

PHASE CORRELATION-BASED ILLUMINATION-INSENSITIVE IMAGE MATCHING FOR TERRAIN- RELATED APPLICATIONS

This thesis is presented for the degree of Doctor of Philosophy

Xue Wan

Imperial College London, Department of Earth Science and
Engineering

October 2015

Declaration of Originality

I, Xue Wan, declare that this thesis is my original work and I have clearly referenced all sources used in the work in both text and references.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

The Earth surface is monitored by air-borne and space-borne sensors locally, regionally and globally and at short time intervals. The capability of multi-temporal image matching is essential for many terrain-related applications, such as glacier motion monitoring, landslide localisation and earthquake deformation assessment. Robust image matching using time-varying remotely sensed data is, however, not always achievable because images are acquired separately, under different illumination conditions. For remotely sensed images of natural landscapes, such as mountains, deserts and glaciers, the majority of image patterns is produced by topographic features, such as shading and shadows. The appearances of these image patterns can be greatly altered by illumination variation, especially local illumination variation caused by solar position change, and thus parts of the images may lose correspondence for accurate matching.

This research provides a thorough mathematical analysis of the impact of local illumination variation via Phase Correlation (PC) cross power spectrum in the Fourier frequency domain and thus proves the illumination-insensitive property of PC-based image matching algorithms. Specifically, the relationship between PC fringe appearance and sun angle variation is investigated. Two improved PC based approaches, ADCF-PC (Absolute Dirichlet Curve Fitting based Phase Correlation) and PLSF-PC (Piecewise Least Square Fitting based Phase Correlation), are presented and compared for illumination-insensitive image matching. Finally, three original approaches to explore terrain related applications have been developed including: vision-based UAV navigation, illumination-insensitive change detection and DEM generation from multi-temporal remotely sensed images.

Acknowledgements

I would like to thank Dr Jianguo Liu for guiding me into the PhD research field; for supporting my new ideas; for helping me whenever I was stuck; for encouraging me when I was not confident; for inspiring and encouraging me to become better; for being a nice supervisor and a good friend.

I also want to thank my co-supervisor Dr Philippa Mason for her support and suggestion for my PhD work; for the careful proof reading of my thesis; for always being enthusiastic every time I talked to you.

I would like to thank my group to help, support and inspire me during the PhD research. I would like to thank Dr Hongshi Yan for helping me to enhance my programming skills and Dr Gareth Morgan for helping me establish my experiments. I acknowledge Dr Jianguo Liu, Dr Hongshi Yan and Dr Gareth Morgan for establishing the software PCIAS and allow me to develop further investigation based on prior research in our group. I am proud to be one member of this group, and I am happy to work with you guys.

I would like to thank all my colleagues and friends in ESE department Imperial College to make my PhD life so colourful and wonderful; I shall never forget the Friday Fribes, the Snowdonia fieldtrip, the badminton playing, the Christmas quiz and all the happy moments we spent together.

I would like to thank my family for supporting me during my PhD; My mother and father, for their endless love and contribution to me; my grandfather and grandmother, for taking care of me since my childhood; my father-in-law and mother-in-law, for their completely support to my work; and finally my husband, Xiaozhi, for being my best friend, for sharing our thoughts, hopes and dreams, and for loving, respecting and always being there for me.

Table of Contents

Declaration of Originality.....	2
Copyright Declaration.....	3
Abstract.....	4
Acknowledgements	5
Table of Contents.....	6
Chapter 1 Introduction.....	9
1.1 Background.....	9
1.2 Multi-temporal remote sensing image matching.....	10
1.3 Illumination variation type.....	13
1.4 Research objectives	15
1.5 The structure of the thesis.....	16
Chapter 2 Illumination-insensitive Image Matching	19
2.1 Feature-based approaches	19
2.1.1 Pre-processing	19
2.1.2 Feature extraction algorithms	20
2.2 Area-based approaches.....	24
2.2.1 Normalised Cross Correlation (NCC)	25
2.2.2 Mutual Information (MI).....	26
2.2.3 Phase Correlation (PC)	26
Chapter 3 Phase Correlation Decomposition	28
3.1 Phase Correlation	28
3.2 Phase Correlation matrix decomposition.....	30
3.3 PC matrix decomposition experiments	32
3.4 Summary.....	36
Chapter 4 The Impact of Local Illumination Change to Phase Correlation.....	37
4.1 The <i>SAI</i> Space	37
4.2 Fourier transform of <i>SAI</i>	42
4.3 Effects of zenith angle variation on Phase Correlation matrix	45
4.4 Effects of azimuth angle variation on Phase Correlation matrix	48

4.5 Effects of reflectance rate variation on Phase Correlation matrix	56
4.6 Illumination-insensitive PC matching.....	57
4.6.1 Sub-pixel PC matching.....	57
4.6.2 Robustness of Absolute Dirichlet function Curve Fitting Phase Correlation (ADCF-PC) to illumination variation	59
4.6.3 Illumination-insensitive Piecewise Least Square Fitting Phase Correction (PLSF-PC) for sub-pixel image alignment.....	60
4.7 Experiments.....	62
4.7.1 Effect of illumination variation on the PC matrix.....	62
4.7.2 Sub-pixel image matching results and comparison.....	72
4.8 Summary	80
Chapter 5 Autonomous Vision-based UAV Navigation in a Changing Illumination Environment	81
5.1 Vision-based UAV navigation	81
5.2 Proposed PC-based autonomous UAV navigation method	84
5.2.1 Geometric distortion rectification.....	85
5.2.2 Query and search image generation.....	86
5.2.3 Robust PC based image matching.....	87
5.2.4 Position estimation quality assessment	90
5.2.5 Risk assessment and flight route adjustment	91
5.3 UAV navigation experiments	92
5.3.1 DEM simulated images	92
5.3.2 Analogue terrain model experiments	100
5.3.3 Demonstration of potential applications	104
5.4 Summary	112
Chapter 6 Illumination Invariant Change Detection.....	113
6.1 Remote sensing change detection	113
6.1.1 Change types.....	113
6.1.2 Problems in remote sensing change detection.....	115
6.2 Related Research	117
6.3 Principle of Phase Correlation based illumination-insensitive change detection....	119
6.4 PC-Based change detection algorithm under variant illumination conditions.....	123
6.4.1 Pixel-wise PC based dense matching	124
6.4.2 Initial change map generation	125
6.4.3 Change type identification.....	126

6.4.4 Small object change detection.....	129
6.5 Experiments.....	133
6.5.1 DEM hill shading image.....	133
6.5.2 Natural terrain model image	144
6.5.3 Example applications	149
6.6 Summary.....	159
Chapter 7 DEM Generation Based on Multi-temporal Remotely Sensed Images	160
7.1 Illumination-insensitive stereo matching	160
7.2 PC-based disparity map generation from time variant remotely sensed imagery ..	162
7.2.1 Stereo rectification.....	162
7.2.2 Illumination-insensitive image alignment.....	164
7.2.3 PC based pixel wise dense matching.....	165
7.3 Experiments.....	166
7.3.1 Hill shading images	166
7.3.2 Analogue terrain model images under daily illumination conditions.....	174
7.3.3 Application example: Martian surface 3D reconstruction using HiRISE images.....	178
7.4 Summary.....	184
Chapter 8 Summary	185
8.1 Summary of research	185
8.2 Original contributions	187
8.3 Implications for multi-temporal remotely sensed image processing.....	192
8.4 Future work	193
8.4.1 Forward-view on-board UAV navigation.....	193
8.4.2 Multi-spectral bands based change detection.....	193
8.4.3 Disparity map refinement.....	194
List of Figures.....	195
List of Tables	204
List of Abbreviations	205
List of Symbols.....	207
Reference.....	210

Chapter 1 Introduction

1.1 Background

Remote sensing refers to the techniques which acquire an objective record of the spectrum relating to the physical properties and chemical composition of Earth surface materials [1]. Since the launch of the first remote sensing satellite, Landsat 1, in 1972, remote sensing has been developed into an essential tool for earth environment monitoring and applied widely in various fields name a few, landuse/cover mapping, topographic survey, geological study and planetary exploration.

Nowadays, many Earth Observation (EO) satellites are in orbit to monitor the Earth surface globally at short time intervals and at high spatial resolutions. There are also airborne sensors, such as fixed wing aircrafts, balloons and UAVs (Unmanned Aerial Vehicles), providing local and regional image data. To analysis and extract thematic information from such large scale data collection of remotely sensed images for particular applications, a key research issue is to develop effective technology that provides the capability of multi-temporal image matching. The task sounds simple, but robust high precision image matching between multi-temporal EO images is not always achievable, because images acquired in different seasons and at different times may be subject to significant variations in solar illumination, which changes the hill shading greatly, making it difficult to find correspondence between images.

This study focuses on developing an illumination-insensitive image matching algorithm which can be applied in terrain related applications. Specifically, the applications that motivate my research are three fold:

- Vision based UAV navigation. UAVs can reach places and areas which are difficult or impossible to access on ground [2], and they can be flexibly

deployed whenever there is a need. Light and low cost UAV systems are suitable for the evaluation of geo-hazards [3, 4]. Without a human pilot, UAV autonomous navigation becomes a crucial issue. This thesis is aimed at achieving autonomous UAV navigation by using high-medium resolution satellite images as references for comparison with real time on-board imaging. While satellite images are, in general, acquired in the local morning time around 9-10 am, illumination-insensitive image matching becomes essential for the UAV flying in afternoon. There are quite a few algorithms to tackle the problem but a fast and reliable illumination-insensitive image matching method is yet to be developed.

- Multi-temporal remote sensing change detection. As one of the most active research fields in multi-temporal image analysis, change detection identifies either the appearance variation or the object motion on the Earth surface, for example landslides, earthquakes, glacier retreat and volcanic activities. However, Illumination angle change results in topographic shade and shadow variations. This research is aimed at pinpointing real ground changes; conventional change detection will fail because of shadow changes between images captured under different sun illumination angles.
- Multi-temporal stereo matching for DEM (Digital Elevation Model) reconstruction. Conventional stereo images are captured simultaneously but standard stereo images may not be available in certain areas or at certain times. Thus, the capability for generating DEMs from images captured at different times, seasons and years and by different sensors, will expand available the data sources for this purpose.

1.2 Multi-temporal remote sensing image matching

Image matching is a fundamental step in the transformation of multi-temporal and multi-sensor remote sensing images into one coordinate system, so the images can be compared and further analysed. The core algorithm in image matching should find correspondence between images taken at different time or by different sensors. However, this is not always achievable because the appearance of remote

sensing images acquired at different times and seasons can be very different. An example is demonstrated in Figure 1.1, which shows two satellite images of Tasman glacier acquired on 31st October 2006 and 11th April 2013. Although the two images cover exactly the same area, the shading patterns in the two images are so different that it is even not easy to find the correspondence visually.

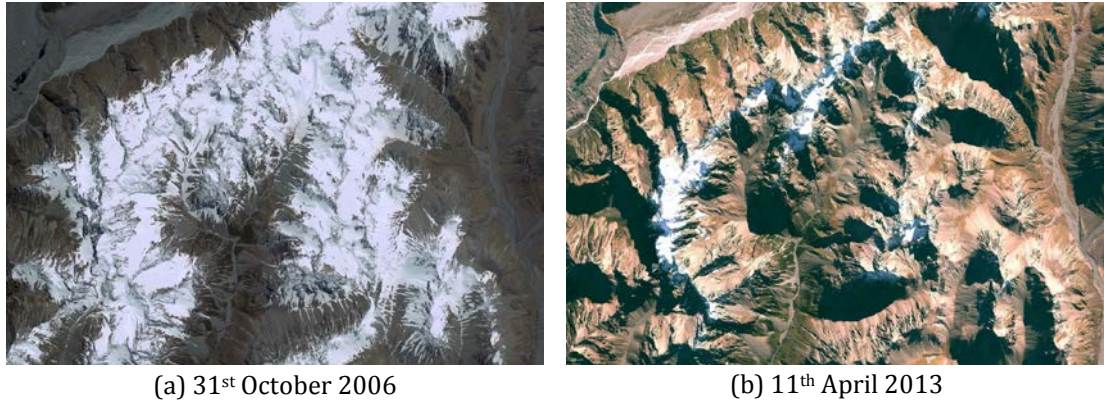


Figure 1.1 Multi-temporal satellite images of Tasman Glacier

There are several potential causes for the loss-correspondence between multi-temporal remote sensing images, and in this thesis, they are divided into three types: change in object surface, sensor-related distortion and illumination.

- Change in object surface.

Multi-temporal EO remotely sensed image acquisition may be separated by a long period of months and years, during which time various changes in object space may have happened on Earth surface, as illustrated in Figure 1.1, where (a) was taken during winter, so that the snow area is much larger than that in (b) taken during summer. Another object space change is motion, the movement of a target within a scene, e.g. a vehicle. The changes in object space result in low correspondence rates in the changed areas while the unchanged areas can still have high rate of correspondence. This is the key principle to detect change in multi-temporal images, which is one of the application objectives of this research.

- Sensor related distortion

If multi-temporal images are taken by the same sensor but with different geometric attitudes, the shapes of the image objects will be distorted; such geometric distortion caused by camera motion can be regarded as a *global* image distortion. A more complex geometric distortion is introduced when different sensors are used, since the imaging geometry, focal length and field of view may all be different. For instance, most satellites sensors operate in push-broom scanning mode while aerial images are in central projection mode. The sensor-related distortions are usually corrected based on known sensor parameters, imaging geometry and platform 3D (3 dimension) status by data suppliers.

- Illumination change

This is significant factor causing the loss of correspondence in multi-temporal image matching. It can globally alter the image contrast and brightness, as shown in Figure 1.1, where (b) appears brighter and has higher contrast than (a). More seriously, the change in illumination can cause variations in local spatial patterns between two images, as demonstrated by Figure 1.1; there are deep shadows and shading along the steep mountain slopes in (b), while in (a), such effects are not in significant.

To find the real object changes and consistency information in multi-temporal remotely sensed images, the effects of sensor-related distortions and illumination conditions need to be removed. The geometric rectification problem has been well studied and effectively dealt with [5-7]. For most remotely sensed images, the geometric distortions caused by camera motion and multiple sensors can be rectified by applying a low-dimensional transformation models: similarity, affine and perspective [8], or more complex 3D rational functions [9, 10]. The geometric correction to make the images conforming to a given map projection can be done without image matching, but it cannot assure precise image matching at pixel level for effective detection of the object changes. Alternatively, a more sophisticated geometric correction based on pixel-wise co-registration [11, 12] can be applied to remove the irregular geometric distortions at every pixel. Compared to global

geometric registration, the pixel-wise co-registration more thoroughly eliminates the geometric discrepancies between multi-temporal and multi-sensor images. However, any algorithm based on image matching is subject to the effects of illumination and if the images for matching are taken under different illumination conditions, the task becomes difficult and challenging. Thus, overcoming the effects introduced by illumination variation is an important and useful research topic for multi-temporal remote sensing image matching.

1.3 Illumination variation type

The Sun is the principal illumination source for optical EO remote sensing. We can consider the Sun as a single point light source emitting parallel sunlight rays. The sunlight is reflected by the Earth surface, and the reflected radiation is measured by passive sensors to generate optical remotely sensed images. This imaging procedure can only take place during daytime when the Earth surface is illuminated by the sun.

For multi-temporal remote sensing, the light source varies in two aspects: strength (radiation intensity) and position. Solar radiation intensity determines the overall brightness of images and its variation is a global linear transformation to image intensity. An example of multi-temporal images produced by variant solar strength is demonstrated in Figure 1.2 in which, (a) was taken under a weaker solar strength than (b).

From Figure 1.2, the variation of solar strength causes differing global brightness in images, i.e. Figure 1.2(a) is darker than Figure 1.2(b) however, the local pattern and texture are not altered. Thus, the change in solar strength is named as *global illumination variation*. Apart from solar strength, another variable factor is solar position, which means that sunlight comes from different angles. The solar position is defined by zenith and the azimuth angles. As defined in Figure 1.3, *zenith* is the angular height of the sun in the sky measured from the vertical and *azimuth*, the direction of the sunlight as given by a compass bearing.

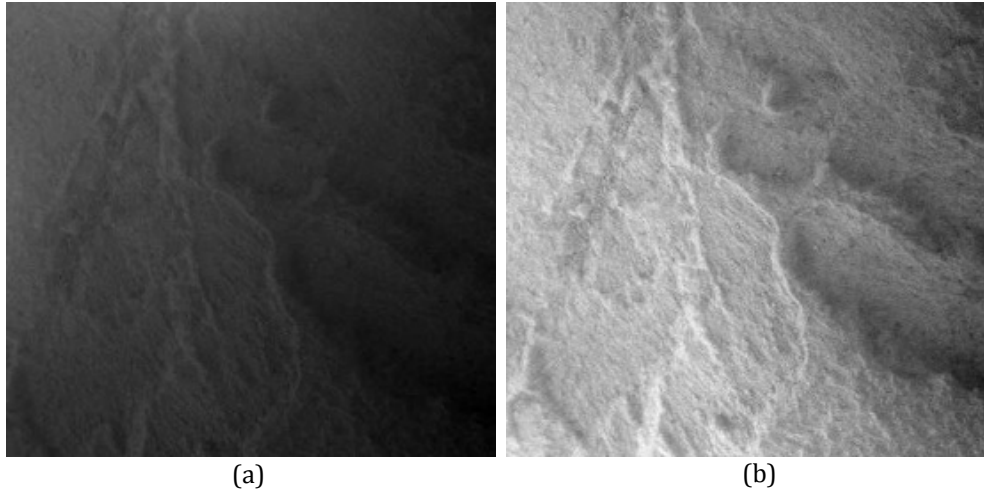


Figure 1.2 Two hill images taken under different solar strengths; (a) was taken under a weaker solar strength while (b) a strong solar strength.

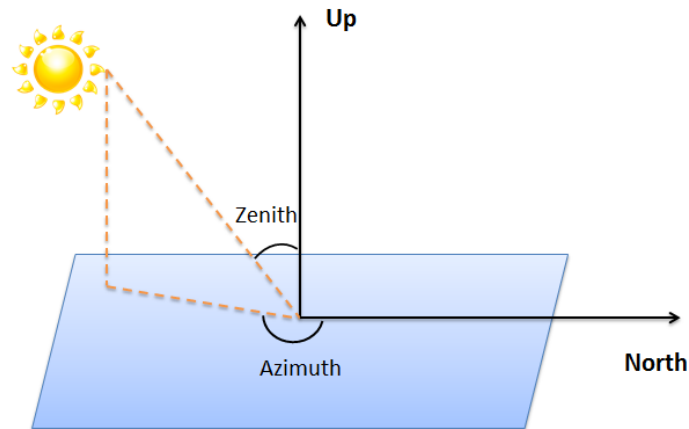


Figure 1.3 Definition of zenith angle and azimuth angle

The illumination variation caused by solar position change is named as *local illumination variation*. Compared to the effect of solar strength variation on image global brightness, local illumination variation not only changes image brightness but also alters the spatial pattern/texture of the image by changing the shapes and orientations of topographic shading and shadows. An example of local illumination variation is shown in Figure 1.4: two satellite images of a mountain terrain were taken in different seasons. The local patterns are altered greatly by the change in solar illumination angle; while the shaded areas in Figure 1.4 (a) are in general oriented north-south, those in (b) are orientated east-west direction.

Most image matching algorithms are robust to the linear transform of image intensity caused by global illumination variations, but few of them are truly robust to the changes in DN (Digital Number) value distribution caused by local illumination variation. High accuracy of matching is especially hard to achieve in high relief, mountainous areas [13] which are of the interest in geological study. Compared to low relief areas, high relief areas are more sensitive to local illumination variation, because of complex topographic patterns. In this thesis, we will focus on local illumination-insensitive multi-temporal image matching in high relief terrain.



(a) April 2007



(b) January 2007

Figure 1.4 Shading and shadow variation caused by imaging under different solar positions in a mountainous area. (a) and (b) are Google Earth images taken in different seasons.

1.4 Research objectives

From the above introduction, local illumination variation is one of the major challenges in multi-temporal image matching. This thesis focuses on the analysis of

local illumination variation in remotely sensed images of terrain, especially high relief areas. The aim is to quantify the effect of local illumination variation in relation to topographic relief in remote sensing images, and then develop a robust matching algorithm for multi-temporal image alignment as well as pixel-wise co-registration for applications of vision based UAV navigation, change detection and DEM generation.

Specifically, the study will accomplish the following objectives:

1. Investigate the impact of local illumination variation (azimuth and zenith) to Phase Correlation (PC). PC is an image matching algorithm based on frequency domain operations.
2. Improve the PC based image matching algorithm to achieve sub-pixel matching accuracy under locally illumination variations.
3. Develop a robust frame-reference localisation algorithm and test it for UAV navigation. The algorithm is insensitive to local illumination change and can tolerate image distortion caused by camera 3D motion.
4. Develop a robust illumination invariant change detection algorithm which is able to detect object temporal changes and motion from multi-temporal images taken in different solar illumination angles.
5. Accomplish an illumination-insensitive DEM generation methodology to derive DEM data from multi-temporal remote sensing images.

1.5 The structure of the thesis

The thesis is structured as follows:

Chapter 2 provides a review of image matching algorithms and discuss their robustness to local illumination change;

Chapter 3 presents the mathematical derivation of Phase Correlation decomposition and validation experiments;

Chapter 4 investigates the effect of illumination angle variation in terms of azimuth angle and zenith angle on Phase Correlation matrix via mathematical derivation, and then two approaches to improve Phase Correlation algorithm are introduced to achieve illumination-insensitive image matching for image alignment and dense matching;

Chapter 5 introduces a PC based frame-reference UAV navigation scheme followed by results of experiments of vision-based UAV navigation;

Chapter 6 presents an illumination robust change detection algorithm and the test results of four datasets including appearance change, motion detection and camouflage detection;

Chapter 7 demonstrates an illumination-insensitive disparity estimation approach and experiments of DEM generation from multi-temporal remote sensing images;

Chapter 8 is a summary of the study and major contributions and suggestions for future work.

Some of the content in Chapter 4 has been published in *IEEE Transactions on Geoscience and Remote Sensing* by Xue Wan, Jian Guo Liu, and Hongshi Yan as "The Illumination Robustness of Phase Correlation for Image Alignment." 10(2015):5746-5759.

Much of the content in Chapter 5 has been submitted to *ISPRS Journal of Photogrammetry and Remote Sensing* (Minor correction) by Xue Wan, Jian Guo Liu, Hongshi Yan and Gareth Morgan as "Illumination Invariant Image Registration for Autonomous UAV Navigation based on Optical Sensing".

Much of the content in Chapter 3 and Chapter 7 have been presented in *Photogrammetry Computer vision-PCV conference 2014, 5th-7th September 2014, Zurich, Switzerland*, and then published in *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* by Xue Wan, Jianguo Liu, and Hongshi Yan as "Phase Correlation based Local Illumination-

invariant Method for Multi-Temporal Remote Sensing Image Matching" 1 (2014): 365-372.

Some work in Chapter 6 has been presented in *IGARSS Conference 2015, 26th-31st July, 2015, Milan, Italy* by Xue Wan, Jian Guo Liu, Hongshi Yan and Gareth Morgan as "Automatic Mars Rover Detection Based on Multi-Temporal HiRise Imagery".

Chapter 2 Illumination-insensitive Image Matching

Illumination invariant matching algorithms have been studied for years. The aim of these studies so far is either finding unchanged patterns under different lighting conditions or eliminating the illumination-sensitive patterns. These algorithms can be broadly categorised into two approaches: feature-based and area-based.

2.1 Feature-based approaches

2.1.1 *Pre-processing*

In remote sensing image matching, radiometric correction, also named as contrast enhancement, the adjustment of image DN (Digital Number) values to achieve radiometrical normalisation serves as a pre-processing to remove the sun illumination variation between multi-temporal images [1].

An image histogram can be regarded as a probability density function which describes the global grey value distribution in the image. By adjusting the image histogram to comply a particular distribution, the radiometric normalisation can be achieved. The radiometric correction methods includes Linear Contrast Enhancement (LCE), logarithmic and exponential contrast enhancement, histogram equalization[14], histogram matching and Gaussian stretch, Balance Contrast Enhancement Technique (BCET) [15], etc. Detailed description and summary of contrast enhancement algorithms can be found in [1].

Radiometric correction is widely used in remote sensing image pre-processing for effectively remove the global brightness variation. However, most radiometric correction methods do not hold in the case of local illumination, because local sun angle change results in a locally transformation between two remote sensing images which cannot be eliminated via global image contrast enhancement.

Other radiometric correction approaches transform the images into low-dimensional linear subspaces, which are less sensitive to illumination change. Shashua and Riklin-Raviv proposed a Quotient Image (QI) [16], a ratio image depends only on surface albedo, to achieve illumination invariance. Then several improved algorithms based on QI have been proposed [17, 18]. In [19, 20], illumination cones were used to represent image under all possible illumination conditions. In [21], an illumination invariant colour space is introduced based on peak spectral responses.

These radiometric methods usually require a training dataset and are based on known lighting conditions, which is not practical in our study. Moreover, by removing or eliminating illumination variations, the quality of image inevitably declines. This is proved by the experiment as results in [22]: the accuracy of image matching may degrade if the illumination normalization algorithms are applied to force a normalized illumination condition. This quality declination is acceptable for object recognition, but can result in lost-correspondence in small window matching for pixel-wise disparity estimation within DEM generation.

2.1.2 Feature extraction algorithms

Edges in digital images are lines that change sharply or have discontinuities in intensity, and are caused by the following effects [23]:

- Change of material or material properties
- Discontinuities in elevation and orientation of surface (breaks in slope or aspect)

Corners are intersection points of two edges. Both corners and edges are distinctive features in images which can be robustly extracted. The foundation of feature-based illumination-invariant image matching is under the assumption that distinctive features maintain their positions and shape geometry under different imaging conditions.

Most corner and edge detection algorithms are based on the spatial domain, which detect the sharply changed patterns based on image brightness. Some commonly used edge detection algorithms include the Canny operator [24], zero-crossing operator [25] and Mar operator [26]. Much research has been focused on the generation of edge maps, sorts of illumination-normalized images based on features, for illumination-insensitive image matching. For instance, edge maps have been applied to face recognition under variant light conditions [27-29]. In [30], skylines are extracted from images to match with the contours of digital elevation models. In [31], an adaptive Sobel filter has been applied to generate edge images for template matching.

Currently, the most popular feature detection operator is SIFT (Scale-invariant Feature Transform) [32], proposed by Lowe. SIFT is invariant to scale, rotation and illumination intensity change, which enables robust image matching. Several improvements have been made to the SIFT operator to enhance its robustness of matching. One of the most widely accepted improved algorithms is SURF (Speed Up Robust Features) [33], which claims to be faster and more robust than SIFT. Studies using SIFT-related algorithms for face recognition under varying illumination can be found in [34, 35].

These corner and edge detection algorithms work quite well in computer vision image matching and several of them have been reported to be robust under illumination change. However, computer vision images are, in general, scenes from daily life; the majority of features in these images resulted from changes of material or material properties. These features are less sensitive to illumination change than those resulted from discontinuities in surface. If the edges are caused by a change of object material, the edge boundaries of the object remain the same though the illumination condition changes. An example is shown in Figure 2.1, which is a tested illumination variant stereo pair from Middlebury stereo datasets [36]. The edges in Figure 2.1(a) and Figure 2.1(b) are caused mainly by a change of object type, such as the dividing line between aloe and pot, and the related edge

images detected by the Canny operator shown in Figure 2.1(c) and (d) have similar appearance though under different illumination conditions.



(a) Illumination 1



(b) Illumination 2



(c) Canny edge detection of (a)



(d) Canny edge detection of (b)

Figure 2.1 Illumination test pair from Middlebury stereo datasets with the edge detection results using Canny operator

In remotely sensed images of natural landscapes, such as mountains, deserts and glaciers, edge features caused by spectral changes between different object classes are limited. The majority edges in these images are the result of shading and shadows of macro and micro topography, corresponding to terrain elevation and orientation discontinuities. The appearance of the natural landscapes in images is therefore more sensitive to illumination change. As shown in Figure 2.2 for example, two hill shading images are generated under different azimuth angles ($60^\circ, 180^\circ$) and zenith angles ($50^\circ, 35^\circ$). The edge detection results using the Canny

operator to extract the edges from Figure 2.2(a) and Figure 2.2(b) are presented in Figure 2.2(c) and Figure 2.2(d) respectively, which demonstrate that features like edges and corners are not truly invariant to illumination angle change, since the number of similar features may decrease significantly if the lighting directions between the matched images are very different. The reason is that feature-based matching algorithms depend on local grey value distribution that can be greatly altered by the change of illumination orientation such as solar position. This can be seen in Figure 2.2; vertical edges are predominant in Figure 2.2(c) while horizontal edges in Figure 2.2(d). It has been reported that both SIFT and SURF, though robust to the global intensity changes in terms of image brightness and contrast, are not truly robust under considerable changes of illumination direction (azimuth and zenith angles of lighting source) [21, 37].

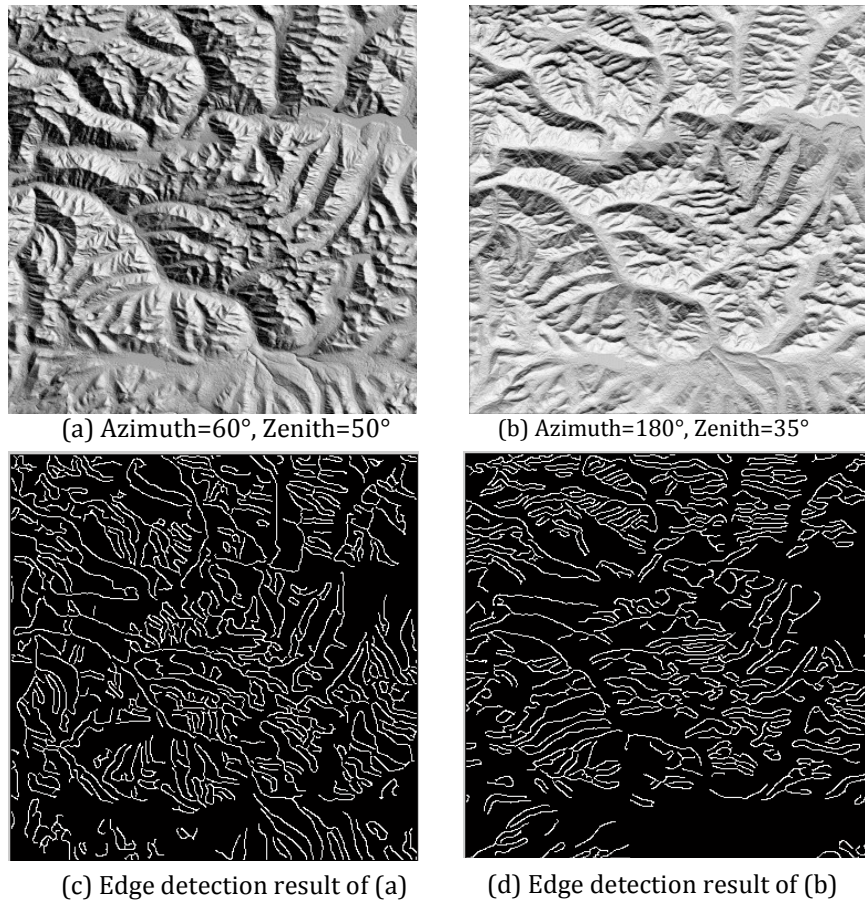


Figure 2.2 Two hill shading images under varying local illumination with the edge detection results using the Canny operator

In conclusion, feature-based approaches can tolerate the global illumination change, but are not robust to large local illumination change because as the local patterns will be altered greatly under in such circumstances.

Other approaches seek features in the frequency domain to achieve illumination normalization [38]. High frequency information, which is less sensitive to illumination variation, can then be effectively enhanced by filtering [39, 40] and thus generate a normalised image after transforming back to the spatial domain. This can also be achieved by wavelet analysis and DCT (Discrete Cosine Transform) [38], [41, 42].

The problems associated with these methods are similar to those of the spatial domain based feature matching approach. They are not truly invariant to illumination angle changes that alter the local texture depending on topographic reliefs. In terms of enhancing high frequency information such as edges and corners, the approach bears similarity to feature-based matching.

2.2 Area-based approaches

Area-based image matching approaches, known as *template matching*, are another branch of illumination-insensitive image matching approaches. Area correlation based algorithms, such as NCC (Normalised Cross Correlation), MI (Mutual Information) [43] and PC (Phase Correlation) [44], directly match two areas in the reference and target images without feature extraction.

One disadvantage of area-based approaches is that they cannot tolerate complex geometric distortions, because the similarity measurement is based on windows [5]. Thus, area-based approaches are not applicable to direct registration between two images which have large view angle difference. However, for airborne and spaceborne remote sensing, the attitude of the imaging platform is very stable with sensor view angle nadir, in most cases, or off nadir with a given oblique view angle in direction perpendicular to or along the flight direction. For small UAV imaging, although the platform attitude can be affected by wind and air turbulence, the

irregular oblique view can be rectified to the nadir view using on-board recorded orientation data. Area based matching approaches are widely applied in disparity map estimation, because after epipolar rectification, the motion between the matching windows in the stereo pair is a simple translation shift.

In contrast to feature based algorithms, area based matching algorithms are usually independent of global grey value distribution and thus are more robust to illumination variation. This is because although the the brightness and contrast of local patterns can be greatly altered by changing global illumination; the global spatial pattern of image textures remains the same.

2.2.1 Normalised Cross Correlation (NCC)

NCC is the most widely used similarity measurement for area-based methods. The basic principle of this approach is to seek the maximum similarity between a reference and a target images. Specifically, correlation coefficient values are calculated consecutively between window pairs of the reference image and the target image that roam within a given neighbourhood in the reference image; the registration is achieved at the position where correlation coefficient value is the maximum among the consecutively calculations. Several approaches were proposed to improve NCC matching performances, such as SSDA (Sequential Similarity Detection algorithm) [45] which is faster than NCC. A modified NCC based approach, ANCC (Adaptive Normalised Cross Correlation), is proposed in [46] to reduce the flattening effect and normalise different brightness factors in each colour channels caused by change of lighting conditions.

NCC has been proved to be invariant to affine transformation change in image radiometry [47, 48], because the normalisation within the window pairs compensates for the brightness difference between the reference image and the target image [36]. This means that NCC is independent of global illumination change. It is however not invariant to local illumination changes, because the image patterns within the matching window can be greatly altered by the change of solar position and thus the correlation coefficient value becomes very low.

2.2.2 Mutual Information (MI)

Another well-known area based approach is MI, an image matching algorithm based on entropy. Suppose that there are two unregistered images I_1 and I_2 , and their mutual information $MI(I_1, I_2)$ can be calculated by the equation:

$$MI(I_1, I_2) = H(I_1) + H(I_2) - H(I_1, I_2) \quad (2.1)$$

where $H(I_1)$ and $H(I_2)$ are the entropy of image I_1 and I_2 respectively, and $H(I_1, I_2)$ is the joint entropy of image I_1 and I_2 .

The MI based matching is under the assumption that the information redundancy is the maximal if the two images are perfectly aligned [49]. Then, registration can be achieved by seeking the maximum MI value via roaming, just as in NCC based matching. Mutual information is a powerful matching criterion for its property of data independence, i.e. no constraints on imaging modality and no pre-processing needed [49].

Studies in comparison of state-of-the-art algorithms for image alignment have shown the robustness of MI to image radiometric transformations [50]. This property of MI makes it a popular matching cost function for small window matching in disparity map estimation, including local stereo matching [51], global matching [52] and Semi-Global Matching (SGM) [53]. However, the high calculation complexity in MI makes it less favourable for fast real time matching and large window matching in remotely sensed image.

2.2.3 Phase Correlation (PC)

Phase Correlation is an image matching algorithm based on Fourier shift property, which states that a translation shift between two similar images generates a linear phase difference in the Fourier frequency domain [44]. This phase difference can be presented by a cross power spectrum, and then estimated directly in the frequency domain or resolved in the spatial domain via IFT (Inverse Fourier Transform).

As a none-iterative method, Phase Correlation can directly calculate the image shifts without roaming search, as performed by NCC and MI. The matching speed is thus fast enough to meet the demand of real-time image matching.

Phase spectrums reflect the structure of images and are less influenced by illumination conditions as compared to spatial based approaches. The importance of phase information in image analysis has been proved in [54]; the phase information alone is sufficient to reconstruct a signal. Experiment results [55, 56] have demonstrated the superior robustness of frequency based approaches for image classification and object recognition under illumination varying conditions. Research has shown that phase correlation matching can achieve sub-pixel accuracy in estimation of translation, rotation and scale change between matched images and it is robust to random noise [57]. Over several years, our research group has developed professional C++ software package based on PC, named PCIAS, for a variety of remote sensing applications [58] including narrow and wide baseline DEM generation, pixel-pixel precise image co-registration, change detection, camouflaged moving target detection and super resolution, etc. PC is robust to variations of image contrast and brightness, the global illumination change, but whether it is invariant to local illumination change, especially sun angle change, is yet to be investigated. In the following chapters, the robustness of PC to local illumination change will be proved and demonstrated via mathematical analysis and image matching experiments, and then it will be applied to terrain related applications.

Chapter 3 Phase Correlation Decomposition

3.1 Phase Correlation

Phase Correlation is a frequency-based image matching algorithm based on the shift property of Fourier Transform (FT) [44]. Suppose that there is a translation shift (a, b) between two identical images $I_1(x, y)$ and $I_2(x, y)$

$$I_1(x, y) = I_2(x - a, y - b) \quad (3.1)$$

then,

$$F_1(u, v) = F_2(u, v)e^{-i(au+bv)} \quad (3.2)$$

where $F_1(u, v)$ and $F_2(u, v)$ are the Fourier transforms of the two images $I_1(x, y)$ and $I_2(x, y)$.

Phase correlation is defined as the normalized cross power spectrum $Q(u, v)$ between the Fourier transforms of the two images.

$$Q(u, v) = \frac{F_1(u, v)F_2^*(u, v)}{|F_1(u, v)F_2^*(u, v)|} = e^{-i(au+bv)} \quad (3.3)$$

where $*$ stands for complex conjugate. $F_1(u, v)F_2^*(u, v)$ is a pixel to pixel multiplication.

$Q(u, v)$ is a complex 2D array with its absolute value ranging between 0 and 1. Obviously, if the two images are identical without shifts, $a = 0$ and $b = 0$, then $I_1(x, y) = I_2(x, y)$, and thus the cross power spectrum $Q(u, v) = 1$. In case there is random noise in the images for matching, the absolute value of the cross power spectrum will be lower than 1.

An example of Phase Correlation based image matching is shown in Figure 3.1. A small translation, ($a = 10, b = 10$), between the two identical images, shown in Figure 3.1(a) and Figure 3.1(b), in horizontal and vertical directions. As $Q(u, v)$ is a complex array, it is shown by phase angle map in Figure 3.1(d). The density and orientation of fringes are determined by the direction and magnitude of the shift: $\sqrt{a^2 + b^2}$ and a/b .

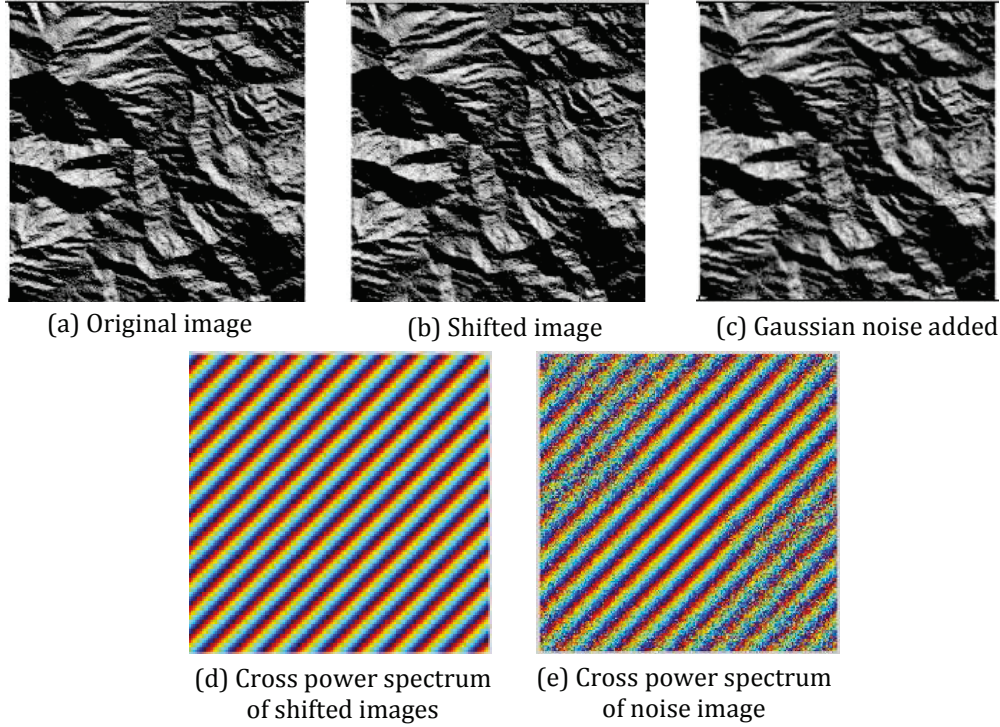


Figure 3.1 Example of phase correlation between two images with translational shift and added noise. Image (b) is from image (a) by translational shift of 10 pixels in both horizontal and vertical directions and image (c) is produced by adding Gaussian noise to (a). (d)-(e) Phase correlation cross power spectrum without and with random noise.

The shifts (a, b) can be resolved at integer level via IFT (Inverse Fourier Transform) to convert $Q(u, v)$ to a Dirac delta function $\delta_0(x, y)$ [44].

$$\text{FT}^{-1}(Q(u, v)) = \delta_0(x - a, y - b) \quad (3.4)$$

The shift can also be resolved directly in the frequency domain with sub-pixel accuracy by unwarping and rank one fitting of the fringe patterns in the cross power spectrum [12, 59-61].

By adding noise to the original image, shown in Figure 3.1(c), the interference fringes in cross power spectrum will be blurred with degraded quality, as shown in Figure 3.1(e). The density and orientation of fringes remain the same, and are only determined by the translational shift values in horizontal and vertical directions. This means that the derivation of translational shift values (a, b) using the phase correlation is robust to random noise.

The rotation and scale differences between matched images for matching can also be estimated by transferring the Fourier spectra to a log-polar plane after applying a high pass filter [57]. The peak location in the log-polar plane indicates the rotation and scale value while the peak value indicates the translation value. As an improvement, in [62], a grey level edge map was used instead of image function in log-polar Fourier Transform and then a normalized gradient correlation was introduced for greater tolerance to scale change as large as 6.

Foroosh [63] demonstrated the robustness of Phase Correlation by image matching experiment using two SPOT satellite images under different illumination conditions and he concluded that the global illumination change will only alter the magnitude (correlation quality) of PC but not change the orientation and density of PC fringes which is the key information of image shift. Thus phase correlation is robust to the change of global illumination. However, as far as we know, there is lack of rigorous study on the effects of local illumination change on Phase Correlation. The following sections involve the investigation of the robustness of Phase Correlation to local illumination variation caused by the sunlight angle change.

3.2 Phase Correlation matrix decomposition

In §3.1, the two images are supposed to be identical with a simple linear shift, however the two images are not identical if they are taken under different lighting conditions. Consider two images $I_1(x, y)$ and $I_2(x, y)$ of different lighting conditions and without shifts, their Fourier Transforms are $F_1(u, v)$ and $F_2(u, v)$ respectively. The cross power spectrum Q_1 of $F_1(u, v)$ and $F_2(u, v)$ is

$$Q_1(u, v) = \frac{F_1(u, v) F_2^*(u, v)}{|F_1(u, v) F_2^*(u, v)|} \quad (3.5)$$

Q_1 represents the texture difference between the two images resulted from different illumination, and contains no information of translational 2D motion between $I_1(x, y)$ and $I_2(x, y)$. We name Q_1 as the illumination impact matrix. If the two images are identical, then $Q_1 = 1$, and the phase angle of Q_1 , $\emptyset = 0$. The more similar the two images are, the closer the phase angle $\emptyset$ approaches to 0.

Suppose a new image $I_1'(x, y)$ is generated by shifting image $I_1(x, y)$ by (a, b) in x and y directions: $I_1'(x, y) = I_1(x - a, y - b)$. According to Equation (3.3), the cross power spectrum Q_2 between the original image $I_1(x, y)$ and shifted image $I_1'(x, y)$ can be expressed as

$$Q_2(u, v) = e^{-i(au+bv)} \quad (3.6)$$

Q_2 corresponds to the translational 2D motion of the same image, and is named as the translation matrix. This means that if the 2D motion parameters of images are known, Q_2 is fixed no matter how image texture varies as a result of illumination variation.

We now discuss the PC matrix Q_3 between $I_1'(x, y)$ and $I_2(x, y)$

$$Q_3(u, v) = \frac{F_1'(u, v) F_2^*(u, v)}{|F_1'(u, v) F_2^*(u, v)|} = \frac{F_1(u, v) e^{-i(au+bv)} F_2^*(u, v)}{|F_1(u, v) e^{-i(au+bv)} F_2^*(u, v)|} \quad (3.7)$$

Given that $F_1(u, v) e^{-i(au+bv)} F_2^*(u, v)$ is a pixel to pixel multiplication, the order of multipliers can be altered, thus

$$Q_3(u, v) = \frac{F_1(u, v) F_2^*(u, v)}{|F_1(u, v) F_2^*(u, v)|} \frac{e^{-i(au+bv)}}{|e^{-i(au+bv)}|} \quad (3.8)$$

According to Euler's formula

$$e^{-i(ua+vb)} = \cos(au + bv) - i \sin(au + bv) \quad (3.9)$$

Thus $|e^{-i(ua+vb)}| = 1$. Taking Equation (3.5) and Equation (3.6) into Equation (3.8),

$$Q_3 = \frac{F_1(u, v) F_2^*(u, v)}{|F_1(u, v) F_2^*(u, v)|} e^{-i(au+bv)} = Q_1 Q_2 \quad (3.10)$$

Equation (3.10) indicates that the PC matrix Q_3 of two images with a translational shift and under different illumination conditions can be decomposed into two independent matrices: the illumination impact matrix Q_1 corresponding to texture difference caused by illumination variation, and the translation matrix Q_2 resulting from the 2D shift between the two images. This indicates the potential illumination invariance of PC; though the illumination impact exists, the fringe density and orientation in Q_3 is dominated by Q_2 as $\sqrt{a^2 + b^2}$ and a/b .

For image matching, Q_1 and Q_2 are unknown and only Q_3 can be calculated directly from phase correlation. For illumination-insensitive image matching, the aim is to extract the translation matrix (Q_2) from the PC matrix Q_3 . The illumination invariant matching can be achieved by suppressing or eliminating the impact of Q_1 from Q_3 .

3.3 PC matrix decomposition experiments

As proved in §3.2, the PC matrix Q_3 can be decomposed into two independent matrices: the illumination impact matrix Q_1 and the translation matrix Q_2 . To validate this relationship, an example of cross power spectrum maps: Q_1 , Q_2 and Q_3 was plotted and shown in Figure 3.2.

In Figure 3.2, hill shading images (a) and (b) were generated from a DEM under different azimuth angles and Figure 3.2(b) was shifted 10 pixels to the right and downwards. Figure 3.2 (c)-(e) present the PC fringes of Q_1 , Q_2 and Q_3 . Here Q_1 was generated between Figure 3.2(a) and the same image as (b) but without shift, which corresponds to the impact of azimuth angle change; Q_2 is the PC matrix generated from two identical images of Figure 3.2(a) with the same shift as that between Figure 3.2(a) and Figure 3.2(b); and Q_3 is the PC matrix generated from Figure 3.2(a) and Figure 3.2(b) taken under different azimuth angles and with a shift. In this case, the pure shift matrix $Q_2 = e^{-i(ua+vb)}$, where $a=10$ and $b=10$. The fringe density and orientation in Q_3 shown in Figure 3.2(e) are the same as that

shown in Q_2 in Figure 3.2(d) as suggested in §3.2. Figure 3.2(f) shows the pixel-wise product of Q_1Q_2 , which is nearly identical to Q_3 in Figure 3.2(e) as proved by Equation (3.10).

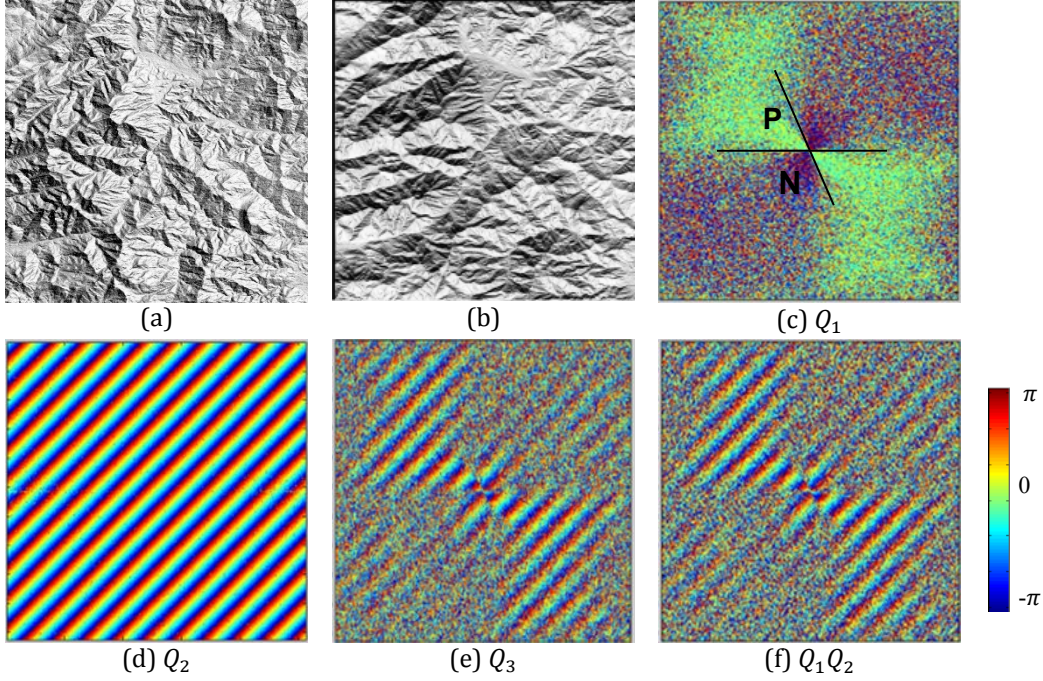


Figure 3.2 PC matrix decomposition to the illumination impact matrix and the translation matrix. (a) DEM hill shading image (azimuth=60°, zenith=45°). (b) DEM hill shading image (azimuth=180°, zenith=45°), and shifted by 10 pixels to the right and downwards. (c) Illumination impact PC matrix Q_1 . N: negative correlation; P: positive correlation. (d) Translational PC matrix Q_2 . (e) Phase correlation matrix of (a) and (c): Q_3 . (f) Pixel to pixel multiplication of (c) and (d): Q_1Q_2 .

To further assess the difference between Q_1Q_2 and Q_3 , an error estimation function E can be defined according to Equation (3.10) as

$$E = Q_3/Q_1 - Q_2 \quad (3.11)$$

Noting that the Q_1 , Q_2 and Q_3 are complex matrices which contain real parts and imaginary parts, which need to be calculated separately. The final error deviation γ is evaluated by the root mean square error of both the real parts errors and the imaginary parts errors.

$$\gamma = \sqrt{RMSE(R_E)^2 + RMSE(I_E)^2} \quad (3.12)$$

where R_E and I_E represent the real part and the imaginary part of E respectively. $RMSE$ is the root mean square error.

We tested the four datasets of different scenery and with different azimuth angle differences ranging from 0° to 300° . As shown in Figure 3.3, the four datasets include two DEM hill shading images and two images with homogeneous and directional textures. The hill shading images of mountainous regions, Figure 3.3(a) and Figure 3.3(b), are generated from ASTER GDEM (Global Digital Elevation Model), with 30 m spatial resolution and 7-14 m elevation accuracy, using ER Mapper image processing software. The texture images, Figure 3.3(c) and Figure 3.3(d), are downloaded from the photometric dataset at the website of TextureLab, Heriot-Watt University. The dataset includes images from different texture surfaces and taken under all different azimuth and zenith angles. The light source for this dataset is a fixed desk-lamp (12V, max 0.45A) and the images were taken using a Vosskuhler CCD 1300LN digital camera.

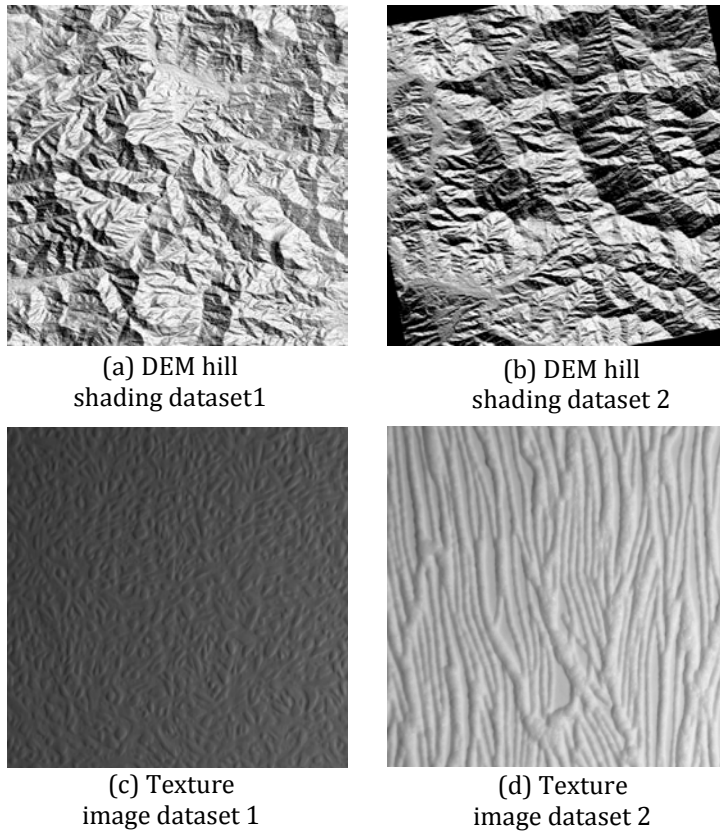


Figure 3.3 Four different testing datasets for PC matrix decomposition

For each of the four datasets, the illumination impact matrix Q_1 , the translation matrix Q_2 and the PC matrix Q_3 were derived from a matching window of 512

×512 pixels and then, the error deviation γ defined by (3.11) and (3.12) was calculated. As presented in Figure 3.4, the blue and red polylines represent γ of the hill shading image datasets and the green and purple polylines γ of the texture image datasets.

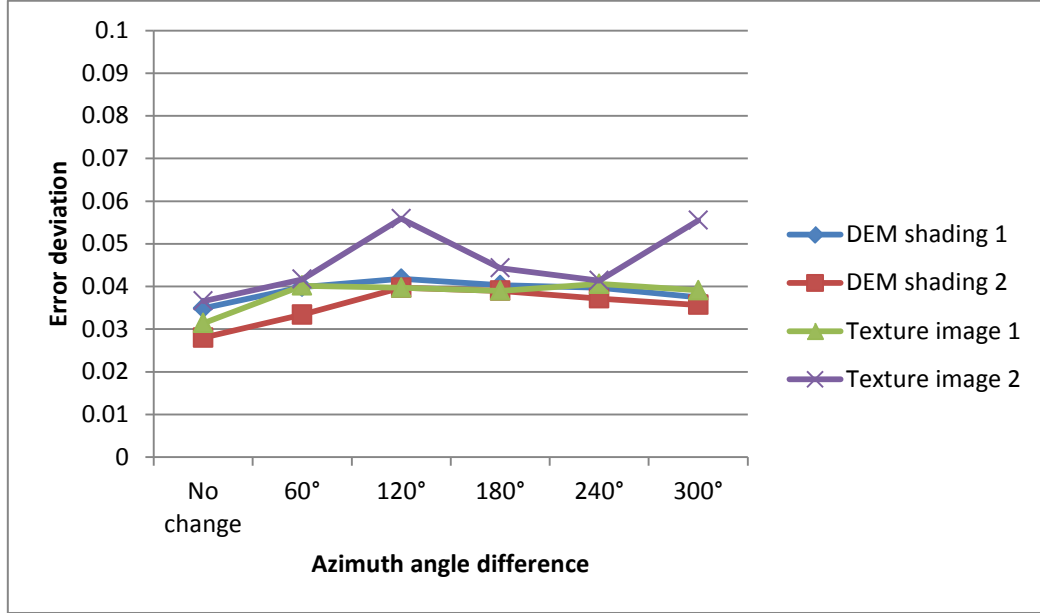


Figure 3.4 Error deviations with illumination condition variation

Figure 3.4 indicates that even without illumination change, there is still a small deviation, 0.03-0.04, between the real PC matrix Q_3 and the theoretical translation matrix Q_2 . These errors are intrinsic in PC matching, due to non-overlapped regions in the two images, aliasing, periodicity assumption and wide-band random noise [63], and independent of illumination conditions. If the error deviation values calculated with various illumination differences are very close to the initial error deviation values without illumination change, then our analytical solution of PC decomposition is proved by the experiments. The blue, red and green polylines in Figure 3.4 are pretty flat with small fluctuations in the range of [0.03, 0.04]; the largest deviation error is 0.06 occurring in the purple polyline of the texture image dataset 2, shown in Figure 3.3(d), with a dominant textural direction. Obviously for all the four datasets, γ is small enough to be negligible and Q_3/Q_1 can be regarded as the same as Q_2 , the translation matrix without illumination impact. These

experiments demonstrate that our theoretical derivation of PC matrix decomposition is correct.

The largest deviation error of texture image dataset 2 occurred when azimuth angle difference is 120° and the illumination directions are about perpendicular to the direction of the dominant texture. In this case, as shown in Figure 3.5, the image texture is almost completely reversed. The experiment indicates that the images with dominant textual orientation are more sensitive to azimuth variations especially when illumination direction is perpendicular to the dominant textures.

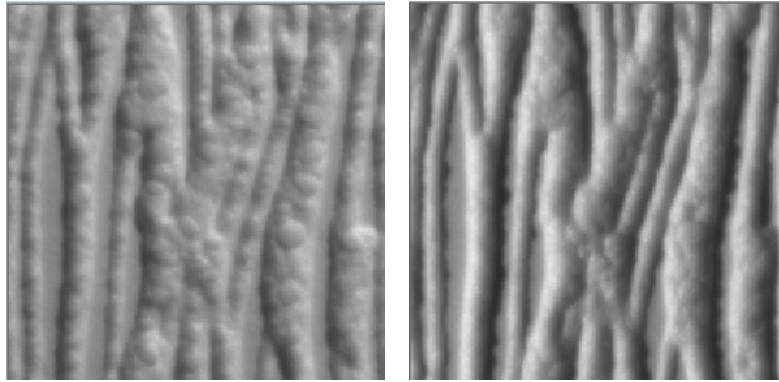


Figure 3.5 Real dataset 2 with azimuth difference of 120°

3.4 Summary

In this chapter, the impacts of illumination variation and image translation to PC matrix have been investigated. Through mathematic derivation, it has been proved that the PC matrix is a result of pixel-wise multiplication of an illumination impact matrix and a translational shift matrix. This PC decomposition relationship is further demonstrated by experiment using simulated hill shading images and texture images under different illumination variation.

The mathematical analysis and experiments in this chapter pave the way to the investigation on how Q_1 is affected by illumination conditions as presented in the next chapter.

Chapter 4 The Impact of Local Illumination Change to Phase Correlation

4.1 The *SAI* Space

For simplicity, we must first consider a terrain surface that satisfies a Lambertian model, which means that the surface is an ideal diffusely and consistently reflecting surface. Under given solar illumination L , azimuth angle τ and zenith angle σ , the intensity of hill shading image $I(x, y)$ can be expressed as [64]

$$I(x, y) = L \frac{p \cos \tau \sin \sigma + q \sin \tau \sin \sigma + \cos \sigma}{(p^2 + q^2 + 1)^{1/2}} \quad (4.1)$$

where

$$p = \frac{\partial}{\partial x} V_H(x, y)$$

$$q = \frac{\partial}{\partial y} V_H(x, y)$$

$V_H(x, y)$ is the map of height or DEM.

The Surface Normal vector is determined by slope S , the gradient angle of the terrain surface, ranging from 0° to 90° , and aspect A , the direction maximum gradient (given as a compass bearing), ranging from 0° to 360° , illustrated in Figure 4.1. The slope angle is defined as

$$\tan S = \sqrt{p^2 + q^2} \quad (4.2)$$

The aspect angle is defined as

$$\tan A = \frac{p}{q} \quad (4.3)$$

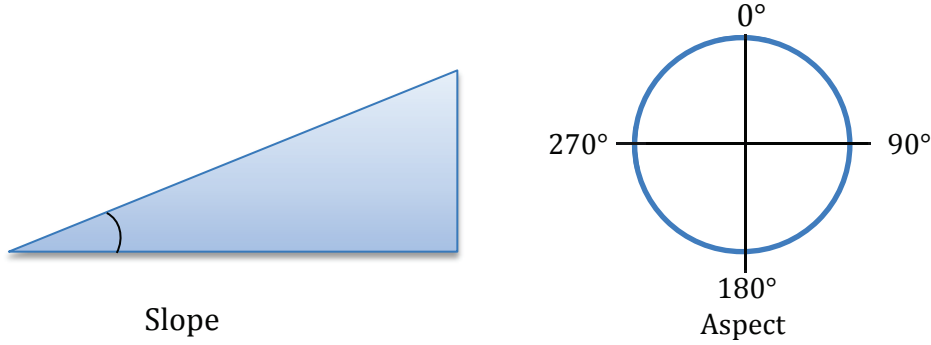


Figure 4.1 Definition of slope and aspect

Conversely,

$$p = \cos A \tan S \quad (4.4)$$

$$q = \sin A \tan S$$

If we take Equation (4.4) into Equation (4.1), the relationship between the terrain surface brightness (intensity) I and surface slope and aspect is [65]

$$I = L(\cos \sigma \cos S + \sin \sigma \sin S \cos(\tau - A)) \quad (4.5)$$

For 8 bits raster data, L is given as 255 but it can be set to 1 and then, Equation (4.5) is expressed in a normalised form without L and the outmost brackets. Compared to real illumination conditions in aerial and satellite images, this model ignores scattering sky light and overcast shadows relating to the geometry of topography as well as the spectral variation of different terrain surface materials. Most software packages for automatic generation of simulated hill shading images from DEMs are based on the Lambertian model of ideal diffusing surfaces, such as in ArcGIS, ERMapper and MATLAB. In our experiments, the Lambertian model has been applied to simulate hill shading images from DEM data under various solar illumination conditions in quantitative analysis of their effect on image matching.

According to Equation (4.5), the intensity of a hill shading image is determined by two factors, the topography (DEM), defined by slope S and aspect A , and the illumination orientation, defined by zenith σ and azimuth τ . In order to separate

the effects of topography from that of illumination, we propose a 3D space coordinates system defined by Slope, Aspect and Intensity and denote as *SAI*. The *SAI* space facilitates our investigation of the effects of illumination on hill shading images. While every pixel in a hill shading image derived from a DEM has its own intensity, slope and aspect, pixels of the same slope and aspect share the same intensity forming a 3D spatial distribution as a curved surface, named as *SAI* surface. For a given illumination geometry to the DEM, the slope and aspect are the variables while the zenith and azimuth are constants. Thus, the Equation (4.5) in normalized form can be re-written as,

$$I(S, A) = m \cos S + n \sin S \cos(\tau - A) \quad (4.6)$$

where $m = \cos\sigma$, $n = \sin\sigma$

In Equation (4.6), m , n and τ are constant coefficients, and the intensity I of each pixel is determined by the slope and aspect at the pixel's position.

An important property of *SAI* space is that the shape of the *SAI* surface is determined only by illumination conditions and is independent of the DEM. According to Equation (4.6), the intensity I is a trigonometric function of variables S and A , and the shape of *SAI* surface is uniquely determined by the coefficients m and n which are calculated from the zenith and azimuth angles. This means that under the same solar zenith and azimuth angles, the shape of the surface is invariant or unique for the whole range of slopes and aspects while for a DEM that covers only partial ranges of slopes and aspects, the theoretical *SAI* surface is partially populated but the shape of the *SAI* surface is unaltered. This property is clearly illustrated in Figure 4.2. From two different DEMs (Figure 4.2 (a)-(b)), hill shading images (Figure 4.2 (c)-(d)) under the same illumination condition of $\sigma=60^\circ$, $\tau=315^\circ$ are generated. Figure 4.2(e) shows the *SAI* space in which the blue surface represents the theoretical *SAI* surface for the given zenith and azimuth angles while the red and green dots plot the *SAI* surfaces of images Figure 4.2(c)-(d)

respectively; both the red and the green surfaces conform exactly to the blue surface in different parts of it.

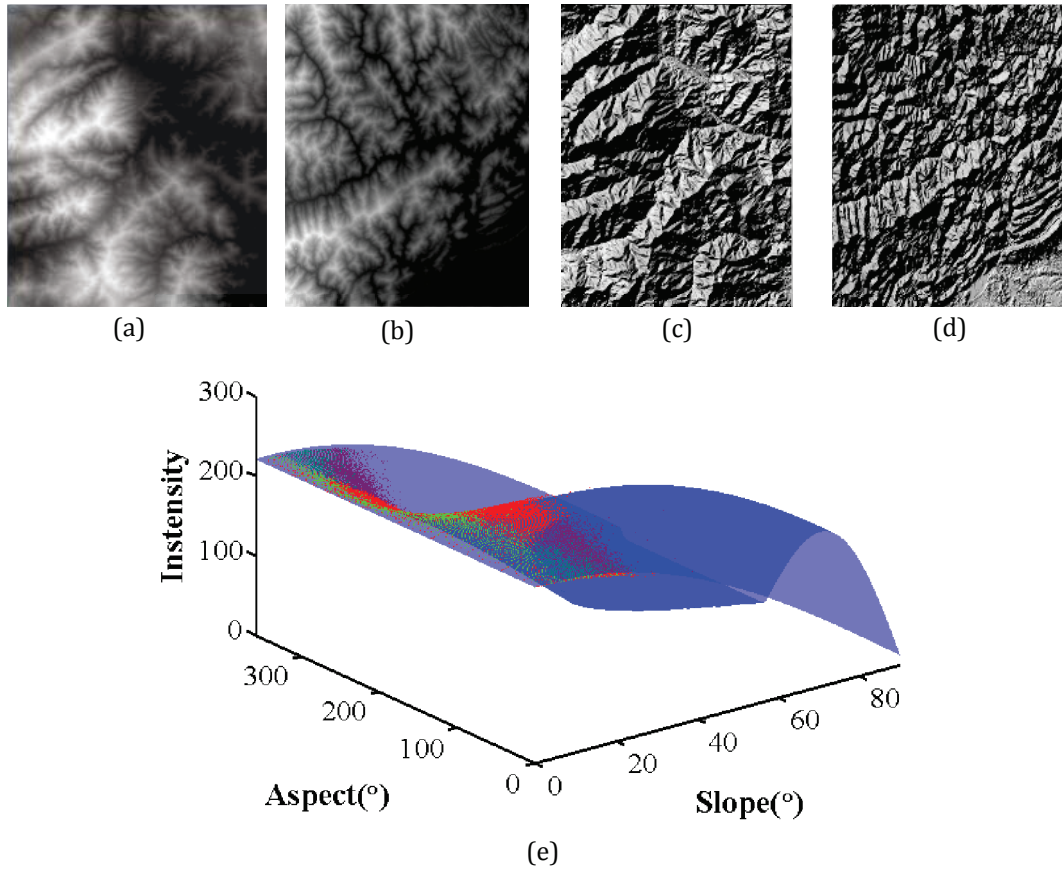


Figure 4.2 Hill shading images produced from two DEMs (a) and (b) under the same illumination condition ($\sigma=60^\circ$, $\tau=315^\circ$). (c)-(d): Hill shading image from (a) and (b) respectively. (e): The 3D view of *SAI* surface in which, the red dots are from (c) and the green from (d), and they both cover parts of the theoretical surface under the given illumination shown as the blue surface.

On the other hand, under different illumination conditions, the shape of *SAI* surfaces will be altered even when the DEM remains the same. As shown in Figure 4.3, (a) and (c) are generated from the same DEM, but under different illumination conditions: $\sigma=60^\circ$, $\tau=315^\circ$ for (a); and $\sigma=45^\circ$, $\tau=125^\circ$ for (c). The corresponding *SAI* surfaces presented in Figure 4.3(b) and Figure 4.3(d) are of different shapes as a result of the differing illumination conditions. Therefore, judging from 3D shapes of the *SAI* surfaces, it is possible to group the images under the same illumination condition.

Another key property of *SAI* space is that the shape of a *SAI* 3D surface is invariant to image translation and rotation, because such image transformation is independent of the relationship between image intensity and illumination conditions. As shown in Figure 4.4, a hill shading image under the illumination condition: $\sigma = 60^\circ, \tau = 315^\circ$ is presented in Figure 4.4(a) and its *SAI* surface in Figure 4.4(b). Figure 4.4(a) was then shifted 10 pixels to the right and downwards as shown in Figure 4.4(c) together with its *SAI* surface in Figure 4.4(d). The shape of *SAI* surface in Figure 4.4(d) remains unchanged compared to Figure 4.4(b). Furthermore, the initial image Figure 4.4(a) was rotated 90° clockwise, and the resultant image Figure 4.4 (e) and the corresponding *SAI* surface Figure 4.4(f) again show that the shape of *SAI* surface is unchanged. In summary, by introducing *SAI* space, we are able to focus on the effect of variable illumination conditions on hill shading images regardless of the DEM coverage, linear translation and rotation.

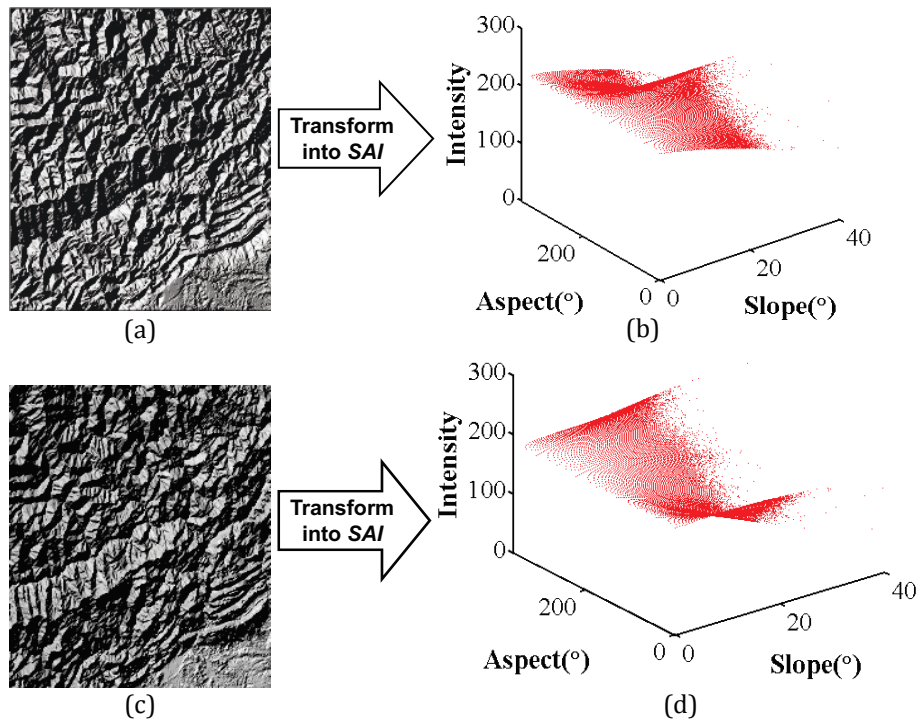


Figure 4.3 Hill shading images under different illumination conditions. (a)-(b): $\sigma=60^\circ$ and $\tau=315^\circ$, (c)-(d): $\sigma=45^\circ$ and $\tau=125^\circ$

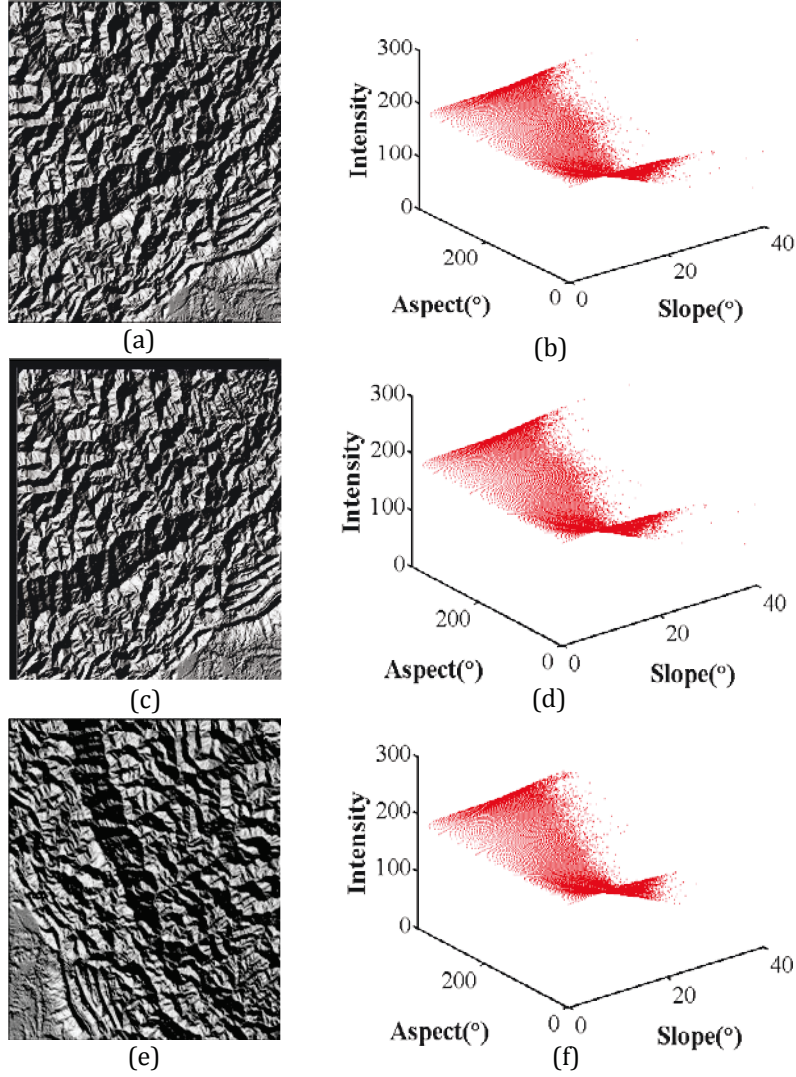


Figure 4.4 The effects of image shift and rotation on *SAI* space. (a) Original hill shading image, (b) *SAI* surface of (a), (c) 10 pixel shifted right and downwards from (a), (d) *SAI* surface of (c), (e) 90° degree rotation from (a) and (f) *SAI* surface of (e).

The shape change of the *SAI* surface indicates that changes in illumination geometry can be differentiated by the *SAI* surface shapes. The *SAI* coordinates system will be used to investigate the illumination invariant property of the PC algorithm in the next section.

4.2 Fourier transform of *SAI*

To analyse the effect of illumination angle change on phase correlation in later sections, we derive the Fourier transform of image intensity in *SAI* space.

From Equation (4.6), the image intensity of hill shading images are the sum of two terms $L_1 = m \cos S$ and $L_2 = n \sin S \cos(\tau - A)$: $I(S, A) = L_1 + L_2$.

According to the linearity property of Fourier transform,

$$FT(I) = FT(L_1) + FT(L_2) \quad (4.7)$$

The Fourier spectrums of L_1 and L_2 can be calculated separately, and then added together.

Firstly, we derive the Fourier spectrums of L_1 . Given that

$$\cos S = \frac{1}{\sqrt{1 + p^2 + q^2}} \quad (4.8)$$

Then,

$$L_1 = \frac{m}{\sqrt{1 + p^2 + q^2}} \quad (4.9)$$

Taking the Taylor series expansion of $I(x, y)$ at $p = 0, q = 0$,

$$L_1 = m \left[1 - \frac{(p^2 + q^2)}{2!} + \frac{9(p^2 + q^2)^2}{4!} \dots \right] \quad (4.10)$$

For general cases of real terrain surfaces, $p, q \ll 1$, and so the higher order terms can be ignored [64], then

$$L_1 \approx m \left[1 - \frac{(p^2 + q^2)}{2} \right] \quad (4.11)$$

Let

$$F_{H1}(\omega, \theta) = FT \left(1 - \frac{(p^2 + q^2)}{2} \right) \quad (4.12)$$

where ω is the angular frequency of the Fourier component and θ the compass direction with respect to the 0° -axis.

Thus the Fourier spectrum of L_1 can be expressed as

$$F_{L1}(\omega, \theta) = m F_{H1}(\omega, \theta) \quad (4.13)$$

We now analyse the Fourier spectrum of L_2 . According to the relationship between S , A and p , q

$$\begin{aligned}\sin S &= \frac{\sqrt{p^2 + q^2}}{\sqrt{1 + p^2 + q^2}} \\ \cos A &= \frac{q}{\sqrt{p^2 + q^2}} \\ \sin A &= \frac{p}{\sqrt{p^2 + q^2}}\end{aligned}\tag{4.14}$$

Then, L_2 becomes

$$L_2 = \frac{n(p \cos \tau + q \sin \tau)}{\sqrt{1 + p^2 + q^2}}\tag{4.15}$$

Taking the Taylor expansion and ignoring the higher order terms,

$$L_2 \approx n(p \cos \tau + q \sin \tau) \left(1 - \frac{(p^2 + q^2)}{2}\right)\tag{4.16}$$

Since p and q are partial derivatives of the height map, $V_H(x, y)$, their Fourier spectrums F_p and F_q can be expressed as

$$\begin{aligned}F_p &= \cos \theta F_{H2}(\omega, \theta) \\ F_q &= \sin \theta F_{H2}(\omega, \theta)\end{aligned}\tag{4.17}$$

where $F_{H2}(\omega, \theta)$ is the Fourier spectrum of $V_H(x, y)$.

The Fourier spectrum $F_{L_2}(\omega, \theta)$ becomes

$$F_{L_2}(\omega, \theta) = n(\cos \theta \cos \tau + \sin \theta \sin \tau) F_{H2}(\omega, \theta) * F_{H1}(\omega, \theta)\tag{4.18}$$

where $*$ denotes convolution.

Since both $F_{H2}(\omega, \theta)$ and $F_{H1}(\omega, \theta)$ are determined by topographic vector p , q , let

$$F_{H3}(\omega, \theta) = F_{H2}(\omega, \theta) * F_{H1}(\omega, \theta)\tag{4.19}$$

where $R_{H3}(\omega, \theta)$ relates only to the topography. Equation (4.19) is simplified as,

$$F_{L_2}(\omega, \theta) = n \cos(\theta - \tau) F_{H_3}(\omega, \theta) \quad (4.20)$$

Take Equation (4.20) and Equation (4.13) into Equation (4.7), The Fourier spectrum $FT(I)$ of image I becomes

$$FT(I) = m F_{H_1}(\omega, \theta) + n \cos(\theta - \tau) F_{H_3}(\omega, \theta) \quad (4.21)$$

4.3 Effects of zenith angle variation on Phase Correlation matrix

In this section, the effect of zenith angle change on the phase correlation matrix will be analysed. We consider two hill shading images, I_1 and I_2 , which are generated from a DEM with the same azimuth angle but different zenith angles, σ_1 and σ_2 . According to Equation (4.6), the SAI surfaces of I_1 and I_2 are

$$\begin{aligned} I_1(S, A) &= m_1 \cos S + n_1 \sin S \cos(\tau - A) \\ I_2(S, A) &= m_2 \cos S + n_2 \sin S \cos(\tau - A) \end{aligned} \quad (4.22)$$

where $m_1 = \cos \sigma_1$, $m_2 = \cos \sigma_2$, $n_1 = \sin \sigma_1$ and $n_2 = \sin \sigma_2$

The function variables S and A vary in the same way Equation (4.22) but the coefficients of the intensity function are different i.e. $\cos \sigma_1 \neq \cos \sigma_2$, $\sin \sigma_1 \neq \sin \sigma_2$. Thus the major effect of zenith angle change is to change the image intensity. As illustrated in Figure 4.5, the SAI surfaces of different zenith angles shift vertically along the intensity axis but there are no horizontal shifts along the aspect and slope axes; here, the red SAI surface, $\sigma_1=60^\circ$ and $\tau=150^\circ$, is above the green SAI surface, $\sigma_1=45^\circ$ and $\tau=150^\circ$. Indeed, depending on the amount of zenith angle variation the shape of SAI surface alters slightly.

We now analyse the effect of zenith angle change on PC image matching. According to Equation (4.21), the Fourier spectrums of I_1 and I_2 are

$$\begin{aligned} F_{I_1}(\omega, \theta) &= m_1 F_{H_1}(\omega, \theta) + n_1 \cos(\theta - \tau) F_{H_3}(\omega, \theta) \\ F_{I_2}(\omega, \theta) &= m_2 F_{H_1}(\omega, \theta) + n_2 \cos(\theta - \tau) F_{H_3}(\omega, \theta) \end{aligned} \quad (4.23)$$

Equation (4.23) indicates that the change of zenith does not affect the phase information of the F_I , but changes its magnitude.

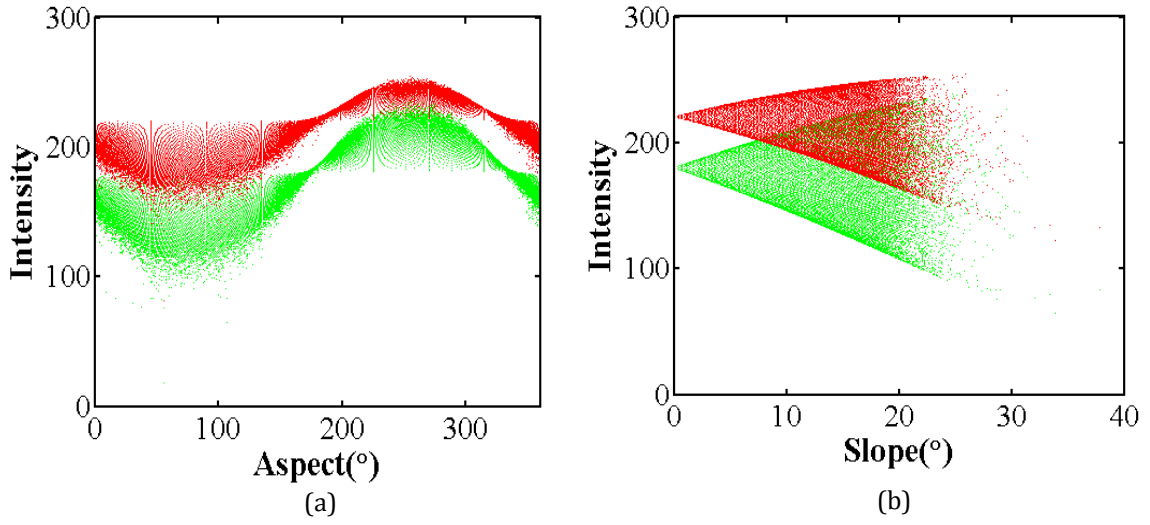


Figure 4.5 3D shapes in *SAI* space of images in aspect view, (a), and in slope view, (b), with the same azimuth angle (150°) but different zenith angles of 60° (red dots) and 45° (green dots).

The phase correlation $Q(\omega, \theta)$, the normalized cross power spectrum, of Equation (4.23) is

$$Q(\omega, \theta) = \frac{F_{I_1}(\omega, \theta)F_{I_2}^*(\omega, \theta)}{|F_{I_1}(\omega, \theta)F_{I_2}^*(\omega, \theta)|} \quad (4.24)$$

The image PC fringe patterns in *SAI* space will be affected by $Q(\omega, \varphi)$ in terms of phase angle $\phi(\omega, \theta)$ defined as

$$\phi(\omega, \theta) = \arctan\left(\frac{Q_I(\omega, \theta)}{Q_R(\omega, \theta)}\right) \quad (4.25)$$

where $Q_R(\omega, \theta)$ and $Q_I(\omega, \theta)$ are the real part and the imaginary part of $Q(\omega, \theta)$ respectively.

Take Equation (4.23) into Equation (4.24), then

$$Q(\omega, \theta) = \frac{M_{R2} + M_C}{M_{R1}} \quad (4.26)$$

where the denominator of $Q(\omega, \theta)$ is a real matrix: $M_{R1} = |F_{I_1}(\omega, \theta)F_{I_2}^*(\omega, \theta)|$ and the numerator of $Q(\omega, \theta)$ is a summation of a real matrix,

$$M_{R2} = m_1 m_2 |F_{H1}(\omega, \theta)|^2 + n_1 n_2 \cos^2(\theta - \tau) |F_{H1}(\omega, \theta)|^2 \quad (4.27)$$

and a complex matrix,

$$M_C = m_1 n_2 G(\omega, \theta) + m_2 n_1 G^*(\omega, \theta) \quad (4.28)$$

where $G(\omega, \theta) = \cos(\theta - \tau) F_{H1}(\omega, \theta) F_{H3}^*(\omega, \theta)$.

We can then decompose the $Q(\omega, \theta)$ defined in Equation (4.24) into a real matrix $Q_a(\omega, \theta)$, and a complex matrix $Q_b(\omega, \theta)$ that comprises both real parts and imaginary parts.

$$Q(\omega, \theta) = Q_a(\omega, \theta) + Q_b(\omega, \theta) \quad (4.29)$$

We then have

$$Q_a(\omega, \theta) = \frac{M_{R2}}{M_{R1}} \quad (4.30)$$

$$Q_b(\omega, \theta) = \frac{M_C}{M_{R1}} \quad (4.31)$$

For the reason that only $Q_b(\omega, \theta)$ contains imaginary parts, the imaginary parts of $Q(\omega, \theta)$ can be expressed as

$$Q_I(\omega, \theta) = I(Q_b(\omega, \theta)) = I\left(\frac{M_C}{M_{R1}}\right) \quad (4.32)$$

where I represents the imaginary part operator.

If there is no zenith angle difference,

$$\sin(\sigma_2 - \sigma_1) = 0 \quad (4.33)$$

then,

$$m_1 n_2 = m_2 n_1 \quad (4.34)$$

Matrix M_C becomes a real matrix, and $Q_b(\omega, \theta) = 0$. Then there will be no zenith angle impact on the image phase correlation matrix.

If $\sigma_2 - \sigma_1 \neq 0$, the value of $Q_I(\omega, \theta)$ will increase with increasing difference between the zenith angles σ_2 and σ_1 , according to Equation (4.28) and Equation (4.32). Because $Q_I(\omega, \theta)^2 + Q_R(\omega, \theta)^2 = 1$ according to the definition of phase

correlation, the value $Q_R(\omega, \theta)$ will decrease with the increase of $Q_I(\omega, \theta)$ and thus the phase angle $\emptyset(\omega, \theta)$ defined by Equation (4.25) will increase. The increasing $\emptyset(\omega, \theta)$ means that the effect of zenith change on the phase correlation matrix becomes more significant. This effect, however, does not introduce a modulated periodical phase variation because, according to Equation (4.23). The change of zenith does not introduce consistent phase shifts but only results in a magnitude change to the Fourier spectrum. This magnitude change is essentially controlled by $\emptyset(\omega, \theta)$ which varies with the local slope S and aspect A of topography and it is therefore irregular because terrain topography is irregular. The effect is similar to adding random noise to the cross power spectrum of image phase correlation. Consequently, the principal trend of fringe orientation and the density of the image phase correlation matrix are not altered. As long as the ‘noise’ level is not high enough to overwhelm the fringes of the cross power spectrum, the PC algorithm remains effective for accurate image matching.

4.4 Effects of azimuth angle variation on Phase Correlation matrix

In this section, the effects of azimuth angle variation on phase correlation matrix are investigated. According to §4.3, a changing zenith angle will alter the magnitude of Fourier spectrum of the hill shading image, and thus the following discussion of azimuth angle variation is based on different cases of zenith angle σ .

Case 1: small zenith angles

When zenith angles are very small ($<6^\circ$), the value of L_2 will approach 0 because $\sin \sigma$ is close to 0. Thus $I(S, A)$ can be approximated as

$$I(S, A) = m \cos S \quad (4.35)$$

As L_1 is dominant in this case, according to Equation (4.13), the SAI Fourier spectrum $F(\omega, \theta)$ becomes

$$F(\omega, \theta) = mF_{H1}(\omega, \theta) \quad (4.36)$$

Equation (4.36) indicates that the Fourier spectrum is only a function of zenith angle σ , given that $m = \cos\sigma$. This means that the distribution of image intensity values is not sensitive to sun azimuth angle changes when zenith angle is small. When zenith angle is close to 0, the two images with different azimuth angles can be regarded as nearly identical, and the cross power spectrum of illumination impact matrix Q_1 is close to 0. For $\sigma = 0$, there is therefore no definition of azimuth and nor its effect.

Thus in the cases of small zenith angle σ , the phase correlation between two images with different azimuth angles has a high correlation confidence. The example shown in Figure 4.6 illustrates that in the case of very small zenith angles (5°), despite very different azimuth angles ($60^\circ, 120^\circ$), the related *SAI* surfaces plotted in Figure 4.6(d) are largely similar and so are the intensity distributions of the two images. The PC cross power spectrum of Figure 4.6(a) and Figure 4.6 (b) is presented in Figure 4.6(c). As there is no image shift between the two images, the PC cross power spectrum Q_3 in Figure 4.6(c) closely approximates the illumination impact matrix Q_1 , in which most values are 0 though there are some non-zero noise values caused by minor illumination differences.

Case 2: large zenith angles

With an increase in zenith angles, $\cos\sigma$ becomes small, the effect of L_1 becomes negligible while that of L_2 becomes dominant. We therefore can ignore the L_1 term in the case of large zenith angles, so the expression of $I(S, A)$ can be simplified as,

$$I(S, A) = n \sin S \cos(\tau - A) \quad (4.37)$$

Consider that two hill shading images are generated from a DEM with the same zenith angle but different azimuth angles, τ_1 and τ_2 , then, according to Equation (4.37), in the same notation:

$$\begin{aligned} I_1(S, A) &= n \sin S \cos(\tau_1 - A) \\ I_2(S, A) &= n \sin S \cos(\tau_2 - A) \end{aligned} \quad (4.38)$$

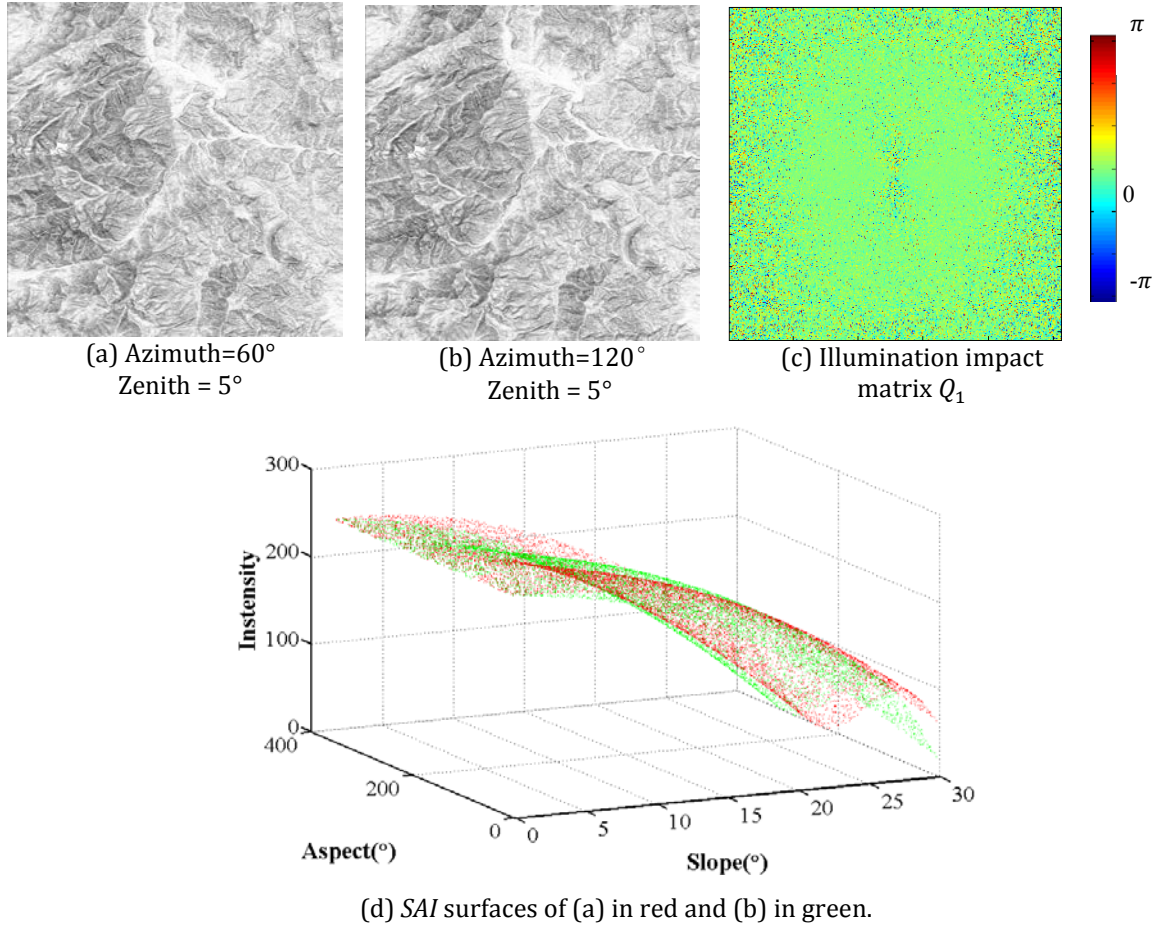


Figure 4.6 Illumination impact matrix and SAI surfaces between two hill shading images with small zenith angles but very different azimuth angles.

Equation (4.38) implies that the change in illumination azimuth angle introduces a translational shift, $d = \tau_1 - \tau_2$, between the SAI surfaces of the two images along the aspect axis only. An example is shown in Figure 4.7(a), there is $150^\circ - 125^\circ = 25^\circ$ shift along the aspect axis between the red surface ($\sigma = 70^\circ, \tau_1 = 125^\circ$) and the green surface ($\sigma = 70^\circ, \tau_2 = 150^\circ$). In the slope view, the surfaces are of the same shape without shifts and thus fully overlapped (Figure 4.7(b)). The azimuth angle variation results in truncation at one side and extension to the other side of the surface in aspect direction but no effect in slope direction. If we shift the red surface of τ_1 to the left in order to match the green surface of τ_2 , the two surfaces are identical in the shared parts within the aspect range $[d, 2\pi - d]$ and this is the major correlation between the two images. As SAI surfaces are periodic, the

remaining small parts of the red and green surfaces can also be correlated, which is the minor correlation.

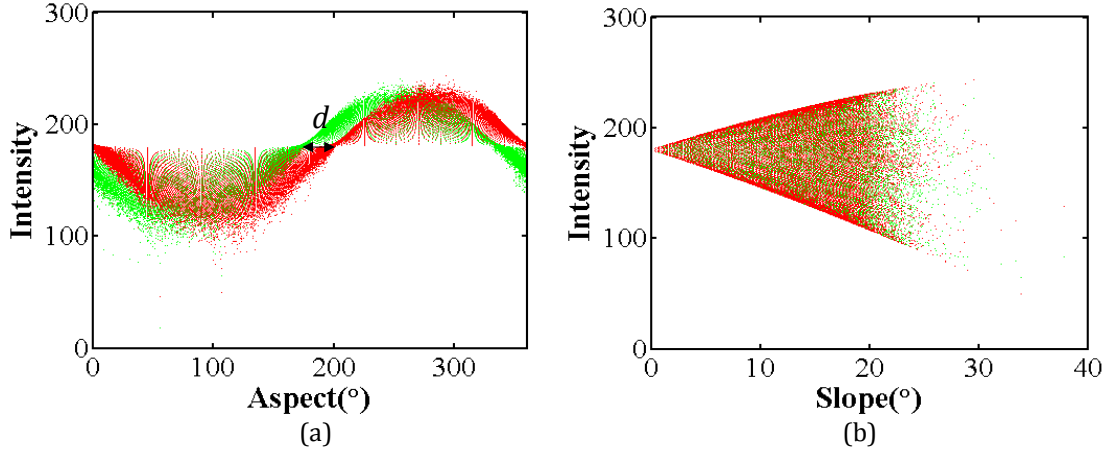


Figure 4.7 3D shapes in SAI surfaces of images in aspect view, (a), and slope view, (b), with the same zenith angle ($\sigma = 70^\circ$) but different azimuth angles of $\tau_1 = 125^\circ$ (red dot) and $\tau_2 = 150^\circ$ (green dot).

According to Equation (4.20), the Fourier spectrum of image I_1 and I_2 can be expressed as

$$\begin{aligned} F_{I_1}(\omega, \theta) &= n \cos(\theta - \tau_1) F_{H3}(\omega, \theta) \\ F_{I_2}(\omega, \theta) &= n \cos(\theta - \tau_2) F_{H3}(\omega, \theta) \end{aligned} \quad (4.39)$$

The illumination impact phase correlation matrix $Q_1(\omega, \theta)$ of $F_{I_1}(\omega, \theta)$ and $F_{I_2}(\omega, \theta)$ is

$$Q_1(\omega, \theta) = \frac{F_H(\omega, \theta) F_{H3}^*(\omega, \theta) n^2 \cos(\theta - \tau_1) \cos(\theta - \tau_2)}{|F_H(\omega, \theta) F_{H3}^*(\omega, \theta) n^2 \cos(\theta - \tau_1) \cos(\theta - \tau_2)|} \quad (4.40)$$

Since $R_{H3}(\omega, \theta) R_{H3}^*(\omega, \theta) = 1$, Equation (4.40) becomes,

$$Q_1(\omega, \theta) = \frac{\cos(\theta - \tau_1) \cos(\theta - \tau_2)}{|\cos(\theta - \tau_1) \cos(\theta - \tau_2)|} \quad (4.41)$$

$Q_1(\omega, \theta)$ can be either positive or negative, as determined by $\cos(\theta - \tau_1) \cos(\theta - \tau_2)$ according to the equations below:

$$Q_1(\omega, \theta) = \begin{cases} 1, & \tau_2 - \frac{3\pi}{2} < \theta < \tau_1 - \frac{\pi}{2}, \tau_2 - \frac{\pi}{2} < \theta < \tau_1 + \frac{\pi}{2} \\ -1, & \text{else} \end{cases} \quad (4.42)$$

This equation shows that the effect of azimuth angle variation is to introduce negative correlation in the illumination matrix $Q_1(\omega, \theta)$ with respect to the ranges of azimuth angles τ_1 and τ_2 . If the direction θ_i of a point in $Q_1(\omega, \theta)$ falls into the ranges of $[\tau_2 - \frac{3\pi}{2}, \tau_1 - \frac{\pi}{2}]$ or $[\tau_2 - \frac{\pi}{2}, \tau_1 + \frac{\pi}{2}]$, this point will be positively correlated. This is in the major correlation part in Figure 4.7. Otherwise, the point will be negatively correlated, which is the minor correlation in Figure 4.7.

Knowing the illumination impact matrix $Q_1(\omega, \theta)$ and the translational matrix $Q_2(\omega, \theta) = e^{-i(au+bv)}$, we can then work out the PC matrix of two images with a translational shift and under different illumination conditions by taking Equation (4.42) into Equation (3.10).

$$Q_3(\omega, \theta) = \begin{cases} e^{-i(au+bv)}, & \tau_2 - \frac{3\pi}{2} < \theta < \tau_1 - \frac{\pi}{2}, \tau_2 - \frac{\pi}{2} < \theta < \tau_1 + \frac{\pi}{2} \\ -e^{-i(au+bv)}, & \text{else} \end{cases} \quad (4.43)$$

Equation (4.43) shows that although reversed fringes are introduced by azimuth angle variation, the fringe density $\sqrt{a^2 + b^2}$ and fringe orientation b/a remain unchanged.

To validate this theory, an example is shown in Figure 4.8 where two images are taken under different azimuth $\tau_1 = 60^\circ$ and $\tau_2 = 180^\circ$ with the same zenith $\sigma = 65^\circ$. Figure 4.8(a) is a plot of $\cos(\theta - \tau_1)\cos(\theta - \tau_2)$ in which the critical angles defining positive and negative ranges are $-90^\circ, -30^\circ, 90^\circ$ and 150° in the range of $[-\pi, \pi]$. As shown in Figure 4.8(b), a cross power spectrum $Q_3(\omega, \theta)$ is generated from these two images, here $Q_1(\omega, \theta) = Q_3(\omega, \theta)$ as $Q_2(\omega, \theta) = 1$ when there is no translational shift between the two images. The positive correlation, $Q_1(\omega, \theta) = 1$, and the negative correlation $Q_1(\omega, \theta) = -1$ are presented by a 2π wrapped cross power spectrum in ranges of $[-\pi, \pi]$. The positive correlation sectors are labelled as 'P', the negative correlation sectors are labelled as 'N' and the 0° axis denotes 0° azimuth pointing northward in Figure 4.8(b). The compass angle value θ in Figure 4.8(b) is positive clockwise from the 0° -axis. Thus, the angular ranges of positive correlation are $[90^\circ, 150^\circ]$ and $[-90^\circ, -30^\circ]$, which exactly match the critical angles

of positive and negative ranges in Figure 4.8(a) based on the theoretical model of the effect of azimuth angle variation on Phase Correlation matrix.

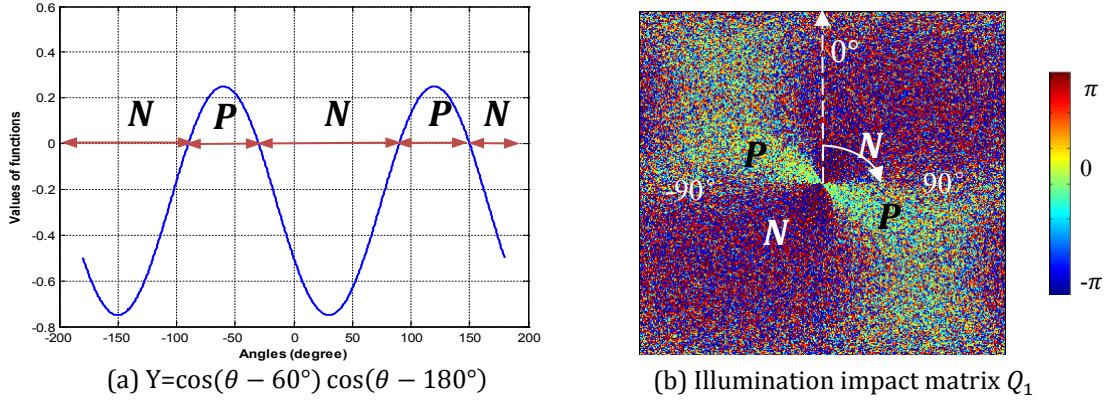


Figure 4.8 The theoretical plot and the real illumination impact matrix between two hill shading images with different azimuth angles, 60° and 180° , and the same zenith angle 65° .

A special case in azimuth angle variation is when the angular difference equals 180° (Figure 4.9), for example, 60° shown in the red curve and 240° the green curve. Note that the two curves are reversed at every position along the angle axis, the multiplication of the two curves, $\cos(\theta - \tau_1) \cos(\theta - \tau_2)$, is constantly negative. This means that $Q_1(\omega, \theta) = -1$ and that there is only negative correlation in this situation.

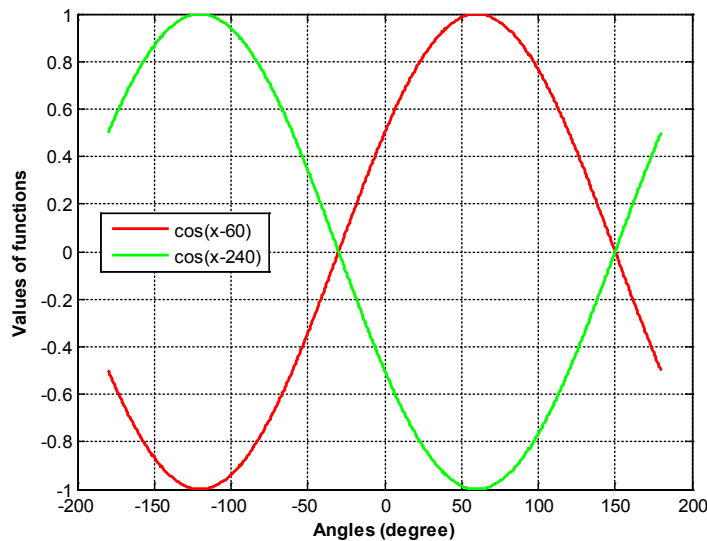


Figure 4.9 An example of the effect of 180° azimuth angle difference, 60° and 240° , on illumination impact matrix Q_I .

Case 3: General cases

In the two cases above, we analysed the effect of L_1 and L_2 on the illumination matrix $Q_1(\omega, \theta)$, separately, by setting the zenith angle σ to two extremes and this produces two different correlation patterns

$$Q_1 = \begin{cases} P, & \sigma \rightarrow 0^\circ \\ NP, & \sigma \rightarrow 90^\circ \end{cases} \quad (4.44)$$

Here P is the positive correlation in the PC matrix while NP is the combination of negative and positive correlation in the PC matrix.

In the case of small zenith angles, the term L_1 is dominant in $I(x, y)$, which results in the illumination impact matrix $Q_1(\omega, \theta) \approx 1$. This means that PC matching is almost unaffected by azimuth angle variation. In this case, there is only positive correlation in the PC matrix. In the case of large zenith angles, the term L_2 is dominant in $I(x, y)$. Depending on the azimuth angle difference, the illumination impact matrix $Q_1(\omega, \theta)$ equals either 1 or -1, which indicates negative and positive correlation, in different angle sectors. According to $Q_1(\omega, \theta)$, there is a combination of positive and negative correlation in $Q_1(u, v)$, but this has no effect on the translation PC matrix $Q_2(u, v)$ and thus the fringe density and orientation relating translation remains the same in the PC matrix $Q_3(u, v)$. Phase Correlation is therefore robust to illumination variation for image matching.

A more general case is that zenith angle σ is neither too small nor too large, when there is a combination of P and NP correlation co-existing in the illumination impact matrix Q_1 and with different values of $L_1(\sigma)$ and $L_2(\sigma)$.

$$Q_1 = L_1(\sigma)P + L_2(\sigma)NP \quad (4.45)$$

According to Equation (4.45), the strength of NP and P in the illumination impact matrix Q_1 is determined by the values of L_1 and L_2 , which are functions of zenith angle σ . The relationship between L_1 , L_2 and σ is shown by Figure 4.10. When σ is small, as discussed in §4.4 *Case 1* the value of L_2 is close to 0, and there is only positive correlation in PC matching because L_1 is dominant and image intensity is

insensitive to azimuth angle changes. With increasing zenith angle σ , the values of L_2 increase and those of L_1 gradually decline toward zero. Thus, the strength of negative correlation increases with the increase of NP correlation of the L_2 term and the decrease of P correlation of the L_1 term. When the zenith angle approaches 90° , the value of the L_1 is close to zero, and there is only NP correlation in Q_1 . Judging from the whole range of zenith angle variation, the positive correlation caused by L_1 is dominant in most cases, and only when the zenith angle is very large, will L_2 be dominant. This explains why the positive correlation is usually the major correlation while negative correlation is often the minor correlation.

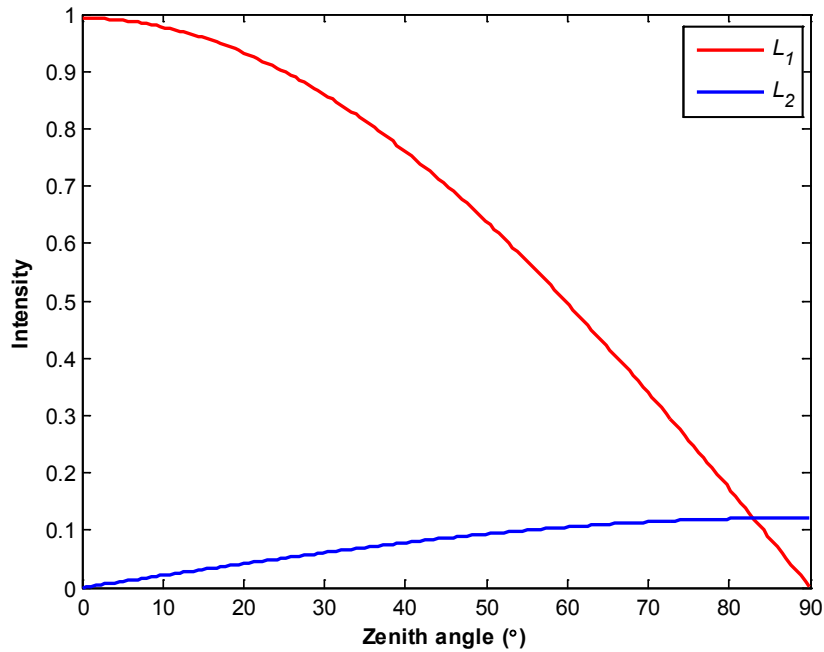


Figure 4.10 L_1 term and L_2 term variation with the increase of zenith angles ($S=7^\circ, \tau=60^\circ$)

The above discussion reveals a joint effect of zenith and azimuth angular variation on the PC matrix. The angular sector of NP is determined by azimuth angle difference, according to Equation (4.43). The strength of NP , however, is enhanced by increasing zenith angle. This means that the zenith angle serves as a weighting factor in the illumination impact matrix; the larger the zenith angle is, the more obvious the negative correlation will appear. In other words, zenith angle variation

modulates the effect of azimuth angle variation by changing the strength of negative correlation.

4.5 Effects of reflectance rate variation on Phase Correlation matrix

The previous analysis of illumination variation on Phase Correlation is under the assumption of Lambertian surface model. However, a terrain surface in the real world is only an approximation of a Lambertian surface. An optical image is taken in a particular spectral range and the image intensity of any ground position is decided not only by the solar irradiance E on the terrain surface but also by the spectral reflectance of the terrain surface ρ , i.e. $L = \rho E$ in Equation (4.1). In this case, for a given slope and aspect in the *SAI* space, there can be scattered points corresponding to the intensities of the same slope and aspect but different reflectance values. This will make the *SAI* surface 'cloudy', as shown in Figure 4.11(b), which is much less smooth than that shown in Figure 4.11 (a). In fact, such a 'cloudy' *SAI* surface is composed of several parallel *SAI* surfaces of the same shape but of differing intensity levels as decided by reflectance ρ .

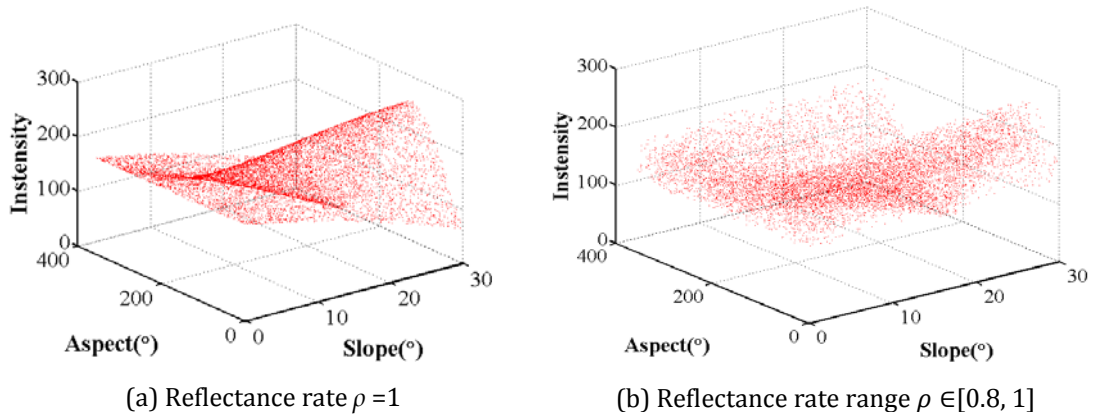


Figure 4.11 (a) *SAI* surface of constant reflectance; (b) *SAI* surface of varying reflectance in range of 0.8-1.

Then, we consider the situation of two *SAI* surfaces containing spectral variation under two different illumination conditions. As the two *SAI* surfaces are generated from the same DEM with the same spectral reflectance distribution, the cloud shape of the two *SAI* surfaces are very similar. Thus, the two 'cloudy' *SAI* surfaces

overlap in the same way as the case of a Lambertian surface. For comparison, Figure 4.12 shows two *SAI* surfaces with spectral variation and under the same illumination conditions, as in Figure 4.7. Although the *SAI* 3D shapes become thick and ‘cloudy’, the effects of spectral variation are identical in each of the two *SAI* surfaces in Figure 4.12. The shift between the two *SAI* surfaces still equals to $d = \tau_1 - \tau_2$, the same as in Figure 4.7.

In conclusion, we can consider an image with spectral variation as an assembly of many spectrally homogenous image patches of different reflectance levels. A spectrally homogenous image patch means all the pixels within the patch have identical or very similar reflectance. Thus, our analysis based on the single *SAI* surface of an ideal Lambertian surface is valid for multiple *SAI* surfaces of images with spectral variation.

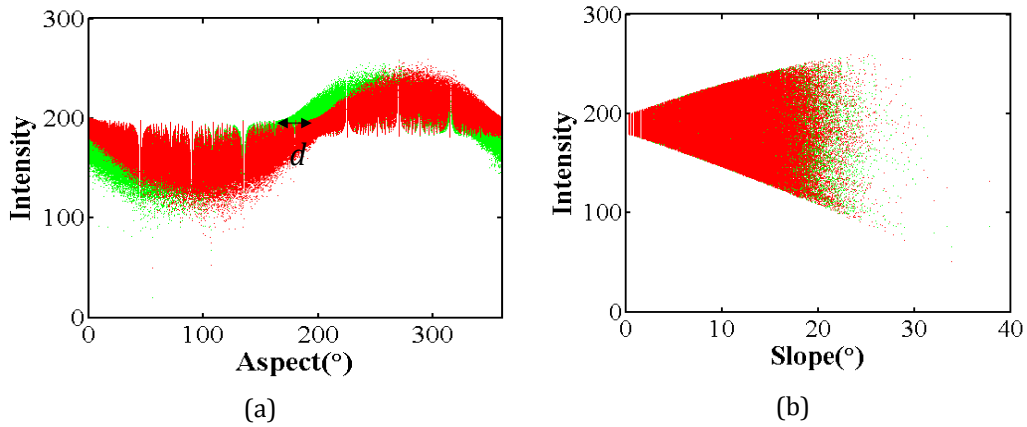


Figure 4.12 3D shapes in *SAI* space of images in aspect view, (a), and in slope view, (b), with spectral variation taken under the same zenith angle (70°) but different azimuth angles of 125° (red dots) and 150° (green dots)

4.6 Illumination-insensitive PC matching

4.6.1 Sub-pixel PC matching

Several approaches have been proposed to achieve sub-pixel PC matching, and they fall into two categories: *frequency domain based PC* and *IFT spatial domain based PC*.

As indicated in §3.1, the integer translation shift between two images can be estimated by the peak location of Dirac delta function δ_0 that is the IFT of the cross

power spectrum. When a sub-pixel image shift exists, the Dirac delta function δ_0 becomes a Dirichlet function δ_1 , a down-sampled Dirac function [60].

In the first category, sub-pixel location is determined by fitting the peak of the Dirichlet function δ_1 in the spatial domain by IFT [66]. Proved by Foroosh [60], Dirichlet function is an approximation of a *Sinc* function, but in practice, a Gaussian function [66] is often adequate and efficient to fit the peaks. As the Gaussian function is non-negative, the absolute value of Dirichlet function is used for the peak fitting. A Gaussian function is applied to the points around the peak at a matching window position, defined as

$$G_{cfx} = \lambda_1 e^{-\frac{(x-a)^2}{2\lambda_2^2}} + \lambda_3, \quad G_{cfy} = \lambda_4 e^{-\frac{(y-b)^2}{2\lambda_5^2}} + \lambda_6 \quad (4.46)$$

Where $\lambda_1, \lambda_2, \lambda_3$ and $\lambda_4, \lambda_5, \lambda_6$ are the coefficients to be determined by the selected peak points.

This ADCF-PC (Absolute Dirichlet function Curve Fitting-Phase Correlation) method is proved to be robust to white Gaussian noise [66]. Whether it is robust to illumination variation will be discussed in the following.

In the second category, sub-pixel shifts are directly solved in the frequency domain by PC fringe analysis. The translation matrix of PC, Q_2 , is a rank one matrix, and can be decomposed as the product of two dominant singular vectors $Q_x(a)$ and $Q_y(b)$.

$$Q_2 = e^{-i(ua+vb)} = e^{-iua}e^{-ivb} = Q_x(a)Q_y(b) \quad (4.47)$$

The sub-pixel displacement can then be estimated directly in the frequency domain using either Singular Value Decomposition (SVD) and then Least Square Fitting (LSF) [59], or 2D fitting technique with Quick Maximum Density Power Estimator (QMDPE) [12, 67]. The frequency-based methods can reach very high sub-pixel accuracy (1/20-1/50 pixels), as the sub-pixel displacement is directly solved via unwrapped PC fringes. The robustness of these approaches degrades

when illumination angles are significantly different, and thus an improved PC-SVD algorithm is proposed as described later.

4.6.2 Robustness of Absolute Dirichlet function Curve Fitting Phase Correlation (ADCF-PC) to illumination variation

According to §4.4, the illumination variation yields a combination of positive and negative correlation (*NP* phenomenon) in the PC matrix Q_3 , as shown in Figure 4.13(a) and in Figure 4.13(b), the corresponding δ_1 function calculated by IFT has both positive and negative peaks. According to Equation (4.43), the positive and negative fringe patterns in Figure 4.13(a) can be expressed as $Q_3 = e^{-i(ua+vb)}$ and $Q'_3 = -e^{-i(ua+vb)}$. Taking the absolute value of the Dirichlet function $\delta_1(x, y)$, the negative peak is merged with the positive peak as shown in Figure 4.13(c).

$$D_{abs}(x, y) = |\delta_1(x, y)| \quad (4.48)$$

To prove that the illumination variation does not alter the orientation and density of the fringes, a new cross power spectrum Q_{abs} is then generated by transforming the $D_{abs}(x, y)$ back to the frequency domain by FT, as shown in Figure 4.13 (d).

$$Q_{abs} = FT(D_{abs}(x, y)) \quad (4.49)$$

Figure 4.13(d) demonstrates that only one fringe patterns exists after the negative fringes are reversed by taking the absolute value of $\delta_1(x, y)$. Only when the orientation and density of the positive and negative fringe patterns are the same, can the above operation enables to convert the negative fringes to match the positive fringes.

From the above analysis, the negative correlation, introduced by illumination variation, can be corrected by using the absolute value of Dirichlet function, and thus the ADCF-PC is proved to be robust to illumination variation. The illumination invariant property of the ADCF-PC will be further demonstrated and illustrated through image matching experiment results. It should be noted that the initial purpose of using the absolute value of Dirichlet function was to provide a better

fitting to a Gaussian function, not for illumination-invariance. On the basis of this study of illumination effects on the PC matrix, this approach is proven to be illumination-invariant.

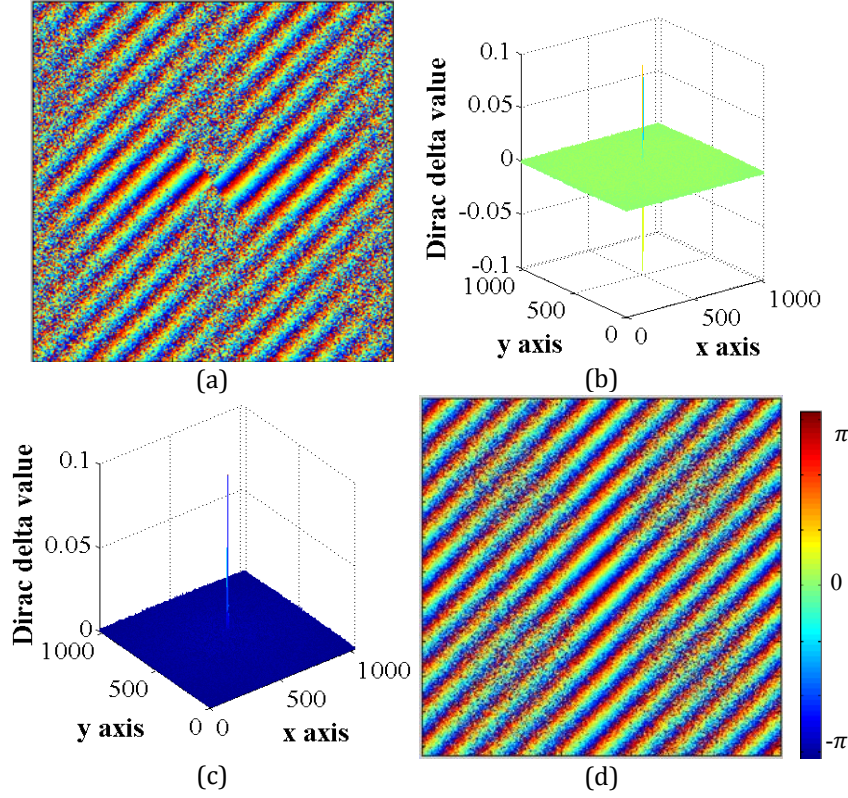


Figure 4.13 ADCF-PC algorithm robustness to illumination variation: (a) PC fringes between images of different azimuth angles (60° and 120°), (b) δ_I function derived from IFT of (a), (c) The absolute value of δ_I in (b) and (d) The power spectrum fringes of (c) via FT.

4.6.3 Illumination-insensitive Piecewise Least Square Fitting Phase Correction (PLSF-PC) for sub-pixel image alignment

As indicated earlier, the frequency-based PC can achieve higher matching accuracy than the spatial domain based approach. The SVD of Q'_3 will generate two extra singular vectors $Q'_x(a)$ and $Q'_y(b)$ alongside $Q_x(a)$ and $Q_y(b)$ which are derived from Q_2 , as shown in Figure 4.14(a) by the blue dots corresponding to the shift in x direction. With two dominant singular vectors in each direction correspond to image shift, LSF failed to achieve a proper fit to the true slopes to correctly resolve the shift in x direction as illustrated by the red line in Figure 4.14 (a).

To solve this problem, an algorithm called Piecewise Least Square Fitting PC (PLSF-PC) is proposed, which automatically fits a single dominant singular vector in one direction for LSF instead of two singular vectors. The proper fitting range is determined by a best-fit estimation as follows. The LS Fitting begins using the first 30% points from the original in dominant singular vector results in the initial fitting PL_{f0} .

$$(PL_{f0}, rmse_{f0}) = LSF(x_0, y_0) \quad (4.50)$$

Where LSF is the Least Square Fitting function, x_0, y_0 are fitting variables from the dominant singular vector in SVD, and $PL_{f0}, rmse_{f0}$ are the result function and root mean square error from the fitting.

Then, we gradually increase the numbers of points for fitting, as

$$(PL_{fn}, rmse_{fn}) = LSF(x_n, y_n), x_n = x_0 + n\Delta x \quad (4.51)$$

where $\Delta x = 0.1$.

The LSF fitting will end when the root mean square error $rmse_{fn}$ exceeds a given threshold T_{LSF} ($T_{LSF} = 0.1$). The result of PLSF-PC is shown in Figure 4.14 (b); the piecewise LSF ensures correct fitting based on the dominant singular vectors $Q_x(a)$.

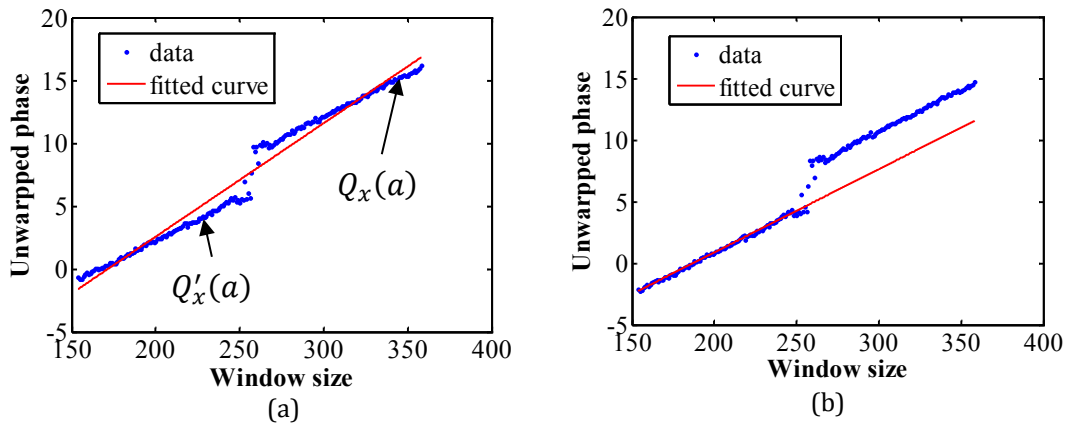


Figure 4.14 Original SVD and PLSF-PC in x direction, (a): SVD unwrapping and LSF result of PC of images taken under different illumination conditions. (b): PLSF-PC unwrapping and LSF result.

4.7 Experiments

4.7.1 *Effect of illumination variation on the PC matrix*

To verify and test the mathematical analysis of the illumination invariant property of the PC image matching algorithm presented in the previous section, investigation has been carried out by experiments using simulated hill shading images generated from DEMs with different azimuth and zenith angles. This experiment focuses on mountainous terrains because high relief regions are more sensitive to illumination angle variation than flat areas. DEM datasets of a mountainous region are selected from the ASTER GDEM which has been described in detailed in §3.3. We used ERMapper image processing software to generate the hill shading images from the DEM.

(1) Zenith angle variation

In this experiment, the DEM is given different zenith angles of illumination from 0° to 90° and identical azimuth angle 210° , to generate six hill shading images as shown in Figure 4.15. These images demonstrate that the main effect of zenith angle change is the change of both image brightness and contrast. The images are darker with large zenith angles than with small zenith angles. There are considerable variations in local grey value distributions in these images. For instance, the topographic shadows become shorter and brighter with the decrease of zenith angles but the orientations of the shadows remain approximately the same with the change of zenith angles.

The PC algorithm was tested using all possible zenith angle combinations. Figure 4.16 shows the results of using Figure 4.16(a) as the reference and the remaining five as target images. The target images were shifted 10 pixels to the right and downwards from the reference. The cross power spectrums are shown in Figure 4.16.

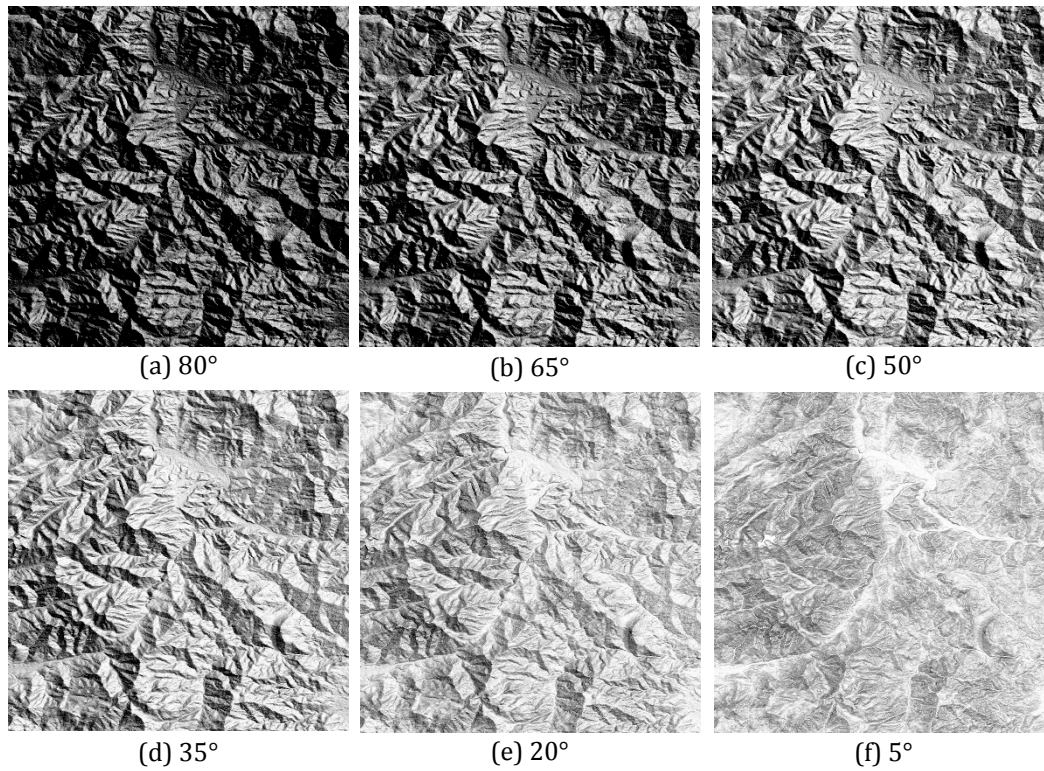


Figure 4.15 DEM simulated hill shading images with 210° azimuth angle and different zenith angles. The number under each image indicates the zenith angle of the image. (Image size: 1024×1024 pixels)

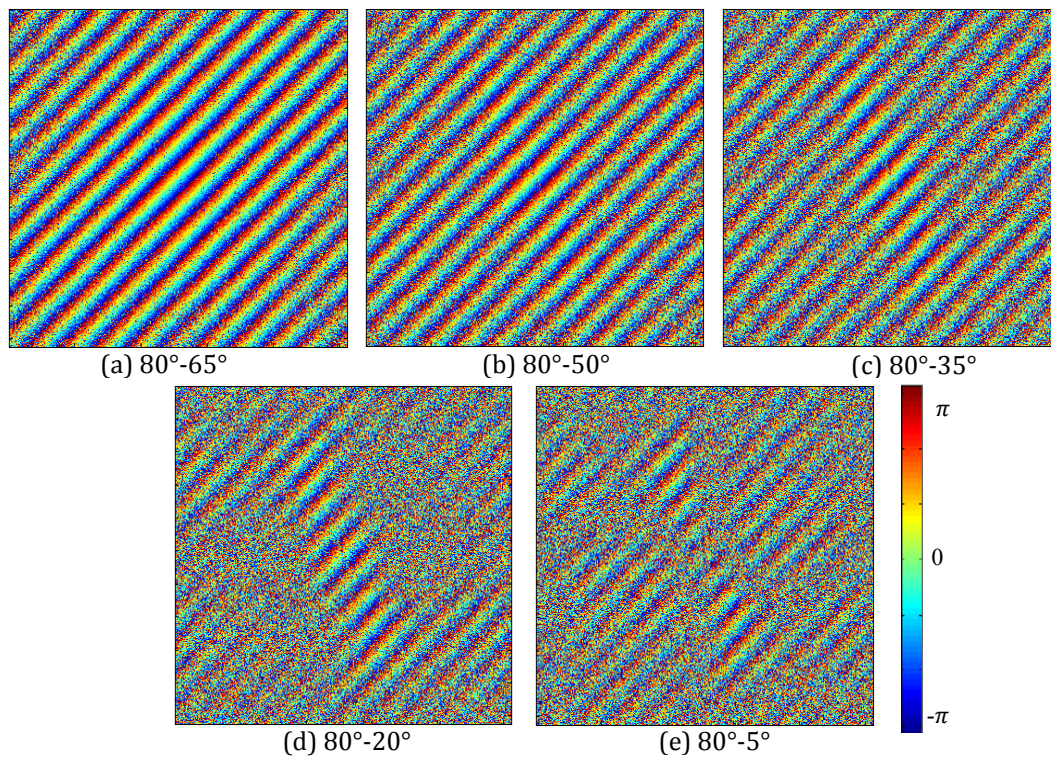


Figure 4.16 Phase correlation power spectra of image pairs with different zenith angles. The numbers under each spectrum indicate the zenith angles of the two images for matching.

The results reveal that the change of zenith angle blurs the phase correlation fringes globally but the orientation and density of fringes remain unchanged. This is consistent with the mathematical analysis discussed in §4.4: the variation of zenith angle degrades the image correlation, which is similar to the effect of adding random noise to an image. The greater the zenith angle difference is, the more blurred the fringes become.

To evaluate the quality of phase correlation fringes subject to different zenith angle differences, the peak δ_0 values for all the matching pairs are calculated as shown in Figure 4.17, closes to a nearly linear trend: the peak δ_0 values decrease with the increase of zenith angle differences corresponding to the more and more blurred phase correlation fringes (Figure 4.16).

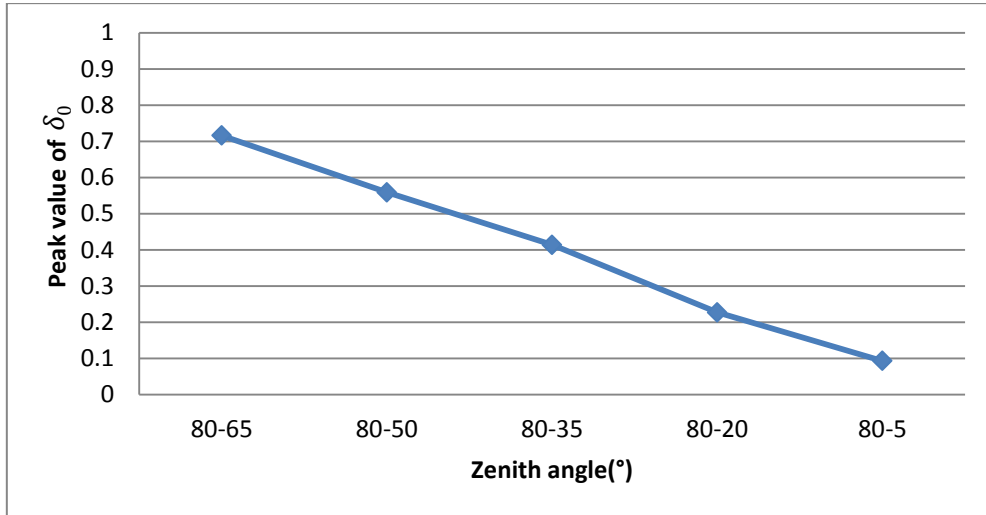


Figure 4.17 The peak values of δ for PC matching with different zenith angles.

(2) Azimuth angle variation

In this experiment, the effect of azimuth angle change on image matching is investigated. Hill shading images are generated every 60° as shown in Figure 4.18: six hill shading images with azimuth angles ranging from 60° to 360° and at the same zenith angle (35°).

Although the change in azimuth angles does not alter the global image brightness, it does alter local image contrast and terrain texture/pattern orientation. Since all

the images are generated with the same zenith angle, the mean grey values of each image are similar. However, with the change of illumination direction, the local contrast patterns of topographic textures are altered between the images. When the difference in azimuth angle equals 180° , the textural contrast of topography between the two images is entirely reversed, for example, Figure 4.18(a) and Figure 4.18(d).

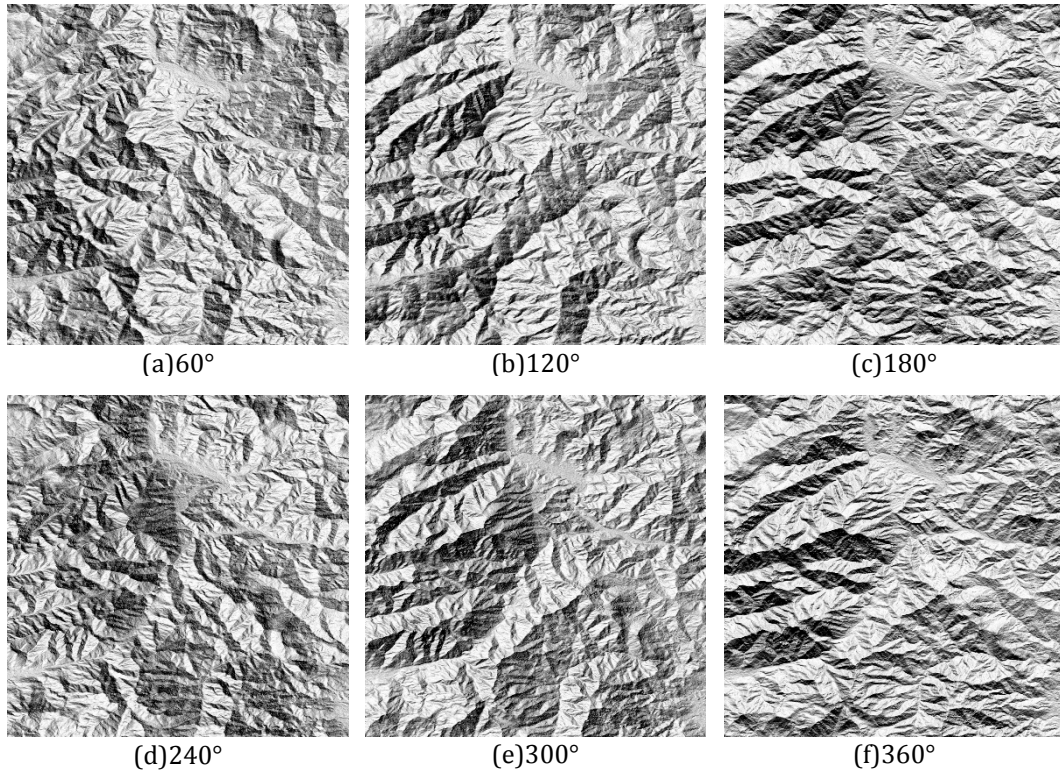


Figure 4.18 DEM simulated hill shading images with different azimuth angles at the same zenith angle (35°). The number under each image indicate the azimuth angle of the image. (Image size: 1024×1024 pixels)

All possible azimuth angle combinations between the six images have been tested. Figure 4.19 shows an example of the cross-power spectra of the phase correlation for image matching using the first image in Figure 4.18 ($\tau=60^\circ$) as the reference image and the remaining five as target images. The target images were shifted 10 pixels to the right and downwards from the reference in the matching experiment. In general, the dominant features in the cross power spectrum are the fringes corresponding to the image shift but these fringes are divided into angular sectors

of positive and negative correlation, corresponding to the azimuth angle difference, $[\tau_1, \tau_2]$, between the two images for matching, as proved by Equation (4.43).

Considering Figure 4.19(a) as an example, the azimuth angles of the two images are 60° and 120° ; the range of positive correlation sectors is 240° , $2\pi - 2d$, whilst the range of negative correlation sectors is 120° , $2d$. Since the zenith angle is relatively small, according to Figure 4.10, *NP (Negative Correlation)* phenomenon is not very strong, as can be seen in Figure 4.19(a), where the fringe quality of negative correlation is lower than that of positive correlation; and the same is true for the other four fringe maps. Figure 4.19(c) is a special case of azimuth angles 60° - 240° ; here the positive correlation area is $2\pi - 2d = 0$, and there remains negative correlation only. The blurred fringes in Figure 4.19(c) are not caused by illumination angles but by the orientation of terrains.

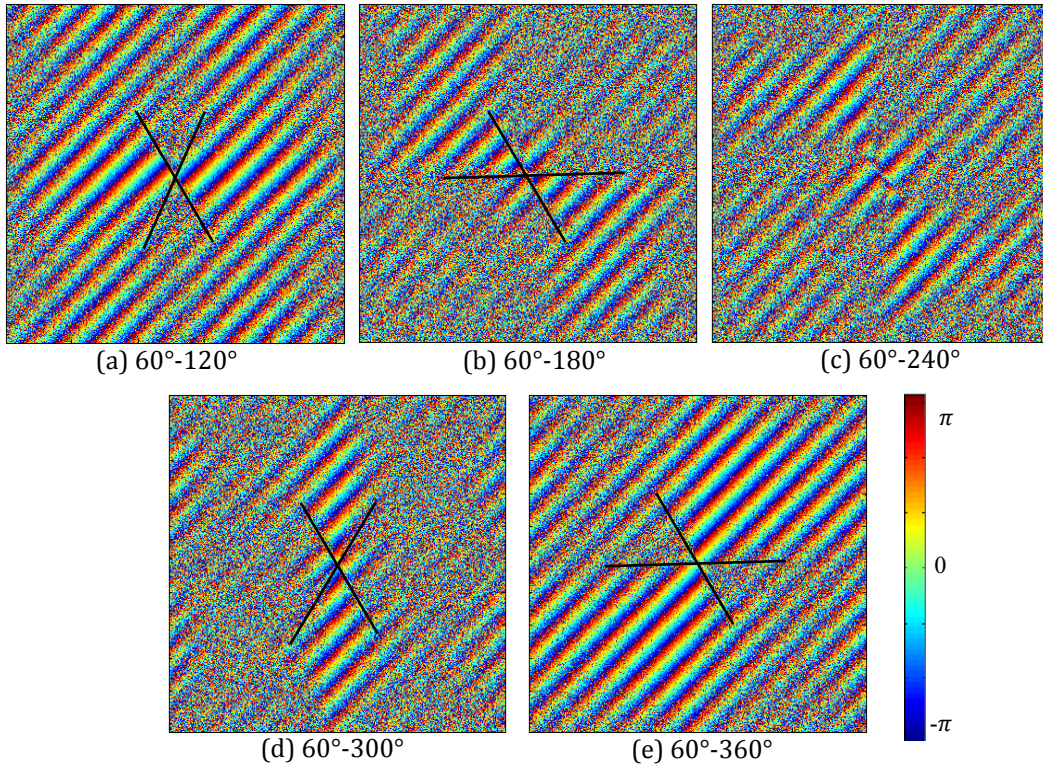
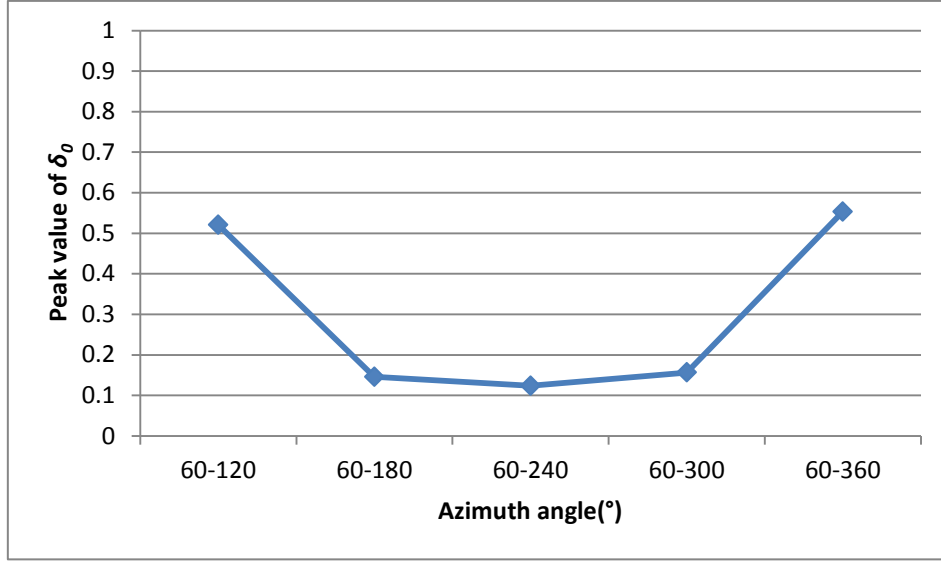


Figure 4.19 Phase correlation power spectrums of image pairs with different azimuth angles. The numbers under each image indicate the azimuth angles of the two images for matching.

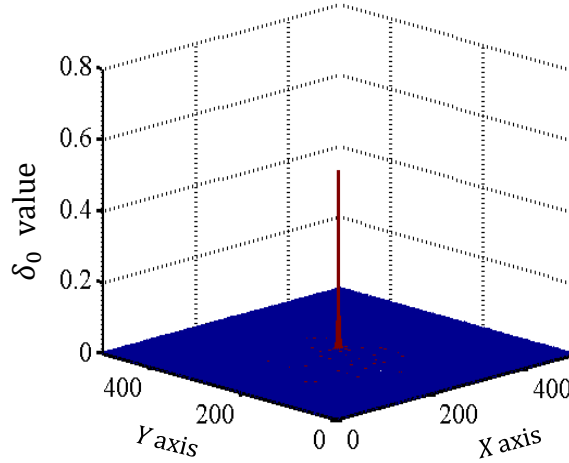
The densities and orientations of PC fringes are the same in Figure 4.19. This confirms that azimuth angle change in illumination and image translational shift

are two independent processes. The azimuth angle change does not alter the density and orientation of phase correlation fringes which are determined by the translational shift in vertical and horizontal directions between the two images for matching. As the azimuth angle variation does not introduce any other patterns and the rest of the fringes are still clear and with the same density and orientation; the shift in horizontal and vertical direction can still be fairly accurately resolved from the phase correlation matrix.

To evaluate the overall quality of phase correlation fringes in Figure 4.19, which are partially degraded by the variation of azimuth angle, the value of δ_0 was calculated via inverse Fourier transform for every power spectrum. The peak value of Dirac delta function δ_0 is close to 1 for high quality clear fringes but decreases with the degradation of fringe quality although the peak position remains unchanged. As illustrated in Figure 4.20(a), the δ_0 peak value decreases with increasing in azimuth angle difference between the two images for matching, and it drops to its lowest when the difference in azimuth angle reaches 180° (60° - 240°) because, in this case, there is no sector of major correlation as shown in Figure 4.19(c). The smaller the azimuth angle difference is, the larger the area of major correlation and, hence, the higher the δ_0 peak value. While the image shift is identified by the peak position of δ_0 rather than its value, the PC image matching is effective and invariant to the azimuth angle variation. Figure 4.20(b) shows the distinctive peak of δ_0 generated by the azimuth angles of 60° and 120° , which demonstrates that even under the condition of significantly altered grey value distribution, the PC method is still able to correctly identify the image shift between the pair of images at integer (1 pixel) level.



(a) The values of δ_0 peaks



(b) 3D display of δ in PC

Figure 4.20 (a) Values of δ_0 peaks of PC matching with different azimuth angles. (b) A 3D illustration of δ_0 peak for the image matching case of azimuth angle 60° and 120° .

Then, we keep the same azimuth angle difference, 60° - 120° , for every image pair but gradually increasing zenith angles from 5° to 80° . Six image pairs were generated, and Figure 4.21 shows the first image pair: (a) the reference image with azimuth angle 60° and zenith angle 80° , and (b) the target image with azimuth angle 120° and the same zenith angle. For each image pair, the target image was shifted 10 pixels horizontally rightward and 10 pixels vertically downward to generate PC fringes, shown in Figure 4.22.

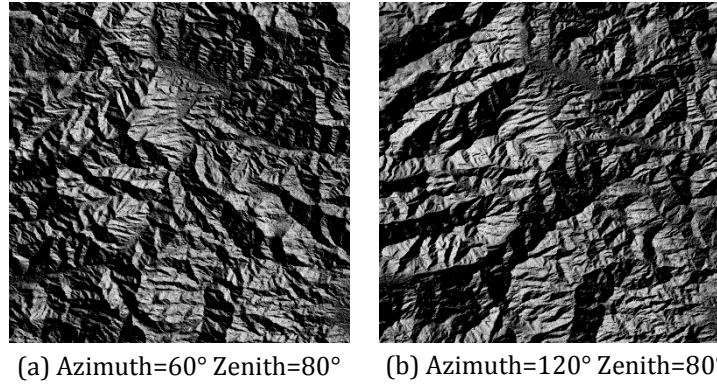


Figure 4.21 Image pair under different illumination conditions for PC based matching. (Image size:1024×1024pixels)

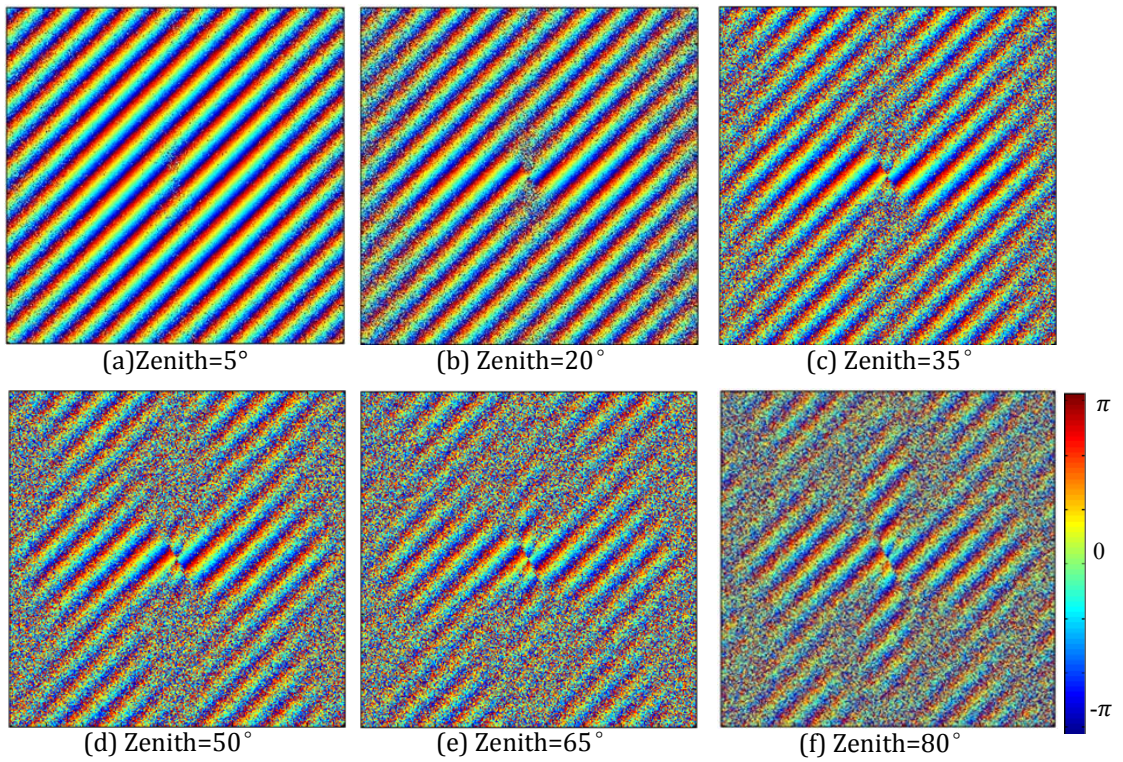


Figure 4.22 PC cross power spectrums generated from two DEM shading images with azimuth angles of 60° and 120° and various zenith angles, ranging from 5° to 80°.

Figure 4.22(a) shows that in the case of very small zenith angle, despite of 60° azimuth angle difference, the PC power spectrum only shows strong positive correlation and no negative correlation owing to the dominance of the term L_1 according to analysis in the §4.4 *case 1*. Then, with the increase in zenith angles (Figure 4.22(b)-(f)), the negative correlation as the result of azimuth angle difference, becomes increasingly stronger. This verifies Equation (4.45) that with the increase in zenith angle, the *NP* phenomenon becomes significant, and the

strength of both positive correlation and negative correlation becomes weaker compared to the positive correlation shown in Figure 4.22(a). However, the dividing lines of negative and positive correlation sectors remain the same, and these are determined by the azimuth angle difference according to Equation (4.42). When zenith angle reaches 80° , the negative correlation is at its strongest. According to the analysis in §4.4 case 2, for large zenith angles, the L_2 term is dominant, so that the *NP* phenomenon becomes most significant. From Figure 4.22, it can be concluded that azimuth angle variation results in the combination of negative and positive correlation, but value of zenith angle determines the strength of the azimuth variation impact effect.

To assess phase correlation quality, IFT was performed to convert the cross power spectrum in Figure 4.22 to corresponding δ_0 functions. As indicated in §4.2, in cases of the existence of negative correlation, there are two peaks in δ_0 function: a positive peak for positive correlation and a negative peak for negative correlation. The values of peaks, ranging from 0 to 1, indicate the level of the correlation. The positive and negative correlation δ_0 peak of Figure 4.22 are calculated and shown in Figure 4.23, which demonstrate that the negative peak values increase gradually with the increase in zenith angle. According to §4.4, the increasing trend of negative peak values is caused by the increase of the contribution of the term L_2 in Equation (4.45), that is a monotonically increasing function of zenith angle. This can be confirmed by comparing Figure 4.23 with Figure 4.10 where the negative peak values show the same trend as L_1 values in Figure 4.10. It can also be seen in Figure 4.22, that the strength of negative correlation fringes become stronger with the increase in zenith angle.

In contrast, positive peak values shown in Figure 4.23 illustrate that the strength of positive correlation decreases significantly in the increase of zenith angle. There are two reasons for the decrease in positive peak values. Since the sum of all elements in δ_0 function equals 1, the increased strength of negative correlation will suppress the positive correlation, and thus the values of positive peaks go down.

The other reason is that the values of term L_2 decrease with the increase in zenith angle as shown in Figure 4.10 and so is the contribution of L_2 to positive correlation. The phenomenon is also shown in Figure 4.22; the quality of positive correlation fringes decreases considerably with the increase in zenith angle.

In spite of the increasing in negative correlation, the overall correlation value, the sum of the negative peak and the positive peak, decrease with increase in zenith angle, as indicated in Figure 4.23. This implies that the image matching robustness would decline with the increase in zenith angle.

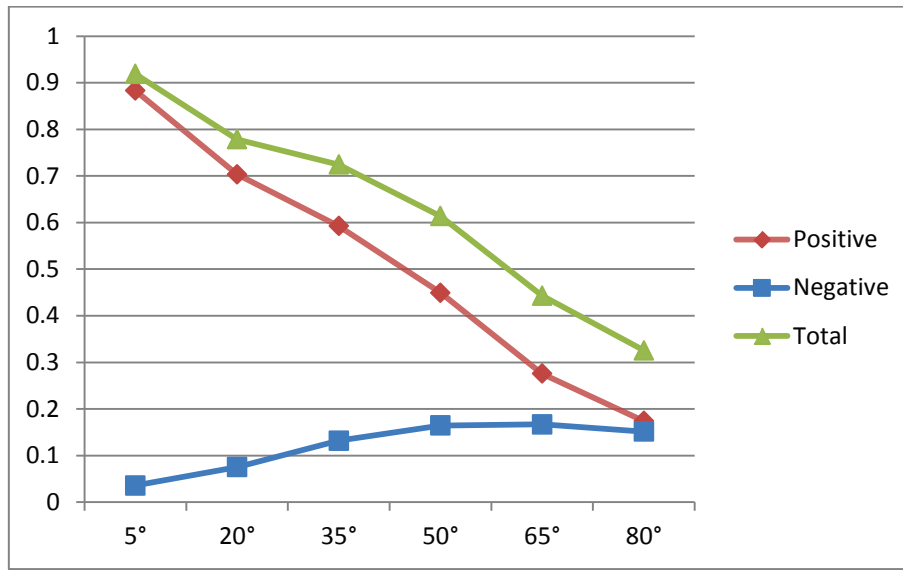


Figure 4.23 Positive, negative and total peak values of the Dirac delta function in respect to the increase of zenith angle.

In summary, the experimental results in this section verified the analysis in §4.4 that PC cross power spectrum comprises the positive correlation contributed by the L_1 term in Equation (4.45) and the combination of negative and positive correlation (NP) contributed by the L_2 term in Equation (4.45). The ratio of NP and P is determined by zenith angle, while the sectors of negative and positive correlations in NP phenomenon are determined by azimuth angle difference. With the increase in zenith angle, the negative correlation increases gradually with rapid declination of positive correlation and thus the overall correlation decreases significantly.

4.7.2 Sub-pixel image matching results and comparison

(1) Azimuth angle variation

Firstly, the robustness of the proposed PC based image matching algorithms towards azimuth angle variation is tested. The six hill shading images shown in Figure 4.18 were used for image alignment experiment. To test the sub-pixel image matching accuracy, the image shown in Figure 4.18(a) was shifted 5.5 pixels rightward and downward by image up-sampling. Matching accuracy is assessed by the average value of shifts in x and y directions:

$$v = (|v_x| + |v_y|)/2 \quad (4.52)$$

Where $v_x = x - x'$, $v_y = y - y'$, (x, y) is the measured frame shift, and (x', y') is the actual frame shift. In this experiment, $x' = 5.5$, $y' = 5.5$.

Three PC-based matching algorithms SVD, ADCF-PC and PLSF-PC were used for image alignment, with a 512×512 pixels image window. For comparison, the same image pairs of different azimuth angles in Figure 4.18 were matched using SURF, NCC and MI. The matching accuracy is also assessed by the deviation v . If v is larger than 10 pixels, the matching is considered as 'failure', labelling 'F' in Table 4.1. For NCC and MI, the reference image was regarded as a template window, searching for the range of 20 pixels in vertical and horizontal in a target image, and the final location is determined by the peak value of NCC or MI coefficients. This experiment is under the known image translation, so the searching range can be determined. In the real situation of image alignment, the translation between the two images is unknown, so the searching time could be much longer since the searching range is unknown. The processing times of the six algorithms were calculated, and the processor is Intel® Core™ i7-3770 CPU @ 3.40GHz.

Table 4.1 Image alignment performances of six matching algorithms under azimuth angle variation

Algorithm (ν)\Azimuth($^\circ$)	60-120	60-180	60-240	60-300	60-360	Time(s)
SVD-PC	3.62	7.35	3.29	0.64	0.06	0.61
ADCF-PC	0.02	0.15	0.33	0.20	0.10	0.72
PLSF-PC	0.03	0.05	0.03	0.07	0.03	0.65
SURF	F	F	F	F	F	1.93
NCC	0.5	F	F	F	0.5	38.27
MI	0.5	0.5	0.5	0.5	0.5	201.03

Table 4.1 shows that SVD-PC is not very robust to azimuth angle variation; it failed to reach sub-pixel precision in most cases except for 60°-300° and 60°-360°. For the failure cases, there are two groups of inversed fringe patterns in the PC matrix and thus LSF failed to fit one line properly, as explained in Figure 4.14(a). ADCF-PC succeeded in all the azimuth angle variation cases, and the average matching accuracy is 0.16 pixels. If the difference of azimuth angle is less than 60°, ADCF-PC can reach 0.02 pixel matching accuracy. The matching accuracy of ADCF-PC decreases with increasing azimuth angle difference, and the lowest matching accuracy appeared in the case of 60°-240° or, the azimuth angle difference equals to 180°. In this case, the negative fringes are predominant in the PC matrix, corresponding to the largest texture difference between the two images. PLSF-PC succeeded in all the cases and achieved higher matching accuracy than ADCF-PC, averaging 0.05 pixels. An interesting observation of PLSF-PC is that the matching accuracy is quite consistent for all illumination variation cases. For instance, the accuracy of 60°-120° is almost the same as that of 60°-240° for PLSF-PC.

The matching accuracy of SURF indicates that the algorithm is sensitive to the variation of azimuth illumination angle and it is not azimuth invariant at all. With matching deviation $\nu > 10$ pixels in all cases, SURF failed to find correct corresponding points when the two images for matching are taken under illumination of different azimuth angles. As reviewed in §2.1, the feature-based matching algorithms rely mainly on the local grey value distribution. If the local grey value distribution is altered considerably, as with the change of illumination azimuth, the feature-based matching algorithms fail.

NCC only succeeded in the cases of small azimuth variation, 60° - 120° and 60° - 360° , but failed in large azimuth variation cases. NCC coefficients indicate that the azimuth angle variation reduces the image correlation considerably. MI succeeded in all azimuth angle variations image matching cases with consistent accuracy. However, both NCC and MI cannot achieve sub-pixel matching accuracy without discrete interpolation. In contrast, PC-based approaches can directly achieve sub-pixel matching accuracy as phase is a continuous variable in the frequency domain.

Another limitation of these two area-based approaches is matching speed. Even based on prior knowledge of searching ranges, the image matching speed of NCC and MI are significantly lower than SURF and PC which can determine the translation shift between images without roaming searching.

(2) Zenith angle variation

Here the robustness of the six algorithms towards zenith angle variations is tested. The hill shading images for alignment are shown in Figure 4.15. The image shift and experiment settings are the same as azimuth variation cases.

Image matching using SVD-PC, ADCF-PC, PLSF-PC, SURF, NCC and MI algorithms was then carried out for every possible pair of the images in Figure 4.15. Table 4.2 shows the matching accuracy for the image pairs with 80° zenith as reference and the remaining images as target images.

Table 4.2 Image alignment performances of six matching algorithms under zenith angle variation

Algorithm (v)\Zenith($^\circ$)	80-65	80-50	80-35	80-20	80-5
SVD-PC	0.006	0.01	0.02	0.05	1.45
ADCF-PC	0.03	0.03	0.01	0.13	0.55
PLSF-PC	0.006	0.01	0.04	0.05	1.33
SURF	0.02	0.16	0.66	F	F
NCC	0.5	0.5	0.5	0.5	2.5
MI	0.5	0.5	0.5	0.5	0.5

Compared to azimuth angle change, SVD-PC is more robust to zenith angle change. Without azimuth angle variation, there is no negative fringe pattern that may cause inaccurate LSF in SVD-PC approach. It can achieve 1/150 pixel matching accuracy when the zenith angle difference is less than 15° and, even if the zenith angle

difference is as large as 60° , the matching accuracy can still reach 1/20 pixel. The largest matching error of PC is 1.45 pixels when the zenith difference is 75° . PLSF-PC has similar performance to SVD-PC and the largest error also occurred for the case of 75° zenith angle difference. ADCF-PC cannot achieve as high accuracy as SVD based approaches when zenith angle difference is small but it is more robust when the zenith angle difference is large. The matching accuracy is within 1 pixel when the zenith difference is 75° , which is higher than the results of SVD-PC and PLSF-PC.

The data indicate that the three PC-based approaches are all invariant to zenith angle change. According to the analysis in §4.3, if the azimuth angles are the same, the effect of zenith angle variation is likened to the addition noise to Phase Correlation cross power spectrum. Both SVD frequency domain based and spatial domain based approaches are both robust to image noise.

As indicated in Table 4.2, the SURF matching algorithm succeeds when the zenith angle difference is within 45° . The accuracy of matching equals PC. As zenith differences increases, the robustness of the algorithm decreases drastically. When the zenith difference is greater than 60° , SURF starts to fail in finding adequate corresponding points for effective image matching. SURF is less sensitive to zenith angle change than to azimuth angle change because zenith angle variation mainly changes the global image contrast and brightness without altering the local grey value distribution too much. Most feature-based matching algorithms have certain tolerance to the image global contrast and brightness differences. When the zenith angle difference becomes greater, however, the local grey value distribution changes more significantly and SURF algorithm becomes unreliable.

The data in Table 4.2 demonstrate that NCC and MI algorithms are both able to achieve integer level accuracy in all cases, except NCC, for the case of 80° - 5° with 2 pixels deviation. These template-based algorithms are generally robust to zenith angle change but, as indicated previously, they cannot directly achieve sub-pixel matching accuracy.

In general, nearly all tested algorithms perform better in zenith angle variation experiment than in azimuth angle variation experiment and thus have higher tolerance to zenith angle variation.

(3) Daily illumination variation

Finally, the case of changing illumination azimuth angle and zenith angle jointly is tested. Hill shading images are generated from a DEM based on the daily Sun azimuth and zenith variation on 1st June. The azimuth and zenith angles are calculated every two hours from 8:00 to 16:00 hours as shown in Table 4.3. The simulated hill shading images under these illumination conditions are shown in Figure 4.24. The data indicate that the azimuth angle variation during the time interval can be as large as 177°; this means that the shadows can appear in opposite directions, changing the local terrain patterns greatly. The largest zenith angle difference is 36° and consequently the images taken at noon time around 12:00 are brighter than other images, as we know from common sense. The local terrain patterns are also altered by the change of zenith angle but less significantly in comparison with the case of azimuth angle change.

Image matching was then carried out using SVD-PC, ADCF-PC, PLSF-PC, SURF, NCC and MI algorithms for comparison. The matching assessment is shown in Table 4.4 and Figure 4.25. The data indicate that PLSF-PC is the most robust to daily illumination changes for image matching, as compared to PLSF-PC and ADCF-PC. The matching accuracy of PLSF-PC is high, with matching deviation v mainly well below 0.1 pixel and in a range of 0.01-0.1 pixels. The sub-pixel matching accuracy is unsatisfactory for both SVD-PC in the case of 8:00-12:00 and the ADCF-PC in the case of 8:00-14:00.

Table 4.3 Solar azimuth and zenith angle variation in on 1st June

Illumination angle(°)\Time	08:00	10:00	12:00	14:00	16:00
Azimuth	89.89	114.86	173.14	239.21	266.87
Zenith	55.24	33.20	19.07	29.98	51.60

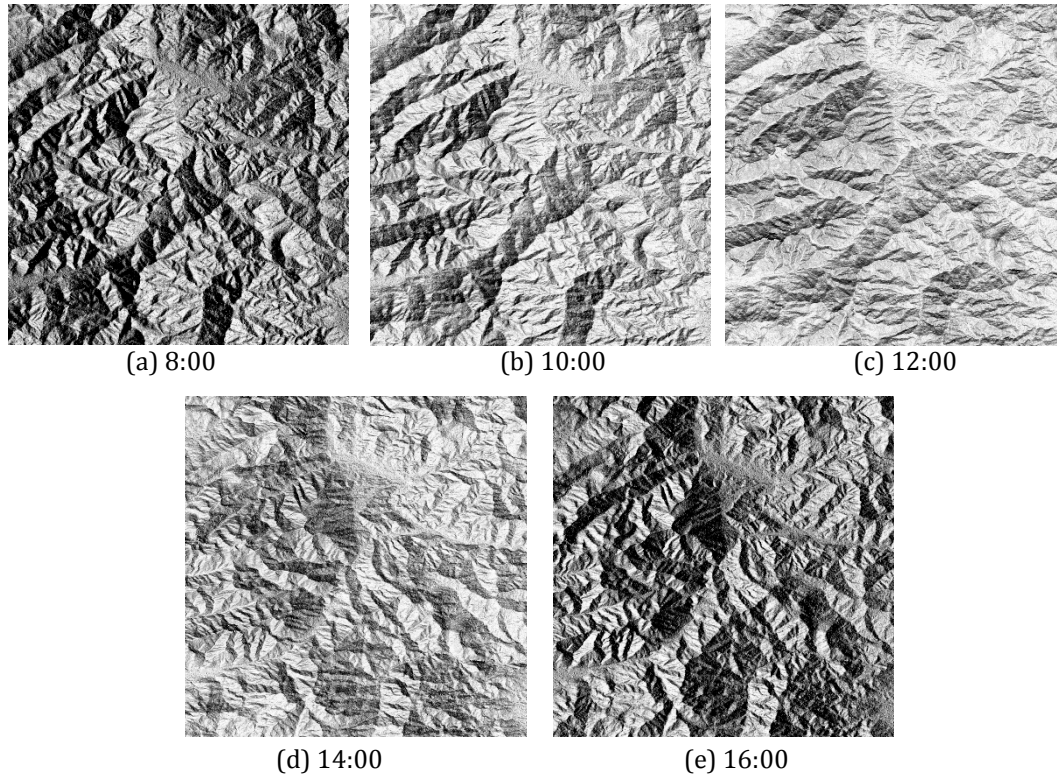


Figure 4.24 DEM simulated hill shading images under solar illumination conditions of different time on a day. The numbers under each image indicate the imaging times. (Image size: 1024×1024 pixels)

Table 4.4 Image alignment performances of six matching algorithms under daily illumination angle variations (pixel)

Algorithm (v)\Time	8:00-10:00	8:00-12:00	8:00-14:00	8:00-16:00
SVD-PC	0.03	0.76	0.11	0.02
ADCF-PC	0.04	0.16	0.59	0.05
PLSF-PC	0.01	0.10	0.06	0.01
SURF	0.22	F	F	F
NCC	0.5	F	F	F
MI	0.5	0.5	0.5	0.5

In contrast, the SURF performed poorly for matching the images taken under daily solar illumination variation, as evidenced by the data in Table 4.4 and Figure 4.25. The only successful case is the matching between 8:00 and 10:00; matching fails if the time interval is more than two hours. The SURF is generally not robust to daily illumination change. The matching performance of NCC is similar to that of SURF, which fails during a time interval of more than two hours. MI succeeds in all image matching cases under daily change of illumination but it cannot directly achieve sub-pixel accuracy.

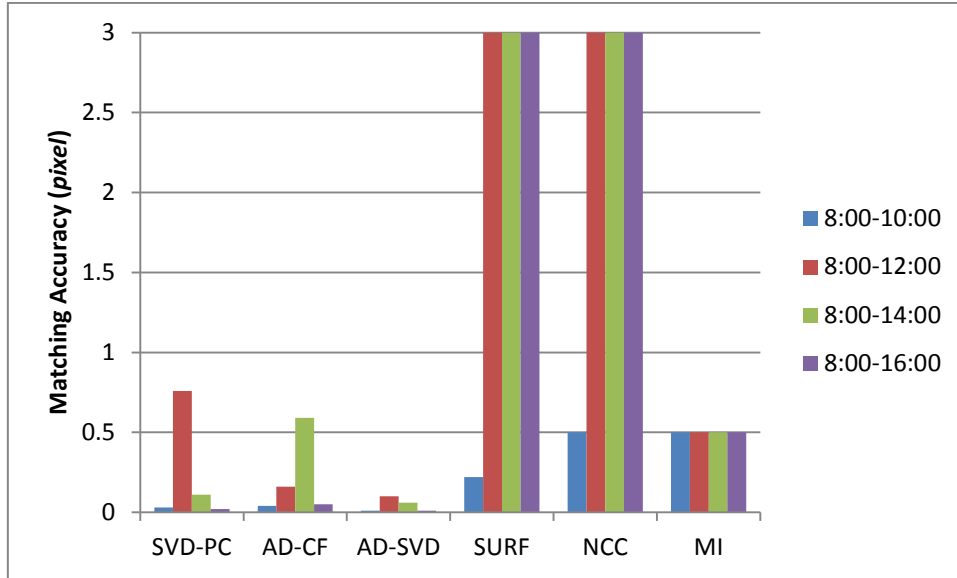


Figure 4.25 Matching accuracy of hill shading images under daily illumination variation using six matching algorithms (matching window: 512×512 pixels). The accuracy is set as 3 pixels for all failure cases.

Furthermore, the robustness of image matching using various sizes of matching window was tested in an experiment focusing on the comparison of PLSF-PC and ADCF-PC since the previous experiments revealed that only these two algorithms can robustly achieve sub-pixel matching accuracy under illumination variant conditions. The window size is digitally reduced from 512×512 pixels to 32×32 pixels. Figure 4.24(a) and Figure 4.24(e) in were used as reference image and target image, with a shift of 4.5 pixels in x and y direction. A match is considered as ‘correct’ if v is within 1 pixel. The comparison of matching correct rates of ADCF-PC and PLSF-PC is presented in Figure 4.26.

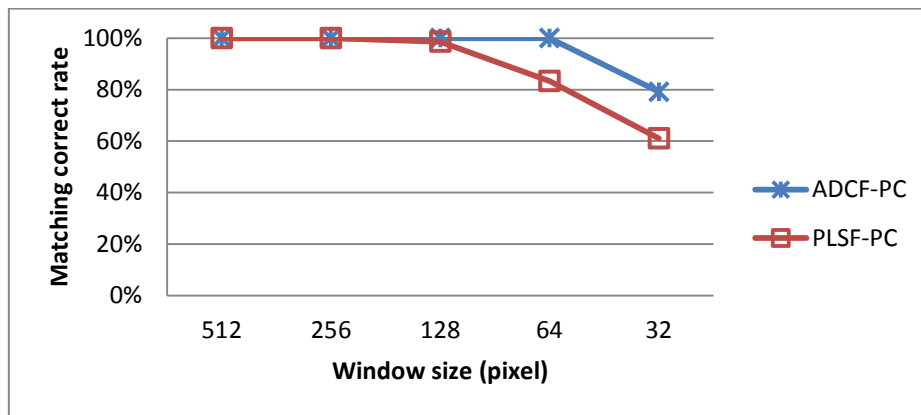


Figure 4.26 Image matching robustness comparison between ADCF-PC and PLSF-PC

The data in Figure 4.26 demonstrated that both ADCF-PC and PLSF-PC have nearly 100% matching correct rate when the matching window is larger than 128×128 pixels but PLSF-PC becomes less robust when the matching window size drops to 64×64 pixels and 32×32 pixels. The matching correct rate of PLSF-PC at window size 32×32 is only 61% indicating that reliable matching cannot be achieved using a small matching window. In contrast, the matching correct rate of ADCF-PC indicates better tolerance to small matching windows.

Figure 4.27 shows an example of disparity estimation from images under different zenith angles. The size of matching window is 32×32 pixels. With different azimuth and zenith angles, the images in Figure 4.27(a) and Figure 4.27(b) are significantly different in texture but the ADCF-PC still presents a distinctive peak in δ_1 function, as shown in Figure 4.27(c). The PLSF-PC algorithm failed to achieve a reliable and accurate estimation of disparity as illustrated in Figure 4.27(d). The robustness of LSF in SVD declines significantly when the number of fitting points becomes too small when the algorithm becomes vulnerable to noise and unstable in a small matching window.

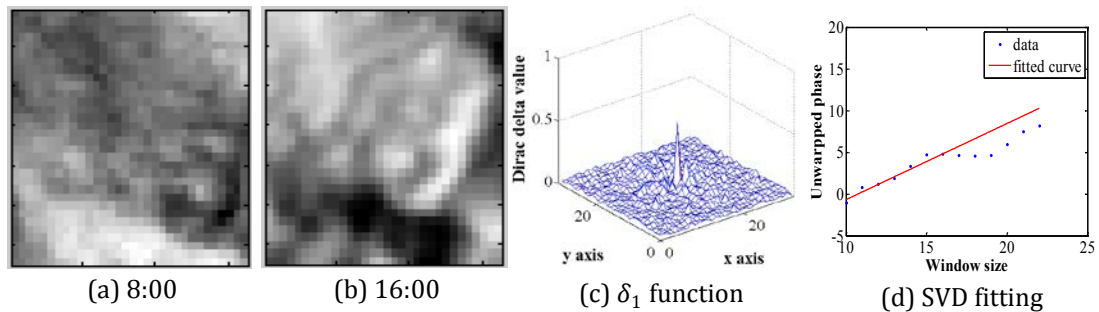


Figure 4.27 Performances of ADCF-PC and PLSF-PC using a small matching window. (a) and (b) are two hill shading images in a 32×32 pixels matching window. (c) is the plot of δ_1 function from the IFT of the PC matrix of images (a) and (b). (d) is the plot of SVDs derived from PC matrix of images (a) and (b) and the LSF lines.

The above experiments demonstrate that PLSF-PC is able to achieve high sub-pixel matching accuracy when a large matching window is applied (larger than 128×128 pixels). Thus, PLSF-PC is suitable for illumination-insensitive image alignment, for example in vision-based UAV navigation (Chapter 5), using real-time UAV images to match a reference satellite image acquired at different time. However, the well-

established ADCF-PC algorithm is more robust for small matching windows (e.g. 32×32 pixels). Thus, ADCF-PC will be used for change detection (Chapter 6) and pixel-wise disparity estimation (Chapter 7), for which a small matching window is required.

4.8 Summary

In this chapter, the effects of varying the local illumination, azimuth variation and zenith angle variation, on PC (phase correlation) matrix have been investigated. Through mathematic derivation, it is proved that the illumination impact matrix is under a joint effect of azimuth angle variation and zenith angle variation; the negative correlation caused by the azimuth angle difference between the two images for matching becomes more significant with the increase of zenith angle. If the azimuth angles are the same, the effect of zenith angle alone is like adding random noise to the phase correlation cross power spectrum. Consequently, the inversed fringes caused by azimuth angle difference and increased noise caused by zenith angle variation appear in the PC matrix but the orientation and density of the PC fringes are not altered. This is the key factor for PC based image matching algorithms to be robust to illumination angle variation, such as the absolute Dirichlet Curve Fitting based PC (ADCF-PC) algorithm. Furthermore, it is proposed that a Piecewise Least Square Fitting based PC (PLSF-PC) matching algorithm that can achieve higher sub-pixel precision for illumination-insensitive image alignment but it becomes unstable when the matching window is small.

Chapter 5 Autonomous Vision-based UAV Navigation in a Changing Illumination Environment

5.1 Vision-based UAV navigation

UAVs (Unmanned Aerial Vehicles), also known as UAS (Unmanned Aerial System), have been developed very rapidly for ever-widening applications, particularly in the last decade. To achieve full autonomy, the movement of a UAV needs to be self-monitored and self-controlled from a starting point to an end point, known as 'autonomous navigation'. GPS is the most widely used positioning system for UAV navigation but in some circumstances, such as military operations, GPS signals may not be available or can be jammed or sabotaged [68]. Moreover, for planetary exploration, since GPS service is unavailable, autonomous navigation becomes a crucial issue to UAVs. An UAV flight path usually contains many stops allowing the UAV to acquire on-board images and the stop interval depends on the required image overlapping rates. When GPS data is unavailable or unreliable, vision-based navigation, utilizing images from on-board cameras for localisation, is a promising alternative for navigation.

Vision-based UAV navigation can be divided into two categories: the frame-frame and the frame-reference approaches. One of the popular frame-frame methods is SLAM (Simultaneous Localisation and Mapping) [69], which estimates the camera motion by image matching between the current and the previous frames. Without absolute positioning information, however, these methods suffer from a drift problem which can be moderately reduced by applying filters, such as Kalman [70]

and EK (Extended Kalman) filters [71, 72], but this relies on the assumption of loop closures which means that UAVs need to fly on the same area twice [73]. Recent research on SLAM also takes the DEM as a constraint for avoiding error accumulation, known as the geo-SLAM [74], but the large amount of computation makes it unsuitable for real-time high speed UAV navigation. Another approach that achieves UAV navigation by image matching between the real time UAV images captured by an on-board camera and pre-installed reference images, is called frame-reference method. One of the advantages of the frame-reference approach is that the localisation error does not accumulate with time because the position of the camera is calculated independently at each UAV stop. Moreover, by geo-referencing UAV images and reference images, not only can a current UAV position be determined, but also the next UAV position can be predicted because the UAV images cover certain front areas depending on the UAV view angles. The localisation of front areas is of great importance for autonomous navigation, because it allows vehicles to know 'where to go'.

Image matching for the frame-reference approach is challenging because the on-board real time images are not necessarily acquired at the same time under the same illumination and with the same imaging setting as pre-installed reference images. To achieve robust performance, variation of illumination, scale and imaging geometry are problems to be overcome in frame-reference approaches [75]. By knowing the flight altitude and camera parameters, roll, pitch and yaw, the scale difference and geometric distortions can be rectified but the appearance of the two images can be very different because of different illumination conditions. This issue is especially serious in mountainous areas because image matching depends largely on topographic features, which can be greatly altered by changes in solar position. For example, the shadows on a mountain slopes may appear on opposite sides if the reference image was acquired in the morning whilst the UAV was flying in the afternoon.

Several approaches have been proposed to achieve UAV navigation using robust image matching algorithms. Edges are also useful features for UAV image matching; Rodriguez and Aggarwal [76] introduced a cliff map for the image matching between the aerial images and the reference 3D data. Shang and Shi [77] used the boundary of runway for autonomous UAV safe landing. However, as indicated in Chapter 2, edges are not truly illumination and scale invariant.

Other approaches based on area correlation formulate another brunch of vision-based UAV navigation. Omead and Talep [78] proposed a visual odometer to determine the position of UAVs by SSD (Sum of Squared Differences) template matching. An auxiliary template was used to calculate the rotation, scale and grey value normalisation between templates. Lin and Medioni [73] used MI to find the correspondences between the UAV frame and the reference map. The general perception that area-based matching is slower than the feature-based matching makes it less widely used in vision-based navigation thus far. The searching for optimal matching in frame-reference is often time-consuming making real-time navigation not always achievable. Following the study of the illumination invariant property of PC detailed in Chapters 4-5, a Phase Correlation (PC) based UAV navigation approach in illumination change environment is introduced in this chapter. The illumination invariant Phase Correlation matching algorithms will be used to achieve frame-reference based UAV navigation. As a non-iterative method, Phase Correlation can directly calculate the image shifts without roaming search and therefore operates fast enough to meet the requirements of real-time navigation.

The research in this chapter focuses on the following objectives:

1. Develop a robust frame-reference localisation algorithm for UAV navigation in local illumination variation environments.
2. The proposed UAV navigation system need to have high tolerance to image distortion caused by camera 3D motion because imaging from agile small UAVs can

result in severe geometric distortion. Fast geometric correction for UAV images is an important issue needing to be addressed for vision-based UAV navigation.

3. The UAV navigation system should be able to predict the next UAV position based on the current on-board UAV images and the planned flight route.

4. The UAV navigation system should be safe guarded by a reliable risk alarming algorithm in case the UAV deviates from the planned route, and should perform self-correction.

5.2 Proposed PC-based autonomous UAV navigation method

We assume that the field of view (FOV) of each UAV image covers at least two stops, the current stop and the next stop. Under this assumption, a fast and robust UAV navigation workflow can be designed to determine the current UAV position as well as the next UAV position as shown in Figure 5.1.

According to the UAV mission plan, two search images, one based on the current stop and the other based on the next stop, are extracted from pre-installed reference images. For UAV images, the platform could be affected by many factors, and so a geometric rectification needs to be done before the generation of two query images for current position and the next position from on-board UAV images. The two query images are then localised on the two search images independently via a PC based robust image matching algorithm. If the matching results are reliable via matching quality assessment, the current and the next UAV position are determined. A risk assessment is carried out by the comparison between the calculated UAV position and the corresponding mission planned position. If the deviation between the calculated position and the planned position is smaller than a given threshold, the UAV will keep its current flight status. Otherwise, the UAV needs to adjust its position according to the estimated deviation.

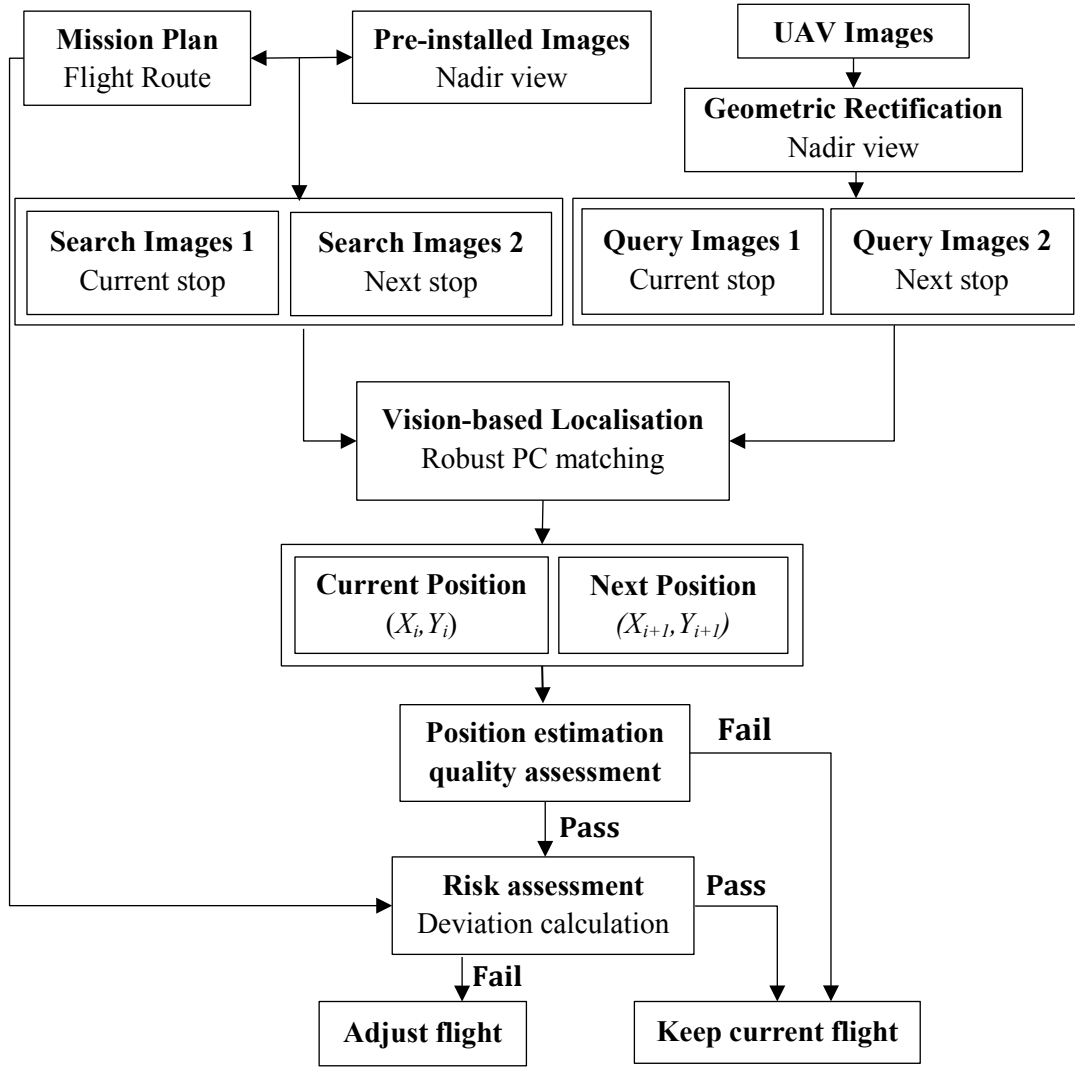


Figure 5.1 PC-based UAV navigation Workflow

5.2.1 Geometric distortion rectification

Owing to its light weight, the UAV's 3D status can easily be affected by weather conditions, such as wind and airflow, so nadir imaging cannot be ensured. The reference images are usually geo-rectified satellite or aerial images which provide a nadir view. It is therefore essential to rectify the oblique view of UAV images to a nadir view. A UAV is free to rotate in 3D, thus a rotation matrix R_M can then be determined,

$$R_M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.1)$$

Where α, β and γ are the rotation angles of camera from reference axes x, y and z , shown in Figure 5.2.

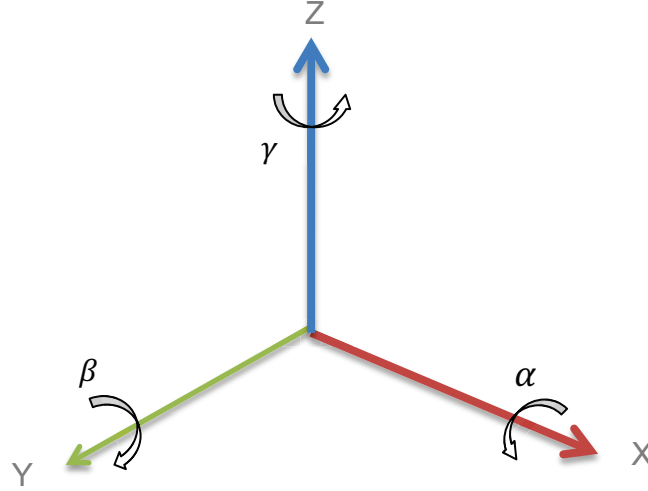


Figure 5.2 3D rotation angles of UAV

The three angles are usually termed as yaw, pitch and roll which can be precisely measured by an IMU (Inertial Measurement Unit) sensor. A perspective transformation via the rotation matrix R rectifies UAV images from an oblique view to the vertical view.

5.2.2 Query and search image generation

After geometric rectification, a UAV image is transformed to the same view as the reference image. To ensure fast real-time navigation, instead of using the whole scene of the UAV image, two query images, q_i and q_{i+1} , are extracted for localisation (i is the current stop number). Here, q_i covers the geometric central area of the UAV image, which corresponds to the current position of the UAV, and q_{i+1} covers the forward view area according to the mission plan, which contains the next position of UAV.

$$\begin{aligned} q_i &= \text{imcrop}(I_u, w_u/2, h_u/2, R_1) \\ q_{i+1} &= \text{imcrop}(I_u, w_u/2 + bx, h_u/2 + by, R_1) \end{aligned} \quad (5.2)$$

Where w_u and h_u are the width and height of the UAV image I_u , and R_1 is the size (number of lines and columns) of the query image. The image crop algorithm *imcrop* crops a square image with size defined by R_1 from the current UAV image at position relative to the image centre. Variables bx and by are determined based on the assumed distance and direction between stop i and stop $i+1$ in the UAV image.

Then, two search images, s_i and s_{i+1} , are extracted from the reference image according to the mission plan and UAV positions for current stop i and the next stop $i+1$.

$$\begin{aligned} s_i &= \text{imcrop}(I_r, X'_i, Y'_i, R_2) \\ s_{i+1} &= \text{imcrop}(I_r, X'_{i+1}, Y'_{i+1}, R_2) \end{aligned} \tag{5.3}$$

where (X'_i, Y'_i) and (X'_{i+1}, Y'_{i+1}) are the planned UAV positions in stop i and $i+1$, and R_2 is the image size.

The search images are larger than the query images, $R_2 > R_1$, to ensure robust image matching results. The relationship between the query and search images and the fly path is illustrated in Figure 5.3; here the query images are shown in red squares and search images in green squares, and the planned UAV flight path is denoted as black solid lines. If the actual UAV path exactly follows the planned path, the query images will be located to the centre of search images. However, if the UAV has a drift from the planned path, the query images will have an offset from the centre of the search image. In other words, the image shifts between the query images and the search images indicate the drift between the planned and the actual paths.

5.2.3 Robust PC based image matching

According to §3.1, the image shifts (a, b) can be directly calculated at integer level via IFT (Inverse Fourier Transform) to convert Q to a Dirac delta function. This allows the design of a fast PC-based scanning approach, in which the query image

is used as a scanning window roaming from top-left to bottom-right of a given searching range in the search image.

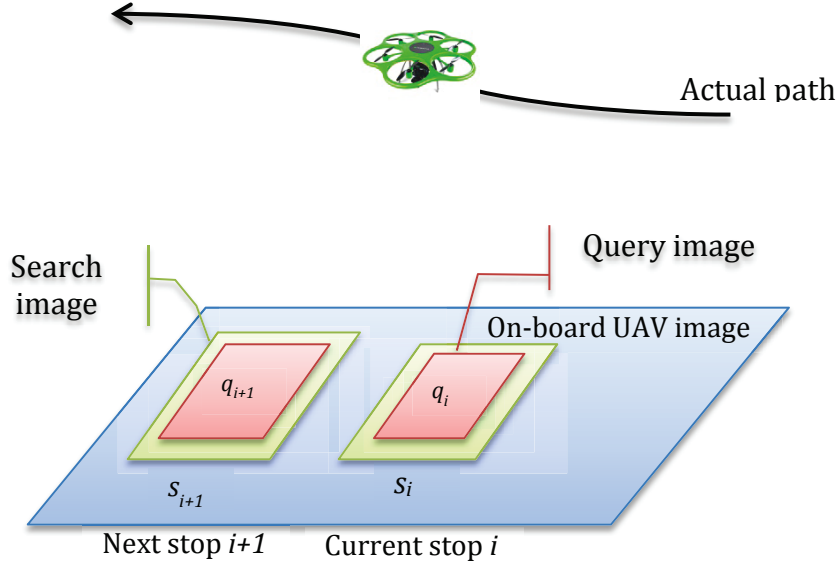


Figure 5.3 Query and search image generation

The total number of image matching N occasions is determined by the size difference between the search image s_i , the query image q_i , and the search interval L as

$$N = \frac{H - h}{L} \times \frac{W - w}{L} \quad (5.4)$$

where H and W are the height and width of s_i , and h and w are the height and width of q_i .

To speed up the matching, the scanning based on the query image is not pixel by pixel, but step by step. The step interval is generally between 30 and 50 pixels. In each step, the UAV position is estimated using fast PC matching. Several UAV candidate positions are calculated at the completion of the PC scanning.

$$[(x_{i1}, y_{i1}), (x_{i2}, y_{i2}), \dots (x_{iN}, y_{iN})] = PC_{sc}(q_i, s_i, L) \quad (5.5)$$

where $(x_{i1}, y_{i1}), (x_{i2}, y_{i2}), \dots (x_{iN}, y_{iN})$ are candidate UAV positions and PC_{sc} is the fast PC scanning algorithm.

The candidate UAV positions estimated via PC scanning may include incorrect positions which are usually of low matching quality. The peak value of Dirac delta function δ_0 indicates the quality of PC matching from 0 to 1. If the two images are identical, the peak value of δ_0 equals 1. Thus incorrect positions can be eliminated by a threshold of δ_0 peak value. The comparison between δ_0 functions of a wrong position and a correct position is illustrated in Figure 5.4 from which, it can be seen that the distribution of δ_0 function resulted from a wrong position match is very noisy without a clear peak, while that from a correct position match has a distinctively high peak value.

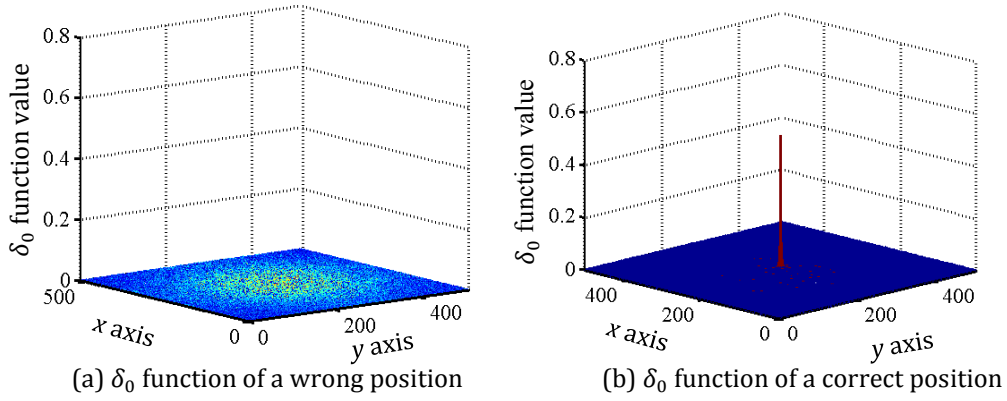


Figure 5.4 δ_0 function based wrong estimation elimination

The final current UAV position (X_i, Y_i) is then calculated as the weighted average of n ($n \geq 1$) candidate positions which have largest δ_0 peak values. The weight is proportional to the δ_0 peak value, i.e. a position which has a larger δ_0 peak value will have a higher weight in the final position estimation.

$$(X_i, Y_i) = \left[\frac{\sum_{j=1}^n \delta_j(x_{ij})x_{ij}}{\sum_{j=1}^n \delta_j(x_{ij})}, \frac{\sum_{j=1}^n \delta_j(y_{ij})y_{ij}}{\sum_{j=1}^n \delta_j(y_{ij})} \right] \quad (5.6)$$

where δ_j is the normalised δ_0 function peak value of the j^{th} candidate position (x_{ij}, y_{ij}) of n candidates and $n < N$. In this paper, n is set as $N/3$.

In the same way, the next UAV position (X_{i+1}, Y_{i+1}) can be calculated using the query image q_{i+1} and the corresponding search image s_{i+1} .

As described above, to ensure robust, reliable and rapid navigation, the proposed algorithm uses a PC-scan approach to generate several candidate positions, followed by a δ_0 function based final position determination. A large search interval is used to shorten the searching time. Within a range of half the scanning window size, the PC algorithm can in fact calculate image shifts directly without roaming. The roaming scanning is intended to make the navigation scheme robust to very large shifts between the UAV and the reference images. In this way, the PC algorithm can tolerate fairly large search intervals.

5.2.4 Position estimation quality assessment

As described in the previous section, the current (X_i, Y_i) and the next (X_{i+1}, Y_{i+1}) UAV positions can be determined using the robust PC matching scheme. Ideally, the calculated UAV positions should match the intended corresponding positions in the planned flight route in the mission map but, in reality, there are always discrepancies between the calculated and planned UAV positions. This is caused mainly by i) inaccurate image matching, and, ii) UAV drift away from the planned flight route. It is essential to evaluate the final matching quality for a UAV position and to accept only those estimated UAV positions which have good matching quality for the risk assessment in the next stage.

From the previous analysis, the peak value of δ_0 indicates the quality of PC image matching for each of the n candidate positions, thus the final δ_0 peak value of the estimated UAV position $\overline{\delta_0}$ can be calculated by averaging the δ_0 peak values of all the n candidate positions from the candidate,

$$\overline{\delta_0} = \frac{1}{n} \sum (\delta_{01}, \delta_{02}, \dots, \delta_{0n}) \quad (5.7)$$

and this average value is used as an assessment of the position estimation quality. If $\overline{\delta_0}$ is larger than a given threshold, the position estimation derived from image matching results is deemed as reliable and can be passed to the next stage for risk assessment and flight adjustment. Otherwise, the estimated positions are not good

enough for flight route refinement but can still be used to carry on the flight to the next round of position estimation.

5.2.5 Risk assessment and flight route adjustment

After matching quality assessment, the estimated UAV positions need to go through a ‘risk’ assessment. The deviations between the actual UAV path and the planned UAV route are calculated in x and y direction for current and next stops as follows

$$\begin{aligned} e_x &= |X_i - X'_i| \\ e_y &= |Y_i - Y'_i| \end{aligned} \tag{5.8}$$

where (X'_i, Y'_i) are the planned UAV position.

If the deviations e_x and e_y are smaller than the given safety distance, the UAV will keep its current flight route. Otherwise, an alert will be issued and the UAV needs to adjust its flight direction based on the deviation vector (e_x, e_y) .

In summary, a PC based approach which achieves fast and robust UAV navigation in illumination changing environment is proposed here. The approach has the following merits compared to other UAV navigation methods. Firstly, it is insensitive to illumination change, especially sun angle change. Even with an afternoon UAV flight time while satellite images are usually acquired in the morning, the UAV position can still be reliably tracked. The image matching is also robust to geometric distortions caused by camera motion. Secondly, as a non-iterative image matching algorithm, PC is able to achieve fast localisation using large search interval instead of pixel to pixel searching. This enables real-time navigation in the case of high speed flying. Moreover, as every PC matching in our approach is independent and can be done in parallel, the speed of localisation can be further enhanced using parallel computing. The approach also estimates the next UAV position which is crucial for target detection and obstacle avoidance. Finally, risk alerts will be issued if the UAV is drifting from the planned path and

the UAV can adjust its positions using the calculated deviation vector between the actual UAV positions and the planned route.

5.3 UAV navigation experiments

5.3.1 DEM simulated images

To verify and test the proposed PC based navigation algorithm, simulated reference and UAV images are generated from the five simulated hill shading images, which has been shown in Figure 4.24, were generated from DEMs under changing diurnal illumination conditions. Figure 4.24(b) was set as the reference image because satellite images are usually acquired around the local times of 10:30 to maintain solar synchronisation. The simulated UAV images are generated by cropping small areas in the Figure 4.24(a), (c), (d) and (e) based on planned flight routes in the reference image. For each hill shading image, 15 simulated UAV images of 600×700 pixels are extracted. In total, there are 60 UAV images along 4 flight routes on the reference image.

Figure 5.5 illustrates the differences between the reference image and the four UAV images of the same area at different times of the day. In the morning, the direction of shading are northwest then gradually changing to north around 12:00 and finally in east in the afternoon at 16:00. The variations of local patterns make it challenging to match the UAV images taken in the afternoon to the reference image taken in the morning.

A UAV, flying at a lower altitude, may acquire images with a higher resolution than the reference images. To investigate the effect of resolution differences on image matching accuracy, we down-sampled the reference images to a series of different scale ratios compared to the UAV images from $1/4$ (0.25) to $1/1$ (1). Here $1/4$ or 0.25 means that 1 reference image pixel covers the area of 16 UAV image pixels. We then tested a simulated UAV path with 7 stops to match the reference images at different scales with the UAV images taken at different time. The localisation accuracy is shown in Table 5.1 and illustrated in Figure 5.6. The localisation

accuracy is within 1 pixel in the scale range from 0.375 to 1, but is degraded considerably at the scale ratio 0.25. In this case, approximately 1.5 pixels localisation accuracy is still acceptable for the UAV images taken at 8:00, 12:00 and 16:00.

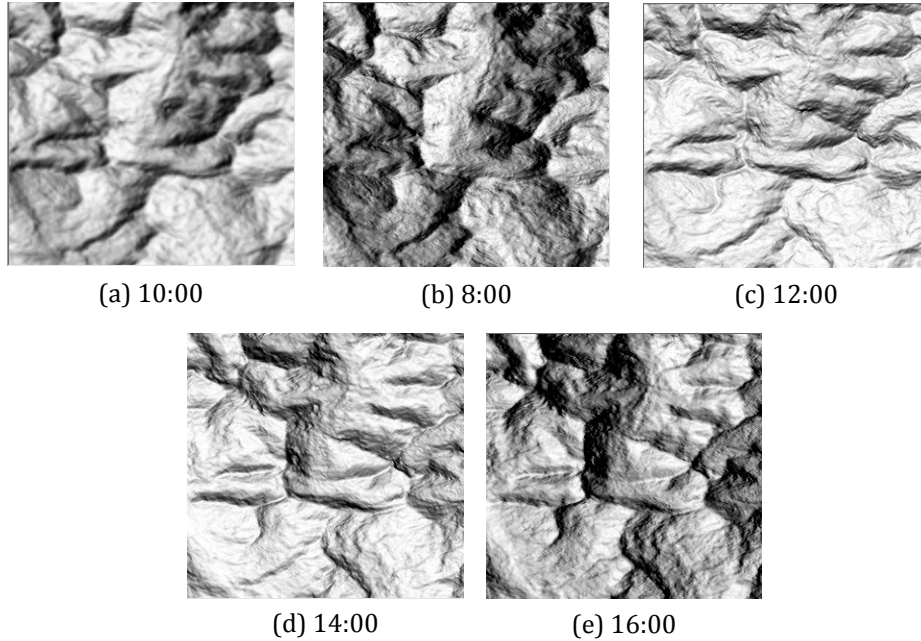


Figure 5.5 Comparison of reference image and UAV images of one same scene: (a) is the reference image (scale=0.5) and other four images are UAV images, acquired at (b) 10:00, (c) 12:00, (d) 14:00 and (e) 16:00.

Table 5.1 Localisation accuracies at different scales (pixel)

Scale Time	0.25	0.375	0.500	0.625	0.750	0.875
08:00	1.49	0	0	0	0	0
12:00	1.52	0.02	0.04	0	0	0
14:00	2.93	1.41	1.11	0.84	1	0.77
16:00	1.58	0.47	0.40	0.39	0.6	0.53

In Figure 5.6, the polylines of 14:00 and 16:00 are consistently above the polylines of 8:00 and 12:00 at all the scales, indicating that 2 hours interval between a reference image and a UAV image ensures high robustness to the scale difference between them. In the following experiment, we set the image scale ratio as 0.5, for which localisation accuracy is largely within 1 pixel. In real situations, the higher resolution UAV images can always be down sampled to match the resolution of a

reference image. The above investigation provides an assessment of the tolerance of scale differences.

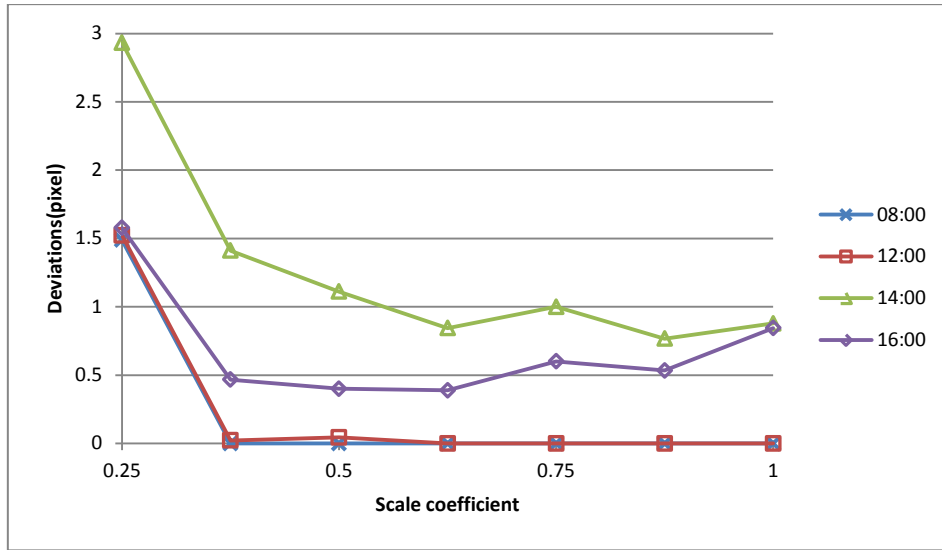


Figure 5.6 Localisation accuracy with the increase of search intervals

An optimal search interval in PC based scanning ensures high processing efficiency without drastically degrading the accuracy of localisation. To investigate this, a simulated UAV path of images taken at 14:00 was tested to match the reference image taken at 10:00 at different searching intervals. The size of the query image is 300×300 pixels and that of the search image 450×450 pixels. As the search interval L is a factor relating to the size of the query image, we set the initial $L = 0.05R_1$, which equals to 15 pixels. The test was then carried out by gradually increasing L to $0.35 R_1$, which corresponds to 105 pixels. The average localisation accuracy in x and y direction and the localisation speed are listed in Table 5.2 and shown in Figure 5.7.

Figure 5.7 indicates that the localisation accuracy is stable within 1.5 pixels until the search interval reaches 85 pixels when the localisation error suddenly jumps more than 5 pixels. This is because when the search interval is very large, only one PC match is available at each stop and so robust location estimation based on multiple candidate positions is not possible. Moreover, with the increased of search interval, the overlaps between query and search images decrease

significantly; PC matching becomes unreliable when the overlap is smaller than $\frac{1}{4}$ size of the matching window. The graph in Figure 5.7 implies that while the search interval is smaller than 75 pixels (~ 0.25 of the query image size R_1), the localisation accuracy is insensitive to search interval variations.

Table 5.2 Localisation accuracy and speed with increase of search intervals

Interval (pixel)	15	25	35	45	55	65	75	85	95	105
Accuracy (pixel)	1.26	1.23	1.32	1.27	1.1	1.3	1.43	6.63	6.63	6.63
Speed (s/stop)	2.26	0.7	0.45	0.27	0.15	0.14	0.1	0.05	0.03	0.03

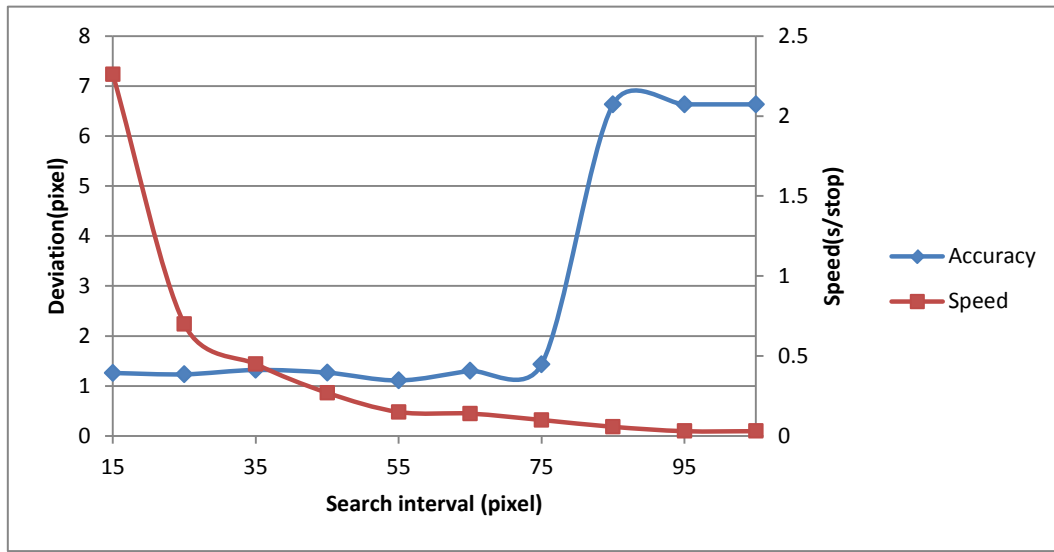


Figure 5.7 Localisation accuracy and speed with the increase of search intervals

The time cost of localisation decreases with increasing of search interval, as a larger search interval means fewer PC image matching operations. When the search intervals are greater than 35 pixels, approximately $0.1 R_1$, the localisation speed is greater than 0.5 seconds per stop. Fast and robust UAV navigation demands a balanced compromise between accuracy and speed and, according to this experiment, the optimal search interval is 45~65 pixels.

The experimental settings for the proposed PC based UAV navigation approach are therefore determined, as shown in Table 5.3.

Table 5.3 Experimental settings for UAV navigation

R_1 (pixel)	R_2 (pixel)	Search interval (pixel)	Scale
300	450	50	0.5

(2) Daily illumination variation experiments

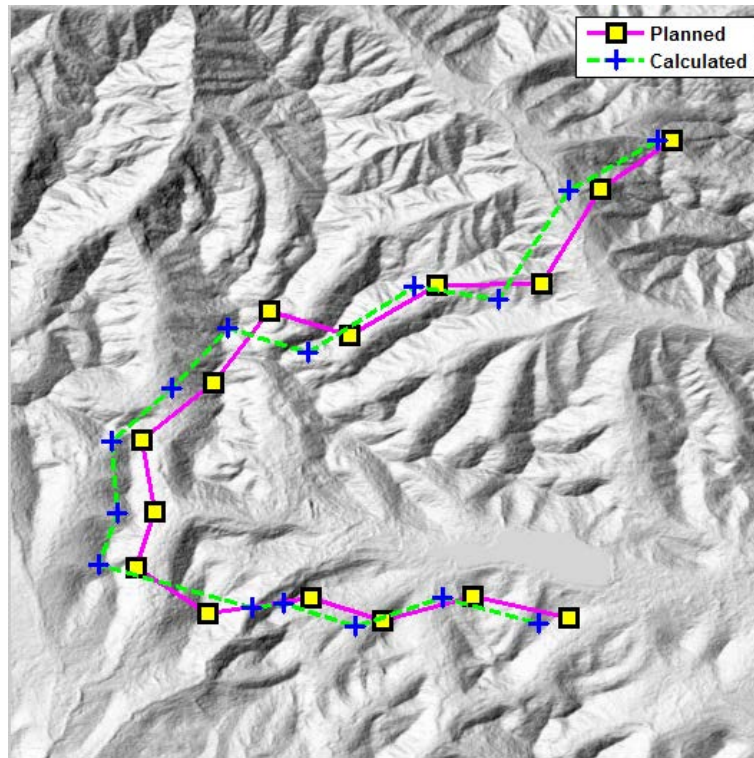
Using the experimental settings in Table 5.3, the robustness of proposed PC based navigation algorithm under daily illumination variations has been tested. To assess the performance of our PC algorithm in comparison with other vision based navigation approaches, experiments were also carried out using NCC and MI algorithms. For fair comparison, the search interval for NCC and MI is set as 5 pixels while PC is 50 pixels. This is because PC is a non-iterative method and has been proven robust to large search intervals, while NCC and MI are iterative matching methods and their performance, in terms of both accuracy and processing speed, declines significantly at large search intervals. Other settings are the same for the three algorithms.

The localisation speed per stop and average accuracy of NCC, MI and PC are shown in Table 5.4. The data show that the localisation accuracy NCC is not acceptable when the time difference between the UAV images and reference image is 4 hours; the average localisation accuracy in x direction is as large as 46.33 pixels, making it unsuitable for navigation, although its processing speed, 1.16 seconds per stop, is about adequate. The MI has higher localisation accuracy than NCC but the error can still be as large as 25.2, and 10.5 pixels in x and y direction respectively in the case of 14:00. The speed of MI is slow, 20 seconds per stop, which is inadequate for real-time UAV navigation. In contrast, PC is robust to different illumination and has the highest localisation accuracy, within 1.5 pixels. The localisation speed is 0.2 second per stop, which is nearly 6 times and more than 20 times faster than NCC and MI. The speed is adequate for real-time navigation of high speed UAVs.

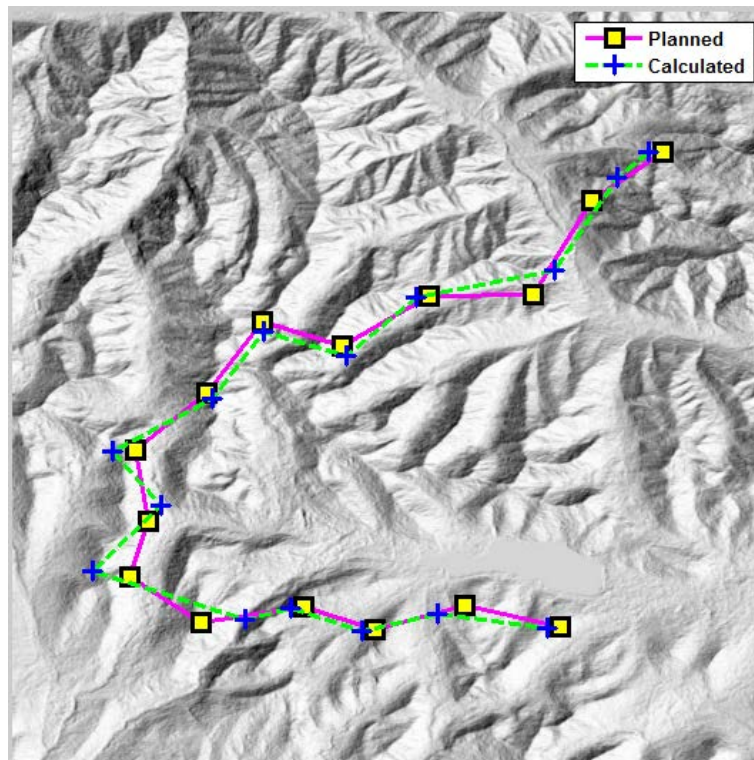
Figure 5.8 illustrates the navigation results of the 14:00 UAV path which is the case of the worst performance of all the three algorithms tested, as shown in Table 5.4. The NCC result shows clear gaps between the calculated path and the planned path. MI performs better than NCC, but there are still large errors at several stops, such as at the first and the last stops. The PC result presents the two paths completely overlapped as the localisation error is no greater than 1.42 pixels. Clearly, the PC based UAV navigation method has a superior performance in comparison with NCC and MI methods for its robustness to daily sunlight angle variation, high localisation accuracy and speed. It is adequate for navigation at high speed UAVs while the other two are not.

Table 5.4 Comparison of localisation average accuracy and speed with different matching algorithms

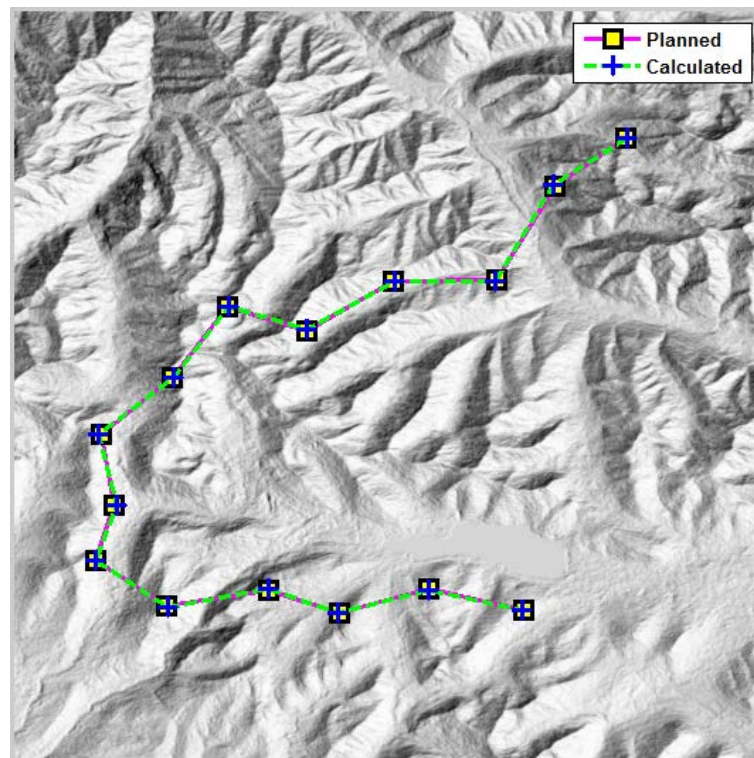
	8:00(pixel)		12:00(pixel)		14:00(pixel)		16:00(pixel)		Speed (s/stop)
	e_x	e_y	e_x	e_y	e_x	e_y	e_x	e_y	
NCC	5.6	5.27	2.80	2.33	46.33	8.41	56.7	15.50	1.16
MI	5.70	5.50	3.33	2.33	25.20	10.46	14.7	8.96	20.52
PC	0	0	0.04	0.04	1.42	0.82	0.47	0.80	0.20



(a)



(b)



(c)

Figure 5.8 Localisation comparison using (a) NCC, (b) MI and (c) PC when UAV flies at 14:00.

In Figure 5.8, the UAV followed the planned flight route exactly only in the case of the PC result but, in reality, UAV may drift away from the planned path. It is essential to monitor the discrepancy between the actual and planned paths and correct the off-course error, as quickly as possible. In the next experiment, we simulated a case that the UAV drifted 25 pixels north and 30 pixels east from a planned route at stop 10, to test whether the proposed navigation scheme is able to detect the flight deviations and adjust the off-course UAV back on-course. The off-course tolerance was set as 20 pixels and, if the deviation from a planned path is greater than the tolerance, a warning will be issued with the calculated deviations by the navigation algorithm. Figure 5.9 presents the experiment result with stop 10 marked by a red circle together with a warning box in which the flight deviation values at stop 10 are correctly estimated and the UAV can then adjust its flight route back to the planned route accordingly.

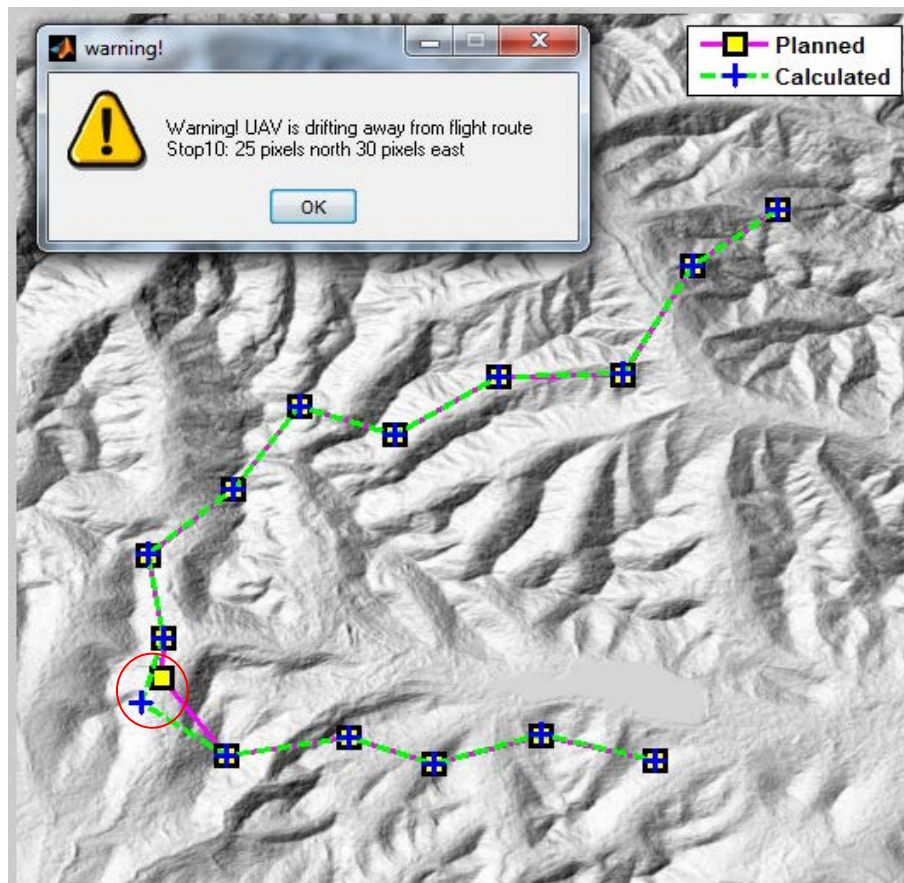


Figure 5.9 Flight deviation observation results.

Furthermore, flight deviations of 25 pixels north and 30 pixels east were added to every stop to test robustness of the flight deviation estimation, and the results are shown in Table 5.5. The flight deviations were quite accurately estimated for all the stops, and the largest error is 1.3 pixels in y direction at stops 5 and 6. The average accuracy for flight deviation estimation is 0.03 pixels in x direction and 0.45 pixels in y direction. The results indicate that the proposed navigation scheme is very effective to control the UAV flight on-course and at high precision.

Table 5.5 Detected flight deviations for every stop

Stop	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x(pixel)	24.3	25.3	25	25	25	25	25	25	25	25	25	25	25	25	25
y(pixel)	29.3	30.3	29	29	28.7	28.7	29.7	30	30	30	30	29	30	29	29

5.3.2 Analogue terrain model experiments

A miniature analogue terrain model of Snowdonia with the surface elevation range of 50 mm was imaged using a Nikon D70s SLR camera, during the morning of 25th May 2014 at 10:30am, and this image used as the reference image in this experiment. The camera was set at a distance of $H=136$ cm and with a tilt angle $\alpha = 17.2^\circ$ from vertical. A series of simulated UAV images were then acquired in the afternoon using an iPhone 5S from a distance of approximately 50 cm, with a nadir view. The proposed PC based navigation algorithm was then tested by localisation of the afternoon smartphone images over the morning Nikon camera image. In this experiment, the nadir view simulated UAV images are rectified to a tilted view, to match the reference image taken by camera. The tested images are shown in Figure 5.10.

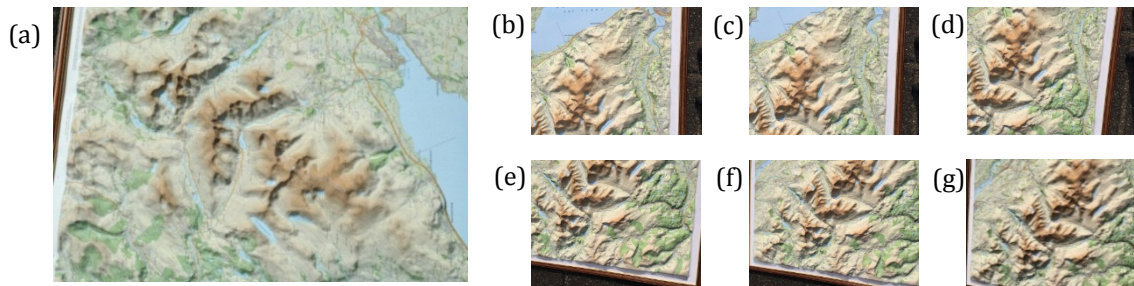
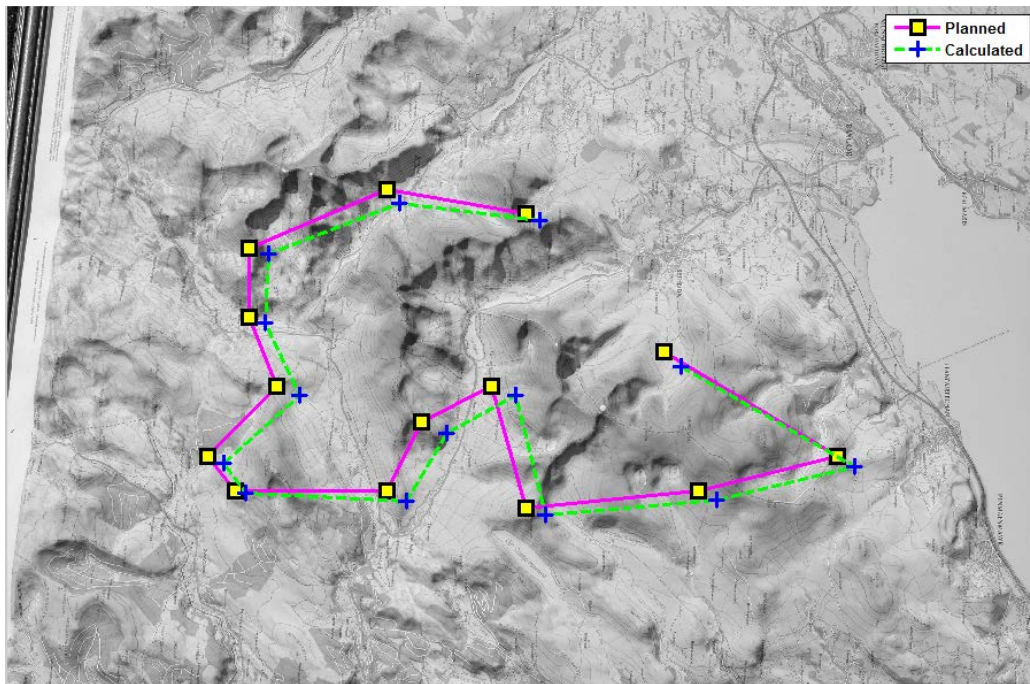


Figure 5.10 (a) Camera image taken in the morning at 10:30 (reference) (b)-(g): smart phone images taken in sequence, in the afternoon at 17:00. The size of camera image is 3008×2000 pixels and of the smartphone images is 3264×2448 pixels.

A flight route with several stop points was defined arbitrarily on the reference image. The aim was to correctly follow this route using the simulated UAV images. For comparison, the PC based algorithm was tested together with MI and NCC algorithms. The size of query image is 500×500 pixels and of the search image is 650×650 pixels; the other experiment settings remain the same as the experiments in §5.3.1. The simulated UAV navigation results of MI, NCC and the PC algorithm are shown in Figure 5.11. This figure demonstrates that neither MI nor NCC is able to accurately match the smartphone images and the reference image under different illumination conditions; there are large deviations between the planned and calculated UAV paths. This is because the change in solar position greatly alters the local grey value distribution, and the correlation coefficients between corresponding areas become low. In contrast, as the PC-based algorithm is insensitive to the local illumination variation, its accuracy of localisation is so high that the two poly-lines representing the planned and calculated paths match precisely.



(a)

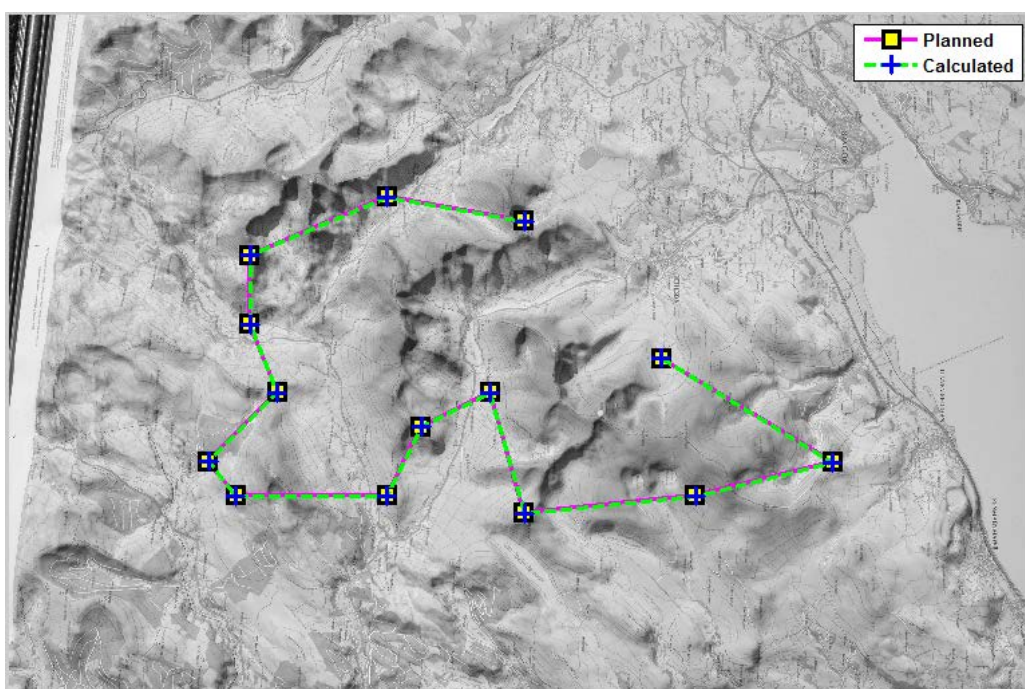
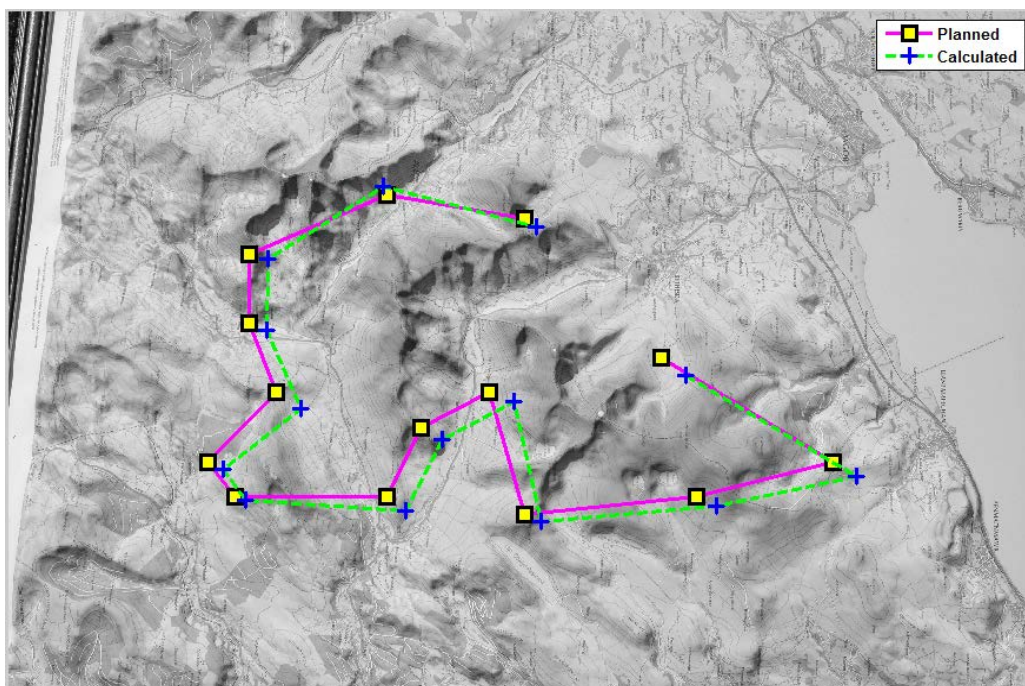


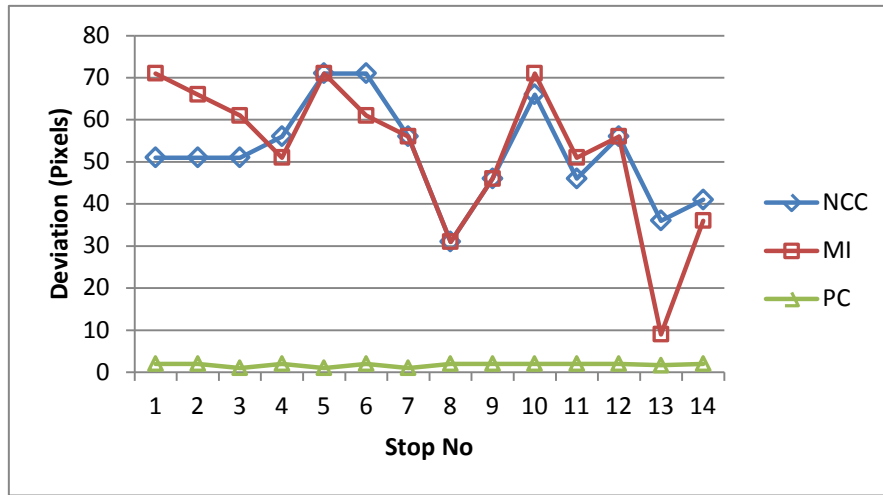
Figure 5.11 Comparison experiment results of UAV navigation produced by (a) MI, (b) NCC and (c) PC algorithms.

The localisation accuracies and speeds of the three methods are compared in Table 5.6 and plotted in Figure 5.12. Compared with MI and NCC, the PC algorithm is very robust and accurate for all 14 stops, achieving average localisation accuracies of

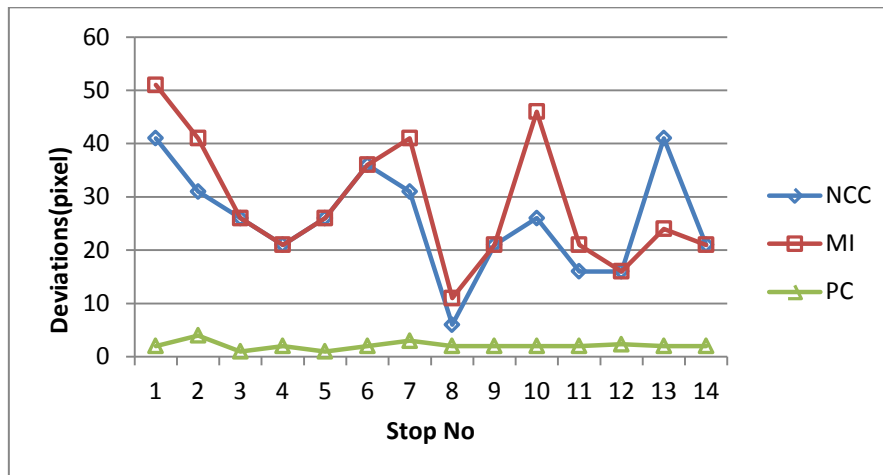
1.76 pixels and 2.09 pixels in x and y direction respectively (Figure 5.11), which are 10-30 times more accurate than the best performance by the other two methods. The speed of the PC algorithm is also much faster than the other two, 0.18 seconds per stop in comparison with 11.46 seconds for NCC and 130.82 seconds for MI.

Table 5.6 Comparison of the accuracy and speed of the MI, NCC and PC algorithm

Algorithm	e_x (pixel)	e_y (pixel)	Speed (seconds/stop)
NCC	52.07	25.64	11.46
MI	52.64	28.71	130.82
PC	1.76	2.09	0.18



(a)



(b)

Figure 5.12 UAV localisation accuracy comparison of MI, NCC and PC algorithms for each stop in (a) in x direction, figure (b) in y direction. The x axis represent UAV stop.

5.3.3 Demonstration of potential applications

(1) UAV navigation on other planet

For the reason that there is no GPS system on other planet, such as Mars, vision based approach is a good alternative for UAV navigation. In this experiment, two HiRISE (High Resolution Imaging Science Experiment) images were used, one as reference image and the other to generate a simulated UAV image sequence. The spatial resolution of HiRISE imagery is approximately 25 cm and they are taken by a push-broom imaging scanner.

The metadata provided by NASA/JPL/University of Arizona (Table 5.7) show that the sun illumination angle difference between the two HiRISE images is 56° in azimuth and 5° in zenith as a result of different Martian daily times and seasons. The illumination differences lead to considerable variation in the direction and length of shadows between these two HiRISE images. Moreover, the shadows in HiRISE images are darker than those in satellite images of the Earth's surface because of the significantly lower atmospheric density on Mars. This adds difficulties in image matching because areas in shadow are likely to become featureless and this making correlation ineffective.

In this experiment, image ESP_012820_1780 was down sampled and used as the reference image, as shown in Figure 5.13(a). A total of 15 UAV images were extracted from ESP_016644_1780, as demonstrated in Figure 5.13(b)-(i). Figure 5.13 reveals that the image textures in this area are spatially repetitive and form highly similar patterns; another challenge for the vision based navigation on Mars. The experimental settings are the same as §5.3.1, shown in Table 5.3. The localisation results are shown in Figure 5.14.

Table 5.7 Observation statistics of the two HiRISE images

HiRISE number	Acquisition date	Local Mars time	Solar azimuth	Solar zenith	Emission angle
ESP_012820_1780	21 Apr. 2009	3:24PM	151.9°	54°	26.1°
ESP_016644_1780	13 Feb. 2010	3:00PM	208.0°	39°	0.9°

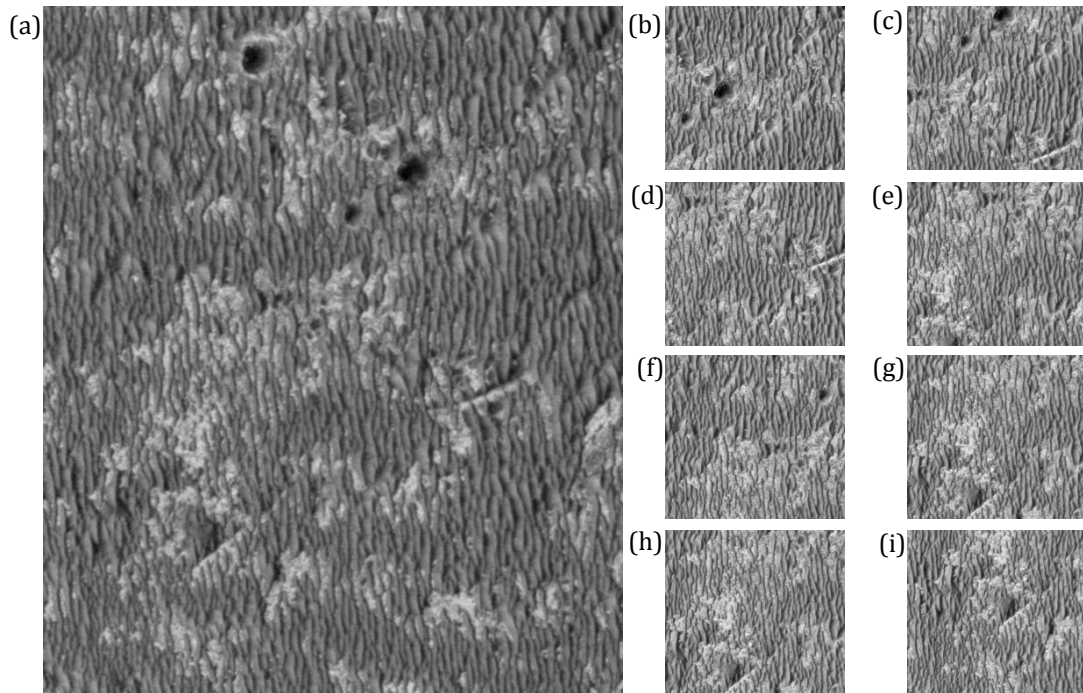


Figure 5.13 (a) Reference image generated from ESP_012820_1780 and (b)-(i) simulated UAV images on Mars generated from ESP_016644_1780. (Reference image size: 1098×1302pixels, UAV image size: 500×600 pixels)

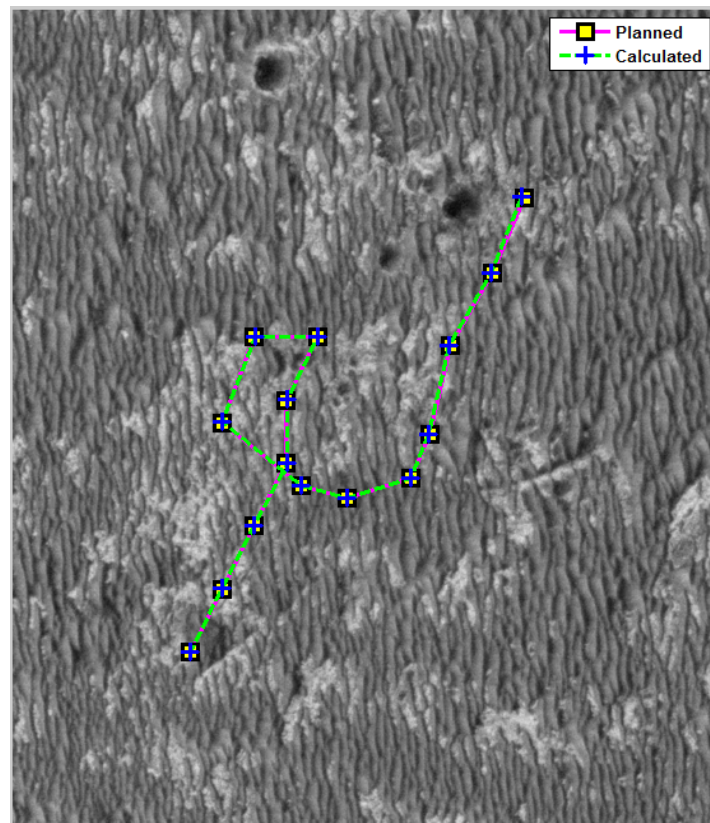


Figure 5.14 UAV navigation result on Mars.

The overall localisation accuracy in the x direction is 1.11 pixels and 0.83 pixel in the y direction, demonstrating a high navigation accuracy. The speed of navigation is 0.12 second per stop. Thus the proposed PC based navigation approach is potentially capable of GPS-denial real-time UAV navigation on Mars.

(2) UAV navigation in Maling Mountain

In this experiment, a real UAV image series is traced on a Ziyuan-3(ZY-3) satellite image. The UAV images, shown in Figure 5.15, were acquired by Gatewing UAV from Trimble at Maling Mountain in Shandong Province, China. The size of the UAV image is 3648×2736 pixels, the imaging time is 11:27 on 01st Feb 2013, the flight altitude is 350 m on average, and the image resolution is approximately 6 cm.

The reference image (Figure 5.16) was acquired by the Ziyuan-3 sensor on board, a Chinese high-resolution remote sensing satellite launched on 9th Jan. 2012. The footprints of Figure 5.15(a)-(c) are also demonstrated in Figure 5.16. The imaging time of the ZY-3 image used is 10am on 15th November 2012. There is a 1.5 hour difference and 2.5 months temporal separation between the UAV images and the ZY-3 image. The resolution of the ZY-3 image is approximately 2.1 m, which means 35 times scale difference between the UAV images and the ZY-3 image. The scale difference is demonstrated in Figure 5.17 by houses seen in the ZY-3 image and in a UAV image. Because of great scale difference, it is very hard to recognise that the two images cover the same area and ground features even after careful visual inspection.

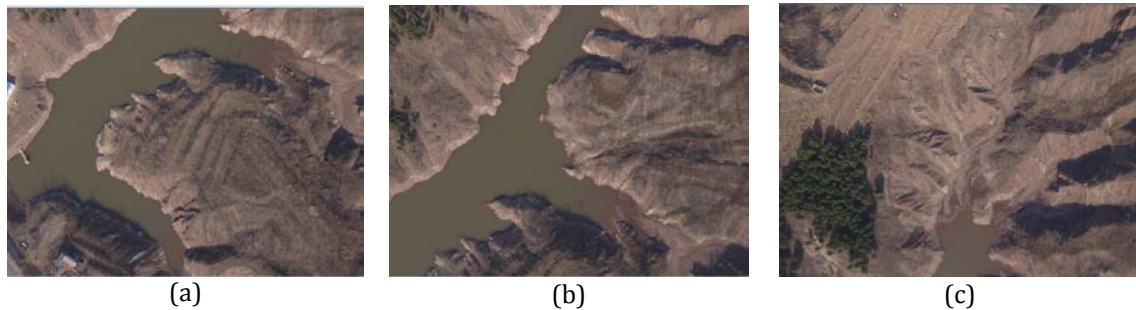


Figure 5.15 (a)-(c) Gatewing UAV images of Maling Mountain, China

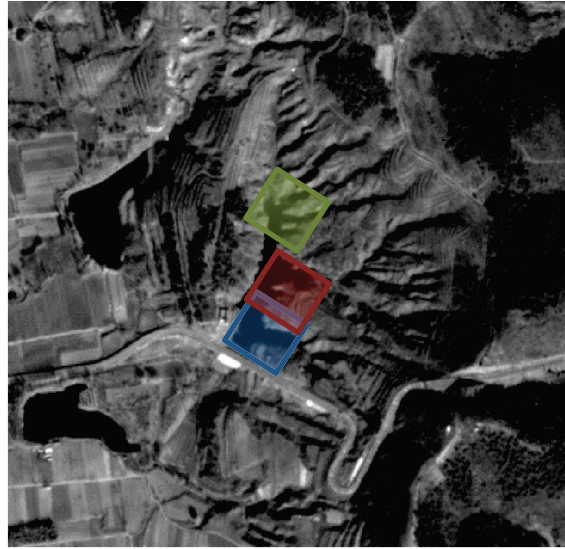
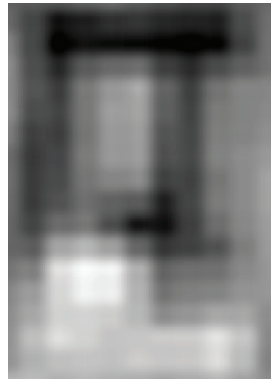
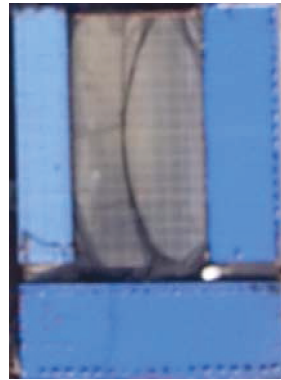


Figure 5.16 Reference image (Ziyuan-3 satellite image). The footprints of Figure 5.15 (a)-(c) images as blue, red and green polygons on this image.



(a) ZY-3 image



(b) UAV image

Figure 5.17 A house seen by ZY-3 image and UAV image

Another challenge is that ZY-3 and UAV are of different imaging geometry. ZY-3 uses a triple array CCD in push-broom imaging mode, while the UAV uses a camera (central projection). Moreover, the UAV as a light, instable and low altitude platform, often acquires images at irregular oblique view angles. The severe image distortion could be fully rectified by pixel-wise co-registration but the processing speed is too slow for real-time navigation. On the other hand, as an area-based approach, the PC algorithm has a certain level of tolerance to geometric distortion and thus can operate with the UAV images roughly rectified to the vertical view using the on-board Position Orientation System (POS). Figure 5.18 illustrates such

an example of geometric rectification converting an oblique view UAV image to conform to the vertical view of the ZY-3 image.

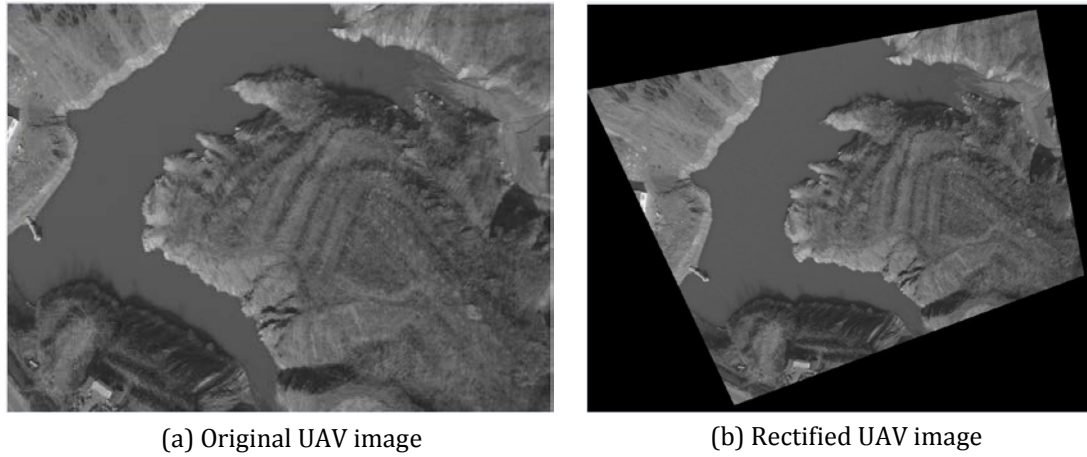


Figure 5.18 Geometric rectification of a UAV image using POS data

After the approximate geo-rectification, 9 UAV images were traced in the ZY-3 image using the proposed approach. The experimental settings are the same as §5.3.1 shown in Table 5.3, and the localisation results are shown in Figure 5.19.

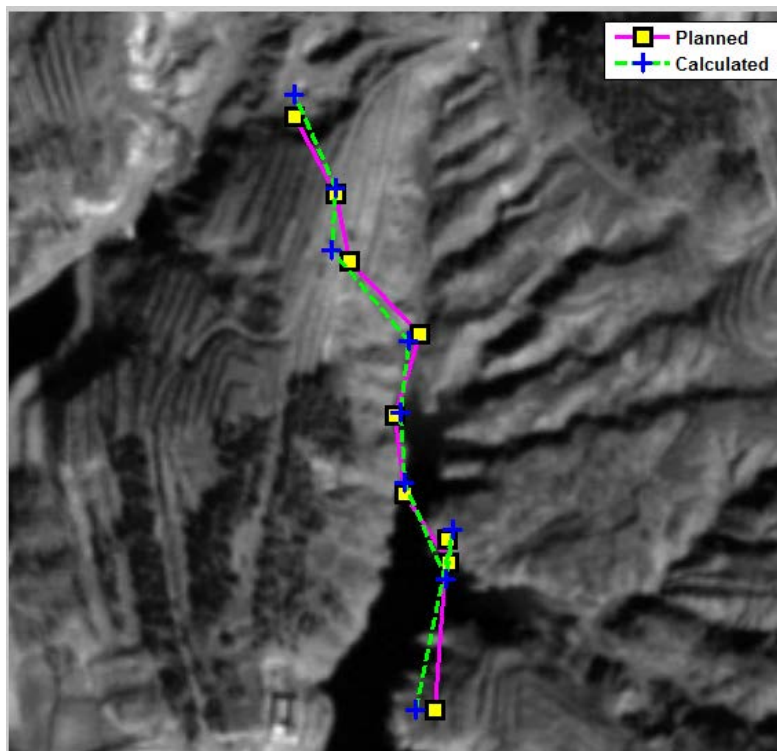


Figure 5.19 UAV navigation result in Maling Mountain.

The calculated and the planned UAV paths in Figure 5.19 do not perfectly overlap but the UAV locations are correct with an average error of 15 pixels from the true positions. The results confirm that even with very large scale difference and being subject to severe image distortion, the proposed navigation approach is still able to find the approximate location of UAV to conduct effective navigation without GPS.

(3) UAV navigation in Cranfield

A Further UAV navigation experiment was carried out using an aerial image from Bluesky and a Geoeye satellite image of Cranfield, UK, as shown in Figure 5.20. This experiment tests the capability and robustness of the PC algorithm for UAV navigation in urban environments. The two images were acquired under very different sun illumination as evidenced by different directions of shadows in the two images; the sunlight is from the SW (South West) in the aerial image and from the SE (South East) in the satellite image.

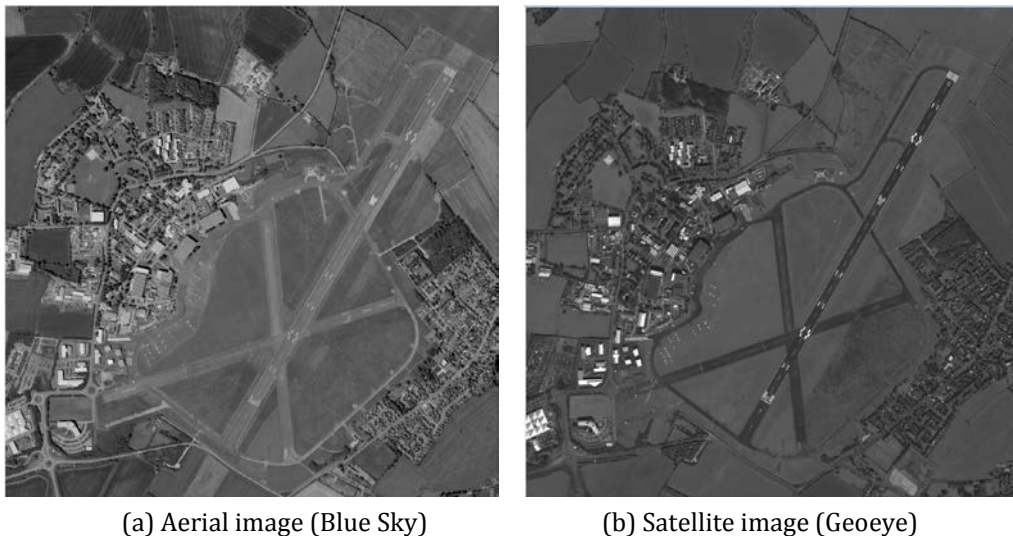
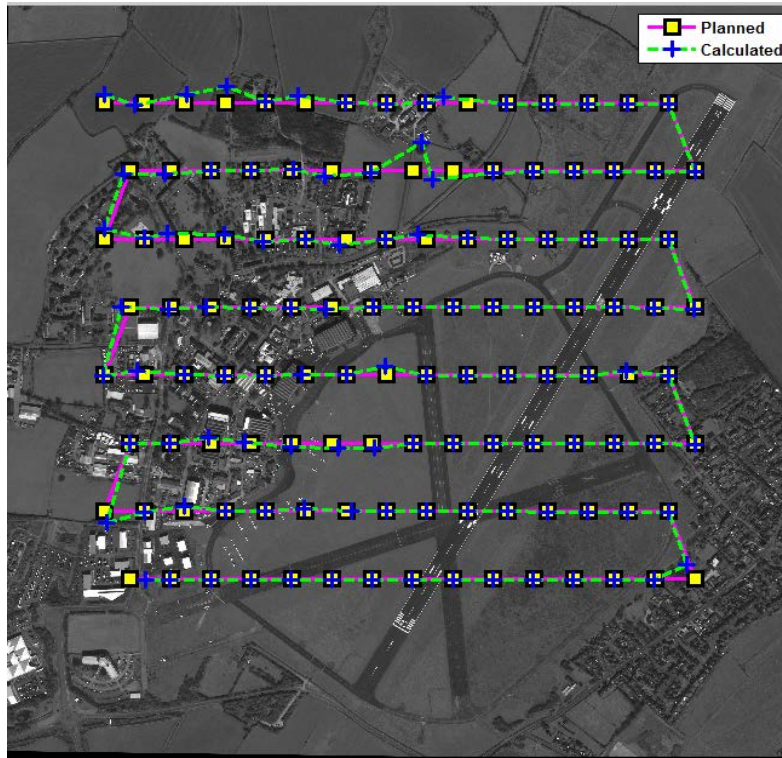


Figure 5.20 Aerial image and satellite image for navigation. (Image size:3930×3922 pixels)

A total of 120 simulated UAV images (1000×1000 pixels) were generated from the aerial image to formulate a typical photogrammetry survey flight route. The overlap between images is 60% in flight direction and 30% between adjacent flight lines.

The UAV navigation results using various R_1 values, heights and widths of query images, are presented in Figure 5.21, using the same notion as before. In Figure 5.21(a), small query images (300×300 pixels) were used and the average localisation accuracy is 8.41 in x direction, and 8.44 in y direction. In this case, the high resolution of the original image means that the small query images are not very robust for reliable and accurate localisation, because if the areas covered by the query images are too small, they are more likely to be featureless. In contrast, if large query images are used (800×800 pixels), as shown in Figure 5.21(b), the calculated UAV position matches the corresponding planned UAV positions precisely for most stops. The average localisation accuracy is 2.71 pixels in the x direction, and 3.16 pixels in the y direction, which equals to 1.24m and 1.45m in x and y directions. The comparison of Figure 5.21(a) and Figure 5.21(b) demonstrates that for high resolution aerial images, the query image needs to be large enough to ensure reliable localisation in high accuracy.



(a) $R_1=300$ pixels



(b) $R_1=800$ pixels

Figure 5.21 UAV navigation results in Cranfield, UK showing the effect of varying query image size on localisation accuracy.

Even using a large query image, there are still a few stops which are subject to the matching failure. The localisation errors caused by matching failures are unavoidable in vision-based navigation for various reasons, such as the lack of distinctive features, environmental changes, new buildings, seasonal changes in farmlands, etc., which may also lead to the low correspondence between the query and the search image. The localisation error propagation, however, can be avoided; otherwise the next UAV positions will all be drifted away affected by one localisation error. It can be confirmed from Figure 5.21 that using the proposed method, the localisation errors will not propagate because every stop is calculated independently. This is one of the important advantages of a frame-reference based approach.

Moreover, the proposed navigation approach is able to detect the matching failures by position estimation quality assessment, described in §5.2.4, and thus the matching failures can be differentiated from the true deviations of UAV flight if it

drifts away from the planned route. In the proposed method, only the candidate positions with high matching confidence will be passed for risk assessment and then flight route adjustment, while the matching failures will not trigger a risk warning. In this way, the effect of matching failures can be reduced to a minimum.

The three examples of Mars, Maling Mountain and Cranfield demonstrate several potential applications of the proposed UAV navigation algorithm. It is reliable and effective for UAV navigation, not only in mountainous areas, but also in urban areas and in terrains showing highly similar and repetitive textures. It is robust to daily and seasonal differences of sun illumination between the real time UAV images and reference images. It can also tolerate large scale differences and severe geometric distortion.

5.4 Summary

This chapter introduces a Phase Correlation based UAV navigation algorithm in a GPS denied environment. Based on the analysis of local illumination effects on the Phase Correlation matrix, we proposed an illumination robust navigation approach. The proposed method is able to determine the current UAV position and to predict the next position. It issues a risk warning if the UAV is off-course from the planned flight route. Five different datasets, including simulated hill shading images, analogue terrain model images and real remotely sensed images, were used to test the robustness, accuracy and computing efficiency of the algorithm in comparison with NCC and MI methods. The PC method is the only one among the three tested which is capable of real-time, high precision UAV navigation based on consistently accurate localisation under various sun illumination conditions.

Chapter 6 Illumination Invariant Change Detection

6.1 Remote sensing change detection

With a growing archive number of repeat coverage earth observation images, change detection based on multi-temporal remotely sensed images has been an active research topic and an important application field. The large volume of remotely sensed images has demonstrated their great potential in long term and short time-lap environmental monitoring. For instance, glacier motion [79, 80], deforestation [81, 82], Land use/cover change [83], landslide localisation [84, 85] and ground deformation measurement [86], river flow from ice debris [87], and sea ice monitoring [88].

6.1.1 Change types

The change information can be categorised into the three types: motion, new object and texture change.

(1) Motion

Motion refers to the position change of an object from one image to another while the appearance of the object remains the same. Object motion takes place on Earth surface from time to time and at different scales and speeds. Tracing these object motions plays an important role in environmental monitoring such as geohazard prediction. Large scale motion includes tectonic plate movement, sand dune shift and cloud motion. Motion can also happen on a small scale, for instance, in vehicle and ship motion, and river debris flow. These small scale object motions can be detected and tracked from space or aerial images owing to the high image spatial resolution and short temporal base-line of data acquisition. Figure 6.1 shows an example of river flow estimation from stereo image pair of ASTER with 15m

ground resolution [89]. The river flow velocities were calculated by motion tracking of ice debris shown in Figure 6.1.

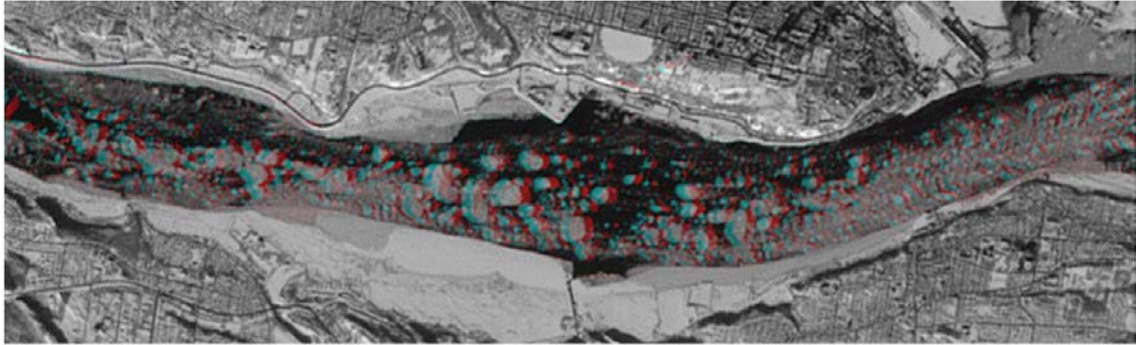
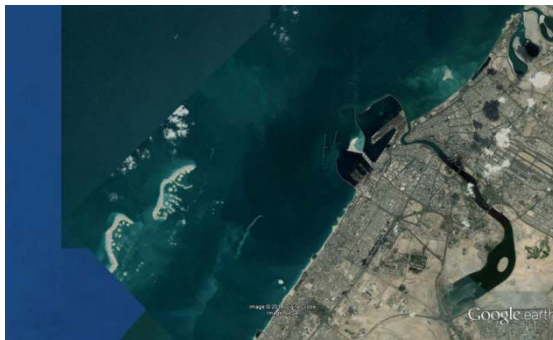


Figure 6.1 An example of river flow from ice debris from ASTER nadir and back-looking images shown by red-cyan modes over Lawrence River, Quebec City, Canada[89].

(2) New object

‘New objects’, refers to those features which have been created either by natural process or human activities. This type of change can be large and significant, like the example in Figure 6.2 showing the urban development in Dubai from 2005 to 2012; and they can be small and difficult to detect, such as locating the position of a rover on the Martian surface in multi-temporal satellite images.



(a) June 2005



(b) April 2012

Figure 6.2 City appearance change in Dubai from June 2005 to April 2012 from Google Earth

(3) Texture change

Appearance change may also be caused by image texture change. An example of image texture change is camouflage that hides a target by using a very similar texture to the background to confuse an observer. Camouflage is a common practice in military operations. Figure 6.3 demonstrates an example of camouflage

using a pair of terrain model images. Though visually similar, one small pattern in Figure 6.3(a) is replaced by another similar pattern in Figure 6.3(b), as revealed by the detailed images in Figure 6.3(c) and (d).

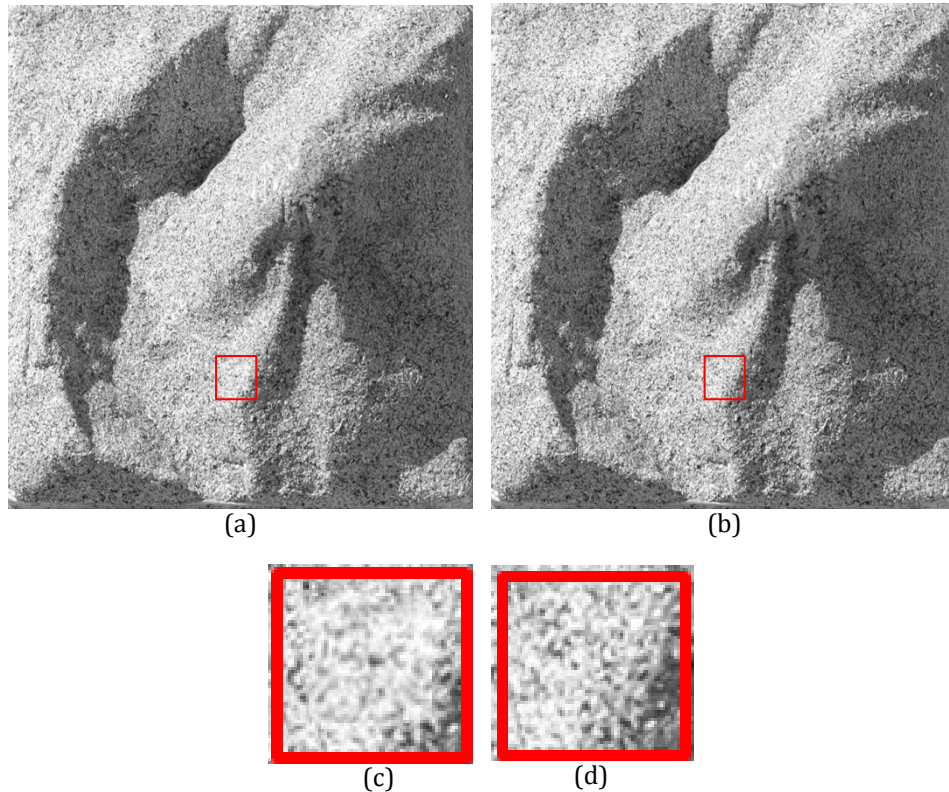


Figure 6.3 An example of camouflaged changes between two images of a terrain model, (c) and (d) are enlarged images of red box areas in (a) and (b).

In summary, *object motion* does not introduce image appearance change, but *new object* and *image texture change* results in image appearance changes at different levels. The divisions between the three types of change are however not always clear. Object motion can sometimes be accompanied by appearance change, for example, a sand dune moves from one time to another but the shape and texture of the sand dune may also be altered.

6.1.2 Problems in remote sensing change detection

The goal of multi-temporal change detection is to extract the pixels which correspond to the above described two types of changes on Earth's surface in an image sequence. Since the remotely sensed images used for change detection are acquired separately, many factors may also cause changes including differences in

spatial resolution and radiometry, illumination variation and camera motion [90]. Thus, the objective of remote sensing change detection is to pinpoint the *real* change of environment while suppressing the above *false* changes caused by observation factors.

This chapter focuses on local illumination variation caused by the changing solar position-driven variations of shading and shadows between the two earth observation images. Significant changes in local patterns, caused by the change in sun angles, may 'hide' the real changes in an image sequence. The study of [91] demonstrated that a 6° difference in solar zenith angle is sufficient to decrease the accuracy of change detection by introducing great radiometric difference, especially in high relief areas.

Several models have been proposed to remove the effects of illumination variation, based on known DEM and sun illumination angles, such as the cosine model [92], C model [93] and rotation model [94]. The correction results largely depend on the accuracy of the DEM that often results in over correction in dark, steep sloped areas [82]. Another solution to minimize the effect of sun illumination, suggested in [90], is to use satellite images acquired on the same calendar date, but this limits the available data for change detection.

Other sensor related distortions may also decrease the accuracy of change detection. As stated in §1.2, if the geometric calibration is carried out based on pixel-wise co-registration, illumination-insensitive dense matching becomes an essential technology.

Thus, the core image matching algorithm for illumination invariant change detection should be able to achieve the following:

- Perform robust and accurate pixel-wise geometric registration if the two remotely sensed images are in different imaging geometry.
- Detect the real change information while suppressing the effects of local illumination change.

- Classify motion and appearance change.

6.2 Related Research

Figure 6.4 presents a generic data processing for most remote sensing change detection. It includes five main steps: data acquisition, pre-processing, change information extraction, change information analysis and final output.

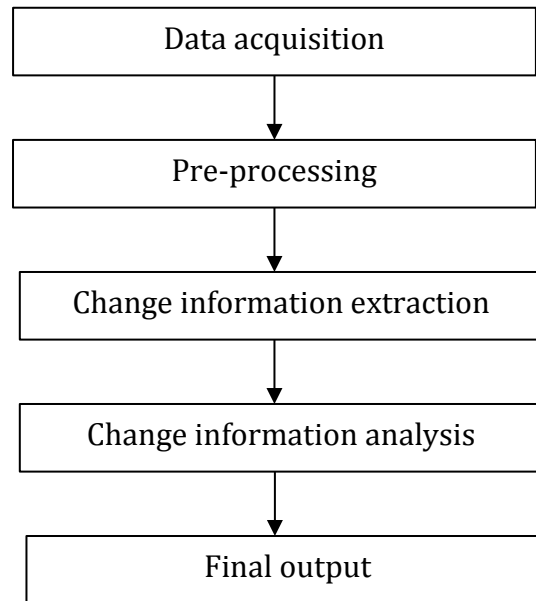


Figure 6.4 Remote sensing change detection process

The change information extraction is the core step in the data processing and some commonly used pixel-wise operations used for this purpose extraction are briefly described here.

(1) Image differencing

Image differencing, the subtraction of multi-temporal image pixel by pixel, produces a new image S that represents the DN value variations between the two images.

$$S = |I_1 - I_2| \quad (6.1)$$

Where I_1 and I_2 are two input images.

The values of unchanged pixels are likely to be near zero while values of changed pixels should be large. A threshold can be set and the pixels with difference value exceeding that threshold will be labelled as 1 in change map C otherwise labelled as '0'.

$$C = \begin{cases} 1, & D_k > thr \\ 0, & D_k \leq thr \end{cases} \quad (6.2)$$

Where thr is the threshold for change.

Image differencing is a simple and effective change detection method and has been applied in a variety of applications [95, 96]. Raw image differencing is however very sensitive to pixel mis-registration and radiometric differences between images [97], as well as to random image noise. Several improved differencing based approaches have been therefore proposed, for example normalised image differencing [97] and radiometrically normalized image differencing. The normalisation image I_{nor_1} is calculated as follows

$$I_{nor_1} = a_k + b_k(I_1 - \bar{I}_1) \quad (6.3)$$

Where a_k and b_k are coefficients and $\bar{I}_1$ is the mean value of I_1 .

As we discussed in §2.1.1, the radiometric distortions caused by local illumination change cannot be removed by normalization based approaches.

(2) Image ratio

Image ratio is another pixel-wise change detection approach, which uses image division instead of subtraction. A ratio image R is produced from two images.

$$R = I_1/I_2 \quad (6.4)$$

The principle of image ratio based change detection is similar to that of differencing. The ratio values of unchanged areas are most likely to be around 1 while the areas subject to change will have values either much higher than 1 or less

than 1 in positive decimals in the ratio image. Again, the change information can be extracted by setting high and low thresholds.

Studies show that the ratio technique performs better than differencing at finding the real changes whilst suppressing the effects of shadowing [1]. As with image differencing, image ratio is sensitive to local illumination variation.

(3) Correlation analysis

Correlation based approaches achieve change detection by finding the least similar pixels. NCC is one of the most widely used similarity measurements for change detection [98, 99]. The correlation coefficient of NCC ranges from -1 to 1, and the largest absolute value of the coefficient represents the highest similarity. Compared to image differencing and image ratio, NCC is less sensitive to global radiometric effects and to mis-registration. As indicated in §2.2.1, local illumination changes will reduce NCC making it unable to detect the real change information.

In this chapter, the ability of phase correlation to achieve change detection in conditions of local illumination variation is explored.

6.3 Principle of Phase Correlation based illumination-insensitive change detection

From the analysis and experiments in previous chapters, there are several merits to Phase Correlation in illumination-insensitive change detection:

- The robustness of the PC algorithm to local illumination variation in Chapter 4 has been proved. The change of sun illumination angle does not change the orientation and density of PC fringes, and thus following the inverse Fourier transformation from the PC cross power spectrum, the Dirac delta function δ_0 peak is distinctive for image shift estimation at sub-pixel level accuracy. This property enables illumination invariant motion detection

- The accurate pixel-wise dense matching under different illumination conditions presented in §4.7.2 has demonstrated the capability of PC based geometric rectification for change detection. The image sequences do not need to be precisely co-registered before applying phase correlation because it is able to calculate the image shift between the matching window in both the reference and target images and then generate a geometrically rectified target image via pixel-wise PC scanning. This means that the pre-processing and change detection stages can be merged into one process.

In this section, the ability of PC to detect the motion and appearance change under illumination change conditions is investigated. An approach to differentiate the change information into motion and appearance change from the PC change detection result is thus introduced.

For change detection, all the pixels in images can be generalised into the four scenarios as illustrated in Figure 6.5, in which the corresponding Dirac delta function δ_0 performance is also presented by examples.

Case 1 No change

When there is no change in two images taken under different illumination conditions, δ_0 has the highest peak value at (0,0) position as indicated in Chapter 4.

Case 2 Large appearance change

When there a large appearance change ($> \frac{1}{2}$ size of the image matching window) in two images, δ_0 becomes very noisy indicating that the two regions in the matching window have a low correspondence rate. In this case, the δ_0 peak value is too low to be distinctive and consequently, the image shift calculated from δ_0 is a random number rather than 0.

Case 3 Small appearance change

When the appearance change is smaller than $\frac{1}{2}$ size of the matching window, the remaining part in the matching windows can still be correlated by Phase

Correlation. In this case, the peak value of δ_0 drops considerably owing to partial de-correlation, caused by small appearance change, but the peak is distinctive enough at the (0,0) location reflecting 0 image shift in x and y direction.

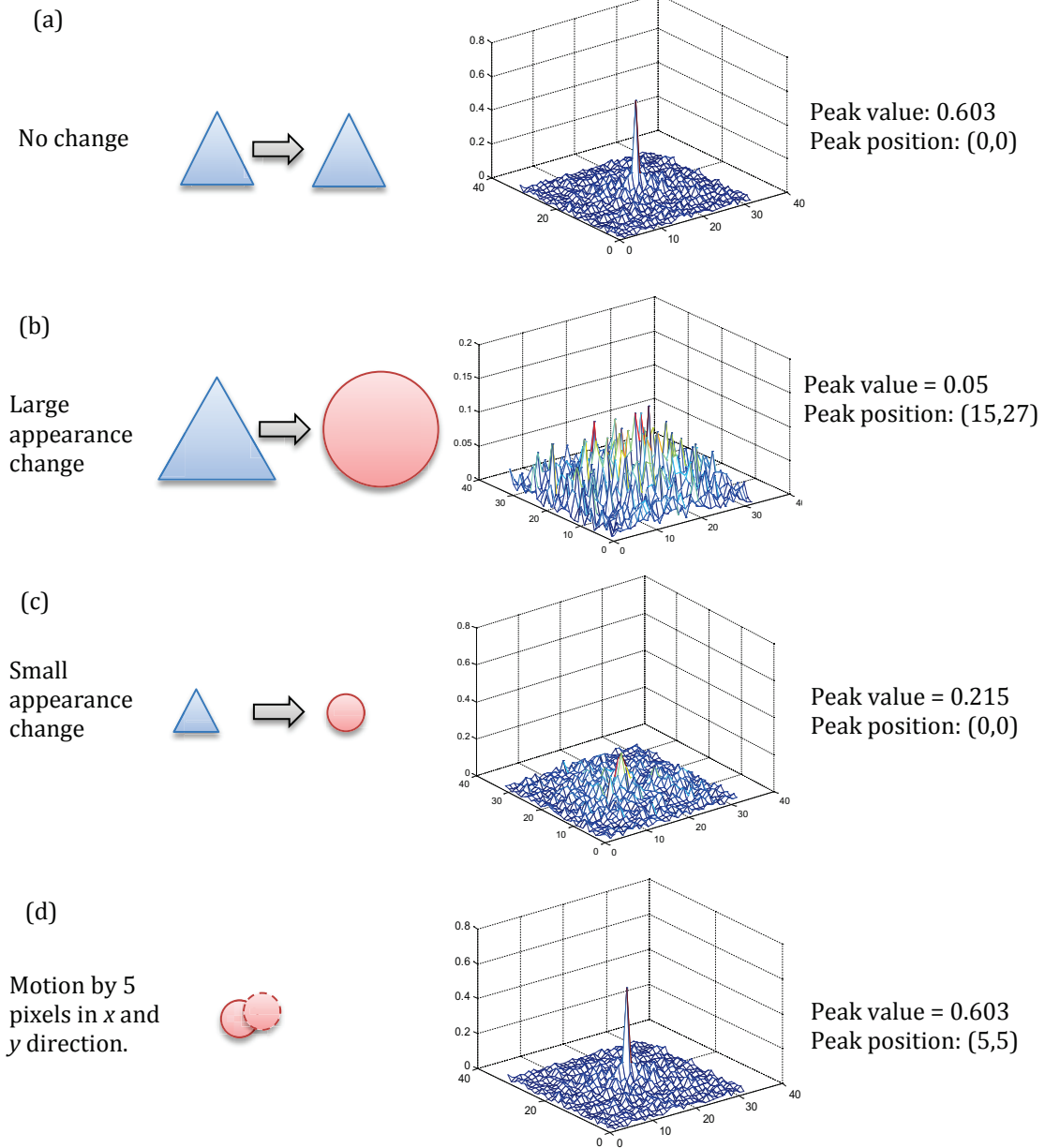


Figure 6.5 Peak values and peak positions of Dirac delta function δ_0 in four different scenarios in change detection.

Case 4 Motion

When there is object motion within the matching window of PC, the related δ_0 will have a clear peak because the appearance of the moving object remains the same and the location of the maximum peak value indicates the object shift in x and y direction.

In summary, δ_0 from PC matching will have a clear peak in (0,0) position when there is no change; the peak value of δ_0 decreases for appearance change depending on the areal size of the change; and the object motion shifts the location of δ_0 peak away from (0,0) position depending on the motion value. The combination of δ_0 peak values and peak locations can therefore detect the change and classify different types of change as characterised in Figure 6.6.

<i>Very Small Disparity</i> <i>High δ_0 Peak</i> <i>No Change</i>	<i>Large Disparity</i> <i>High δ_0 Peak</i> <i>Motion</i>
<i>Large Disparity</i> <i>Very Low δ_0 Peak</i> <i>Appearance Change</i>	<i>Small Disparity</i> <i>Low δ_0 Peak</i> <i>Small Appearance Change</i>

Figure 6.6 Basic criterion for PC based change detection

A pixel will be labelled as ‘no change’ for a very small disparity value with a high δ_0 peak; as ‘motion’ for a large disparity value with a high δ_0 peak; as ‘appearance variation’ for a large disparity value with a very low δ_0 peak, and as ‘small appearance variation’ for a small disparity value with a low δ_0 peak. The last scenario is not entirely unique to small appearance change, it may also be caused by other reasons such as image noise. This special situation will be further discussed in §6.4.4.

6.4 PC-Based change detection algorithm under variant illumination conditions

The process flow chart of PC based illumination-insensitive change detection is shown in Figure 6.7.

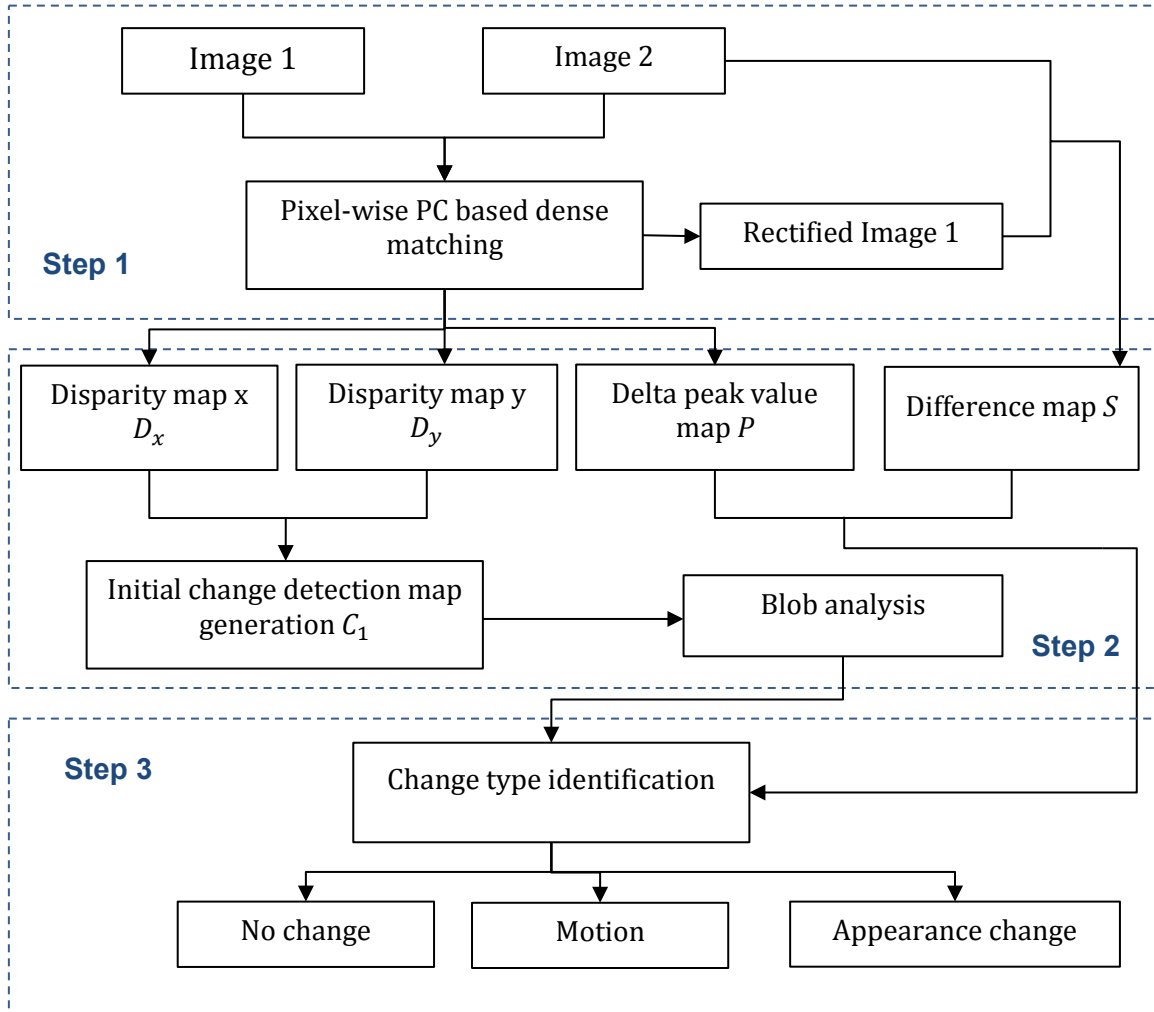


Figure 6.7 Flow chart of PC-based illumination-insensitive change detection

The two multi-temporal remotely sensed images are firstly rectified by pixel-wise PC based dense matching, which results in three outputs: disparity maps in x and y direction, D_x and D_y , and δ_0 peak value map P . Meanwhile, a difference map S is generated by subtraction between the rectified image 1 and initial image 2. The initial change detection map C_1 is generated by segmenting D_x and D_y into several change areas through ‘blob’ analysis. For each change area, change type is

identified based on both δ_0 peak value map P and difference map S . After this, the motion and appearance changes between images 1 and 2 are extracted and labelled. The system is composed of three main steps: pixel-wise PC based dense matching, initial change map generation and change type identification.

6.4.1 Pixel-wise PC based dense matching

In §5.2.1, the geometric rectification is achieved by a perspective transformation based on known 3D attitudes of the image. This geometric rectification can coarsely remove the global geometric distortions caused by camera motion but cannot eliminate the irregular distortions caused by different imaging geometry and other factors. A more precise and rigid geometric rectification is achieved by pixel-wise image co-registration based on the x and y disparities between the corresponding pixels in the two images.

A sequence of multi-temporal images is firstly aligned by Phase Correlation based image matching, and the matching window is usually the whole image. Then a pixel-wise co-registration [12] is carried out by a small corresponding window pair that scans throughout the images, pixel-by-pixel, to estimate the disparity values in x and y directions, $D_x(i, j)$ and $D_y(i, j)$, between the window's central pixel $I_1(i, j)$ in image 1 and pixel $I_2(i, j)$ in image 2, as shown in Figure 6.8.

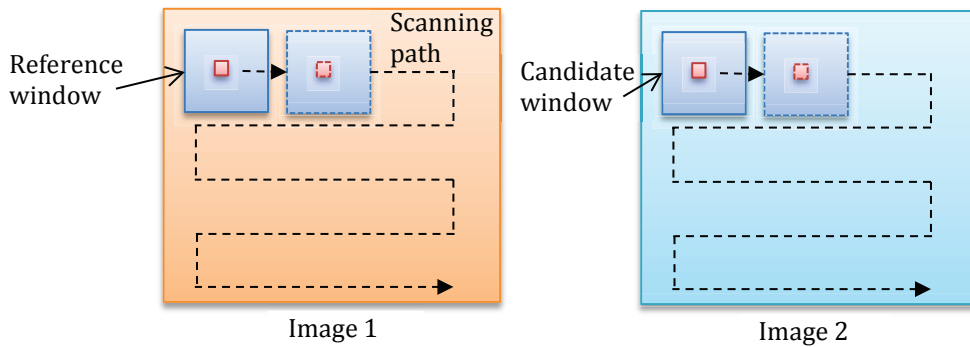


Figure 6.8 Pixel-wise PC based dense matching

Meanwhile the corresponding delta peak value $P(i, j)$ is also calculated through δ_0 based PC matching. Finally, disparity maps D_x and D_y are generated to rectify image 1 to the reference image 2 pixel-to-pixel. After rectification, a difference

image S is generated by taking the absolute values of subtracting the rectified image 1 from the original image 2. The final outputs of pixel-wise PC based matching are two disparity maps, D_x and D_y , a δ_0 peak value map P and a difference image S .

6.4.2 Initial change map generation

As the flowchart in Figure 6.9 shows, an initial change map C_1 is generated in three key steps: binary change map generation based on thresholding, morphological closure, and determination of areas subject to change based on blob analysis.

As discussed in §6.3, disparity values are small for areas of no change, approximately to 0, and are large in case of motion or large appearance change. Thus the change map C_1 (Figure 6.9(c)) can be generated by extracting the large values in two disparity maps, D_x and D_y as shown in Figure 6.9(a) and Figure 6.9(b).

$$C_1(i, j) = \begin{cases} 1, & D_x(i, j) > T_d \cup D_y(i, j) > T_d \\ 0, & \text{else} \end{cases} \quad (6.5)$$

where T_d is the threshold for disparity map.

Given that the large values in disparity maps only indicate large appearance change or motion, small regions in change map C_1 need to be either merged into large areas or removed. A morphological closing is applied to fill the gaps between small regions in C_1 , and the result is illustrated in Figure 6.9(d). Compared to the initial binary map shown in Figure 6.9(c), some small white patches have been eliminated in Figure 6.9(d). The statistics (size and shape) of every ‘blob’ region in the change map C_1 are then calculated, and the ‘blob’ areas with region size smaller than the threshold T_b are not considered as areas of change. In this way, the isolated noise pixels in the change map C_1 are removed. The selected ‘blob’ regions are then defined by red rectangles in Figure 6.9(e). The boundary rectangle of each ‘blob’ is expanded 10 pixels from the outmost pixels for later analysis.

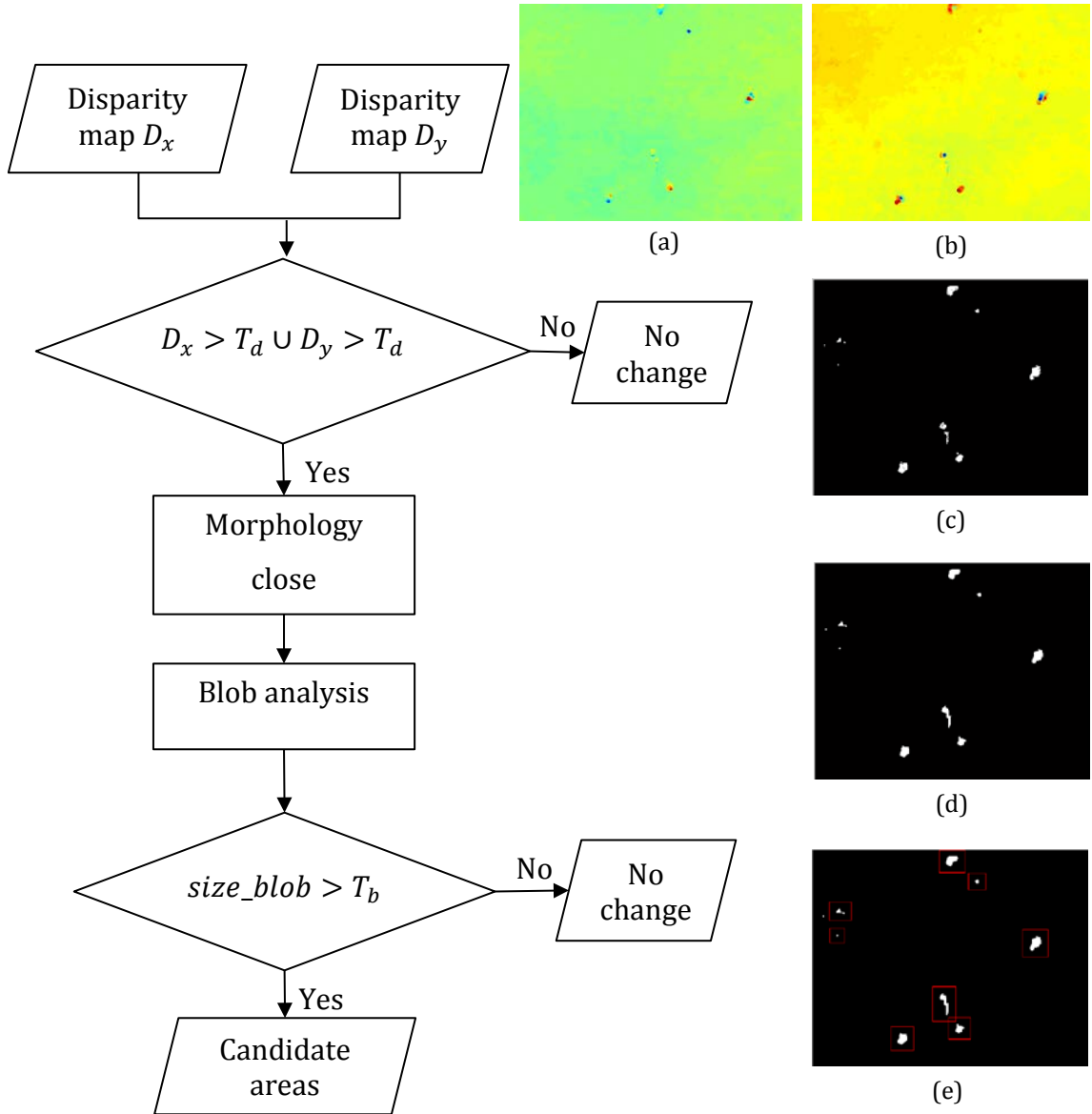


Figure 6.9 Flow chart of initial change map generation: (a) and (b) Disparity map in x and y direction respectively; (c) The binary change map generated by thresholding of disparity, in which white areas indicate 'change' against black background for 'no change'; (d) The change map after morphological closure; and (e) Candidate changed areas marked by red squares determined by blob analysis.

6.4.3 Change type identification

From the previous step, an initial change detection map has been generated with several candidate areas marked by red boundary boxes, as shown in Figure 6.9(e). The next step is to remove the false change areas based on δ_0 peak value map P and difference map S and, then, classify the candidate change areas into three different change types: motion, new object and texture change. For each candidate change region, the change type determination process is shown in Figure 6.10.

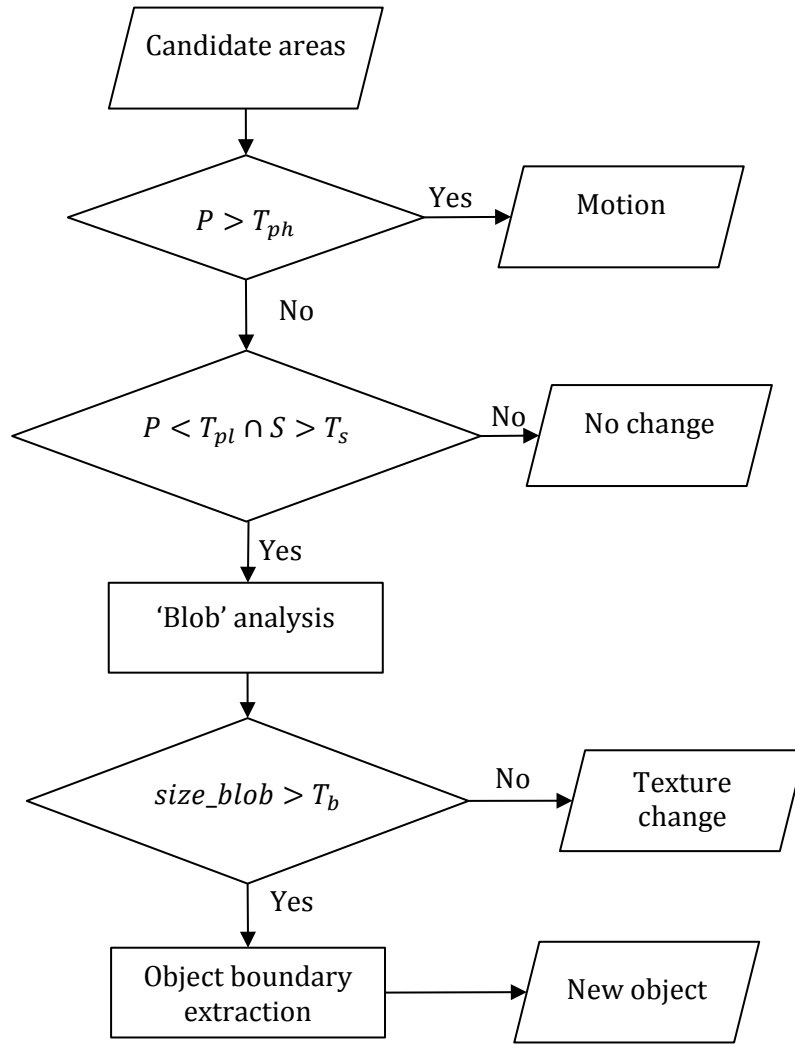


Figure 6.10 Flowchart of change type determination

A high-low thresholding scheme is proposed to differentiate ‘motion’ and ‘appearance change’. According to the analysis in §6.3, ‘motion’ produces high δ_0 peaks while ‘appearance change’ produces low δ_0 peaks. Two thresholds, T_{ph} as high threshold and T_{pl} as low threshold are introduced for the change type identification.

Motion identification is carried out as follows: in the current candidate area, the pixels with δ_0 peak values larger than the high threshold T_{ph} will be marked. If the marked pixels in the candidate areas account for more than 80%, this area will be labelled as ‘motion’. The reason for motion identification in the first step of the

approach, outlined in Figure 6.10, is that object motion leads to false appearance change around the moved object and to the so called ‘hallo effect’ [100]. It is assumed that every candidate region of change only contains one change type, and thus if the candidate change region is assigned to ‘motion’, it cannot proceed to the next stage and has no chance of being labelled as ‘appearance change’.

For the candidate areas which do not belong to ‘object motion’, the following judgment is carried out to check whether they belong to ‘appearance change’ or not. If the pixel (i, j) in the candidate area meet the following requirement

$$P(i, j) < T_{pl} \text{ and } S(i, j) > T_s \quad (6.6)$$

where T_s is the threshold for difference map, it will be marked. If the total number of marked pixels is more than 80%, this candidate area will be labelled as ‘Appearance change’.

The selection of T_{ph} and T_{pl} has a direct effect on change detection results. If T_{ph} is set a very high value, for example 0.95, the identified ‘motion’ areas will be limited. However, if T_{ph} is set a low value, some ‘appearance change’ areas may be assigned to ‘motion’. The same principle works for the selection of T_{pl} . The optimal selection of T_{ph} and T_{pl} varies between different EO data, while in this thesis, the suggested value of T_{ph} and T_{pl} are 0.8 and 0.3 respectively.

Finally, we differentiate change types of ‘new object’ and ‘texture change’ (e.g. camouflage) based on the ‘blob’ size of the change detection map generated according to Equation (6.6) and determine precise boundary box of change areas. A new ‘object change’ area in the change detection map is a compact and continuous patch with clear boundary, as shown in Figure 6.11(b) while for ‘texture change’(Figure 6.11(c)), the change detection map contains scattered small regions (Figure 6.11(d)) without obvious boundary. Through ‘blob’ analysis, if a ‘blob’ size is larger than the threshold T_b , this candidate area will be marked as ‘new object’ and a new boundary box will be drawn based on the ‘blob’. Otherwise,

the candidate area will be marked as ‘texture change’, and the boundary box will remain the same as determined in the initial change map generation.

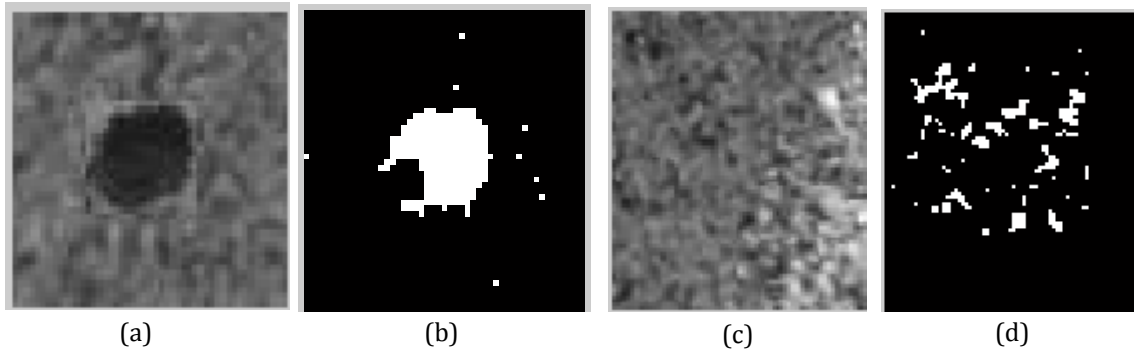


Figure 6.11 Precise boundary identification for new object and texture changes determined by logical combination of difference image and δ_0 peak map: (a)-(b) the new object appearance change and related binary change detection image generated; (c)-(d) the texture change and the related binary change detection image.

6.4.4 Small object change detection

For change detection using multi-temporal remotely sensed images, small object change detection can be a challenge because the images may be subject to significant variation in illumination conditions that overwhelm the subtle change information of small objects. An example shown in Figure 6.12, where images (a) and (b) are under different illumination conditions, and (b) contains a small object (5×6 pixels) that is not presented in (a). The textural patterns of the two images are altered greatly by the change of sunlight angles, and thus it is difficult to pinpoint the real change (the object) from changes caused by illumination variation.

According to the analysis in §6.3, changes of small object type does not produce large values in the disparity maps because the remaining areas in the matching window are without change and can still be correlated. As shown in Figure 6.13(a), the disparity map generated by PC dense matching between Figure 6.12(a) and (b) is smooth and flat without large (distinctive) values and cannot indicate small object changes.

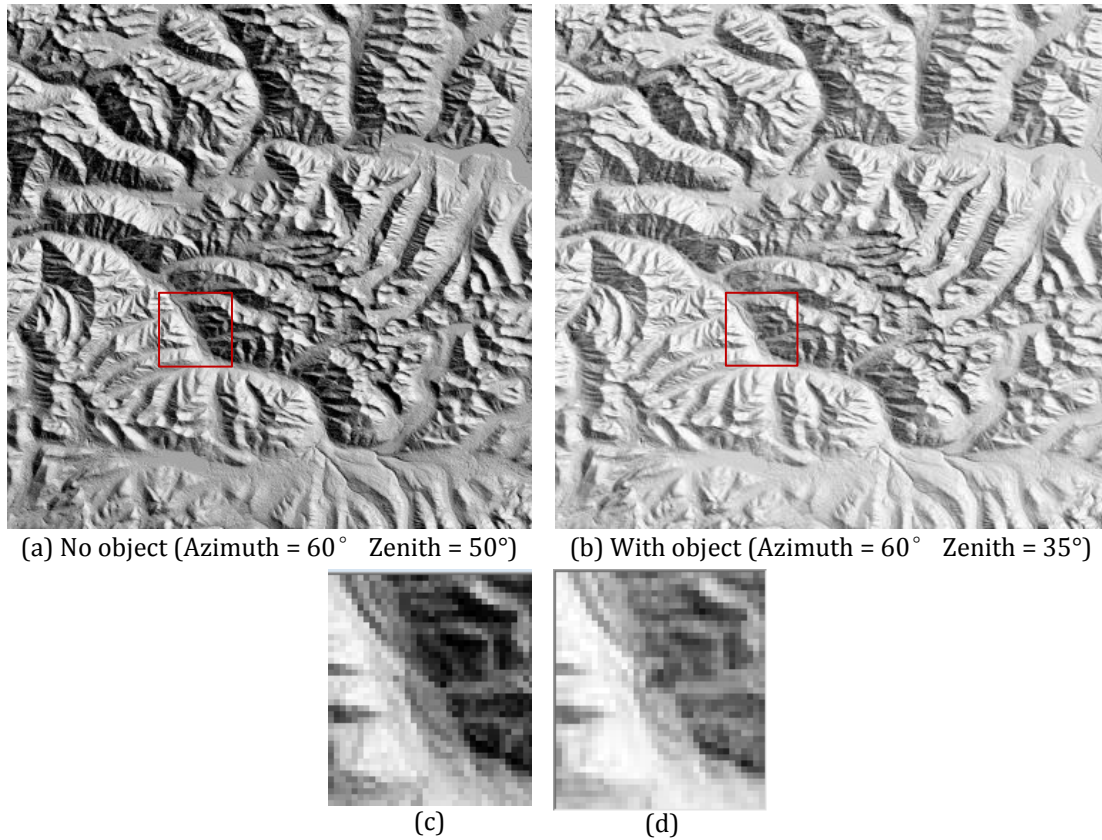


Figure 6.12 A case of small object change detection in an illumination change environment; (c)-(d) are sub-scene images of areas marked by red boxes above.

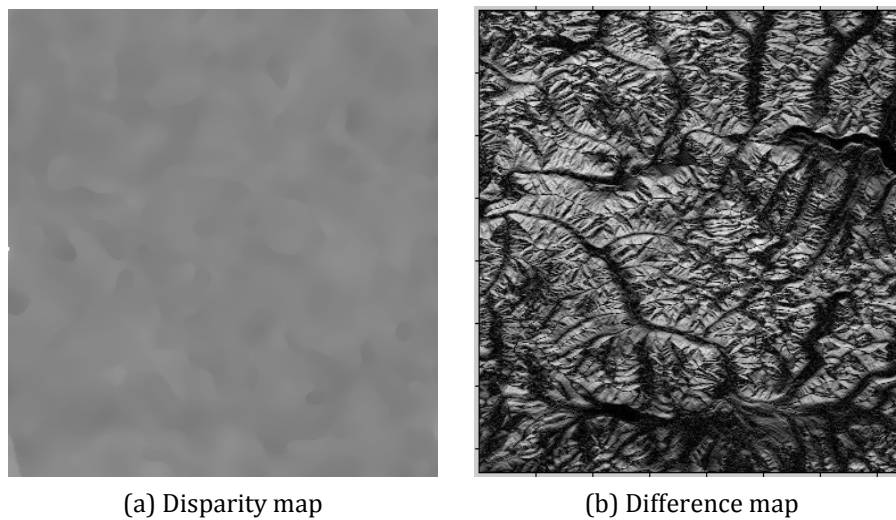


Figure 6.13 Disparity map and difference map of Figure 6.12(a) and Figure 6.12(b).

The difference map between Figure 6.12(a) and Figure 6.12(b) is shown by Figure 6.13(b). Although small object change can produce large values in the difference map, the phenomena can be easily overwhelmed by the changes of spatial patterns,

such as shades and shadows, as the result of the variation of illumination conditions. Thus, the difference map alone cannot distinguish between the small object change and illumination change.

Indeed, small object changes cause the decrease of δ_0 peak values because the changes reduce the similarity rate of the matching areas. Many other reasons may also cause the declination of δ_0 peak values, such as image noise, image geometric distortions, severe illumination variation etc. A small object change detection approach based on a combination of δ_0 peak map and difference map, is therefore proposed. Specifically, the small object change detection is based on the following assumptions:

- The small object is the only change source. If there is a large area of change within the study area, δ_0 value change caused by the large area appearance change will be much stronger than the value change caused by small object change. The change information caused by a small object would be negligible compared to the large appearance change.
- The size of the small object is roughly known. By judging the size of changing areas, some irrelevant candidate areas can be eliminated.
- The illumination variation is limited. Since the imaging time of satellite images is fixed, the seasonal variation of solar azimuth and zenith angles is not very significant. According to the analysis of the existing satellite images, the azimuth and the zenith angle differences are no greater than 90° and 30° respectively in middle latitude regions on the earth.

Based on the above assumptions, the proposed small object change detection approach is outlined in the flowchart in Figure 6.14.

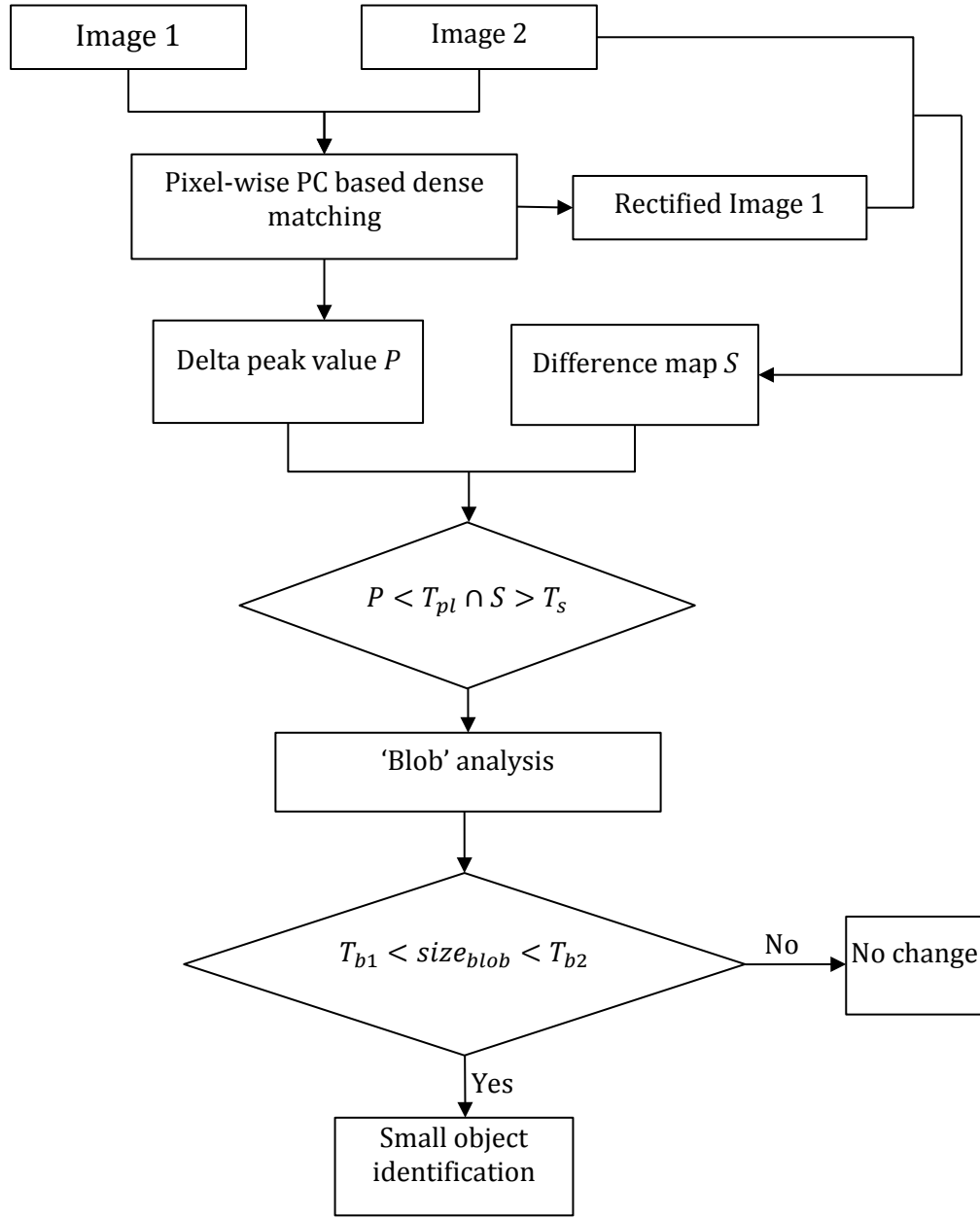


Figure 6.14 Flow chart for small object change detection

After the pixel-wise PC based dense matching, as described in §6.4.1, the δ_0 peak value map P and the difference map S are generated. According to the previous analysis, the small object change areas have low values in P and high value in S , and then the initial change detection map C_1 is generated as:

$$C_1(i,j) = \begin{cases} 1, & P < T_{pl} \cap S > T_s \\ 0, & \text{else} \end{cases} \quad (6.7)$$

Then, ‘blob’ analysis is performed on the change detection map C_1 . Based on known object size, the ‘blob’ areas not matching the size within a particular range will be removed. ‘Blob’ analysis is a key step in the small object change detection approach, which filters out the large dark shaded areas and small isolated patches of image noise, and the final remaining ‘blobs’ should be highly relevant to small object changes.

6.5 Experiments

6.5.1 DEM hill shading image

In this section, DEM hill shading images are used to test the robustness of the proposed PC based illumination-insensitive change detection approach. The generation of hill shading images has been described in detail in §3.3. All four types of change were tested, including appearance change, 3D terrain elevation change, object motion and small object change.

(1) Appearance change detection

In this experiment, images of a case of 2D appearance change are generated, as shown in Figure 6.15. The two hill shading images are under different illumination conditions ($\sigma_1=50^\circ$, $\tau_1=60^\circ$ and $\sigma_2=50^\circ$, $\tau_2=240^\circ$) and a black square of 50 pixels in width and height was placed in Figure 6.15(b) marked by a red circle. Due to the azimuth angle difference, the topographic shades and shadows vary significantly between the two images in Figure 6.15.

The first step, i.e. the pixel-wise PC based dense matching using a matching window of 32×32 pixels, produced two disparity maps in x and y direction as shown in Figure 6.16(a) and Figure 6.16(b) respectively.

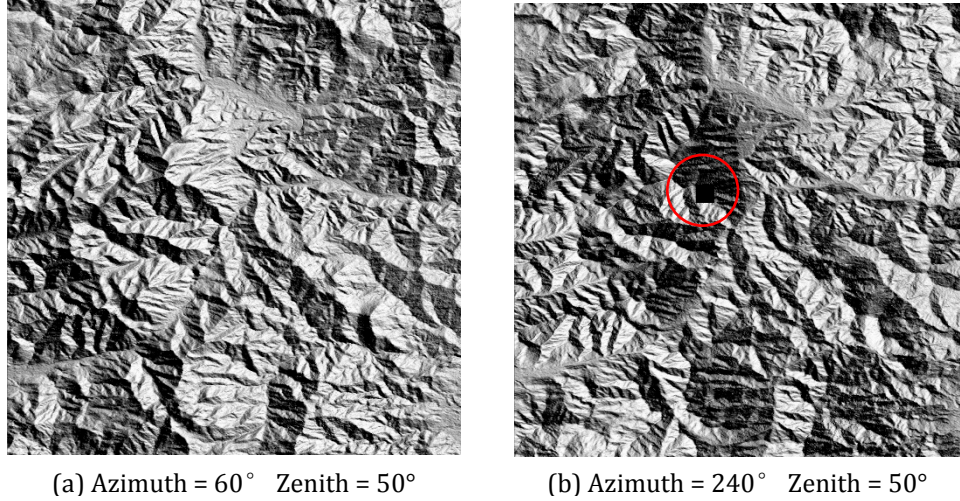


Figure 6.15 Appearance change dataset under different local illumination conditions. The appearance change area is marked by a red circle. (Image size: 1024×1024pixels)

According to §6.3, the appearance change areas would have large values in the disparity maps while the unchanged area very small values, approximately zero. It can be seen in Figure 6.16 that the appearance change areas have distinctively larger values than other areas but it can also be seen in Figure 6.16 that the disparity values of some unchanged areas are not close to zero; these are shown as some fluctuations in the background areas of Figure 6.16(a) and Figure 6.16(b). This is caused by the low correspondence, as a result of illumination variation. Although the illumination invariant property of PC has been proved in Chapter 4, the image matching correct rates decline with the decrease of matching window size, according to the image matching experiment in §4.7.2. For a small matching window with limited image spatial information, the matching performance degrades and the disparity values may not equal zero for areas without true changes. This lost correspondence phenomena, caused by illumination variation appears randomly in the disparity maps, and does not cause large areas of discontinuities as caused with true appearance change. The appearance change areas can thus be detected by extracting the areas populated with large disparity values; then even the local image texture patterns change greatly as a result of the azimuth angle variation.

A threshold of $T_d = 5$ pixels was then applied to the initial change detection map Figure 6.16 to produce Figure 6.17(a). In the 'blob' analysis, any 'blob' smaller than the threshold T_b were removed and any adjacent 'blobs' were merged using morphological closure as shown in Figure 6.17(b).

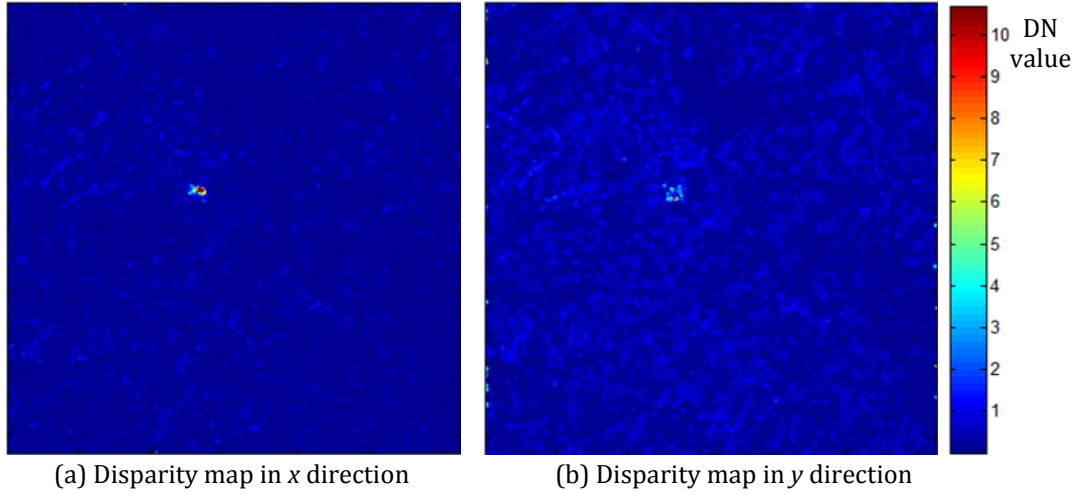


Figure 6.16 Disparity maps generated from PC based dense matching

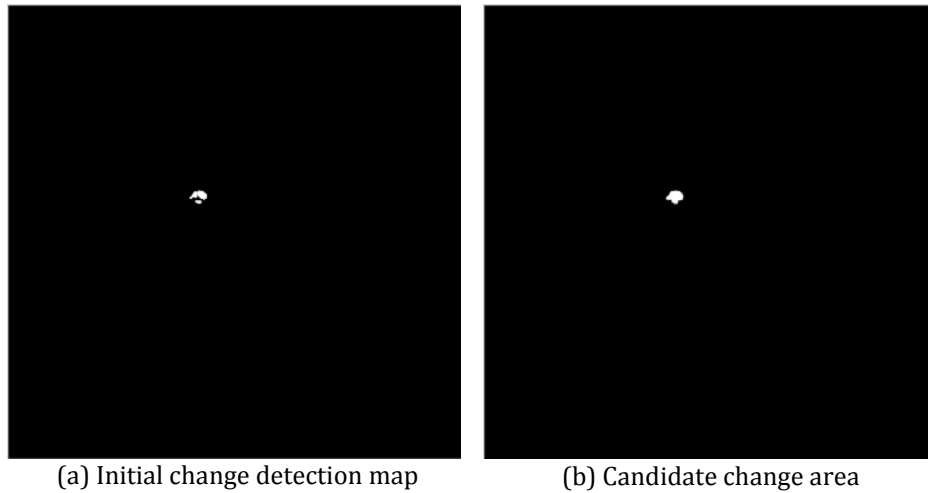


Figure 6.17 (a) Initial change detection map and (b) the candidate change area generation after morphological closure

In the next step, the candidate areas are to be assigned to one of the following types: object motion, new object and texture change by change type identification, based on the δ_0 peak value of the candidate areas. The final result is shown in Figure 6.18; the type of the change was set as 'New object' because the majority of pixels in the candidate area have low δ_0 peak values.

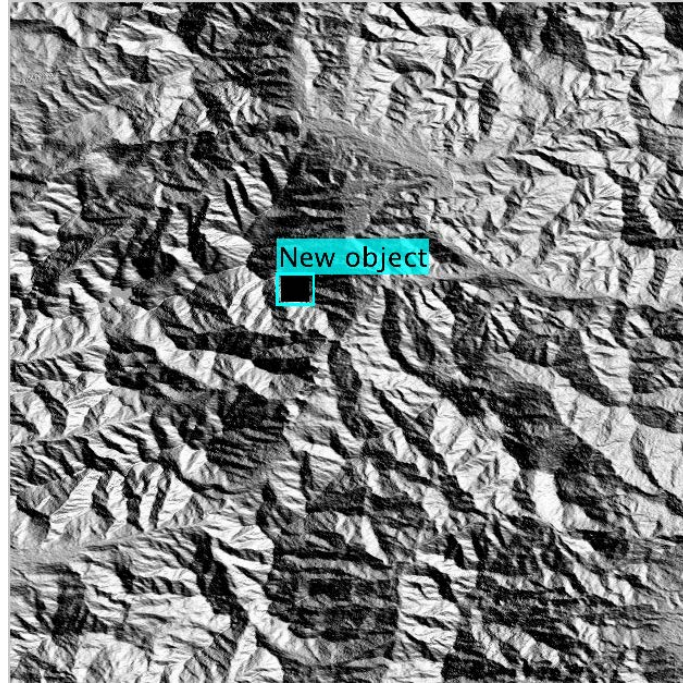


Figure 6.18 Final appearance change detection result

To test the robustness of the proposed illumination-insensitive PC based change detection approach, further experiments with different azimuth and zenith angle variations were carried out. The results of the 2D appearance change detection under different azimuth angles are shown in Figure 6.19 and the results of zenith angle variation are shown in Figure 6.20.

In all the tested illumination variation cases presented in Figure 6.19 and Figure 6.20, the appearance change of the black box has been successfully detected. The change detection results are better in Figure 6.20 than those in Figure 6.19. The blue boxes cover exactly the correct range of appearance change in Figure 6.20 while the blue boxes in Figure 6.19 sometimes do not cover the exact boundary of the area of change. Moreover, in Figure 6.19(b), there is an incorrectly detected false change area. This is because the azimuth angle variation may produce in negative correlation in the PC cross power spectrum while zenith angle variation does not. Thus the proposed approach is more robust to zenith angle change than azimuth angle change.

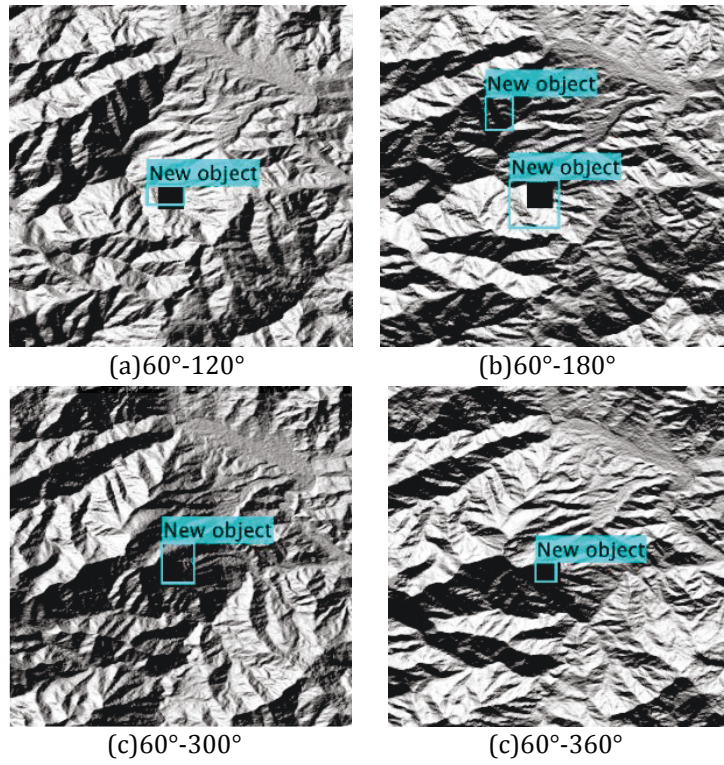


Figure 6.19 2D appearance change detection results under same zenith angle (50°) and different azimuth angles (shown below the images).

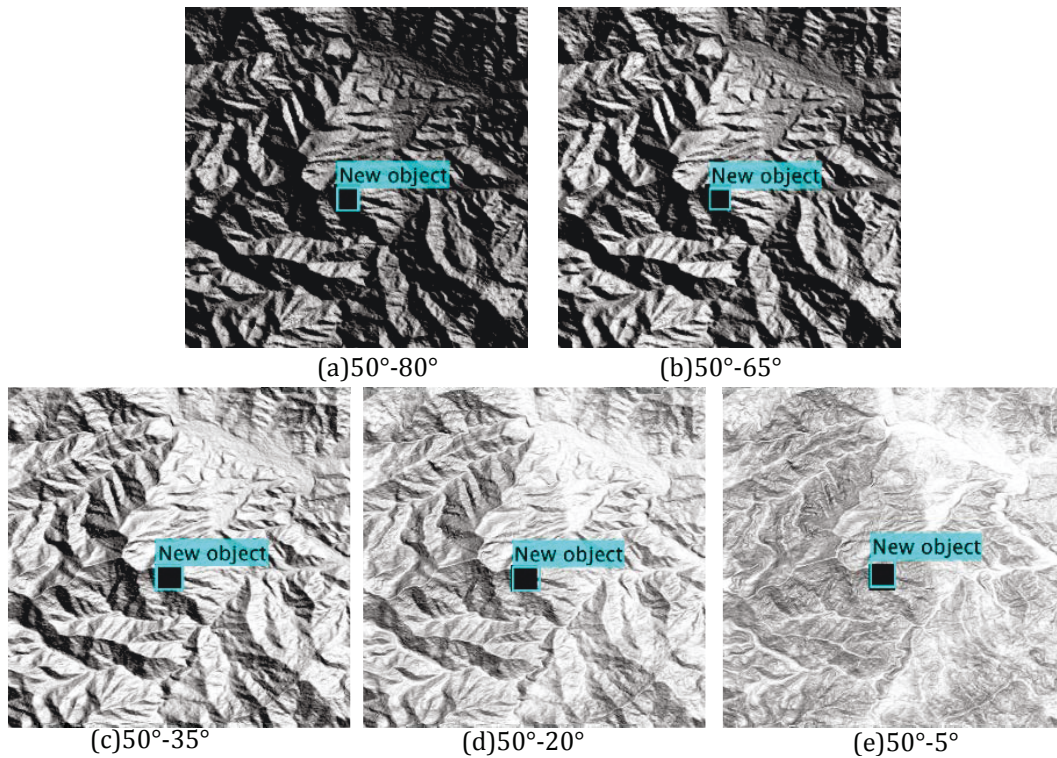


Figure 6.20 2D appearance change detection results under same azimuth angle (60°) and different zenith angles shown below the images.

(2) 3D change detection

In this experiment, we test the robustness of proposed change detection approach to detect 3D elevation change in a terrain model. The 3D change detection dataset was generated as following:

Firstly, a small area in the original DEM, shown in Figure 6.21(a), was slightly modified to generate Figure 6.21(b). The modified area is marked by red boxes in the two images (a) and (b), and enlarge displayed in (c) and (d).

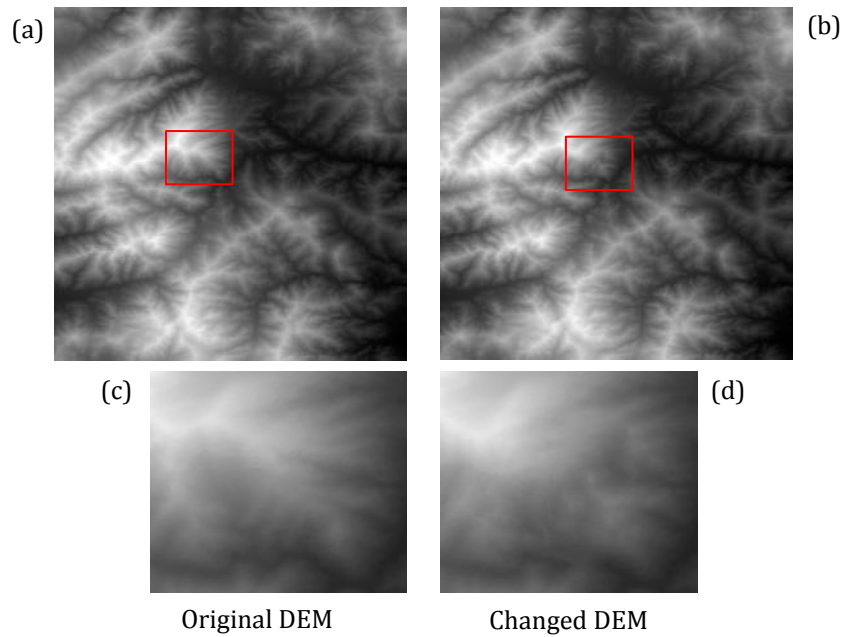


Figure 6.21 DEM dataset for 3D change detection, and the red boxes in (a) and (b) indicate 3D change in area in (c) and (d).

Two hill shading images under different daily illumination conditions, at 8:00 and 16:00, were then generated from the DEMs in Figure 6.21(a) and Figure 6.21(b). Compared to the 2D change, shown in Figure 6.15, this 3D change is visually more difficult to identify, because there is no distinctive alteration of the DN values within the changed area. Moreover, under different local illumination conditions, the local pattern change introduced by the variation of azimuth and zenith angles may overwhelm the real 3D change information.

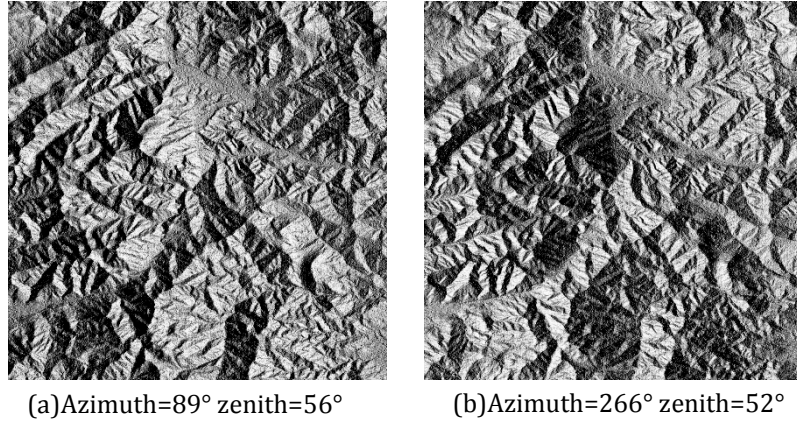


Figure 6.22 Hill shading images with 3D change under different daily illumination conditions at 8:00 and 16:00.

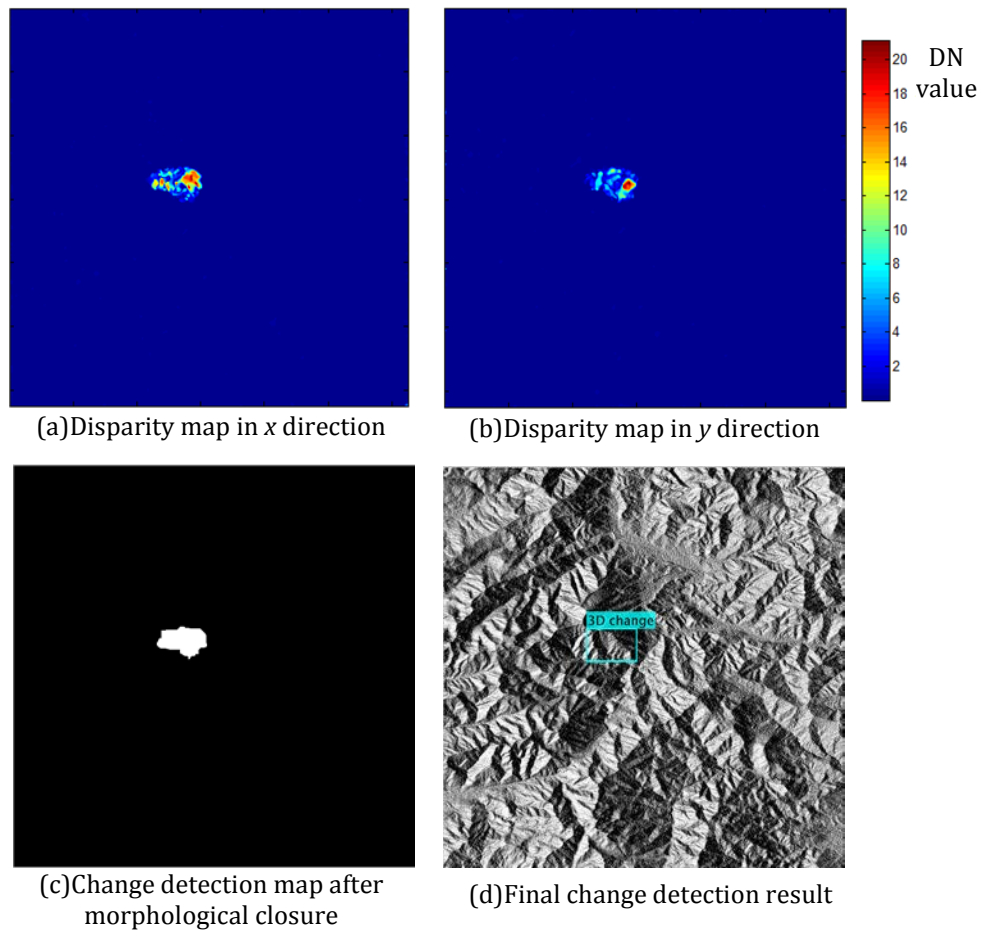


Figure 6.23 PC based illumination-insensitive 3D change detection results

The dataset shown in Figure 6.22 was then processed by the proposed PC based change detection approach, and the results are shown in Figure 6.23. The disparity maps generated by pixel-wise PC based dense matching shown in Figure 6.23(a)

and (b) clearly indicate the areas of 3D change, and the change detection map after morphological closure in figure (c) clearly shows the area of 3D change. The final result of change detection, shown in Figure 6.23(d), shows the correctly extracted area of 3D change. This experiment demonstrates that although 3D change is visually difficult to detect, the proposed PC based change detection approach achieve it effectively.

(3) Object motion

This section presents the experimental results of detecting another change type: object motion, under daily illumination variation. The dataset includes hill shading images generated under different azimuth and zenith angles listed in Table 4.3, as illustrated in Figure 6.24. A small image patch (representing a car) of 57 pixels width and 54 pixels height, labelled by red squares in Figure 6.24, was inserted into the two hill shading images, at different positions to simulate object motion. The car was shifted 5 pixels from its original position to the right and downward from Figure 6.24 (a) to (b).

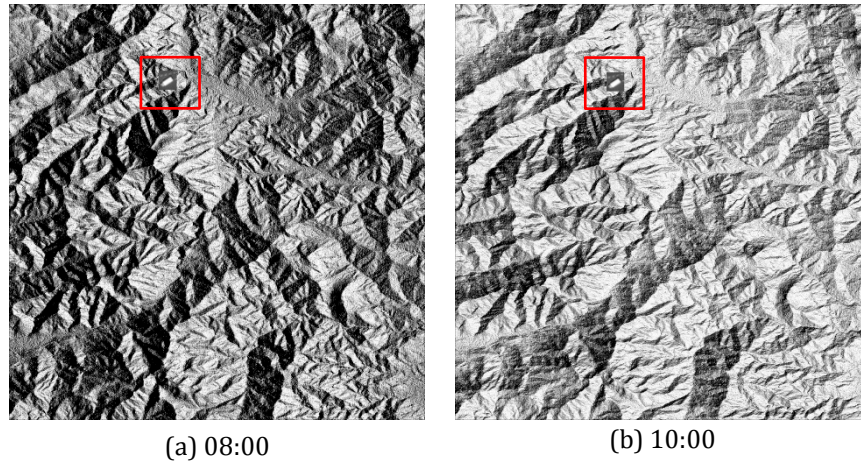


Figure 6.24 Object motion detection dataset, (a) is a hill shading image with a car added and (b) is a hill shading image with the car shifted

The dataset was then processed by the proposed change detection algorithm, and the object motion detection result is shown in Figure 6.25. The calculated object motion is 4.9393 pixels and 4.973 pixels in x and y direction shown in the

annotation box. Figure 6.25 demonstrated that the proposed approach is not only able to extract the moving object but also to accurately estimate its motion.

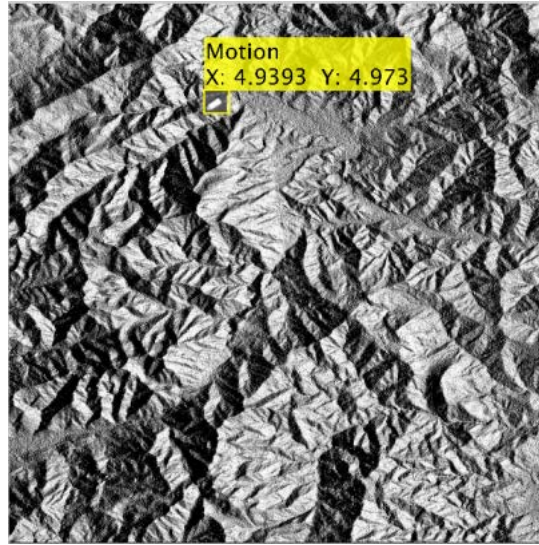


Figure 6.25 Object motion detection result. The first line of the annotation text shows the change type and the second line of the annotation text presents the estimated object motion in x and y direction respectively.

Three further object motion detection experiments were carried out using the hill shading images from images acquired at different daily times, and the true object motion remains the same as 5 pixels in x and y direction. The accuracy of object motion estimation was calculated as the absolute difference between the estimated motion and the true motion in x and y direction, e_x and e_y , as shown in Table 6.1.

Table 6.1 Object motion detection accuracy under daily illumination variation

Time Accuracy	08:00-10:00	08:00-12:00	08:00-14:00	08:00-16:00
e_x (pixel)	0.06	0.01	0.65	0.19
e_y (pixel)	0.03	0.02	0.04	0.20

The data in Table 6.1 indicate that when the time interval is less than 4 hours, the estimation of object motion can reach $\sim 1/20$ pixel accuracy. The proposed approach has the largest error, 0.65 pixels in x direction in the interval of 08:00-14:00, because in this case, both azimuth and zenith varied greatly, 150° and 26° respectively. Although the object motion estimation accuracy declines with increasing in time interval, in the four tested daily illumination variation cases, the

moving object is correctly extracted and the motion estimation errors are less than 1 pixel.

(4) Small object change detection

Two hill shading images were generated under different illumination conditions: $\tau_1=120^\circ$, $\sigma_1=50^\circ$ and $\tau_2=60^\circ$, $\sigma_2=50^\circ$, as shown in Figure 6.26. A small object with size of 10×10 pixels was inserted into Figure 6.26(b). Since the dense matching window size was 32×32 pixels, this small object does not cause significant change in disparity maps. The approach described in §6.4.4 was then carried out to detect the small object change.

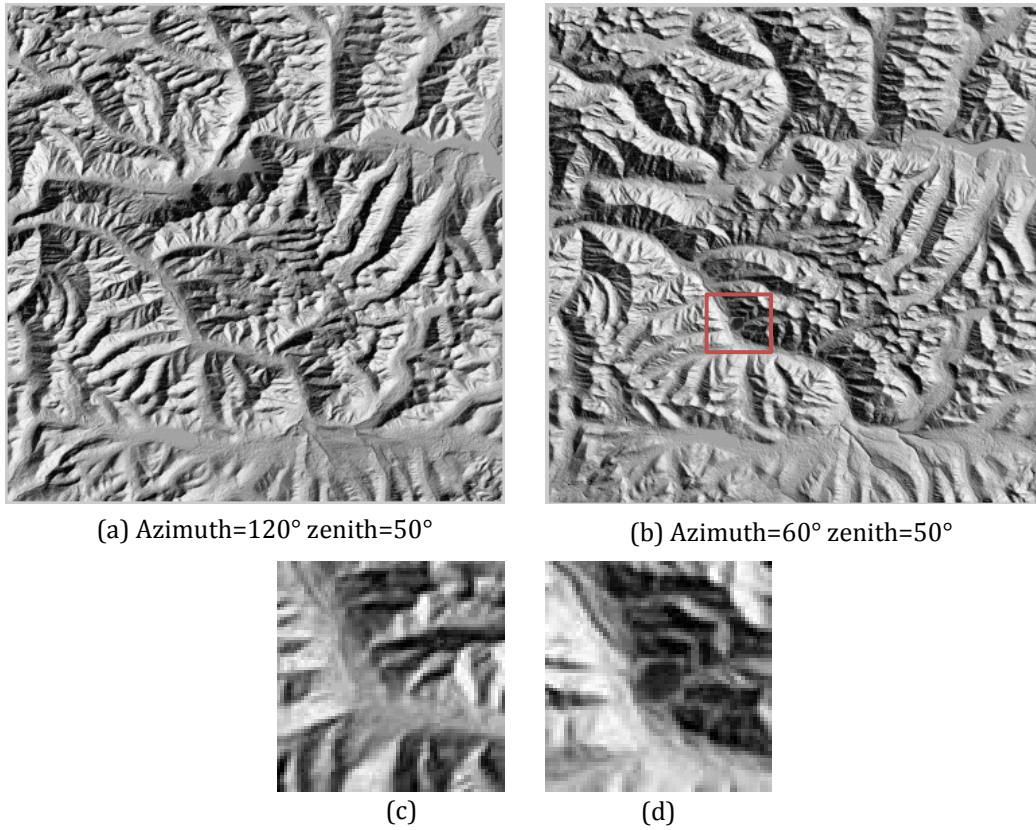


Figure 6.26 Two hill shading images under different illumination conditions (Image size: 368×368 pixels). A 10×10 pixels small object, labelled by a red box, has been inserted to (b). (c) and (d) are sub-scene images of areas marked by the red box.

The δ_0 peak function map P generated from the PC based pixel-wise dense matching is shown in Figure 6.27 displayed in pseudo colour mode with the values ranging from 0 to 1. An area with obviously lower δ_0 values than surrounding

areas in Figure 6.27 pinpoint to the small object change because the correlation level within this area is reduced by the change.

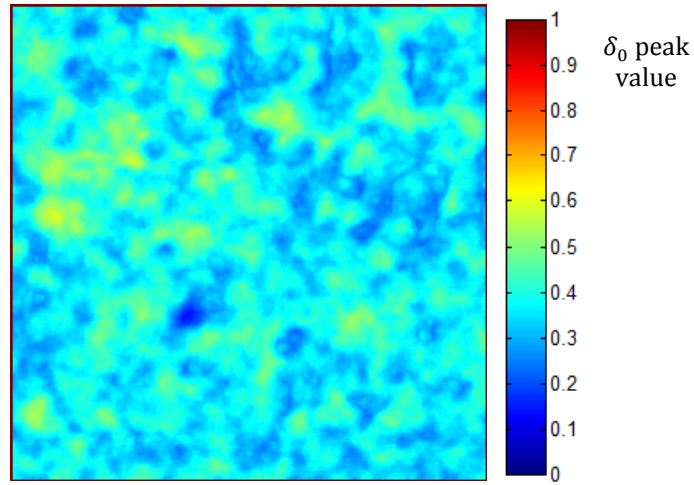


Figure 6.27 δ_0 map generated by PC based pixel-wise dense matching of images shown in Figure 6.26(a) and (b).

The detection result of this small object change is shown in Figure 6.28, in which the detected small object is marked by a red box. Figure 6.28 demonstrates that the proposed approach is able to correctly detect small object changes under illumination variation environment.

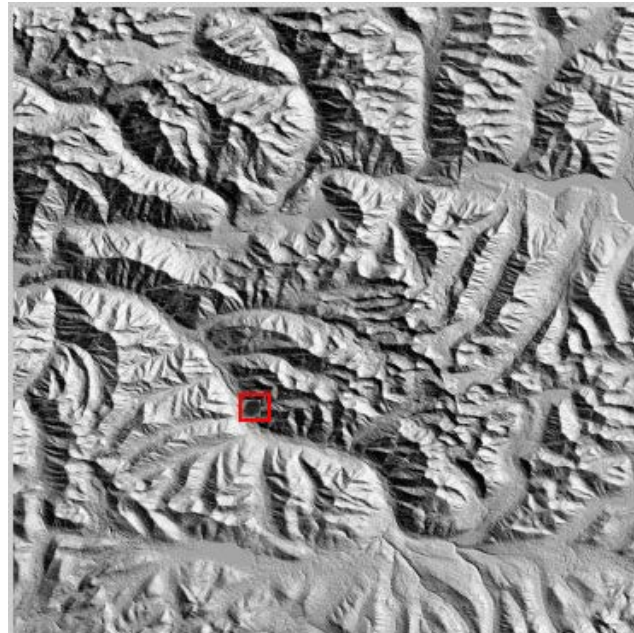


Figure 6.28 Small object change detection results by the proposed PC based approach. The changed area is marked by a red box.

6.5.2 Natural terrain model image

(1) Simple terrain model

A miniature terrain model with a surface elevation range of 50mm was imaged from a distance of $H = 1295$ mm using a Nikon D70s camera which was set tilted in an angle $\alpha = 70^\circ$ and with compass direction $\beta = 175^\circ$, to best avoid shadow blockage. The experiment was carried out on 24th April 2013 in the South Kensington campus of Imperial College London. The two terrain images, taken at 09:47 and 15:29 are shown in Figure 6.29.

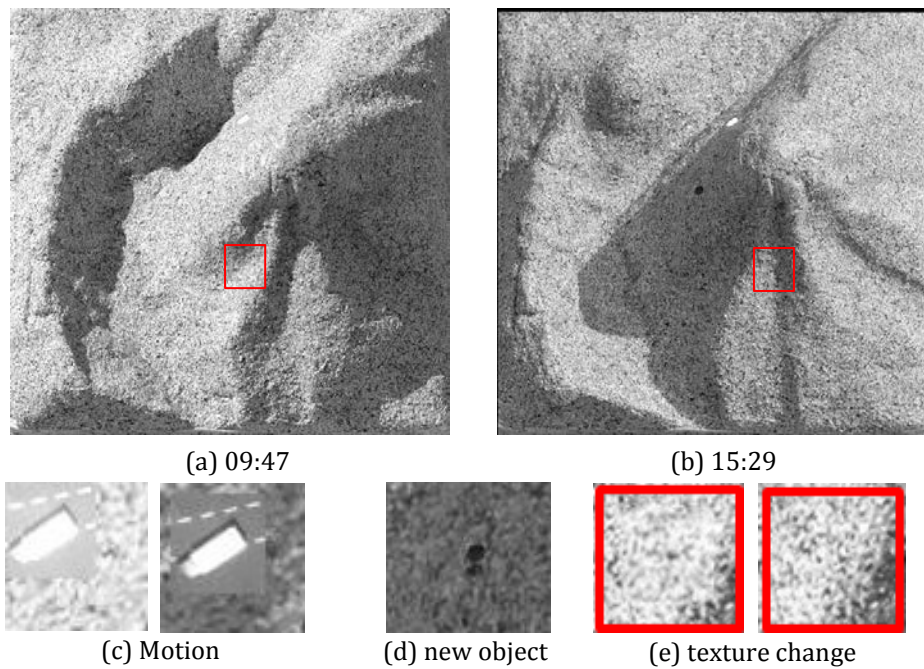


Figure 6.29 Real sun shading images taken at 09:47 and 15:29 (Image size 428×440 pixels). Three types of changes are implanted and the red box marks the camouflage area. The enlargements of three change targets are shown in (c)-(e).

Three types of changes have then been added to the image pair in Figure 6.29, including:

- A small image, 38×36 pixels, of a car shifted 5 pixels rightward and downward in (b), and detailed in (c);
- A small object of 10×10 pixels in (b) giving a as small appearance change, and detailed in (d);

- A simulated camouflage change by extracting a small textural patch from one position in (a) and pasting it to another position in (b) indicated by the red boxes, and detailed in (e).

The experiment dataset was processed by the proposed PC based change detection algorithm and the results are shown in Figure 6.30, in which the detected change areas are marked and defined by boxes in different colours: with yellow for object motion, cyan for new object and green for camouflage, together with annotation. All the three types of changes were successfully detected and extracted. The motion of the moving object was accurately estimated with only 0.08 pixel and 0.01 pixel deviation from the true value of motion in x and y direction.

The experimental results in Figure 6.30 demonstrate that the proposed PC based change detection algorithm is able to detect and discriminate all three types of change and to accurately estimate the distance of object motion under very different illumination conditions.

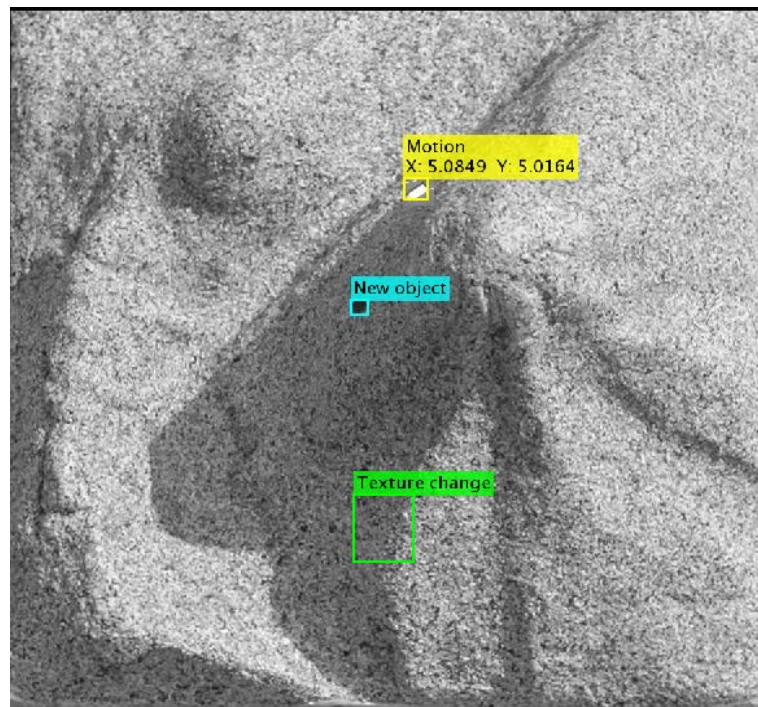


Figure 6.30 Multiple type change detection results. The change areas are marked by boundary boxes and annotation text above the box show the change type identified.

(2) Snowdonia terrain model

To test in more complex terrain, the Snowdonia terrain model, as described in §5.3.2 was used to generate terrain images under different daily illumination conditions, as shown in Figure 6.31.

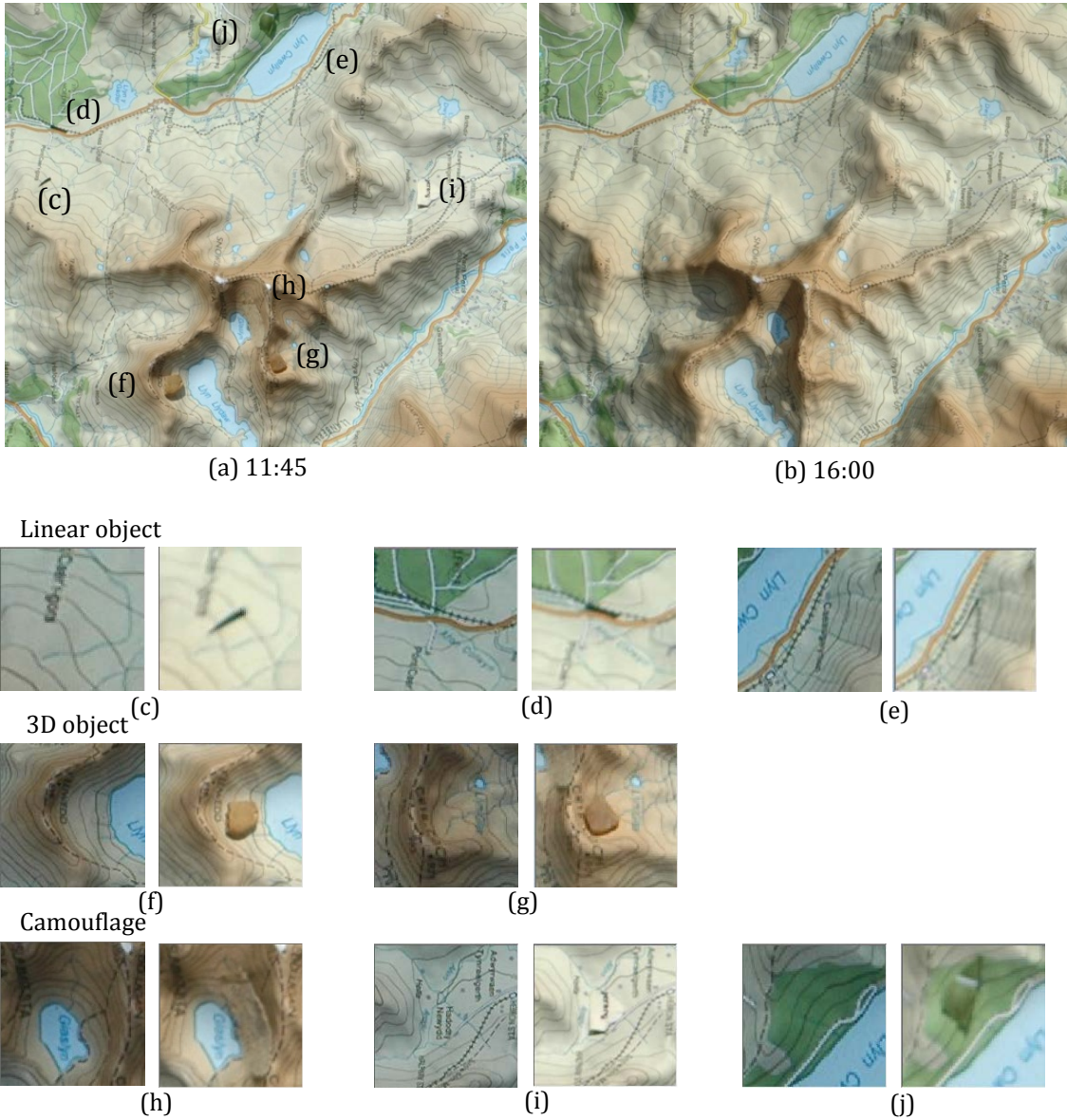


Figure 6.31 Images (694×828 pixels) of Snowdonia model taken at 11:45 and 16:00 with appearance changes. The enlargements of changing targets are shown in (c)-(j).

Instead of inserting changes in images, test objects were placed on the model and moved manually; some of them at different time periods in this experiment. The pair of images for testing appearance changes was taken at 11:45 in the morning and at 16:00 in the afternoon as shown in Figure 6.31(a) and Figure 6.31 (b). There are eight areas with appearance difference between the two images of three types:

linear object, 3D object and camouflage, as shown in Figure 6.31(c) and Figure 6.31(j). The proposed PC based change detection was then carried out and the results are shown in Figure 6.32.

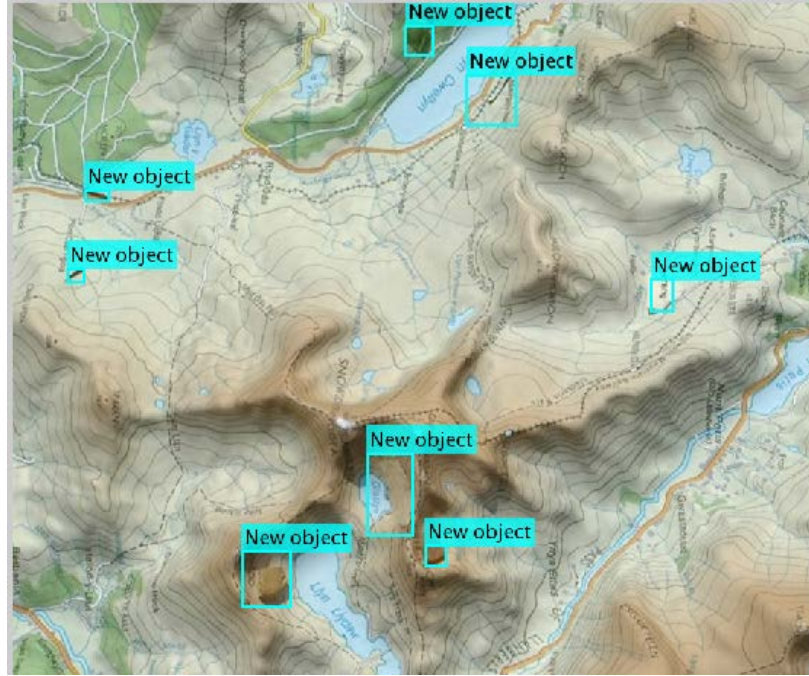


Figure 6.32 Appearance change detection result of Snowdonia model image. The changed areas are marked by cyan boxes and the annotation texts above the boxes indicate the change type.

In Figure 6.32, eight appearance change areas between Figure 6.31(a) and Figure 6.31(b) have all been extracted by the proposed PC algorithm. For linear object changes, two linear objects shown in Figure 6.31(c) and Figure 6.31(d) have been correctly extracted and defined, but the defining box of the linear object shown in Figure 6.31(e) is slightly larger than the actual size. This is because the width of this linear object is approximately 1 pixel, which does not produce significant decrease in δ_0 peak value. The three 3D objects have all been correctly detected. For the camouflage detection, all the three camouflage areas have been correctly defined and extracted.

To test detection and motion estimation of moving objects, two images of the Snowdonia model taken at 11:45 and 14:30 were used as shown in Figure 6.33. The three linear targets and two 3D targets of the last experiment were manually moved before the image Figure 6.33(b) was taken. The change detection results

are shown in Figure 6.34. The five motion areas have been correctly detected and extracted with the motion values accurately estimated although the true motion values are not known precisely. The experiment demonstrates the robustness of the proposed approach to detect multiple object motions under different daily illumination conditions.

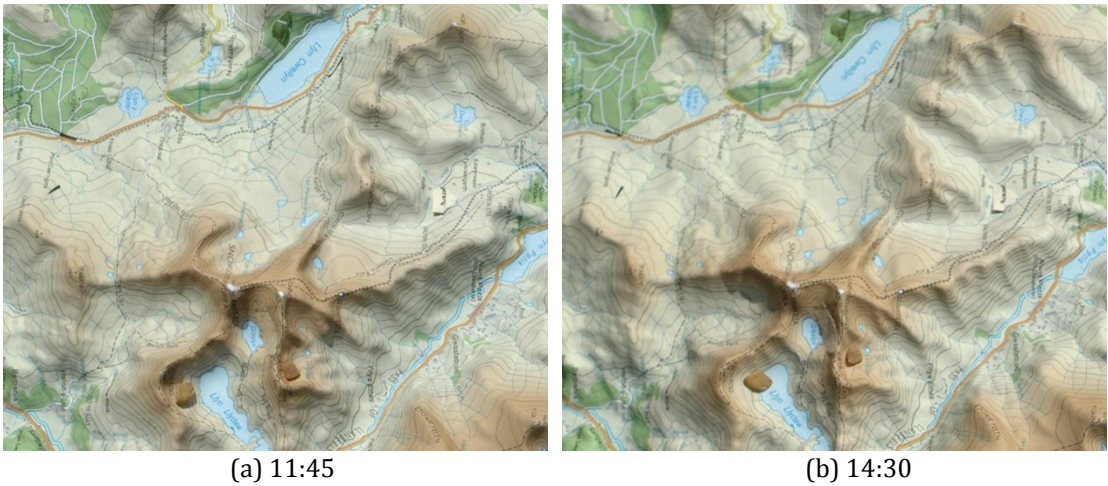


Figure 6.33 Snowdonia model images taken at 11:45 and 14:30 with motion change.

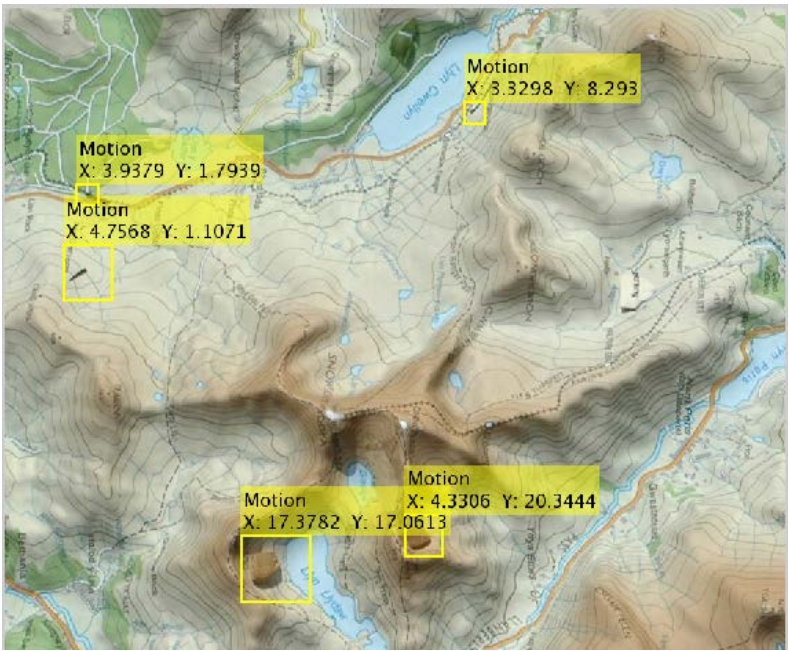


Figure 6.34 Motion Detection results of Snowdonia model images. The detected motion areas are marked by yellow boxes and the annotation texts above the boxes indicate the change type and the estimated motion values.

6.5.3 Example applications

In §6.5.1 and §6.5.2, the robustness of PC based change detection to local illumination change has been demonstrated by experiments using DEM simulated hill shading images and images of terrain models taken in real illumination conditions. In this section, the robustness of the proposed approach is tested using different satellite images under various illumination conditions.

(1) Glacier monitoring of the Hubbard Glacier, Alaska

As mentioned in §6.1, glacier monitoring is of great importance to the environment and climate research. Illumination variation imposes an obstacle to change detection in glacier areas mainly in high latitude regions near poles. In such regions, the sun zenith angle is much greater (i.e. very low sun angle) than that in low latitude areas during satellite imaging time. This means that topographic shadows of high relief glacial terrains can be very long and may darken and obscure the ‘real’ changes. High latitude regions are subject to high sun zenith together with great seasonal variation and thus the terrain shadow patterns change greatly with seasons altering the appearance of multi-temporal images. An example is provided in Figure 6.35, which two Landsat 8 panchromatic band images of the Hubbard glacier, in Alaska, USA; it is the largest calving glacier in North America. The centre latitude and longitude of the study area are 59°59'31.22"N and 139°29'38.30"W, as shown in Figure 6.36. Compared to the simulated hill shading images shown in Figure 6.15, which only contains the effects of shade, the real satellite images in Figure 6.35 also show long overcast shadows. The images in Figure 6.35 were taken in February and April respectively, so the sun illumination angles are different, as shown in Table 6.2. The sun azimuth angles of the two images are similar, because the imaging times are approximately 10am for sun-synchronous satellites. There is, however, a 16° sun zenith angle difference owing to the seasonal change of solar position. Since the zenith angles are relatively large during winter at high latitude, even 16° zenith angle difference will result in great variation of glacier shadows, as shown in Figure 6.35. The

subtraction and ratio images between Figure 6.37(a) and Figure 6.37(b), and the change detection shown in Figure 6.38, confirm that the change information is overwhelmed by the shade and shadow variation under different illumination condition.

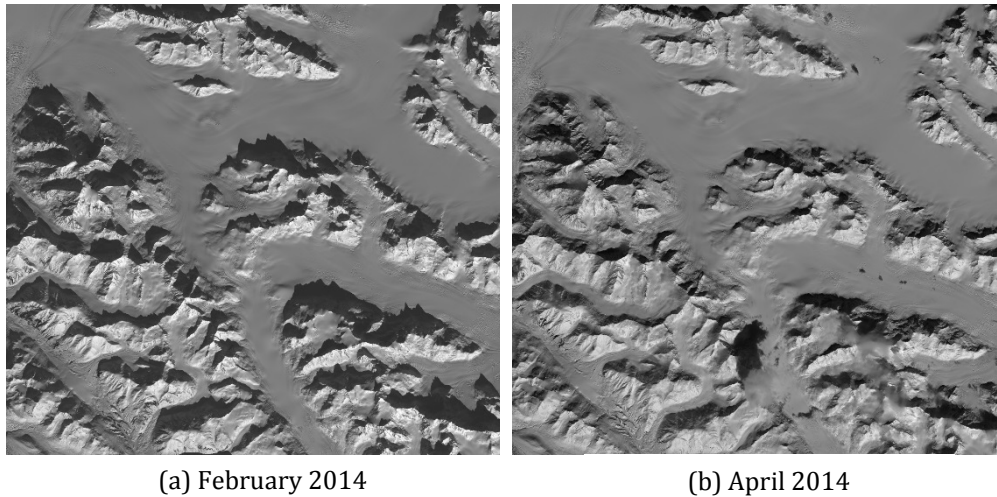


Figure 6.35 Lansat 8 panchromatic band image of Hubbard glacier taken in February and April 2014 (Image size 1226×1118 pixels, Image resolution: 15m)

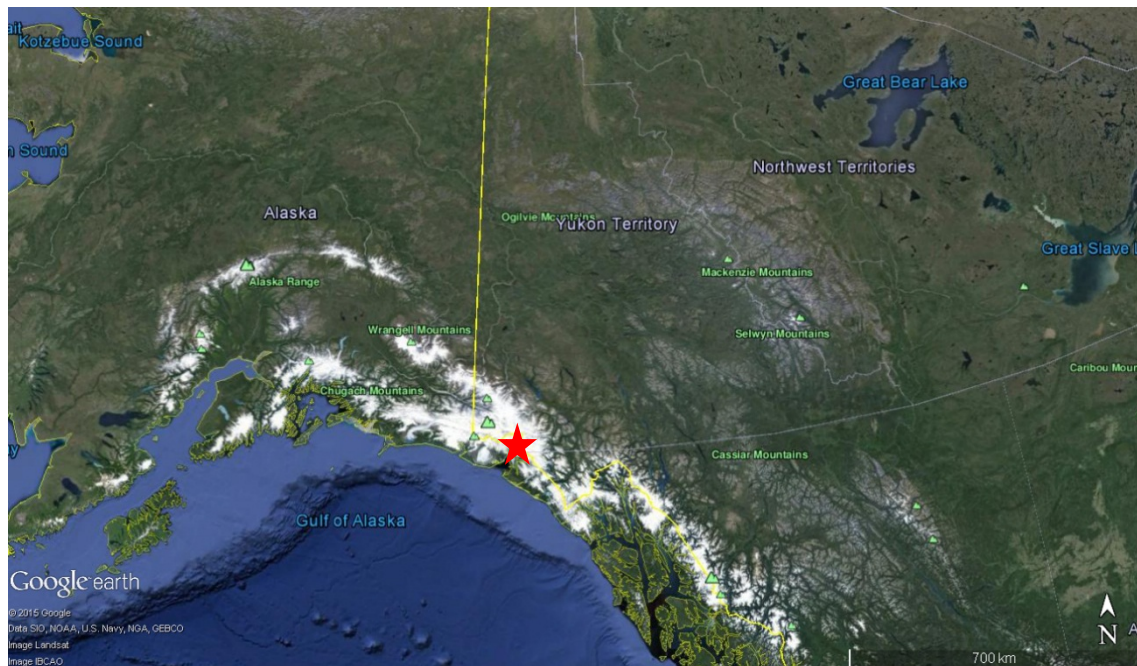
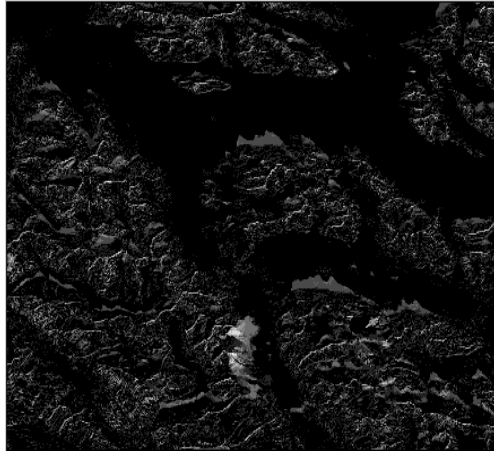


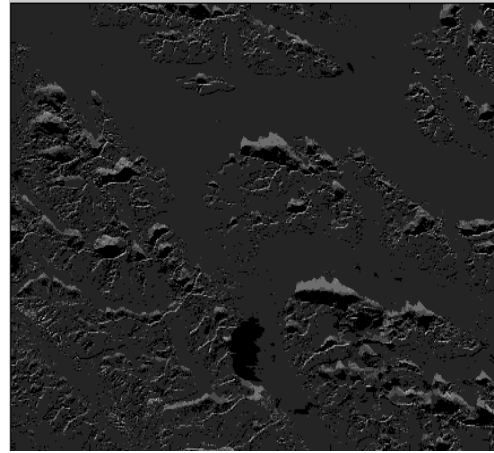
Figure 6.36 Location of Hubbard glacier in Google Earth shown by the red star (Centre latitude: 59°59'31.22"N, Centre Longitude: 139°29'38.30"W)

Table 6.2 Acquisition date and illumination conditions of images in Figure 6.35

	Acquisition date	Solar zenith(°)	Solar azimuth(°)
Image (a)	10 April 2014	52.7	163.9
Image (b)	28 February 2014	68.8	163.5



(a) Subtraction image



(b) Ratio image

Figure 6.37 Subtraction and ratio images of Figure 6.35(a) and (b)

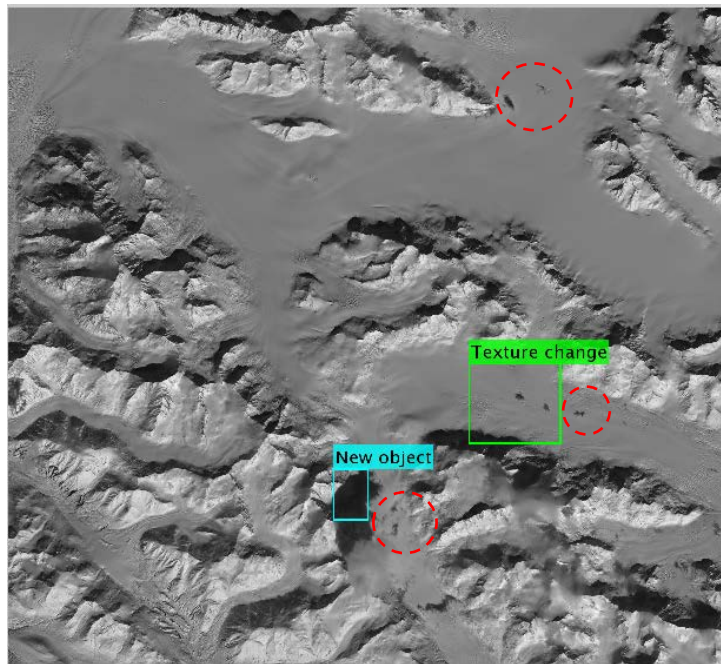


Figure 6.38 Change detection results of Hubbard glacier based on Landsat 8 images. The change areas are marked by square boxes and the annotation texts above the boxes indicate the change type. The red circles drawn in dashed line show some mis-detected change areas.

Figure 6.38 shows the change detection result using the proposed PC based illumination-insensitive change detection algorithm; two change areas were extracted and classified as ‘New object’ and ‘Texture change’ respectively. The

reason for the area marked by a green box labelled as 'Texture change' is that the sizes of new object is too small and was recognized as texture change by the proposed algorithm. By the comparison of Figure 6.35(a) and (b), these two areas represent real changes of new objects, rather than appearance change caused by illumination variation. The changes in shading and shadows shown in Figure 6.37 are almost completely eliminated by the proposed approach, further demonstrating the illumination invariant property of the PC based algorithm. The area of change defined by the green box is, however, imprecise and some small changes remain undetected, as marked by red circles. This indicates that although distinctive changes can be extracted, the proposed approach is not yet powerful enough to detect small-scale subtle real environmental changes from the overwhelming changes in image appearance resulted caused by drastic variation of solar zenith in high relief areas in high latitude regions.

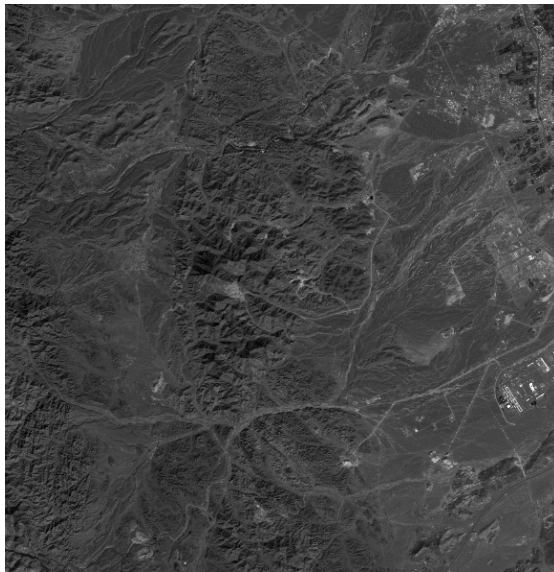
(2) Change detection in Sohar, Sultanate of Oman

This example is focused on the detection of manmade changes such as urban development, new roads, mining and quarrying. A low latitude area near Sohar in Oman was selected for this experiment. Sohar is a coastal town surrounded by mountains shown in Figure 6.39, and centre latitude and longitude of the area are 24°33'10.37"N and 56°43'55.06"E.

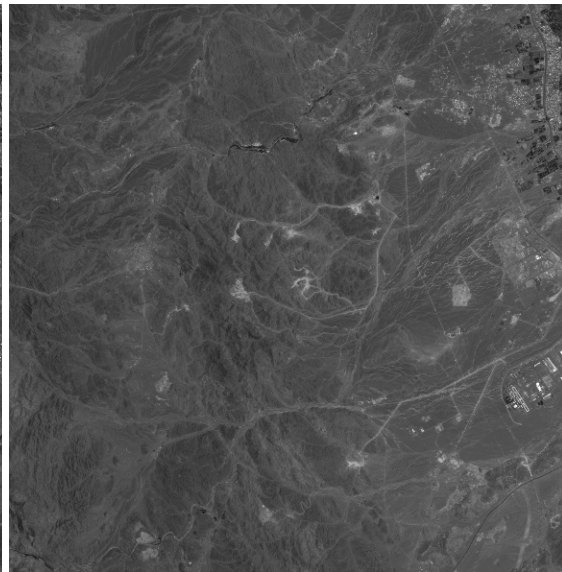
Two Landsat 8 images of Sohar acquired in January and July of 2014, shown in Figure 6.40, were used for this change detection experiment. The data acquisition date and sun illumination angles are listed in Table 6.3. The sun illumination change greatly from January to July in Sohar; the sun zenith angle difference between the acquisition dates is 31° and the sun azimuth angle difference is 63°. The effect of local illumination change can be seen clearly in Figure 6.40 in which the direction of topographic shading and shadows are different between the two images. As demonstrated in Figure 6.41, the shading changes caused by illumination variation, and the real environmental changes are combined in the difference image and ratio image.



Figure 6.39 Location of study area, Sohar. (Centre latitude: 24°33'10.37"N, Centre Longitude: 56°43'55.06"E)



(a) January 2014



(b) July 2014

Figure 6.40 Landsat 8 Panchromatic band image of Sohar acquired in January and July 2014 (Image size 1516×1487 pixels, image resolution:15m)

Table 6.3 Acquisition date and illumination conditions of Figure 6.40(a) and (b)

	Acquisition date	Solar zenith(°)	Solar azimuth(°)	Path row number
Image (a)	2014/01/03	52.63	151.94	LC81590432014003LGN00
Image (b)	2014/06/12	21.22	88.63	LC81590432014163LGN00

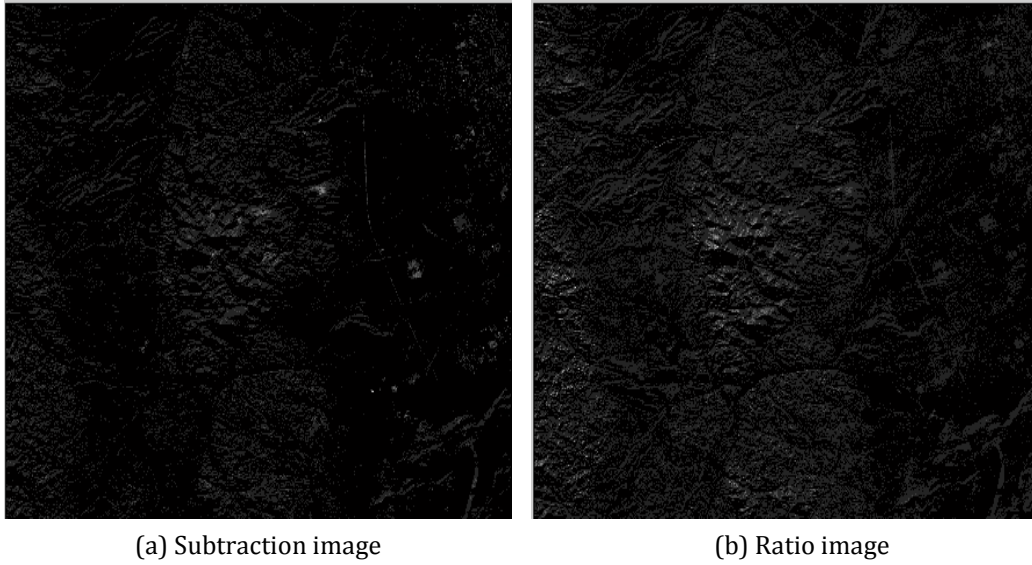


Figure 6.41 Subtraction and ratio images of Figure 6.40(a) and Figure 6.40 (b).

The change detection result produced by the proposed PC algorithm is shown in Figure 6.42. Visual comparison between Figure 6.40(a) and Figure 6.40(b) confirms that the five detected change areas in Figure 6.42 are all real changes and likely manmade. Despite considerable variation of illumination conditions, no areas of topographic pattern change were incorrectly identified as environmental change areas. This demonstrates the robustness of the proposed approach to local illumination change.

It should be noticed that the white path marked by a yellow box in Figure 6.42 is not really an object motion. The zoom in images of the yellow box (Figure 6.43) reveals that the white patch in Figure 6.43(b) appears much larger than that in Figure 6.43(a). These two white patches in Figure 6.43(a) and Figure 6.43 (b) are considered to represent the same object, and because of the expansion of the white patch in Figure 6.43(b) the centre of this patch has shifted slightly in comparison with the centre of the white patch in Figure 6.43(a). The patch could represent a quarry in expansion toward the east and in this sense, it can be considered as an 'object motion'.

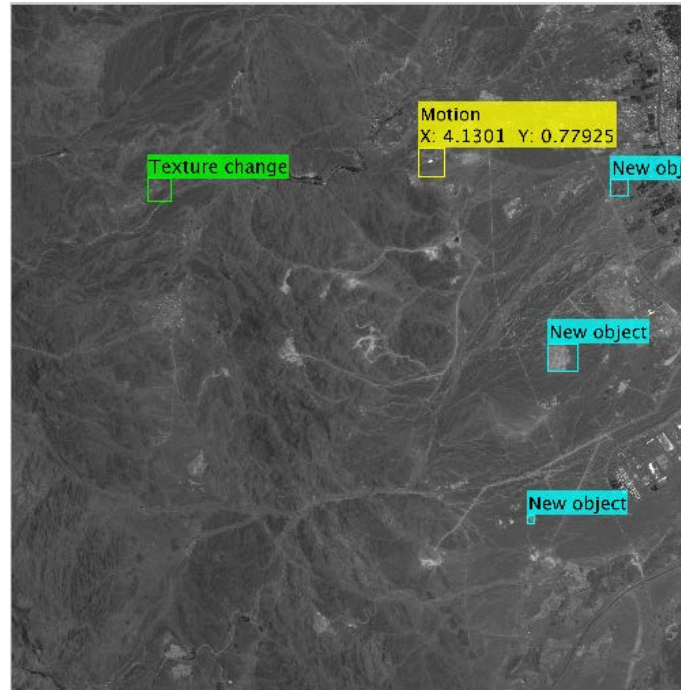


Figure 6.42 Change detection results of Sohar from January to July 2014 based on Landsat 8 images. The change areas are marked by boxes and the annotation texts above the boxes indicate the change type.

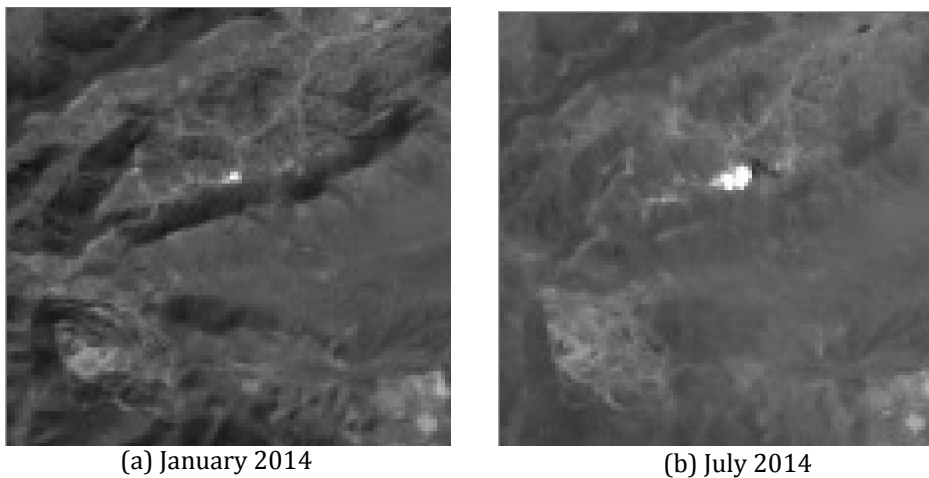


Figure 6.43 Zoom in images of the yellow box area in Figure 6.42

(3) Rover detection on Mars

In this experiment, three HiRISE images captured the Opportunity rover in different positions were used as shown in Figure 6.44 in which the rover positions are marked by red box and enlarged to show the rover. The Opportunity rover is 2.3m wide and 1.6m long and in the 25cm resolution HiRISE images, the size of the rover is approximately 6×9 pixels.

According to the observation statistics provided by NASA/JPL/University of Arizona, shown in Table 6.4, the illumination conditions vary considerably in the three HiRISE images, with the largest azimuth angle difference of 56.1° between ESP_012820_1780 and ESP_016644_1780, and the largest zenith angle difference of 8° between ESP_011765_1780 and ESP_016644_1780.

Table 6.4 Observation statistics of the three HiRISE images

HiRISE image ID	Acquisition date	Local Mars time	Solar azimuth	Solar zenith	Original resolution	Emission angle
ESP_011765_1780	29 Jan.2009	3:46 PM	171.3°	57°	27.6 cm/pixel	12.5
ESP_012820_1780	21 Apr.2009	3:24 PM	151.9°	54°	29.6 cm/pixel	26.1
ESP_016644_1780	13 Feb. 2010	3:00 PM	208.0°	49°	27.0 cm/pixel	0.9

The effects of illumination variation on the HiRISE images are illustrated in Figure 6.45 (a-c); there are significant variation of the direction and length of shadows vary in the high relief areas. As shown in Table 6.4, the emission angle, the oblique angle between the HiRISE camera and the Mars surface normal are very different: ESP_016644_1780 is nearly nadir view, while ESP_012820_1780 has a oblique angle of 26.1° . The emission angle difference leads to very different geometric distortions between the two images. The combined effects of original resolution, emission angle and weather condition at the imaging times result in the variations of SNR (Signal-Noise-Ratio) as illustrated in Figure 6.45(c-e); ESP_012820_1780 clearly has an obviously lower SNR than the other two images.

In this experiment, the three images in Figure 6.44 were used as reference and target in turn to formulate three image pairs. The reference images and target images are generated around the rover, with 922 pixels width and 1410 pixels height, equivalent to the searching range of 230m width and 252m length. As the rover is in quite different places in the three images, the small reference and target images ensure the rover is only in the target image of each pair.

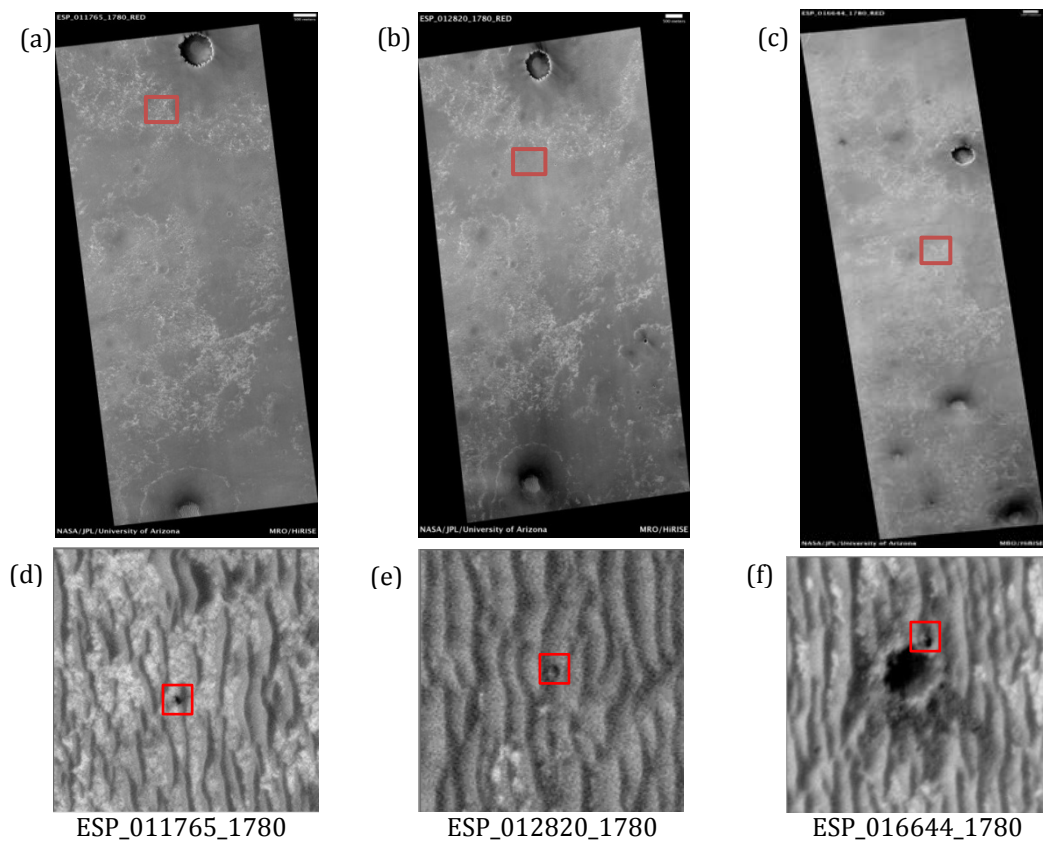


Figure 6.44 (a), (b) and (c) are three HiRISE images captured the Opportunity rover in different positions, indicating by the red boxes, image (d), (e) and (f) represent the images the enlargement of the red box areas in (a), (b) and (c).

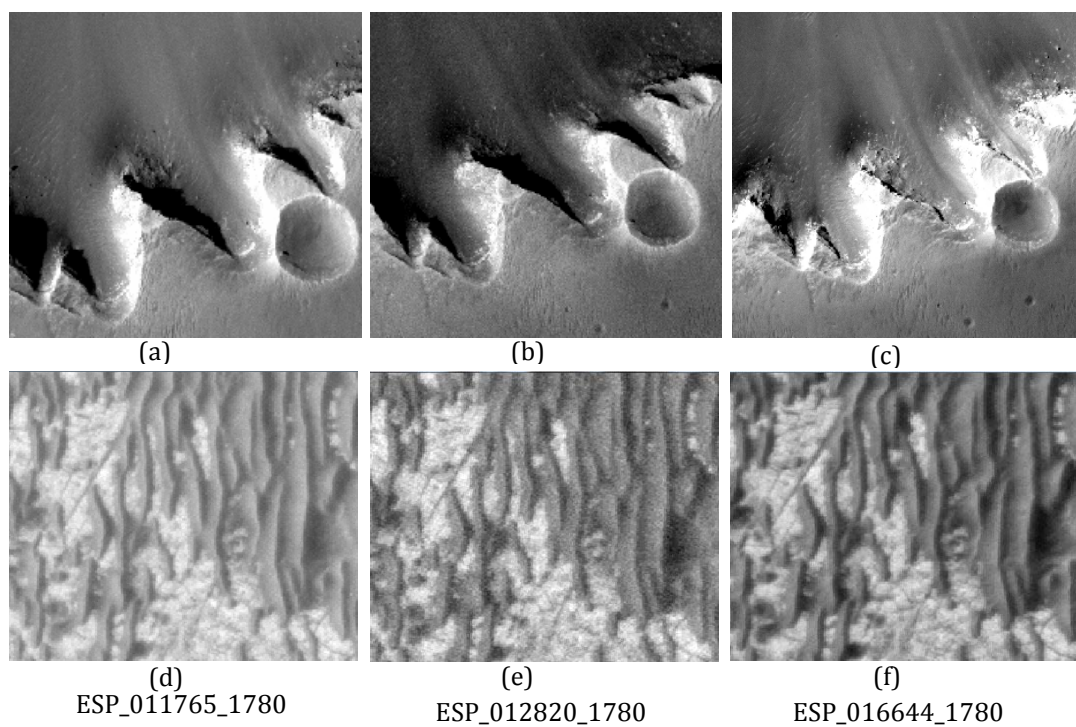


Figure 6.45 Illumination difference and SNR difference in the three HiRISE images

The proposed small object change detection approach, described in §6.4.4 was applied then to the three image pairs. For illustration, the rover detection using the image pair of ESP_012820_1780 and ESP_016644_1780 is presented here in detail. The two images were taken under quite different illumination conditions and have large SNR difference. Figure 6.46 shows the δ_0 peak map generated by PC scanning in Figure 6.46(a) and the difference image between the rectified target image and the reference image in image Figure 6.46(b). The combination of low values (blue) in the δ_0 peak map and high values (red) in the difference map, based on the approach proposed in §6.4.4, enabled a correct detection of the rover position.

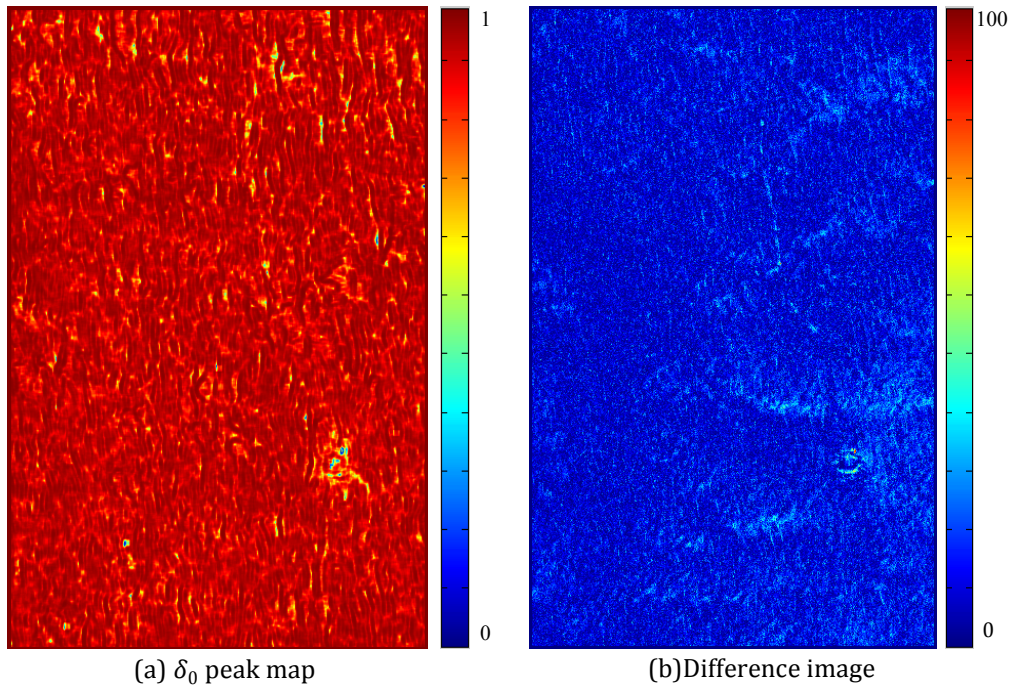


Figure 6.46 δ_0 peak map and difference image of ESP_012820_1780 and ESP_016644_1780.

The final detection results locating the Opportunity Rover in different positions in the three image pairs are shown in Figure 6.47. The final rover positions estimated by this approach are marked with red circles. In the three cases, although under different illumination conditions, irregular geometric distortions and different SNR levels, the positions of the Opportunity Rover have been correctly detected. The experimental results demonstrate the effectiveness and robustness of the proposed small object change detection approach under substantial illumination variation and geometric distortion.

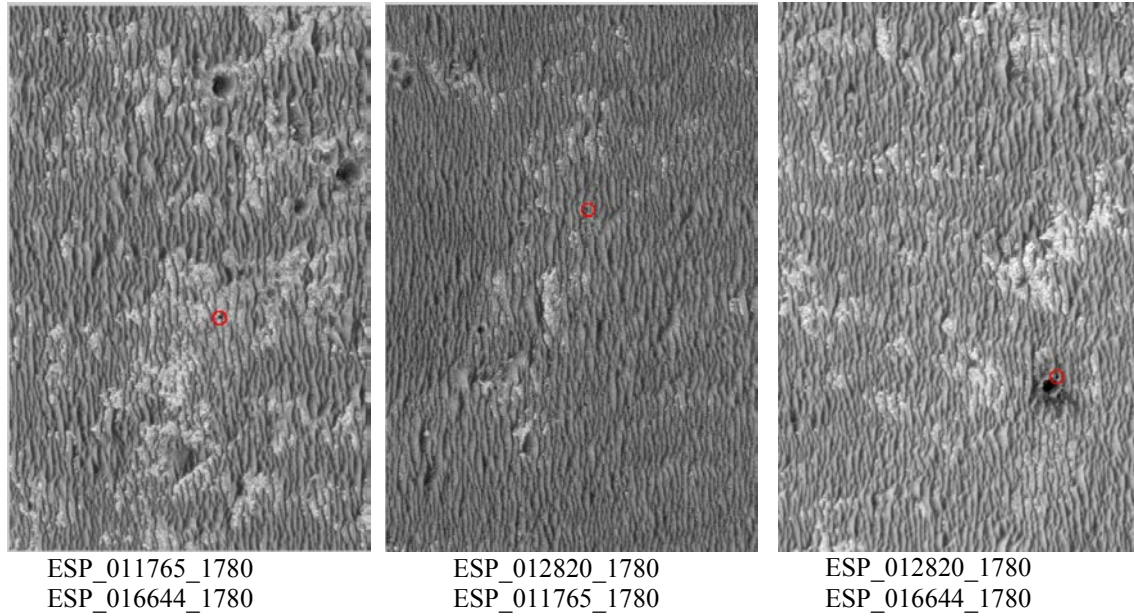


Figure 6.47 Rover Detection results using three stereo pairs

6.6 Summary

In this chapter, a PC based change detection algorithm is proposed. It has been proven to be insensitive to illumination variation in a variety of difference situations, and is able to rectify geometric distortions caused by camera motion. The pixel-wise co-registration using PC scanning is able to eliminate the geometric distortions between the images. Based on the illumination-insensitive property of PC, the disparity map and the δ_0 peak map generated from PC based image dense matching can be considered indicators for change. A comprehensive package of experiments using DEM simulated hill shading images and images of terrain models acquired under real illumination conditions, as well as satellite images, has been conducted to rigorously assess the capability of the algorithm. The experimental results demonstrate that the proposed PC based change detection approach is able to robustly detect real changes in object space regardless of illumination variation. Moreover, the proposed approach can also differentiate appearance change from object motion, and can correctly calculate the motion with sub-pixel accuracy.

Chapter 7 DEM Generation Based on Multi-temporal Remotely Sensed Images

7.1 Illumination-insensitive stereo matching

Presented in either raster or vector format, Digital Elevation Model (DEM) is a digital 3D terrain surface model generated from elevation data. There are many techniques to generate DEM [101-103], and one frequently used technique is using remotely sensed data, which two optical EO images are taken at different angles to formulate a stereo pair. The stereo pair is firstly rectified to obey the epipolar geometry, which the pairs of conjugate epipolar lines are collinear and parallel to one of the image axes. Then, pixel-wise image correlation, also named as dense matching, is carried out within the epipolar stereo pair to estimate the disparity value at each pixel. Disparity maps generated from dense matching can be transformed into DEM using the equations below:

$$\begin{aligned} Z &= \frac{fB}{d} \\ X &= \frac{xZ}{f} \\ Y &= \frac{yZ}{f} \end{aligned} \tag{7.1}$$

where (x, y) is the plane coordinate in 2D image, (X, Y, Z) is the 3D coordinate in object space, f is the focal length, B is the baseline and d is the disparity value.

Stereo matching, or dense matching, has long been studied in computer vision and photogrammetry and numerous algorithms have been proposed and tested [104]. They are basically under the following assumptions [104]:

- Smoothness of surfaces. The ideal disparity maps should be smooth except for edges and boundaries of surfaces.
- Radiometrically similar. The corresponding pixels in stereo pairs need to be identical.
- Camera calibration and epipolar geometry. The geometric distortion within the stereo images needs to be rectified to obey the epipolar geometry, and thus the disparities values only reflect depth information.

In this chapter, we agree with the smoothness and epipolar assumption, but we try to expand the boundary of the radiometrically similarity assumption for a wider application and to demonstrate that even when the radiometric appearance of two stereo images are not very similar, disparity maps can still be generated.

Radiometric difference will not happen when stereo pairs are taken consecutively by the same sensor at the same time. In the age of big data, however, numerous EO data are acquired separately from multiple sensors. The stereo pair may not necessarily be formed by the same sensor at the same time, and thus it may contain radiometrically difference caused by illumination variation. Thus, the illumination-insensitive dense matching becomes an essential technique.

Although several illumination robust dense matching algorithms have been proposed in computer vision society [46, 53], these studies mainly focused on global illumination change or limited local illumination variation, which means that the stereo images still have to be radiometrically similar in order to carry on dense matching algorithms. Illumination-insensitive dense matching is especially hard to achieve because topographic features are very sensitive to local illumination change according to the analysis in §2.1.

For years, our research group has developed a robust PC sub-pixel image DEM reconstruction software [12, 67] based on SVD [59] and QMDPE 2D fitting [60]. The sub-pixel PC matching algorithm has been successfully applied for DEM generation from very narrow baseline stereo images [61] via accurate estimation of sub-pixel disparities. The Phase correlation based stereo matching in terms of illumination invariant conditions however, to the best of our knowledge, has not been rigorously studied yet.

7.2 PC-based disparity map generation from time variant remotely sensed imagery

Based on the analysis of illumination invariant property of PC in Chapter 4, a PC based illumination invariant stereo matching approach is shown in Figure 7.1, which includes three key steps: stereo rectification, illumination-insensitive image alignment and ADCF-PC based pixel wise dense matching.

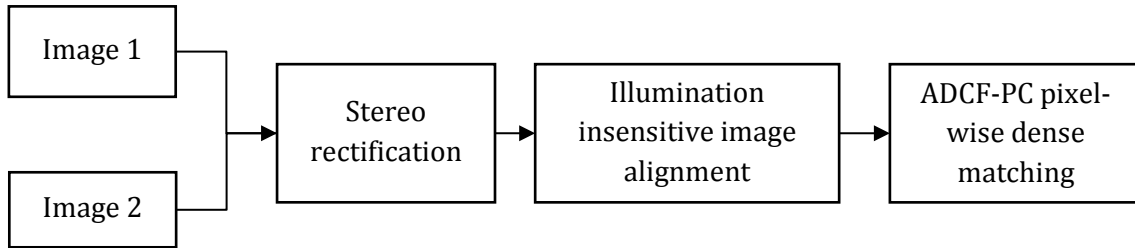


Figure 7.1 Flowchart of proposed disparity map generation approach

7.2.1 Stereo rectification

As indicated in §7.1, most stereo matching algorithms are based on the assumption of camera calibration and epipolar geometry, because if lens distortions and geometric distortions exists, the pixel shifts calculated from dense matching will not purely reflects the depth or 3D information.

For human eye 3D vision, the images captured by two eyes are aligned in the same homography plane to form a stereo vision. However, in reality, the two images taken from two perspective angles for disparity map estimation may not aligned in the same homography plane, as the case shown in Figure 7.2 that the two images

have different attitude angles. In this case, a stereo calibration is firstly carried out. stereo calibration is aimed at the rectification of the initial intersected epipolar lines to parallel lines, as demonstrated in Figure 7.2, so that only disparity values in x direction remains. Another merit of stereo calibration is simplifying the search for corresponding points to one single direction rather than 2D search directions.

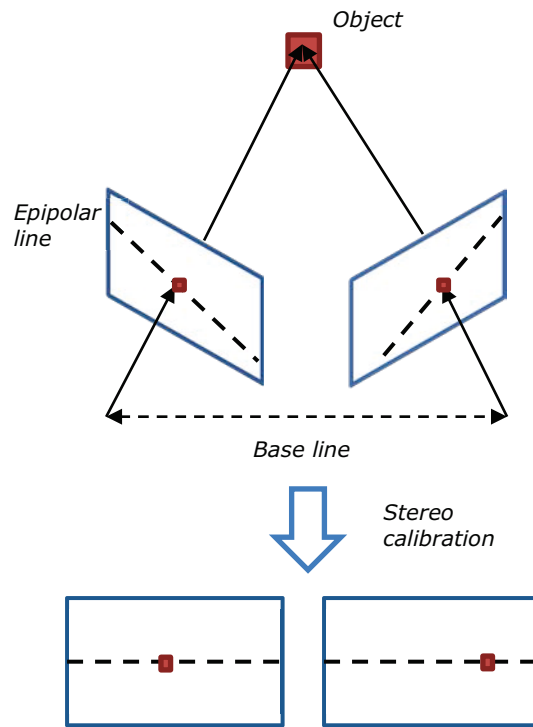


Figure 7.2 Concept of stereo calibration

Stereo calibration has been well studied in computer vision and photogrammetry [105-107], and the key step in stereo calibration is to find adequate corresponding points by feature matching for fundamental matrix estimation. However, as mentioned in §2.1, the feature based matching sometimes do not work for illumination variation cases, so that the number of corresponding points matched by feature matching may not be enough for robust fundamental matrix estimation. This approach used two strategies for illumination-insensitive stereo calibration. For the simulation image experiments, before the generation of illumination variant stereo pairs for 3D reconstruction, two images from different perspective angles but under the same illumination angles were firstly generated and the

fundamental matrix can be correctly calculated through feature based matching. Then we change the illumination conditions, but keep the two view angles as before so the fundamental matrix will not change. Thus, the fundamental matrix calculated under the same illumination condition can be used for stereo calibration for image pair under various illumination conditions. If the image pair under the same illumination condition is not available, 20-30 interest points in the stereo pair were manually selected to ensure an accurate stereo calibration result.

7.2.2 Illumination-insensitive image alignment

After the first step, the initial image pair has been calibrated to obey epipolar geometry so that the pixel shift values in y direction should ideally equals to zero and only the pixel shift values in x direction will be used for disparity map generation. However, the current disparity map is still not purely reflects depth information, it also contains the information of base line, which is the distance of the sensors. The effect of baseline to the stereo pair is adding a constant shifting value to all disparity values, which can be considered as a global image shift within the stereo pair.

For most dense matching algorithms, the effect of baseline do not need to be removed, because the image matching is based on the search of whole epipolar line so the corresponding points will be ultimately found out no matter how large the baseline is. However, PC based dense matching is directly window matching without roaming search, and thus the effect of baseline needed to be removed. According to the research in [108], the errors of PC based matching grow with the increase of measured disparity values, so that the corresponding matching windows needs to be moved as close as possible to ensure a higher matching accuracy. Thus, an illumination-insensitive image alignment proposed is required to remove the global image shift within the stereo pair. We used an illumination-insensitive Piecewise Least Square Fitting Phase Correction (PLSF-PC) based image alignment algorithm which is proposed and detailed described in §4.6.3.

7.2.3 PC based pixel wise dense matching

In this step, the rectified stereo images are matched pixel by pixel via small matching window. The disparity map is generated by PCIAS [109], a C++ software package developed by Dr. H. Yan and Dr. J.G. Liu in our research group in Department of Earth Science and Engineering, Imperial College London. The stereo matching algorithm of PC based sub-pixel disparity map generation is also conducted in Matlab by Gareth Morgan [108].

The pixel wise disparity map estimation from illumination variant stereo pair is based on an ADCF-PC (Absolute Dirichlet function Curve Fitting based Phase Correlation) approach. The robustness of ADCF-PC to illumination variation has been proved in §4.6.2. By using absolute value of Dirichlet function δ_1 , the negative correlation can be corrected to positive. In ADCF-PC, a curve fitting is applied to the absolute Dirichlet function δ_1 in spatial domain. This may not achieve as high sub-pixel accuracy as the PLSF-PC for large window image alignment but it enables quickly small window image matching for disparity estimation.

The image matching experiments in §4.7.2 demonstrated that ADCF-PC approach is more robust than PLSF-PC approach when using the small matching windows. For pixel-wise disparity estimation, the matching window size is usually 32×32 pixels, or 16×16 pixels. When using small matching window, PLSF-PC may fail to achieve a reliable and accurate estimation of disparity as there is not enough points, however, the ADCF-PC still presents a distinctive peak in δ_1 . Thus, ADCF-PC is used as the scanning method for illumination-insensitive disparity map generation.

To reduce the matching errors caused by the increasing disparity, the disparity values introduced by baseline has been removed in §7.2.2. Here a multi-level, coarse to fine image pyramid procedure is applied to generate disparity map at lower resolution firstly and then refined it [108]. PC based matching is firstly performed based on two down-sampling stereo images to generate a low resolution disparity map. In the next image pyramid, the image matching is carried

out based on a higher resolution stereo images but the matching window is relocated based on the initial low resolution disparity map. This procedure ensures that every time the input image matching windows from the stereo images are within one pixel positions, which proved to effectively avoid the matching errors caused by disparity estimation above one pixel [108].

After PC based scanning, an initial disparity map is generated, which may contain 'holes' or 'dirty' areas owing to the matching failures. These areas are removed by a Median Shift Propagation (MSP) technique [12]. The refined disparity map can be converted into DEM using forward intersection based on precise camera parameters.

7.3 Experiments

In this section, we will test our ADCF-PC PC based illumination-insensitive stereo matching algorithm using four different datasets. For comparison, two widely used stereo matching algorithms, SGM (Semi-global Matching) and SAD (Sum of Absolute Difference), were also applied on the same datasets.

Based on the image matching experiments in §4.7.2, the chosen size of matching window for PC is 32×32 pixels for a robust illumination matching performance. The window size for SGM and SAD is 17×17 pixels as the best performance choice in most tested cases in the tested experiments. The above window sizes will be used for all cases in the following experiments.

7.3.1 Hill shading images

(1) Multi-zenith angles

Five hill shading images, shown in Figure 7.3, were generated by ER Mapper image processing software under same azimuth angle (60°) and various zenith angles. Firstly, the DEM was displayed in a 3D perspective view with hill shading under a given illumination direction and the image such formulated was captured as the position 1 image shown in Figure 7.3(a). Then, this DEM 3D perspective display was slightly tilted to formulate an image of another view position, the position 2.

With slightly different view angles, the images captured in these two positions formulate a stereo pair with a baseline equivalent to ~ 80 pixels in x direction and no shift in y direction. Finally, change the zenith angle in the view of position 2 and capture images of different zenith angles but the same imaging geometry. In this way, four hill shading images in position 2 under different zenith angles were generated, as shown in Figure 7.3 (b)-(e).

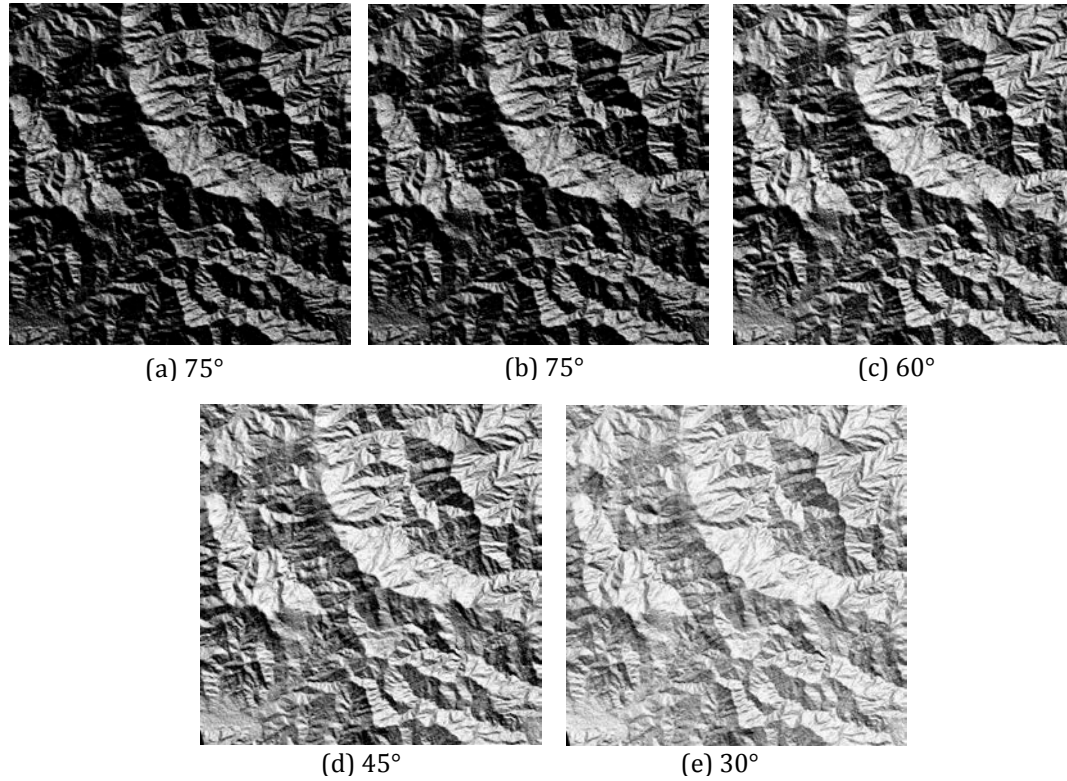


Figure 7.3 Hill shading images of different zenith angles and the same azimuth angle (60°). (a) The hill shading image simulated in position 1. (b)-(e) Hill shading images of different zenith angles simulated in position 2. (Image size: 990×890 pixels)

With Figure 7.3(a) as the reference image and Figure 7.3(b)-(e) as the target images, four varying zenith angle stereo pairs were formulated for DEM generation test using the ADCF-PC stereo matching algorithm. The four disparity maps are shown in Figure 7.4(b)-(e), together with the original DEM (Figure 7.4(a)) for comparison. SGM and SAD based stereo matching algorithms were also applied to the four stereo pairs, and the generated disparity maps are shown in Figure 7.5 and Figure 7.6 respectively.

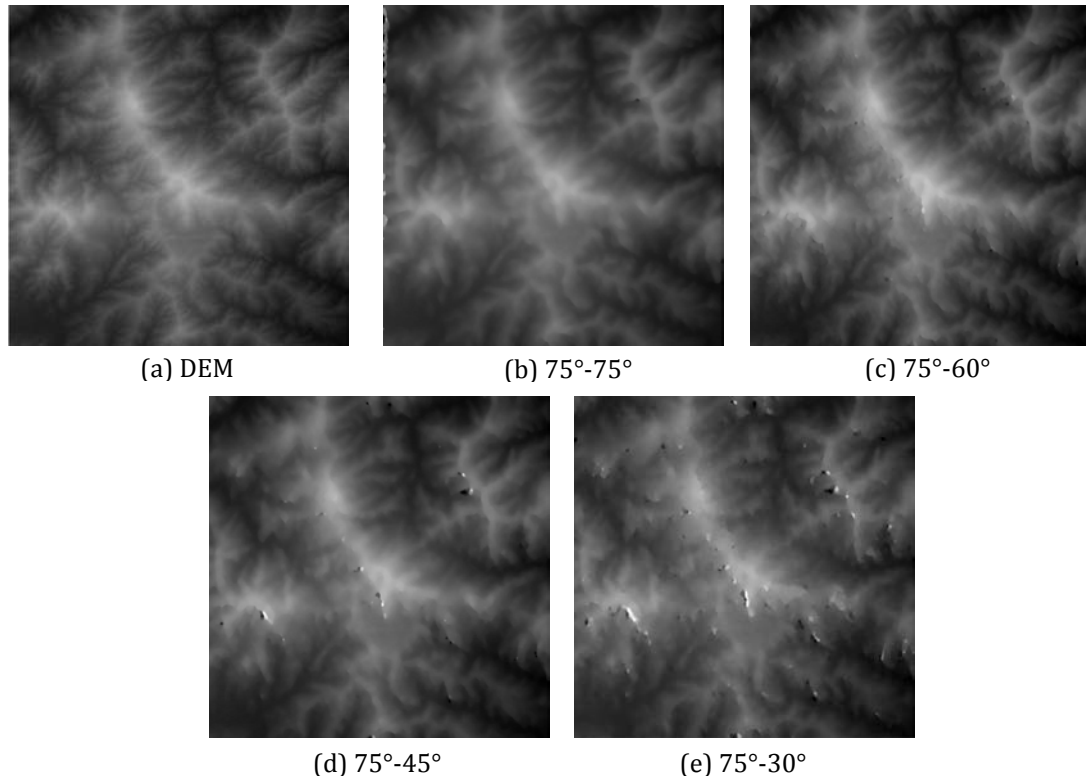


Figure 7.4 Disparity maps generated from the four multi-zenith angle stereo pairs using the ADCF-PC image matching algorithm in comparison with the original DEM data. (a) Original DEM. (b)-(e) Disparity maps.

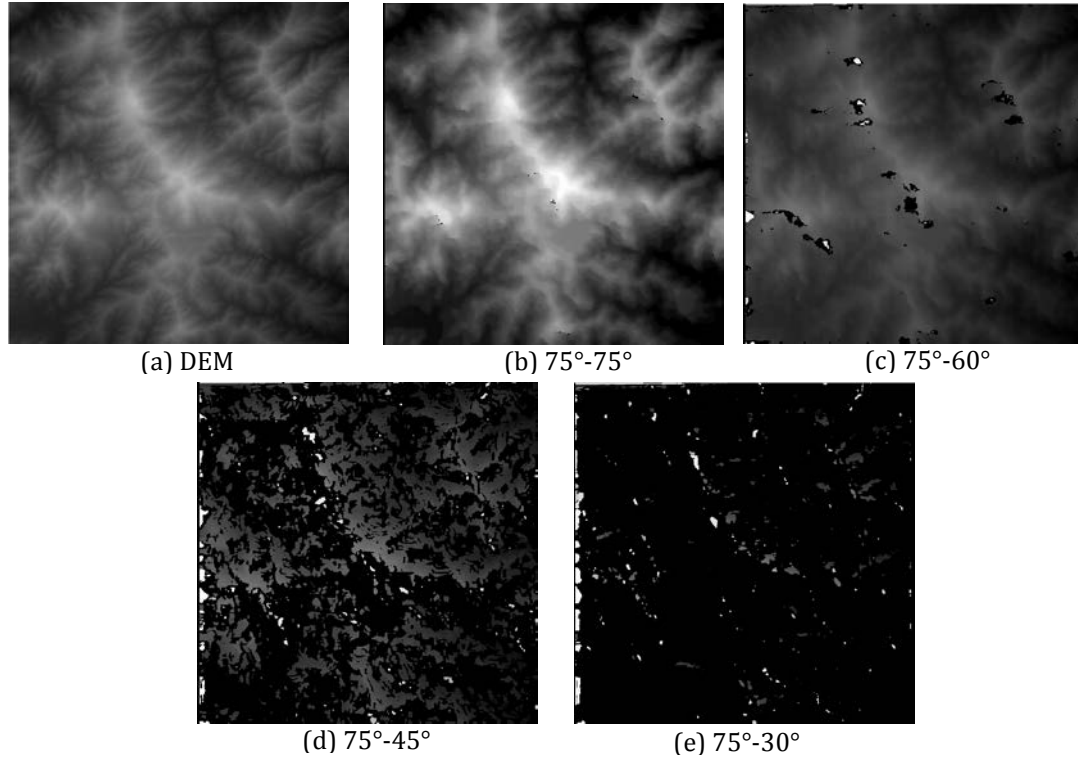


Figure 7.5 Disparity maps generated from the four multi-zenith angle stereo pairs using the SGM algorithm in comparison with the original DEM data. (a) Original DEM. (b)-(e) Disparity maps.

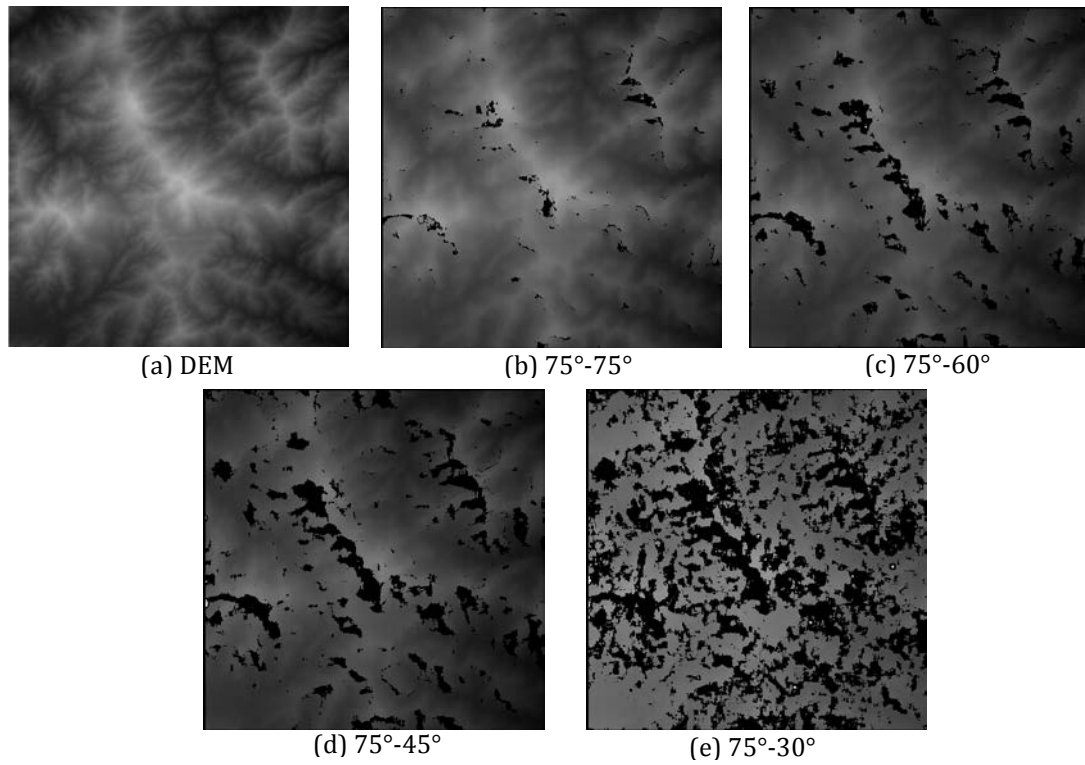


Figure 7.6 Disparity maps generated from the four multi-zenith angle stereo pairs using the SAD algorithm in comparison with the original DEM data. (a) Original DEM. (b)-(e) Disparity maps.

The disparity maps in Figure 7.4 demonstrate that the ADCF-PC image matching algorithm is able to correctly reconstruct the 3D shape of DEM under all tested zenith angle combinations though the quality of disparity maps declines slightly with the increasing zenith angle difference. The areas of ‘dirty’ pixels in Figure 7.4(b)-(e) are mostly distributed along the mountain peaks, where the shading areas are too dark to contain enough image textures within the small matching window to enable effective correlation between the reference and the target images for successful disparity estimation.

Although SGM performed well when there is no illumination difference as shown in Figure 7.5(b), the quality of disparity maps drops drastically with the increase of zenith angle difference. The algorithm fails to generate usable disparity maps when the zenith angle difference is greater than 15° as shown in Figure 7.5(d)-(e).

The SAD algorithm performed more robustly than SGM for large zenith angle difference, such as the case of zenith angle combination of 75° - 45° however, it

became incompetent when the zenith angle difference reaches 35° as shown in Figure 7.6(e). Besides, the quality of the disparity map generated by SAD from the stereo pair of the same illumination angle (Figure 7.6(b)) is lower than those produced by ADCF-PC (Figure 7.4 (b)) and SGM (Figure 7.5(b)). The experimental results shown in Figure 7.4, Figure 7.5 and Figure 7.6 visually demonstrated that ADCF-PC is significantly more robust than SGM and SAD to zenith angle variation, especially in the cases of large zenith angle changes.

For a quantitative comparison between the disparity maps generated by the three algorithms, the NCC coefficients between the original DEM and the disparity maps generated by the three algorithms under different zenith angle combinations were calculated as shown in Table 7.1 and Figure 7.7. The reason for choosing NCC coefficient instead of other pixel-wise evaluation methods, such as RMSE (Root Mean Square Error), is that the disparity maps derived from DEM simulated hill shading stereo images have lower image resolution than the original DEM and thus a pixel-wise evaluation method is deemed “unfair”. The NCC as an area based similarity measurement is insensitive to spatial resolution difference and is therefore appropriate for the assessment.

Table 7.1 NCC coefficients between generated disparity maps and the original DEM using the three algorithms

Zenith Algorithm	75°-75°	75°-60°	75°-45	75°-30°
ADCF-PC	0.9825	0.9821	0.967	0.9484
SGM	0.9414	0.6948	0.214	0.0402
SAD	0.7006	0.5382	0.441	0.2116

Table 7.1 and Figure 7.7 present convincing data to confirm that the ADCF-PC algorithm has consistent robust performance for DEM generation under illumination zenith variation and is significantly superior to the other two algorithms, SGM and SAD, in comparison. For the stereo pair of identical illumination condition, the ADCF-PC achieved the highest NCC coefficient among the three and the NCC=0.9825 indicates that the ADCF-PC generated disparity map is nearly the same as the original DEM. The NCC coefficients of ADCF-PC decline

slightly with the increase of zenith angle difference, but even the lowest correlation value is nearly 0.95, which indicates that the disparity maps have been re-constructed from multi-zenith angle stereo image pairs with high fidelity. In contrast, for SGM and SAD, the NCC coefficients decrease drastically with the increasing zenith angle difference that the both coefficients drop below 0.5 when the zenith angle is larger than 20° . The evaluation confirms that ADCF-PC stereo matching algorithm is able to reconstruct the 3D shape of DEM robustly and accurately under zenith angle variation when other widely used algorithms are deemed failure.

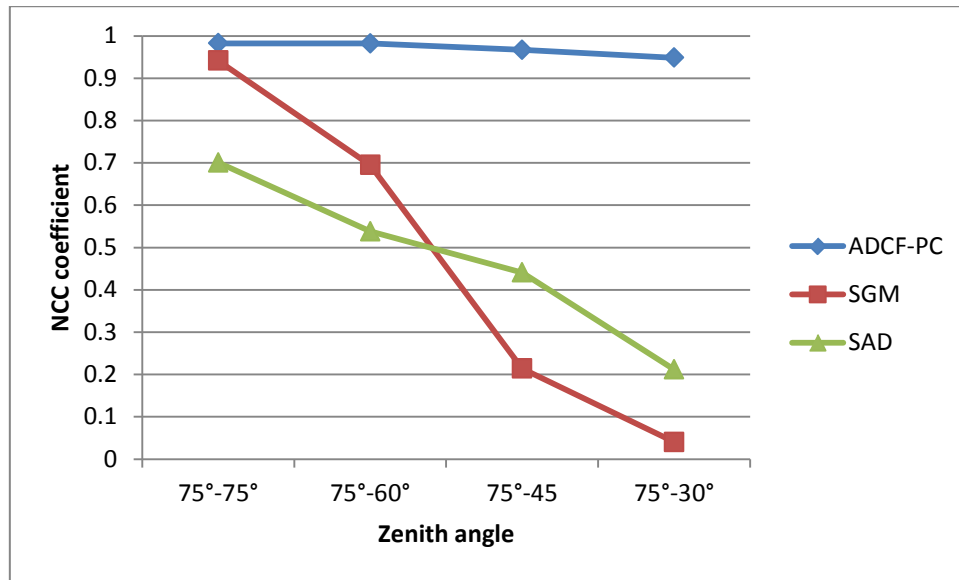


Figure 7.7 NCC coefficient trends of the three algorithms with the increase of zenith angle difference.

(2) Seasonal change of solar illumination

In this experiment, a stereo pair of two hill shading images according to the real sunlight angles in winter and summer was simulated from a DEM. The sunlight angles are calculated according to the latitude (51°) and longitude (0°) of London in 1st January and 1st June. Figure 7.8(a) and (b) are the simulated winter and summer hill shading images. The difference of azimuth angle is 21° and the zenith difference is 42° . Judging from the related image histograms, shown in Figure 7.8(c) and (d), the grey value distributions of the two shading images are very different as the result of the sun illumination angle change.

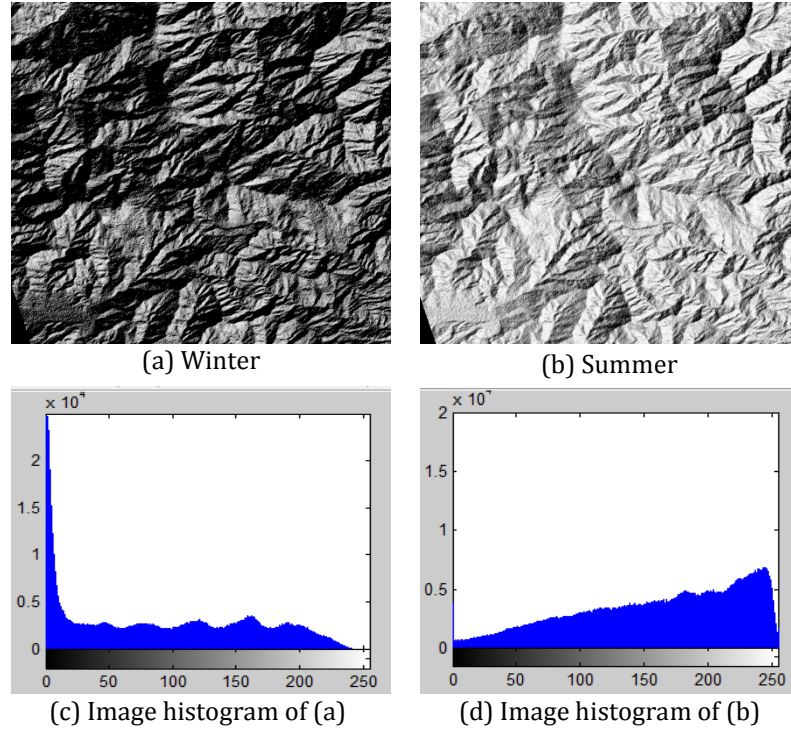


Figure 7.8 Simulated winter and summer hill shading images. (a) The simulated winter hill shading image ($\tau_1=151^\circ$, $\sigma_1=79^\circ$). (b) The simulated summer hill shading image ($\tau_2=130^\circ$, $\sigma_2=37^\circ$). Image size (990×890 pixels) (c) Image histogram of (a). (d) Image histogram of (b).

ADCF-PC based stereo matching was applied to the stereo pair shown in Figure 7.8 to derive the disparity map shown in Figure 7.9(a). For comparison, the disparity maps generated by SGM and SAD are shown in Figure 7.9(b) and (c) respectively. The NCC coefficients between the disparity maps and the original DEM were then calculated to assess the disparity map quality as shown in Table 7.2.

The disparity map generated by the ADCF-PC shown in Figure 7.9(a) fully recovers the topography without visually observable errors though the spatial details are slightly degraded, which is inevitable for data re-generation. The NCC coefficient between the ADCF-PC derived disparity map and the original DEM is very high, $NCC=0.9904$, in Table 7.2. SGM is not robust to the illumination variation in the case as the disparity map shown in Figure 7.9(b) and the NCC coefficient value (0.0418) to the original DEM is very low. SAD is able to reconstruct parts of topography but failed in many scattered parts as shown in Figure 7.9(c). The NCC coefficient of SAD value, $NCC=0.5812$, is slightly higher than SGM but significantly lower than ADCF-PC.

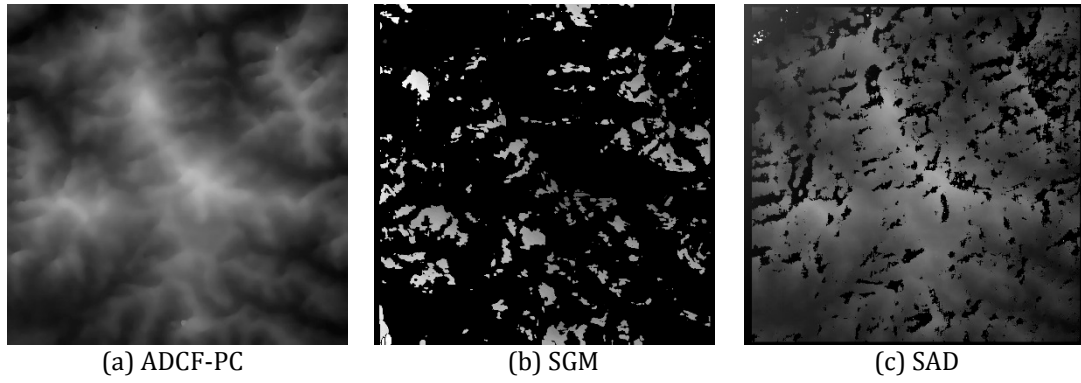


Figure 7.9 Disparity maps generated from winter and summer stereo pair images using the three matching algorithms

Table 7.2 Disparity map quality assessment by NCC coefficients

Algorithm	PC	SGM	SAD
NCC coefficient	0.9904	0.0418	0.5812

To further illustrate the matching accuracy of ADCF-PC, pixel profiles of original DEM and the disparity map generated by ADCF-PC are plotted along a horizontal line and vertical line in Figure 7.10. The shapes of the disparity map and the original DEM match nearly exactly only with slight missing of details. This is because the spatial resolution of disparity maps has to be lower than the original DEM, as the hill shading images are generated from the DEM dataset through a process involving down-sampling.

The stereo matching experiments using DEM hill shading images in comparison with other state-of-the-art algorithms have demonstrate the superior robustness of the ADCF-PC algorithm to different illumination conditions, including large zenith angle change and seasonal solar illumination change.

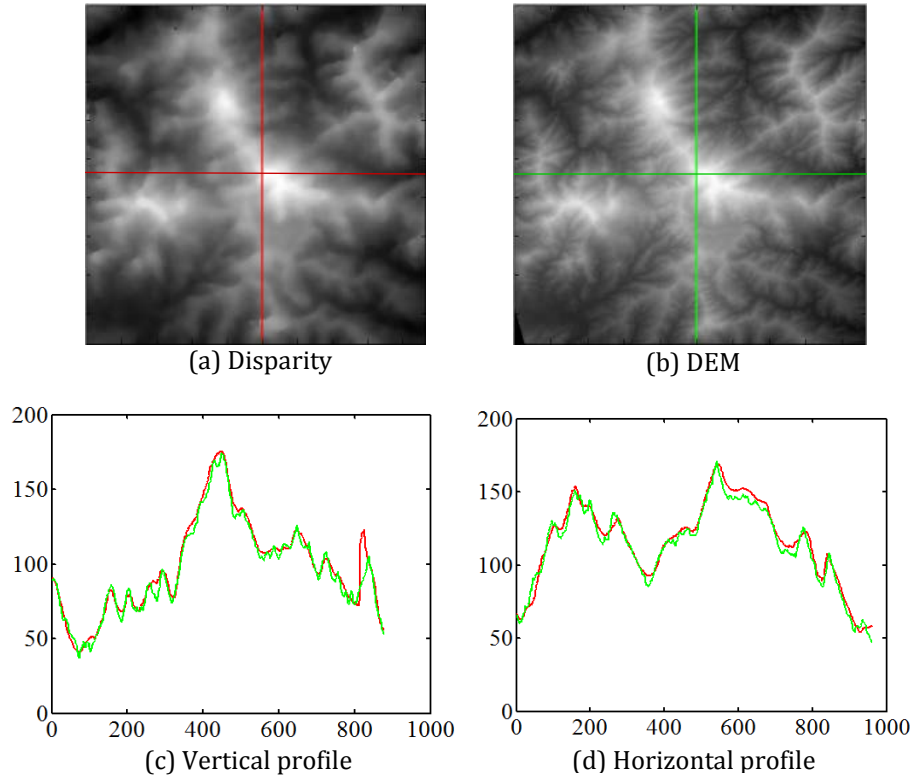


Figure 7.10 Pixel profiles of generated disparity map and original DEM. (a) Disparity map. (b) DEM. (c) Pixel profile in vertical direction. (d) Pixel profile in horizontal direction.

7.3.2 Analogue terrain model images under daily illumination conditions

In this experiment, images were taken from the Snowdonia hill model (shown in Figure 7.11) under daily illumination variation. The experiment configuration has been described in §5.3.2, and stereo images are taken by a binocular setting with two Nikon D70 cameras as shown in Figure 7.11 from 10:30 in the morning till 17:00 in the afternoon. Then, one image taken in 10:30 from camera 1 and another image taken in 17:00 from camera 2 were used to formulate an unconventional illumination variant stereo pair, shown in Figure 7.12.

Compared to solar seasonal change at the same local time, solar daily illumination change is more challenging for DEM generation, because the azimuth angle variation can be as large as 117° in daily solar illumination variation, which is significantly greater than the limited seasonal azimuth change. The large azimuth angle variation results in the local pattern variation as shown in Figure 7.12; the directions and shapes of shades and shadows are greatly altered from Figure

7.12(a) to Figure 7.12(b). According to the discussion in §4.4, solar azimuth angle variation introduce a combination of negative and positive correlation in PC cross power spectrum. Though the negative correlation can be corrected by using the absolute value of δ_1 , the stereo matching robustness and accuracy will inevitably degrade as shown by the image matching experiments using small matching window in §4.7.2. The stereo pair in Figure 7.12 was matched by three stereo matching algorithms, SGM, SAD and ADCF-PC for disparity map generation, and the disparity maps are shown in Figure 7.13.



Figure 7.11 Stereo configuration to take terrain images under real sun illumination.



(a) 10:30 (azimuth=114°,zenith=45°)

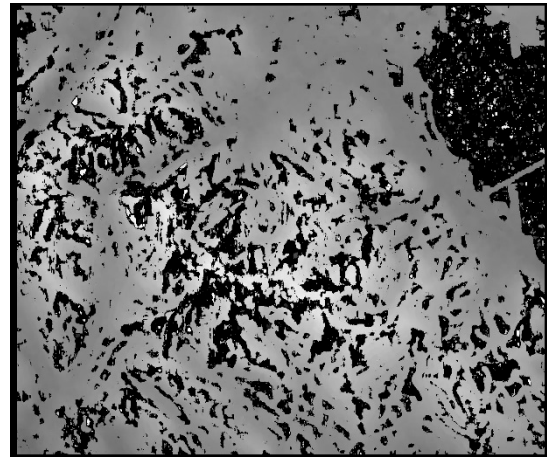


(b) 17:00 (azimuth=261°,zenith=56°)

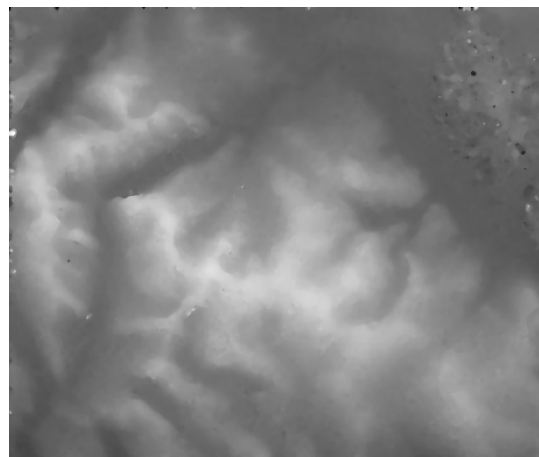
Figure 7.12 Stereo pair of Snowdonia terrain model images taken in 10:30 and 17:00 (image size: 2236×1923 pixels)



(a) SGM



(b) SAD



(c) ADCF-PC

Figure 7.13 Disparity maps generated from SGM, SAD and ADCF-PC using Snowdonia model hill shading images

The disparity maps generated by SGM and SAD shown in Figure 7.13(a) and (b) both contain large areas of matching failure and demonstrate that both SGM and SAD are not robust to daily illumination change in this tested dataset. It is interesting to note that SGM performed slightly better than SAD this time unlike the case in §7.3.1. Although matching errors exist in the mountain areas, SGM succeeded in the low relief areas where the local patterns are not greatly altered by the daily change of illumination. Moreover, SGM performed better than SAD in the featureless water area of a lake in the upper right of the scene. Unlike the DEM simulated hill shading images which only contain topographic features, the Snowdonia terrain model images also contain typical map features which are not altered by the change of illumination. For example, different colour in the model represents different objects; mountains are yellow, vegetation is green and roads are orange. There are also text annotations of the names of the towns and places of interest. These features are independent of illumination conditions, and will be beneficial to the stereo image matching. Thus, the performance of SGM for illumination variant stereo matching is improved compared to its performance in §7.3.1.

Though considerably improved with the help of illumination invariant texture features, the performance of SGM is simply not comparable to ADCF-PC that is able to correctly reconstruct the 3D shape of the Snowdonia model as in Figure 7.13(c). Overall, the disparity map generated by ADCF-PC is smooth and continuous without obvious large patches of matching failures. Scattered matching errors can be observed mainly in the featureless area in the upper right corner of the scene. Even though, compared to the same areas in Figure 7.13 (a) and (b), the disparity map generated by ADCF-PC is the smoothest with least distinctive discontinuities. This is owing to the coarse-to-fine pyramid processing that has been implemented in the ADCF-PC algorithm to achieve accurate and robust disparity estimation via a multi-resolution approach.

For 3D visualization of the disparity map data generated by the ADCF-PC algorithm, Figure 7.12(a) was draped on the disparity map of Figure 7.13(c) to formulate a 3D perspective view of the Snowdonia terrain model as shown in Figure 7.14. Despite of some small matching errors appearing near the image boundary, the overall 3D shape of Snowdonia model has been correctly reconstructed by ADCF-PC.

This experiment has verified the robustness of the ADCF-PC algorithm to daily solar illumination change in real condition rather than simulation. The results confirmed the capability of ADCF-PC to generate 3D data using remote sensing images taken at different daily time.

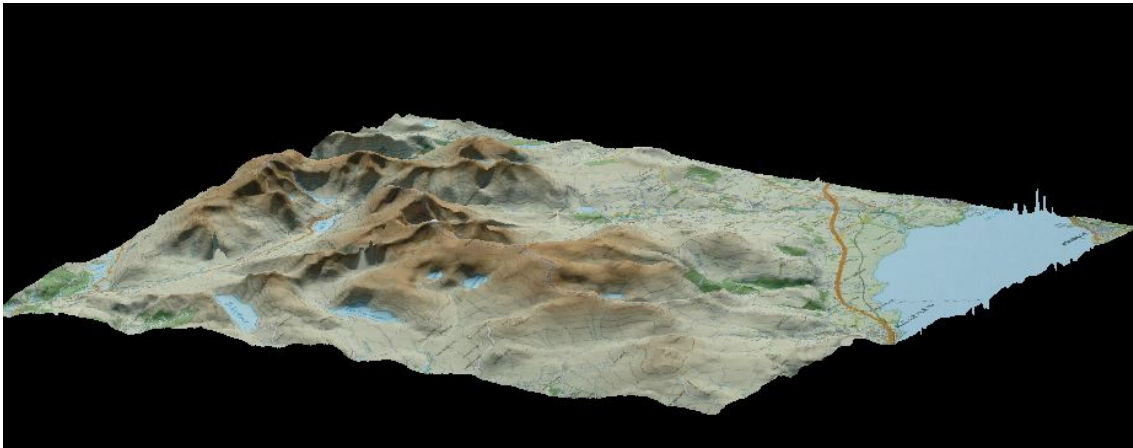


Figure 7.14 The disparity map generated by PC covered by Figure 7.12(a)

7.3.3 Application example: Martian surface 3D reconstruction using HiRISE images

This experiment explores the application potential of the ADCF-PC algorithm to construct 3D model of Mars surface using multi-temporal stereo images taken by Mars' satellite, HiRISE. The 3D data generation of Mars surface using HiRISE images has to rely on multi-temporal stereo image pairs as the sensor system, a bush-broom scanner, does not have stereo imaging capability. We shall demonstrate two examples of Martian surface 3D reconstruction using multi-temporal HiRISE images: a mountain in southwest Melas Chasma and Victoria Crater.

(1) Southwest Melas Chasma

The full image size of a HiRISE image is approximately 40,000×80,000 pixels and the image size may vary owing to the emission angle. To demonstrate the 3D capability of the ADCF-PC stereo matching algorithm under different local illumination angles, a sub-scene of a multi-temporal stereo pair from two HiRISE images was generated as shown in Figure 7.15 and the related observation statistics are in Table 7.3.

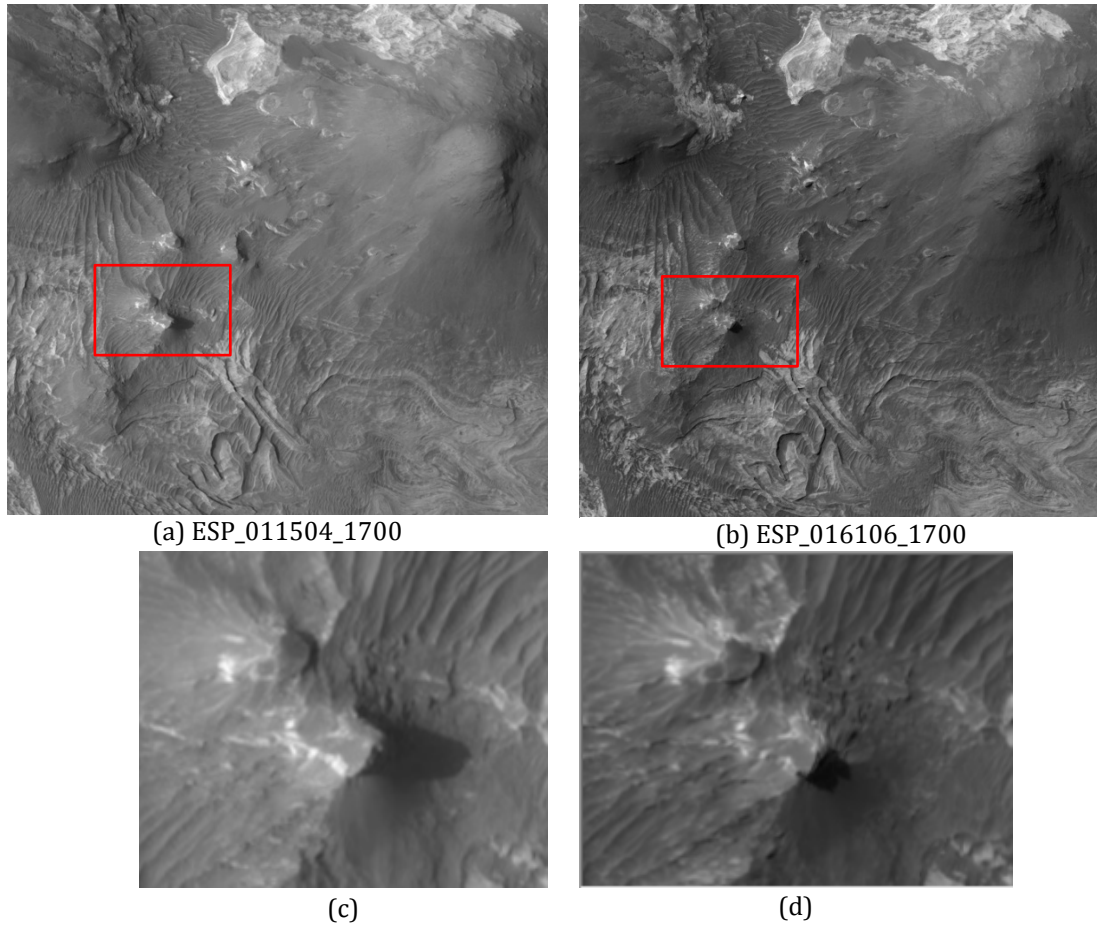


Figure 7.15 (a)-(b) are stereo pairs HiRISE images in Southwest Melas Chasma (image size: 1266×1214 pixels), and (c)-(d) are the enlargement images showing the significant change of shadow caused by local illumination change.

Table 7.3 Observation statistics of the HiRISE stereo images shown in Figure 7.15

	ESP_011504_1700	ESP_016106_1700
Acquisition date	08 Jan 2009	02 Jan 2010
Mars local time	4:01 PM	2:52 PM
Emission angle	60°	49°
Solar azimuth	181.7°	207.1°
Solar zenith	60°	45°

The two stereo images were taken under different sun light angles. According to the statistics in Table 7.3, the solar azimuth angle difference between the two images is 26° and solar zenith angle difference is 15° . The impact of the different sun illumination angles can be clearly seen in Figure 7.15(c)-(d). As Figure 7.15(a) was taken under a relatively large solar zenith angle, the mountain shadows appeared much longer than those in Figure 7.15(b).

Another challenge for HiRISE image stereo matching is that the predominant textures in the images are topographic features, and these topographic features are more sensitive to local illumination change compared to illumination independent texture features as we discussed in §2.1. Thus, a robust illumination-insensitive stereo matching approach is essential for multi-temporal HiRISE image stereo matching.

Figure 7.16 shows, in pseudo colour mode, the disparity map generated from the multi-temporal stereo pair in Figure 7.15 using the ADCF-PC stereo matching algorithm. To demonstrate the 3D reconstruction result, Figure 7.17 presents a 3D perspective view with Figure 7.15(a) draped on the disparity map of Figure 7.16.

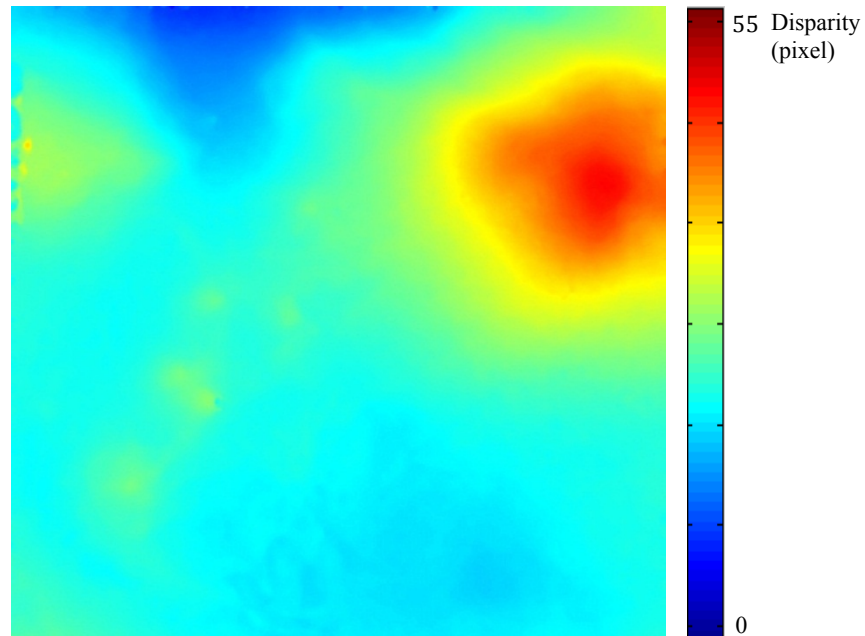


Figure 7.16 Disparity map generated from Figure 7.15 by PC based dense matching and shown in pseudo colour mode.

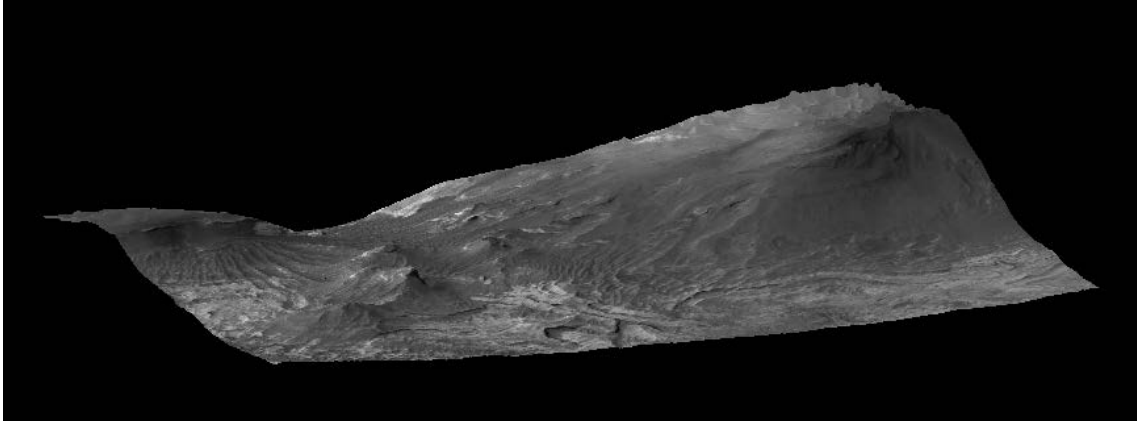


Figure 7.17 3D perspective view of Figure 7.15(a) draped on the disparity map in Figure 7.16.

As shown in Figure 7.16 and Figure 7.17, the 3D Martian surface of the chosen scene has been fairly well reconstructed. The disparity values in plain areas are low and smooth, while those in mountainous areas are high and distinctive. This experiment confirms the potential of the ADCF-PC image matching algorithm for 3D Martian surface DEM generation from multi-temporal, varying illumination stereo images.

(2) Victoria Crater

In this experiment, two HiRISE images of Victoria Crater, shown in Figure 7.18(a) and (b), were chosen for stereo matching. The observation statistics of the two images are in Table 7.4; apart from a half hour acquisition time difference, the two HiRISE images were taken in different months, February and April. The combination of daily and seasonal illumination variation results in 57° azimuth difference and 5° zenith difference between the two images. The effects of azimuth difference can be clearly seen in Figure 7.18: shadows may appear at opposite sides of the crater edge in the two images.

Table 7.4 Observation statistics of the two HiRISE images in Figure 7.18

	ESP_012820_1780	ESP_016644_1780
Acquisition date	21 Apr. 2009	13 Feb. 2010
Mars local time	3:24 PM	3:00 PM
Emission angle	26.1°	0.9°
Solar azimuth	151.9°	208.0°
Solar zenith	54°	39°

Compared to Figure 7.15, the shadows in Figure 7.18 are darker because of the stronger solar strength, and this adds another difficulty in stereo matching as the shadows larger than the matching window may not contain adequate features for matching. Indeed, with thin atmosphere, deep shadows of Martian terrain often contain no features at all.

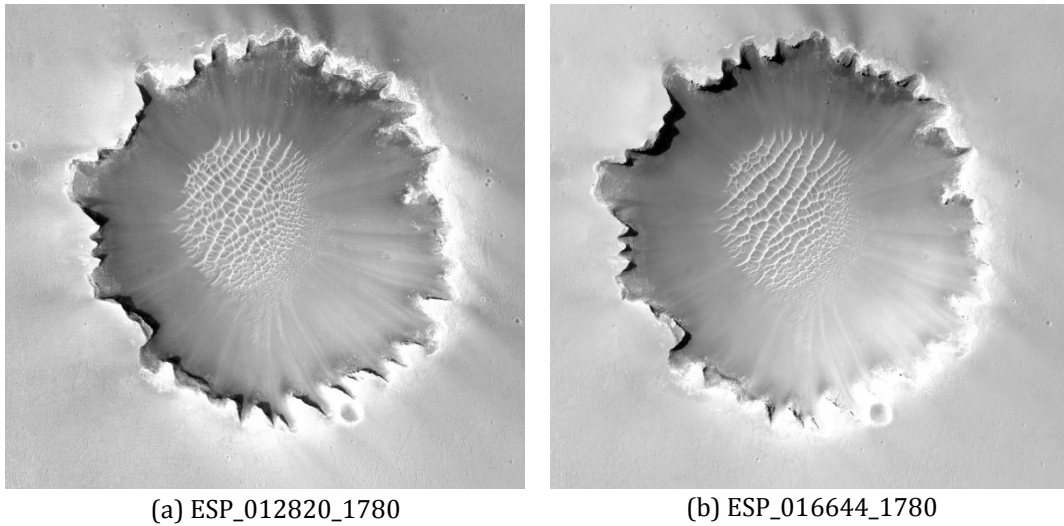


Figure 7.18 Stereo HiRISE image for 3D reconstruction of Victoria Crater. (Image size:4442×4178 pixels)

ADCF-PC stereo matching was then carried out using Figure 7.18(a) and (b) as left and right images. The disparity map generated is shown in Figure 7.19 in pseudo colour and a 3D perspective of the disparity map draped on by the image Figure 7.18(a) is shown in Figure 7.20.

Despite of large illumination variation, especially in azimuth angle, the experiment results shown in Figure 7.19 and Figure 7.20 demonstrate that the 3D shape of Victoria Crater is effectively reconstructed. However, there are several matching failure areas such as a small distinctive bulge feature marked by a red circle in Figure 7.19; the area should be flat and smooth according to Figure 7.18. This matching failure is likely owing to the image matching failure in last few levels of image pyramid.

Another problem in the disparity map is that some crater edges, especially in shadows, are a bit lumpy which should be sharp ridges. This is because the image

matching failed in the shadow areas where there is limited or no information of features within the matching window. Unlike the shadows in EO images, the shadows in Mars images can be almost completely black due to the low atmosphere density. This can be a serious problem for stereo matching when using small matching windows.

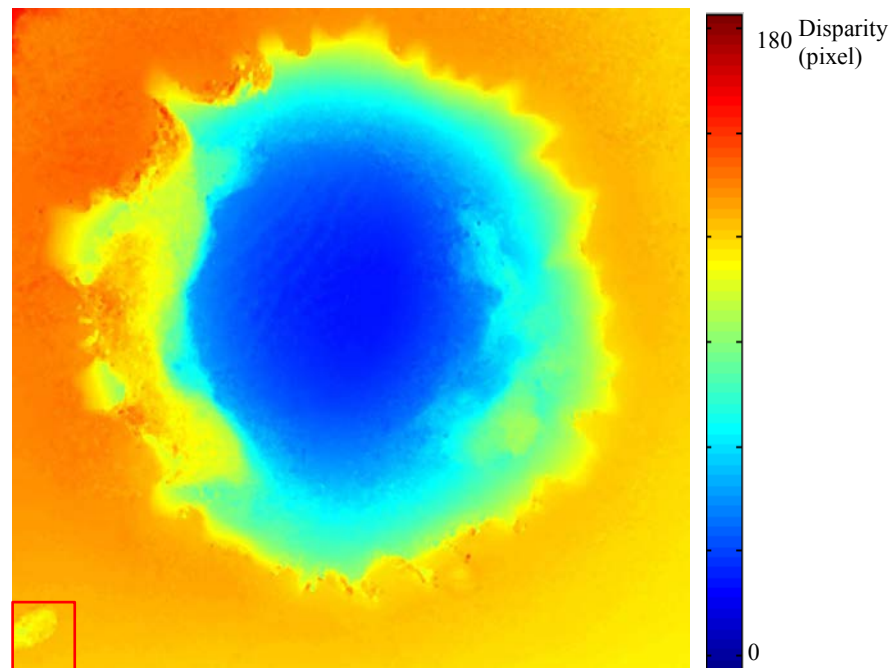


Figure 7.19 disparity map shown in pseudo colour mode

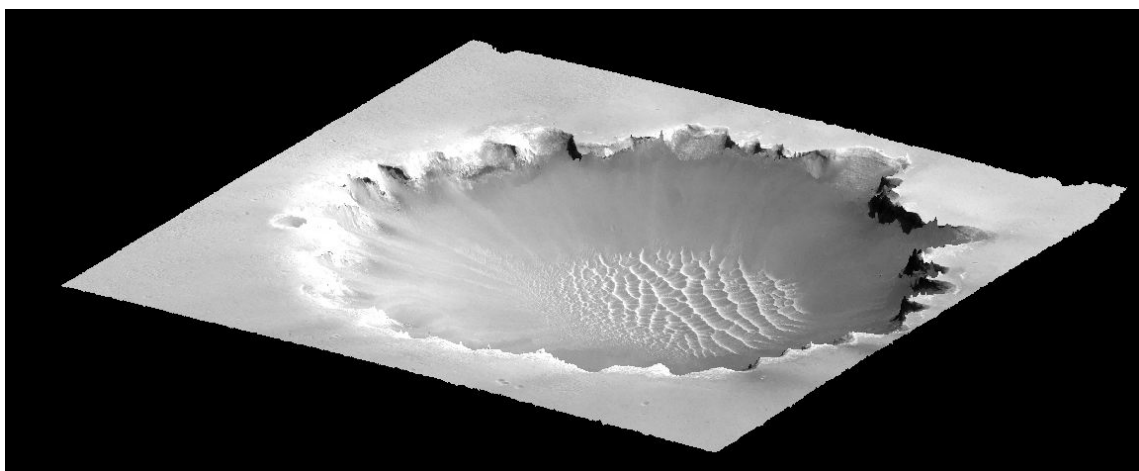


Figure 7.20 3D perspective view of Figure 7.18(a) draped on the disparity map in Figure 7.19.

The two stereo matching experiments in this section demonstrate the potential of PC based stereo matching algorithm to correctly reconstruct the 3D Martian

surface under daily and seasonal solar illumination change. However, the issues of featureless areas and deep shadows need to be resolved by higher radiometric resolution of sensors and further improvement of matching algorithms in order to further improve the 3D data quality.

7.4 Summary

After a detailed description of the proposed ADCF-PC dense image matching approach, several examples have been presented to demonstrate the 3D reconstruction from multi-temporal remote sensing images using the ADCF-PC in comparison with two other most popular in comparison with algorithms, SGM and SAD, and in application potential. The ADCF-PC is the core algorithm of the proposed approach for stereo matching based on its ability of illumination invariance as mathematically proved and experimentally demonstrated in Chapter 4. Stereo matching experiments using four different datasets of stereo images including DEM simulated hill shading images, stereo images of Snowdonia terrain model and HiRISE images of Mars have demonstrated that compared with SGM and SAD, the PC based stereo matching algorithm is the only one highly robust to both daily and seasonal solar illumination variation. While most state-of-the-art stereo matching are based on the assumption of radiometrically similarity, our experiments demonstrated that the ADCF-PC based dense matching algorithm is capable of producing disparity map with high accuracy from unconventional stereo pairs of images taken under significantly different illumination conditions with very difference image appearance.

Chapter 8 Summary

8.1 Summary of research

Illumination variation, especially local illumination variation, as caused by solar position change, is one of the key issues in multi-temporal remotely sensed image matching. Local image patterns, terrain shading and shadows, may be greatly altered by local illumination variation under different sun angles. This makes multi-temporal image matching challenging because parts of images may lose correspondence because of the changes in spatial pattern. This research provides a direct illumination-insensitive image matching solution based on Phase Correlation and explores its applications for autonomous vision-based UAV navigation, change detection and DEM generation from multi-temporal remote sensing.

Chapter 1 discussed the technical challenges in multi-temporal remotely sensed matching, and then divided the illumination change into two types: global illumination change caused by solar intensity and local illumination change caused by variation of solar position. As most image matching algorithms are robust to the linear transformation of image intensity caused by global illumination change but few of them remain robust to the change of local grey value distribution caused by local illumination variation. The following research is then focused on the local illumination-insensitive image matching.

Chapter 2 is a comprehensive review of the state-of-the-art in image matching algorithms of two major types: feature-based and area-based approaches. The discussion particularly focuses on the robustness of these algorithms to local illumination change. As feature-based approaches are generally based on local DN value distributions that may be subject to significant changes in multi-temporal images, they are not very robust to local illumination variations. In comparison,

area-based approaches, such as NCC, MI and PC, utilize the global DN value distribution and they should therefore be more robust to illumination variation. As NCC and MI are iterative image matching methods, while PC is able to directly calculate the image shifts without roaming search, PC is more suitable for fast real-time image matching. This research focuses on the investigation of local illumination invariance of PC.

Chapter 3 proved, by mathematic derivation, that the PC (Phase Correlation) matrix is produced by pixel-wise multiplication of an illumination impact matrix and a translational shift matrix. This relationship, revealed by PC decomposition is further demonstrated by experiments using DEM simulated hill shading images and texture images under various illumination conditions. The key issue is to investigate the effect of local illumination variation in azimuth and zenith to PC matrix.

Chapter 4 revealed that the illumination impact matrix is under a joint effect of azimuth angle and zenith angle variation, based on the proposed theory of *SAI* space. While the fringe patterns in a PC matrix are partially inversed, caused by azimuth angle difference, and with increased noise caused by zenith angle variation, the orientation and density of the PC fringes are not altered. This is the key factor for PC based image matching algorithms to be robust to illumination angle variation. The negative correlation is corrected by using the absolute value of Dirichlet function, so the illumination invariance of widely used ADCF (Absolute Dirichlet Curve Fitting) based PC was proved. Furthermore, a PLSF (Piecewise Least Square Fitting) based PC matching algorithm was proposed, which can achieve higher sub-pixel precision for illumination-insensitive image alignment but it becomes unstable when the matching window is small, in the case for generating disparity maps.

Based on the illumination invariant of two PC-based matching algorithms, ADCF-PC and PLSF-PC, three terrain related applications, based on multi-temporal remotely sensed image matching are presented in chapter 5-7.

Chapter 5 proposed an autonomous UAV navigation scheme, based on the registration of on-board UAV images in a sequence along a flight path to a pre-installed reference satellite image, using the PC based image alignment. UAV experiments using five different datasets, including simulated hill shading images and real remotely sensed images, demonstrated the superior robustness, accuracy and computing efficiency of the proposed PC based UAV navigation approach in comparison with NCC and MI based methods.

Chapter 6 proposed a PC based change detection algorithm under illumination-variant conditions. Experiments using simulated hill shading images and remotely sensed images demonstrated that the proposed approach is able to robustly detect the actual object change despite significant illumination variation and to correctly differentiate appearance changes and object motions. The motion can be estimated at sub-pixel accuracy.

Chapter 7 introduced a illumination-insensitive DEM generation algorithm based on both PLSF-PC and ADCF-PC for multi-temporal remotely sensed data. Stereo matching experiments using four different datasets of stereo images including DEM hill shading images, stereo images of a Snowdonia terrain model and HiRISE images of Mars demonstrated that, compared to two popular image matching algorithms, SGM and SAD, the ADCF-PC image matching algorithm is more robust to illumination change regardless of whether the change is daily or seasonal.

8.2 Original contributions

1. This research provides a thorough mathematical analysis the effects of local illumination variations on PC cross power spectrum in the Fourier frequency domain and thus has proved its illumination-insensitive property, as summarised and illustrated in Figure 8.1. Specifically, three original theoretical contributions to demonstrate phase correlation properties are presented in this research.

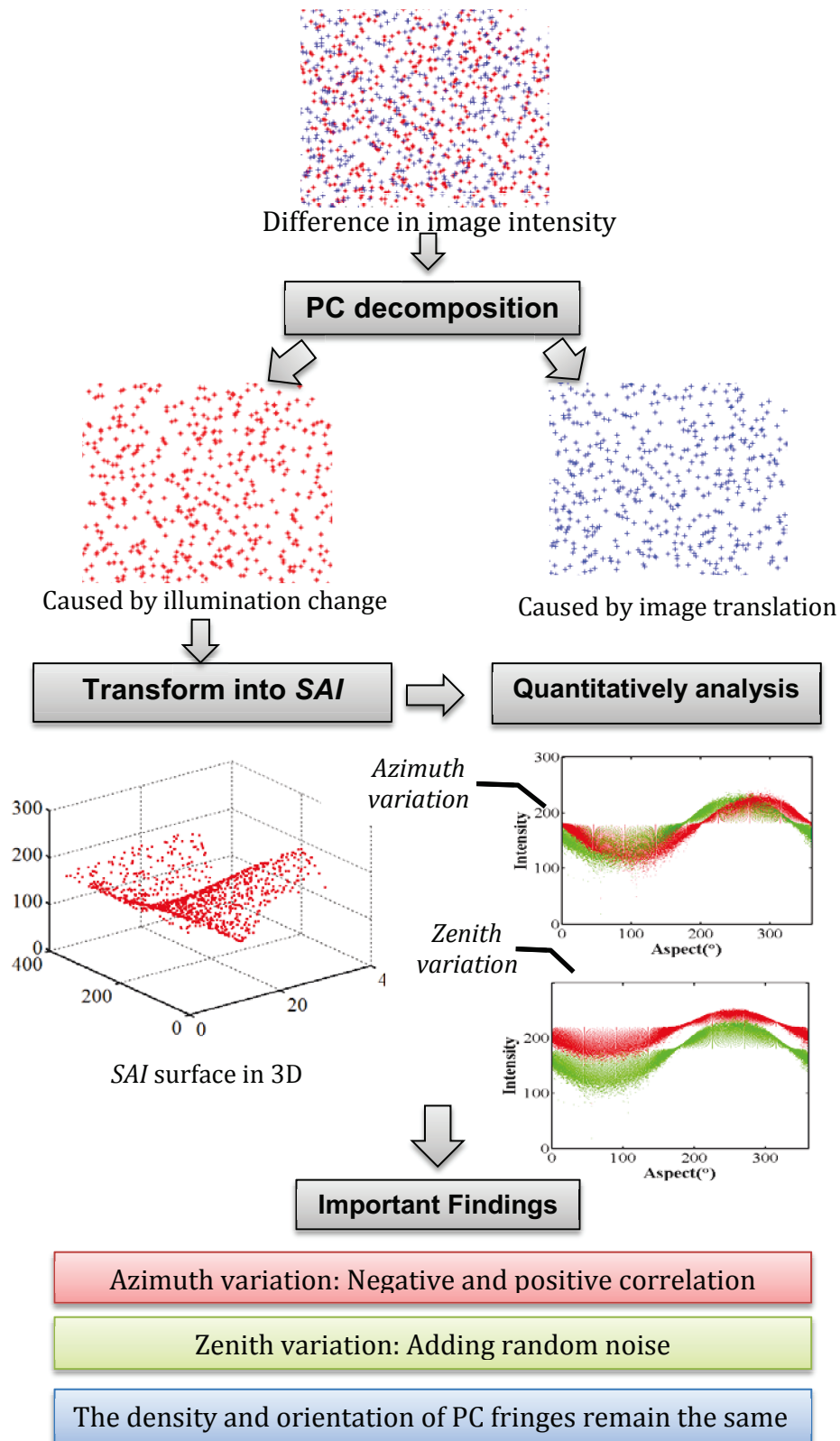


Figure 8.1 Theoretical contributions to investigate the effect of local illumination variation on PC cross power spectrum.

- Phase Correlation matrix decomposition. As demonstrated in Figure 8.1, in the image spatial domain, the image appearance changes caused by illumination variation between the two images for matching (red dots) and those caused by image translation (blue dots) are mixed up. My research has proved that a phase correlation matrix of two similar images with a translation between them can be decomposed as a pixel-wise multiplication of an illumination impact matrix and a translation matrix. The effects of illumination and image translation can thus be distinguished from each other and investigated separately.
- *SAI* space theory. To better present and analyse the illumination impact, an auxiliary 3D space named *SAI* (Slope, Aspect and Intensity) space is proposed. As proved in this thesis, the *SAI* data distribution of a terrain surface image is independent of topography and image translation and only depends on illumination conditions. The shape of *SAI* 3D surface is determined only by illumination conditions and it is therefore an effective tool for the analysis of illumination effect on Phase Correlation. By transforming image intensity values and the related topographic slopes and aspects into *SAI* space, the effects of azimuth and zenith angle on PC can be analysed either separately or jointly.
- Effects of azimuth and zenith angle change on the PC matrix. Using *SAI* as tool, it has been proved that the change in azimuth angle produces a horizontal shift along the aspect axis of *SAI* space while the morphology of the *SAI* surfaces remains the same. The change in zenith angle mainly causes a vertical shift along the intensity axis while the morphology of the *SAI* surfaces largely remain unchanged though not entirely identical, as illustrated in Figure 8.1. Further quantitative analysis and experiments led to three main findings:
 - Azimuth angle variation introduces negative correlation as well as positive correlation, and the angular sectors of negative correlation

are determined by the azimuth angle difference between the two images for matching.

- The effect of zenith angle variation can be likened to the addition of random noise to PC cross power spectrum, while the strength of negative correlation, introduced by the azimuth angle difference, is enhanced by the increase of zenith angle.
- Although the azimuth angle change introduces negative correlation and the zenith angle change adds noise, the orientation and density of the fringe patterns in the PC cross power spectrum in the Fourier frequency domain remain largely the same. This is the key mathematical property which makes PC robust to illumination change in image matching.

2. The above theoretical findings of the effect of local illumination angle change on PC cross power spectrum confirm that regardless of any illumination variation, the orientation and density of PC fringes remain the same, and thus the main issue is the PC fringe reversion caused by negative correlation. Then, two improved PC based matching algorithms: ADCF-PC and PLSF-PC were investigated and compared.

- Firstly, it has been proved that the widely used ADCF-PC algorithm is illumination robust because the negative correlation is inversed by using the absolute value of Dirichlet function in spatial domain.
- Secondly, a PLSF-PC algorithm is proposed, which can achieve higher sub-pixel accuracy for image matching by a piecewise function fit to either negative correlation or positive correlation in frequency domain.
- Experiments using simulated hill shading images generated under various illumination conditions have demonstrated that, compared to ADCF-PC, PLSF-PC is able to reach higher sub-pixel matching accuracy when a large matching window is applied. PLSF-PC is therefore used in illumination-insensitive image alignment, whilst ADCF-PC, being more

robust in small window matching, is used for illumination-insensitive stereo matching.

3. Three original approaches to explore the three terrain related applications - vision-based UAV navigation, illumination-insensitive change detection and DEM generation from multi-temporal remotely sensed images have been developed:

- An illumination-insensitive vision-based UAV navigation scheme. A key advantage of the system is that the illumination invariant property of PC enables effective navigation despite of daily/seasonal variation of solar illumination. This means that a UAV fly in the afternoon can be navigated by matching the real time UAV taken images with a reference satellite image routinely taken in the morning. A fast PC based scanning approach is proposed to speed up the image matching while maintaining robustly accurate localisation. Moreover, a Dirac Delta δ_0 function based registration quality assessment followed by a risk alarming algorithm has been designed and implemented to enable the UAV self-correction when the UAV deviates from the planned route.
- Illumination-invariant change detection. Illumination variation induces cluttered changes of topographic patterns/textures that hinder the detection of true object changes. The proposed change detection approach is based on pixel-wise ADCF-PC matching algorithm; the actual object changes can be detected from illumination variation induced changes of topographic patterns/textures by a joint judgment using the disparity map, the Dirac delta δ_0 peak map and the difference map. The change areas detected are automatically classified into three types: motion, new object and texture change, according to their values in disparity and the Delta δ_0 peak maps.
- A PC based illumination-insensitive DEM generation method from multi-temporal remotely sensed data. Two images taken with a baseline and under different illumination conditions can still formulate an effective

stereo pair for DEM generation using illumination-insensitive image matching algorithms. In the proposed approach, the two stereo images are firstly aligned, using the PLSF-PC, and then pixel-wise dense matching is carried out using ADCF-PC. Experiments using DEM simulated hill shading images and remotely sensed images demonstrated that the combination of these two illumination-insensitive PC algorithms enables robustly high quality stereo matching in an illumination varying environment.

8.3 Implications for multi-temporal remotely sensed image processing

1. The proposed PC based frame-reference matching approach can be applied to UAV navigation for military operations in GPS denied circumstances. It also has potential for planetary exploration, such as on Mars, where GPS is unavailable. Even for civilian tasks on Earth, our vision-based UAV navigation approach can still serve as an alternative for assisted navigation method when GPS signal is weak or unavailable.
2. The work of illumination-insensitive pixel-wise dense matching enables more flexible, reliable and effective change detection without the constraints of consistent time/date/month/season on multi-temporal remote sensing images. Monthly or seasonal changes can be therefore robustly detected and classified. This research has demonstrated the application potential in precise environmental change detection in small time intervals, such as seasonal changes of glaciers, small moving object tracking and monitoring of urban expansion.
3. Current DEM generation methods are in general based on simultaneous stereo image pairs but such data may not be available for generating DEMs of historic topography, while such DEM data are of vital importance for some earth science and environmental research. Archived historical imagery data, such as aerial photos, were often taken not in stereo mode but unconventional stereo image pairs can be formulated from multi-temporal and multi-sensor imagery data. In this case, the illumination-insensitive PC based dense matching

approach is particularly valuable for its ability to generate DEMs from these unconventional stereo pairs of multi-temporal, multi-sensor remotely sensed data.

In this age of 'big data' with rapidly increasing volumes of remotely sensed imagery data acquired by a myriad different sensors at ever shorter time intervals, the ability to utilize multi-source remotely sensed data enriches and enhances the EO capabilities.

8.4 Future work

8.4.1 Forward-view on-board UAV navigation

Forward view on-board imaging will enable a larger FOV compared to the nadir-view, so that the navigation can predict more and farther stops. It also benefits the detection of distant targets. Forward view images are, however, subject to more severe geometric distortions than nadir-view images, and so a more rigorous and versatile geometric transformation is required for a fast and accurate frame-reference matching.

The search images for navigation used in our current navigation scheme are the central parts of the on-board UAV images, which may contain featureless areas. In further development, we will focus on automatically identifying the areas of interest areas to generate search images with distinctive features for a more robust frame-reference matching.

8.4.2 Multi-spectral bands based change detection

In this study, the change detection algorithm is only applied to panchromatic images and thus the change detection excludes an important domain: spectral changes. The change detection will be more effective and richer in information if our current approach is expanded to include multi-spectral images. This will be a direction of further development.

8.4.3 Disparity map refinement

Although the experiments in Chapter 7 demonstrated that the PC based stereo image matching method is able to correctly reconstruct the 3D terrain shapes under illumination variant conditions, the resulting disparity maps may sometimes contain errors above the acceptance levels for real applications and further refinement is therefore required.

The disparity refinement used in the ADCF-PC image matching algorithm is based on image filtering. While the areas of matching errors may not be effectively removed using a small filtering window, a large filtering window will surely degrade topographic edges and the sharpness of the terrain features. Future research will aim at a targeted approach to automatically identify and remove the matching errors while preserve the true terrain edges and features.

List of Figures

Figure 1.1 Multi-temporal satellite images of Tasman Glacier.....	11
Figure 1.2 Two hill images taken under different solar strengths; (a) was taken under a weaker solar strength while (b) a strong solar strength.....	14
Figure 1.3 Definition of zenith angle and azimuth angle	14
Figure 1.4 Shading and shadow variation caused by imaging under different solar positions in a mountainous area. (a) and (b) are Google Earth images taken in different seasons.....	15
Figure 2.1 Illumination test pair from Middlebury stereo datasets with the edge detection results using Canny operator	22
Figure 2.2 Two hill shading images under varying local illumination with the edge detection results using the Canny operator.....	23
Figure 3.1 Example of phase correlation between two images with translational shift and added noise. Image (b) is from image (a) by translational shift of 10 pixels in both horizontal and vertical directions and image (c) is produced by adding Gaussian noise to (a). (d)-(e) Phase correlation cross power spectrum without and with random noise.....	29
Figure 3.2 PC matrix decomposition to the illumination impact matrix and the translation matrix. (a) DEM hill shading image (azimuth=60°, zenith=45°). (b) DEM hill shading image (azimuth=180°, zenith=45°), and shifted by 10 pixels to the right and downwards. (c) Illumination impact PC matrix Q_1 . N: negative correlation; P: positive correlation. (d) Translational PC matrix Q_2 . (e) Phase correlation matrix of (a) and (c): Q_3 . (f) Pixel to pixel multiplication of (c) and (d): Q_1Q_2	33
Figure 3.3 Four different testing datasets for PC matrix decomposition.....	34
Figure 3.4 Error deviations with illumination condition variation.....	35
Figure 3.5 Real dataset 2 with azimuth difference of 120°.....	36
Figure 4.1 Definition of slope and aspect.....	38
Figure 4.2 Hill shading images produced from two DEMs (a) and (b) under the same illumination condition ($\sigma=60^\circ$, $\tau=315^\circ$). (c)-(d): Hill shading image from (a) and (b) respectively. (e): The 3D view of <i>SAI</i> surface in which, the red dots are from	

(c) and the green from (d), and they both cover parts of the theoretical surface under the given illumination shown as the blue surface.....	40
Figure 4.3 Hill shading images under different illumination conditions. (a)-(b): $\sigma=60^\circ$ and $\tau=315^\circ$, (c)-(d): $\sigma=45^\circ$ and $\tau=125^\circ$	41
Figure 4.4 The effects of image shift and rotation on <i>SAI</i> space. (a) Original hill shading image, (b) <i>SAI</i> surface of (a), (c) 10 pixel shifted right and downwards from (a), (d) <i>SAI</i> surface of (c), (e) 90° degree rotation from (a) and (f) <i>SAI</i> surface of (e).	42
Figure 4.5 3D shapes in <i>SAI</i> space of images in aspect view, (a), and in slope view, (b), with the same azimuth angle (150°) but different zenith angles of 60° (red dots) and 45° (green dots).	46
Figure 4.6 Illumination impact matrix and <i>SAI</i> surfaces between two hill shading images with small zenith angles but very different azimuth angles.	50
Figure 4.7 3D shapes in <i>SAI</i> surfaces of images in aspect view, (a), and slope view, (b), with the same zenith angle ($\sigma = 70^\circ$) but different azimuth angles of $\tau_1=125^\circ$ (red dot) and $\tau_2=150^\circ$ (green dot).	51
Figure 4.8 The theoretical plot and the real illumination impact matrix between two hill shading images with different azimuth angles, 60° and 180° , and the same zenith angle 65°	53
Figure 4.9 An example of the effect of 180° azimuth angle difference, 60° and 240° , on illumination impact matrix Q_I	53
Figure 4.10 L_1 term and L_2 term variation with the increase of zenith angles ($S=7^\circ, \tau=60^\circ$)	55
Figure 4.11 (a) <i>SAI</i> surface of constant reflectance; (b) <i>SAI</i> surface of varying reflectance in range of 0.8 - 1.....	56
Figure 4.12 3D shapes in <i>SAI</i> space of images in aspect view, (a), and in slope view, (b), with spectral variation taken under the same zenith angle (70°) but different azimuth angles of 125° (red dots) and 150° (green dots)	57
Figure 4.13 ADCF-PC algorithm robustness to illumination variation: (a) PC fringes between images of different azimuth angles (60° and 120°), (b) δ_I function derived from IFT of (a), (c) The absolute value of δ_I in (b) and (d) The power spectrum fringes of (c) via FT.....	60

Figure 4.14 Original SVD and PLSF-PC in x direction, (a): SVD unwrapping and LSF result of PC of images taken under different illumination conditions. (b): PLSF-PC unwrapping and LSF result.....	61
Figure 4.15 DEM simulated hill shading images with 210° azimuth angle and different zenith angles. The number under each image indicates the zenith angle of the image. (Image size: 1024×1024pixels)	63
Figure 4.16 Phase correlation power spectrums of image pairs with different zenith angles. The numbers under each spectrum indicate the zenith angles of the two images for matching.....	63
Figure 4.17 The peak values of δ for PC matching with different zenith angles.	64
Figure 4.18 DEM simulated hill shading images with different azimuth angles at the same zenith angle (35°). The number under each image indicate the azimuth angle of the image. (Image size: 1024×1024pixels)	65
Figure 4.19 Phase correlation power spectrums of image pairs with different azimuth angles. The numbers under each image indicate the azimuth angles of the two images for matching.....	66
Figure 4.20 (a) Values of δ_o peaks of PC matching with different azimuth angles. (b) A 3D illustration of δ_o peak for the image matching case of azimuth angle 60° and 120°	68
Figure 4.21 Image pair under different illumination conditions for PC based matching. (Image size:1024×1024pixels).....	69
Figure 4.22 PC cross power spectrums generated from two DEM shading images with azimuth angles of 60° and 120° and various zenith angles, ranging from 5° to 80°.....	69
Figure 4.23 Positive, negative and total peak values of the Dirac delta function in respect to the increase of zenith angle.....	71
Figure 4.24 DEM simulated hill shading images under solar illumination conditions of different time on a day. The numbers under each image indicate the imaging times. (Image size: 1024×1024 pixels)	77
Figure 4.25 Matching accuracy of hill shading images under daily illumination variation using six matching algorithms (matching window: 512×512 pixels). The accuracy is set as 3 pixels for all failure cases.....	78
Figure 4.26 Image matching robustness comparison between ADCF-PC and PLSF-PC.....	78

Figure 4.27 Performances of ADCF-PC and PLSF-PC using a small matching window. (a) and (b) are two hill shading images in a 32×32 pixels matching window. (c) is the plot of δ_l function from the IFT of the PC matrix of images (a) and (b). (d) is the plot of SVDs derived from PC matrix of images (a) and (b) and the LSF lines.	79
Figure 5.1 PC-based UAV navigation Workflow	85
Figure 5.2 3D rotation angles of UAV	86
Figure 5.3 Query and search image generation	88
Figure 5.4 δ_o function based wrong estimation elimination	89
Figure 5.5 Comparison of reference image and UAV images of one same scene: (a) is the reference image (scale=0.5) and other four images are UAV images, acquired at (b) 10:00, (c) 12:00, (d) 14:00 and (e) 16:00	93
Figure 5.6 Localisation accuracy with the increase of search intervals	94
Figure 5.7 Localisation accuracy and speed with the increase of search intervals	95
Figure 5.8 Localisation comparison using (a) NCC, (b) MI and (c) PC when UAV flies at 14:00.	98
Figure 5.9 Flight deviation observation results	99
Figure 5.10 (a) Camera image taken in the morning at 10:30 (reference) (b)-(g): smart phone images taken in sequence, in the afternoon at 17:00. The size of camera image is 3008×2000 pixels and of the smartphone images is 3264×2448 pixels.	100
Figure 5.11 Comparison experiment results of UAV navigation produced by (a) MI, (b) NCC and (c) PC algorithms.	102
Figure 5.12 UAV localisation accuracy comparison of MI, NCC and PC algorithms for each stop in (a) in x direction, figure (b) in y direction. The x axis represent UAV stop.	103
Figure 5.13 (a) Reference image generated from ESP_012820_1780 and (b)-(i) simulated UAV images on Mars generated from ESP_016644_1780. (Reference image size: 1098×1302pixels, UAV image size: 500×600 pixels)	105
Figure 5.14 UAV navigation results on Mars.	105
Figure 5.15 (a)-(c) Gatewing UAV images of Maling Mountain, China	106
Figure 5.16 Reference image (Ziyuan-3 satellite image). The footprints of Figure 5.15 (a)-(c) images as blue, red and green polygons on this image.	107

Figure 5.17 A house seen by ZY-3 image and UAV image	107
Figure 5.18 Geometric rectification of a UAV image using POS data.....	108
Figure 5.19 UAV navigation result in Maling Mountain.....	108
Figure 5.20 Aerial image and satellite image for navigation. (Image size:3930×3922 pixels)	109
Figure 5.21 UAV navigation results in Cranfield, UK showing the effect of varying query image size on localisation accuracy.....	111
Figure 6.1 An example of river flow from ice debris from ASTER nadir and back-looking images shown by red-cyan modes over Lawrence River, Quebec City, Canada[89].....	114
Figure 6.2 City appearance change in Dubai from June 2005 to April 2012 from Google Earth.....	114
Figure 6.3 An example of camouflaged changes between two images of a terrain model, (c) and (d) are enlarged images of red box areas in (a) and (b).	115
Figure 6.4 Remote sensing change detection process	117
Figure 6.5 Peak values and peak positions of Dirac delta function δ_o in four different scenarios in change detection.....	121
Figure 6.6 Basic criterion for PC based change detection.....	122
Figure 6.7 Flow chart of PC-based illumination-insensitive change detection	123
Figure 6.8 Pixel-wise PC based dense matching.....	124
Figure 6.9 Flow chart of initial change map generation: (a) and (b) Disparity map in x and y direction respectively; (c) The binary change map generated by thresholding of disparity, in which white areas indicate ‘change’ against black background for ‘no change’; (d) The change map after morphological closure; and (e) Candidate changed areas marked by red squares determined by blob analysis.	126
Figure 6.10 Flowchart of change type determination.....	127
Figure 6.11 Precise boundary identification for new object and texture changes determined by logical combination of difference image and δ_o peak map: (a)-(b) the new object appearance change and related binary change detection image generated; (c)-(d) the texture change and the related binary change detection image.....	129

Figure 6.12 A case of small object change detection in an illumination change environment; (c)-(d) are sub-scene images of areas marked by red boxes above.	130
Figure 6.13 Disparity map and difference map of Figure 6.12(a) and Figure 6.12(b).	130
Figure 6.14 Flow chart for small object change detection	132
Figure 6.15 Appearance change dataset under different local illumination conditions. The appearance change area is marked by a red circle. (Image size: 1024×1024pixels)	134
Figure 6.16 Disparity maps generated from PC based dense matching.....	135
Figure 6.17 (a) Initial change detection map and (b) the candidate change area generation after morphological closure.....	135
Figure 6.18 Final appearance change detection result.....	136
Figure 6.19 2D appearance change detection results under same zenith angle (50°) and different azimuth angles (shown below the images).....	137
Figure 6.20 2D appearance change detection results under same azimuth angle (60°) and different zenith angles shown below the images.....	137
Figure 6.21 DEM dataset for 3D change detection, and the red boxes in (a) and (b) indicate 3D change in area in (c) and (d).....	138
Figure 6.22 Hill shading images with 3D change under different daily illumination conditions at 8:00 and 16:00.....	139
Figure 6.23 PC based illumination-insensitive 3D change detection results	139
Figure 6.24 Object motion detection dataset, (a) is a hill shading image with a car added and (b) is a hill shading image with the car shifted	140
Figure 6.25 Object motion detection result. The first line of the annotation text shows the change type and the second line of the annotation text presents the estimated object motion in x and y direction respectively.....	141
Figure 6.26 Two hill shading images under different illumination conditions (Image size: 368×368 pixels). A 10×10 pixels small object, labelled by a red box, has been inserted to (b). (c) and (d) are sub-scene images of areas marked by the red box.	142
Figure 6.27 δ_0 map generated by PC based pixel-wise dense matching of images shown in Figure 6.26(a) and (b).....	143

Figure 6.28 Small object change detection results by the proposed PC based approach. The changed area is marked by a red box.	143
Figure 6.29 Real sun shading images taken at 09:47 and 15:29 (Image size 428×440 pixels). Three types of changes are implanted and the red box marks the camouflage area. The enlargements of three change targets are shown in (c)-(e).	144
Figure 6.30 Multiple type change detection results. The change areas are marked by boundary boxes and annotation text above the box show the change type identified.....	145
Figure 6.31 Images (694×828 pixels) of Snowdonia model taken at 11:45 and 16:00 with appearance changes. The enlargements of changing targets are shown in (c)-(j).	146
Figure 6.32 Appearance change detection result of Snowdonia model image. The changed areas are marked by cyan boxes and the annotation texts above the boxes indicate the change type.	147
Figure 6.33 Snowdonia model images taken at 11:45 and 14:30 with motion change....	148
Figure 6.34 Motion Detection results of Snowdonia model images. The detected motion areas are marked by yellow boxes and the annotation texts above the boxes indicate the change type and the estimated motion values.	148
Figure 6.35 Lansat 8 panchromatic band image of Hubbard glacier taken in February and April 2014 (Image size 1226×1118 pixels, Image resolution: 15m).....	150
Figure 6.36 Location of Hubbard glacier in Google Earth shown by the red star (Centre latitude: 59°59'31.22"N, Centre Longitude: 139°29'38.30"W)	150
Figure 6.37 Subtraction and ratio images of Figure 6.35(a) and (b)	151
Figure 6.38 Change detection results of Hubbard glacier based on Landsat 8 images. The change areas are marked by square boxes and the annotation texts above the boxes indicate the change type. The red circles drawn in dashed line show some mis-detected change areas.....	151
Figure 6.39 Location of study area, Sohar. (Centre latitude: 24°33'10.37"N, Centre Longitude: 56°43'55.06"E)	153
Figure 6.40 Lansat 8 Panchromatic band image of Sohar acquired in January and July 2014 (Image size 1516×1487 pixels, image resolution:15m).....	153
Figure 6.41 Subtraction and ratio images of Figure 6.40(a) and Figure 6.40 (b).	154

Figure 6.42 Change detection results of Sohar from January to July 2014 based on Landsate 8 images. The change areas are marked by boxes and the annotation texts above the boxes indicate the change type.	155
Figure 6.43 Zoom in images of the yellow box area in Figure 6.42.....	155
Figure 6.44 (a), (b) and (c) are three HiRISE images captured the Opportunity rover in different positions, indicating by the red boxes, image (d), (e) and (f) represent the images the enlargement of the red box areas in (a), (b) and (c).	157
Figure 6.45 Illumination difference and SNR difference in the three HiRISE images	157
Figure 6.46 δ_0 peak map and difference image of ESP_012820_1780 and ESP_016644_1780.	158
Figure 6.47 Rover Detection results using three stereo pairs.....	159
Figure 7.1 Flowchart of proposed disparity map generation approach.....	162
Figure 7.2 Concept of stereo calibration.....	163
Figure 7.3 Hill shading images of different zenith angles and the same azimuth angle (60°). (a) The hill shading image simulated in position 1. (b)-(e) Hill shading images of different zenith angles simulated in position 2. (Image size: 990×890 pixels)	167
Figure 7.4 Disparity maps generated from the four multi-zenith angle stereo pairs using the ADCF-PC image matching algorithm in comparison with the original DEM data. (a) Original DEM. (b)-(e) Disparity maps.	168
Figure 7.5 Disparity maps generated from the four multi-zenith angle stereo pairs using the SGM algorithm in comparison with the original DEM data. (a) Original DEM. (b)-(e) Disparity maps.	168
Figure 7.6 Disparity maps generated from the four multi-zenith angle stereo pairs using the SAD algorithm in comparison with the original DEM data. (a) Original DEM. (b)-(e) Disparity maps.	169
Figure 7.7 NCC coefficient trends of the three algorithms with the increase of zenith angle difference.....	171
Figure 7.8 Simulated winter and summer hill shading images. (a) The simulated winter hill shading image ($\tau_1=151^\circ$, $\sigma_1=79^\circ$). (b) The simulated summer hill shading image ($\tau_2=130^\circ$, $\sigma_2=37^\circ$). Image size (990×890 pixels) (c) Image histogram of (a). (d) Image histogram of (b).	172
Figure 7.9 Disparity maps generated from winter and summer stereo pair images using the three matching algorithms.....	173

Figure 7.10 Pixel profiles of generated disparity map and original DEM. (a) Disparity map. (b) DEM. (c) Pixel profile in vertical direction. (d) Pixel profile in horizontal direction.....	174
Figure 7.11 Stereo configuration to take terrain images under real sun illumination.	175
Figure 7.12 Stereo pair of Snowdonia terrain model images taken in 10:30 and 17:00 (image size: 2236×1923 pixels).....	176
Figure 7.13 Disparity maps generated from SGM, SAD and ADCF-PC using Snowdonia model hill shading images.....	176
Figure 7.14 The disparity map generated by PC covered by Figure 7.12(a).....	178
Figure 7.15 (a)-(b) are stereo pairs HiRISE images in Southwest Melas Chasma (image size: 1266×1214 pixels), and (c)-(d) are the enlargement images showing the significant change of shadow caused by local illumination change.....	179
Figure 7.16 Disparity map generated from Figure 7.15 by PC based dense matching and shown in pseudo colour mode.....	180
Figure 7.17 3D perspective view of Figure 7.15(a) draped on the disparity map in Figure 7.16.	181
Figure 7.18 Stereo HiRISE image for 3D reconstruction of Victoria Crater. (Image size:4442×4178 pixels)	182
Figure 7.19 disparity map shown in pseudo colour mode	183
Figure 7.20 3D perspective view of Figure 7.18(a) draped on the disparity map in Figure 7.19.	183
Figure 8.1 Theoretical contributions to investigate the effect of local illumination variation on PC cross power spectrum.....	188

List of Tables

Table 4.1 Image alignment performances of six matching algorithms under azimuth angle variation.....	73
Table 4.2 Image alignment performances of six matching algorithms under zenith angle variation.....	74
Table 4.3 Solar azimuth and zenith angle variation in on 1 st June.....	76
Table 4.4 Image alignment performances of six matching algorithms under daily illumination angle variations (pixel)	77
Table 5.1 Localisation accuracies at different scales (pixel).....	93
Table 5.2 Localisation accuracy and speed with increase of search intervals.....	95
Table 5.3 Experimental settings for UAV navigation.....	96
Table 5.4 Comparison of localisation average accuracy and speed with different matching algorithms	97
Table 5.5 Detected flight deviations for every stop	100
Table 5.6 Comparison of the accuracy and speed of the MI, NCC and PC algorithm.....	103
Table 5.7 Observation statistics of the two HiRISE images	104
Table 6.1 Object motion detection accuracy under daily illumination variation	141
Table 6.2 Acquisition date and illumination conditions of images in Figure 6.35	151
Table 6.3 Acquisition date and illumination conditions of Figure 6.40(a) and (b)	153
Table 6.4 Observation statistics of the three HiRISE images	156
Table 7.1 NCC coefficients between generated disparity maps and the original DEM using the three algorithms	170
Table 7.2 Disparity map quality assessment by NCC coefficients.....	173
Table 7.3 Observation statistics of the HiRISE stereo images shown in Figure 7.15.....	179
Table 7.4 Observation statistics of the two HiRISE images in Figure 7.18.....	181

List of Abbreviations

3D	3 Dimension
ADCF-PC	Absolute Dirichlet Curve Fitting Phase Correlation
ANCC	Adaptive Normalised Cross Correlation
BCET	Balance Contrast Enhancement Technique
DEM	Digital Elevation Model
DN	Digital Number
EK	Extended Kalman
EO	Earth Observation
FOV	Field of View
FT	Fourier Transform
GDEM	Global Digital Elevation Model
HiRISE	High Resolution Imaging Science Experiment
IFT	Inversed Fourier Transform
IMU	Inertial Measurement Unit
LCE	Linear Contrast Enhancement
MI	Mutual Information
MSP	Median Shift Propagation
N	Negative correlation
NCC	Normalised Cross Correlation
NP	A combination of positive and negative correlation
P	Positive correlation
PC	Phase Correlation
POS	Position Orientation System
QI	Quotient Image

QMDPE	Quick Maximum Density Power Estimator
RMSE	Root Mean Square Error
SAD	Sum of Absolute Difference
SAI	Slope Aspect Intensity
SE	South East
SGM	Semi-Global Matching
SIFT	Scale Invariant Feature Transform
SLAM	Simultaneous Localisation and Mapping
SNR	Signal Noise Ratio
SSD	Sum of Squared Differences
SSDA	Sequential Similarity Detection Algorithm
SURF	Speeded Up Robust Features
SVD	Singular Value Decomposition
SW	South West
UAS	Unmanned Aerial System
UAV	Unmanned Aerial Vehicle
ZY-3	Ziyuan-3

List of Symbols

I_1, I_2	Matched image pair
MI	Mutual information function
H	Entropy of image
(a, b)	Translation shift between matched images
F	Fourier transform function
(x, y)	Coordinates in the image spatial domain
(u, v)	Coordinates in the frequency domain
Q	Normalised cross power spectrum
$*$	Complex conjugate
δ_0	Dirac delta function
$\emptyset$	Phase angle
Q_1	Illumination impact matrix
Q_2	Translation matrix
E	Error estimation function
$RMSE$	Root mean square error
R	Real part of a complex number
I	Imaginary part of a complex number
γ	Error deviation
L	Solar illumination
τ	Azimuth angle
σ	Zenith angle
(p, q)	First order partial derivatives of DEM
V_H	Map of height or DEM
S	Terrain slope angle

A	Terrain aspect angle
ω	Angular frequency of the Fourier component
θ	Compass direction with respect to 0°
L_1, L_2	Functions of zenith angle to determine the strength of NP
E	Solar irradiance
ρ	Spectral reflectance
δ_1	Dirichlet function
G_{cf}	Gaussian fitting function
$\lambda_1, \dots, \lambda_6$	Coefficients for Gaussian fitting function
Q_x, Q_y	Dominant singular vector decomposed from Q_2
D_{abs}	Absolute value of Dirichlet function
Q_{abs}	Cross power spectrum generated by D_{abs}
PL_f	Piecewise least square fitting function
LSF	Least square fitting function
T_{LSF}	Threshold for piecewise least square fitting
v	Matching accuracy
v_x, v_y	Matching accuracies in x and y direction
R_M	Rotation matrix
(α, β, γ)	Rotation angles of camera from reference axes x, y and z
q	Query image
(w, h)	Width and height of image
$imcrop$	Image cropping algorithm
R_1	Size of query image
R_2	Size of search image
(b_x, b_y)	Distance of two stops in on-board UAV images in x and y direction
(PC_{sc})	Fast PC scanning algorithm
(X_i, Y_i)	Calculated UAV position in stop i

(X_i', Y_i')	Planned UAV position in stop i
(e_x, e_y)	Deviations between calculated and planned UAV positions in x and y directions
S	Difference image
C	Change map
thr	Threshold for change
I_{nor_1}	Normalised image
R	Ratio image
D_x, D_y	Disparity maps in x and y directions
P	δ_0 peak map
T_d	Threshold for disparity map
T_b	Threshold for 'blob' area
T_{ph}	Threshold for motion detection based on δ_0 peak map
T_{pl}	Threshold for appearance change based on δ_0 peak map
T_s	Threshold for appearance change based on difference map
T_{b1}, T_{b2}	Threshold for 'blob' area in small object change detection
B	Baseline

Reference

- [1] J. G. Liu and P. Mason, *Essential image processing and GIS for remote sensing*: John Wiley & Sons, 2013.
- [2] C. Pascual, F. C. Juan, M. Ivan, M. Carol, O. Miguel, M. Luis, *et al.*, "Computer Vision Onboard UAVs for Civilian Tasks," *Journal of Intelligent and Robotic Systems*, vol. 54, 2008.
- [3] H. Bendea, P. Boccardo, S. Dequal, F. Giulio Tonolo, D. Marenchino, and M. Piras, "Low cost UAV for post-disaster assessment," pp. 3-11, 2008.
- [4] S. M. Adams and C. J. Friedland, "A Survey of Unmanned Aerial Vehicle (UAV) Usage for Imagery Collection in Disaster Research and Management," in *Proceedings of the Ninth International Workshop on Remote Sensing for Disaster Response*, 2012.
- [5] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, pp. 977-1000, 2003.
- [6] T. Toutin, "Review article: Geometric processing of remote sensing images: models, algorithms and methods," *International Journal of Remote Sensing*, vol. 25, pp. 1893-1924, 2004.
- [7] L. G. Brown, "A survey of image registration techniques," *ACM computing surveys (CSUR)*, vol. 24, pp. 325-376, 1992.
- [8] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *Image Processing, IEEE Transactions on*, vol. 14, pp. 294-307, 2005.
- [9] M. Madani, "Real-time sensor-independent positioning by rational functions," in *Proceedings of ISPRS workshop on direct versus indirect methods of sensor orientation*, 1999, pp. 25-26.
- [10] A. Okamoto, "Orientation and construction of models. III- Mathematical basis of the orientation problem of one-dimensional central perspective photographs," *Photogrammetric Engineering and Remote Sensing*, vol. 47, pp. 1739-1752, 1981.
- [11] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *IJCAI*, 1981, pp. 674-679.
- [12] J. G. Liu and H. Yan, "Phase correlation pixel-to-pixel image co-registration based on optical flow and median shift propagation," *International Journal of Remote Sensing*, vol. 29, 2008.
- [13] G. Jianya, S. Haigang, M. Guorui, and Z. Qiming, "A review of multi-temporal remote sensing data change detection algorithms," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 757-762, 2008.
- [14] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, *et al.*, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, pp. 355-368, 1987.
- [15] L. J. Guo, "Balance contrast enhancement technique and its application in image colour composition," *Remote Sensing*, vol. 12, pp. 2133-2151, 1991.
- [16] A. Shashua and T. Riklin-Raviv, "The quotient image: Class-based re-rendering and recognition with varying illuminations," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, pp. 129-139, 2001.

- [17] H. Wang, S. Z. Li, and Y. Wang, "Face recognition under varying lighting conditions using self quotient image," in *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, 2004, pp. 819-824.
- [18] Y. Zhang, J. Tian, X. He, and X. Yang, "MQI based face recognition under uneven illumination," in *Advances in Biometrics*, ed: Springer, 2007, pp. 290-298.
- [19] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible illumination conditions?," *International Journal of Computer Vision*, vol. 28, pp. 245-260, 1998.
- [20] A. S. Georgiades, D. Kriegman, and P. Belhumeur, "Illumination cones for recognition under variable lighting: Faces," in *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, 1998, pp. 52-58.
- [21] W. Maddern, A. D. Stewart, C. McManus, B. Upcroft, W. Churchill, and P. Newman, "Illumination Invariant Imaging: Applications in Robust Vision-based Localisation, Mapping and Classification for Autonomous Vehicles," in *Proc. of Workshop on Visual Place Recognition in Changing Environments, IEEE International Conference on Robotics and Automation (ICRA)*, 2014.
- [22] H. Han, S. Shan, L. Qing, X. Chen, and W. Gao, "Lighting aware preprocessing for face recognition across varying illumination," in *Computer Vision-ECCV 2010*, ed: Springer, 2010, pp. 308-321.
- [23] H. G. Barrow and J. M. Tenenbaum, "Interpreting line drawings as three-dimensional surfaces," *Artificial intelligence*, vol. 17, pp. 75-116, 1981.
- [24] J. Canny, "A computational approach to edge detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 8, pp. 679-698, 1986.
- [25] R. M. Haralick, "Digital step edges from zero crossing of second directional derivatives," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, pp. 58-68, 1984.
- [26] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)*, vol. 207, pp. 187-217, 1980.
- [27] R. Brunelli and T. Poggio, "HyperBF Networks for Real Object Recognition," in *IJCAI*, 1991, pp. 1278-1285.
- [28] V. Govindaraju, D. B. Sher, R. K. Srihari, and S. N. Srihari, "Locating human faces in newspaper photographs," *IEEE*, pp. 549-554, 1989.
- [29] O. Arandjelovic, "Gradient Edge Map Features for Frontal Face Recognition under Extreme Illumination Changes," in *BMVC*, 2012, pp. 1-11.
- [30] G. Baatz, O. Saurer, K. Köser, and M. Pollefeys, "Large scale visual geo-localization of images in mountainous terrain," in *Computer Vision-ECCV 2012*, ed: Springer, 2012, pp. 517-530.
- [31] C. Gräßl, T. Zinßer, and H. Niemann, "Illumination insensitive template matching with hyperplanes," in *Pattern Recognition*, ed: Springer, 2003, pp. 273-280.
- [32] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, pp. 91-110, 2004.
- [33] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," *Springer*, pp. 404-417, 2006.
- [34] J. Krizaj, V. Struc, and N. Pavesic, "Adaptation of SIFT features for face recognition under varying illumination," *IEEE*, pp. 691-694, 2010.

- [35] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, "On the use of SIFT features for face authentication," in *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, 2006, pp. 35-35.
- [36] H. Hirschmuller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, 2007, pp. 1-8.
- [37] A. J. Glover, W. P. Maddern, M. J. Milford, and G. F. Wyeth, "FAB-MAP+ RatSLAM: appearance-based SLAM for multiple times of day," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, 2010, pp. 3507-3512.
- [38] W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 36, pp. 458-466, 2006.
- [39] S. Du and R. Ward, "Wavelet-based illumination normalization for face recognition," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, 2005, pp. II-954-7.
- [40] S. L. Wijaya, M. Savvides, and B. Vijaya Kumar, "Illumination-tolerant face verification of low-bit-rate JPEG2000 wavelet images with advanced correlation filters for handheld devices," *Applied optics*, vol. 44, pp. 655-665, 2005.
- [41] M. Savvides, B. Kumar, and P. K. Khosla, "" Corefaces"-Robust Shift Invariant PCA Based Correlation Filter for Illumination Tolerant Face Recognition," *IEEE Computer Society*; 1999, vol. 2, p. 834, 2004.
- [42] M. Savvides, B. V. Kumar, and P. K. Khosla, "Robust shift-invariant biometric identification from partial face images," in *Defense and Security*, 2004, pp. 124-135.
- [43] A. Papoulis and S. U. Pillai, "Probability, random variables and stochastic processes with errata sheet," *New York, NY, McGraw-Hill Education*, 2002.
- [44] C. Kuglin, "The Phase Correlation Image Alignment Method," in *Proc. Int. Conference Cybernetics Society*, 1975, pp. 163-165.
- [45] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration," *Computers, IEEE Transactions on*, vol. 100, pp. 179-186, 1972.
- [46] Y. S. Heo, K. M. Lee, and S. U. Lee, "Illumination and camera invariant stereo matching," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008, pp. 1-8.
- [47] J. Liu and J. M. Moore, "Hue image RGB colour composition. A simple technique to suppress shadow and enhance spectral signature," *International Journal of Remote Sensing*, vol. 11, pp. 1521-1530, 1990.
- [48] O. Faugeras, B. Hotz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, *et al.*, "Real time correlation-based stereo: algorithm, implementations and applications," Inria1993.
- [49] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *Medical Imaging, IEEE Transactions on*, vol. 16, pp. 187-198, 1997.
- [50] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *International journal of computer vision*, vol. 24, pp. 137-154, 1997.
- [51] G. Egnal, "Mutual information as a stereo correspondence measure," *Technical Reports (CIS)*, p. 113, 2000.
- [52] J. Kim, V. Kolmogorov, and R. Zabih, "Visual correspondence using energy minimization and mutual information," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, 2003, pp. 1033-1040.

- [53] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, pp. 807-814.
- [54] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, pp. 529-541, 1981.
- [55] M. Savvides, B. V. Kumar, and P. K. Khosla, "Eigenphases vs eigenfaces," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, 2004, pp. 00810-00813 Vol. 3.
- [56] J. Heo, M. Savvides, and B. Vijayakumar, "Illumination tolerant face recognition using phase-only support vector machines in the frequency domain," in *Pattern Recognition and Image Analysis*, ed: Springer, 2005, pp. 66-73.
- [57] B. S. Reddy and B. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *Image Processing, IEEE Transactions on*, vol. 5, pp. 1266-1271, 1996.
- [58] J. G. Liu, H. Yan, and G. Morgan, "PCIAS Subpixel Technology," *Measurement and Control*, vol. 45, pp. 207-211, 2012.
- [59] W. Hoge, "A subspace identification extension to the phase correlation method," *IEEE transactions on medical imaging*, vol. 22, pp. 277-280, 2003.
- [60] H. Foroosh, J. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 11, pp. 188-200, 2002.
- [61] G. L. K. Morgan, J. G. Liu, and H. Yan, "Precise subpixel disparity measurement from very narrow baseline stereo," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, pp. 3424-3433, 2010.
- [62] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, and T. Stathaki, "Robust FFT-based scale-invariant image registration with image gradients," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, pp. 1899-1906, 2010.
- [63] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *Image Processing, IEEE Transactions on*, vol. 11, pp. 188-200, 2002.
- [64] P. Kube and A. Pentland, "On the imaging of fractal surfaces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 10, pp. 704-707, 1988.
- [65] H. Wiechel, "Theorie und Darstellung der Beleuchtung von nicht gesetzmässig gebildeten Flächen mit Rücksicht auf die Beiezeichenaus," ed: Civllingenieur, 1878.
- [66] V. Argyriou and T. Vlachos, "A Study of Sub-pixel Motion Estimation using Phase Correlation," in *BMVC*, 2006, pp. 387-396.
- [67] J. Liu and H. Yan, "Robust phase correlation methods for sub-pixel feature matching," *Proc. 1st Annual Conference of Systems Engineering ...*, 2006.
- [68] J. Zhang, W. Liu, and Y. Wu, "Novel Technique for Vision-Based UAV Navigation," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 47, pp. 2731-2741, 2011.
- [69] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *Robotics & Automation Magazine, IEEE*, vol. 13, pp. 108-117, 2006.
- [70] M. Bosse, "A vision augmented navigation system for an autonomous helicopter," 1997.
- [71] S. Merhav and Y. Bresler, "On-line vehicle motion estimation from visual terrain information Part I: Recursive image registration," *Aerospace and Electronic Systems, IEEE Transactions on*, pp. 583-587, 1986.

- [72] Y. Bresler and S. Merhav, "On-line vehicle motion estimation from visual terrain information Part II: Ground velocity and position estimation," *Aerospace and Electronic Systems, IEEE Transactions on*, pp. 588-604, 1986.
- [73] Y. Lin and G. Medioni, "Map-enhanced UAV image sequence registration and synchronization of multiple image sequences," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, 2007, pp. 1-7.
- [74] P. Lothe, S. Bourgeois, F. Dekeyser, E. Royer, and M. Dhome, "Towards geographical referencing of monocular slam reconstruction using 3d city models: Application to real-time accurate vision-based localization," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 2882-2889.
- [75] W. Jan, M. Oliver, S. Christian, and F. T. Gert, "An integrated GPS/MEMS-IMU navigation system for an autonomous helicopter," *Aerospace Science and Technology*, vol. 10, 2006.
- [76] J. J. Rodriguez and J. Aggarwal, "Matching aerial images to 3-D terrain maps," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, pp. 1138-1149, 1990.
- [77] J. Shang and Z. Shi, "Vision-based runway recognition for uav autonomous landing," *International Journal of Computer Science and ...*, 2007.
- [78] A. Omead, K. Takeo, and F. Keisuke, "A visual odometer for autonomous helicopter flight," *Robotics and Autonomous Systems*, vol. 28, 1999.
- [79] E. Berthier, H. Vadon, D. Baratoux, Y. Arnaud, C. Vincent, K. Feigl, *et al.*, "Surface motion of mountain glaciers derived from satellite optical imagery," *Remote Sensing of Environment*, vol. 95, pp. 14-28, 2005.
- [80] V. Kaufmann and R. Ladstädter, "Quantitative analysis of rock glacier creep by means of digital photogrammetry using multi-temporal aerial photographs: two case studies in the Austrian Alps," in *8th International Conference on Permafrost*, 2003, p. 530.
- [81] P. V. Potapov, S. A. Turubanova, M. C. Hansen, B. Adusei, M. Broich, A. Altstatt, *et al.*, "Quantifying forest cover loss in Democratic Republic of the Congo, 2000-2010, with Landsat ETM+ data," *Remote Sensing of Environment*, vol. 122, pp. 106-116, 2012.
- [82] B. Tan, J. G. Masek, R. Wolfe, F. Gao, C. Huang, E. F. Vermote, *et al.*, "Improved forest change detection with terrain illumination corrected Landsat images," *Remote Sensing of Environment*, vol. 136, pp. 469-483, 2013.
- [83] Z. Zhang, X. Wang, X. Zhao, B. Liu, L. Yi, L. Zuo, *et al.*, "A 2010 update of National Land Use/Cover Database of China at 1: 100000 scale using medium spatial resolution satellite images," *Remote Sensing of Environment*, vol. 149, pp. 142-154, 2014.
- [84] K. Cheng, C. Wei, and S. Chang, "Locating landslides using multi-temporal satellite images," *Advances in Space Research*, vol. 33, pp. 296-301, 2004.
- [85] P. Lu, A. Stumpf, N. Kerle, and N. Casagli, "Object-oriented change detection for landslide rapid mapping," *Geoscience and Remote Sensing Letters, IEEE*, vol. 8, pp. 701-705, 2011.
- [86] S. Leprince, F. Ayoub, Y. Klingert, and J.-P. Avouac, "Co-registration of optically sensed images and correlation (COSI-Corr): An operational methodology for ground deformation measurements," in *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, 2007, pp. 1943-1946.
- [87] A. Kääb and T. Prowse, "Cold - regions river flow observed from space," *Geophysical Research Letters*, vol. 38, 2011.

- [88] A. Münchow and H. Melling, "Ocean current observations from Nares Strait to the west of Greenland: Interannual to tidal variability and forcing," *Journal of Marine Research*, vol. 66, pp. 801-833, 2008.
- [89] A. Kääb and S. Leprince, "Motion detection using near-simultaneous satellite acquisitions," *Remote Sensing of Environment*, vol. 154, pp. 164-179, 2014.
- [90] R. S. Lunetta and C. D. Elvidge, *Remote sensing change detection: environmental monitoring methods and applications*: Taylor & Francis Ltd, 1999.
- [91] A. Thomson, "A multi-temporal comparison of two similar Landsat Thematic Mapper images of upland North Wales, UK," *International Journal of Remote Sensing*, vol. 13, pp. 947-955, 1992.
- [92] P. Teillet, B. Guindon, and D. Goodenough, "On the slope-aspect correction of multispectral scanner data," *Canadian Journal of Remote Sensing*, vol. 8, pp. 84-106, 1982.
- [93] C. DUGUAY and E. LEDREW, "Estimating surface reflectance and albedo from Landsat-5 Thematic Mapper over rugged terrain," *Photogrammetric Engineering and Remote Sensing*, vol. 58, pp. 551-558, 1992.
- [94] B. Tan, R. Wolfe, J. Masek, F. Gao, and E. F. Vermote, "An illumination correction algorithm on Landsat-TM data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, 2010, pp. 1964-1967.
- [95] R. Weismiller, S. Kristof, D. Scholz, P. Anuta, and S. Momin, "Change detection in coastal zone environments.[by Landsat MSS data analysis," 1977.
- [96] D. Williams and M. Stauffer, "Monitoring gypsy moth defoliation by applying change detection techniques to Landsat imagery," 1978.
- [97] C. Riordan, "Non-urban to Urban Land Cover Change Detection Using Landsat Data, Summary Report of the Colorado Agricultural Research Experiment Station," *Fort Collins. Colorado*, 1980.
- [98] J. C. Coiner, "Using Landsat to monitor changes in vegetation cover induced by desertification processes," in *International Symposium on Remote Sensing of Environment, 14 th, San Jose, Costa Rica*, 1980, pp. 1341-1351.
- [99] Y. Yasuoka, T. MIYAZAKI, Y. IIKURA, and S. OTOMA, "Detection of land-cover change from remotely sensed images using spectral signature similarity," in *Asian Conference on Remote Sensing, 9 th, Bangkok, Thailand*, 1988.
- [100] R. Wang, F. Bunyak, G. Seetharaman, and K. Palaniappan, "Static and moving object detection using flux tensor with split gaussian models," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on*, 2014, pp. 420-424.
- [101] K. Deilami and M. Hashim, "Very high resolution optical satellites for DEM generation: a review," *European Journal of Scientific Research*, vol. 49, pp. 542-554, 2011.
- [102] X. Liu, "Airborne LiDAR for DEM generation: some critical issues," *Progress in Physical Geography*, vol. 32, pp. 31-49, 2008.
- [103] M. Crosetto, "Calibration and validation of SAR interferometry for DEM generation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 57, pp. 213-227, 2002.
- [104] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision*, vol. 47, pp. 7-42, 2002.

- [105] C. Loop and Z. Zhang, "Computing rectifying homographies for stereo vision," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*, 1999.
- [106] V. Papadimitriou and T. J. Dennis, "Epipolar line estimation and rectification for stereo image pairs," *Image Processing, IEEE Transactions on*, vol. 5, pp. 672-676, 1996.
- [107] J. Oh, "Novel Approach to Epipolar Resampling of HRSI and Satellite Stereo Imagery-based Georeferencing of Aerial Images," The Ohio State University, 2011.
- [108] G. I. K. Morgan, "Sub-pixel Stereo Image-matching for Precise 3-d Surface Measurement," Imperial College London (University of London), 2010.
- [109] J. G. Liu, H. Yan, and G. Morgan, "PCIAS Subpixel Technology," *Measurement and Control*, vol. 45, 2012.