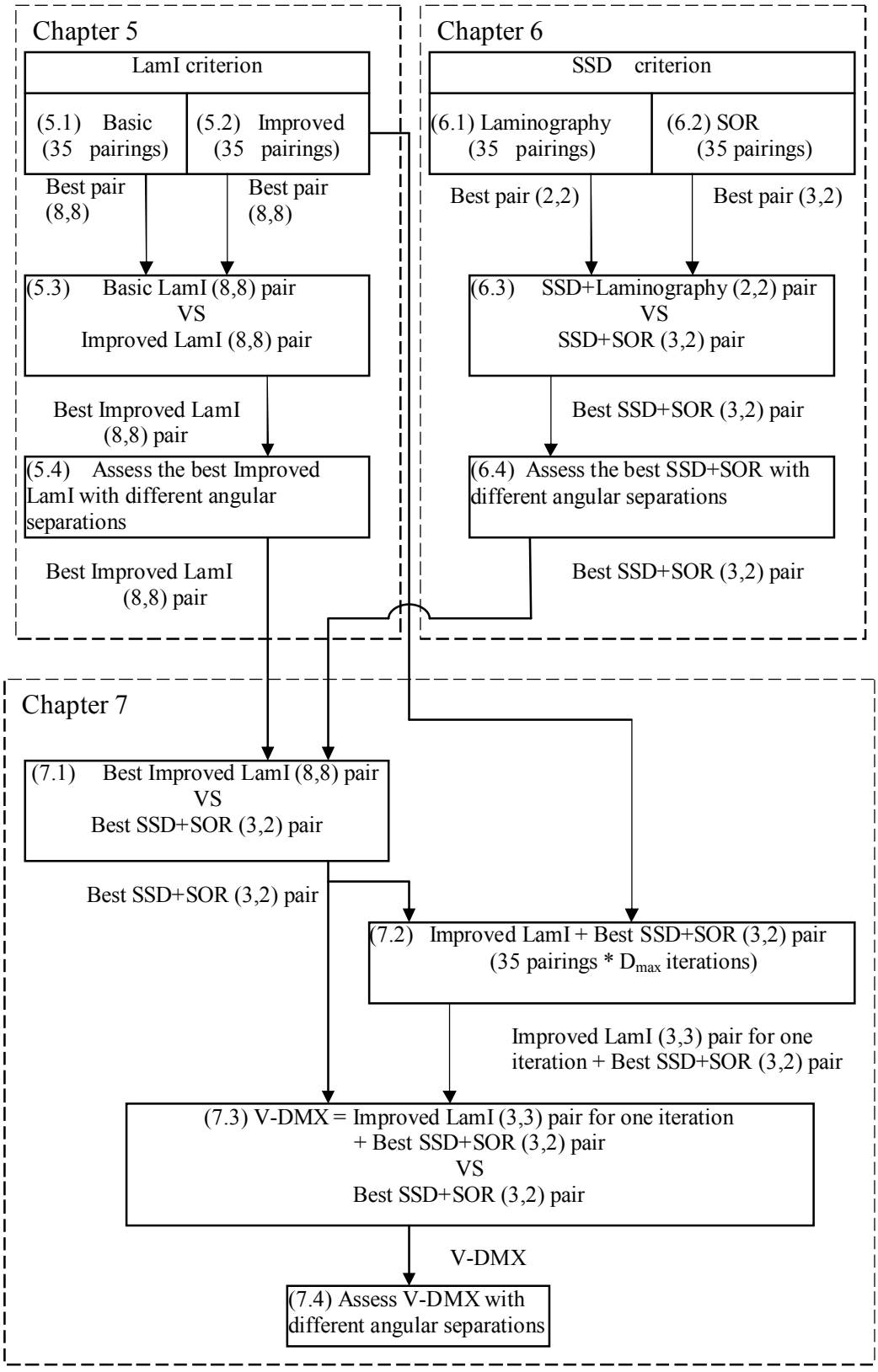


Figure 4.2 Experimental plan and results overview. The (nInput, nVote) pairings for an angular separation of 4° is taken as an example for illustration.

The LamI criterion is a novel measure which is bespoke to the laminography matching cost calculation method. However, the SSD criterion can be applied to both cost calculation methods. When using SOR method, the SSD matching cost is computed

between the reference images and the input images, in which all of them are real X-ray images. However, the incorporation of SSD into the laminography cost calculation method is uncommon because the matching cost is computed between the reference images (real X-ray images) and the laminogram. It was considered that the intensity fluctuation in the neighbouring pixels of a "focused" feature in laminograms is normally small due to the averaging effect produced during the laminogram generation. This effect tends to stabilise the focused feature and may increase its matching likelihood. The matching cost calculation equation for using SSD with the laminography method is:

$$\text{COSTFINAL}(x,y,m,D) = \sum_{(u,v)\in W_m(x,y)} [I_m(u,v) - L(u,v,m,D)]^2 \ . \qquad \text{Equation 4.2}$$

It should be noted that this equation is different from the cost calculation of using SSD with SOR method, as presented in Section 3.3.2. After determining the optimum combinations of nInput and nVote for the SOR and laminography based methods, both optimum results are compared to find out the best method when using SSD criterion. Empirical results demonstrably revealed the SOR method is superior over the laminography method. An investigation into the angular performance was conducted on the SOR method.

Chapter 7 reports the experiment results and analysis to support the development of V-DMX. Section 7.1 compares the best pairing of the Improved LamI with the best pairing of the SSD based method. The comparative result indicates that SSD based method has a better overall performance except for overlapping and discontinuity image regions. In contrast, the Improved LamI shows a relative superior in these regions. So, efforts were devoted to determine the best way to combine both different but balancing advantages afforded by the criteria. Research findings are reported in Section 7.2. Section 7.3 compares the performance of V-DMX with the best performance of SSD, to demonstrate the advantage of V-DM algorithm. Section 7.4 presents the experiment to assess the performance of V-DMX with different angular separations.

Chapter Five Experiments and analysis: Laminogram intensity (LamI) criterion

The experiment results and analysis concerning the LamI criterion are presented in this chapter. These experiments are designed to be part of a larger group of interrelated experiments as discussed in Chapter 4. The experiment plan, Figure 4.2, illustrates the relationship of the experiments to the final algorithm structure. The objective of the experiments is to initially determine the optimum number of nInput and nVote required when employing the LamI criterion with the rank based voting method to generate intermediary images. Two different versions of the criterion, namely the Basic LamI and the Improved LamI, are reported and discussed. The differences between the criteria are explained in Chapter 3. An extended study of the Improved LamI has been undertaken to establish the optimum number of nInput and nVote as a function of the increasing angular separation between the input X-ray images.

5.1 Experiment employing the Basic LamI criterion

Figure 5.1 and 5.2 are arranged to enable the convenient comparison of a family of NE (number of errors) curves for different n-m pairings (n input images and m voting images) recorded for target images at 12^0, 8^0, 4^0, 0^0, -4^0,-8^0, -12^0 angles. The NE curves are produced by applying the Basic LamI on the integration of the laminography matching cost calculation and the rank voting approach. It should be noted that the discrete data is represented in "line graph mode" to aid the visual identification of relative trends in the data. The symbol nVote1 indicates one voting image is employed. Also, the GT image of target angle 12° and the resultant synthetic images produced by 2-1, 8-1 and 8-8 pairings are presented in Figure 5.3 for comparative discussion.

5.1.1 Result: NE curves

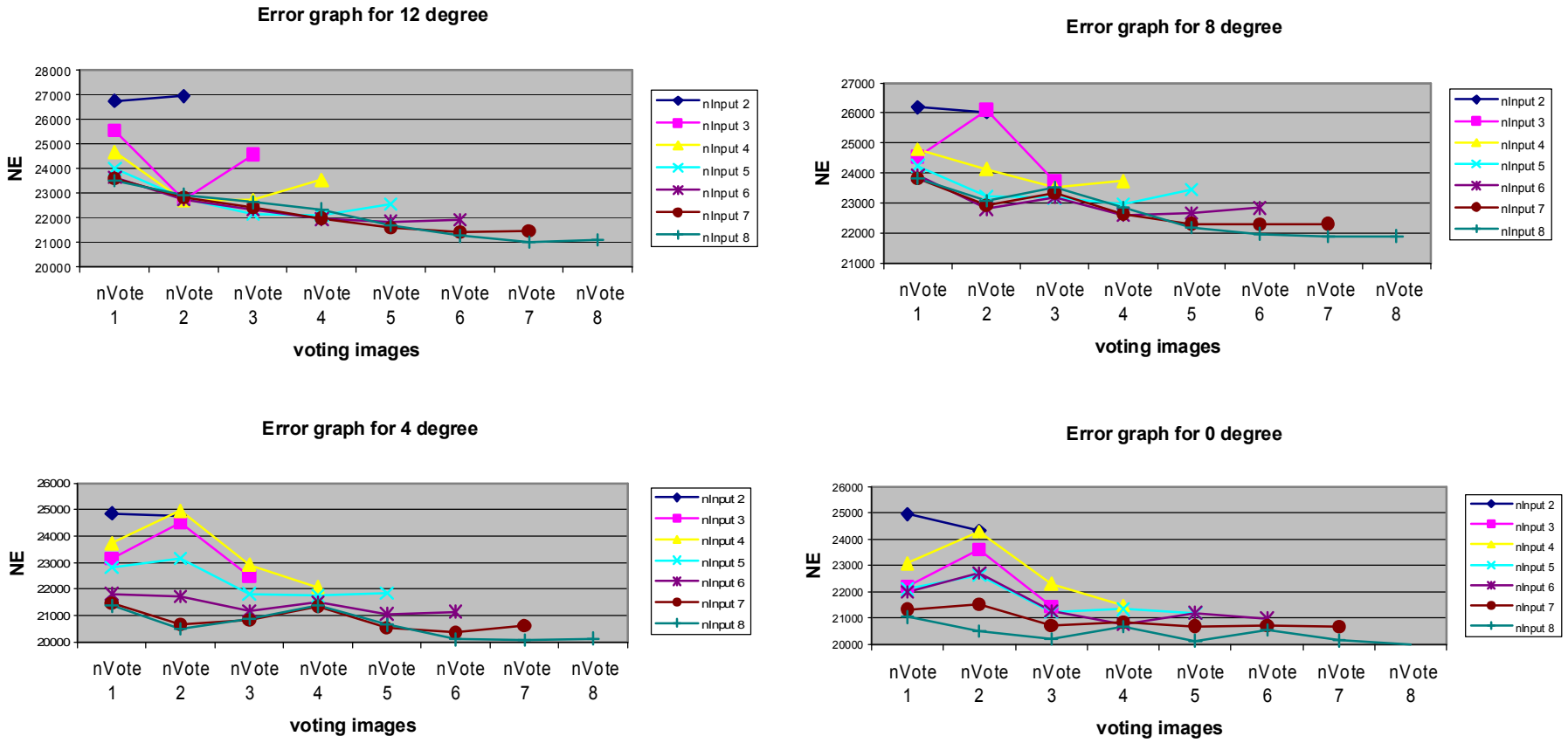


Figure 5.1 NE plots for different n-m pairings produced by using the Basic LamI on laminography matching cost calculation and rank voting approach, to synthesis target images at 12°(left top), 8°(right top), 4°(left bottom) and 0°(right bottom), where the angular separation between adjacent detector images is 4^{o}.

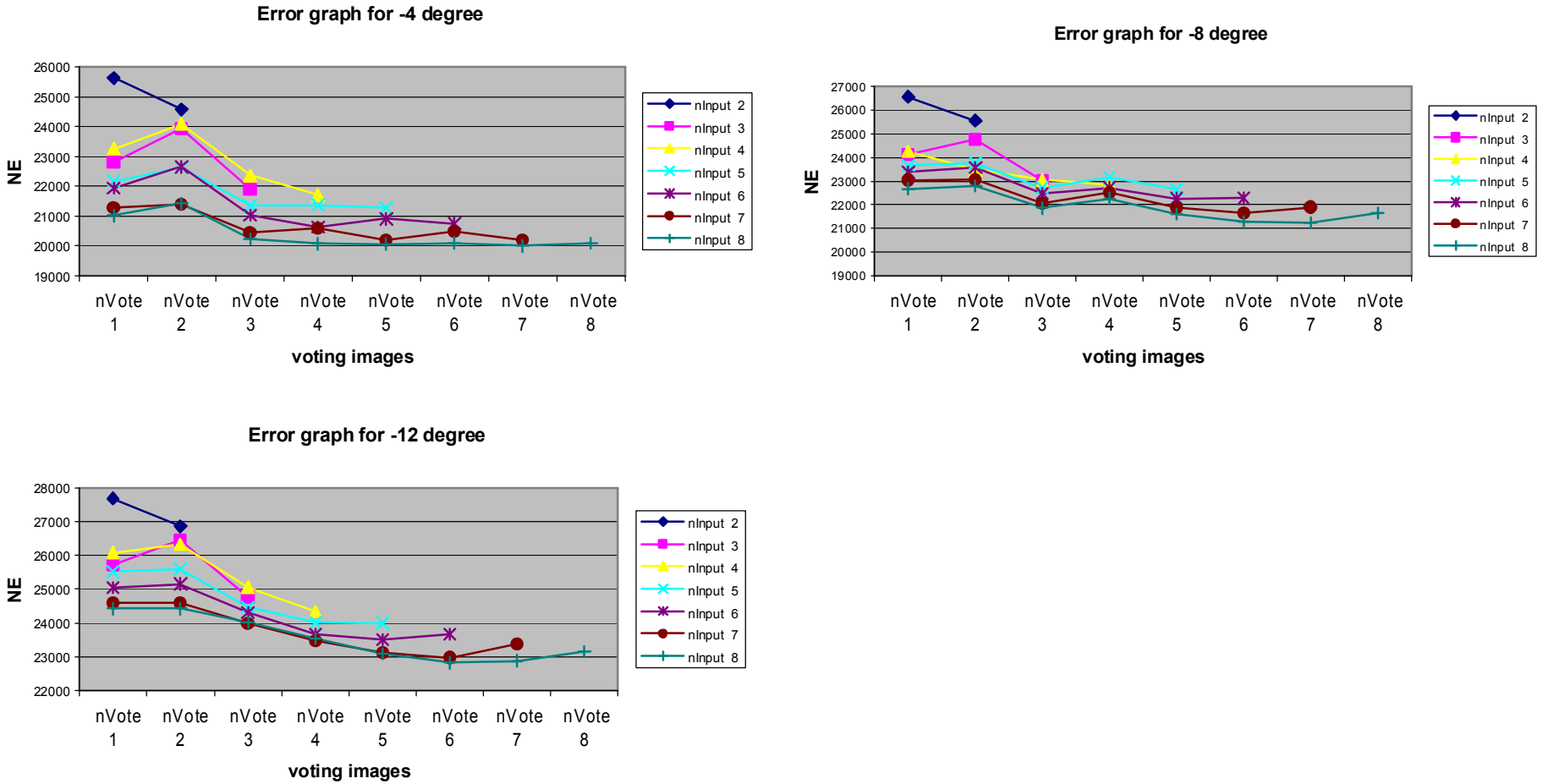


Figure 5.2 NE plots for different n-m pairings produced by using the Basic LamI on laminography matching cost calculation and rank voting approach, to synthesis target images at -4°(left top), -8°(right top) and -12°(left bottom), where the angular separation between adjacent detector images is 4°.

5.1.2 Result: Images

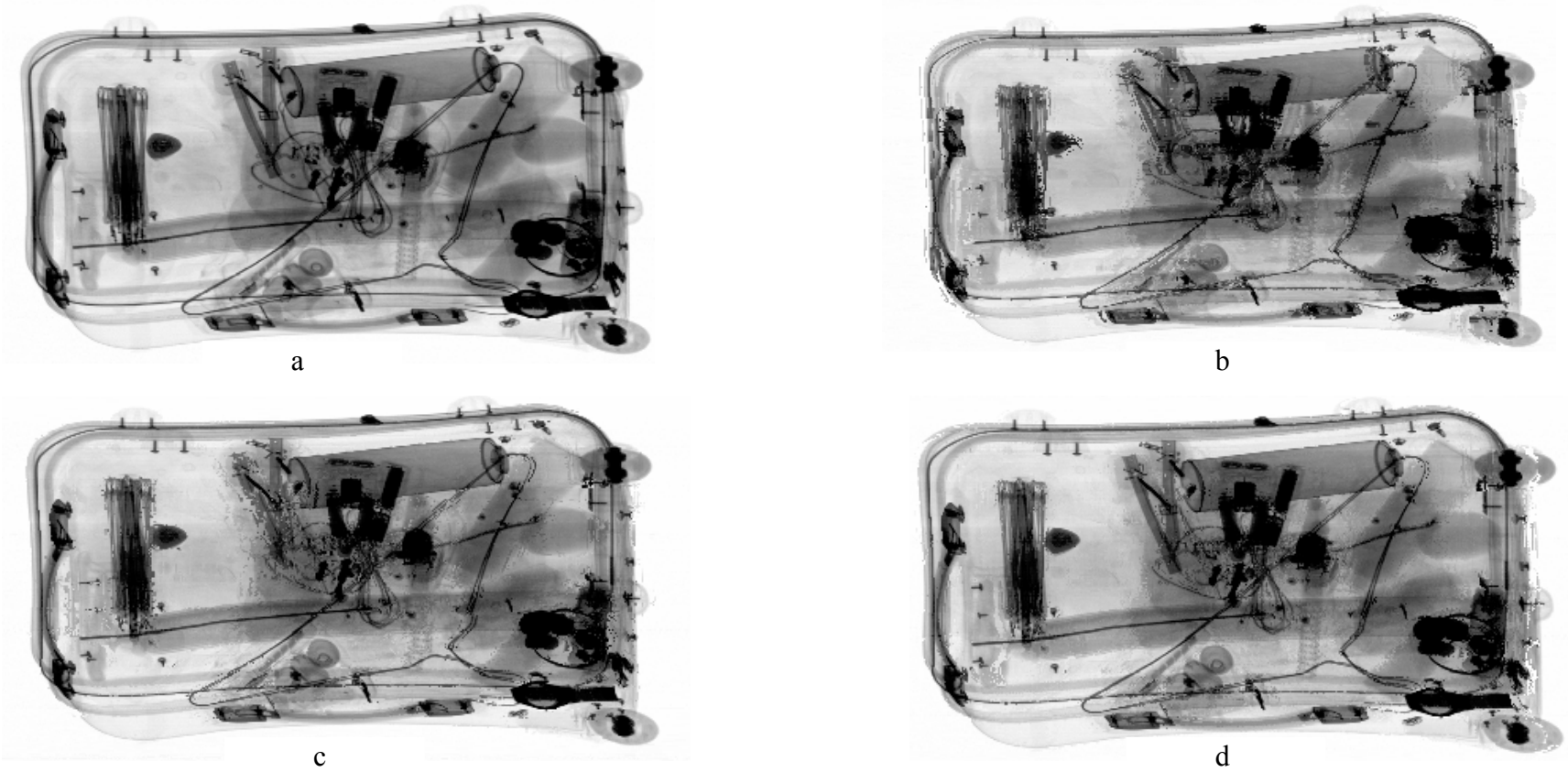


Figure 5.3 (a) The ground truth (GT) image at 12° and the resultant synthetic images produced by (b) 2-1, (c) 8-1 and (d) 8-8 pairings of nInput and nVote.

5.1.3 Analysis

As nVote and nInput increases there is a significant reduction in the NE. While there are a few outliers, the decreasing error trend can be easily observed. As can be deduced that the 8-8 pairing (i.e, involving 8 input images and 8 voting images) exhibited the best performance for all target images as can be appreciated by the visual inspection of the example of the resultant synthetic images arranged in Figure 5.3. Figure 5.3 (b) is the result of 2-1 pairing, which is equivalent to the traditional stereo matching method. As nInput increases for a constant nVote, the quality of the synthetic image is improved, see Figure 5.3 (c). A further improvement is evident when nVote increases, see Figure 5.3 (d). It can be concluded from these observations that the Basic LamI criterion performance increases as the nInput and nVote increased. To illustrate the impact of the increasing nInput and nVote, following four case studies are organised.

Case study 1: Positive impact of increasing nInput with fixed nVote

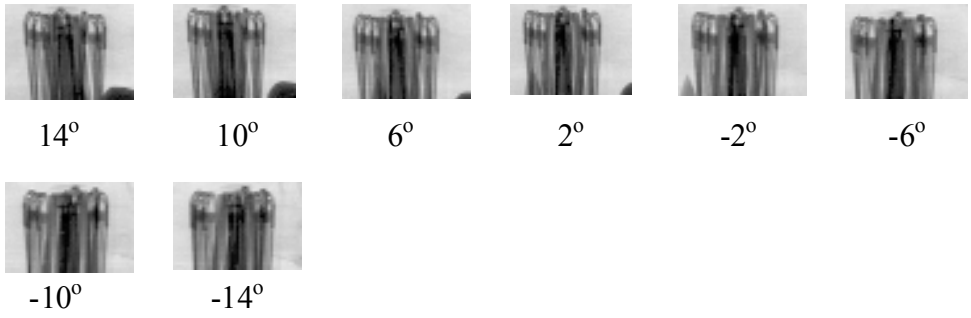


Figure 5.4 Regions of interest in input images over the range of 14° to -14°.

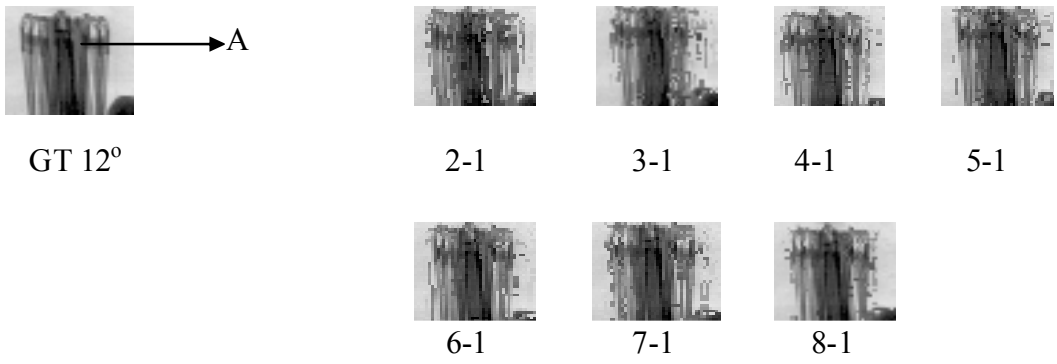
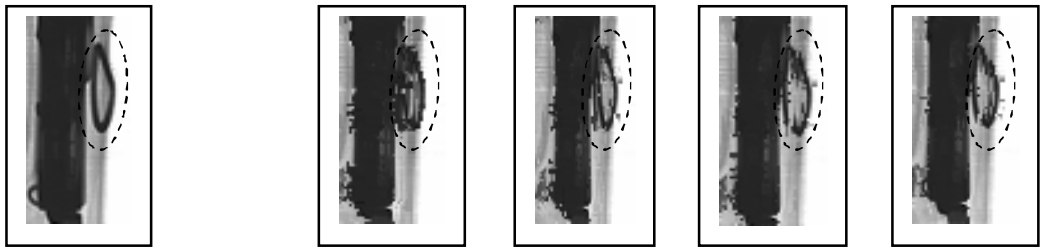


Figure 5.5 Illustration of the advantage (i.e. increase in quality) of increasing nInput by comparing the region of interest in the ground truth (GT) image and in the synthetic images produced by different n-m pairings.

To enable an effective discussion, the regions of interest (ROI) in all the input images are arranged in Figure 5.4. Also, the correlated ROI in the GT image at 12° as well as in

the resultant synthetic images for different n-m pairings are organised in Figure 5.5. It can be seen from Figure 5.4 that the feature-A (section of the umbrella with higher intensity) is not overlapped by other object in all input images. However, there is a big intensity fluctuation in the immediate vicinity of feature-A. As described in Section 3.7.1.1, the result of pixels with high intensity value is likely to be affected by neighbouring pixels with low intensity value. This is illustrated by the synthetic images that are arranged on the right hand side of the Figure 5.5. When only two input images are utilised, the synthetic quality is low. When more input images are included, the apparent defects are reduced thus the synthetic quality for feature-A is improved. In addition, it was observed that increasing the number of input images shows a great potential for enhancing the reproduction of repeating features. Figure 5.6 illustrates an example of this observation.



GT 12° 2-1 3-1 4-1 5-1

Figure 5.6 Improvement in repeating feature matching (object edges in this as a result of increasing nInput.(image set 2).

Case study 2: Negative impact of increasing nInput with fixed nVote

The ROI in all the input images are arranged in Figure 5.7, and the ROI in the GT image at 12° as well as in the resultant synthetic images for different n-m pairings are organised in Figure 5.8. When nInput increases from 2 to 5, the quality of feature-A (portion of a curved object feature) improved because there are no dark pixels in the immediate vicinity of the feature-A in the input images 14^o, 10^o, 6^o, 2^o, -2^o and -6^o. However, a new darker object (highlighted with black circle in Figure 5.7) appears in image -14^o and -10^o, has caused the Basic LamI criterion to perform incorrectly on feature-A (see 7-1 and 8-1 images).

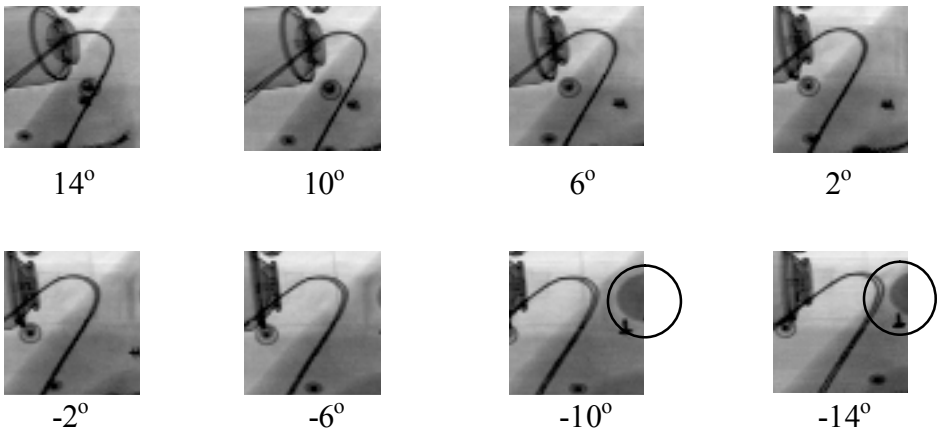


Figure 5.7 Regions of interest in input images over the range of 14° to -14°.

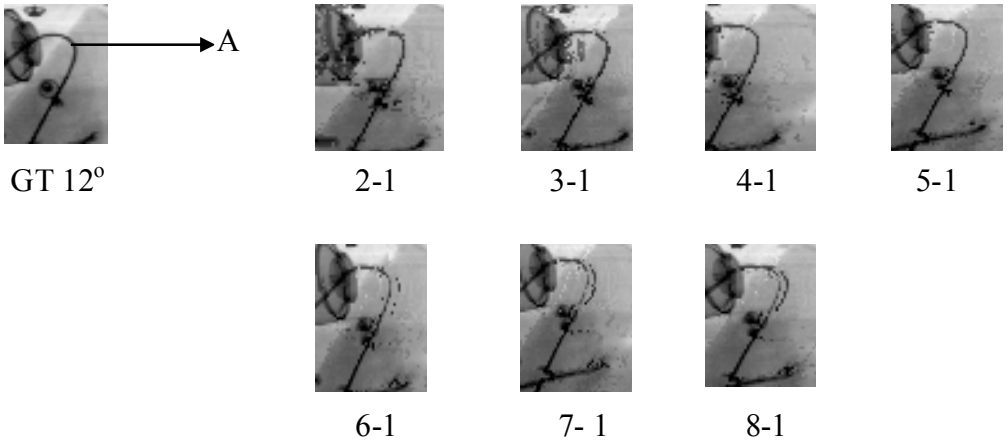


Figure 5.8 Illustration of the disadvantage (i.e. decrease in quality) of increasing nInput by comparing the region of interest in the ground truth (GT) image and in the synthetic images produced by different n-m pairings.

Case study 3: Positive impact of increasing nVote with fixed nInput

The issue described in the previous case can be alleviated by increasing the nVote from 1 to 8, as illustrated by Figure 5.9. It can be appreciated from the figure that the incorrect duplication of the curved object is partly rectified when *nVote* is increased from 7 to 8.

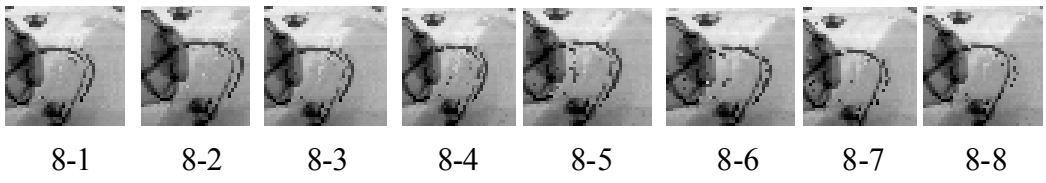


Figure 5.9 Improvement in object edges as a result of increasing nVote for 8-1 (left most image), 8-2, 8-3, 8-4, 8-5, 8-6, 8-7 and 8-8 (right most image) pairing.

When the voting image is either 14^o or 10^o or 6^o or 2^o or -2^o or -6^o input image, the search for feature-A to the right hand direction will be affected by the new darker object that appeared in -10^o and -14^o images, leading to incorrect matching. However, when the voting image is -10^o, the search is only affected by input image at -14^o, so it can process feature-A better than with the other six voting images. The quality of the feature-A is further improved when the voting image is -14^o because the search is limited to the left hand direction with the darker object on the right side of the feature-A.

Case study 4: Negative impact of increasing nVote with fixed nInput

The synthetic result of feature-A (section of the object with higher intensity) will be very good when only 14^o input image in Figure 5.10 is used as the voting image. This is because the search for correspondence is only performs to the right hand direction and there is no feature with intensity lower than the feature-A. On the other hand, when other input images are used as the voting image, the search to the left hand direction will encounter the dark area (i.e, feature-B) will cause some pixels of feature-A to be matched wrongly. In this case, adding more voting images would produce an inferior image quality, as shown in the 8-8 pairing result in Figure 5.11.

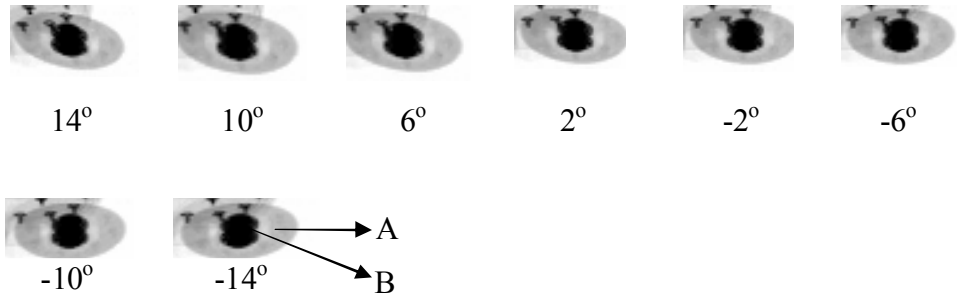


Figure 5.10 Regions of interest in input images over the range of 14^o to -14^o.

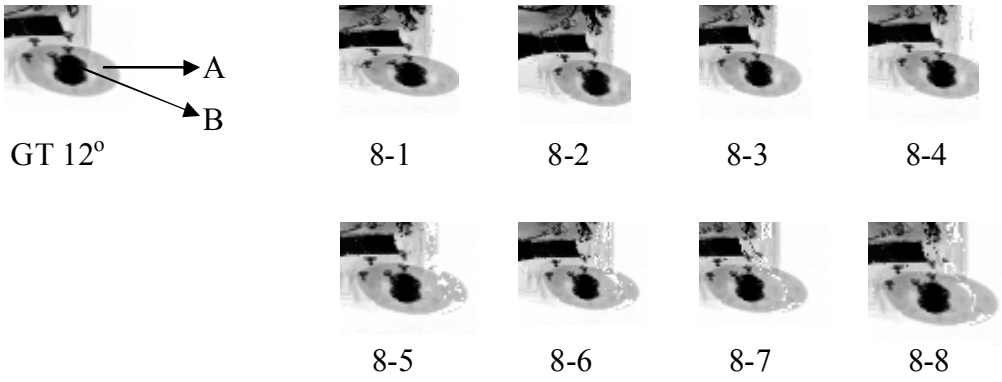


Figure 5.11 Illustration of the disadvantage (i.e. decrease in quality) of increasing nInput by comparing the region of interest in the ground truth (GT) image and in the synthetic images produced by different n-m pairings.

In summary, increasing either nInput or nVote can produce a positive and or negative impact on the resultant synthetic images. When only two input images are available, the quality of the generated target images is low because LamI does not handle pixels with a high intensity very effectively when they are in the vicinity of low intensity pixels. In practice, this is a very common scenario in luggage images; hence the performance when employing two input images is limited. Adding more input images can ease the problem, although there is a chance that adding more input images may generate undesirable artefacts. In practice, it is impossible to know in advance which image should be employed as the voting image(s) to achieve the best result. However, the hypothesis, as supported by the empirical results, of exploiting more images has on balance more advantages than disadvantages.

5.2. Experiment employing the Improved LamI criterion

The objective of this experiment is to determine the optimum number of input images (nInput) and number of voting images (nVote) when using the Improved LamI criterion with the rank based voting method to generate intermediary images.

Despite the success achieved by the 8-8 rank based method, the image smearing produced by a darker object when in the vicinity of a lighter object remains a problematic aspect of the laminographic techniques employed in this work. This section presents the experiments results produced by the improved version of laminogram intensity criterion, which dynamically selects nInput employed for each iteration. Figure 5.12 and 5.13 are arranged to enable the convenient comparison of a family of NE curves for different n-m pairings recorded for target images at -12^0, -8^0, -4^0, 0^0, 4^0,

8^0, 12^0 angles. The NE curves are produced by using the laminography matching cost calculation together with the rank based disparity voting approach. A linear trend is superimposed on the discrete empirical data to aid the visual comparison of relative trends in the data. The symbol nVote1 indicates one voting image is employed. For comparative discussion, the GT image of target angle 12° and the resultant synthetic images produced by 2-1, 8-1 and 8-8 pairings are presented in Figure 5.14.

5.2.1 Result: NE curves

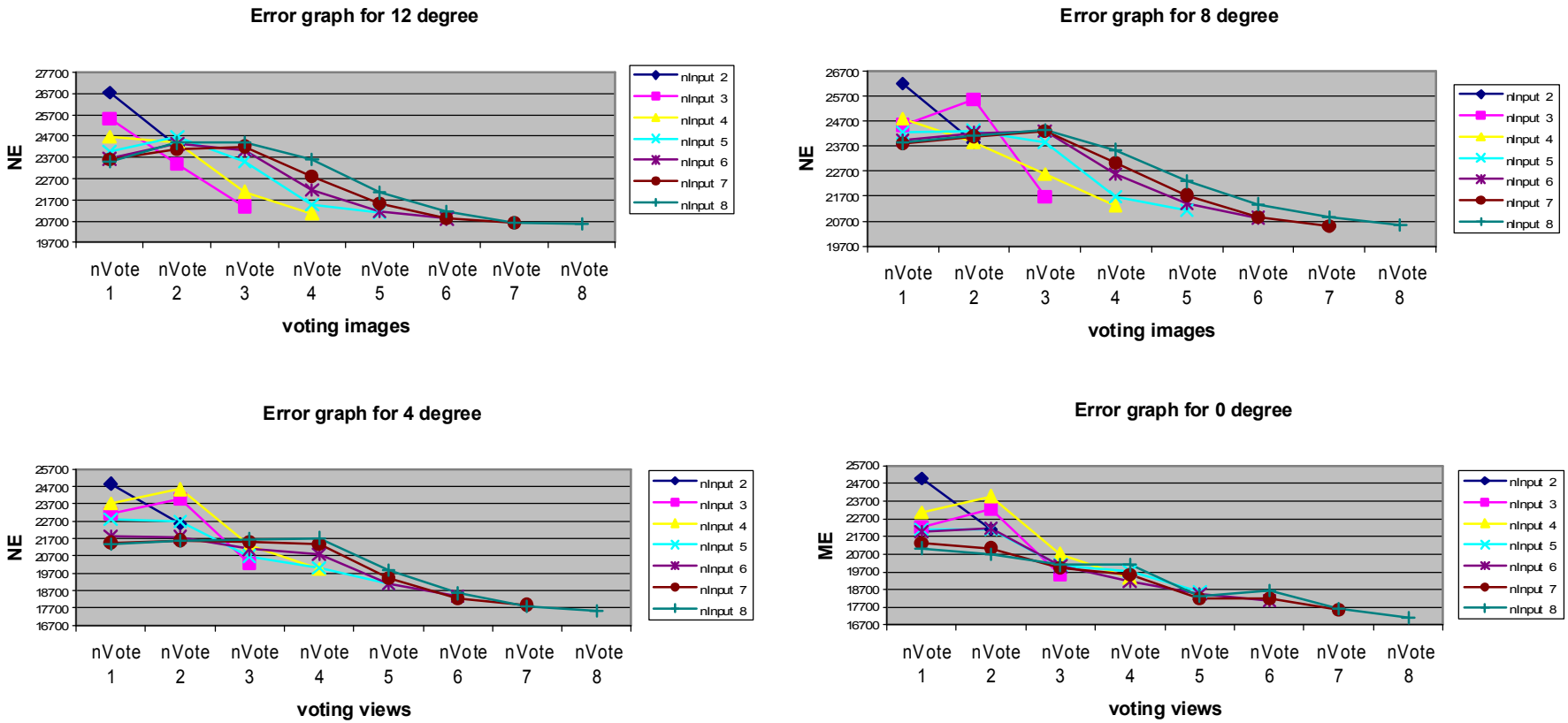


Figure 5.12 NE plots for different n-m pairings produced by using the Improved LamI on laminography matching cost calculation and rank voting approach to synthesis target images at 12° (left top), 8°(right top), 4°(left bottom) and 0° (right bottom), where the angular separation between adjacent detector images is 4°.

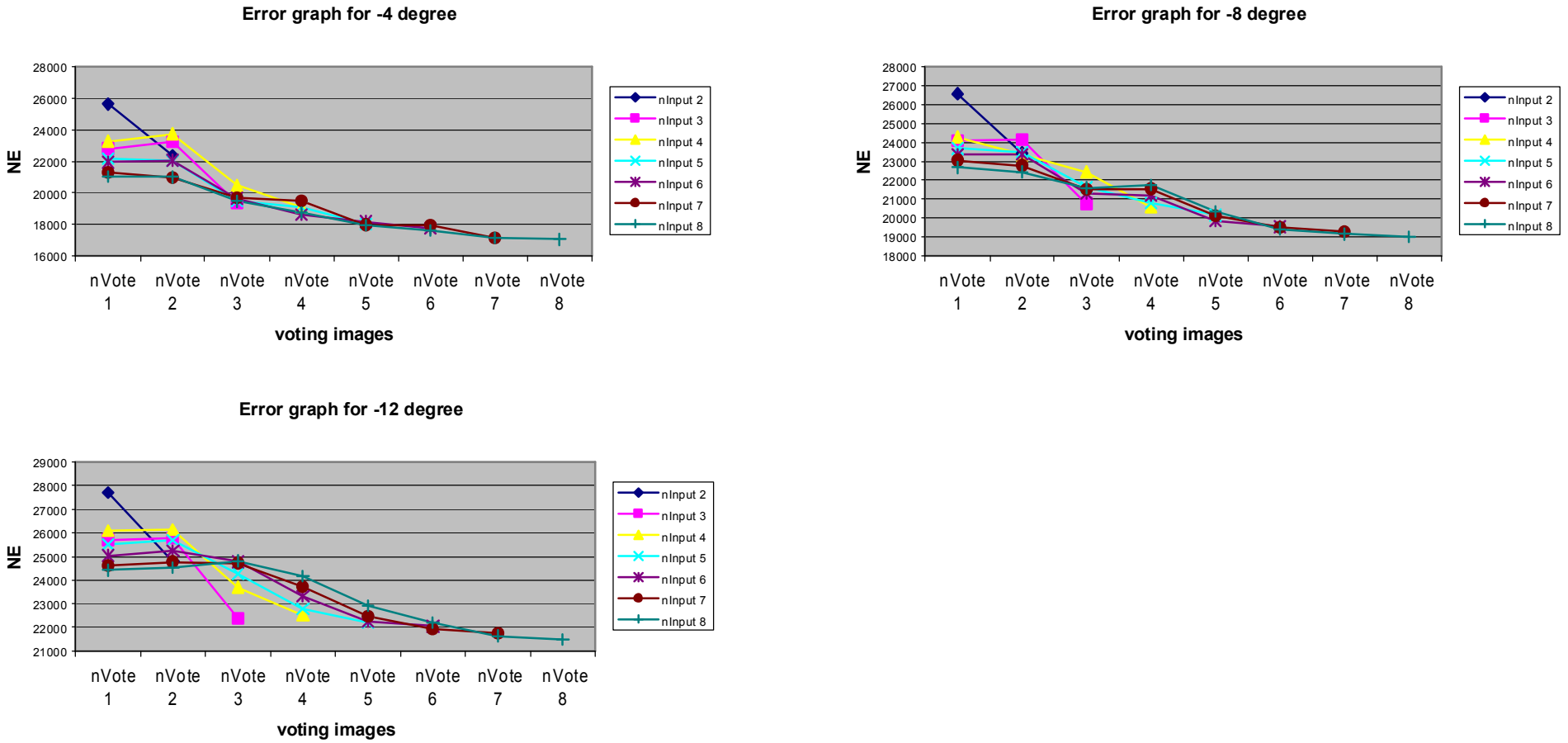
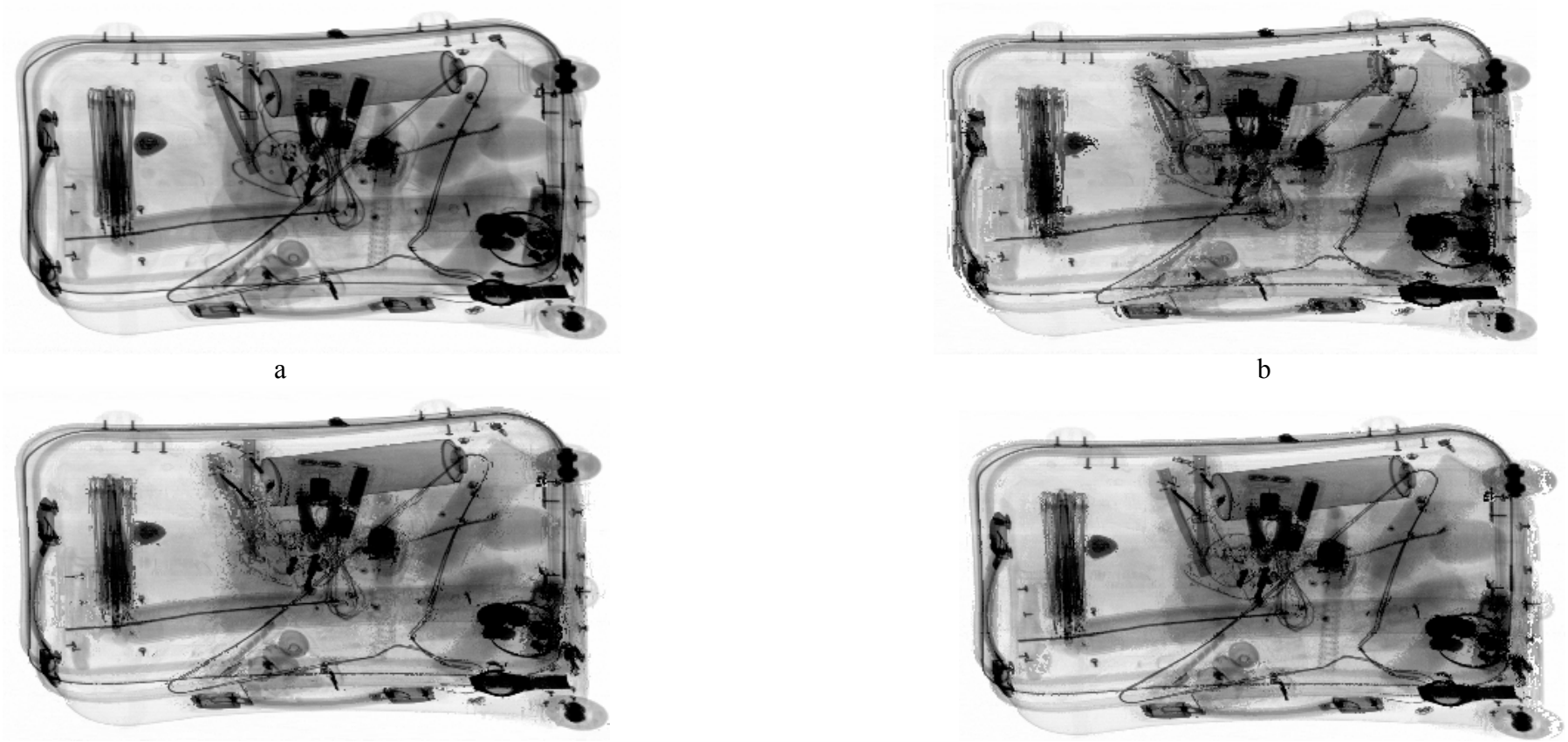


Figure 5.13 NE plots for different n-m pairings produced by using the Improved LamI on laminography matching cost calculation and rank voting approach to synthesis target images at -4° (left top), -8°(right top) and -12°(left bottom), where the angular separation between adjacent detector images is 4^o

5.2.2 Result: Images



a b c d

Figure 5.14 (a) The ground truth (GT) image at 12° and the resultant synthetic images produced by (b) 2-1, (c) 8-1 and (d) 8-8 pairings of nInput and nVote.

5.2.3 Analysis

As evidenced by Figures 5.12, 5.13 and 5.14 the improved LamI inherits the same trend from the Basic LamI method reported in Section 5.1. The increased number of nInput and nVote, produces improved results.

5.3 Comparative study of Basic LamI and Improved LamI

The objective of this experiment is to compare the performance of the Basic LamI with respect to the Improved LamI.

Using the rank based voting, both LamI criteria achieved the best performance when 8-8 pairing is employed. Figure 5.15 depicts the NE plots of the two LamI criteria produced by the 8-8 pairing for target angles 12°, 8°, 4°, 0, 4°, -8° and -12°. As can be deduced from the highlighted image regions in Figure 5.16 the Improved LamI exhibits a general improvement in the image quality.

5.3.1 Result: NE curves

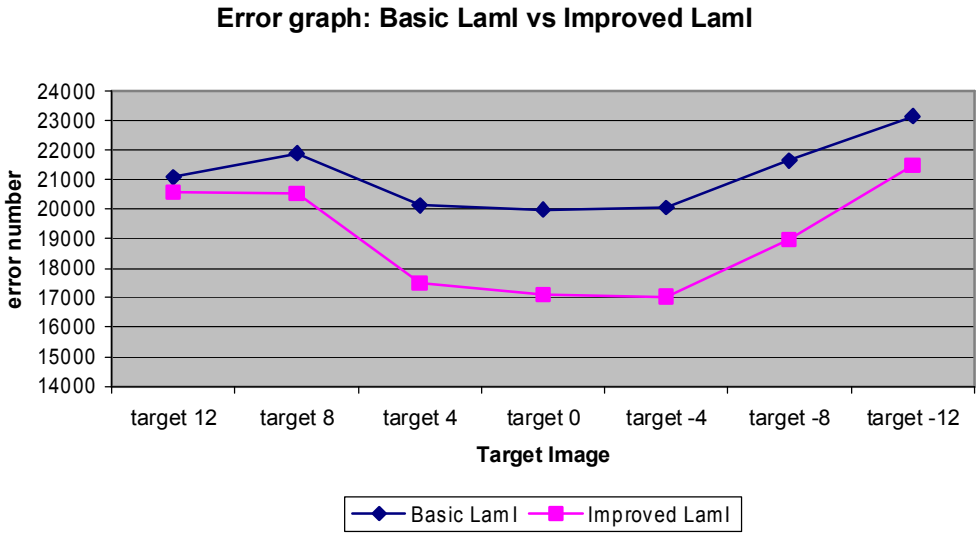


Figure 5.15 Comparison of NE plots of rank based voting methods produced by the 8-8 pairing for target angle 12°, 8°, 4°, 0, -4°, -8° and -12° using the Basic LamI and the Improved LamI criteria.

5.3.2 Result: Images

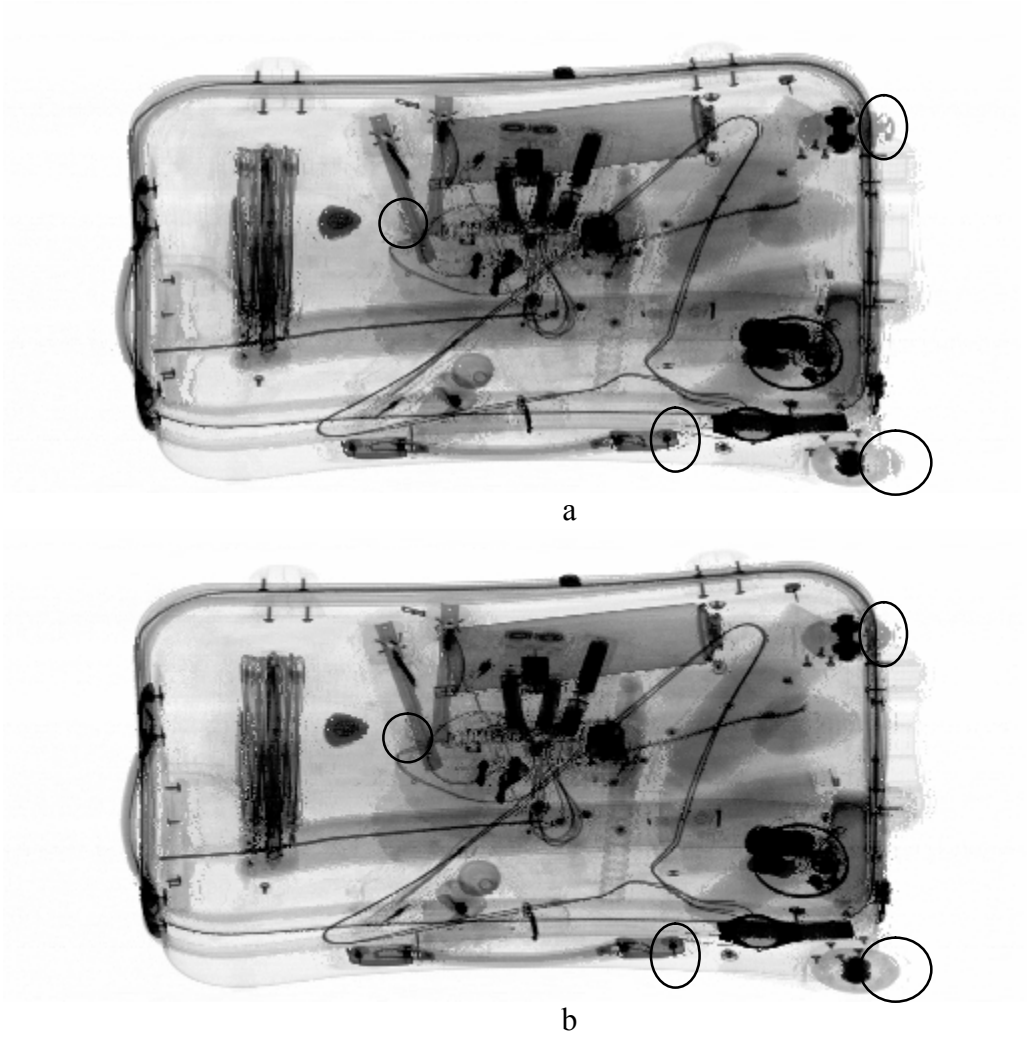


Figure 5.16 Synthetic images produced using detector images separated by 4° for target angle 0° for produced by 8-8 pairing of the Basic LamI (a) and the Improved LamI (b) criteria.

5.3.3 Analysis

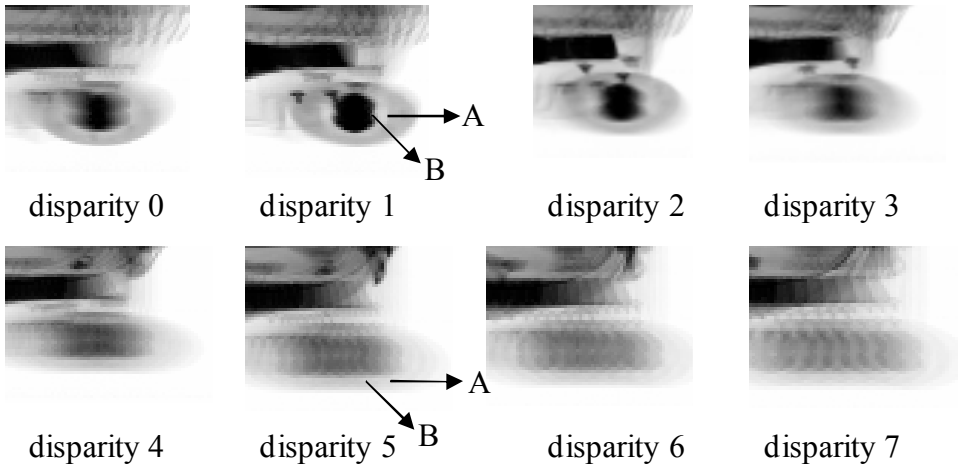


Figure 5.17 Laminograms of disparity 0 to disparity 7 used for synthesising target image at angle 0°.

Figure 5.16 shows the target image 0^o generated by Basic LamI and Inproved LamI. Figure 5.17 depicts the scene where the Basic LamI is prone to error because the adjacent feature (i.e, feature B) has a lower intensity. Consider feature-A has a disparity value of 1, so it focused in the laminogram with unity disparity value. However, due to smearing artefacts introduced by feature-B, the Basic LamI criterion will not choose disparity 1 but disparity 5 instead because the intensity of the pixel in disparity 5 is lower. Using the Improved LamI version, the problem may be circumvented because most of the feature-B would have been matched and removed before handling feature-A. Figure 5.18 presents the synthetic result for target image 0° produced by different iterations using the Improved LamI criterion. The pixels that have been matched in the first iteration (see Ite #1 image in Figure 5.18) are removed from all the input images. Then, a set of new laminograms is produced as arranged in Figure 5.19. It can be appreciated that after removing the pixels that have been matched in the first iteration, the degree of smearing over feature-A is reduced for all subsequent laminograms, which mitigates the mismatching encountered by the Basic LamI method. As the iteration continues, more pixels from feature-B will be matched and removed; eventually all of the pixels will be removed. The net result would be, ideally, "zero" smearing of feature-A and therefore an increased likelihood of a correct match. The performance of LamI is illustrated in Figure 5.20.

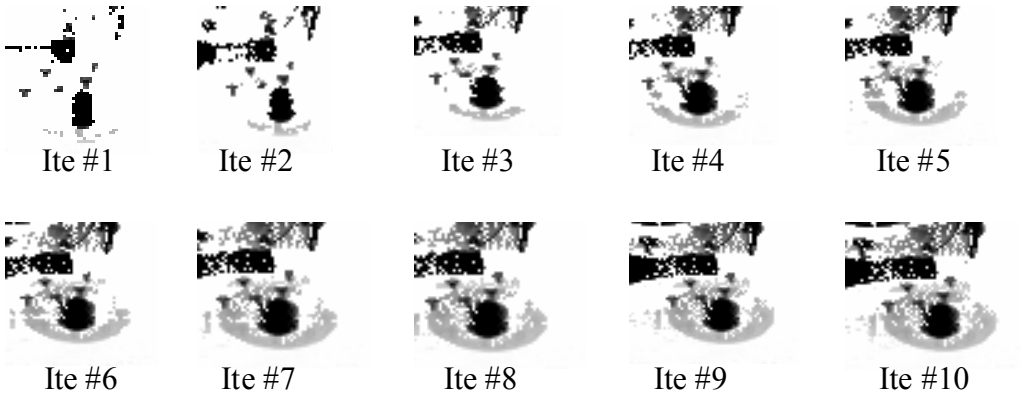


Figure 5.18 Synthesised images of target image 0° produced by different iterations (Ite) in ascending order where the top left image represents the first iteration result.

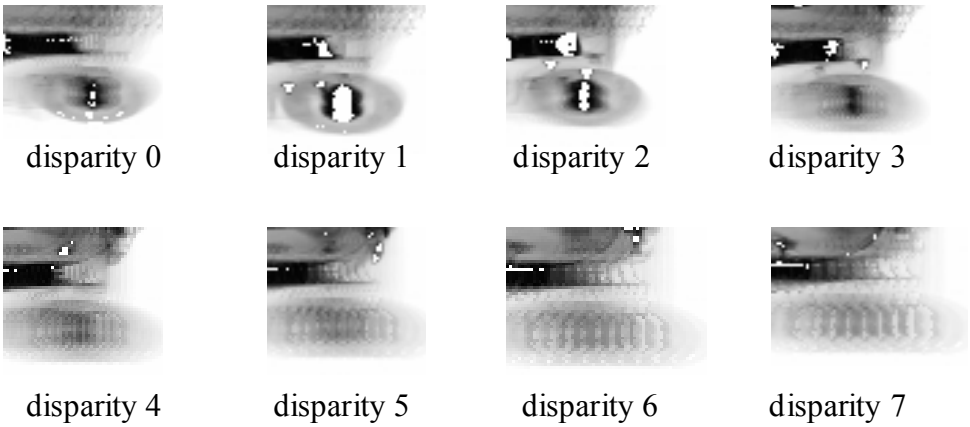


Figure 5.19 Laminograms for disparity 0 to disparity 7 employed for synthesising a target image at angle 0° after the first iteration.

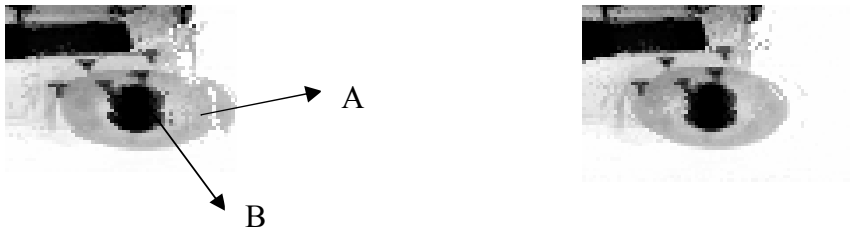


Figure 5.20 Region of interest in the synthetic images produced using detector images separated by 4° for target angle 0°, produced by the Basic LamI (left) and the Improved LamI (right) criteria.

5.4 Effect of increasing angular separation

The objective of this experiment is to investigate the performance of the Improved LamI criterion as a function of nInput and nVote and the angular separation of the input images. Employing the integration of the Improved LamI and the rank based voting, the investigation is extended to include increased angular separation. In this thesis, the angular separations considered are 4 ^{o}, 6 ^{o}, 8 ^{o} and 10^{o}. The increase in angular separation causes a concomitant decrease in the nInput and the number of different pairings for image synthesis. The relationship of angular separation, nInput and number of pairings are described by Table 4.1.

Figure 5.21 is arranged to illustrate the effect of increasing angular separation on the NE calculation for different n-m pairings. Besides, the resultant synthetic images for target angle 12° produced by 2-1 and 8-8 pairings are presented in Figure 5.22 for discussion.

5.4.1 Result: NE curves

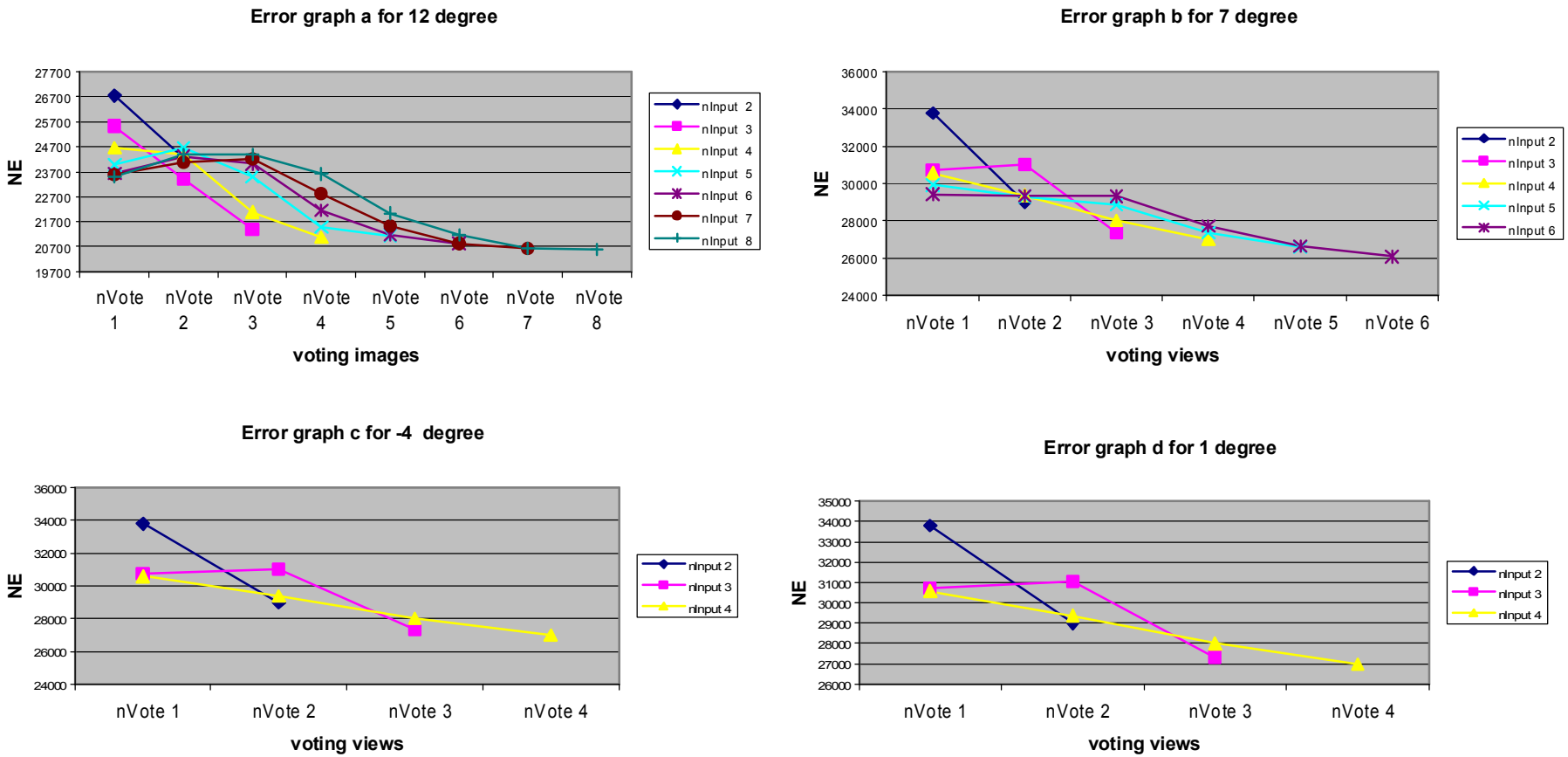


Figure 5.21 NE plots for different n-m pairings produced by using the Improved LamI on laminography matching cost calculation and rank based voting approach for (a)4°, (b)6°, (c)8°and (d10° angular separation between adjacent detector images where the respective target angle is 12°, 7°, -4° and 1°.

5.4.2 Result: Images

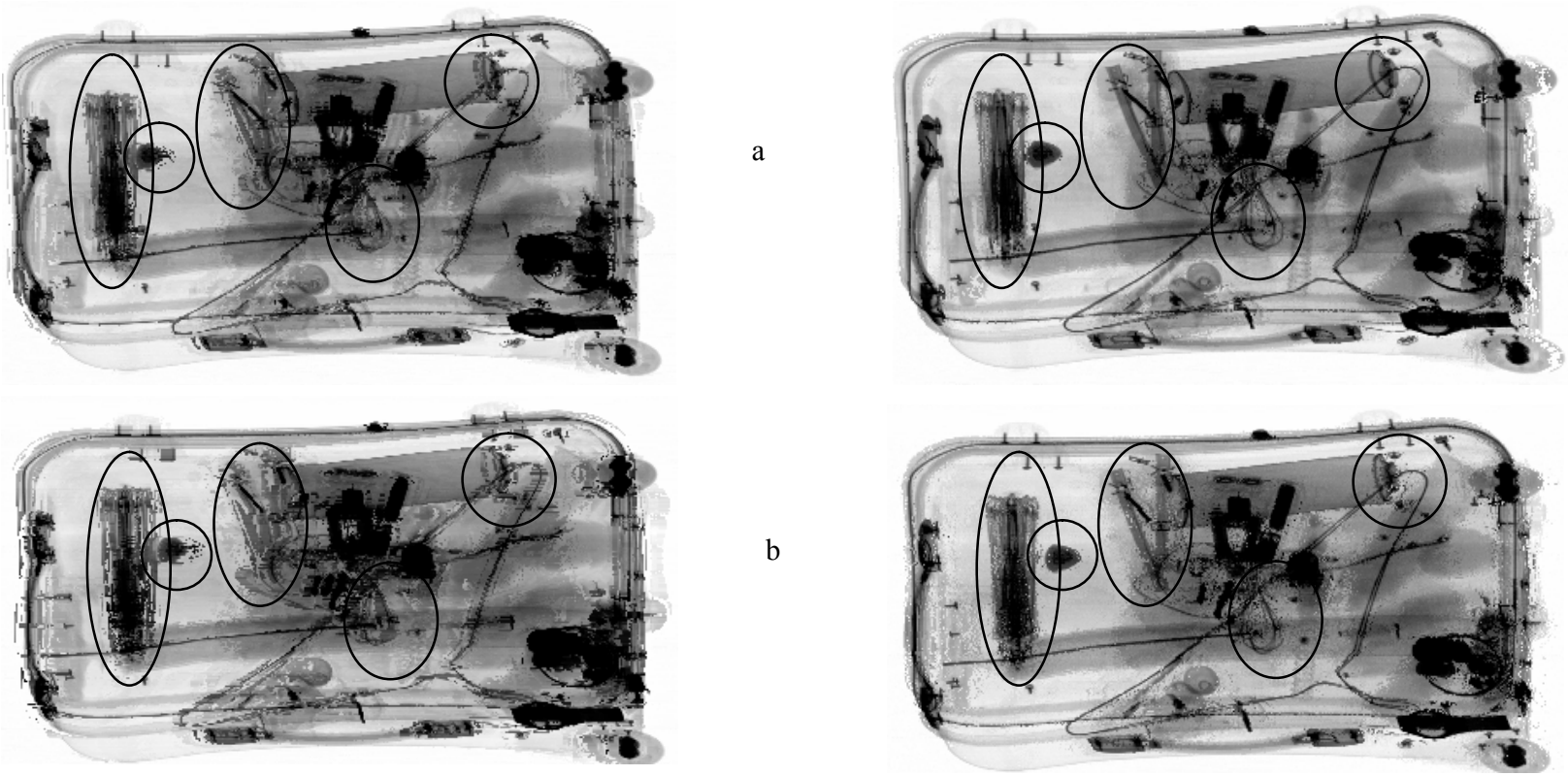


Figure 5.22a Examples of synthetic images produced by 2-1 and the 8-8 pairing for (a)4^o, (b)6^o angular separation between adjacent detector images that are produced by using the Improved LamI on laminography matching cost calculation and rank based voting approach. The respective target angle is 12^o, 7^o.

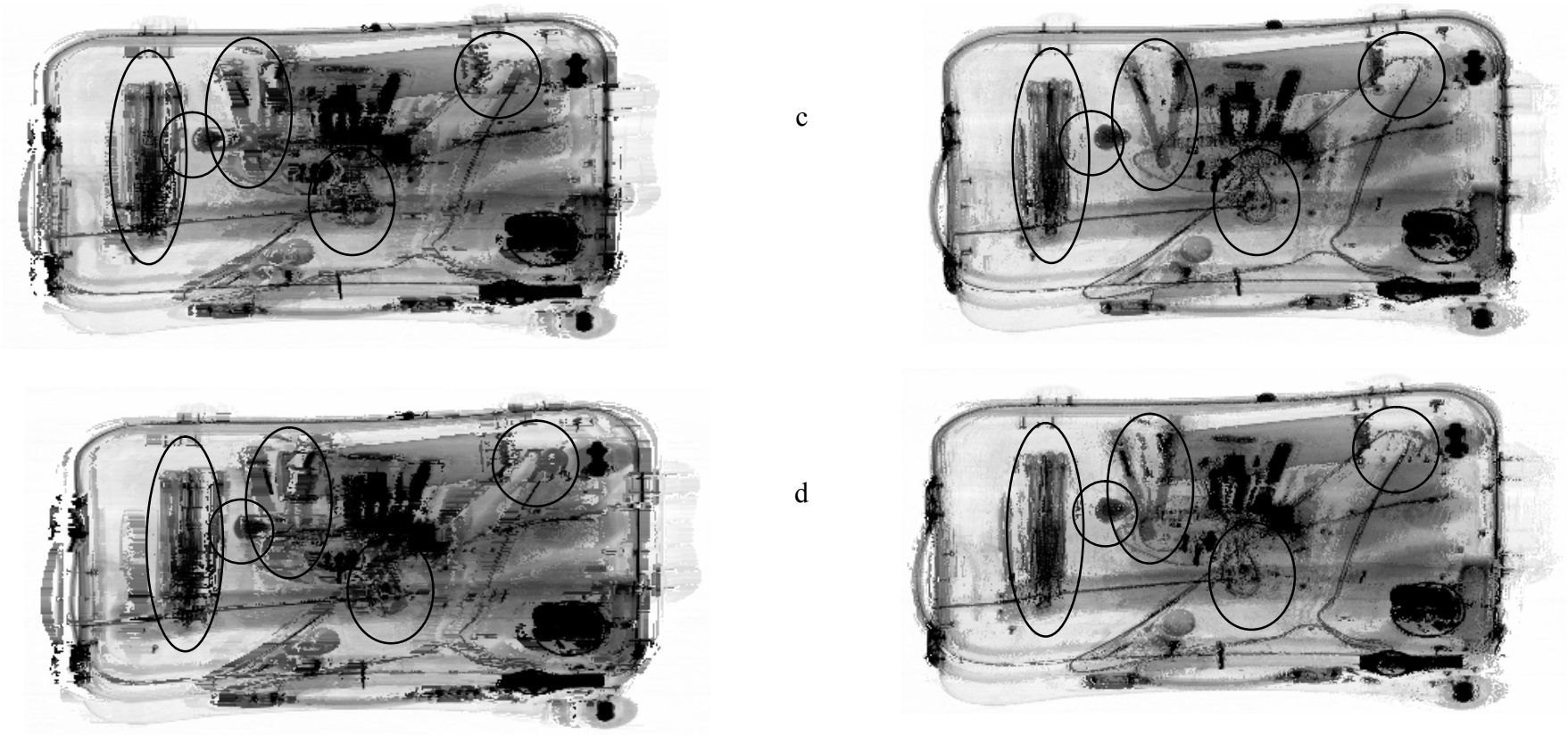


Figure 5.22b Examples of synthetic images produced by 2-1 and the 8-8 pairing for (c)8^{o}and (d)10^o angular separation between adjacent detector images that are produced by using the Improved LamI on laminography matching cost calculation and rank based voting approach. The respective target angle is -4^o and 1^o.

5.4.3 Analysis

It can be appreciated from the NE plots in Figure 5.21 that all angular separations exhibited the same trend, i.e, the NE reduces as the nInput and nVote increases. This finding is supported by the visual inspection of the resultant synthetic images arranged in Figure 5.22 by comparison of the highlighted regions for each angular separation.

5.5 Interim conclusion

The investigation into two different versions of the LamI criterion is reported. Both versions achieved the best result when 8-8 pairing is employed. As described in Section 5.1.3 and Section 5.2.3, while the increase in nInput and nVote offers a better solution, it may introduce unwanted artefacts during the synthesis process. Nonetheless, it was hypothesised and supported by the empirical results that the Improved LamI criterion could exploit the utility afforded by employing more images to produce an improved result.

Chapter Six Experiments and analysis: Sum of squared differences (SSD) criterion

The experimental results and analysis concerning the SSD criterion are presented in this chapter. These experiments are designed to be part of a larger group of interrelated experiments as discussed in Chapter 4. The experiment plan, Figure 4.2, illustrates the relationship of the experiments to the final algorithm structure. The objective of the experiments described in this chapter is to initially determine the optimum number of nInput and nVote when the SSD criterion is incorporated into (a) sum of reciprocals (SOR) matching cost calculation and rank voting and (b) laminography matching cost calculation and rank voting methods. For convenience of discussion, the former is termed the SOR-Rank voting method and the latter is termed the Lam-Rank voting method. Empirical analysis is included to demonstrably compare the performance of these two competing methods. Finally, the optimum number of nInput and nVote is studied as a function of the increasing angular separation between the input X-ray images. The SSD window size is set to 5 x 5 pixels for all the experiments.

6.1 SOR-Rank voting experiment

The objective of the experiments is to determine the optimum number of input images (nInput) and voting images (nVote) when SOR-Rank voting method is used to generate intermediary images. Figure 6.1 and 6.2 are arranged to enable the convenient comparison of a family of NE (number of errors) curves for different n-m pairings recorded for the target images at 12^0, 8^0, 4^0, 0^0, -4^0, -8^0 and -12^0 angles. The NE curves are produced by applying the SSD criterion in the SOR matching cost calculation and the rank voting approach. The symbol nVote1 indicates one voting image is involved. Besides, the GT image at 12° and the resultant synthetic images produced by 2-1, 3-2 and 8-8 pairings are presented in Figure 6.3 for comparative demonstration.

6.1.1 Result: NE curves

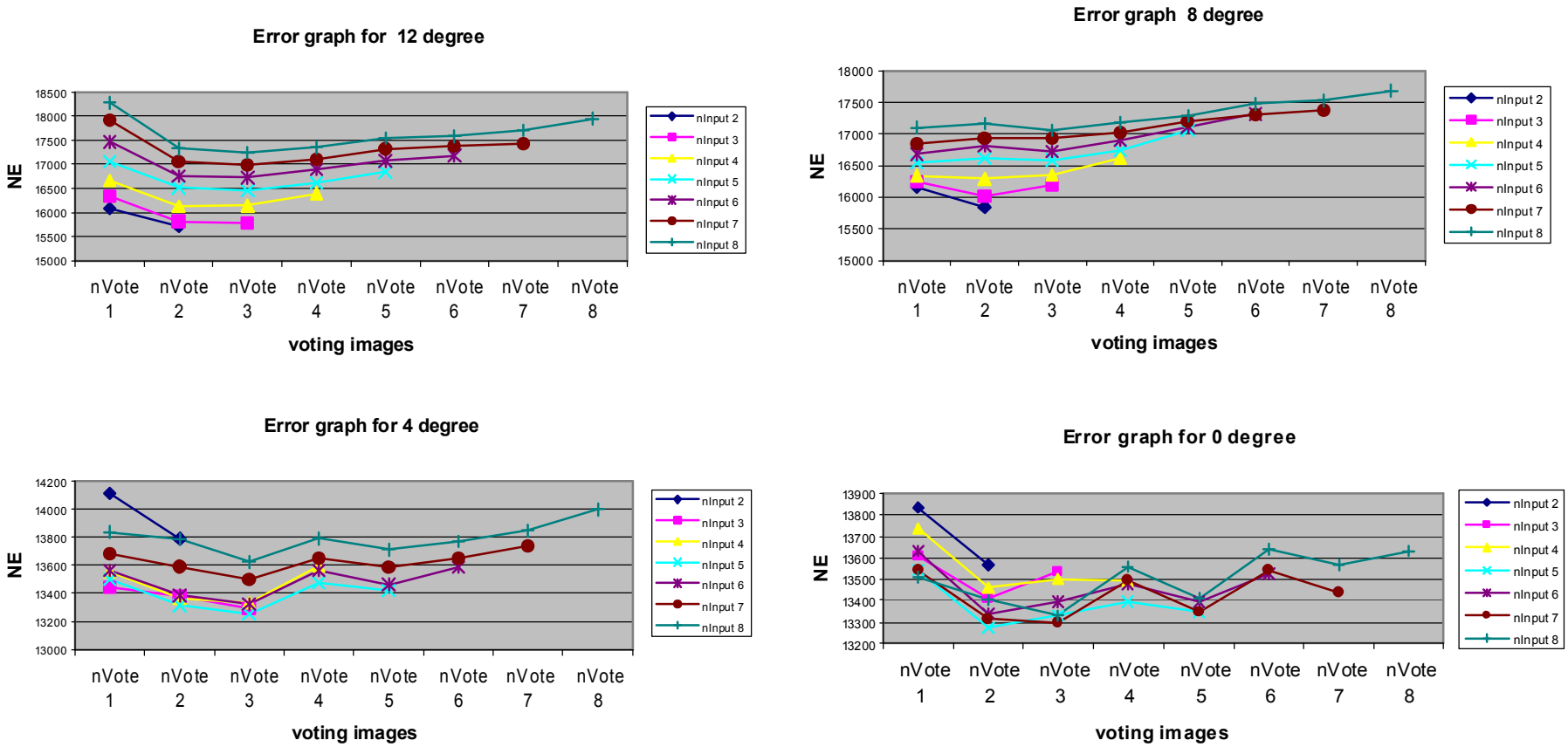


Figure 6.1 NE plots for different n-m pairings produced by using the SSD criterion on the SOR-Rank voting method to synthesis target images at 12° (left top), 8°(right top), 4°(left bottom) and 0° (right bottom), where the angular separation between adjacent detector images is 4°.

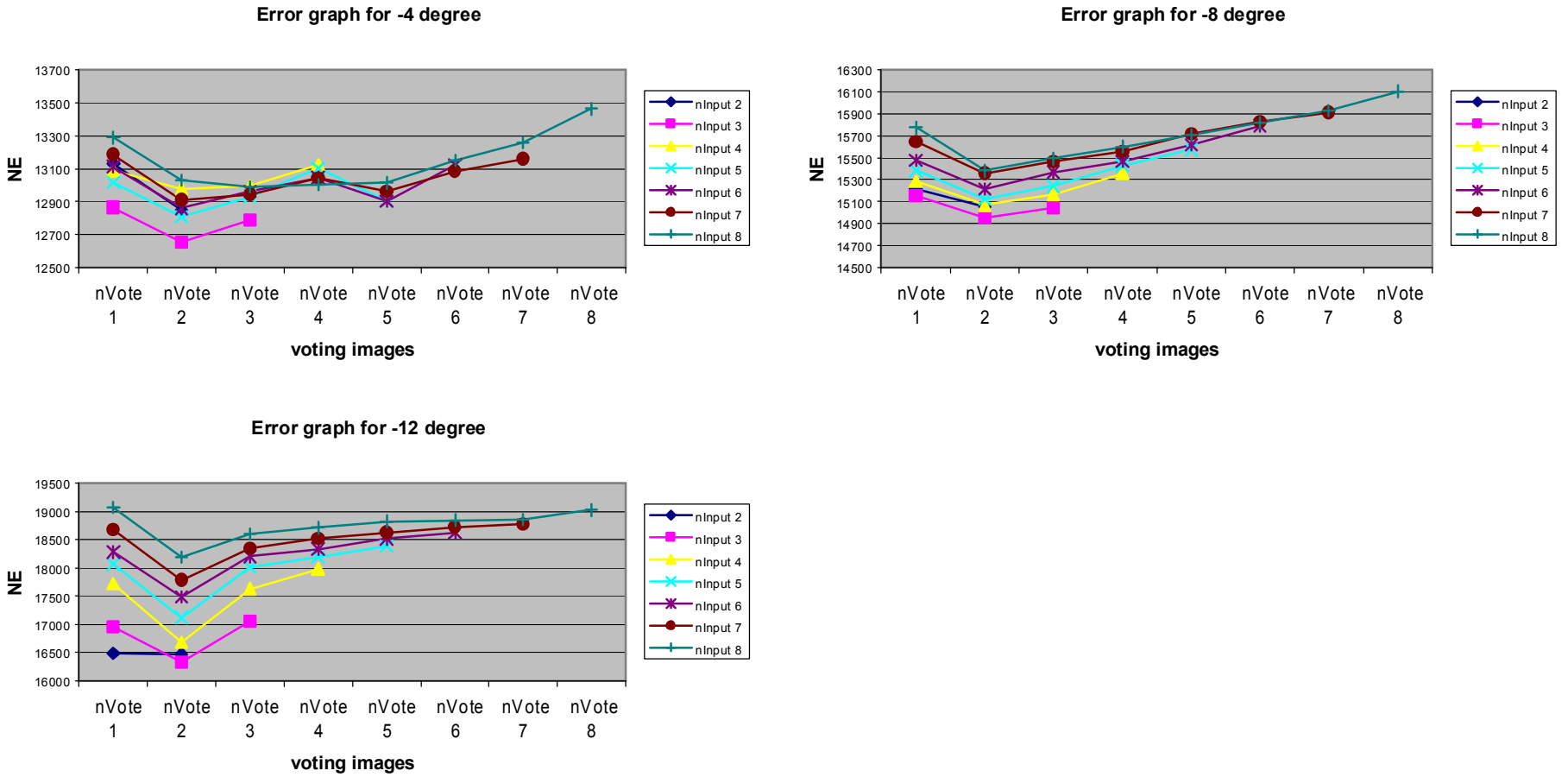


Figure 6.2 NE plots for different n-m pairings produced using the SSD criterion on the SOR-Rank voting method to synthesis target images at -4° (left top), -8°(right top) and -12°(left bottom), where the angular separation between adjacent detector images is 4°.

6.1.2 Result: Images

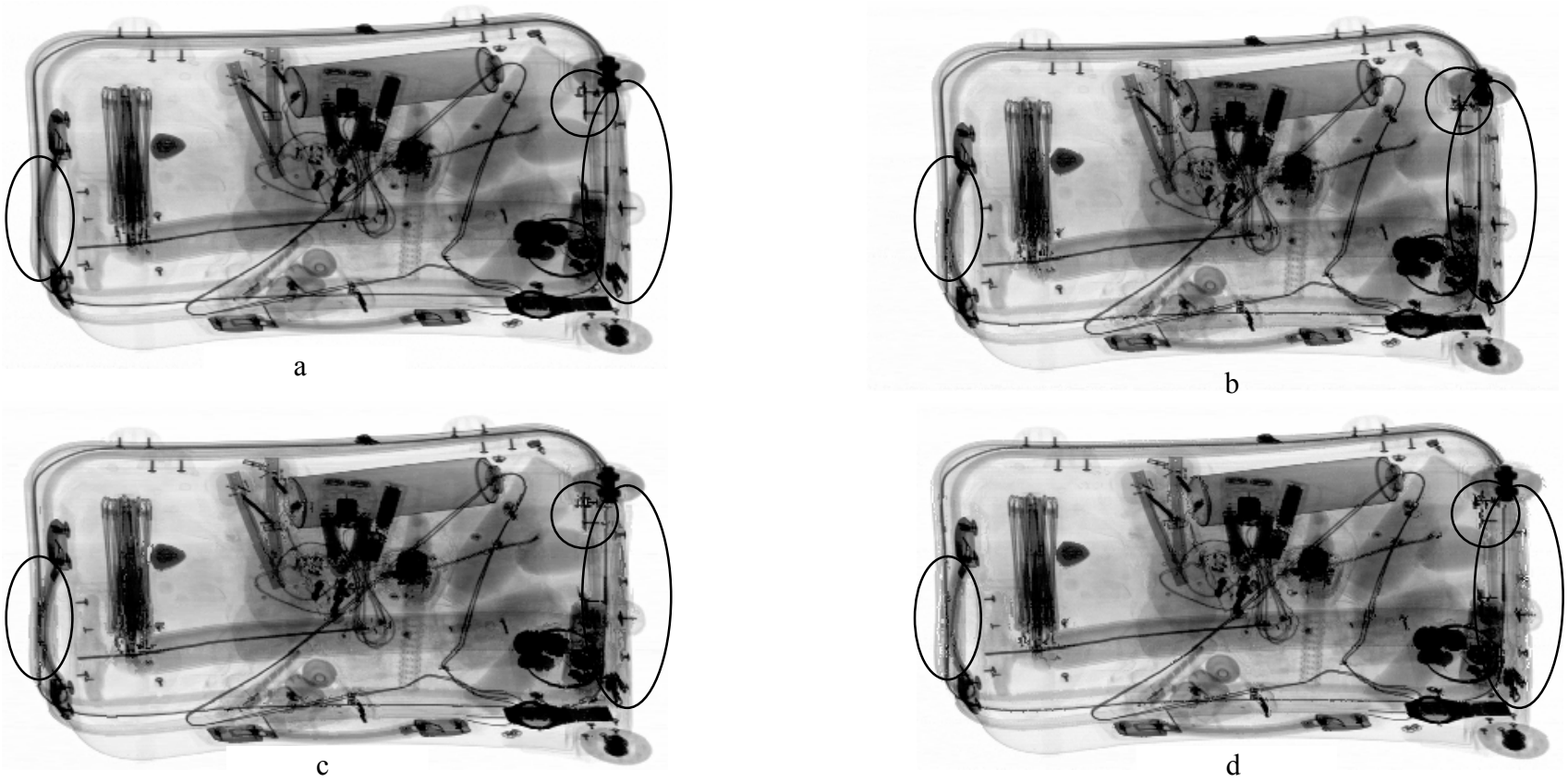


Figure 6.3 (a) The ground truth (GT) image at 12° and the synthetic images produced by (b) 3-2, (c) 2-1 and (d) 8-8 pairings of nInput and nVote.

6.1.3 Analysis

The basic trend is that the NE increases when nInput increases. Employing two voting images produced improved performance in comparison to a single voting image. However, the NE increases as the nVote continues to increase. Considering all target images, the algorithm produced the best result with the 3-2 pairing of nInput and nVote. This is supported by an example of the synthetic images that are arranged in Figure 6.3. Figure 6.3(a) is the GT image while the other three images are the resultant images using (b) 3-2 pairing, (c) 2-1 pairing and (d) 8-8 pairing. The 2-1 pairing uses the least number of input images and voting images, which represents the typical result produced by the basic stereo matching method. While this pairing uses the minimum amount of processing time, the result is inferior to the best 3-2 pairing result. This signifies the success and importance of the developed V-DMX algorithm. On the other hand, when all images are employed for cost calculation and voting (i.e, 8-8 pairing in this example), the resultant images are not better than the images produced by the best 3-2 pairing. This can be appreciated by comparing Figure 6.3(b) and Figure 6.3(d). This desirable outcome has demonstrated that only three input images are required for producing sufficiently high quality images without using all the available detector images. Four case studies are presented to illustrate the effect of changing nInput and nVote, individually.

Case study 1: Positive impact of increasing nInput with fixed nVote

The example presented here concerns the synthesis of the target image at -10°. To enable an effective discussion, the regions of interest (ROI) in the input images are arranged in Figure 6.4. Besides, the same ROI in the ground truth (GT) image as well as in the resultant synthetic images for different n-m pairings are organised in Figure 6.5.

Figure 6.4 Regions of interest in input images over the range of 12° to -12°.

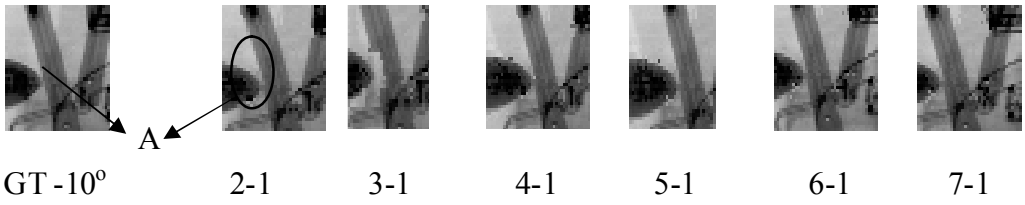


Figure 6.5 Comparison of the region of interest (ROI) in ground truth (GT) image -10° and ROI in the synthetic images produced by different n-m pairings, the increase in image quality produced by increasing nInput is evident.

A careful observation on the ROI extracted from the sequence of input images in Figure 6.4 revealed that feature-A is not overlapped in the first reference image -8° but is overlapped in the second reference image -12°. If only two input images are involved, feature-A will not be generated in the resultant synthetic image. This can be appreciated by inspecting 2-1 image in Figure 6.5. By adding the third input image -4° where the feature-A is not occluded assists the identification of the correct disparity for feature-A. This effect is evident on the 3-1 image in Figure 6.5. It is interesting to note that the images in Figure 6.5 indicate that any further addition of input images has no significant impact on synthesising feature-A. Nonetheless, the increase in nInput has provided additional support to the matching of repeating features, as exemplified by the features highlighted by black dotted circles in Figure 6.6.

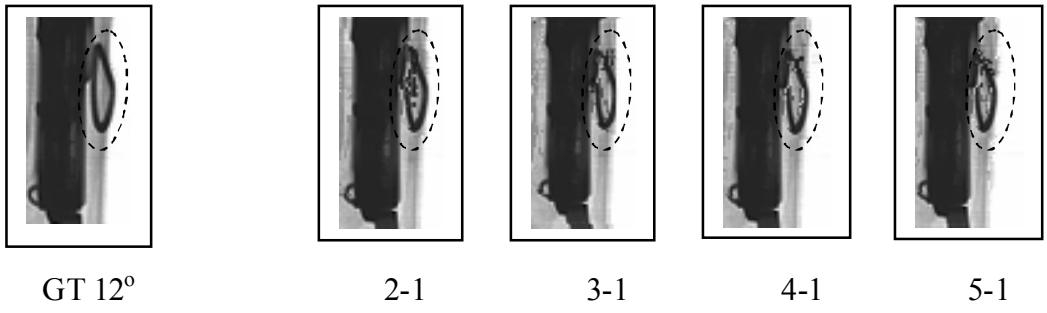


Figure 6.6 Improvement in repeating feature matching as a result of increasing nInput. (image set three).

Case study 2: Negative impact of increasing nInput with fixed nVote

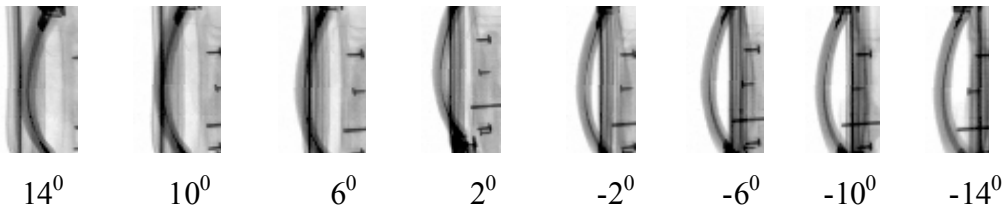


Figure 6.7 Regions of interest in input images over the range of 14° to -14°.

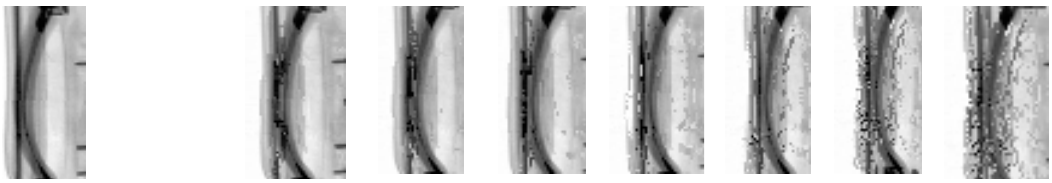


Figure 6.8 Comparison of the region of interest (ROI) in ground truth (GT) image -12° and ROI in the synthetic images produced by different n-m pairings, the decrease in image quality produced by increasing nInput is evident.

In this example, the target image is at 12^o, first reference image is at 14^o and the second reference is at 10^o. The ROI in all input images are arranged in Figure 6.7, and the ROI in the GT image as well as in the resultant synthetic images for different n-m pairings are organised in Figure 6.8. Comparatively, the image quality produced by using 2 to 5 input images can be considered equivalent. However, as soon as the nInput increases to 6 or 7 or 8, the image quality degrades significantly. This observation is due to the large angular separation between the reference images and the additional 6^{th} input image at -6^o (or 7^{th} input image at -10^o or the 8^{th} input image at -14^o). Normally, the larger the angular separation between two images, the more disparate the image content becomes. This general implication can be appreciated by a visual comparison between the reference images and the additional input images at -8^o or at -10^o or -14^o. The large dissimilarity in the neighbourhood pixels of the two SSD windows involved contribute to the localised failure. It is concluded that the matching potential realised by the application of the SSD criterion is limited by the image dissimilarity which in turn is exacerbated by increased angular separation between the input images.

Case study 3: Positive impact of increasing nVote with fixed nInput

Consider the problem illustrated by the example in Figure 6.8 where the image quality degrades as the nInput increases. Empirical results in Figure 6.9 are arranged to demonstrate that the problem can be eased by increasing the number of voting images. This improved quality increases from left to right i.e. 8-1 is inferior to the 8-8 pairing.

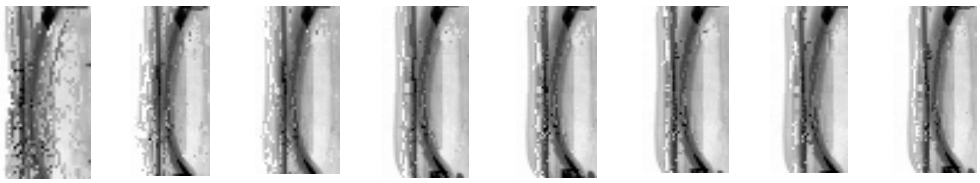


Figure 6.9 The region of interest in the synthetic images for target angle 12^o produced by 8-1, 8-2, 8-3, 8-4, 8-5, 8-6, 8-7 and 8-8 pairings.

Case study 4: Negative impact of increasing nVote with fixed nInput

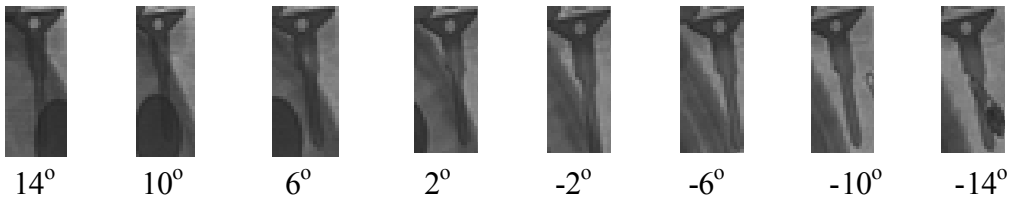


Figure 6.10 Regions of interest in input images over the range of 14^o to -14^o.(image set 2)

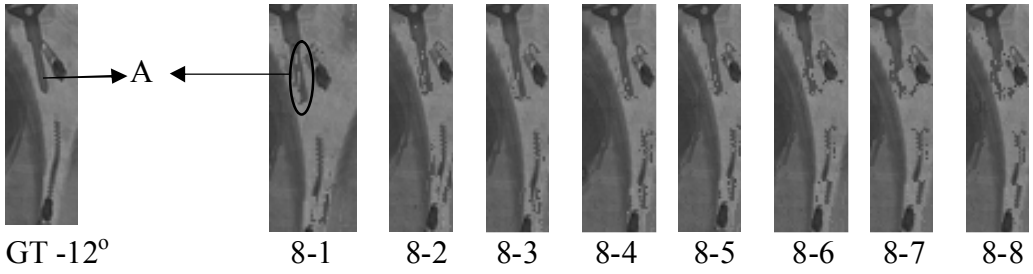


Figure 6.11 Comparison of the region of interest (ROI) in ground truth (GT) image -12° and ROI in the synthetic images produced by different n-m pairings.(image set two)

The target angle for this example is angle -12^o. The ROI in all the input images are arranged in Figure 6.10. By comparing the feature-A in the GT image and in the resultant synthetic images produced by nVote = 1 to 5, it was considered to be comparable in terms of image quality. However, as the nVote increases to 6 or 7 or 8, feature-A becomes significantly distorted (see the 8-6, 8-7 and 8-8 pairings in Figure 6.11). This is due to the increase in the overlapping in the ROI exhibited by the 6th, 7th and 8th voting images (i.e, 14^o, 10^o, 6^o input images), which increases the matching ambiguity.

In summary, when only two input images are involved, the SSD criterion is capable of handling most parts of the images very well. However, the SSD criterion is known to suffer from errors produced by discontinuity, repeating features, and overlapping

features. When the nInput increases to 3, an apparent improvement to the common problems were observed but simultaneously, the features that two input images can handle well were compromised. Nonetheless, it was considered that the utility afforded by employing three input images has conveyed an overall improvement to the image quality. As the nInput increased to more than 3, the contribution to address the common problems is marginal. On the other hand, the image quality reduced. This effect occurs because the increase in angular separation between the images (e.g, between the first and the fourth image) produces a concomitant increase in the dissimilarity in the SSD window content. It is concluded that the optimum number of nInput is three for the imaging geometry under consideration in this programme of research.

Since nVote is either the subset or the full set of nInput, the nVote can only be either 1 or 2 or 3. It was observed that while using one voting image can offer a good solution, SSD is still susceptible to the common problems discussed in the preceding text. The situation can be improved when using two images to vote for the result. Any further addition of voting images tends to impair the result produced by using 2 voting images. In comparison with the LamI criterion (as described by the Chapter 5), which favours more images, the empirical results produced by the SSD criterion has demonstrated the conflicting requirements of increasing the number of input images, as exemplified by the graphs in Figure 6.1 and 6.2. Thus, it is important to establish the best n-m pairing to optimise the utility afforded by the multiple images. In conclusion the empirical evidence supports, on balance, a 3-2 pairing for optimum performance.

6.2 Lam-Rank voting experiment

The objective of the experiments is to determine the optimum number of input images (nInput) and voting images (nVote) when Lam-Rank method is used to generate intermediary images.

The method described in the previous section computes SSD matching costs between the voting images and the input images prior to combining the results using the SOR concept. However, when the SSD criterion is applied to the laminography matching cost calculation method, the computation of SSD costs refer to inherently dissimilar (in terms of intensity) voting images and laminograms. In other words, the SSD criterion operates on the raw data when using SOR method, while it operates on the laminograms (the combined) data when using laminography method. The rationale of the method is

associated with the in focus feature offered by the laminograms. When an object is in focus, the intensity variation in the neighbourhood window would be less fluctuating because the effect of other object is minimally averaged, making the in focus object to be more stable.

Figures 6.12 and 6.13 are arranged to enable the convenient comparison of a family of NE curves for different n-m pairings recorded for target images at 12^0, 8^0, 4^0, 0^0, -4^0, -8^0 and -12^0 angles. The NE curves are produced by applying SSD criterion to the integration of the laminography matching cost calculation and the rank voting approach. The symbol nVote1 indicates one voting images is involved. Besides, the GT image at 12° and the resultant synthetic images produced by 2-1, 2-2 and 8-8 pairings are presented in Figure 6.14 for comparative discussion.

6.2.1 Result: NE curves

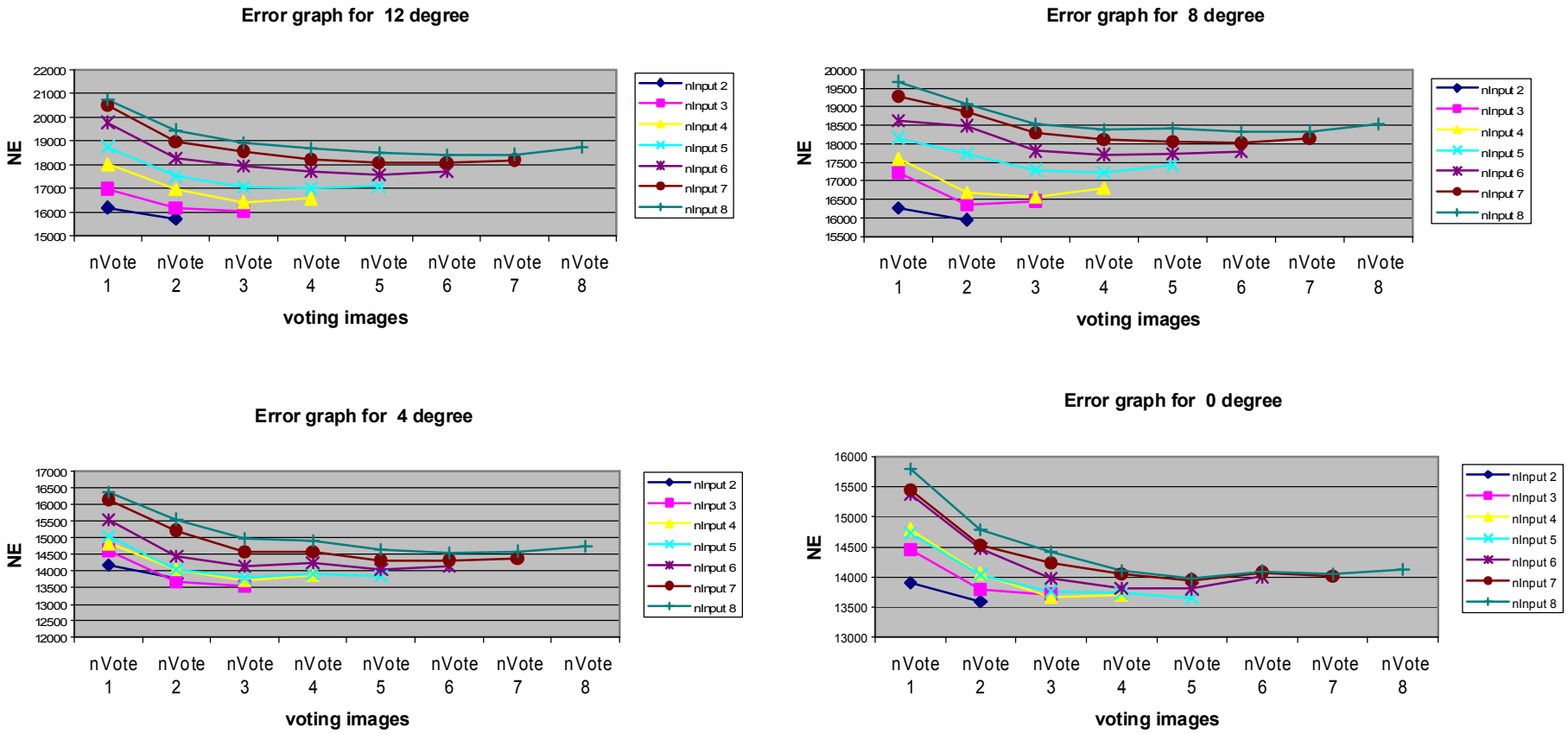


Figure 6.12 NE plots for different n-m pairings produced by using the SSD criterion on Lam-Rank method to synthesis target images at 12° (left top), 8°(right top), 4°(left bottom) and 0° (right bottom), where the angular separation between adjacent detector images is 4°.

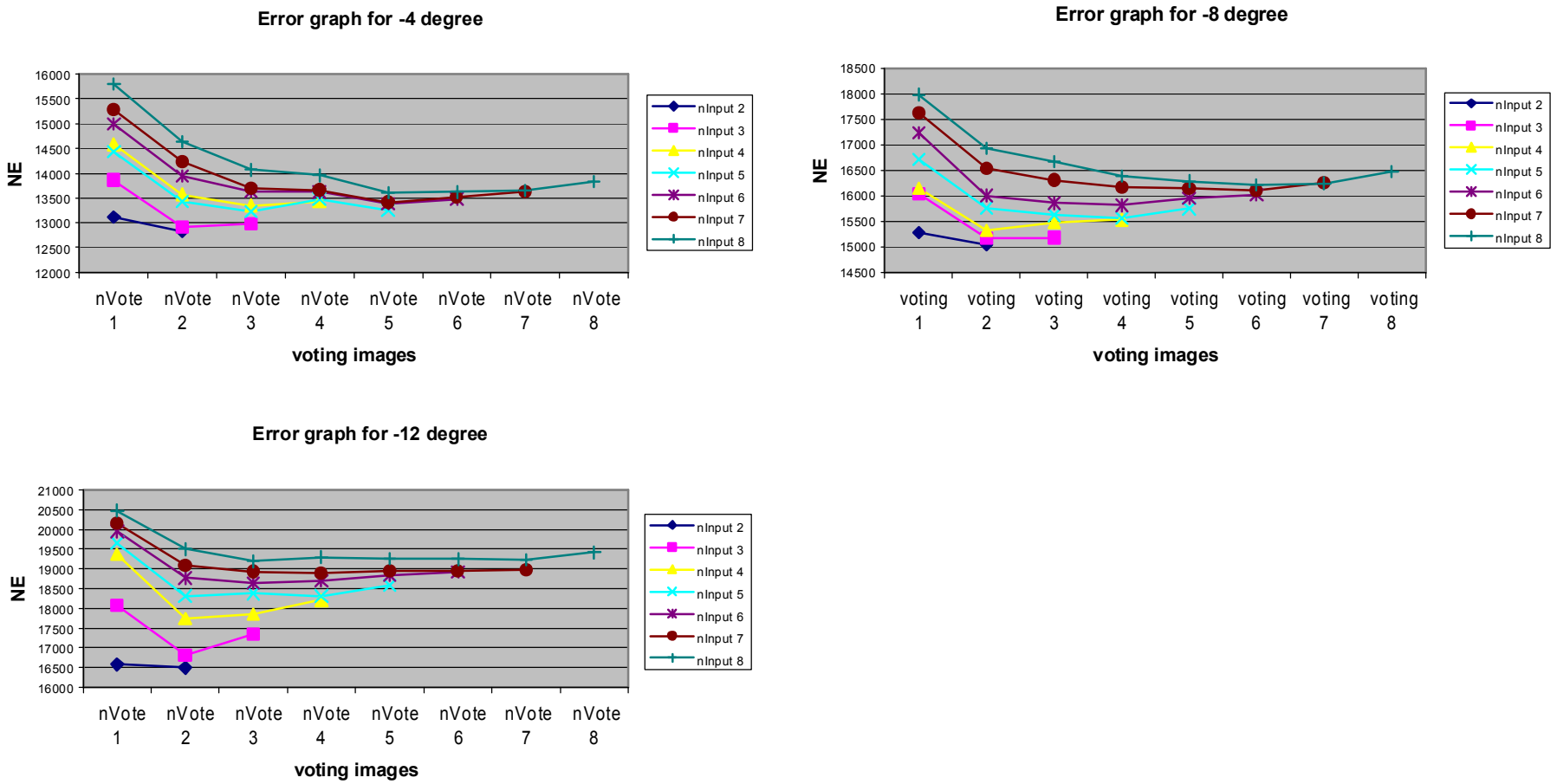


Figure 6.13 NE plots for different n-m pairings produced by using the SSD criterion on Lam-Rank method to synthesis target images at -4°(left top), -8°(right top) and -12°(left bottom), where the angular separation between adjacent detector images is4°.

6.2.2 Result: Images

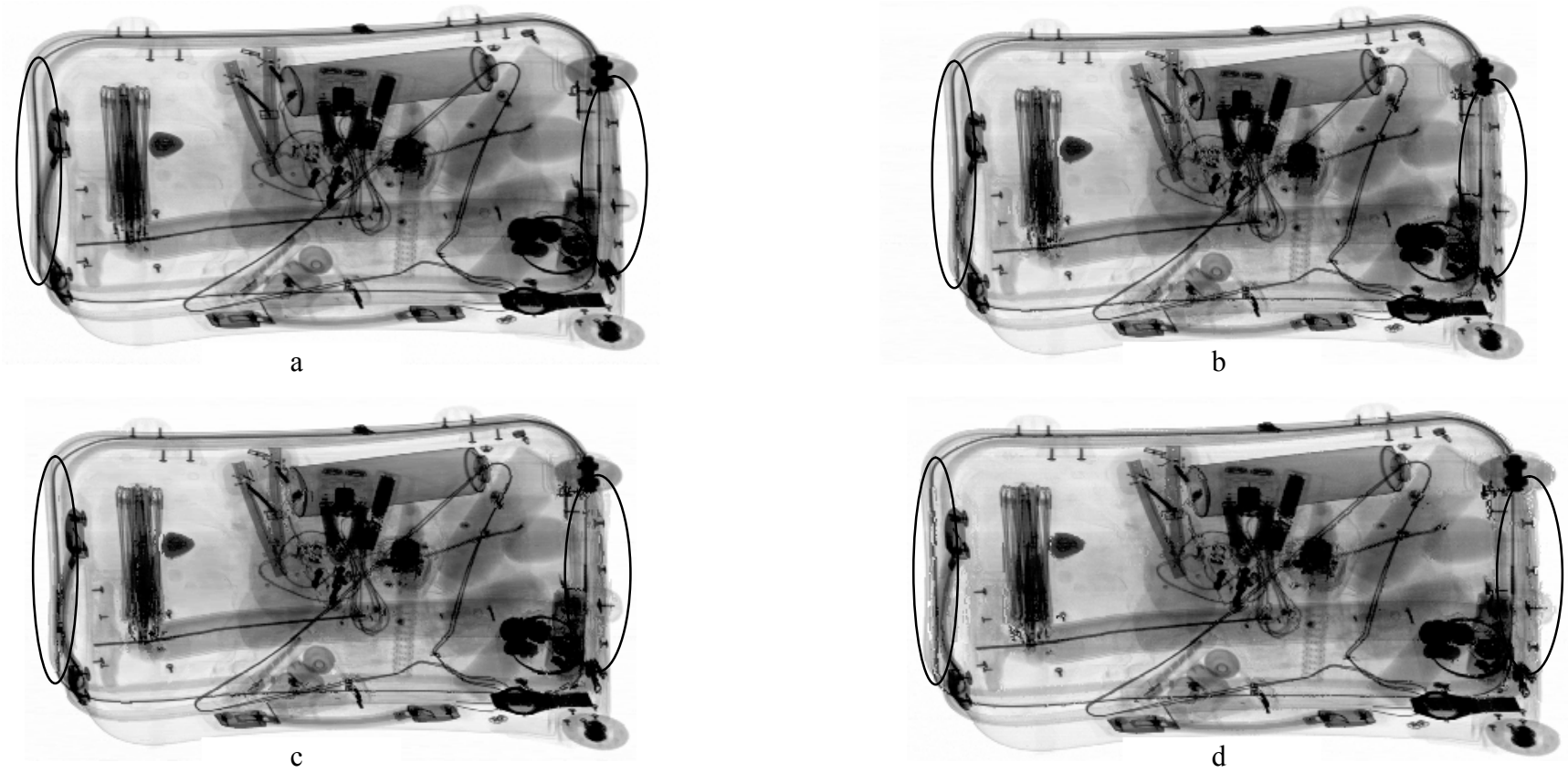


Figure 6.14 (a) The ground truth (GT) image at 12° and the synthetic images produced by (b) 2-2, (c) 2-1 and (d) 8-8 pairings of nInput and nVote.

6.2.3 Analysis

The basic trend observed is that fewer input images perform better, while an increased number of voting images improve the performance. It is evident that the combination for the best performance is the 2-2 pairing. This is supported by an example of the synthetic images that are arranged in Figure 6.14. Figure 6.14(a) is the GT image while the other three images are the resultant images using (b) 2-2 pairing, (c) 2-1 pairing and (d) 8-8 pairing. By visually comparing the highlighted image regions in Figure 6.14(c) and Figure 6.14(d), it can be appreciated that the 2-2 pairing outperformed the 2-1 pairing result, as expected. However, if the input images and the voting images are each increased to eight, then the quality of the target is not radically impaired, as can be appreciated by comparing Figure 6.14(b) and Figure 6.14(d). The result has once again validated the rationale of the research programme that compelling synthetic images can be generated without employing all the detector images.

6.3 Comparative study of SOR-Rank and Lam-Rank methods

This experiment is a comparative study of the SOR matching cost calculation and the laminography matching cost calculation approaches. Figure 6.15 and Figure 6.16 report the synthetic results produced by the 3-2 pairing of the SOR approach and the 2-2 pairing of the laminography approach.

6.3.1 Result: NE curves

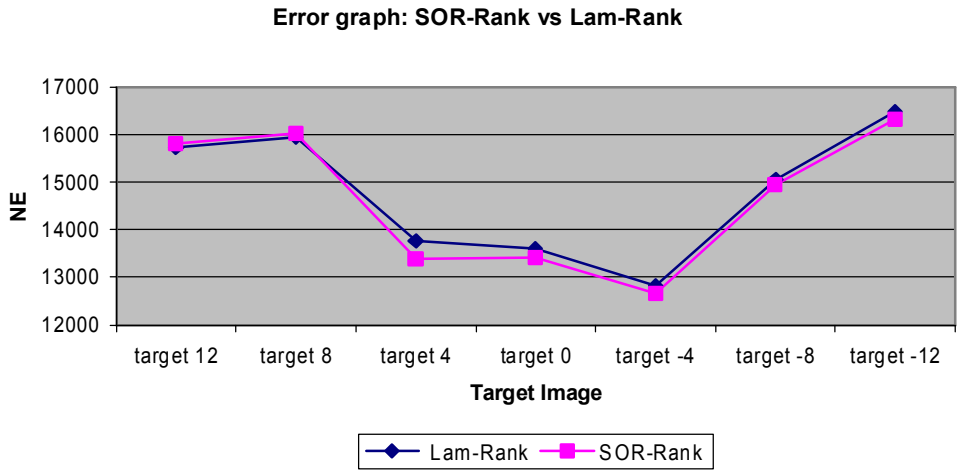
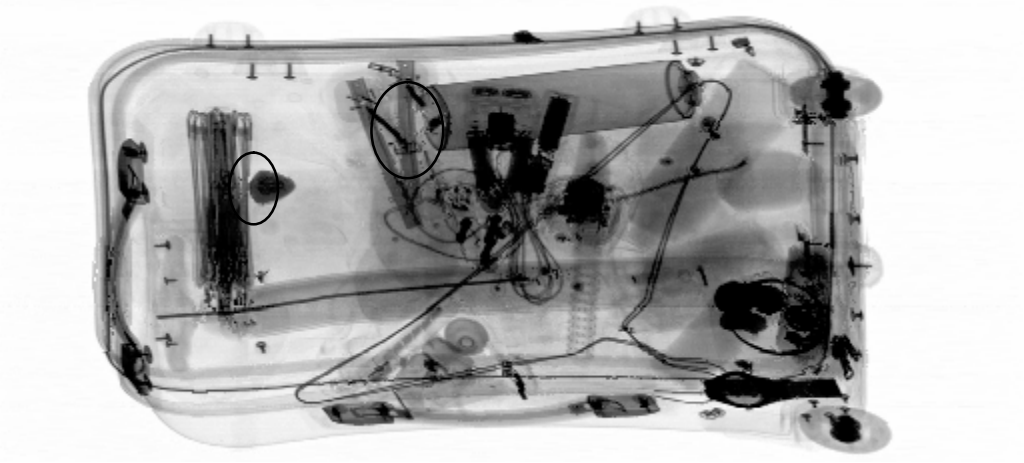
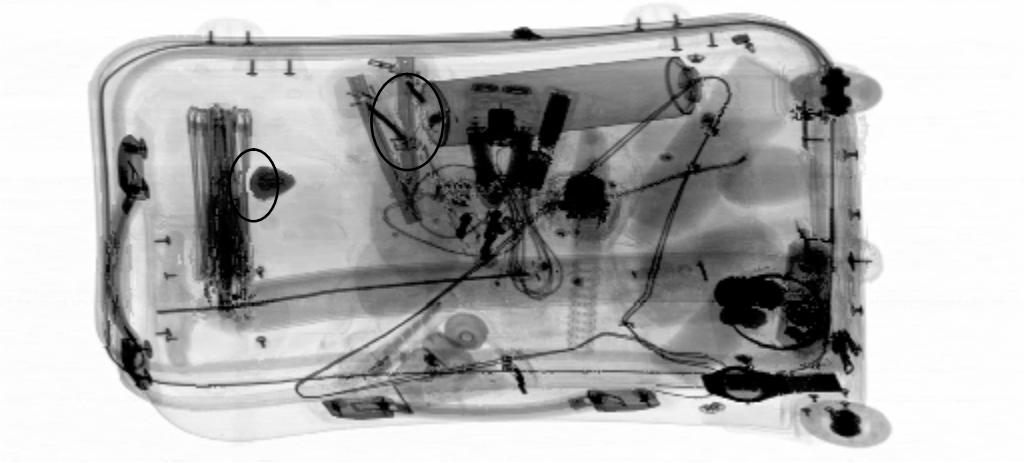


Figure 6.15 Comparison of NE plots of rank based voting methods produced by the 3-2 pairing for the SOR matching cost calculation, and the 2-2 pairing of the laminography matching cost calculation methods, for target angle 12°, 8°, 4°, 0, -4°, -8° and -12° using the SSD criterion.

6.3.2 Result: Images



a



b

Figure 6.16 Synthetic images produced using perspective images separated by 4° for target angle 12° produced by 3-2 pairing of the SOR matching cost calculation (a) and the 2-2 pairing of the laminography matching cost calculation methods (b) using the SSD criterion.

6.3.3 Analysis

As described in Section 3.3.2, SOR is designed to highlight the low matching cost between images, which produced better results than the laminography based method. This can be appreciated by comparing the curves in Figure 6.15 and the highlighted image regions in the resultant synthetic images in Figure 6.16.

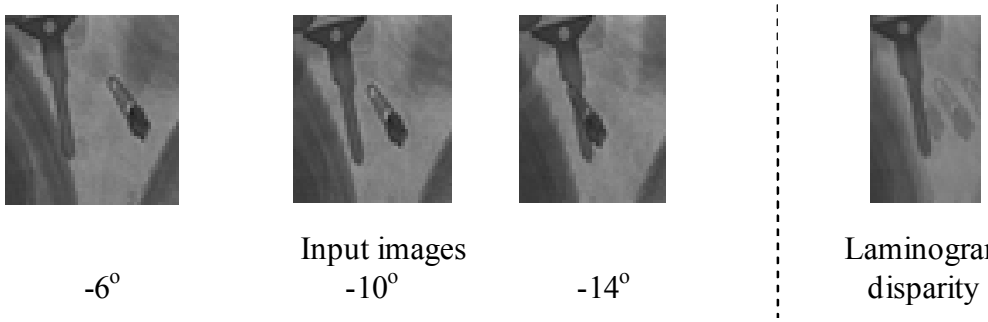


Figure 6.17 Region of interest in the input images -6° -10° and -14° and in the laminograms with disparity.(image set 2).

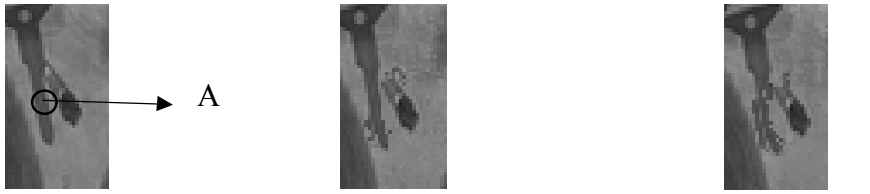


Figure 6.18 Region of interest in the ground truth (GT) image at -12° and the synthetic images produced by using the SSD on 3-2 SOR-Rank voting method and, by using the SSD on 2-2 Lam-Rank voting method. (image set 2).

To enable a detailed discussion with an example, Figure 6.17 depicts correlated regions of interest (ROI) in input images at -6°, -10° and -14°, as well as in laminogram with disparity 9. Figure 6.18 shows the same ROI in the GT target image and synthetic images produced by the SOR-Rank voting method using 3-2 pairing and the Lam-Rank voting method using 2-2 pairing. It was observed that the SOR based approach could handle feature-A better than the laminography based approach.

Feature-A is overlapped by an adjacent structure in the second reference image (image -12°) but it is not subject to overlapping in the first reference image (image -10°) and the next input image -6° in the image sequence. Using the SSD criterion, the matching cost on the feature-A between image -10° and image -14° is high, and the matching cost between image -10° and image -6° is low. When aggregating the two costs, the reciprocal nature of the SOR approach highlighted the low cost between -10° and -6°, thus generated the feature-A partially right (see Figure 6.18).

It can be deduced from Figure 6.17 that feature-A is focused on laminogram 9 and its neighbourhood intensity value has changed in comparison with the first reference image. This led to a high matching cost that is relatively higher than the SOR approach when using the SSD criterion on feature-A between image 4° and laminogram 9. As a result,

the laminography approach is inferior to SOR approach to synthesise feature-A, as presented in Figure 6.18.

In conclusion, the SSD criterion has a better performance when using SOR matching cost calculation than laminography based approach.

6.4 Effect of increasing the angular separation

The objective of this experiment is to investigate the performance of the optimum setting of nInput and nVote for the SOR-Rank voting method as the angular separation between the input images is increased. In the research programme, the angular separations considered are 4^{o}, 6^{o}, 8^{o} and 10^{o}. The increase in angular separation causes a concomitant decrease in the nInput and the number of different pairings for image synthesis. The relationship of angular separation, nInput and number of pairings are described in Chapter 5 Table 5.1.

Figure 6.19 is arranged to illustrate the effect of increasing the angular separation on the NE calculation for different n-m pairings. The target angle for angular separation of 4°, 6°, 8° and 10° is -4°, -11°, -12° and -8°, respectively. To enable a visual demonstration, the resultant synthetic images for target angle 12° produced by the 3-2 pairing and the worst pairing for each angular separation are presented in Figure 6.20.

6.4.1 Result: NE curves

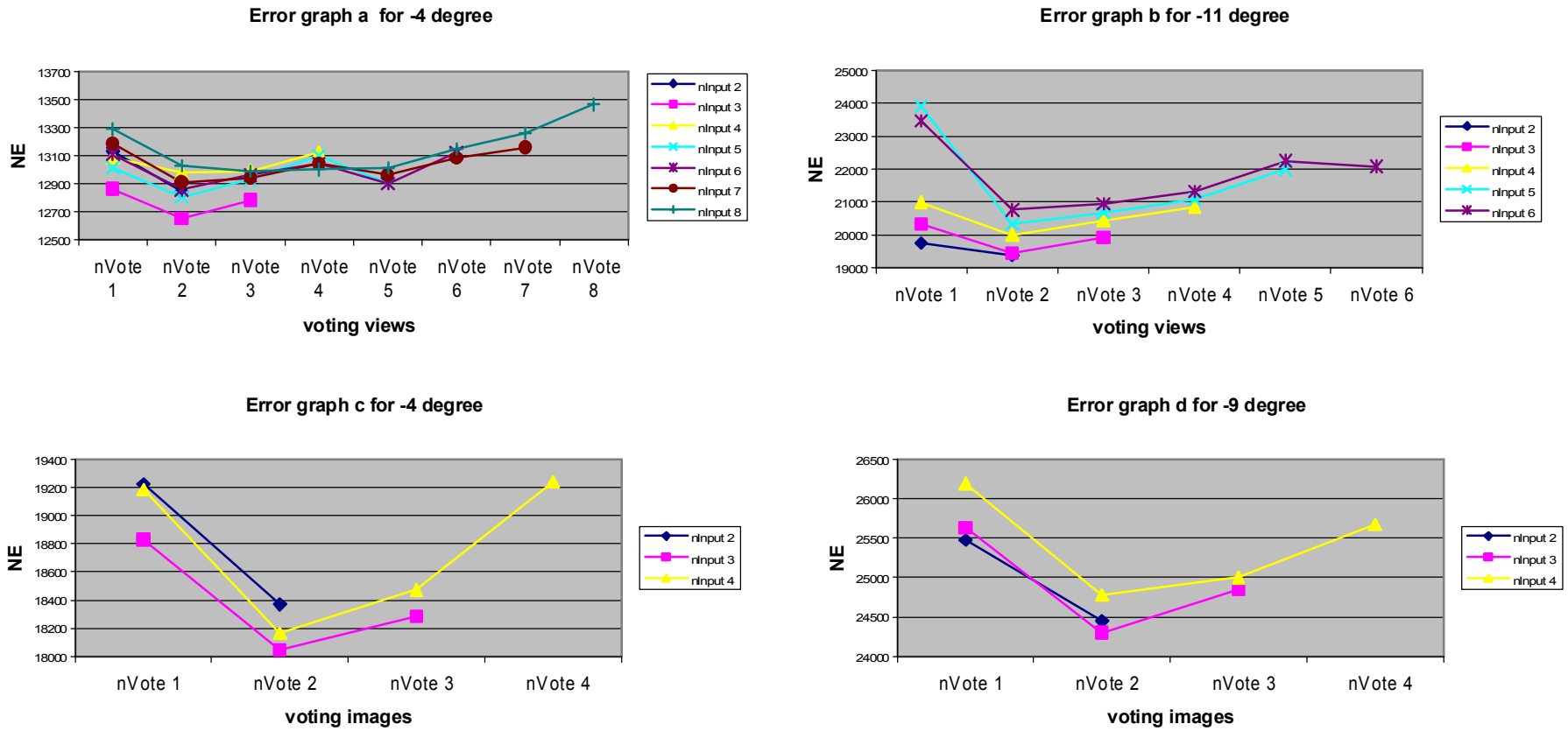


Figure 6.19 NE plots for different n-m pairings produced by using SSD criterion on the SOR-Rank voting method for (a)4°, (b)6°, (c)8°and (d)10° angular separation between adjacent detector images where the respective target angle is -4°, -11°, -4° and -9°.

6.4.2 Result: Images

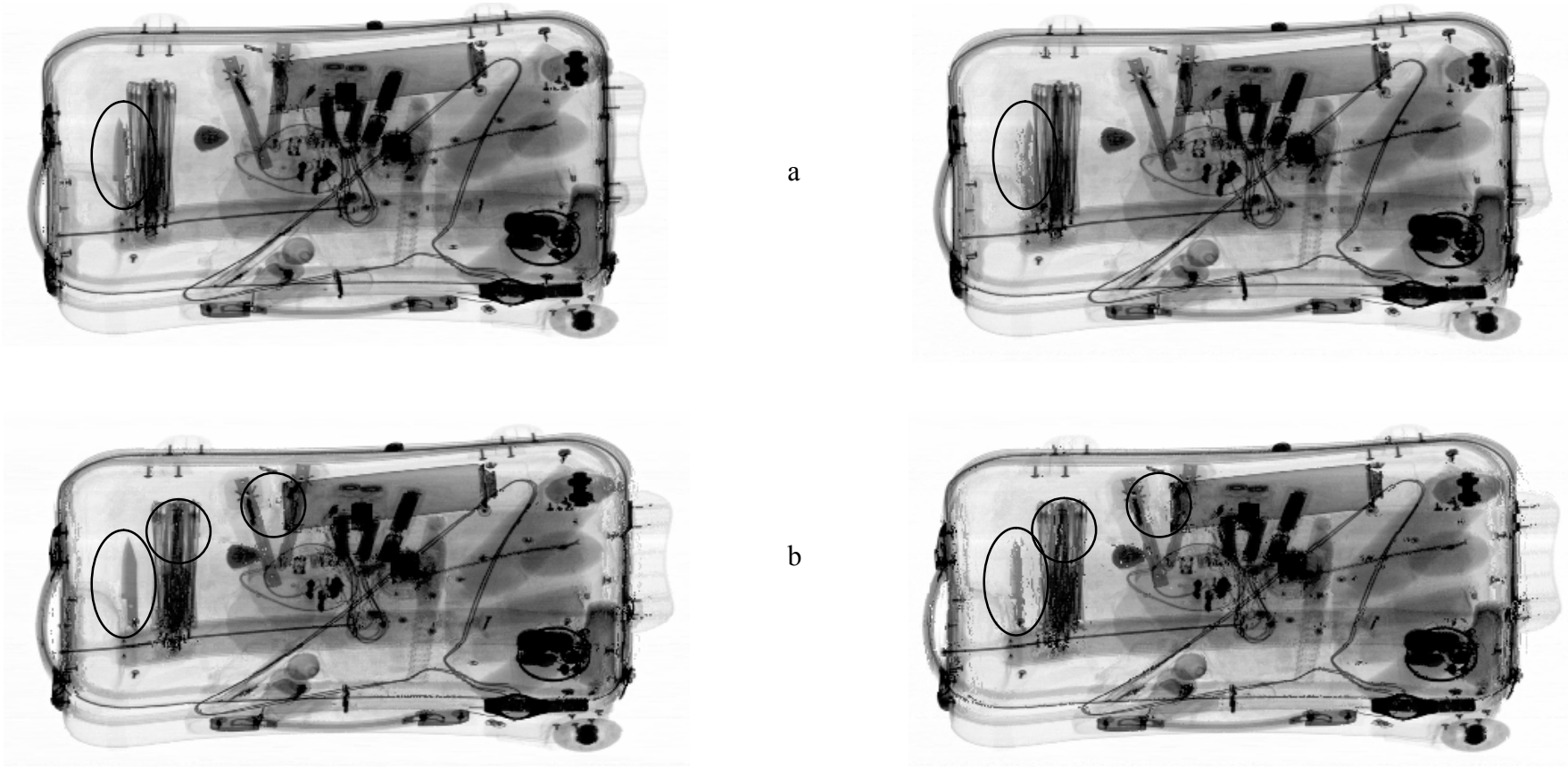


Figure 6.20a Examples of synthetic images produced by the 3-2 pairing (left column) and the worst pairing (right column) for (a)4^o, (b)6^o angular separation between adjacent detector images that are produced by using SSD criterion on the SOR-Rank voting method. The respective target angle is -4^o, -11^o.

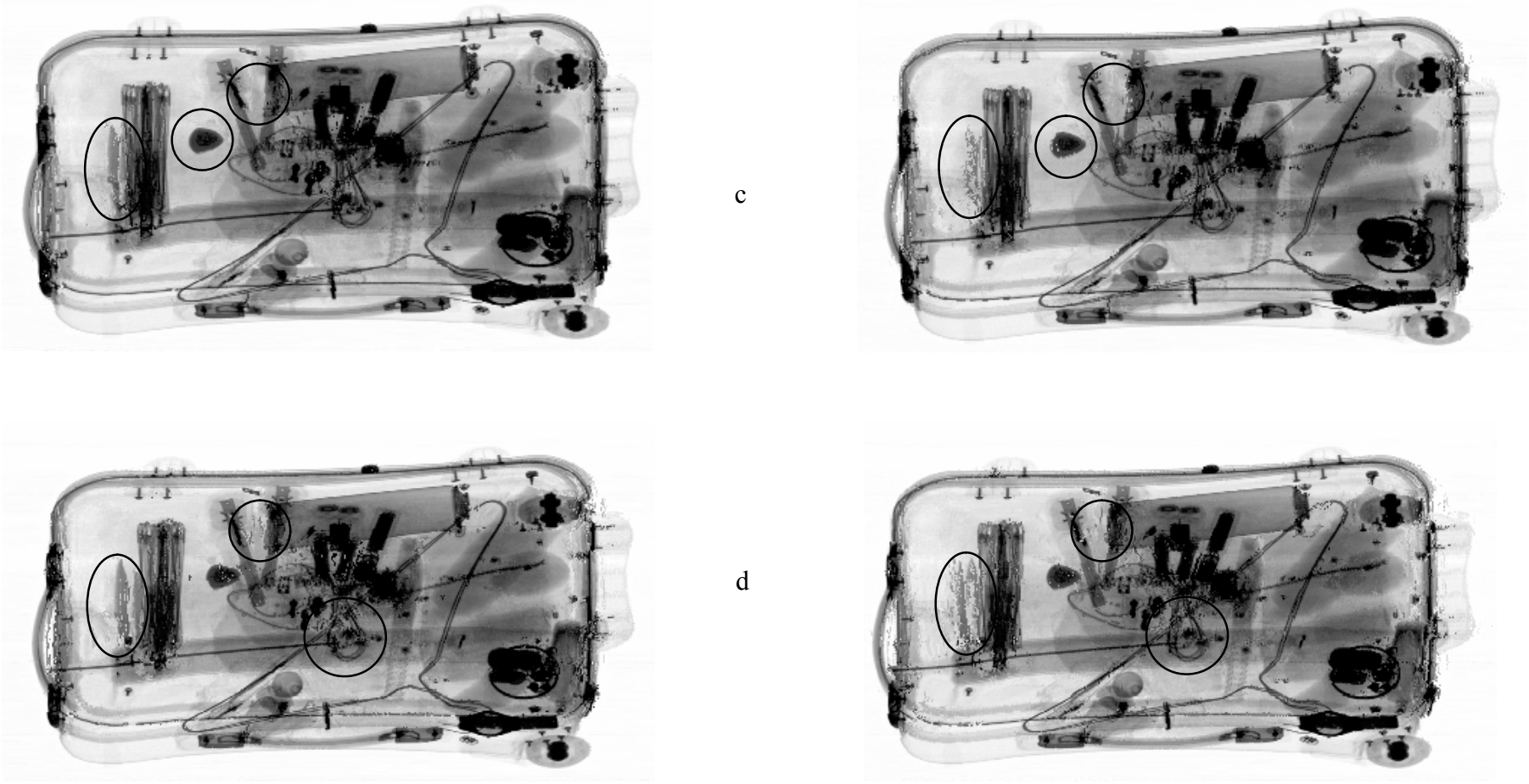


Figure 6.20b Examples of synthetic images produced by the 3-2 pairing (left column) and the worst pairing (right column) for (c)8^o and (d)10^o angular separation between adjacent detector images that are produced by using SSD criterion on the SOR-Rank voting method. The respective target angle is -4^o and -9^o.

6.4.3 Analysis

It can be appreciated from Figure 6.19 that the NE plots produced by all angular separations exhibited the same trend, i.e, 3-2 has the lowest NE. This finding is supported by the visual inspection of the resultant synthetic images arranged in Figure 6.20. The images arranged on the left column are produced by the best 3-2 pairing while the images on the right column are produced by the worst pairing for each respective angular separation.

Generally as the angular separation of the beams is increased the resultant images become more disparate necessitating an increase in the pixel disparity limit, the size of disparity window becomes bigger which increase the number of false match. It can be appreciated from the synthetic images presented in Figure 6.20 that the image fidelity is degraded as the X-ray beam angle increases.

6.5 Interim conclusion

In conclusion, the empirical evidence indicated the best pairing for the SOR-Rank voting method and the Lam-Rank voting method is 3-2 and 2-2. This encouraging result revealed that the algorithm does not require all the detector images to produce sufficiently high quality synthetic images. The outcome of the comparative study has proven that the SOR approach is a better manner to compute matching cost when using SSD criterion. The capability of the SOR-Rank voting method was investigated as a function of four angular separations, which are 4°, 6°, 8° and 10°. For all angular separations, the 3-2 pairing outperformed other possible n-m pairings. The best result is obtained for 4º and the worst result at 10º. This finding was expected as increasing the separation of the X-ray beams also increases the maximum disparity which in turn increases the potential for spurious pixel matches.

Chapter Seven Development of the V-DMX algorithm

This Chapter reports the development and empirical investigation of a dual criteria, which combines the complementary aspects of the Improved LamI and the SSD criteria; each criterion is presented and discussed independently in the preceding Chapters Five and Six, respectively.

7.1 Comparative study of LamI criterion and SSD criterion

Experiments reported in Chapters 5 &6 established the optimum pairing of input images and voting images for the Improved LamI and SSD criteria. For the SSD criterion, three input images and two voting images with SOR matching cost calculation (BEST-SSD) has the best performance. For the Improved LamI criterion, maximum number of input images and voting images (BEST-LamI) produced the best result. A comparative analysis of the BEST-SSD and BEST-LamI is presented in the following text and figures.

7.1.1 Result: NE curves

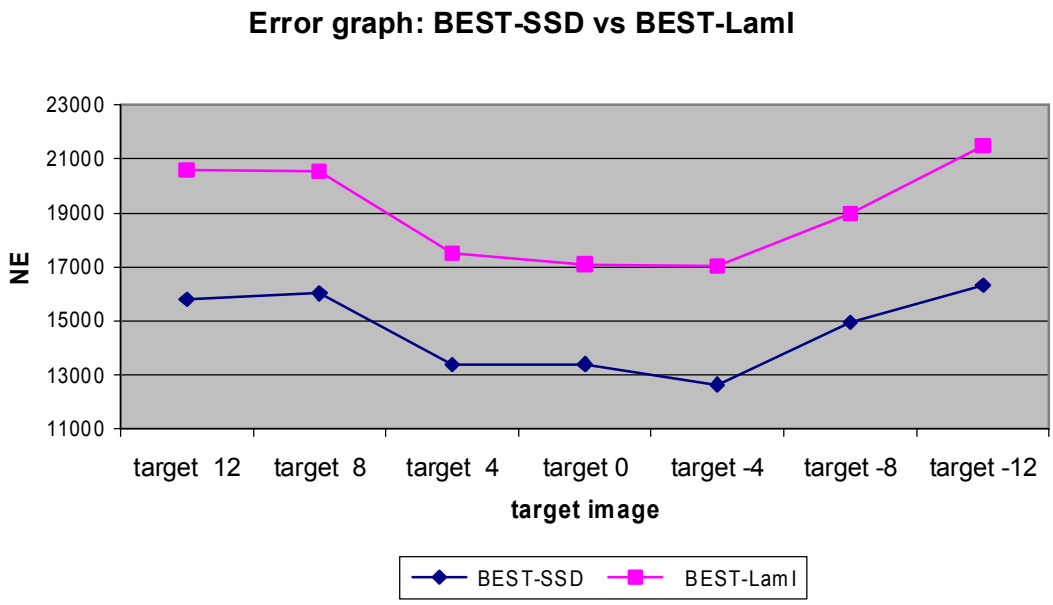
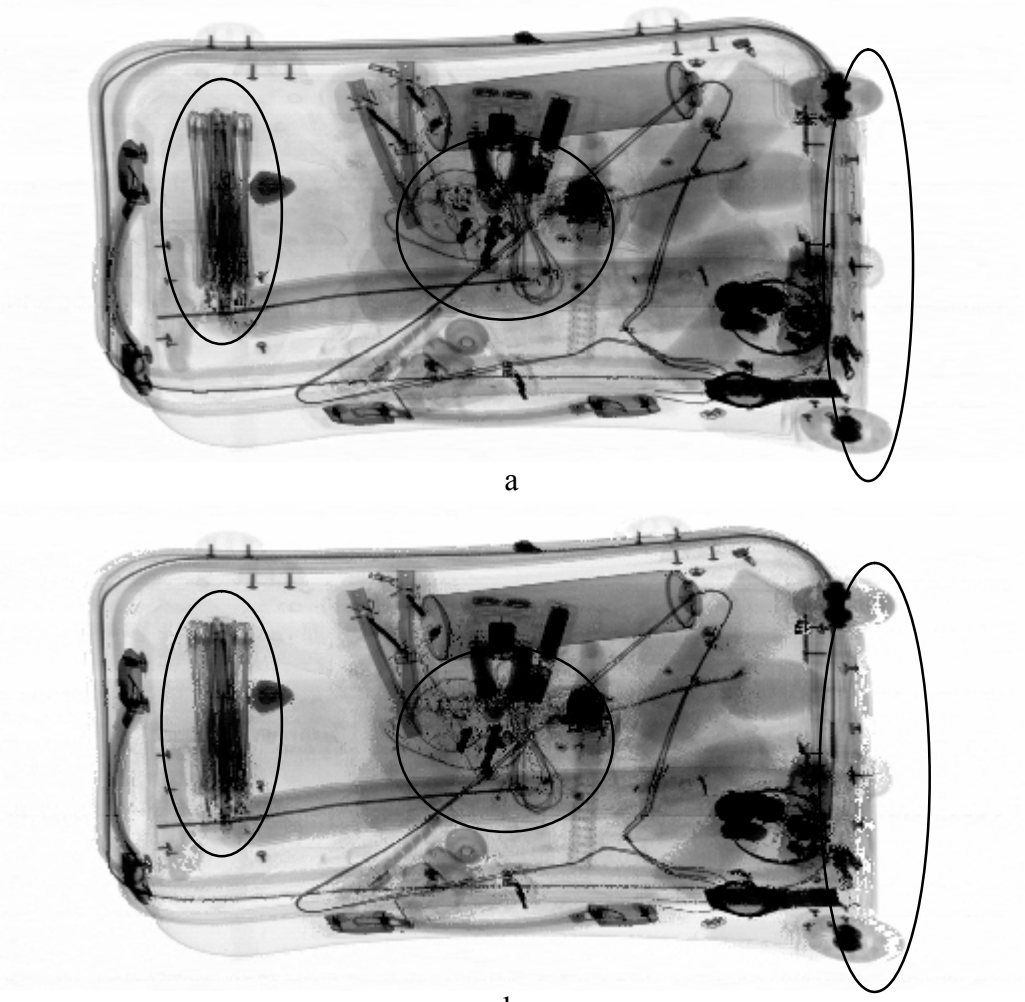


Figure 7.1 Comparisons of the BEST-SSD and BEST-LamI among seven target images.

7.1.2 Result: Images



a

b

Figure 7.2 (a) target image 12^o, created by the BEST-SSD and (b) target image 12^o, created by the BEST-LamI.

7.1.3 Analysis

It can be appreciated from Figure 7.1 and the highlighted image regions in Figure 7.2 that BEST-SSD generally outperforms BEST-LamI; however the following example illustrates the image feature that latter handles better than former.

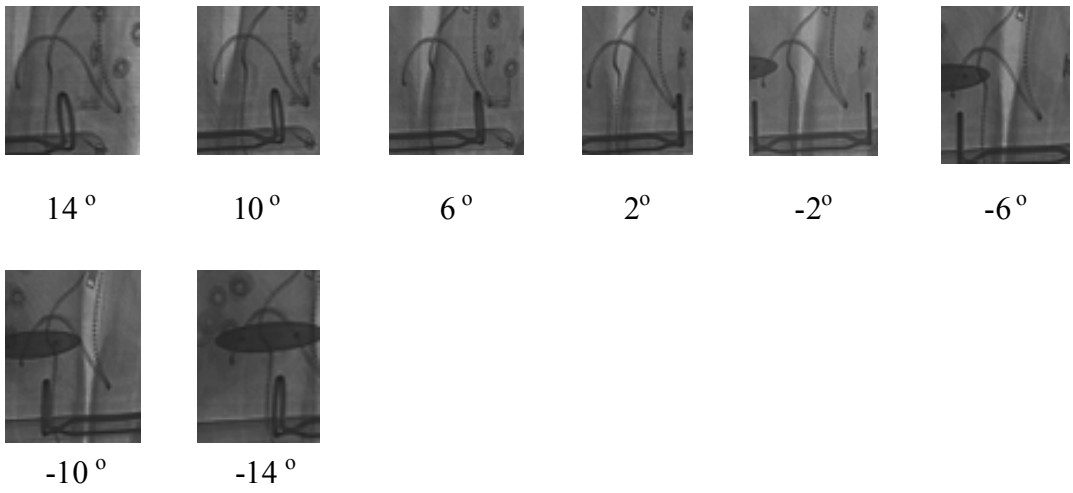


Figure 7.3 Region of interest in input images over the range 14^o to -14^o that are separated by an angular separation of 4^0.(image set 2).

Figure 7.4 Region of interest in the ground truth (GT) at -4^0; generated by the BEST-SSD and the BEST-LamI methods.(image set 2).

It can be observed from Figure 7.3 that, region A is located in a discontinuity area. Consequently the window based SSD criterion does not handle the intersecting linear features well, as can be appreciated from the resultant images presented in Figure 7.4. On the other hand the Improved LamI criterion, which is not window based, performs better than the SSD criterion.

7.2 Establishing the V-DMX

The comparative study reported in Section 7.1 shows that the BEST-SSD has a better overall performance except for overlapping and discontinuity regions. The V-DMX algorithm is investigated to take the advantage afforded by the Improved LamI to improve the BEST-SSD's performance. The success of this improvement relies upon developing a method to search for when and where the Improved LamI outperforms the BEST-SSD. The search result is then used to replace the synthetic pixels produced by the BEST-SSD with the Improved LamI's synthetic pixels. The method adopted in this research programme is based on an iteration process of Improved LamI and a NE

comparison in each iteration. Results from Chapter 5 revealed that the BEST-LamI employs the maximum nInput and nVote to produce the best synthetic images. This finding is based on comparing the final synthetic images produced by all the possible pairings of nInput and nVote. In each cycle of the iteration process, only a number of synthetic pixels are created. By comparing the NE of these pixels, other Improved LamI pairing(s) may outshine the BEST-LamI pairing. Therefore, in order to determine the optimum setting to integrate the Improved LamI with BEST-SSD, the performance of all possible pairings for the Improved LamI criterion should be reinvestigated. It should be noted that the iteration process is solely applied on the Improved LamI criterion while the synthetic images generated by the BEST-SSD is utilised as the NE comparison benchmark. The flowchart presented in Figure 7.5 describes algorithmic processes in this experiment.

The first process is to generate the target image using the BEST-SSD method. Then, for each iteration of the Improved LamI criterion, calculate the NE of the generated synthetic pixels, only includes those pixels whose intensity are assigned in the current iteration. A section of the results is tabulated in Table 7.1 where the complete table for all iterations can be found in Appendix B. The third process is to calculate the NE of the BEST-SSD's synthetic pixels, which have the same x,y position with those synthetic pixels generated by the Improved LamI in the current iteration. Table 7.2 tabulates a section of the result where the complete table is organised in Appendix B. Once all the iterations are completed, the final process compares the NE generated by all the possible pairings of Improved LamI and the BEST-SSD. Table 7.3 tabulates a section of the result where the NE of the BEST-SSD is subtracted from the NE of the Improved LamI, the complete table is organised in Appendix B. (Pair-n-m presents pairing with n input images and m voting images in following tables, and tables in Appendix B.)

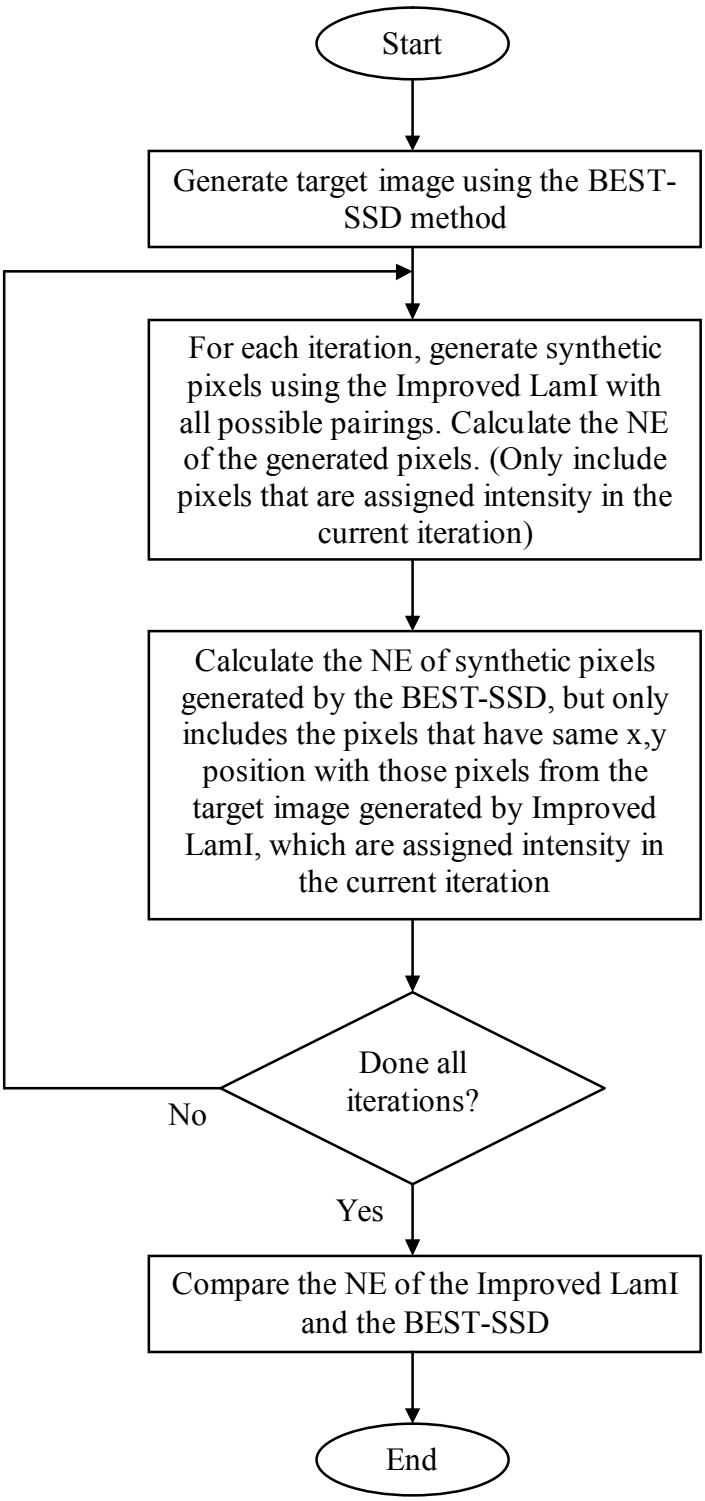


Figure 7.5 Flowchart of the development of V-DMX.

7.2.1 Result: NE tables

	view-2-1	view-2-2	view-3-1	view-3-2	view-3-3	view-4-1	view-4-2	view-4-3	view-4-4	view-5-1	view-5-2	view-5-3	view-5-4	view-5-5
iteration1	20255	2178	2059	1762	1363	1350	1206	1079	982	974	865	798	754	725
iteration2	0	1804	18008	8470	2681	18423	13182	5282	3052	18820	15688	8468	4363	3192
iteration3	0	2148	0	4195	2109	0	3548	3298	2344	0	2584	3736	3408	2631

	view-6-1	view-6-2	view-6-3	view-6-4	view-6-5	view-6-6	view-7-1	view-7-2	view-7-3	view-7-4	view-7-5	view-7-6	view-7-7
iteration1	735	697	660	631	615	604	619	602	575	558	546	541	536
iteration2	19067	16720	10804	5940	3877	3163	19253	17537	12383	7564	4816	3624	3191
iteration3	0	2057	3781	3978	3388	2870	0	1680	3769	4442	3782	3209	2875

	view-8-1	view-8-2	view-8-3	view-8-4	view-8-5	view-8-6	view-8-7	view-8-8
iteration1	542	523	500	484	478	470	465	458
iteration2	19569	18217	13808	9137	5650	4088	3452	3142
iteration3	0	1368	3572	4517	4423	3724	3223	2951

Table 7.1 The NE of the newly synthesised pixels in 3 iterations produced by the Improved LamI criterion using all possible pairings.(target 12^o)

	view-2-1	view-2-2	view-3-1	view-3-2	view-3-3	view-4-1	view-4-2	view-4-3	view-4-4	view-5-1	view-5-2	view-5-3	view-5-4	view-5-5
iteration1	12957	2296	1989	1757	1518	1405	1259	1169	1102	1056	953	889	857	832
iteration2	0	1539	11054	6102	2462	11888	9105	4499	2896	12570	10889	6614	3932	3008
iteration3	0	1749	0	2983	1884	0	2374	2685	2154	0	1778	2836	2982	2455

	view-6-1	view-6-2	view-6-3	view-6-4	view-6-5	view-6-6	view-7-1	view-7-2	view-7-3	view-7-4	view-7-5	view-7-6	view-7-7
iteration1	816	779	745	726	711	701	688	669	644	631	622	620	615
iteration2	13020	11783	8102	5070	3598	3003	13240	12328	9090	6100	4306	3417	3057
iteration3	0	1352	2734	3313	3081	2674	0	1060	2620	3480	3299	2926	2643

	view-8-1	view-8-2	view-8-3	view-8-4	view-8-5	view-8-6	view-8-7	view-8-8
iteration1	602	584	566	554	549	541	536	529
iteration2	13384	12692	9865	6919	4868	3786	3259	2982
iteration3	0	910	2484	3428	3640	3324	2986	2757

Table 7.2 The NE of the BEST-SSD"s synthetic pixels that have the same position with the newly matched pixels in three iterations by the improved LamI pairings. (target 12^o)

	view-2-1	view-2-2	view-3-1	view-3-2	view-3-3	view-4-1	view-4-2	view-4-3	view-4-4	view-5-1	view-5-2	view-5-3	view-5-4	view-5-5
iteration1	-7298	118	-70	-5	155	55	53	90	120	82	88	91	103	107
iteration2	0	-265	-6954	-2368	-219	-6535	-4077	-783	-156	-6250	-4799	-1854	-431	-184
iteration3	0	-399	0	-1212	-225	0	-1174	-613	-190	0	-806	-900	-426	-176

	view-6-1	view-6-2	view-6-3	view-6-4	view-6-5	view-6-6	view-7-1	view-7-2	view-7-3	view-7-4	view-7-5	view-7-6	view-7-7
iteration1	81	82	85	95	96	97	69	67	69	73	76	79	79
iteration2	-6047	-4937	-2702	-870	-279	-160	-6013	-5209	-3293	-1464	-510	-207	-134
iteration3	0	-705	-1047	-665	-307	-196	0	-620	-1149	-962	-483	-283	-232

	view-8-1	view-8-2	view-8-3	view-8-4	view-8-5	view-8-6	view-8-7	view-8-8
iteration1	60	61	66	70	71	71	71	71
iteration2	-6185	-5525	-3943	-2218	-782	-302	-193	-160
iteration3	0	-458	-1088	-1089	-783	-400	-237	-194

Table 7.3 The NE difference between the BEST-SSD and the Improved LamI (NE of the BEST-SSD - NE of the Improved LamI) for newly synthesised pixels for three iterations.(target 12^{o})

7.2.2 Result: Images

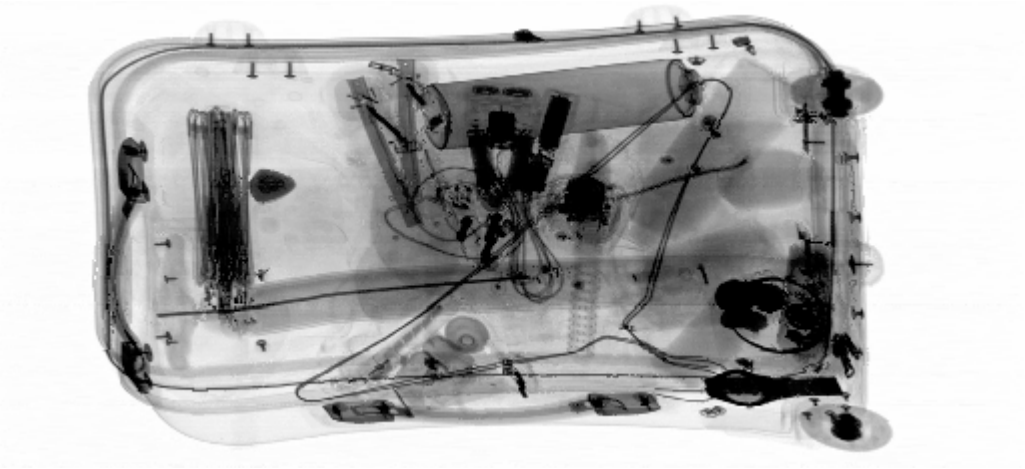


Figure 7.6 Target image 12^o generated by the BEST-SSD method.

Figure 7.7 Target image 12^{o}generated by the Improved LamI using 3-3 pairing in iteration 1.

7.2.3 Analysis

Table 7.3 shows the comparison of the NE produced by the Improved LamI pairings and the BEST-SSD. A positive number denotes that, among the pixels generated of the target image in related iteration, the BEST-SSD generates more error than the Improved LamI, and vice-versa. The objective of this experiment is to explore how Improved LamI can be used to improve the performance of BEST-SSD, so a bigger number in the table indicates a bigger room for improvement. It can be observed that only iteration one contains positive data. This is the case for all possible pairings. This observation indicates that the Improved LamI created less error than the BEST-SSD method only when it operates in the first iteration. Also, it can be deduced from Table 7.3 that the pairing 3-3 has biggest positive number, which means this pairing produces the least NE compare to the NE of BEST-SSD.

In conclusion, the application of Improved LamI using 3-3 pairing in iteration 1 has the potential to improve the quality of target image created by the BEST-SSD. Owing to this, the V-DMX is developed to initially apply the 3-3 pairing of Improved LamI for one iteration to generate a portion of the synthetic image, and followed by using the BEST-SSD method to synthesise the remaining pixels. Figure 7.6 shows the target image 12^o generated by the BEST-SSD method, and Figure 7.7 shows the target image 12^o generated by the Improved LamI using 3-3 pairing in iteration 1.

7.3 Comparative study of V-DMX and BEST-SSD

The objective of this experiment is to compare the performance of the V-DMX and the BEST-SSD.

7.3.1 Result: NE curves

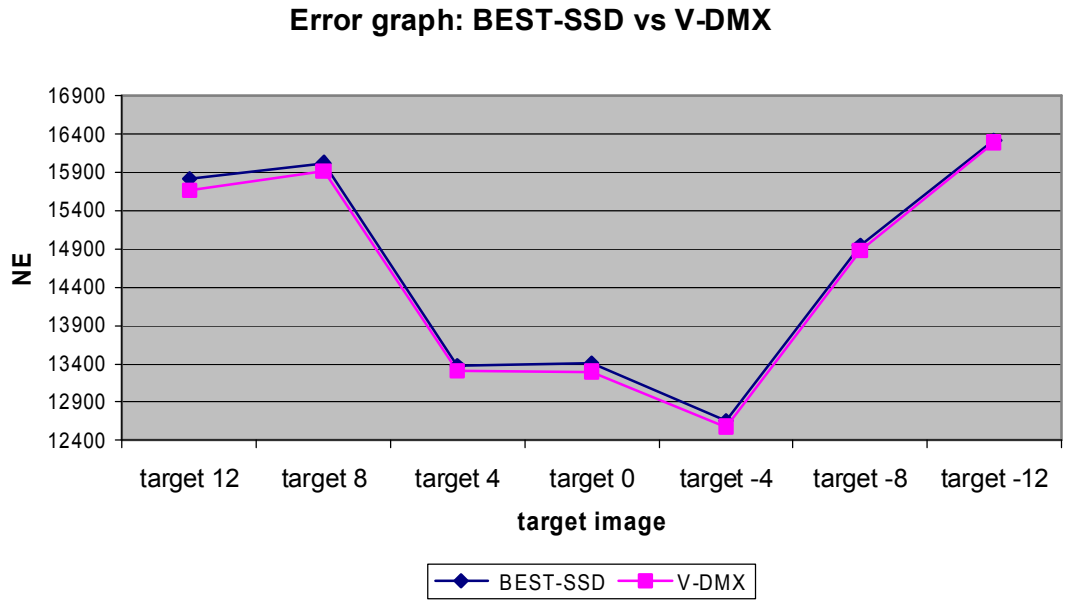


Figure 7.8 The NE comparison of the V-DMX with BEST-SSD generated by seven target images with an angular separation between adjacent detector images of 4°.

7.3.2 Result: Images

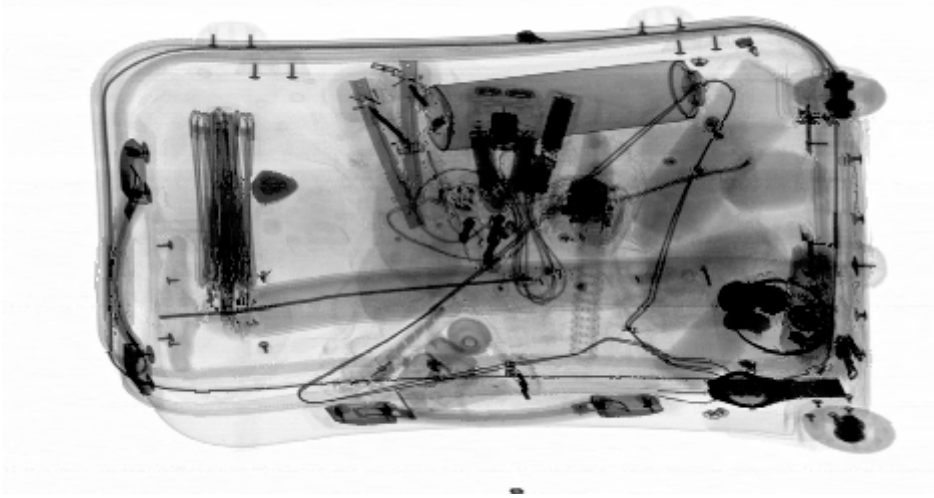


Figure 7.9 Target image 12^o created by the BEST-SSD method.

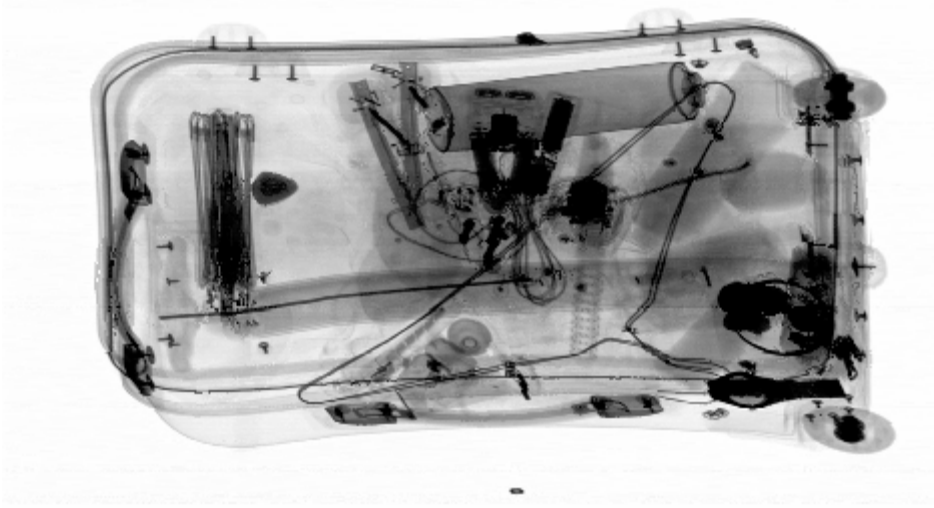
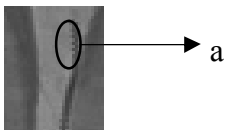


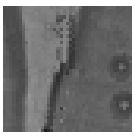
Figure 7.10 Target image 12^o created by the proposed V-DMX method.

7.3.3 Analysis

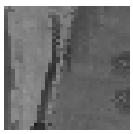
The performance of the proposed V-DMX is compared with BEST-SSD. Figure 7.9 and 7.10 show the target image 12^o generated by BEST-SSD and V-DMX. The NE plots are arranged in Figure 7.8. NE for other angular separation is presented in Appendix C. It can be deduced from the plots that the V-DMX has offered a slight increase in the overall performance over the BEST-SSD method. The improvement can be appreciated by referring to the example reported in the following text.



GT -8°

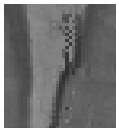


BEST-SSD



Improved LamI

Improved LamI iteration one



Improved LamI iteration one + BEST-SSD

Figure 7.11 Region of interest in GT -8° and in target images generated by BEST-SSD, Improved LamI, Improved LamI iteration one, Improved LamI iteration one+ BEST-SSD.(image set 2).

It is obvious from the Figure 7.11 that the BEST-SSD struggled to handle region-A because region-A is located in discontinuity area. However, the Improved LamI can handle the region-A very well. When the Improved LamI iteration one is employed, region-A is partially generated, offering an improvement to the BEST-SSD method.

7.4 Effect of increasing angular separation

The objective of this experiment is to investigate the performance of the V-DMX algorithm as a function of the angular distribution of the X-ray beams employed to produce the input images.

7.4.1 Result: NE curves

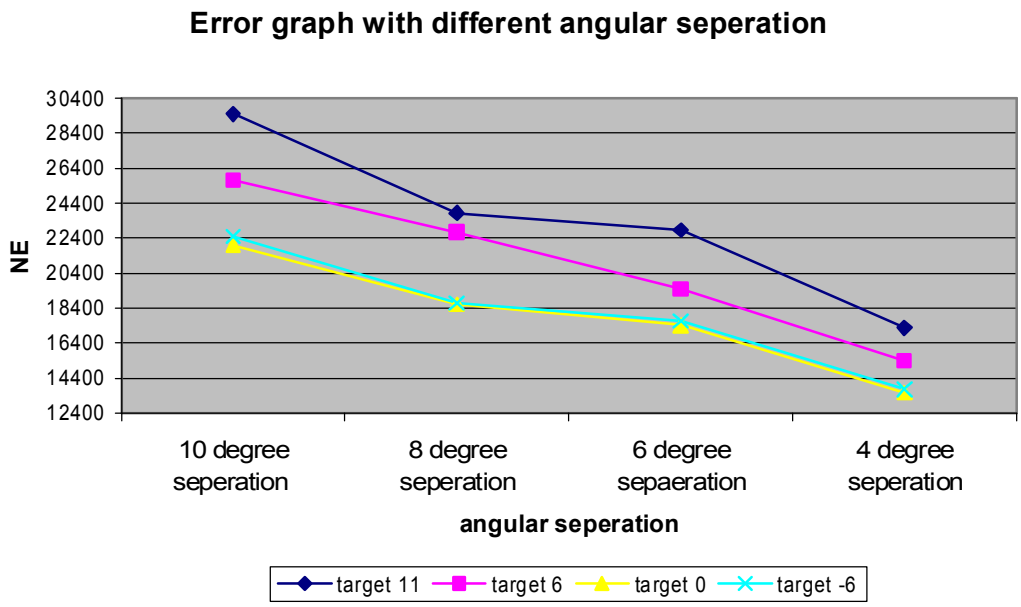


Figure 7.12 Error number of the V-DMX for different angular separations of the X-ray beams

7.4.2 Result: Images

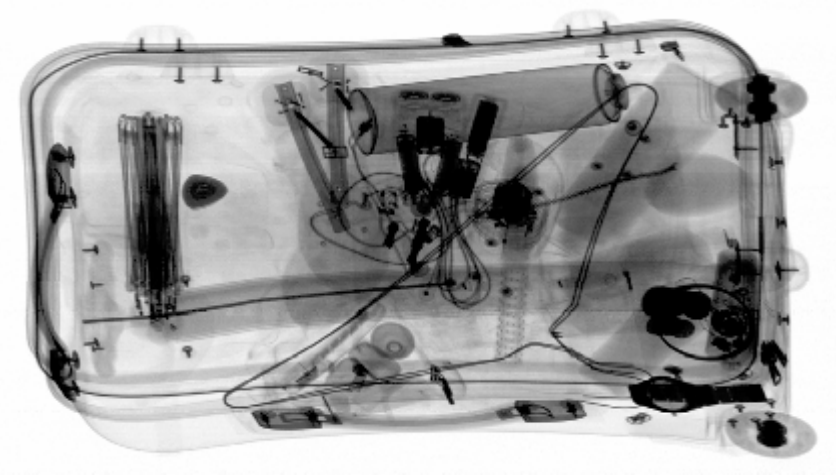
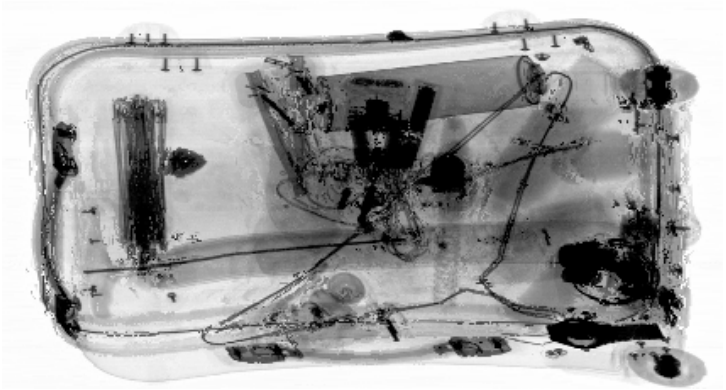
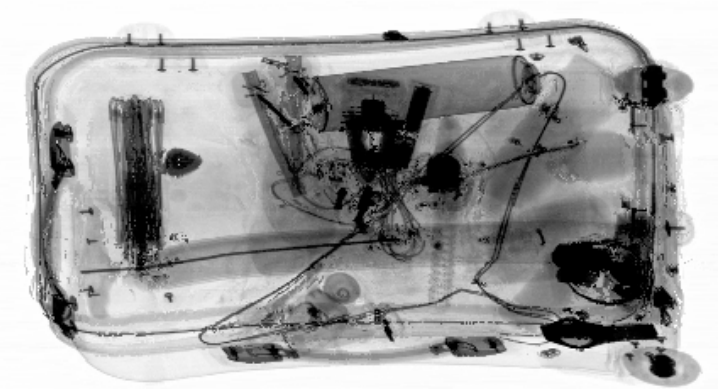


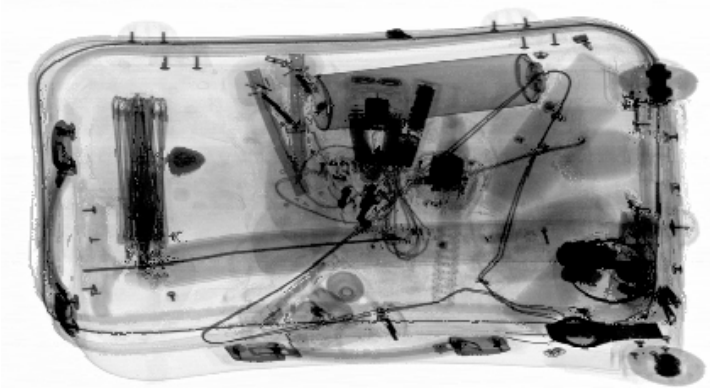
Figure 7.13 GT image 11°.



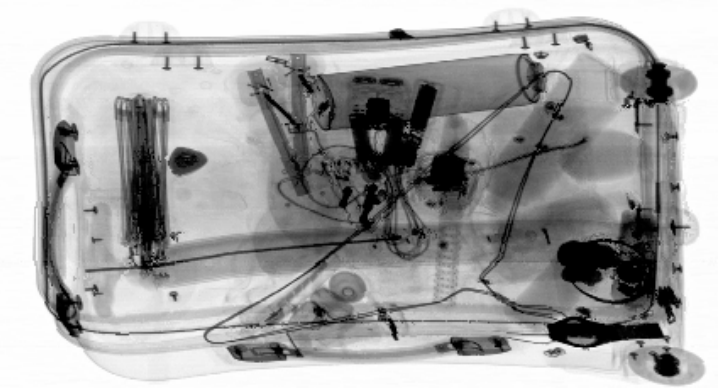
Target 11^{o} created by 10^{o} angular separation



Target 11^{o} created by 8^{o} angular separation



Target 11^{o} created by 6^{o} angular separation



Target 11^{o} created by 4^{o} angular separation

Figure 7.14 Target image generated by V-DMX with different angular separation.

7.4.3 Analysis

As the angles are increased, the resultant imagery becomes increasingly disparate and the performance of the algorithm is decreased as can be observed in Figure 7.12 to Figure 7.14. The visual quality of the synthetic imagery produced at an angular separation of 6^{o} is observed consistently to produce acceptable kinetic imagery; the decreasing quality of the synthetic imagery with increasing X-ray beam angle is also observed. However, the precise relationship between the visual quality of the synthetic imagery and its ultimate utility when incorporated into a sequence views is beyond the scope of this thesis, and is subject to an established human factors research programme, see Chapter 8.

Chapter Eight Summary, conclusion and future work

8.1 Summary

This research programme was instigated to investigate and develop an image synthesis algorithm capable of introducing high quality intermediary images between successive pairs of practical (or detector derived) images in a sequence of transmission X-ray images. The motivation is to establish the technological basis for a practical KDEX imaging system. One of the fundamental requirements of such an image synthesis technique is to develop matching algorithms to solve the correspondence problem. However, the correspondence problem is ill posed, exacerbated by the transparency property of X-ray images, which makes them fundamentally different from visible light images and presents additional considerations and complexities for image synthesis; for example the inapplicability of uniqueness, smoothness that are commonly applied to reflected visible light images. To address these issues, existing and new matching techniques have been proposed and integrated with new algorithmic procedures in an attempt to meet the overall aim of image synthesis in a KDEX scanning environment. Here the narrow angular separation of the order of 1º between successive views in the image sequences are dictated by constraints concerning perception of smooth motion, kinetic depth effect and binocular stereoscopic fusion. To this end it has been demonstrated that intermediary images may be successfully synthesised by processing adjacent perspective images produced by X-ray beams separated by angular increments of 6º. Therefore a time multiplexed sequence alternating between synthetic and X-ray sensor views may be produced, a key requirement if commercial machines are to be produced.

The conventional SSD matching criterion has been experimentally evaluated and found to degrade significantly in discontinuity and overlap image regions but perform well for other regions. To address this limitation, a new criterion LamI based on laminography has been developed. Referenced to ground truth images, improvements are seen with this technique where SSD fails; although in the general case the technique underperforms SSD. Further, through the use of information from multiple images significant enhancements can be made. This study will enable further refinement to rationalise and optimise the current algorithmic approach, and is of practical significance when considering it is to be employed in human factors studies, funded by

the US Department of Homeland Security (DHS) Human Factors Programme, under an established rolling grant entitled, “kinetic depth imaging for security screening”.

The additional information made available by the multiple views utilised in this research programme is usefully implemented to enhance the treatment of repeating features and the handling of discontinuity and overlapping features.

8.1.1 Algorithm

The algorithm employs matching cost computation, voting for disparity, synthetic pixel generation and void filling. Images are categorised into five groups: detector images, target images, reference images, input images, and voting images. Different images play different roles in the process.

In stereo matching, for one pixel with position (x,y) in image I_1, the matching cost calculation only requires to consider the “other” perspective image, so each disparity has one matching cost. However, with multiple images, there are multiple matching costs and each disparity value has several matching costs; these costs require aggregation to obtain the final matching cost. Two matching cost computation methods are proposed in this project: laminography and sum of reciprocals.

The concept of summing multiple images, with appropriate geometric registration, to produce an image of a planar section of an object is the basis of laminography. By appropriate processing of the input images, a set of laminograms, representing contiguous depth planes is produced. The laminogram intensity is the average value of the intensities contributed by all the images. The resultant matching cost calculation is equal to summing, with equal weighting, all the matching costs for a given disparity. The outcome of such a scheme is the loss of details concerning the individual matching costs. To preferentially emphasize low matching costs, the sum of reciprocals is proposed. The number of input images included in the process is *nInput*.

A voting mechanism is proposed to perform multiple matching consistency voting using the multiple detector images. The number of voting images is equal to *nVote*, where $nInput \supseteq nVote$. A voting process is implemented after the matching cost computation stage. This process performs first a rank order procedure to the matching cost values before summing those values together.

Once the disparities are determined, both intensity and position (x,y) of the synthetic pixels may be computed. The process begins with employing the disparities of the first reference image to synthesise intermediate target pixels. The disparities of each pixel in the first reference image enabled the algorithm to determine its correspondence in the second reference image. Upon the completion of using all pixels of the first reference image for image synthesis, the algorithm can identify which pixels in the second reference image are not matched by the pixels of the first reference image. The algorithm will subsequently utilise the disparities of the unmatched pixels to synthesise target pixels that are not generated by the first reference image. The latter process is designed to partially cope with the multiple correspondences property in X-ray images. In other words, the latter process enabled some pixels in the first reference image to have two disparities (i.e. two correspondences). While this method only provided a partial solution, the empirical evidence has supported the rationale of the method.

After the synthetic pixel generation process, it is still possible that some pixels in the target image have not been assigned an intensity value; such "empty" pixels are termed voids. The possible causes of voids include errors in the disparity information table and inadequate treatment of multiple correspondences. The final process assigns intensity values to the voids by look back to it's possible correspondent point in the first reference image, then select the point that have maximum score that correspond to the disparity that move to the void.

8.1.2 Algorithm criteria

To calculate the matching cost, appropriate criteria have been developed and applied. A novel criterion based on laminogram intensity is proposed in this project. This criterion utilises depth information contained in laminograms. The criterion exploits a basic property of laminograms concerning the reduction in intensity values which comprise the in focus portions of the laminogram. Simplistically stated, the "in focus" pixels exhibit the lowest relative intensities in comparison to the pixels with same position in other laminograms. This criterion is not good at handling pixels when there are other pixels that have lower intensity value around it. An improved version is developed to dynamically select input images base on iteration level to ease the effect of low intensity value pixels on high intensity value pixels. Experiments are conducted to explore laminogram intensity criterion, how many input images and voting images have best performance.

The SSD criterion is well known in stereo matching world for its utility and reliability. Experiments have been conducted to explore all the possible combinations of input images, voting images, matching cost calculation method, experiments show that, for SSD, three input images with two voting images and SOR matching cost calculation exhibits the best overall performance in comparison to other combinations.

A comparative study of SSD and improved LamI is reported. SSD produces overall better performance in comparison to Improved LamI. However, it was observed that for some features, such as discontinuity areas and overlapping regions, Improved LamI outperforms SSD. The final version of the algorithm termed V-DMX is designed to combine the SSD criterion and the improved laminogram intensity criterion.

Independently, each criterion offers different but important solutions to address the problem. The criteria are implemented as a complementary pair, mutually reinforcing the performance of the algorithm. Each criterion is run independently to generate disparity information for reference images, and the disparity information is integrated together in the synthetic pixel generation process.

8.2 Conclusions

In terms of the original aims and objectives of this research program, there has been significant progress. An algorithm has been developed that takes practical multiple-view X-ray images and generates synthesised intermediate views.

The experiments reported investigate the fidelity of the synthesised images as a function of the number of input images, number of voting images, criteria, and X-ray beam angular separation. A performance error number, based on error with respect to a ground truth image, indicates the quality of synthesised images.

Within the constraints of the experimental approach taken, these appear to support the feasibility of generating intermediary images from practical images with an angular separation of up to 6^{o}. Since the notional requirement for image separation is 1^{o}, this holds the prospect of a substantial reduction in hardware complexity. Such encouraging results lead to the conclusion that this work should form the basis of further investigation and optimization outlined in the following section.

One of difficulties encountered in the work undertaken is deriving a meaningful criteron to quantify the fidelity of the synthesized image. In the absence of access to more sophisticated human factors related assessment (ultimately the image is expected to be viewed by a human operator), the techniques employed involved a pixel error count and a qualitative visual inspection with reference to a ground truth practical image. These may not be the most applicable indicators when the image is viewed in the context of an operating KDEX system.

It should also be noted that whilst good results have been obtained using the intensity of reference image to synthesis target images, this work has focused on the correspondence problem thus leaving intensity determination an area for further work.

The development of the final algorithm involved examination of various techniques independently from which separate conclusions can be drawn.

8.2.1 Laminogram intensity criterion

The laminogram intensity criterion, designed to extract in focus content from laminograms, here the empirical evidence indicates that this criterion performs best with the maximum number of input images and the maximum number of voting images. Increasing the number of input images tends to produce improved laminographic data and produce less ambiguous results. With more voting images, more images are involved in the disparity consistency voting.

In the context of this work the laminogram criterion is not reliable for the processing of relatively high intensity pixel values when lower intensity image structures produce “smearing” artefacts in the “processing vicinity”. To accommodate this scenario the improved version of laminogram intensity criterion dynamically selects input images on each iteration to create new laminograms, the result has been significantly improved.

8.2.2 SSD criterion

SSD is successfully applied by the synthetic algorithm. SSD performs best with three input images and two voting images with SOR matching cost calculation. However, it does not perform well for repeating features, overlapping and discontinuity regions, although with three input images and two voting images, performance enhancement can

be observed in comparison to standard stereo matching, which utilises two input images and one voting image.

8.2.3 Voting based dual criteria multiple X-ray images synthesis (V-DMX)

A comparison of the best performance pair of SSD and the best performance pair of Improved LamI reveals that the SSD has better performance than the improved laminogram intensity criterion. Interestingly, for some features exhibiting discontinuity and overlap regions, the improved laminogram intensity criterion performs better than SSD. A voting based dual criteria multiple X-ray images synthesis (V-DMX) has been successfully developed to exploit the complimentary advantages offered by implementing both criteria. Various implementations of the two criteria are investigated; experiment results indicate that running improved laminogram intensity criterion with three input images, three voting images, with iteration one followed by SSD employing three input images and two voting images has the best performance.

The performance of V-DMX is compared with the best performance of SSD; experiment results show that, V-DMX has enhanced performance relative to the lone use of the SSD criterion.

8.2.4 Final conclusion

Empirical evidence has demonstrated that intermediary images may be successfully synthesised by processing adjacent detector images produced by X-ray beams separated by angular increments of up to 6º. Therefore, a time multiplexed sequence alternating between synthetic and X-ray sensor views may be produced. This encouraging result has formed a basis for further research to optimise the current algorithmic approach using dual-energy X-ray data.

8.3 Future work

The approach developed in this research is still in its infancy. A number of interdependent areas require further development work to improve the image synthesis techniques presented in this research. The future work is proposed as follows.

1. **Improvement in matching algorithm using dual-energy X-ray data**

   Measurements taken in two different X-ray energy windows will produce two sets of data, i.e, low-energy and high-energy X-ray data, which is the basis for

materials discrimination. It is hypothesised that the independent and or combined treatment of the low-energy and the high-energy X-ray data could enhance the algorithmic approach developed in this research programme. An investigation is planned to establish the efficacy of this hypothesis.

2. **Human factors evaluation to support algorithm development**

   The success of the approach adopted is dependent upon the algorithms being of sufficiently high quality to not destroy the flow of 3D visual information to the observer. The validity of such work is dependent upon psychological human factors considerations. It is planned to test future imagery in collaboration with the US Dept. of Homeland Security (DHS) within the context of an established funded collaboration. Thus image sequences containing X-ray sensor images and synthetic images can be used to establish if kinetic and stereo depth cues are adversely affected by the inclusion of synthetic images.

3. **Extension of the synthesis algorithm to colour encoded X-ray images**

   The future development of the image synthesis technique to compute intermediary views is to be extended to the colour encoded dual-energy (materials discriminating) X-ray data produced by the experimental machine. A novel approach utilising basis material subtraction calibration is planned in collaboration with the HOSDB. The future development of the image synthesis technique will be extended to examine the problem of producing matched colour images by applying basis material subtraction (BMS) calibration [8]. The benefit of including the BMS calculations will be theoretically assessed and if appropriate empirically demonstrated. The performance of the algorithms will continue to be tested by quantitative comparison of the computed images with the sensor views being replaced.

4. **Application of AI techniques to adaptively select images for processing**

   The current algorithm depends on conducting a series of experiments to predetermine the optimum number of input images for the image synthesis process. This approach may limit the full potential of the algorithm during its actual field test. The solution to this limitation is to investigate various AI techniques to adaptively select images for processing.

5. **Improvement in intensity assignment**

   Selecting the correct intensity to generate a target image is very important. This research programme has adopted an approximation, but effective, approach to assign synthetic pixels with the intensity of its correspondence points in the

reference images. Thus, there is scope to seek for a quantitative method to decide the intensity of target image using more sophisticated algorithms.

6. **Choose more than one disparity for one pixel based on the rank voting values**

   The developed algorithm employs the disparity information of the second reference images and the void filling process, in an attempt, to partially solve the multiple correspondence problems associated with X-ray images. The practical limitation of the proposed method is the selection of only one "winning" disparity for each pixel under consideration. Since the rank voting values are correlated with the probability of the disparities, a statistical method may be exploited to choose more than one disparities for each pixel.

7. **Laminogram pre-processing**

   Pre-processing of the laminograms to reduce the effect of focus artefacts has the potential to significantly improve the current algorithm.

Reference

1. Evans J.P.O, Robinson M, “Line-scan imaging in 3D”, 2000, Patent Int. pub. No. WO 99/17555.

2. Hon HW, Evans J.P.O, “Multiple view line-scan imaging”, 2002, IEE Proc. Optoelectronics, 149, 2, 45-50.

3. Evans J.P.O, Hon HW, “Dynamic stereoscopic X-ray imaging”, 2002, Journal of NDT&E, 35, 337-345.

4. Evans J.P.O, Robinson M, “Design of a stereo. X-ray imaging system using a single X-ray source", 2000, Journal of NDT&E, 33, 5, 325-32.

5. Hon H.W, “The modelling of multiple beam X-ray systems using visible light”, 2000, NTU, PhD Thesis.

6. Proffitt D. R., Hecht H., Rock T. and Schubert J., “Stereo Kinetic Effect and its Relation to the Kinetic Depth Effect”, Journal of Experimental Psychology: Human Perception and Performance”, volume 18, No.1, pp. 3-21, 1992..

7. Evans J.P.O, “Stereoscopic imaging using folded linear dual-energy X-ray detectors”, 2002, IOP Journal Meas. Sc. & Tech, 13, 1388-1397.

8. Wang T.W, Evans J.P.O, “Materials identification using a stereoscopic dual-energy X-ray system”, 2003, accepted for IEE Proc. VISP.

9. Evans J.P.O, Robinson M, “Detecting, improving and characterising material in 3D space”, UK Patent No. 232 9817.

10. Wang T.W, “Target materials discrimination using stereoscopic dual-energy X-ray images”, 2002, NTU, PhD Thesis.

11. Seitz S.M and Dyer C.R, “Physically-valid view synthesis by image interpolation”, 1995, In Proc. IEEE Workshop on Representations of Visual Scenes, 18-25.

12. Chen, S.E and Willams, L, “View interpolation for image synthesis”, 1993, Proc. SIGGRAPH 93, In Computer Graphics, 279-288.

13. Evans J.P.O, Robinson M and Godber S X, “A new stereoscopic x-ray imaging technique using a single x-ray source: theoretical analysis “, Non-Destructive Testing Evaluation Int. 29 27–35,1996.

14. Evans J.P.O and Robinson M, "Design of a stereoscopic X-ray imaging system using a single X-ray source" 2000, J. NDT&E Int, 33, 5, 325-332

15. A. S. Sobania, Dipl.Ing, “The automatic extraction of 3D information from stereoscopic dual-energy X-ray images”, 2003, NTU, PhD Thesis.

16. A. Michette, S. Pfauntsch, "X-rays - The First Hundred Years", John Wiley & Sons, March 1997.

17. D.G. Brown, R.F. Wagner, "Physics and Statistics of Medical Imaging", Journal of Digital Imaging, Vol. 2(4), pp. 194-211, November 1989.

18. K. Mori, L. Hasegawa, J. Toriwaki, H. Anno, K. Katada, "Recognition of Bronchus in Three-dimensional X-ray CT Images with Application to Virtualised Brochoscopy System" Pattern Recognition, Proceedings of the 13th International Conference, Vol. 3, pp. 528-532, 1996.

19. D.G. Grown, R.J. Jennings, "Detection and Discrimination Using X-radiation", Neural Networks International Conference, Vol. 4, pp. 2503-2506, 1997.

20. C. Neubauer, "Intelligent X-ray Inspection for Quality Control of Solder Joint", IEEE Transactions on Component Packing and Manufacturing Technology, Part C, Vol. 20(1), pp. 111-120, April 1997.

21. M. Annis, P. Bjorkholm, D. Schafer, "Automatic Detection of Explosives Using X-ray Imaging", Access Security Screening: Challenges & Solutions, ASTM STP 1127, pp. 68-81, 1992.

22. Evans J.P.O, M. Robinson, S.X. Godber, "The Development of 3D

(Stereoscopic) Imaging Systems for Security Applications", 29th IEEE International Carnahan Conference, pp. 505-511, October 1995.

23. N.C. Murray, R.J. Lacey, P.H. Mason, "Exploitation of X-ray Technology for the Detection of Contraband - Aviation Security Applications", European Conference on Security and Detection, Conference Publication No. 437, pp. 13-18, April 1997.

24. R. Benjamiin, S. Prakoonwit, "Automatic X-ray Screening of Aircraft Hold Luggage", European Conference on Security and Detection, Conference Publication No. 437, pp. 5-9, April 1997.

25. Evans J.P.O, Robinson. M, “The Development of 3D X-ray equipment for security Systems”, Contraband and Cargo Inspection Technology International Symposium, Sponsored by: Office of National Drug Control Policy and National Institute of Justice, pp.99-103, Washington, D.C, U.S.A, 28-30 Oct. 1992.

26. Evans J.P.O, Robinson. M, “The development of 3D X-ray systems for airport security applications”, SPIE Proceeding: Applications of Signal and Image Processing in Explosives Detection Systems, Vol. 1824, pp. 171-182, Boston Massachusetts, U.S.A, 16-17 Nov. 1992.

27. Evans J.P.O, Godber S.X, Robinson. M, “Stereoscopic X-ray systems for airport security applications”, 4th European Workshop on Three-Dimensional Television, pp. 97-101, Rome, Italy, Oct. 1993.

28. Evans J.P.O, Godber S.X, Robinson. M, “Three-dimensional X-ray imaging techniques”, SPIE Proceeding: Symposium on Electronic Imaging Science and Technology in Stereoscopic Displays and Applications V, Vol. 2177, pp. 161-165, California, U.S.A, Feb. 1994.

29. Evans J.P.O, Robinson. M, Godber S.X, “A new stereoscopic imaging technique using a single X-ray source: theoretical analysis”, Journal of Non-Destructive Testing and Evaluation International (NDT&E), pp. 27-35, Feb. 1996.

30. E. E. Christensen, T. S. Curry, and J. Nunnally, “An Introduction to the Physics of Diagnostic Radiology”. Philadelphia: Lee & Febiger, 1972.

31. Laprade G.L, “The Manual of Photogrammetry”, Fourth Edition, American Society of Photogrammetry, pp. 519-544, 1980.

32. R. Spottiswoode, N.L. Spottiswoode, C. Smith, "Basic principles of the three-dimensional film", Journal of the SMPTE, Vol. 59, pp. 249-286, October 1952

33. A. Valyus, "Stereoscopy", The Focal Press, 1966.

34. J.W. Chan, Evans J.O.P, Y.S. Yong, "The Development of an Integrated Stereoscopic Dual-energy X-ray Camera", PREP 2002 Fourth Conference, April 2002.

35. L. Tychsen, "Adler’s Physiology of the Eye", Ninth Edition, Mosby-year Book Inc, pp. 773-853, 1992.

36. Evans J.P.O, “Development of a 3-D X-ray system”, Nottingham Trent University, PhD Thesis, 1993.

37. Evans J.P.O, Robinson. M, “A binocular stereoscopic X-ray imaging technique using folded array linear X-ray detectors”, Institute of Physics: Journal of Measurement Science and Technology, No. 13, pp. 1388-1397, July 2002.

38. Friend. D.B, Jones. A, “A stereoscopic television system for reactor inspection”, C.E.G.B. Research Division, March 1980.

39. Jones. A, “Some theoretical aspects of the design of stereoscopic television systems”, C.E.G.B. Research Division, March 1980.

40. Chan J. W, Evans J. P. O, Yong Y. S. and Monteith A, “Wire transfer function analysis for castellated dual-energy X-ray detectors”, Applied Optics, Optical Society of America, Vol. 43, Issue 35, pp. 6413-6420, December 2004.

41. Evans J. P. O, Liu Y, Chan J. W, “Depth from motion 3D X-ray imaging for

security screening", IEE International symposium on Imaging for crime detection and prevention (ICDP 2005), pp. 5-8, London, ISBN 0-86431-535-0, June 2005.

42. HW, Evans J.P.O, "Multiple view line-scan imaging", IEE Proc. Optoelectronics, 149, 2, 45-50, 2002.

43. Evans J. P. O. and Robinson M, "3D Imaging with line-scanning", Patent PCT: WO 99/17555, 1999.

44. Evans J.P.O, "Kinetic depth effect X-ray imaging for security screening", IEE: Visual Information Engineering VIE 2003, Univ. of Surrey, IEE 495, pp. 69-73, ISSN 0537-9989, July 2003.

45. Evans J.P.O. and Robinson M, "Detecting improving and characterising material in 3D space", Patent GB 2360685, 2001.

46. Z. des Plantes, Eine neue Methode zur Differenzierung in der Rontgenographie, Acta Radio., vol. 13 (1932) 182.

47. A. R. Kalukin and V. Sankaran, "Three-dimensional visualization of multilayered assemblies using x-ray laminography," IEEE Trans. Comp., Packag., Manufact. Technol. A, vol. 20, pp. 361–366, Sept. 1997.

48. S. Gondrom and S. Schropfer, "Digital computed laminography and tomosynthesis—Functional principles and industrial applications," in Proc. Int. Symp. Comput. Tomography Ind. Applicat. Image Process. Radiology, Berlin, Germany, Mar. 1999.

49. S. M. Rooks, B. Benhabib, and K. C. Smith, "Development of an inspection process for ball-grid-array technology using scanned-beam x-ray laminography," IEEE Trans. Comp., Packag., Manufact. Technol. A, vol. 18, pp. 851–861, Dec. 1995.

50. Raparia, D, Alessi, J, Kponou, A, "The Algebraic Reconstruction Technique(ART)", American Physical Society, Particle Acceleration Meeting,

May 12-16, 1997, abstract #2P.57, May 1997

51. Dobbins, James T., III; Godfrey, Devon J, "Digital x-ray tomosynthesis: current state of the art and clinical potential", Physics in Medicine and Biology, Volume 48, Issue 19, pp. R65-R106 (2003).

52. Reiter, H. Maisl, M. Zhou, J. Gondrom, S, "Computed laminography a X-ray method for image analysis of solder joints", IEE Colloquium on Techniques for the Inspection of Bonded Structures, 1997.

53. S. Mattoccia, F. Tombari, and L. Di Stefano, "Stereo vision enabling precise border localization within a scanline optimization framework", ACCV 2007.

54. O. Woodford, P. Torr, I. Reid, and A. Fitzgibbon, "Global stereo reconstruction under second order smoothness priors", CVPR 2008.

55. D. Min and K. Sohn, "Cost aggregation and occlusion handling with WLS in stereo matching", IEEE TIP 2008.

56. F. Tombari, S. Mattoccia, and L. Di Stefano, "Segmentation-based adaptive support for accurate stereo correspondence", PSIVT 2007.

57. S. Mattoccia, F. Tombari, and L. Di Stefano, "Stereo vision enabling precise border localization within a scanline optimization framework", ACCV 2007.

58. Scharstein D, Szeliski R, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", 2002, IJCV, vol. 47, no. 1/2/3, pp. 7-42.

59. Dhond U, Aggarwal J, "Structure from stereo | a review", 1989, IEEE Transactions on Systems, Man and Cybernetics, 19(6).

60. O'Neill M, Denos M, "Automated system for coarse-to-fine pyramidal area correlation stereo matching", 1996, Image and Vision Computing, 14(3):225-236.

61. Kanade T, Okutomi M, "A Stereo Matching Algorithm with an Adaptive

Window: Theory and Experiments", 1994, IEEE Pattern Analysis and Machine Intelligence, vol. 16, no. 9, pp. 920–932.

62. Boykov Y, Veksler O, Zabih R, "A Variable Window Approach to Early Vision", 1998, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 12, pp. 1283–1294.

63. Veksler O, "Stereo Correspondence with Compact Windows via Minimum Ratio Cycle", 2002, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 12, pp. 1654–1660.

64. Veksler O, "Fast Variable Window for Stereo Correspondence using Integral Images", 2003, IEEE Conf. Computer Vision and Pattern Recognition Proc, vol. 1, pp. 556–561.

65. Lotti J.L, Giraudon G, "Adaptive window algorithm for aerial image stereo", 1994, International Conference on Pattern Recognition Proc, volume A, pp. 701–703.

66. Kanade, T, Okutomi.M, "A stereo matching algorithm with an adaptive window: Theory and experiment", 1994, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 9.

67. Kostkova, J, Sara, R, "Stratified Dense Matching for Stereopsis in Complex Scenes", 2003, BMVC Proc, pp. 339-348.

68. Geiger D, Ladendorf , B, Yuille A, "Occlusions and binocular stereo", 1992, European Conference on Computer Vision Proc, pp. 425–433.

69. Fusiello A, Roberto V, Trucco E, "Efficient stereo with multiple windowing", 1997, IEEE Conference on Computer Vision and Pattern Recognition Proc, pp. 858–863.

70. Satyajit A A, Nasser K, Mihai N, "Stereo matching via selective multiple windows", 2007, Journal of Electronic Imaging, vol. 16, issue 1, p. 013012.

71. Kang S. B, Szeliski R, Jinxjang C, “Handling Occlusions in Dense Multi-View Stereo”, 2001, IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 103–110.

72. Vekseler O, “Stereo matching by compact windows via minimum ratio cycle”, 2001, International Conference on Computer Vision Proc, volume I, pp. 540-547.

73. Bleyer, M. Gelautz, M., “A Layered Stereo Algorithm Using Image Segmentation and Global Visibility Constraints”, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 59, nr. 3, pp. 128-150, 2005.

74. Xu Y, Wang D, Feng T, Shum H.Y, “Stereo Computation using Radial Adaptive Windows”, 2002, Pattern Recognition Proc, vol. 3, pp. 595–598.

75. Bhat D.N, Nayar S.K, “Ordinal Measures for Visual Correspondence”, 1996, Columbia Univ, Computer Science, tech. rep. CUCS- 009-96.

76. Dinkar N. Bhat ,Shree K. Nayar, “Ordinal Measures for Image Correspondence”, 1998, IEEE Transactions on pattern analysis and machine intelligence, VOL. 20, NO. 4.

77. Y. Shan and Z. Zhang, “Corner guided curve matching and its application to scene reconstruction”, In Proc. IEEE conference on Computer Vision and Pattern Recognition, pp I:796–803, 2000.

78. T. Tuytelaars and L. Van Gool, “Wide Baseline Stereo based on Local, Affinely invariant Regions”, British Machine Vision Conference, pp. 412-422, 2000.

79. K.Mikolajczyk and C.Schmid, “An affine invariant interest point detector”, ECCV, vol. 1, 128–142, 2002.

80. A. Baumberg, “Reliable Feature Matching Across Widely Separated Views”, Intl. Conf. on Comp, Vis, pp. 774-781, 2000.

81. Birchfield S, and Tomasi C, “Depth discontinuities by pixel-to-pixel stereo”, 1999, International Journal of Computer Vision, 35(3):269–293.

82. Scharstein D, Szeliski R, “Stereo matching with nonlinear diffusion”, 1998, International Journal of Computer Vision, 28(2):155–174.

83. Birchfield S, Tomasi C, “Multiway cut for stereo and motion with slanted surfaces”, 1999, the 7th Int. Conf. on Computer Vision, Kerkyra, Greece Proc, pp. 489–495.

84. M. Bleyer and M. Gelautz, “A layered stereo algorithm using image segmentation and global visibility constraints”, ICIP 2004

85. Y. Taguchi, B. Wilburn, and L. Zitnick, “Stereo reconstruction with mixed pixels using adaptive over-segmentation”, CVPR 2008.

86. L. Hong and G. Chen, “Segment-based stereo matching using graph cuts”, In Proc. CVPR, volume 1, 2004.

87. Y.Wei and L. Quan, “Region-based progressive stereo matching”, In Proc. CVPR, volume 1, pp 106–113, 2004.

88. H. Hirschmüller, “Stereo vision in structured environments by consistent semi-global matching”, CVPR 2006, PAMI 30(2):328-341, 2008.

89. H. Tao, H.S. Sawhney, and R. Kumar, “A Global Matching Framework for Stereo Computation,” Proc. Int’l Conf. Computer Vision, 2001

90. V. Kolmogorov and R. Zabih, “Computing visual correspondence with occlusions using graph cuts”, In International Conference for Computer Vision, pp508.515, 2001.

91. Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts”, PAMI, 23(11):1222–1239, 2001.

92. V. Kolmogorov and R. Zabih, “Multi-camera scene reconstruction via graph-cuts”, In ECCV 2002, volume 3, pp 82–96, 2002.

93. O. Veksler, “Stereo correspondence by dynamic programming on a tree”, CVPR

2005.

94. L. Wang, M. Liao, M. Gong, R. Yang, and D. Nistér, “High-quality real-time stereo using adaptive cost aggregation and dynamic programming”, 3DPVT 2006.

95. O. Faugeras and R. Keriven, “Complete dense stereovision using level set methods,” in Proc. 5th European Conf. on Computer Vision, pp. 379–393, 1998.

96. R. T. Collins, “A space-sweep approach to true multi-image matching,” in Proc. Computer Vision and Pattern Recognition Conf, pp. 358–363, 1996.

97. S. M. Seitz and C. R. Dyer, “Photorealistic scene reconstruction by voxel coloring,” in Proc. Computer Vision and Pattern Recognition Conf, pp. 1067–1073, 1997.

98. P. Fua and Y. G. Leclerc, “Object-centered surface reconstruction: Combining multi-image stereo and shading,” Int. J. Computer Vision, vol. 16, pp. 35–56, 1995.

99. Mview website, http://vision.middlebury.edu/mview/.

100. C. Dyer, “Volumetric scene reconstruction from multiple views”, In L. S. Davis, editor, Foundations of Image Understanding, pp. 469.489. Kluwer, 2001.

101. SM. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski, “A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms”, CVPR 2006, vol. 1, pp 519-526.

102. B. Jia, Y.J. Zhang, X.G. Lin, "Stereo matching using both orthogonal and multiple image pairs", Acoustics, Speech, and Signal Processing, 2000 Vol 4. 2000 IEEE International Conference, pp. 2139-2142, 2000.

103. V. Ferrari, T. Tuytelaars, L.J. Van Gool, “Wide-baseline multiple-view correspondences”, in: IEEE Conference on Computer Vision & Pattern Recognition (CVPR), vol. 1, Wisconsin, Madison, USA, June 2003, pp. 718-728.

104. S. Tran and L. Davis, “3D surface reconstruction using graph cuts with surface constraints”, ECCV 2006, vol. II, pp. 219-231.

105. V. Kolmogorov and R. Zabih, “Multi-camera scene reconstruction via graph cuts”, In ECCV, vol. III, pp. 82.96, 2002.

106. Watanabe, M. Ohta, Y, “Cooperative integration of multiple stereo algorithms” Computer Vision, 1990. Proceedings, Third International Conference on Publication Date: 4-7 Dec 1990, pp: 476-480.

107. M. Goesele, B. Curless, and S. Seitz, “Multi-view stereo revisited”, In CVPR, pp 2402 – 2409, 2006..

108. C. Strecha, R. Fransens, and L. V. Gool, “Combined depth and outlier estimation in multi-view stereo”, In CVPR, pp 2394 – 2401, 2006.

109. D. Bradley, T. Boubekeur, and W. Heidrich, “Accurate multi-view reconstruction using robust binocular stereo and surface meshing”, CVPR 2008.

110. P. Merrell, A. Akbarzadeh, L. Wang, P. Mordohai, J.-M. Frahm, R. Yang, D. Nister, and M. Pollefeys, “Real-time visibility-based fusion of depth maps”, In ICCV, 2007.

111. R. Szeliski, “A multi-view approach to motion and stereo”, In CVPR, vol. 1, pp. 157.163, 1999.

112. Roger Y. Tsai, “Multiframe image point matching and 3-d surface reconstruction”, IEEE Trans. Pattern Anal. Machine Intell, PAMI-S (No.4). March 1983.

113. T. Kanade, M. Okutomi, and T. Nakahara, “A Multiple-baseline Stereo Method” Proceedings of the 1992 DARPA Image Understanding Workshop, pp 409-426 1992.

114. C. Hernandez and F. Schmitt, “Silhouette and stereo fusion for 3D object modelling”. CVIU, 96(3):367.392, 2004.

115. M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. Seitz, “Multi-view stereo for community photo collections”, ICCV 2007.

116. G. Vogiatzis, P. Torr, and R. Cipolla, “Multi-view stereo via volumetric graph-cuts”, In CVPR, pp. 391.398, 2005.

117. J. Pons, R. Keriven, and O. Faugeras, “Modelling dynamic scenes by registering multi-view image sequences”, In CVPR, 2005.

118. D. Marr and T. Poggio, “Cooperative Computation of Stereo Disparity”, pp 432–461. Artificial intelligence at MIT: Expanding Frontiers. MIT Press, Cambridge, MA, USA, 1996.

119. Chan. J.W, “Investigation into an integrated stereoscopic dual-energy X-ray camera”, 2004, NTU, PhD Thesis.

Appendix A Image set two

Image set two 14°

Image set two 12°

Image set two 10°

Image set two 8°

Image set two 6^{o}

Image set two 4^{o}

Image set two 2°

Image set two 0°

Image set two -2^{o}

Image set two -4^{o}

Image set two -6^{o}

Image set two -8^{o}

Image set two -10^{o}

Image set two -12^{o}

Image set two -14^{o}

Appendix B Tables for establishing the V-DMX experiment

The NE of the newly synthesised pixels in each iterations (Ite) produced by the Improved LamI criterion using all possible pairings.(target -12^{o})

	Ite1	Ite2	Ite3	Ite4	Ite5	Ite6	Ite7	Ite8	Ite9	Ite10	Ite11	Ite12	Ite13	Ite14
Pair-2-1	20255	0	0	0	0	0	0	0	0	0	0	0	0	0
Pair-2-2	2178	1804	2148	2237	2379	2305	1948	1361	846	505	188	37	14	1
Pair-3-1	2059	18008	0	0	0	0	0	0	0	0	0	0	0	0
Pair-3-2	1762	8470	4195	2245	1072	607	406	191	66	69	33	12	13	5
Pair-3-3	1363	2681	2109	1623	1302	1174	1070	1037	833	774	610	461	364	772
Pair-4-1	1350	18423	0	0	0	0	0	0	0	0	0	0	0	0
Pair-4-2	1206	13182	3548	1120	401	218	101	80	35	26	12	2	5	0
Pair-4-3	1079	5282	3298	1991	1232	853	637	508	430	320	273	228	174	432
Pair-4-4	982	3052	2344	1605	1347	1024	849	811	679	585	522	453	422	1534
Pair-5-1	974	18820	0	0	0	0	0	0	0	0	0	0	0	0
Pair-5-2	865	15688	2584	527	208	85	53	49	12	14	3	0	1	0
Pair-5-3	798	8468	3736	1833	942	543	483	341	271	225	180	111	123	244
Pair-5-4	754	4363	3408	2051	1290	849	661	474	431	373	358	278	254	1037
Pair-5-5	725	3192	2631	1849	1273	991	779	670	573	502	442	395	335	2082
Pair-6-1	735	19067	0	0	0	0	0	0	0	0	0	0	0	0
Pair-6-2	697	16720	2057	385	134	67	32	15	15	3	0	0	0	0
Pair-6-3	660	10804	3781	1388	715	429	330	244	167	107	96	65	38	122
Pair-6-4	631	5940	3978	2233	1159	664	469	381	324	312	243	180	171	700
Pair-6-5	615	3877	3388	2150	1279	932	660	549	431	367	341	289	247	1478
Pair-6-6	604	3163	2870	1996	1342	955	711	637	485	419	397	390	324	2333

lte1	lte2	lte3	lte4	lte5	lte6	lte7	lte8	lte9	lte10	lte11	lte12	lte13	lte14	lte15
Pair-7-1	619	19253	0	0	0	0	0	0	0	0	0	0	0	0
Pair-7-2	602	17537	1680	220	103	41	20	12	5	1	1	0	0	0
Pair-7-3	575	12383	3769	1206	506	331	216	154	127	85	63	48	25	63
Pair-7-4	558	7564	4442	2061	901	572	408	325	283	226	181	152	140	454
Pair-7-5	546	4816	3782	2320	1286	775	559	466	386	340	256	223	213	1048
Pair-7-6	541	3624	3209	2237	1339	920	670	480	463	422	368	343	271	1806
Pair-7-7	536	3191	2875	2097	1346	882	715	574	444	451	361	350	328	2666
Pair-8-1	542	19569	0	0	0	0	0	0	0	0	0	0	0	0
Pair-8-2	523	18217	1368	214	62	32	19	4	0	3	0	1	0	0
Pair-8-3	500	13808	3572	1021	427	243	148	125	101	74	42	49	22	48
Pair-8-4	484	9137	4517	1809	839	490	373	287	225	177	147	109	88	267
Pair-8-5	478	5650	4423	2436	1227	696	450	375	345	267	238	161	151	751
Pair-8-6	470	4088	3724	2316	1396	854	592	503	352	325	269	274	270	1414
Pair-8-7	465	3452	3223	2215	1448	942	676	509	435	360	335	294	284	2166
Pair-8-8	458	3142	2951	2058	1389	936	655	547	424	393	371	304	312	2911

The NE of the BEST-SSD's synthetic pixels that have the same position with the newly matched pixels in each iteration by the improved LamI pairings. (target -12^{o})

Ite1	Ite2	Ite3	Ite4	Ite5	Ite6	Ite7	Ite8	Ite9	Ite10	Ite11	Ite12	Ite13	Ite14	Ite15
Pair-2-1	12957	0	0	0	0	0	0	0	0	0	0	0	0	0
Pair-2-2	2296	1539	1749	1602	1614	1461	1237	836	537	250	116	31	11	1
Pair-3-1	1989	11054	0	0	0	0	0	0	0	0	0	0	0	0
Pair-3-2	1757	6102	2983	1344	626	334	199	98	31	39	17	7	3	5
Pair-3-3	1518	2462	1884	1368	967	795	653	587	448	468	337	270	207	456
Pair-4-1	1405	11888	0	0	0	0	0	0	0	0	0	0	0	0
Pair-4-2	1259	9105	2374	728	229	137	51	44	16	10	5	2	4	0
Pair-4-3	1169	4499	2685	1588	823	542	408	341	251	178	132	140	112	231
Pair-4-4	1102	2896	2154	1406	1057	711	623	526	393	322	299	243	228	898
Pair-5-1	1056	12570	0	0	0	0	0	0	0	0	0	0	0	0
Pair-5-2	953	10889	1778	322	106	60	37	21	11	6	1	2	0	0
Pair-5-3	889	6614	2836	1217	615	328	264	203	137	113	76	49	68	135
Pair-5-4	857	3932	2982	1693	1000	559	411	312	271	239	210	175	144	606
Pair-5-5	832	3008	2455	1672	1064	747	531	459	355	316	248	224	188	1271
Pair-6-1	816	13020	0	0	0	0	0	0	0	0	0	0	0	0
Pair-6-2	779	11783	1352	223	69	42	22	9	7	4	3	0	0	0
Pair-6-3	745	8102	2734	933	407	244	181	121	100	68	58	32	18	60
Pair-6-4	726	5070	3313	1675	794	402	306	225	212	181	155	104	112	363
Pair-6-5	711	3598	3081	1847	1047	680	436	367	269	237	210	180	165	886
Pair-6-6	701	3003	2674	1783	1142	753	516	454	336	239	249	210	189	1430

lte1	lte2	lte3	lte4	lte5	lte6	lte7	lte8	lte9	lte10	lte11	lte12	lte13	lte14	lte15
Pair-7-1	688	13240	0	0	0	0	0	0	0	0	0	0	0	0
Pair-7-2	669	12328	1060	121	60	18	8	5	6	2	0	0	0	0
Pair-7-3	644	9090	2620	764	300	165	124	89	66	60	41	26	19	45
Pair-7-4	631	6100	3480	1477	593	329	251	186	158	129	112	91	74	240
Pair-7-5	622	4306	3299	1935	1029	538	393	292	215	205	153	139	126	609
Pair-7-6	620	3417	2926	2026	1156	734	483	348	283	245	208	172	173	1085
Pair-7-7	615	3057	2643	1935	1218	716	534	407	324	279	216	194	169	1547
Pair-8-1	602	13384	0	0	0	0	0	0	0	0	0	0	0	0
Pair-8-2	584	12692	910	118	40	14	10	3	0	2	0	0	0	0
Pair-8-3	566	9865	2484	650	227	144	82	67	62	44	32	27	19	34
Pair-8-4	554	6919	3428	1340	515	273	222	160	112	88	89	67	54	156
Pair-8-5	549	4868	3640	1897	926	452	279	214	184	146	114	102	95	441
Pair-8-6	541	3786	3324	1997	1115	707	454	347	236	196	172	147	146	787
Pair-8-7	536	3259	2986	1936	1292	758	535	362	337	247	218	182	160	1186
Pair-8-8	529	2982	2757	1830	1254	792	513	400	332	290	261	206	186	1590

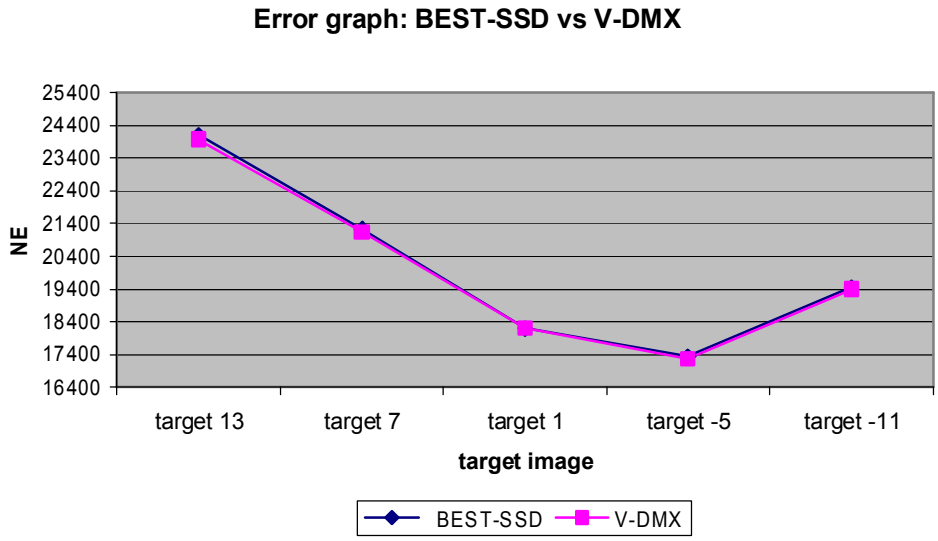
The NE difference between the BEST-SSD and the Improved LamI (i.e, NE of the Improved LamI – NE of the BEST-SSD) for newly synthesised pixels for each iterations. (target -12^{o})

Ite1	Ite2	Ite3	Ite4	Ite5	Ite6	Ite7	Ite8	Ite9	Ite10	Ite11	Ite12	Ite13	Ite14	Ite15
Pair-2-1	-7298	0	0	0	0	0	0	0	0	0	0	0	0	0
Pair-2-2	118	-265	-399	-635	-765	-844	-711	-525	-309	-255	-72	-6	-3	0
Pair-3-1	-70	-6954	0	0	0	0	0	0	0	0	0	0	0	0
Pair-3-2	-5	-2368	-1212	-901	-446	-273	-207	-93	-35	-30	-16	-5	-10	0
Pair-3-3	155	-219	-225	-255	-335	-379	-417	-450	-385	-306	-273	-191	-157	-316
Pair-4-1	55	-6535	0	0	0	0	0	0	0	0	0	0	0	0
Pair-4-2	53	-4077	-1174	-392	-172	-81	-50	-36	-19	-16	-7	0	-1	0
Pair-4-3	90	-783	-613	-403	-409	-311	-229	-167	-179	-142	-141	-88	-62	-201
Pair-4-4	120	-156	-190	-199	-290	-313	-226	-285	-286	-263	-223	-210	-194	-636
Pair-5-1	82	-6250	0	0	0	0	0	0	0	0	0	0	0	0
Pair-5-2	88	-4799	-806	-205	-102	-25	-16	-28	-1	-8	-2	2	-1	0
Pair-5-3	91	-1854	-900	-616	-327	-215	-219	-138	-134	-112	-104	-62	-55	-109
Pair-5-4	103	-431	-426	-358	-290	-290	-250	-162	-160	-134	-148	-103	-110	-431
Pair-5-5	107	-184	-176	-177	-209	-244	-248	-211	-218	-186	-194	-171	-147	-811
Pair-6-1	81	-6047	0	0	0	0	0	0	0	0	0	0	0	0
Pair-6-2	82	-4937	-705	-162	-65	-25	-10	-6	-8	1	3	0	0	0
Pair-6-3	85	-2702	-1047	-455	-308	-185	-149	-123	-67	-39	-38	-33	-20	-62
Pair-6-4	95	-870	-665	-558	-365	-262	-163	-156	-112	-131	-88	-76	-59	-337
Pair-6-5	96	-279	-307	-303	-232	-252	-224	-182	-162	-130	-131	-109	-82	-592
Pair-6-6	97	-160	-196	-213	-200	-202	-195	-183	-149	-180	-148	-180	-135	-903

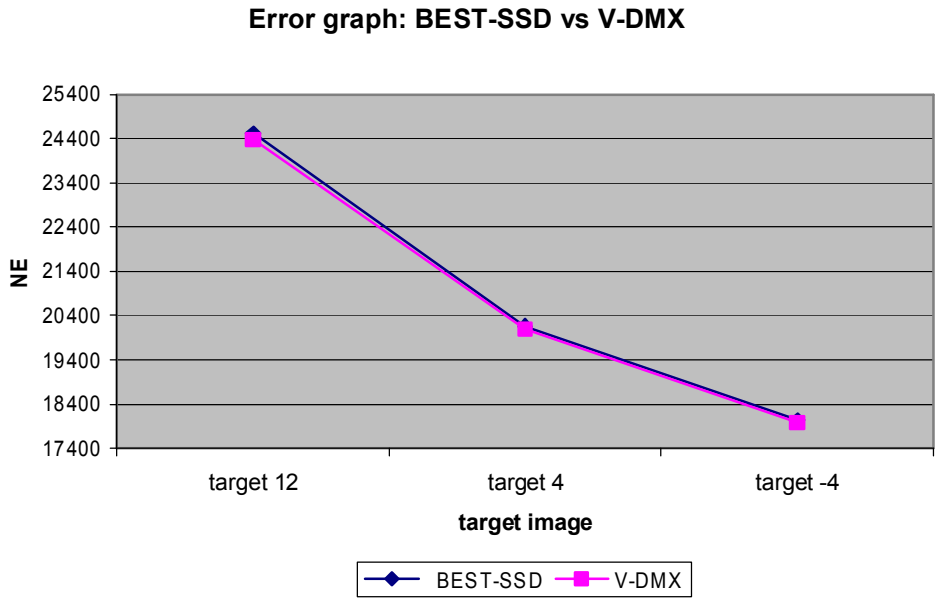
lte1	lte2	lte3	lte4	lte5	lte6	lte7	lte8	lte9	lte10	lte11	lte12	lte13	lte14	lte15
Pair-7-1	69	-6013	0	0	0	0	0	0	0	0	0	0	0	0
Pair-7-2	67	-5209	-620	-99	-43	-23	-12	-7	1	1	-1	0	0	0
Pair-7-3	69	-3293	-1149	-442	-206	-166	-92	-65	-61	-25	-22	-22	-6	-18
Pair-7-4	73	-1464	-962	-584	-308	-243	-157	-139	-125	-97	-69	-61	-66	-214
Pair-7-5	76	-510	-483	-385	-257	-237	-166	-174	-171	-135	-103	-84	-87	-439
Pair-7-6	79	-207	-283	-211	-183	-186	-187	-132	-180	-177	-160	-171	-98	-721
Pair-7-7	79	-134	-232	-162	-128	-166	-181	-167	-120	-172	-145	-156	-159	-1119
Pair-8-1	60	-6185	0	0	0	0	0	0	0	0	0	0	0	0
Pair-8-2	61	-5525	-458	-96	-22	-18	-9	-1	0	-1	0	-1	0	0
Pair-8-3	66	-3943	-1088	-371	-200	-99	-66	-58	-39	-30	-10	-22	-3	-14
Pair-8-4	70	-2218	-1089	-469	-324	-217	-151	-127	-113	-89	-58	-42	-34	-111
Pair-8-5	71	-782	-783	-539	-301	-244	-171	-161	-161	-121	-124	-59	-56	-310
Pair-8-6	71	-302	-400	-319	-281	-147	-138	-156	-116	-129	-97	-127	-124	-627
Pair-8-7	71	-193	-237	-279	-156	-184	-141	-147	-98	-113	-117	-112	-124	-980
Pair-8-8	71	-160	-194	-228	-135	-144	-142	-147	-92	-103	-110	-98	-126	-1321

Appendix C NE for comparative study of V-DMX and BEST-SSD experiment

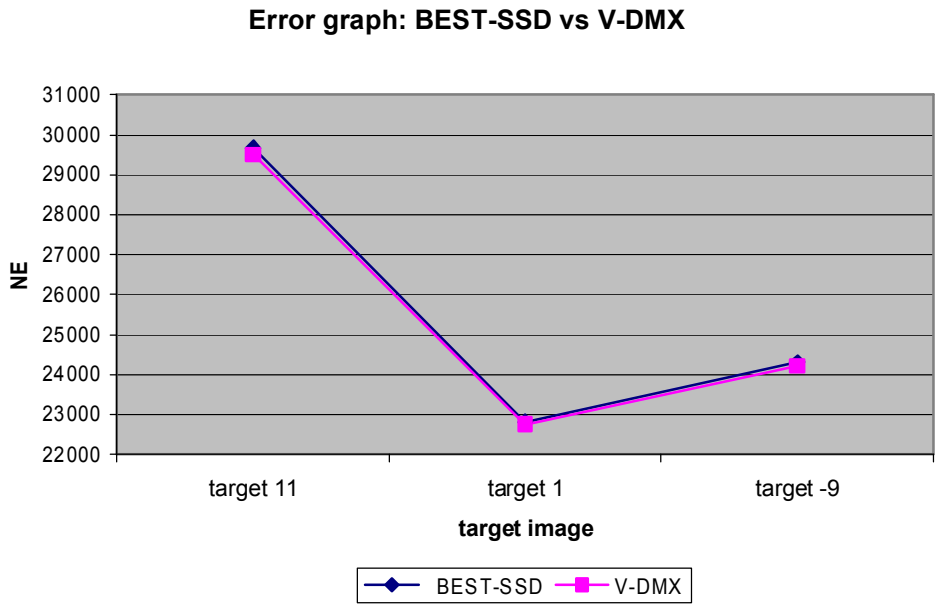
The NE comparison of the V-DMX with BEST-SSD generated by seven target images with an angle separation between adjacent detector images of 6°.



The NE comparison of the V-DMX with BEST-SSD generated by seven target images with an angle separation between adjacent detector images of 8°.



The NE comparison of the V-DMX with BEST-SSD generated by seven target images with an angle separation between adjacent detector images of 10°.



Appendix D List of publications

Evans J.P.O, Liu, Y, Chan, J.W, “Depth from motion 3D X-ray imaging for security screening”. In: IEEE International Symposium on Imaging for Crime Detection and Prevention, London, England, 7-8 June 2005.

Evans, J.P.O, Liu, Y, Chan, J.W, and DOWNES, D, “View synthesis for depth from motion 3D X-ray imaging" . Pattern Recognition Letters. Vol 27 (15) , pp. 1863-1873. 2006.