A New Asymmetrical Corner Detector (ACD) for a Semi-automatic Image Co-registration Scheme

Submitted to Imperial College, UK
for the degree of Master of Philosophy

By

Lisha Xie

Department of Earth Science and Engineering,
Imperial College, London, UK

October 2007

Abstract

Co-registration of multi-sensor and multi-temporal images is essential for remote sensing applications. In the image co-registration process, automatic Ground Control Points (GCPs) selection is a key technical issue and the accuracy of GCPs localization largely accounts for the final image co-registration accuracy. In this thesis, a novel Asymmetrical Corner Detector (ACD) algorithm based on auto-correlation is presented and a semi-automatic image co-registration scheme is proposed.

The ACD is designed with the consideration of the fact that asymmetrical corner points are the most common reality in remotely sensed imagery data. The ACD selects points more favourable to asymmetrical points rather than symmetrical points to avoid incorrect selection of flat points which are often highly symmetrical. The experimental results using images taken by different sensors indicate that the ACD has obtained excellent performance in terms of point localization and computation efficiency. It is more capable of selecting high quality GCPs than some well established corner detectors favourable to symmetrical corner points such as the Harris Corner Detector (Harris and Stephens, 1988).

A semi-automatic image co-registration scheme is then proposed, which employs the ACD algorithm to extract evenly distributed GCPs across the overlapped area in the reference image. The scheme uses three manually selected pairs of GCPs to determine the initial transformation model and the overlapped area. Grid-control and non-maximum suppression methods are used to secure the high quality and spread distribution of GCPs selected. It also involves the FNCC (fast normalised cross-correlation) algorithm (Lewis, 1995) to refine the corresponding point locations in the input image and thus the GCPs are semi-automatically selected to proceed to the polynomial fitting image rectification. The performance of the proposed co-registration scheme has been demonstrated by registering multi-temporal, multi-sensor and multi-resolution images taken by Landsat TM, ETM+ and SPOT sensors. Experimental results show that consistent high registration accuracy of less than 0.7 pixels RMSE has been achieved.

Keywords: Asymmetrical corner points, image co-registration, ACD

Acknowledgements

Firstly and most importantly, I would like to express my deepest gratitude to my supervisor Dr. JianGuo Liu for his support, guidance, knowledge and encouragement during the study. The work would never have been done without his constant assistance. I have enormously benefited from his invaluable scientific and spiritual support and continuous guidance.

I also would like to express thanks to the Department of Earth Science and Engineering, Imperial College London, especially the faculty, postgraduates students and staff for their support.

My gratitude extends to Dr. Hongshi Yan and Dr. Philippa J, Mason from Imperial College London, for their valuable comments, academic discussions and encouragement during the research.

Many particular thanks must go to my friends, Xiaolin Lu, Ran Qi, Matthew Rhodes, Huiyun Lu, Ioannis Fourniadis, Gareth Morgan, Qin Wang for their continuous support and friendship at Imperial College London.

Finally, I would like to thank my dearest parents very much for their love, kindness, encouragement, and understanding during my years of studies. Without their love and support, the work would not have been possible. I wish to express my deep gratitude to them for offering countless opportunities of facing various challenges in the world.

To my parents,

Yonghong Zhang and Anren Xie

Contents

Abstract

Acknowledgements

Contents i

List of Figures iii

List of Tables . ……ix

Chapter 1 Introduction 1

1.1 The main objective of the study 2

1.2 The structure of the thesis 2

Chapter 2 Review and background knowledge 4

2.1 Image co-registration: general procedure 4

2.2 Identification of Ground Control Points (GCPs) 9

2.3 Interest Point Detector 13

2.4 Auto-correlation function 17

Chapter 3 A New Asymmetrical Corner Detector (ACD) 24

3.1 Symmetrical and asymmetrical interest points 24

3.2 Harris Corner Detector and its property 26

3.3 Mathematical property and behaviour of the Asymmetrical Corner Detector (ACD) 33

3.4 Experimental results 39

Chapter 4 A semi-automatic image co-registration scheme.....83

4.1 The scheme.....83

4.2 Image co-registration experimental results.....87

Chapter 5 Discussion and conclusions.....102

5.1 Original contributions and major results of the research.....102

5.2 Methodology development.....103

5.3 Conclusions.....105

5.4 Critical appraisal of the research.....106

5.5 Future work.....107

References:.....108

Appendix: Kernel C codes of the ACD.....118

List of Figures

Figure 1. 1: Workflow of the investigation process.....3

Figure 2. 1: (a) flat region (b) edge region (c) corner region.....19

Figure 2. 2: Three Gaussian distributions each with a mean of zero and standard deviations of 0.5, 1.0 and 2.0, respectively.....20

Figure 2. 3: Three 2D Gaussian filters each with standard deviations of 0.5(a), 1.0(b) and 2.0(c), respectively. 21

Figure 2. 4: Principle of 1% cut-off in contrast enhancement 23

Figure 3. 1: Five types of interest points, they are: L-corner, Y-junction, T-junction, Arrow-junction and X-junction.....25

Figure 3. 2: Grey dots denote the interest points selected in different resolution images. 26

Figure 3. 3: The characteristics of the Harris Corner Detector. λ_1 and λ_2 vary between 0 and 100. The fine lines display the values of the Harris Corner Detector (*R*). 28

Figure 3. 4: The isoline distribution of the Harris Corner Detector.....29

Figure 3. 5: The isoline distributions of the response function *R* with different α values. (a) $\alpha = 0.06$ (b) $\alpha = 0.1$ (c) $\alpha = 0.15$ (d) $\alpha = 0.2$.....32

Figure 3. 6: The characteristics of the SCD. λ_1 and λ_2 vary between 0 and 100. The fine lines display the values of the SCD. 35

Figure 3. 7: The isoline distribution of the SCD. 35

Figure 3. 8: The plot of the relation between the ACD and fixed $\lambda_1 + \lambda_2$ summation varying between 0 and 100. The curves display the values of the ACD. 37

Figure 3. 9: The characteristics of peak areas in the ACD isolines. 38

Figure 3. 10: Manually selected interest points plotted on (a) the Harris Corner Detector *R* contour chart; (b) the ACD contour chart. The red dots are the interest points selected from a SPOT image and the black ones are extracted from a Landsat ETM+ image. (The values of each isoline in both charts are different) 40

Figure 3. 11: Interest points displayed as the white dots. White boxes denote zoom-in sub areas displayed in Figure 3.12. (a) Initial Landsat-7 ETM+ visible band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Results obtained using the ACD.....47

Figure 3. 12: Zoom-in sub areas. Interest points showed as the white dots. Red circles highlight the points for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD. 48

Figure 3. 13: Further zoom-in observation of interest points selected using the ACD and the Harris Corner Detector. The red circles locate the very good quality interest point which is detected only by the ACD. The blue circles denote the points selected by both the ACD and the Harris Corner Detector with window dynamic threshold while the Harris Corner Detector with global threshold failed. 49

Figure 3. 14: Interest points displayed as the white dots. (a) Initial SPOT panchromatic band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window

dynamic threshold. (d) Interest points detected using the ACD. Zoom-in sub areas are displayed in white squares.. 54

Figure 3. 15: Interest points displayed as the white dots. Red circles highlight the points for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD.....55

Figure 3. 16: Interest points displayed as the white dots. (a) Initial Landsat-7 ETM+ NIR band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. White square highlights the zoom-in area for comparison. 59

Figure 3. 17: Interest points displayed as the white dots. Red circles highlight the areas for comparison. (a) Initial zoom-in image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD. 59

Figure 3. 18: Interest points displayed as the white dots. (a) Initial Landsat-7 ETM+ visible band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. White squares denote the zoom-in areas for comparison. 64

Figure 3. 19: Interest points displayed as the white dots. Red circles highlight the points for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD.....66

Figure 3. 20: Further enlargement comparison of interest points detected. The blue circles denote the typical mountainous interest points selected by both the ACD and

the Harris Corner Detector with window dynamic threshold while the Harris Corner Detector with global threshold failed. 66

Figure 3. 21: Interest points displayed as the white dots. (a) Initial SPOT multi-spectral band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. Zoom-in sub areas are displayed in white squares. 71

Figure 3. 22 White dots denote the interest points selected using the Harris Corner Detector and the ACD. The red circles highlight the areas for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD. The blue circles denote several high quality interest points selected only by the ACD while both the Harris Corner Detectors with the global and window dynamic thresholds failed. .73

Figure 3. 23: Interest points displayed as the white dots. (a) Initial SPOT panchromatic band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. Zoom-in sub areas are displayed in white squares. 79

Figure 3. 24: White dots denote the interest points selected using the Harris Corner Detector and the ACD. The red circles highlight the areas for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD. 81

Figure 3. 25: Further zoom-in observation of interest points selected using the ACD. The red circles locate the very good quality interest points which are detected only by the ACD while both the Harris Corner Detector with global and window dynamic thresholds failed. The blue circles denote the points selected by both the ACD and the Harris Corner Detector with window dynamic threshold. The white circles highlight the points selected by all the three methods. 82

Figure 4. 1: Flow chart of the semi-automatic image co-registration scheme. 86

Figure 4. 2: (a) the input image taken in 1984 (b) the reference image taken in 1992. The red dashed square boxes highlight the overlapped area in both images.. 88

Figure 4. 3: The white dots denote the interest points selected using the ACD in the overlapped area of the reference image. 89

Figure 4. 4: 113 precisely matched pairs of GCPs detected in the image pair. The left one is the input image and the right one is the reference image. Red dashed square boxes denote the overlapped areas. 90

Figure 4. 5: The registration result and three zoom-in sub scenes. The reference image and co-registered image are displayed in green and red, respectively 91

Figure 4. 6: (a) the input image obtained using Landsat TM sensor in 1989 (b) the reference image taken in 2000 using Landsat-7 ETM+ sensor. The red dashed polygons highlight the overlapped areas in both images……..93

Figure 4. 7: The white dots denote the interest points selected using the ACD in the overlapped area (red dashed polygon) of the reference image 94

Figure 4. 8: 72 GCP pairs extracted from the image pair. The left one is the input image and the right one is the reference image. Red dashed polygons denote the overlapped areas……....95

Figure 4. 9: The registration result is presented in a colour composite with the reference image displaying in green and co-registered image displaying in red, respectively. The two red squares highlight the zoom-in sub scenes for a close look..96

Figure 4. 10: (a) the input image taken by Landsat-7 ETM+ sensor (b) the reference image taken by SPOT sensor. The red squares highlight the overlapped areas in both images. The white squares denote the zoom-in sub areas of the reference image displayed in Figure 4.11……....98

Figure 4. 11: The white dots denote the interest points selected using the ACD in the overlapped area of the reference image (four zoom-in sub images as white boxes highlighted in Figure 4.10b).....99

Figure 4. 12: 123 precisely matched GCP pairs detected in the image pair. The left one is the input image and the right one is the reference image. Red squares denote the overlapped areas……..100

Figure 4. 13: The co-registration result and four zoom-in sub scenes. The co-registered image and reference image are displayed in green and red, respectively. The area displaying in yellow denotes the overlapped area.....101

List of Tables

Table 3. 1: Wavebands and their characteristics of Landsat-7 ETM+ sensor.....41

Table 3. 2: Wavebands and their characteristics of SPOT sensor.....42

Table 3. 3: Numbers of points and computation time of interest point detection on satellite images (mountainous area) obtained using the Harris Corner Detector and the ACD. 45

Table 3. 4: Numbers of points and computation time of interest point detection on a Landsat ETM+ Near Infrared band image (mountainous and urban areas) obtained using the Harris Corner Detector and the ACD. 57

Table 3. 5: Numbers of points and computation time of interest point detection on satellite images (mountainous, urban and sea areas) obtained using the Harris Corner Detector and the ACD. 65

Table 3. 6: Numbers of points and computation time of interest point detection on a SPOT Panchromatic band image (flat urban area and field area along coast) obtained using the Harris Corner Detector and the ACD. 75

Chapter 1

Introduction

As the economy and technology are globally developing, remotely sensed data becomes more and more widely being implemented in a variety of application fields. Simply stated, remote sensing is the practice of deriving information about the Earth's land and water surfaces using images acquired from an overhead perspective, using electromagnetic radiation in one or more regions of the electromagnetic spectrum, reflected or emitted from the Earth's surfaces.

An important principle underlying the use of remotely sensed data is that different objects on the Earth's surface and in the atmosphere reflect, absorb, transmit or emit electromagnetic energy in different proportions, and that such differences allow these components to be identified. Sensors mounted on aircraft or satellite platforms are used to record the magnitude of the energy flux reflected from or emitted by objects on the Earth's surface to generate digital imagery data after essential calibration and corrections.

In order to measure the change over time using multi-temporal or multi-sensor image sets, in the case of optical imagery, will require correction for atmospheric variability, and it will also be necessary to co-register the images forming the multi-temporal sequence or multi-sensor image sets to a common geographic coordinate system. And the images are spatially matched so that corresponding pixels in the images correspond to the same physical region of the scene being imaged (Dai and Khorram, 1999). It includes two major steps:

- Establish a mathematical transformation model using ground control points (GCPs);
- Carry out an image re-sampling based on the transformation model.

In commonly used image processing packages, image co-registration is usually done manually. In these methods, GCPs are located visually on both the input image and reference image by choosing characteristic and easily recognizable points such as road junctions, river attribute joints, and building corners, etc. In order to get reasonably good registration results, a large number of ground control points must be selected over the whole image. This is a tedious, labour-intensive work and can only achieve limited accuracy. Therefore, there is a critical need to develop an automatic technique to accurately co-register multi-temporal and multi-sensor images with little manual supervision. The key technique for current automatic image co-registration techniques is an automatic GCPs selection algorithm while the core of such an algorithm is often an interest point detector.

1.1 The main objective of the study

This study aims to analyse current existing automatic GCPs selection algorithms, evaluate their merits and drawbacks, and finally develop an effective technique to achieve the automatic GCPs selection process and subsequently carry out high-accuracy image co-registration. The workflow illustrated in Figure 1.1 gives a brief description of the methods and approaches employed to carry out this investigation.

1.2 The structure of the thesis

This thesis presents details of my investigation of the proposed Asymmetrical Corner Detector (ACD), the semi-automatic image co-registration scheme, and the methodology which is used to achieve the research. It is organised as follows.

Chapter 2: Summary and discussion of the information theoretical aspects of interest point detector and a literature review of the current existing interest point detectors.

Chapter 3: The mathematical definition and property, the behaviour of the proposed Asymmetrical Corner Detector (ACD) and the comparison with the Harris Corner Detector are discussed together with the experimental results of the ACD.

Chapter 4: The outline of a semi-automatic image co-registration scheme and co-registration results between multi-temporal, multi-sensor and multi-resolution images.

Chapter 5: The discussion and conclusions.

Appendix : Kernel C codes of the ACD.

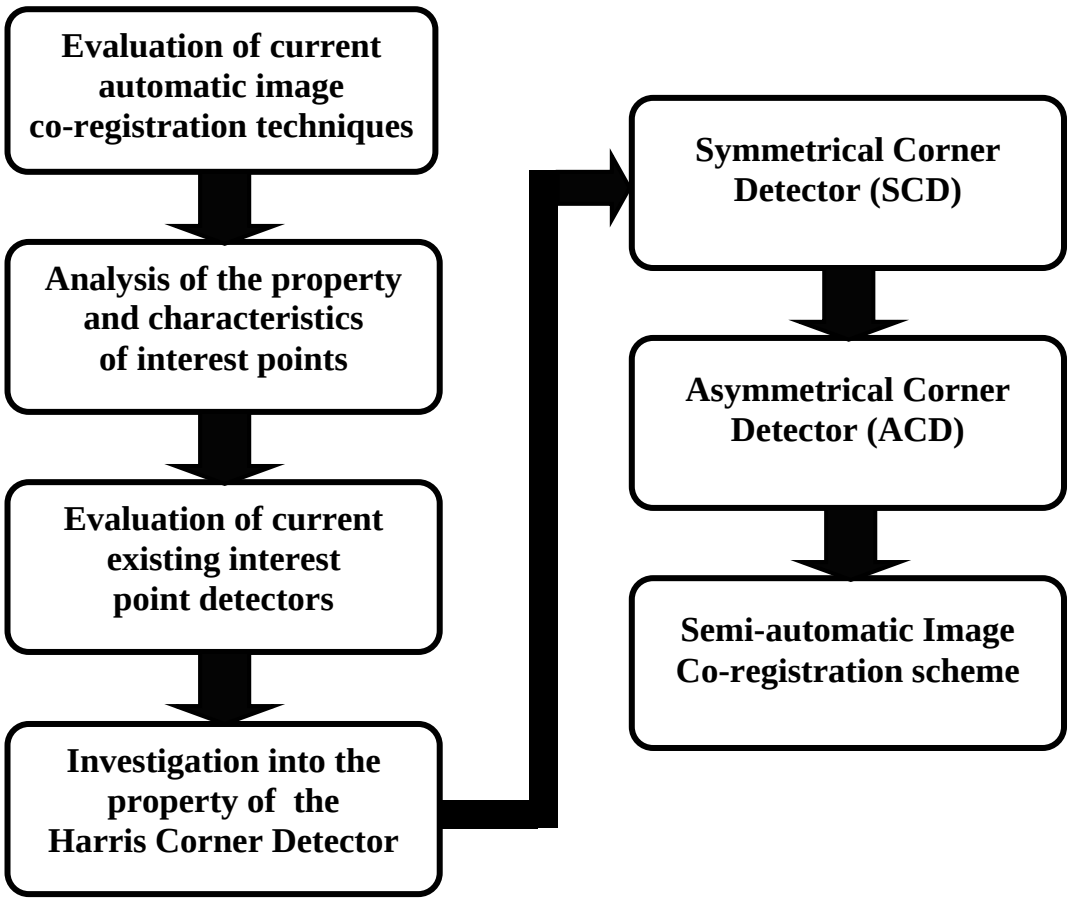


Figure 1. 1: Workflow of the investigation process

Chapter 2

Review and background knowledge

Image co-registration, as an essential pre-processing for various remote sensing applications, is required for higher accuracy and efficiency than ever before. In response to the reality of large quantity of data and various sensor systems, many novel image co-registration algorithms and strategies have been developed focusing on the four main registration steps including feature detection, feature matching, transform model estimation, and re-sampling and transformation. Among these, the automatic GCP selection is a key challenging task for the automation of image co-registration.

In this chapter, a comprehensive review of image co-registration approaches as well as the relevant GCP selection techniques is embedded in an introduction to the essential background knowledge for the technical development of the proposed Asymmetrical Corner Detector (ACD) and the semi-automatic image co-registration scheme reported in this thesis.

2.1 Image co-registration: general procedure

Image-to-image registration is the translation and rotation alignment process by which two images of like geometry and of the same geographic area are positioned coincident with respect to one and another so that corresponding elements of the same ground area appear in the same place on the registered images (Chen and Lee, 1992).

Theoretically, the majority of the image co-registration (image-to-image registration) methods consist of three steps:

- ***GCP selection and transformation model.***

Ground Control Points (GCPs) are located in the input image and reference image to mathematically model the geometric distortions (Jensen, 2005). The locations of the output pixels are deriving from location information provided by GCPs.

Normally, the geometric transformations employed are based on Euclidean geometry and are defined in a Cartesian space using real co-ordinates. And as the position of a pixel in an image has to be presented in integers of line and column while the transformation may not be one to one exactly, some pixels in the input image may be overlapped or take decimal positions in the output image. In other words, using input-to-output mapping can result in output raster pixels with no output value (Wolberg, 1990) or holes. In order to avoid this situation, inverse mapping, also known as output-to-input mapping, is often employed.

Suppose transformation M is an output-to-input mapping that maps pixel (i, j) to (k,l). For each pixel located at (i, j) in the output image, (k,l) is computed from $M(i, j)$, giving the corresponding positions in the original input image, and at this location, the pixel value is drawn and assigned to the output pixel at (i, j). (Liu, 2002)

After the correspondence has been established, the types and parameters of transformation model should be determined and estimated. In principle, they should correspond to the assumed geometric distortion, the method of image acquisition, and required accuracy of the co-registration. Polynomial approximation is commonly used to derive transformation M. Let (k,l) represents the location in the original input image corresponding to the output location (i, j), the general form of the polynomial for k and l are

$$\begin{aligned} k &= Q(i, j) = q_0 + q_1 i + q_2 j + q_3 i^2 + q_4 ij + q_5 j^2 + ... \\ l &= R(i, j) = r_0 + r_1 i + r_2 j + r_3 i^2 + r_4 ij + r_5 j^2 + ... \end{aligned} \tag{2.1}$$

In the above polynomials, k, l and i, j are known from GCPs and the least squares solution for the coefficients of the polynomial can thus be derived. This approach is proved to be very effective and accurate for satellite images. From Eq. 2.1, for n GCPs, we have:

$$Q = \begin{bmatrix} q_0 \\ q_1 \\ \dots \\ q_n \end{bmatrix} \quad K = \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix} \quad M = \begin{bmatrix} 1 & i_1 & j_1 & \dots \\ 1 & i_2 & j_2 & \dots \\ \dots & \dots & \dots & \dots \\ 1 & i_n & j_n & \dots \end{bmatrix} \tag{2.2}$$

The least squares solution for Q is

$$Q = (M^t M)^{-1} M^t K \tag{2.3}$$

Similarly, for R

$$R = (M^t M)^{-1} M^t L \tag{2.4}$$

Thus, the corresponding position (k, l) in the input image can be calculated from the location (i, j) in the output image and the transformation model M derived from Eq. 2.1 and Eq. 2.2.

For moderate distortions in a relatively small area of an image, a first-order, six-parameter, affine (linear) polynomial approximation (Goshtasby, 1986) is usually used to generate the transformation model M . This type of transformation can model six parameters of geometric distortion in the remotely sensed imagery data including the translation and scale change in two directions, skew and rotation (Novak, 1992; Buiten and Van Putten, 1997). Depending on the distortion in the imagery data and the degree of topographic relief displacement in the area, higher-order polynomial approximation is required to geometrically rectify data. In principle, the higher the order of the polynomial, the more closely it can model the geometric distortions. Higher order polynomials can produce a more accurate fit but require more GCPs and computational complexity. For instance, in the case of same sensor or same resolution image co-registration, a linear polynomial approximation is sufficient while in the

multi-sensor or multi-resolution case, the order increases to quadratic or cubic (Goshtasby, 1987) to guarantee the co-registration accuracy.

- ***Image re-sampling based on the transformation model.***

This step is to decide how to best estimate the values and positions of pixels in the rectified output image based on the transformation model M . In most cases, (k,l) is not an integer co-ordinate and there is no pixel value for this point. Therefore, re-sampling, an interpolation procedure, is used to extract the pixel value for the pixel located at (k,l) from the input image so as to assign it to the pixel at (i, j) in the output image. The most efficient strategy from a computational perspective is to assign the value of the input image pixel nearest to position (k,l) to the pixel at position (i, j) in the output image. This function is called the nearest neighbour interpolation (Liu, 2002).

A more accurate method is bilinear interpolation which assigns output pixel values by interpolating pixel values in two orthogonal directions in the input image. It basically fits a plane to the four pixel values nearest to the desired position in the input image and then computes a new pixel value based on the weighted distances to these points. A higher-order interpolation method, called cubic convolution, calculates values for output pixels based on 16 input pixels surrounding the desired pixel with weighted distances. Besides these methods, cubic B-splines (Hou and Andrews, 1978), quadractic splines (Toraichi *et al*., 1988; Dodgson, 1997), Catmull-Romcardinal splines (Thevenaz *et al*., 1998) belong to the most commonly used interpolation methods.

Actually, nearest neighbour interpolation is not used in most cases because it sometimes introduces a small error into rectified image. The output image may be offset spatially by up to 0.5 pixels, causing a jagged or blocky appearance. Even though the bilinear interpolation is outperformed by higher-order methods in terms of accuracy and visual appearance of the rectified image, it offers the best trade-off between accuracy and computational complexity and thus it is the most commonly used method. Cubic interpolation is recommended when the geometric transformation

involves a significant enlargement. Therefore, bilinear interpolation method is used to carry out re-sampling in the proposed semi-automatic image co-registration scheme.

- ***Evaluation of the image co-registration accuracy***

Estimation of co-registration accuracy is essential to assess the performance of co-registration algorithms used and it can help to make right choice of particular methods for the process. Co-registration errors may be introduced into the co-registration process in each of its stages. It is a common approach to evaluate the partial accuracy in the early stage in order to improve the final co-registration accuracy.

The accuracy of GCPs localization and correspondence largely accounts for the final co-registration accuracy. Serious mistakes in this stage usually lead to failure of the co-registration. It is very important to determine how well the polynomial approximation derived from the initial GCPs account for the geometric distortion in the input image before applying the transformation model to generate the rectified output image. For this purpose, the error at each GCP needs to be observed so as to anticipate the final accuracy of image co-registration. The method used most often is the root mean square error (RMSE) estimation (Ton and Jain, 1989), which is the standard deviation of the difference between actual positions of the GCPs and their calculated positions (i.e., after rectified) (Campbell, 2002).

Suppose a transformation model M based on polynomial approximation is derived from a set of corresponding GCP pairs identified in the input image and reference image. We first identify a GCP in the original input image and its column and row coordinates are (i_{orig}, j_{orig}) .Then, using the transformation model M and the corresponding GCP position (x, y) in the reference image, we can calculate the estimated co-ordinates (i, j) in the input image space (Jensen, 1995). And the RMSE can be computed using Eq. 2.5

$$RMSE = \sqrt{(i - i_{orig})^2 + (j - j_{orig})^2} \tag{2.5}$$

Where i_{orig} and j_{orig} are the original row and column co-ordinates of the GCP in the input image and i and j are the estimated co-ordinates. The square root of the

standard deviations represents a measure of the accuracy of this GCP. Therefore, we always compute RMSE for all GCPs to see which GCP exhibits the greatest error and generate the total RMSE. If the total RMSE exceeds a pre-defined threshold, it is common to delete some GCPs with high RMSE or refine the GCP positions. In such a way, RMSE is simple and effective assessment for selection of high quality GCPs (Bentoutou *et al*., 2005). The RMSE is used to evaluate the accuracy of the GCPs in the semi-automatic image co-registration scheme discussed in Chapter 4.

2.2 Identification of Ground Control Points (GCPs)

GCPs are of vital importance to the accuracy of above described image co-registration procedure, which is the focus of this study. In manual co-registration methods, GCPs can be located with precision and accuracy on accurate maps and are also easily located on digital images. In theory, GCPs could be as small as a single pixel and easily identified against its background. However, most GCPs are likely spectrally distinct areas as small as a few pixels in real images depending on the spatial resolution of the image such as the intersections of major roads, edges of land-cover parcels, stream junctions etc. Although locating this kind of points seems to be an easy task, in fact, it is quite difficult to select individual points with high accuracy. Normally, it is relatively easy to locate a few good GCPs precisely but, in some areas, it is rather difficult to locate adequate high quality GCPs that evenly spread in an image. Bernstein (1983) presented that the registration error decreases as the number of GCPs is increased. However, the quality of GCPs accuracy may degrade as their number increases as we usually select the best and easiest ones first. In addition, GCPs should be well spread throughout the image with good coverage near image frame margins. If a large number of GCPs are concentrated in a few regions of the image, the transformation model M derived from these GCPs may not be accurate to the whole image. The reason is that polynomial approximation produces a more accurate fit for areas immediately surrounding GCPs while geometric errors could be significant in large distances from the GCPs (Gibson and Power, 2000). As a result, the GCP selection process is a labour-intensive, repetitive, tedious work with limited accuracy.

A critical need is therefore raised to develop automatic techniques that require little or no operator supervision to select high quality GCPs.

During the past decades, numerous approaches have been proposed to deal with this problem. Generally, the majority of automatic GCP selection methods consist of the following two steps:

- ***Feature extraction***

In this step, salient and distinctive features are automatically selected and their point representatives are subsequently used as GCPs to rectify images. Normally, distinctive objects which are frequently spread over the image, easily recognized, and can be identified in both the reference image and input image are detected.

Region-like features are usually selected such as high contrast closed-boundary regions of an appropriate size (Goshtasby *et al*., 1986; Flusser and Suk, 1994), water body (Holm, 1991), forests (Sester *et al*., 1998), urban areas (Hsieh *et al*., 1992; Roux, 1996) etc. After the detection, the regions are often represented by their centres of gravity which are invariant with respect to rotation, scaling and stable under random noise and grey level variation. In addition, several line feature extraction methods are developed such as classical methods like canny detector (Canny, 1986) and the Laplacian of Gaussian based detector (Marr and Hildreth, 1980). Besides, Li *et al*., (1992), Li *et al*., (1995), Moss and Hancock (1997), and Wang and Chen (1997), selected object contours, roads and coastal lines and described the correspondence using the pairs of line segment ends or middle points.

However, the most widely used features are point features because they can be directly employed as GCPs to establish the relationship between the reference image and the input image. The techniques for point feature selection consist of methods that locate line intersections (e.g. road crossings) (Growe and Tonjes, 1997), centroids of regions (Ton and Jain, 1989), and local curvature discontinuities detected using the Gabor wavelets (Zheng and Chellapa, 1993). However, the core algorithms of point feature detectors follow the identification of corners. Considerable efforts have been contribute to developing precise, robust and effective methods to detect corners (Kitchen and Rosenfeld, 1982; Harris and Stephens, 1988; Horaud *et al*., 1990;

Heitger *et al*., 1992; Förstner, 1994; Smith and Brady, 1997). As this research involves in developing a robust and effective new technique for corner point detection, several widely used corner detectors will be further reviewed in detail in the next section §2.3.

- ***Feature matching***

In this step, the correspondence between the features detected in the reference and input images is established and subsequently used to generate the transformation model. Generally, image intensity values in close neighbourhood, feature spatial distribution or feature descriptors are applied to match the features. The classical representatives of feature matching methods are the normalized cross-correlation (NCC) and its modifications (Pratt, 1991).

Normalized Cross-Correlation (NCC) is widely used to automatically optimize point positions from one image to the other. Suppose we get the fixed GCP positions in the reference image, the optimized co-ordinates of this GCP in the input image are decided when $R(u,v)$ reaches the maximum. $R(u,v)$ is the normalized cross-correlation between the input image and reference image computed in a calculation window roaming in a searching window centred at the roughly selected position of the GCP in the input image. The Normalized Cross-Correlation is defined as:

$$R(u,v)=\frac{\sum_{x,y}[I_r(x,y)-\overline{I_r}][I_i(x-u,y-v)-\overline{I_i}]}{\{\sum_{x,y}[I_r(x,y)-\overline{I_r}]^2\sum_{x,y}[I_i(x-u,y-v)-\overline{I_i}]^2\}^{1/2}} \tag{2.6}$$

Where $I_r(x,y)$ and $I_i(x,y)$ are the grey value functions of the reference image and input image. $\overline{I_r}$ and $\overline{I_i}$ are the mean grey values of all the pixels in the calculation window in the reference and input images, respectively. $(x-u,y-v)$ denotes the co-ordinates of the pixel in the searching window in the input image.

The computing efficiency for NCC can be significantly improved via the Fast Normalized Cross-Correlation (FNCC) algorithm (Lewis, 1995) by computing the items in the denominator of the above formula with a lookup table that holds integral (running sum) of image columns of calculation window width. In a searching window of size M^2 and moving calculation window of size N^2, NCC requires $N^2(M-N+1)^2$

additions and $2N^2(M-N+1)^2$ multiplications, while FNCC only needs N^2 additions and $N^2+N^2(M-N+1)^2$ multiplications. Avoiding all the repeated operations in calculation of the denominator of NCC, the FNCC algorithm speeds up the processing by 5 to 10 times (depending on the searching window size and calculation window size) (Liu and Ma, 2004).

Generalized versions of Normalized Cross Correlation are often used for geometrically more deformed images (Hanaizumi and Fujimura, 1993; Berthilsson, 1998). There are also several variants of correlation-like methods such as edge-based correlation (Anuta, 1970; Wie and Stein, 1977), Sequential Similarity Detection Algorithm (SSDA) (Barnea and Silverman, 1976), sum of squared differences similarity measure (Wolberg and Zokai, 2000), and increment sign correlation (Kaneko *et al*., 2002; Kaneko *et al*., 2003). Besides, template matching (Fonseca and Manjunath, 1996), phrase correlation (Castro and Morandi, 1987; Chen *et al*., 1994; Reddy and Chatterji, 1996), mutual information (Viola and Wells, 1997; Thevenaz and Unser, 1998), Gauss-Newton (Sharma and Pavel, 1997) and Levenberg-Marquardt (Sawhney and Kumar, 1999) are also proposed for feature matching. In addition, a large group of matching methods are based on the relaxation approach which labels each feature from the input image with the label of a feature from the reference image (Ranade and Rosenfeld, 1980; Wang *et al.,* 1983; Mitra and Murthy, 1991; Li *et al*., 1992; Cheng, 1996). Pyramid approach is used to reduce the computational cost due to the large image size (Vanderbrug and Rosenfeld, 1977; Zheng and Chellapa, 1993; Turcajova and Kautsky, 1996; Wang and Chen, 1997; Althof *et al*., 1997; Shinagawa and Kunii, 1998).

Considering its simplicity and computing efficiency, the FNCC is applied to refine the GCP positions in the input image in the semi automatic image co-registration scheme presented in Chapter 4.

2.3 Interest Point Detector

Interest points are pixels that exhibit some singularity with respect to their neighbourhood or the points which represent distinct features and are quite robust to signal noise ratio. In addition, these points are invariant to image geometry and well perceived by human observers. The characteristics of interest points are quite close to the property of the manually selected GCPs. Consequently, several interest point detectors are considered as effective algorithms to select GCPs automatically. They can effectively extract typical interest points such as L-corners, T-junctions, and Y-junctions for certain purposes.

Current methods can be classified into three categories: contour based, parametric model based and intensity based methods (Schmid *et al*., 2000). The following discussion gives a brief review of the current interest point detectors and focuses on some widely used intensity based methods.

Contour based methods

Contour based methods extract contours first and then search for maximal curvature or inflexion points along the contour chains or achieve some polygonal approximation and then search for intersection points (Schmid *et al*., 2000).

Interest points are extracted from planar curves which are observed as the changes such as junctions and endings in curvature (Asada and Brady, 1986). Mokhtarian and Mackworth (1986) developed a similar method to detect inflexion points from planar curves. By calculating the coefficients of B-splines, the points with the local maxima of curvature are detected (Medioni and Yasumoto, 1987). Horaud *et al*. (1990) first extracted contour chains using canny edge detector and then extracted line segment from the contour chains and the intersections of grouped line segments are finally regarded as interest points. High curvature points and intersection points along ridges and troughs are detected as interest points (Shi and Tomasi, 1994). This method selects minima points in the image surface instead of maximum points and it is proved to detect easily tracking points which are less likely to lie on occluding contours of objects. However, the computational expense is very high requiring long processing time. Pikaz and Dinstein (1994) proposed to decompose noisy digital curves into a

minimal number of convex and concave sections. The detection of the interest points is then based on properties of pairs of sections that are determined in an adaptive manner, rather than on properties of single points that are based on a fixed-size neighbourhood. Rattarangsi and Chin (1992) presented a transform converting the scale space into a tree in order to detect objects with multiple-size features and noisy boundaries, and then a parsing scheme is established to locate a corner with maximum curvature at a majority of scales.

Another corner detector based on curvature scale space (CSS) is developed to detect corner point at a high scale of the CSS and tracked through multiple lower scales to improve localization (Mokhtarian and Suomela, 1998). A further improvement is proposed by Zhong and Liao (2007) that convolves the curvature of a planar curve with a Gaussian kernel directly and forms a Direct Curvature Scale Space (DCSS). According to the behaviour of isolated single and double corner models and the specific properties of the models, corners are detected in a multi-scale sense. Corners are detected by sliding set of three rectangles along the curve and counting the number of contour points lying in each rectangle (Masood and Sarfraz, 2006).

Parametric Model methods

Parametric model methods consist in determining the value of the model's parameter which best fit the underlying intensity pattern (Schmid *et al*., 2000). They can provide sub-pixel accuracy, but are limited to specific types of the interest points (i.e. L-corners). In addition, the complexity of this type of detectors can be quite high depending on the approach taken to optimize the parameters.

Guiducci (1988) proposed an algorithm based on the characteristics of the corner such as the amplitude, aperture angle and the smoothness of the wedge which best approximates the image gray-level gradients and curvatures in a region surrounding the corner point. Rohr (1992) developed an analytic junction model convolved with a Gaussian filter to extract interest point. It can obtain very precise location of corner points but the quality of the approximation depends on the initial position estimation. An extension of Rohr's method is proposed by Deriche and Blaszka (1993). The corner detector is based on the second directional derivative orthogonal to the gradient and an exponential model is used instead of Gaussian smoothing filter in order to

reduce the computational expense. The image region has to be quite large to guarantee the convergence however, a clustered image is likely to contain several kinds of signals making convergence difficult. Laganiere and Vincent (2002) combined the two detectors (Rohr, 1992; Deriche and Blaszka, 1993) together by detecting corner shapes from intensity patterns. A wedge model that characterizes corners by their orientation and angular width is employed in this approach. Each pixel location is examined by comparing its surrounding circular neighbourhood with the ideal corner model. Baker *et al*., (1998) proposed an algorithm to automatically construct detectors for arbitrary parametric features in which realistic multi-parameter feature models, optical and sensing effects are used to maximize the robustness. A unified approach based on a template deformation framework presented to detect, classify and reconstruct general junctions is developed by Parida *et al*., (1998).

Intensity based methods

Intensity based methods calculate a measure that indicates the presence of an interest point directly from the grey-values (Schmid *et al*., 2000). An interest point always has a well-defined position and can be robustly detected. So it can be a corner but it can also be, for example, an isolated point of local intensity maximum or minimum. In practice, however, most so-called corner detectors are sensitive not specifically to corners, but to interest points in general which have a high degree of variation in all directions. Therefore, considerable efforts have been contribute to developing precise, robust and effective methods to detect corners.

The first signal based corner detector is developed by Moravec (1977). It measures the differences of grey-values in a window that is shifted in several directions around a central position. However, it is anisotropic, noisy and too biased to edges since the variation is computed along four directions only. Beaudet (1978) employed the second derivative of the signal as the measure of interest points. Similarly, Kitchen and Rosenfeld (1982) proposed an interest point detector that searches for the curvature maxima on isophotes of the signal. Dreschler and Nagel (1982) tried locating interest points with local extrema of the Gaussian curvature. The points with the steepest gray value slope along the line connecting the local extrema are selected as interest points. Interest points are located if the image differences along the contour direction is above a given threshold. Meanwhile, Nagel proved that his method and Kitchen-Rosenfeld's

approach are actually identical. A fuzzy logic based corner detector which works like a pattern classification is developed by Lee and Bien (1996). It is based on the Bayesian classifier and the probability density function that is calculated using the means of fuzzy logic. However, the requirement of a proper training dataset makes it less practical. Trajkovic and Hedley (1998) proposed another corner detector based on the variation of image intensity along arbitrary lines passing through the point under investigation within a neighbourhood of the point. A corner is detected if the variation of image intensity along lines in all orientations is high. The variation is found using first derivatives to avoid being sensitive to noise. Furthermore, a multi-grid approach is employed to improve computing efficiency. The shortcoming of this detector is its relatively poor accuracy though it is very fast.

One widely used branch of intensity based interest point detection methods is based on auto-correlation matrix (Harris and Stephens, 1988; Förstner and Gulch, 1987; Tomasi and Kanade, 1991; Cooper *et al*., 1993; Förstner, 1994). The 2×2 matrix averages derivatives of the signal in a window around a point. It captures the structure of the neighbourhood. If the matrix is of rank two, that is both of its two eigenvalues are large, an interest point is detected. A matrix of rank one indicates an edge and a matrix of rank zero indicates a homogeneous region.

To avoid using discrete directions and discrete shifts, Harris and Stephens (1988) improved the corner detector proposed by Moravec (1977) using an auto-correlation matrix. Cooper *et al.*, (1993) first measure the contour direction locally and then compute image differences along the contour direction. An interest point is detected when the image differences along the contour direction is sufficient compared to the noise. Förstner (1994) uses the auto-correlation matrix to divide images into three categories – flat region, contours and interest points. Interest points are further classified into junctions and circular features by using the ratio of the eigenvalues. Then interest point locations are determined by minimizing a function of the local gradient field. An operator based on the intuitive notion of symmetry is proposed by Reisfeld *et al*., (1995). The magnitude and direction of the derivatives of neighbourhood points composed the measure of symmetry strength and then, points with high symmetry are selected as interest point. Smith and Brady (1997) developed a low-level feature detector called SUSAN (Smallest Univalue Segment Assimilating

Nucleus). Interest points are detected based on the size, centroid and second moment of the image region which has a similar brightness to the centre. An improved version of SUSAN which employs a multi-threshold strategy rather than one global threshold is proposed by Zhou *et al*. (2004). In order to improve the efficiency, the search space for corner candidates is limited to the pixels with the intensity gradient magnitudes superior to a fixed threshold.

Among all the interest point detectors reviewed above, the Harris Corner Detector (Harris and Stephens, 1988) is well known as one of the most stable and robust corner detectors in image processing applications especially for Ground Control Points (GCPs) selection (Zhang *et al*., 1995; Schmid *et al*., 2000; Jung and Lacroix, 2001; Zhang *et al*., 2004; Bentoutou *et al*., 2005). However, the design of the Harris Corner Detector overlooked an important fact that the asymmetrical corner points are the far more common case than the symmetrical corner points in most images of land surface. My study will focus on this issue to develop an Asymmetrical Corner Detector based on the investigation of mathematical property and behaviour of the Harris Corner Detector as detailed in Chapter 3.

2.4 Auto-correlation function

Several interest point detectors (Harris and Stephens, 1988; Förstner and Gulch, 1987; Tomasi and Kanade, 1991; Cooper *et al*., 1993; Förstner, 1994) are based on a matrix related to the auto-correlation function. The local auto-correlation function measures the local changes of the signal. The measure is obtained by correlating a patch with its neighbouring patches, which is achieved by moving a calculation window from its original position to different directions. In the case of an interest point which represents distinctive features, the auto-correlation function is high in all shift directions (Schmid *et al*., 2000). Given a shift $(\Delta x, \Delta y)$ from a pixel P_{x_k, y_k}, the auto-correlation function is defined as:

$$f(x,y) = \left[\sum_{W} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2 \right] \tag{2.7}$$

where $I(x_k, y_k)$ denotes the image grey value at pixel P_{x_k, y_k}.

Considering a first-order approximation based on the Taylor expansion below:

$$I(x_k + \Delta x, y_k + \Delta y) \approx (I(x_k, y_k) + (I_x(x_k, y_k) I_y(x_k, y_k)) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \tag{2.8}$$

Combining Eq. 2.7 and Eq. 2.8 gives

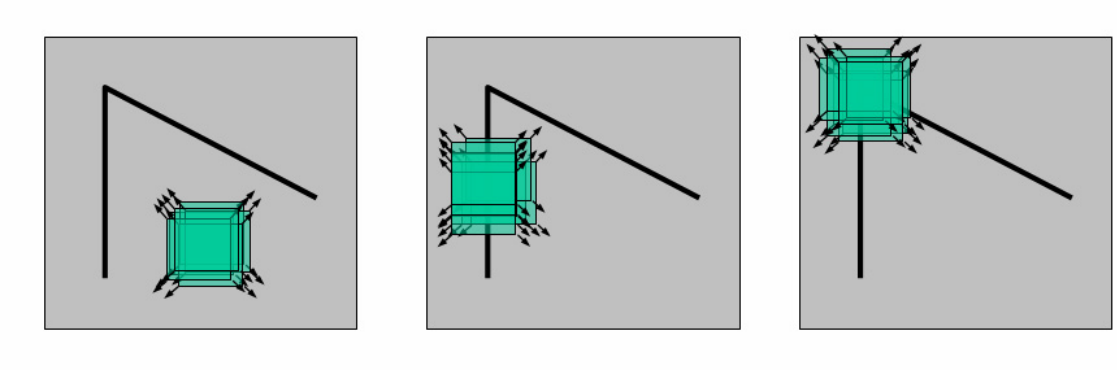
$$\begin{aligned} f(x, y) &= \sum_W ((I(x_k, y_k) + (I_x(x_k, y_k) I_y(x_k, y_k)) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix})^2 \\ &= (\Delta x \Delta y) \begin{bmatrix} \sum_W (I_x(x_k, y_k))^2 & \sum_W I_x(x_k, y_k) I_y(x_k, y_k) \\ \sum_W I_x(x_k, y_k) I_y(x_k, y_k) & \sum_W (I_y(x_k, y_k))^2 \end{bmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ &= (\Delta x \Delta y) A(x, y) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \end{aligned} \tag{2.9}$$

The above Eq. 2.9 shows that the auto-correlation function can be approximated by matrix $A(x, y)$. This matrix captures the structure of the neighbourhood. The eigenvalues of matrix A, λ_1 and λ_2, are proportional to the principal curvatures of the local auto-correlation function and form a rotation invariant description of matrix A. Base on the magnitudes of the eigenvalues, the following inferences can be made as shown in Figure 2.1 (Schmid *et al.*, 2000).

(1) If $\lambda_1 \approx 0$ and $\lambda_2 \approx 0$, the matrix is of rank zero, there are no features of interest at pixel $P_{x,y}$. It is in the flat region.

(2) If $\lambda_1 \approx 0$ and λ_2 is some large positive values, the matrix is of rank one and an edge is found where no changes is along the edge direction.

(3) If λ_1 and λ_2 are both large, distinct positive values, the matrix is a full rank matrix. Significant changes are found in different directions and a corner is found.



(a) **(b)** **(c)**

Figure 2. 1: (a) flat region (b) edge region (c) corner region

Gaussian Filter

Evaluation of gradients obtained using Eq. 2.7 is sensitive to noise in the input data. It is therefore desirable to smooth the image by some low pass filters before computing the gradients. Typically, a Gaussian filter (Gonzalez and Woods, 2002) can be applied in the pre-processing step because the random noise and the nature signal response usually have a probability density function of Gaussian distribution.

A Gaussian or normal distribution gives the probability of observing a value x with a mean value $\bar{x}$ is defined as:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\bar{x})^2}{2\sigma^2}} \tag{2.10}$$

The standard deviation σ for a normal distribution defines the range of the variable for which the function $p(x)$ drops by a factor of $e^{-0.5}$ or 0.607 of its maximum value. In other words, 60.7% of the values of a normally distributed variable lie within one standard deviation of the mean. The so-called standard Gaussian distribution is given by taking $\bar{x} = 0$ in a general Gaussian distribution. It is evident from Figure 2.2 that the probability decreases away from the mean in a systematic and symmetrical trend.

Mathematically, applying a Gaussian filter to an image is the same as convolving the image with a Gaussian or normal distribution. It is specifically in two-dimensional as defined:

$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2.11}$$

where σ is the standard deviation of the Gaussian distribution.

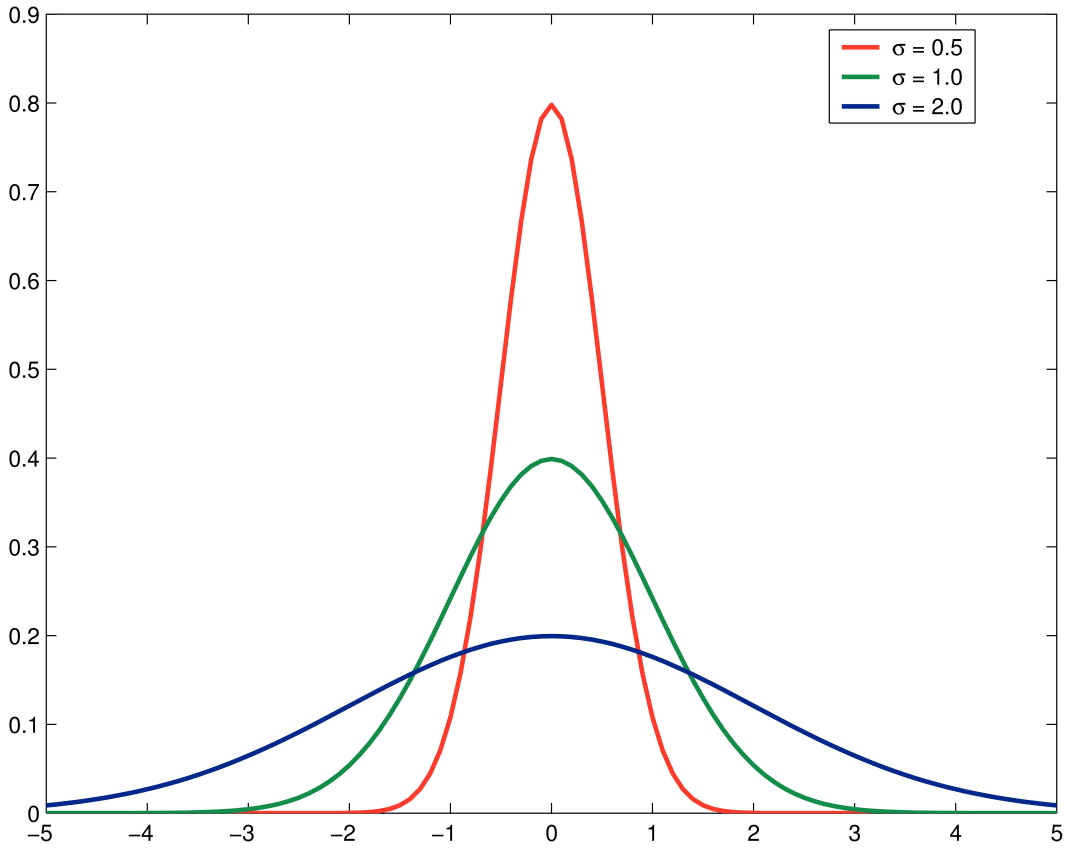
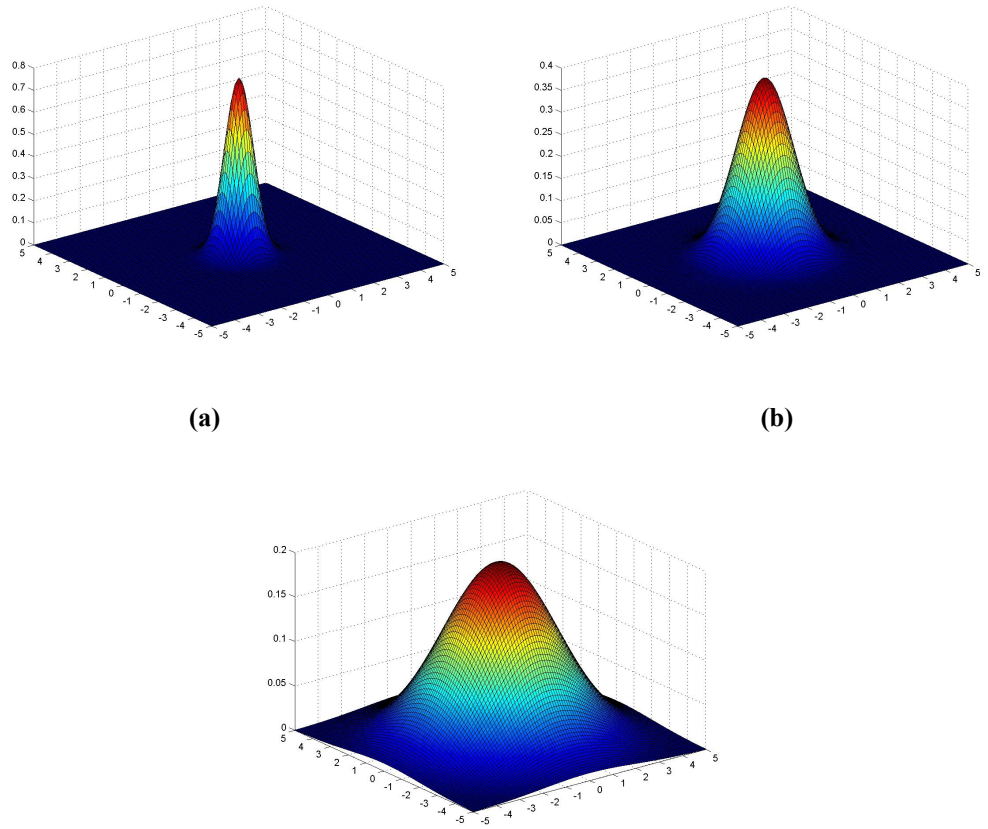


Figure 2. 2: Three Gaussian distributions each with a mean of zero and standard deviations of 0.5, 1.0 and 2.0, respectively.

When applied in two dimensions, the equation produces a surface whose contours are concentric circles with a Gaussian distribution from the centre point. Pixels where this distribution is non-zero are used to build a convolution matrix, which is applied to the original image. Each pixel's value is set to a weighted average of that pixel's neighbourhood. The origin pixel receives the heaviest weight (having the highest

Gaussian value), and neighbourhood pixels receive decreasing weights with the increase of their distances to the origin pixel. In theory, the distribution at every point on the image will be non-zero, meaning that the entire image would need to be included in the calculations for each pixel. In reality, when computing a discrete approximation of the Gaussian function, pixels outside of approximately 3σ are small enough to be considered effectively zero. Thus, pixels outside of that range can be ignored. Typically, an image filtering program needs only calculate a matrix with dimensions $(6\sigma+1)\times(6\sigma+1)$ to ensure all relevant pixels are accounted for.



(a) **(b)**

(c)

Figure 2. 3: Three 2D Gaussian filters each with standard deviations of 0.5(a), 1.0(b) and 2.0(c), respectively.

Three-dimension views of two-dimensional Gaussian filters with standard deviations (σ) of 0.5, 1.0 and 2.0 are displayed in Figure 2.3, respectively. The Gaussian filter with standard deviation of 0.5 is used in the Harris Corner Detector and also the proposed Asymmetrical Corner Detector (ACD) to smooth the image before computing the gradients. Its convolution kernel is given:

$$G = \begin{pmatrix} 0.0113 & 0.0838 & 0.0113 \\ 0.0838 & 0.6193 & 0.0838 \\ 0.0113 & 0.0838 & 0.0113 \end{pmatrix} \tag{2.12}$$

Contrast Enhancement

The auto-correlation matrix which represents the nature of the pixel's neighbourhood is based on the local contrast. As a consequence, the auto-correlation based point detection methods always fail to achieve satisfied results in low contrast imagery data. Therefore, contrast enhancement is essential before carrying out auto-correlation calculation for remotely sensed digital imagery data. *Linear scale* which expands the original input brightness values to make use of the total dynamic range of the video display (0 to 255) is widely used as the equation listed below (Gonzalez and Woods, 2002; Liu, 2002):

$$y = 255 \times \frac{x - Min(x)}{Max(x) - Min(x)} \tag{2.13}$$

Where y denotes the grey value assigned to the pixel in the output image while x is the original grey value of the pixel in the input image. $Max(x)$ and $Min(x)$ denote the maximum and minimum grey value of the input image.

In addition, a few pixels (often represent noise) may occupy wide value range at the low and high ends of the histogram of a digital image. In this case, setting a proper cut-off at the both ends to saturate a contrast enhancement is necessary to make an effective usage of the dynamic range of a display system. Normally, cut-off is given as the percentage of total number of pixels in an image. In the Harris Corner Detector and also the proposed Asymmetrical Corner Detector (ACD), 1% and 99% are set as the cut-off limits at the low and high ends of the histogram of an image. Then, the image is such stretched as to set 1% of the total number of pixels in the image in the low end of the histogram to 0 and 1% pixels in the high end of the histogram to 255 (Figure 2.4).

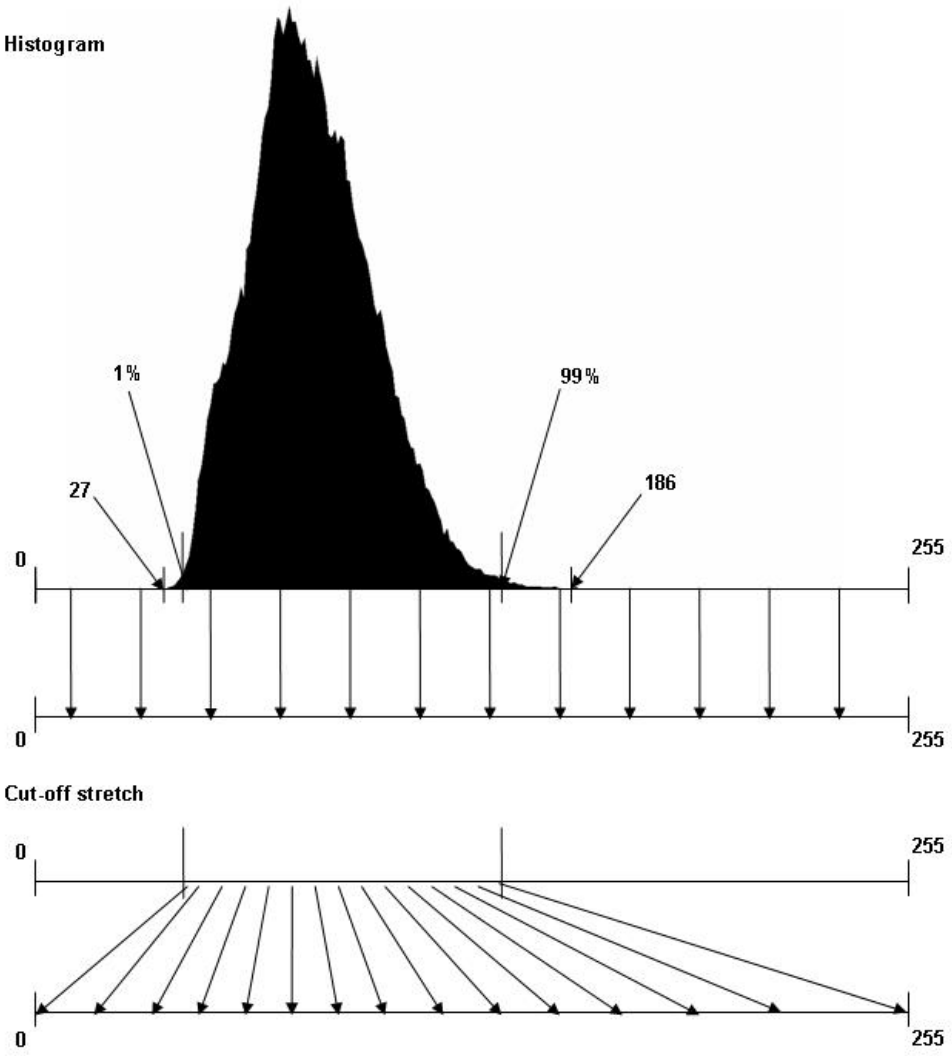


Figure 2. 4: Principle of 1% cut-off in contrast enhancement

Chapter 3

A New Asymmetrical Corner Detector (ACD)

In this chapter, an in-depth analysis of the auto-correlation based Harris Corner Detector is presented first. Recognising its drawbacks and limitations, a new Asymmetrical Corner Detector (ACD) is then proposed together with the investigation of its mathematical property and behaviour for critical assessment.

3.1 Symmetrical and asymmetrical interest points

To assess the principle and the performance of the auto-correlation based Harris Corner Detector, let us look into the nature of interest points and the difference between symmetrical and asymmetrical interest points first.

An interest point is a point in the image which in general can be characterized as follows:

- It is relatively distinctive and has a clearly defined position in image.
- It can be mathematically defined.
- The local image structure around the interest point is rich in terms of local information contents. Information content measures the distinctiveness of the interest point.

- It is stable under local and global perturbations in the image domain, including deformations as those arising from perspective transformations as well as illumination/brightness variations, such that the interest points can be reliably detected independently to changes in imaging conditions.

For a two-dimension grey-level image, interest points can also be simply stated as any point in the image for which the signal changes above an interest level in two dimensions. In real-life images, we generally consider five types of distinctive points such as L-corners, T-junctions, Y-junctions, Arrow-junctions and X-junctions as typical interest points for particular applications. Examples of these types of interest points are shown in Figure 3.1.

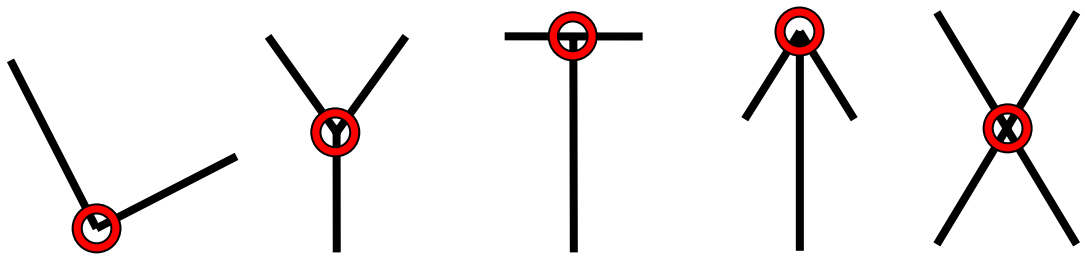
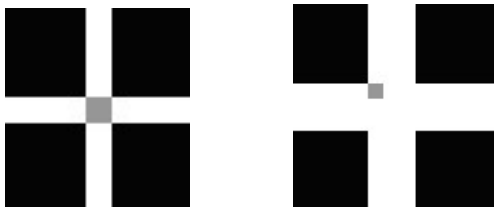


Figure 3. 1: Five types of interest points, they are: L-corner, Y-junction, T-junction, Arrow-junction and X-junction.

All the interest points selected from real images can be considered into two categories: symmetrical and asymmetrical interest points. Symmetrical interest points are the points with same changes in corresponding directions in their neighbourhood like X-junctions or isolated points and asymmetrical interest points are the points with different changes in their neighbourhood such as L-corners, T-junctions, Y-junctions and Arrow-junctions. In real images taken by satellite sensors or aerial cameras, typical points on the edge are normally selected as interest points such as Y-junction and T-junction which are hardly symmetrical. Therefore, ideal symmetrical interest points are quite rare while asymmetrical interest points represent a more common case.



(a) **(b)**

Figure 3. 2: Grey dots denote the interest points selected in different resolution images. (a) symmetrical interest point selected in a low resolution image (b) asymmetrical interest point selected in a high resolution image

In addition, resolution should also be considered as an important factor on the symmetry of interest points. Due to the diversity of sensors mounted on various satellite platforms, images of different spatial resolution are obtained. As shown in Figure 3.2, the grey dots denote the interest points selected. Obviously, in the low resolution image (a), the X-junction corner has only one pixel and it is entirely symmetrical. However, if we consider this corner in a higher resolution image (b), in that corner, there are 9 pixels and the pixel which can be selected as an interest point is not entirely symmetrical. To conclude, symmetrical interest point is only a special case of the general case of asymmetrical interest points. An effective corner detector must be particularly robust to detect asymmetrical interest points.

3.2 Harris Corner Detector and its property

Harris and Stephens (1988) improved the corner detector proposed by Moravec (1977) using an auto-correlation matrix. The use of discrete directions and discrete shifts is thus avoided. In the past decades, several comparisons with other interest point detectors have demonstrated that Harris Corner Detector gets the best results under various different imaging conditions. It is one of the well-known interest point detectors which are robust and stable for automatic selection of Ground Control Points (Zhang *et al*., 1995; Schmid *et al*., 2000; Jung and Lacroix, 2001; Zhang *et al*., 2004; Bentoutou *et al*., 2005). However, the design of the Harris Corner Detector overlooked an important fact that the asymmetrical corner points are far more common than the

symmetrical corner points in most images of land surface. Therefore, let us look into the mathematical property and behaviour of the Harris Corner Detector first.

Given auto-correlation matrix *A*, the response function of the Harris Corner Detector is defined as:

$$R = Det(A) - \alpha \times Trace(A)^2 \qquad (\alpha = 0.04\sim0.06) \tag{3.1}$$

Where $Det(A)$ and $Trace(A)$ are the determinant and trace of Matrix A, respectively. Considering the eigenvalues of Matrix A, λ_1 and λ_2, we can re-write Eq. 3.1 as:

$$R = \lambda_1 \times \lambda_2 - \alpha(\lambda_1 + \lambda_2)^2 \tag{3.2}$$

The value of this function for each pixel is a measure of the interest strength or the level of full rankness.

In actual data processing, Harris first computes the gradients of the image by [-1 0 1] mask. According to the evaluation that gradients are sensitive to noise, Gaussian filter is therefore employed to smooth the image before computing the gradients. It efficiently eliminates noise such as isolated points.

$$I_s = I * G \tag{3.3}$$

Where

$$G(x,y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (\sigma = 0.5)$$

σ is the standard deviation of the Gaussian distribution.

Then, Matrix A averages gradients of the image in a window *W* around a point $P_{x,y}$ given by

$$A(x,y) = \begin{bmatrix} \sum_W (I_x(x_k,y_k))^2 & \sum_W I_x(x_k,y_k) I_y(x_k,y_k) \\ \sum_W I_x(x_k,y_k) I_y(x_k,y_k) & \sum_W (I_y(x_k,y_k))^2 \end{bmatrix} \tag{3.4}$$

Where (x_k, y_k) denotes the points P_{x_k, y_k} in the window W around point $P_{x,y}$, $I_x(x_k, y_k)$ and $I_y(x_k, y_k)$ are the first derivatives along x and y directions at point P_{x_k, y_k}.

Additionally in the Matrix A, a Gaussian is used to weight the derivatives summed over the window to make the response function isotopic.

In order to understand the property of the Harris Corner Detector, the variation in the detector's value along with the change of the two eigenvalues λ_1 and λ_2 is analysed. Figure 3.3 shows the curves of fixed summation of the two eigenvalues ($\lambda_1 + \lambda_2$) in relation to the variation between λ_1 and the Harris Corner Detector value (R). Obviously, in each curve, R reaches its peak value when $\lambda_1 = \lambda_2$ as red crosses shown. On the other hand, when λ_1 are significantly different from λ_2, R becomes negative and indicates typical edges as red circles displayed.

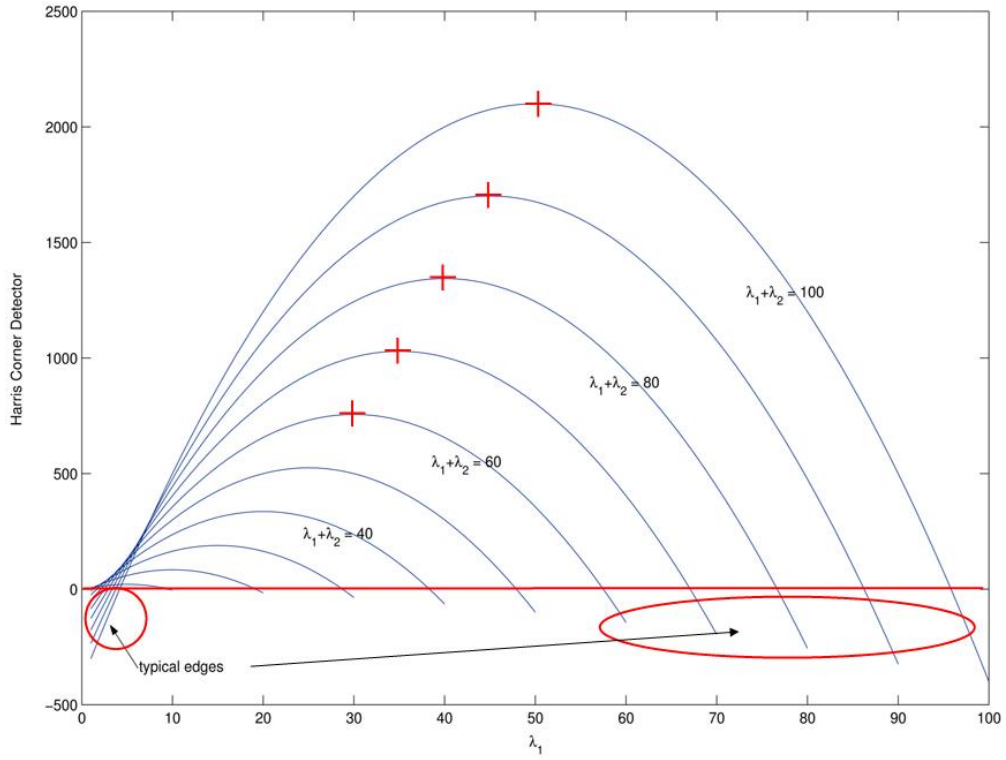


Figure 3. 3: The characteristics of the Harris Corner Detector. λ_1 and λ_2 vary between 0 and 100. The fine lines display the values of the Harris Corner Detector (R).

Pixels with the interest strength superior to a pre-defined threshold will be selected as interest points. In the literature, there are a few different settings of the threshold.

Schmid *et al.,* (2000) used 1% of the maximum observed interest point strength as the threshold which is widely used in the literature to select interest points. A fixed threshold of 1000 is employed to carry out interest point selection process (Mikolajczyk and Schmid, 2002). Marinkovic and Hanssen (2004) specified a window threshold which is 25 times larger value of the maximum intensity inside a search window and the window sizes for different areas such as urban area and non-urban area are varied. Gao *et al.*, (2004) divided the image into subsections and extracted one corner point from each subsection. Besides, in order to limit the number of detected corners, adjusted thresholds are used by Zhu *et al.,* (2005). Weinheimer *et al.,* (2006) applied an empirically determined threshold of 700 to achieve a reasonable number of points. Therefore, a fixed threshold is used to help evaluate the property of the Harris Corner Detector.

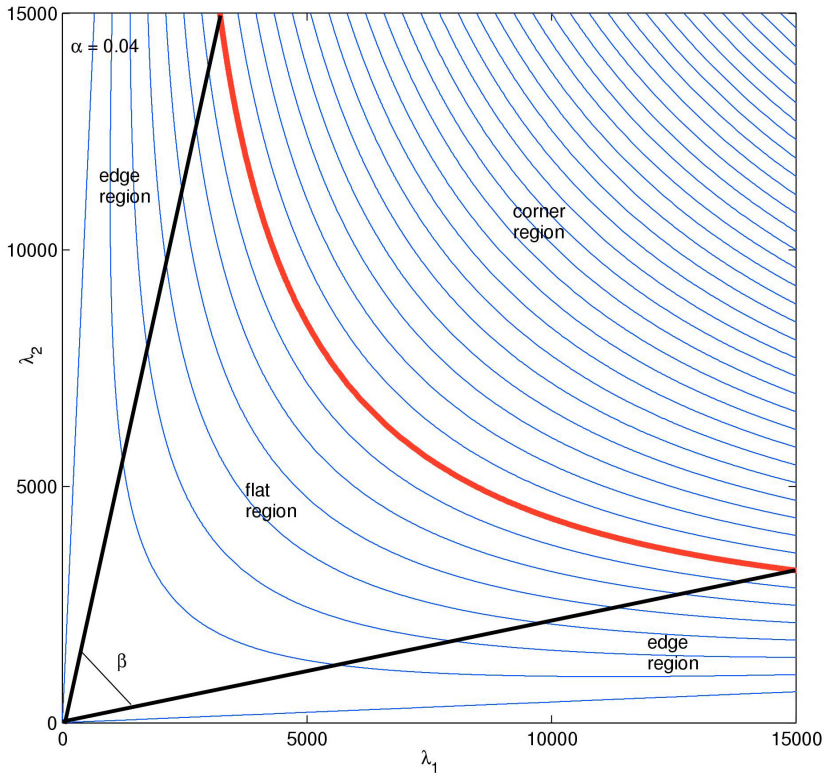


Figure 3. 4: The isoline distribution of the Harris Corner Detector

Contours of response function R plotted against the variation of λ_1 and λ_2 are presented in Figure 3.4. The red solid line is set as the fixed threshold of the Harris Corner Detector. For typical images obtained by satellite sensors and aerial cameras, without contrast enhancement, the range of λ_1 and λ_2 is usually (0, 15000) as the two

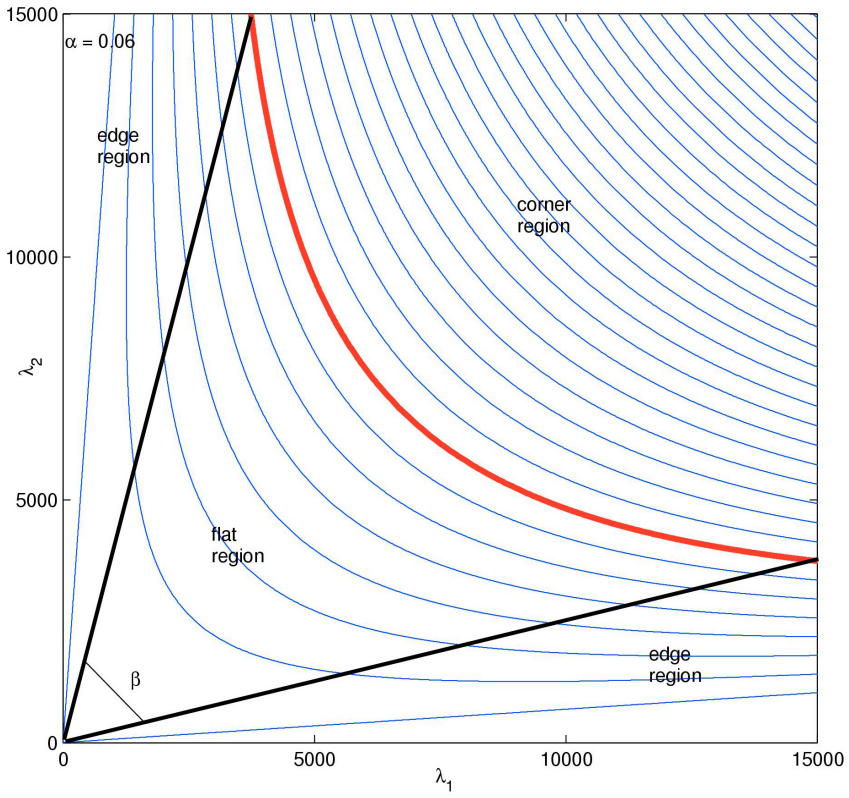
axes shown in Figure 3.4. As a consequence, the critical isoline angle β is defined as the angle between the two lines connecting the origin and the intersections between the threshold curve and the λ_1 and λ_2 axes. The chart can thus be divided into a flat region, two edge regions and a corner region. As the contours of *R* are curved toward the origin of the $\lambda_1 \sim \lambda_2$ coordinate system, the Harris Corner Detector obviously favourites ideal symmetrical points when two eigenvalues are the same. Here we summarise the key phenomena presented in Figure 3.4 as follows:

(1) When λ_1 and λ_2 are both large, the response function *R* is positive and large and the pixel is in the corner region.

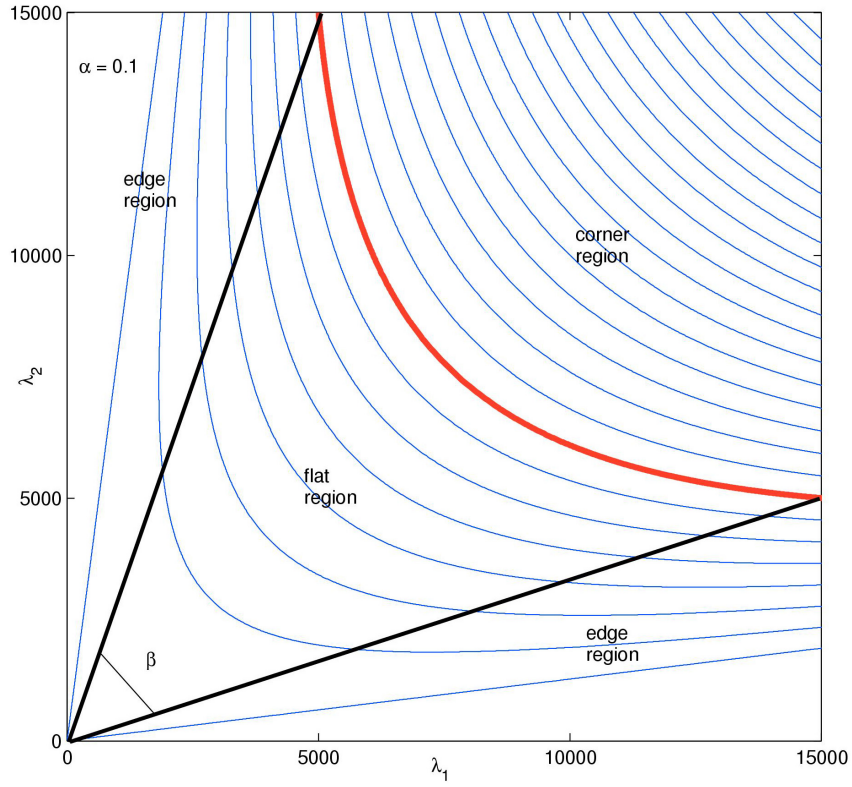
(2) When one of λ_1 and λ_2 is large and the other one is small, the response function R is negative and the pixel is in either of the two edge regions.

(3) When both of λ_1 and λ_2 are small, the response function R is positive but small and the pixel is in the flat region.

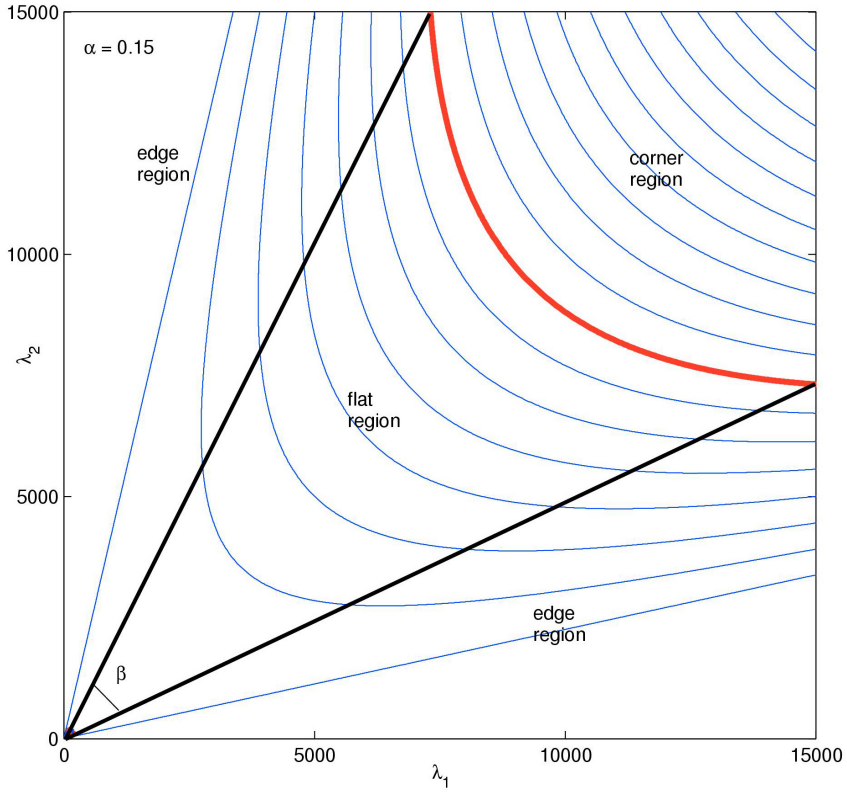
However, the performance of the response function *R* is largely dependent on the empirical parameter α. In order to prove this assumption, different values are set to α and results indicate that the behaviour of the response function *R* changes. Comparisons of response function *R* with different α values are presented in Figure 3.5. The values of α are set to 0.06, 0.1, 0.15 and 0.2, respectively. With the increase in α value, the isoline angle β for a fixed threshold of *R* decreases and the contours of response function *R* increasingly curved toward the origin of the $\lambda_1 \sim \lambda_2$ coordinate system. In other words, the function is increasingly becoming a measurement of symmetry rather than magnitude. It becomes very eliminative to obvious asymmetrical corner points and can only select nearly perfect symmetrical points. It is prone to select nearly symmetrical pixels in the flat region rather than partially asymmetrical pixels on the edge regions. Therefore, the empirical parameter α brings in ambiguity to the response function *R* and with a fixed threshold of the response function *R*, the performance of the Harris Corner Detector depends on two interfering parameters: α and the fixed threshold.


15000
α = 0.06
edge
region
corner
region
10000
λ2
5000
flat
region
edge
region
β
0
0
5000
10000
15000
λ1

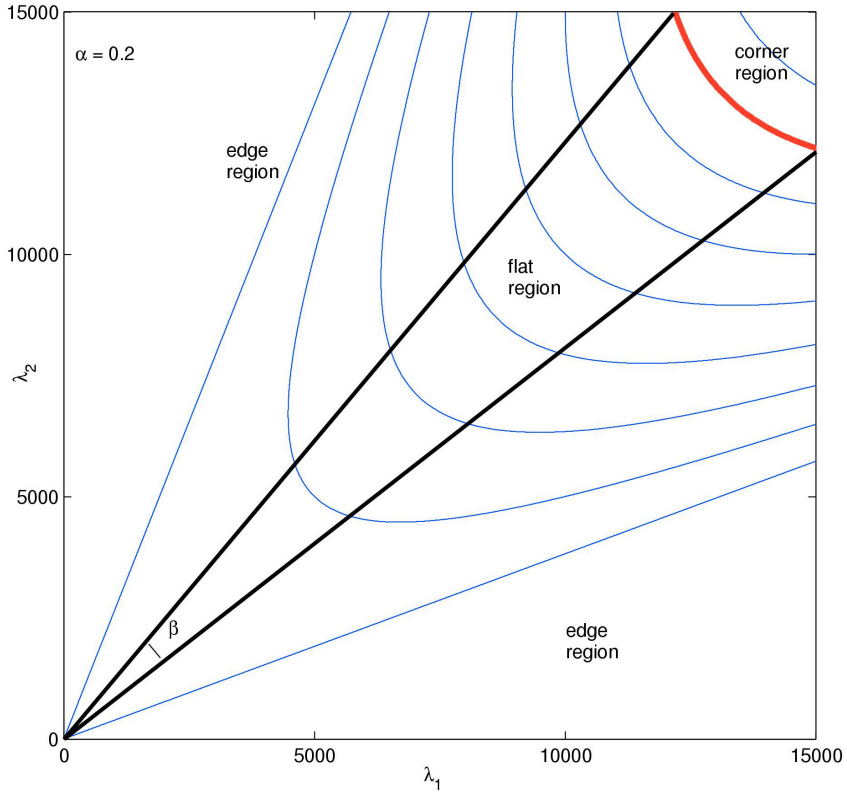
(a)


15000
α = 0.1
edge
region
corner
region
10000
λ2
5000
flat
region
edge
region
β
0
0
5000
10000
15000
λ1

(b)



(c)



(d)

Figure 3. 5: The isoline distributions of the response function *R* with different α values. (a) $\alpha = 0.06$ (b) $\alpha = 0.1$ (c) $\alpha = 0.15$ (d) $\alpha = 0.2$

In addition, another response function *R* based on the eigenvalues of Matrix *A* but without an empirical parameter has been proposed as listed below.

Haralick (1993): $$R = 1 - (\frac{\lambda 1 - \lambda 2}{\lambda 1 + \lambda 2})^2 = \frac{4Det(A)}{Trace^2(A)} \quad (3.5)$$

In order to look into the nature of the response function *R* defined in Eq. 3.5, we analyzed the relationship between the λ_1 , λ_2 and the response function *R*, as summarised below:

(1) On the edge region where one of λ_1 and λ_2 is large and the other one is small, the response function *R* is nearly zero.

(2) However, when both λ_1 and λ_2 are large or small, the response function *R* gets the value of one.

From the above two inferences, we can draw a conclusion that this detector is not effective to distinguish the pixels in the corner region and flat region. Though the response function *R* has normalised value of [0, 1], it is not as effective as the Harris Corner Detector. A robust, stable and effective single user control parameter corner detector is still in searching.

3.3 Mathematical property and behaviour of the Asymmetrical Corner Detector (ACD)

Based on the analysis of the behaviour of eigenvalues λ_1 and λ_2, we propose a robust interest point detector which can effectively select asymmetrical interest points.

3.3.1 Investigation of function $(\lambda_1 - \lambda_2)^2$ and a symmetrical corner detector (SCD)

For the eigenvalues λ_1 and λ_2 of auto-correlation Matrix *A*, we have:

$$(\lambda_1 - \lambda_2)^2 \geq 0 \tag{3.6}$$

Expanding the Eq. 3.6, then gives:

$$\lambda_1^{\,2} + \lambda_2^{\,2} \geq 2\lambda_1\lambda_2 \tag{3.7}$$

Thus, we get

$$0 \leq \frac{2\lambda_1\lambda_2}{\lambda_1^{\,2} + \lambda_2^{\,2}} \leq 1 \tag{3.8}$$

From the Eq. 3.8, we have the following inferences

(1) If $\lambda_1 = \lambda_2$, then $\frac{2\lambda_1\lambda_2}{\lambda_1^{\,2} + \lambda_2^{\,2}} = 1$

(2) If $\lambda_1 \neq \lambda_2$, then $\frac{2\lambda_1\lambda_2}{\lambda_1^{\,2} + \lambda_2^{\,2}} < 1$

Consider a function

$$SCD = \frac{2\lambda_1\lambda_2}{\lambda_1^{\,2} + \lambda_2^{\,2}} Min(\lambda_1, \lambda_2) \tag{3.9}$$

then, we have

$$0 \leq \frac{2\lambda_1\lambda_2}{\lambda_1^{\,2} + \lambda_2^{\,2}} Min(\lambda_1, \lambda_2) \leq Min(\lambda_1, \lambda_2) \tag{3.10}$$

In order to analyze the property of this function, we plot Figure 3.6 and Figure 3.7 in the same way as Figure 3.3 and Figure 3.4. The SCD reaches a sharp peak when $\lambda_1 = \lambda_2$ for each fixed summation $\lambda_1 + \lambda_2$ as red crosses shown in Figure 3.6. All the contours of fixed SCD in relation to the variation of λ_1 and λ_2 in Figure 3.7 have a sharp corner at the position of $\lambda_1 = \lambda_2$ pointing to the origin of the $\lambda_1 \sim \lambda_2$ coordinate system like arrows, i.e. the β angle is small for all the levels of fixed SCDs. This indicates that the SCD is a very strong symmetrical corner point detector but unfortunately it is very exclusive to asymmetrical corner points and edges.

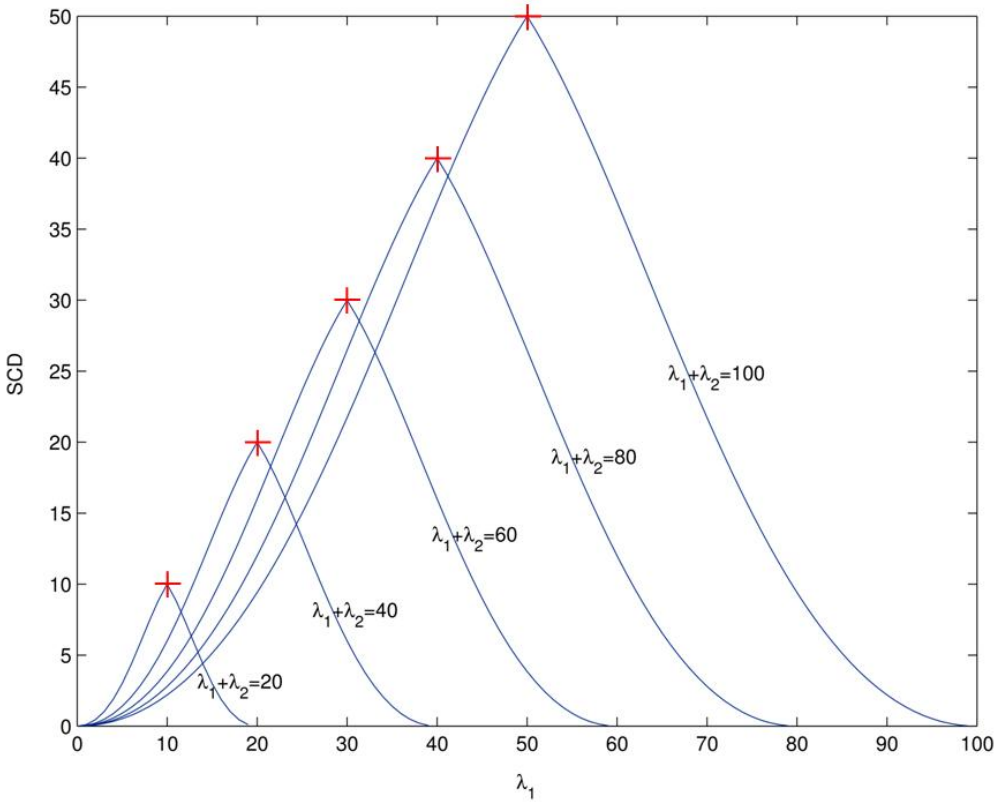


Figure 3. 6: The characteristics of the SCD. λ_1 and λ_2 vary between 0 and 100. The fine lines display the values of the SCD.

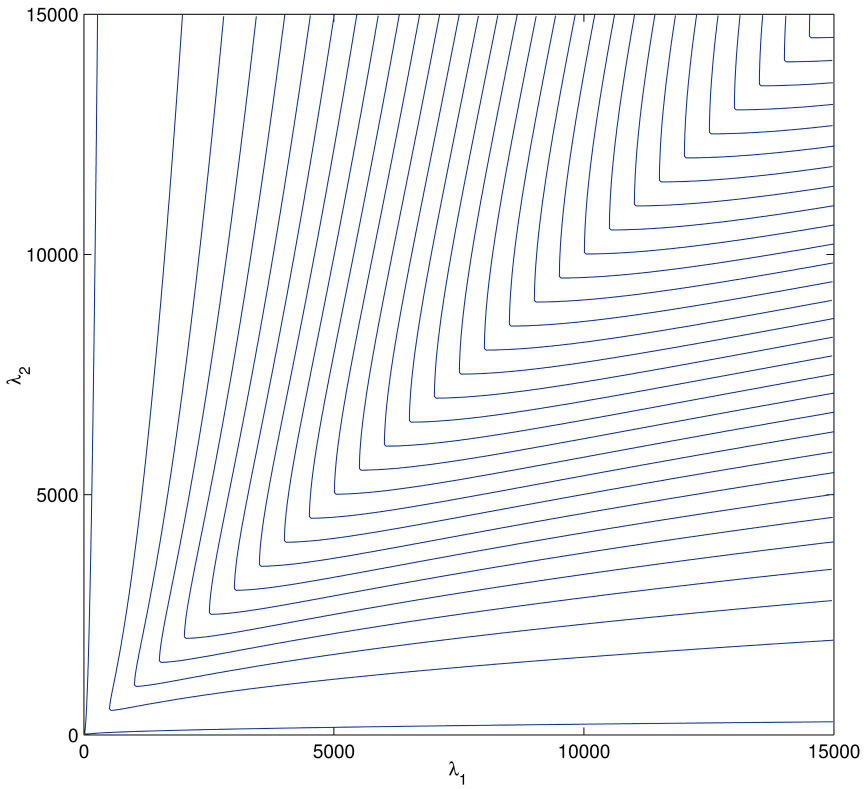


Figure 3. 7: The isoline distribution of the SCD.

Experiment results have proved that the SCD is very prone to select points with two very close eigenvalues and consequently it is more likely to select symmetrical flat points than asymmetrical corner points. The results are quite similar to those obtained using the Harris Corner Detector when the value of α is greater than 0.1. Therefore, a further improvement of response function is required.

3.2.2 An asymmetrical corner detector (ACD)

In order to improve the SCD to a detector which can select asymmetrical interest points, we use $Max(\lambda_1, \lambda_2)$ instead of $Min(\lambda_1, \lambda_2)$ in Eq. 3.9. Thus, definition of a new function which we call asymmetrical corner detector (ACD) is given:

$$ACD = \frac{2\lambda_1\lambda_2}{\lambda_1^{\,2} + \lambda_2^{\,2}} Max(\lambda_1, \lambda_2) \tag{3.11}$$

Obviously,

$$0 \le \frac{2\lambda_1\lambda_2}{\lambda_1^{\,2} + \lambda_2^{\,2}} Max(\lambda_1, \lambda_2) \le Max(\lambda_1, \lambda_2) \tag{3.12}$$

To investigate the property of this function, the plots of the relation between ACD and fixed $\lambda_1 + \lambda_2$ summation varying between 0 and 100 are presented in Figure 3.8. This function has a rather "funny" shape; it reaches twin peaks when both λ_1 and λ_2 are relatively large but slightly different from each other as red crosses displayed. The twin peaks are symmetrical to the position of $\lambda_1 = \lambda_2$ where the ACD value drops into a sharp trough. This indicates that the ACD reaches the maximum at the asymmetrical corner points rather than the ideal symmetrical corner points.

The ACD contours plotted against the variation of λ_1 and λ_2 in Figure 3.9 illustrate the same phenomena. The hip-like curves have a sharp peak at the position of $\lambda_1 = \lambda_2$ pointing away from the origin of the $\lambda_1 \sim \lambda_2$ coordinate system. The ACD is therefore more favourable to asymmetrical corner points than to perfect symmetrical corner points. In particular, the function is more critical to flat region than the Harris Corner Detector. It is obvious that the isoline angle β of the ACD is wide and the isoline

distribution of the ACD can again be divided into one flat region, two edge regions and one corner region but within the corner region, it can be further divided into a symmetrical corner region and two asymmetrical regions (Figure 3.9). The red crosses in Figure 3.9 denote the asymmetrical corner points where the ACD reaches the local maxima with two eigenvalues which are relatively large but slightly different from each other. The red shadowed areas in Figure 3.9 show the asymmetrical combinations of λ_1 and λ_2 that are most likely selected by the ACD and the green shadowed area present the highly symmetrical points where $\lambda_1 = \lambda_2$.

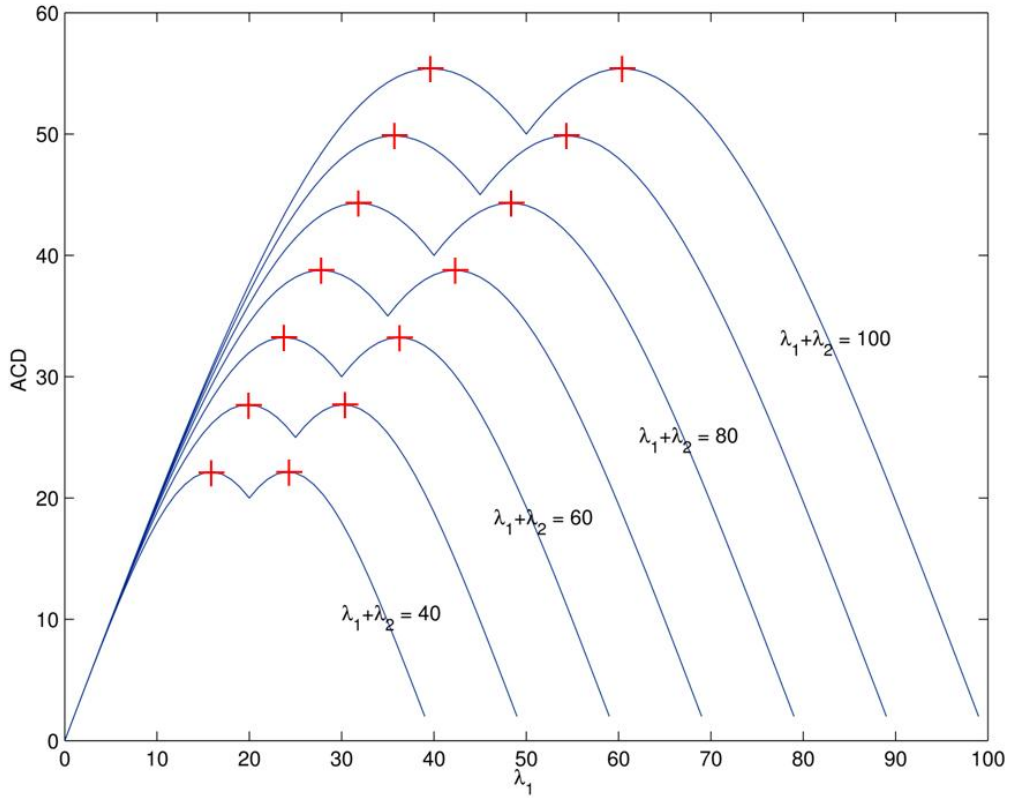


Figure 3. 8: The plot of the relation between the ACD and fixed $\lambda_1 + \lambda_2$ summation varying between 0 and 100. The curves display the values of the ACD.

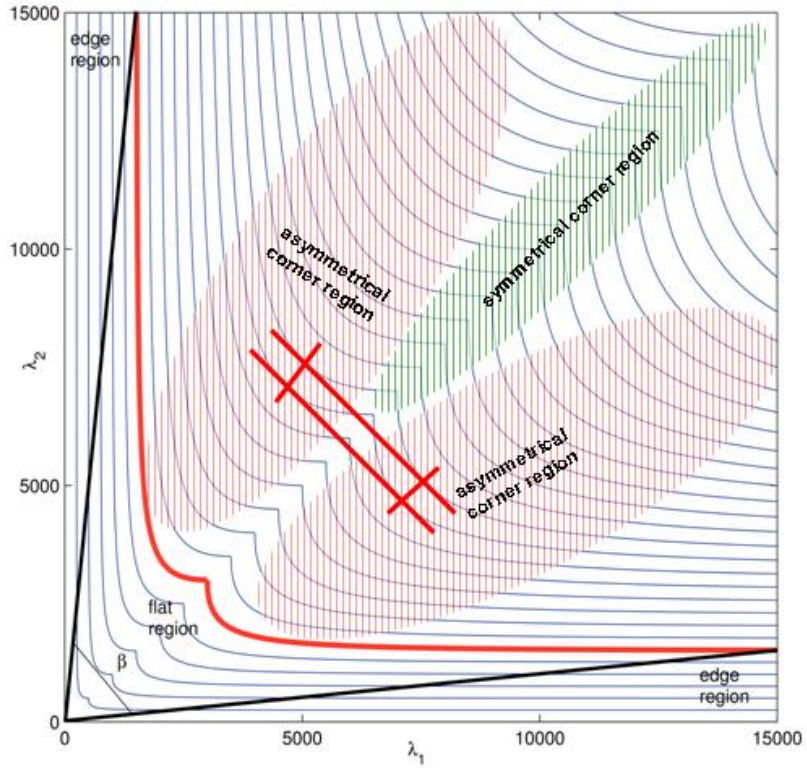


Figure 3. 9: The characteristics of peak areas in the ACD isolines.

The characteristics of the three regions are summarised as below.

(1) In the corner region, the value of the ACD is large because λ_1 and λ_2 are both large but the maximum values are not at the ideal symmetrical positions where $\lambda_1 = \lambda_2$. Based on the maxima of the ACD, this region can be further divided into two asymmetrical corner regions where ACD reaches maximum and a symmetrical region in between the two.

(2) In the edge regions, if one of the two eigenvalues is very small, the ACD value will be small. If one eigenvalue is large, the ACD value increases rapidly when the difference between the two eigenvalues decreases.

(3) In the flat region, λ_1 and λ_2 are both very small and the value of the ACD is also small.

3.4 Experimental results

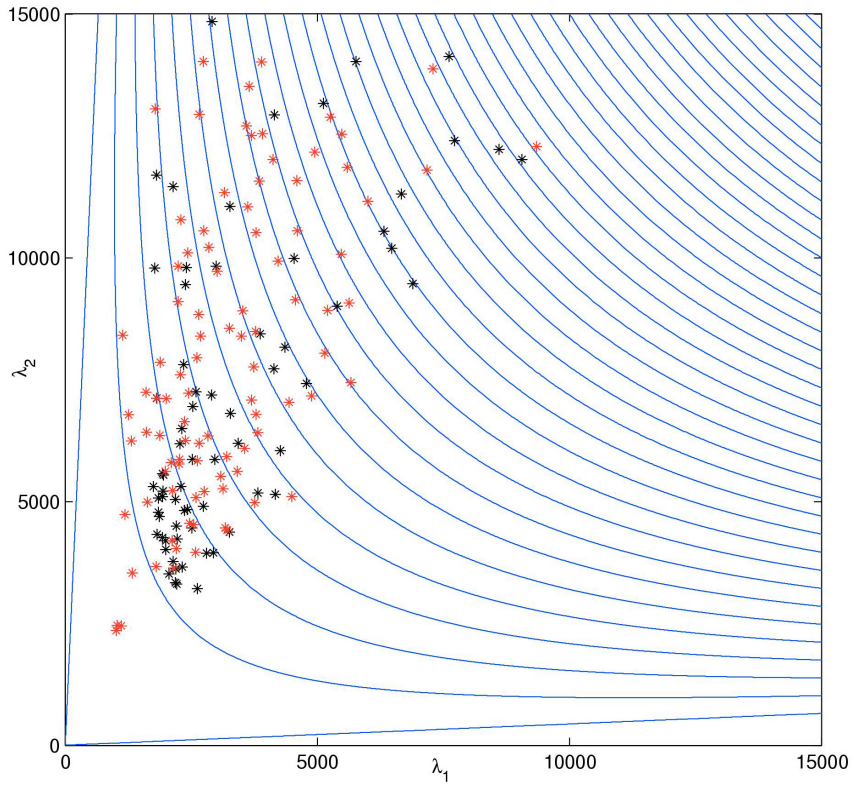
3.4.1 Comparison between the ACD with the Harris Corner Detector

Unlike the Harris Corner Detector, the ACD does not have an empirical parameter and it operates with a threshold set by user as to be discussed later. Asymmetrical corner points are far more common than symmetrical corner points in remotely sensed images of land surface; the design and the property of the ACD are just in response to this reality.

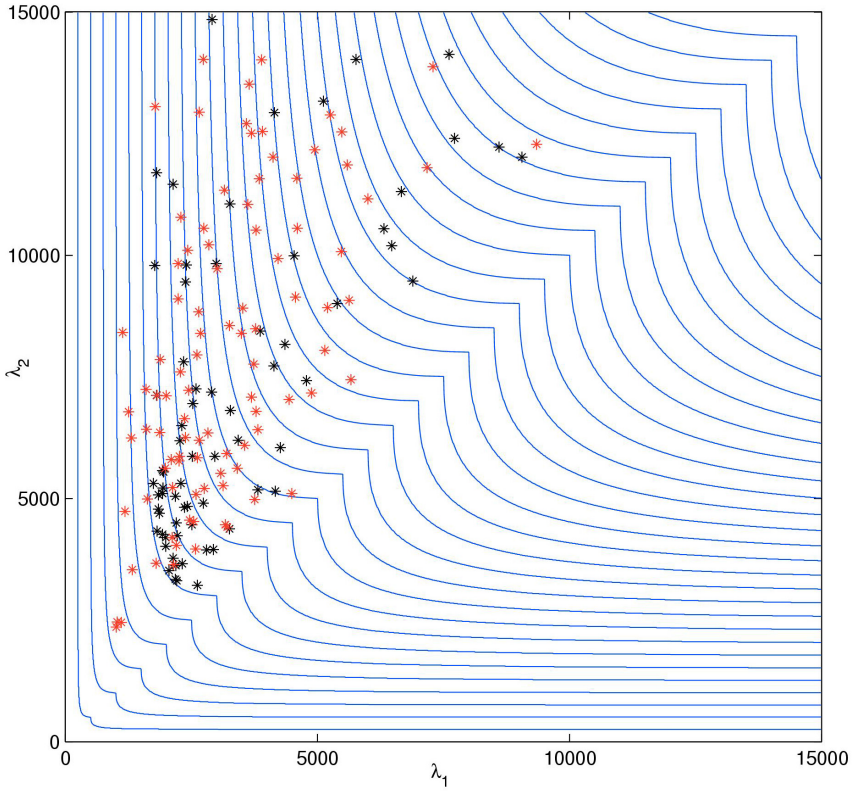
I manually selected large amount of interest points from two satellite images taken by Landsat-7 Enhanced Thematic Mapper plus (ETM+) sensor and Satellite Pour l'Observation de la Terre (SPOT) sensor, and then calculated the values of the response function R of the Harris Corner Detector and that of the ACD. These selected interest points are evenly distributed across the whole images and represent typical distinctive features in different types of areas including mountainous and urban areas. All the manually selected interest points are plotted in both the Harris R function contour chart (Figure 3.10a) and the ACD function contour chart (Figure 3.10b), in which the red dots denote the interest points selected from the SPOT image and the black ones from the Landsat ETM+ image.

As shown in Figure 3.10 (a) and (b), it is obvious, without any deliberation, that these manually selected interest points are rarely ideal symmetrical and therefore most of them are away from the maximum region in the central part of the R contour chart but just fall in the maximum value region in the ACD contour chart. The pixels with two relatively large but not equal eigenvalues are more likely to be selected by the ACD. We therefore expect that the ACD will over perform the Harris Corner Detector in detection of interest points on remote sensing images.

Note: The figures are generated using Matlab 7.0. When calculating eigenvalues of an auto-correlation matrix, Matlab automatically assigns the smaller eigenvalue to λ_1 and the larger eigenvalue to λ_2. Therefore, all the points fall into the upper half of contour graphs.



(a)



(b)

Figure 3. 10: Manually selected interest points plotted on (a) the Harris Corner Detector *R* contour chart; (b) the ACD contour chart. The red dots are the interest points selected from a SPOT image and the black ones are extracted from a Landsat ETM+ image. (The values of each isoline in both charts are different)

3.4.2 Experimental results and assessment

Several images with various types of ground objects, taken by Landsat-7 Enhanced Thematic Mapper plus (ETM+) sensor and Satellite Pour l'Observation de la Terre (SPOT) sensor were used to evaluate the performance of the ACD in comparison with the Harris Corner Detector.

Landsat-7 ETM+ and SPOT sensors

In order to effectively assess the interest point detection performance of the ACD on remote sensing imagery data, let us look into the characteristics and parameters of the Landsat-7 ETM+ and SPOT sensors and platforms first.

The ETM+ sensor is mounted on the Landsat-7 platform located in a near polar, sun synchronous, circular orbit with an altitude of 705km, orbit inclination of 98.2 degrees and a 16-day repeat cycle. The scanning mechanism of the ETM+ builds up an image pixel-by-pixel, line-by-line by utilising a cross-track scanning action. It has seven channels covering the visible, near- and mid-infrared, and the thermal infrared (Table 3.1). In addition, a 15 m resolution panchromatic band is also included to improve the spatial resolution and enhance the resolving power. All the data are quantised onto a 0 ~ 255 range.

Table 3. 1: Wavebands and their characteristics of Landsat-7 ETM+ sensor

Band	*Spectral (μm)*	*Description*	*Spatial resolution (m)*	*Main application*
1	0.45 ~ 0.52	Visible (blue~green)	28.5	soil, vegetation
2	0.52 ~ 0.60	Visible (green)	28.5	vegetation
3	0.63 ~ 0.70	Visible (red)	28.5	road, bare soil, vegetation
4	0.75 ~ 0.90	Near-infrared	28.5	soil
5	1.55 ~ 1.75	Mid-infrared	28.5	road, bare soil, water
6	10.4 ~ 12.5	Thermal-infrared	60	
7	2.08 ~ 2.35	Mid-infrared	28.5	rock, mineral
8	0.52 ~ 0.90	Panchromatic	15	

The SPOT system is placed in a near polar, sun-synchronous orbit at an altitude of 832km and an inclination of 98.7 degrees. The orbital repeat period is 26 days and equatorial crossing time of 10.30 A.M. It carries two identical High Resolution Visible (HRV) imaging sensors with push broom scanners inside. HRV can be operated in two modes. In panchromatic mode, the sensor is sensitive across a broad spectral band from 0.51 to 0.73 μm with a spatial resolution of 10 m. In the multispectral mode, the HRV senses three spectral regions: green, red and near-infrared at a spatial resolution of 20 m (Table 3.2). The SPOT system is designed to provide data for land-use studies, assessment of renewable resources, and exploration of geological resources.

Table 3. 2: Wavebands and their characteristics of SPOT sensor

Band	*Spectral (μm)*	*Description*	*Spatial resolution (m)*
1	0.50 ~ 0.59	Multispectral (green)	20
2	0.61 ~ 0.68	Multispectral (red)	20
3	0.79 ~ 0.89	Multispectral (infrared)	20
4	0.51 ~ 0.73	Panchromatic	10

Data Description

Two datasets are used to evaluate the performance of the ACD compared with the Harris Corner Detector.

The first study area is in the southern part of Almeria Province, south Spain (Example 1 ~ 5). It is composed of mountainous, urban and field area along the coast. Part of the mountainous area is semi-arid and thinly vegetated and the spectral reflectance is dominated by soil and outcrops of rock. The other part of the mountainous area, which is closed to the urban area, is covered by various kinds of vegetation. Therefore, visible green, red bands (band 2, 3) and near-infrared band (band 4) taken by Landsat-7 ETM+ sensor which are widely applied for soil, vegetation and roads observation are suitable to evaluate the point detection performance of the ACD. In addition, two images from SPOT multispectral and panchromatic bands were involved in the evaluation of the performance.

The other dataset is provided by ERMapper 7.0 (Example 6). The study area is composed of flat urban area, vegetation, and field area along the coast. This area contains characteristic and distinctive features such as complex intersections of major roads, edges of land-cover parcels, etc. A high spatial resolution panchromatic band image taken by SPOT sensor is used to evaluate the performance of the ACD compared with the Harris Corner Detector.

Parameter setting

In all the experiments, the parameters of the Harris Corner Detector are set to the standard default values recommended by the author (Harris and Stephens, 1988). The α parameter of the Harris Corner Detector response function R is set to 0.04 and the global threshold is set to 1% of the maximum observed interest strength (R value) in the whole image (Schmid *et al.*, 2000). Normally, global threshold may yield small number of interest points, which are not evenly distributed over the whole image. Therefore, we also used the window based dynamic local threshold which is the 1% of the maximum observed R value in a fixed size window of 41×41.

For the ACD, several synthetic images containing typical features are used to determine an effective and reasonable threshold. In these synthetic images, there are a few points representing ideal distinctive features that can be regarded as good quality interest points. The ACD values of these points are calculated and subsequently, the smallest ACD value for these assumed good interest points is set as the ACD fixed threshold. As such, the global ACD threshold is set to 7200, which is used in the following interest point detection experiments and also the proposed semi-automatic image co-registration scheme.

Contrast enhancement is applied to secure the effectivity of interest point detection. Same Gaussian with the standard deviation σ of 0.5 is used for both detectors to smooth the image and weight the derivatives summed over the window. In addition, in order to prevent interest points clustering in a small area or assure the sparseness of interest points selected, non-maximum suppression with a 7×7 mask is recommended. The method extracts the central pixel if its interest strength is the largest one among its neighbourhood in the given 7×7 window.

(1) Mountainous area (Landsat-7 ETM+ sensor)

A 512×512 Landsat-7 ETM+ visible band (band 2) image of southeast Spain was used to test the performance of the ACD. The spatial resolution of the image is 28.5 metres. As shown in Table 3.3 and Figure 3.11, 2038 interest points were selected using the Harris Corner Detector with global threshold, the number is increased to 2473 with a window dynamic threshold while for the ACD with a fixed threshold of 7200, 2627 interest points were extracted.

Experimental results have indicated that the Harris Corner Detectors with both the global threshold and window dynamic threshold are not capable to detect adequate interest points in a mountainous area as zoom-in images displayed in Figure 3.12(1). Points which fell into the red circles are typical and distinctive interest points along the mountain ridges. Although the mountainous area has relatively low contrast features, the ACD still effectively selected these good quality asymmetrical interest points as clearly presented in a further enlargement in Figure 3.13. The red circles highlight the typical asymmetrical interest point which is selected only by the ACD. The blue circles denote the points selected by both the ACD and the Harris Corner Detector with window dynamic threshold while the Harris Corner Detector with global threshold failed.

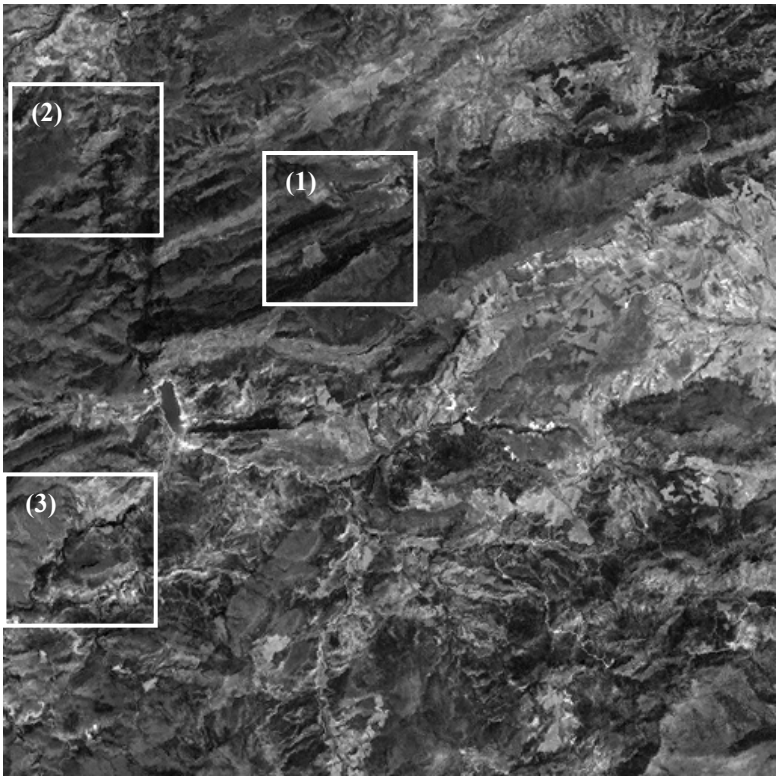
In Figure 3.12(2), the Harris Corner Detector with window dynamic threshold yields large number of well spread interest points but many of them are of rather low quality without obvious contrast to their neighbourhoods as highlighted in the red circles.

On the other hand, as shown in Figure 3.12(3) in another zoom-in display, the Harris Corner Detector with window dynamic threshold and global threshold failed to select all the good quality interest points as highlighted by the red circles. In comparison, the ACD extracted reasonable quantity of evenly distributed good quality interest points.

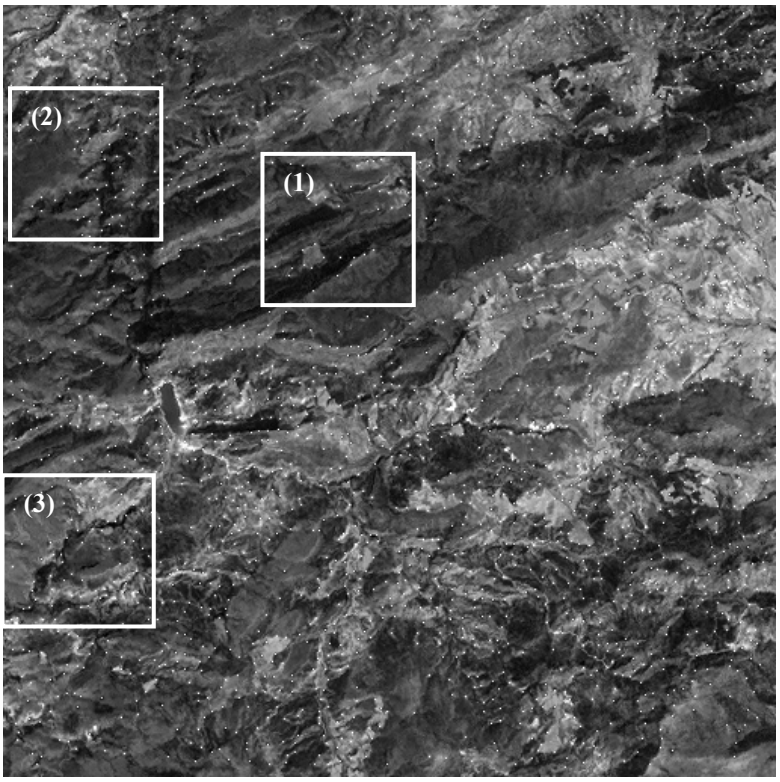
From the above comparison, it is obvious that the ACD can effectively select high quality interest points in the mountainous area and eliminates points with low contrast to the neighbourhoods. Additionally, the ACD showed better performance in terms of computation time as listed in Table 3.3.

Table 3. 3: Numbers of points and computation time of interest point detection on satellite images (mountainous area) obtained using the Harris Corner Detector and the ACD.

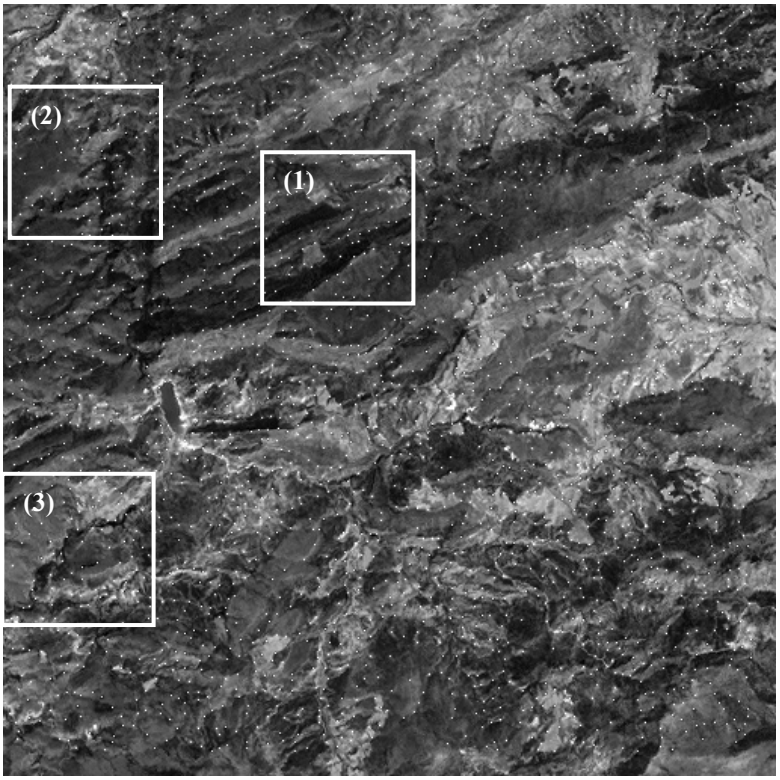
Image			Landsat ETM+ Visible	SPOT Panchromatic
Size(pixel)			512 × 512	1024 × 1024
Resolution(m)			28.5	10
Harris Corner Detector	*Global Threshold*	*Number of Points*	2038	5737
		Computation Time(s)	1.21	3.78
	Window Dynamic Threshold	*Number of Points*	2473	9511
		Computation Time(s)	1.35	6.48
ACD		*Number of Points*	2627	8076
		Computation Time(s)	1.17	3.58



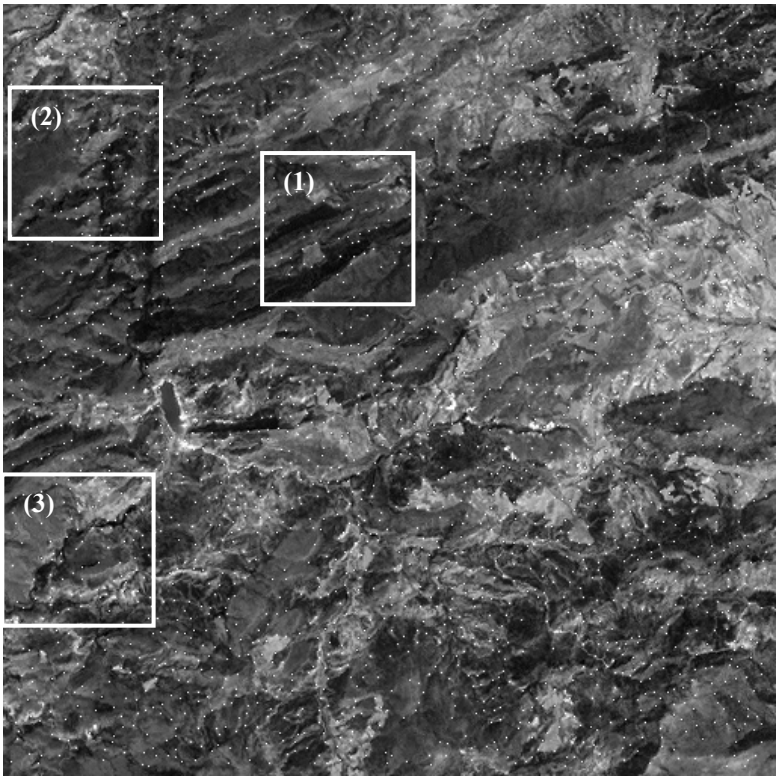
(a)



(b)

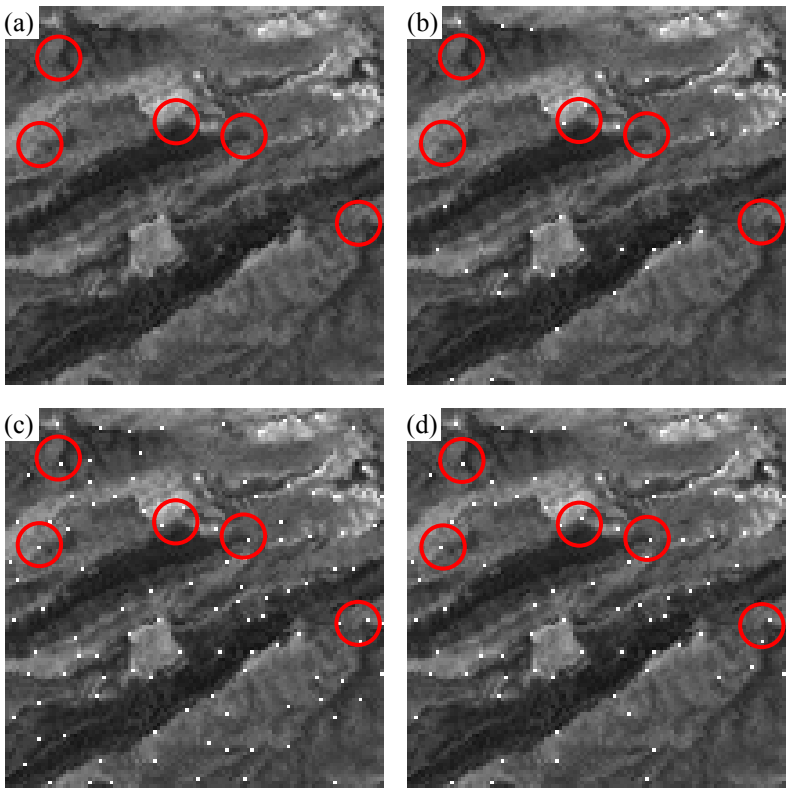


(c)

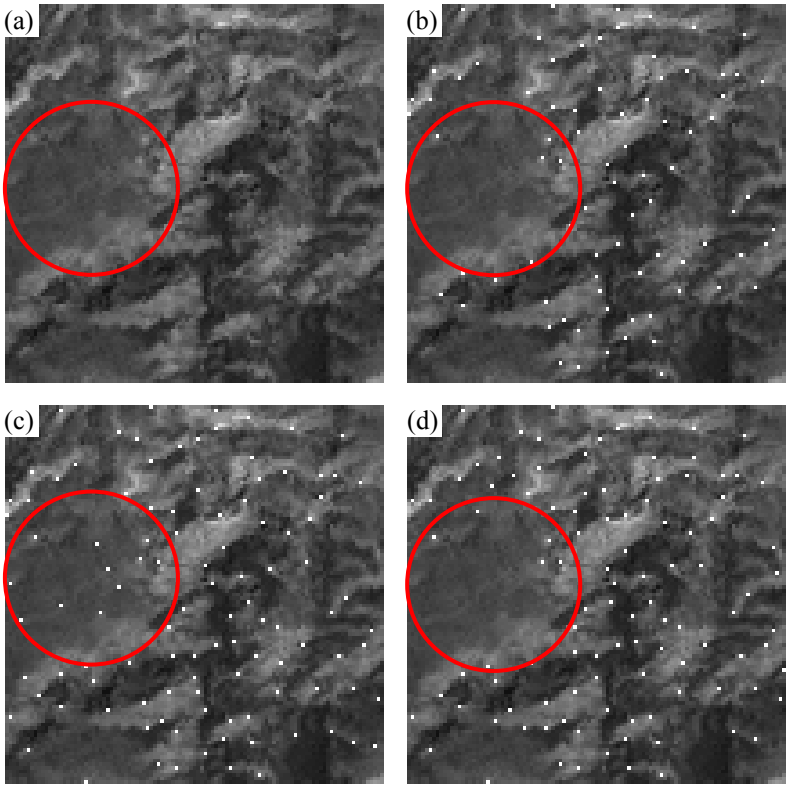


(d)

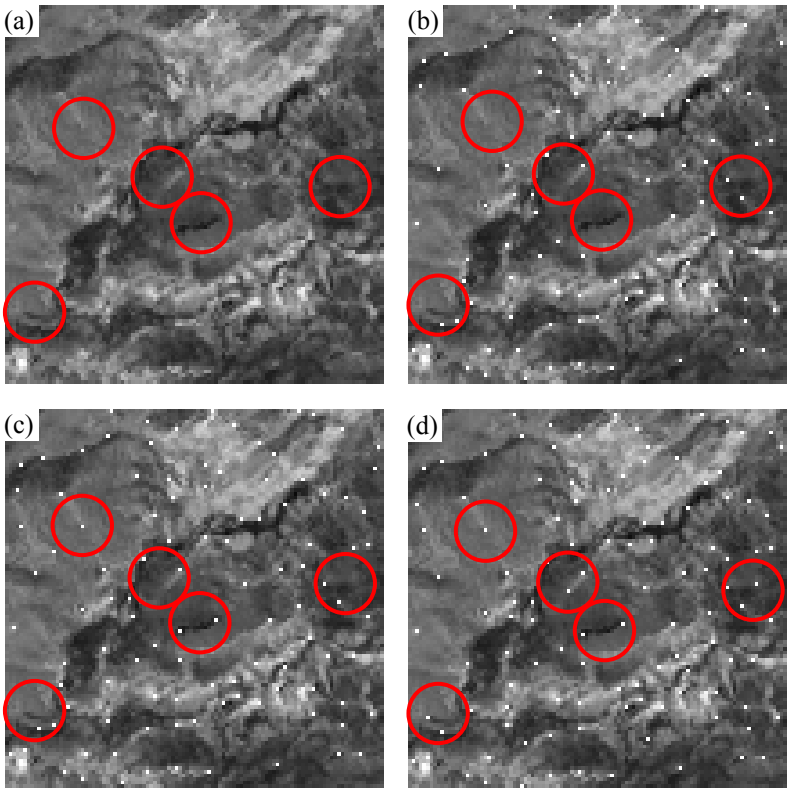
Figure 3. 11: Interest points displayed as the white dots. White boxes denote zoom-in sub areas displayed in Figure 3.12. (a) Initial Landsat-7 ETM+ visible band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Results obtained using the ACD.



(1)



(2)



(3)

Figure 3. 12: Zoom-in sub areas. Interest points showed as the white dots. Red circles highlight the points for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD.

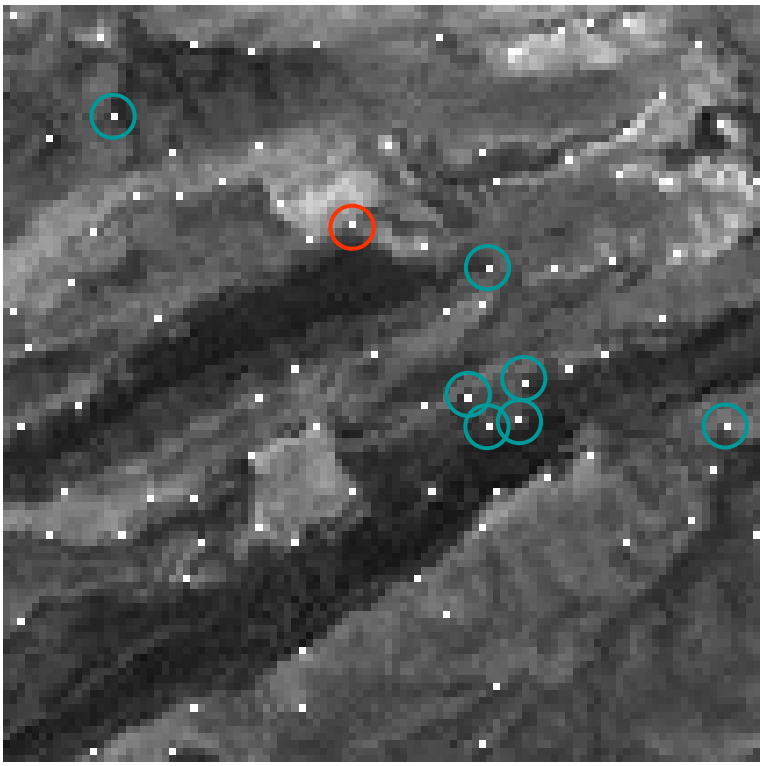


Figure 3. 13: Further zoom-in observation of interest points selected using the ACD and the Harris Corner Detector. The red circles locate the very good quality interest point which is detected only by the ACD. The blue circles denote the points selected by both the ACD and the Harris Corner Detector with window dynamic threshold while the Harris Corner Detector with global threshold failed.

(2) Mountainous area (SPOT sensor)

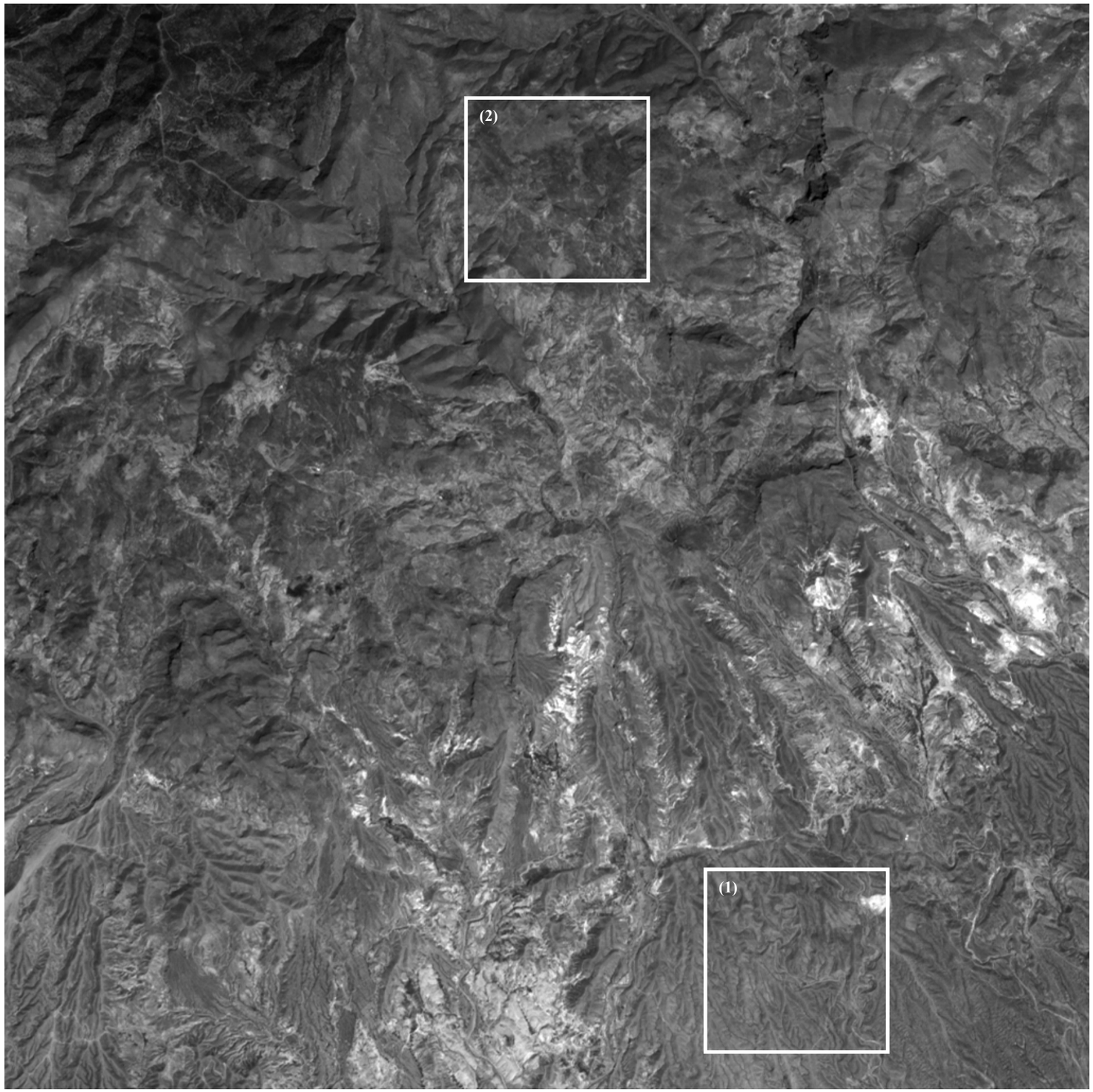
We also used a 1024×1024 SPOT panchromatic band image of southeast Spain to evaluate the performance of the ACD on mountainous area. This image has a higher spatial resolution which is 10 metres. As presented in Figure 3.14 b,c and Table 3.3, the Harris Corner Detector selected 5737 and 9511 interest points with global and window dynamic thresholds, respectively. The ACD detected 8076 interest points as shown in Figure 3.14d.

In this case, the Harris Corner Detector with global threshold is unfortunately not efficient to select adequate interest points especially for the peak points along the ridges as zoom-in images displayed in Figure 3.15(1). As the resolution becomes higher, a large number of symmetrical interest points transform to asymmetrical interest points. It therefore largely restrains the performance of the Harris Corner Detector. The three red circles highlight the points which have relative low contrast but are good quality interest points. However, these highly asymmetrical interest points are failed to be detected by the Harris Corner Detector with global threshold.

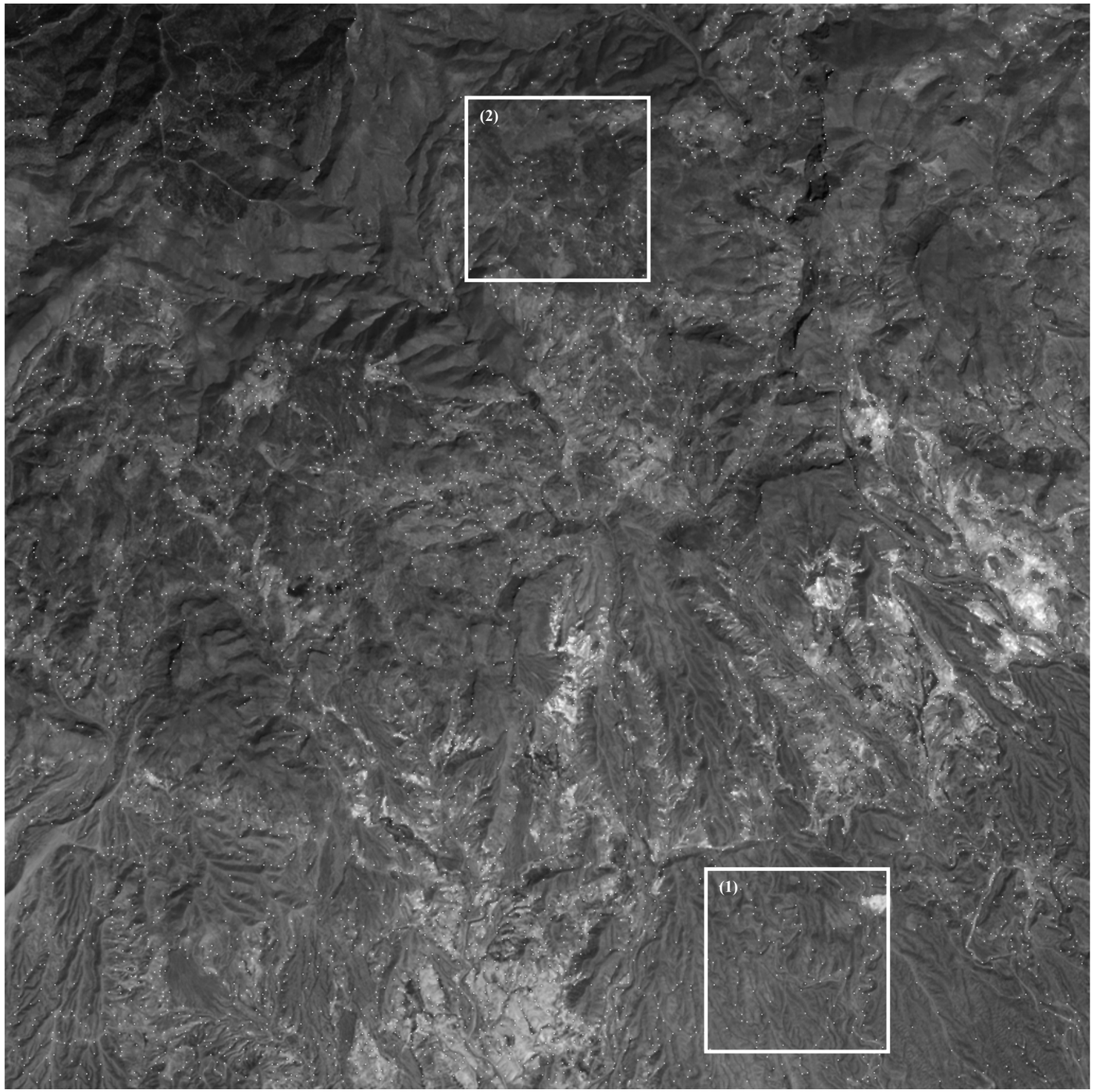
On the other hand, although the Harris Corner Detector with window dynamic threshold can select plenty of interest points along the mountain ridges, it still extracted large number of flat but highly symmetrical interest points as the zoom-in sub area displayed in Figure 3.15(2). In addition, several typical edge points highlighted by the upper red circle are selected by the Harris Corner Detector with window dynamic threshold.

The performances have shown that the Harris Corner Detector is not effective to select high quality Ground Control Points (GCPs). In comparison, the ACD successfully selected adequate and well distributed interest points with good quality and effectively eliminated symmetrical flat interest points in the mountainous area.

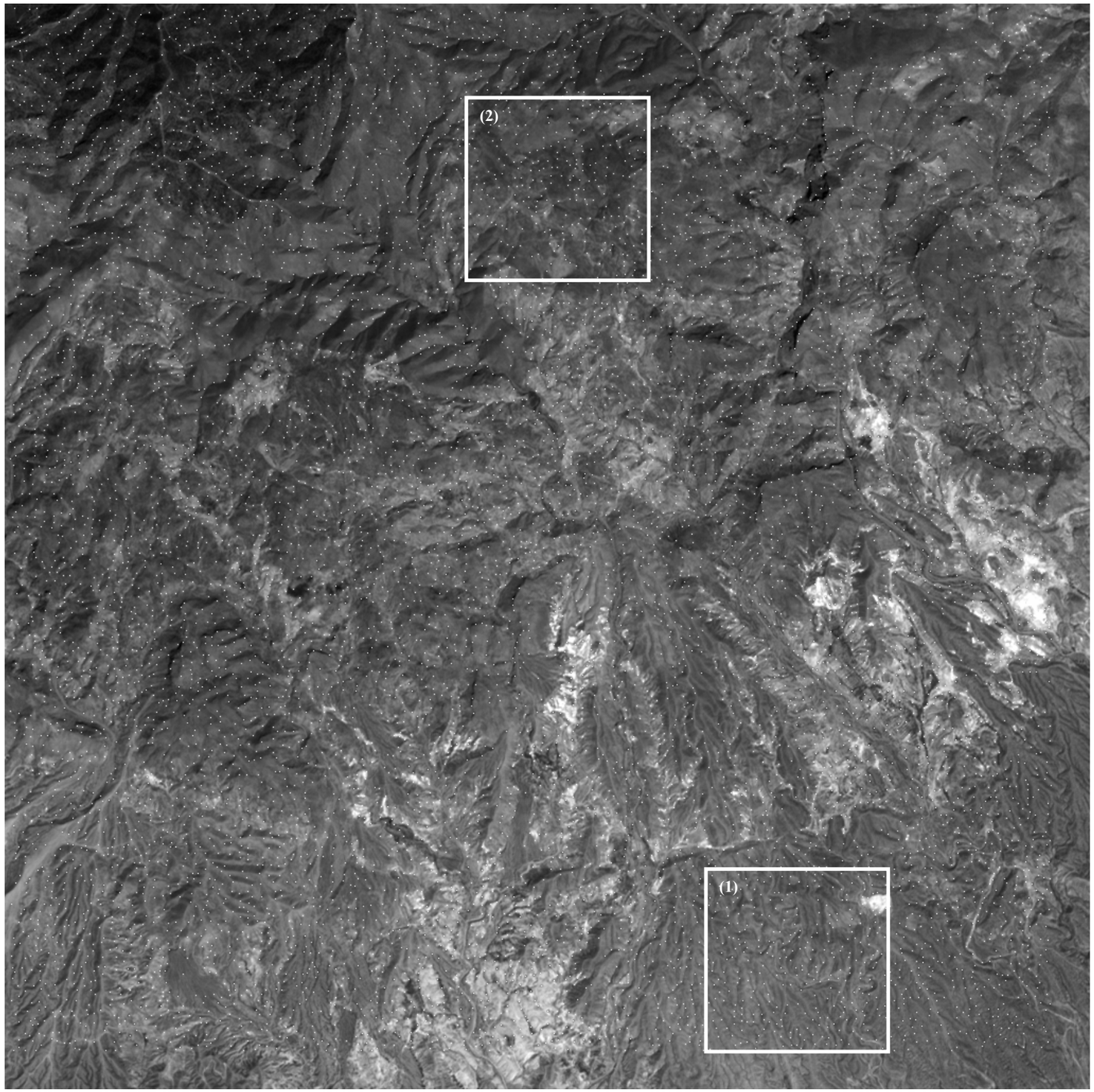
As shown in Table 3.3, even though the computational loading increases with the growing size of the images, comparison has demonstrated that the ACD offers the best performance in the computational efficiency as well as effectiveness of interest point selection.



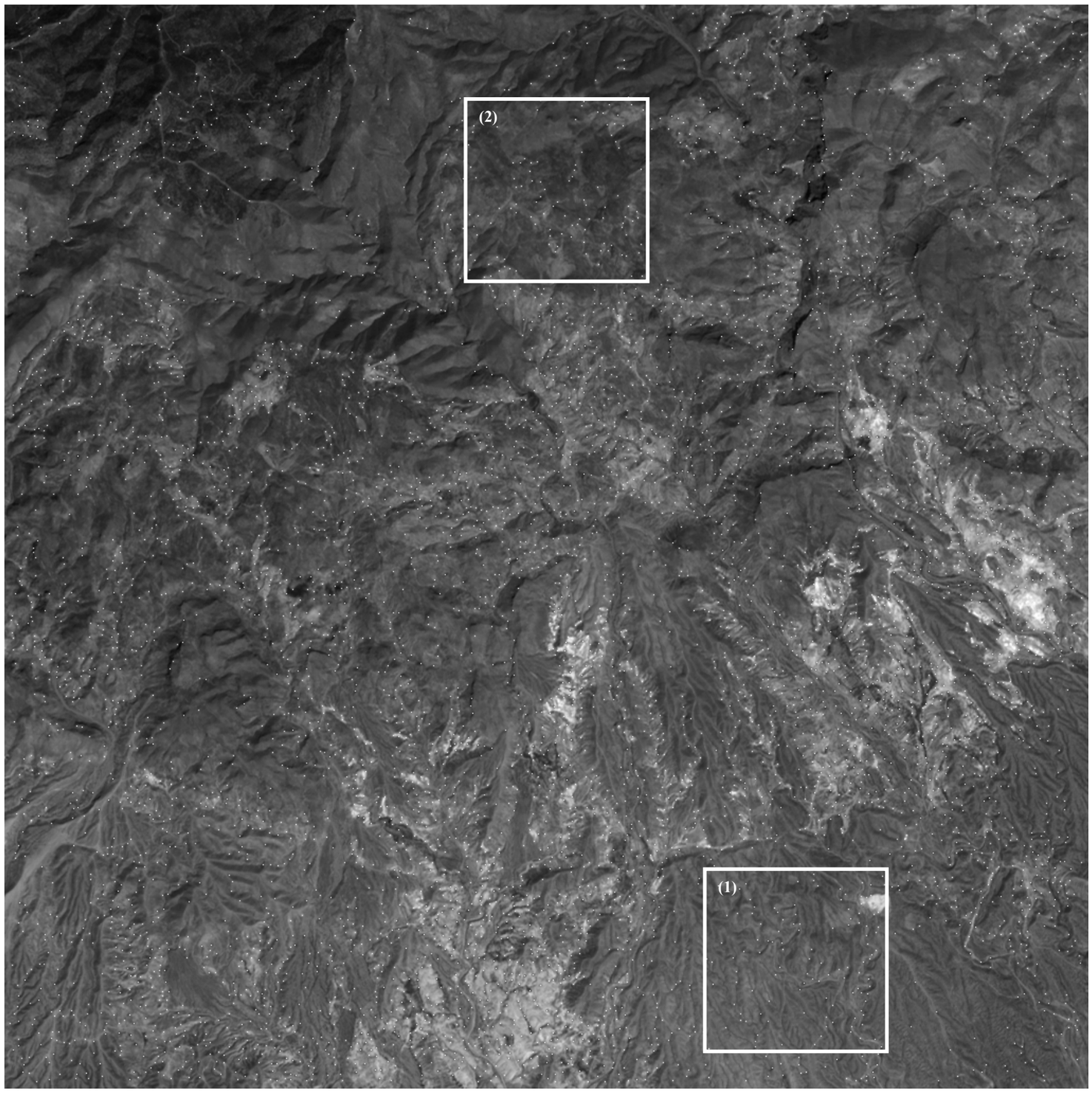
(a)



(b)

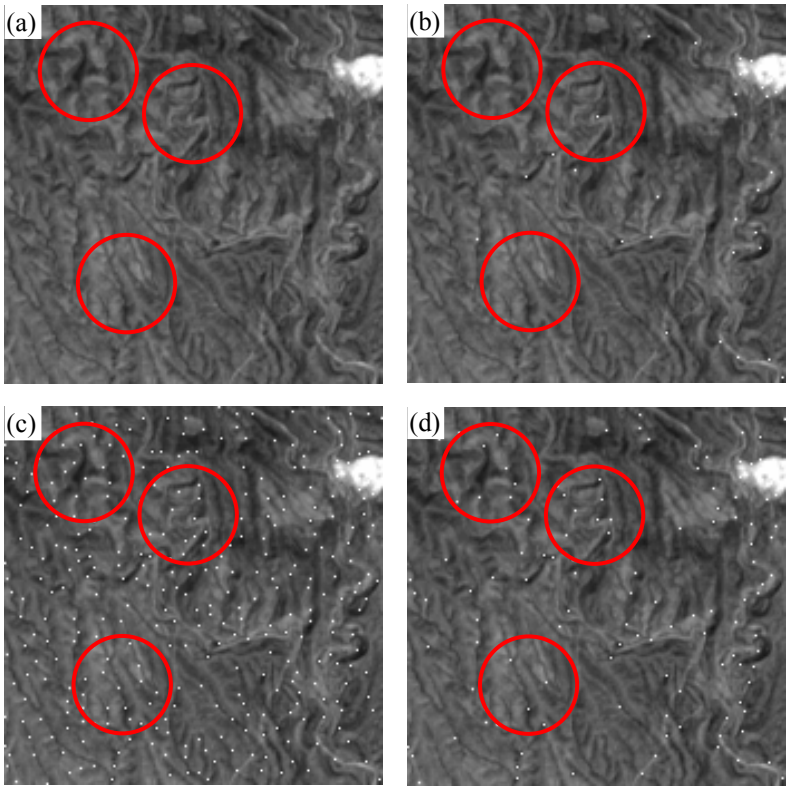


(c)

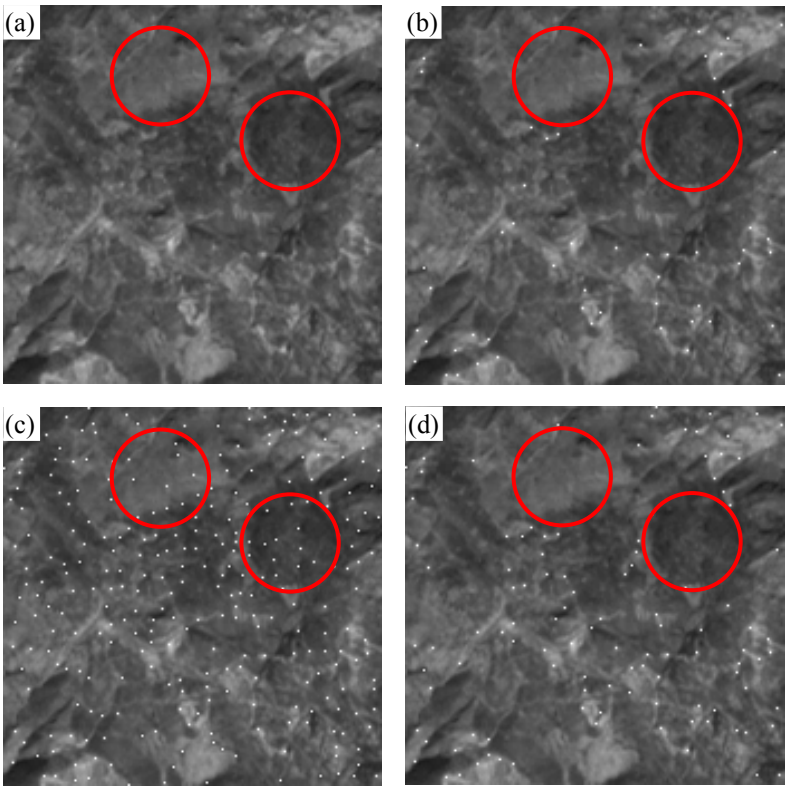


(d)

Figure 3. 14: Interest points displayed as the white dots. (a) Initial SPOT panchromatic band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. Zoom-in sub areas are displayed in white squares.



(1)



(2)

Figure 3. 15: Interest points displayed as the white dots. Red circles highlight the points for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD.

(3) Mountainous area and urban area (Landsat-7 ETM+ sensor)

A 512×512 Landsat-7 ETM+ Near-Infrared band image composed of both mountainous and urban areas was used to assess the efficiency and effectiveness of the ACD. As compared in Table 3.4, the Harris Corner Detector extracted 1123 interest points with global threshold, the number of interest points is increased to 2067 when using window dynamic threshold.

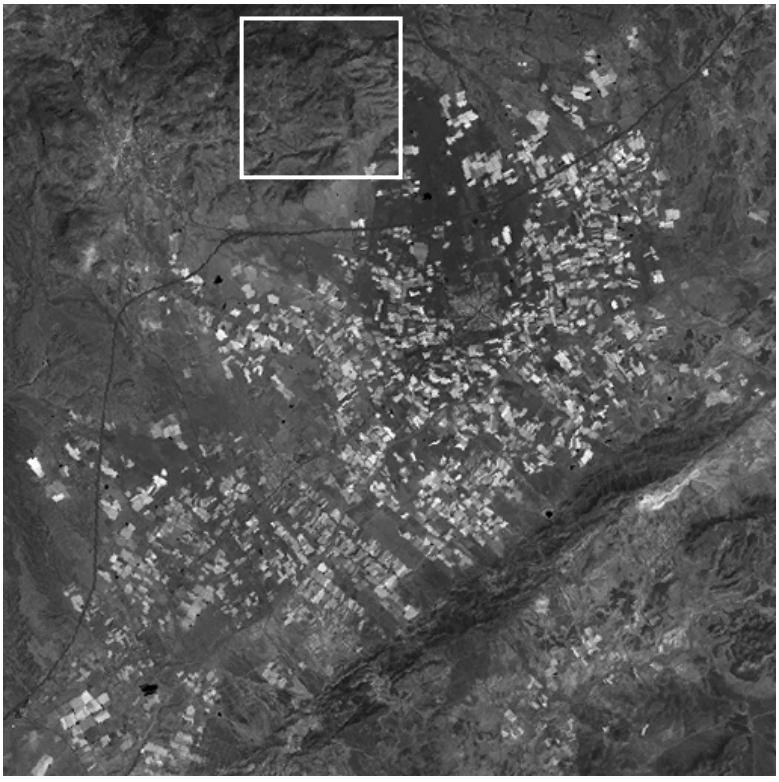
As illustrated in Figure 3.16b, the Harris Corner Detector with global threshold failed to extract evenly distributed points. Normally, the manmade features have higher contrast to their neighbourhood compared to the mountainous area. Consequently, the interest points selected are concentrated in the urban area. Few interest points in the mountainous area are extracted as shown in the zoom-in images in Figure 3.17b. In this case, the Harris Corner Detector with window dynamic threshold performed much better in interest point selection (Figure 3.16c).

The ACD efficiently selected good quality interest points in both urban and mountainous areas and effectively controlled the balance between the point clustering in the urban area and sparseness in the mountainous area. In comparison, the ACD again demonstrated best performance in both interest point detection and computational efficiency. 2104 interest points are selected which are evenly distributed across the whole image (Figure 3.16d and Figure 3.17d) in 1.14 seconds.

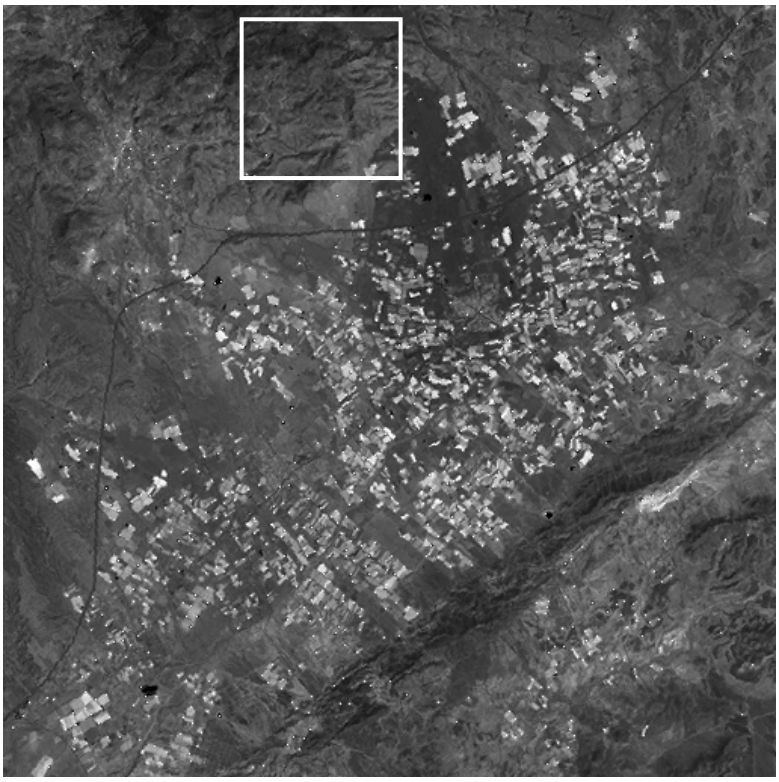
One reason for the failure of the Harris Corner Detector is that the thresholds used are 1% of the maxima of the observed interest strength in the whole image or in the fixed size windows. In fact, if there is a very typical corner feature in the area, the threshold will become quite high and consequently the detector is too critical to detect many of the good quality interest points. Due to the uncertainty of α, the actual range of the response function R of the Harris Corner Detector cannot be pre-determined and the detector often operates with an ineffective threshold. It therefore largely restricts its performance of interest point detection.

Table 3. 4: Numbers of points and computation time of interest point detection on a Landsat ETM+ Near Infrared band image (mountainous and urban areas) obtained using the Harris Corner Detector and the ACD.

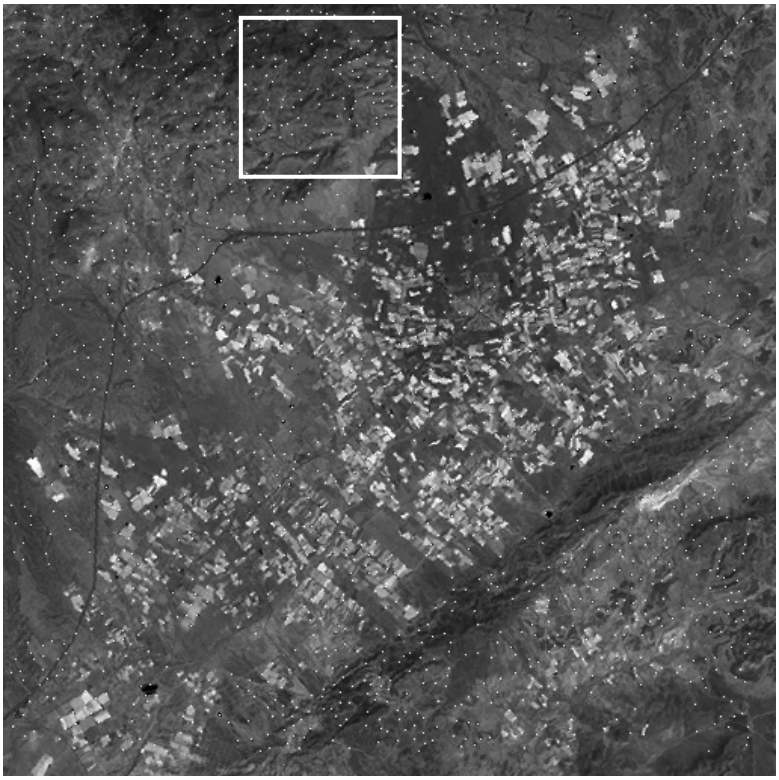
Image			Landsat-7 ETM+ Near-Infrared
Size(pixel)			512 × 512
Resolution(m)			28.5
Harris Corner Detector	*Global Threshold*	*Number of Points*	1123
		Computation Time(s)	1.20
	Window Dynamic Threshold	*Number of Points*	2067
		Computation Time(s)	1.34
ACD		*Number of Points*	2104
		Computation Time(s)	1.14



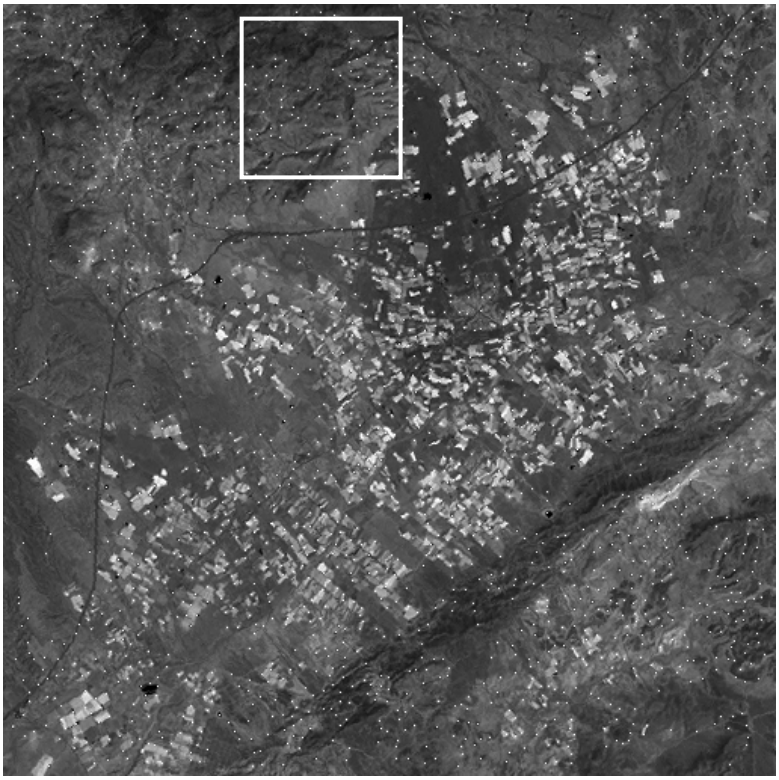
(a)



(b)



(c)



(d)

Figure 3. 16: Interest points displayed as the white dots. (a) Initial Landsat-7 ETM+ NIR band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. White square highlights the zoom-in area for comparison.

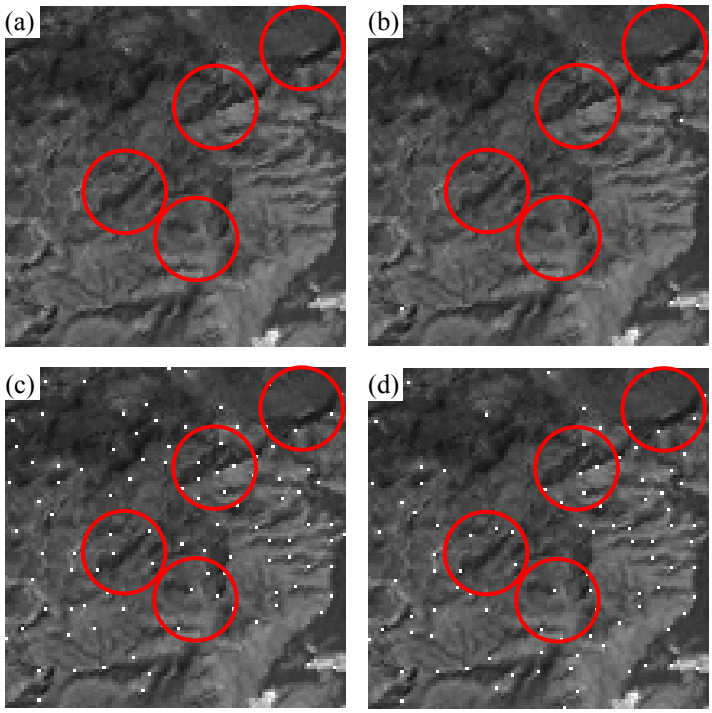


Figure 3. 17: Interest points displayed as the white dots. Red circles highlight the areas for comparison. (a) Initial zoom-in image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD.

(4) Combined area of mountainous, urban and sea (Landsat-7 ETM+ sensor)

To further evaluate the robustness of the ACD, a 512×512 Landsat-7 ETM+ visible band (band 3) image composed of mountainous, urban and sea areas in the Almeria area, southeast Spain was tested. Depending on the spectral characteristics of the visible red band, this image contains distinctive and easily recognized features which can be applied for road observation, soil and vegetation analysis.

As shown in Figure 3.18d, the ACD successfully selected 1347 interest points which are well distributed across the whole image, no matter in what landscape and land covers except the flat sea area. The selected interest points include representative building corners, road intersections in the urban area and ridge peak points in the mountainous area. Although the mountainous area has a lower contrast compared to the urban area, the ACD still achieved a satisfied amount of interest points as the zoom-in images presented in Figure 3.19 and successfully avoided point clustering in the urban area. In particular, there is no flat point selected in the sea by the ACD with a fixed threshold of 7200.

As presented in Table 3.5 and Figure 3.18b-c, with global and window dynamic thresholds, the Harris Corner Detector extracted 974 and 2588 interest points, respectively. As the two zoom-in comparisons displayed in Figure 3.19, in the mountainous area with relatively lower contrast compared to the urban area, the Harris Corner Detector with global threshold failed to select adequate typical interest points e.g. inflexions and ridge intersections as highlighted by the red circles. In this case, the Harris Corner Detector with window dynamic threshold performs much better. A further enlargement comparison of the interest points selected in the mountainous area is displayed in Figure 3.20. The blue circles locate several very good quality mountains interest points which are selected by both the ACD and the Harris Corner Detector with window dynamic threshold while the Harris Corner Detector with global threshold failed. However, the Harris Corner Detector with window dynamic threshold selected large amount of poor quality flat points in the sea area.

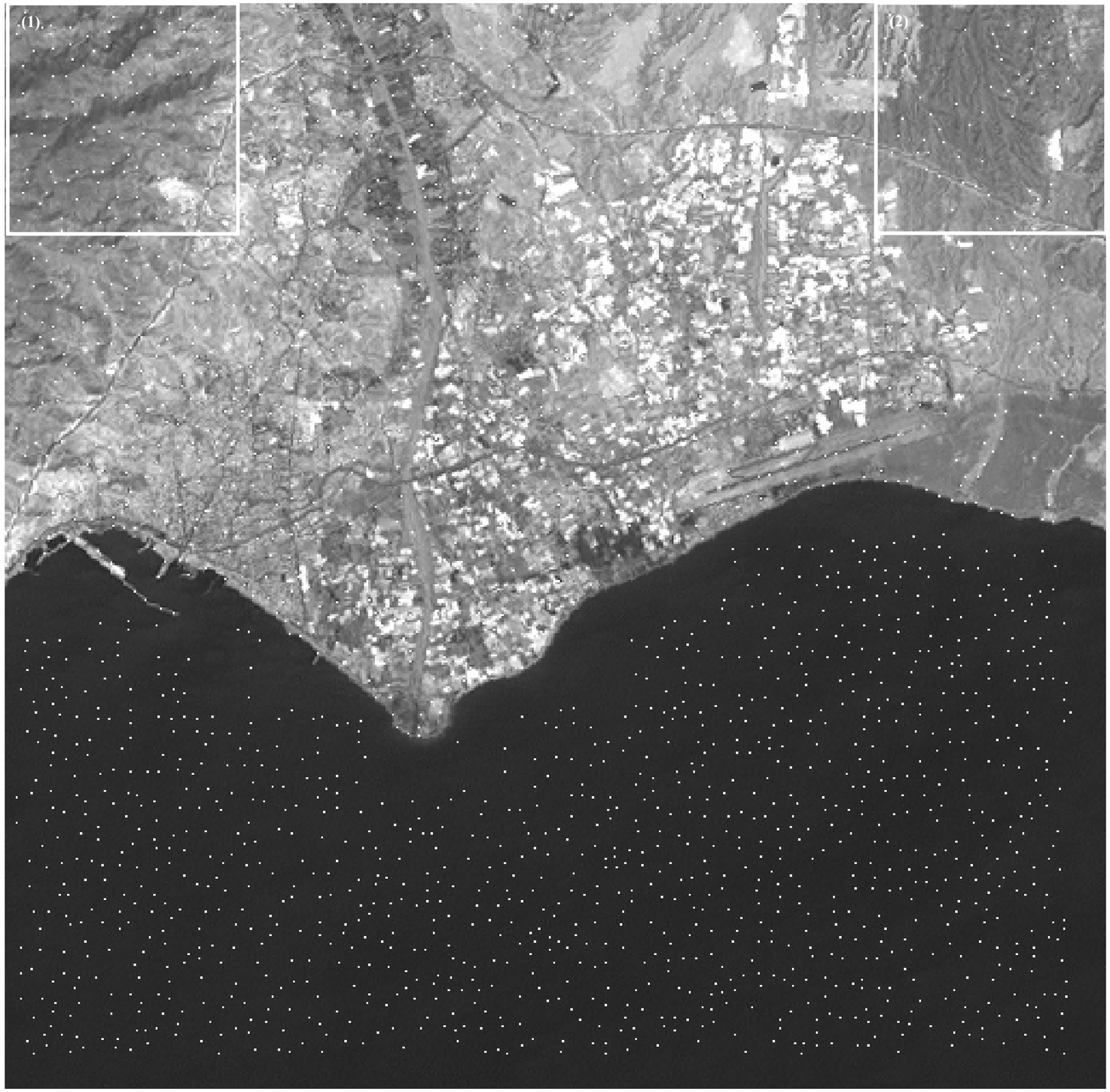
All in all, the results have indicated that the ACD is superior to the Harris Corner Detector in terms of interest point detection in the whole image.



(a)



(b)



(c)

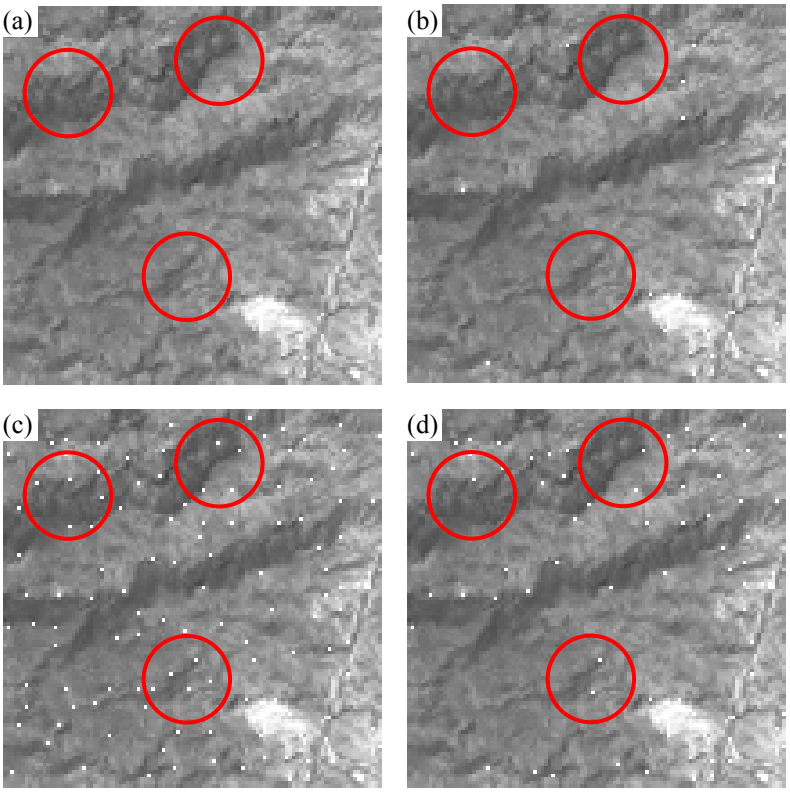


(d)

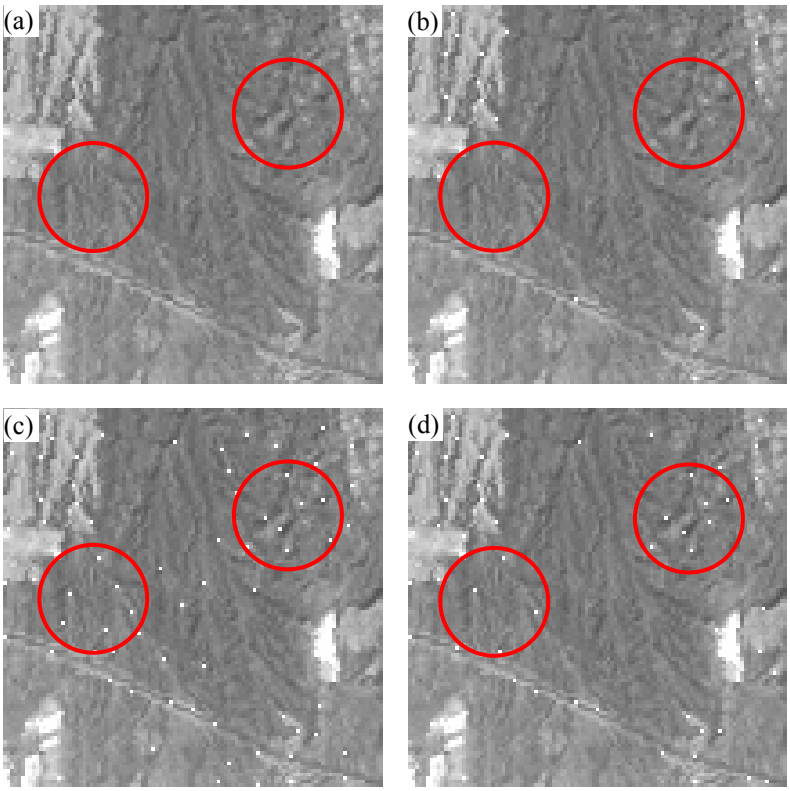
Figure 3. 18: Interest points displayed as the white dots. (a) Initial Landsat-7 ETM+ visible band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. White squares denote the zoom-in areas for comparison.

Table 3. 5: Numbers of points and computation time of interest point detection on satellite images (mountainous, urban and sea areas) obtained using the Harris Corner Detector and the ACD.

Image			Landsat ETM+ Visible	SPOT Multispectral
Size(pixel)			512 × 512	512 × 512
Resolution(m)			28.5	20
Harris Corner Detector	*Global Threshold*	*Number of Points*	974	772
		Computation Time(s)	1.27	1.64
	Window Dynamic Threshold	*Number of Points*	2588	7267
		Computation Time(s)	1.56	3.67
ACD		*Number of Points*	1347	1720
		Computation Time(s)	1.12	1.57



(1)



(2)

Figure 3. 19: Interest points displayed as the white dots. Red circles highlight the points for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD.

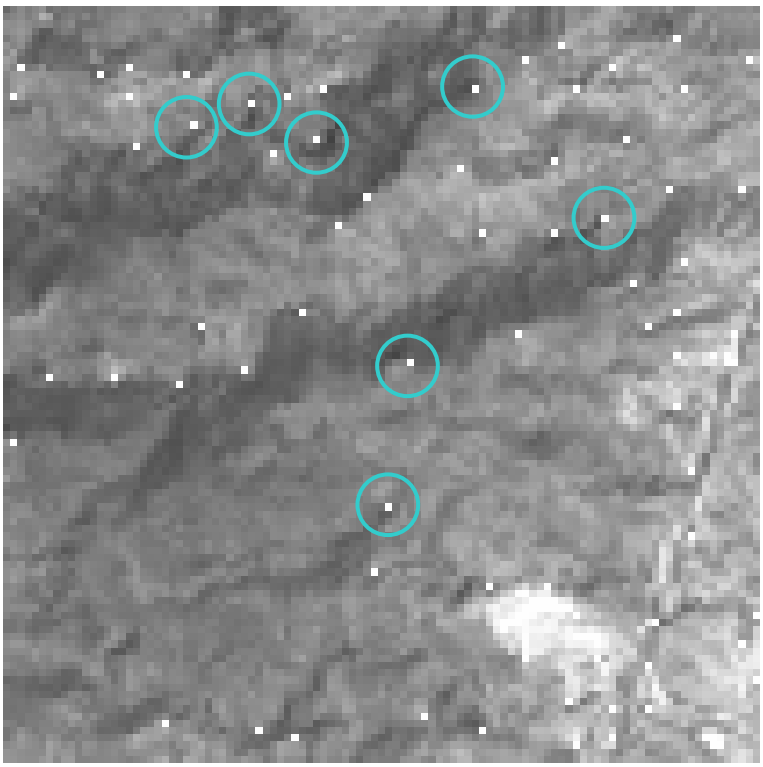


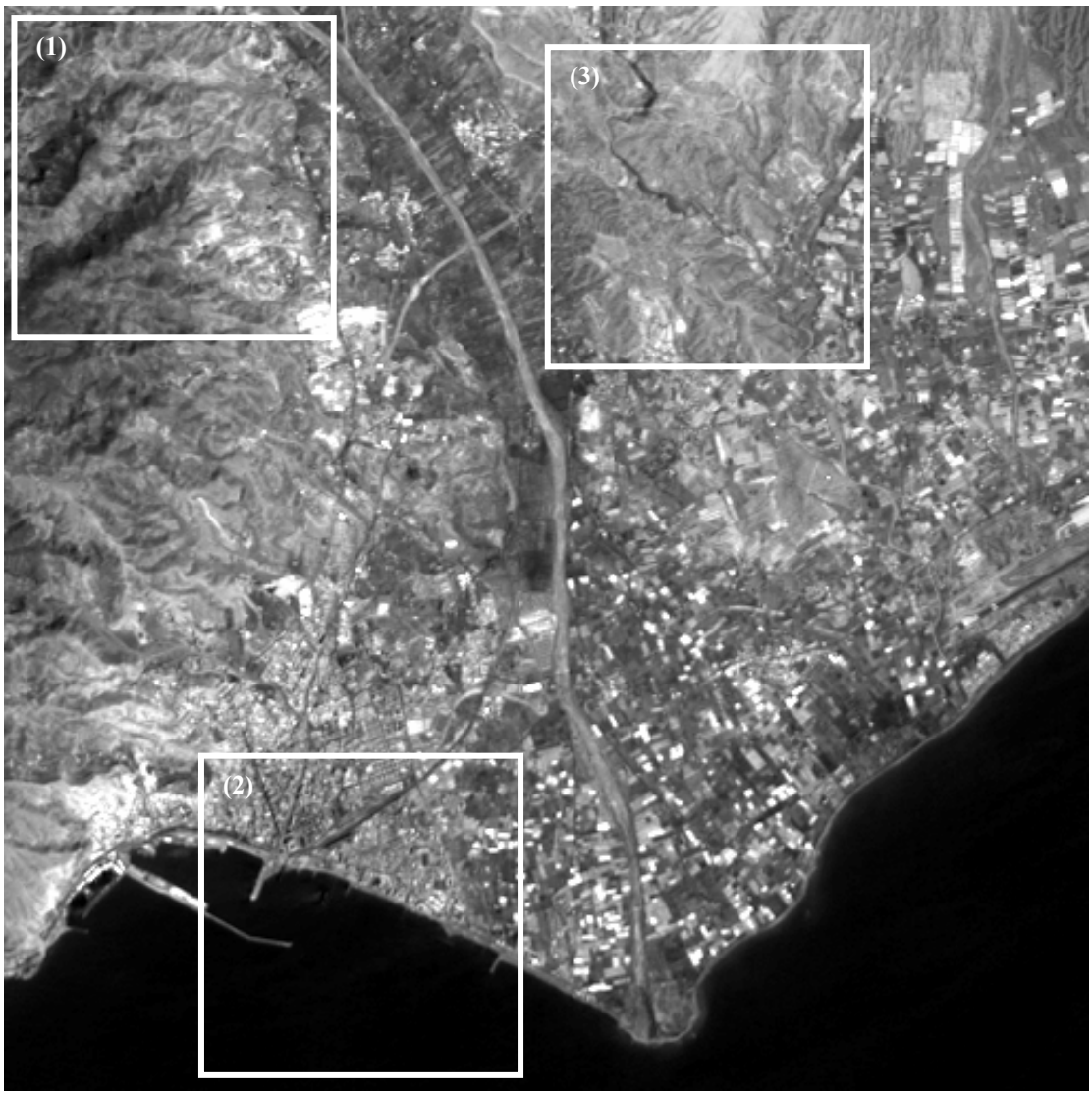
Figure 3. 20: Further enlargement comparison of interest points detected. The blue circles denote the typical mountainous interest points selected by both the ACD and the Harris Corner Detector with window dynamic threshold while the Harris Corner Detector with global threshold failed.

(5) Combined area of mountainous, urban and sea (SPOT sensor)

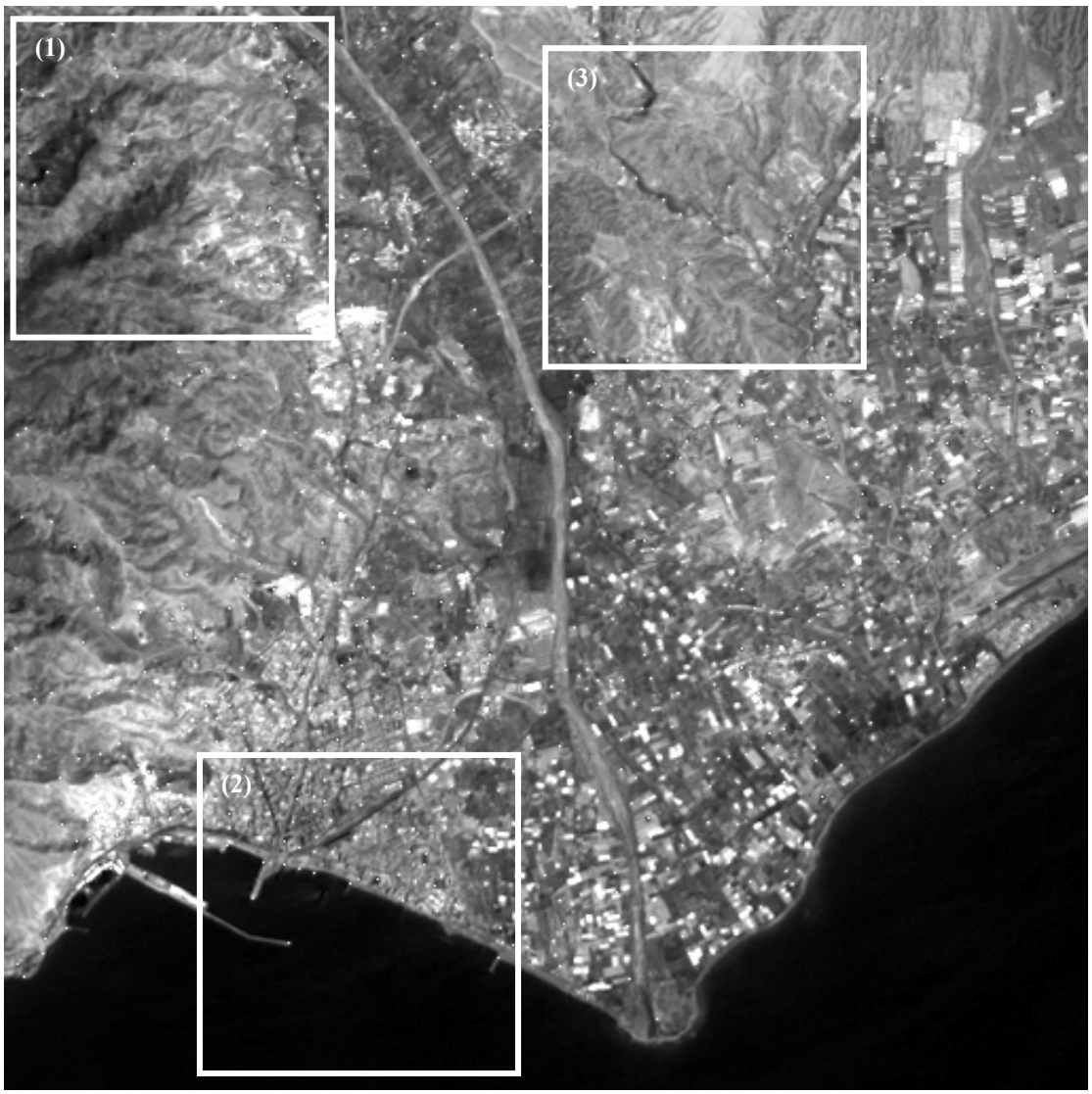
We selected a 512×512 sub scene from a SPOT multispectral band image of south Spain to further evaluate the performance of the ACD on combined area of mountainous, urban and sea. The spatial resolution of the image is 20 metres. This area contains large amount of characteristic interest points representing typical features of different land covers such as road junctions and building corners in the urban area, manmade infrastructure along the coast, and crossings of the mountain ridges, etc.

As listed in Table 3.5, 772 and 7267 interest points were selected using the Harris Corner Detector with global and window dynamic thresholds, respectively. In the mountainous area, the Harris Corner Detector with global threshold again performed insufficiently (Figure 3.22(1)). Even though the mountainous area contains many distinctive features such as the peak points along the ridges or crossings of the ridges, it still has lower contrast than the urban area leading to the failure of the interest point detection. In this case, the Harris Corner Detector with window dynamic threshold performed much better in both the mountainous and urban areas. Plenty of interest points in the mountainous area are detected though not all of them are of good quality. Meanwhile, it extracted poor quality points in the flat sea area because the dynamic threshold is a relative measurement compared to the other pixels in the fixed size window as displayed in the zoom-in comparison in Figure 3.22(2). In addition, zoom in displays in Figure 3.22(2)-(3) have demonstrated some very good quality interest points which are detected only by the ACD while both the Harris Corner Detector with global and window dynamic thresholds failed (highlighted by blue circles).

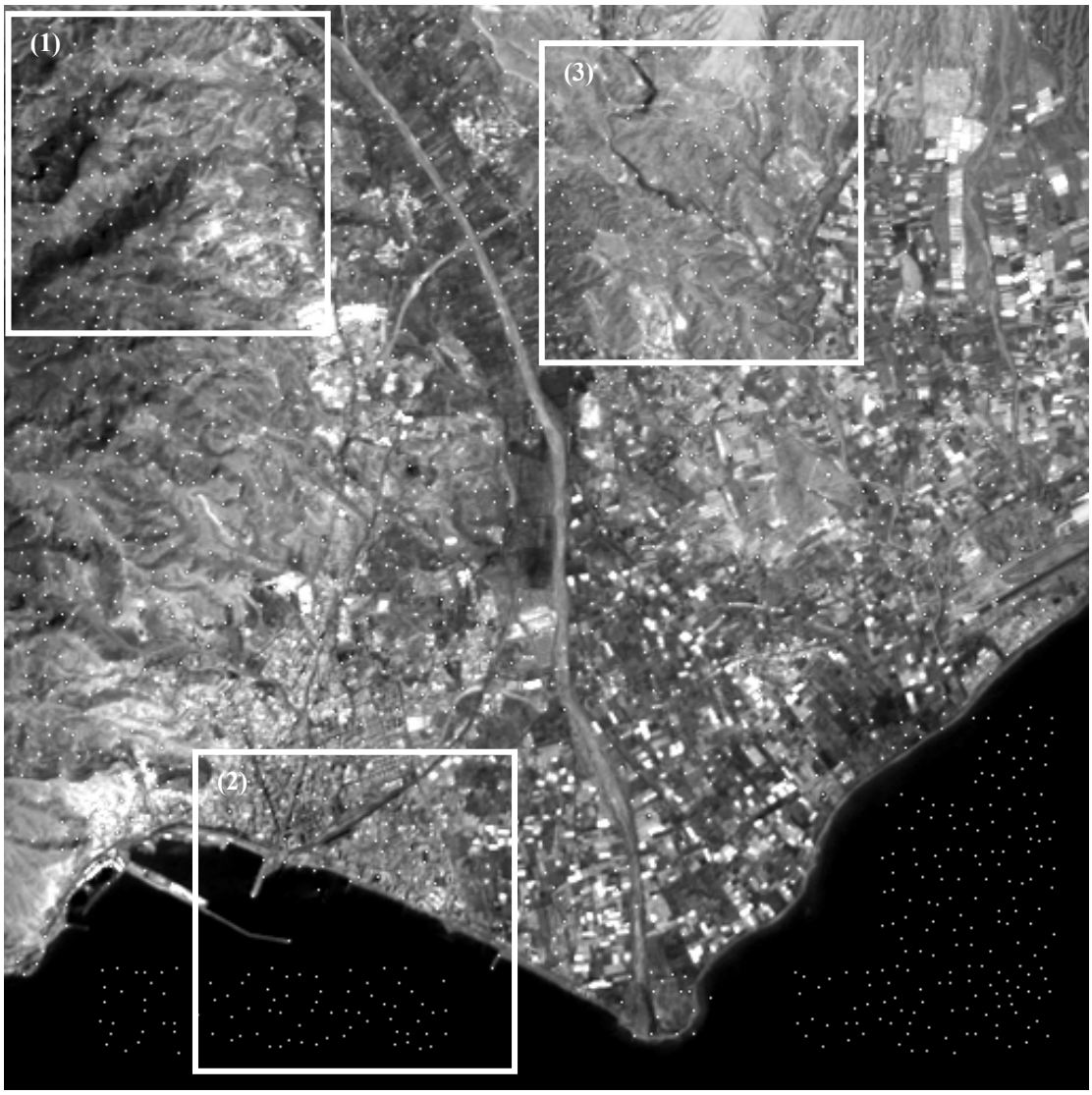
In contrast, the ACD performed much better than the Harris Corner Detector with 1720 interest points selected in a very short computational time of 1.57 seconds. The points are well distributed across the whole image, no matter in what landscape and land covers, and with no poor quality interest points extracted in the sea area (Figure 3.21d).



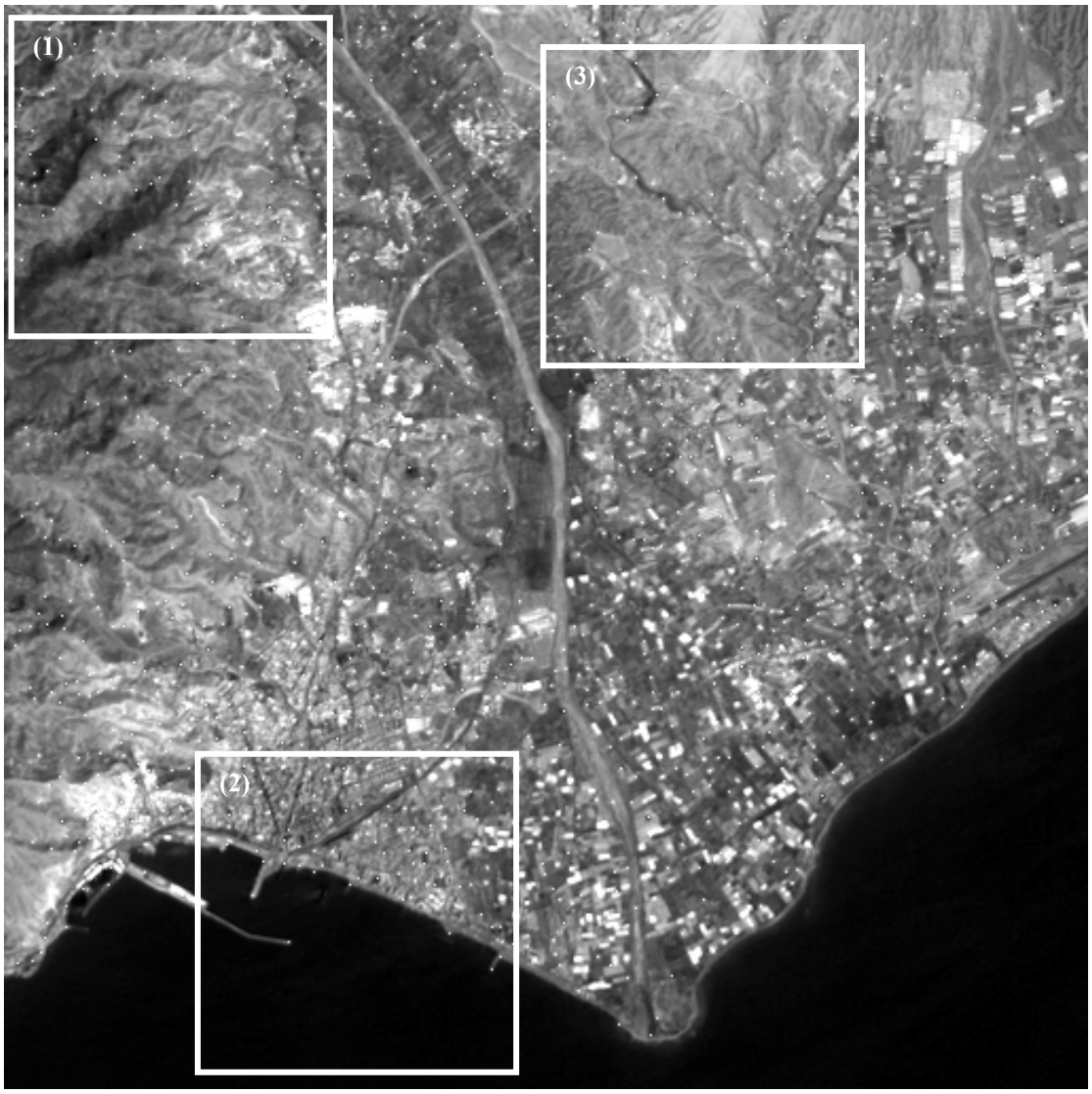
(a)



(b)

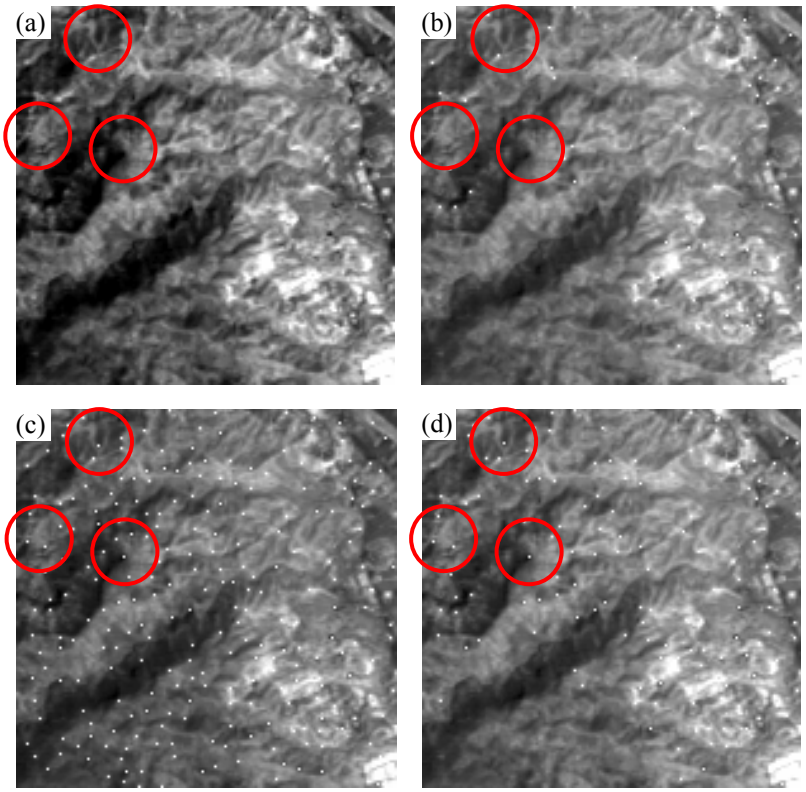


(c)

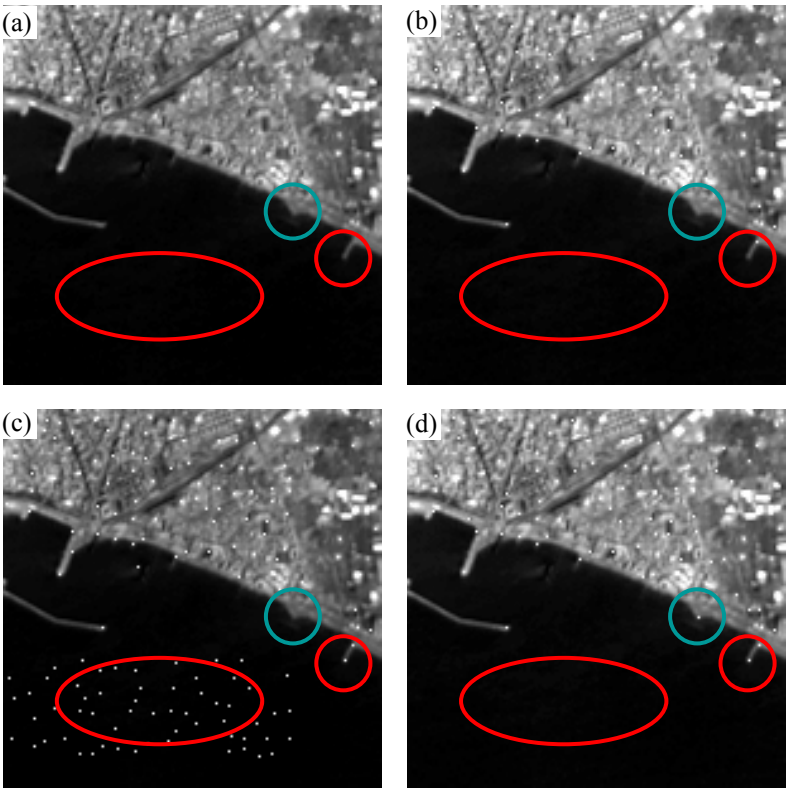


(d)

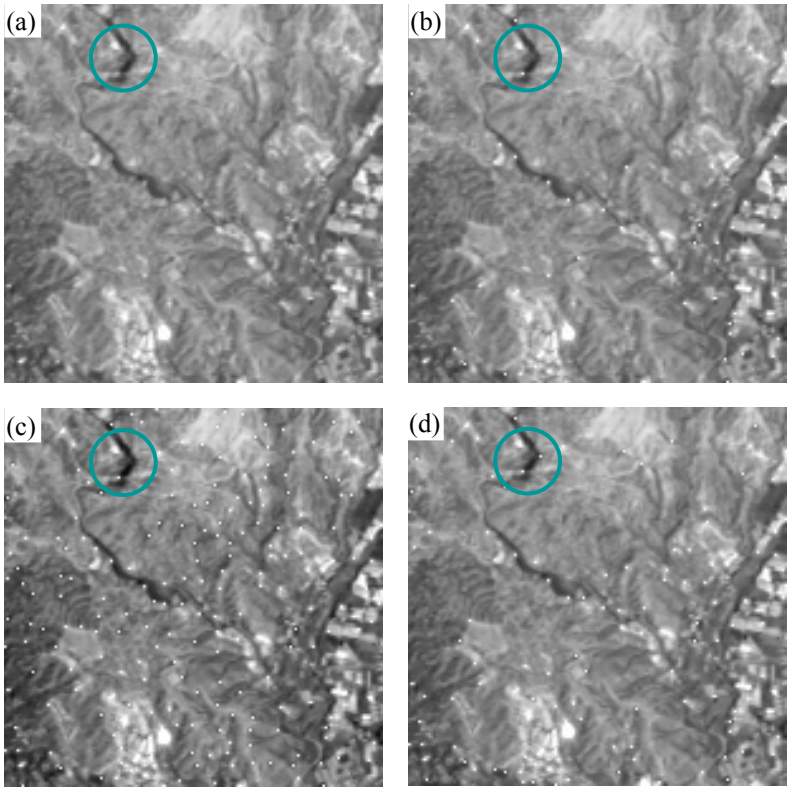
Figure 3. 21: Interest points displayed as the white dots. (a) Initial SPOT multi-spectral band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. Zoom-in sub areas are displayed in white squares.



(1)



(2)



(3)

Figure 3. 22: White dots denote the interest points selected using the Harris Corner Detector and the ACD. The red circles highlight the areas for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD. The blue circles denote several high quality interest points selected only by the ACD while both the Harris Corner Detectors with the global and window dynamic thresholds failed.

(6) Flat urban and field area along coast (SPOT sensor)

A 1500×1500 SPOT panchromatic band image composed of both sea and urban areas was used to evaluate the efficiency and effectiveness of the ACD. Its spatial resolution is 10 metres.

As shown in Table 3.6, the Harris Corner Detector extracted 7708 and 18176 interest points using global and window dynamic thresholds, respectively. As displayed in Figure 3.23b, the Harris Corner Detector with global threshold failed to extract evenly distributed points. The manmade features normally have higher contrast to their neighbourhood compared to the area along the coast. The Harris Corner Detector with window dynamic threshold selected large amount of poor quality interest points in the flat sea area (Figure 3.23c). As indicated before, the window dynamic threshold is essentially a relative measurement compared to the other pixels in the window. Therefore, if the threshold is calculated in a water area containing no distinctive features, it will be too low to eliminate poor quality interest points.

The ACD extracted 15352 high quality interest points which are quite evenly distributed across the whole image (Figure 3.23d). The advantage of the ACD is particularly demonstrated by the fact that no points in the sea are selected with a simple fixed threshold of 7200. Zoom in observation indicates that the best corner points are selected by both methods but the ACD is more critical and effective to select high quality corner points than the Harris Corner Detector, as nearly all the corner points in this image are asymmetrical.

Figure 3.24 (2)-(3) are the close observation of complex road junctions which are almost arrow-junctions and highly asymmetrical. The Harris Corner Detector with global threshold failed to detect them though road intersections are often the first choice for manually selected high quality GCPs. In addition, the Harris Corner Detector again can not select interest points representing distinctive features in relief areas as illustrated in Figure 3.24 (4).

In the zoom-in images displayed in Figure 3.24(1), individual interest points selected using these three methods are compared. The ACD selected best quality interest points

as clearly demonstrated in a further enlargement of the area shown in Figure 3.25. The red circles locate the very good quality interest points which are detected only by the ACD while both the Harris Corner Detector with global and window dynamic thresholds failed. The blue circles denote the points selected by both the ACD and the Harris Corner Detector with window dynamic threshold. The white circles highlight the points selected by all the three methods. Obviously, several high quality interest points are only selected by the ACD in Figure 3.25 which is a very small area, we can consequently assume that the ACD offers much more effective selection for interest points than the Harris Corner Detector in the whole image.

For the whole experiment we tried to tune the parameters for the best performance of the Harris Corner Detector to ensure fair comparisons. However, the ACD always performed obviously better.

Table 3. 6: Numbers of points and computation time of interest point detection on a SPOT Panchromatic band image (flat urban area and field area along coast) obtained using the Harris Corner Detector and the ACD.

Image			SPOT Panchromatic
Size(pixel)			1500 × 1500
Resolution(m)			10
Harris Corner Detector	*Global Threshold*	*Number of Points*	7708
		Computation Time(s)	7.83
	Window Dynamic Threshold	*Number of Points*	18176
		Computation Time(s)	18.54
ACD		*Number of Points*	15352
		Computation Time(s)	7.41



(a)



(b)

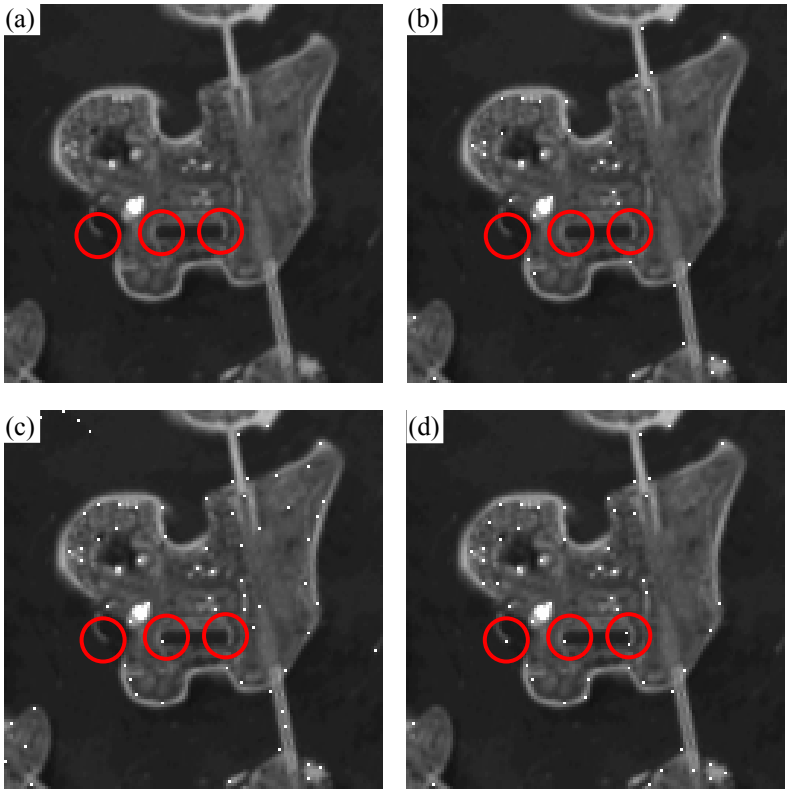


(c)

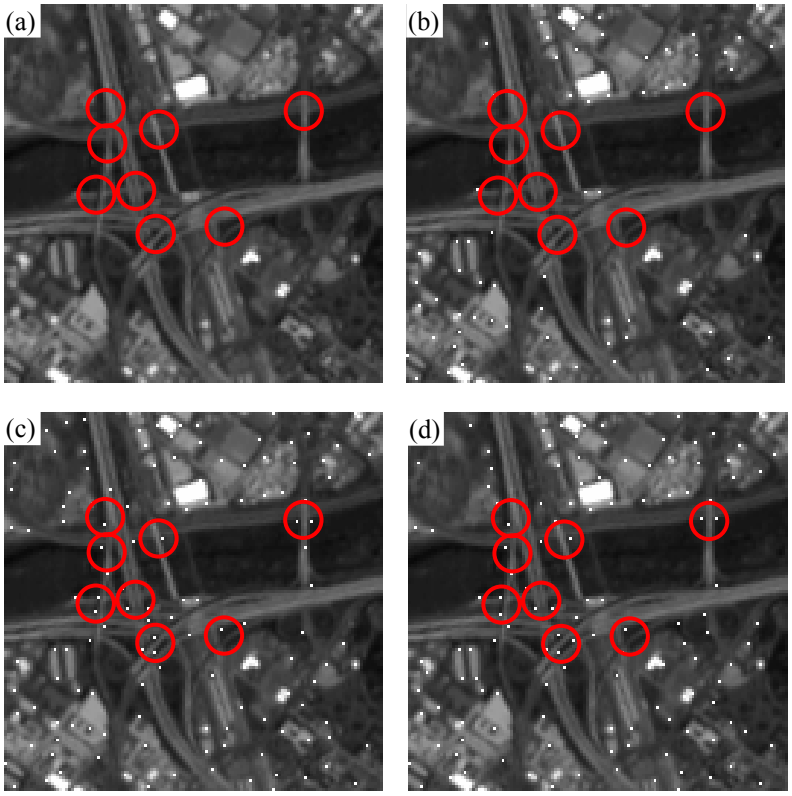


(d)

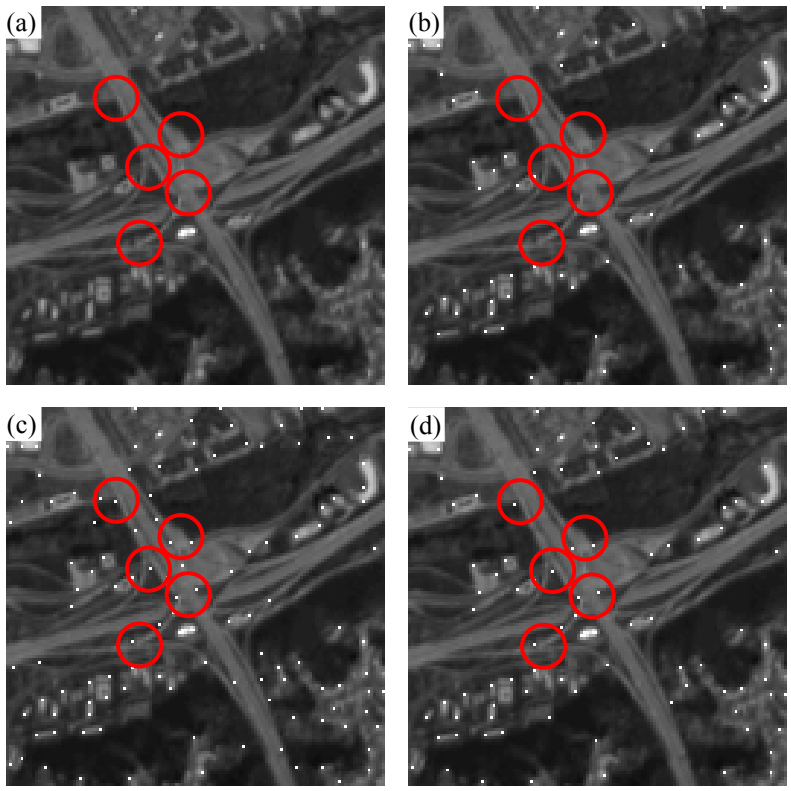
Figure 3. 23: Interest points displayed as the white dots. (a) Initial SPOT panchromatic band image. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points extracted using the Harris Corner Detector with window dynamic threshold. (d) Interest points detected using the ACD. Zoom-in sub areas are displayed in white squares.



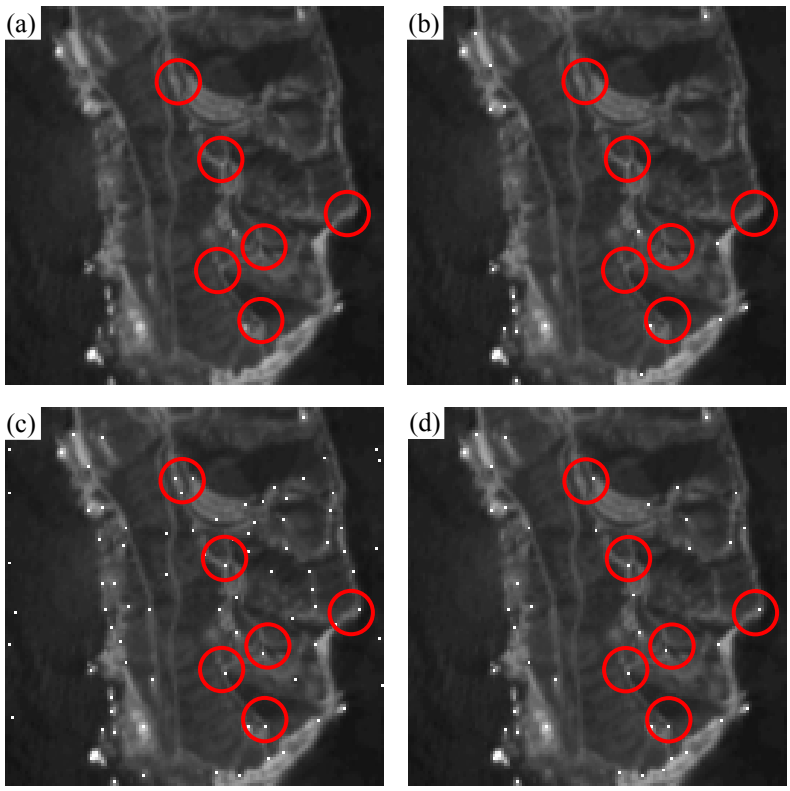
(1)



(2)



(3)



(4)

Figure 3. 24: White dots denote the interest points selected using the Harris Corner Detector and the ACD. The red circles highlight the areas for comparison. (a) Initial zoom-in images. (b) Interest points selected using the Harris Corner Detector with global threshold. (c) Interest points detected using the Harris Corner Detector with window dynamic threshold. (d) Interest points extracted using the ACD.

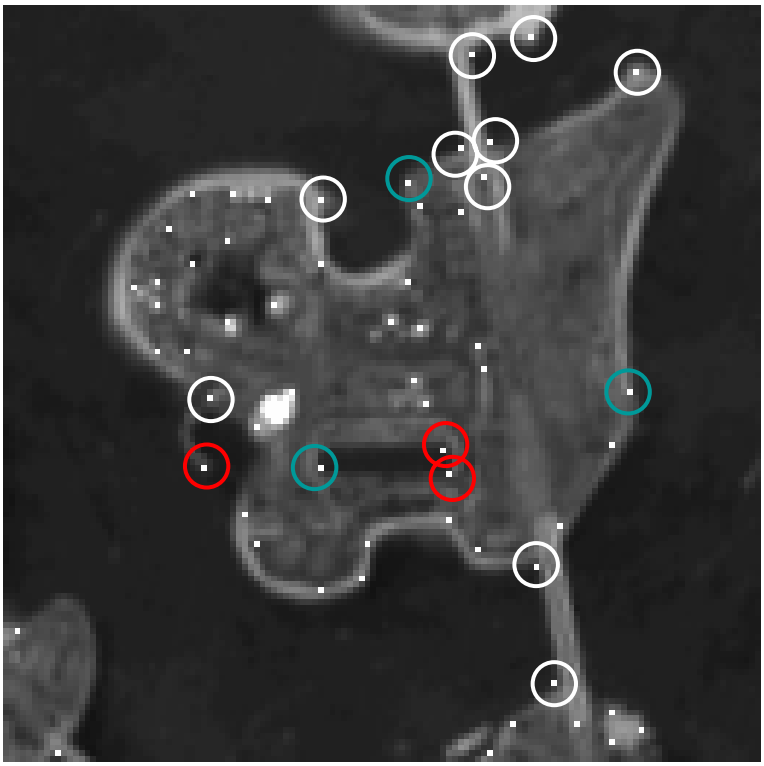


Figure 3. 25: Further zoom-in observation of interest points selected using the ACD. The red circles locate the very good quality interest points which are detected only by the ACD while both the Harris Corner Detector with global and window dynamic thresholds failed. The blue circles denote the points selected by both the ACD and the Harris Corner Detector with window dynamic threshold. The white circles highlight the points selected by all the three methods.

Chapter 4

A semi-automatic image co-registration scheme

In this chapter, a semi-automatic image co-registration scheme is developed. Comprehensive explanation of each step is subsequently presented with all-round image co-registration experimentation of multi-sensor, multi-temporal and multi-resolution images taken by by Landsat Thematic Mapper (TM), Landsat-7 Enhanced Thematic Mapper Plus (ETM+), and Satellite Pour l'Observation de la Terre (SPOT) sensors.

4.1 The scheme

In commonly used image processing packages, image co-registration is usually done manually. In these methods, Ground Control Points (GCPs) are located visually on both the input and reference images by choosing large amount of characteristic and easily recognizable points such as road junctions, river attribute joints, and building corners, etc to establish the correspondence between the input and reference images. Based on the GCPs selected, a precise transformation model is subsequently generated and together with a proper order polynomial fitting including linear, quadratic or cubic, accurate image co-registration is carried out. The manually methods are based on the identification of control points and the selection of polynomial fittings.

In this research, based on the ACD algorithm, a semi-automatic image co-registration scheme is developed with the sequential six steps (Figure 4.1).

- ***Initial transformation model*:** three GCP pairs are manually selected to determine the overlapped area and establish an initial linear (affine) transformation model characterising shift and rotation between the input and reference images. This type of transformation can model six kinds of geometric distortion in the remotely sensed imagery data including the translation and scale changes in two directions, skew and rotation. The key point of this step is that the three initial GCP pairs should be well spread in the whole image scene and of good quality.

- ***Interest point extraction*:** the ACD of a fixed threshold (7200) is employed to locate interest points in the overlapped area of the reference image as candidates of GCPs. In this step, a pre-defined threshold and non-maximum suppression are applied. The non-maximum suppression extracts the central pixel if its interest strength is the largest among its neighbourhood in the window of a given size. The purpose of this method is to avoid point clustering in a small area and assure the sparseness of interest points selected in the reference image.

- ***Grid-control*:** apply a grid to divide the image into sub areas and in each area extract user defined number of points from all the selected interest points. This simple method assures that the final selection of interest points are evenly distributed in the overlapped area in the reference image and the total number of interest points selected is controlled. It also guarantees the good quality of the selected GCPs as they are the points with the highest ACD values within the grid. Different size grids are automatically selected based on the complexity of the image co-registration. For example, the overlapped area is divided into 144 grids for a same sensor, multi-temporal image co-registration. For a multi-temporal, multi-sensor, multi-resolution image co-registration, the number of grids is increased to 256.

- ***Candidate GCP pairs:*** The roughly corresponding GCPs in the input image are calculated based on the initial transformation model.

- ***GCP refinement using FNCC*****:** The Fast Normalized Cross-Correlation (FNCC) (Lewis, 1995) algorithm is employed to refine a GCP position in the input image based on the maximum local feature correlation calculated in a correlation window centred at the GCP position in the reference image roaming in a search window around the initial position of the corresponding GCP in the input image. In addition, when we apply the FNCC algorithm to register images taken by different satellites and sensors, due to the geometric discrepancy between different sensors, large search window and calculation window are required to ensure the localization accuracy of the GCP pairs in the reference and input images.

- ***Image rectification using ERMapper 7.0 Geocoding Wizard*****:** Based on the GCP pairs obtained from the above several steps, a transformation model is generated. Considering the geometric discrepancy between the input and reference images, three different order polynomial fittings (linear, quadratic and cubic) are manually selected to generate high accuracy image co-registration. Subsequently, image re-sampling using either the bilinear or cubic interpolation method is carried out to produce the co-registered image.

This semi-automatic image co-registration scheme only requires a user to manually select three initial GCP pairs and all the following process is automatic. The FNCC is a relatively slow process. Consequently, in order to make this semi-automatic image co-registration scheme operate at a high speed and achieve required accuracy, it is essential to select a limited but adequate number of high quality GCPs to minimise the loading of the FNCC. The ACD serves this purpose the best as shown in the experiment examples presented in the next section.

Actually, there could be a different approach that divides the image into sub areas first and then sets a dynamic contrast dependent window threshold to select interest points.

And in each sub area, the largest one or two interest points can be selected as GCPs. However, this approach could force to select poor quality interest points in the areas containing no distinctive features such as the water area and it is therefore disregarded.

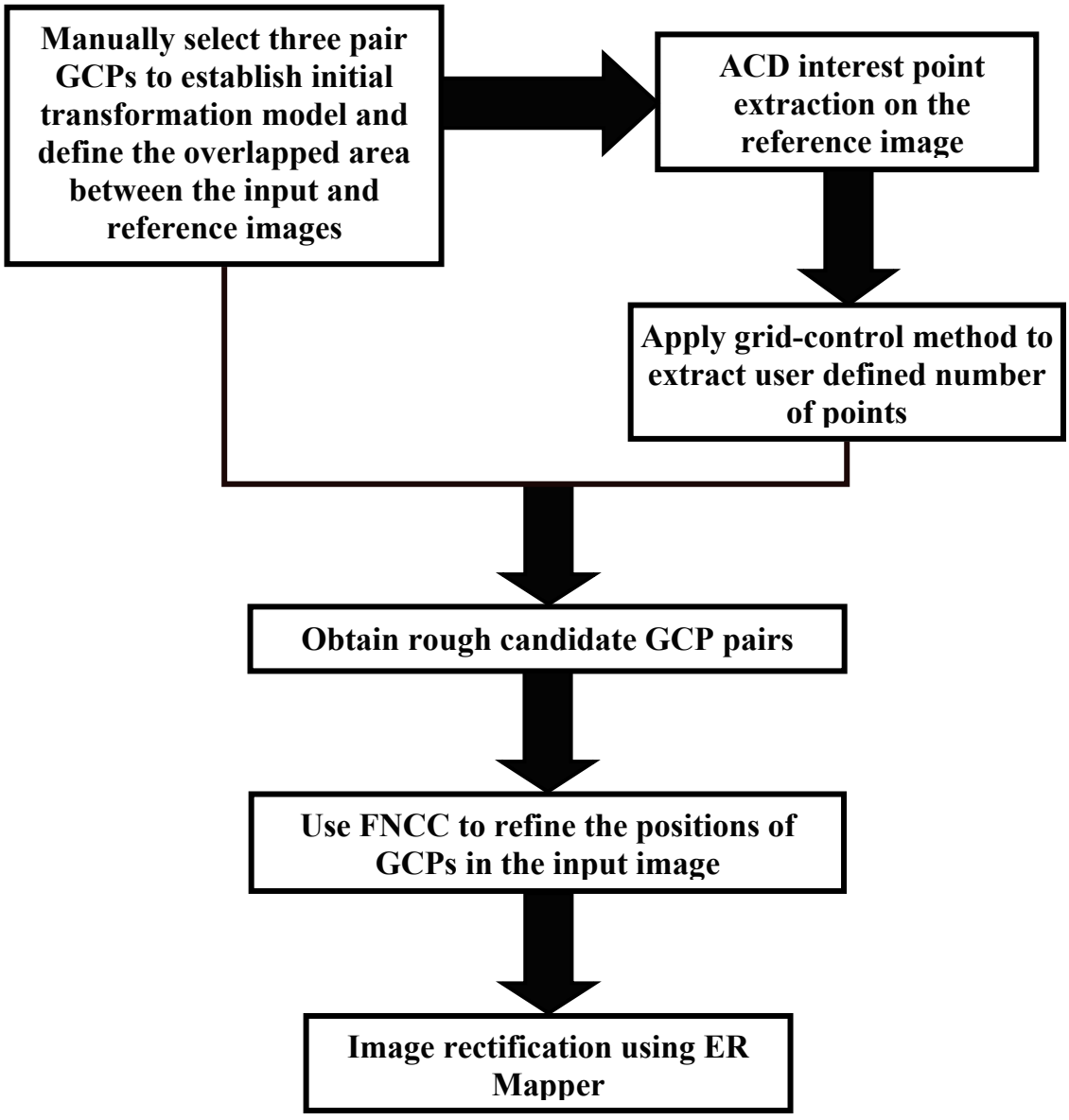


Figure 4. 1: Flow chart of the semi-automatic image co-registration scheme.

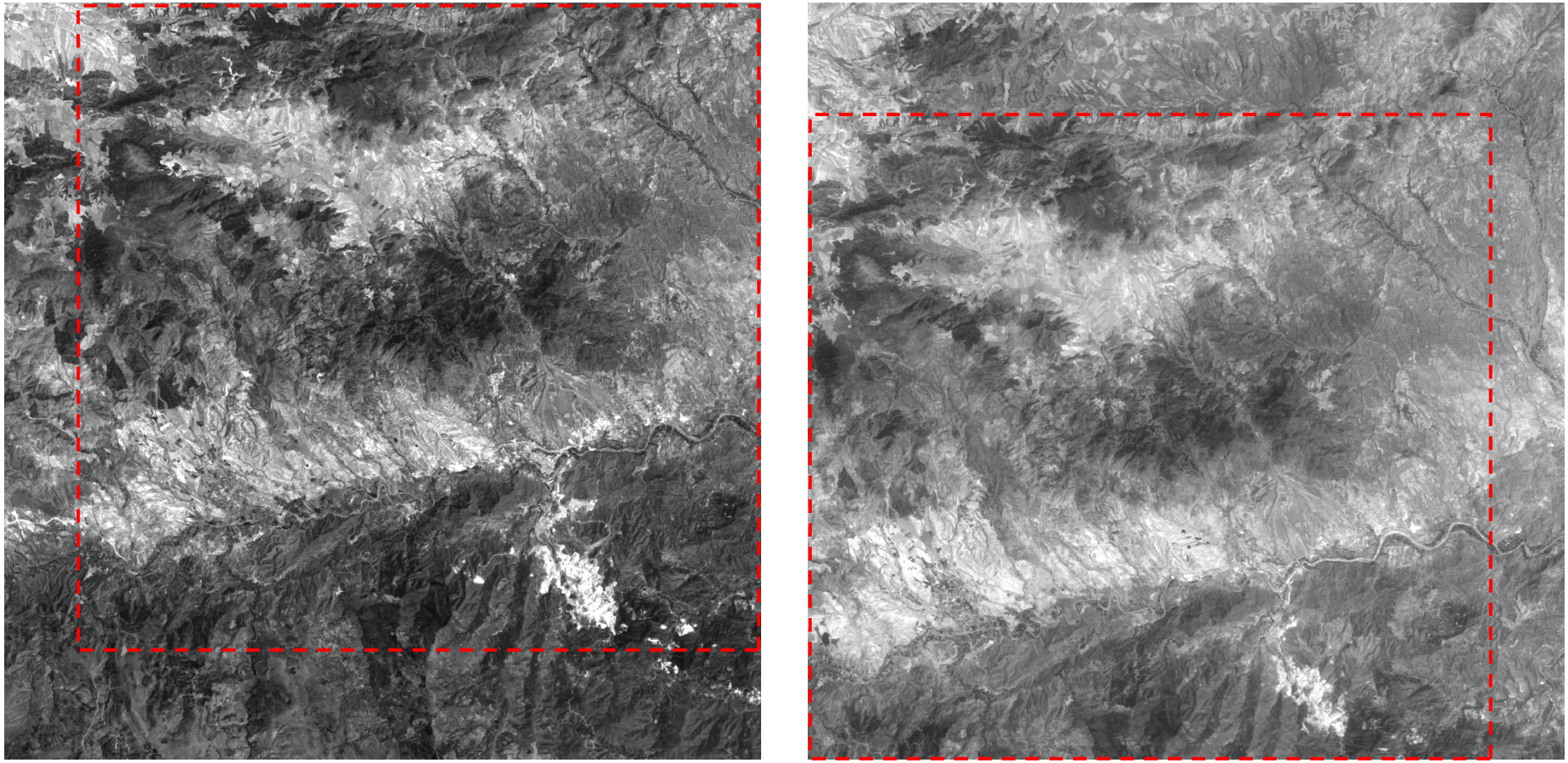
4.2 Image co-registration experimental results

The proposed semi-automatic registration approach was tested for the co-registration of multi-temporal, multi-sensor and multi-resolution satellite images taken by Landsat Thematic Mapper (TM), Landsat-7 Enhanced Thematic Mapper Plus (ETM+), and Satellite Pour l'Observation de la Terre (SPOT) sensors.

(1) Co-registration of the same sensor multi-temporal images

Two 1024×1024 sub-scenes extracted from Landsat TM visible band images in southeast Spain were used to test the effectiveness of the proposed image co-registration scheme for multi-temporal image co-registration. As displayed in Figure 4.2, the reference image and input image were obtained in 1992 and 1984, respectively. Three pairs of GCPs were manually selected to determine the overlapped area and to establish the initial linear transfermation model. A default parameters set was used; which includes the fixed global threshold of 7200, the standard deviation $\sigma = 0.5$ for the Gaussian Filter and a 7×7 non-maximum suppression mask. The ACD automatically selected 4469 interest points denoted by the white dots in the overlapped area of the reference image (Figure 4.3).

The overlapped area is then divided into 144 grids and the pixels with the highest ACD values in each of the grid areas were finally selected as GCPs. The rough positions of the corresponding GCPs in the input image were computed from the initial transformation model to form the candidate GCP pairs. The FNCC algorithm was then applied with a 25×25 calculation window roaming in a 33×33 search window to refine the GCP locations in the input image. To guarantee the high accuracy of GCPs, those GCP pairs with NCC value lower than a pre-defined threshold (e.g. 0.7) were eliminated. In this case, 113 precisely matched pairs of GCPs were selected from the initial 144 candidates. These finally selected GCPs achieved 0.54 pixels average RMSE in a linear polynomial transformation model. These 113 GCP pairs are evenly distributed in the overlapped area as shown in Figure 4.4. The final registration result is illustrated in a colour composite displaying the 1992 image in green and the co-registered 1984 image in red as shown in Figure 4.5. The co-registration quality is self-evident in this case; there is hardly any mismatching even by zoom in close look.



(a) **(b)**

Figure 4. 2: (a) the input image taken in 1984 (b) the reference image taken in 1992. The red dashed square boxes highlight the overlapped area in both images.

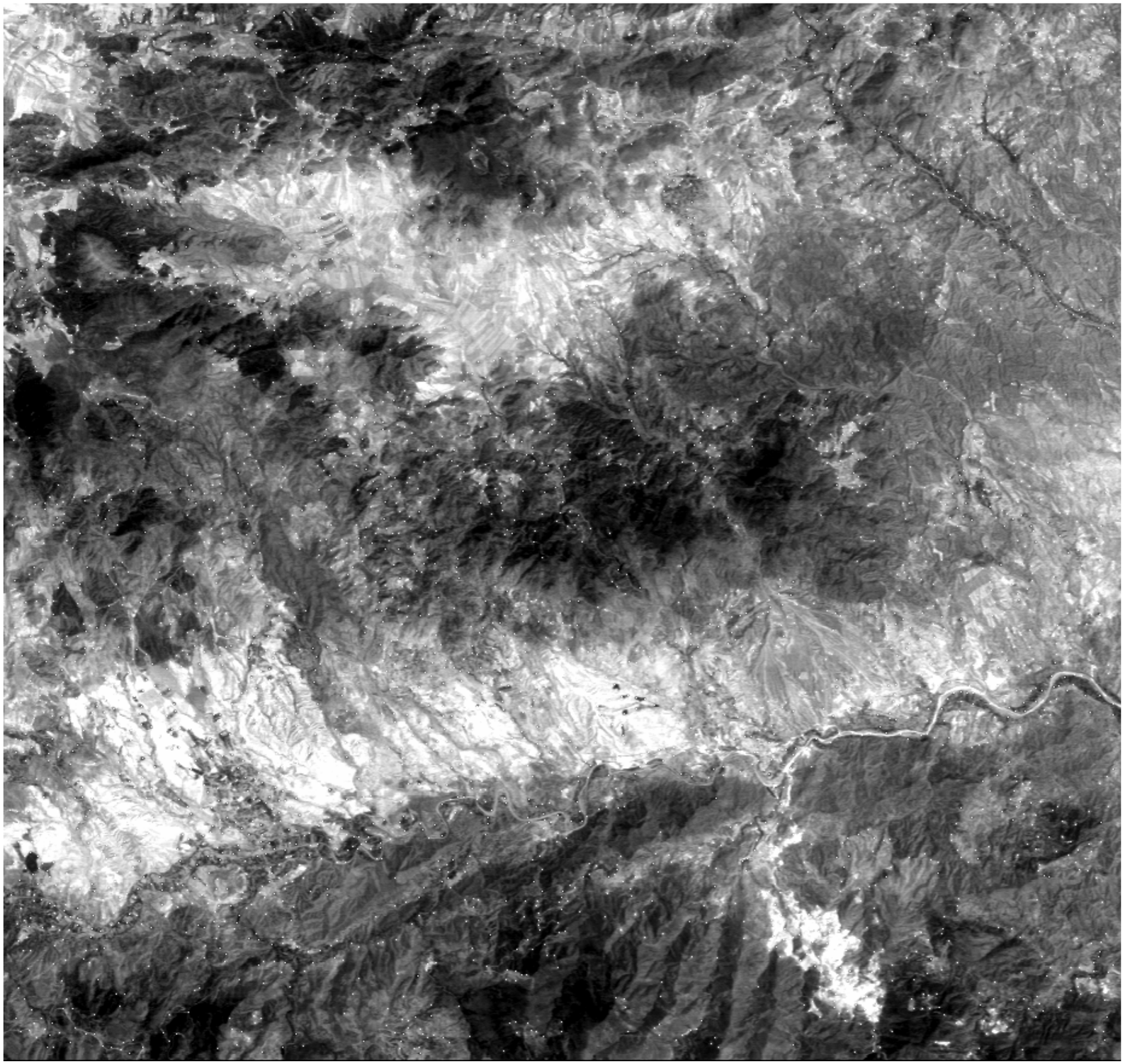


Figure4. 3: The white dots denote the interest points selected using the ACD in the overlapped area of the reference image**.**

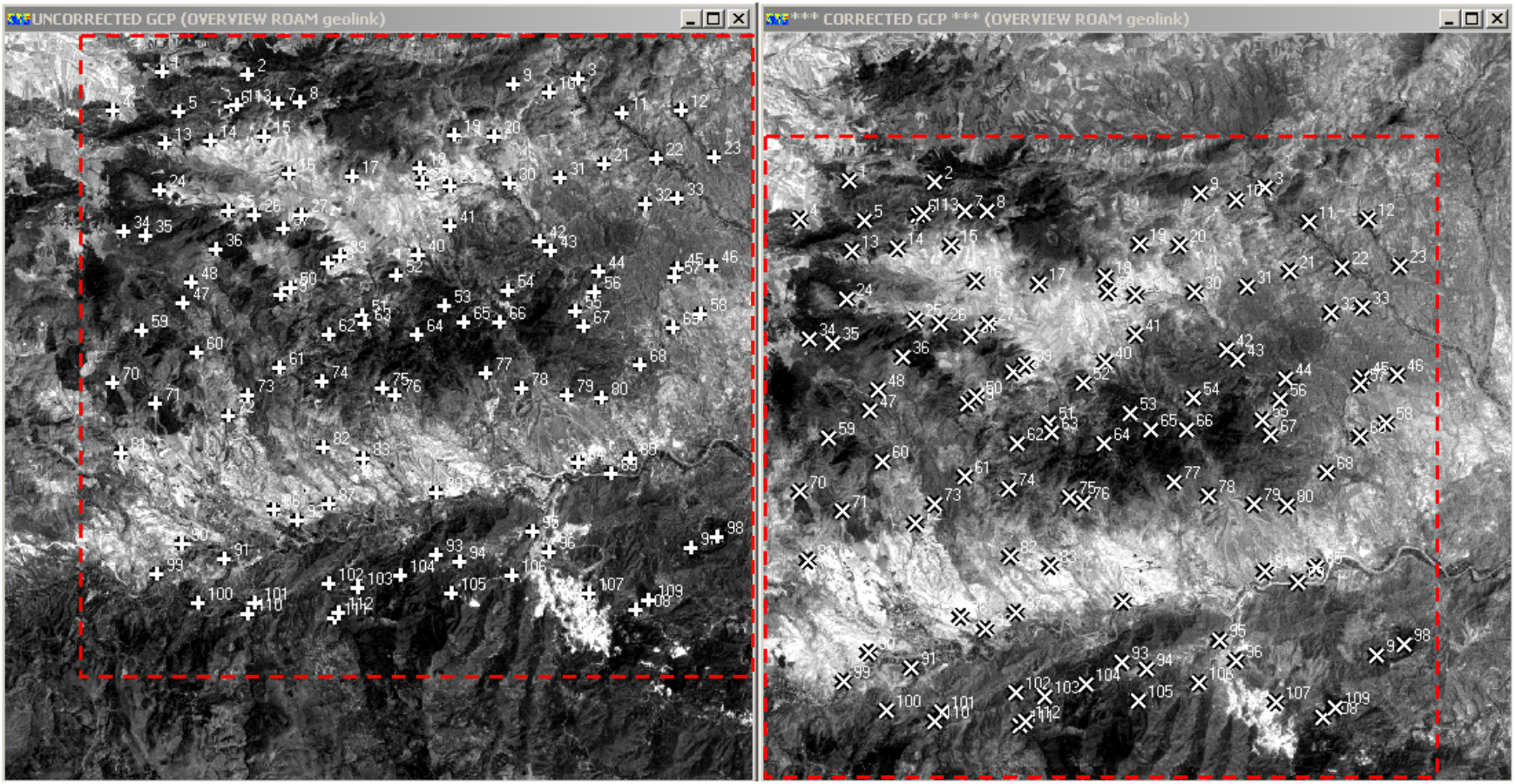
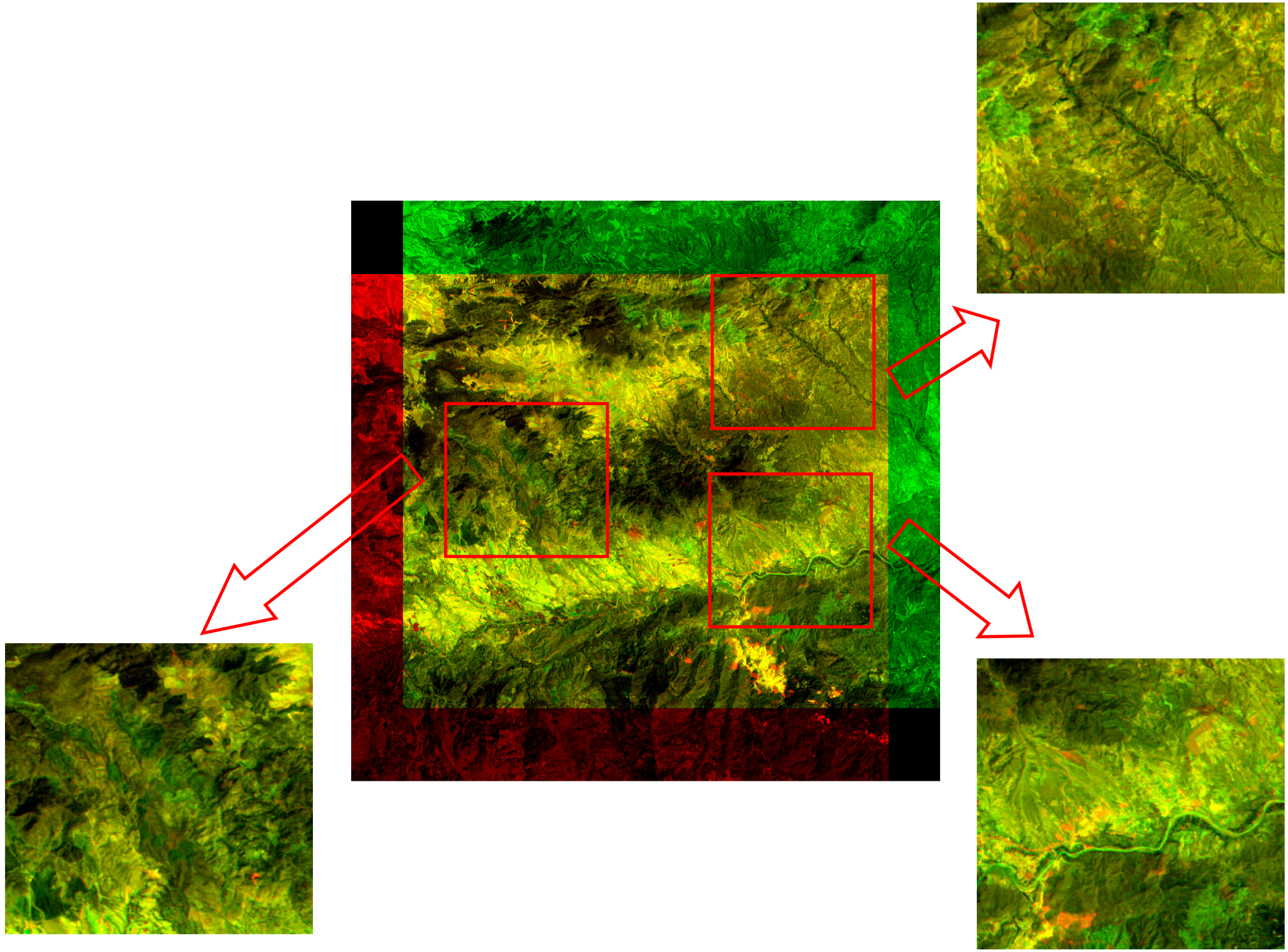


Figure 4. 4: 113 precisely matched pairs of GCPs detected in the image pair. The left one is the input image and the right one is the reference image. Red dashed square boxes denote the overlapped areas.



Figuer 4. 5: The registration result and three zoom-in sub scenes. The reference image and co-registered image are displayed in green and red, respectively

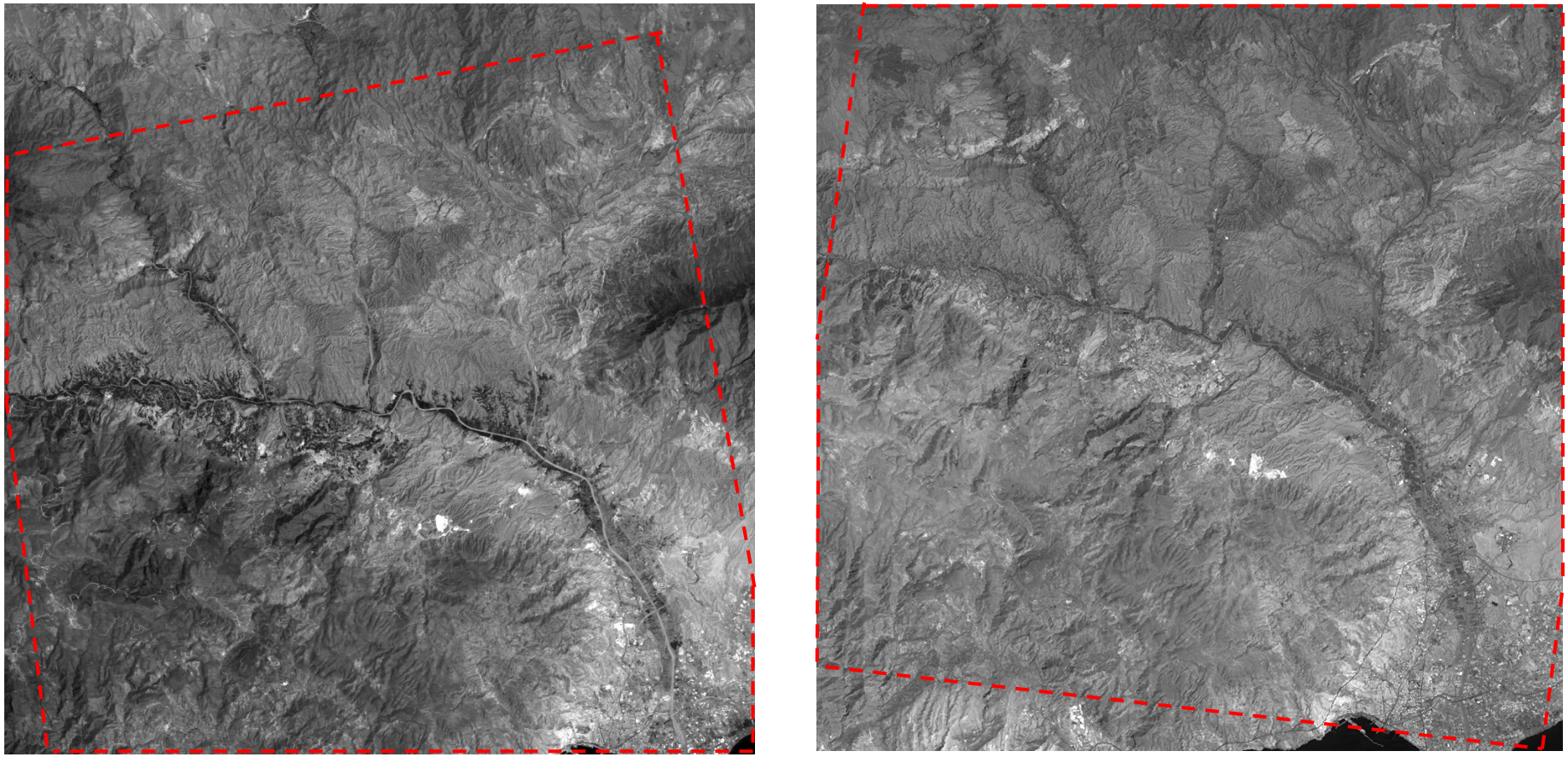
(2) Co-registration of multi-sensor and multi-resolution images

- *Landsat TM and ETM+ sensor images*

Two images of Almeria Province in the southeast Spain taken by Landsat TM and ETM+ sensors were used to evaluate the performance of the semi-automatic image co-registration scheme for multi-sensor images. A sub-scene of 1024×1024 pixels from the 30 metre resolution TM visible band image obtained in 1989 was used as the input image (Figure 4.6a). The middle-infrared band (band 5) image of 1024×1024 pixels taken by Landsat-7 ETM+ sensor in 2000 was used as the reference image (Figure 4.6b) and its spatial resolution is 28.5 metres. The ETM+ band 5 image is recognised as the best spectral band for distinguishing different types of land covers including road, soil, vegetation and water. In addition, there is an approximately 10 degree rotation between the initial reference image and the input image.

Three GCP pairs were selected to determine the overlapped area, as delineated in red dashed polygons in Figure 4.6, and establish the initial affine transformation model. Then, the ACD, with the same parameters setting as the first example, successfully detected 10869 interest points as white dots displayed in Figure 4.7. Due to the slightly different spatial resolution and spectral differences between the two images, larger number of GCPs is needed to ensure the required image co-registration accuracy. Therefore, the overlapped area in the reference image was divided into 196 grids. In the FNCC processing, a 25×25 calculation window roaming in a 35×35 search window was applied to refine the GCP locations in the input image. The pre-defined NCC threshold of 0.7 was used to eliminate poor quality GCP pairs. Finally, 72 highly matched and evenly distributed GCP pairs were obtained with an average RMSE of 0.67 pixels as displayed in Figure 4.8.

The input image is then transformed using the transformation model and interpolated into a co-registered image with a spatial resolution of 28.5 metres. Red-Green colour composite method was used to evaluate the co-registration result with the reference image displaying in green and the co-registered image displaying in red (Figure 4.9). Several zoom-in close looks demonstrate that the co-registration accuracy for this multi-sensor and multi-resolution image pair is of satisfactory with precise matching in both the urban and mountainous areas.



(a) **(b)**

Figure 4. 6: (a) the input image obtained using Landsat TM sensor in 1989 (b) the reference image taken in 2000 using Landsat-7 ETM+ sensor. The red dashed polygons highlight the overlapped area in both images

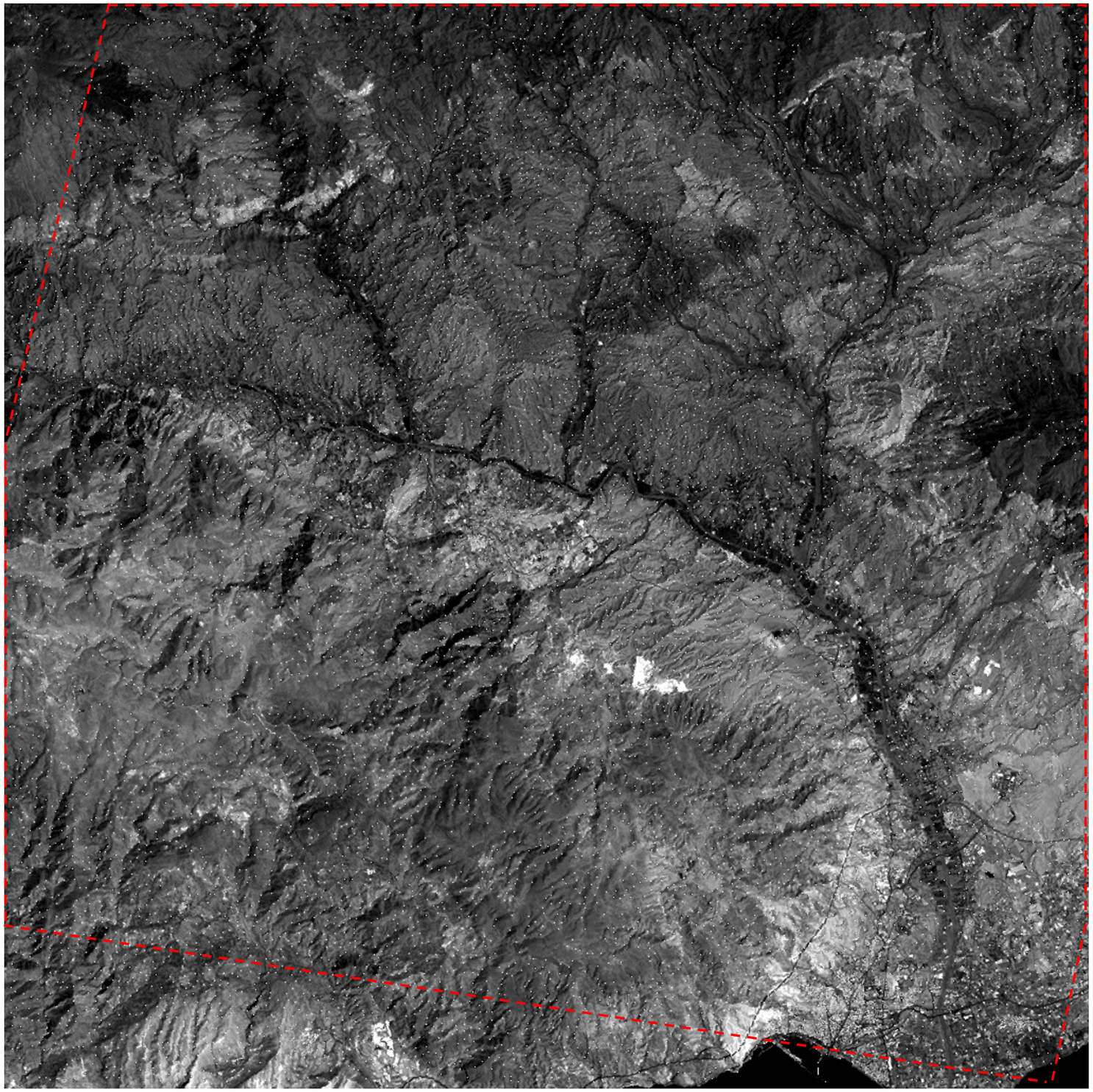


Figure 4. 7: The white dots denote the interest points selected using the ACD in the overlapped area (red dashed polygon) of the reference image

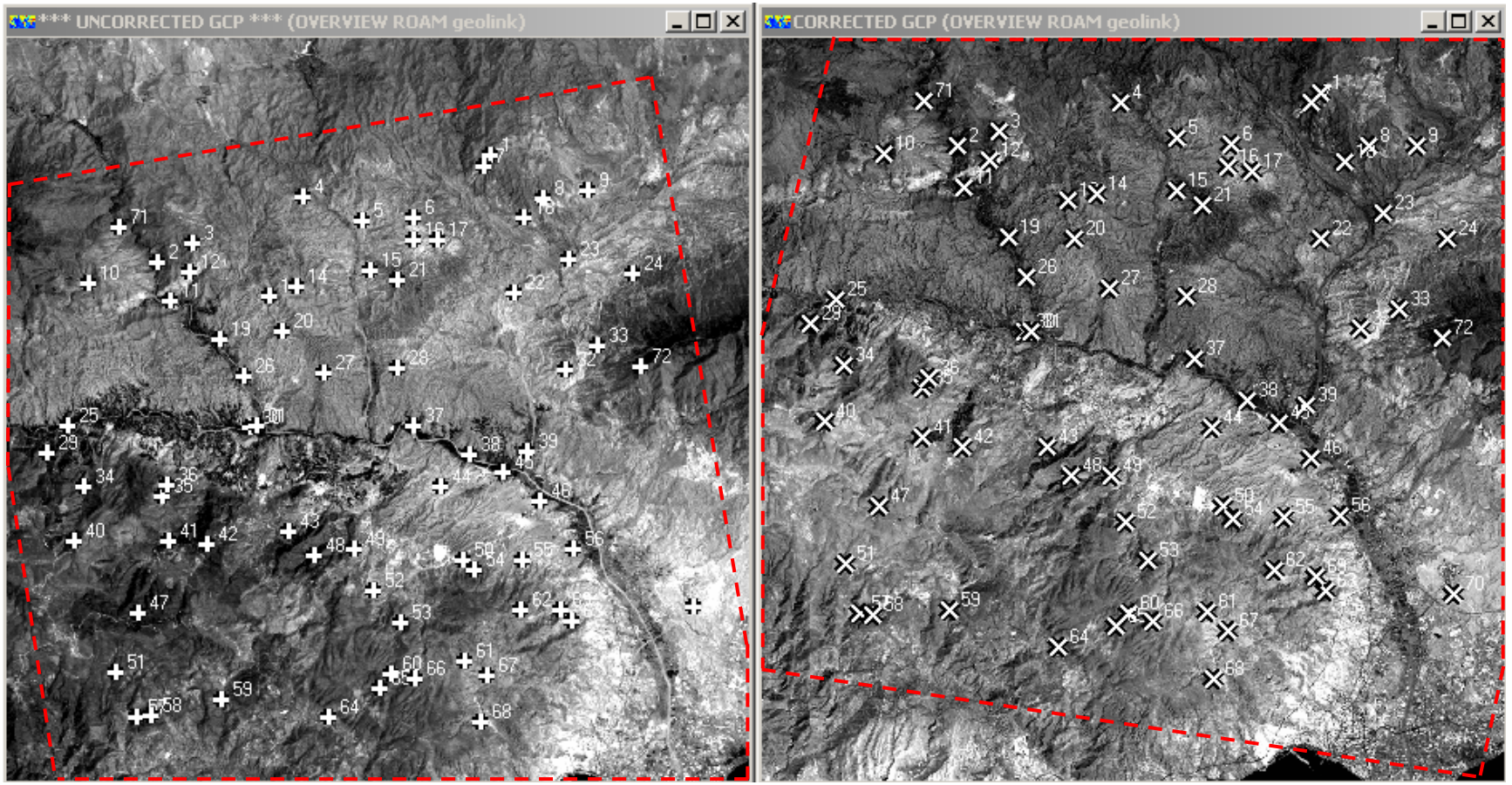


Figure 4. 8: 72 GCP pairs extracted from the image pair. The left one is the input image and the right one is the reference image. Red dashed polygons denote the overlapped areas.

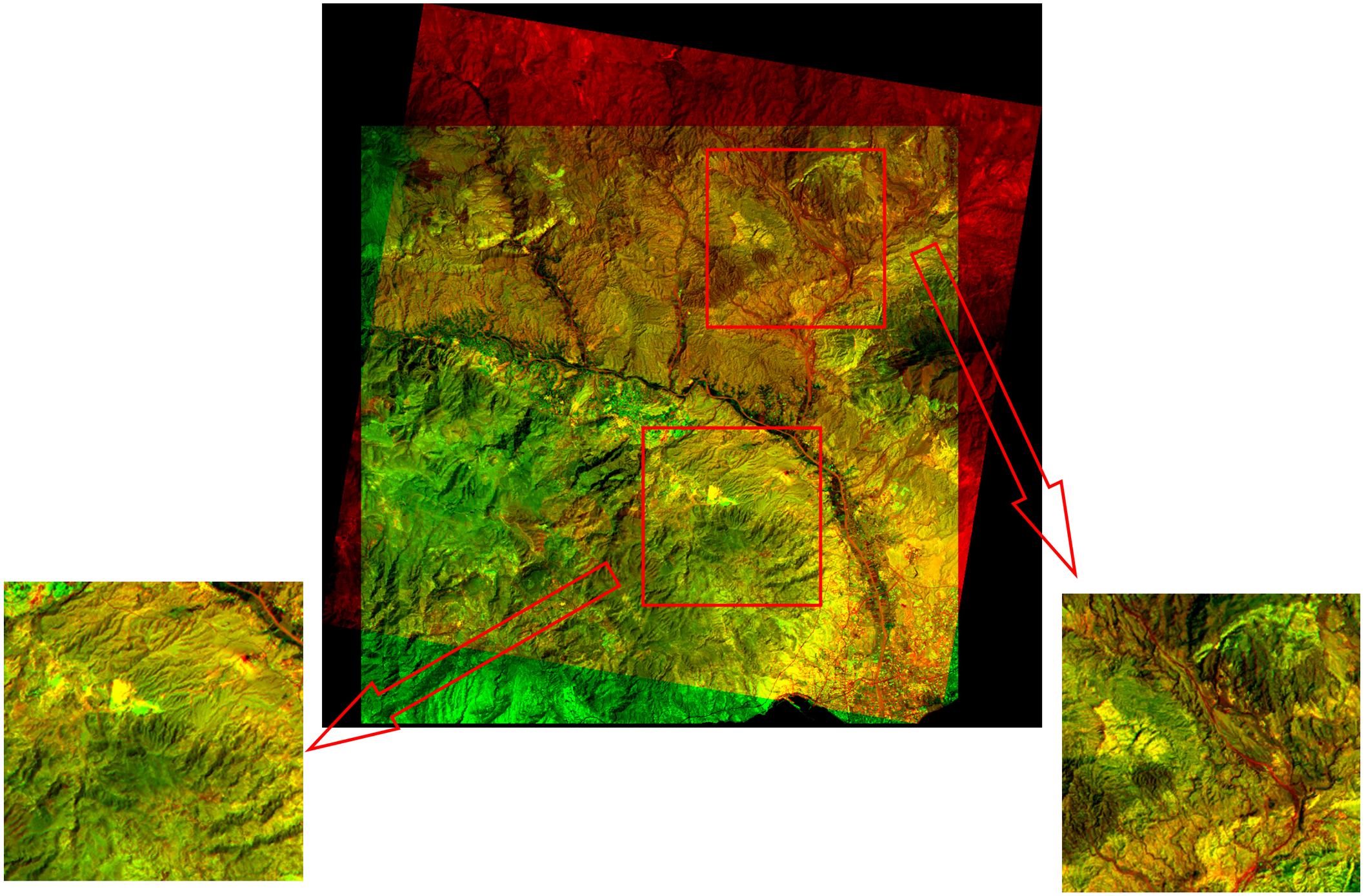


Figure 4. 9: The registration result is presented in a colour composite with the reference image displaying in green and co-registered image displaying in red, respectively. The two red squares highlight the zoom-in sub scenes for a close look.

- *SPOT and Landsat-7 ETM+ sensor images*

A 10 m resolution SPOT panchromatic band image and a 28.5 m resolution Landsat-7 ETM+ image of Almeria Province were used to assess the semi-automatic image registration scheme for multi-sensor/resolution images. A sub-scene of 1200×1200 pixels from the Landsat-7 ETM+ multispectral band (band 2) was used as the input image (Figure 4.10a). A sub-scene of 3000×3000 pixels SPOT panchromatic band image was used as the reference image (Figure 4.10b). There is an approximately 9.5 degree rotation between the reference image and input image.

After initial set up based on the three manually selected GCP pairs, the ACD algorithm, with the same default parameters setting as before, automatically extracted 51666 interest points from the overlapped area of the reference image. Due to the limit of printing resolution, four zoom-in sub areas representing various land covers in both mountainous and urban areas are selected to provide a close look at the interest points detected as white dots displayed in Figure 4.11. Considering the geometric discrepancy between the two images taken by different sensors (Landsat-7 ETM+ is an across-track two way scanner while SPOT is a push-broom scanner) is more significant than the case of the same sensor images, more GCPs are required to carry out higher order polynomial fitting so as to achieve high accuracy image co-registration. Therefore, the overlapped area in the reference image was divided into 256 grids. The window size in the FNCC processing was increased to 30×30 for the calculation roaming in a 41×41 search window in order to compensate the offsets caused by rotation, resolution difference and projection distortion between the two images. After eliminating poor quality GCP pairs with NCC < 0.7, 123 precisely matched pairs of GCPs were finally obtained (Figure 4.12) with an average RMSE of 0.64 pixels in a cubic polynomial transformation model.

The final co-registration result is illustrated as a colour composite in Figure 4.13. The co-registered Landsat-7 ETM+ image with the same spatial resolution (10 m) as the reference image is displayed in green and the reference image is displayed in red. The enlargements of four sub areas (Figure 4.13) evenly distributed in the overlapped area demonstrate very good registration quality with no recognizable mismatching in both mountainous and urban areas.

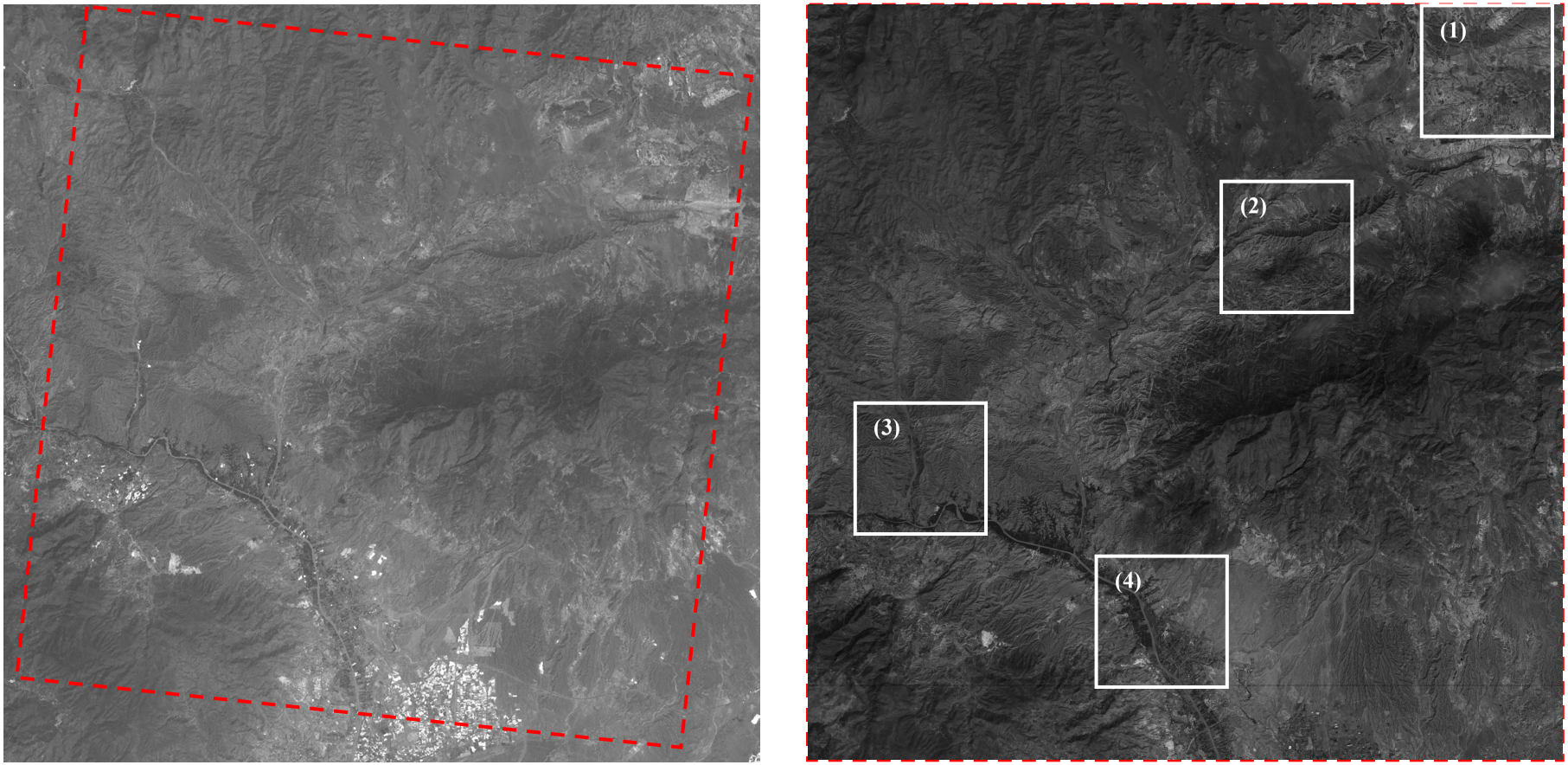


Figure 4. 10: (a) the input image taken by Landsat-7 ETM+ sensor (b) the reference image taken by SPOT sensor. The red squares highlight the overlapped areas in both images. The white squares denote the zoom-in sub areas of the reference image displayed in Figure 4.11.

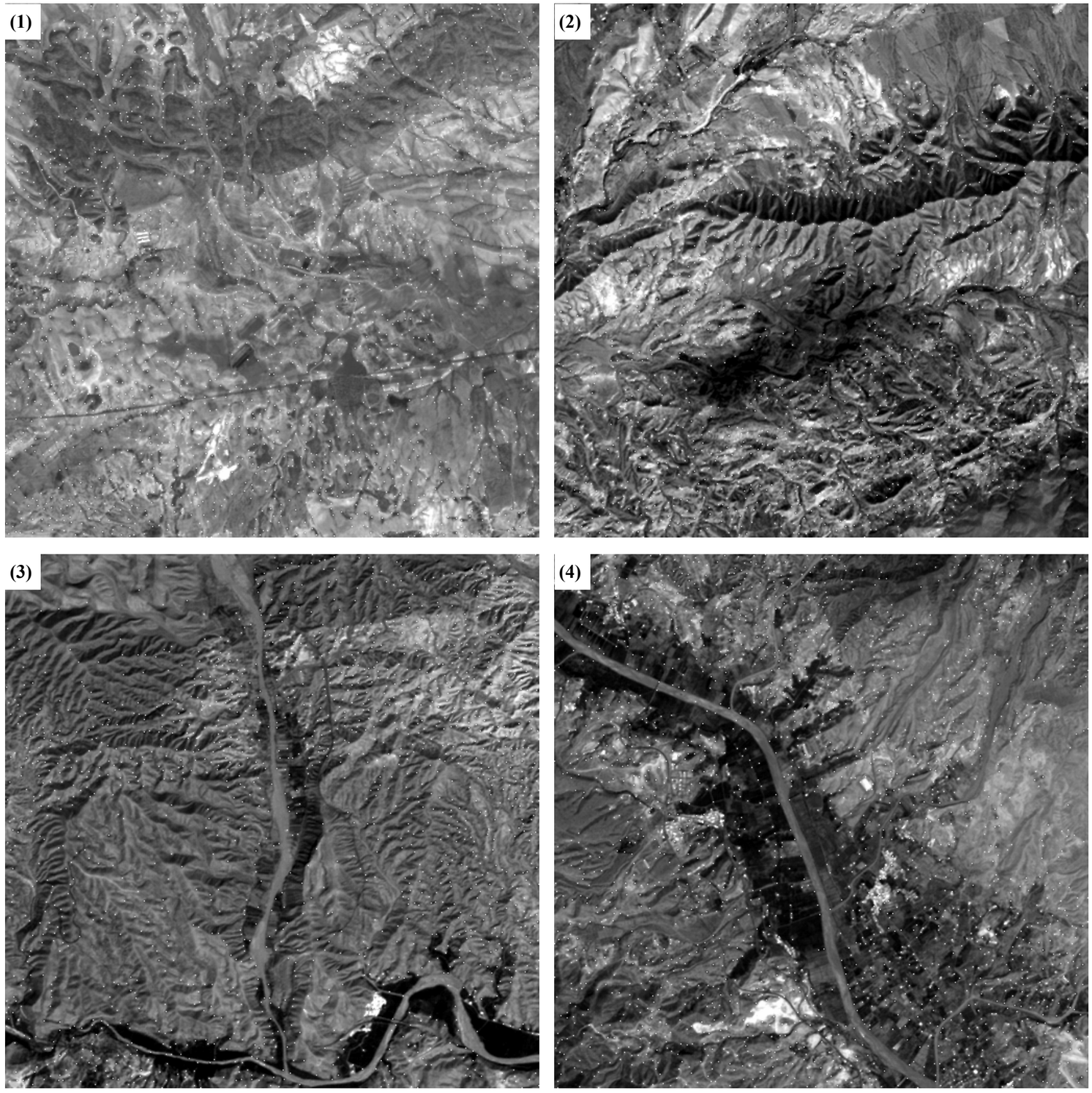


Figure 4. 11: The white dots denote the interest points selected using the ACD in the overlapped area of the reference image (four zoom-in sub images as white boxes highlighted in Figure 4.10b).

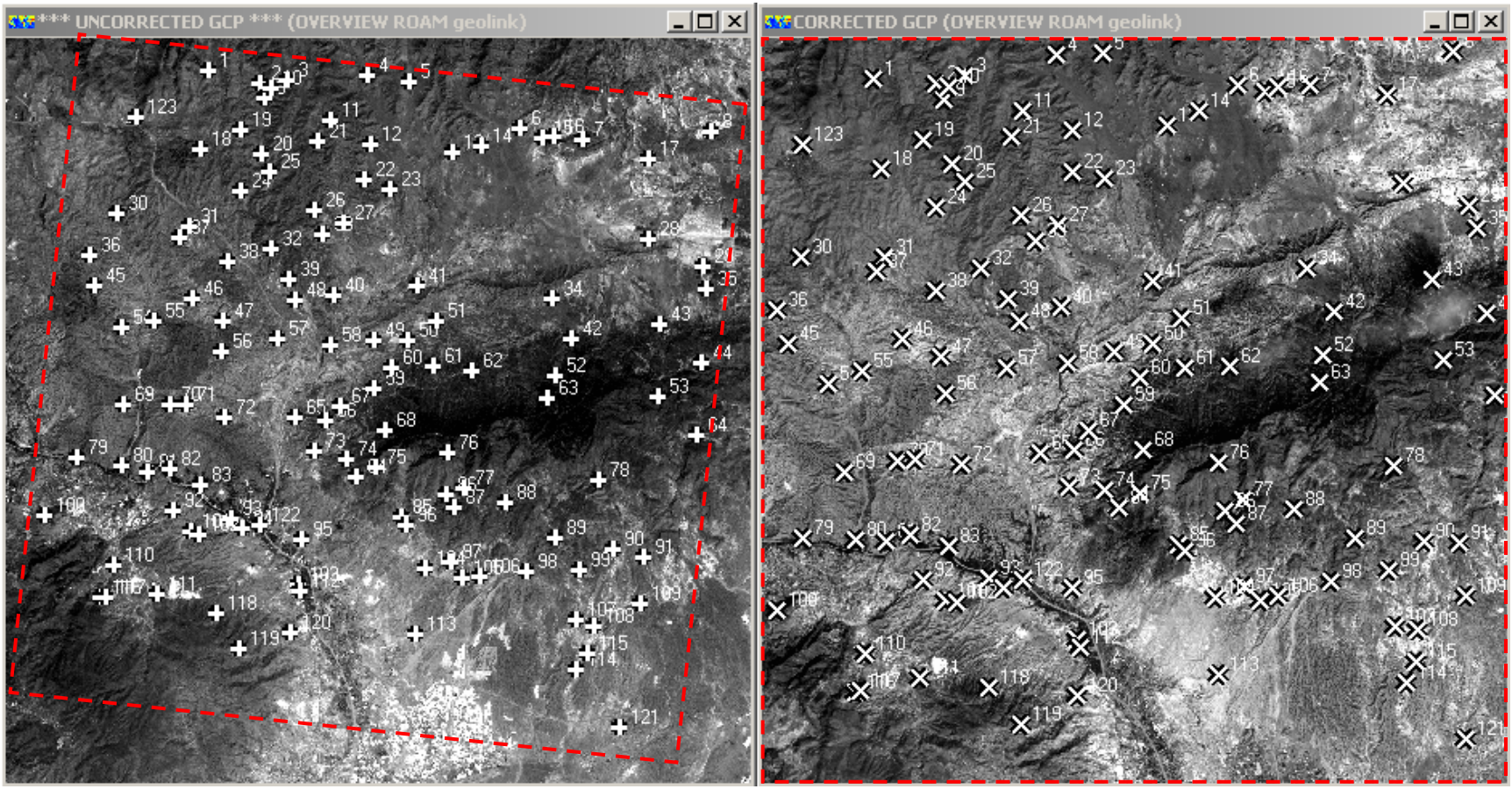


Figure 4. 12: 123 precisely matched GCP pairs detected in the image pair. The left one is the input image and the right one is the reference image. Red squares denote the overlapped areas

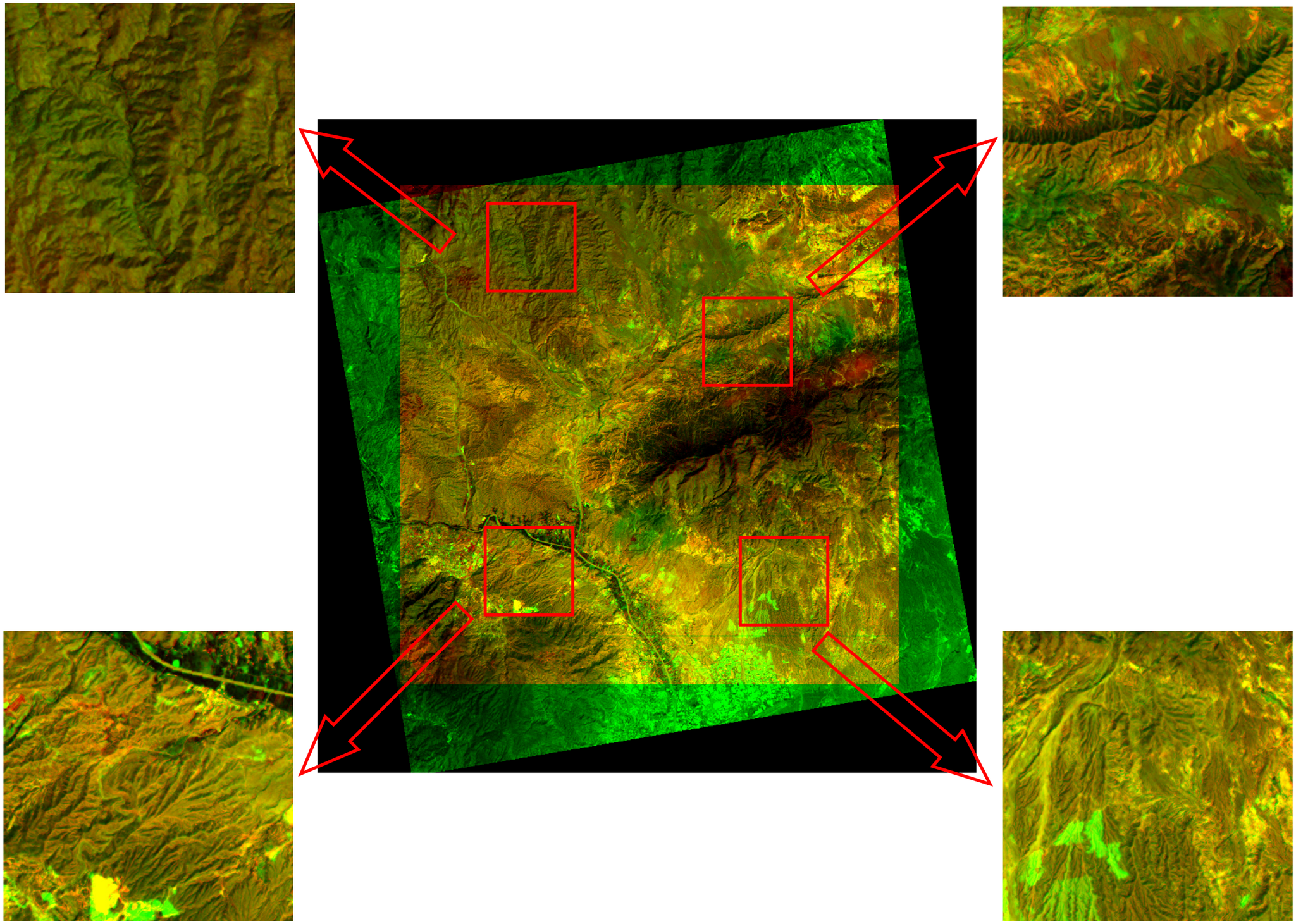


Figure 4. 13: The co-registration result and four zoom-in sub scenes. The co-registered image and reference image are displayed in green and red, respectively. The area displaying in yellow denotes the overlapped area.

Chapter 5

Discussion and conclusions

This chapter outlines the original contributions and major results of the research, and summarizes the methodology developed and the main conclusions obtained together with some suggestions for further research.

5.1 Original contributions and major results of the research

As an essential process for the remote sensing applications, co-registration of images taken by different sensors with various spatial resolutions requires high accuracy and efficiency. The selection of Ground Control Points (GCPs) is a key factor that accounts for both the co-registration accuracy and the processing automation and it is therefore a highly desired and challenging research.

One of the major original contributions of this research is that a novel Asymmetrical Corner Detector (ACD) algorithm is developed aiming to extract good quality Ground Control Points (GCPs) automatically from remotely sensed imagery data. The ACD is designed based on the auto-correlation matrix, with the consideration of the fact that the asymmetrical corner points are the most common reality in remote sensing images. It selects interest points more favourable to asymmetrical points rather than ideal symmetrical points to avoid selecting flat points which are often highly symmetrical. Experiments of several remote sensing images taken by Landsat ETM+ and SPOT sensors have proved that the ACD robustly achieves better performance in terms of point localization and computational time than the well established Harris Corner

Detector (Harris and Stephens, 1988) which is widely used and highly regarded in the literature. While the Harris Corner Detector contains an empirical parameter which brings ambiguity to its response function, the ACD has a clear and definite mathematical form without any empirical parameters.

In addition, a semi-automatic image co-registration scheme is thus proposed. The ACD algorithm is employed to extract evenly distributed GCPs across the overlapped image area in the reference image. The scheme uses three manually selected pairs of GCPs to determine the initial transformation model and the overlapped area. And then, grid-control and non-maximum suppression methods are used to secure the high quality and spread distribution of GCPs selected. The scheme also involves the Fast Normalized Cross-Correlation (FNCC) algorithm (Lewis, 1995) to refine the rough point localization in the input image obtained using the initial transformation model. The performance of the proposed registration scheme has been demonstrated by registering two multi-temporal Landsat TM images, two multi-resolution images taken by Landsat ETM+ and TM sensors, and multi-sensor images taken by SPOT and Landsat ETM+ sensors. Experimental results have indicated that this co-registration scheme is reliable to achieve consistent high registration accuracy of less than 0.7 pixels RMSE.

5.2 Methodology development

In the investigation, the methodology developed for interest point detection and image co-registration can be listed as follows:

(1) *Investigation of the nature of interest points based on characteristic models and manually selected Ground Control Points (GCPs).*

The primary characteristics of interest points have been investigated and five typical types of interest points which are widely recognized as good quality Ground Control Points (GCPs) in two-dimension grey-level remote sensing images are presented and compared. In addition, an in-depth investigation of the nature of manually selected GCPs from different remote sensing images is carried out.

(2) *Definition of the symmetry of interest points.*

The symmetry of interest points is defined on the basis of the thorough comparison and property investigation of the typical interest points. The essential factors influencing the point symmetry have been analyzed, such as the spatial resolution and the type of the interest point.

(3) *Analysis of the mathematical property and behaviour of the auto-correlation based Harris Corner Detector*.

One main investigation of this research is the analysis of the auto-correlation based Harris Corner Detector. The property and behaviour of this most widely used corner detector have been numerically analysed via the investigation of the relationship between the eigenvalues of the auto-correlation matrix *A* and the response function *R*. The ambiguity caused by the empirical parameter α in the response function *R* is revealed by isoline distribution graphs of different α settings.

(4) *Definition of the Asymmetrical Corner Detector (ACD) and investigation of its characteristics*.

In order to develop a robust corner detector which can overcome the limitations and drawbacks of the Harris Corner Detector, the function $(\lambda_1 - \lambda_2)^2$ has been first investigated and two corner detectors based on this function have been designed. The contour graphs are used to analyze their mathematical properties and effectiveness of corner detection. The design of the Asymmetrical Corner Detector is finalised with the approval of comprehensive observation on contour plots and isoline distribution.

(5) *Development of the semi-automatic image co-registration scheme*.

As a new feature for the semi-automatic image co-registration scheme, a grid-control method is introduced aiming to extract user-defined number and evenly distributed GCPs in the overlapped area of the reference image and also guarantee the quality of the selected GCPs. Non-maximum suppression is also applied with the ACD to avoid point clustering. The FNCC algorithm is used to refine GCP locations in the input image, then achieve high accuracy image co-registration in high computing efficiency.

5.3 Conclusions

With the methodology developed as summarized above, this study presents a useful algorithm to automatically detect high quality GCPs from remotely sensed imagery data, and a semi-automatic image co-registration scheme is thus proposed. The main conclusions of this research are as below.

- In remotely sensing imagery data, asymmetrical corner points (e.g. L-corners, T-junctions, Y-junctions and Arrow-junctions) represent the more common reality while the symmetrical interest point (e.g. X-junctions) is only a special case. Besides, along with the continuous improvement in spatial resolution, more and more symmetrical interest points in low resolution images will be transformed to asymmetrical interest points in high resolution images.

- Among several existing interest point selectors, the auto-correlation based Harris Corner Detector is a robust and stable technique for automatic GCPs selection. However, the empirical parameter α brings in ambiguity to the response function R employed in the Harris Corner Detector. In addition, the detector is prone to select highly symmetrical corner points which largely restrict its performance to extract high quality GCPs which are more often asymmetrical in remotely sensed earth observation images.

- An Asymmetrical Corner Detector (ACD) is thus designed based on the auto-correlation matrix, similar to the Harris Corner Detector but without any empirical parameters in the response function and with the consideration of the fact that the asymmetrical corner point is the most common reality in remote sensing images. The ACD selects interest points more favourable to asymmetrical points rather than ideal symmetrical points to avoid selecting flat points which are often highly symmetrical. The ACD obtains better performance in terms of point localization and computational time than the Harris Corner Detector;

- A semi-automatic image co-registration scheme is proposed. It is composed of six sequential stages: (a) Initial transformation model; (b) Interest point detection; (c) Grid-control; (d) Candidate GCP Pairs; (e) GCP refinement using FNCC, (f) Image rectification. Image co-registration experiments of multi-temporal, multi-sensor, and multi-resolution images taken by different satellite sensors have indicated that consistent high registration accuracy of less than 0.7 pixels RMSE has been achieved.

5.4 Critical appraisal of the research

This research presents an effective algorithm to extract asymmetrical corner points from the real remote sensing images obtained using aerial cameras or sensors mounted on the satellite platforms. A novel response function based on the auto-correlation matrix is thus designed without any empirical parameter contained. With a simple fixed threshold, the proposed Asymmetrical Corner Detector (ACD) is effective and efficient to select adequate high quality interest points in the images, no matter in what kind of land covers. The experimentation results have demonstrated that the ACD is more robust to select the asymmetrical interest points which can be used as Ground Control Points (GCPs) than the Harris Corner Detector.

However, like the Harris Corner Detector or other auto-correlation matrix based interest point detectors, the ACD is not invariant to significant scale changes. A multi-scale framework can be therefore added to the detector to avoid the degradation due to the presence of scale changes.

Another major contribution of this study is the development of a semi-automatic image co-registration scheme. The approach has been tested for the co-registration of multi-temporal, multi-sensor and multi-resolution satellite images taken by Landsat Thematic Mapper (TM), Landsat-7 Enhanced Thematic Mapper Plus (ETM+), and Satellite Pour l'Observation de la Terre (SPOT) sensors. Consistent high registration accuracy of less than 0.7 pixels RMSE has been achieved. In the scheme, the Fast Normalised Cross-Correlation (FNCC) algorithm (Lewis, 1995) is employed to refine the rough point localization in the input image obtained using the initial transformation

model. However, for the image pairs which have significantly different spectral characteristics, the grey-level cross-correlation based methods can not achieve satisfied matching results. Some other methods based on the structure similarity can be applied to improve the matching performance and guarantee the image co-registration accuracy.

5.5 Future work

The novel Asymmetrical Corner Detector (ACD) and semi-automatic image co-registration scheme developed in this research have the potential of a wide scope of remote sensing applications in high accuracy image co-registration. To fully explore this potential, the effectiveness and efficiency of the ACD need to be further evaluated using images of different land covers taken by other satellite sensors, such as high spatial resolution sensors, Quickbird-II and IKONOS.

The development of the ACD has put a light on corner detector design based on the investigation of manually selected GCPs. Study of the properties of manually selected GCPs and the corresponding mathematical models can lead to the design of variant corner detectors depending on the types of land covers.

The semi-automatic image co-registration scheme can be extended into other types of image co-registration, such as medical image co-registration. In addition, different matching methods other than FNCC, such as phase correlation, can be employed to improve the accuracy of feature matching and computational efficiency.

References:

Althof, R.J., Wind, M.G.J., and Dobbins, J.T., 1997. A rapid and automatic image registration algorithm with subpixel accuracy. *IEEE Transactions on Medical Imaging*, 16, 308-316.

Anuta, P.E., 1970. Spatial registration of multispectral and multitemporal digital imagery using Fast Fourier Transform. *IEEE Transactions on Geoscience Electronics*, 8, 353-368

Asada, H., and Brady, M., 1986. The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 1, 2-14.

Baker, S., Nayar S.K., and Murase, H., 1998. Parametric feature detection. *International Journal of Computer Vision*, 27, 1, 27-50.

Barnea, D.I., and Silverman, H.F., 1976. A class of algorithms for fast digital image Registration. *IEEE Transactions on Computing*, 21, 179-186.

Beaudet, P. R., 1978. Rotationally invariant image operators. *Proceedings of the 4th International Joint Conference Pattern Recognition*, Tokyo, Japan, pp.579-583.

Bentoutou, Y., Taleb, N., Kpalma, K., and Ronsin, J., 2005. An Automatic Image Registration for Applications in Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 9, 2127-2137.

Bernstein, R., 1983. Image Geometry and Rectification, *Manual of Remote Sensing (2nd Edition)*(pp. 873-922), American Society of Photogrammetry.

Berthilsson, R., 1998. Affine correlation. *Proceedings of the International Conference on Pattern Recognition ICPR'98*, Brisbane, Australia, pp. 1458-1461.

Buiten, H.J., and van Putten, B., 1997. Quality Assessment of Remote Sensing Registration – Analysis and Testing of Control Point Residuals. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52, 2, 57-73.

Campbell, J.B., 2002. Introduction to Remote Sensing. 3rd Edition. The Guildford Press, New York.

Canny, J., 1986. A computational approach to edge detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8, 679-698.

Castro, E.D., and Morandi, C., 1987. Registration of translated and rotated images using finite Fourier transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9, 700-703.

Chen, L., and Lee, L., 1992. Progressive generation of control frameworks for image registration. *Photogrammetric Engineering and Remote Sensing*, 58, 9, 1321-1328.

Chen, Q., Defrise, M., and Deconinck, F., 1994. Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16, 1156-1168.

Cheng, F.H., 1996. Point pattern matching algorithm invariant to geometrical transformation and distortion, *Pattern Recognition Letters*, 17, 1429-1435.

Cooper, J., Venkatesh, S., and Kitchen, L., 1993. Early jump-out corner Detectors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15, 8, 823-828.

Dai, X., and Khorram, S., 1999. A feature-based image registration algorithm using improved chain-code representation combined with invariant moment. *IEEE transactions on Geoscience and Remote Sensing*, 37, 5, 2351-2362.

Deriche, R., and Blaszka, T., 1993. Recovering and characterizing image features using an efficient model based approach. *Proceedings of the Conference on Computer Vision and Pattern Recognition*, New York, USA, pp.530-535.

Dodgson, N.A., 1997. Quadratic interpolation for image resampling, *IEEE Transaction on Image Processing*, 6, 1322-1326.

Dreschler, L., and Nagel, H.H., 1982. Volumetric model and 3D trajectory of a moving car derived from monocular TV frame sequences of a street scene. *Computer Graphics and Image Processing*, 20, 3, 199-228.

Flusser, J., and Suk, T., 1994. A moment-based approach to registration of images with affine geometric distortion. *IEEE Transactions on Geoscience and Remote Sensing*, 32, 382-387.

Fonseca, L.M.G., and Manjunath, B.S., 1996. Registration techniques for multisenor remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*, 62, pp.1049-1056.

Förstner, W., 1994. A framework for low level feature extraction. *Proceedings of the 3rd European Conference on Computer Vision*, Stockholm, Sweden, pp.383-394.

Förstner, F., and Gulch, E., 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Intercommission Conference on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland, pp.281-305.

Gao, S., Xiao, Y., and Hu, S.H., 2004. A comparison of two similarity measures in intensity-based ultrasound image registration. *Proceedings of the 2004 International Symposium on Circuits and Systems, ISCAS' 04*, Vancouver, Canada, 4, IV-61-4.

Gibson, P.J., and Power, C.H., 2000. Introductory Remote Sensing: Digital Image Processing and Applications. Routledge, New York.

Gonzalez, R.C., and Woods, R.E., 2002. Digital Image Processing. 2nd Edition. Pearson Education Publishing Company.

Goshtasby, A., 1986. Piecewise linear mapping functions for image registration. *Pattern Recognition*, 19, 459-466.

Goshtasby, A., 1987. Piecewise cubic mapping functions for image registration. *Pattern Recognition*, 20, 525-533.

Goshtasby, A., Stockman, G.C., and Page, C.V., 1986. A region-based approach to digital image registration with subpixel accuracy. *IEEE Transactions on Geoscience and Remote Sensing*, 24, 390-399.

Growe, S., and Tonjes, R., 1997. A knowledge based approach to automatic image registration. *Proceedings of the IEEE International Conference on Image Processing ICIP'97*, Santa Barbara, California, USA, pp.228-231.

Guiducci, A., 1988. Corner characterization by differential geometry techniques. *Pattern Recognition Letters*, 8, 5, 311-318.

Hanaizumi, H., and Fujimura, S., 1993. An automated method for registration of satellite remote sensing images. *Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS'93*, Tokyo, Japan, pp.1348-1350.

Haralick, R.M., and Shapiro, L.G., 1993. Compute and Robot Vision Volume II. Addison-Wesley Longman Publishing Company.

Harris, C., and Stephens, M., 1988. A combined corner and edge detector. in *Alvey Vision Conference*, pp.147-151.

Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E., and Kuebler, O. 1992. Simulation of neural contour mechanism: From simple to end-stopped cells. *Vision Research*, 32, 5, 963-981.

Holm, M., 1991. Towards automatic rectification of satellite images using feature based matching. *Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS'91*, Espoo, Finland, pp. 2439-2442.

Horaud, R., Skordas, T., and Veillon, F., 1990. Finding geometric and relational structures in an image. *Proceedings of the 1st European Conference on Computer Vision*, 374-384.

Hou, H.S., and Andrews, H.C., 1978. Cubic splines for image interpolation and digital filtering. *IEEE Transactions on Acoustic, Speech and Signal Prcessing*, 26, 508-517.

Hsieh, Y.C., McKeown, D.M., and Perlant, F.P., 1992. Performance evaluation of scene registration and stereo matching for cartographic feature extraction. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 14, 214-237.

Jensen. J. R., 1995. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition. Pearson Education Publishing Company.

Jung, II-Kyun., and S, Lacroix., 2001. A robust interest points matching algorithm. *Proceedings of the eighth IEEE international conference on computer vision ICCV 2001*, Vancouver, Canada, 2, pp. 538-543.

Kaneko, S., Murase, I., and Igarashi, S., 2002. Robust image registration by increment sign correlation. *Pattern Recognition*, 35, 2223-2234.

Kaneko, S., Satoh, Y., and Igarashi, S., 2003. Using selective correlation coefficient for robust image registration. *Pattern Recognition*, 36, 1165-1173.

Kitchen, L., and Rosenfeld, A., 1982. Gray-level corner detection. *Pattern Recognition Letters*, 1, 95-102.

Laganiere, R., and Vincent, T., 2002. Wedge-based corner model for widely separated view matching. *Proceedings of the 16th International Conference on Pattern Recognition*, 3, pp. 672-675.

Lee, K., and Bien, Z., 1996. A gray-level corner detector using fuzzy logic. *Pattern Recognition Letters*, 17, 9, 939-950.

Lewis, J. P., 1995. Fast Normalized Cross-Correlation. *Vision Interface*, pp.120-123.

Li, H., Manjunath, B.S., and Mitra, S.K., 1995. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4, 320-334.

Li, S.Z., Kittler, J., and Petrou, M., 1992. Matching and recognition of road networks from aerial images. *Proceedings of the Second European Conference on Computer Vision ECCV'92*, St Margherita, Italy, pp. 857-861.

Liu, J.G., 2002. Lecture Notes of Image Processing and GIS. Department of Earth Science and Engineering, Imperial College London, United Kingdom.

Liu, J.G., and Ma, J., 2004. Imageodesy on MPI & GRID for Co-seismic Shift Study Using Satellite Optical Imagery. *Proceedings of the UK e-Science All Hands Meeting*, Nottingham, United Kingdom, pp. 232-239.

Marinkovic, P., and Hanssen, R., 2004. Advanced InSAR coregistration using point clusters. P*roceedings of Geoscience and Remote Sensing Symposium, IGARSS '04*, 1, 489-492.

Marr, D., and Hildreth, E., 1980. Theory of edge detection. *Proceedings of the Royal Society of London*, B 207, 187-217.

Masood, A., and Sarfraz, A., 2006. A Novel Corner Detector Approach using Sliding Rectangles. *IEEE International Conference on Computer Systems and Applications*, pp. 621-626.

Medioni, G., and Yasumoto, Y., 1987. Corner detection and curve representation using cubic B-splines. *Computer Vision, Graphics and Image Processing*, 39, 267-278.

Mikolajczyk, K., and Schmid, C., 2002. An affine invariant interest point detector. In *Proceeding of the 7th Europe Conference on Computer Vision*, Copenhagen, Denmark, 2350, 128-142.

Mitra, R.S., and Murthy, N.N., 1991. Elastic maximal matching. *Pattern Recognition*, 24, 747-753.

Mokhtarian, F., and Mackworth, A., 1986. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 20, 12, 1376-1381.

Mokhtarian, F., and Suomela, R., 1998. Robust image corner detection through curvature scale space. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 20, 12, 1376-1381.

Moravec, H. P., 1977. Towards automatic visual obstacle avoidance. *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, Cambridge, Massachusetts, USA, p.584.

Moss, S., and Hancock, E.R., 1997. Multiple line-template matching with EM algorithm. *Pattern Recogntion Letters*, 18, 1283-1292.

Novak, K., 1992. Rectification of Digital Imagery. *Photogrammetric Engineering and Remote Sensing*, 58, 3, 339-344.

Parida, L., Geiger, D., and Hummel, R., 1998. Junctions: Detection, classification and reconstruction. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 20, 7, 687-698.

Pikaz, A., and Dinstein, I., 1994. Using Simple decomposition for smoothing and feature point detection of noisy digital curves. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 16, 8, 808-813.

Pratt, W.K., 1991. Digital Image Processing. 2nd Edition. Wiley, New York.

Ranade, S., and Rosenfeld, A., 1980. Point pattern matching by relaxation. *Pattern Recognition*, 12, 269-275.

Rattarangsi, A., and Chin, R. T., 1992. Scaled-Based Detection of Corners of Planar Curves. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 14, 4, 430-449.

Reddy, B.S., and Chatterji, B.N., 1996. An FFT-based technique for translation, rotation and scale-invariant image registration, *IEEE Transactions on Image Processing*, 5, 1266-1271.

Reisfeld, D., Wolfson, H., and Yeshurun, Y., 1995. Context-free attentional operators: The generalized symmetry transform. *International Journal of Computer Vision*, 14, 2, 119-230.

Rohr, K., 1992. Recognizing corners by fitting parametric models. *International Journal of Computer Vision*, 9, 3, 213-230.

Roux, M., 1996. Automatic registration of SPOT images and digitized maps. *Proceedings of IEEE International Conference on Image Processing ICIP'96*, Lausanne, Switzerland, pp.625-628.

Sawhney, H.S., and Kumar, R., 1999. True multi-image alignment and its applications to mosaicing and lens distortion corrections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21, 235-243.

Schmid, C., Mohr, R., and Bauckhage, C., 2000. Evaluation of Interest Point Detectors. *International Journal of Computer Vision*, 27, 2, 151-172.

Sester, M., Hild, H., and Fritsh, D., 1998. Definition of ground control features for image registration using GIS data. *Proceedings of the Symposium on Object Recogniton and Scene Classification from Multispectral and Multisensor Pixels*, CD-ROM, Columbus, Ohio.

Sharma, R.K., and Pavel, M., 1997. Multisensor image registration, *Proceedings of the Society for Information Display XXVIII*, 951-954.

Shi, J., and Tomasi, C., 1994. Good features to track. *Proceedings of the Conference on Computer Vision and Pattern Recognition*, Seattle, Washington, USA, pp.593-600.

Shinagawa, Y., and Kunii, T.L., 1998. Unconstrained automatic image matching using multiresolution critical-point filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 994-1010.

Smith, S.M., and Brady, J.M., 1997. SUSAN—A new approach to low level image processing. *International Journal of Computer Vision*, 23, 1, 45-78.

Thevenaz, P., Ruttimann, U.E., and Unser, M., 1998. A pyramidal approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing*, 7, 27-41.

Thevenaz, P., and Unser, M., 1998. An efficient mutual information optimizer for multiresolution image registration. *Proceedings of the IEEE International Conference on Image Processing ICIP'98*, Chicago, USA, IL, pp. 833-837.

Tomasi, C., and Kanade, T., 1991. Detection and tracking of point features. Technical report, Carnegie Mellon University.

Ton, J., and Jain, A.K., 1989. Registering Landsat images by point matching. *IEEE Transactions on Geoscience and Remote Sensing*, 27, 642-651.

Toraichi, K., Yang, S., and Mori, R., 1988. Two-dimensional spline interpolation for image reconstruction. *Pattern Recognition*, 21, 275-284.

Trajkovic, M., and Hedley, M., 1998. Fast corner detection. *Image and Visual Computing*, 16, 2, 75-87.

Turcajova, R., and Kautsky, J., 1996. A hierarchical multiresolution technique for image registration. *Proceedings of SPIE Mathematical Imaging: Wavelet Applications in Signal and Image Processing*, Colorado.

Vanderbrug, G.J., and Rosenfeld, A., 1977. Two stage template matching. *IEEE Transactions on Computers*, 26, 384-393.

Viola, P., and Wells, W.M., 1997. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24, 137-154.

Wang, W.H., and Chen, Y.C., 1997. Image registration by control points pairing using the invariant properties of line segments. *Pattern Recognition Letters*, 18, 269-281.

Wand, W.H., Sun, H., Yadas, S., and Rosenfeld, A., 1983. Some experiments in relaxation image matching using corner features. *Pattern Recognition*, 16, 167-182.

Wie, P.N., and Stein, M., 1977. A landsat digital image rectification system. *IEEE Transactions on Geoscience Electronics*, 15, 130-136.

Weinheimer, J., Qi, X.J., and Qi, J., 2006. Towards a Robust Feature-based Watermarking Scheme. *Proceedings of 2006 IEEE International Conference on Image Processing*, Atlanta, GA, USA, pp.1401-1404.

Wolberg, G., 1990. Digital Image Warping. 1st Edition. John Wiley – IEEE Computer Society, New York.

Wolberg, G., and Zokai, S., 2000. Image registration for perspective deformation recovery. *SPIE Conference on Automatic Target Reocogniton X*, Orlando, Florida, USA, p.12.

Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.T., 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78, 87-119.

Zhang, P., Guan, Z., and Wang, X., 2004. Algorithm for feature point in sequence image extraction and matching. *Geomatics and information Science of Wuhan University*, 29, 4, 329-332.

Zheng, Z., and Chellapa, R., 1993. A computational vision approach to image registration. *IEEE Transactions on Image Processing*, 2, 311-325.

Zhong, B., and Liao, W., 2007. Direct Curvature Scale Space: Theory and Corner Detection. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 29, 3, 508-512.

Zhou, D.X., Liu, Y.H., and Cai, X.P., 2004. An efficient and robust corner detection algorithm, *Proceeding of the 5th world congress on intelligent control and automation*, Hangzhou, P.R.China, pp. 4020-4024.

Zhu, Q., Avidan, S., and Cheng, Kwang-Ting., 2005. Learning a sparse, corner-based representation for time-varying background modelling, in *Proceeding of the 10th IEEE International Conference on Computer Vision*, Beijing, P.R.China, 1, pp. 678-685.

Appendix:

Kernel C codes of the ACD

ACD.h

```
// store image pixel greyvalue
typedef struct _IMAGE_VECTOR
{
	float Data;
}IMAGE_VECTOR, *INPUT_VECTOR;

// store ACD value
typedef struct _COARSE_OPERATOR1
{
	float index;

}COARSEOPERATOR1,*COARSE_OPERATOR1;

// grid-control
typedef struct _WPoints
{
	int *data;
	int e;

}WPoints,*WPOINTS;

// calculate ACD value
int ExecuteFeaturePointsACD(INPUT_VECTOR X, int Rows, int Cols, float c_min,
float r_min, float c_max, float r_max, int *Featurepoints, int grid_num);

//maximum-suppression
int MaximumSuppression(COARSE_OPERATOR1 P1, int nRows, int nCols, float
min_c, float min_r, float max_c, float max_r, int *Featurepoints);

//Grid-control
BOOL Gridpoints(COARSE_OPERATOR1 P1, int nRows, int nCols, int points_num,
float min_c, float min_r, float max_c, float max_r, int *Featurepoints, int *points,
int num);

BOOL Weight(WPOINTS P, int n, int m, int num, COARSE_OPERATOR1 G, int
num_points, int cols, int *Points);

// Polynomials transformation
BOOL Transformation(float *A, float *B, float *C, int size, int basesize);

// Interpolation
float Resample(INPUT_VECTOR X, int col, float x, float y);

void CreateGauss(float sigma, float **pdKernel, int *pnWindowSize);
```

```
void GaussianSmooth(int rmin, int rmax, int cmin, int cmax, int col, int row,
INPUT_VECTOR X, INPUT_VECTOR Y, float sigma);
```

ACD.cpp

1% Cut-off Contrast Enhancement

```
int max, min, i;
int cut_off = int ((Row_base * Col_base)/100);

typedef struct _Histogram
{
	int num;
}Histogram,*HISTOGRAM;

HISTOGRAM histogram = (HISTOGRAM)malloc(256 * sizeof(Histogram));
if (histogram==NULL)
{
	AfxMessageBox("malloc failed");
	return;
}

for (i=0;i<256;i++)
{
	histogram[i].num = 0;
}

int greyvalue = 0;

for (r=0; r<nRows_base; r++)
{
	for (int c=1; c<nCols_base-Col_base2 + 1; c++)
	{
		greyvalue = BaseImageVector[r * nCols_base + c].Data;
		histogram[greyvalue].num ++;
	}
}

int cutoff = 0;

// 1%
for (i=0;i<256;i++)
{
	cutoff = cutoff + histogram[i].num;
	if (cutoff >= cut_off) break;
}

min = i;
cutoff = 0;

for (i=255;i>0;i--)
{
	cutoff = cutoff + histogram[i].num;
	if (cutoff >= cut_off) break;
}
```

```
max = i;
for (r=0; r<nRows_base;r++)
{
	for (int c=1; c<nCols_base-Col_base2 + 1;c++)
	{
		if (BaseImageVector[r * nCols_base + c].Data > max)
		{
			BaseImageVector[r * nCols_base + c].Data = 255;
		}
		else if (BaseImageVector[r * nCols_base + c].Data < min)
		{
			BaseImageVector[r * nCols_base + c].Data = 0;
		}
		else
		{
			BaseImageVector[r * nCols_base + c].Data = 255 *
(BaseImageVector[r * nCols_base + c].Data - min)/(max - min);
		}
	}
}

if (histogram!=NULL)
     free(histogram);
```

Gaussian Filter

```
void CreatGauss(float sigma, float **pdKernel, int *pnWidowSize)
{
	LONG i;

	int nCenter;

	float dDisX;
	float dDisY;

	float dValue;
	float dSum;
	dSum = 0;

	*pnWidowSize = int(1+ 2*floor(3*sigma));

	nCenter = (*pnWidowSize)/2;

	*pdKernel = new float[(*pnWidowSize)*(*pnWidowSize)];

	for (i=0;i<(*pnWidowSize);i++)
	{
		for(int j=0;j<(*pnWidowSize);j++)
		{
			dDisX = float(i - nCenter);
			dDisY = float(j - nCenter);
			dValue = float(exp(-
(0.5)*(dDisX*dDisX+dDisY*dDisY)/(sigma*sigma))/(sqrt(2*3.1415926)*sigma));
			(*pdKernel)[i * (*pnWidowSize) + j] = dValue;
			dSum+=dValue;
		}
```

```
        }
        for (i=0;i<(*pnWidowSize);i++)
        {
                for(int j=0;j<(*pnWidowSize);j++)
                {
                        (*pdKernel)[i *(*pnWidowSize) + j]/=dSum;
                }
        }
}

// X- input Y- output Z-xdirection convolution output

void GaussianSmooth(int rmin, int rmax, int cmin, int cmax, int col, int row,
INPUT_VECTOR X, INPUT_VECTOR Y, float sigma)
{
        int x, y;

        int i, j;

        int nWindowSize;

        int nLen;

        float *pdKernel;

        float dDotMul;

        float dWeightSum;

        CreatGauss(sigma, &pdKernel, &nWindowSize);

        nLen = nWindowSize/2;

        for (y=(rmin-4);y<(rmax+4);y++)
        {
                for(x=(cmin-4);x<(cmax+4);x++)
                {
                        if (x>(cmin+nLen) && x<(cmax-nLen) && y>(rmin+nLen)
&& y<(rmax-nLen))
                        {
                                dDotMul = 0;
                                dWeightSum = 0;
                                for (i=(-nLen);i<=nLen;i++)
                                {
                                        for (j=(-nLen);j<=nLen;j++)
                                        {
                                                dDotMul+=float(X[(y+j) * col +
(i+x)].Data* pdKernel[(j+nLen) * (nWindowSize) + (i+nLen)]);
                                                dWeightSum += pdKernel[(j+nLen) *
(nWindowSize) + (i+nLen)];
                                        }
                                }
                        }
                        Y[y*col+x].Data = dDotMul/dWeightSum;
                }
        }
```

```
        delete []pdKernel;
        pdKernel = NULL;
}
```

Non-Maximum Suppression

```
int MaximumSuppression(COARSE_OPERATOR1 P1, int nRows, int nCols, float
min_c, float min_r, float max_c, float max_r, int *featurePoints)
{
   // window size 7*7

        int c,r,c1,r1,e2;
        e2=0;

        for (r=int(min_r+3);r<int(max_r-3);r++)
        {
                for (c=int(min_c+3);c<int(max_c-3);c++)
                {
                        if (P1[r * nCols + c].index>7200)
                        // 7200 is the threshold for ACD
                        {
                                int e1=0;
                                for (r1=r-3;r1<=r+3;r1++)
                                {
                                        for (c1=c-3;c1<=c+3;c1++)
                                        {
                        if (P1[r * nCols + c].index > P1[r1 * nCols + c1].index) e1++;
                                        }
                                }
                                if (e1==48)
                                {
                                        e2++;
                                        featurePoints[e2*2+0]=c;
                                        featurePoints[e2*2+1]=r;
                                }
                        }
                }
        }
    return e2;
}
```

Grid-Control

```
BOOL Gridpoints(COARSE_OPERATOR1 P1, int nRows, int nCols, int points_num,
float min_c, float min_r, float max_c, float max_r, int *Featurepoints, int *points,
int num)
{
        int col[50], row[50],col_num,row_num;

        col[0]=int(min_c);
        row[0]=int(min_r);
        col_num=int((max_c-min_c)/num);
        row_num=int((max_r-min_r)/num);

        for (int i=1;i<=num;i++)
```

```
	{
		col[i]=col[i-1]+col_num;
	}

	for (int j=1;j<=num;j++)
	{
		row[j]=row[j-1]+row_num;
	}

	WPOINTS Points = (WPOINTS)malloc(num*num*sizeof(WPoints));

	if(Points==NULL)
	{
		AfxMessageBox("malloc failed");
		return -1;
	}

	int c,r;
	c=r=0;

	for (j=0;j<num;j++)
	{
		for (int ii=0;ii<num;ii++)
		{
			Points[ii].data=(int*)malloc(sizeof(int)*col_num*row_num);
			Points[ii].e=0;
			int e=Points[ii].e;
			for(i=0;i<points_num;i++)
			{
				c=points[i*2+0];
				r=points[i*2+1];
				if (r<row[j+1]&&r>row[j])
				{
					if (c<col[ii+1]&&c>col[ii])
					{
						Points[ii].data[e*2+0]=c;
						Points[ii].data[e*2+1]=r;
						e++;
						Points[ii].e=e;
					}
				}
			}
			if (e!=0)
			{
				Weight(Points,ii,j,num,P1,e,nCols,Featurepoints);
			}
			if (Points[ii].data!=NULL)
				free(Points[ii].data);
		}
	}
	num=num*num;
	return TRUE;
}
```

```
BOOL Weight(WPOINTS P, int n, int m, int num, COARSE_OPERATOR1 G, int
num_points, int cols, int *Points)
{
	int c,r,c1,r1;

	c=P[n].data[0];
	r=P[n].data[1];

	for(int i=1;i<num_points;i++)
	{
		c1=P[n].data[i*2+0];
		r1=P[n].data[i*2+1];

		if(G[r * cols + c].index > G[r1 * cols + c1].index) continue;
		else
		{
			c=c1;
			r=r1;
		}
	}

	Points[(m*num+n)*2+0]=c;
	Points[(m*num+n)*2+1]=r;

	return TRUE;
}
```

ACD

```
int ExecuteFeaturePointsACD(INPUT_VECTOR X, int Rows, int Cols, float c_min,
float r_min, float c_max, float r_max, int *Featurepoints, int grid_num)
{

//Gaussian convolution

	INPUT_VECTOR
Ix=(INPUT_VECTOR)malloc(Rows*Cols*sizeof(IMAGE_VECTOR));

	if(Ix==NULL)
	{
		AfxMessageBox(" malloc failed");
		return -1;
	}

	INPUT_VECTOR
Iy=(INPUT_VECTOR)malloc(Rows*Cols*sizeof(IMAGE_VECTOR));

	if(Iy==NULL)
	{
		AfxMessageBox(" malloc failed");
		return -1;
	}

	INPUT_VECTOR
Ixy=(INPUT_VECTOR)malloc(Rows*Cols*sizeof(IMAGE_VECTOR));
```

```
	if(Ixy==NULL)
	{
		AfxMessageBox(" malloc failed");
		return -1;
	}

	int c,r;

	for (r=int(r_min-4);r<int(r_max+4);r++)
	{
		for (c=int(c_min-4);c<int(c_max+4);c++)
		{
			float dx,dy;
			dy = X[(r+1) * Cols+ c].Data - X[(r-1) * Cols +c].Data;
			dx = X[r * Cols+ c+1].Data - X[r * Cols +c-1].Data;
			Ix[r * Cols + c].Data = dx*dx;
			Iy[r * Cols + c].Data = dy*dy;
			Ixy[r * Cols + c].Data = dx*dy;
		}
	}

	INPUT_VECTOR
Ix2=(INPUT_VECTOR)malloc(Rows*Cols*sizeof(IMAGE_VECTOR));

	if(Ix2==NULL)
	{
		AfxMessageBox(" malloc failed");
		return -1;
	}

	INPUT_VECTOR
Iy2=(INPUT_VECTOR)malloc(Rows*Cols*sizeof(IMAGE_VECTOR));

	if(Iy2==NULL)
	{
		AfxMessageBox(" malloc failed");
		return -1;
	}

	INPUT_VECTOR
Ixy2=(INPUT_VECTOR)malloc(Rows*Cols*sizeof(IMAGE_VECTOR));

	if(Ixy2==NULL)
	{
		AfxMessageBox(" malloc failed");
		return -1;
	}

	float Sigma=0.5; // standard deviation of Gaussian Convolution

	int min_r = int(r_min);
	int max_r = int(r_max);
	int min_c = int(c_min);
	int max_c = int(c_max);

	GaussianSmooth(min_r,max_r,min_c,max_c,Cols,Rows,Ix,Ix2,Sigma);
```

```
	GaussianSmooth(min_r,max_r,min_c,max_c,Cols,Rows,Iy,Iy2,Sigma);

	GaussianSmooth(min_r,max_r,min_c,max_c,Cols,Rows,Ixy,Ixy2,Sigma);

	if (Ix!=NULL)
		free(Ix);
	if (Iy!=NULL)
		free(Iy);
	if (Ixy!=NULL)
		free(Ixy);

//calculate ACD Value

	COARSE_OPERATOR1
Index=(COARSE_OPERATOR1)malloc(Rows*Cols*sizeof(COARSEOPERATOR1));

	if(Index==NULL)
	{
		AfxMessageBox(" malloc failed");
		return -1;
	}

	for (r = 1; r< Rows; r++)
		for (c = 1; c< Cols; c++)
		{
			Index[r * Cols + c].index = 0;
		}

//Lapack package to calculate eigenvalues

	integer ncol;
	integer lwork;
	char jobu,jobvt;
	integer info2 = 0;
	integer lda = 10;
	ncol = 2;
	lwork = 10 * ncol;

	jobu = 'N';
	jobvt = 'N';
	double u[4];
	double vt[4];
	double eig[2];
	double work[20];


	for (r=int(r_min-2);r<int(r_max+2);r++)
	{
		for (c=int(c_min-2);c<int(c_max+2);c++)
		{
			int c1,r1;
			c1=r1=0;

			float q1,q2,q3,q4;
			q1=q2=q3=q4=0;
```

```
                        for (r1=r-1;r1<=r+1;r1++)
                        {
                                for (c1=c-1;c1<=c+1;c1++)
                                {
                                        q1 = q1 + Ix2[r1 * Cols + c1].Data;
                                        q2 = q2 + Ixy2[r1 * Cols + c1].Data;
                                        q3 = q2;
                                        q4 = q4 + Iy2[r1 * Cols + c1].Data;
                                }
                        }

                        double Q[4];
                        Q[0]=q1;
                        Q[1]=q2;
                        Q[2]=q3;
                        Q[3]=q4;

                        dgesvd_(&jobu, &jobvt, &ncol, &ncol, Q, &ncol, eig, u, &ncol,
vt, &ncol, work, &lwork, &info2);

                        if (eig[0]!=0&&eig[1]!=0)
                        {
                                Index[r * Cols + c].index =
float(2*eig[0]*eig[1]*max(eig[0],eig[1]))/(eig[0]*eig[0]+eig[1]*eig[1]);

                        }
                }
        }

        if (Ix2!=NULL)
                free(Ix2);
        if (Iy2!=NULL)
                free(Iy2);
        if (Ixy2!=NULL)
                free(Ixy2);

        int e;

        int *feature_points=(int*)malloc(2*Rows*Cols*sizeof(int));

        e=MaximumSuppression(Index,Rows,Cols,c_min,r_min,c_max,r_max,featu
re_points);

        Gridpoints(Index,Rows,Cols,e,c_min,r_min,c_max,r_max,Featurepoints,fea
ture_points,grid_num);

        if (Index!=NULL)
                free (Index);

        if (feature_points!=NULL)
                free(feature_points);

        int e0=grid_num*grid_num;
        return e0;
}
```